

PATRÍCIA DE MELO ABRITA BASTOS

ENSAIOS SOBRE O RENDIMENTO RURAL E URBANO, 1981-2007

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

PATRÍCIA DE MELO ABRITA BASTOS

ENSAIOS SOBRE O RENDIMENTO RURAL E URBANO, 1981-2007

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: 16 de outubro de 2009.

Prof. Adriano Provezano Gomes
(Co-orientador)

Prof. Sidney Martins Caetano

Prof. Luiz Eduardo de Vasconcelos
Rocha

Prof. Evaldo Henrique da Silva

Prof^ª. Sílvia Harumi Toyoshima
(Orientadora)

A Deus, a minha família, minha mãe
Rita, meu pai João e ao meu noivo Gilnei.

“Noções do Tempo”.

Diz o preguiçoso: “Amanhã farei”.

Exclama o fraco: “Amanhã, terei forças”.

Assevera o delinquente: “Amanhã, regenero-me”.

É imperioso reconhecer, porém, que a criatura, adiando o esforço pessoal, não alcançou, ainda, em verdade, a noção real do tempo.

Quem não aproveita a benção do dia, vive distante da glória do século.

Emmanuel

Psicografado por Francisco Cândido Xavier

AGRADECIMENTOS

Agradeço primeiro a Deus, pois no momento que o indaguei, sua resposta foi rápida e reconfortante. A crença na sua existência e nas suas criações são fontes instigantes para meu crescimento intelectual na tentativa de compreender um mínimo de sua inteligência.

Aos meus pais, que sempre acreditaram no meu potencial, oferecendo-me segurança para conduzir a dissertação com tranquilidade. Seus estímulos começaram cedo, proporcionando contato com revistas científicas e incentivando a realização de pequenos trabalhos científicos, onde os prêmios recebidos são méritos deles. A eles devo minha vida e minha alegria de viver.

Ao meu namorado Gilnei que sempre me achou muito capaz, cuja admiração (que é recíproca) alimenta meu ego e me faz acreditar que sempre posso mais nos estudos. Somente através do seu despreendimento em passar seus conhecimentos de programação que foi possível chegar aos resultados econométricos de fronteira apresentados nesta dissertação. Sendo assim, fica registrado aqui o meu mais profundo agradecimento.

À minha orientadora, que permitiu realizar a pesquisa que já havia programado e, por conseguinte, pelo esforço na orientação da pesquisa.

Aos membros da banca, que foram precisos nas suas considerações, contribuindo para a clareza dos ensaios.

Por fim, aos professores do Departamento de Economia pela indiscutível contribuição na minha formação profissional e aos amigos pelas boas risadas.

SUMÁRIO

AGRADECIMENTOS	iv
SUMÁRIO.....	v
LISTA DE FIGURAS	vii
LISTA DE TABELAS	viii
RESUMO	x
ABSTRACT	xii
2.1. Introdução.....	7
2.1.1. Problema e justificativa do trabalho.....	7
2.1.2. Hipóteses e Objetivos	12
2.1.2.1. Hipótese	12
2.1.2.2. Objetivos	13
2.2. Referencial Teórico	13
2.2.1. Teoria do Capital Humano	13
2.2.2. Externalidades negativas	16
2.2.3. Rural e Urbano	18
2.3. Metodologia.....	19
2.3.1. Equações de Rendimento	19
2.3.2. Seleção das variáveis para a elaboração da equação de rendimento	21
2.3.2. Métodos empregados.....	28
2.3.3. Fonte de dados.....	31
2.3.4. Tratamento dos dados.....	32
2.3.5. Deflator da renda	34
2.4. Resultados e discussão	34
2.4.1. Estatística descritiva	34
2.4.2. Limiar da Educação.....	44
2.4.3. Decomposição do Diferencial de Salário com Correção do Viés de Seleção do Salário na Análise de Regressão.....	46
2.4.3.1. Evolução da Decomposição do Diferencial de Salário.....	52
2.5. Conclusão	55
3.1. Introdução.....	57
3.1.1 Problema e justificativa do trabalho	57
3.1.2. Hipótese e objetivos	63
3.1.2.1. Hipótese	63
3.1.2.2. Objetivos	63

3.2. Referencial Teórico	64
3.2.1. Eficiência.....	64
3.3. Metodologia.....	72
3.3.1. Desigualdade	72
3.3.1.1. Índice de Gini.....	72
3.3.1.2. Índices de Theil.....	73
3.3.2. Eficiência.....	75
3.3.2.1. Análise envoltória de dados - DEA.....	75
3.3.2.2. Índice Malmquist de produtividade	77
3.3.3. Fonte de dados.....	80
3.3.4. Tratamento dos dados.....	81
3.3.4.1. Seleção das variáveis	81
3.4. Resultados e discussão	81
3.4.1. Medidas de desigualdade (Gini e Theil).....	81
3.4.2. Estudo da eficiência do trabalho	87
3.4.3. Malmquist.....	91
3.4.4. Análise de eficiência, 1981 a 2007.....	96
3.5. Conclusão	101
APÊNDICE 1	117
Resultados complementares	117
APÊNDICE 2	124
EXEMPLO DO ALGORITMO DESENVOLVIDO PARA O ANO DE 2007.....	124

LISTA DE FIGURAS

Figura 2.1: População do Meio Urbano e Meio Rural Brasileiro.....	36
Figura 2.2: Proporção de Indivíduos Economicamente Ativos no Brasil.	37
Figura 2.3: Proporção de Indivíduos Ocupados em Atividades não Agrícolas no Brasil. ...	38
Figura 2.4: Média de Anos de Estudo dos Indivíduos Entre 15 e 70 Anos.....	39
Figura 2.5: Rendimento Médio Advindo do Trabalho Principal.....	40
Figura 2.6: Rendimento Médio Familiar (R\$).....	41
Figura 2.7: Rendimento do Trabalho Principal no meio Rural.	42
Figura 2.8: Rendimento do Trabalho Principal no meio Urbano.	42
Figura 2.9: Média de Horas no Trabalho Principal.	43
Figura 2.10: Idade em anos.	44
Figura 2.11: Escolha do Limiar da Educação.....	45
Figura 2.12: Evolução do Limiar da Educação.	46
Figura 2.13: Evolução do Retorno dos Anos de Estudo.....	53
Figura 2.14: Evolução do Retorno da Posição na Ocupação (empregador).....	54
Figura 2.15: Evolução do Retorno de Cor ou Raça (Branca).....	54
Figura 2.16: Evolução do Retorno por Gênero (Homem).....	55
Figura 3.1. Fronteira de Produção.	66
Figura 3.2. Eficiência Técnica dos estados brasileiros, 1981.....	98
Figura 3.3. Eficiência Técnica dos estados brasileiros, 2007.....	98
Figura 3.4: Variação da Eficiência Técnica (VET) dos estados brasileiros, 1981-2007.....	99
Figura 3.5: Variação Tecnológica (VTC) dos estados brasileiros, 1981-2007.	100
Figura 3.6: Variação da Produtividade Total dos Fatores dos estados brasileiros, 1981-2007.....	100

LISTA DE TABELAS

Tabela 1. Evolução da População Rural e Urbana Brasileira.....	2
Tabela 2.1. Descrição das variáveis utilizadas e seus efeitos esperados sobre as equações, de participação e de rendimentos.....	27
Tabela 2.2. Determinantes do rendimento em 1995.....	48
Tabela 2.3. Equação de seleção 1995.....	49
Tabela 2.4. Determinantes do rendimento em 2007.....	50
Tabela 2.5. Equação de seleção 2007.....	51
Tabela 3.1. Distribuição de Renda dos Estados Brasileiros.....	84
Tabela 3.2. Distribuição de Renda da Área Rural dos Estados Brasileiros.....	85
Tabela 3.3. Distribuição de Renda da Área Urbana dos Estados Brasileiros.....	86
Tabela 3.4. Distribuição da Eficiência do Brasil para os Anos 1981, 1990 e 1992.....	88
Tabela 3.5. Distribuição da Eficiência do Brasil para os anos 1999, 2001 e 2007.....	90
Tabela 3.6. Cálculo do índice de Malmquist para os estados brasileiros, 1981 a 1990.....	93
Tabela 3.7. Cálculo do índice de Malmquist para os estados brasileiros, 1992 a 1999.....	94
Tabela 3.8. Cálculo do índice de Malmquist para os estados brasileiros, 2001 a 2007.....	95
Tabela A.1. Análise descritiva das principais variáveis utilizadas na pesquisa.....	117
Tabela A.2. Análise descritiva das principais variáveis utilizadas na pesquisa.....	118
Tabela A.3. Análise descritiva das principais variáveis utilizadas na pesquisa.....	118
Tabela A.4. Análise descritiva das principais variáveis utilizadas na pesquisa.....	119
Tabela A.5. Análise descritiva das principais variáveis utilizadas na pesquisa.....	119
Tabela A.6. Análise descritiva das principais variáveis utilizadas na pesquisa.....	120
Tabela A.7. Análise descritiva das principais variáveis utilizadas na pesquisa.....	120
Tabela A.8. Análise descritiva das principais variáveis utilizadas na pesquisa.....	121
Tabela A.9. Análise descritiva das principais variáveis utilizadas na pesquisa.....	121

Tabela A.10. Análise descritiva das principais variáveis utilizadas na pesquisa.	122
Tabela A.11. Análise descritiva das principais variáveis utilizadas na pesquisa.	122
Tabela A.12: Análise descritiva das principais variáveis utilizadas na pesquisa.	123

RESUMO

BASTOS, Patrícia de Melo, M.sc., Universidade Federal Viçosa, Outubro de 2009, **Ensaio sobre o rendimento rural e urbano, 1981-2007**. Orientadora: Silvia Harumi Toyoshima. Co-orientadores: João Eustáquio de Lima e Adriano Provezano Gomes.

O objetivo deste trabalho foi analisar a evolução das diferenças do rendimento do trabalho entre o meio rural e meio urbano, devido às transformações do emprego no campo e na cidade, no período 1981-2007. Para isso foram elaborados dois artigos, tendo como base os microdados da Pesquisa Nacional por Amostra de Domicílio (PNAD) do Instituto Brasileiro de Geografia e Estatística (IBGE). O primeiro artigo estimou os determinantes dos rendimentos das pessoas ocupadas residentes no meio rural e no meio urbano, separadamente, por meio da equação de rendimento, uma metodologia de estatística paramétrica. O segundo calculou a fronteira de eficiência do trabalho da população dos estados brasileiros, onde também se fez o corte entre o rural e urbano, a partir da Análise Envoltória dos Dados – DEA –, caracterizada como técnica não paramétrica. Verificou-se que o rural apresenta características quanto à renda ainda desfavorável frente ao urbano. Em média o trabalhador rural, independentemente do ramo de atividade, agrícola ou não-agrícola, recebe a metade da renda do trabalhador urbano. Entretanto, em relação a outros aspectos como horas trabalhadas, limiar da educação e do impacto da posição na ocupação no retorno da renda do trabalhador residente no rural observou-se convergência em 2007, caracterizando maior aproximação destes espaços. O rural também se aproximou do urbano nas análises de eficiência, sendo que nas Regiões Sudeste, Sul e Centro-Oeste o diferencial de eficiência diminuiu expressivamente, de cerca de 100% para próximo de 30%. Coerente com tal resultado, constatou-se aumento da produtividade mais expressivo no rural por meio do índice de *Malmquist*. Os resultados sugerem, assim, que as políticas públicas deveriam dar maior atenção ao rural brasileiro no que se refere a renda do

trabalho para que os trabalhadores sejam remunerados de acordo com a atividade que desempenha, amenizando a segmentação de mercado, principalmente no norte e nordeste onde a eficiência do trabalho urbano é quase o dobro do rural.

ABSTRACT

BASTOS, Patrícia de Melo, Msc., Federal University of Viçosa, October 2009, **Essays about income in urban and rural areas, 1981-2007**. Adviser: Silvia Harumi Toyoshima. Co- Advisers: João Eustáquio de Lima e Adriano Provezano Gomes.

This paper objective was to analyze the evolution of work income differences between rural and urban areas due to the transformations in labor, both at the field and the city, in the period from 1981 to 2007. Two articles were prepared using the microdata from the National Household Survey (PNAD) by the Brazilian Institute of Geography and Statistic (IBGE). The first article used the earnings equation, a parametric statistic methodology, to estimate, separately, the income determinants of occupied resident persons from the rural and the urban areas. The second paper used the Data Envelopment Analysis/DEA, characterized as non parametric technique, to compute the labor efficiency frontier from Brazilian states population, where it has also done the cohort between the rural and urban. It was verified the rural presents unfavorable income characteristics comparatively to the ones from the urban area. On the average, the rural worker receives half of the income than the urban worker does, in either activity sector, agricultural or non agricultural. Regarding to others aspects as worked hours, however, threshold of education and occupation position impact in the income return of the rural resident worker it was observed convergence in 2007. This fact characterized very important approximation of the two areas, rural and urban. The rural is also close to the urban in terms of efficiency analysis. In the Southeast, South and Western-Center regions efficiency differential decreased expressively, about 100% to around 30 %. An expressive and increased productivity was noted in the rural area through the Malmquist rate. The results then suggested public policies should focus on Brazilian rural area regarding to work income so that workers would be remunerated

according to their activities. This could decrease the market segmentation, mainly in the North and Northeast Regions where urban labor efficiency is almost the double of rural area.

1. INTRODUÇÃO

As transformações ocorridas no meio rural brasileiro, nas últimas décadas, têm alterado o tipo de trabalho e o rendimento das pessoas ocupadas no campo. Em princípios da década de 1980, verificou-se no Brasil a perda da importância das atividades agrícolas na geração de emprego e renda no meio rural. Dentre as causas principais dessas mudanças, destaca-se a modernização agrícola que vem substituindo mão-de-obra pelo capital, com a disseminação da mecanização, promovendo o processo denominado de mercantilização do tempo livre, onde as famílias rurais procuram alternativas para complementar sua renda e viabilizar a permanência no campo.

De acordo com Graziano e Del Grossi (2001), em 1998, o total das rendas não-agrícolas do Brasil ultrapassou o montante das rendas agrícolas recebido pelos moradores rurais. Não obstante, em artigo mais recente, Graziano e Del Grossi (2006) verificaram que as ocupações agrícolas começaram a retomar sua importância onde se observou a recuperação das rendas de origem agrícola. Porém, o avanço das ocupações não-agrícolas é um fato consolidado no meio rural. Este crescimento é devido, principalmente, à expansão dos serviços públicos; ao aumento da demanda por educação, saúde, transporte, dentre outros bens e serviços; ao aumento da demanda por terras para uso não-agrícola, por parte das empresas que não querem assimilar as externalidades negativas dos centros urbanos; e, ao aumento da demanda da população urbana de alta renda por uma segunda residência, em busca de áreas de lazer.

Quanto às transformações no meio urbano, segundo Morley (2001), destaca-se a maior homogeneização da educação. E, segundo Barros *et al.* (2007) e Hoffmann e Simão (2005), a consequente diminuição da desigualdade socioeconômica da população, uma vez que no urbano a educação é o principal determinante do rendimento. Por outro lado, tem-se o agravamento das externalidades negativas geradas pelo rápido crescimento das cidades.

Para se defenderem dessas, os indivíduos incorrem em custos, denominados “*defensive expenditures*”, por exemplo, para contratar seguranças e vigias noturnos, comprar remédios devido à poluição, gastos que reduzem a atratividade das cidades. Deste modo, o rendimento do trabalho no meio urbano teria que superar os custos que o indivíduo teria que arcar na cidade para tornar a análise custo-benefício da migração positiva.

Muitos especialistas, como Neder (2003), acreditam que as transformações no meio rural estão viabilizando a permanência da população no campo. A tabela 1 abaixo corrobora essa assertiva. No princípio do século XX, a população era predominantemente rural. Atualmente, cerca de 80% da população brasileira é urbana, contra apenas 20% rural. Todavia, verifica-se que na última década houve um crescimento de cerca de apenas 5% na população urbana contra 12% da penúltima, indicando o arrefecimento do êxodo rural.

As áreas pertencentes ao meio rural e ao urbano são periodicamente redefinidas. O IBGE, por exemplo, a cada 10 anos, faz nova delimitação entre perímetro urbano e rural, quando realiza os censos demográficos¹. Pode-se observar que, nas últimas décadas, houve um estreitamento das áreas rurais, na medida em que novas cidades foram “aparecendo”, em alguns casos devido a interesses como o aumento da arrecadação de impostos, e não devido a uma mudança na estrutura da região.

Neste contexto, Diniz e Silveira (2009) afirmam que a queda na população rural de 31,8 milhões de pessoas, em 1993, para 27,6 milhões de pessoas, em 2006, se deve em grande parte à atualização das malhas municipais de seus perímetros urbanos. Ou seja, sinaliza-se uma estabilização da população rural, onde muito dos avanços populacional do urbano é meramente metodológica, não se verificando de fato a migração.

Tabela 1. Evolução da População Rural e Urbana Brasileira.

Ano	População urbana (Percentual)	População rural (Percentual)
1940	30.24	69.76
1950	36.16	63.84
1960	45.08	54.92
1970	55.94	44.06
1980	67.59	32.41
1991	75.59	24.41
2000	81.23	18.77

Fonte: Elaborado pela autora a partir do SIDRA, IBGE.

¹ Em 1938, através do Decreto-Lei 311, ficou estabelecido que as cidades (sedes de municípios) e as vilas (sedes de distritos) são urbanas e que o restante do território é rural. Desde então cabe às prefeituras definir, através de lei, os perímetros urbanos e rurais.

De acordo com Barros (2004), em 2002 cerca de 75% da renda das famílias brasileiras eram derivadas do trabalho, sendo que mais da metade tinham o trabalho como sua única fonte de renda, colocando o funcionamento do mercado de trabalho em lugar de destaque na determinação do crescimento, pobreza e desigualdade de renda no país.

Em 1995, observou-se que, dentre as pessoas ocupadas em algum tipo de trabalho no país, 76% estavam empregadas em atividades não-agrícolas. No meio rural, somente 25% estavam ocupadas em atividades não-agrícolas, contra 91% no urbano. Já, em 2007, o percentual de pessoas ocupadas em atividades não-agrícolas aumentou para 83%, sendo que a população do meio rural que desempenhavam alguma atividade não-agrícola se elevou para 31%, do meio urbano para 94%. Esses resultados fundamentam as mudanças que estão ocorrendo no rural. Ou seja, ao longo dos anos verifica-se o aumento das ocupações não-agrícolas no meio rural. Ademais, nota-se não somente o crescimento do número de ocupações não-agrícolas, mas também da importância dessas atividades no meio rural, conforme supramencionado.

A delimitação clara entre o trabalho agrícola ser efetuado no meio rural e o trabalho não-agrícola no meio urbano, assim, tem-se modificado nas últimas décadas, o que levanta inúmeras questões: a) qual o atual perfil ocupacional no meio rural e no urbano? b) quais são os determinantes da desigualdade no rendimento do trabalho em ambos os espaços? c) houve redução das desigualdades no rendimento do trabalho entre o campo e a cidade, em vista do aumento de atividades não-agrícolas no primeiro? d) houve redução das desigualdades entre o nível de escolaridade entre o meio rural e urbano? e) houve aumento da eficiência do rendimento do trabalho em ambos os espaços? Em suma, são muitas questões importantes a serem respondidas, as quais se tornaram objeto do presente estudo.

O objetivo geral da dissertação foi, assim, estudar as alterações ocorridas no rendimento do trabalho no meio rural contrapondo-as com o meio urbano, no período 1981-2007. A comparação entre estes dois espaços visou compreender quão próximos se encontram estes espaços no que concerne ao tema em questão.

Para o cumprimento do objetivo, a presente dissertação está composta por dois ensaios, uma vez que comporta diferentes estudos em torno do mesmo tema. O primeiro ensaio analisou as modificações nos rendimentos das pessoas ocupadas residentes no meio rural e no meio urbano, no período selecionado, por meio da equação de rendimento, uma metodologia de estatística paramétrica. E, o segundo estudou a evolução da eficiência do trabalho da população dos estados brasileiros, onde também se fez o corte entre o rural e

urbano, a partir da Análise Envoltória dos Dados – DEA² –, caracterizada como técnica não paramétrica.

No que concerne ao primeiro ensaio, a análise dos principais determinantes dos rendimentos teve como base a Teoria do Capital Humano (Becker, 1964), que supõe que a educação e a experiência são as principais variáveis que determinam a renda do indivíduo. Essa teoria sugere, a partir de amplas pesquisas, que ao investir em educação o indivíduo tem o retorno em aumento do salário. Isso acontece porque ao aumentar a quantidade de anos de estudo, é esperado um aumento da qualificação que se reverte em vantagens na aquisição de novas aprendizagens para o mercado de trabalho. Igualmente, quanto mais anos de experiência maior a tendência para que o trabalhador acumule informações que se traduzem em maior eficiência do trabalho. Nota-se que o conceito de capital humano é uma analogia ao conceito de capital econômico³. O indivíduo ao ser dotado de educação e experiência estaria de posse do capital humano.

Por outro lado, existem críticas à Teoria do Capital Humano, como as de Gandra (2002) e Langoni (1973), de que indivíduos igualmente produtivos, com mesmos anos de estudo e mesma experiência, são remunerados de forma distinta. Ou seja, argumentam que existem outros fatores que determinam o rendimento do trabalho, que se relacionam com segmentação de mercado e fatores de discriminação.

O primeiro ensaio, assim, incorporou outras variáveis além da educação e experiência no estudo, comportando três partes que se encontram descritas abaixo.

Inicialmente, verificou-se a existência de segmentação de mercado nos meios rural e urbano, por meio de uma equação de rendimento conjunta. Em outras palavras, procurou-se avaliar se o local de residência, rural ou urbano, teve impacto na renda do indivíduo. Partiu-se da hipótese que devido à aproximação do rural em relação ao urbano, no que se refere aos ramos de atividade, se observaria uma redução do diferencial de renda entre os dois espaços.

Em seguida, buscou-se a partir da estimação de equações de rendimento para o rural e o urbano, separadamente, verificar as mudanças ocorridas ao longo dos anos nos determinantes da renda, dando ênfase a variável educação e a posição na ocupação.. A posição na ocupação⁴ foi utilizada como uma *proxy* para a posse de capital. Dado que o

² *Data Envelopment Analysis*.

³ Capital econômico: elemento produtor de riqueza que rende lucros periódicos para quem o possui.

⁴ Segundo o dicionário da PNAD, a Posição na Ocupação refere-se se a pessoa é empregador, empregado ou conta-própria.

rural sofreu mudanças ao longo dos anos na composição da população ocupada, o resultado esperado é que a posição na ocupação não exerça expressivo diferencial na explicação da variação da renda no rural. Ou seja, espera-se que o fato de ser empregador não represente expressiva diferença de renda em relação aos trabalhadores e conta-própria.

Por fim, comparou-se a evolução do limiar da educação para o rural e urbano. O limiar da educação capta uma alteração não-linear do efeito de anos de estudos sobre os rendimentos, ou seja, identifica a partir de qual ano de estudo, o indivíduo tem um aumento de rendimento proporcionalmente maior do que os anos de estudo anteriores. Trata-se realmente de um limiar, porque o aumento é brusco para o acréscimo de apenas um ano de estudo a mais. Neste sentido, buscou-se verificar se com o aumento de ocupações no meio rural, o limiar da educação deste último se aproximou àquele do meio urbano.

Cabe destacar que para a estimação de todas as equações de rendimento os dados da PNAD foram tratados como amostra complexa, para não se obter estimativas viciadas das variâncias da amostra, o que poderia afetar o teste de significância das variáveis. A fim de corrigir o Viés de Seleção, foi aplicado o Procedimento de Heckman que é feito em dois estágios: no primeiro estima-se o modelo Probit, para se obter a variável λ ; e, no segundo, caso o λ seja estatisticamente significativo, ele é utilizado como variável explicativa para estimação da equação de rendimento.

O segundo ensaio, que compreende a análise de eficiência, procurou demonstrar que devido ao aumento de atividades não-agrícolas no rural, verificou-se um aumento da eficiência no meio rural, aproximando-a à do meio urbano.

A definição de eficiência, segundo a teoria econômica refere-se à otimização de recursos, pela utilização máxima dos recursos existentes para satisfazer as necessidades e os desejos de indivíduos e organizações (Pindyck e Rubinfeld 1994).

Quando se trata da eficiência do trabalho, objeto do presente estudo, supõe-se que as famílias e indivíduos, sendo agentes econômicos que agem com racionalidade, procurarão extrair o máximo de lucro do trabalho que exercem, dado o tempo de dedicação e respectiva qualificação. E, de acordo com a Teoria do Capital Humano de Becker (1964), como mencionado acima, a remuneração é equivalente ao custo e tempo dedicado na obtenção de qualificação.

No entanto, segundo Barros *et al.* (2007), tem-se observado que dentro de um mesmo estado ou município persistem disparidades espaciais de remuneração entre trabalhadores com idênticas características produtivas, ou seja, qualificação semelhante,

dando origem à segmentação de mercado. As mais notórias são as existentes entre as áreas urbanas e as rurais, onde trabalhadores urbanos recebem uma remuneração acima da dos trabalhadores rurais com idênticas características observáveis em postos de trabalho similares. Não obstante, as transformações que aconteceram no meio rural, como consolidação de ocupações não-agrícolas, imprimem uma maior dinamização na economia local, colocando este espaço mais próximo ao meio urbano. Neste sentido, espera-se uma tendência de redução da disparidade da eficiência do rendimento do trabalho devido a um provável arrefecimento da segmentação de mercado (rural/urbano).

Outrossim, existe também a segmentação setorial que diz respeito a trabalhadores com idênticas características produtivas que, por exercerem atividades em setores diferentes, agrícolas e não-agrícolas, são remunerados de forma distinta pelo mercado. Assim, dado que as atividades não-agrícolas, geralmente, exigem maior qualificação, o aumento de pessoas ocupadas em atividades não-agrícolas no campo, em princípio, deveria ter ocasionado a elevação da eficiência no rural.

Para a verificação, assim, da alteração na eficiência do trabalho em ambos os espaços, o rendimento do trabalho foi considerado como produto e as horas trabalhadas e os anos de estudo como insumos.

Para estimar a eficiência do rendimento do trabalho foi calculada a eficiência técnica baseada na estimação da fronteira empírica, por meio de programação linear. O foco de análise continuou sendo a comparação entre o meio rural e o urbano. Porém, para abranger as diferenças regionais e totalizar o número de unidades necessárias para o cálculo do método DEA, foram considerados os estados brasileiros, comparando-os em termos de eficiência dos seus meios rurais e urbanos.

Com o objetivo de corroborar os resultados encontrados na análise de eficiência foi calculado o Índice de Malmquist. Este índice permite analisar se houve deslocamentos da unidade em relação à fronteira tecnológica ou deslocamento da própria fronteira tecnológica devido a evolução da tecnologia. Em outras palavras, observar se as unidades ficaram mais eficientes. Deste modo, permite verificar se houve melhora na eficiência do rendimento do trabalho nos estados, relativa às áreas rural e urbana, ao longo dos anos.

