

GILNEI COSTA SANTOS

INSERÇÃO E RENDIMENTOS NO MERCADO DE TRABALHO BRASILEIRO

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

VIÇOSA
MINAS GERAIS – BRASIL
2009

**Ficha catalográfica preparada pela Seção de Catalogação e
Classificação da Biblioteca Central da UFV**

T

S237i
2009 Santos, Gilnei Costa, 1981-
Inserção e rendimentos no mercado de trabalho brasileiro /
Gilnei Costa Santos. – Viçosa, MG, 2009.
xii, 114f. : il. ; 29cm.

Inclui apêndices.

Orientador: Rosa Maria Oliveira Fontes.

Dissertação (mestrado) - Universidade Federal de Viçosa.

Referências bibliográficas: f. 71-78

1. Economia - Métodos estatísticos. 2. Mercado de trabalho
- Estatísticas. 3. Produtividade do trabalho.

4. Desenvolvimento econômico - Efeito da educação.

I. Universidade Federal de Viçosa. II. Título.

CDD 22. ed. 331

GILNEI COSTA SANTOS

INSERÇÃO E RENDIMENTOS NO MERCADO DE TRABALHO BRASILEIRO

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

APROVADA: 13 de julho de 2009.

Prof. Dr. Luiz Eduardo de Vasconcelos
Rocha

Prof. Dr. Sidney Martins Caetano

Prof. Dr. Maurinho Luiz dos Santos

Prof. Dr. João Eustáquio de Lima
(Co-orientador)

Prof^a. Dra. Rosa Maria Olivera Fontes
(Orientadora)

A Deus, a minha família, minha mãe Neida, meu pai Gilberto, meus irmãos Gilneana, Carlos Henrique, Giovanni, Sônia, minha namorada Patrícia e também a Hannicka e Branca, por acreditarem em mim.

“Comece pelo começo”.

Allan Kardec

AGRADECIMENTOS

A Deus, pela saúde e alento nos momentos difíceis.

Aos meus pais, pela confiança e apoio incondicional.

Aos meus irmãos, Carlos Henrique (Cacá), Giovanni (Macaco), Sônia (Soninha) pelos nossos momentos de descontração e especialmente à minha irmã Gilneana (Teté) pelo seu carinho, Amor, e pelos incentivos constantes.

A minha futura esposa Patrícia por ser meu combustível, meu alicerce, minha vida. Se Deus quiser em Outubro de 2010 sai o casamento.

A Hannicka e Branca por ser a alegria da família.

Aos meus familiares por acreditarem em mim.

A minha orientadora Prof. Rosa Fontes, por sua gentileza, calma, por sempre estar pronta a me ajudar e por sua confiança no meu potencial.

Aos meus parceiros no mestrado: Leonardo, Françoise, Nayana, Larissa, Tatiana, Vitor, João Guilherme, Gustavo, Gisele, Jamilsen e principalmente ao Luckas, Chrystian e Vinícius pelos bons momentos.

Gostaria de agradecer também à banca que avaliou esta dissertação, professores: João Eustáquio de Lima, Luiz Eduardo de Vasconcelos Rocha, Maurinho Luiz dos Santos e Sidney Martins Caetano, pelas sugestões e críticas que enriqueceram o trabalho. Enfatizando meu profundo agradecimento ao Prof. Maurinho Luiz dos Santos por sua grande contribuição no referencial teórico.

Por fim, aos professores do Departamento de Economia pela indiscutível contribuição na minha formação profissional.

SUMÁRIO

LISTA DE TABELAS	vii
LISTA DE FIGURAS	viii
RESUMO.....	ix
ABSTRACT	xi
1. INTRODUÇÃO.....	1
1.1. Considerações Iniciais.....	1
1.2. O Problema e sua Importância.....	6
1.2.1. Mercado de Trabalho e Capital Humano	7
1.2.2. Mercado de Trabalho e as Características Não Ligadas à Produtividade	7
1.2.2.1. Renda e Mercado de Trabalho Feminino	8
1.2.2.2. Renda e Trabalho entre Raças	10
1.2.3. Mercado de Trabalho no Meio Rural.....	11
1.3. Objetivos	14
2. REFERENCIAL TEÓRICO.....	16
2.1. Decisão de Participação no Mercado de Trabalho.....	16
2.2. Determinantes do Rendimento do Trabalho	21
3. METODOLOGIA.....	23
3.1. Medidas de Desigualdade	23
3.1.1. Índice de Gini	23
3.1.2. Índices de Theil	24
3.2. Medidas de Pobreza	25
3.2.1. Índices FGT	25
3.3. Procedimento de Heckman	28
3.4. Descrição das Equações e Variáveis Seleccionadas.....	30
3.5. O Efeito Limiar da Educação.....	33
4. RESULTADOS E DISCUSSÃO	36
4.1. Desigualdade, Pobreza e Determinantes da Participação e Rendimentos no Mercado de Trabalho Brasileiro	36
4.1.1. Estatísticas Descritivas dos Dados da Amostra.....	36
4.1.2. Análise da Desigualdade da Renda no Mercado de Trabalho Brasileiro entre 2002 e 2007.....	38

4.1.3. Análise da Pobreza no Mercado de Trabalho Brasileiro entre 2002 e 2007	41
4.1.4. Participação no Mercado de Trabalho Brasileiro e Equações de Rendimento	44
4.1.5. Equação de rendimento para o Brasil 2002 e 2007	50
4.2. Determinantes da Participação e Rendimentos no Mercado de Trabalho Rural Brasileiro	54
4.2.1. Estatísticas Descritivas dos Dados da Amostra	54
4.2.2. Equação de Participação para o Meio Rural Brasileiro	57
4.2.3. Equação de Rendimento para o Meio Rural do Brasil	62
5. CONSIDERAÇÕES FINAIS	67
6. REFERÊNCIAS	71
APÊNDICE 1	79
Resultados Complementares	79
APÊNDICE 2	90
1.1. Características Básicas	90
1.2. O Plano Amostral	91
1.2.1. Construção do Desenho Amostral da PNAD	92
1.2.2. Análise Estatística em Amostras Complexas	94
1.2.3. Restrições da Pesquisa Nacional por Amostra de Domicílios	96
APÊNDICE 3	97
Exemplo do Algoritmo Desenvolvido para a Análise do Mercado de Trabalho Rural brasileiro 2007	97

LISTA DE TABELAS

Tabela 1.1. Características do Mercado de Trabalho Brasileiro nos anos de 1940 e 1980. ...	2
Tabela 1.2. Características do Mercado de Trabalho Brasileiro nos anos de 1980-2007.....	3
Tabela 4.1. Estatísticas Descritivas das Variáveis de Interesse, Brasil 2002-2007.....	37
Tabela 4.2. Equação de Participação no Mercado de Trabalho Brasileiro, 2002.....	46
Tabela 4.3. Equação de Participação no Mercado de Trabalho Brasileiro, 2007.....	47
Tabela 4.4. Equação de Rendimento para o Brasil 2002.....	51
Tabela 4.5. Equação de Rendimento para o Brasil 2007.....	52
Tabela 4.6. Estatísticas Descritivas das Variáveis de Interesse no Meio Rural, Brasil 2002-2007.....	56
Tabela 4.7. Equação de Participação no Mercado de Trabalho do Meio Rural, Brasil 2002.....	58
Tabela 4.8. Equação de Participação no Mercado de Trabalho do Meio Rural, Brasil 2007.....	59
Tabela 4.9. Equação de Rendimento para o Meio Rural, Brasil 2002.....	63
Tabela 4.10. Equação de Rendimento para o Meio Rural, Brasil 2007.....	64
Tabela A.1.1. Teste de Heterocedasticidade para o Mercado de Trabalho, Brasil 2002.....	79
Tabela A.1.2. Teste de Heterocedasticidade para o Mercado de Trabalho, Brasil 2007.....	80
Tabela A.1.3. Equação de Participação no Mercado de Trabalho Brasileiro, 2002.....	81
Tabela A.1.4. Equação de Rendimento, Brasil 2002.....	82
Tabela A.1.5. Equação de Participação no Mercado de Trabalho Brasileiro, 2007.....	83
Tabela A.1.6. Equação de Rendimento Utilizando o Procedimento de Heckman, Brasil 2007.....	84
Tabela A.1.7. Equação de Participação no Mercado de Trabalho do Meio Rural, Brasil 2002.....	85
Tabela A.1.8. Equação de Participação no Mercado de Trabalho do Meio Rural, Brasil 2007.....	86
Tabela A.1.9. Equação de Rendimento para o Meio Rural, Brasil 2002.....	87
Tabela A.1.10. Equação de Rendimento para o Meio Rural, Brasil 2007.....	88
Tabela A.2.1. Características da Amostra da PNAD.....	91

LISTA DE FIGURAS

Figura 1.1. Alguns Indicadores do Mercado de Trabalho Brasileiro nos Anos 2000.	4
Figura 1.2. Taxa de Atividade da População Brasileira em Relação ao Percentual de Desempregados, 2001-2007.....	5
Figura 2.1. Alocação do tempo do indivíduo entre trabalho e lazer.....	17
Figura 2.2. Decisão de oferta de trabalho a partir de uma solução de canto.	19
Figura 2.3. Determinando o salário reserva a partir da equação de Slutsky.....	20
Figura 3.1. Retorno dos anos de estudo sobre a razão de log da renda e total de horas trabalhadas, sem instrução ou menos de um ano é o controle.	35
Figura 4.1. Índices de desigualdade para o Brasil e grupos de interesse, 2002.....	39
Figura 4.2. Índices de desigualdade para o Brasil e grupos de interesse, 2007.....	40
Figura 4.3. Índices de Pobreza no Brasil em 2002.	42
Figura 4.4. Índices de Pobreza no Brasil em 2007.	43
Figura 4.5. Contribuição dos Grupos de Interesse para a Pobreza Brasileira, 2002.	43
Figura 4.6. Contribuição dos Grupos de Interesse para a Pobreza Brasileira, 2007.	44
Figura 4.7. Rendimento médio de todos os trabalhos para homens e mulheres no meio rural, 2007.	65
Figura A.2.1. Construção do plano amostral da PNAD.	93

RESUMO

SANTOS, Gilnei Costa, M.Sc. Universidade Federal Viçosa, Julho de 2009. **Inserção e rendimentos no mercado de trabalho brasileiro.** Orientadora: Rosa Maria Olivera Fontes. Co-orientadores: Silvia Harumi Toyoshima e João Eustáquio de Lima.

Vários são os fatores que afetam os rendimentos de um indivíduo, como a distribuição de renda no país, o estoque de capital humano ou questões não relacionadas à produtividade, como o sexo, raça ou a região em que a pessoa reside. No entanto, deve-se considerar que boa parte da renda dos indivíduos é proveniente do trabalho. Neste sentido, os fatores citados, ao menos indiretamente, são afetados pelo mercado de trabalho. O objetivo da presente pesquisa consiste em analisar o comportamento do mercado de trabalho brasileiro, nos anos recentes, sobre a perspectiva tanto da entrada do indivíduo no mercado como sob a ótica dos rendimentos obtidos, utilizando variáveis relacionadas ao gênero, raça, região geográfica e vulnerabilidade, como por exemplo número de crianças na família ou mãe solteira. A base de dados são os microdados da PNAD nos anos de 2002 e 2007, cuja característica de amostra complexa foi considerada, para não viesar os estimadores. A metodologia adotada é o procedimento de Heckman, visando evitar o viés de seletividade. Os resultados indicaram traços de segmentação no mercado de trabalho brasileiro e no mercado de trabalho rural, por gênero e raça ou cor. Em termos de cor ou raça, apesar de não se verificar empecilhos à entrada no mercado rural, observa-se que os negros têm, em média, rendimentos mais baixos que os brancos. Por outro lado, para o mercado de trabalho como um todo, negros e pardos apresentaram maior probabilidade de estarem economicamente ativos do que brancos, contudo com retornos consideravelmente menores do que o dos brancos, sugerindo que os pardos e negros possuem salários reserva mais baixos. Quando o mercado é analisado pelo gênero, os resultados mostram que, apesar das

mulheres possuírem um nível de qualificação médio maior que o dos homens, observa-se que o rendimento destes tende significativamente mais elevado do que o do sexo feminino para os indivíduos situados no meio rural. O mesmo ocorre quando considera-se o mercado de trabalho sem o corte do rural. Ademais, a desigualdade de renda é mais elevada entre as mulheres, assim como os níveis de pobreza, quando se considera o mercado de trabalho como um todo. Uma explicação para isso está relacionada às atividades desenvolvidas por homens e mulheres. Historicamente ocupações de baixo rendimento, como o trabalho doméstico, estão associadas ao sexo feminino, o que é verificado na presente pesquisa, onde a grande maioria dos indivíduos ocupados em trabalhos domésticos são mulheres. Analisando-se os diferenciais entre mães solteiras e as demais mulheres, verifica-se que as primeiras têm maior probabilidade de estarem no mercado rural, mas com rendimentos menores do que as demais mulheres, sugerindo que o salário reserva das mães solteiras é consideravelmente menor. Estes resultados apontam para um agravamento na situação das mulheres no mercado de trabalho rural. Ou seja, além das indicações de discriminação por gênero, existe ainda uma propagação do que se pode chamar de hiato de renda entre mães solteiras e as demais mulheres no meio rural. Finalmente, quanto à variável educação, cada ano adicional investido em qualificação eleva a probabilidade de o indivíduo estar empregado no mercado de trabalho como um todo. No mercado de trabalho rural, cada ano adicional aumenta a probabilidade das mulheres estarem empregadas, enquanto que para os homens a qualificação formal não mostra significância estatística. Relacionando educação e rendimentos, os resultados sugerem que investimentos em qualificação formal é uma das principais fontes de geração de renda, principalmente para os homens. Outrossim, a *proxy* de experiência apresentou uma trajetória parabólica, sugerindo que o capital humano deprecia-se ao longo do tempo.

ABSTRACT

SANTOS, Gilnei Costa, M.Sc., Universidade Federal de Viçosa, July, 2009. **Participation and earnings in the brazilian labor market.** Adviser: Rosa Maria Olivera Fontes. Co-Advisers: Silvia Harumi Toyoshima and João Eustáquio de Lima.

There are several factors that affect the income of an individual, such as distribution of income in the country, the stock of human capital or questions not related to productivity, such as sex, race or region where the person resides. However, one should consider that much of the income of individuals is derived from the work. Accordingly, the factors cited above, at least indirectly, are affected by the labor market. The objective of this research is to examine the dynamics of the Brazilian labor market in recent years both from the perspective of the individual's entry in the market and from the point of view of income, using variables related to gender, race, geographic region and vulnerability, such as number of children in the family or single mother condition. The database is the PNAD microdata for the years 2002 and 2007, which respects the feature of complex sample and does not to bias the results. The methodology is the Heckman procedure, to avoid the selectivity bias. The results showed traces of segmentation in the Brazilian labor market as a whole and in the rural labor market, by gender and race or color. In terms of color or race, although there was not detected an impediment to entrance into the rural labor market, it appears that blacks had earnings lower than the whites. Moreover, for the labor market as a whole, the black and brown are more likely to be economically active than white, but with significantly lower returns than the whites, suggesting that the brown and black people tend to get lower reservation wages. When the market is analyzed by gender, the results had showed that although women have a higher average skill level than men, it is observed that the male earnings tend to be higher than the female for individuals located

in rural areas, the same happens for Brazilian labor market as a whole. Moreover, inequality of income is higher among women, as well as the levels of poverty, when one considers the labor market as a whole. One explanation for this is related to the activities undertaken by men and women. Historically low-income occupations such as domestic work are associated with female sex, which is confirmed in this study, where the great majority of individuals employed in domestic work are women. Looking up the differences between single mothers and other women, the first are more likely to be in the rural market, suggesting that the reservation wage of single mothers is considerably smaller. These results point to a worsening in the situation of women in the rural work. That is, beyond the indications of discrimination by gender, there is a spread of what can be called the gap of income among single mothers and other women in rural areas. Finally, as the education variable, each additional year invested in qualification increases the probability of the individual being employed in the labor market as a whole. In the rural labor market, each additional year of education increases the probability for women, while for men it was not statistically significant. Relating education and earnings, the results suggest that investments in formal qualification are one of the main sources of generation of income, mainly for the men. Likewise, the proxy of experience presented a parabolic trajectory, suggesting that the human capital depreciates over time.

1. INTRODUÇÃO

1.1. Considerações Iniciais

O mercado de trabalho brasileiro tem apresentado modificação contínua ao longo dos anos, sendo assim esta seção apresenta um histórico do mercado de trabalho brasileiro em que foi possível verificar três comportamentos distintos no período analisado. Primeiramente observou-se a estruturação do mercado de trabalho, mesmo que incompleta, entre os anos de 1930 a início de 1980. Posteriormente, ocorreram alterações na estrutura do mercado de trabalho reduzindo o avanço do processo inicial, dentre outros motivos, devido ao período de estagflação seguido da abertura comercial dos anos 90. Nos anos 2000, os resultados mostram indícios ainda pequenos de um novo perfil do mercado de trabalho brasileiro.

De acordo com Mendonça (2003), a partir dos anos 1940, ocorreu no Brasil uma maior estruturação do mercado de trabalho, que durou até o fim dos anos de 1970. As principais características deste período foram a elevação do assalariamento formal e a redução das taxas de desemprego, sendo esse período coincidente com o ciclo de industrialização do Brasil.

Por estruturação, Pochmann (2006) considera o desenvolvimento de ocupações mais homogêneas e uma maior organização do mercado de trabalho urbano, em que prevalecem as ocupações em empresas privadas, estatais e na administração pública e em termos de remuneração verifica-se o trabalho assalariado regularizado. Exemplificando a estruturação do mercado de trabalho brasileiro, Pochmann (2001) demonstra que, de cada

dez postos de trabalhos criados, entre 1930 e 1980, oito eram assalariados e desses, sete tinham registro formal.

A Tabela 1 apresenta a estrutura do mercado de trabalho brasileiro nos anos de 1940 e 1980. Ao longo dessas décadas, a população economicamente ativa (PEA) cresceu aproximadamente 67%, e a população economicamente ativa ocupada atingiu 97,2%.

Tabela 1.1. Características do Mercado de Trabalho Brasileiro nos anos de 1940 e 1980.

	1940	1980
PEA	100%	100%
PEA ocupada	93,70%	97,20%
Conta-própria	29,80%	22,10%
Sem remuneração	19,60%	9,20%
Assalariado	42,00%	62,80%
- Com registro	12,10%	49,20%
- Sem registro	29,90%	13,60%
Desempregados	6,30%	2,80%
Trabalho precário*	55,70%	34,10%

*Segundo Pochmann (2006), o trabalho precário corresponde à soma dos desempregados, sem remuneração e conta-própria.

Fonte: Pochmann (2006), adaptado pelo autor.

Neste período de intensa industrialização e da institucionalização das relações do trabalho no Brasil, o percentual de desempregados caiu consideravelmente, chegando em 1980 a 2,80% dos indivíduos economicamente ativos. Soma-se a isso o fato de que o assalariamento formal aumentou em aproximadamente 38 p.p., atingindo, em 1980, 49,2% dos assalariados, além da redução do trabalho precário.

A partir da década de 1980, ocorre uma expansão da População Economicamente Ativa (PEA), época em que ocorre alterações na estrutura do mercado de trabalho. Dentre outros motivos, de acordo com Pochmann (2001), o processo de desregulamentação do mercado de trabalho brasileiro toma impulso devido à crise da dívida externa que marcou o início de um período de relativa estagnação econômica atrelado a uma crescente inflação. Outrossim, a abertura comercial do país no início da década de 90, sem a devida cobertura dos produtores internos, fragilizou ainda mais o mercado de trabalho brasileiro. Camargo *et al.* (1999) destacam como efeito desta desestruturação o inchaço do setor terciário seguido da expansão da informalidade, redução do emprego industrial, precarização das atividades, além do aumento da demanda por qualificação da mão de obra, o que acaba por gerar ocupações heterogêneas, peculiares das economias subdesenvolvidas.

A Tabela 1.2 corrobora as afirmativas anteriores. Os resultados demonstram que de 1980 a 2000 ocorreu um incremento nas atividades consideradas precárias, e em 2000 os grupos de conta-própria, sem remuneração e desempregados somaram 40% da PEA. Outra característica interessante consiste no percentual de trabalhadores sem registros, ou seja, o grau de informalidade da economia que chegou a 20,9% em 2000. Um dos fatores que explicam o crescimento tão acentuado da informalidade é a reestruturação dos setores produtivos, com a redução de ocupações formais e a abertura comercial desordenada (DEDECCA e BALTAR, 1997). Ou seja, apesar da crescente PEA nos anos 90, boa parte dessa população não foi absorvida pelo setor formal nacional, sendo destinada a ocupações menos articuladas onde prevalece o mercado de trabalho informal. É importante destacar ainda que o nível de desemprego em 2000 no Brasil alcançou 15% da população economicamente ativa, um salto de mais de 12 p.p. em vinte anos.

Tabela 1.2. Características do Mercado de Trabalho Brasileiro nos anos de 1980-2007**.

	1980	2000	2001	2002	2003	2004	2005	2006	2007
PEA	100%	100%	100%	100%	100%	100%	100%	100%	100%
PEA ocupada	97,2%	85%	90.6%	91.1%	90.5%	91.4%	91%	91.8%	91.8%
Conta-própria	22,10%	19,10%	23.84%	23.98%	24.07%	23.76%	23.79%	23.69%	23.52%
Sem remuneração	9,20%	6,30%	6.93%	6.68%	6.38%	6.34%	6.16%	5.54%	5.37%
Assalariado	62,80%	57,20%	56.07%	56.32%	56.03%	57.26%	56.90%	58.28%	59.50%
- Com registro	49,20%	36,30%	34.22%	34.28%	34.85%	35.47%	35.70%	37.00%	38.67%
- Sem registro	13,60%	20,90%	21.85%	22.05%	21.18%	21.79%	21.20%	21.28%	20.84%
Desempregados	2,80%	15%	9.34%	9.16%	9.73%	8.90%	9.31%	8.42%	8.15%
Trabalho precário*	34,10%	40,40%	40.12%	39.82%	40.18%	38.99%	39.26%	37.64%	37.05%

*Segundo Pochmann (2006), o trabalho precário corresponde à soma dos desempregados, sem remuneração e conta-própria.

** Os resultados desta tabela podem apresentar certa diferenciação das Tabelas 1.1 e 1.2 devido à fonte dos dados, assim como à metodologia utilizada em cada pesquisa.

Fonte: Elaboração do autor a partir dos microdados das PNADs de 2001-2007.

Os resultados da Tabela 1.2 mostram que as características que marcaram o mercado de trabalho brasileiro em 1980 perderam força durante os anos 2000. Nesse período, ocorre pequena melhora dos indicadores do mercado de trabalho, conforme mostrado na Figura 1.1. A taxa de assalariamento no período teve um leve incremento, chegando a quase 60% da PEA ocupada, valor próximo ao obtido em 1980. No entanto, o número de trabalhadores com carteira assinada em 2007 foi consideravelmente menor do que em 1980, 49,20% contra 38,67%.

Apesar da oscilação do desemprego entre 2001 e 2007, quando são observados apenas os extremos, os resultados mostram redução na taxa de desemprego no período, chegando em 2007 a 8,15% da PEA. Por fim, a precarização das atividades se reduziu ao patamar de 37,05% em 2007. Apesar de não superar o resultado de 1980, de 34,10%, foi o segundo melhor valor nos mais de 60 anos analisados.

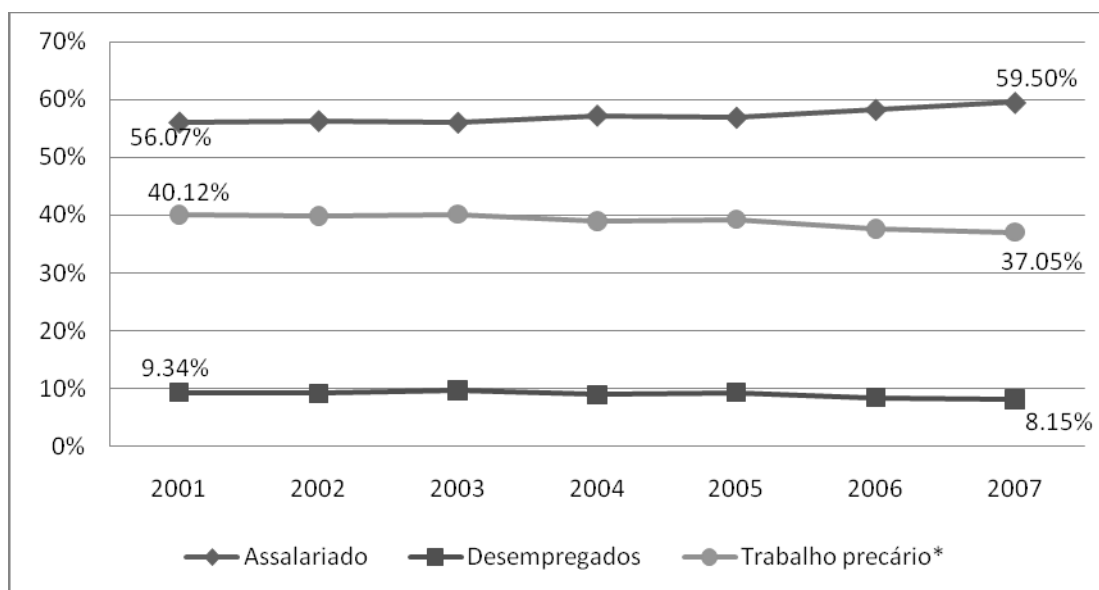


Figura 1.1. Alguns Indicadores do Mercado de Trabalho Brasileiro nos Anos 2000.

Fonte: Elaboração do autor a partir dos microdados das PNADs de 2001-2007.

Outro tipo de análise que pode ser feita a partir dos dados do mercado de trabalho nos anos 2000, conforme Figura 1.2, consiste em relacionar a taxa de atividade do mercado de trabalho¹ com o percentual de desempregados. Desta forma, é possível verificar o nível de oferta de mão de obra em relação ao patamar dos desempregados, determinando em que anos o mercado de trabalho se encontrava numa situação mais confortável.

¹ A taxa de atividade do mercado de trabalho corresponde ao percentual de indivíduos economicamente ativos em relação a toda a população.

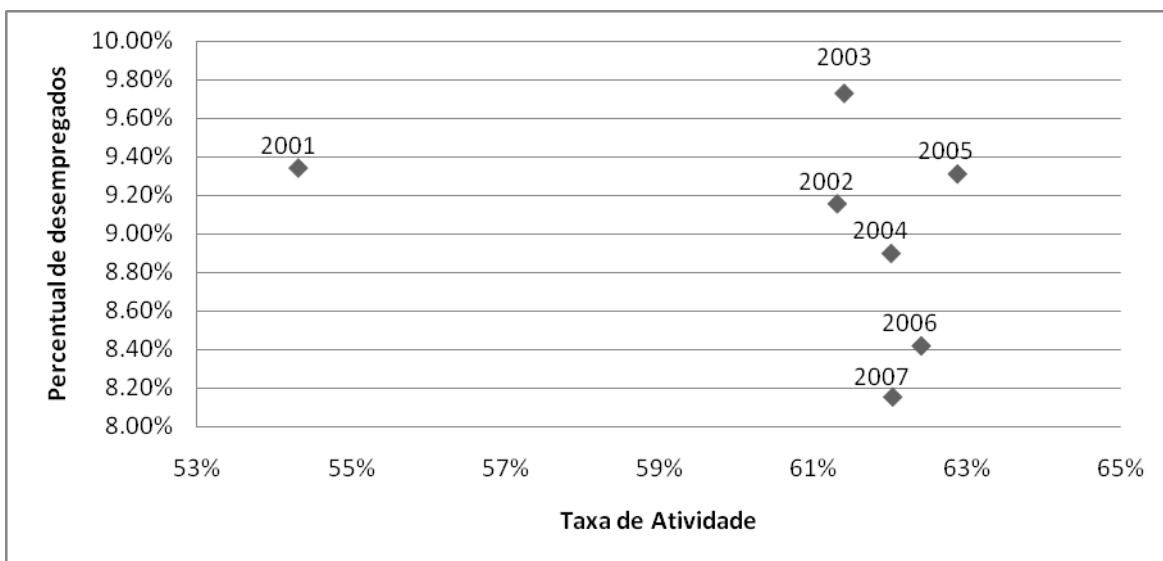


Figura 1.2. Taxa de Atividade da População Brasileira em Relação ao Percentual de Desempregados, 2001-2007.

Fonte: Elaboração do autor a partir dos microdados das PNADs 2001-2007.

Apesar de, em 2001, existir a menor taxa de atividade do período, ou seja, o menor número de pessoas economicamente ativas em relação ao total da população, o percentual de desempregados no Brasil nesse ano é o segundo maior, ficando abaixo apenas de 2003. Segundo Ramos (2007), o fator gerador dessa pressão foi o racionamento de energia no período. No outro extremo, encontra-se o ano de 2007, com um dos maiores níveis de atividade, próximo a 62%, e, apesar disso, a taxa de desemprego é a menor da série, próxima a 8% da PEA, mostrando que capacidade de geração de empregos e absorção de mão de obra pelo mercado em 2007 é maior do que em 2001.

Diante deste panorama e considerando que o indivíduo per se é peça fundamental no mercado, uma vez que a oferta de trabalho é definida pela decisão pessoal, e ademais a própria demanda por trabalho é determinada, dentre outros fatores, pela capacidade produtiva do ofertante, Ramos (2007) diz:

"... a importância fundamental do mercado reside no fato de ser o local onde os indivíduos transacionam, aos preços assim determinados, o seu principal – e, na maioria das vezes, único – ativo, que é a sua capacidade laboral. Daí a importância desse mercado para assuntos ligados à pobreza, equidade e bem-estar."

A próxima seção apresenta os objetos de estudo do presente trabalho, que são os fatores que determinam a entrada do indivíduo no mercado de trabalho brasileiro, como os ligados à capacidade produtiva ou aqueles não relacionados a esta. Ademais, após o indivíduo encontrar-se no mercado, o trabalho analisa quais os fatores que afetam a renda

do trabalho. De forma geral, Gandra (2002) verifica que atributos como raça, sexo, idade, etnia e credenciais formais são variáveis determinantes no processo de seleção de trabalhadores e determinação de salários no Brasil.

1.2. O Problema e sua Importância

São inúmeros os fatores que afetam os rendimentos de um indivíduo, como a distribuição de renda no país, as condições macroeconômicas em determinado período, o estoque de capital humano, ou seja, o nível de escolaridade e/ou experiência, assim como os não relacionados à produtividade, como os diferenciais de remuneração por gênero, raça ou pela região em que a pessoa reside. No entanto, deve-se considerar que boa parte da renda dos indivíduos é proveniente do trabalho. Como destaca IPEA (2009), no Brasil os rendimentos do trabalho representaram, em 2007, cerca de 41,7% da renda nacional. Neste sentido, os fatores citados, ao menos indiretamente, são afetados pelo mercado de trabalho.

O Brasil é um dos países com maior desigualdade de renda do mundo, dependendo da medida empregada², o que impacta no comportamento dos níveis de pobreza e justifica a existência de uma vasta literatura sobre sua evolução e determinantes, onde se encontram trabalhos como os de Morley (2001) e Ribeiro (2006) para o caso da América Latina, e especificamente para o Brasil como, Barros e Mendonça (1993, 1995a e b), Hoffmann (1989, 1992, 1996, 1998, 2000), Teixeira (2006) e Barros *et al.* (2007), entre outros. Dentre os determinantes desta disparidade de renda, além de fatores históricos, o nível educacional da população merece destaque. Outros fatores explicativos seriam a influência do mercado de trabalho, por meio de suas características potenciais de discriminação e segmentação, e os diferenciais de remuneração regional.

Nas próximas subseções segue um breve histórico sobre a literatura nacional e internacional, que tratam do tema determinantes da inserção e rendimentos no mercado de trabalho, assim como a forma como o trabalho pretende contribuir com o desenvolvimento do tema no caso específico do mercado de trabalho brasileiro, enfatizando também o mercado de trabalho rural.

² Como o índice de Gini, por exemplo.

1.2.1. Mercado de Trabalho e Capital Humano

Ramos e Vieira (2001) apontam o alto poder explicativo da escolaridade sobre a renda em vários países latino-americanos. A contribuição da escolaridade é ainda maior no Brasil, segundo IPEA-CAIXA (2007). A grande influência da educação sobre a desigualdade salarial brasileira ocorre por dois motivos: a elevada desigualdade educacional entre trabalhadores e a elevada sensibilidade dos salários em relação ao nível educacional. Esse mesmo pensamento é compartilhado por Morley (2001), verificando que no Brasil os retornos por um alto nível educacional são maiores do que em outros países de mesmo grau de desenvolvimento, como os asiáticos. Segundo o mesmo autor, no Brasil existe uma política de educação pró-desigualdade, na medida em que a mesma prioriza investimentos em ensino universitário. Nesse sentido, no caso brasileiro, a qualificação formal é uma das principais fontes de geração de renda.

Em consonância, Dhareshwar (2001) relata que “nenhum país conseguiu desenvolvimento sustentado sem investir substancialmente e eficientemente na educação e na saúde de seu povo. Países em desenvolvimento geralmente vêm investindo mais recursos públicos na educação”. Segundo Bonelli (2002), ao se considerar que a escolaridade é um dos principais determinantes da renda dos indivíduos, a transmissão de escolaridade entre gerações é um mecanismo de perpetuação da pobreza: famílias pobres e com baixa escolaridade tendem a educar menos seus filhos e aumentam as chances de que estes estejam abaixo da linha da pobreza.

1.2.2. Mercado de Trabalho e as Características Não Ligadas à Produtividade

A questão da discriminação e/ou segmentação no mercado de trabalho analisa a situação desfavorável a que estão sujeitas variáveis não relacionadas à produtividade, como raça, gênero e região geográfica. Ocorre discriminação salarial, por exemplo, se mulheres ou negros recebem menos do que homens ou não negros que possuem as mesmas aptidões, condições e atribuições no mercado de trabalho. Segundo Cambota (2006), a discriminação no mercado de trabalho é uma das principais fontes da desigualdade de rendimentos entre raça e gênero no Brasil e, portanto, da pobreza desses indivíduos. Henriques (2001) calcula, com base na PNAD de 1999, que embora negros e pardos representem 45% da

população brasileira, eles correspondem a 64% da população pobre e a 69% da população indigente. Soares (2000) estima que, por causa da discriminação, homens negros ganham em torno de 5% a 20% menos que homens brancos. Já as chances de as mulheres saírem da condição de pobreza também são reduzidas pela discriminação no mercado de trabalho, na medida em que mulheres com as mesmas capacidades produtivas que os homens recebem salários menores. Leme e Wajnman (2000) estimam que no Brasil os homens ganhem em média 40% mais que as mulheres, o que afeta a oferta de trabalho feminino, comprometendo o papel das mulheres como provedoras de renda. Neste sentido, raça e gênero são variáveis importantes no estudo da pobreza e da desigualdade de renda e com potencial de influência na inserção e nos rendimentos do mercado de trabalho.

1.2.2.1. Renda e Mercado de Trabalho Feminino

A crescente participação das mulheres no mercado de trabalho, observada a partir do século XX, foi intensificada no Brasil a partir da década de 70. O período de elevado crescimento brasileiro proporcionou oportunidade de maior inclusão das mulheres no mercado de trabalho e atualmente existem localidades em que o contingente feminino já supera o masculino em termos de participação no mercado. Segundo Soares e Izaki (2006), enquanto, em 1977, a taxa de participação feminina no mercado era de 32%, essa já atingia 39% em 1989, alcançando 46% em 2001 e 48% em 2006, de acordo com dados da Pesquisa Nacional por Amostra de Domicílios (PNAD). Estes resultados demonstram consideráveis mudanças na estrutura do mercado de trabalho brasileiro, sendo favorável ao sexo feminino ao menos em termos de inserção no mercado.

Alguns fatores que podem ter afetado a incorporação tão acentuada das mulheres no mercado de trabalho brasileiro são, conforme destacam Cirino e Lima (2008), a crescente taxa de urbanização e industrialização da economia brasileira, além da abertura comercial na década de 90, que prejudicou os níveis de emprego e renda e possivelmente incentivou as mulheres a entrar no mercado de trabalho para complementar a renda familiar. Essa hipótese advém do conceito de trabalhador adicional, que pressupõe que parte da decisão das mulheres de ofertar trabalho está relacionada à situação do cônjuge no mercado, ou

seja, a mulher inativa casada tende a ofertar trabalho, ao menos temporariamente, quando o marido se encontra desempregado³, de forma a suprir o déficit na renda familiar.