A dissertação é dividida além desta introdução em mais dois capítulos que constituem os dois ensaios analisados.

2. EVOLUÇÃO DOS DETERMINANTES DA RENDA RURAL DO BRASIL, 1995 a 2007

2.1. Introdução

2.1.1. Problema e justificativa do trabalho

Neste ensaio foi apresentada e analisada a evolução dos determinantes da renda proveniente do trabalho das pessoas residentes no meio rural e meio urbano, de 1995 a 2007. A base de dados utilizada foram os microdados da Pesquisa Nacional por Amostra de Domicílio (PNAD) do Instituto Brasileiro de Geografia e Estatística (IBGE). Partiu-se da hipótese de que as transformações no campo e na cidade afetaram a renda dessas pessoas, alteraram o seu perfil ocupacional modificando os condicionantes da renda.

Como destacado, as análises pertinentes têm como base as transformações do meio rural. Entre os fatores que explicam estas mudanças, destaca-se a modernização agrícola que, desde 1980, vem permutando trabalho por capital, promovendo o processo, denominado por especialistas de mercantilização do tempo livre, onde as famílias rurais procuram alternativas para complementarem sua renda. Neste sentido, as atividades como a produção de doces caseiros, artesanato, cultivo de flores, ecoturismo, dentre outras, são transformadas em fontes de renda, alterando o perfil ocupacional destas famílias e, deste modo, viabilizando a permanência no campo.

Além da modernização, a agricultura brasileira passou, na década de 1990, por transformações estruturais importantes, tais como, redução de incentivos do governo, abertura comercial, sobrevalorização do câmbio, queda do preço dos produtos agrícolas, que impactaram de forma negativa a agricultura, diminuindo sua importância quanto à

geração de empregos. De acordo com Graziano e Del Grossi (2001), em 1998 o total das rendas não-agrícolas do Brasil superou a soma das rendas agrícolas recebida pelos moradores rurais. Ou seja, no século atual as atividades agropecuárias já não respondem pela maior parte da renda da população rural. Diniz e Silveira (2009) observaram que, no final dos anos 1990, a renda do trabalho não agrícola chegou a responder por 42% da renda total, contra apenas 38% da renda de atividades agrícolas.

No meio urbano, as mudanças não são tão expressivas quanto no rural em termos de ocupação (ramo de atividade). Não obstante, algumas características quanto aos determinantes da renda foram destacadas em algumas pesquisas, como a homogeneização da educação (Morley, 2001). Ademais, a análise do meio urbano imprime credibilidade aos resultados. Em outras palavras, dificulta cometer erros como, por exemplo, concluir que os resultados encontrados são particulares ao rural e não uma mudança geral, apresentada nos dois espaços.

Diante do exposto acima, o presente ensaio investiga se essas mudanças no perfil ocupacional provocaram alterações nos determinantes da renda ao longo dos anos. Para este propósito estimam-se equações de rendimento que objetivam identificar a relação do salário com seus determinantes. Estas equações revelam a sensibilidade do salário diante das variações de seus determinantes, medindo precisamente o impacto deles. Em outras palavras, permite saber quanto que, em média, um ano a mais de estudo representa de acréscimo no salário do trabalhador. Outro exemplo, mostra se o fato do trabalhador ser branco ou negro interfere no seu rendimento. Segundo a literatura especializada⁵, anos de estudo e idade, características produtivas, são os principais determinantes da renda. Outros fatores, como ramo de atividade, posição na ocupação, cor e raça, características não-produtivas, são determinantes secundários. Ambos podem variar ao longo dos anos diante de mudanças no perfil ocupacional.

É importante ressaltar que o presente ensaio enfatiza o estudo do retorno da educação e o impacto da posição na ocupação ao longo dos anos no meio rural e meio urbano. As demais variáveis foram incluídas para controlar os seus efeitos sobre a participação do indivíduo no mercado de trabalho e sobre a sua remuneração, visando à correta especificação do modelo.

⁵ Para acompanhamento de uma extensa revisão bibliográfica dos determinantes de renda ao longo dos anos no Brasil vide Coelho e Corseuil (2002)

Neste âmbito, seguem abaixo alguns resultados já levantados por estudiosos sobre determinantes de renda. Nestes estudos a metodologia usualmente empregada é o de mínimos quadrados ordinários - MQO. No presente ensaio, ao invés do MQO utiliza-se o método máxima pseudo-verossimilhança a fim de corrigir o viés de seletividade. Esta metodologia está sendo difundida recentemente no Brasil, sendo assim na presente revisão são apresentados trabalhos cujo método é o MQO, e no final os que utilizam máxima pseudo-verossimilhança.

O primeiro trabalho no Brasil foi o de Langoni (1973) que retratou as relações entre diferenciais de salário e nível educacional, idade, gênero, setor de atividade e região de residência. A investigação mostra que essas características são determinantes dos diferenciais salariais, destacando que disparidades educacionais existentes entre os trabalhadores constituem-se no principal fator determinante desses diferenciais.⁶

No rural, esse diferencial de educação é mais expressivo. Dabos e Psacharopoulos (1991) estimam que, em 1980, a taxa média de retorno à educação de homens na área rural é de 3,4 pontos percentuais superior à observada na área urbana.

Hoffmann e Simão (2005), a partir dos dados do Censo Demográfico de 2000, estimaram equações de rendimento para avaliar o efeito de vários fatores (idade, sexo, escolaridade, posição na ocupação, cor, tempo semanal de trabalho *etc.*) sobre o rendimento das pessoas ocupadas em Minas Gerais. Na especificação das equações de rendimento, mostraram que é inapropriado admitir que o efeito da escolaridade possa ser captado por uma função linear, uma vez que ocorre substancial aumento da taxa de retorno da escolaridade a partir do limiar de 10 anos de estudo. Para contornar tal problema, os autores utilizaram o efeito limiar da educação, que será adotado no presente ensaio.

No que se refere às características não-produtivas, destaca-se a posição na ocupação. Segundo estimativas de Ney e Hoffmann (2003) os resultados apresentados em seu estudo são suficientemente robustos para reafirmar a grande importância do capital físico como determinante do rendimento das pessoas ocupadas no rural em 2000. A renda esperada dos empregadores é bem maior do que a dos conta própria e empregados, e uma condição fundamental para alguém pertencer àquela posição ocupacional, principalmente à de grande empregador, é possuir uma quantidade relativamente grande de meios de produção, entre eles a terra.

⁶ A partir da década de 1980, quando algumas instituições que possuíam computador de grande porte passam a ter acesso às pesquisas domiciliares do IBGE, é que as investigações desse tema são retomadas. Sendo mais disseminadas em meados da década de 1990 quando IBGE passou a disponibilizá-los em CD-ROM.

Em consonância, trabalhos como o de Hoffmann e Simão (2005) demonstraram que em 2000 no setor agrícola a posição na ocupação exercia maior influência que a variável educação na explicação da renda, ao contrário dos demais setores. Foram apontadas diferenças importantes quando são ajustadas equações separadas por setor, (agrícola, indústria e serviços) cabendo destacar que, no setor agrícola, a variável que mais contribuiu para explicar as variações no logaritmo do rendimento das pessoas ocupadas foi a posição na ocupação, superando até mesmo a contribuição da escolaridade, que é a variável mais importante nos setores industrial e de serviços.

Sendo assim, o aumento das atividades não-agrícolas poderá afetar os determinantes da renda rural provocando um crescimento da importância da educação na explicação da renda do campo, assim como uma diminuição da importância da posição na ocupação. Diferentemente da agricultura, outros setores da economia geralmente demandam mão-de-obra mais especializada, ou seja, pessoas que em média possuem anos de estudos a mais que as ocupadas em atividade agrícola. Contudo, pesquisas como as de ROCHA *et al.* (2005) relatam que as novas ocupações não-agrícolas em Minas Gerais são muitas vezes de baixa qualificação como a de pedreiro e a de empregada doméstica, comparativamente, no geral ainda são mais especializadas. Resultados semelhantes foram encontrados para São Paulo por Kageyama (2003).

Aliado a estes fenômenos tem-se fatos já consolidados. Estudos, como os de Rocha (2000) e de Barros e Mendonça (1995), demonstraram que há uma grande disparidade na alocação da renda segundo a localização do domicílio, quando se compara os urbanos e os rurais. Em sintonia, os dados levantados por Hoffmann (2000) mostravam que, em 1998, o rendimento médio familiar *per capita* na área rural (R\$ 102,9) correspondia a apenas 35% do rendimento médio na área urbana (R\$ 292,4).

Menezes *et al.* (2002) verificaram a extensão dos diferenciais regionais de rendimentos sobre o mercado de trabalho entre as regiões metropolitanas de São Paulo (RMSP) e de Salvador (RMS) utilizando o procedimento de Heckman. Os resultados mostraram um importante diferencial regional de rendimentos, para todos os agrupamentos de trabalhadores analisados, em favor dos trabalhadores da RMSP. Tais diferenciais de rendimento são decorrentes das sinergias vinculadas à maior concentração e aglomeração econômica da RMSP, relativamente à RMS. Observou-se ainda um diferencial significativo de rendimentos, também em favor dos trabalhadores da RMSP.

Como destacado acima, a maioria dos estudos utilizavam o MQO. As pesquisas relatadas abaixo, utilizaram o método máxima pseudo-verossimilhança, mostrando a importância do emprego do método. Quando se utiliza os microdados da PNAD o método de Mínimos Quadrados Ordinários (MQO) torna-se inadequado, uma vez que a amostra não se distribui aleatoriamente, ou seja, não se trata de uma amostra independente e identicamente distribuída (iid), o que gera tendenciosidade nas estimativas. Outra correção, trata-se de contornar o viés de seleção amostral por meio do procedimento de Heckman (1979)⁷.

Segundo Moura⁸ (2008), a incorporação do desenho amostral da PNAD é um ganho adicional na literatura empírica para o Brasil, a qual não tem sido considerada na estimação de diversos modelos econômicos. Neste mesmo trabalho, o autor também corrige o problema de viés de seleção amostral através do procedimento de Heckman que surge devido ao fato de que alguns indivíduos escolhem não trabalhar pois o salário de mercado é fixado abaixo do seu salário de reserva. Estas correções corroboraram os testes feitos e influenciou consideravelmente na mensuração correta dos retornos.

Carvalho *et al.* (2006) também corrigiram o viés de seleção através do mesmo procedimento, bem como destacaram a diferença entre a amostragem simples e a amostragem complexa. A primeira traz o pressuposto que qualquer indivíduo tem igual chance de ser selecionado, já a segunda é caracterizada por um plano amostral complexo elaborada considerando a estratificação, conglomeração, probabilidades desiguais de seleção em um ou mais estágios e calibração da amostra através de pesos amostrais.

Resende e Wyllie (2006) investigaram econometricamente os retornos para educação no Brasil, tomando como referência dados da Pesquisa sobre Padrão de Vida do IBGE. As equações de rendimentos para homens e mulheres foram consideradas mediante o procedimento para correção do efeito de seletividade amostral concebido por Heckman, incluindo uma medida de qualidade para educação. Os resultados obtidos revelaram que os retornos para educação ficariam superestimados se não fosse considerada a seletividade amostral.

Diante da revisão bibliográfica exposta acima se pode constatar que são muitas as pesquisas sobre determinantes de renda no Brasil. Não obstante, não se verificou estudo

⁷ Quando a equação-objeto possui erro de especificação incorpora-se uma equação adicional para determinar a ocorrência, ou não, do evento e contornar o viés de seletividade.

⁸ Neste trabalho o autor estima a taxa interna de retorno (TIR), um conceito central da teoria do capital humano, desenvolvido na análise desta decisão dos investimentos dos agentes em capital humano.

específico de diferenças entre o rural e o urbano. Assim, serão elaboradas equações de rendimento divididas para cada segmento, urbano e rural, ao longo dos anos. Neste âmbito, pretendeu-se contribuir para a discussão das diferenças entre o mercado de trabalho urbano e rural brasileiro a fim de fomentar idéias para o desenvolvimento econômico e social específico destes setores censitários, uma vez que a diferença de renda rural e urbana se constitui em indicador de desenvolvimento social. Os resultados econométricos da pesquisa trazem expressiva confiabilidade ao abarcar conjuntamente o limiar da educação, o plano amostral e a amostra seletiva.

2.1.2. Hipóteses e Objetivos

2.1.2.1. Hipótese

Parte-se da hipótese de que se as transformações no campo e nas cidades afetaram de forma expressiva a renda das pessoas residentes nestes locais e alteraram o seu perfil ocupacional. É esperado, então, um aumento da similaridade entre os condicionantes da renda do meio rural e meio urbano brasileiro. Ou seja, espera-se, por exemplo, que um ano a mais de estudo influencie da mesma forma a renda do trabalhador rural e do trabalhador urbano.

Espera-se evolução positiva no limiar da educação no rural, devido às novas ocupações que necessitam de maior nível educacional, contrastando com as ocupações agrícolas. Em outras palavras, espera-se que o efeito diploma do rural seja o mesmo do urbano.

Se há uma interseção cada vez maior entre o rural e urbano então é esperado que os diferenciais salariais entre o rural e o urbano arrefeçam, uma vez que os mesmos passam a ter mais características de similaridade quanto às ocupações.

2.1.2.2. Objetivos

Objetivo geral:

Analisar a evolução dos determinantes do rendimento advindo do trabalho das pessoas residentes no rural e no urbano frente à disseminação de novas ocupações no campo, cotejando os resultados do meio rural com o que se verifica na cidade.

Objetivos específicos:

- traçar o perfil ocupacional (agrícolas e não agrícolas) dos moradores rurais e urbanos para os anos que a pesquisa abarca;
- analisar os determinantes da desigualdade de rendimento no Brasil, bem como a evolução destes ao longo dos anos;
- determinar o limiar do urbano e rural, comparando os resultados encontrados;
- aplicar técnicas econométricas avançadas, como a mostra seletiva de Heckman e incorporação do plano amostral da PNAD.

2.2. Referencial Teórico

2.2.1. Teoria do Capital Humano

A Teoria do Capital humano começou a ser analisada a partir de Adam Smith (1776), no livro “Uma investigação sobre a natureza e as causas da riqueza das nações” e Alfred Marshall (1920), em “Princípios de Economia”. Porém, as análises do investimento em educação e treinamento desses estudiosos não incluíam em suas discussões a produtividade.

Foi com Gary Becker (1964), Jacob Mincer (1958) e Theodore Schultz (1961) que se tem a maior contribuição para explicar os determinantes dos diferenciais salariais.

Theodore W. Schultz, entre outros, iniciaram uma exploração pioneira nas implicações do investimento do capital humano nas questões econômicas. Schultz coordenou uma grande pesquisa em escala mundial desenvolvida pela UNESCO, na década de 1950. O objetivo era realizar um levantamento acerca da situação do ensino de

países pobres e ricos de todos os continentes do planeta. Os resultados desta pesquisa mostraram que quanto maior a riqueza de um país, maior o nível de instrução de sua população, e melhor, conseqüentemente, a sua situação de ensino. Schultz concluiu que a Educação era o fator que fazia a diferença no que diz respeito à riqueza de um país, pois alguns dos países mais pobres possuíam grande volume de recursos e riquezas naturais. Em outras palavras, verificou-se uma relação positiva entre escolarização de uma nação e sua riqueza.

Mincer (1958) estudou o efeito da experiência do trabalhador na determinação dos salários. Seu modelo analisa como diferenças na experiência podem influenciar nas desigualdades de renda. Schultz (1961), por sua vez, analisa os impactos da educação no rendimento das pessoas, abordando as capacidades adquiridas pelos indivíduos como fator de aumento de produtividade. Os pensamentos desses autores também eram compartilhados por Becker (1962).

De acordo com Gary Becker (1962), capital humano é qualquer atividade que implique custo no período corrente e que aumente a produtividade no futuro. É o conjunto das habilidades do indivíduo ligadas à capacidade produtiva, e incorporadas no conhecimento e qualificação para determinadas tarefas.

A relação positiva entre rendimento e escolaridade é um dos padrões empíricos mais bem estabelecidos na literatura de economia do trabalho e da educação. Ou seja, vários estudos⁹ reforçam a evidência de que a escolaridade eleva a renda individual.

Essas habilidades produtivas, supramencionadas, podem ser inatas ou adquiridas por meio de educação e treinamento. Sendo assim, as pessoas investem em educação para elevar o fluxo futuro de renda esperada. Esse fluxo seria determinado pelo melhor uso alternativo de sua capacidade produtiva, ou seja, ofertar trabalho na atividade que lhe proporciona maior rendimento. Uma diferença entre capital físico e capital humano refere-se ao fato de não existir um mercado para esse último. Isso acontece porque capital humano inalienável, indissociável do indivíduo, dificultando o seu investimento.

⁹Para a teoria do Capital Humano, quanto maior for o grau de escolaridade, maior serão as oportunidades no mercado de trabalho, isto dentro de um meio econômico equilibrado, havendo uma relação de causalidade, onde educação explica as oportunidades.

2.2.1.1. Críticas à Teoria do Capital Humano

Embora a teoria do capital humano esteja consolidada, ela ainda possui algumas críticas. As primeiras delas foram instigadas por John Stuart Mill, em seu livro *Princípios de Economia Política com Algumas de suas Aplicações à Filosofia Social*. Para Mill, as diferenças dos salários não refletem, necessariamente, desigualdades produtivas relacionadas à educação e treinamento. Em sua análise, Mill enfatiza a existência de barreiras à entrada nas ocupações de altos salários.

Diante deste debate, tem-se a divisão em neoclássicos e não-neoclássicos. Os seguidores da primeira corrente acreditam que o mercado funciona relativamente bem e as imperfeições de mercado, a não ser quando impostas pelo governo, possuem um papel secundário na explicação dos diferenciais de salários. De acordo com a análise neoclássica do capital humano, os salários estariam relacionados à produtividade do indivíduo, de forma que trabalhadores igualmente produtivos (mesmo nível educacional e experiência) receberiam rendimentos iguais.

Os da segunda ressaltam o papel fundamental das imperfeições de mercado que impedem a mobilidade dos trabalhadores dos postos de trabalho com baixos salários para aqueles com altos salários. Segundo Langoni (1973), Thurow (1975) e Grandra (2004) atributos como raça, sexo, região geográfica, setor censitário (rural e urbano) e ramo de atividade (atividades-agrícolas e não agrícolas) são também variáveis determinantes no processo de seleção de trabalhadores. Sexo e raça, por exemplo, estariam relacionadas à discriminação. Hoffmann e Leone (2004) verificaram que as ocupações menos valorizadas e tradicionalmente femininas do mercado de trabalho continuam se reproduzindo, implicando a persistência de nichos ocupacionais, como o do emprego doméstico. Quanto a cor, a mais discriminada é a raça negra que, assim como a população feminina ocupada, tende a ocupar cargos de baixa remuneração. As demais variáveis determinam a segmentação¹⁰ do mercado de trabalho e, por conseguinte, os diferenciais salariais.

¹⁰Por *segmentação* entende-se a diferença de remuneração entre dois postos de trabalho distintos, ocupados por trabalhadores idênticos, com base em características do posto de trabalho, apesar das produtividades daquelas combinações de posto e trabalhador serem idênticas. Assim, por exemplo, diz-se haver segmentação regional quando dois postos iguais, ocupados por trabalhadores iguais, geram remunerações distintas em diferentes regiões geográficas do país. Analogamente, o significado é semelhante para segmentação setorial ou entre segmentos formais e informais do mercado de trabalho (Ferreira, 2000).

Atualmente esses aspectos têm ganhado notório destaque na tentativa de explicar a alta desigualdade de rendimentos nos países da América Latina.

Não obstante, o arcabouço analítico neoclássico (ortodoxo), é consolidado e dominante. Trata-se de modelos mais desenvolvidos e elaborados. Por outro lado, os argumentos levantados pela literatura não-neoclássica são passíveis de ser tratados com o instrumental analítico tradicional. Neste sentido, o presente ensaio faz uso da teoria neoclássica do capital humano adaptando, a ela, a teoria não-neoclássica. Isso acontece, ao acrescentar nas equações de rendimento estimadas variáveis referentes a características não-produtivas¹¹. A teoria do capital humano aplicada à análise de regressão fornece um arcabouço para avaliar como decisões individuais influenciam nos retornos dos rendimentos.

2.2.2. Externalidades negativas

Nas primeiras décadas do século XX a sociedade brasileira se configurava como amplamente rural. Em decorrência ao expressivo crescimento da população verificado no país e ao processo de industrialização, entre 1940 e 1980, observou-se uma inversão da distribuição populacional entre as áreas rurais e urbanas, provocando o esvaziamento das áreas rurais, o crescimento desordenado de grandes cidades e a formação de centros metropolitanos. De acordo com Marx (1994), na fase seminal da industrialização, não se tinha pleno domínio sobre a energia gerada para o funcionamento das máquinas, sendo assim a sede das indústrias eram nos locais de fácil acesso às forças motrizes de origem animal ou natural como quedas d'água, moinhos de vento, ou seja, no campo. Somente após o controle da técnica a indústria tornou-se espacialmente independente, podendo se instalar nos centros urbanos. Neste momento, segundo Graziano (1999), consumou-se a separação entre cidade e campo.

Apesar da teoria de Malthus (1996) subestimar o papel do progresso tecnológico, ela foi a primeira a identificar, indiretamente, a diferença entre desenvolvimento e crescimento desordenado. Um crescimento populacional sem planejamento pode ser considerado um entrave ao desenvolvimento, na medida em que gera sérios problemas de poluição e violência, induz processos migratórios e contribui enormemente para o inchaço das grandes cidades. Embora o processo de urbanização criasse as economias

¹¹Maiores detalhes podem ser acompanhados na metodologia.

de escala, segundo a teoria de Marshall (1920), com consequentes ganhos de produtividade, quando se dá de forma rápida e desorganizada, gera-se também as externalidades negativas que diminuem o bem-estar social e elevam as desigualdades.

O rápido crescimento das cidades após a revolução industrial deu origem ao que se denominou de êxodo rural. As cidades ganharam destaque pelos benefícios que trouxeram, como a maior oferta e concentração de bens e serviços públicos e privados que geram eficiência e eficácia para a economia. Ou seja, a cidade ao dispor uma logística que oferecia competitividade econômica, tornou-se atrativa para os moradores rurais. Aliada a essa atratividade urbana, teve-se a modernização agrícola que, ao permutar trabalho por capital, reduziu significativamente a utilização da mão-de-obra disponível das famílias rurais. Neste âmbito, as cidades tornaram-se cada vez mais populosas. De acordo com Friedmann (2007), este processo teve início desde princípios de 1800, situando-se na primeira década do séc XXI pela primeira vez acima dos 50%, em nível mundial.

Contudo, essa atratividade urbana vem se arrefecendo. Segundo Martins *et al.* (2005), muitas cidades de maior dimensão dos países desenvolvidos foram ao longo das décadas de 1960 e 1970 atingidas por um processo de reestruturação setorial marcado pela crise das atividades industriais e dos modos de organização que favoreceram o crescimento nas décadas anteriores.

Por outro lado, com as novas ocupações que estão se consolidando no rural, tem-se a viabilização da permanência das pessoas no campo, ou até mesmo um movimento reverso de migração, uma vez que o urbano não oferece atratividade expressiva. De acordo com Matos *et al.* (2004), as mudanças que vêm ocorrendo nas áreas rurais brasileiras rompem o mito do esvaziamento do campo considerado inevitável por alguns pesquisadores no passado. Hoje, mesmo que ainda exista algum êxodo rural, é notória a tendência de recuperação de parte expressiva das áreas rurais no país.

Segundo Chapuis (1991), Kageyama (2003) e Graziano (1999), o meio rural não pode mais ser estudado como um universo fechado, direcionado estritamente para atividades agrícolas. De acordo com Saraceno (1994), a agricultura era o meio de vida predominante no período pré-industrial, no qual a taxa de atividade agrícola era um bom indicador da ruralidade, o que não se pode dizer do período pós-industrial.

Essa ruptura do conceito de rural vinculada à agricultura é devido ao crescimento das atividades não-agrícolas nas áreas delimitadas como rurais representando um dos mais

importantes fenômenos dos anos 1990, tanto nas regiões desenvolvidas como EUA e Europa, como na América Latina, de modo geral. No Brasil, os especialistas do projeto RURBANO e demais pesquisadores vêm estudando este fenômeno, onde muitos autores, como Neder (2003) e Rocha (2005), concluem que essas novas ocupações não-agrícolas servem como alternativa de trabalho para as famílias rurais viabilizando sua permanência no campo. Em suma, essas novas atividades geraram um dinamismo para a região rural fazendo com que o mercado de trabalho se aquecesse gerando mais postos de trabalhos, o que por sua vez podem repercutir nos diferenciais salariais.

2.2.3. Rural e Urbano¹²

Existem duas abordagens que definem rural e urbano, a dicotômica que se refere ao campo como algo que se opõe à cidade e a de *continuum* que enfatiza a aproximação entre o espaço rural e a realidade urbana. De acordo com Blume (2004), os primeiros debates e reflexões da corrente dicotômica eram embasados em estudos das diferenças entre comunidade e sociedade, de Ferdinand Tönnies. De acordo com os clássicos Marx e Weber, antes da disseminação do capitalismo urbano industrial, originou-se um conflito entre duas realidades sociais diferentes: a urbana, símbolo de incorporação do capitalismo e do progresso da técnica, e a rural, refratária ao capitalismo, refúgio da aristocracia decadente e de antigas relações e formas de vida.

Sorokin & Zimmermann (1929) foram os primeiros a introduzir a perspectiva do *continuum* rural e urbano. Essa corrente considera o rural e o urbano como pólos extremos em uma escala de gradação. Mais tarde, Redfield (1947) observou a intensificação das relações rurais e urbanas, com o intuito de evidenciar uma maior relação entre os espaços opostos, diluindo ainda mais as diferenças verificadas pelos clássicos nas primeiras décadas do século XX. A obra clássica de Sorokin & Zimmermann (1929) sugere uma série de traços essenciais na diferenciação dos espaços urbano e rural. As definições clássicas do conceito de rural partem de uma enumeração de vários aspectos da realidade que seriam indicadores da situação do local estudado. Os traços que caracterizariam a sociedade rural têm como base a dimensão econômica, caracterizando-se por desenvolver um determinado tipo de atividade: a produção de alimentos por meio da criação de plantas e de animais. A esta atividade econômica estão vinculadas todas as características que

¹² Esta seção é baseada no artigo de Reis (2006).

definem o rural, como a diferença ambiental, uma vez que sofrem influência direta e constante do contato com a natureza e das condições climáticas. O meio rural seria também mais arremido à técnica, uma vez que se trabalha com organismos vivos e com forças da natureza, que não podem ser inteiramente controlados. A população rural tende, também, a adquirir características semelhantes por se envolverem nas mesmas funções, tornando-se mais homogêneos, pois não sofrerem os problemas de uma intensiva divisão do trabalho. Ao contrário, o meio urbano é sinônimo de desigualdade.