Outro fator que pode ter impulsionado a entrada das mulheres no mercado foi a considerável diminuição da taxa de fecundidade. No ano de 2006, o Instituto Brasileiro de Geografia e Estatística (IBGE)⁴ registrou a mais baixa taxa de fecundidade no País, cerca de 2 filhos por mulher, inferior à taxa de 1970, que se encontrava em torno de 5,8 filhos por mulher. Outro fator é a evolução no nível de qualificação da mulher. De acordo com o IBGE, a escolaridade feminina passou de 3,9 anos de estudo, em média, em 1977 para 7 anos em 2006.

Essas mudanças nos indicadores de gênero, para o caso brasileiro, apontam para uma nova formatação das relações sociais – as mulheres estão participando mais do mercado de trabalho, da chefia das famílias e as taxas de fecundidade diminuíram. Nesse cenário, é relevante a análise dos rendimentos e salários, dado o amplo reconhecimento no que diz respeito à igualdade de direitos e de oportunidades para homens e mulheres. O estudo dos rendimentos femininos também é importante devido à presença de um grande número de mulheres que sustentam seus lares como chefes de família e de muitas outras que colaboram com seus rendimentos em épocas de desemprego e de queda do rendimento familiar (CHAVES, 2002). Entretanto, as disparidades ainda existem nas relações de gênero. De acordo com o relatório de desenvolvimento humano (PNUD, 2005), de 1992 a 2003, no Brasil, a proporção de mulheres desempregadas foi sempre superior à de homens na mesma condição.

Ainda sobre questões relacionadas a gênero, no cenário internacional, em pesquisa sobre o Reino Unido, Davies e Joshi (1998) discutem a desigualdade de rendimento entre homens e mulheres e os impactos sobre a família e os níveis de pobreza, enfatizando a participação feminina no mercado de trabalho. Neste mesmo estudo, os autores demonstram que, em 1990, o rendimento feminino foi um importante fator para manter a família fora dos níveis de pobreza, mas, por outro lado, famílias chefiadas por mulheres tenderam a apresentar maior grau de pobreza do que as chefiadas por homens. Em outro estudo, Entwisle *et al.* (1995) analisam o mercado de trabalho rural na China através da visão de “*family businesses*”, enfocando o papel da mulher neste mercado. Os autores constatam que, para o ano de 1989, os negócios familiares onde existia elevado número de

³ Estudos como os de Kohara (2008) para o caso do mercado de trabalho do Japão, Serneels (2002) em estudo sobre a Etiópia, entre outros, como o de Oliveira e Oliveira (2006) para o Brasil, discutem esse tema.

⁴ Dados do banco SIDRA (Sistema IBGE de Recuperação Automática).

mulheres tendiam a permanecer pequenos. Neste sentido, as atividades das mulheres predominantes neste meio figuravam entre jardinagem e criação de animais, ou seja, atividades secundárias.

Em estudo sobre o mercado de trabalho nos Estados Unidos, Pettersen e Morgan (1995) argumentam que a segmentação ocupacional entre homens e mulheres explica a maior parcela da desigualdade de renda, contribuindo ainda para a segmentação no próprio local de trabalho. Ademais, os autores verificam que o hiato de renda entre os sexos aumenta de acordo com o status da ocupação dentro das firmas. Voltando ao mercado de trabalho brasileiro, Posthuma e Lombardi (1996) chegam a resultados semelhantes, concluindo que, nas empresas, as mulheres, de modo geral, são destinadas a atividades que demandam pouca capacitação e responsabilidade e, conseqüentemente, com menor remuneração.

1.2.2.2. Renda e Trabalho entre Raças

Outra forma potencial de discriminação refere-se à raça ou cor, sendo a principal representante a raça negra que, assim como a população feminina empregada, tende a ocupar cargos de baixa remuneração. Segundo PNUD (2005), ao longo do período de 1992 a 2003, para o mesmo grupo etário, a taxa de desemprego da população negra foi, em média, 23% superior à da população branca. Mulheres e negros brasileiros são os grupos mais sub-representados pelos indicadores sociais. Analisando-se os dados do IBGE e do IPEA, fica claro que a população negra está em situação mais crítica em relação ao desemprego, educação e renda e, quando as mulheres são incluídas, a situação parece se agravar, principalmente para as mulheres negras.

No Brasil, 51% do total da população é branca, e a renda média dos domicílios chefiados por brancos é cerca de 130% maior que a dos encabeçados por negros (IBGE, 2006⁵). Além disso, os negros estão associados aos 70% dos domicílios mais pobres, aumentando a probabilidade de suas escolhas educacionais serem feitas com recursos limitados, o que posteriormente restringe suas oportunidades no mercado de trabalho.

Browne (1997) analisa os diferenciais de rendimento entre raça, mais especificamente entre mulheres negras e brancas nos Estados Unidos no ano de 1989. A autora destaca que, dentre os fatores que geram os diferenciais, encontram-se

⁵ Dados do banco SIDRA (Sistema IBGE de Recuperação Automática).

primeiramente o capital humano e os conflitos na tomada de decisão de entrada no mercado de trabalho, como o cuidado com filhos pequenos e/ou cuidados domésticos. Um segundo fator seria a estrutura industrial na qual estas mulheres estão inseridas, ou seja, áreas em que a oportunidade de emprego é menos favorável. E, por fim, verifica-se a segmentação do mercado, no qual os negros tenderiam a assumir empregos de baixa remuneração.

Segundo Neal e Johnson (1996), ao menos um terço ou até a metade dos diferenciais de renda entre negros e brancos nos Estados Unidos se deve à discriminação. Os autores argumentam que os diferenciais de renda entre brancos e negros estão intimamente ligados ao *background* familiar, uma vez que estes afetam os custos do investimento na educação das crianças. Ademais, os mesmos autores afirmam que parte do hiato de renda está relacionado ao ambiente escolar a que está sujeita a criança. Em resumo, os autores concluem que a discriminação no mercado de trabalho americano está associada às desvantagens que os jovens negros têm em adquirir capital humano enquanto crianças. Para os autores, investimentos em capital humano, experiência e ocupação após a adolescência são escolhas endógenas que afetam o rendimento, mas são restringidas pelo nível inicial de capital humano.

De modo geral para os estudos de diferenciais da renda e seleção no mercado de trabalho entre raças e sobre gênero, Lovell (1994) afirma que os diferenciais de renda e prestígio na ocupação entre raça e gênero, para o Brasil, aparecem a partir de cinco anos ou mais de escolaridade. Relacionado a isso, Hoffmann e Leone (2004) destacam que as ocupações menos valorizadas e tradicionalmente femininas do mercado de trabalho continuam se reproduzindo, implicando persistência de nichos ocupacionais, como o emprego doméstico. Complementarmente, Hasenbalg (1992) afirma que a raça e o sexo marcam a oportunidade de o indivíduo se inserir no mercado de trabalho e que as mulheres são canalizadas para algumas atividades econômicas de menor remuneração e responsabilidade. Por outro lado, os negros, de modo geral, apresentam uma gama maior de atividades, porém com empregos de baixa remuneração, baixa qualificação e reduzido prestígio social.

1.2.3. Mercado de Trabalho no Meio Rural

Todas as características discutidas anteriormente apresentam dinâmicas distintas quando consideradas dentro do meio rural ou urbano brasileiro. Sendo assim, no capítulo 4

seção 4.2, o trabalho terá como foco principal analisar o mercado de trabalho rural brasileiro. No presente estudo, a análise do mercado de trabalho rural foi realizada utilizando como objeto de estudo os indivíduos residentes em áreas rurais⁶, sendo o meio rural compreendido, de acordo com a metodologia da PNAD (2007), pelos indivíduos residentes em áreas que não abrangem as cidades (sedes municipais), as vilas (sedes distritais) ou as áreas urbanas isoladas. A situação rural então abrange toda a área situada fora desses limites. Neste mesmo sentido, segundo IBGE (2009), a tipologia para a diferenciação de espaço rural e agricultura é como se segue: “As diferenças entre agrícola e rural podem ser bem definidas, onde o primeiro termo refere-se a um setor da atividade econômica, enquanto a noção de rural é territorial e espacial...”.

A escolha deste corte se justifica devido às profundas diferenças relacionadas aos dois meios, como, por exemplo, as condições de vida no meio rural, com famílias mais numerosas, condições de moradia mais precárias, considerável importância das transferências governamentais para composição da renda familiar (DINIZ *et al.*, 2008), além da estrutura do mercado de trabalho rural. Nesta estrutura, mesmo com a modernização nas relações de trabalho e profissionalização das atividades ocorridas, principalmente a partir da década de 90, ainda se encontram, em grande proporção, atividades como as de pequeno produtor, meeiro, trabalho sem remuneração, agricultura familiar, entre outras, que são distintas das atividades assalariadas comumente encontradas no meio urbano (SOUZA e MACHADO, 2004).

Além disso, as particularidades de cada meio afetam a composição do mercado de trabalho, onde se apontam os maiores salários do meio urbano, a estrutura organizacional do trabalho mais aprimorada, além do acesso mais adequado ao sistema judiciário e à maior concorrência no mercado de trabalho urbano frente ao rural.

Por fim, em termos da população residente no meio rural, Lima (2002) verifica que na década de 90 ocorre crescimento significativo dessa população à taxa de 0,90% a.a. com incremento total, entre 1992 e 1999, de aproximadamente 1,4 milhões de pessoas, concomitantemente a uma redução de 1,70% a.a. das atividades agrícolas. Isso significa dizer que as famílias residentes no meio rural têm apresentado uma tendência de diversificação nas atividades exercidas, com ênfase em atividades não agrícolas. Todos os fatores apresentados caracterizam o meio rural como uma complexa e intrincada fonte de estudos que necessita ser mais explorada.

⁶ Estes indivíduos não necessariamente se dedicam a atividades rurais clássicas como a agricultura.

Conforme evidenciado anteriormente, o mercado de trabalho é uma importante fonte de geração de renda dos indivíduos. Ademais, o mesmo pode ser um fator que amplifica as desigualdades de rendimento em uma economia ao, por exemplo, promover a seleção de indivíduos, não por características que tenham real efeito sobre seu desempenho, mas sim por questões discriminatórias, como o gênero ou a raça, ou de segmentação. As questões levantadas acima consistirão o cerne deste estudo. Sendo assim, a hipótese central da pesquisa é a de que, no mercado de trabalho brasileiro contemporâneo, representado aqui pelos os anos de 2002 e 2007, ainda existem traços de segmentação e discriminação, tanto de forma geral como, especificamente, no meio rural.

Pelo exposto, a presente pesquisa tem o intuito de contribuir com o tema de mercado de trabalho e sua dinâmica, no que concerne à entrada do indivíduo no mercado, assim como os determinantes da renda dos que já se encontram nele. Além disso, o trabalho foca o estudo dos diferenciais de gênero e raça e também os efeitos na renda e na probabilidade de entrada no mercado de mães e pais solteiros, de homens e mulheres casadas e os efeitos do número de crianças na família.

Por outro lado, esta pesquisa pretende contribuir também com a forma de manipulação da base de dados, ao criar variáveis⁷ que não fazem parte do questionário oficial da PNAD, suprimindo algumas lacunas da pesquisa. Por exemplo, boa parte dos estudos sobre mercado de trabalho utiliza uma variável *dummy* para analisar a presença de criança na família, por ser a única presente na PNAD relacionada ao tema, sendo que o mais adequado seria considerar o número de crianças em cada família, uma vez que quanto maior o número de crianças na família, maior sua vulnerabilidade⁸ e, em consequência, maior a necessidade de os pais estarem no mercado de trabalho. Em outras palavras, a utilização da variável *dummy* subestimaria, ao menos nesse caso, o efeito da criança sobre a entrada no mercado de trabalho dos pais, bem como sobre a renda do trabalho. Neste contexto, desenvolveu-se a variável número de crianças na família.

⁷ As variáveis foram criadas utilizando-se técnicas de construção de algoritmos, tendo como suporte para o desenvolvimento outras variáveis presentes na pesquisa. No Apêndice é apresentado um exemplo do algoritmo criado para este estudo.

⁸ O sentido da palavra vulnerabilidade neste estudo está relacionado às maiores dificuldades que, por exemplo, uma família com muitas crianças pequenas enfrenta no mercado ou, por outro lado, uma família em que o chefe é uma mãe solteira. Em outras palavras, refere-se à capacidade de certas particularidades do indivíduo ou do meio em serem fatores de inclusão ou exclusão social, seja na forma de seleção no mercado de trabalho, ou por meio da renda.

Outro exemplo é a criação das variáveis mãe e pai solteiros, ou seja, famílias com crianças pequenas⁹ em que não existe a presença de cônjuge. Esta variável possibilita abordar um tema de grande relevância quando se considera o mercado de trabalho, principalmente para o caso de mães solteiras, que é a entrada no mercado de trabalho com salários abaixo do que se obteria em outra situação. Infere-se que, devido à necessidade de prover o sustento dos filhos pequenos, a entrada no mercado da mãe solteira, nestas condições, tende a ter um efeito perpetuador do ciclo de pobreza. Em outras palavras, o baixo salário obtido pela mãe pode não ser suficiente para propiciar, por exemplo, um bom nível educacional para os filhos, ou até mesmo nutricional dessas crianças, o que acelera a entrada dessas crianças no mercado, novamente com baixos salários e baixa qualificação.

Por fim, o desenvolvimento das rotinas computacionais na presente pesquisa possibilitou que as análises pudessem ser realizadas em termos de famílias ao invés apenas de domicílios ou indivíduos, que são uma unidade de grande importância para a implementação de políticas públicas.

1.3. Objetivos

O objetivo geral do presente estudo consistiu em analisar o comportamento do mercado de trabalho brasileiro sobre o foco tanto da entrada no mercado quanto do rendimento, utilizando variáveis relacionadas ao gênero, raça, região geográfica, vulnerabilidade, entre outras. Adicionalmente, foram utilizados indicadores de concentração de renda e pobreza como forma de subsidiar as análises das equações de rendimento e participação para o caso da população como um todo. O estudo foi dividido em uma análise para o mercado brasileiro como um todo e outra para o mercado de trabalho rural. Em ambos, a análise foi dividida para homens e mulheres, com o intuito de se identificar mais eficientemente os efeitos ou impactos das variáveis consideradas sobre os grupos de homens e mulheres separadamente. O período de análise são os anos de 2002 e 2007, o que possibilita uma análise atual do mercado de trabalho brasileiro.

O objetivo geral será alcançado por meio dos seguintes objetivos específicos:

⁹ Entende-se por criança pequena os indivíduos de 0 a 14 anos.

- (i) Determinar os fatores que influenciam a participação de homens e mulheres no mercado de trabalho brasileiro como um todo e verificar os determinantes do rendimento do trabalho desses indivíduos no mercado de trabalho brasileiro;
- (ii) Determinar os fatores que influenciam a participação feminina e masculina no mercado de trabalho rural e verificar os determinantes do rendimento do trabalho para homens e mulheres no mercado de trabalho rural;
- (iii) Analisar as alterações na concentração da renda e na pobreza no período entre os anos de 2002 e 2007, desagregando a análise para os grupos de interesse, como gênero, raça, região geográfica.

Além das considerações iniciais e das considerações finais, este trabalho foi subdividido em mais três capítulos: o primeiro consiste numa análise dos modelos teóricos utilizados, seguido da metodologia empregada e finalizando tem-se a discussão dos principais resultados.

2. REFERENCIAL TEÓRICO

2.1. Decisão de participação no mercado de trabalho

Esta seção descreve os fundamentos teóricos que corroboram a participação do indivíduo no mercado de trabalho. A base teórica parte de um modelo neoclássico estático de renda-lazer, ou de oferta de trabalho adaptado de Varian (1992), Killingsworth (1983) e Bloom e Northrup (1973), em que a decisão de oferta de trabalho de um indivíduo é determinada por uma função de utilidade descrita da seguinte forma:

$$\text{Max } U(G, L) \quad (2.1)$$

Sujeito a:

$$P_G G = WH + V \quad (2.2)$$

$$T = H + L \text{ ou } H = T - L \quad (2.3)$$

substituindo 2.3 em 2.2 tem-se:

$$P_G G = W(T - L) + V$$

ou seja, o objetivo do agente é maximizar a utilidade (U), sendo argumentos desta função (U) as quantidades de bens (G) e horas de lazer (L), em que ambos são bens e sendo assim o indivíduo sempre prefere mais a menos. A maximização está sujeita a uma restrição orçamentária, cujo total de bens adquiridos $P_G G$ deve ser igual à renda não-oriunda do

trabalho (V)¹⁰, e a renda do trabalho determinada por (WH), em que (W) é a taxa salarial do indivíduo e (H) são as horas destinadas ao trabalho. A maximização está sujeita ainda a uma restrição de tempo, em que o número de horas disponíveis do indivíduo é dado por (T), o número total de horas trabalhadas (H) é definido como $H = T - L$ e (L) significa o total de horas dispensadas com lazer. Nesse sentido, o indivíduo deve alocar seu tempo disponível entre horas trabalhadas, que lhe permitam adquirir os bens (G), e horas de lazer¹¹.

Destarte, podem ser desenvolvidas graficamente as argumentações anteriores:

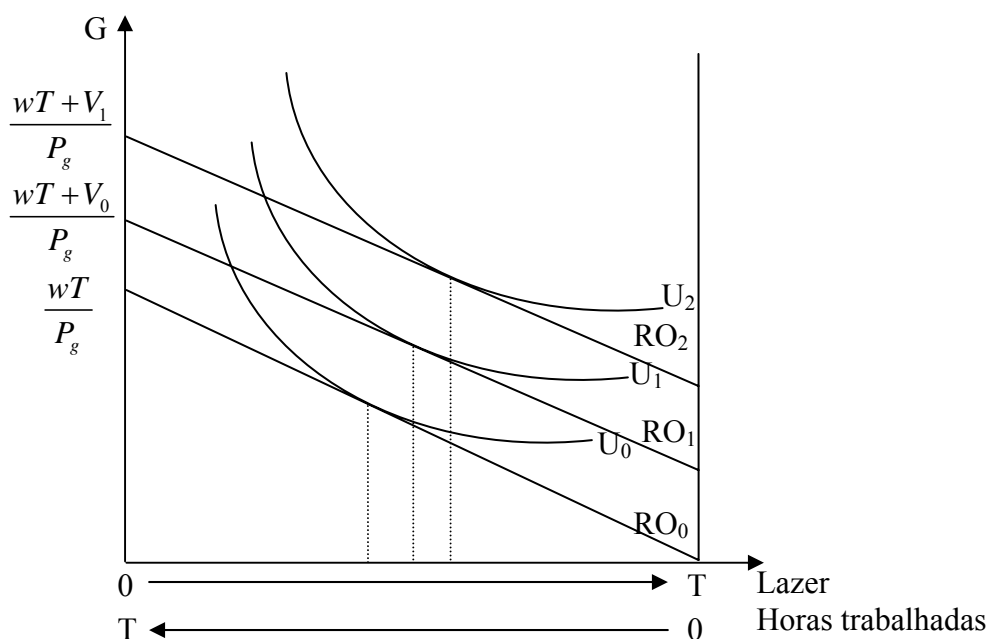


Figura 2.1. Alocação do tempo do indivíduo entre trabalho e lazer.

De acordo com a Figura 2.1, a inclinação da restrição orçamentária (RO_i) é dada por $\frac{dG}{dL} = -\frac{W}{P_g}$, em que $\frac{W}{P_g}$ representa o salário real. A curva de restrição orçamentária apresentará deslocamentos caso a renda não proveniente do trabalho se altere, podendo essa característica ser verificada nas curvas RO_0 , RO_1 e RO_2 , com $RO_2 > RO_1 > RO_0$.

O nível máximo de utilidade será alcançado no ponto de tangência da curva de indiferença U_i com a restrição orçamentária RO_i . Neste ponto, as inclinações de ambas as

¹⁰ Renda que pode advir de alguma transferência governamental como o Programa Bolsa Família, ou de alguma forma de investimento como, os rendimentos de juros da poupança.

¹¹ Considera-se a ausência de poupança, empréstimos ou pagamentos de impostos neste modelo.

curvas são iguais, ou seja, resolvendo o problema de maximização condicionada, conforme Killingsworth (1983) obtém-se¹²:

$$TMS_{G,L} = \frac{dG}{dL} = \frac{Um_g L}{UM_g G} = \frac{W}{P_g}, \text{ com } G \text{ e } L > 0 \text{ ou } H \text{ e } L > 0 \quad (2.4)$$

A utilidade marginal do lazer por unidade monetária gasta com lazer deve ser igual à utilidade marginal gerada pelos bens consumidos sobre o preço desses bens. Em outras palavras, a taxa marginal de substituição (TMS) de quantidades de bens (G) por horas de lazer (L) é igual aos preços relativos. Desta forma, quando a $TMS_{G,L}$, que representa o salário reserva do indivíduo, for igual ao salário real oferecido no mercado $\left(\frac{W}{P_g}\right)$, o indivíduo faz a escolha do montante de horas destinadas ao trabalho e ao lazer. É importante destacar que, quanto mais elevado o salário de mercado, $\left(\frac{W}{P_g}\right)$, mais horas serão destinadas ao trabalho em sacrifício do lazer. Por outro lado, quanto menor o salário de mercado, tanto menor serão as horas dedicadas ao trabalho.

A decisão do indivíduo de entrar no mercado de trabalho parte, no entanto, de uma solução de canto, em que $L = T$ e $H = 0$, o que quer dizer que o agente está disposto a ofertar zero horas de trabalho. Pressupõe-se, então, que a satisfação de uma hora a mais de lazer é maior do que o preço relativo do lazer e, sendo assim, o agente não oferta trabalho, ou seja, não participa da PEA¹³. Pode-se afirmar, então, que, no caso da solução de canto, o salário reserva do indivíduo é maior do que o ofertado no mercado, sendo este a $TMS_{G/L}$.

Então, $TMS_{G,L} > \frac{W}{P_g}$.

¹² Considerando que a condição de segunda ordem é satisfeita.

¹³ População Economicamente Ativa.

Graficamente:

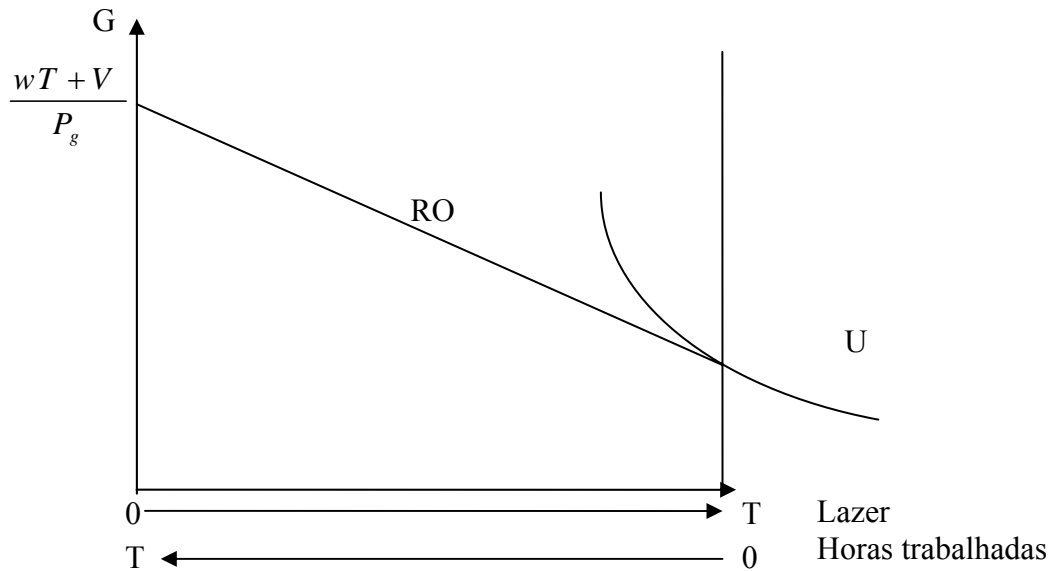


Figura 2.2. Decisão de oferta de trabalho a partir de uma solução de canto.

A regra de decisão de participação no mercado será: o agente participa da força de trabalho se o salário oferecido no mercado superar seu salário reserva.

Adicionalmente pode-se utilizar a equação de Slutsky para obter algumas conclusões adicionais a respeito da oferta individual de trabalho, de acordo com Varian (1992) a equação pode ser definida como¹⁴:

$$\frac{dL}{dW} = \frac{\partial L}{\partial W} + \frac{\partial L}{\partial V} [T - L] \quad (2.5)$$

Considerando que o lazer (L) é um bem normal, é possível dividir o efeito total, ou a demanda por lazer $\left(\frac{dL}{dW}\right)$ em um efeito substituição negativo $\left(\frac{\partial L}{\partial W}\right)$ uma vez que incrementos no salário (W) tornam o lazer mais caro. E o efeito renda positivo $\left(\frac{\partial L}{\partial V}\right)$ porque ganhos de renda provocam aumento na demanda por lazer, por último $[T - L]$ é um

¹⁴ A demonstração completa da equação de Slutsky pode ser obtida em Varian (1992) pág. 145-146.

número positivo por definição. Mas como o interesse é determinar a oferta de trabalho pode-se reescrever a equação (2.5) em termos de horas trabalhadas, ou seja¹⁵:

$$\frac{dH}{dW} = \frac{\partial H}{\partial W} - \frac{\partial H}{\partial V} H \quad (2.6)$$

Da equação (2.6) é possível retirar algumas conclusões importantes. Primeiro, para um indivíduo que maximiza a utilidade aumentos salariais apresentam um efeito direto na oferta de trabalho, e um efeito compensatório no sentido inverso, devido ao aumento de renda. Ou seja, o aumento na taxa salarial tende a aumentar a oferta de trabalho do indivíduo, uma vez que o lazer tornou-se mais caro, isso porque com mais horas trabalhadas pode-se consumir mais, Bloom e Northrup (1973). Concomitantemente, ocorre redução na oferta de trabalho porque a elevação na taxa salarial aumenta a riqueza do indivíduo e conseqüentemente impacta positivamente na demanda por lazer. Segundo, a partir da equação de Slutsky em (2.6) é possível determinar o salário reserva do indivíduo, considerando que se $H=0$ o efeito renda desaparece e apenas variações na taxa salarial afetarão a oferta de trabalho, sendo assim pode-se determinar o salário reserva diagramaticamente.

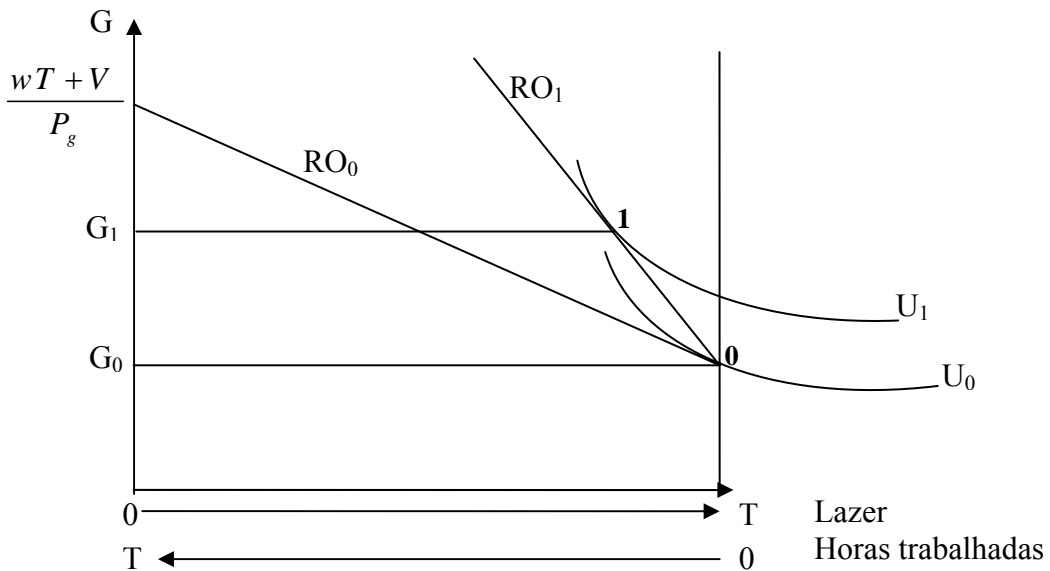


Figura 2.3. Determinando o salário reserva a partir da equação de Slutsky.

¹⁵ Neste caso o efeito substituição será sempre positivo, e o efeito renda será negativo no caso de horas trabalhadas ser um bem inferior.

A baixos níveis salariais o indivíduo é indiferente entre trabalhar ou não o que é demonstrado pelo ponto “0” da Figura 2.3, essa área é representada pelo retângulo (OG₀OT) e é conhecida como salário reserva. Por definição a restrição orçamentária RO₀ é tangente a curva U₀ quanto H=0. Realizando uma rotação na restrição orçamentária, ou aumentando os salários, se verifica que RO₀ é representada agora por RO₁ e o indivíduo passa a ofertar horas de trabalho, com o equilíbrio se encontrando no ponto “1”, deve-se lembrar que como H=0, ignora-se a compensação do efeito renda. Então conclui-se que, a partir da equação de Slutsky em 2.6 e pela Figura 2.3, a todos os salários abaixo do salário reserva o indivíduo prefere não trabalhar e a todos os salários que estão acima do salário reserva, este mesmo indivíduo decide ofertar horas de trabalho.

Assim, como destacam Scorzafave e Menezes Filho (2001), o salário reserva exerce papel fundamental na determinação da entrada ou não no mercado de trabalho.

No âmbito da teoria neoclássica, diversos fatores podem afetar a oferta de trabalho. Como exemplo, indivíduos com crianças em casa tendem a ter um salário reserva menor que os demais, uma vez que as necessidades dos filhos pequenos são maiores que as dos demais membros da família¹⁶. Por outro lado, no caso da mulher, espera-se que o salário reserva tenha correlação positiva com o número de filhos. Porém se considerar as mães solteiras, espera-se que seu salário reserva seja menor, quanto maior for o número de filhos.

É importante destacar também que outros fatores além dos já apontados podem afetar a oferta de trabalho dos indivíduos como por exemplo, de acordo com Bloom e Northrup (1973), força física do trabalhador, disposição para o trabalho, condições de trabalho, proximidade ao local de trabalho, ambiente de trabalho agradável e outros.

2.2. Determinantes do rendimento do trabalho

Para a análise dos determinantes dos rendimentos, utiliza-se a teoria do capital humano, baseada na abordagem de Mincer (1974). Esta visão parte do pressuposto de que os rendimentos individuais em qualquer período correspondem ao retorno em nível de suas habilidades (estoque de capital humano incorporado e acumulado pelo indivíduo ao longo do tempo). Em termos de estimativa econométrica, a forma funcional proposta por Jacob

¹⁶ Pode-se considerar que a mesma análise é válida também para o caso da presença de idosos na família. No entanto, esta hipótese é um juízo de valor.

Mincer para estimar a taxa de retorno da instrução e experiência, de acordo com Chiswick e Mincer (2003), pode ser especificada como:

$$\ln Y_i = \ln Y_0 + \beta_1 S_i + \beta_2 X_i + \beta_3 X_i^2 + \beta_4 S_i X_i + u_i \quad (2.7)$$

$i = 1, 2, \dots, n$

em que:

$\ln Y_0$ = logaritmo natural dos rendimentos do indivíduo sem instrução;

$\ln Y_i$ = logaritmo natural dos rendimentos do indivíduo i ;

S_i = anos de escolarização ou instrução formal do indivíduo i ;

X_i = anos de experiência no mercado de trabalho do indivíduo i ; e

u_i = termo de erro aleatório.

No presente estudo, esta abordagem foi ampliada no que pode ser definido como um modelo híbrido da teoria do capital humano, por utilizar outras variáveis além das consideradas na equação 2.7, o que será descrito no próximo capítulo.

3. METODOLOGIA

3.1. Medidas de Desigualdade

3.1.1. Índice de Gini

Os indicadores de desigualdade e de pobreza foram utilizados nesta pesquisa com intuito de auxiliar na exemplificação da estrutura sócio-econômica a que estão inseridos os indivíduos da amostra analisada no Brasil nos anos de 2002 e 2007. Ou seja, as seções que tratam da concentração da renda e pobreza apresentarão uma visão mais geral dos problemas que enfrentam, em termos da distribuição da renda, os indivíduos tratados nas análises dos determinantes de inserção e renda no mercado de trabalho.

Para analisar a distribuição da renda, utilizou-se primeiramente o índice de Gini. Segundo Hoffmann (1998), uma de suas principais vantagens é a possibilidade de sua associação com a curva de Lorenz, que representa a desigualdade de uma distribuição, além de obedecer à condição de Pigou-Dalton¹. O índice varia de zero a um. Quando uma única pessoa se apropria de toda a renda, o índice assume valor unitário, representando a concentração máxima da renda em uma dada população. Quando a renda é perfeitamente distribuída entre a população, o índice assume valor zero. O índice de Gini pode ser representado pela seguinte fórmula:

¹ Essa condição estabelece que o valor de uma medida de desigualdade ou de pobreza deve aumentar quando for feita uma transferência regressiva de renda, ou seja, transferência de renda de um indivíduo mais pobre para outro menos pobre, ocorrendo o inverso para transferências progressivas de renda.

$$Gini = \frac{2 \sum_{i=1}^N ir_i}{N \sum_{i=1}^N r_i} - 1 - \frac{1}{N} \quad (3.1)$$

em que N é o número de indivíduos na distribuição, i é o número de ordem do indivíduo após a ordenação pelo valor do rendimento, e r_i é o valor do rendimento da i -ésima pessoa.

3.1.2. Índices de Theil

Os índices são compostos por duas medidas: a de redundância, denominada de T de Theil e a de L de Theil. A medida de redundância (T de Theil) é mais sensível a alterações na desigualdade dentro de grupos de rendimentos mais elevados, enquanto o L de Theil é mais sensível a alterações na desigualdade dentro dos grupos de renda mais baixa, segundo Hoffmann (1998). O T de Theil pode ser representado da seguinte forma:

$$T = \sum_i y_i \log \frac{x_i}{\mu} \quad (3.2)$$

em que y_i é a fração da renda apropriada pelo i -ésimo indivíduo na distribuição, i é o número de ordem do indivíduo após a ordenação pelo valor do rendimento, x_i é a renda da i -ésima pessoa, com $i = 1, \dots, n$, e seja μ a renda média. Segue abaixo a representação do L de Theil:

$$L = -\frac{1}{n} \sum_i \log \frac{x_i}{\mu} \quad (3.3)$$

em que n é o número de pessoas consideradas na análise, i é o número de ordem do indivíduo após a ordenação pelo valor do rendimento, x_i é a renda da i -ésima pessoa, com $i = 1, \dots, n$, e seja μ a renda média.

Essa característica dos indicadores permite analisar a contribuição dos segmentos da sociedade no comportamento da distribuição da renda.

3.2. Medidas de pobreza

3.2.1. Índices FGT

Um novo marco no desenvolvimento das medidas de pobreza é o trabalho de Foster, Greer e Thorbecke (1984), que apresenta três etapas para o cálculo dessas medidas. A primeira etapa consiste em fixar o valor monetário das linhas de pobreza (z_p), cuja determinação em estudos que utilizam a variável renda pode se dar de duas maneiras. Na primeira, utiliza-se o salário mínimo como referência para fixar a linha de pobreza; e na segunda, a linha é determinada exogenamente em função do nível de consumo das famílias, por meio de pesquisas de orçamento familiar como em Rocha (1997). Na segunda etapa das medidas de pobreza, deve-se, a partir da linha de pobreza, dividir os indivíduos em pobres e não pobres. Como último passo, agrega-se a distância dos pobres à linha de pobreza, de forma a se dar mais ou menos peso aos indivíduos relativamente mais pobres da população. Os índices absolutos de pobreza de Foster, Greer e Thorbecke (FGT) podem ser calculados utilizando o grau de aversão à pobreza igual a 0, 1 e 2, respectivamente. A fórmula a seguir representa esse índice:

$$FGT^\alpha = \frac{1}{n} \sum_{i=1}^q \left(\frac{z_p - y_i}{z_p} \right)^\alpha \quad (3.4)$$

em que n é o número de indivíduos na população; q é o número de indivíduos abaixo das linhas de pobreza (z_p); y_i é a renda do indivíduo i ; e, por fim, α é o grau de aversão à pobreza. Se α for igual a zero, o índice passa a representar apenas a extensão da pobreza, não captando a intensidade, como demonstra a equação a seguir:

$$FGT^0 = \frac{q}{n} \quad (3.5)$$

no caso de α igual a 1, o índice passa a incorporar a medida de extensão FGT^0 , que é a intensidade da pobreza. O índice FGT^1 , apresentado a seguir, conhecido como hiato da

pobreza, corresponde ao valor médio da distância das rendas dos pobres em relação à linha de pobreza.

$$FGT^1 = \frac{1}{n} \sum_{i=1}^q \left(\frac{z_p - y_i}{z_p} \right)^1 \quad (3.6)$$

O índice FGT^1 não é capaz de considerar os efeitos na mudança da distribuição da renda entre os pobres. Para resolver esse problema, basta assumir α igual a 2, ou seja, atribuir mais peso para indivíduos com menor renda. O índice FGT^2 , descrito a seguir, corresponde à distância média quadrática dos pobres da linha de pobreza.

$$FGT^2 = \frac{1}{n} \sum_{i=1}^q \left(\frac{z_p - y_i}{z_p} \right)^2 \quad (3.7)$$

À medida que se passa de FGT^0 para FGT^2 , os índices atribuem maior peso aos indivíduos mais pobres da população. No índice FGT^0 , todos os indivíduos situados abaixo da linha de pobreza entram com pesos idênticos. Para os índices FGT^1 e FGT^2 , os indivíduos são ponderados de forma proporcional à distância e ao quadrado da distância de suas rendas em relação à linha de pobreza. No presente estudo, a linha de pobreza selecionada corresponde a meio salário mínimo (R\$ 190,00), de acordo com o valor de setembro de 2007 (PNAD, 2007)¹⁷.

Uma propriedade importante dos índices FGT^α é a possibilidade de decompô-los o que possibilita determinar a importância de cada grupo para determinar a pobreza total. Aqui, os índices de pobreza calculados para o Brasil (N), foram decompostos para os (k) grupos, segundo as macro-regiões do Brasil, Nordeste, Sul, Centro-Oeste, Norte e Sudeste; o sexo, em masculino e feminino, a cor ou raça em pardos, brancos ou negros; e de acordo com a situação censitária em urbano e rural. Retomando a simbologia utilizada nas equações acima, indica-se a renda da i -ésima pessoa do r -ésimo grupo com y_{ri} , com $i =$

¹⁷ Essa mesma linha é adotada para as análises do ano de 2002, uma vez que, os valores deste ano foram deflacionados de acordo com o INPC – restrito para valores compatíveis aos recebidos em setembro de 2007.

$1, \dots, n_r$ e $r = 1, \dots, k$. A participação do r -ésimo grupo na população total pode ser representada da seguinte forma:

$$\pi_r = \frac{n_r}{N} \quad (3.8)$$

Admitindo-se que a linha de pobreza (z) é a mesma para todos os grupos e que, dentro de cada grupo as rendas estão ordenadas de maneira que:

$$y_{r1} \leq y_{r2} \leq \dots \leq z \leq \dots \leq y_m \quad (3.9)$$

e considerando que na equação (3.10), r_q é o número de pobres dentro do grupo r . A partir desses dados, pode-se calcular o índice para cada grupo:

$$FGT_r^\alpha = \frac{1}{n_r} \sum_{i=1}^{r_q} \left(\frac{z - y_{ri}}{z} \right)^\alpha \quad (3.10)$$

Sendo assim, o índice para toda a população pode ser estimado pela seguinte equação:

$$FGT^\alpha = \frac{1}{N} \sum_{r=1}^k \sum_{i=1}^{r_q} \left(\frac{z - y_{ri}}{z} \right)^\alpha \quad (3.11)$$

Ou de outra forma:

$$FGT^\alpha = \sum_{r=1}^k \pi_r \cdot FGT_r^\alpha \quad (3.12)$$

A expressão (3.12) acima demonstra que, fixada a distribuição da em k grupos, alterações na renda das pessoas que causarem um aumento do índice FGT_r^α dentro de cada um ou mais grupos levarão a um aumento da medida de pobreza em toda a população.

3.3. Procedimento de Heckman

Como destacam Hoffmann e Kassouf (2005), o procedimento de Heckman tornou-se muito popular nas últimas décadas. Este autor desenvolveu um método relativamente simples para corrigir o possível problema de seletividade amostral¹⁸, consistindo o procedimento em estimar duas equações. A primeira delas determina a decisão do indivíduo em participar ou não do mercado de trabalho por meio de um modelo Probit em que se obtém a razão inversa de Mills. Essa equação é conhecida como equação de seleção e é definida a seguir.

Considerando L_i^* a diferença dos salários oferecidos no mercado e o de reserva do indivíduo i e sabendo que L_i^* é uma variável latente, ou seja, não observável, tem-se:

$$L_i^* = \alpha' Z_i + \mu_i, \quad (3.13)$$

em que Z é o vetor de variáveis exógenas que afetam a decisão de participar ou não da força de trabalho, sendo μ o termo de erro aleatório. Mesmo L_i^* sendo não observável, é possível verificar se determinado agente trabalha ou não, então a tomada de decisão é feita como segue:

$$\begin{aligned} L_i &= 1 & \text{se} & L_i^* > 0 \\ L_i &= 0 & \text{se} & L_i^* \leq 0 \end{aligned} \quad (3.14)$$

Se o indivíduo trabalha ($L_i=1$), o salário reserva desse agente é menor do que o oferecido no mercado. Em contrapartida, se o indivíduo não trabalha, seu salário de reserva se mostra maior do que o de mercado.

Considerando agora W o logaritmo dos salários, tem-se:

$$W_i = \beta' X_i + v_i \quad (3.15)$$

¹⁸ Maiores detalhes sobre o problema de seletividade amostral consultar: Heckman (1979), Hoffmann e Kassouf (2005), Resende e Wyllie (2006).

em que X representa o vetor de variáveis explicativas que afetam o nível de rendimento e v o termo de erro aleatório.

De acordo com Hoffmann e Kassouf (2005), considerando que μ_i e v_i têm uma distribuição normal bivariada com média zero e desvio padrão σ_μ e σ_v e coeficiente de correlação ρ e que W_i é observada apenas quando L_i for maior que zero, seu valor esperado será definido como:

$$E(W_i | L_i^* > 0) = \beta' X_i + \rho \sigma_v \lambda_i \quad (3.16)$$

em que λ é a razão inversa de Mills, dada por:

$$\lambda_i = \frac{\phi\left(\frac{\alpha' Z_i}{\sigma_u}\right)}{\Phi\left(\frac{\alpha' Z_i}{\sigma_u}\right)} \quad (3.17)$$

e ϕ e Φ são, respectivamente, as funções normais padrão de densidade de probabilidade e acumulada.

Fica claro, pelo apresentado, que a estimação direta da equação de rendimento (3.15) apenas com os indivíduos economicamente ativos geraria viés de especificação, conhecido como viés de seletividade. A inclusão da razão inversa de Mills na equação (3.16) leva em consideração explicitamente a decisão do indivíduo de participar ou não do mercado de trabalho, de tal forma que esta variável elimina o viés de seletividade e, por conseguinte, proporciona estimativas consistentes dos parâmetros da equação de rendimento.

Devido à característica de amostragem da base de dados adotada (PNAD)¹⁹, a estimação do modelo de seleção amostral foi realizada por Pseudo-Máxima Verossimilhança ao invés do Método de Dois Estágios. A estimação por Máxima Verossimilhança pode gerar estimadores inconsistentes caso a variância do erro seja

¹⁹ O plano amostral das PNADs não é uma amostragem aleatória simples, mas sim uma amostragem complexa.

heterocedástica, segundo Greene (2003). Assim, realizou-se o teste para cada variável explicativa do modelo.

Observa-se ainda, sobre o modelo de seleção amostral, que a equação de participação é na verdade um modelo *probit* que procura analisar os fatores que influenciam na probabilidade de um indivíduo se encontrar no mercado de trabalho. Destaca-se também que, para evitar problemas de multicolinearidade, as variáveis explicativas consideradas na equação de rendimentos são um subconjunto daquelas consideradas na equação de participação.

3.4. Descrição das Equações e Variáveis Seleccionadas

A equação de seleção foi definida como:

$$Z_i = \alpha_0 + \alpha_1 \text{renda_fa}_i + \alpha_2 \text{fam_comp}_i + \alpha_3 \text{crianc}_i + \alpha_4 D_{\text{gêneroi}} + \alpha_5 \text{IDAD}_i + \alpha_6 \text{IDAD}_i^2 + \alpha_7 D_{\text{cor}_i} + \alpha_8 \text{EDU}_i + \alpha_9 \text{ELEDU}_i + \alpha_{10} D_{\text{região}_i} + \alpha_{11} D_{\text{casado}_i} + \alpha_{12} D_{\text{pm_solteira}_i} + u_i \quad (3.18)$$

em que:

Z_i = Variável *dummy* concernente à condição de atividade do indivíduo i , ou seja, economicamente ativo (1) ou não economicamente ativo (0);

renda_fa_i = Representa o rendimento familiar mensal do indivíduo i ;

fam_comp_i = Número de componentes na família do indivíduo i ;

crianc_i = Consiste no número de crianças na família com idade de 0 a 14 anos;

$D_{\text{gêneroi}}$ = Variável *dummy* para gênero, em que a categoria de controle é o sexo feminino;

IDAD_i = Idade do indivíduo i em dezenas de anos (segundo metodologia proposta em Hoffmann e Simão, 2005);

IDAD_i^2 = Idade ao quadrado em dezenas de anos;

EDU_i = Anos de estudo do indivíduo i ;

ELEDU_i = Efeito Limiar dos anos de estudo do indivíduo i ²⁰;

²⁰ O modo de construção e definição dessa variável se encontra na seção 3.5 sobre efeito limiar da educação.

$D_{região_i}$ = Conjunto de variáveis *dummy* para descrever as regiões geográficas do Brasil, divididas em Centro-Oeste, Sudeste, Norte e Sul, sendo a região Nordeste o controle;

D_{cor_i} = Conjunto de variáveis *dummy* para descrever cor ou raça dos indivíduos, divididas em Preta e Indígena, Parda e Amarela. A cor Branca foi utilizada como controle. O agrupamento da cor ou raça foi realizado devido à baixa representatividade dos amarelos e indígenas na amostra;

D_{casado_i} = Variável *dummy* para indivíduos casados. A condição para determinar este tipo de união consistiu em considerar as famílias onde foram encontradas a pessoa de referência e o cônjuge;

$D_{pm_solteira_i}$ = Variável *dummy* para indivíduo que foi considerado pai ou mãe solteiro. Esta variável foi construída admitindo como condição a existência de criança de 0 a 14 anos na categoria de filho na família e a ausência de cônjuge na mesma, ou seja, as famílias selecionadas consistiam da pessoa de referência, filhos e os demais membros excetuando o cônjuge; e

μ_i = termo de erro aleatório.

A equação de rendimento foi especificada da seguinte forma:

$$\ln_renda_i = \beta_0 + \beta_1 D_{gêneroi} + \beta_2 IDAD_i + \beta_3 IDAD_i^2 + \beta_4 D_{cor_i} + \beta_5 EDU_i + \beta_6 ELEDU_i + \beta_7 D_{urbanoi} + \beta_8 D_{região_i} + \beta_9 D_{casado_i} + \beta_{10} D_{pm_solteira_i} + \beta_{11} \lambda_i + v_i \quad (3.19)$$

em que:

\ln_renda = logaritmo natural dos rendimentos de todos os trabalhos do indivíduo i sobre o total de horas trabalhadas do indivíduo i ;

λ_i = a razão inversa de Mills; e

v_i = termo de erro aleatório.

As demais variáveis seguem as mesmas especificações supramencionadas.

É necessário destacar alguns aspectos dos modelos e das variáveis selecionadas, quais sejam:

- i) Em toda análise econométrica realizada foi utilizado o fator de expansão da amostra, disponibilizado junto com as PNADs.
- ii) A utilização da variável condição de atividade pode gerar questionamentos a respeito desta escolha como *proxy* para oferta de trabalho, sendo assim, é cabível apresentar alguns argumentos. O objetivo de considerar a variável condição de atividade e não a condição na ocupação procura ser o mais coerente possível com o referencial teórico exposto. Ou seja, ao analisar os indivíduos economicamente ativos e não economicamente ativos, leva-se em consideração a opção do indivíduo de ofertar trabalho mesmo que este não esteja ocupado no momento da entrevista. Apesar de desocupado, deve-se ter em mente que esta condição é temporária e que o agente já exerceu sua escolha de ofertar horas de trabalho, e que tomou providências efetivas de procura (PNAD, 2007). Por outro lado, caso fosse considerada a variável condição de ocupação como *proxy* para a oferta de trabalho, estar-se-ia omitindo parte da amostra que decide ofertar trabalho, mas não encontra emprego.
- iii) A escolha da forma funcional log para a equação de rendimentos é devida à expectativa de que a renda não varie linearmente com as variáveis relacionadas à produtividade e à experiência. Autores como Mincer (1974 e 1993), Chiswick (2003), Hoffmann (1996, 1998), Ney e Hoffmann (2003), Teixeira (2006), dentre outros, apoiam a utilização desta forma funcional, mesmo quando foram incorporadas as variáveis discriminatórias. De qualquer maneira, a forma funcional adotada não apresentaria graves problemas, uma vez que a amostra utilizada é consideravelmente elevada, como será verificado adiante.
- iv) Para analisar os rendimentos, foram desconsideradas as pessoas que não declararam a renda. Na PNAD, a não declaração da renda é computada como R\$ 999.999.999.999²¹, ou seja, devem-se desconsiderar estes valores para evitar vieses nas análises.
- v) A variável idade da pessoa foi medida em dezenas de anos, e também o quadrado dessa variável, tendo em vista que Y não varia linearmente com a

²¹ IBGE (2007).

idade (NEY e HOFFMANN, 2003; HOFFMANN e SIMÃO, 2005). A idade é medida em dezenas de anos apenas para evitar que os coeficientes sejam muito pequenos. Esta duas variáveis foram incluídas no modelo com o intuito de se medir a experiência do indivíduo e o impacto sobre a entrada no mercado de trabalho e os rendimentos. Os resultados demonstrarão a idade média em que o trabalhador tende a receber o retorno máximo no ciclo de vida.

- vi) Quanto aos anos de estudo, foram desconsideradas as observações com valor igual a 17 (não determinados ou sem declaração) por motivos similares ao item (ii); os anos variaram de 1 a 16 anos de estudo, sendo este último os com 15 anos ou mais de estudo, PNAD (2006). Um fator restritivo da variável educação, para esta base de dados, é a impossibilidade de se determinar a qualidade do ensino. Como destacam Behrman e Birdsall (1983, *apud* RESENDE e WYLLIE, 2006), a desconsideração do fator qualidade tenderia a levar a uma superestimação do retorno para educação. Ademais, *a priori*, espera-se que existam diferenças em termos de ensino público e privado (neste caso dependendo se o mesmo é de nível fundamental, médio ou superior) e em termos regionais.
- vii) Foram excluídas da análise pessoas que não declararam a cor e as pessoas que não se enquadram como membros da família.

3.5. O Efeito Limiar da Educação

O efeito limiar da educação pode ser considerado como o impacto crescente dos anos de estudo sobre o rendimento a partir de determinado valor, ou seja, o valor da escolaridade a partir do qual a taxa de retorno da escolaridade torna-se maior (Hoffmann e Simão, 2005). O efeito limiar pode ser descrito como se segue:

$$S^* = D(S - L) \tag{3.20}$$

em que:

S^* = Efeito Limiar da Educação;

S = Anos de Estudo do indivíduo;

L = Limiar da educação, que no presente artigo foi determinado em 6 anos de estudo, seguindo metodologia adotada por Ney e Hoffmann (2003)²²; e

D = uma variável *dummy* que assume valor zero para $S \leq L$ e assume valor 1 para $S > L$.

Considerando que $S = EDU_i$ e $S^* = ELEDU_i$, e K as demais variáveis, e coeficientes, da equação de rendimento, o valor do logaritmo dos rendimentos de todos os trabalhos sobre o total de horas trabalhadas ficará como:

$$\ln_renda_i = K + \beta_5 EDU_i + \beta_6 ELEDU_i \quad (3.21)$$

Quando $S \leq L$, tem-se $D = 0$ e a equação se reduz a:

$$\ln Y_i = K + \beta_5 EDU_i \quad (3.22)$$

ou seja, desconsiderando o limiar da educação, cada ano adicional de estudo está associado a um aumento de $[\exp(\beta_5) - 1]100$ no rendimento dos indivíduos.

Quando $S > L$, tem-se $D = 1$, e a expressão (3.21) se tornará:

$$\ln Y_i = K - \beta_6 ELEDU + (\beta_5 + \beta_6) EDU \quad (3.23)$$

Sendo assim, após o nível de limiar, cada ano a mais de escolaridade provocará retorno nos rendimentos das pessoas em $[\exp(\beta_5 + \beta_6) - 1]100$.

²² O primeiro critério foi estimar as equações com diversos níveis de limiar e a partir daí verificar o ajustamento do modelo para cada um. Como todos os modelos testados, com exceção do modelo com 1 ano de estudo para o limiar, apresentaram bom ajustamento, partiu-se para o segundo critério de determinação, ou seja, estimar os modelos com *dummies* para os anos de estudo de 2 a 16 anos, tendo como controle os sem instrução ou com menos de um ano. A partir daí, pôde-se determinar o efeito marginal sobre o rendimento. A escolha é feita no nível onde se verifica incremento considerável da renda, ou seja, onde existe uma quebra na sequência de retornos. A Figura 3.1 apresenta detalhes dos resultados para o efeito limiar.

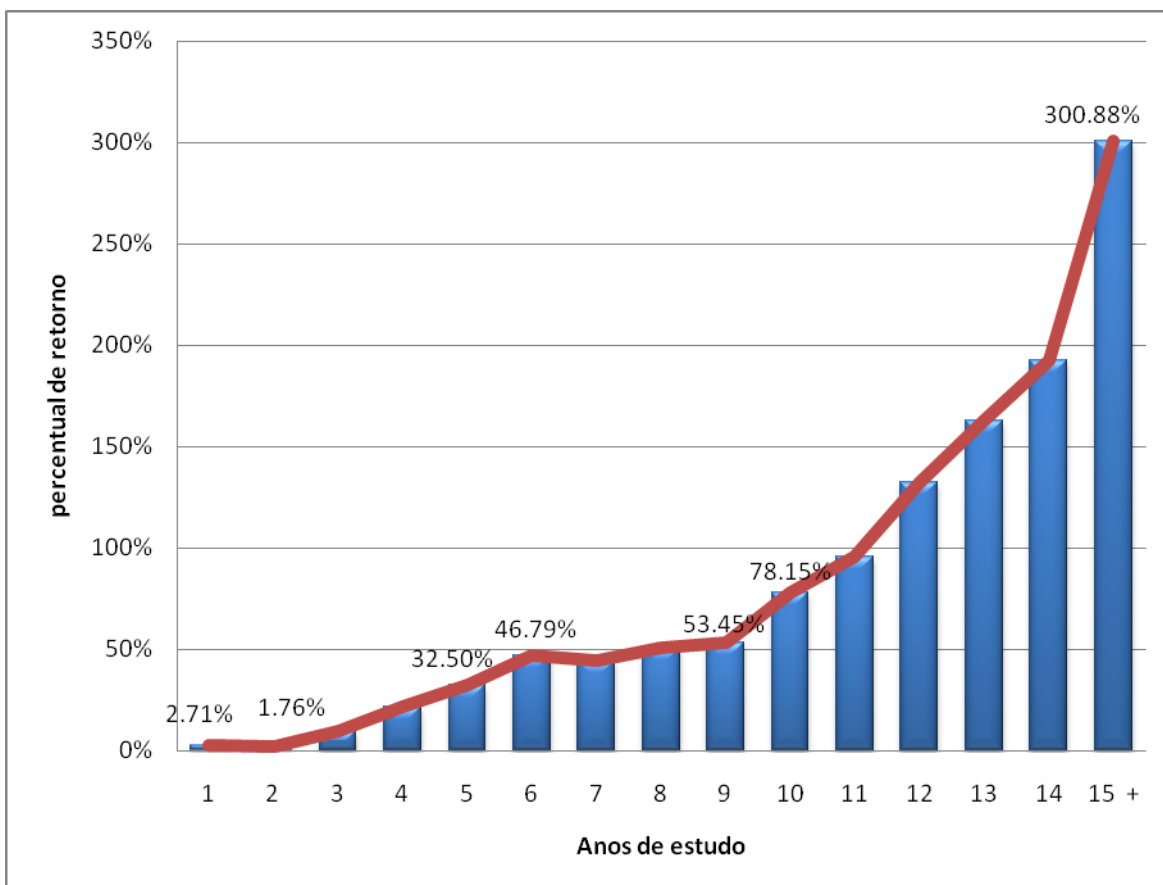


Figura 3.1. Retorno dos anos de estudo sobre a razão de log da renda e total de horas trabalhadas, sem instrução ou menos de um ano é o controle.²³

Fonte: Elaboração do autor a partir dos microdados da PNAD 2007.

²³ Gráfico idêntico foi desenvolvido para o ano de 2002 mas foi suprimido por apresentar padrão semelhante.

4. RESULTADOS E DISCUSSÃO

4.1. Desigualdade, Pobreza e Determinantes da Participação e Rendimentos no Mercado de Trabalho Brasileiro

A presente subseção tem como foco analisar a desigualdade, pobreza e alguns fatores que determinam a entrada do indivíduo no mercado de trabalho e gerar equações de rendimentos para o Brasil nos anos de 2002 e 2007, com o intuito de identificar os fatores que mais impactam na renda e analisar como estes fatores se alteraram na década, dando especial atenção aos resultados relacionados à raça e ao gênero. Adicionalmente, verifica-se se a desconsideração da PNAD como uma amostra complexa gera viés sobre os resultados. Segundo Resende e Wyllie (2006), boa parte “dos estudos econométricos acerca de retornos para educação ignoram o desenho amostral de pesquisas... como a PNAD”. Ratificando estes autores, Silva e Pessoa (2002) afirmam que análises de regressão com a PNAD “muitas vezes são feitas por analistas que trabalham fora da agência produtora dos dados e frequentemente utilizam para a modelagem pacotes estatísticos que se baseiam em hipóteses válidas somente quando os dados são obtidos por meio de Amostras Aleatórias Simples com Reposição (AASC)”.

4.1.1. Estatísticas Descritivas dos Dados da Amostra

As pessoas economicamente ativas no Brasil, entre 15 e 70 anos, representaram cerca de 71% da população considerada no ano de 2002, elevando-se em 2007 para 72%. Separando a análise de acordo com o sexo, verifica-se que a taxa de atividade entre os homens é consideravelmente maior do que entre as mulheres: cerca de 83% contra 60%.

Algumas características importantes podem ser retiradas da Tabela 4.1. Por exemplo, os indivíduos no mercado de trabalho apresentaram incremento considerável na renda média²⁴ num montante de 7,30% entre 2002 e 2007, enquanto a inflação, medida pelo INPC, passou do patamar de 12,53% a.a. em 2002 para 5,16% a.a. em 2007²⁵. Torna-se importante destacar que, apesar de a renda dos homens ser mais elevada do que a das mulheres, estas apresentaram ganhos médios no período maiores do que o sexo oposto. Por exemplo, em 2002 uma mulher recebia em média R\$ 608 por mês, enquanto a renda masculina era de R\$ 962; já em 2007, os homens recebiam R\$ 1.022 e as mulheres R\$ 675. Sendo assim, o ganho de renda para o sexo masculino foi em torno de 6,23%, enquanto para o feminino foi de 11%.

A renda familiar alcançou em 2007 o patamar de R\$ 1.881,63, com as famílias tendo em média quatro membros, dois destes membros crianças entre 0 e 14 anos de idade. Existe predominância de indivíduos residentes na região Sudeste e Nordeste. A cor ou raça predominante foi a branca com 54% da população em 2002, seguida dos pardos ou amarelos com 39,6% e pretas ou indígenas com 6,2%. Em 2007, a ordenação das raças permanece a mesma em termos da quantidade de representantes. Contudo, os brancos apresentam redução enquanto as demais têm aumento.

Na amostra considerada, menos da metade dos indivíduos eram homens, 48%, e este patamar permanece o mesmo em 2007. Porém a taxa de atividade entre o sexo masculino foi de 83% em 2002 e 2007, e para as mulheres esse percentual foi menor, com 57% em 2002 e próximo a 61% em 2007, mostrando que, apesar da menor taxa de atividade, as mulheres vêm aumentando sua inserção no mercado de trabalho. Deve-se considerar que estes resultados são influenciados, pelo menos em parte, pelos diferenciais de salário reserva entre homens e mulheres, uma vez que, por hipótese, espera-se que as mulheres com filhos pequenos e/ou com muitos filhos tenham um salário reserva maior que o dos homens na mesma condição. Conseqüentemente, o número de homens economicamente ativos tende a ser maior que o de mulheres.

Considerando agora os indivíduos em relações conjugais estáveis, ou seja, casados, os homens apresentaram um percentual maior de indivíduos nesta situação entre 2002 e

²⁴ As rendas para o período foram deflacionadas de acordo com o INPC para valores reais de setembro de 2007.

²⁵ Dados obtidos junto ao Banco de dados Sidra(Sistema IBGE de Recuperação Automática), disponível em: <<http://www.ibge.gov.br>>

2007, 78,6% no último ano contra 70,2% das mulheres, destacando-se a redução das taxas em ambos os grupos. É interessante considerar também o percentual de pais ou mães solteiras. A mulher é a principal representante, correspondendo a 10% do total nesta situação em 2002. Para 2007, existe um arrefecimento, e a taxa cai para 9% das mulheres. Por outro lado, o pai solteiro não ultrapassa 3% da população masculina.

Levando em consideração agora a jornada de trabalho de homens e mulheres, os resultados mostram que enquanto o homem tende a trabalhar em patamares próximos à jornada de 40 horas semanais, mais precisamente 35 horas em 2002 e 34,6 em 2007, as mulheres chegam a valores próximos a 20 horas, ou seja, metade da jornada cumprida pelo homem. Deve-se considerar, entretanto que, apesar de o número de horas trabalhadas pelos homens, entre 2002 e 2007, ser maior do que o das mulheres, eles tiveram redução na jornada em 2,14%, e elas apresentaram elevação em 5%.

Quanto às variáveis que representam o capital humano, verifica-se que a idade média na amostra esteve próxima dos 36 anos. Outrossim, os dados mostram um reduzido nível de qualificação, sendo os anos de estudo médio em torno de 7,7 anos, ou seja, ensino fundamental incompleto. Ademais, realizando o corte por gênero, verifica-se que as mulheres apresentam em média nível de qualificação mais elevado do que os homens: 8,3 anos dos homens em 2007 contra 8,7 anos de estudo para as mulheres. Sendo assim, se os critérios relacionados à capacitação do indivíduo, em termos de anos de estudo, são uma boa *proxy* para a determinação do nível de renda, então, esperava-se que os rendimentos para o sexo feminino fossem superiores ou muito próximos aos obtidos pelos homens, o que não foi verificado pelos resultados anteriores.

Continuando a análise comparativa entre os rendimentos de homens e mulheres, o seguinte artifício ajudará a obter resultados mais apurados sobre os diferenciais de renda destes dois grupos, ou seja, ao invés de se comparar apenas os rendimentos brutos recebidos, pode-se utilizar a razão de rendimentos pelo total de horas trabalhadas. No caso dos homens, a média de reais recebidos por horas trabalhadas é de R\$ 7,37/h em 2007; para as mulheres o valor é de R\$ 8,53/h, mostrando que o rendimento recebido por elas em termos de horas trabalhadas é consistente com o tempo investido em qualificação. No entanto, as equações de rendimento poderão aprimorar ainda mais essas análises, uma vez que será possível, por exemplo, identificar como a decisão de investimento em qualificação formal (ou seja, anos de estudo) impactará nos retornos monetários de homens e mulheres.

Tabela 4.1. Estatísticas descritivas das variáveis de interesse, Brasil 2002-2007*.

Variáveis	TODOS				HOMENS				MULHERES			
	2002		2007		2002		2007		2002		2007	
	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.
Economicamente ativos	0.706	0.002	0.717	0.002	0.837	0.002	0.832	0.001	0.583	0.002	0.609	0.002
Renda de Todos os Trabalhos	815.00	8.96	874.37	9.35	962.70	10.97	1,022.81	11.53	608.83	7.93	675.12	7.83
Log da Renda de Todos os Trabalhos	5.894	0.007	6.372	0.006	6.024	0.007	6.500	0.007	5.691	0.008	6.185	0.007
Pai ou mãe solteiros	-	-	-	-	0.031	0.001	0.030	0.001	0.100	0.001	0.092	0.001
Situação conjugal (casado)	-	-	-	-	0.805	0.002	0.786	0.002	0.715	0.002	0.702	0.002
Rendimento Familiar	1,738.52	18.63	1,881.63	19.83	1,783.05	19.59	1,929.39	20.65	1,696.83	18.70	1,836.85	19.97
Número de horas todos os trabalhos (semana)	26.84	0.08	26.98	0.07	35.42	0.11	34.66	0.10	18.84	0.09	19.79	0.08
Número de Componentes na Fam.	3.855	0.008	3.662	0.008	3.948	0.009	3.725	0.009	3.767	0.008	3.602	0.008
Núm. crianças 0 a 14 anos na família	1.771	0.007	1.678	0.007	1.793	0.007	1.692	0.007	1.752	0.006	1.666	0.006
Proporção de famílias com crianças	0.487	0.002	0.434	0.002	0.475	0.002	0.422	0.002	0.499	0.002	0.445	0.002
Gênero (Masculino)	0.483	0.001	0.483	0.001	-	-	-	-	-	-	-	-
Idade (em dezenas de anos)	3.58	0.00	3.69	0.00	3.55	0.00	3.65	0.00	3.62	0.00	3.73	0.00
Cor preta ou indígena	0.062	0.001	0.084	0.001	0.065	0.001	0.086	0.002	0.060	0.001	0.082	0.001
Cor amarela ou pardas	0.396	0.003	0.415	0.003	0.406	0.003	0.426	0.003	0.387	0.003	0.404	0.003
Cor branca	0.542	0.003	0.502	0.003	0.529	0.003	0.488	0.003	0.553	0.003	0.514	0.003
Educação em anos de estudo	7.752	0.028	8.517	0.029	7.572	0.031	8.305	0.033	7.920	0.028	8.715	0.028
Região Centro-Oeste	0.072	0.001	0.072	0.001	0.072	0.001	0.073	0.001	0.071	0.001	0.071	0.001
Região Sudeste	0.449	0.002	0.439	0.003	0.446	0.002	0.435	0.003	0.452	0.002	0.442	0.003
Região Norte	0.055	0.001	0.075	0.003	0.055	0.001	0.076	0.003	0.055	0.001	0.074	0.003
Região Sul	0.154	0.001	0.150	0.002	0.155	0.001	0.151	0.002	0.153	0.001	0.149	0.002
Região Nordeste	0.270	0.002	0.264	0.003	0.271	0.002	0.264	0.003	0.269	0.002	0.264	0.003

Fonte: Elaboração do autor baseada nos microdados da PNAD 2002-2007.

* Em que Prop. significa Proporção e Desv. Pad. significa Desvio Padrão.

4.1.1. Análise da Desigualdade da Renda no Mercado de Trabalho Brasileiro entre 2002 e 2007

Muitos autores constataram que a desigualdade de renda vem apresentando contínua redução nos anos 2000. Segundo Barros *et al.* (2007), entre 2001 e 2005, a desigualdade de renda se reduziu consideravelmente e de maneira contínua, independentemente da medida considerada, alcançando em 2005 o nível mais baixo das três últimas décadas.

As Figuras 4.1 e 4.2 mostram os resultados da desigualdade brasileira para indivíduos entre 15 e 70 anos, nos anos de 2002 e 2007. A desigualdade de renda no Brasil ainda é bastante elevada, mas, contudo constataram-se reduções no período. Em 2002, o índice de Gini se encontrava em 0,60, cinco anos depois, no ano de 2007, este valor tem redução de 5%, caindo para 0,57. Para os índices de Theil, o comportamento é análogo, com ambos diminuindo no período.

Partindo para as análises dos grupos de interesse, em termos de cor ou raça, a desigualdade, independentemente do indicador utilizado, é maior para os brancos, tanto em 2002 quanto em 2007. No outro extremo, encontram-se os pretos ou indígenas com o menor nível de concentração de renda. Porém, conforme poderá ser verificado nas próximas seções, isso se deve em parte aos baixos rendimentos recebidos por estes indivíduos em relação aos demais.

Quanto à localização do domicílio, de acordo com o índice de Gini, a desigualdade é maior nas áreas rurais do que em áreas urbanas, tanto em 2002 quanto em 2007. Os resultados mostram também que, quando a análise é realizada baseada nos índices de Theil, o meio rural apresenta características importantes. De acordo com o T e L de Theil, a desigualdade de renda no meio rural concentra-se principalmente entre as rendas mais elevadas da distribuição e que os rendimentos dos mais pobres no rural brasileiro são bastante homogêneos. Ademais, apenas entre os mais pobres do meio rural existe uma elevação na concentração de renda entre 2002 e 2007, válido para qualquer grupo de análise.

Considerando agora as macrorregiões do Brasil, os resultados são compatíveis, de modo geral, com a literatura brasileira sobre o tema, em que a região Nordeste detém os maiores indicadores de concentração de renda. Por exemplo, para o índice de Gini, enquanto este chega a 0,62 em 2007 para o Nordeste, no Sudeste ele não ultrapassa 0,52,

ou seja, o nível de desigualdade de renda entre os nordestinos é cerca de 19% maior do que nos estados de Minas Gerais, São Paulo, Rio de Janeiro e Espírito Santo em conjunto. É importante destacar ainda a composição da desigualdade entre Nordeste e Sudeste: enquanto no primeiro a desigualdade é maior entre as rendas mais elevadas (T-Theil), para o Sudeste ocorre o contrário, isto é, a desigualdade é mais profunda entre os indivíduos de renda mais baixa.

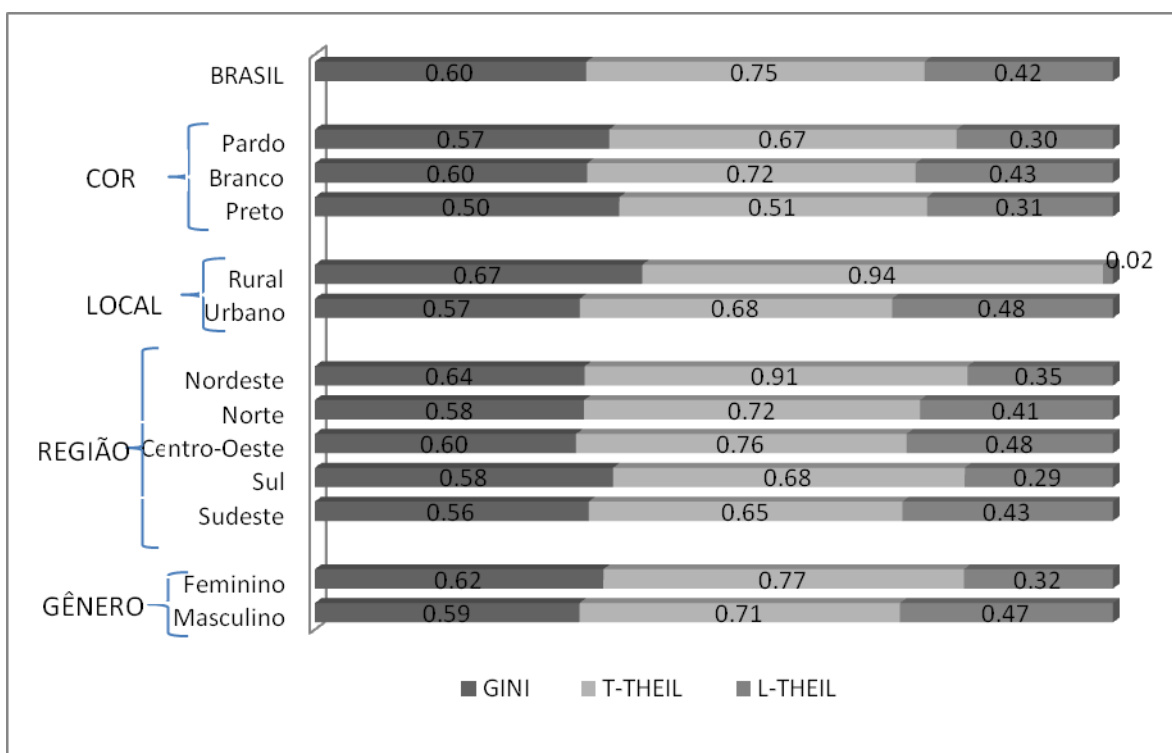


Figura 4.1. Índices de desigualdade para o Brasil e grupos de interesse, 2002.