Nas últimas décadas, contudo, tornou-se cada vez mais complexo delimitar o que é rural e o que é urbano. Do ponto de vista da organização da atividade econômica, as cidades não podem mais ser identificadas apenas com a atividade industrial, nem os campos com a agricultura. A partir dessas formulações teóricas, as definições de rural e urbano foram alvos de intensos debates. De acordo com Blume (2004), diante dos embates teóricos, das crescentes indefinições e também pela crescente urbanização, o rural deixou de ser atrativo como objeto de pesquisa, pois começou a ser aceito que a urbanização do campo era uma questão de tempo.

Contudo, segundo Reis (2006) debates que permeiam os diferentes níveis de relacionamento entre o rural e o urbano ressurgiram, diante da rapidez e magnitude das mudanças técnicas, econômicas, sociais e ambientais que atingiram estes espaços a partir da segunda metade do século XX. Para Graziano (2001), as mudanças verificadas no meio rural brasileiro, a partir da década de 1980, com o aumento da urbanização do campo e o surgimento de novas dinâmicas que apontam para uma nova ruralidade, muitas vezes distante do contexto das atividades agropecuárias, proporcionaram a retomada e a intensificação desses debates, diante de sua crescente importância.

2.3. Metodologia

2.3.1. Equações de Rendimento

Esta seção tem como finalidade explicar a principal equação utilizada no presente ensaio, pela qual se dará a análise da evolução dos determinantes do rendimento advindo do trabalho das pessoas residentes no rural e no urbano frente à disseminação de novas ocupações no campo e na cidade. Essa equação de rendimento ou equação de salários é denominada pela estatística de regressão que é uma importante ferramenta econométrica

para se estudar a dependência de uma variável, o regressando, em relação a uma ou mais variáveis explanatórias. Muito utilizada para demonstrar os efeitos de variáveis econômicas e sócio-demográficas sobre o rendimento.

Optou-se pela escolha das equações de rendimento semelhante às utilizadas por Hoffman (2000) e Morley (2001). Nestas, além de se incorporar variáveis da teoria do capital humano, utilizadas nas funções de salários por Jacob Mincer (1958 e 1974) e Becker (1964), considera-se o fato de que o mercado de trabalho pode apresentar barreiras à mobilidade dos trabalhadores. E, estas não são reflexos apenas de diferentes níveis de educação e habilidades por parte do trabalhador. Ou seja, nestas equações também se considera as características não produtivas¹³.

Apesar de serem muito utilizadas, as equações de rendimento possuem algumas particularidades que dificilmente podem ser contornadas, quais sejam: heterocedasticidade, erros de especificação e viés de seletividade.

Segundo Hoffman (2000), a primeira é devido, principalmente, à assimetria na distribuição da variável renda. Segundo o Relatório de Desenvolvimento Humano, de 1999, do Programa das Nações Unidas para o Desenvolvimento (PNUD), somente a África do Sul e Malawi têm um grau de desigualdade maior que do Brasil, segundo o Índice de Gini, de forma que é esperado que as variâncias dos erros não sejam constantes. Para respeitar tal propriedade da renda emprega-se o logaritmo neperiano.

De acordo com o mesmo pesquisador, os erros de especificação se dão devido à omissão de variáveis relevantes como ambição, tino comercial, capacidade empresarial, dentre outros, cuja mensuração é praticamente impossível. Sendo assim, os valores dos coeficientes de determinação (R^2), em equações de rendimento, são normalmente baixos, cerca de 0,50, pois as rendas das pessoas dependem de elementos aleatórios e de características pessoais. Não obstante, considerando que se trata de uma amostra expressiva, esse problema tende a ser minimizado.

Quanto ao viés de seleção, este será removido por meio do Procedimento de Heckman¹⁴ (1979). Este viés de seleção é devido ao fato de abarcar apenas os ocupados com renda positiva, excluindo desempregados e inativos. Caso se proceda à estimação da função proposta, os resultados podem apontar conclusões viesadas e errôneas.

¹³ Maiores detalhes sobre características produtivas e não produtivas serão apresentados na seção subsequente.

¹⁴ Será demonstrada na subseção 2.3.3. a metodologia do Procedimento de Heckman.

Foram utilizados dados em *cross-section*, pois nas PNADs a amostra não pode ser controlada uma vez que a base de dados segue o princípio da anonimidade; ademais, a amostra não permanece a mesma durante os anos. Outrossim, é comum na literatura o uso de *cross-sections* para não captar influências dos ciclos econômicos como acontece quando se usa dados em painel. De acordo com Mincer (1993), os rendimentos ao longo do ciclo de vida tendem a aumentar de forma crescente na primeira fase da vida produtiva do trabalhador, a aumentar de forma decrescente na segunda fase e a se manter praticamente constante na última fase. Tal comportamento refere-se, contudo, aos perfis de rendimentos de grupos homogêneos, livres de flutuações e de tendências ao longo do tempo. Sendo assim, para observar tais perfis, empiricamente, é necessário realizar análises de dados *cross-section*, uma vez que dados longitudinais dos rendimentos médios dos grupos poderão ser afetados pelas tendências e ciclos da economia.

2.3.2. Seleção das variáveis para a elaboração da equação de rendimento

Utilizou-se como variável dependente, o logaritmo da razão do rendimento de todos os trabalhos das pessoas ocupadas e das horas trabalhadas. É relevante destacar que foram retirados da amostra os indivíduos que não declararam suas rendas nas análises pertinentes à equação de rendimento, mas mantendo-os na análise descritiva.

Segundo Hoffmann e Simão (2005), informações sobre rendimento obtidas por meio de questionários, como nos Censos ou nas PNADs, produzem valores que subestimam a renda efetiva, já que as pessoas (especialmente as relativamente ricas) subdeclaram seus rendimentos, e vários tipos de renda real são omitidos. As famílias também possuem rendas provenientes de outras fontes como aluguéis e juros, que não foram computadas devido a sua baixa importância relativa e porque o objetivo principal é analisar o rendimento advindo do trabalho.

Isto posto, empregou-se o logaritmo neperiano da renda dividido pelas horas trabalhadas, uma vez que esta variável tem como característica uma distribuição assimétrica.

Variáveis explicativas – características produtivas

Como supramencionado, é esperado que trabalhadores com maiores escolaridade e experiência sejam mais eficientes e recebam maiores remunerações. Nesse caso, as diferenças de remuneração são apenas a tradução das desigualdades preexistentes em produtividade, sendo essa parcela da desigualdade em remuneração apenas revelada pelo mercado de trabalho.

A idade é incluída na regressão como *proxy* da experiência, ou seja, com o passar dos anos espera-se que os indivíduos se tornem mais especializados e eficientes nas funções que exercem e por isso tenham um retorno financeiro correspondente a tal empenho. Optou-se em medir a idade em dezenas de anos para tornar os coeficientes maiores. Também se utilizou o quadrado dessa variável, uma vez que a renda não varia linearmente com a idade, tendo um comportamento de U invertido. Esta curva revela a depreciação do capital humano, ou seja, a produtividade do indivíduo após atingir um ápice começa a recuar. Se os parâmetros para a idade e idade ao quadrado forem indicados por β_1 e β_2 , respectivamente, deve-se ter $\beta_1 > 0$ e $\beta_2 < 0$ e então o valor esperado da renda será máximo quando a idade da pessoa for igual a $-\beta_1/(2\beta_2)$;

Para captar o efeito da educação, que se trata de uma *proxy* de produtividade, utilizou-se a variável anos de estudo. Nos dicionários da PNAD, ela é dividida em 14 categorias até 1990, quais sejam: (1) sem instrução e menos de um ano; (2) um ano; (3) dois anos; (4) três anos; (5) quatro anos; (6) cinco anos; (7) seis anos; (8) sete anos; (9) oito anos; (10) nove a onze anos; (11) doze ou mais; (12) não determinado; e, (13) não declarado. De 1996 a 2006, mudam-se as categorias a partir da dez: (10) nove anos; (11) dez anos; (12) onze anos; (13) doze anos; (14) treze anos; (15) quatorze anos, (16) quinze anos ou mais; e, (17) não determinados e sem declaração. Optou-se pela exclusão da categoria doze e treze nos anos até 1995, e dezessete de 1996 a 2006 para não causar viés nos resultados por se tratar de uma variável contínua que indica a taxa de retorno, ou seja, o acréscimo percentual no rendimento esperado produzido por cada ano adicional de estudo.

Outra variável utilizada foi o efeito limiar da educação. Segundo Hoffmann e Simão (2005), tem-se um limite a partir do qual os anos de estudo começam a impactar de forma mais expressiva no rendimento. Segundo Ney e Hoffmann (2004) este valor deve se situar em 10 anos de estudo.

$$S^* = Z(S - \lambda) \quad (2.1)$$

onde:

S^* = Efeito Limiar da Educação;

S = Anos de Estudo do indivíduo;

λ = Limiar da educação;

Z = é uma variável *dummy* que assume valor zero para $S \leq \lambda$ e valor 1 para $S > \lambda$.

Considerando que $S = \text{anosdeestudo}_i$, $S^* = ELEDU_i$ e K as demais variáveis das equações de rendimentos, o valor do logaritmo dos rendimentos de todos trabalhos ficará como:

$$\ln Y_i = K + \beta_4 \text{anosdeestudo}_i + \beta_5 ELEDU_i \quad (2.2)$$

Quando $S \leq \lambda$, teremos $Z = 0$ e a equação se reduz a:

$$\ln Y_i = K + \beta_4 \text{anosdeestudo}_i \quad (2.3)$$

Ou seja, desconsiderando o limiar da educação, cada ano adicional de estudo está associado a um aumento de $[\exp(\beta_4) - 1]100\%$ no rendimento dos indivíduos;

Quando $S > \lambda$, temos $Z = 1$, e a expressão (2.3) se tornará:

$$\ln Y_i = K - \lambda\beta_5 + (\beta_4 + \beta_5)S \quad (2.4)$$

Sendo assim, após o nível de limiar cada ano a mais de escolaridade provocará retorno nos rendimentos das pessoas em $[\exp(\beta_4 + \beta_5) - 1]100\%$.

Para o cálculo do limiar do meio rural e urbano foram construídas *dummies* para cada ano de estudo e rodada a regressão de rendimento. O limiar é escolhido quando se tem um aumento brusco de retorno para cada ano adicional de estudo seguido de um aumento crescente.

Variáveis explicativas – características não-produtivas

Segundo Barros *et al.* (2007), nem todas as diferenças em remuneração resultam dessas diferenças intrínsecas de produtividade entre trabalhadores, reveladas apenas pelo mercado de trabalho. Boa parte delas ocorre entre trabalhadores perfeitamente substituíveis no processo de produção, isto é, entre aqueles que, se trocassem entre si os postos que ocupam, não alterariam o nível da produção em nenhum deles. Nesse caso, temos que o mercado de trabalho remunera de forma diferenciada trabalhadores com a mesma produtividade intrínseca e, portanto, gera desigualdades. O mercado gera desigualdade quando remunera de forma diferenciada homens e mulheres, brancos e negros com mesma produtividade. Neste sentido, foram incluídas *dummies* de gênero e cor ou raça para captar se no rural mineiro existe discriminação.

Outra razão para explicar as disparidades em remuneração do trabalho são as que resultam de segmentação na remuneração de trabalhadores com a mesma produtividade. Como a divisão de setores apresentada no presente artigo abarca agrícolas e não-agrícolas, utilizamos uma *dummy* para identificar se no rural existe segmentação. A partir do coeficiente estimado de uma variável explanatória binária podemos obter a diferença percentual entre o rendimento esperado na categoria tomada como base e o rendimento da categoria para a qual aquela variável binária assume valor 1. Adicionalmente, foram utilizadas *dummies* para captar a segmentação regional devido a heterogeneidade brasileira, onde a variável controle será o Nordeste.

Por fim, para a mensuração de uma das mais importantes variáveis responsáveis pela desigualdade de rendimento no rural, a posição na ocupação, utiliza-se duas *dummies* para identificar se as pessoas eram empregadores ou conta-própria.

De acordo com Corrêa (1998), as variáveis incluídas devem captar o efeito do treinamento e da experiência das pessoas (idade e escolaridade), as discriminações (cor, sexo e setor de atividade) e as diferenças de posse de propriedade e riqueza (posição na ocupação). É importante ressaltar que esses efeitos são captados não de forma perfeita, uma vez que não se dispõe de uma medida da qualidade do ensino e que a posição na ocupação é uma *proxy* não adequada para a posse de capital.

Segue abaixo, o modelo de equação estimado:

Equação de participação:

$$\begin{aligned} cond_ativid_i = & \beta_0 + \beta_1 idaded_i + \beta_2 idade2_i + \beta_3 anosest_i + \beta_4 efeito\limiar_i + \beta_5 num_comp_fam_i + \\ & \beta_6 rend_fam_i + \beta_7 brancaD_i + \beta_8 amar_pardD_i + \beta_9 sudesteD_i + \beta_{10} centro - oesteD_i + \\ & \beta_{11} norteD_i + \beta_{12} sulD_i + \beta_{13} urbanoD_i + \beta_{14} masculinoD_i \end{aligned} \quad (2.5)$$

Em que as variáveis:

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

$\beta_1 idaded_i$ = Idade do indivíduo i em dezenas de anos (segundo metodologia proposta em Hoffmann e Simão, 2005);

$\beta_2 idade2_i$ = Idade ao quadrado em dezenas de anos;

$\beta_3 anosest_i$ = Anos de estudo do indivíduo i ;

$\beta_4 efeito\limiar_i$ = Efeito Limiar dos anos de estudo do indivíduo i ¹⁵;

$\beta_5 num_comp_fam_i$ = Número de componentes na família do indivíduo i ;

$\beta_6 rend_fam_i$ = Rendimento familiar mensal do indivíduo i ;

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

$\beta_{13} urbanoD_i$ = Variável *dummy* para setor censitário, sendo a variável de controle o rural;

$\beta_{14} masculinoD_i$ = Variável *dummy* para gênero, em que a categoria de controle é o sexo feminino;

$D_{região_i}$ = 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;

μ_i = termo de erro aleatório.

¹⁵ O modo de construção e definição dessa variável se encontra na seção “3.5 efeito limiar da educação”.

Equação de rendimento:

$$\ln \text{rendimento}_i = \beta_0 + \beta_1 \text{idaded}_i + \beta_2 \text{idade2}_i + \beta_3 \text{anosest}_i + \beta_4 \text{eledu}_i + \beta_5 \text{emprD}_i + \beta_6 \text{cpD}_i + \beta_7 \text{brancaD}_i + \beta_8 \text{amar_pardD}_i + \beta_9 \text{nagricolaD}_i + \beta_{10} \text{sudesteD}_i + \beta_{11} \text{centro - oesteD}_i + \beta_{12} \text{norteD}_i + \beta_{13} \text{sulD}_i + \beta_{14} \text{urbanoD}_i + \beta_{15} \text{masculinoD}_i$$

(2.6)

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;

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

$\beta_1 \text{idaded}_i$ = Idade do indivíduo i em dezenas de anos;

$\beta_2 \text{idade2}_i$ = Idade ao quadrado em dezenas de anos;

$\beta_3 \text{anosest}_i$ = Anos de estudo do indivíduo i ;

$\beta_4 \text{efeitolimiar}_i$ = Efeito Limiar dos anos de estudo do indivíduo i ¹⁶;

$\beta_5 \text{emprD}_i$ e $\beta_6 \text{cpD}_i$ = Conjunto de variáveis *dummy* para identificar a posição na ocupação, quais sejam: empregado, empregador e conta-própria. Empregado foi utilizada como variável de 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 preta foi utilizada como controle. O agrupamento da cor ou raça foi realizado devido à baixa representatividade dos amarelos e indígenas na amostra;

$\beta_9 \text{nagricolaD}_i$ = Variável *dummy* para indicar o ramo de atividade, agrícola ou não-agrícola. Agrícola é a variável de controle.

$\beta_{13} \text{urbanoD}_i$ = Variável *dummy* para setor censitário, sendo a variável de controle o rural;

$\beta_{14} \text{masculinoD}_i$ = Variável *dummy* para gênero, em que a categoria de controle é o sexo feminino;

¹⁶ O modo de construção e definição dessa variável se encontra na seção “3.5 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;

v_i = termo de erro aleatório.

Tabela 2.1. Descrição das variáveis utilizadas e seus efeitos esperados sobre as equações, de participação e de rendimentos.

Variável	Descrição das variáveis	Sinal esperado	
		Participação	Rendimentos
Idade	Idade do indivíduo	+	+
idade2	Idade ao quadrado	-	-
Anosest	Nº de anos de estudos	+	+
Eledu	Efeito Limiar da Educação	+	+
Empr	=1 se a posição na ocupação do indivíduo é empregador	n.a.	+
Cp	=1 se a posição na ocupação do indivíduo é conta-própria	n.a.	+
Branca	=1 se o indivíduo é branco	+	+
amar_pard	=1 se o indivíduo é pardo ou amarelo	+	-
nagricola	=1 se o ramo de atividade é não-agrícola	n.a.	-
Norte	=1 se o indivíduo mora na região Norte	+	-
nordeste	=1 se o indivíduo mora na região Nordeste	+	-
centro-oeste	=1 se o indivíduo mora na região Centro-oeste	+	+ ou -
Sul	=1 se o indivíduo mora na região Sul	+	+ ou -
Urbano	=1 se o indivíduo mora em cidades ou vilas	+	+
masculino	=1 se o sexo é masculino	+	+
num_comp_fam	Nº componentes na família	-	n.a.

Fonte: Elaborada pela autora.

Nota: n.a. corresponde aos casos não aplicáveis, ou seja, a variável não faz parte da equação.

Sinais positivos das variáveis na tabela 2.1 na coluna de rendimento indicam que o indivíduo que possui a característica em pauta, tem impactos positivos em seu rendimento. Sendo o contrário para sinais negativos. Já os sinais positivos nas equações de participação, indicam que o indivíduo que possui tal característica tem maior probabilidade de estar (participar) no mercado de trabalho.

2.3.2. Métodos empregados

Geralmente, regressa-se o salário em função de variáveis explicativas somente das pessoas ocupadas, desconsiderando-se desempregados e inativos. No entanto, os desempregados e inativos podem não possuir empregos remunerados em decorrência de apresentarem salários de reserva mais elevados que os salários de mercado. A escolha entre participar ou não do mercado de trabalho é tomada pelos trabalhadores em um processo de maximização dos seus benefícios. O trabalhador decide entrar (participar) ou não no mercado com base no salário reserva. Ao considerar apenas as pessoas ocupadas, a amostra pode não ser necessariamente, representativa de toda a população, havendo, por conseguinte, um viés de seleção, que conduzirá a conclusões falsas. Neste sentido, a amostra de indivíduos que participam do mercado de trabalho deve ser considerada como não-aleatória, onde se faz necessária a correção do viés de seletividade amostral. Essa correção se fará com base no procedimento de Heckman.

Outro problema em relação à amostra é o esquema de seleção da amostra da PNAD. Não se trata de uma amostra aleatória simples, pois é estratificada e conglomerada com um, dois ou três estágios de seleção, de acordo com o tipo de estrato. Seria muito custoso para o IBGE realizar uma pesquisa amostral baseada em um desenho amostral simples. Por isso, segundo utilizam-se desenhos amostrais complexos para controlar tais custos. De acordo com Lumley (2004), muitos estudos que utilizam os microdados da PNAD não consideram o desenho amostral complexo ao estimar a variância, na Chromy e Abeyasekara (2005), construção de intervalos de confiança e de testes de hipóteses. Essa negligência pode gerar estimativas viesadas e invalidar os testes de hipóteses usuais. Quando imprecisos, os resultados podem acarretar em mudanças tanto quantitativa quanto qualitativa, ao alterar a significância dos parâmetros estimados.

Como destacado acima, para obter resultados não viesados incorpora-se o plano amostral e a correção do viés de seleção da informação dos salários, visando melhorar a

qualidade das estimativas das equações de rendimento. Para tal propósito, o método dos Mínimos Quadrados Ordinários (MQO) torna-se inadequado. Para abarcar ambas as correções, utiliza-se o método máxima pseudo-verossimilhança.

2.3.3. Procedimento de Heckman

Heckman (1979) propõe o método de dois estágios para contornar o problema de seleção. Inicialmente, estima-se uma equação de participação por meio do Modelo Probit, para observar qual a probabilidade de uma pessoa participar do mercado de trabalho, seja esta pessoa ocupada, desempregada ou inativa. Em seguida, com o uso dos estimadores encontrados, calcula-se lambda (λ), denominada de razão inversa de Mills. Essa nova variável é incorporada à equação de rendimento para que corrija o viés de seleção. É interessante observar que para a equação de participação no mercado de trabalho, são considerados todos os indivíduos e, para a equação principal, somente os que auferem algum rendimento.

Matematicamente, o método é assim realizado:

A equação de participação ou de decisão entre estar, ou não, no mercado de trabalho é dada por:

$$L_i^* = \gamma' Z_i + u_i \quad , \quad (2.7)$$

Onde: L^* = variável latente que indica a participação no mercado de trabalho

Z_i = vetor do conjunto de variáveis que explicam a participação no mercado de trabalho.

Embora L^* seja latente, a variável L pode ser observada, em que:

$$\begin{aligned} L_i = 1 & \quad se \quad L_i^* > 0 \\ L_i = 0 & \quad se \quad L_i^* \leq 0. \end{aligned}$$

A equação acima, obtida pelo método Probit¹⁷, estima a probabilidade de o indivíduo estar ocupado (L) em função de um conjunto de variáveis (Z). L tem valor 1

¹⁷ No modelo probit assume-se que a variável latente L^* possui distribuição normal padronizada (GREENE, 2003). A partir desta hipótese é possível que a probabilidade do indivíduo estar ocupado esteja sempre situado entre 0 e 1.

quando o indivíduo participa do mercado de trabalho, e 0, caso contrário. Nesta etapa, investigam-se como as características socioeconômicas influenciam na decisão do indivíduo participar do mercado de trabalho, como por exemplo, educação, idade, sexo, cor, condição familiar dentre outras.

No que concerne à variável dependente da equação de rendimento tem-se:

$$W_i = \beta' R_i + v_i, \quad (2.8)$$

Onde: R_i = vetor de variáveis explicativas do salário.

W = salários, e é observada quando L^* for maior que zero, ou seja, quando os indivíduos estão ocupados e, por conseguinte, auferem rendimento. Partindo-se do pressuposto de que u_i e v_i apresentam distribuição normal, com médias zero e desvios-padrões σ_u e σ_v , com correlação igual a ρ , tem-se:

$$\begin{aligned} E(W_i | W_i \text{ é observado}) &= E(W_i | L_i^* > 0) = E(W_i | u_i > -\gamma' Z_i), \\ &= \beta' R_i + E(v_i | u_i > -\gamma' Z_i) = \beta' R_i + \rho \sigma_v \lambda_i(\alpha_u), \end{aligned} \quad (2.9)$$

em que:

$$\lambda_i(\alpha_u) = \frac{\phi\left(\frac{\gamma' Z_i}{\sigma_u}\right)}{\Phi\left(\frac{\gamma' Z_i}{\sigma_u}\right)}$$

$$\alpha_u = \frac{\gamma' Z}{\sigma_u},$$

em que ϕ é a função densidade de probabilidade normal padrão e Φ , a função de distribuição normal acumulada. A função $\lambda_i(\alpha_u)$ é denominada razão inversa de Mills que será considerada variável explicativa na regressão principal a fim de eliminar o viés de seletividade amostral. Logo, a partir da equação 1, a regressão pode ser representada por:

$$W | L > 0 = \beta' R_i + \beta_\lambda \lambda_i(\alpha_u) + \varepsilon_i. \quad (2.10)$$

A relação entre a variável L e os coeficientes associados às variáveis explicativas possui uma relação não-linear, impossibilitando a estimação através do método (MQO),

uma vez que estes estimadores não mantêm suas propriedades assintóticas (GREENE, 2003).

Como solução, Silva *et al.* (2002) indica o método “Máxima Pseudo-verossimilhança” (MPV). O referido método é geralmente empregado para ajustes de modelos paramétricos, quando se consideram o plano amostral e os pesos no processo de inferência com dados de amostras complexas. Esse método também corrige a seletividade da amostra, realizado através da estimação das equações de participação e rendimentos. Esse processo não envolve mais duas etapas, sendo a estimação das duas regressões concomitantemente. Para simplificar, segue abaixo um esquema simples do processo.

1. Estima-se uma equação de participação através do modelo Probit, que permitirá obter, por meio da determinação da razão inversa de Mills, uma nova variável. Quando $L = 1$, significa dizer que o indivíduo aceitou exercer uma atividade econômica remunerada. Por outro lado, quando $L = 0$, é porque os salários de reserva encontram-se acima dos salários oferecidos no mercado, de maneira que o indivíduo prefere manter-se fora do mercado de trabalho.

2. Calcula-se a razão inversa de Mills.

3. Logo após, introduz-se a razão inversa de Mills (λ) nas equações de rendimento como variável explicativa para corrigir possíveis problemas de viés de seleção amostral. Este procedimento permitirá obter estimadores consistentes para os parâmetros das equações de salários. Isto posto, estima-se por meio do método dos mínimos quadrados ponderados, uma equação de (logaritmo neperiano dos) rendimentos. Admite ser normalmente distribuído com média zero e desvio padrão igual à unidade.

2.3.3. Fonte de dados

Os dados da pesquisa foram obtidos no Instituto Brasileiro de Geografia e Estatística (IBGE) extraídos da Pesquisa Nacional por Amostra a Domicílio (PNAD). Na medida em que essa pesquisa não disponibiliza em seus resultados finais informações compatíveis com a divisão do setor rural proposta no presente trabalho, fez-se necessário

acessar os microdados da PNAD. Esse procedimento, além da complexidade em manipular uma ampla base de dados, necessita da utilização de pacotes computacionais, neste caso o SPSS, para a obtenção dos dados agrupados da população compatíveis com a divisão proposta pela pesquisa.

O uso da PNAD em pesquisas sociais vem crescendo nos últimos anos, na medida em que serve de fonte de dados robusta, utilizada por muitos pesquisadores que visam nortear políticas públicas como o acesso à educação básica, o combate à discriminação e segmentação no mercado de trabalho, a redução da pobreza e desigualdade de renda, dentre outras.

2.3.4. Tratamento dos dados

Tendo em vista o relevante período que a pesquisa abarca, 1995 a 2007, fez-se necessário um amplo estudo da metodologia adotada pela PNAD para garantir comparabilidade no período apresentado. Ao longo dos anos, algumas variáveis foram recodificadas, outras tiveram suas categorias estendidas, como a variável anos de estudo. A variável cor e raça, por exemplo, somente em 1984 aparece no questionário da PNAD, e só na década de 1990 incorporou-se a raça indígena. Dentre as seleções mais importantes realizadas por meio das rotinas computacionais do SPSS destacam-se: a escolha do meio rural e urbano e a exclusão de indivíduos cuja renda, anos de estudo, ramo de atividade e posição na ocupação, não foram declarados nas PNADs.