Fonte: Elaborado pelo autor baseado nos microdados da PNAD 2002.

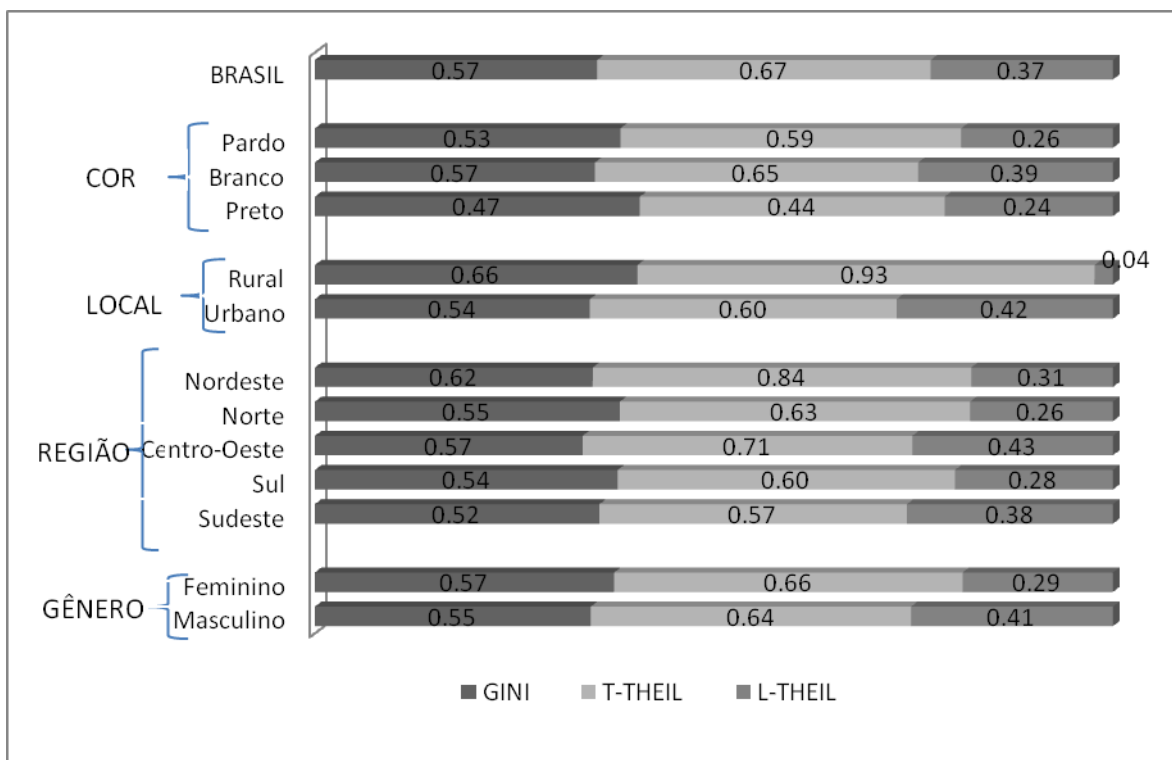


Figura 4.2. Índices de desigualdade para o Brasil e grupos de interesse, 2007.
 Fonte: Elaborado pelo autor baseado nos microdados da PNAD 2007.

Finalmente, as análises de gênero mostram que a desigualdade é maior para as mulheres. O índice de Gini passa de 0,62 em 2002 para 0,57 em 2007, para os homens o índice alterna de 0,59 para 0,55. Este resultado pode ser explicado ao menos parcialmente pelos diferenciais de ocupação entre homens e mulheres. De um total de 6.730.198 pessoas ocupadas em atividades domésticas em 2007, 93,78% eram mulheres, ou seja, 6.311.937 de pessoas. Adicionalmente, se forem considerados apenas os indivíduos em atividade doméstica que não obtiveram retorno financeiro (sem renda), do total de 28.383 pessoas nessa situação em 2007, 88% são mulheres, de acordo com os dados da PNAD (2007). Caso sejam consideradas apenas as mulheres no mercado de trabalho, os resultados mais uma vez mostram que parcela substancial, aproximadamente 16,4%, encontra-se em atividades de baixa remuneração como o trabalho doméstico.

Outrossim, a desigualdade de renda entre homens e mulheres tem comportamento distinto. No caso dos homens, a desigualdade tende a ser mais marcante entre os rendimentos mais baixos, e para as mulheres isso ocorre entre as rendas mais altas. Conforme mostrado anteriormente, a estrutura de ocupação das mulheres pode ser um fator que influencia estes resultados, isso porque o rendimento feminino tende a ser mais homogêneo para rendas mais baixas, em outras palavras, é nivelado por baixo.

4.1.3. Análise da Pobreza no Mercado de Trabalho Brasileiro entre 2002 e 2007

O índice FGT^0 mede a extensão da pobreza, o FGT^1 captura a intensidade da pobreza, ou seja, a distância média do indivíduo à linha de pobreza e o FGT^2 representa a desigualdade entre pobres, ou a distância média quadrática do indivíduo à linha de pobreza. As Figuras 4.3 e 4.4 mostram os resultados para essa família de índices.

Os indicadores de pobreza FGT representados nas Figuras 4.3 e 4.4 apresentam a mesma tendência de redução que os índices de desigualdade abordados na seção anterior. Da população de 15 a 70 anos de idade, utilizada na amostra, e considerando a linha de pobreza como sendo meio salário mínimo (R\$ 190,00) em valores de setembro de 2007, aproximadamente 21,7% dos indivíduos foram considerados pobres, com este montante se reduzindo para 19,9% em 2007. Estes indivíduos pobres estavam, em média, com rendimentos 15% abaixo da linha de pobreza em 2002 e 12,9% em 2007, em outras palavras, os pobres no Brasil ficaram menos pobres de 2002 para 2007.

Desagregando a análise primeiramente pela cor ou raça do indivíduo, verificou-se que os mais afetados pela pobreza são os pardos, independentemente do ano de referência, seguidos pelos negros. De acordo com as Figuras 4.5 e 4.6, apesar de a raça negra apresentar o segundo maior grau de pobreza, a contribuição desse grupo para a pobreza total é muito pequena, cerca de 6% em 2002 e 8% em 2007, fato este que pode ser explicado pela baixa representatividade dos negros na população total. Um resultado interessante é o hiato de pobreza (FGT^1) dos brancos no ano de 2007, apesar de cerca de 14,7% desta população ser considerada pobre em 2007, estes indivíduos estão muito próximos de ultrapassar a barreira da pobreza, ou seja, o FGT^1 assume o valor de 0,099 ou 9,9%.

As Figuras 4.5 e 4.6 também mostram os resultados para a localização do domicílio. Apesar de o meio rural apresentar indicadores de pobreza extremamente elevados com 55% da população residente em estado de pobreza em 2002 e 51,5%, no ano de 2007, as áreas urbanas apresentam peso maior para explicar a pobreza Urbano-Rural com 55% contra 45% para áreas rurais no FGT^0 .

Os resultados para as macrorregiões do Brasil não mostram nenhuma surpresa em relação ao que já é conhecido na literatura sobre o tema, ou seja, o Nordeste brasileiro é a região mais pobre do Brasil. De acordo com as Figuras 4.3 e 4.4, a extensão da pobreza teve leve queda entre 2002 e 2007, passando de 41,1% para 39,2% de pessoas pobres no

Nordeste. Além disso, os pobres do Nordeste estavam distantes da linha de pobreza em cerca de 28,2% em 2002 e 25,5% em 2007, o que quer dizer que os pobres nesta região recebiam em média R\$ 136,00 em 2002 (71,8% de R\$ 190,00) e R\$ 141,55 em 2007 (74,5% de R\$ 190,00).

Finalmente, na relação de gênero verifica-se que em 2007 o montante de homens pobres foi praticamente a metade do número de mulheres na mesma situação: 27,2% contra 14,5%. Das mulheres consideradas pobres em 2007, elas eram consideravelmente mais pobres do que os homens na mesma situação²⁶, isso porque a intensidade de pobreza para as mulheres era 0,185, enquanto para os homens foi de 0,087. O sexo feminino foi o principal contribuinte para a pobreza do Brasil, comparativamente aos homens, como mostrado nas Figuras 4.5 e 4.6.

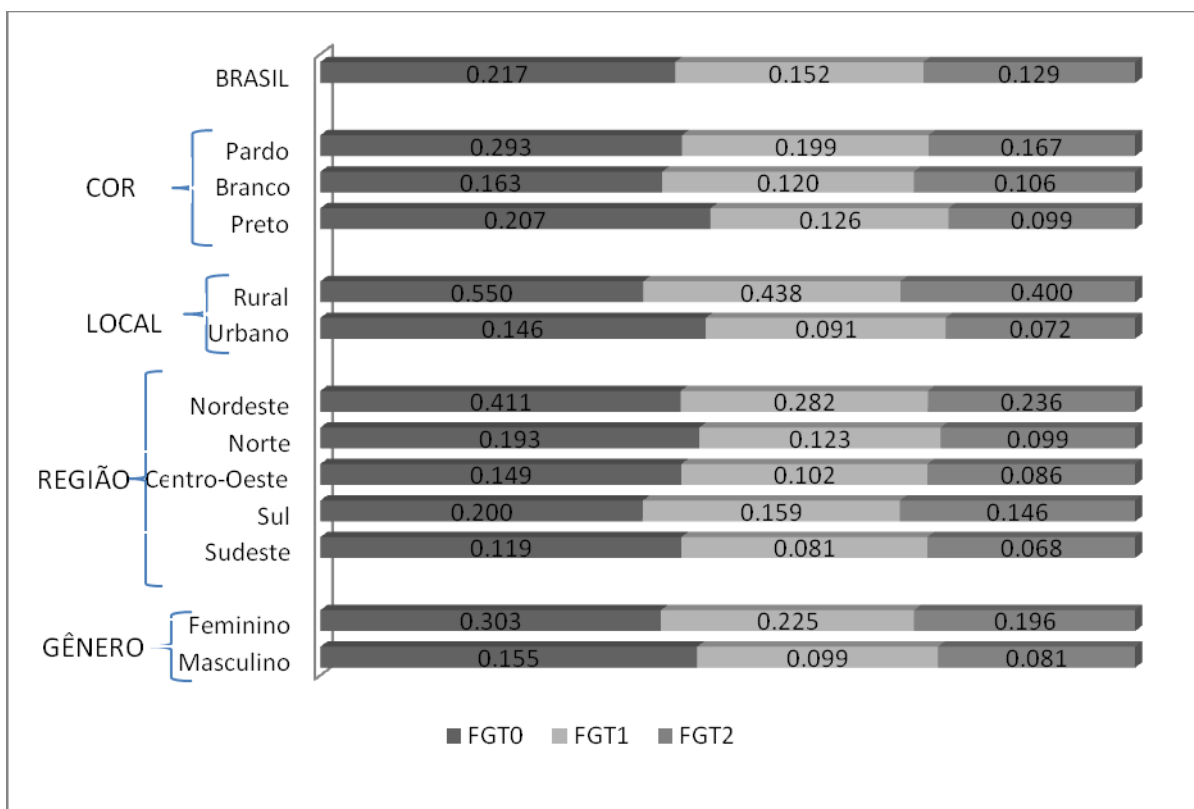


Figura 4.3. Índices de Pobreza no Brasil em 2002.

Fonte: Elaboração do autor baseada nos microdados da PNAD 2002.

²⁶ FGT¹.

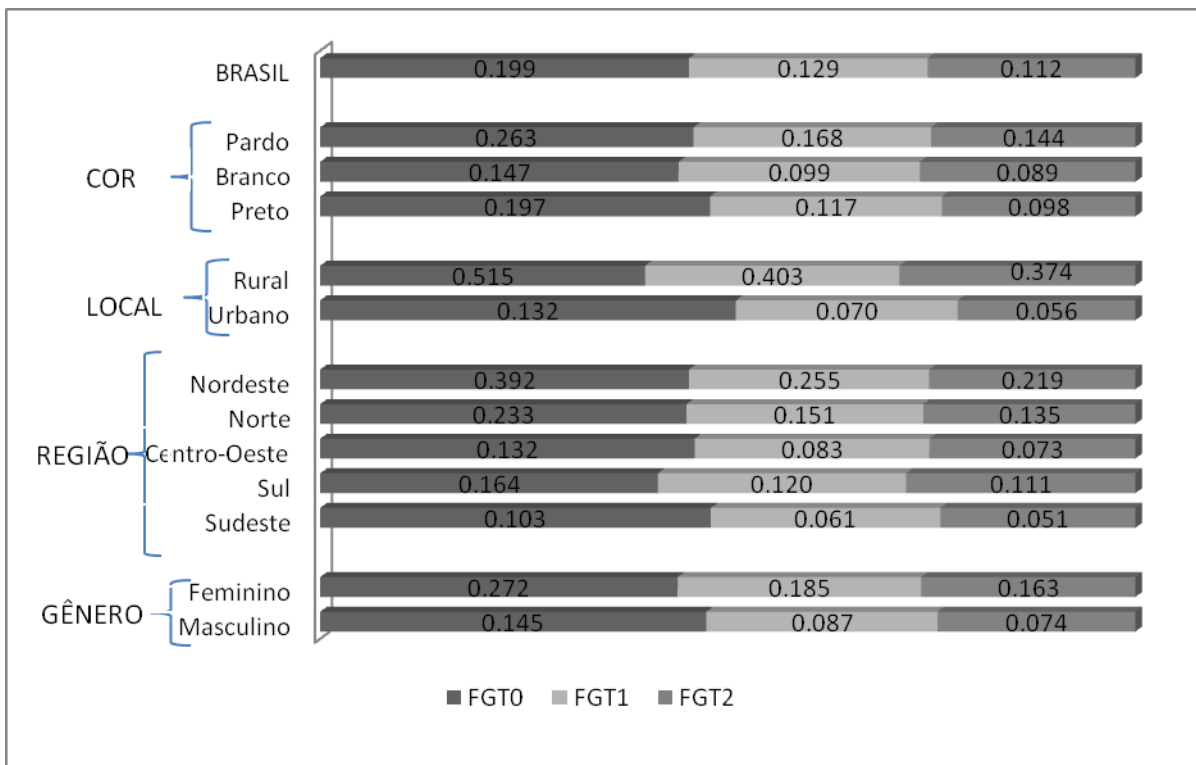


Figura 4.4. Índices de Pobreza no Brasil em 2007.
 Fonte: Elaboração do autor baseada nos microdados da PNAD 2007.

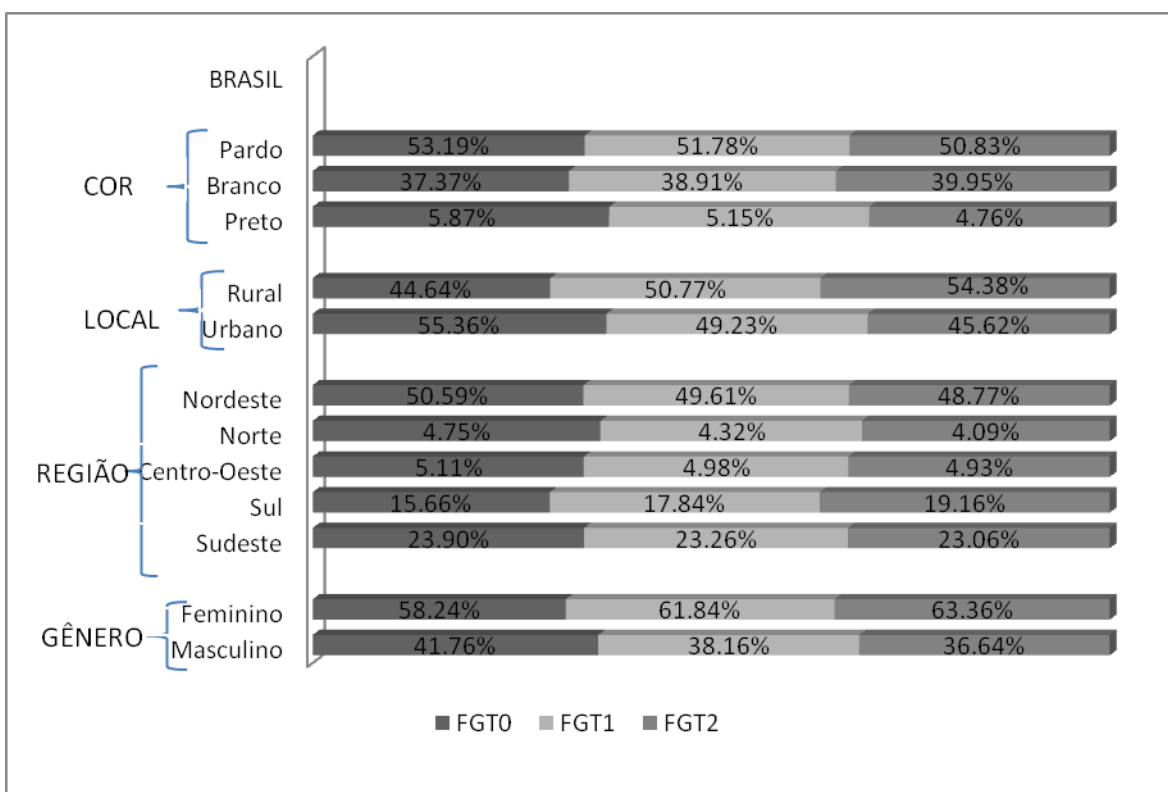


Figura 4.5. Contribuição dos Grupos de Interesse para a Pobreza Brasileira, 2002.
 Fonte: Elaboração do autor baseada nos microdados da PNAD 2002.

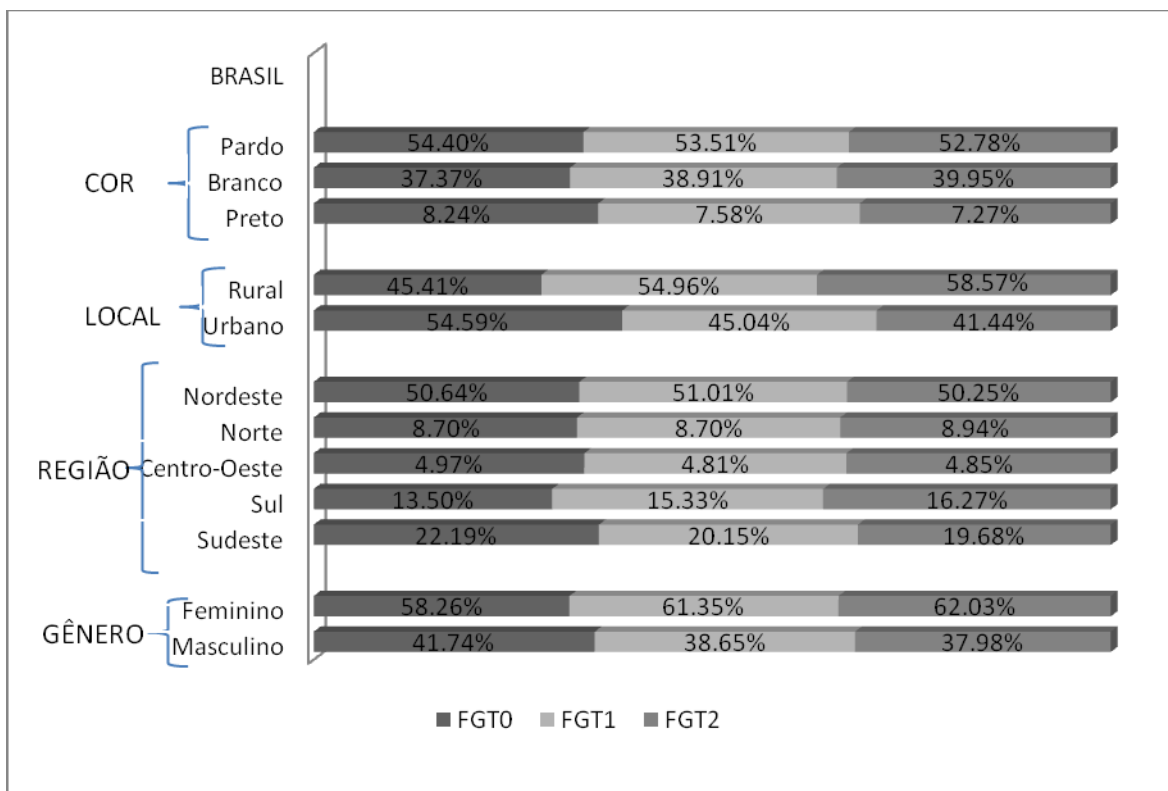


Figura 4.6. Contribuição dos Grupos de Interesse para a Pobreza Brasileira, 2007.
 Fonte: Elaboração do autor baseada nos microdados da PNAD 2007.

4.1.4. Participação no Mercado de Trabalho Brasileiro e Equações de Rendimento

As Tabelas 4.2 e 4.3 apresentam os resultados para a equação de participação. A análise é feita a partir de três modelos, ou seja, uma equação para todos os indivíduos, uma para os homens e outra para mulheres: a primeira consiste em verificar os efeitos das variáveis selecionadas sobre a entrada no mercado de trabalho para a população como um todo²⁷. Basicamente este primeiro modelo é utilizado para evidenciar a necessidade de se tratar o mercado de trabalho de forma distinta, ou seja, modelos separados para homens e mulheres. As demais frentes de análise consistem em verificar como as variáveis selecionadas impactam o sexo masculino e o feminino. A estatística MEFF²⁸ indica que, caso as estimativas desconsiderassem o plano amostral complexo, todas as variáveis teriam suas variâncias subestimadas tanto em 2002 como em 2007. Sendo assim, a consideração

²⁷ Este modelo será tratado nas tabelas como “Todos”, ou seja, toda a população sem distinção de gênero.

²⁸ *Misspecification effect*. Mais detalhes sobre a estatística MEFF podem ser encontradas no apêndice 2.

do plano amostral mostrou-se necessária para evitar o viés do desenho amostral²⁹. Quanto à significância estatística, verificou-se que o coeficiente da variável região geográfica Norte não foi estatisticamente significativo para determinar a entrada do indivíduo no mercado de trabalho³⁰ em nenhum dos modelos. No entanto, para o ano de 2007, no modelo de homens, ele apresentou significância estatística a 10%. Os demais coeficientes das variáveis foram significativos, sendo a grande maioria ao nível de 1%. A estatística de significância global (F) do modelo indica que a equação de participação descreve adequadamente os determinantes da probabilidade de o indivíduo se encontrar economicamente ativo, sendo estatisticamente significativa a 1%.

Antes de iniciar a análise propriamente dita dos coeficientes estimados, deve-se destacar que para o tratamento prático do modelo de seleção é necessário considerar os efeitos marginais das variáveis que na estimação do modelo probit não são constantes³¹. Por isso, não se podem interpretar diretamente os coeficientes, segundo Long e Freese (2006).

Para o primeiro modelo nos anos de 2002 e 2007, que compreende toda a amostra, verifica-se que o fato de o indivíduo ser do sexo masculino eleva a probabilidade de ele se encontrar no mercado de trabalho em aproximadamente 36 p.p. em 2002 e 31,4 p.p. em 2007. Este resultado apresenta duas características principais, primeiro tende a demonstrar certa segmentação no mercado de trabalho rural, como destacado por Gandra (2002), e também confirma o apontado pela análise das estatísticas descritivas (Tabela 4.1). Segundo, a redução nestes valores mostra que a probabilidade de entrada da mulher no mercado de trabalho tem aumentado, o que também é indicado nas estatísticas descritivas em que a taxa de atividade feminina se eleva de 2002 para 2007.

²⁹ As Tabelas A.1.1. e A.1.2. no apêndice 1 apresentam os resultados para os testes de heterocedasticidade.

³⁰ A 10% de significância.

³¹ Como é tradicionalmente utilizado na literatura nacional e internacional, os efeitos marginais foram calculados no ponto médio da amostra.

Tabela 4.2. Equação de participação no mercado de trabalho brasileiro, 2002¹.

Variáveis explicativas	TODOS			HOMENS			MULHERES		
	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF
constante	-3.348	-	1.85	-2.625	-	1.91	-2.672	-	1.81
Gênero (Masculino)	1.055	0.358	1.78	-	-	-	-	-	-
Pai ou Mãe solteiros	0.197	0.066	1.60	-0.074*	-0.017*	1.52	0.165	0.066	1.36
Situação conj. (casado)	-0.050	-0.018	1.66	0.207	0.051	1.60	-0.338	-0.134	1.48
Rendimento Familiar	-5.E-05	-2.E-05	13.09	9.E-05	2.E-05	8.26	6.E-05	2.E-05	6.08
Num. Comp. na Fam.	-3.5E-3ns	-1E-3ns	2.22	-0.11	-0.03	1.69	-0.03	-0.01	1.79
N. de crianças na Fam.	0.004ns	0.002ns	1.73	0.193	0.044	1.50	-0.0314	-0.012	1.93
Idade (dez. de anos)	1.766	-	1.68	2.127	-	1.80	1.532	-	1.59
idade^2 (dez. de anos)	-0.223	-	1.73	-0.267	-	1.83	-0.204	-	1.62
Cor Amarela ou pardas	0.054	0.019	1.65	0.110	0.02	1.41	0.123	0.049	1.34
Cor preta ou indígena	0.107	0.037	1.52	0.114	0.025	1.44	0.231	0.091	1.37
Educação em anos	0.038	0.014	2.01	-0.005**	-0.001**	2.20	0.043	0.017	1.88
Efeito Limiar da Edu.	0.130	0.046	2.23	-0.029	-0.007*	1.99	0.057	0.023	1.80
Região Sudeste	0.025**	0.009**	2.87	-0.114	-0.026	2.41	0.050	0.020	2.24
Região Sul	0.170	0.058	2.52	0.083	0.019	2.25	0.203	0.081	2.24
Região Centro-Oeste	0.105	0.036	2.69	0.076*	0.017*	2.07	0.023ns	0.0ns	2.33
Região Norte	-0.012ns	-0.00ns	3.80	-0.04ns	-0.01ns	2.29	-0.07ns	-0.0ns	2.49
Num. obs	221,199			108,999			112,200		
Pop. Considerada	99,804,804			49,408,764			50,396,040		
Estatística F (Prob)	0.00			0.00			0.00		

Fonte: Elaboração do autor baseada nos microdados da PNAD 2002.

¹ M.E. significa *Marginal Effect* ou Efeito Marginal. MEFF significa *Misspecification Effect*.

* valor significativo a 5% de significância.

** valor significativo a 10% de significância.

ns, valor não significativo a 10% de significância.

Os valores não assinalados são significativos a 1% de significância.

Tabela 4.3. Equação de participação no mercado de trabalho brasileiro, 2007¹.

Variáveis explicativas	TODOS			HOMENS			MULHERES		
	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF
constante	-3.314	-	2.60	-2.940	-	2.11	-2.858	-	1.82
Gênero (Masculino)	0.945	0.314	2.70	-	-	-	-	-	-
Pai ou Mãe solteiros	0.184	0.060	1.43	0.025ns	0.006ns	1.41	0.182	0.071	1.34
Situação conj. (casado)	-0.050	-0.017	1.74	0.234	0.057	1.49	-0.273	-0.107	1.55
Rendimento Familiar	5.E-05	2.E-05	34.58	9.E-05	2.E-05	18.92	5.E-05	2.E-05	34.08
Num. Comp. na Fam.	-7.1E-02	-2.4E-02	3.11	-0.12	-0.03	2.11	-0.04	-0.02	2.61
N. de crianc. na Fam.	0.046	0.016	2.33	0.187	0.043	1.58	-0.0411	-0.0163	2.77
Idade (dez. de anos)	1.851	0.63	1.90	2.195	0.50	1.69	1.608	0.64	1.62
idade^2 (dez. anos)	-0.237	-0.08	1.85	-0.275	-0.06	1.64	-0.213	-0.08	1.65
Cor Amarela pardas	0.101	0.034	1.47	0.105	0.02	1.46	0.104	0.041	1.41
Cor preta ou indígena	0.149	0.049	1.58	0.070	0.016	1.35	0.204	0.080	1.55
Educação em anos	0.037	0.013	2.40	0.021	0.005	2.24	0.049	0.020	1.81
Efeito Limiar da Edu.	0.039	0.013	4.24	-0.052	-0.012	2.78	0.057	0.023	3.60
Região Sudeste	0.033*	0.011*	3.66	-0.082	-0.019	2.84	0.119	0.047	2.76
Região Sul	0.166	0.055	3.87	0.056*	0.012*	2.54	0.245	0.096	2.83
Região Centro-Oeste	0.066	0.022	3.26	0.042ns	0.009**	2.36	0.085	0.034	2.66
Região Norte	0.008ns	0.003ns	4.05	0.047**	0.011**	3.26	-0.023ns	-0.009ns	2.88
Num. obs	238,449			118,054			120,395		
Pop. Considerada	113,490,535			56,275,537			57,214,998		
Estatística F (Prob)	0.00			0.00			0.00		

Fonte: Elaboração do autor baseada nos microdados da PNAD 2002.

¹ M.E. significa *Marginal Effect* ou Efeito Marginal. MEFF significa *Misspecification Effect*.

* valor significativo a 5% de significância.

** valor significativo a 10% de significância.

ns, valor não significativo a 10% de significância.

Os valores não assinalados são significativos a 1% de significância.

Quanto à situação conjugal, para os modelos compostos de homens e o de mulheres, a variável apresentou elevada significância estatística. No primeiro, o fato de o homem ser casado eleva sua probabilidade de se encontrar no mercado de trabalho em cerca de 5 p.p., e esse valor varia muito pouco entre os dois anos. Contudo, para as mulheres casadas, a probabilidade de estar economicamente ativa é negativa, ou seja, as casadas têm probabilidade de 13,4 p.p. a menos de estar no mercado de trabalho que as demais mulheres em 2002 e 10,7 em 2007, e esse último valor mostra que as casadas têm

aumentada a probabilidade de entrar no mercado relativamente a 2002. Estes resultados podem ser justificados pelas características culturais de diversas economias ocidentais, como é o caso brasileiro, em que o homem assume o papel de provedor da família, enquanto a mulher tende a se dedicar ao bem-estar da família. Um fator que pode acentuar esse efeito é o fato de o homem receber, em média, aproximadamente 1,5 vezes mais que as mulheres em 2007. É interessante destacar ainda a precarização das atividades desenvolvidas por mulheres em relação aos homens. De acordo com os dados da PNAD 2007, enquanto de um total de 42.153.850 homens casados e economicamente ativos, apenas 5,44% exercem atividades como empregado doméstico ou alguma atividade não remunerada, 24,85% das mulheres, de um total de 26.384.490, com as mesmas características, ou são empregadas domésticas ou exercem atividades não remuneradas (dona de casa).

Considerando agora pais ou mães solteiras, os resultados apontam que, para os homens, o fato de não haver um cônjuge na família não afeta a probabilidade de o indivíduo se encontrar no mercado de trabalho em 2007, mas para 2002 os pais solteiros apresentam a probabilidade de reduzir sua participação no mercado em 1,7 p.p.. No caso das mães solteiras, existe impacto considerável, aumentando a probabilidade de estas estarem no mercado de trabalho em 6,6 p.p. e 7,1 p.p. no período. Este fato pode afetar de forma negativa os rendimentos destas mulheres, uma vez que a necessidade de prover o sustento dos filhos deve reduzir o salário reserva delas, impulsionando um hiato na renda entre mães solteiras e as demais mulheres no mercado. Esta hipótese será testada a partir da análise das equações de rendimento.

A variável rendimento familiar, apesar de apresentar valores estatisticamente diferentes de zero, mostrou efeitos marginais extremamente pequenos, como, por exemplo, em 2007 para os homens, pois a cada R\$ 1.000 a mais no rendimento familiar o homem tende a aumentar em 2 p.p. a probabilidade de estar no mercado de trabalho, sendo a mesma relação verificada para as mulheres. O fato de o rendimento familiar, no presente estudo, apresentar relação positiva com a probabilidade de o indivíduo estar economicamente ativo pode, ao menos em parte, estar ligado à rede social familiar, ou seja, famílias mais ricas tendem a se relacionar com outras no mesmo padrão de vida. Desse modo, o acesso ao mercado de trabalho pode ser facilitado. No entanto, este resultado vai de encontro ao esperado no modelo teórico, isto é, uma vez que o indivíduo realiza um *tradeoff* entre trabalho e lazer, era esperado que quanto maior o rendimento familiar, maior

a demanda por lazer do indivíduo. Na literatura sobre determinantes da entrada no mercado de trabalho para o caso feminino, como, por exemplo, em Hoffmann e Kassouf (2005) ou Scorzafave e Menezes Filho (2001), a relação destas variáveis é negativa.

A quantidade de membros na família afeta negativamente a probabilidade tanto do homem quanto da mulher de estarem empregados, permanecendo estes valores praticamente estáveis entre 2002 e 2007, sendo o efeito relativamente pequeno. O número de crianças na família apresentou resultados interessantes para os casos dos homens e mulheres. Um filho adicional na família eleva a probabilidade de o homem se encontrar no mercado de trabalho em 4,4 p.p. em 2002 e 4,3 p.p. 2007, mostrando que historicamente o indivíduo do sexo masculino se comporta como provedor da família. Contudo, para as mulheres, filhos adicionais reduzem a probabilidade de participação na força de trabalho em aproximadamente em 1,2 e 1,6 p.p., comprovando a hipótese de que as mulheres são compelidas a abdicar, ao menos por determinado período de tempo, do mercado de trabalho rural em prol do bem-estar dos seus descendentes. Este resultado sugere que o salário reserva da mulher tende a aumentar nas famílias em que estão presentes crianças.

A idade e a idade ao quadrado apresentaram sinais condizentes com os encontrados na literatura sobre o tema em todos os modelos, como, por exemplo, em Resende e Wyllie (2006). Ou seja, observou-se um formato de U invertido na interação entre elas, de tal forma que o retorno sobre a experiência apresenta um ponto de máximo a partir do qual o retorno da renda é decrescente. O comportamento destas variáveis reflete a depreciação do capital humano ao longo do ciclo de vida. No presente estudo, o ponto de máximo encontra-se entre 40 anos de idade para os homens e 37 para as mulheres em média em qualquer ano analisado, sendo os valores próximos aos encontrados por Hoffmann e Kassouf (2005) para as mulheres.

Quanto à cor, os resultados indicam que o fato de os indivíduos serem pretos, pardos, indígenas ou amarelos eleva a probabilidade de estarem empregados. Isso ocorre tanto para homens quanto para mulheres nos dois anos analisados. O efeito é sentido principalmente para mulheres negras ou indígenas, com 9 p.p. a mais de probabilidade de estarem no mercado de trabalho que as brancas³² em 2002 e 8 p.p. em 2007. Este fato pode demonstrar apenas que os indivíduos desta cor ou raça necessitam desenvolver mais atividades ou que

³² Apesar de este resultado aparentar uma conquista para os indivíduos desta cor ou raça, estes resultados não determinam se o indivíduo exerce uma atividade formal ou precária, ademais não está relacionado aos rendimentos. Mais detalhes, PNAD (2007) notas metodológicas.

têm um salário reserva menor do que os brancos³³, o que será discutido na equação de rendimento.

Cada ano adicional investido em educação formal eleva a probabilidade de as mulheres participarem do mercado de trabalho em 1,70 p.p., no ano de 2002 e 2 p.p. em 2007, demonstrando que o investimento em capital humano é fator de importância para a entrada das mulheres no mercado de trabalho brasileiro. No entanto, no que concerne aos homens, a qualificação formal do indivíduo apresenta baixa significância estatística em 2002, e o valor do efeito marginal é negativo, contrariando o esperado segundo a teoria do capital humano, pois, em 2007, apesar de o sinal ser positivo, o efeito em si é muito pequeno, 0,5 p.p..

Por fim, os resultados para as regiões geográficas guardam semelhanças estreitas com o trabalho de Hoffmann e Kassouf (2005). Por exemplo, a região Sul apresenta a maior probabilidade de entrada no mercado de trabalho em relação ao Nordeste, em qualquer modelo ou ano considerado.

4.1.5. Equação de Rendimento para o Brasil 2002 e 2007

As Tabelas 4.4 e 4.5 mostram os resultados dos impactos das variáveis explicativas sobre a razão do logaritmo dos rendimentos e horas trabalhadas. Inicialmente verifica-se que a desconsideração da estrutura do plano amostral geraria resultados viesados da variância, uma vez que a estatística MEF³⁴ (Tabelas 4.5 e 4.6) apresentou valores positivos para todas as variáveis. A estatística F, ou seja, o ajustamento global do modelo, segue o encontrado para o modelo probit³⁵ nas Tabelas 4.3 e 4.4. A respeito da razão inversa de Mills, variável “*Lambda*”, esta também foi estatisticamente significativa a 1%, mostrando que sua inclusão era necessária para evitar o viés de seletividade.

³³ O que pode ser novamente um indicador de segmentação no mercado.

³⁴ *Misspecification Effect*.

³⁵ Isso ocorre porque a equação de seleção e a equação de rendimento são estimadas conjuntamente no procedimento de Heckman por pseudo máxima verossimilhança.

Tabela 4.4. Equação de rendimento para o Brasil 2002¹.

Variáveis explicativas	TODOS			HOMENS			MULHERES		
	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF
constante	-3.184	-95.8%	2.44	-0.926	-60.38%	2.09	-0.324	-27.6%	2.35
Gênero (Masculino)	0.507	66.10%	2.26	-	-	-	-	-	-
Pai ou Mãe solteira	0.039	3.97%	1.32	0.014ns	1.44%ns	1.35	-0.076	-7.33%	1.21
Situação conj. (casado)	0.045	4.6%	1.65	0.057	5.8%	1.59	0.131	14.0%	1.69
Idade (dez. de anos)	1.137	-	2.33	0.304	-	2.20	0.083*	-	2.14
idade^2 (dez. de anos)	-0.124	-	2.45	-0.017	-	2.32	-0.013	-	2.24
Cor amarela ou pardas	-0.120	-11.3%	1.75	-0.149	-13.87%	1.71	-0.162	-14.9%	1.42
Cor preta ou indígena	-0.117	-11.1%	1.48	-0.156	-14.48%	1.40	-0.206	-18.6%	1.49
Educação (anos)	0.091	9.49%	2.88	0.087	9.11%	2.50	0.057	5.84%	2.47
Efeito Limiar (edu)	0.172	30.04%	2.52	0.147	26.39%	2.08	0.152	23.22%	1.98
Região Sudeste	0.441	55.47%	4.36	0.437ns	54.7%ns	4.05	0.380	46.21%	3.13
Região Sul	0.399	49.02%	4.05	0.330	39.08%	4.43	0.271	31.17%	2.62
Região Centro-Oeste	0.438	54.90%	3.92	0.393	48.18%	3.31	0.374	45.35%	2.62
Região Norte	0.313	36.80%	4.45	0.315	37.05%	3.43	0.293	34.10%	2.64
Lambda	0.691	-	5.10	-0.770	-	4.88	-0.806	-	3.64
Num. obs	221,199			108,999			112,200		
Pop. Considerada	99,804,804			49,408,764			50,396,040		
Estatística F (Prob)	0.00			0.00			0.00		

Fonte: Elaboração do autor baseada nos microdados da PNAD 2002.

¹ A variável dependente considerada é o logaritmo do rendimento de todos os trabalhos sobre o total de horas trabalhadas. M.E. significa *Marginal Effect* ou Efeito Marginal. MEFF significa *Misspecification Effect*.

* valor significativo a 5% de significância.

** valor significativo a 10% de significância.

ns, valor não significativo a 10% de significância.

Os valores não assinalados são significativos a 1% de significância.

Tabela 4.5. Equação de rendimento para o Brasil 2007¹.

Variáveis explicativas	TODOS			HOMENS			MULHERES		
	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF
Constante	-0.443	-35.80%	9.12	-0.360	-30.25%	5.09	0.054ns	5.55%ns	6.50
Gênero (Masculino)	0.129	13.78%	8.45	-	-	-	-	-	-
Pai ou Mãe solteira	-0.037	-3.66%	1.63	0.003ns	0.32%ns	1.26	-0.049	-4.78%	1.34
Situação conj. (casado)	0.074	7.7%	1.64	0.059	6.1%	1.61	0.093	9.8%	2.18
Idade (dez. de anos)	0.293	-	7.86	0.291	-	5.19	0.131*	-	5.43
idade^2 (dez. de anos)	-0.016	-	8.11	-0.016	-	5.40	0.004ns	-	5.63
Cor amarela ou pardas	-0.134	-12.55%	1.96	-0.139	-12.95%	1.69	-0.129	-12.13%	1.42
Cor preta ou indígena	-0.129	-12.07%	1.68	-0.138	-12.92%	1.42	-0.119	-11.22%	1.39
Educação (anos)	0.063	6.45%	3.71	0.069	7.11%	2.84	0.046	4.74%	2.97
Efeito Limiar (edu)	0.140	22.45%	2.94	0.145	23.79%	2.22	0.149	21.56%	1.94
Região Sudeste	0.389	47.48%	7.06	0.412ns	50.98%ns	5.57	0.339	40.33%	3.74
Região Sul	0.348	41.59%	5.46	0.372	45.03%	4.22	0.293	34.06%	3.37
Região Centro-Oeste	0.415	51.49%	5.99	0.435	54.43%	4.65	0.371	44.97%	3.08
Região Norte	0.306	35.83%	8.36	0.318	37.38%	6.30	0.282	32.57%	4.43
Lambda	-0.497	-	17.04	-0.597	-	14.64	-0.579	-	11.87
Num. obs	238,449			118,054			120,395		
Pop. Considerada	113,490,535			56,275,537			57,214,998		
Estatística F (Prob)	0.00			0.00			0.00		

Fonte: Elaboração do autor baseada nos microdados da PNAD 2007.

¹A variável dependente considerada é o logaritmo do rendimento de todos os trabalhos sobre o total de horas trabalhadas. M.E. significa *Marginal Effect* ou Efeito Marginal. MEFF significa *Misspecification Effect*.

* valor significativo a 5% de significância.

** valor significativo a 10% de significância.

ns, valor não significativo a 10% de significância.

Os valores não assinalados são significativos a 1% de significância.

Os resultados para diferença de rendimentos entre homens e mulheres mostraram variações de grande magnitude entre 2002 e 2007. Comparativamente às mulheres, os indivíduos do sexo masculino tendiam a apresentar rendimento cerca de 66% mais elevado em 2002. Já em 2007 ocorreu uma queda significativa, com o efeito se reduzindo para 13,8%. Comprovando esta análise, Santos *et al.* (2008) destacam que para os anos 2000 existe uma tendência de perpetuação desse diferencial. Este resultado demonstra que o *gap* por gênero ainda é elevado e pode ser um indicador de discriminação no mercado de trabalho, devendo-se considerar, contudo, que o rendimento médio por horas trabalhadas é maior para as mulheres.

O fato de o homem ser considerado pai solteiro não afeta seus rendimentos em nenhum dos dois anos, mas para as mulheres os valores são estatisticamente significativos e elevados. Em 2002, a mãe solteira apresentava em média rendimentos 7% menores que as demais mulheres, e em 2007 há um arrefecimento e o valor reduz para 4,8%. Isto é um indício de que as mães solteiras têm salário reserva menor que as demais devido à maior necessidade de garantir o sustento de seus descendentes. Adicionalmente, das mães solteiras, 48% são mulheres negras ou indígenas em 2007 (PNAD, 2007).

Outra questão relevante são os rendimentos dos indivíduos casados. Por exemplo, o homem casado tende a apresentar rendimentos cerca de 6% maiores do que os demais homens, tanto em 2002 quanto 2007. Para as mulheres, o efeito é ainda maior, 14% em 2002 e aproximadamente 10% em 2007, mostrando que relacionamentos estáveis auxiliam a composição da renda individual.

Ratificando as análises feitas anteriormente sobre cor e raça, fica evidenciado, pelas Tabelas 4.5 e 4.6, que o salário reserva de indivíduos pretos ou indígenas e pardos ou amarelos é menor do que o dos indivíduos brancos. Isso se justifica porque pretos ou indígenas tendem a apresentar rendimentos de todos os trabalhos 11% menores do que os brancos em 2002 e 12% em 2007, e os amarelos ou pardos cerca de 11,3% e 12,55%. Contudo, deve-se considerar que, em termos de anos de estudo médio, os brancos têm aproximadamente 7,51 anos em 2007 contra 6,56 dos negros e 6 anos dos pardos. Pode-se inferir então que os diferenciais de renda entre negros e brancos estão mais vinculados aos desvios educacionais do que à discriminação racial propriamente dita.

Considerando os dados da variável “anos de estudo” e o limiar da educação, verifica-se que, até 10 anos de escolaridade, um ano adicional representa acréscimo, em média, de 9,11%, no rendimento dos homens em 2002 e de 7,11% em 2007. No caso das mulheres, o efeito da decisão de investimento na qualificação formal (educação) é menor, com valores variando de 5,84% em 2002 para 4,74% em 2007. Estes resultados mostram que a educação vem perdendo força na determinação da renda do trabalho. Todavia, levando em consideração os efeitos do limiar, ou dito de outra forma, o bônus sobre a qualificação, a partir de 10 anos de estudo, o acréscimo passa a ser próximo de 24% para os homens e de 21% para as mulheres no ano de 2007. Santos *et al.* (2008) afirmam que “anos de estudo e limiar apresentam redução contínua entre 2002 e 2006, percebe-se que a educação, embora extremamente importante para explicar o nível de renda, tem sua importância reduzida”. Quanto à experiência, ou seja, através de sua *proxy* (idade), esta

apresentou o sinal desejado para idade e idade ao quadrado, mostrando a relação em formato de U invertido, representando uma trajetória parabólica para experiência, evidenciando que o capital humano se deprecia com o passar do tempo. Uma exceção foi o modelo para mulheres em 2007, em que a variável idade^2 não apresentou significância estatística.

4.2. Determinantes da Participação e Rendimentos no Mercado de Trabalho Rural Brasileiro

A presente subseção tem como foco analisar alguns fatores que determinam a entrada do indivíduo no mercado de trabalho do meio rural do Brasil e gerar equações de rendimentos nos anos de 2002 e 2007. Novamente é dada atenção especial aos resultados relacionados à raça e ao gênero. As demais características como o desenho amostral da PNAD também são testadas.

4.2.1. Estatísticas Descritivas dos Dados da Amostra

A Tabela 4.6 apresenta as principais estatísticas descritivas do meio rural brasileiro em 2002 e 2007. Como se vê, as pessoas economicamente ativas situadas no meio rural do Brasil representaram cerca de 78% da população considerada em 2002, tendo esse percentual redução em 2007, caindo para 76,7%. Os indivíduos no mercado de trabalho rural possuíam rendimento médio que figurou entre R\$ 277,38 e 357,42. A renda familiar foi de R\$ 777,35 em 2002 para R\$ 994,58 em 2007, com as famílias tendo em média quatro membros. É importante destacar que o montante de famílias com crianças de zero a quatorze anos reduziu muito de 2002 para 2007. No primeiro ano, a taxa era de 56%, ou seja, em 56% das famílias existiam, em média, duas crianças entre zero e quatorze anos de idade, e em 2007 o valor se reduz para 51%. Há predominância dos indivíduos na região Nordeste e Sudeste, com 45,7% e 21,3% em 2007. No mesmo ano, a cor ou raça predominante foram os pardos ou amarelos com 54% da população, seguida dos brancos com 39% e preta ou indígena com 7%.

Em termos de gênero no meio rural, a predominância é do sexo masculino com 52% da população, montante que pouco se alterou no período destes, 91% em 2002 e 89% em 2007 eram economicamente ativos. Por outro lado, das mulheres consideradas na amostra, em torno de 64% no primeiro ano e 63% no último estavam no mercado de trabalho. Adicionalmente, para o grupo de mulheres, cerca de 7% são mães solteiras,

enquanto apenas 3% dos homens estavam nesta situação em 2007, estes valores apresentaram leve incremento em relação ao ano de 2002. Mais de 80% da população se encontra em algum tipo de situação conjugal estável, ou seja, casada.

Em média, a quantidade de horas semanais trabalhadas por homens no meio rural é consideravelmente maior que a das mulheres, ou seja, em 2002, as horas semanais trabalhadas por homens ultrapassaram as 40 horas semanais, reduzindo-se para 37 horas no fim do período. Para as mulheres, a variação é de 17 para 16 horas. Em termos de jornada de trabalho, pode-se dizer que o homem tende a trabalhar o expediente completo, enquanto as mulheres trabalham metade do expediente. Estes resultados são influenciados, pelo menos em parte, pelos diferenciais de salário reserva entre homens e mulheres, ou seja, espera-se que as mulheres casadas com filhos pequenos e/ou com muitos filhos tenham um salário reserva maior que o dos homens na mesma condição. Ou, como destaca Becker (1985), para o caso feminino, além deste efeito, espera-se também queda nos retornos do trabalho influenciado pela redução no tempo gasto no mercado devido, por exemplo, a períodos de gestação ou nos meses em que a criança é ainda muito pequena. Os resultados anteriores podem, pelo menos em parte, explicar o diferencial de rendimentos entre mulheres e homens. No entanto, deve-se destacar que no meio rural existem mais atividades manuais, que demandam considerável esforço físico. Dessa maneira, os homens tendem a ter maior inserção nessas atividades do que as mulheres.

Para as mulheres, a renda média mensal de todos os trabalhos era de R\$ 121,57 em 2002, elevando-se para R\$ 179,59 em 2007, enquanto para os homens o valor é significativamente maior, sendo R\$ 466,40, ou seja, 2,75 vezes o recebido pelas mulheres situadas no meio rural.

Quanto às variáveis que representam o capital humano, verifica-se que a idade média amostral passou de 36 para 37 anos. Os dados também sugerem um reduzido nível de qualificação, sendo os anos de estudo, em média, em torno de 4,8 em 2002 e 5,7 anos em 2007, ou seja, ensino fundamental incompleto. Os baixos valores para estas variáveis podem corroborar o pequeno rendimento médio para todos os trabalhos: R\$ 277,38 e R\$ 357,00.

Realizando o corte por gênero, verifica-se que as mulheres, em média, apresentam nível de qualificação, em anos de estudo, mais elevado do que os homens, 5,4 anos dos homens contra 6 anos de estudo para as mulheres no fim do período. Sendo assim, se os critérios relacionados à capacitação do indivíduo, em termos de anos de estudo, são uma

boa *proxy* para a determinação do nível de renda, então, espera-se que os incrementos na renda para o sexo feminino sejam superiores ou muito próximos aos obtidos pelos homens. Esta hipótese poderá ser analisada nas próximas seções. No entanto, as diferenças no salário reserva de homens e mulheres podem afetar a análise.

Tabela 4.6. Estatísticas descritivas das variáveis de interesse no meio rural, Brasil 2002-2007*.

Variáveis	TODOS				HOMENS				MULHERES			
	2002		2007		2002		2007		2002		2007	
	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.
Economicamente ativos	0.783	0.005	0.767	0.004	0.911	0.003	0.894	0.003	0.641	0.008	0.626	0.007
Renda de Todos Trabalhos	277.38	7.38	357.42	10.37	375.15	9.92	466.72	13.55	121.57	5.24	179.59	8.01
Log da Renda de Todos os Trabalhos	5.250	0.023	5.823	0.020	5.359	0.021	5.927	0.020	4.901	0.037	5.525	0.031
Pai ou mãe solteiros	-	-	-	-	0.022	0.001	0.029	0.002	0.064	0.003	0.068	0.003
Situação conjugal (casado)	-	-	-	-	0.852	0.004	0.823	0.004	0.838	0.004	0.815	0.004
Rendimento Familiar	777.35	15.29	994.58	24.67	787.92	15.24	1,011.34	27.41	765.54	16.32	975.97	23.68
Núm. horas todos os trabalhos (semana)	29.589	0.262	27.459	0.231	40.617	0.282	37.501	0.298	17.269	0.313	16.293	0.251
Número de Componentes na Fam.	4.374	0.032	4.083	0.027	4.431	0.034	4.121	0.030	4.310	0.031	4.039	0.026
Núm. crianças 0 a 14 anos na família	2.088	0.021	1.956	0.019	2.106	0.022	1.972	0.020	2.071	0.021	1.940	0.019
Proporção de famílias com crianças	0.563	0.006	0.510	0.005	0.543	0.006	0.490	0.006	0.586	0.006	0.532	0.006
Gênero (Masculino)	0.528	0.002	0.526	0.002	-	-	-	-	-	-	-	-
Idade (em dezenas de anos)	3.59	0.01	3.68	0.01	3.57	0.01	3.66	0.01	3.61	0.01	3.70	0.01
Cor preta ou indígena	0.049	0.003	0.069	0.004	0.054	0.003	0.069	0.004	0.044	0.003	0.068	0.004
Cor amarela ou pardas	0.508	0.011	0.538	0.010	0.512	0.011	0.550	0.010	0.503	0.011	0.524	0.010
Cor branca	0.443	0.010	0.394	0.010	0.434	0.011	0.381	0.010	0.453	0.011	0.408	0.010
Educação em anos de estudo	4.783	0.058	5.684	0.063	4.543	0.059	5.405	0.065	5.051	0.061	5.995	0.067
Região Centro-Oeste	0.061	0.005	0.060	0.005	0.063	0.005	0.062	0.005	0.059	0.005	0.058	0.005
Região Sudeste	0.237	0.012	0.213	0.011	0.234	0.012	0.210	0.011	0.239	0.012	0.217	0.011
Região Norte	0.012	0.002	0.108	0.013	0.013	0.002	0.111	0.013	0.011	0.002	0.104	0.013
Região Sul	0.187	0.011	0.162	0.010	0.186	0.011	0.161	0.010	0.188	0.011	0.164	0.010
Região Nordeste	0.504	0.014	0.457	0.015	0.505	0.014	0.456	0.015	0.503	0.014	0.457	0.015

Fonte: Elaboração do autor baseada nos microdados da PNAD 2002-2007.

* Em que Prop. significa Proporção e Desv. Pad. significa Desvio Padrão.

4.2.2. Equação de Participação para o Meio Rural Brasileiro

A presente seção descreve algumas características que impactam a probabilidade de o indivíduo, situado no meio rural brasileiro, encontrar-se economicamente ativo. Esta análise é realizada a partir de três frentes ou modelos, ou seja, uma equação para todos os indivíduos, uma para os homens e outra para mulheres: a primeira consiste em verificar os efeitos das variáveis selecionadas sobre a entrada no mercado de trabalho rural para a população como um todo³⁶. Basicamente, este primeiro modelo é utilizado para evidenciar a necessidade de se tratar o mercado de trabalho de forma distinta, ou seja, modelos separados para homens e mulheres. As demais frentes de análise consistem em verificar como as variáveis selecionadas impactam o sexo masculino e o feminino.

Iniciando a análise sobre a participação do indivíduo no mercado de trabalho rural, pode-se afirmar que, caso as estimativas não levassem em conta o plano amostral complexo da PNAD, todas as variáveis teriam suas variâncias subestimadas, de acordo com as Tabelas 4.7 e 4.8, estatística MEFF. Sendo assim, a consideração do plano amostral mostrou-se necessária para evitar o viés do desenho amostral.

Quanto à significância estatística, verificou-se que alguns coeficientes das variáveis não foram estatisticamente diferentes de zero³⁷ em alguns modelos. Um exemplo são os das macrorregiões do país, que apresentam significância estatística para o modelo de mulheres, mas não para o modelo de homens ou todos, justificando o tratamento de modelos diferenciados. Os demais coeficientes foram altamente significativos, principalmente ao nível de 1%. A estatística de significância global (F) para os três modelos corroborou os resultados anteriores, sendo estatisticamente significativa a 1%, sugerindo que os modelos descrevem adequadamente os determinantes da probabilidade de entrada do indivíduo no mercado de trabalho rural brasileiro.

³⁶ Este modelo será tratado nas tabelas como “Todos”, ou seja, toda a população sem distinção de gênero.

³⁷ A 10% de significância.

Tabela 4.7. Equação de participação no mercado de trabalho do meio rural, Brasil 20021.

Variáveis explicativas	TODOS			HOMENS			MULHERES		
	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF
Constante	-3.093	-	2.34	-1.958	-	2.00	-2.634	-	1.84
Gênero (Masculino)	1.505	0.490	2.57	-	-	-	-	-	-
Pai ou Mãe solteira	0.247	0.073	1.57	0.078ns	0.012ns	1.29	0.049ns	0.019ns	1.05
Situação conj. (casado)	-0.024ns	-0.008ns	1.71	0.309	0.059	1.45	-0.421	-0.165	1.25
Rendimento Familiar	2.E-04	7.E-05	13.3	4.E-04	6.E-05	18.60	2.E-04	7.E-05	6.24
Num. Comp. na Fam.	-7.1E-02	-2.3E-02	2.15	-0.13	-0.02	1.92	-0.02**	-0.01**	1.59
N. de crianças na Fam.	0.089	0.028	1.57	0.191	0.031	1.22	0.01ns	0.005ns	1.65
Idade (dez. de anos)	1.513	0.48	2.29	1.669	0.27	2.35	1.385	0.53	1.54
idade^2 (dez. de anos)	-0.180	-0.01	2.13	-0.197	-0.03	2.24	-0.172	-0.07	1.42
Cor Amarela ou pardas	0.103	0.033	1.80	0.100	0.02	1.38	0.119	0.046	1.71
Cor preta ou indígena	0.266	0.078	1.72	0.314	0.043	1.64	0.281	0.111	1.49
Educação (anos)	0.018*	0.006*	1.42	0.004ns	0.001ns	1.82	0.030**	0.012**	1.65
Efeito Limiar (edu)	0.017ns	0.005ns	1.52	-0.065	-0.011	1.63	0.035*	0.013*	1.28
Região Sudeste	-0.138*	-0.045*	3.06	-0.240	-0.043	2.32	-0.069ns	-0.026ns	2.03
Região Sul	0.114*	0.036*	2.52	0.005ns	0.001ns	2.35	0.174	0.068	2.15
Região Centro-Oeste	-0.096**	-0.032**	2.64	0.016ns	0.003ns	1.68	-0.218*	-0.081*	2.61
Região Norte	0.125ns	0.038ns	2.87	0.180ns	0.026ns	2.52	0.066ns	0.026ns	1.59
Num. obs	24,462			14,647			9,815		
Pop. Considerada	12,178,583			7,306,557			4,872,026		
Estatística F (Prob)	0.00			0.00			0.00		

Fonte: Elaboração do autor baseada nos microdados da PNAD 2002.

¹ M.E. significa *Marginal Effect* ou Efeito Marginal. MEFF significa *Misspecification Effect*.

* valor significativo a 5% de significância.

** valor significativo a 10% de significância.

ns, valor não significativo a 10% de significância.

Os valores não assinalados são significativos a 1% de significância.

Tabela 4.8. Equação de participação no mercado de trabalho do meio rural, Brasil 2007¹.

Variáveis explicativas	TODOS			HOMENS			MULHERES		
	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF
Constante	-3.969	-	1.83	-2.828	-	2.35	-3.597	-	1.45
Gênero (Masculino)	1.518	0.495	2.20	-	-	-	-	-	-
Pai ou Mãe solteira	0.271	0.081	1.60	0.097ns	0.015ns	1.59	0.251	0.099	1.36
Situação conj. (casado)	0.052ns	0.017ns	1.65	0.298	0.057	1.53	-0.211	-0.082	1.51
Rendimento Familiar	3E-05ns	1E-05ns	56.22	9.E-05ns	1.E-05ns	80.94	2E-05ns	7E-06ns	27.47
Num. Comp. na Fam.	-4.5E-02	-1.5E-02	1.59	-0.10	-0.02	2.35	0.01ns	0.00ns	1.43
N. de crianças na Fam.	0.035	0.011	1.05	0.128	0.022	1.42	-0.0451	-0.0175	1.13
Idade (dez. de anos)	1.910	0.62	1.61	2.084	0.35	2.90	1.749	0.68	1.30
idade^2 (dez. de anos)	-0.226	-0.07	1.57	-0.246	-0.04	2.72	-0.211	-0.08	1.28
Cor Amarela ou pardas	0.011ns	0.004ns	1.36	0.030ns	0.01ns	1.43	0.000ns	0.000ns	1.34
Cor preta ou indígena	0.032ns	0.010ns	1.14	0.004ns	0.001ns	0.99	0.051ns	0.020ns	1.31
Educação (anos)	0.039	0.013	1.39	0.058	0.010	1.43	0.041	0.016	1.28
Efeito Limiar (edu)	0.036	0.012	1.48	-0.074	-0.012	1.82	0.064	0.025	1.23
Região Sudeste	0.031ns	0.010ns	2.09	-0.029ns	-0.005ns	4.38	0.071ns	0.027ns	1.65
Região Sul	0.160**	0.050*	3.64	0.004ns	0.001ns	4.37	0.249	0.098	2.78
Região Centro-Oeste	-0.008ns	-0.003ns	2.17	0.014ns	0.002ns	3.92	-0.037ns	-0.014ns	2.12
Região Norte	-0.026ns	-0.008ns	3.09	0.054ns	0.009ns	3.88	-0.097ns	-0.037ns	2.53
Num. obs	29,994			17,535			12,459		
Pop. Considerada	14,701,310			8,558,957			6,142,353		
Estatística F (Prob)	0.00			0.00			0.00		

Fonte: Elaboração do autor baseada nos microdados da PNAD 2007.

¹M.E. significa *Marginal Effect* ou Efeito Marginal. MEFF significa *Misspecification Effect*.

* valor significativo a 5% de significância.

** valor significativo a 10% de significância.

ns, valor não significativo a 10% de significância.

Os valores não assinalados são significativos a 1% de significância.

Para o primeiro modelo, que compreende toda a amostra e considerando os dois anos, verifica-se que o fato de o indivíduo ser do sexo masculino eleva a probabilidade de ele se encontrar no mercado de trabalho rural em aproximadamente 49 p.p. Este resultado tende a demonstrar que existe um *gap* na seleção do mercado de trabalho rural, como destacado por Gandra (2002).

Quanto à situação conjugal, no modelo para toda população, esta variável não apresentou significância estatística a 10% em nenhum dos anos. No entanto, para os

modelos compostos de homens e de mulheres, a variável tem elevada significância estatística. No primeiro, o fato de o homem ser casado eleva sua probabilidade de se encontrar no mercado de trabalho rural em cerca de 5,9 p.p. em 2002, reduzindo-se um pouco em 2007, chegando a 5,7; para as mulheres casadas, a probabilidade de estarem economicamente ativas é negativa, ou seja, os efeitos têm redução acentuada no período. As casadas tinham probabilidade de 16,5 p.p. a menos de estarem no mercado de trabalho rural que as demais mulheres, caindo em 2007 esse valor para 8,2. Estes resultados para homens e mulheres podem ser justificados pelas características culturais de diversas economias ocidentais, sendo este o caso brasileiro, em que de certo modo o homem assume o papel de provedor da família, enquanto a mulher tende a se dedicar ao bem-estar da família. Um fator que pode acentuar esse efeito é o fato de o homem receber, em média, no meio rural brasileiro, aproximadamente 2,75 vezes a mais que as mulheres.

Considerando agora pais ou mães solteiras no meio rural, os resultados apontam que, para os homens e mulheres em 2002, o fato de não haver um cônjuge na família não afeta a probabilidade de o indivíduo se encontrar no mercado de trabalho rural. Por outro lado, para as mães solteiras em 2007 existe impacto considerável, aumentando a probabilidade de estas estarem no mercado de trabalho rural em 10 p.p. Isto pode afetar de forma negativa os rendimentos destas mulheres, uma vez que a necessidade de prover o sustento dos filhos deve reduzir o salário reserva delas, impulsionando um hiato na renda entre mães solteiras e as demais mulheres no mercado. Esta hipótese foi testada a partir da análise das equações de rendimento.

A variável rendimento familiar, tanto para homens quanto para mulheres, tende a aumentar a oferta de trabalhos destes grupos em 2002, porém o impacto é maior para as mulheres. E no caso dessas, a cada R\$ 1.000,00 de incremento no rendimento familiar, a probabilidade de estas se encontrarem economicamente ativas aumenta em 7 p.p., para os homens é próximo a 6 p.p. A explicação nesse caso está relacionada ao que foi considerado no caso do mercado de trabalho como um todo, ou seja, o *background* familiar geraria efeitos sobre a entrada dessas pessoas no mercado. Para o ano de 2007, não se verificou significância estatística em nenhum dos modelos.

A quantidade de membros na família afeta negativamente a probabilidade de o homem estar empregado em 2 p.p., sendo que este efeito permaneceu fixo nos dois anos. No caso feminino, só se encontrou significância estatística entre membros na família e na

probabilidade de a mulher estar empregada no ano de 2002, reduzindo sua probabilidade de entrada em 1 p.p.

O número de crianças na família apresentou resultados interessantes para os casos dos homens e mulheres. Um filho adicional na família eleva a probabilidade de o homem se encontrar no mercado de trabalho rural em 3,1 p.p. em 2002, reduzindo-se em 2007 para 2,2 p.p., mostrando que historicamente o indivíduo do sexo masculino se comporta como provedor da família. Contudo, para as mulheres, filhos adicionais reduzem a probabilidade de participação na força de trabalho em aproximadamente 1,75 p.p. em 2007³⁸. Este resultado comprova a hipótese de que as mulheres são compelidas ou incentivadas a abdicar, ao menos por determinado período de tempo, do mercado de trabalho rural em prol do bem-estar dos seus descendentes. Por outro lado, este resultado pode estar ligado também às possíveis transferências governamentais, como as provenientes do Bolsa Família, porque se o montante de mulheres que recebe essas transferências for consideravelmente maior do que o de homens, isso geraria efeitos sobre os resultados, visto que, com menos pressão sobre as mulheres (em termos de geração de renda), elas podem permanecer um tempo maior fora do mercado. Este resultado sugere que o salário reserva da mulher tende a aumentar nas famílias onde estão presentes crianças.

A idade e a idade ao quadrado apresentaram sinais condizentes com os encontrados na literatura sobre o tema, em todos os modelos, como, por exemplo, em Resende e Wyllie (2006). Ou seja, novamente observou-se um formato de U invertido na interação entre elas, de tal forma que o retorno sobre a experiência apresenta um ponto de máximo a partir do qual o retorno da renda é decrescente. O comportamento destas variáveis reflete a depreciação do capital humano ao longo do ciclo de vida. A título de exemplo, o ponto de máximo em 2007 encontra-se entre 42 anos de idade para os homens e 41 para as mulheres.

Em 2002, a cor ou raça foi um fator de seleção no meio rural. Os resultados indicam que os homens negros têm mais probabilidade de estar no mercado de trabalho que os brancos em 4,3 p.p.. Para os pardos, o valor não alcança os 2 p.p.³⁹ e para as mulheres os valores são ainda maiores para as negras, ultrapassando os 11 p.p. em relação às brancas e às pardas, cujo valor é de 4,6 p.p. Este fato pode demonstrar apenas que os indivíduos

³⁸ Em 2002, a variável não pode ser considerada estatisticamente diferente de zero a 10%.

³⁹ Apesar de estes valores aparentarem uma conquista para os negros ou pardos, estes resultados não determinam se o indivíduo exerce uma atividade formal ou precária, ademais não estão relacionados aos rendimentos. Mais detalhes em PNAD 2002 e 2007, nas notas metodológicas.

desta cor ou raça necessitam desenvolver mais atividades ou que têm um salário reserva menor do que os brancos⁴⁰, o que será discutido na equação de rendimento. Em 2007, não existe efeito de seleção por cor ou raça dos indivíduos para entrada no mercado de trabalho no meio rural brasileiro. Ou seja, nenhum dos coeficientes apresentou significância estatística para cor ou raça nos três modelos considerados.

No ano de 2002, cada ano adicional investido em educação formal eleva a probabilidade de as mulheres participarem do mercado de trabalho rural em 1,20 p.p. No entanto, no que concerne aos homens, a qualificação formal do indivíduo não se mostra significativa para este ano. Em 2007, o efeito é significativo e positivo para os dois grupos, com valores mais elevados para as mulheres, demonstrando que o investimento em capital humano é fator de importância para a entrada de homens e mulheres no mercado de trabalho rural.

4.2.3. Equação de Rendimento para o Meio Rural do Brasil

Analisando as Tabelas 4.9 e 4.10, que mostram os impactos das variáveis explicativas sobre a razão do logaritmo dos rendimentos de todos os trabalhos e o total de horas trabalhadas no mês, verifica-se que a desconsideração da estrutura do plano amostral geraria resultados viesados da variância dos regressores, uma vez que a estatística MEFF⁴¹ tem valores positivos para todas as variáveis. Quanto à significância estatística, boa parte das variáveis foi estatisticamente diferente de zero a 1% de significância, e a estatística F, ou seja, a que capta o ajustamento global do modelo, segue o encontrado para o modelo probit⁴² nas Tabelas 4.8 e 4.7. A respeito da razão inversa de Mills, variável “Lambda”, esta também foi estatisticamente diferente de zero em todos os modelos, mostrando que sua inclusão era necessária para evitar o viés de seletividade.

Comparativamente às mulheres, os indivíduos do sexo masculino tendem a apresentar rendimento cerca de 21% mais elevado em 2002. Em 2007 essa variável não foi importante para determinar o diferencial de rendimentos entre homens e mulheres no meio rural. Apesar da maior qualificação das mulheres, em termos de média de anos de estudo, elas tem rendimentos consideravelmente inferiores aos dos homens, refutando a hipótese de que a maior qualificação do sexo feminino tenderia, pelo menos, a deixá-las com

⁴⁰ O que pode ser novamente um indicador de segmentação no mercado.

⁴¹ *Misspecification Effect*.

⁴² Isso ocorre porque a equação de seleção e a equação de rendimento são estimadas conjuntamente pelo procedimento de Heckman.

rendimentos próximos aos obtidos pelos homens. Porém, deve-se enfatizar que no meio rural a demanda por trabalhos manuais tende a ser elevada e, neste caso, por questões fisiológicas, a produtividade masculina é superior, o que pode confirmar o maior rendimento.

Tabela 4.9. Equação de rendimento para o meio rural, Brasil 2002¹.

Variáveis explicativas	TODOS			HOMENS			MULHERES		
	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF
Constante	-0.483	-38.33%	2.65	-0.511	-40.03%	1.75	-0.180ns	-16.4%ns	1.92
Gênero (Masculino)	0.233	20.79%	3.09	-	-	-	-	-	-
Pai ou Mãe solteira	-0.071**	-6.81%**	1.37	0.007ns	0.71%ns	1.30	-0.012ns	-1.24%ns	1.13
Situação conj. (casado)	0.082	8.5%	1.55	0.053*	5.5%*	1.46	0.226	25.3%	1.33
Idade (dez. de anos)	0.139	-	1.96	0.060ns	-	1.71	-0.060ns	-	1.73
idade^2 (dez. de anos)	-0.007ns	-	1.94	0.001ns	-	1.70	0.026ns	-	1.72
Cor amarela ou pardas	-0.122	-11.48%	2.33	-0.117	-11.04%	1.80	-0.141*	-13.13%*	1.86
Cor preta ou indígena	-0.118*	-11.14%*	1.82	-0.086*	-8.23%*	1.76	-0.233*	-20.82%*	1.37
Educação (anos)	0.071	7.40%	1.75	0.070	7.26%	1.59	0.071	7.40%	1.98
Efeito Limiar (edu)	0.017*	9.22%*	1.71	0.029*	10.41%*	1.60	-0.001ns	7.25%ns	1.61
Região Sudeste	0.387	47.25%	5.50	0.378ns	45.90%ns	4.28	0.407	50.30%	2.67
Região Sul	0.397	48.80%	5.11	0.408	50.37%	4.21	0.328	38.85%	2.33
Região Centro-Oeste	0.486	62.66%	4.93	0.425	52.99%	3.83	0.654	92.34%	3.07
Região Norte	0.100ns	10.47%ns	6.68	0.086ns	8.93%ns	5.04	0.107ns	11.24%ns	6.08
Lambda	-0.830	-	4.69	-1.087	-	5.18	-0.982	-	2.94
Num. obs	24,462			14,647			9,815		
Pop. Considerada	12,178,583			7,306,557			4,872,026		
Estatística F (Prob)	0.00			0.00			0.00		

Fonte: Elaboração do autor a partir dos microdados da PNAD 2002.

¹ A variável dependente considerada é o logaritmo do rendimento de todos os trabalhos sobre o total de horas trabalhadas. M.E. significa *Marginal Effect* ou Efeito Marginal. MEFF significa *Misspecification Effect*.

* valor significativo a 5% de significância.

** valor significativo a 10% de significância.

ns, valor não significativo a 10% de significância.

Os valores não assinalados são significativos a 1% de significância.

Tabela 4.10. Equação de rendimento para o meio rural, Brasil 2007¹.