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). 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 erros-padrões, variância, quantis, percentis, análise de regressão, dentre outros. De acordo com Silva *et al.* (2002) a PNAD é caracterizada por um plano amostral complexo por ser elaborada considerando a estratificação, conglomeração, probabilidades desiguais de seleção em um ou mais estágios e calibração da amostra através de pesos amostrais. Estas características são devido a limites de custos e a restrições de tempo da pesquisa, uma vez que o território nacional abrange uma área de 8.514.876 Km², com um total de 191.790.900 habitantes (IBGE, 2007).

A estratificação da PNAD 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. 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 censo.

Para os demais 27 estratos¹⁸ 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. De acordo com 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 observações de estratos com PSU único em estrato na mesma UF com maior número de observações¹⁹.

Diante destas considerações, é preciso considerar a PNAD como amostra complexa e não como observações de uma amostragem aleatória simples, ou seja, observações independentes e identicamente distribuídas (IID). 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, o que torna imprescindível a consideração do desenho amostral para obter estimativas não viciadas das variâncias da amostra.

Testes de Kish (1965), conhecido como DEFF (*design-effect*), e o de Skinner, Holt e Smith (1989), denominado estatística MEFF (*misspecification effect*), comparam a estimativa da variância do parâmetro obtida considerando o plano amostral com outra, do

¹⁸ 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.

¹⁹ Este método foi utilizado por Guimarães (2007).

mesmo modelo, desconsiderando peso, conglomerado e estratificação²⁰, reiterando a importância de se considerar a PNAD como amostra complexa.

Para considerar tal particularidade utilizaram-se dois tipos de metodologias: linearização, e replicação. Dentre os métodos de linearização, destaca-se o método de linearização de Taylor, que aproxima os estimadores de interesse que são não-lineares por estimadores lineares para obter a variância. Concernente aos métodos de replicação tem-se o procedimento de *bootstrapping* desenvolvido por Efron (1979), que consiste em reamostrar os dados com reposição em número suficiente de vezes gerando uma estimativa empírica da distribuição amostral da estatística. Guimarães (2007) o método de *bootstrap* é frequentemente usado quando existe grande dificuldade de se calcular a estimativa do erro padrão de um estimador. Outro método é o de *Jackknife* (*Jackknife Replication*). 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. Ademais, a replicação de *Jackknife* não requer estratificação explícita.

2.3.5. Deflator da renda

Para comparação de valores monetários de diferentes períodos utilizou-se como deflator o Índice Nacional de Preço ao Consumidor (INPC) restrito (mesmo deflator adotado pelo Instituto Brasileiro de Geografia e Estatística (IBGE) nas divulgações da síntese de indicadores da PNAD). Optou-se pelo INPC por ser um índice de custo de vida com grande abrangência geográfica; ademais seu uso é comum, permitindo confrontar resultados.

2.4. Resultados e discussão

2.4.1. Estatística descritiva

Nesta seção é apresentada uma análise descritiva comparativa da escolaridade, experiência de trabalho, idade, ramo de atividade, dentre outras variáveis utilizadas para

²⁰ A análise do MEFF é similar ao DEFF.

caracterizar os grupos de trabalhadores no meio rural e no meio urbano do Brasil. É importante destacar que se incorpora o plano amostral para o desenvolvimento destas estatísticas, uma vez que se utilizam os microdados da PNAD. Para melhor visualização apresentam-se Figuras com as médias ao longo do tempo, cujas Tabelas de origem, que englobam medidas de variância, encontram-se em apêndice. Outrossim, a fim de facilitar o reconhecimento das variáveis segue-se a nomenclatura adotada nos dicionários da PNAD.

Segundo metodologia do IBGE, denomina-se área urbana quando se trata de cidades. De acordo com Diniz e Silveira (2009), a distinção entre rural e urbano no Brasil é de caráter legal-administrativo não guardando relação direta com processos e variáveis demográficas, notadamente a densidade populacional, como ocorre nos países centrais. Assim, a queda na população rural de 31,8 milhões de pessoas, em 1993, para 27,6 milhões de pessoas, em 2006, se deve a atualização das malhas municipais de seus perímetros urbanos. Ou seja, as áreas rurais se tornaram cidades, geralmente em regiões mais próximas a grandes centros urbanos. Por serem mais dinâmicas em termos ocupacionais, em geral, concentravam expressiva parte das ocupações não-agrícolas. Neste contexto, pode-se inferir que parte da redução de ocupações não-agrícolas tenha sido influenciada pelas mudanças metodológicas do IBGE. Tais alterações podem ser corroboradas pela Figura 2.1 abaixo, onde se percebe claramente que o recuo da população rural só ocorreu entre 1999 e 2001.

Em 1938, através do Decreto-Lei 311, ficou estabelecido que as cidades (sedes de municípios) e as vilas (sedes de distritos) são urbanas e que o restante do território é rural. Desde então cabe às prefeituras definir os perímetros urbanos, onde há municípios que são inteiramente urbanos, como a cidade do Rio de Janeiro que ocupa exatamente o mesmo território que o município. Ou seja, mesmo quando as pessoas não estão migrando elas deixam de ser residentes no meio rural porque o local agora é designado de meio urbano.

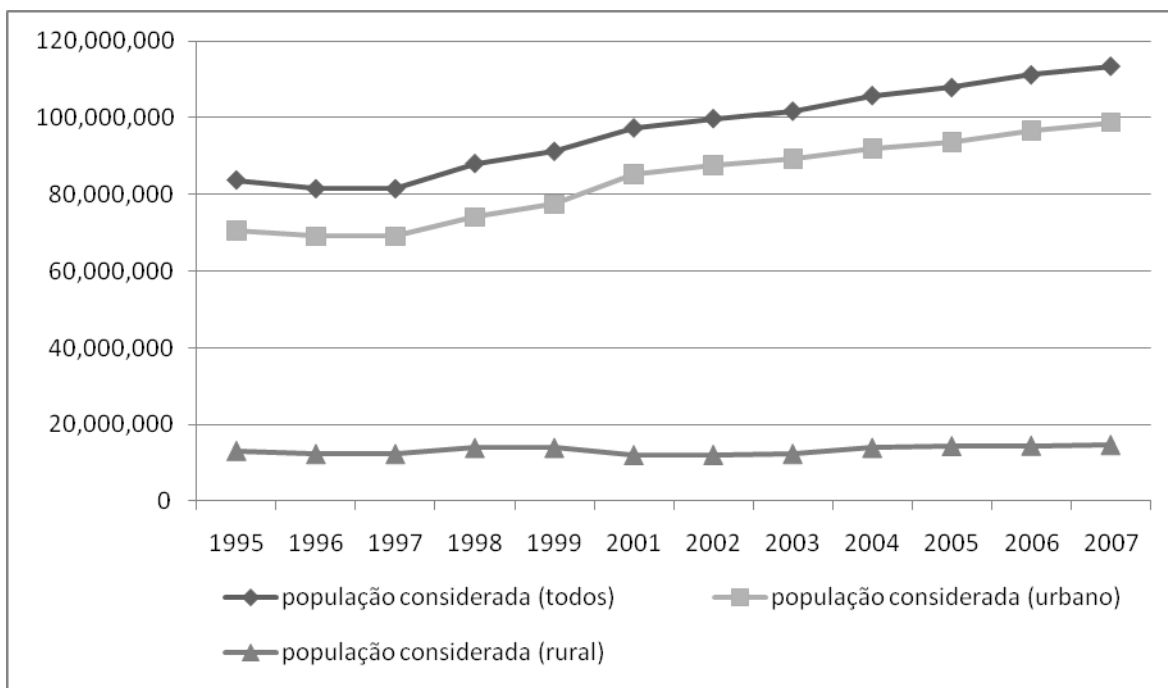


Figura 2.1: População do Meio Urbano e Meio Rural Brasileiro.

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1996 a 2007, IBGE.

Segundo Rocha *et al.* (2005), a partir de 1992 o IBGE passou a adotar nova metodologia concernente ao conceito de trabalho. Para as PNADs anteriores a esse período, considerava-se ocupada (economicamente ativa) a pessoa de 10 anos e mais que realizou alguma atividade remunerada ou não-remunerada por pelo menos 15 horas na semana de referência. A nova metodologia considera como economicamente ativa todas as pessoas de 10 anos e mais que exerceram alguma atividade por pelo menos 1 hora na semana de referência, incluindo as que se dedicam exclusivamente ao autoconsumo e, ou, a autoconstrução. Nesta pesquisa considera-se a nova metodologia.

A Figura 2.2 apresenta a evolução do número de pessoas economicamente ativa para o Brasil como um todo e também divididos entre urbano e rural. De acordo com os resultados encontrados é possível verificar que a proporção relativa de indivíduos economicamente ativos no rural supera a do urbano, onde se pode inferir que tal comportamento é devido ao fato do meio rural absorver pessoas que trabalham sem a contrapartida financeira, como por exemplo, na agricultura de subsistência. Ademais, estudos de Freeman (1987), Murphy e Welch (1992), afirmam que um indivíduo deve a cada ano decidir se vai para a escola e investe em educação ou vai para o mercado de trabalho em período integral, sendo a segunda opção mais escolhida no meio rural. Neste contexto, pode-se inferir que há um maior número de filhos trabalhando no rural enquanto

no urbano os mesmos estão estudando em busca de retorno financeiro pelos anos de estudo dedicados. Outrossim, culturalmente os filhos de famílias rurais tendem a seguir o negócio da família representado pela agricultura familiar. Segundo a Secretaria do Planejamento e Gestão – SEPLAG (2009), taxas elevadas de ocupação da População Economicamente Ativa é um indicativo de uma economia forte e sustentável.

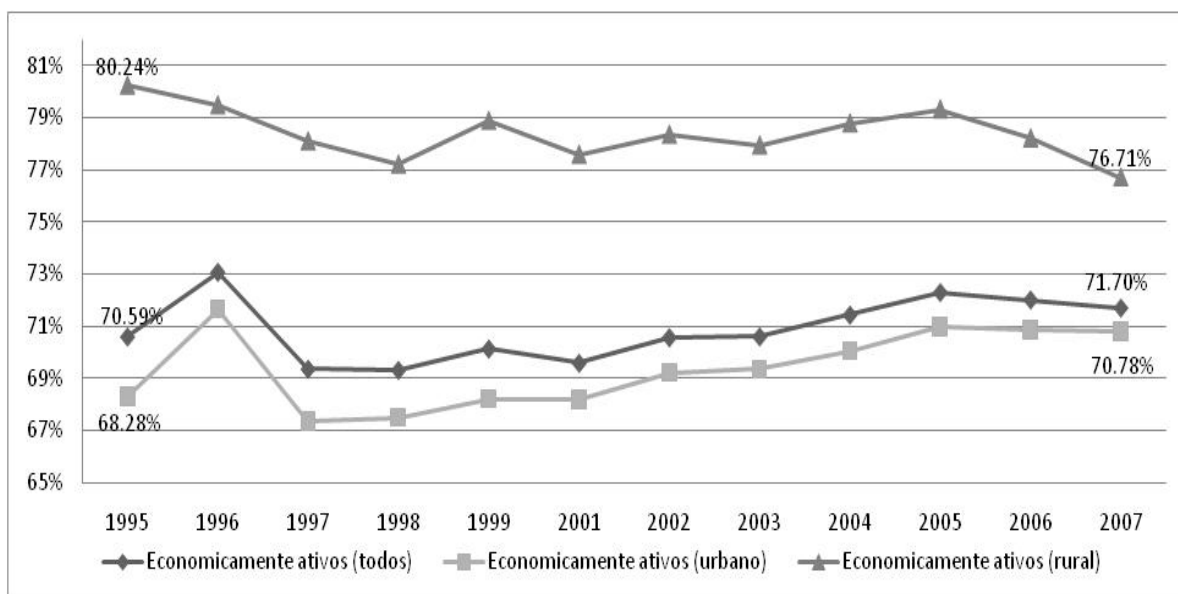


Figura 2.2: Proporção de Indivíduos Economicamente Ativos no Brasil.

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1996 a 2007, IBGE.

A população economicamente ativa quando ocupada desenvolve atividades agrícolas e não-agrícolas. Neste contexto, a Figura 2.3 representa as pessoas ocupadas em atividades não agrícolas. É interessante acrescentar que as demais pessoas estão ocupadas em atividades agrícolas, uma vez que a pesquisa não abarca aposentados ou pensionistas. No meio urbano, a porcentagem de pessoas ocupadas em atividades não-agrícolas é cada vez mais alta, chegando em 2007 a 71%. Concernente ao meio rural é relevante notar que ao longo da década de 1990 o aumento de pessoas ocupadas em atividades não-agrícolas aumentou de forma contínua, mas no período de troca de década verifica-se uma queda e em seguida permanece a tendência de crescimento. Tal fenômeno pode ser explicado, em parte, pela metodologia do IBGE que a cada 10 anos, quando se realiza os censos demográficos, são redefinidas as áreas rurais e urbanas, onde se pode observar o estreitamento das áreas rurais, na medida em que novas cidades “surtem”, em alguns casos devido a interesses, como o aumento da arrecadação de impostos, e não devido a uma mudança na estrutura da região.

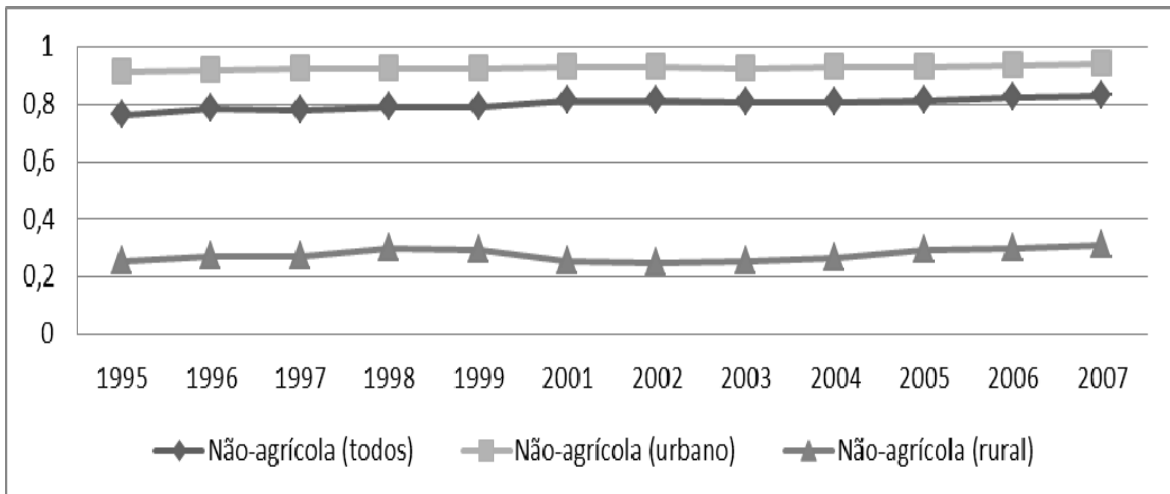


Figura 2:3. Proporção de Indivíduos Ocupados em Atividades não Agrícolas no Brasil.

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1996 a 2007, IBGE.

Educação é considerada a variável que mais impacta na renda do indivíduo, sendo mais expressiva em atividades não-agrícolas. Neste contexto, muitos estudos econométricos acerca dos retornos para educação, como os de Psacharopoulos (1981) e Ashenfelter *et al.* (1999), têm sido realizados em diferentes países. Psacharopoulos e Patrinos (2004) destacam que a América Latina/Caribe e a África Subsaariana são as regiões com maior retorno, para todos os níveis educacionais. No Brasil merece destaque os estudos de Menezes *et al.* (2004), entre outros

A Figura 2.4 demonstra a média de anos de estudo para o Brasil como um todo e separados entre urbano e rural. Verifica-se que essa variável cresce na mesma proporção entre os trabalhadores urbanos e rurais, indicando a existência de atividades no campo capazes de absorver mão-de-obra qualificada. Pela teoria do capital humano as pessoas investem em educação quando esperam retorno financeiro. Ou seja, essas novas ocupações estimulam a busca por qualificação uma vez que o retorno financeiro compensa tal investimento, demonstrado no Figura 2.13 da próxima seção. Contudo, ainda persiste o diferencial de anos de estudo entre o urbano e o rural.

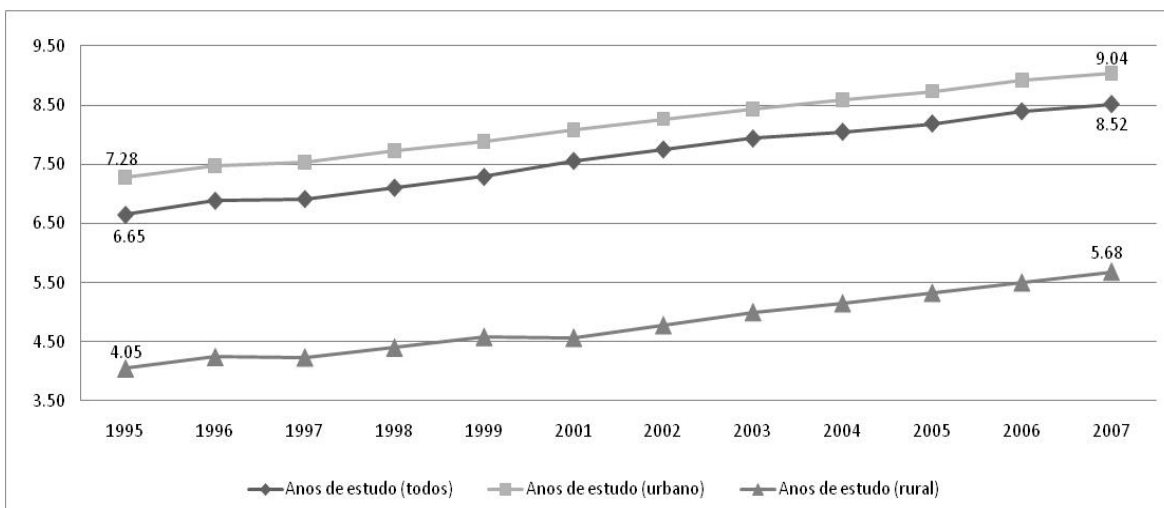


Figura 2.4: Média de Anos de Estudo dos Indivíduos Entre 15 e 70 Anos.

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1996 a 2007, IBGE.

Em relação à renda, a Figura 2.5 apresenta a evolução do rendimento médio per capita advindo do trabalho. Nesse, pode-se constatar que para o rural o rendimento do trabalho principal aumentou ao longo dos anos da mesma década, sendo mais expressivo na década de 2000. Cotejando as Figuras 2.3 e 2.5, nota-se um pouco de similaridade na linha de tendência dos dois, ou seja, quando a renda média do trabalho principal aumenta, verifica-se crescimento das ocupações não-agrícolas, sendo o contrário também verdadeiro. Para o urbano, este crescimento aconteceu somente a partir de 2001, tendo um pico expressivo em 2002 que mantém durante os demais anos considerados, cujo período é marcado pelo governo de Luiz Inácio Lula da Silva. Tal governo é marcado por expressivos reajustes no salário mínimo. De acordo com pesquisa do Departamento Intersindical de Estatística e Estudos Socioeconômicos (Dieese), o salário mínimo acumulou reajuste real de 45% entre 2002 a 2006. Outrossim, segundo o IBGE houve aumento da renda disponível das famílias em 2002 devido, em grande parte, à elevação do rendimento médio do trabalho em 6,7% e do crescimento do pessoal ocupado em 3,0%, além de outros fatores como aumento dos benefícios sociais e transferências correntes líquidas.

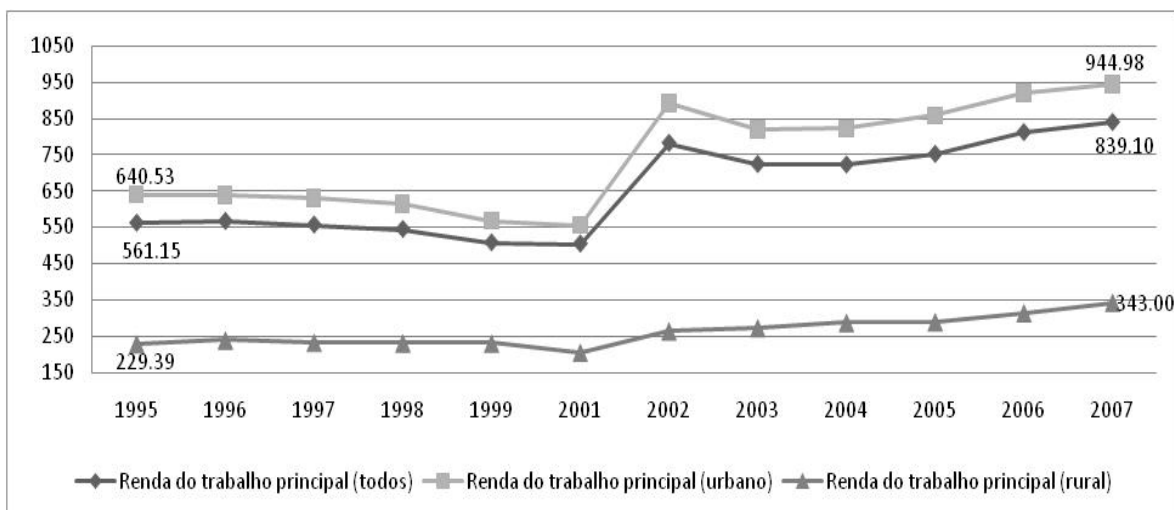


Figura 2.5: Rendimento Médio Advindo do Trabalho Principal

Fonte: Elaborado pela autora a partir dos microdados da PNAD – 1996 a 2007, IBGE.

A Figura 2.6 apresenta o rendimento familiar médio ao longo dos anos. A partir desta, verifica-se que o rendimento familiar diminuiu ao longo dos anos, recuperando no final do período. Tal comportamento pode ser justificado pela inserção cada vez mais tarde dos jovens no mercado de trabalho devido ao aumento do tempo destinado à qualificação, que pode ser comprovado pelo aumento dos anos de estudo (Figura 2.4). Por outro lado, no rural o comportamento é oposto, cujo aumento pode estar relacionado ao aumento das novas ocupações capazes de absorver esta mão de obra. Este acréscimo do rendimento familiar mostra-se diretamente relacionado com o aumento da viabilidade de permanência no campo. Ou seja, o rural está oferecendo maiores oportunidades de emprego para toda família, embora o salário seja menor do que no urbano.

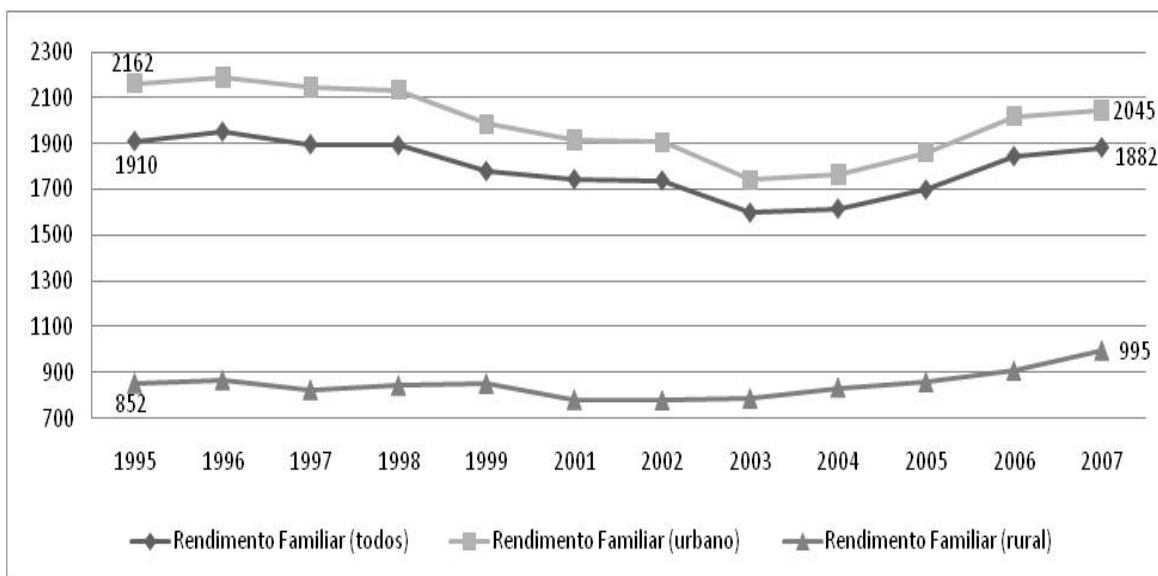


Figura 2.6: Rendimento Médio Familiar (R\$).

Fonte: Elaborado pela autora a partir dos microdados da PNAD 1995 a 2007, IBGE.

As Figuras 2.7 e 2.8 abaixo apresentam a evolução da renda por setor de atividade, agrícola e não-agrícola, sendo a primeira para o meio rural e a segunda para o meio urbano. Mesmo diante do desenvolvimento do rural em termos de dinâmica de ocupação e elevação dos anos de estudo, persiste a elevada disparidade de salários entre o rural e urbano. Em 2007, este diferencial ultrapassa 100%. Estas figuras revelam a grande segmentação de mercado existente. Mesmo analisando somente a renda advinda de atividades não-agrícolas, a renda do urbano (977,36) é quase o dobro do rural (528,22) em 2007.

Pode-se observar que, tanto no rural quanto no urbano, as pessoas ocupadas em atividades agrícolas possuem menores rendimentos. Tal fenômeno é perfeitamente explicado na literatura nos trabalhos de Kageyama (2003) e Schneider (2003), uma vez que tais atividades absorvem mão-de-obra pouco qualificada.

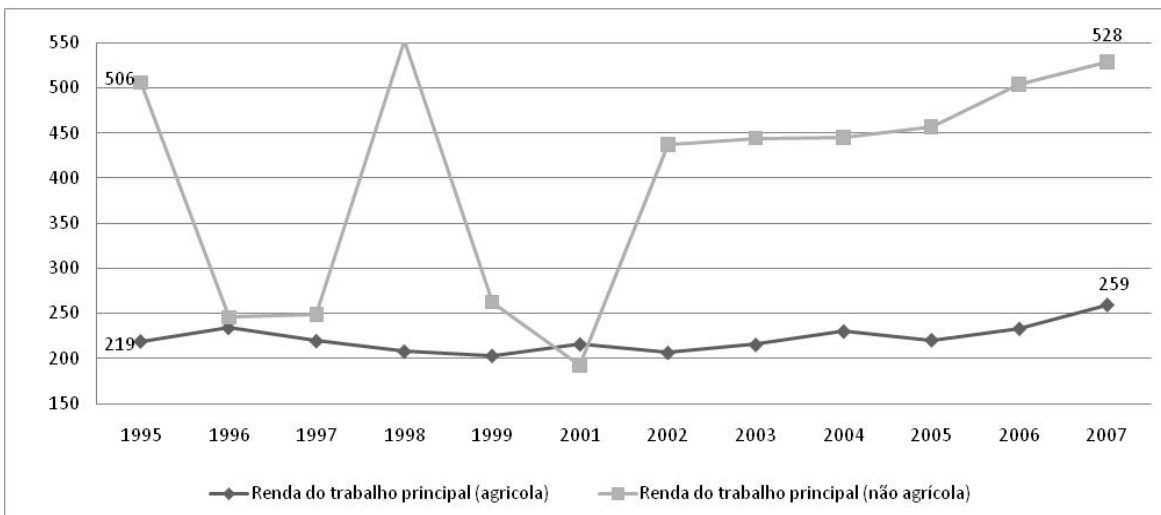


Figura 2.7: Rendimento do Trabalho Principal no meio Rural.

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1995 a 2007, IBGE.

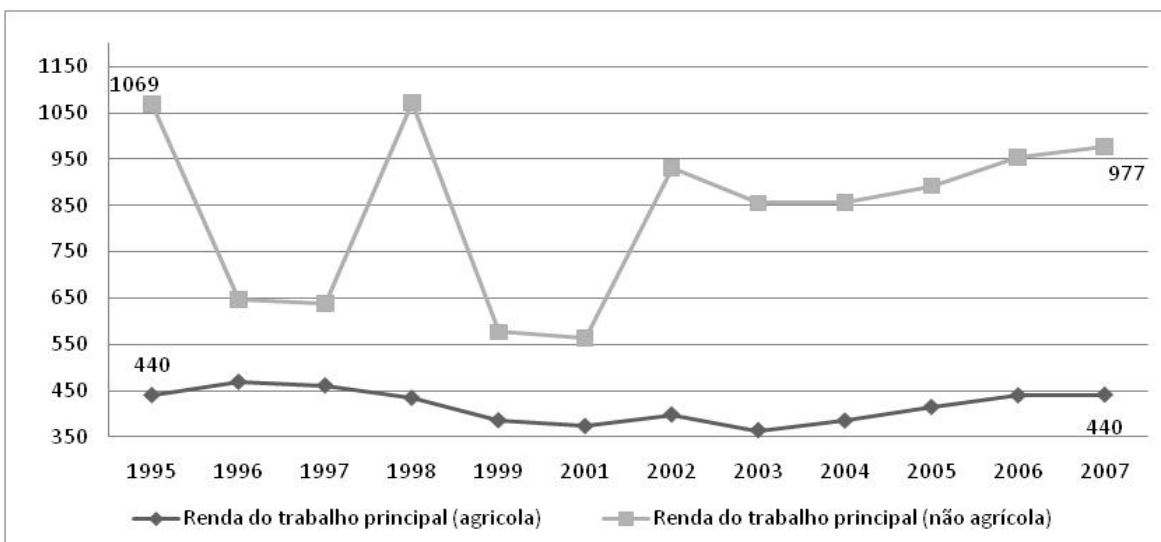


Figura 2.8: Rendimento do Trabalho Principal no meio Urbano.

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1995 a 2007, IBGE.

Fazendo uma digressão de renda, analisam-se horas trabalhadas e longevidade. A Figura 2.9, apresenta a evolução da média de horas trabalhadas²¹.

²¹ Estão sendo consideradas somente as pessoas que declararam à PNAD todas as perguntas concernentes às variáveis que a presente pesquisa abarca, ou seja, a média da idade sofre impacto das pessoas que não responderam a todas as perguntas, bem como as horas trabalhadas. Tal característica é válida para todos os Figuras apresentados.

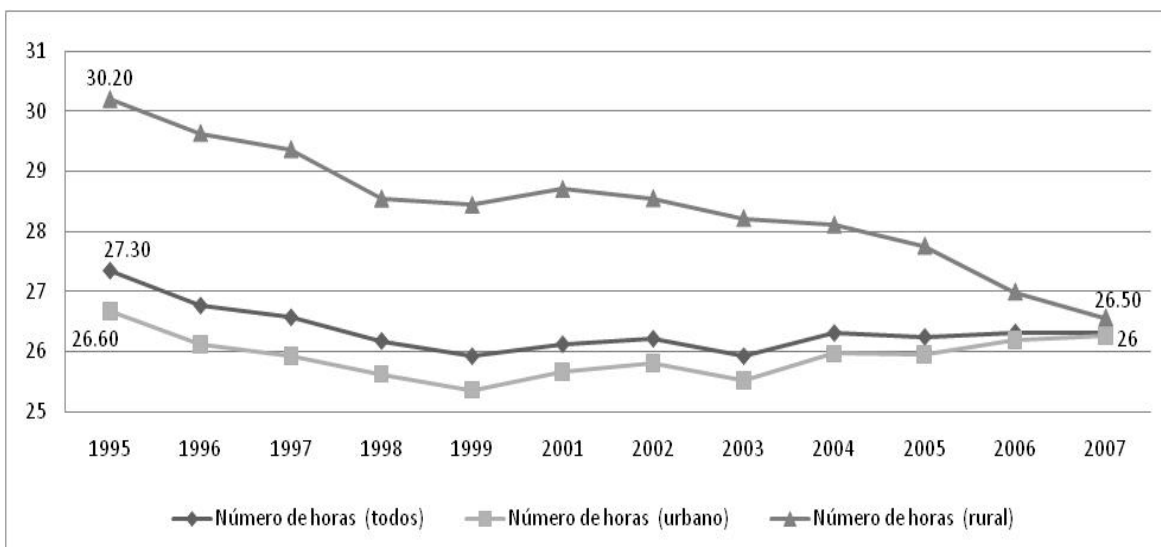


Figura 2.9: Média de Horas no Trabalho Principal.

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1996 a 2007, IBGE.

Embora a agropecuária brasileira tenha passado por processo de modernização nas relações de trabalho, ainda predominam no campo inserções trabalhistas como pequeno produtor, meeiro, trabalho sem remuneração, entre outras bastante distintas do emprego assalariado presente nos centros urbanos. O meio rural sempre se mostrou menos rígido quanto às leis do trabalho, ou seja, no rural a jornada de trabalho é relativamente maior que no urbano e os indivíduos também se dedicam à agricultura de subsistência. Mas como verificado em figura abaixo, há o encontro destas médias em 2007.

A Figura 2.10 demonstra a evolução da média de idade dos espaços que a pesquisa abarca. Quanto à longevidade, o rural e o urbano caminham juntos, dado que a média da idade é similar para o urbano e para o rural. Mesmo longe do estresse e da violência inerentes à vida na cidade, a média de anos no rural é a mesma desta.

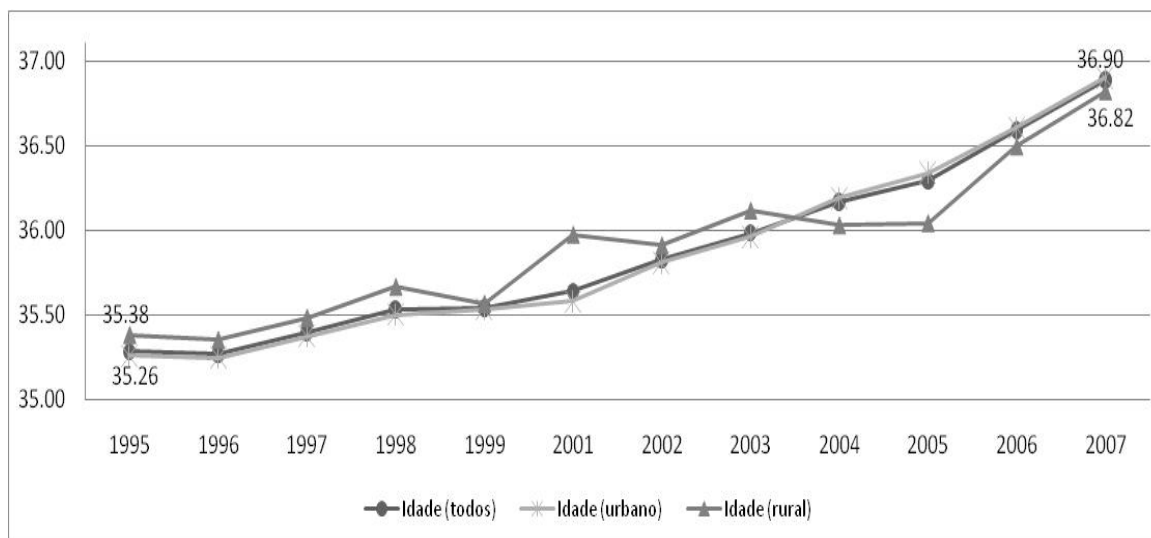


Figura 2.10: Idade em anos.

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1996 a 2007, IBGE.

Os resultados descritivos acima demonstraram que no final do período estudado as horas trabalhadas no rural e urbano se igualaram e a variável anos de estudo sofreu acréscimos tanto no urbano quanto no rural. Contudo, o diferencial deste último indicador continua bastante expressivo. Estudos indicam que essas variáveis são determinantes na geração da renda, o que explica o grande diferencial da renda do trabalho entre esses dois espaços geográficos. Ou seja, pode-se inferir que o problema não se concentra mais na legislação do trabalhador rural no que concerne às horas trabalhadas, nem na segmentação de mercado, típico entre regiões menos desenvolvidas e mais desenvolvida.

Em suma, pode-se inferir que o diferencial existente entre a renda do trabalho urbano e rural é devido ao grau de instrução.

2.4.2. Limiar da Educação

De acordo com Hoffmann e Simão (2005), quando a taxa de retorno da educação aumenta a partir de certo nível de escolaridade, deve-se fazer uso de uma função de rendimentos que possa captar esse aumento brusco. Do contrário, o uso de uma equação de rendimentos em que logaritmo do rendimento é uma função linear da escolaridade, sem o efeito limiar, constitui erro de especificação, que pode levar a conclusões errôneas ao comparar a taxa de retorno da escolaridade em diferentes situações (entre regiões ou entre setores de atividade).

Neste contexto, foram desenvolvidas equações de rendimento com variáveis binárias para cada ano de estudo, com a finalidade de encontrar o limiar da educação para cada setor censitário (urbano e rural). Depois de encontrado este limiar, exclui-se todas as variáveis binárias de educação, pois apenas com a inclusão deste parâmetro (limiar), consegue-se a redução de cerca de 95% da falta de ajustamento. (Hoffmann e Kassouf, 2005).

A Figura 2.11, referente a 1995, que representam este procedimento estão dispostos abaixo e os demais em apêndice. O limiar é escolhido quando se tem um aumento brusco de retorno para cada ano adicional de estudo seguido de um aumento crescente. Pode-se verificar que o limiar para o urbano (a) é de 11 anos e para o rural (b) 7 anos. Os coeficientes das duas variáveis associadas à escolaridade permitem calcular as taxas de retorno por ano de estudo quando ela é inferior ao limiar da educação.

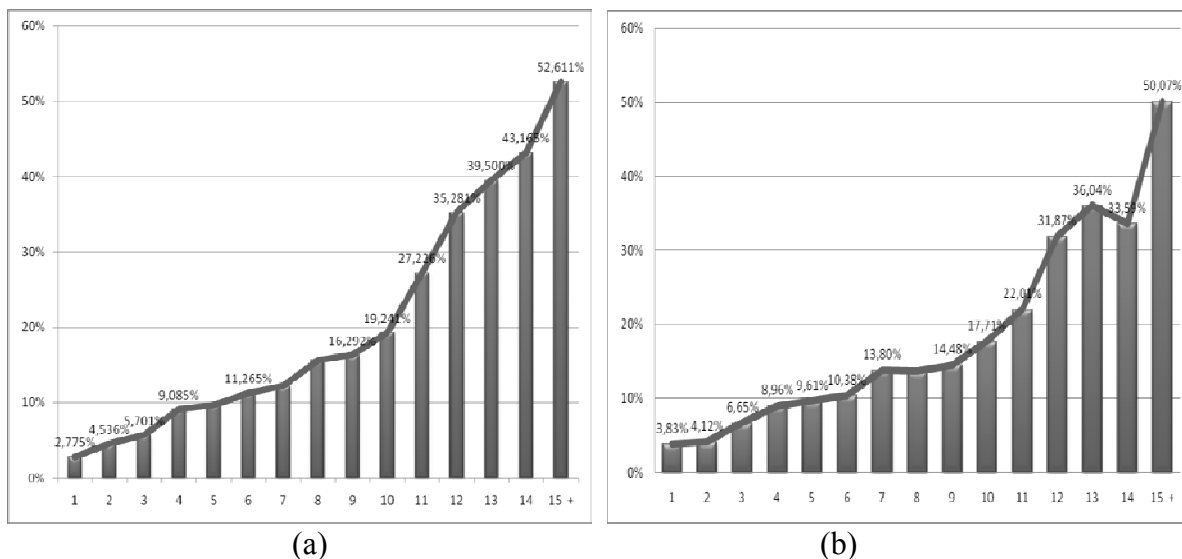


Figura 2.11: Escolha do Limiar da Educação

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1996 a 2007, IBGE.

Como esperado, os limiares da educação são diferentes para o urbano e para o rural, comprovando a hipótese de que para cada setor censitário existe um determinado limiar. No urbano, o efeito diploma é mais fácil de ser observado, na medida em que o limiar é na maioria das vezes 11 anos, ou seja, quando o indivíduo completa o segundo grau. Já no rural, o limiar gira em torno de 7 anos, cujo período é quase suficiente para obter o diploma do ensino básico. Cabe salientar que este limiar muda ao longo dos anos, principalmente no rural, que em 2007 passa a ser 12 anos (Figura 2.12). Detalhes da evolução deste limiar ao longo dos anos que a pesquisa abarca podem ser acompanhados pela Figura 2.12 abaixo.

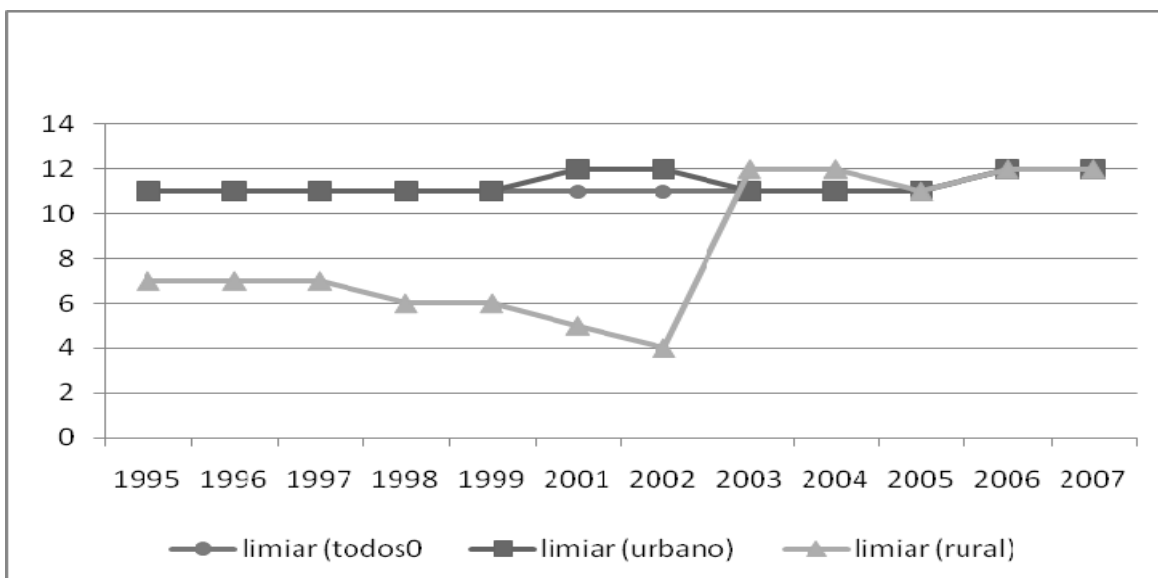


Figura 2.12: Evolução do Limiar da Educação.

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1996 a 2007, IBGE.

Esta evolução demonstra que embora o urbano e o rural apresentem níveis de escolaridade diferentes, o limiar tornou-se o mesmo a partir de 2003. Isso indica que no rural, comparativamente, estão sendo destinados retornos bem maiores a pessoas com grau de instrução maior. Antes, o pico era a partir dos 7 anos, agora são 12, ou seja, o grande diferencial passa a ser a partir do diploma do ensino básico e não fundamental. Deste modo, o investimento em educação no meio rural torna-se mais compensatório.

2.4.3. Decomposição do Diferencial de Salário com Correção do Viés de Seleção do Salário na Análise de Regressão

Esta seção tem por finalidade testar a existência de segmentação dos trabalhadores segundo local de residência, rural e urbano, através da inclusão da variável *dummy* para urbano. Em algumas pesquisas, o foco da análise é a avaliação das características dos indivíduos na determinação do salário, enquanto outras têm como objetivo mensurar diferenciais entre grupos de trabalhadores de acordo com suas características sociodemográficas. Nesta pesquisa, priorizou-se o segundo objetivo.

Para tanto, apresenta-se o procedimento em dois estágios, onde se estima um modelo Probit para a participação no mercado de trabalho. Trata-se do modelo apresentado na Tabela 2.3. Considera-se uma amostra com trabalhadores economicamente ativos e não economicamente ativo, a partir da qual é calculada a variável lambda (ou razão inversa de Mills) e utilizada como variável explicativa para estimação da equação de salários (modelo

da Tabela 2.2). É importante ressaltar que o coeficiente da variável lambda é estatisticamente significativo, indicando que a inclusão é necessária para correção do viés de seleção da informação do salário. Pode-se verificar pelo MEFF, cujos resultados foram na maior parte afastado de 1, que a incorporação do plano amostral é preciso. Do contrário, as medidas estariam subestimadas, impedindo o resultado de medidas concretas que propiciam um maior conhecimento do mercado de trabalho brasileiro.

Nas equações de participação no mercado de trabalho, Tabelas 2.3 e 2.5, são encontradas algumas características que impactam sobre a probabilidade do indivíduo encontrar-se economicamente ativo. Esta análise é realizada a partir da comparação do modelo rodado para 1995 e para 2007. As variáveis apresentaram expressiva significância estatística, principalmente ao nível de 1%, bem como a estatística de significância global (F) para os dois modelos. Os efeitos marginais das variáveis foram convertidos, uma vez que em modelos probit não se pode interpretar diretamente os coeficientes. Para o modelo com todos os indivíduos, o fato do indivíduo ser do sexo masculino elevava a probabilidade deste encontrar-se no mercado de trabalho (ou economicamente ativo) em 33,5 p.p. em 1995, permanecendo no mesmo patamar em 2007, 31,5 p.p.

Passando para a análise dos componentes da família, embora esta variável seja estatisticamente significativa seu impacto na probabilidade do indivíduo estar economicamente ativo é mínima em 1995. Já em 2007, cada membro adicional na família passa a reduzir a probabilidade do indivíduo encontrar-se economicamente ativo, porém ainda não é expressiva. As demais variáveis também impactam de maneira pouco expressiva, cujos resultados são acompanhadas nas Tabelas 2.3 e 2.5.

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.

Tabela 2.2. Determinantes do rendimento em 1995*.

Variáveis explicativas	TODOS		
	COEF.	M.E.	MEFF
Constante	0.223 (0.000)	- -	2.56 -
Gênero (Masculino)	0.253 (0.000)	28.85% (0.000)	3.85 -
Idade (em dezenas de anos)	0.441 (0.000)		2.44 -
idade^2 (em dezenas de anos)	-0.052 (0.000)		2.75 -
Cor Amarela ou pardas	-0.019 (0.000)	-1.91% (0.000)	1.35 -
Cor Branca	0.032 (0.000)	3.26% (0.000)	1.36 -
Educação em anos de estudo	0.027 (0.000)	2.78% (0.000)	1.52 -
Efeito Limiar da Educação	0.051 (0.000)	5.27% (0.000)	1.83 -
Conta-própria	0.011 (0.000)	1.12% (0.000)	2.71 -
Empregador	0.127 (0.000)	13.58% (0.000)	2.06 -
Não-agrícola	0.141 (0.000)	15.14% (0.000)	3.13 -
Urbano	0.013 (0.017)	1.27% (0.017)	2.52 -
Sudeste	0.105 (0.000)	11.10% (0.000)	4.20 -
Sul	0.101 (0.000)	10.68% (0.000)	2.45 -
Centro	0.079 (0.000)	8.17% (0.000)	1.63 -
Norte	0.062 (0.000)	6.39% (0.000)	1.05 -
Lambda	1.673 (0.000)	- -	12.54 -
Num. Obs		184,672	
Pop. Considerada		83,746,508	
Estatística F (Prob)		0.00	

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1995, IBGE.

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

* Todos se refere a soma de Urbano e Rural.

Tabela 2.3. Equação de seleção 1995*

Variáveis explicativas	TODOS		
	COEF.	M.E.	MEFF
Constante	-2.381 (0.000)	- -	1.64 -
Gênero (Masculino)	0.939 (0.000)	0.335 (0.000)	6.19 -
Número de Componentes na fam.	0.009 (0.000)	0.003 (0.000)	1.60 -
Rendimento Familiar	-0.00004 (0.000)	-0.00002 (0.000)	14.86 -
Idade (em dezenas de anos)	1.298 (0.000)		2.46 -
idade^2 (em dezenas de anos)	-0.167 (0.000)		2.48 -
Cor Amarela ou pardas	-0.010 (0.569)	-0.004 (0.569)	1.55 -
Cor Branca	-0.053 (0.002)	-0.019 (0.002)	1.53 -
Educação em anos de estudo	0.027 (0.000)	0.010 (0.000)	1.73 -
Efeito Limiar da Educação	0.121 (0.000)	0.045 (0.000)	1.19 -
Urbano	-0.195 (0.000)	-0.070 (0.000)	2.54 -
Sudeste	0.005 (0.734)	0.002 (0.734)	2.77 -
Sul	0.116 (0.000)	0.042 (0.000)	2.35 -
Centro	0.049 (0.019)	0.018 (0.019)	2.49 -
Norte	0.020 (0.246)	0.007 (0.246)	1.19 -
Num. Obs		184,672	
Pop. Considerada		83,746,508	
Estatística F (Prob)		0.00	

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1995, IBGE.

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

* Todos se refere a soma de Urbano e Rural.

Tabela 2.4. Determinantes do rendimento em 2007*.

Variáveis explicativas	TODOS		
	COEF.	M.E.	MEFF
Constante	0.790 (0.000)	120.36% (0.000)	1.47 -
Gênero (Masculino)	0.042 (0.000)	4.15% (0.000)	0.98 -
Idade (em dezenas de anos)	0.090 (0.000)		1.01 -
idade^2 (em dezenas de anos)	-0.007 (0.000)		1.03 -
Cor Amarela ou pardas	-0.006 (0.025)	-0.62% (0.025)	2.25 -
Cor Branca	0.029 (0.000)	2.93% (0.000)	2.00 -
Educação em anos de estudo	0.014 (0.000)	1.43% (0.000)	2.55 -
Efeito Limiar da Educação	0.028 (0.000)	2.79% (0.000)	2.09 -
Conta-própria	-0.007 (0.001)	-0.73% (0.001)	2.96 -
Empregador	0.102 (0.000)	10.16% (0.000)	1.69 -
Não-agrícola	0.079 (0.000)	7.94% (0.000)	4.57 -
Urbano	0.026 (0.000)	2.63% (0.000)	2.11 -
Sudeste	0.082 (0.000)	8.52% (0.000)	2.69 -
Sul	0.077 (0.000)	7.96% (0.000)	1.98 -
Centro	0.097 (0.000)	10.14% (0.000)	1.56 -
Norte	0.070 (0.000)	7.28% (0.000)	2.74 -
Lambda	-0.049 (0.000)	- -	0.63 -
Num. Obs		238,447	
Pop. Considerada		113,489,497	
Estatística F (Prob)		0.00	

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1995, IBGE.

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

* Todos se refere a soma de Urbano e Rural.

Tabela 2.5. Equação de seleção 2007*

Variáveis explicativas	TODOS		
	COEF.	M.E.	MEFF
Constante	-3.334 (0.000)	- -	1.65 -
Gênero (Masculino)	0.955 (0.000)	0.315 (0.000)	1.81 -
Número de Componentes na fam.	-0.051 (0.000)	-0.017 (0.000)	1.69 -
Rendimento Familiar	0.000013 (0.000)	0.00000 (0.000)	4.35 -
Idade (em dezenas de anos)	1.907 (0.000)		1.62 -
idade^2 (em dezenas de anos)	-0.243 (0.000)		1.63 -
Cor Amarela ou pardas	-0.058 (0.000)	-0.020 (0.000)	1.47 -
Cor Branca	-0.109 (0.000)	-0.037 (0.000)	1.61 -
Educação em anos de estudo	0.051 (0.000)	0.017 (0.000)	1.66 -
Efeito Limiar da Educação	0.036 (0.000)	0.012 (0.000)	1.62 -
Urbano	-0.190 (0.000)	-0.062 (0.000)	2.34 -
Sudeste	0.081 (0.000)	0.027 (0.000)	2.56 -
Sul	0.204 (0.000)	0.066 (0.000)	2.79 -
Centro	0.123 (0.000)	0.040 (0.000)	1.60 -
Norte	0.033 (0.098)	0.011 (0.098)	2.71 -
Num. Obs		238,447	
Pop. Considerada		113,489,497	
Estatística F (Prob)		0.00	

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1995, IBGE.

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

* Todos se refere a soma de Urbano e Rural.

Em seguida, estima-se a equação de rendimento pelo método pseudo maximaverossimilhança apresentados nas Tabelas 2.2 e 2.4. De acordo com os modelos estimados acima, verifica-se que ainda predomina diferença salarial entre o urbano e o

rural. Em 1995 essa diferença era 1,27% passando em 2007 para 2,63%, levando a questionar o porquê da elevação do diferencial entre urbano e rural, diante do desenvolvimento do último. Mesmo que os resultados referentes à média da renda do urbano seja quase o dobro, o fato de residir no rural não impacta da mesma forma na renda. Ou seja, não se verifica expressiva segmentação de mercado. Como demonstrado nas seções anteriores, o número de ocupações não agrícolas aumentou, bem como a média de anos de estudo. Era de se esperar, assim, o arrefecimento deste diferencial, o que não foi comprovado.

Quanto às análises das macroregiões do país, em 1995 a região Sudeste prevalece como o local em que o trabalhador pode obter uma remuneração mais elevada comparativamente à região nordeste em mais de 11%. Em 2007 a região Centro-Oeste toma o posto do Sudeste com o diferencial de rendimentos para o nordeste figurando em 10,14%. Para ambos os anos pode-se verificar que o trabalhador situado na região Nordeste, que é o grupo de controle, tende a apresentar os menores rendimentos.

2.4.3.1. Evolução da Decomposição do Diferencial de Salário

Esta seção tem por finalidade apresentar as equações de rendimento para o rural e o urbano separadamente, bem como uma regressão para todo conjunto, totalizando três modelos. Para melhor detectar a evolução das diferenças dos determinantes ao longo dos anos optou-se por apresentar figuras, cujos gráficos permitem melhor visualização.