Variáveis explicativas	TODOS			HOMENS			MULHERES		
	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF
Constante	-0.48**	-38%**	6.10	-0.12ns	-12%ns	14.49	-1.01**	-63.7%**	4.85
Gênero (Masculino)	0.011ns	1.07%ns	6.50	-	-	-	-	-	-
Pai ou Mãe solteira	-0.129*	-12.1%*	2.04	-0.059ns	-5.7%ns	1.47	-0.16**	-14.6%**	1.49
Situação conj. (casado)	0.027ns	2.7%ns	1.44	0.014ns	1.4%ns	2.01	0.026ns	2.6%ns	1.47
Idade (dez. de anos)	0.301	-	5.27	0.145ns	-	13.88	0.531**	-	4.33
idade^2 (dez. de anos)	-0.025*	-	5.27	-0.008ns	-	13.85	-0.05**	-	4.28
Cor amarela ou pardas	-0.101	-9.58%	1.49	-0.098	-9.37%	1.41	-0.112	-10.56%	1.01
Cor preta ou indígena	-0.06**	-5.8%**	2.06	-0.079*	-7.58%*	1.79	-0.01ns	-0.96%ns	1.22
Educação (anos)	0.051	5.24%	2.09	0.051	5.25%	1.93	0.043	4.36%	2.26
Efeito Limiar (edu)	0.032	8.64%	1.71	0.030	8.46%	1.45	0.053	5.47%	1.74
Região Sudeste	0.449	56.74%	4.17	0.462ns	58.8%ns	3.63	0.418	51.90%	2.04
Região Sul	0.547	72.81%	4.30	0.558	74.63%	3.68	0.535	70.74%	2.46
Região Centro-Oeste	0.532	70.30%	4.22	0.526	69.28%	3.36	0.530	69.93%	2.11
Região Norte	0.527	69.40%	8.24	0.522	68.62%	8.68	0.512	66.81%	3.66
Lambda	-0.325*	-	8.34	-0.530*	-	35.05	-0.19ns	-	5.74
Num. obs	29,994			17,535			12,459		
Pop. Considerada	14,701,310			8,558,957			6,142,353		
Estatística F (Prob)	0.00			0.00			0.00		

Fonte: Elaboração do autor a partir dos microdados da PNAD 2007.

¹ A variável dependente considerada é o logaritmo do rendimento de todos os trabalhos sobre o total de horas trabalhadas. M.E. significa *Marginal Effect* ou Efeito Marginal. MEFF significa *Misspecification Effect*.

* valor significativo a 5% de significância.

** valor significativo a 10% de significância.

ns, valor não significativo a 10% de significância.

Os valores não assinalados são significativos a 1% de significância.

A Figura 4.7 procura elucidar o diferencial por gênero existente entre os rendimentos no meio rural brasileiro. Os retornos em reais de homens e mulheres no meio rural brasileiro são apresentados para três níveis de formação: ensino fundamental completo, ensino médio completo e superior completo. Como se vê, a remuneração masculina é consideravelmente maior que a feminina, mesmo com níveis semelhantes de qualificação. Para homens e mulheres com nível superior, no meio rural, a renda do primeiro grupo chega a ser 52% mais elevada do que das mulheres.

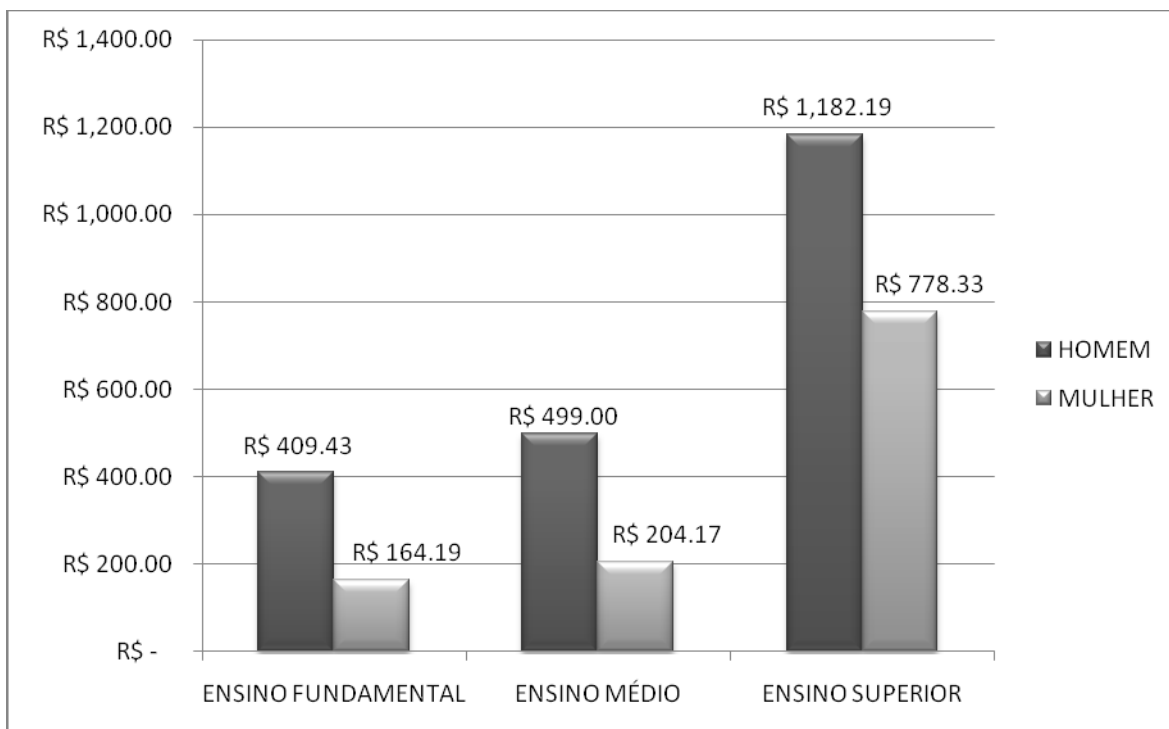


Figura 4.7. Rendimento médio de todos os trabalhos para homens e mulheres* no meio rural, 2007.

Fonte: Elaboração do autor baseada nos microdados da PNAD 2007.

*Ensino superior completo foi considerado para indivíduos com 15 anos de estudo ou mais.

Em 2002 e 2007, o fato de o indivíduo ser considerado pai solteiro não afeta os retornos do trabalho. Por outro lado, apesar de em 2002 a variável mãe solteira não ser diferente de zero ao nível de significância de 10%, para 2007 ela tende a receber 14,7% a menos que as demais mulheres. Este resultado confirma a hipótese de que mães solteiras têm salário reserva consideravelmente menor do que as demais mulheres. Outrossim, as mães solteiras trabalham em média 19 horas por semana contra 16 das demais mulheres, de acordo com os dados da PNAD (2007), que sugere que estas mulheres, além de apresentarem salário reserva menor, devem realizar mais atividades para garantir o sustento dos descendentes.

No meio rural, em 2002, as mulheres casadas tem retornos maiores do que os homens casados: enquanto para elas o retorno é de aproximadamente 25%, para o sexo oposto o efeito é de 5,5%. Em 2007, os resultados não foram estatisticamente diferentes de zero.

Diferentemente da participação no mercado de trabalho rural, a cor ou raça do indivíduo influenciam, em praticamente todos os modelos e nos dois anos, os rendimentos recebidos de todos os trabalhos. Este efeito é sentido tanto para homens como para

mulheres, excetuando-se as mulheres pretas ou indígenas em 2007. Homens pardos ou amarelos tendem a apresentar rendimentos 11,4% menores do que os brancos em 2002, sendo o diferencial arrefecido no fim do período, para 9,37%, enquanto entre os pretos ou indígenas o efeito gira em torno de -8,23% em 2002 e 7,60% em 2007. É interessante observar que este diferencial é maior entre as mulheres: para as pretas em 2002 o diferencial chegava a 21% em relação às brancas e para as pardas era de 13,13%. Por estes resultados, verifica-se certa segmentação no mercado de trabalho rural quanto à cor ou raça, tanto para homens como para mulheres. Contudo, deve-se considerar que, em termos de anos de estudo médio, os brancos têm aproximadamente 6,12 anos contra 4,97 e 4,70 anos dos pretos ou indígenas e pardos ou amarelos em 2007. Para as mulheres, as brancas apresentam 6,6 anos de estudo enquanto as pretas ou indígenas e pardas ou amarelas ficam com 5,61 e 5,28 anos. Pode-se dizer então que os diferenciais de rendimentos entre negros e brancos ou pardos e brancos seriam decorrentes mais do desnível educacional do que da discriminação propriamente dita.

Em 2002, os retornos da variável educação foram maiores para as mulheres do que para os homens: 7,40% contra 7,26%. No entanto, o bônus sobre a qualificação, ou seja, o efeito limiar não apresentou significância estatística para as mulheres a 10%. Em 2007, tanto a educação quanto o efeito limiar foram estaticamente significativos para homens e mulheres. Neste ano, o impacto da educação se inverte, sendo maior para os homens até 5 anos de escolaridade, um ano adicional representando acréscimo, em média, de 5,25%, no rendimento dos homens no meio rural. Levando em consideração os efeitos do limiar, a partir de 6 anos de estudo, o acréscimo passa a ser próximo de 8,46%. Por outro lado, esses efeitos são menores para as mulheres, mesmo elas tendo um nível de escolaridade maior que o dos homens. Por exemplo, até 5 anos de escolaridade, um ano adicional representa acréscimo de apenas 4,36% na renda da mulher. Considerando o limiar, ou seja, a partir de 6 anos de estudo, o acréscimo é em torno de 5,47%, muito próximo da metade do que ocorre para o sexo oposto. O comportamento da experiência seguiu o esperado mostrando a relação em formato de U invertido e que o capital humano se deprecia com o passar do tempo.

5. CONSIDERAÇÕES FINAIS

O mercado de trabalho brasileiro passou por um ciclo de estruturação entre os anos de 1940 a 1970, podendo ser comparado com os mercados do mundo desenvolvido, onde se verificou expansão da taxa de assalariamento, do emprego formal e o desemprego foi reduzido. No entanto, entre a década de 1980 e início dos anos 2000 o Brasil vivenciou o outro extremo: taxa de informalidade crescente, nível elevado de desemprego e precarização das atividades. Entre 2001 e 2007 os dados mostraram que existe uma tendência de estabilização desse ciclo com uma leve retomada do padrão de estruturação do mercado.

Em meio a todas essas mudanças no mercado de trabalho brasileiro, encontra-se o indivíduo como peça fundamental desse mercado. A presente pesquisa focou o mercado de trabalho no Brasil, no sentido de verificar quais os fatores que mais impactam a seleção dos indivíduos e também sobre os rendimentos do trabalho, levando em consideração não apenas variáveis ligadas à produtividade, como também aquelas com capacidade de gerar discriminação e/ou segmentação, além de alguns fatores que afetam a própria decisão do indivíduo em ofertar trabalho, como é o caso de pais ou mães solteiras, a situação conjugal ou as famílias com crianças pequenas, o que por fim tende a afetar os rendimentos.

A literatura internacional mostra que os traços de discriminação e segmentação no mercado de trabalho não são um problema exclusivo do Brasil. No entanto, para o caso brasileiro, um montante considerável de toda a renda gerada no país provém do mercado de trabalho. Assim, trabalhos que procuram estudar as características que impactam na entrada e no rendimento dos indivíduos no mercado são importante ferramenta para realização de políticas públicas que tenham como foco equalizar esses aspectos. Ao longo

do trabalho, realizou-se uma análise criteriosa da dinâmica do mercado de trabalho brasileiro como um todo e também do mercado de trabalho rural. Torna-se importante destacar que os dados, quando tratados separadamente, ou seja, homens e mulheres, mostraram que o comportamento de algumas variáveis varia bastante, tanto em termos de sinais encontrados, como em magnitude, como é o caso de pais e mães solteiras, ou a situação conjugal. Além disso, o correto tratamento da PNAD, ou seja, a consideração do desenho amostral foi imprescindível para obter estimativas corretas das variâncias nos modelos. Por outro lado, o procedimento de Heckman, por meio da razão inversa de Mills, evitou o viés de seletividade. Ademais, a criação das variáveis, número de crianças na família, pai e mãe solteiros e homens e mulheres casadas possibilitou abordar temas não encontrados no questionário oficial da PNAD.

A respeito dos resultados, verificou-se que a amplitude dos microdados da PNAD, ou seja, as mais de 30.000 observações em um dos estudos e 290.000 no outro foram um importante fator para se conseguir alta significância estatística das variáveis e resultados mais robustos. Outra questão relevante é o elevado diferencial de rendimentos entre o mercado como um todo e o rural: a renda média no primeiro se aproxima dos R\$ 900,00 em 2007, enquanto no rural o rendimento médio é pouco maior que R\$ 350,00⁴³ no mesmo ano, evidenciando a relativa precariedade das ocupações no meio rural. Contudo, esses rendimentos apresentaram aumentos durante a década, e para o mercado de trabalho como um todo, o incremento de renda real foi de 7,3% e no meio rural de 29%, mostrando certa tendência de equalização ao longo dos anos nos rendimentos entre meio rural e os demais. Neste mesmo sentido, deve-se destacar que a desigualdade de renda e a pobreza se reduziram no período de 2002 a 2007, e esta tendência já havia sido apontada por outros autores.

Tanto para o mercado como um todo como para o meio rural, foram identificados traços de segmentação e discriminação no mercado de trabalho. Analisando-se os resultados para raça, enfocando o mercado como um todo, os pardos e negros têm maior probabilidade de entrada no mercado do que os brancos para os dois anos. No mercado rural, a raça não é um fator de seleção dos indivíduos em 2007. Por outro lado, nos dois segmentos do mercado de trabalho analisados, o rendimento dos negros ou pardos é menor do que o obtido pelos brancos. Apesar disso, deve-se considerar que o nível de qualificação

⁴³ Deve-se levar em consideração que os resultados são de períodos distintos, sendo assim aspectos macroeconômicos podem afetar as comparações. Outro ponto a ser destacado é o diferencial na amostra.

médio dos brancos é mais elevado do que o das demais raças, o que pode estar relacionado às oportunidades que os brancos têm em relação às demais raças. Sendo assim, pelo menos parte do diferencial de rendimentos está ligado ao *gap* de qualificação.

Quando o mercado de trabalho é analisado pela perspectiva do gênero, verificou-se que por qualquer medida considerada as mulheres se encontram em situação menos favorável do que os homens, excetuando o caso das mulheres no mercado de trabalho como um todo, que têm rendimento por hora de trabalho maior que o homem. No mercado de trabalho rural em 2002, por exemplo, as mulheres tenderiam a receber 21% a menos do que os homens, enquanto no mercado de trabalho brasileiro como um todo em 2007, o diferencial de salários de homens e mulheres é de 14%. Uma possível explicação para isso consiste no fato de que as mulheres, de modo geral, apresentam períodos fora do mercado de trabalho devido, por exemplo, à gestação ou aos cuidados de crianças pequenas. E também no meio rural, devido às características do meio, mesmo com qualificação menor do que a mulher, a produtividade do homem tende a ser maior devido à existência de muitas atividades manuais, como a dos “boias-frias”.

Levando-se em consideração as variáveis ligadas a fatores que afetam a própria decisão do indivíduo em ofertar trabalho e, conseqüentemente, a renda, os resultados mostram que o fato de as mulheres serem casadas reduz a probabilidade de participação no mercado de trabalho nos dois segmentos analisados e em todos os anos, enquanto para os homens o fato de se ter um cônjuge eleva a probabilidade de estar no mercado. Este resultado pode indicar uma característica cultural no Brasil, ou seja, a de que o homem é historicamente o provedor da família. Outro resultado que corrobora a ideia acima é a de que um filho adicional na família eleva a probabilidade de o homem de estar no mercado de trabalho. Para as mulheres, filhos adicionais reduzem a probabilidade de participação na força de trabalho. Ou seja, enquanto os homens tem redução no salário reserva quando têm mais filhos, as mulheres têm elevação. Verificou-se também que as mães solteiras apresentam um hiato de renda em relação às demais mulheres no mercado e que estas mulheres tendem a trabalhar mais horas semanais do que as outras.

Finalmente, quanto às variáveis relacionadas aos investimentos em capital humano, os resultados indicaram que em ambos os mercados a experiência mostra uma trajetória parabólica, evidenciando que o capital humano se deprecia ao longo do tempo. Os resultados também indicaram que os retornos sobre o nível de qualificação são um importante fator para geração de renda tanto no mercado rural como no mercado de

trabalho brasileiro. Além disso, o bônus sobre a qualificação, ou seja, o efeito limiar da educação gera efeitos consideráveis no nível de rendimentos, principalmente no caso dos homens.

6. REFERÊNCIAS

BARROS, R. P., MENDONÇA, R. 1993. “A Evolução do Bem-Estar e da Desigualdade no Brasil desde 1960”, in Teixeira, E. C. Desenvolvimento Agrícola na década de 90 e no século XXI. Viçosa.

BARROS, R. P., MENDONÇA, R. 1995(a). “Os Determinantes da Desigualdade no Brasil”, in IPEA, Rio de Janeiro.

BARROS, R. P., MENDONÇA, R. 1995(b). “Bem-estar, pobreza e desigualdade de renda: uma avaliação da evolução histórica e das disparidades regionais”, in IPEA, Rio de Janeiro.

BARROS, R. P. de, CARVALHO M. de, FRANCO, S. ,MENDONÇA, R. 2007. “A Queda Recente da Desigualdade de Renda no Brasil”, in HENRIQUES, (org). IPEA, Rio de Janeiro.

BEHRMAN J.R. and BIRDSALL N. 1983. “The Quality of Schooling: Quantity Alone is Misleading”, American Economic Review, 73(5), pp. 928-946.

BERNDT, E. 1996. “The Practice of Econometrics . Classic and Contemporary”, Addison-Wesley Publishers.

BONELLI, R. 2002. “Crescimento, Desigualdade e Educação: Notas para uma Resenha com Referência ao Brasil”. Economia Aplicada, vol.6, nº 4.

BLOOM, G. F., NORTHRUP, H. R. 1973. **Economics of Labor Relations**. Irwin-Dorsey Limited. Homewood, Illinois. P. 771.

BROWNE, I. 1997. "Explaining the Black-White Gap in Labor Force Participation among Women Heading Households". *American Sociological Review*, Vol. 62, No. 2 (April), pp. 236-252.

CAMARGO, J. M.; NERI, M.; REIS, M. C. 1999. "Emprego e Produtividade no Brasil na Década de Noventa". Fundação Getúlio Vargas, texto para discussão n. 434, Rio de Janeiro. Disponível em: <[http://www.fgv.br/ibre/cps/artigos/Textos_Discus/Emprego e Produtividade no Brasil na Decada de Noventa.pdf](http://www.fgv.br/ibre/cps/artigos/Textos_Discus/Emprego_e_Produtividade_no_Brasil_na_Decada_de_Noventa.pdf)>.

CHISWICK, B. R. and MINCER, J. 2003. "Experience and the Distribution of Earnings. University of Illinois at Chicago and IZA Bonn", Available in: <<http://ssrn.com/abstract=435260>>.

CIRINO, J. F.; LIMA, J. E. 2008. "Participação Feminina no Mercado de Trabalho: uma Análise de Decomposição para o Brasil e a Região Metropolitana de Belo Horizonte", In: XIII Seminário sobre Economia Mineira, 2008, Diamantina. XIII Seminário sobre Economia Mineira. Belo Horizonte : Cedeplar/UFMG. v. único.

COCHRAN, W.G. 1977. "Sampling Techniques", 3rd Edition. New York: John Wiley & Sons.

CORRÊA, A. M. C. J. 1998. "Distribuição de Renda e Pobreza na Agricultura Brasileira (1981-1990)", Editora UNIMEP, Piracicaba. 260 p.

DAVIES, H.; JOSHI, H. 1998. "*Gender and Income Inequality in the UK 1968-1990: The Feminization of Earnings or of Poverty?*" *Journal of The Royal Statistical Society. Series A (Statistics in Society)*, Vol. 161, No. 1, pp. 33-61.

DEDECCA, C. S.; BALTAR, P. E. de A. 1997. “Mercado de Trabalho e Informalidade nos Anos 90”. Estudos Econômicos, São Paulo, V.27, N. Especial, pág. 65-84.

DEL GROSSI, M. E., GRAZIANO, J. S. 2002. “O Uso das PNADs para Áreas Rurais”. Rio de Janeiro: IPEA, Texto para Discussão 874, Abril de.

DINIZ, B. P. C. et al. 2008. Perfis dos Rendimentos e dos Orçamentos Familiares Brasileiros: o que Diferencia o Rural? *Texto para Discussão (no prelo)*, Brasília: IPEA.

EFRON, B. 1979. “Bootstrap Methods: Another Look at the Jackknife”. **Annals of Statistics**, v.7, n.1, p. 1-26, Jan.

ENTWISLE, B.; HENDERSON, G. E.; SHORT, S. E.; BOUMA, J.; FENGYING, Z. 1995. “Gender and Family Businesses in Rural China”. *American Sociological Review*, Vol. 60, No. 1 (Feb), pp. 36-57.

GANDRA, R. M. 2002. “O debate Sobre a Desigualdade de Renda no Brasil: da Controvérsia dos anos 70 ao Pensamento Hegemônico nos anos 90”. Dissertação de Mestrado. Niterói (RJ): UFF.

GRAZIANO DA SILVA, J. 1998. “A Nova Dinâmica da Agricultura Brasileira”. Campinas: UNICAMP. IE,.

GRAZIANO DA SILVA, J., DEL GROSSI, E. 2001 “O Novo Rural Brasileiro: uma atualização para 1992-98”. IE/Unicamp.

GREENE, W. H. 2003 “Econometric Analysis”, New York: Pearson, 1026p.

GUIMARÃES, P. W. 2007. “Variação de Renda Familiar, Desigualdade e Pobreza no Brasil”. Tese (doutorado) – Universidade Federal de Viçosa.

GUJARATI, D. N. 2000 “Econometria Básica”. Terceira Edição. São Paulo: Macro Books.

HASENBALG, C. 1992. “A Pesquisa das Desigualdades Raciais no Brasil”. In: SILVA, N.V.; HASENBALG, C. (orgs.) “Relações Raciais no Brasil Contemporâneo”. Rio de Janeiro: Rio Fundo Editora.

HECKMAN, J. J. 1979. “*Sample Selection Bias as a Specification Error*”. *Econometrica*. Vol. 47, No. 1, Jan., pp. 153-161.

HOFFMANN, R. 1989. “Evolução da Distribuição de Renda no Brasil, entre Pessoas e entre Famílias, 1979 e 1986”, in SEDLACEK, G. L., BARROS, R. P. (eds). Mercado de Trabalho e Distribuição de Renda: uma Coletânea. Rio de Janeiro: IPEA/INPES,.

_____. 1992. “Vinte Anos de Desigualdade e Pobreza na Agricultura Brasileira”, *Revista de Economia e Sociologia Rural*, v. 30, n.2, p. 97-113, abr./jun.

HOFFMANN, R. and SCAMPINI, P. J. 1996. “Desigualdade e Pobreza na Agricultura do Estado de Minas Gerais”. *Nova Economia*. Belo Horizonte, v.6, n.2, p.67-84, nov.

HOFFMANN, R. 1998. “Distribuição de Renda: Medidas de Desigualdade e Pobreza”. São Paulo: Edusp.

_____. 2000. “Mensuração da Desigualdade e da Pobreza no Brasil”, in HENRIQUES, R. (eds) *Desigualdade e pobreza no Brasil*. Rio de Janeiro: IPEA.

HOFFMANN, R. LEONE E. and T. 2004 “Participação da Mulher no Mercado de Trabalho e Desigualdade da Renda Domiciliar per capita no Brasil: 1981-2002”. *Nova Economia*. Belo Horizonte. 14 (2) 35-58. maio-agosto.

HOFFMANN, R. and SIMÃO, R. C. S. 2005. “Determinantes do Rendimento das Pessoas Ocupadas em Minas Gerais em 2000: o Limiar no Efeito da Escolaridade e as Diferenças entre Mesorregiões”. *Nova Economia*, v. 15, n. 2, p. 35-62, maio/ago.

HOFFMANN, R. and KASSOUF, A.L. 2005. “Deriving Conditional and Unconditional Marginal Effects in log Earnings Equations Estimated by Heckman’s Procedure”. Applied Economics, Londres, v. 37, n. 11, p. 1303-1311, June.

IBGE, Instituto Brasileiro de Geografia e Estatística. 2007. Available in <<http://www.sidra.ibge.gov.br/bda/tabela/protabl.asp?z=p&o=16&i=P>>.

IPEA-CAIXA. 2007. “As Desigualdades nos Retornos do Ensino Superior no Brasil”. Rio de Janeiro: IPEA.

IPEA. 2009. “Distribuição Funcional da Renda no Brasil: Situação Recente”. Rio de Janeiro: IPEA. Disponível em: <<HTTP://www/ipea.gov.br/sites/>>.

KILLINGSWORTH, M. 1983. **Labor Supply**. Cambridge Surveys of Economic Literature. New York: Cambridge University Press,. P. 493.

KISH, L. 1965. “Survey Sampling”. New York: Wiley.

KOHARA, M. 2008. “The Response of Wives’ Labor Supply to Husbands’ Job Loss”. Osaka School of International Public Policy, Osaka University - OSIPP Discussion Paper: DP-2008-E-007.

KNEIPP, S.M.; Yarandi H.N. 2002. “Complex Sampling Designs and Statistical Issues”, in secondary analysis. West J Nurs Res; 24(5): 552-66.

LIMA, J. R. F. de. 2002. “A Evolução das Rendas e Atividades Não-Agrícolas na Paraíba dos anos 90. 86f. Dissertação (Mestrado em Economia Rural) – Departamento de Economia e Finanças, Universidade Federal da Paraíba, Campina Grande. Disponível em <<http://www.eco.unicamp.br/nea/rurbano/zipados/lima.pdf>>.

LONG, J. S. and FREESE, J. 2006. “Regression Models for Categorical Dependent Variables Using Stata”. Second Edition. College Station Texas.

LOVELL, P. 1994. "Race, Gender and Development in Brazil". *Latin American Research Review*. vol. 29, No 3, pp. 7-35.

MENDONÇA, S. E. A. 2003. "Perspectivas do Mercado de Trabalho para os Próximos Anos". Instituto de Pesquisa Econômica Aplicada – IPEA. Disponível em <http://www.ipea.gov.br/pub/bcmt/mt_22c.pdf>.

MINCER, J. 1974. "Schooling, Experience and Earnings" in New York: National Bureau of Economic Research.

MINCER, J. 1993. "Studies in Human Capital". in Aldershot; Vermont: Edward Elgar, *The Collected essays of Jacob Mincer: v. 1 (Economists of the twentieth century)*.

MORLEY, S. 2001. "The Income Distribution Problem in Latin America and the Caribbean", ECLAC, available in: <<http://www.eclac.cl/cgi-bin>>.

MORETTO, C. F. 2000. "Função Minceriana de Determinação dos Rendimentos Individuais: uma Aplicação do Método de Variáveis Instrumentais". *Teoria e Evidência Econômica*, v. 8, n. 15, p. 47-65.

NEAL, D. A.; JOHNSON, W. R. The Role of Premarket factors in Black-White Wage Differences. *The Journal of Political Economy*, Vol. 104, No. 5 (Oct., 1996), pp.869-895.

NEY, M. G. and HOFFMANN, R. 2003. "Desigualdade de Renda na Agricultura: o Efeito da Posse da Terra", *Economia*, v. 4, n. 1, p. 113-152. NPEC, Niterói, jan./jun..

OLIVEIRA, E. L. de; OLIVEIRA, A. M. H. C. de. 2006. "O Efeito Trabalhador Adicional para Filhos no Brasil". In: XV Encontro Nacional de Estudos Populacionais - ABEP, Caxambú – MG – Brasil, de 18 a 22 de setembro.

PESQUISA NACIONAL POR AMOSTRA DE DOMICÍLIOS. 2007. Brasil. Rio de Janeiro: IBGE, v. 27.

PETTERSEN, T, MORGAN, L. 1995. “Separate and unequal: Occupation-Establishment Sex Segregation and Gender Wage Gap”. In: American Journal of Sociology. Chicago, vol. 101, No. 2.

PNUD. Relatório do Desenvolvimento Humano 2005. Disponível em:
< <http://www.pnud.org.br/rdh/>>. Acesso em: Fev. 2009.

POCHMANN, M. 2001. **O Emprego na Globalização**. São Paulo: Boitempo Editorial.

POCHMANN, M. 2006. “Mercado Geral de Trabalho: O que a de Novo no Brasil?” In: Parcerias Estratégicas, Edição Especial: Análise sobre a Pesquisa Nacional por Amostra de Domicílios (PNAD 2004). Brasília, N.22, pág. 121-144.

POSTHUMA, A. C.; LOMBARDI, M.R. 1996. “Gênero e Exclusão no Novo Paradigma Produtivo”. Caxambu: XX Encontro Anual de ANPOCS.

RAMOS, L. 2007. “O Desempenho Recente do Mercado de Trabalho Brasileiro: Tendências, Fatos Estilizados e Padrões Espaciais”, In Rio de Janeiro: IPEA, Texto para Discussão 1255.

RAMOS, L. and VIEIRA, M. L. 2001. “Desigualdade de Rendimentos no Brasil nas Décadas de 80 e 90: Evolução e Principais Determinantes”, In Rio de Janeiro: IPEA, Texto para Discussão 803.

RESENDE, M. and WYLLIE, R. “Retornos para educação no Brasil: Evidências Empíricas Adicionais”, Economia Aplicada, Ribeirão Preto, v. 10, n. 3, 2006.

RIBEIRO, C. A. C. 2006. “Um Panorama das Desigualdades na América Latina”, Rio de Janeiro, IUPERJ/UCAM. available in < <http://observatorio.iuperj.br/>>.

ROCHA, S. 2002. “A Investigação do Rendimento na PNAD – Comentários e Sugestões à pesquisa nos anos 2000”, in Rio de Janeiro: IPEA, Texto para Discussão 899. Agosto.

SANTOS, G. C., BASTOS, P. M. A., ROCHA, L. E. V. 2008. “Determinantes da Renda do Trabalho no Brasil no Período de 2002 a 2006”, in XLVI Congresso da SOBER, Rio Branco. XLVI Congresso da SOBER.

SCORZAFAVE, L. G., MENEZES FILHO, N. A. 2001. “Participação Feminina no Mercado de Trabalho Brasileiro: Evolução e Determinantes”, Pesquisa e Planejamento Econômico, Rio de Janeiro, v.31, n.3, p. 441-477.

SERNEELS, P. 2002. “The Added Worker Effect and Intrahousehold Aspects of Unemployment”. Centre for the Study of African Economies Working Paper Series, Available in: <<http://www.bepress.com/csae/paper173>>.

SILVA, P. L. do N., PESSOA, D. G. C., LILA, M. F. “Análise Estatística de Dados da PNAD: Incorporando a Estrutura do Plano Amostral”, Ciência Saúde Coletiva, vol.7, no.4, p.659-670. 2002 .

SKINNER, C., Holt, D. and Smith, T. 1989. “Analysis of Complex Surveys”, John Wiley & Sons.

SOUZA, I. V., MACHADO, A. F. 2004. “Curva de Rendimentos: uma Análise no Mercado de Trabalho Urbano e Rural no Brasil (1981/ 99)”. RER, Rio de Janeiro, vol. 42, nº 01, p. 35-54, jan/mar.

TEIXEIRA, W. M. 2006. Equações de Rendimentos e a Utilização de Instrumentos para o Problema de Endogeneidade da Educação. Faculdade de Economia, Administração e Contabilidade (FEA). available in <<http://www.teses.usp.br/teses>>.

VARIAN, H. R. 1992. **Microeconomic Analysis**. Third Edition, Norton International: New York. P. 506.

APÊNDICE 1

Resultados Complementares

Tabela A.1.1. Teste de heterocedasticidade (*hettest Stata 10.1*)
para o mercado de trabalho, Brasil 2002.

Variáveis	Chi2(1)	Prob > chi2
Condição de atividade	0.00	1.0000
Mãe ou Pai solteiro	0.00	0.9981
Situação conjugal (casado)	0.00	0.9916
ln(rendimento de todos os trabalhos)	0.00	1.0000
Número de crianças de 0 a 14 anos na família	0.00	1.0000
Rendimento Familiar	0.00	1.0000
Número de Componentes na Fam.	0.00	0.9986
Gênero (Masculino)	0.00	1.0000
Idade (em dezenas de anos)	0.00	1.0000
idade^2 (em dezenas de anos)	0.00	0.9991
Cor preta ou indígena	0.00	.
Cor Amarela ou parda	0.00	1.0000
Educação em anos de estudo	0.00	0.9963
Efeito Limiar da Educação	0.00	1.0000
Região Sudeste	0.00	0.9987
Região Sul	0.00	.
Região Centro Oeste	0.00	0.9939
Região Norte	1.22E+09	0.0000

Fonte: Elaboração do autor a partir dos microdados da PNAD 2002.

Tabela A.1.2. Teste de heterocedasticidade (*hettest Stata 10.1*)
para o mercado de trabalho, Brasil 2007.

Variáveis	Chi2(1)	Prob > chi2
Condição de atividade	0.00	0.9966
Mãe ou Pai solteiro	0.00	0.9970
Situação conjugal (casado)	0.00	1.0000
ln(rendimento de todos os trabalhos)	0.00	.
Número de crianças de 0 a 14 anos na família	0.00	1.0000
Rendimento Familiar	0.00	1.0000
Número de Componentes na Fam.	0.00	1.0000
Gênero (Masculino)	0.00	0.9879
Idade (em dezenas de anos)	0.00	0.9998
idade^2 (em dezenas de anos)	0.00	0.9977
Cor preta ou indígena	0.00	1.0000
Cor Amarela ou parda	0.00	1.0000
Educação em anos de estudo	0.00	0.9987
Efeito Limiar da Educação	0.00	0.9950
Região Sudeste	0.00	.
Região Sul	0.00	0.9881
Região Centro Oeste	0.00	.
Região Norte	8.93E+08	0.0000

Fonte: Elaboração do autor a partir dos microdados da PNAD 2007.

Tabela A.1.3. Equação de participação no mercado de trabalho Brasileiro, 2002¹.

Variáveis explicativas	TODOS			HOMENS			MULHERES		
	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF
constante	-3.348 (0.000)	- -	1.85 -	-2.625 (0.000)	- -	1.91 -	-2.672 (0.000)	- -	1.81 -
Gênero (Masculino)	1.055 (0.000)	0.358 (0.000)	1.78 -	- -	- -	- -	- -	- -	- -
Pai ou Mãe solteira	0.197 (0.000)	0.066 (0.000)	1.60 -	-0.074 (0.03)	-0.017 (0.035)	1.52 -	0.165 (0.000)	0.066 (0.000)	1.36 -
Situação (casado)	-0.050 (0.000)	-0.018 (0.000)	1.66 -	0.207 (0.000)	0.051 (0.000)	1.60 -	-0.338 (0.000)	-0.134 (0.000)	1.48 -
Rendimento Familiar	-5.E-05 (0.000)	-2.E-05 (0.000)	13.09 -	9.E-05 (0.000)	2.E-05 (0.000)	8.26 -	6.E-05 (0.000)	2.E-05 (0.000)	6.08 -
Num. Comp. na Fam.	-3.5E-03 (0.295)	-1.2E-03 (0.295)	2.22 -	-0.11 (0.000)	-0.03 (0.000)	1.69 -	-0.03 (0.000)	-0.01 (0.000)	1.79 -
N. de crianças na Fam.	0.004 (0.324)	0.002 (0.324)	1.73 -	0.193 (0.000)	0.044 (0.000)	1.50 -	-0.0314 (0.000)	-0.012 (0.000)	1.93 -
Idade (dez. de anos)	1.766 (0.000)	- -	1.68 -	2.127 (0.000)	- -	1.80 -	1.532 (0.000)	- -	1.59 -
idade^2 (dez. de anos)	-0.223 (0.000)	- -	1.73 -	-0.267 (0.000)	- -	1.83 -	-0.204 (0.000)	- -	1.62 -
Cor Amarela	0.054 (0.000)	0.019 (0.000)	1.65 -	0.110 (0.000)	0.02 (0.000)	1.41 -	0.123 (0.000)	0.049 (0.000)	1.34 -
Cor preta ou indígena	0.107 (0.000)	0.037 (0.000)	1.52 -	0.114 (0.000)	0.025 (0.000)	1.44 -	0.231 (0.000)	0.091 (0.000)	1.37 -
Educação em anos	0.038 (0.000)	0.014 (0.000)	2.01 -	-0.005 (0.069)	-0.001 (0.069)	2.20 -	0.043 (0.000)	0.017 (0.000)	1.88 -
Efeito Limiar da Edu.	0.130 (0.000)	0.046 (0.000)	2.23 -	-0.029 (0.000)	-0.007 (0.002)	1.99 -	0.057 (0.000)	0.023 (0.000)	1.80 -
Região Sudeste	0.025 (0.061)	0.009 (0.061)	2.87 -	-0.114 (0.000)	-0.026 (0.000)	2.41 -	0.050 (0.001)	0.020 (0.001)	2.24 -
Região Sul	0.170 (0.000)	0.058 (0.000)	2.52 -	0.083 (0.000)	0.019 (0.000)	2.25 -	0.203 (0.000)	0.081 0.000	2.24 -
Região Centro-Oeste	0.105 (0.000)	0.036 (0.000)	2.69 -	0.076 (0.003)	0.017 (0.002)	2.07 -	0.023 (0.289)	0.009 (0.289)	2.33 -
Região Norte	-0.012 (0.572)	-0.004 (0.573)	3.80 -	-0.042 (0.103)	-0.010 (0.11)	2.29 -	-0.007 (0.756)	-0.003 (0.756)	2.49 -
Num. obs	221,199			108,999			112,200		
Pop. Considerada	99,804,804			49,408,764			50,396,040		

Estatística F (Prob) 0.00 0.00 0.00

Fonte: Elaboração do autor a partir dos microdados da PNAD 2002.