Devido ao número de observações nas amostras utilizadas, quase todos os coeficientes são estatisticamente diferentes de zero ao nível de significância de 1%. É importante lembrar que o rendimento das pessoas é condicionado por características pessoais de mensuração muito difícil (como ambição, tino comercial, etc), englobando também componente aleatório. Porém, considerando o elevado número de observações, tal dificuldade é amenizada.

A estimação de modelos separados para rural e urbano decorre do consenso de que as diferenças não podem ser acomodadas pela simples introdução de variáveis *dummies*. Outrossim, a presente pesquisa tem como objetivo apresentar a evolução do efeito marginal das variáveis estudadas, não se prendendo a justificativa do porque do determinante de cada variável, ou seja, não se procura explicar o comportamento do

mercado de trabalho e sim fazer o cotejamento entre a evolução, principalmente, das variáveis educação e posição na ocupação no meio rural e meio urbano

A Figura 2.13 demonstra a evolução do efeito marginal da variável anos de estudo referentes às três equações de rendimento, total (urbano + rural), rural e urbano. Efeito Marginal dos Anos de educação indicam o acréscimo médio na renda para cada ano adicional de estudo. De 2003 a 2007, o rural e o urbano possuíam praticamente o mesmo limiar, e, considerando a Figura abaixo, possuíam o mesmo retorno. Moura (2008) demonstra que retornos educacionais estão decaindo nos últimos anos, com exceção do nível superior que aponta para um crescimento nesta última década, mas em magnitude menor dos obtidos em diversos estudos recentes. Nesta pesquisa, corroborou-se tal estudo para o rural e urbano.

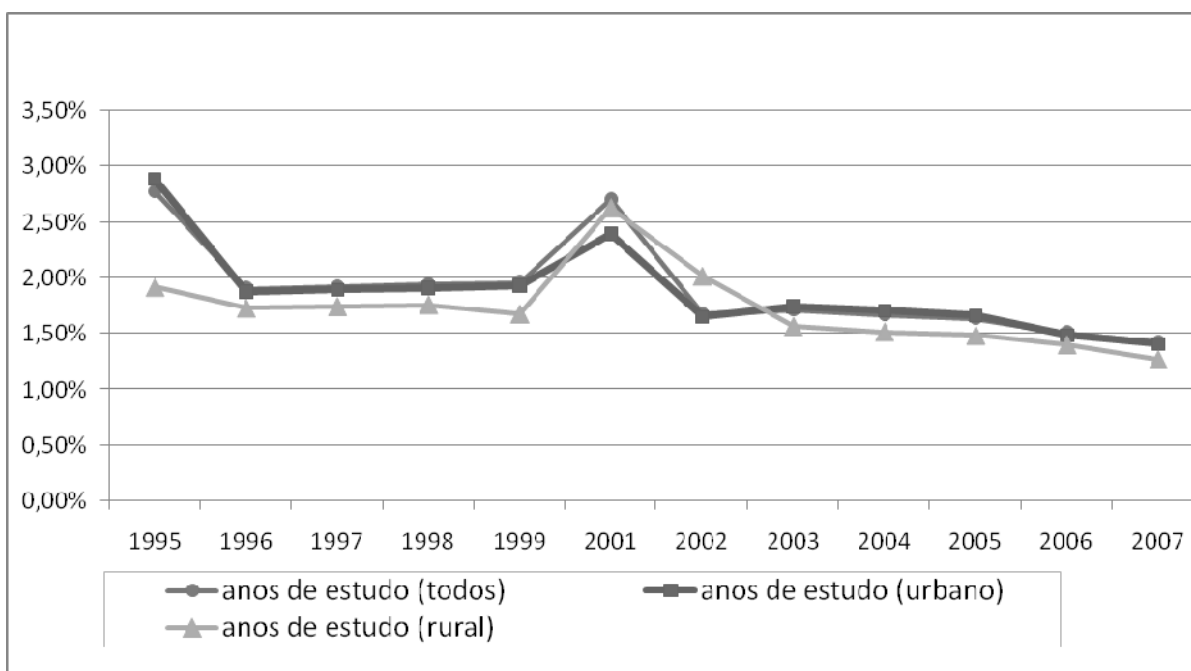


Figura 2.13: Evolução do Retorno dos Anos de Estudo.

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1996 a 2007, IBGE.

A Figura 2.14, representa a evolução do efeito marginal da variável empregador. No meio rural ao contrário do urbano, a posição na ocupação representava um grande diferencial na determinação do rendimento. Observa-se na figura abaixo que, em 2002, este diferencial diminuiu e tornou-se semelhante ao do urbano. Este nivelamento da posição na ocupação é mais um fator de similaridade que o meio rural vem apresentado em relação ao urbano.

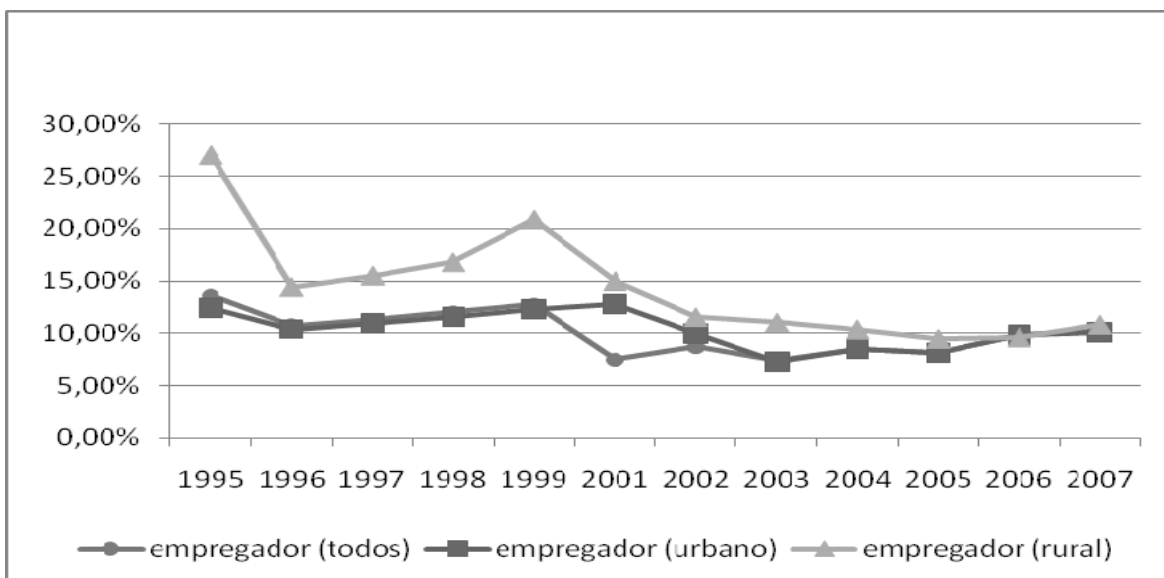


Figura 2.14: Evolução do Retorno da Posição na Ocupação (empregador).

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1996 a 2007, IBGE.

Quanto à discriminação, em termos de gênero e cor, apresentadas na Figura 2.15 e 2.16, estas apresentavam-se mais expressivas no rural. Contudo, em 2007, ambas as variáveis convergiram para um mesmo percentual.

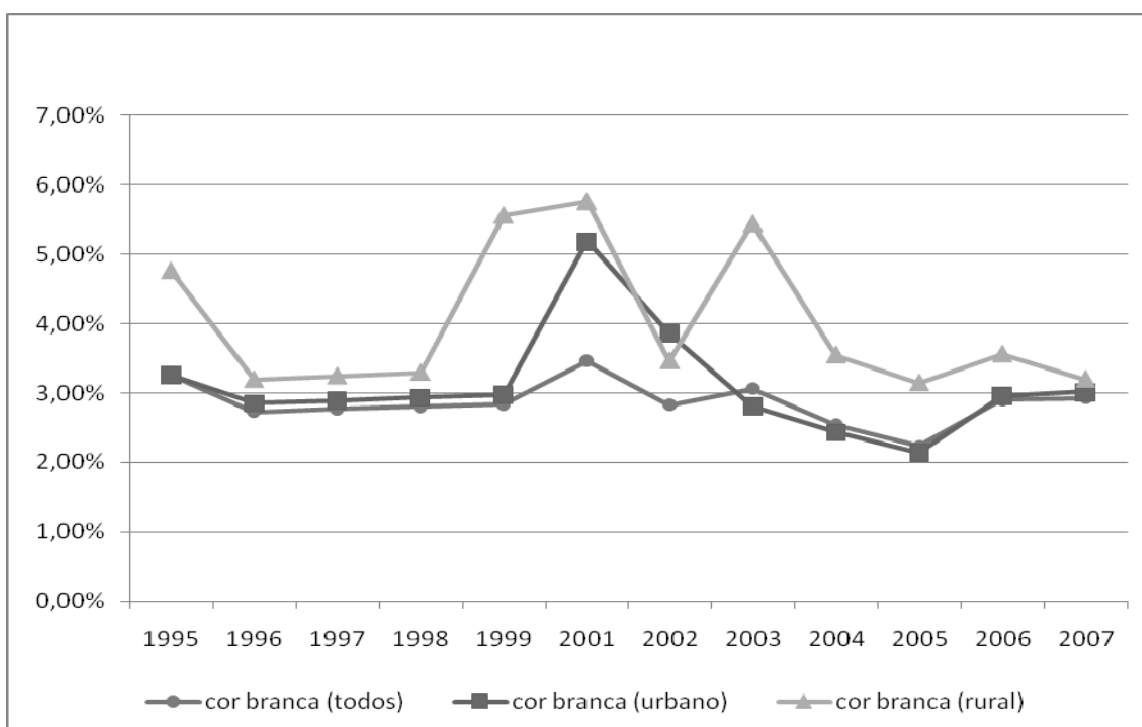


Figura 2.15: Evolução do Retorno de Cor ou Raça (Branca).

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1996 a 2007, IBGE.

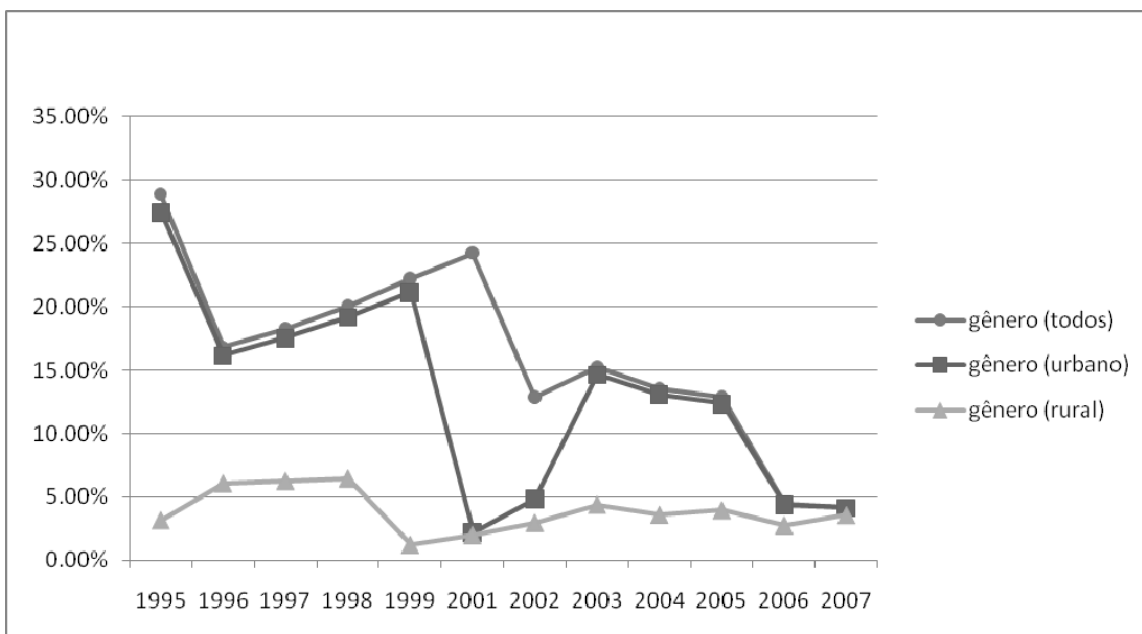


Figura 2.16: Evolução do Retorno por Gênero (Homem).

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1996 a 2007, IBGE.

2.5. Conclusão

A presente pesquisa mostrou a evolução das características pertinentes ao rural e ao urbano, no período 1995-2007, propiciando subsídios para a construção de políticas públicas mais eficazes, capazes de atender às especificidades de cada setor censitário. Ademais, primou pela busca de estatísticas mais consistentes, optando por incorporar o plano amostral, a determinação do limiar de educação e a correção do viés de seletividade pelo método de Heckman.

Diante dos resultados foi possível concluir que o rural apresenta alguns pontos positivos. Na análise descritiva verificou-se que no meio rural a proporção de pessoas economicamente ativas é maior que no urbano. Concomitantemente, observou-se que o rendimento familiar rural cresceu e diminuiu no meio urbano. É interessante destacar que o presente enaio computa apenas o rendimento advindo do trabalho e que não abrange as transferências do governo. Neste contexto, pode-se inferir que a viabilidade de permanência no campo aumentou, e que se deu com base na renda do trabalho.

Quanto aos pontos negativos, o rural apresenta características quanto à renda ainda desfavorável frente ao urbano. Entretanto, em relação a outros aspectos observou-se uma convergência, em 2007, das linhas de tendência das figuras que representam a evolução das variáveis para o urbano e rural. O resultado esperado na presente dissertação era de que as mudanças ocorridas no rural o aproximassem do urbano. Ou seja, esses dois espaços,

antes bem distintos, começariam a ficar mais iguais quanto às características de ocupação. Tal fenômeno foi em grande parte confirmado, pois aumentou a similaridade da quantidade de horas trabalhadas, do limiar da educação e da queda do impacto da posição na ocupação no retorno da renda do trabalhador residente no rural. Tendo em vista a elevação dos anos de estudo e do limiar da educação e da redução da discriminação, no meio rural, era esperado um arrefecimento do diferencial de sua renda em relação ao urbano, o que não ocorreu, no período 1995 a 2007.

Neste sentido, ainda se faz necessário uma maior atenção ao rural brasileiro no que se refere à renda do trabalho para que os trabalhadores sejam remunerados de acordo com a atividade que desempenha, amenizando a segmentação de mercado. Em outras palavras, o que se verificou é que o rural de uma política de arrefecimento do diferencial de renda, ou seja, de políticas voltadas para aumentar a renda do trabalhador rural que ocupa atividades tanto agrícolas quanto não-agrícolas, uma vez que o desenvolvimento rural é um indicador de desenvolvimento de um país. Não obstante, deve-se destacar que houve melhoria na legislação trabalhista do trabalhador rural e progresso das similaridades dos determinantes da renda do rural e urbano.

A agricultura perdeu sua importância quanto à geração direta de emprego, contudo ela ainda é propulsora de emprego. Muitas das novas ocupações dependem indiretamente da renda dos agricultores para o pagamento dos serviços prestados, demonstrando que para se atingir o pleno desenvolvimento rural deve-se incentivar tanto as atividades agrícolas quanto as não agrícolas e principalmente a educação.

O objetivo principal foi a comparação entre rural e urbano. Ao longo da pesquisa vários questionamentos foram propositalmente levantados e esclarecidos, outros involuntariamente apareceram e apenas em parte foram explicados. Sendo assim, o presente ensaio também se constitui em fonte instigante para próximas pesquisas.

3. EVOLUÇÃO DA EFICIÊNCIA DO TRABALHO NO MEIO URBANO E RURAL DO BRASIL, 1981 a 2007

3.1. Introdução

3.1.1 Problema e justificativa do trabalho

A partir de meados do século XVIII, assiste-se a uma nova etapa do sistema capitalista, o qual evoluiu para um estágio mais eficiente de acumulação de capitais. A partir da Revolução Industrial, observaram-se modificações nas configurações espaciais em várias regiões do globo, consumando-se a separação entre as áreas urbanas e rurais.

No Brasil, essa separação espacial foi mais tardia. Nas primeiras décadas do século XX a sociedade brasileira se configurava como amplamente rural e a agricultura permanecia na liderança no que se refere à composição da renda nacional e geração de emprego. Destarte, a partir de 1940 houve um expressivo crescimento da população como um todo, com a aceleração do processo de industrialização, observando-se rápido crescimento das cidades, em grande parte em função do êxodo rural. Na década de 1960, a população urbana superou a rural. A velocidade do processo de urbanização se mostrou superior ao dos países capitalistas mais avançados. De acordo com Brito (2006), em apenas 50 anos, na segunda metade do século XX, a população urbana passou de 19 milhões para 138 milhões, multiplicando-se 7,3 vezes, com uma taxa média anual de crescimento de 4,1%. Tal fenômeno significou, a cada ano, em média, que 2,4 milhões de habitantes eram acrescidos à população urbana.

Do ponto de vista econômico as cidades são relevantes pelos benefícios que trazem como a maior oferta e concentração de bens e serviços públicos e privados que se traduzem em elementos de eficiência. Ou seja, a cidade ao dispor de uma logística que oferece competitividade econômica, tornou-se atrativa para os moradores rurais.

Entretanto, no Brasil, a partir de 1980, verifica-se um arrefecimento do fluxo migratório para as cidades devido ao surgimento de novas ocupações no meio rural. De acordo com Graziano (1999), entre as causas das transformações em curso no meio rural brasileiro, destaca-se a modernização agrícola, que, ao permutar trabalho por capital, promove o processo de mercantilização do tempo livre, onde as famílias rurais procuram alternativas para complementar sua renda. Na medida em que as atividades agrícolas não mais demandam tempo integral, as atividades de uso doméstico, como a fabricação de doces caseiros, artesanato, dentre outras, são transformadas em fonte de renda, alterando o perfil ocupacional destas famílias, de forma a possibilitar sua permanência no campo.

Para Rocha *et al.* (2005), além da modernização, a agricultura brasileira passou, na década de 1990, por transformações estruturais importantes, tais como: redução de incentivos do governo, abertura comercial, sobrevalorização do câmbio, queda do preço dos produtos agrícola, que impactaram de forma negativa a agricultura, diminuindo sua importância quanto à geração de empregos. De acordo com Graziano e Del Grossi (2001), em 1998 o total das rendas não-agrícolas do Brasil ultrapassou o montante das rendas agrícolas recebida pelos moradores rurais. Isso demonstra que no final do século XX as atividades agropecuárias já não respondem pela maior parte da renda da população rural brasileira²², perdendo o lugar de líder que lhe era concedido no princípio deste século.

As atividades não-agrícolas têm remunerações maiores que as atividades agrícolas. Esse diferencial de salário faz com que os trabalhadores mudem de ocupação, mesmo com algum tipo de custo de deslocamento. Dado que no meio rural também habitam atividades não-agrícolas, esse deslocamento não precisa ser realizado entre campo e cidade. Ou seja, não é preciso que haja, necessariamente, êxodo rural.

²² A realidade sócio-espacial torna-se cada vez mais complexa. Kageyama (2003) atenta para a necessidade de discernir rural e agrícola, sendo que o primeiro termo refere-se a um setor da atividade da economia, enquanto a noção de rural é territorial ou espacial, e em seu interior podem desenvolver-se atividades de praticamente todos os setores econômicos. A novidade está na sua inserção cada vez menor, com as práticas e as funções urbanas invadindo cada vez mais o meio rural e famílias agrícolas indo residir nas periferias urbanas.

Êxodo rural, ou Migração rural, é um fenômeno de transferência da população de regiões onde prevalecem atividades não-agrícolas, com uma produtividade do trabalho relativamente mais baixa, para os setores urbanos e industriais da economia capitalista cuja produtividade do trabalho é mais elevada.

O indivíduo faz uma análise custo-benefício. Emigrar em direção às grandes regiões urbanas foi, por muito tempo, a opção que concedia maiores benefícios quando o objetivo era a melhoria do padrão de vida. O migrante é considerado como um indivíduo dotado de racionalidade econômica na decisão de migrar e, portanto, capaz de desenhar os seus caminhos pelo território de uma maneira adequada às necessidades do mercado de trabalho. Considera-se o migrante como empresário de si mesmo que procura a localização ótima para o seu capital humano.

Segundo Harris e Todaro (1970), uma das principais razões que levam os indivíduos a migrar dizem respeito ao âmbito econômico, relacionado à expectativa de um nível mais elevado de rendimento na região de destino. No caso da migração rural-urbana os autores demonstram que os migrantes decidem também com base no diferencial entre a renda na cidade e a renda no campo. De acordo com essa visão, o processo migratório ocorre até o momento em que o diferencial deixa de existir, ou seja, a migração rural-urbana persiste enquanto a renda urbana esperada for superior à renda rural. Os autores concluem que, migração é, portanto, um fenômeno de desequilíbrio diante do modelo de escolhas dos trabalhadores.

Partindo da perspectiva de Harris e Todaro (1970), e considerando que o rural passa por transformações representadas principalmente pelo surgimento de novas ocupações as quais estão viabilizando a permanência da população no campo, é de se esperar que a eficiência do trabalho no meio rural esteja aumentando, estabelecendo, portanto um equilíbrio, ou a diminuição do desequilíbrio, na medida em que não é mais atrativo migrar do rural para o urbano.

Neste contexto, objetiva-se avaliar a performance da renda do trabalho rural em relação ao urbano, dos estados brasileiros, destacando alguns possíveis elementos explicativos das diferenças encontradas entre esses últimos.

A definição de eficiência, segundo a teoria econômica refere-se à otimização de recursos, pela utilização máxima dos recursos existentes para satisfazer as necessidades e os desejos de indivíduos e organizações (Pindyck e Rubinfeld 1994).

De acordo com a Teoria do Capital Humano de Becker (1964), como mencionado acima, a remuneração é equivalente ao custo e tempo dedicado na obtenção de qualificação. No entanto, segundo Barros *et al.* (2007), tem-se observado que dentro de um mesmo estado ou município persistem disparidades espaciais de remuneração entre trabalhadores com idênticas características produtivas, ou seja, qualificação semelhante, dando origem à segmentação de mercado. As mais notórias são as existentes entre as áreas urbanas e as rurais, onde trabalhadores urbanos recebem uma remuneração acima da dos trabalhadores rurais com idênticas características observáveis em postos de trabalho similares.

Não obstante, as transformações que aconteceram no meio rural, como consolidação de ocupações não-agrícolas, imprimem uma maior dinamização na economia local, colocando este espaço mais próximo ao meio urbano. Neste sentido, espera-se uma tendência de redução da disparidade da eficiência do rendimento do trabalho devido a um provável arrefecimento da segmentação de mercado (rural/urbano).

Outrossim, existe também a segmentação setorial que diz respeito a trabalhadores com idênticas características produtivas que, por exercerem atividades em setores diferentes, agrícolas e não-agrícolas, são remunerados de forma distinta pelo mercado. Assim, dado que as atividades não-agrícolas, geralmente, exigem maior qualificação, o aumento de pessoas ocupadas em atividades não-agrícolas no campo, em princípio, deveria ter ocasionado a elevação da eficiência no rural.

Para a verificação, assim, da alteração na eficiência do trabalho em ambos os espaços, o rendimento do trabalho foi considerado como produto e as horas trabalhadas e os anos de estudo como insumos.

Para estimar a eficiência do rendimento do trabalho foi calculada a eficiência técnica baseada na estimação da fronteira empírica, por meio de programação linear. O foco de análise continuou sendo a comparação entre o meio rural e o urbano. Porém, para abranger as diferenças regionais e totalizar o número de unidades necessárias para o cálculo do método DEA, foram considerados os estados brasileiros, comparando-os em termos de eficiência dos seus meios rurais e urbanos.

Com o objetivo de corroborar os resultados encontrados na análise de eficiência foi calculado o Índice de Malmquist. Este índice permite analisar as mudanças na fronteira e em relação à fronteira entre os anos que a pesquisa abarca. Deste modo, permite verificar

se houve melhora na eficiência do rendimento do trabalho nos estados, relativa às áreas rural e urbana, ao longo dos anos.

Como supramencionado, para mensurar se a eficiência no meio rural está aumentando devido à difusão da tecnologia trazida pelas ocupações não agrícolas, ou seja, se em média uma determinada quantidade de horas trabalhadas e de anos de estudo no rural proporcionam a mesma renda que se tem no urbano, utiliza-se o método *DEA*. Trata-se de uma técnica não paramétrica pela qual é possível medir a eficiência relativa do trabalho, comparando a eficiência entre o trabalho rural e o urbano. O uso deste método permite verificar se o trabalho no meio rural tem se tornado eficiente o bastante para viabilizar a permanência no campo.

Segundo Menezes-Filho *et al.* (2004), há uma grande concordância na literatura acerca da existência de diferenças de renda entre os estados brasileiros, principalmente a partir de 1930. A explicação para esse fenômeno compreende uma variedade de fatores: geográficos, históricos, políticos e socioeconômicos. Em termos econômicos, as diferenças de desempenho entre os estados são atribuídas aos diferentes níveis de poupança, de investimentos e de adoção de novas tecnologias. Para se ter uma idéia desta heterogeneidade, em 2000, de acordo com os dados do IBGE, o PIB *per capita* de São Paulo era pouco mais de cinco vezes superior ao do Piauí. Essa constatação, justifica a separações entre estados, além de rural e urbano, uma vez que os estados brasileiros são expressivamente heterogêneos, sendo sua análise conjunta poderiam mascarar realidades distintas do rural e urbano do Brasil.

O método *DEA* tem sido muito abordado em diversas pesquisas devido principalmente ao fato de que as medidas obtidas são comparativas, se baseando na identificação da melhor prática entre as unidades. Porém não foram encontradas pesquisas que tratam especificamente de eficiência do trabalho entre meio rural e urbano. Como estudo adjacente, se pode citar o de Marinho *et al.* (2001), que a partir da elaboração da fronteira estocástica mostraram que a produtividade total dos fatores no Brasil aumentou, principalmente, devido à variação tecnológica, no período entre 1986 e 1998. Além disto, mostraram também que a trajetória de produtividade é influenciada pela abertura econômica e que após 1992 a produtividade dos estados e regiões passou a crescer mais rapidamente. Considerando o transbordamento da tecnologia no país, isto deveria implicar um aumento conjunto do nível de bem-estar de todos os estados.

Marinho e Benegas (2002), analisaram a eficiência técnica das regiões e estados brasileiros no período de 1985 a 1998. Para isso, utilizaram o modelo DEA com base na meta-fronteira e o índice de produtividade total de Malmquist para analisar ganhos de produtividade. Como produto tem-se os PIB's estaduais e insumos, como *proxy* do capital, utilizou-se o consumo não-residencial de energia elétrica; para a variável trabalho, utilizou-se horas trabalhadas anuais extraídas da PNAD. Os resultados mostraram que as regiões Norte, Sudeste e Sul são as mais eficientes no uso da tecnologia de que dispõe. Em relação ao teste de liderança tecnológica a região Sul é quem determina o padrão tecnológico do país e a região Norte que, apesar de ser eficiente no uso da tecnologia a sua disposição, apresenta a maior defasagem tecnológica com relação ao padrão do país que é determinado pela região Sul. Porém, ambas as regiões estão convergindo para o mesmo patamar tecnológico.

Como destacado acima, não foram encontrados estudos específicos do emprego do DEA semelhante à análise que o presente ensaio abarca. Não obstante, segundo Gomes (2008), é expressiva a disseminação desta nova ferramenta para mensuração da eficiência. Somente na área agrícola, foram encontradas 158 referências em periódicos internacionais e nacionais.

Diante desta nova configuração espacial das atividades produtivas estudou-se a evolução das diferenças entre o meio rural e meio urbano em termos de rendimento do trabalho assumindo as transformações nestes respectivos locais. A análise foi feita por estado a fim de obter a fronteira que interliga os pontos referentes às combinações de insumos (horas de trabalho e anos de estudo) capazes de fornecer o maior produto (rendimento do trabalho) segundo o critério da produtividade da análise envoltória dos dados e que possa ser utilizado para o norteamento de políticas públicas. Para tanto, utiliza-se os microdados da Pesquisa Nacional por Amostra de Domicílio (PNAD) do Instituto Brasileiro de Geografia e Estatística (IBGE).

3.1.2. Hipótese e objetivos

3.1.2.1. Hipótese

Se houve difusão da tecnologia trazida pelas ocupações não-agrícolas no meio rural, então é esperado um aumento maior na eficiência do trabalho do rural relativamente à do urbano.

Espera-se também um aumento nos índices de desigualdade de renda no meio rural, devido à disseminação das novas ocupações que necessitam de maior nível educacional quando comparado com as ocupações agrárias e queda da desigualdade para o urbano em face de uma maior homogeneização educacional;

Deslocamento da fronteira de eficiência devido à difusão tecnológica alavancada pelas novas ocupações no rural e a associação entre anos de estudo e ganho de eficiência, uma vez que essas novas ocupações necessitam de mão de obra mais qualificada do que as ocupações tradicionalmente agrícolas;

Aproximação da fronteira eficiente no meio rural em relação ao meio urbano.

3.1.2.2. Objetivos

Analisar a evolução da eficiência dos estados diante do aumento de ocupações não agrícolas no rural, contrapondo os resultados entre a área urbana e rural. Mais especificamente:

- Determinar o comportamento da distribuição da renda do trabalho para as pessoas ocupadas e residentes no meio rural através dos índices de Gini e de Theil.
- Determinar a fronteira de eficiência dos estados brasileiros, utilizando como insumo as horas trabalhadas e a escolaridade e, como produto, o rendimento do trabalho.
- Elaborar o *ranking* de eficiência e mensurar a ineficiência identificando as possíveis causas.

- Verificar, através do Índice de Malmquist, a desagregação de alteração da produtividade em mudança técnica ou tecnológica e mudança no indicador de eficiência.
- Identificação das causas de possíveis alterações na relação produto/insumo

Contribuição do trabalho: Uso de uma metodologia até então não explorada para analisar eficiência dos estados no que concerne ao trabalho. E, principalmente, contrapor os resultados para o meio rural e meio urbano traçando o esboço atual do cenário brasileiro.

3.2. Referencial Teórico

3.2.1. Eficiência

Eficiência é um conceito multidimensional. Pode ser considerada eficiente, por exemplo, uma máquina que desempenha de forma plena o que está descrito em seu manual. Não obstante, o presente ensaio se fundamenta no conceito mais econômico, especificamente na eficiência técnica.

Eficiência Econômica total engloba a eficiência técnica e alocativa. Diz-se que um produtor é tecnicamente eficiente, quando ele só consegue aumentar a produção de um produto diminuindo a produção de outro, ou seja, quando existem desperdícios de insumos. Em outras palavras, eficiência técnica refere-se à produção da maior quantidade de produto possível por quantidade de insumo.

Já a alocativa considera o preço dos insumos em questão, cujas proporções são combinadas de modo a gerar um mínimo custo ou máxima receita (maximização de lucros). Por exemplo, a produção da maior quantidade possível de produtos em um pedaço de terra. O empresário pode decidir dentro de duas combinações tecnicamente eficientes (5 máquinas e 25 trabalhadores) ou (3 máquinas e 35 trabalhadores), aquela cujo preço do insumo permite um lucro maior. Quando o fator capital for mais barato que o fator mão-de-obra, escolhe-se a primeira opção, do contrário, a segunda alternativa.

Não obstante, em muitas atividades, relacionadas à educação e saúde, a mensuração da eficiência alocativa não é possível. Nestes casos, a medida de eficiência técnica é suficiente.

Para analisar a eficiência tem-se dois métodos no que tange ao tipo de fronteira. A abordagem paramétrica²³ envolve a construção da fronteira estocástica a partir de métodos econométricos, e a não-paramétrica relaciona-se com a fronteira de produção empírica baseada em programação linear. A abordagem paramétrica tem como trabalho seminal o estudo de Aigner e Chu (1968). Sua maior limitação concerne na escolha da função de produção, onde se faz necessário a especificação idêntica para todas as organizações. Ao contrário desta, a abordagem não-paramétrica é realizada por meio da comparação de um conjunto de unidades similares, denominadas *decision making units* (DMU), onde se constrói uma fronteira de produção empírica eficiente. As DMU que se encontram sobre a fronteira, portanto, eficientes, têm medida plena de eficiência, 100%, enquanto as DMU localizadas abaixo da fronteira são ineficientes. Esse método é denominado de análise envoltória dos dados - DEA.

Para o entendimento desta técnica é preciso compreender o conceito de fronteira de produção. A fronteira de produção é traçada a partir do Conjunto de Possibilidades de Produção –CPP. O CPP refere-se às combinações possíveis de insumos que uma DMU utiliza para gerar determinado produto diante da tecnologia disponível. Qualquer ponto do CPP indica uma combinação, tecnologicamente viável, de se produzir determinado produto dado os insumos.

$$T(X, Y) = \{(X, Y) \mid \text{é viável transformar } X \text{ em } Y\}$$

Os pontos que indicam a máxima produção, dada uma quantidade fixa de insumos, formam a Fronteira de produção. Em outras palavras, para uma tecnologia dada, caracteriza o mínimo conjunto de insumos necessários para produzir quantidades fixas de produtos (orientação insumo). Ou a máxima produção possível dada uma quantidade fixa de insumos (orientação a produto). Essas duas orientações são possíveis nas análises de DEA. No presente ensaio, utilizar-se-á a orientação a produto, uma vez que o intuito é obter o máximo de rendimento do trabalho, e não reduzir, por exemplo, os anos de estudos.

Assim, todo produtor localizado na fronteira é considerado como tecnicamente eficiente, enquanto os produtores no interior do CPP são considerados ineficientes. Graficamente, a eficiência relativa de uma DMU pode ser avaliada pela distância dela em relação a fronteira. De acordo com Debreu (1951), trata-se da métrica radial que assume um número no intervalo de [0 a 1].

²³ A metodologia da abordagem paramétrica não será desenvolvida, dado que o presente ensaio não utiliza essa abordagem. Maiores detalhes desta técnica paramétrica pode ser encontrada em Aigner e Chu (1968).

O Gráfico 1 representa a fronteira de produção. O campo hachurado neste gráfico, indica todas as combinações possíveis mediante a técnica disponível. Os pontos A e D, embora viáveis, não são combinações eficientes. Já os pontos B e C são eficientes, porque conseguem extrair maior produto dado determinado insumo. Sendo assim, esses pontos formam, junto com outras combinações eficientes, a fronteira de produção. Comparando os pontos A e C, verifica-se que ambos produzem a mesma quantidade de produto. Todavia, a combinação representada pelo ponto C produz este montante com menos quantidade de insumos.

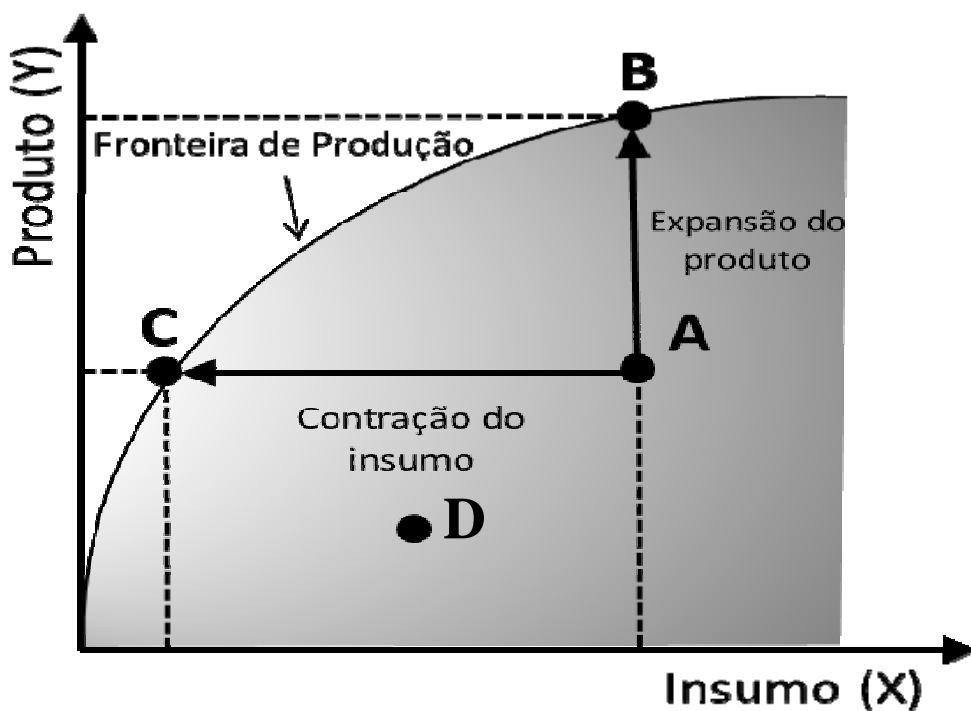


Figura 3.1. Fronteira de Produção.

Fonte: Elaborada pela autora

No enfoque com orientação ao produto, a medida de eficiência expressa a máxima expansão radial da produção para uma quantidade fixa de insumos.

$$E f i c i ê n c i a = M a x \{ \theta \mid (\theta X, Y) \in T(X, Y) \}$$

O valor θ poderá ser maior ou igual a 1. Se $\theta = 1$, a DMU em questão é eficiente. Valores maiores que um indicam que é possível produzir mais mantendo a mesma quantidade de insumos.

3.2.2. Produtividade

O aumento da produtividade, como fonte fundamental do desenvolvimento da sociedade capitalista, é um dos tópicos que sempre mereceram a atenção dos estudiosos. Smith (1999), em seu livro “A Riqueza das Nações”, destacou que o aumento da produtividade é de fundamental importância para o crescimento da riqueza da sociedade.

As economias modernas, apesar de terem passado por profundas transformações, a partir da Revolução Industrial, continuam tendo a produtividade como elemento importante para explicar as diferenças entre o crescimento das regiões. Não obstante, observou-se o aumento da complexidade das análises de produtividade e, por conseguinte, maior complexidade também em sua mensuração.

Uma das formas de se mensurar a produtividade é por meio da relação produto por insumo. Para Grosskopf (1993), a produtividade, em sua abordagem tradicional, pressupõe que a produção observada em todo período seja tecnicamente eficiente. Pelo novo conceito de medida da produtividade, trata-se da mudança na distância do produto observado em relação à fronteira de eficiência (mudança líquida) devido à mudança na eficiência e deslocamento da fronteira de produção (mudança técnica). Segundo Coura (2004) essa nova abordagem para a mensuração da produtividade incorpora a ineficiência e computa mudanças na eficiência.

A produtividade passa a ser analisada em termos de eficiência produtiva. Como supramencionado, utiliza-se a análise de eficiência para entender as diferenças de crescimento entre regiões. Neste contexto, é também apropriado mensurar a eficiência entre a área urbana e rural.

Segundo modelo de Lewis²⁴, o êxodo do campo para as cidades será realizado até o momento em que os salários urbanos não mais excederem o produto médio do campo acrescido dos custos da transferência do campo para a cidade. As divergências de uma taxa salarial decorrem da composição ocupacional da mão-de-obra, em outras palavras, da natureza da ocupação que exige certo nível de especialização. Neste âmbito, o método Análise Envoltória dos Dados – DEA – permitirá avançar no cálculo da eficiência do rural e urbano em termos de eficiência do trabalho.

²⁴ Economista britânico nascido nas Índias Ocidentais. Prêmio Nobel de Economia de 1979.

3.2.3. Teoria Neoclássica da Produção

A teoria neoclássica é baseada no individualismo metodológico e no subjetivismo. Em outras palavras, os fenômenos econômicos são consequência das escolhas que os indivíduos fazem. Eles fazem escolhas com base numa estrutura de meios e fins, sendo os meios limitados e os fins ilimitados. Essa limitação dos meios gera o problema de escolha.

A teoria neoclássica da produção explica a determinação dos preços dos fatores através das propriedades da escolha individual racional. Pressupõe-se que os indivíduos são racionais, logo qualquer meio disponível sempre será alocado para seu fim mais urgente, onde os meios são bens e os fins as preferências.

Na decisão de migrar, o indivíduo faz uma análise custo-benefício. Emigrar em direção às grandes regiões urbanas era a opção que concedia maiores benefícios aos trabalhadores na procura do aumento do poder aquisitivo. Mesmo incorporando o custo de deslocamento, o trabalhador enxergava a possibilidade de obter uma renda maior no meio urbano trabalhando a mesma quantidade de horas que exercia no meio rural. Para Vainer (2005), o migrante é considerado um indivíduo dotado de racionalidade econômica e deste modo escolhe o seu caminho pelo território de uma maneira adequada às necessidades do mercado de trabalho. O referido estudo considera cada migrante um auto-empresário que busca a localização ótima para o seu capital humano.

3.2.3.1 Princípios Básicos da Teoria Neoclássica da Produção

Segue abaixo alguns dos principais conceitos da Teoria Neoclássica da Produção visando o melhor entendimento da Análise de Eficiência apresentada neste ensaio. Uma relação entre os insumos utilizados na obtenção de determinado produto pode ser descrita por uma função de produção. Em outras palavras, estima-se a produção máxima resultante de uma empresa perante uma combinação específica de insumos.

Função de produção:

$$Q = F(K,L) \tag{3.1}$$

Onde: Q = quantidade de produção

L= mão-de-obra

K = capital.

Segundo Watson e Holman (1985) a função de produção descreve a realização de diversas combinações das proporções utilizadas dos insumos para a obtenção de um determinado volume produzido. Pode-se produzir tanto utilizando mais capital e menos mão-de-obra, ou seja, utilização intensiva de capital, quanto com mais mão-de-obra e menos capital, utilização intensiva de mão-de-obra. Como terceira alternativa tem-se combinações intermediárias. É interessante salientar que a equação (1) está associada a uma certa tecnologia, ou seja, um nível de conhecimentos disponíveis que estão relacionados diretamente com o processo produtivo. Tem-se então que o nível de produção é influenciado pelo estado no qual se encontra a tecnologia.

Segundo Pindyck e Rubinfeld (1994), um estado tecnológico refere-se a um determinado estágio do conhecimento sobre os métodos que possibilitam a transformação dos insumos em produto. Com o aperfeiçoamento da técnica, espera-se que se obtenham mais produtos com a mesma quantidade de insumos. Logo, a função de produção se altera de acordo com a evolução da tecnologia. Deste modo a função de produção passa a ser deslocada em relação ao período de tempo. Isto representa o aumento da produtividade no tempo.

A convexidade também é uma propriedade frequentemente associada à tecnologia. Isto significa que se dado dois processos ou técnicas capazes de produzir uma dada quantidade de um bem x , então uma média ponderada (combinação convexa) destes processos, também pertence ao conjunto da tecnologia.

3.2.3.2 Teoria de produção com tecnologia linear²⁵

A programação Linear foi desenvolvida durante a Segunda Guerra Mundial, e era usada como um instrumento de administração das estratégias militares. Hoje, é um complemento para a teoria da produção e pode ser empregado na solução de uma grande quantidade de problemas empresariais que dizem respeito à otimização de recursos escassos. Segundo Ehrlich (1991) a programação linear procura a melhor distribuição possível dos recursos, entre as diversas tarefas ou atividades, de forma a atingir um valor ótimo do objetivo estabelecido.

²⁵ Para maiores detalhes sobre programação linear vide Ehrlich (1991).

Considerada um avanço para o desenvolvimento da Teoria de Produção, ela usa matemática para resolver determinados tipos de problemas. Para Watson e Holman (1979), a palavra programação significa planejamento sistemático ou tomada de decisão; a linearidade refere-se aos relacionamentos tratados, sendo estes representados por retas

Matematicamente, o problema abarca uma função objetivo que apresenta as variáveis a serem maximizadas ou minimizadas, uma vez que o problema de minimização dos custos é o dual da maximização dos lucros. De acordo com Bradley, Hax e Magnanti (1977) a palavra dual significa que um problema de maximização de lucro pode ser convertido matematicamente no dual de minimização de custos. Podem também existir casos onde se encontram muitas soluções ótimas em um problema de programação linear.

Ademais, torna-se necessária a elaboração de um modelo matemático que apresente adequadamente o problema a ser resolvido. Para tanto, a correta definição do problema é possivelmente a etapa mais importante para sua resolução. Outrossim, para que essa função objetivo possa ser matematicamente especificada, as alternativas possíveis para a ocorrência de tal otimização – as chamadas variáveis de decisão envolvidas – deverão ser definidas. Normalmente assume-se que todas essas variáveis possam assumir somente valores positivos cujas variáveis podem estar sujeitas a uma série de restrições representadas por inequações. Essas limitações são relacionadas ao mínimo de capital ou mão-de-obra que uma empresa precisa. O conjunto das limitações define o espaço viável na escolha de alternativas²⁶.

3.2.4 Migração

A teoria do capital humano classifica os movimentos migratórios como um tipo de investimento em capital humano. De acordo com Becker (1993) e Mincer (1994), a decisão de migrar incorre em custos e riscos, tendo por objetivo garantir a elevação dos rendimentos nos períodos subsequentes.

A migração rural começou a ser discutida na literatura econômica nos modelos de economia dual desenvolvidos por Lewis (1954). Este modelo pressupõe que o desenvolvimento econômico causa um diferencial entre a produtividade do trabalho nos meios rural e urbano, atraindo os trabalhadores para as cidades. Supõe-se que o setor rural detém excedentes não produtivos de mão-de-obra, liberando trabalhador ocioso para as

²⁶ Maiores detalhes serão apresentados na metodologia.

cidades. Segundo o referido modelo, quando o excedente capitalista é reinvestido, mais pessoas são empregadas no setor capitalista (urbano) até que o total de força de trabalho do meio rural seja transferido para o setor moderno da economia capitalista.

Não obstante, este modelo perde sustentação quando o processo migratório observado não se arrefece diante das altas taxas de desemprego na cidade. Mais tarde, surge uma explicação econômica para este processo de migração rural-urbana com o modelo de Todaro (1969) e Harris & Todaro (1970). Segundo este modelo, a principal motivação à decisão de migrar do campo para as cidades é o diferencial entre a renda urbana esperada e a renda rural. Mais especificamente, no modelo de Todaro o migrante potencial compara a sua renda rural corrente com a renda que espera auferir na cidade.

A hipótese básica do referido modelo relaciona a probabilidade de o migrante encontrar emprego urbano com a decisão de emigrar. O emigrante age racionalmente baseando-se no cálculo econômico da renda esperada, ou permanente e não na renda imediata após a emigração. Dada a inexistência de pleno emprego, a renda esperada será ponderada pela probabilidade de se obter emprego. Neste sentido, racionalidade econômica da decisão de emigrar pauta-se pela provável renda urbana, equivalente ao valor presente do fluxo de renda esperado. Observa-se a expectativa de renda e não o salário real. A migração do campo para a cidade é o resultado das forças de expulsão e de atração. Esse modelo é, portanto atrativista e baseia-se na escolha racional.

Outra explicação plausível para a migração rural-urbana encontra-se na expansão do modelo de Todaro realizado por Cole e Sanders (1985). Adicionalmente, considera-se a influência do setor urbano de subsistência na propensão do migrante rural em deixar o campo. A entrada no setor de subsistência seria analisada pelo migrante rural como um período durante o qual poderia adquirir capital humano e tentar um emprego no setor urbano moderno, posteriormente. Isso explicaria o movimento de enorme contingentes populacionais do campo para as cidades, apesar dos elevados índices de desemprego no setor urbano moderno.

Os estudos de Todaro (1969) e Cole e Sanders (1985), considerou a renda rural como sendo composta basicamente pela renda do setor agrícola. Esse pressuposto desconsidera a possibilidade de o agricultor ou os membros da sua família empregarem parte do tempo de trabalho disponível em atividades não-agrícolas dentro. Porém, como salientado na introdução, é fato a consolidação das ocupações não-agrícolas no rural. Estudos sobre

estratégia de renda rural em áreas menos desenvolvidas têm mostrado que o trabalhador rural diversifica sua renda com atividades agrícolas e não-agrícolas.

De acordo com Lima (1999) as migrações rurais no Brasil respondem positivamente à renda urbana e negativamente à renda rural, e que o migrante maximiza os retornos ao seu estoques de capital humano mudando-se para áreas em que haja maior remuneração da sua força de trabalho. A economia rural não-agrícola contribuiu para a elevação da renda rural e, portanto, conteve as migrações rurais em áreas de baixa renda. A economia rural não-agrícola, portanto, tem mostrado um potencial de contribuir para a estabilização da renda rural em regiões menos desenvolvidas. Em consequência, essas atividades não-agrícolas devem ser consideradas como um fator de influência na decisão dos trabalhadores rurais em migrar para as zonas urbanas.

3.3. Metodologia

Esta seção apresenta a metodologia utilizada neste segundo ensaio. Inicia-se apresentando os índices de desigualdade, Índice de Gini e Índice de Theil. Esses índices além de permitir analisar diferenças nas desigualdades entre urbano e rural, serve também para validar²⁷ os resultados encontrados no modelo DEA. Logo após apresenta-se o modelo DEA orientação produto, que permitirá analisar as diferenças de eficiência entre o rural e urbano. Por fim, para corroborar os resultados da análise de eficiência, apresenta-se o Índice de Malmquist que permite verificar se houve melhoria na eficiência do rural entre os anos analisados, 1981 a 2007.

3.3.1. Desigualdade

3.3.1.1. Índice de Gini

Para a mensuração da distribuição de renda, utilizou-se uma das principais medidas de desigualdade, denominada de índice de Gini. De acordo com Hoffmann (1998), este índice se destaca pela possibilidade de sua associação com a curva de Lorenz, que representa a desigualdade de uma distribuição. Quando um único indivíduo apropria-se de toda renda, o índice assume valor um, representando a concentração máxima da renda. Se a

²⁷ Na próxima seção será apresentada a justificativa para o uso dos índices de desigualdade para validar o modelo DEA.

renda é perfeitamente distribuída entre a população, o índice assume valor zero. Neste sentido, o índice de Gini que varia de zero a um, podendo ser representado pela seguinte fórmula:

$$Gini = \frac{2 \sum_{i=1}^N ir_i}{N \sum_{i=1}^N r_i} - 1 - \frac{1}{N} \quad (3.2)$$

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.

3.3.1.2. 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 primeira medida é mais sensível a alterações na desigualdade dentro de grupos de rendimentos mais elevados, enquanto a segunda é mais sensível a alterações na desigualdade dentro dos grupos de renda mais baixa. (HOFFMANN, 1998). O T de Theil pode ser representado da seguinte forma:

$$T = \sum_i y_i \log \frac{x_i}{\mu} \quad (3.3)$$

onde 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; e, 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:

$$L = -\frac{1}{n} \sum_i \log \frac{x_i}{\mu} \quad (3.4)$$

onde n é o número de pessoas consideradas na análise; i é o número de ordem do indivíduo após a ordenação pelo valor do rendimento; e, 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. Outra característica importante dos índices de Theil é a possibilidade de decompô-los aditivamente em grupos ou variáveis sócio-demográficas, sendo possível, desta forma, estimar o poder explicativo de cada grupo para desigualdade total (decomposição estática), ou seja, para o T de Theil ou Redundância:

$$R = \sum_g \alpha_g \beta_g \ln \alpha_g + \sum_g \alpha_g \beta_g R_g \quad (3.5)$$

onde, R representa a redundância para toda população de análise; α_g é a razão entre a renda média do grupo g e a renda média global; β_g é a proporção da população ocupada no grupo g ; e R_g é o T de Theil do grupo g , Corrêa (1998). O primeiro termo desta expressão representa a desigualdade entre os grupos considerados o que por sua vez fornecerá o poder explicativo das variáveis sócio-demográficas, enquanto o segundo expressa a desigualdade intragrupos. A decomposição para o L de Theil pode ser expressa da seguinte forma:

$$L = -\sum_g \beta_g \ln \alpha_g + \sum_g \beta_g L_g \quad (3.6)$$

em que: L representa o L de Theil para toda população analisada; α_g é a razão entre a renda média do grupo g e a renda média global; β_g é a proporção da população ocupada no grupo g ; e L_g é o L de Theil do grupo g .

3.3.2. Eficiência

3.3.2.1. Análise envoltória de dados - DEA

A metodologia DEA se tornou uma importante ferramenta nos últimos anos para caracterizar as unidades como eficientes ou ineficientes, além de identificar as variáveis que podem ser alteradas para a melhoria do resultado. Tal relevância se deve ao fato de que as medidas obtidas são comparativas, se baseando na identificação das melhores práticas dentre a base de dados analisada ao fazer uso da fronteira de eficiência. Na literatura foram encontrados muitos estudos do uso de modelos de fronteiras de produção (paramétricas e não paramétricas; determinísticas e estocásticas) em diversas áreas de pesquisa como desempenhos de unidades de ensino e principalmente para agricultura. Entretanto, não foram encontradas referências ao uso específico de modelos DEA em eficiência do trabalho. Dessa forma, este artigo tem como objetivo contribuir para esta classe de estudos.

O DEA consiste em modelos e técnicas de construção de fronteiras de produção e medidas de eficiência relativa, calculada através da distância “radial” da fronteira, que indica a proporção em que devem ser reduzidos os *inputs*, ou aumentados os *outputs*, para que a fronteira eficiente seja alcançada. O conjunto de produção é limitado pela fronteira empírica de produção composta por aquelas DMUs que são eficientes. Assim, uma DMU será eficiente se, comparativamente às demais, obtiver maior produção para determinada quantidade de insumos ou utilizar menos recursos para gerar uma determinada quantidade de produto.

Segundo Gomes e Baptista (2004), objetiva-se encontrar a melhor DMU, seja ela real, ou virtual que se trata da combinação convexa de outras DMU eficientes denominadas *benchmarks*. Quando a DMU eficiente é uma virtual, é porque a DMU real não foi capaz de produzir maiores quantidades de produtos utilizando a mesma ou menor quantidade de insumos.

A determinação das DMUs eficientes envolve programação linear, onde se resolve para cada uma das unidades um sistema de equações lineares definido para mensurar o nível de eficiência de cada DMU.