¹M.E. significa *Marginal Effect* ou Efeito Marginal. MEFF significa *Misspecification Effect*.

Os valores entre parênteses correspondem ao valor-p dos coeficientes.

Tabela A.1.4. Equação de rendimento, Brasil 2002¹.

Variáveis explicativas	TODOS			HOMENS			MULHERES		
	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF
constante	-3.184 (0.000)	-95.8% (0.000)	2.44 -	-0.926 (0.000)	- 60.38%	2.09 -	-0.324 (0.000)	-27.6% (0.000)	2.35 -
Gênero (Masculino)	0.507 (0.000)	66.10% (0.000)	2.26 -	- -	- -	- -	- -	- -	- -
Pai ou Mãe solteira	0.039 (0.000)	3.97% (0.000)	1.32 -	0.014 (0.466)	1.44% (0.593)	1.35 -	-0.076 (0.000)	-7.33% (0.000)	1.21 -
Situação conj. (casado)	0.045 (0.000)	4.6% (0.000)	1.65 -	0.057 (0.000)	5.8% (0.000)	1.59 -	0.131 (0.000)	14.0% (0.000)	1.69 -
Idade (dez. de anos)	1.137 (0.000)	- -	2.33 -	0.304 (0.000)	- -	2.20 -	0.083* (0.003)	- -	2.14 -
idade^2 (dez. de anos)	-0.124 (0.000)	- -	2.45 -	-0.017 (0.000)	- -	2.32 -	0.013 (0.001)	- -	2.24 -
Cor amarela ou pardas	-0.120 (0.000)	- 11.35%	1.75 -	-0.149 (0.000)	- 13.87%	1.71 -	-0.162 (0.000)	-14.9% (0.000)	1.42 -
Cor preta ou indígena	-0.117 (0.000)	- 11.08%	1.48 -	-0.156 (0.000)	- 14.48%	1.40 -	-0.206 (0.000)	-18.6% (0.000)	1.49 -
Educação (anos)	0.091 (0.000)	9.49% (0.000)	2.88 -	0.087 (0.000)	9.11% (0.000)	2.50 -	0.057 (0.000)	5.84% (0.000)	2.47 -
Efeito Limiar (edu)	0.172 (0.000)	30.04% (0.000)	2.52 -	0.147 (0.000)	26.39% (0.000)	2.08 -	0.152 (0.000)	23.22% (0.000)	1.98 -
Região Sudeste	0.441 (0.000)	55.47% (0.000)	4.36 -	0.437 (0.436)	54.7% (0.548)	4.05 -	0.380 (0.000)	46.21% (0.000)	3.13 -
Região Sul	0.399 (0.000)	49.02% (0.000)	4.05 -	0.330 (0.000)	39.08% (0.000)	4.43 -	0.271 (0.000)	31.17% (0.000)	2.62 -
Região Centro-Oeste	0.438 (0.000)	54.90% (0.000)	3.92 -	0.393 (0.000)	48.18% (0.000)	3.31 -	0.374 (0.000)	45.35% (0.000)	2.62 -
Região Norte	0.313 (0.000)	36.80% (0.000)	4.45 -	0.315 (0.000)	37.05% (0.000)	3.43 -	0.293 (0.000)	34.10% (0.000)	2.64 -
Lambda	0.691 (0.000)	- -	5.10 -	-0.770 (0.000)	- -	4.88 -	-0.806 (0.000)	- -	3.64 -

Num. obs	221,199	108,999	112,200
Pop. Considerada	99,804,804	49,408,764	50,396,040
Estatística F (Prob)	0.00	0.00	0.00

Fonte: Elaboração do autor a partir dos microdados da PNAD 2002.

¹ M.E. significa *Marginal Effect* ou Efeito Marginal. MEFF significa *Misspecification Effect*.

Os valores entre parênteses correspondem ao valor-p dos coeficientes.

Tabela A.1.5. Equação de participação no mercado de trabalho Brasileiro, 2007¹.

Variáveis explicativas	TODOS			HOMENS			MULHERES		
	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF
constante	-3.314 (0.000)	- -	2.60 -	-2.940 (0.000)	- -	2.11 -	-2.858 (0.000)	- -	1.82 -
Gênero (Masculino)	0.945 (0.000)	0.314 (0.000)	2.70 -	- -	- -	- -	- -	- -	- -
Pai ou Mãe solteira	0.184 (0.000)	0.060 (0.000)	1.43 -	0.025 (0.428)	0.006 (0.422)	1.41 -	0.182 (0.000)	0.071 (0.000)	1.34 -
Situação conj. (casado)	-0.050 (0.000)	-0.017 (0.000)	1.74 -	0.234 (0.000)	0.057 (0.000)	1.49 -	-0.273 (0.000)	-0.107 (0.000)	1.55 -
Rendimento Familiar	5.E-05 (0.000)	2.E-05 (0.000)	34.58 -	9.E-05 (0.000)	2.E-5 (0.000)	18.92 -	5.E-05 (0.000)	2.E-05 (0.000)	34.08 -
Num. Comp. na Fam.	-7.1E-02 (0.000)	-2.E-02 (0.000)	3.11 -	-0.12 (0.000)	-0.03 (0.000)	2.11 -	-0.04 (0.000)	-0.02 (0.000)	2.61 -
N. de crianc. na Fam.	0.046 (0.000)	0.016 (0.000)	2.33 -	0.187 (0.000)	0.043 (0.000)	1.58 -	-0.0411 (0.000)	-0.0163 (0.000)	2.77 -
Idade (dez. de anos)	1.851 (0.000)	0.63 (0.000)	1.90 -	2.195 (0.000)	0.50 (0.000)	1.69 -	1.608 (0.000)	0.64 (0.000)	1.62 -
idade^2 (dez. anos)	-0.237 (0.000)	-0.08 (0.000)	1.85 -	-0.275 (0.000)	-0.06 (0.000)	1.64 -	-0.213 (0.000)	-0.08 (0.000)	1.65 -
Cor Amarela pardas	0.101 (0.000)	0.034 (0.000)	1.47 -	0.105 (0.000)	0.02 (0.000)	1.46 -	0.104 (0.000)	0.041 (0.000)	1.41 -
Cor preta ou indígena	0.149 (0.000)	0.049 (0.000)	1.58 -	0.070 (0.000)	0.016 (0.000)	1.35 -	0.204 (0.000)	0.080 (0.000)	1.55 -
Educação em anos	0.037 (0.000)	0.013 (0.000)	2.40 -	0.021 (0.000)	0.005 (0.000)	2.24 -	0.049 (0.000)	0.020 (0.000)	1.81 -
Efeito Limiar da Edu.	0.039 (0.000)	0.013 (0.000)	4.24 -	-0.052 (0.000)	-0.012 (0.000)	2.78 -	0.057 (0.000)	0.023 (0.000)	3.60 -
Região Sudeste	0.033 (0.027)	0.011 (0.026)	3.66 -	-0.082 (0.000)	-0.019 (0.000)	2.84 -	0.119 (0.000)	0.047 (0.000)	2.76 -
Região Sul	0.166 (0.000)	0.055 (0.000)	3.87 -	0.056 (0.028)	0.012 (0.025)	2.54 -	0.245 (0.000)	0.096 (0.000)	2.83 -
Região Centro-Oeste	0.066	0.022	3.26	0.042	0.009	2.36	0.085	0.034	2.66

	(0.001)	(0.000)	-	(0.105)	(0.097)	-	(0.000)	(0.000)	-
Região Norte	0.008	0.003	4.05	0.047	0.011	3.26	-0.023	-0.009	2.88
	(0.694)	(0.693)	-	(0.092)	(0.085)	-	(0.286)	(0.286)	-
Num. obs	238,449			118,054			120,395		
Pop. Considerada	113,490,535			56,275,537			57,214,998		
Estatística F (Prob)	0.00			0.00			0.00		

Fonte: Elaboração do autor a partir dos microdados da PNAD 2007.

¹M.E. significa *Marginal Effect* ou Efeito Marginal. MEFF significa *Misspecification Effect*.

Os valores entre parênteses correspondem ao valor-p dos coeficientes.

Tabela A.1.6. Equação de rendimento utilizando o procedimento de Heckman, Brasil 2007¹.

Variáveis explicativas	TODOS			HOMENS			MULHERES		
	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF
constante	-0.443	-35.80%	9.12	-0.360	-30.25%	5.09	0.054	5.55%	6.50
	(0.000)	(0.000)	-	(0.000)	0.000	-	(0.642)	(0.642)	-
Gênero (Masculino)	0.129	13.78%	8.45	-	-	-	-	-	-
	(0.000)	(0.000)	-	-	-	-	-	-	-
Pai ou Mãe solteira	-0.037	-3.66%	1.63	0.003	0.32%	1.26	-0.049	-4.78%	1.34
	(0.000)	(0.000)	-	(0.857)	(0.857)	-	(0.000)	(0.000)	-
Situação (casado)	0.074	7.7%	1.64	0.059	6.1%	1.61	0.093	9.8%	2.18
	(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-
Idade (dez. de anos)	0.293	-	7.86	0.291	-	5.19	0.131	-	5.43
	(0.000)	-	-	(0.000)	-	-	0.003	-	-
Idade^2 (dez. anos)	-0.016	-	8.11	-0.016	-	5.40	0.004	-	5.63
	(0.000)	-	-	(0.000)	-	-	0.461	-	-
Cor amarela ou pardas	-0.134	-12.55%	1.96	-0.139	-12.95%	1.69	-0.129	-12.13%	1.42
	(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-
Cor preta ou indígena	-0.129	-12.07%	1.68	-0.138	-12.92%	1.42	-0.119	-11.22%	1.39
	(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-
Educação (anos)	0.063	6.45%	3.71	0.069	7.11%	2.84	0.046	4.74%	2.97
	(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-
Efeito Limiar (edu)	0.140	22.45%	2.94	0.145	23.79%	2.22	0.149	21.56%	1.94
	(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-
Região Sudeste	0.389	47.48%	7.06	0.412	50.98%	5.57	0.339	40.33%	3.74
	(0.000)	(0.000)	-	(0.411)	(0.411)	-	(0.000)	(0.000)	-
Região Sul	0.348	41.59%	5.46	0.372	45.03%	4.22	0.293	34.06%	3.37
	(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-
Região Centro-Oeste	0.415	51.49%	5.99	0.435	54.43%	4.65	0.371	44.97%	3.08
	(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-
Região Norte	0.306	35.83%	8.36	0.318	37.38%	6.30	0.282	32.57%	4.43
	(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-
Lambda	-0.497	-	17.04	-0.597	-	14.64	-0.579	-	11.87

	(0.000)	-	-	(0.000)	-	-	(0.000)	-	-
Num. obs	238,449			118,054			120,395		
Pop. Considerada	113,490,535			56,275,537			57,214,998		
Estatística F (Prob)	0.00			0.00			0.00		

Fonte: Elaboração do autor a partir dos microdados da PNAD 2007.

¹ M.E. significa *Marginal Effect* ou Efeito Marginal. MEFF significa *Misspecification Effect*.

Os valores entre parênteses correspondem ao valor-p dos coeficientes.

Tabela A.1.7. Equação de participação no mercado de trabalho do meio rural, Brasil 2002¹.

Variáveis explicativas	TODOS			HOMENS			MULHERES		
	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF
Constante	-3.093 (0.000)	-	2.34	-1.958 (0.000)	-	2.00	-2.634 (0.000)	-	1.84
Gênero (Masculino)	1.505 (0.000)	0.490 (0.000)	2.57	-	-	-	-	-	-
Pai ou Mãe solteira	0.247 (0.000)	0.073 (0.000)	1.57	0.078 (0.469)	0.012 (0.447)	1.29	0.049 (0.452)	0.019 (0.453)	1.05
Situação (casado)	-0.024 (0.542)	-0.008 (0.539)	1.71	0.309 (0.000)	0.059 (0.000)	1.45	-0.421 (0.000)	-0.165 (0.000)	1.25
Rendimento Familiar	2.E-04 (0.000)	7.E-05 (0.000)	13.3	4.E-04 (0.000)	6.E-05 (0.000)	18.60	2.E-04 (0.000)	7.E-05 (0.000)	6.24
Num. Comp. na Fam.	-7.1E-02 (0.000)	-2.3E-02 (0.000)	2.15	-0.13 (0.000)	-0.02 (0.000)	1.92	-0.02 (0.057)	-0.01 (0.056)	1.59
N. de crianc. na Fam.	0.089 (0.000)	0.028 (0.000)	1.57	0.191 (0.000)	0.031 (0.000)	1.22	0.0136 (0.389)	0.00522 (0.388)	1.65
Idade (dez. de anos)	1.513 (0.000)	0.48 (0.000)	2.29	1.669 (0.000)	0.27 (0.000)	2.35	1.385 (0.000)	0.53 (0.000)	1.54
idade^2 (dez. de anos)	-0.180 (0.000)	-0.01 (0.000)	2.13	-0.197 (0.000)	-0.03 (0.000)	2.24	-0.172 (0.000)	-0.07 (0.000)	1.42
Cor Amarela	0.103 (0.001)	0.033 (0.001)	1.80	0.100 (0.014)	0.02 (0.015)	1.38	0.119 (0.004)	0.046 (0.003)	1.71
Cor preta ou indígena	0.266 (0.000)	0.078 (0.000)	1.72	0.314 (0.001)	0.043 (0.000)	1.64	0.281 (0.000)	0.111 (0.000)	1.49
Educação (anos)	0.018 (0.015)	0.006 (0.014)	1.42	0.004 (0.770)	0.001 (0.770)	1.82	0.030 (0.008)	0.012 (0.007)	1.65
Efeito Limiar (edu)	0.017 (0.179)	0.005 (0.176)	1.52	-0.065 (0.002)	-0.011 (0.002)	1.63	0.035 (0.026)	0.013 (0.026)	1.28
Região Sudeste	-0.138 (0.004)	-0.045 (0.004)	3.06	-0.240 (0.000)	-0.043 (0.000)	2.32	-0.069 (0.185)	-0.026 (0.182)	2.03
Região Sul	0.114 (0.032)	0.036 (0.027)	2.52	0.005 (0.951)	0.001 (0.951)	2.35	0.174 (0.009)	0.068 (0.009)	2.15

Região Centro-Oeste	-0.096 (0.088)	-0.032 (0.094)	2.64 -	0.016 (0.822)	0.003 (0.819)	1.68 -	-0.218 (0.005)	-0.081 (0.003)	2.61 -
Região Norte	0.125 (0.246)	0.038 (0.225)	2.87 -	0.180 (0.234)	0.026 (0.174)	2.52 -	0.066 (0.555)	0.026 (0.557)	1.59 -
Num. obs	24,462			14,647			9,815		
Pop. Considerada	12,178,583			7,306,557			4,872,026		
Estatística F (Prob)	0.00			0.00			0.00		

Fonte: Elaboração do autor a partir dos microdados da PNAD 2002.

¹M.E. significa *Marginal Effect* ou Efeito Marginal. MEFF significa *Misspecification Effect*.

Os valores entre parênteses correspondem ao valor-p dos coeficientes.

Tabela A.1.8. Equação de participação no mercado de trabalho do meio rural, Brasil 2007¹.

Variáveis explicativas	TODOS			HOMENS			MULHERES		
	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF
constante	-3.969 (0.000)	-	1.83 -	-2.828 (0.000)	-	2.35 -	-3.597 (0.000)	-	1.45 -
Gênero (Masculino)	1.518 (0.000)	0.495 (0.000)	2.20 -	-	-	-	-	-	-
Pai ou Mãe solteira	0.271 (0.000)	0.081 (0.000)	1.60 -	0.097 (0.311)	0.015 (0.281)	1.59 -	0.251 (0.000)	0.099 (0.000)	1.36 -
Situação conj. (casado)	0.052 (0.131)	0.017 (0.134)	1.65 -	0.298 (0.000)	0.057 (0.000)	1.53 -	-0.211 (0.000)	-0.082 (0.000)	1.51 -
Rendimento Familiar	3.E-05 (0.21)	1.E-05 (0.211)	56.22 -	9.E-05 (0.288)	1.E-05 (0.307)	80.94 -	2.E-05 (0.358)	7.E-06 (0.357)	27.47 -
Num. Comp. na Fam.	-4.5E-02 (0.000)	-1.5E-02 (0.000)	1.59 -	-0.10 (0.000)	-0.02 (0.000)	2.35 -	0.01 (0.534)	0.00 (0.534)	1.43 -
N. de crianças na Fam.	0.035 (0.000)	0.011 (0.000)	1.05 -	0.128 (0.000)	0.022 (0.000)	1.42 -	-0.0451 (0.001)	-0.0175 (0.001)	1.13 -
Idade (dez. de anos)	1.910 (0.000)	0.62 (0.000)	1.61 -	2.084 (0.000)	0.35 (0.000)	2.90 -	1.749 (0.000)	0.68 (0.000)	1.30 -
idade^2 (dez. de anos)	-0.226 (0.000)	-0.07 (0.000)	1.57 -	-0.246 (0.000)	-0.04 (0.000)	2.72 -	-0.211 (0.000)	-0.08 (0.000)	1.28 -
Cor Amarela ou pardas	0.011 (0.643)	0.004 (0.643)	1.36 -	0.030 (0.434)	0.01 (0.437)	1.43 -	0.000 (0.994)	0.000 (0.994)	1.34 -
Cor preta ou indígena	0.032 (0.407)	0.010 (0.402)	1.14 -	0.004 (0.942)	0.001 (0.942)	0.99 -	0.051 (0.369)	0.020 (0.369)	1.31 -
Educação (anos)	0.039 (0.000)	0.013 (0.000)	1.39 -	0.058 (0.000)	0.010 (0.000)	1.43 -	0.041 (0.000)	0.016 (0.000)	1.28 -
Efeito Limiar (edu)	0.036 (0.001)	0.012 (0.001)	1.48 -	-0.074 (0.000)	-0.012 (0.000)	1.82 -	0.064 (0.000)	0.025 (0.000)	1.23 -
Região Sudeste	0.031	0.010	2.09	-0.029	-0.005	4.38	0.071	0.027	1.65

	(0.416)	(0.414)	-	(0.728)	(0.733)	-	(0.118)	(0.118)	-
Região Sul	0.160	0.050	3.64	0.004	0.001	4.37	0.249	0.098	2.78
	(0.007)	(0.005)	-	(0.964)	(0.964)	-	(0.000)	(0.000)	-
Região Centro-Oeste	-0.008	-0.003	2.17	0.014	0.002	3.92	-0.037	-0.014	2.12
	(0.865)	(0.866)	-	(0.885)	(0.884)	-	(0.566)	(0.564)	-
Região Norte	-0.026	-0.008	3.09	0.054	0.009	3.88	-0.097	-0.037	2.53
	(0.566)	(0.568)	-	(0.479)	(0.46)	-	(0.091)	(0.087)	-
Num. obs	29,994			17,535			12,459		
Pop. Considerada	14,701,310			8,558,957			6,142,353		
Estatística F (Prob)	0.00			0.00			0.00		

Fonte: Elaboração do autor a partir dos microdados da PNAD 2007.

¹M.E. significa *Marginal Effect* ou Efeito Marginal. MEFF significa *Misspecification Effect*.

Os valores entre parênteses correspondem ao valor-p dos coeficientes.

Tabela A.1.9. Equação de rendimento para o meio rural, Brasil 2002¹.

Variáveis explicativas	TODOS			HOMENS			MULHERES		
	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF
Constante	-0.483	-38.33%	2.65	-0.511	-40.03%	1.75	-0.180	-16.48%	1.92
	(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.456)	(0.577)	-
Gênero (Masculino)	-0.233	-20.79%	3.09	-	-	-	-	-	-
	(0.000)	(0.000)	-	-	-	-	-	-	-
Pai ou Mãe solteira	-0.071	-6.81%	1.37	0.007	0.71%	1.30	-0.012	-1.24%	1.13
	(0.085)	(0.088)	-	(0.908)	(0.908)	-	(0.843)	(0.843)	-
Situação conj. (casado)	0.082	8.5%	1.55	0.053	5.5%	1.46	0.226	25.3%	1.33
	(0.000)	(0.000)	-	(0.033)	(0.033)	-	(0.000)	(0.000)	-
Idade (dez. de anos)	0.139	-	1.96	0.060	-	1.71	-0.060	-	1.73
	(0.000)	-	-	(0.112)	-	-	(0.542)	-	-
idade^2 (dez. de anos)	-0.007	-	1.94	0.001	-	1.70	0.026	-	1.72
	0.164	-	-	(0.807)	-	-	(0.034)	-	-
Cor amarela ou pardas	-0.122	-11.48%	2.33	-0.117	-11.04%	1.80	-0.141	-13.13%	1.86
	(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.002)	(0.002)	-
Cor preta ou indígena	-0.118	-11.14%	1.82	-0.086	-8.23%	1.76	-0.233	-20.82%	1.37
	(0.002)	(0.002)	-	(0.039)	(0.039)	-	(0.002)	(0.002)	-
Educação (anos)	0.071	7.40%	1.75	0.070	7.26%	1.59	0.071	7.40%	1.98
	(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-
Efeito Limiar (edu)	0.017	9.22%	1.71	0.029	10.41%	1.60	-0.001	7.25%	1.61
	(0.038)	(0.038)	-	(0.002)	(0.002)	-	(0.934)	(0.934)	-
Região Sudeste	0.387	47.25%	5.50	0.378	45.90%	4.28	0.407	50.30%	2.67
	(0.000)	(0.000)	-	(0.377)	(0.458)	-	(0.000)	(0.000)	-
Região Sul	0.397	48.80%	5.11	0.408	50.37%	4.21	0.328	38.85%	2.33

	(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-
Região Centro-Oeste	0.486	62.66%	4.93	0.425	52.99%	3.83	0.654	92.34%	3.07
	(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-
Região Norte	0.100	10.47%	6.68	0.086	8.93%	5.04	0.107	11.24%	6.08
	(0.306)	(0.357)	-	(0.362)	(0.436)	-	(0.639)	(0.639)	-
Lambda	-0.830	-	4.69	-1.087	-	5.18	-0.982	-	2.94
	(0.000)	-	-	(0.000)	-	-	(0.000)	-	-
Num. obs	24,462			14,647			9,815		
Pop. Considerada	12,178,583			7,306,557			4,872,026		
Estatística F (Prob)	0.00			0.00			0.00		

Fonte: Elaboração do autor a partir dos microdados da PNAD 2002.

¹M.E. significa *Marginal Effect* ou Efeito Marginal. MEFF significa *Misspecification Effect*.

Os valores entre parênteses correspondem ao valor-p dos coeficientes.

Tabela A.1.10. Equação de rendimento para o meio rural, Brasil 2007¹.

Variáveis explicativas	TODOS			HOMENS			MULHERES		
	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF	COEF.	M.E.	MEFF
constante	-0.480	-38.14%	6.10	-0.129	-12.08%	14.49	-1.015	-63.78%	4.85
	(0.076)	(0.076)	-	(0.663)	(0.663)	-	(0.084)	(0.084)	-
Gênero (Masculino)	0.011	1.07%	6.50	-	-	-	-	-	-
	(0.886)	(0.886)	-	-	-	-	-	-	-
Pai ou Mãe solteira	-0.129	-12.10%	2.04	-0.059	-5.72%	1.47	-0.158	-14.65%	1.49
	(0.003)	(0.003)	-	(0.272)	(0.272)	-	(0.006)	(0.006)	-
Situação conj. (casado)	0.027	2.7%	1.44	0.014	1.4%	2.01	0.026	2.6%	1.47
	(0.159)	(0.159)	-	(0.584)	(0.584)	-	(0.55)	(0.55)	-
Idade (dez. de anos)	0.301	-	5.27	0.145	-	13.88	0.531	-	4.33
	(0.001)	-	-	(0.26)	-	-	(0.012)	-	-
idade^2 (dez. de anos)	-0.025	-	5.27	-0.008	-	13.85	-0.049	-	4.28
	(0.017)	-	-	(0.603)	-	-	(0.055)	-	-
Cor amarela ou pardas	-0.101	-9.58%	1.49	-0.098	-9.37%	1.41	-0.112	-10.56%	1.01
	(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-
Cor preta ou indígena	-0.061	-5.89%	2.06	-0.079	-7.58%	1.79	-0.010	-0.96%	1.22
	(0.063)	(0.06)	-	(0.027)	(0.027)	-	(0.847)	(0.847)	-
Educação (anos)	0.051	5.24%	2.09	0.051	5.25%	1.93	0.043	4.36%	2.26
	(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.001)	(0.001)	-
Efeito Limiar (edu)	0.032	8.64%	1.71	0.030	8.46%	1.45	0.053	5.47%	1.74
	(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-
Região Sudeste	0.449	56.74%	4.17	0.462	58.80%	3.63	0.418	51.90%	2.04
	(0.000)	(0.000)	-	(0.46)	(0.58)	-	(0.000)	(0.000)	-
Região Sul	0.547	72.81%	4.30	0.558	74.63%	3.68	0.535	70.74%	2.46
	(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-

Região Centro-Oeste	0.532	70.30%	4.22	0.526	69.28%	3.36	0.530	69.93%	2.11
	(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-
Região Norte	0.527	69.40%	8.24	0.522	68.62%	8.68	0.512	66.81%	3.66
	(0.000)	(0.000)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-
Lambda	-0.325	-	8.34	-0.530	-	35.05	-0.196	-	5.74
	(0.024)	-	-	(0.045)	-	-	(0.408)	-	-
Num. obs	29,994			17,535			12,459		
Pop. Considerada	14,701,310			8,558,957			6,142,353		
Estatística F (Prob)	0.00			0.00			0.00		

Fonte: Elaboração do autor a partir dos microdados da PNAD 2007.

¹M.E. significa *Marginal Effect* ou Efeito Marginal. MEFF significa *Misspecification Effect*.

Os valores entre parênteses correspondem ao valor-p dos coeficientes.

APÊNDICE 2

Fonte de dados e especificidades da PNAD

1.1. Características Básicas

A fonte de dados do presente trabalho advém dos microdados da Pesquisa Nacional por Amostra de Domicílios (PNAD). Devido ao tratamento não trivial que se deve dedicar ao manejo destes dados, optou-se por abordar nesta seção algumas questões relacionadas à aplicação correta dos métodos estatísticos, além de evidenciar a magnitude e importância da PNAD para o desenvolvimento de pesquisas sobre a população brasileira.

A Pesquisa Nacional por Amostra de Domicílios (PNAD) é fornecida anualmente pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Os microdados das PNADs consistem em dados individuais das principais características socioeconômicas dos indivíduos e famílias, umas de caráter permanente, como as características gerais da população, de educação, trabalho, rendimento e habitação, e outras com periodicidade variável, como as características sobre migração, fecundidade, nupcialidade, saúde, nutrição e outros temas que são incluídos no sistema de acordo com as necessidades de informação para o País (IBGE, 2007). O período de referência das PNADs nos anos noventa e 2000 foi o mês de setembro de cada ano. Além da abrangência temática da PNAD, o tamanho da amostra também consiste em importante fator para garantir a consistência dos resultados, principalmente em termos de inferência estatística. A Tabela 1 a seguir exemplifica a amplitude da pesquisa para alguns anos. As mais de 300.000 pessoas investigadas nas PNADs, em cada ano, consiste em uma das maiores amostras encontradas em pesquisas sobre o Brasil.

Tabela A.2.1. Características da amostra da PNAD.

PNAD	Mês de referência	Total de pessoas	Total de domicílios
1995	Setembro	334.263	102.787
1996	Setembro	331.263	105.059
1997	Setembro	346.269	109.541
1998	Setembro	344.975	112.434
1999	Setembro	352.393	115.654
2001	Setembro	378.837	126.858
2002	Setembro	385.431	129.705
2003	Setembro	384.834	133.255
2004	Setembro	399.354	139.157
2005	Setembro	408.148	142.471
2006	Setembro	410.241	145.547
2007	Setembro	399.964	147.851

Fonte: PNADs 1995-2007.

1.2. O Plano Amostral

Diversos são os métodos probabilísticos de amostragem que podem ser empregados em uma determinada pesquisa. É de se esperar que a escolha do método esteja condicionada a alguns fatores, como abrangência da mesma, bem como pelas restrições impostas ao instituto de pesquisa ou próprio pesquisador, sendo um exemplo o orçamento disponível para realização da pesquisa, entre outros fatores. Segundo Cochran (1977), dentre os métodos mais utilizados destacam-se a amostra aleatória simples, a amostragem estratificada e a amostragem por conglomerado.

A Pesquisa Nacional por Amostra de Domicílios (PNAD) adota um plano amostral estratificado e conglomerado com um, dois ou três estágios de seleção, dependendo do estrato (Silva *et al.*, 2002). Em outras palavras, isto significa que o desenho amostral da pesquisa não é o de uma amostra aleatória simples com reposição (AASC) e o tratamento incorreto do plano amostral geraria resultados viesados para erro-padrão, variância, quantil, percentil, análise de regressão, entre outros. Apesar destas particularidades, vários trabalhos que utilizam a PNAD desconsideram sua forma de amostragem, Silva *et al.* (2002) afirmam que muitas vezes as análises são feitas por pesquisadores que trabalham fora da agência produtora dos dados e freqüentemente utilizam para a modelagem pacotes estatísticos que se baseiam em hipóteses válidas somente quando os dados são obtidos através de amostras aleatórias simples com reposição (AASC).

A PNAD na verdade caracteriza-se por ser realizada a partir de um plano amostral complexo, uma vez que apresenta características que a definem como tal, por exemplo: estratificação, conglomeração, probabilidades desiguais de seleção em um ou mais estágios e calibração da amostra através de pesos amostrais, conforme destacam Silva *et al.* (2002). De acordo com Kneipp e Yarandi (2002), a maior parte das pesquisas conduzidas por institutos governamentais não utilizam o método de amostra aleatória simples, mas um desenho de amostragem complexa. Uma possível explicação consiste nos limites de custos e as restrições de tempo associadas ao montante de dados dos indivíduos em todo território nacional. Esta justificativa é bastante plausível para o caso brasileiro, uma vez que o território nacional abrange uma área de 8.514.876 Km², com um total de 191.790.900 habitantes (IBGE, 2007).

1.2.1. Construção do desenho amostral da PNAD⁴⁴

A PNAD é estratificada em duas etapas. Inicialmente realiza-se uma estratificação que divide o país em 36 estratos naturais, sendo vinte e sete correspondendo às unidades da federação e os nove estratos remanescentes referentes aos municípios das regiões metropolitanas com sede na capital, ou seja, as Regiões Metropolitanas de Belém (PA), Fortaleza (CE), Recife (PE), Salvador (BA), Belo Horizonte (MG), Rio de Janeiro (RJ), São Paulo (SP), Curitiba (PR) e Porto Alegre (RS).

Nos estratos das regiões metropolitanas realiza-se uma nova estratificação por municípios, sendo agrupados (conglomerados) em dois estágios, ou seja, as unidades primárias de amostragem (UPA) são os setores censitários e as unidades secundárias de amostragem (USA) são os domicílios. Os setores são selecionados utilizando-se uma amostragem sistemática com probabilidade proporcional ao tamanho (PPT), onde o tamanho do setor é determinado pelo total de domicílios obtido através do último censo⁴⁵.

Para os demais 27 estratos⁴⁶ o agrupamento é realizado em três estágios. As UPAs são agora os municípios, as USAs são os setores censitários e as unidades terciárias de amostragem (UTA) são os domicílios. Para estes estratos naturais, alguns municípios,

⁴⁴ Esta seção consiste em uma concisa exemplificação da construção do desenho amostral da PNAD e é baseada principalmente no trabalho de Silva *et al.* (2002).

⁴⁵ Para as PNADs da década de 1990, o censo de referência é o de 1991 e para os dados do ano 2000 a referência é o censo do ano 2000.

⁴⁶ Estes estratos são determinados a partir dos municípios que fazem parte das regiões metropolitanas bem como por municípios situados em unidades da federação sem região metropolitana.

conhecidos como auto-representativos devido ao tamanho da população⁴⁷, são incluídos com certeza na amostra. Por outro lado, devem ser considerados também os municípios não auto-representativos, com probabilidade de inclusão proporcional à sua população na época do último censo demográfico. Os municípios desse grupo são conglomerados em estratos por tamanho e proximidade geográfica, buscando formar estratos com população total aproximadamente igual.

Finalizando o processo de seleção da amostra da PNAD, têm-se as unidades domiciliares selecionadas a partir do Cadastro de Projetos de Novas Construções, que inclui projetos habitacionais com mais de 30 domicílios e que porventura surgiram após o último censo realizado. Essas novas construções são estratificadas por municípios, sendo que nesses estratos o plano amostral é conglomerado em um estágio, sendo a unidade primária de amostragem os próprios domicílios, e a seleção é realizada por meio de amostragem sistemática simples⁴⁸. A Figura 1 abaixo apresenta esquematicamente o plano amostral da PNAD.

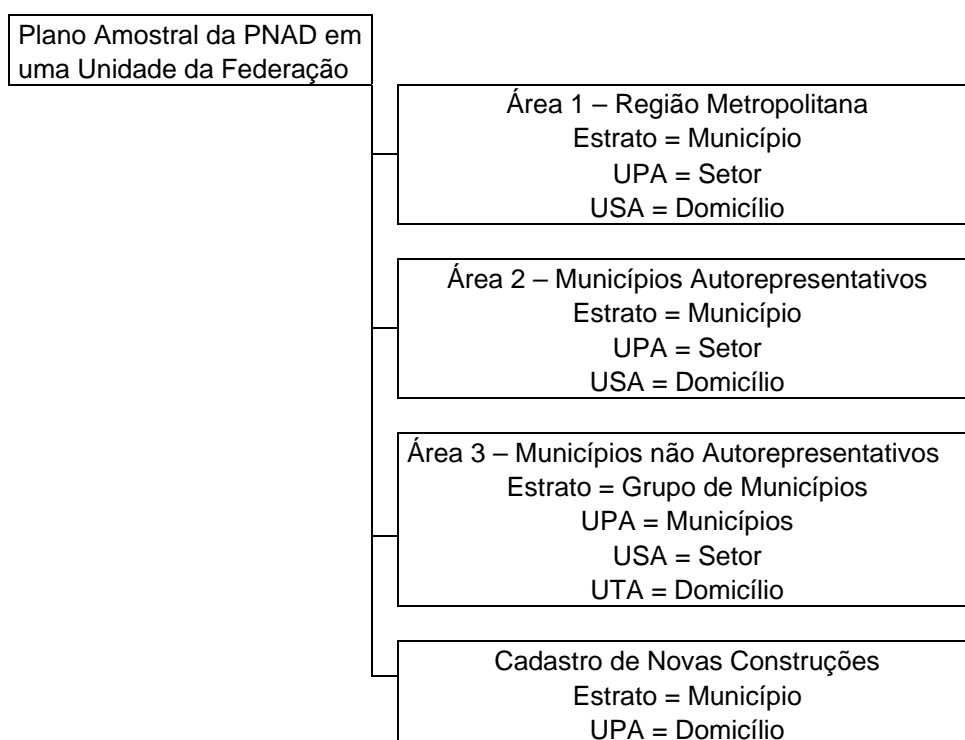


Figura A.2.1. Construção do plano amostral da PNAD.

Fonte: Silva *et al.* (2002).

⁴⁷ Um município é dito auto-representativo se apresentar população que seja maior que 80% do tamanho do estrato estabelecido para Unidade da Federação em questão com base no último censo (IBGE, 2007).

⁴⁸ Deve-se ressaltar que, ao longo de toda a década, são mantidos constantes na PNAD os setores selecionados nos municípios das regiões metropolitanas, auto-representativos e não auto-representativos no início da década.

1.2.2. Análise Estatística em Amostras Complexas

Quando se trata de análise estatística, as informações coletadas a partir de pesquisas com características que definem um plano amostral complexo, como a PNAD não podem ser tratadas como se fossem observações de uma amostragem aleatória simples, ou seja, observações independentes e identicamente distribuídas (IID).

Basicamente os estudos realizados com pesquisas como a PNAD, apresentam dois grupos de análise descritiva. A primeira diz respeito às estimativas pontuais como média, taxas, proporções e totais. Estas são influenciadas apenas pelos pesos amostrais. Outro grupo tem como objetivo a estimação de medidas de dispersão, como variância e desvio-padrão, com o intuito de se calcular intervalos de confiança e realizar testes de hipóteses. Contudo, em pesquisas com desenho amostral complexo a estimação da variância recebe influência não apenas dos pesos amostrais, mas também e conjuntamente da estratificação e conglomeração. Sendo assim, é imprescindível a consideração do desenho amostral para se obter estimativas não viciadas das variâncias da amostra.

Alguns autores desenvolveram métodos que possibilitam verificar o impacto da incorporação do desenho amostral. Kish (1965) desenvolveu um método para avaliar o efeito do plano amostral (EPA) conhecido como DEFF (*design-effect*). O método é definido como a razão da variância obtida considerando o plano amostral e a variância obtida ignorando o plano amostral (isto é, a variância estimada como se a amostra fosse AASC). Valores do DEFF afastados de 1 indicam que ignorar o plano amostral na estimação da variância leva a estimativas viesadas e incorretas. Valores grandes (> 1) do DEFF indicam que o estimador “ingênuo” da variância obtido ignorando o plano amostral complexo leva a subestimar a variância verdadeira do estimador. Valores abaixo de um para o DEFF significam que a variância obtida desconsiderando o EPA superestima a verdadeira variância do estimador e, caso a estatística DEFF seja igual a um, não existe diferença entre as estimativas da variância, segundo Silva *et al.* (2002).

Outra alternativa é a estatística MEFF (*misspecification effect*), desenvolvida por Skinner, Holt e Smith (1989), que compara a estimativa da variância do parâmetro obtida considerando o plano amostral com outra, do mesmo modelo, desconsiderando peso, conglomerado e estratificação⁴⁹.

⁴⁹ A análise do MEFF é similar ao DEFF.

Por fim, dado que um dos objetivos das análises a partir dos microdados das PNADs consiste em estimar intervalos de confiança para a realização de testes de hipóteses e sabendo-se que a PNAD é baseada em um processo de amostragem de domicílios com características que a definem como uma amostra complexa sujeita a erros probabilísticos, deve-se definir o método de estimação das variâncias. Para dados com as características das PNADs, a estimação das variâncias é realizada a partir de duas metodologias: linearização ou replicação.

Dentre os métodos de linearização, o método de linearização de Taylor é um dos mais utilizados⁵⁰. O método consiste em aproximar os estimadores de interesse que são não-lineares por estimadores lineares e utilizá-los para se obter a variância. Quanto aos métodos de replicação, um dos mais conhecidos é o procedimento de *bootstrapping* desenvolvido por Efron (1979). Este consiste em um processo de computação intensiva, ou seja, demanda várias horas para a computação das estimativas. Sinteticamente o método envolve reamostrar os dados com reposição em número suficiente de vezes, com o objetivo de gerar uma estimativa empírica da distribuição amostral da estatística. Em outras palavras, o procedimento pode ser descrito como uma amostragem repetida com reposição feita a partir da própria amostra em questão, obtendo-se a partir daí estimativas do erro-padrão do estimador (Guimarães, 2007). Segundo Chernick (1999), o método de *bootstrap* é frequentemente usado quando existe grande dificuldade de se calcular a estimativa do erro padrão de um estimador. Outro método de replicação que pode ser empregado diz respeito à replicação de *Jackknife*. Segundo Kneipp e Yarandi (2002), a estimação da variância pela replicação de *Jackknife* é preferencialmente utilizada quando os dados são divididos primeiramente em vários grupos. Um grupo pode consistir nos dados de um único PSU ou de um conjunto de PSUs. Cabe ressaltar que a replicação de *Jackknife* não requer estratificação explícita.

No entanto, como destaca Guimarães (2007), uma questão de relevância é que os dados não podem ter estratos com UPA única (também conhecido como *single PSU – Primary Sample Unit, PSU único*). Com PSU único não é possível obter as estimativas da variância, uma vez que é necessário se ter variabilidade para calcular a variância. Em algumas situações, além de PSU único deve-se observar a existência também de PSU com observação única. Para solucionar o problema de PSU único, pode-se agregar as

⁵⁰ Neder (2006) utiliza este método.

observações de estratos com PSU único em estrato na mesma UF com maior número de observações⁵¹.

1.2.3. Restrições da Pesquisa Nacional por Amostra de Domicílios

As PNADs apresentam algumas limitações que devem ser consideradas. O questionário procura captar tanto os rendimentos em dinheiro como em espécie, mas não considera o valor da produção para auto-consumo, que representa componente importante da renda real de pequenos agricultores, de acordo com Graziano da Silva e Del Grossi (2001) e Del Grossi e Graziano (2002). Ademais, a pesquisa até o ano de 2003 não abrangia a área rural da região Norte, não incluindo então o setor agrícola que, apesar de ser relativamente pequeno em termos de atividade, não é desprezível.

A coleta dos dados, por sua vez, tem como base o mês de setembro, não permitindo “que se capte a variedade das atividades agrícolas no país ao longo do ano” (Corrêa, 1998). Estes fatores tendem a subestimar os resultados para os indivíduos ocupados no setor agrícola, afetando os resultados das regressões.

Outra restrição é a subdeclaração das rendas, principalmente as mais elevadas. Hoffmann e Simão (2005) estimam que no Estado de Minas Gerais para o ano 2000 (Censo) os rendimentos declarados apresentam subestimação de cerca de 31% do seu valor real. Apesar de estes fatores promoverem efeitos sobre os resultados, eles não invalidam a análise de dados sobre rendimentos a partir dos questionários da PNAD⁵².

⁵¹ Este método foi utilizado por Guimarães (2007).

⁵² Mais detalhes sobre as restrições das PNADs podem ser encontrados em Hoffmann (1998), Del Grossi e Graziano (2002) e Rocha (2002).

APÊNDICE 3

Exemplo do algoritmo desenvolvido para a análise do mercado de trabalho rural brasileiro 2007

```
*****
***** PREPARACAO DOS DADOS*****
*****
clear

cd "D:\GILNEI\tese\ALGORITMOS"

** LEITURA DAS INFORMACOES DO DESENHO DA AMOSTRA NO ARQUIVO
DE DOMICILIOS**

clear
#delimit;
infix ano 1-4 uf 5-6 controle 5-12 serie 13-15
probmun 94-105 probsetor 109-120
intervalo 121-126 strat 159-165 psu 166-172
using "D:\GILNEI\PNADS\2007\Dados\DOM2007.txt", clear;
#delimit cr

#delimit;
sort controle serie;
format controle %15.0g;
format serie %15.0g;
replace controle = float(controle);
replace serie = float(serie);
#delimit cr

sum

save "D:\GILNEI\PNADS\2007\domicilio07.dta" , replace
```

```
*****
*****
*****
*****
```

```
#delimit;
clear;
use "D:\GILNEI\PNADS\2007\domicilio07.dta";
sort controle serie;
format controle %15.0g;
format serie %15.0g;
replace controle = float(controle);
replace serie = float(serie);
save "D:\GILNEI\PNADS\2007\domicilio07.dta" , replace;
#delimit cr
```

** LEITURA DOS DADOS DAS PESSOAS 2007**

```
clear
```

```
set more off
```

```
#delimit;
infix ano 1-4 uf 5-6 controle 5-12 serie 13-15
sexo 18-18 idade 27-29 cond_dom 30-30 cond_fam 31-31
cor 33-33 horas_total 135-136 horas_p 361-362 horas_s 455-456
horas_o 487-488 edu 668-669 horas_cat 674-674
atividade 675-675 cond_ativid 670-670 cond_ocup 671-671
pos_ocup 672-673 num_fam 32-32
rend_tod 702-713 rend_fon 714-725 rend_fa1 738-749 rend_p 690-701
fam_comp 752-753 areacen 754-754
sitcen 755-755 peso 756-760 pesofam 761-765
using "D:\GILNEI\PNADS\2007\Dados\PES2007.txt";
#delimit cr
```

** JUNCAO DAS INFORMACOES DE DESENHO DA AMOSTRA AO ARQUIVO DE
PESSOAS DA PNAD 2007**

```
#delimit ;
sort controle serie;
format controle %15.0g;
format serie %15.0g;
replace controle = float(controle);
replace serie = float(serie);
merge controle serie using "D:\GILNEI\PNADS\2007\domicilio07.dta";
#delimit cr
```

```
#delimit ;
keep if _merge == 3;
drop _merge;
```

```

#delimit cr

sum

save "D:\GILNEI\PNADS\2007\pess07.dta", replace

*****ARQUIVO FINAL DA PNAD PESSOAS PARA 2007*****

save "D:\GILNEI\PNADS\2007\pess07_FINAL.dta" , replace

*****
*****
*****DECLARANDO O CONJUNTO DE DADOS COMO SENDO DE
AMOSTRA COMPLEXA ANO
2007*****
*****

use "D:\GILNEI\PNADS\2007\pess07.dta", clear

svyset psu [pweight=peso], strata(strat) vce(linearized) singleunit(centered) || _n

svydes, single

save "D:\GILNEI\PNADS\2007\pess07.dta" , replace

use "D:\GILNEI\PNADS\2007\pess07.dta" , clear
gene novo_str = .
gene novo_psu = .
format novo_psu %12.0g
save acum, replace

capture program drop prog1
program define prog1
use "D:\GILNEI\PNADS\2007\pess07.dta" , clear
keep if uf == estado
gene novo_str = strat
gene novo_psu = psu
quietly {
save transf, replace
use acum, clear
append using transf
save acum, replace
}
end

capture program drop prog2
program define prog2
use "D:\GILNEI\PNADS\2007\pess07.dta" , clear
keep if uf == estado

```

```
idonepsu , strata(strat) psu(psu) generate(novo_) quietly { save transf, replace use acum,  
clear append using transf save acum, replace } end
```

```
scalar estado = .
```

```
foreach i in 53 {  
  scalar estado = `i'  
  prog1  
}
```

```
foreach i in 11 12 13 14 15 16 17 21 22 23 24 25 26 27 28 29 31 32 33 35  
41 42 43 50 51 52 {  
  scalar estado = `i'  
  prog2  
}  
end  
drop strat psu  
rename novo_str strat  
rename novo_psu psu
```

**** ROTINA DE ALOCACAO DE ESTRATOS COM UM UNICO PSU EM ESTRATOS
COM MAIOR NUMERO DE OBSERVACOES UTILIZANDO O COMANDO idonepsu
- ANO DE 2007****

```
use "D:\GILNEI\PNADS\2007\pess07.dta" , clear
```

```
set more off
```

```
#delimit;  
idonepsu, strata(strat) psu(psu) generate(new);  
drop strat psu;  
rename newstr strat;  
rename newpsu psu;  
#delimit cr
```

```
svyset psu [pweight=peso], strata(strat) vce(linearized) singleunit(centered) ///| _n
```

```
svydes, single
```

```
save "D:\GILNEI\PNADS\2007\pess07.dta" , replace
```

*****TRATAMENTO DAS VARIÁVEIS*****

*****determinando a condição de atividade em duas variáveis**
economicamente ativa=1**

```

recode cond_ativid (3=.)
recode cond_ativid (1=1) (2=0)

save "D:\GILNEI\PNADS\2007\pess07.dta" , replace

*****
***gerando total de horas trabalhadas na semana***

recode horas_p horas_s horas_o (-1 99 = 0)
recode horas_p horas_s horas_o (. = 0)

#delimit;
gen horas_total =.;
replace horas_total = horas_p+horas_s+horas_o;
gen horas_pmes = 4*(horas_p);
gen horas_totalmes = 4*(horas_total);
#delimit cr

save "D:\GILNEI\PNADS\2007\pess07.dta" , replace

***RETIRANDO OUTLIERS DAS VARIÁVEIS DE RENDA***

use "D:\GILNEI\PNADS\2007\pess07.dta" , clear

recode rend_tod (1.00e+12/2.29e+12=.) (-1=.)
recode rend_fa1 (1.00e+12/2.29e+12=.) (-1=.)
recode rend_p (1.00e+12/2.29e+12=.) (-1=.)
recode rend_fon (1.00e+12/2.29e+12=.) (-1=.)

***gerando o logaritmo da razão da renda principal e horas trabalhadas no mes***

gen ln_rend_horas_p = ln(rend_p/horas_pmes)

***gerando o logaritmo da razão da renda de todos trabalho e horas trabalhadas no mes***

gen ln_rend_tod = ln(rend_tod)

drop if ln_rend_tod <1

gen ln_horas_total = ln(horas_totalmes)

gen ln_rend_ln_horas = (ln_rend_tod/ln_horas_total)

save "D:\GILNEI\PNADS\2007\pess07.dta" , replace

*****
set more off

```

```
***gerando Dummy p/ sexo****
**masculino = 1 ; feminino = 0 ***
```

```
recode sexo (2=1) (4=0)
```

```
*** Gerando idade em dezenas de anos***
```

```
recode idade (999=.)
```

```
gen idade10 = (idade/10)
```

```
*** gerando idade10 ao quadrado*****
```

```
gen idade10q = (idade10^2)
```

```
*** excluindo as observações de anos de estudo sem declaração****
```

```
recode edu (17=.)
```

```
save "D:\GILNEI\PNADS\2007\pess07.dta" , replace
```

```
***CRIANCA NA FAMILIA***
```

```
#delimit
```

```
gen crianc=.
```

```
replace crianc = 1 if cond_fam ==3 & idade<=14;
```

```
replace crianc = 0 if crianc==.;
```

```
#delimit cr
```

```
egen float fam = group(uf controle serie num_fam)
```

```
bysort fam: egen float famcrianc = mean(crianc)
```

```
bysort fam: egen float famcriancnum = sum(famcrianc)
```

```
#delimit
```

```
gen crianca=.
```

```
replace crianca = 1 if famcrianc~=0;
```

```
replace crianca = 0 if famcrianc==0;
```

```
sort fam;
```

```
#delimit cr
```

```
#delimit
```

```
order ano uf controle serie fam_comp
```

```
fam num_fam famcrianc famcriancnum idade cond_fam crianca;
```

```
#delimit cr
```

```
***CRIANCA NA FAMILIA DO SEXO FEMININO***
```

```
#delimit
```

```
gen crianc_f=.
```

```
replace crianc_f = 1 if cond_fam ==3 & idade<=14 & sexo==0;
```

```
replace crianc_f = 0 if crianc_f==.;
#delimit cr
```

```
bysort fam: egen float famcrianc_f = mean(crianc_f)
bysort fam: egen float famcriancnum_f = sum(famcrianc_f)
#delimit
gen crianca_f=.;
replace crianca_f = 1 if famcrianc_f~=0;
replace crianca_f = 0 if famcrianc_f==0;
sort fam;
#delimit cr
```

CRIANCA NA FAMILIA DO SEXO MASCULINO

```
#delimit
gen crianc_m=.;
replace crianc_m = 1 if cond_fam ==3 & idade<=14 & sexo==1;
replace crianc_m = 0 if crianc_m==.;
#delimit cr
```

```
bysort fam: egen float famcrianc_m = mean(crianc_m)
bysort fam: egen float famcriancnum_m = sum(famcrianc_m)
#delimit
gen crianca_m=.;
replace crianca_m = 1 if famcrianc_m~=0;
replace crianca_m = 0 if famcrianc_m==0;
sort fam;
#delimit cr
```

```
#delimit
order ano uf controle serie fam_comp
fam num_fam famcrianc famcriancnum famcriancnum_m famcriancnum_f
sexo idade cond_fam crianca;
#delimit cr
```

```
*****
*****
```

*** criando as regiões do país*****

```
*regiao 1 = CENTRO-OESTE
*regiao 2 = SUDESTE
*regiao 3 = NORTE
*regiao 4 = SUL
*regiao 5 = NORDESTE
```

```
#delimit
recode uf (99=.);
```

```

gen regio =.;
replace regio = 1 if uf >=50 & uf < 88;
replace regio = 2 if uf >=31 & uf <= 35;
replace regio = 4 if uf >=41 & uf <=43;
replace regio = 3 if uf <=17;
replace regio = 5 if uf >=21 & uf <=29;
#delimit cr;

```

```

***GERANDO DUMMY'S PARA AS REGIÕES***
*** REGIAO NORDESTE É O CONTROLE***

```

```

#delimit
gen sudeste =.;
replace sudeste = 1 if regio==2;
replace sudeste = 0 if regio~2;
gen sul =.;
replace sul = 1 if regio ==4;
replace sul = 0 if regio ~4;
gen centro =.;
replace centro = 1 if regio==1;
replace centro = 0 if regio~1;
gen norte=.;
replace norte = 1 if regio ==3;
replace norte = 0 if regio ~3;
tab sudeste;
tab sul;
tab centro;
tab norte;
#delimit cr

```

```

save "D:\GILNEI\PNADS\2007\pess07.dta" , replace

```

```

*****

```

```

***transformando posicao na ocupacao p/ 4 categorias***
*** 1=empregado, 2=conta-propria, 3=empregador, 4=n_remunerados/proprio-
consumo/proprio-uso***

```

```

gen pos_ocup_backup=pos_ocup

```

```

recode pos_ocup (1/4=1) (6=1) (7=1) (9 = 2) (10 = 3) (11 12 13 = 4)

```

```

***criando dummy p/ posicao***
**empregado**

```

```

#delimit
gen empregado =.;
replace empregado = 1 if pos_ocup ==1;

```

```
replace empregado = 0 if pos_ocup > 1 ;
#delimit cr
```

****conta-propria****

```
#delimit
gen cont_propria =.;
replace cont_propria = 1 if pos_ocup ==2;
replace cont_propria = 0 if pos_ocup ~2;
#delimit cr
```

****empregador****

```
#delimit
gen empregador =.;
replace empregador = 1 if pos_ocup ==3;
replace empregador = 0 if pos_ocup < 3 ;
#delimit cr
```

```
#delimit
gen n_remunerado =.;
replace n_remunerado= 1 if pos_ocup ==4;
replace n_remunerado = 0 if pos_ocup < 4 ;
#delimit cr
```

```
save "D:\GILNEI\PNADS\2007\pess07.dta" , replace
```

*****seperacao da cor*****

```
recode cor (9=.)
```

```
#delimit;
gen preta_ind =.;
replace preta_ind = 1 if cor ==4 | cor==0;
replace preta_ind = 0 if preta_ind==.;
#delimit cr;
```

```
#delimit;
gen amarela_pard=.;
replace amarela_pard = 1 if cor >=6;
replace amarela_pard = 0 if cor <6;
#delimit cr;
```

```
#delimit;
gen brancas =.;
replace brancas = 1 if cor ==2;
replace brancas = 0 if cor ~2;
```

```
#delimit cr;
```

```
#delimit;  
gen pardas =.;  
replace pardas = 1 if cor ==8;  
replace pardas = 0 if cor ~=8;  
#delimit cr;
```

```
#delimit;  
gen amarelas =.;  
replace amarelas = 1 if cor ==6;  
replace amarelas = 0 if cor ~=6;  
#delimit cr;
```

```
#delimit;  
gen indigenas =.;  
replace indigenas = 1 if cor ==0;  
replace indigenas = 0 if cor ~=0;  
#delimit cr;
```

```
#delimit;  
gen preta =.;  
replace preta = 1 if cor ==4;  
replace preta = 0 if cor ~=4;  
#delimit cr;
```

```
sum preta brancas pardas amarelas indigenas
```

```
save "D:\GILNEI\PNADS\2007\pess07.dta" , replace
```

```
*****
```

```
****separando em urbano e rural****  
***rural sera o controle****
```

```
#delimit;  
gen urbano = .;  
replace urbano = 1 if sitcen < 4;  
replace urbano = 0 if sitcen > 3;  
tab urbano;  
#delimit cr;
```

```
#delimit;  
gen rural = .;  
replace rural = 0 if sitcen < 4;  
replace rural = 1 if sitcen > 3;  
tab rural;  
#delimit cr;
```

sum rural urbano

save "D:\GILNEI\PNADS\2007\pess07.dta" , replace

CRIANDO DUMMY P/ ATIVIDADE

agricola e o controle

recode atividade (3=.)

```
#delimit;
gen n_agricola = .;
replace n_agricola = 1 if atividade == 2;
replace n_agricola = 0 if atividade < 2;
tab n_agricola;
#delimit cr;
```

```
#delimit;
gen agricola = .;
replace agricola = 0 if atividade > 1;
replace agricola = 1 if atividade == 1;
tab agricola;
#delimit cr;
```

save "D:\GILNEI\PNADS\2007\pess07.dta" , replace

DUMMY PARA REFERÊNCIA NA FAMÍLIA

```
#delimit
gen ref = .;
replace ref = 1 if cond_fam == 1;
replace ref = 0 if cond_fam ~ = 1;
#delimit cr
```

MAE SOLTEIRA OU PAI SOLTEIRO

```
#delimit
gen solteiro_pm = cond_fam;
recode solteiro_pm (1=1) (2=2) (3/8=0);
#delimit cr
```

bysort fam: egen float solteiro = sum(solteiro_pm)

```
#delimit
gen pm_solteiro = .;
bysort fam: replace pm_solteiro = 1 if solteiro == 1 & famcrianc ~ = 0;
bysort fam: replace pm_solteiro = 0 if solteiro ~ = 1 & famcrianc == 0;
```

```
recode pm_solteiro (.=0);
sort fam;
#delimit cr
```

```
#delimit
order ano uf controle serie sexo fam pm_solteiro solteiro_pm solteiro cond_fam
num_fam famcrianc famcriancnum idade crianca;
#delimit cr
```

CRIANDO DUMMY P/ EXIST. DE CASAL

```
bysort fam: egen float casal_i = sum(solteiro_pm)
```

```
#delimit
gen casal=.;
bysort fam: replace casal= 1 if casal_i ==3;
bysort fam: replace casal= 0 if casal_i ~3;
sort fam;
#delimit cr
```

```
#delimit
order ano uf controle serie sexo fam cond_fam casal
num_fam famcrianc famcriancnum idade crianca;
#delimit cr
```

CRIANDO DUMMY'S P/ EDUCACAO PARA TESTAR O EFEITO LIMIAR (SEM INSTRUCAO REFERENCIA)

```
#delimit
gen edu2 =.;
replace edu2 = 1 if edu==2;
replace edu2 = 0 if edu~2;
gen edu3 =.;
replace edu3 = 1 if edu==3;
replace edu3 = 0 if edu~3;
gen edu4 =.;
replace edu4 = 1 if edu==4;
replace edu4 = 0 if edu~4;
gen edu5 =.;
replace edu5 = 1 if edu==5;
replace edu5 = 0 if edu~5;
gen edu6 =.;
replace edu6 = 1 if edu==6;
replace edu6 = 0 if edu~6;
gen edu7 =.;
replace edu7 = 1 if edu==7;
replace edu7 = 0 if edu~7;
gen edu8 =.;
replace edu8 = 1 if edu==8;
```

```

replace edu8 = 0 if edu~=8;
gen edu9 =.;
replace edu9 = 1 if edu==9;
replace edu9 = 0 if edu~=9;
gen edu10 =.;
replace edu10 = 1 if edu==10;
replace edu10 = 0 if edu~=10;
gen edu11 =.;
replace edu11 = 1 if edu==11;
replace edu11 = 0 if edu~=11;
gen edu12 =.;
replace edu12 = 1 if edu==12;
replace edu12 = 0 if edu~=12;
gen edu13 =.;
replace edu13 = 1 if edu==13;
replace edu13 = 0 if edu~=13;
gen edu14 =.;
replace edu14 = 1 if edu==14;
replace edu14 = 0 if edu~=14;
gen edu15 =.;
replace edu15 = 1 if edu==15;
replace edu15 = 0 if edu~=15;
gen edu16 =.;
replace edu16 = 1 if edu==16;
replace edu16 = 0 if edu~=16;
#delimit cr;

```

```

save "D:\GILNEI\PNADS\2007\pess07.dta" , replace

```

```

sum

```

```

***TESTE DAS 15 VAR. DUMMY'S PARA EFEITO LIMIAR DA EDUCAÇÃO***

```

```

#delimit;
svy: heckman ln_rend_tod sexo idade10 idade10q preta pardas edu2 edu3 edu4 edu5 edu6
edu7 edu8 edu9
edu10 edu11 edu12 edu13 edu14 edu15 edu16 sudeste sul centro norte if rural==1 &
idade>=21 & idade<66
, select(cond_ativid = rend_fa1 fam_comp sexo crianca idade10 idade10q preta pardas
edu2 edu3
edu4 edu5 edu6 edu7 edu8 edu9 edu10 edu11 edu12 edu13 edu14 edu15 edu16 sudeste sul
centro norte );
#delimit cr;

```

```

#delimit;
svy: heckman ln_rend_tod sexo idade10 idade10q preta pardas edu eledu
sudeste sul centro norte if rural==1 & idade>=21 & idade<66
, select(cond_ativid = rend_fa1 fam_comp sexo crianca idade10 idade10q preta pardas edu
eledu

```

```
sudeste sul centro norte);  
#delimit cr;
```

```
*****CRIANDO EFEITO LIMAR DA EDU*****
```

```
#delimit;  
gen z=.;  
replace z = 0 if edu <=11;  
replace z = 1 if edu >11;  
gen s= edu - 11;  
gen eledu = z*s;  
#delimit cr;
```

```
***CRIANDO EFEITO LIMAR DA EDU RURAL***
```

```
#delimit;  
gen z=.;  
replace z = 0 if edu <=6;  
replace z = 1 if edu >6;  
gen s= edu - 6;  
gen eledu = z*s;  
#delimit cr;
```

```
***ESTATÍSTICA DESCRITIVA TOTAL***
```

```
#delimit;  
svy linearized : mean rend_tod if rural==1 & idade>=15 & idade<71 ;  
svy linearized : mean ln_rend_tod if rural==1 & idade>=15 & idade<71 ;  
svy linearized : mean idade10 if rural==1 & idade>=15 & idade<71 ;  
svy linearized : mean idade10q if rural==1 & idade>=15 & idade<71;  
svy linearized : mean edu if rural==1 & idade>=15 & idade<71;  
svy linearized : mean eledu if rural==1 & idade>=15 & idade<71;  
svy linearized : mean famcriancnum if rural==1 & idade>=15 & idade<71 & crianca==1;  
svy linearized : mean rend_fal if rural==1 & idade>=15 & idade<71;  
svy linearized : mean fam_comp if rural==1 & idade>=15 & idade<71;  
svy linearized : mean horas_total if rural==1 & idade>=15 & idade<71;  
#delimit cr;
```

```
#delimit  
svy linearized : proportion preta_ind if rural==1 & idade>=15 & idade<71;  
svy linearized : proportion amarela_pard if rural==1 & idade>=15 & idade<71;  
svy linearized : proportion cor if rural==1 & idade>=15 & idade<71;  
svy linearized : proportion sexo if rural==1 & idade>=15 & idade<71;  
svy linearized : proportion regioao if rural==1 & idade>=15 & idade<71;  
svy linearized : proportion cond_ativid if rural==1 & idade>=15 & idade<71;  
svy linearized : proportion crianca if rural==1 & idade>=15 & idade<71;  
#delimit cr
```

```
***ESTATÍSTICA DESCRITIVA HOMENS***
```

```

#delimit;
svy linearized : mean rend_tod if rural==1 & idade>=15 & idade<71 & sexo==1;
svy linearized : mean ln_rend_tod if rural==1 & idade>=15 & idade<71 & sexo==1;
svy linearized : mean idade10 if rural==1 & idade>=15 & idade<71 & sexo==1;
svy linearized : mean idade10q if rural==1 & idade>=15 & idade<71 & sexo==1;
svy linearized : mean edu if rural==1 & idade>=15 & idade<71 & sexo==1;
svy linearized : mean eledu if rural==1 & idade>=15 & idade<71 & sexo==1;
svy linearized : mean famcriancnum if rural==1 & idade>=15 & idade<71 & sexo==1 &
crianca==1;
svy linearized : mean rend_fa1 if rural==1 & idade>=15 & idade<71 & sexo==1;
svy linearized : mean fam_comp if rural==1 & idade>=15 & idade<71 & sexo==1;
svy linearized : mean horas_total if rural==1 & idade>=15 & idade<71 & sexo==1;
#delimit cr;

```

```

#delimit
svy linearized : proportion preta_ind if rural==1 & idade>=15 & idade<71 & sexo==1;
svy linearized : proportion amarela_pard if rural==1 & idade>=15 & idade<71 & sexo==1;
svy linearized : proportion cor if rural==1 & idade>=15 & idade<71 & sexo==1;
svy linearized : proportion sexo if rural==1 & idade>=15 & idade<71 & sexo==1;
svy linearized : proportion regioao if rural==1 & idade>=15 & idade<71 & sexo==1;
svy linearized : proportion cond_ativid if rural==1 & idade>=15 & idade<71 & sexo==1;
svy linearized : proportion pm_solteiro if rural==1 & idade>=15 & idade<71 & sexo==1;
svy linearized : proportion casal if rural==1 & idade>=15 & idade<71 & sexo==1;
svy linearized : proportion crianca if rural==1 & idade>=15 & idade<71 & sexo==1;
#delimit cr;

```

ESTATÍSTICA DESCRITIVA MULHERES

```

#delimit;
svy linearized : mean rend_tod if rural==1 & idade>=15 & idade<71 & sexo==0;
svy linearized : mean ln_rend_tod if rural==1 & idade>=15 & idade<71 & sexo==0;
svy linearized : mean idade10 if rural==1 & idade>=15 & idade<71 & sexo==0;
svy linearized : mean idade10q if rural==1 & idade>=15 & idade<71 & sexo==0;
svy linearized : mean edu if rural==1 & idade>=15 & idade<71 & sexo==0;
svy linearized : mean eledu if rural==1 & idade>=15 & idade<71 & sexo==0;
svy linearized : mean famcriancnum if rural==1 & idade>=15 & idade<71 & sexo==0 &
crianca==1;
svy linearized : mean rend_fa1 if rural==1 & idade>=15 & idade<71 & sexo==0;
svy linearized : mean fam_comp if rural==1 & idade>=15 & idade<71 & sexo==0;
svy linearized : mean horas_total if rural==1 & idade>=15 & idade<71 & sexo==0;
#delimit cr;

```

```

#delimit
svy linearized : proportion preta_ind if rural==1 & idade>=15 & idade<71 & sexo==0;
svy linearized : proportion amarela_pard if rural==1 & idade>=15 & idade<71 & sexo==0;
svy linearized : proportion cor if rural==1 & idade>=15 & idade<71 & sexo==0;
svy linearized : proportion sexo if rural==1 & idade>=15 & idade<71 & sexo==0;
svy linearized : proportion regioao if rural==1 & idade>=15 & idade<71 & sexo==0;
svy linearized : proportion cond_ativid if rural==1 & idade>=15 & idade<71 & sexo==0;

```

```
svy linearized : proportion pm_solteiro if rural==1 & idade>=15 & idade<71 & sexo==0;
svy linearized : proportion casal if rural==1 & idade>=15 & idade<71 & sexo==0;
svy linearized : proportion crianca if rural==1 & idade>=15 & idade<71 & sexo==0;
#delimit cr
```

```
***ESTATÍSTICA DESCRITIVA, TESTANDO SE INDIVÍDUOS IGUAIS SÃO
REMUNERADOS DE FORMA DISTINTA***
```

```
*****POR SEXO*****
```

```
*****ENSINO FUNDAMENTAL COMPLETO*****
```

```
svy linearized : mean rend_tod if rural==1 & idade>=15 & idade<71 & edu==8,
over(sexo)
```

```
*****ENSINO MÉDIO COMPLETO*****
```

```
svy linearized : mean rend_tod if rural==1 & idade>=15 & idade<71 & edu==11,
over(sexo)
```

```
***ENSINO SUPERIOR COMPLETO (CONSIDERANDO 4 ANOS)***
```

```
svy linearized : mean rend_tod if rural==1 & idade>=15 & idade<71 & edu==15,
over(sexo)
```

```
***POR COR OU RAÇA MULHERES***
```

```
*****ENSINO FUNDAMENTAL COMPLETO*****
```

```
svy linearized : mean rend_tod if rural==1 & idade>=15 & idade<71 & edu==8 &
sexo==0, over(cor)
```

```
*****ENSINO MÉDIO COMPLETO*****
```

```
svy linearized : mean rend_tod if rural==1 & idade>=15 & idade<71 & edu==11 &
sexo==0, over(cor)
```

```
*****ENSINO SUPERIOR COMPLETO (CONSIDERANDO 4 ANOS)*****
```

```
svy linearized : mean rend_tod if rural==1 & idade>=15 & idade<71 & edu==15 &
sexo==0, over(cor)
```

```
***ANÁLISE DE REGRESSÃO TOTAL RURAL***
```

```
#delimit;
```

```
svy: heckman ln_rend_horas sexo pm_solteiro casal idade10 idade10q amarela_pard
preta_ind edu eledu
```

```
sudeste sul centro norte if rural==1 & idade>=15 & idade<71
, select(cond_ativid = sexo pm_solteiro casal rend_fa1 fam_comp famcriancnum idade10
idade10q amarela_pard preta_ind edu eledu
sudeste sul centro norte);
#delimit cr;
```

ANÁLISE DE REGRESSÃO HOMENS

```
#delimit;
svy: heckman ln_rend_horas pm_solteiro casal idade10 idade10q amarela_pard preta_ind
edu eledu
sudeste sul centro norte if rural==1 & idade>=15 & idade<71 & sexo==1
, select(cond_ativid = pm_solteiro casal rend_fa1 fam_comp famcriancnum idade10
idade10q amarela_pard preta_ind edu eledu
sudeste sul centro norte);
#delimit cr;
```

*** ANÁLISE DE REGRESSÃO MULHERES***

```
#delimit;
svy: heckman ln_rend_horas pm_solteiro casal idade10 idade10q amarela_pard preta_ind
edu eledu
sudeste sul centro norte if rural==1 & idade>=15 & idade<71 & sexo==0
, select(cond_ativid = pm_solteiro casal rend_fa1 fam_comp famcriancnum idade10
idade10q amarela_pard preta_ind edu eledu
sudeste sul centro norte);
#delimit cr;
```

EFEITO MARGINAL

```
mfx, predict(xb) eydx
mfx, predict (psel)
```

MULTICOLINEARIDADE

```
#delimit
corr ref pm_solteiro casal cond_ativid rend_fa1 ln_rend_tod fam_comp crianca
famcriancnum sexo idade10 idade10q preta pardas
edu eledu sudeste sul centro norte n_agricola empregador cont_propria [fw=peso];
#delimit cr;
```

EFEITO DO PLANO AMOSTRAL

```
estat effects, MEFF DEFF // para analisar o efeito do plano amostral //
```

CHEGANDO A EXISTENCIA DE HETEROCEDASTICIDADE

```

set more on
#delimit
anova cond_ativid [fw=peso];
hettest;
#delimit
anova casal [fw=peso];
hettest;
anova ref [fw=peso];
hettest;
anova pm_solteiro[fw=peso];
hettest;
anova ln_rend_tod [fw=peso];
hettest;
anova famcriancnum [fw=peso];
hettest;
anova rend_fa1 [fw=peso];
hettest ;
anova fam_comp [fw=peso];
hettest;
anova sexo [fw=peso];
hettest;
anova idade10 [fw=peso];
hettest;
anova idade10q [fw=peso];
hettest;
anova preta [fw=peso];
hettest;
anova pardas [fw=peso];
hettest;
anova edu [fw=peso];
hettest;
anova eledu [fw=peso];
hettest;
anova sudeste [fw=peso];
hettest;
anova sul [fw=peso];
hettest;
anova centro [fw=peso];
hettest;
anova norte [fw=peso];
hettest norte;
#delimit cr;

```

```

*****
*****FIM DO ALGORITMO*****
-----

```