O modelo proposto por Charnes, Cooper e Rhodes (1978), com retornos constantes de escala é apresentado a seguir:

$$\begin{aligned}
& \text{MAX } \phi, \lambda \phi, \\
& \text{s.a.} \\
& -\phi y_i + Y \lambda \geq 0, \\
& x_i - X \lambda \geq 0, \\
& \lambda \geq 0.
\end{aligned}$$

Em que:

y_i é um vetor ($m \times 1$) de quantidades de produto da i -ésima DMU;

x_i é um vetor ($k \times 1$) de quantidades de insumo da i -ésima DMU;

Y é uma matriz ($n \times m$) de produtos das n DMU's;

X é uma matriz ($n \times k$) de insumos das n DMU's;

λ é um vetor ($n \times 1$) de pesos;

ϕ é uma escalar que tem valores iguais ou maiores do que 1.

De acordo com Baptista (2002), ϕ^{-1} representa o aumento proporcional nos produtos que a i -ésima DMU pode alcançar mantendo fixa a quantidade de insumo. E $1/\phi$ é o escore de eficiência técnica da i -ésima DMU. Isto posto, o problema é resolvido n vezes, ou seja, uma vez para cada DMU.

Segundo Baptista e Gomes (2004), após executar os modelos DEA, as DMUs podem ser classificadas como tecnicamente eficientes ou ineficientes, sendo necessários a redução e/ou o aumento proporcional de insumos e produtos, respectivamente.

Segundo Golany e Roll (1989), cabe observar que a escolha das DMUs traz o pressuposto de homogeneidade de um grupo, ou seja, essas unidades têm que possuir similaridades para que a análise seja coerente. Em outras palavras, elas têm que realizar tarefas com objetivos similares, atuar sob a mesma condição de mercado e produzir os mesmos produtos a partir dos mesmos insumos, diferindo apenas em intensidade. Logo, a escolha dos estados brasileiros mostra-se consistente.

De acordo com Dyson *et al.* (*apud* Niederauer, 2002), DEA por ser uma técnica de ponto extremo a análise é sensível a ruídos, tais como erros de medição ou valores extremos. Outrossim, à medida que cresce o número de variáveis eleva também a chance de mais unidades alcançarem o desempenho máximo. Ademais, ela estima bem o desempenho “relativo”, mas converge muito vagarosamente para o desempenho “absoluto” porque está baseado em dados observados e não no ótimo ou no desejável. Embora DEA

possua essas limitações supramencionadas, essas características não impactarão na robustez dos resultados, pois não se tem valores extremos e o que deseja é mensurar de forma relativa a eficiência do rural e do urbano.

3.3.2.2. Índice Malmquist de produtividade

O índice de *Malmquist* foi inicialmente proposto por Malmquist (1953) na análise do “comportamento” do consumidor. Entretanto, foram Caves *et al.* (1982) que introduziram o conceito de índice de *Malmquist*, inspirado em sua análise da evolução da produtividade de cada unidade de produção relativamente à evolução do conjunto de unidades em que se insere. O índice de Malmquist, apesar de controlar o efeito da evolução tecnológica, não leva em conta o valor atribuído a cada produto ou recurso. Färe *et al.* (1995) propuseram um índice com essas mesmas características, baseado na composição de escores de eficiência gerados pela Análise Envoltória de Dados, chamado índice de Malmquist-DEA.

O índice de Malmquist orientação produto, calculado pela programação linear, é baseado nos trabalho de Caves *et al.* (1982), Färe *et al.* (1998) e Fried *et al.* (1993). Uma representação funcional da tecnologia é fornecida pela função de distância de produto, ou função de Shephard:

$$D_0^t(x^t, y^t) = \min \left\{ \theta : \left(\frac{y^t}{\theta} \right) \in P^t(x^t) \right\}, t = 1, 2, \dots, T.$$

Essa demonstração indica que para um par de insumos e produtos (x^t, y^t) , pertencentes à tecnologia t , a função distância de produto expressa a máxima expansão proporcional de produtos.

A vantagem na utilização das funções distância reside no fato de que tais funções caracterizam completamente a tecnologia, propriedade bastante útil na construção do modelo DEA para a estimação dos escores de eficiência.

O cálculo do índice Malmquist de produtividade é baseado em quatro funções de distância, quais sejam:

$$D_0^t(x^t, y^t), D_0^{t+1}(x^{t+1}, y^{t+1}), D_0^{t+1}(x^t, y^t), D_0^t(x^{t+1}, y^{t+1})$$

A segunda função acima é definida de forma análoga, utilizando dados do período $t+1$. As demais, por envolver dados do período t e $t+1$ são definidas como:

$$D_0^t(x^{t+1}, y^{t+1}) = \min \left\{ \theta : \left(\frac{y^{t+1}}{\theta} \right) \in P^t(x^{t+1}) \right\}$$

$$D_0^{t+1}(x^t, y^t) = \min \left\{ \theta : \left(\frac{y^t}{\theta} \right) \in P^{t+1}(x^t) \right\}$$

A função distância X relaciona dados t^{+1} com a tecnologia existente no período t , onde o resultado pode ser menor, igual ou maior que 1. Assim, os dados do período t^{+1} podem ser inviáveis para a tecnologia período t . A função de distância x refere-se a dados do período t com a tecnologia existente no período t^{+1} , podendo assumir valores menores, iguais ou maiores que 1.

Diante das funções de distância produto, têm-se o índice de Malmquist:

$$M_0(x^{t+1}, y^{t+1}, x^t, y^t) = \left[\frac{D_0^t(x^{t+1}, y^{t+1})}{D_0^t(x^t, y^t)} \times \frac{D_0^{t+1}(x^{t+1}, y^{t+1})}{D_0^{t+1}(x^t, y^t)} \right]^{1/2}$$

Ou seja, considerando a tecnologia do período t , a distância da DMU até a fronteira de produção considerando os níveis de consumo e de produção no período t e considerando os níveis dos mesmos insumos e produtos no período $t+1$. A primeira razão do índice utiliza a tecnologia do período t para fornecer medidas de mudanças de produtividade. A segunda razão lança mão da tecnologia do período t^{+1} para o cálculo de medidas de mudança de produtividade. Deste modo, tem-se um índice de produtividade relativo ao período t , cujo cálculo baseia-se numa média geométrica de duas razões de funções de distância produto.

A estimação do índice de *Malmquist* pelo uso da análise envoltória de dados, segundo Fare *et al.* (1994), permite decompor a mudança na produtividade total dos fatores em dois componentes: mudanças na eficiência e mudanças na tecnologia. Um valor maior que 1 (um) indica crescimento na produtividade total dos fatores, do período s para o período t , enquanto um valor menor que 1 (um) indica queda na produtividade total dos fatores.

Em sua forma decomposta, o índice de Malmquist permite separar as fontes de variação de produtividade em dois termos:

$$M_o(x_{t+1}, y_{t+1}, x_t, y_t) = \frac{d_o^{t+1}(x_{t+1}, y_{t+1})}{d_o^t(x_t, y_t)} * \left[\left(\frac{d_o^t(x_{t+1}, y_{t+1})}{d_o^{t+1}(x_{t+1}, y_{t+1})} \right) * \left(\frac{d_o^t(x_t, y_t)}{d_o^{t+1}(x_t, y_t)} \right) \right]^{1/2}$$

Nesta decomposição, a razão fora dos parênteses calcula a mudança na eficiência relativa denominada por Färe *et alii* (1994) de “mudança técnica”. Ela verifica se a produção da unidade está se aproximando ou se afastando da fronteira. Já a média geométrica das duas razões dentro dos parênteses mede a mudança na eficiência, ou seja, a mudança na tecnologia entre os dois períodos avaliados em x^t e x^{t+1} . Estes índices parciais são importantes porque permitem verificar se um eventual aumento da eficiência relativa de um ano para o outro deve ser atribuído ao aumento da produtividade da unidade avaliada ou à contração da fronteira de produção.

Os seguintes problemas de programação linear devem ser resolvidos para o cálculo de $M_o(x_{t+1}, y_{t+1}, x_t, y_t)$:

$$\begin{array}{ll} \left[d_0^t(x_t, y_t) \right]^{-1} = \max_{\theta\lambda} \theta & \left[d_0^{t+1}(x_{t+1}, y_{t+1}) \right]^{-1} = \max_{\theta\lambda} \theta \\ st & st \\ -\theta y_{i,t} + Y_t \lambda \geq 0 & -\theta y_{i,t+1} + Y_{t+1} \lambda \geq 0 \\ x_{i,t} - X_t \lambda \geq 0 & x_{i,t+1} - X_{t+1} \lambda \geq 0 \\ \lambda \geq 0 & \lambda \geq 0 \end{array}$$

$$\begin{array}{ll} \left[d_0^t(x_{t+1}, y_{t+1}) \right]^{-1} = \max_{\theta\lambda} \theta & \left[d_0^{t+1}(x_t, y_t) \right]^{-1} = \max_{\theta\lambda} \theta \\ st & st \\ -\theta y_{i,t+1} + Y_t \lambda \geq 0 & -\theta y_{i,t} + Y_{t+1} \lambda \geq 0 \\ x_{i,t+1} - X_t \lambda \geq 0 & x_{i,t} - X_{t+1} \lambda \geq 0 \\ \lambda \geq 0 & \lambda \geq 0 \end{array}$$

Como demonstrado, o índice de Malmquist-DEA é calculado pela média geométrica de dois índices, no qual o primeiro utiliza como referência a fronteira do período t e o segundo a fronteira do período $t+1$. Se o índice de produtividade total de Malmquist (PTF) for maior que um então ocorreu ganho de produtividade. O mesmo ocorre com os componentes do índice, ou seja, se $VET > 1$ então o estado obteve ganho em eficiência técnica; se por outro lado $VTC > 1$, então o estado obteve ganho de variação tecnológica. Vale lembrar que se o índice de produtividade total de Malmquist for maior que a unidade não significa que ocorreu ganho em todos os componentes do índice.

Essa possibilidade de desmembramento do índice de *Malmquist* é muito importante, pois permite entender a origem das alterações na produtividade, ou seja, se um aumento de produtividade é fruto de mudança técnica ou mudança da eficiência, ou das duas simultaneamente, ou deslocamento da unidade em relação à fronteira. De acordo com Nishimizu e Page (1982), quando os ganhos de produtividade estão sendo provenientes, principalmente, em função de deslocamentos na fronteira, estão existindo inovações que proporcionam estes deslocamentos. Por outro lado, ganhos de produtividade mais relacionados ao deslocamento das unidades em relação à fronteira são provenientes da difusão das tecnologias ou de fatores conjunturais.

O índice PTF de Malmquist pode ser decomposto em mais componentes quando existe fatores como imperfeições de mercado que fazem com que as firmas não estejam atuando em escala ótima. Sendo assim, o VET se divide em dois componentes: variação da eficiência técnica pura (VETP) e variação na eficiência de escala (VEE). Se uma firma possui diferentes valores de ET sob RCE e RVE, isso indica que ela apresenta ineficiência de escala, a qual pode ser medida pela diferença entre as ETs com relação a ambas as fronteiras.

3.3.3. Fonte de dados

Os dados utilizados no presente ensaio foram obtidos da Pesquisa Nacional por Amostra de Domicílios (PNAD) realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Mais especificamente utilizou-se os microdados da PNAD. Esta pesquisa foi implantada no Brasil em 1967. Até 1970, foi feita trimestralmente e, a partir de então, os levantamentos passaram a ser anuais sendo realizadas no mês de setembro de cada ano (com exceção dos anos censitários).

São fontes ricas de dados sobre os padrões social e econômico dos indivíduos e permitem o estudo do desenvolvimento sócio-econômico do país. Os temas básicos que integram o questionário são: população, características do domicílio (infra-estrutura), educação, trabalho, rendimento e habitação.

Desde 1981, compreende todas as regiões do Brasil, exceto a área rural da antiga Região Norte (Acre, Amapá, Amazonas, Pará, Rondônia e Roraima). Tocantins é considerado estado da Região Norte apenas a partir de 1992, apesar de ter sido criado em 1988. O desenho amostral estabelecido na pesquisa permite a expansão dos resultados para todo o país, assim como regiões, estados e áreas metropolitanas.

3.3.4. Tratamento dos dados

Tendo em vista o relevante período que a pesquisa abarca, 1981 a 2007, fez-se necessário um estudo da metodologia adotada pela PNAD para garantir comparabilidade no período apresentado. Ao longo dos anos algumas variáveis são recodificadas. Feito o estudo verificou-se a não necessidade de compatibilização das variáveis entre os anos. Somente a variável anos de estudo que teve sua categoria expandida, porém como a medida encontrada trata-se de um total, essa mudança não impactou nos resultados.

Os resultados da PNAD, por ser uma amostra, possuem informações sobre o peso (ponderação) que permite expandir os resultados da amostra para toda população. Neste sentido, foram obtidos o total dos dados em questão considerando o peso amostral.

3.3.4.1. Seleção das variáveis

Horas Trabalhadas: Representam o somatório das horas trabalhadas das pessoas economicamente ativas dos estados brasileiros. Para o computo das horas trabalhadas foi considerado o trabalho principal. Sendo assim, temos uma medida total e não uma média.

Anos de Estudos: Representam o somatório dos anos de estudos das pessoas economicamente ativas dos estados brasileiros. Sendo assim, temos uma medida total e não uma média.

Rendimento do trabalho: Representam o somatório da renda advinda do trabalho das pessoas economicamente ativas dos estados brasileiros. Para o computo foi considerado o trabalho principal. Sendo assim, temos uma medida total e não uma média.

3.4. Resultados e discussão

3.4.1. Medidas de desigualdade (Gini e Theil)

Considerando que o enfoque principal da presente pesquisa é o estudo da eficiência do trabalho no meio rural e meio urbano brasileiro, onde o produto utilizado é o rendimento advindo do trabalho principal, esta seção estuda a distribuição do mesmo. Outrossim, a desigualdade de renda é um dos temas mais discutidos no país, o que é compreensível tendo em vista sua complexidade e os bons resultados que o

desenvolvimento de políticas distributivas pode ter na redução da violência e combate à miséria. Tendo em vista que os insumos e produtos são dados em termos de totalidades de cada Estado no cálculo do DEA, o índice de GINI e Theil vem suprir esta carência do modelo ao mensurar o grau de distribuição do rendimento do trabalho.

A desigualdade de renda no Brasil se encontra entre as maiores do mundo como indicam os trabalhos seminais de Hoffmann e Duarte (1972), Fishlow (1972) e Langoni (1973) e, para as décadas de 1980 e 1990, os de Amadeo *et al.* (1994), Barros, Mendonça e Rocha (1993), Barros, Mendonça e Duarte (1995), Hoffman (1989) e Henriques (2000). Entretanto, alguns estudos recentes como o de Barros *et al.* (2007) verificaram a recente queda desta desigualdade.

A fim de compreender se a redução da desigualdade dos últimos anos foi mais evidente no meio urbano ou no meio rural, utilizou-se na presente pesquisa, a renda per capita do trabalho principal dado que o enfoque é o cálculo da eficiência do trabalho. Segundo Hoffman (1998), muitas vezes acontece subdeclaração dos altos rendimentos, pois algumas pessoas são receosas em dar informações temendo ser prejudicado com o pagamento de altos impostos, dentre outros motivos, fato este responsável por alguns vieses, como subestimação do nível de desigualdade. Apesar das referidas restrições quanto ao rendimento, quando se analisa somente a renda advinda do trabalho, esta limitação torna-se mínima.

A Tabela 3.1 a seguir apresenta o índice de GINI e a família de índices de Theil para cada estado referente aos anos 1981, 1995 e 2007. Pode-se verificar que a desigualdade de rendimento do trabalho, para o Brasil como um todo (rural e urbano), tem se mantido estável com uma ligeira queda nos últimos anos na maioria dos estados, em consonância com o trabalho de Barros *et al.* (2007). De acordo com o Instituto Brasileiro de Geografia e Estatística – IBGE deve-se dar atenção especial à desigualdade regional quando se estuda a desigualdade de renda no Brasil, devido à heterogeneidade das mesmas. Sendo assim, os cálculos são realizados por estado a fim de cobrir tais especificidades regionais. Pode-se verificar, também, um Sul e Sudeste com pior padrão de distribuição de renda do que as demais regiões, embora sejam mais desenvolvidas do que as outras macro regiões do país. O Norte e o Nordeste apresentam um padrão menos desigual dos que as demais, porém, algumas vezes este resultado não pode ser visto de forma positiva. Isso se deve ao fato de que em locais com maior incidência da pobreza tem-se nivelamento por baixo, em outras palavras, são homogeneamente pobres. Segundo Manso *et al.* (2006) a

Região Nordeste, possuidora de 28% da população brasileira, concentra, 49% dos pobres e 55% dos indigentes de todo o país.

De acordo com o estudo de Berni (2007), considerando o rendimento de todas as fontes, o coeficiente de Gini passou de 0,6005, em 1995, para 0,5693, em 2005, ou seja, reduziu-se em 5,20% deste período, em face à estabilização dos preços que repercutiu positivamente no setor informal da economia. Tal valor não destoa do encontrado para o rendimento do trabalho nos estados brasileiros, conforme Tabela 3.1.

Outra medida de desigualdade também aplicada na presente pesquisa trata-se do Índice de Theil, que se divide em redundância T de Theil, 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. Conforme Tabela 3.1, a desigualdade dentro do grupo mais rico está aumentando na maior parte dos estados. Não obstante, ela se revela mais próspera na camada mais baixa da distribuição de renda. Uma das possíveis explicações para esse fenômeno é a estabilização dos preços, dado que a desigualdade no Brasil apresentou seu valor máximo, no período em investigação, no ano de 1995, um ano após a instauração do Plano Real. Logo após deste momento, os indicadores de desigualdade sofreram redução.

Isto posto, procura-se responder se esta queda é verificada tanto no rural quanto no urbano e em qual deles se verifica um decréscimo mais expressivo (Tabelas 3.2 e 3.3). Nestas Tabelas são demonstrados os índices de desigualdade separados para a população rural e urbana, respectivamente. Os estados do norte não estão dispostos na Tabela abaixo porque a PNAD não cobria esses estados nos anos de 1981 e 1995.

Tabela 3.1. Distribuição de Renda dos Estados Brasileiros.

UF	1981			1995			2007		
	GINI	T-THEIL	L-THEIL	GINI	T-THEIL	L-THEIL	GINI	T-THEIL	L-THEIL
Rondônia	0.452	0.412	0.361	0.574	0.654	0.588	0.466	0.423	0.369
Acre	0.482	0.436	0.418	0.569	0.636	0.568	0.564	0.644	0.573
Amazonas	0.496	0.494	0.432	0.543	0.621	0.510	0.466	0.500	0.375
Roraima	0.483	0.429	0.405	0.473	0.424	0.382	0.489	0.453	0.418
Pará	0.538	0.600	0.524	0.576	0.688	0.593	0.494	0.543	0.435
Amapá	0.433	0.319	0.314	0.483	0.418	0.407	0.455	0.407	0.348
Tocantins	-	-	-	0.606	0.922	0.670	0.537	0.597	0.533
Maranhão	0.559	0.662	0.639	0.612	0.787	0.692	0.533	0.584	0.537
Piauí	0.569	0.728	0.739	0.595	0.750	0.676	0.631	0.877	0.814
Ceará	0.603	0.799	0.707	0.604	0.792	0.673	0.537	0.607	0.580
Rio G. do Norte	0.549	0.615	0.539	0.593	0.738	0.639	0.549	0.659	0.573
Paraíba	0.587	0.760	0.631	0.618	0.825	0.711	0.580	0.779	0.621
Pernambuco	0.556	0.678	0.584	0.569	0.699	0.576	0.521	0.622	0.512
Alagoas	0.520	0.591	0.473	0.614	0.842	0.670	0.581	0.971	0.628
Sergipe	0.548	0.585	0.545	0.578	0.689	0.600	0.520	0.582	0.501
Bahia	0.551	0.640	0.553	0.587	0.822	0.614	0.534	0.639	0.533
Minas Gerais	0.562	0.639	0.573	0.581	0.727	0.606	0.505	0.548	0.454
Espírito Santo	0.570	0.744	0.595	0.577	0.685	0.596	0.504	0.572	0.448
Rio de Janeiro	0.555	0.619	0.570	0.560	0.630	0.556	0.514	0.578	0.455
São Paulo	0.522	0.519	0.492	0.535	0.571	0.506	0.492	0.486	0.418
Paraná	0.530	0.573	0.504	0.571	0.657	0.587	0.517	0.562	0.477
Santa Catarina	0.496	0.479	0.439	0.525	0.548	0.490	0.454	0.420	0.353
Rio G. do Sul	0.531	0.577	0.515	0.559	0.651	0.562	0.493	0.500	0.430
Mato G. do Sul	0.524	0.564	0.488	0.567	0.664	0.571	0.553	0.701	0.551
Mato Grosso	0.491	0.506	0.424	0.541	0.591	0.506	0.480	0.472	0.401
Goiás	0.561	0.657	0.573	0.565	0.680	0.566	0.507	0.598	0.451
Distrito Federal	0.561	0.580	0.578	0.562	0.578	0.579	0.587	0.642	0.631

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1996 a 2007, IBGE.

É interessante destacar que muitas das atividades novas do rural estão ligadas ao setor informal da economia. Essas novas ocupações, como artesanato e cultivo de flores, são compostas predominantemente de indivíduos com baixa renda que não tinham como se proteger da perda do poder de compra no período de alta inflação. Porém, este segmento foi favorecido com a estabilização dos preços, impactando positivamente nos índices de desigualdade. Analisando a Tabela 3.2 pode-se dizer que boa parte da queda da desigualdade do Brasil foi influenciada pelo meio rural.

Tabela 3.2. Distribuição de Renda da Área Rural dos Estados Brasileiros.

UF	1981			1995			2007		
	GINI	T-THEIL	L-THEIL	GINI	T-THEIL	L-THEIL	GINI	T-THEIL	L-THEIL
Rondônia	-	-	-	-	-	-	0.425	0.367	0.306
Acre	-	-	-	-	-	-	0.439	0.458	0.359
Amazonas	-	-	-	-	-	-	0.300	0.186	0.166
Roraima	-	-	-	-	-	-	0.435	0.384	0.357
Pará	-	-	-	-	-	-	0.474	0.568	0.398
Amapá	-	-	-	-	-	-	0.426	0.334	0.316
Tocantins				0.687	1.403	0.899	0.445	0.410	0.366
Maranhão	0.522	0.577	0.590	0.626	0.812	0.726	0.553	0.610	0.581
Piauí	0.373	0.279	0.505	0.584	0.956	0.659	0.554	0.618	0.631
Ceará	0.446	0.396	0.413	0.446	0.397	0.367	0.491	0.441	0.498
Rio G. do Norte	0.410	0.322	0.297	0.422	0.346	0.322	0.482	0.670	0.455
Paraíba	0.328	0.226	0.227	0.496	0.492	0.447	0.420	0.337	0.330
Pernambuco	0.411	0.307	0.338	0.475	0.480	0.408	0.405	0.323	0.348
Alagoas	0.389	0.293	0.258	0.393	0.341	0.268	0.412	0.322	0.316
Sergipe	0.507	0.534	0.438	0.426	0.350	0.332	0.414	0.333	0.325
Bahia	0.421	0.340	0.319	0.480	0.656	0.412	0.428	0.422	0.352
Minas Gerais	0.467	0.536	0.386	0.521	0.606	0.483	0.454	0.461	0.381
Espírito Santo	0.545	0.934	0.548	0.521	0.581	0.473	0.429	0.411	0.331
Rio de Janeiro	0.446	0.432	0.366	0.477	0.512	0.395	0.372	0.287	0.249
São Paulo	0.461	0.449	0.378	0.462	0.405	0.366	0.428	0.411	0.326
Paraná	0.493	0.512	0.426	0.497	0.488	0.427	0.543	0.794	0.544
Santa Catarina	0.445	0.380	0.362	0.434	0.346	0.333	0.379	0.280	0.255
Rio G. do Sul	0.516	0.605	0.497	0.515	0.580	0.476	0.484	0.461	0.418
Mato G. do Sul	0.428	0.390	0.330	0.498	0.542	0.450	0.434	0.400	0.350
Mato Grosso	0.455	0.519	0.365	0.446	0.410	0.338	0.384	0.283	0.264
Goiás	0.486	0.504	0.422	0.477	0.540	0.402	0.434	0.532	0.358
Distrito Federal	0.565	0.693	0.573	0.580	0.616	0.613	0.565	0.773	0.561

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1996 a 2007, IBGE.

Quanto ao urbano, grande parte da motivação do estudo da desigualdade tem sido a consequente piora dos problemas sociais, tais como o aumento de criminalidade e da violência urbana, especialmente em grandes metrópoles. Conforme Tabela 3.3, grande parte dos estados apresentou queda da desigualdade. Tal comportamento é reflexo da expansão da oferta de mão-de-obra qualificada, incentivada pelo aumento na rentabilidade privada dos investimentos em educação conforme demonstrado no primeiro artigo. Muitos estudiosos, como Morley (2001), acreditam que o desequilíbrio de rendimento no mercado de trabalho aconteceu devido à alta demanda de mão-de-obra qualificada na fase de expressivo crescimento econômico que teria beneficiado as categorias mais qualificadas,

cuja oferta é mais inelástica, favorecendo ganhos salariais muito acima da média. Com o tempo, esse desvio seria corrigido porque a taxa de crescimento da economia tenderia a atingir um valor mais estável concomitantemente ao crescimento da oferta de mão-de-obra qualificada. A partir das análises realizadas no capítulo anterior é possível concluir que houve grande crescimento dos anos de estudo. Ademais, a taxa de crescimento da economia brasileira já se encontra num patamar mais estável.

Tabela 3.3. Distribuição de Renda da Área Urbana dos Estados Brasileiros.

UF	1981			1995			2007		
	GINI	T-THEIL	L-THEIL	GINI	T-THEIL	L-THEIL	GINI	T-THEIL	L-THEIL
Rondônia	0.452	0.412	0.361	0.574	0.654	0.588	0.466	0.417	0.368
Acre	0.482	0.436	0.418	0.569	0.636	0.568	0.560	0.619	0.566
Amazonas	0.496	0.494	0.432	0.543	0.621	0.510	0.475	0.511	0.388
Roraima	0.483	0.429	0.405	0.473	0.424	0.382	0.484	0.441	0.405
Pará	0.538	0.600	0.524	0.543	0.621	0.510	0.498	0.536	0.442
Amapá	0.433	0.319	0.314	0.473	0.424	0.382	0.456	0.408	0.348
Tocantins	-	-	-	0.576	0.688	0.593	0.540	0.597	0.542
Maranhão	0.568	0.661	0.607	0.483	0.418	0.407	0.519	0.561	0.502
Piauí	0.570	0.678	0.645	0.557	0.658	0.555	0.611	0.807	0.728
Ceará	0.599	0.757	0.698	0.590	0.751	0.637	0.524	0.580	0.539
R. G. do Norte	0.557	0.613	0.564	0.606	0.734	0.678	0.550	0.635	0.579
Paraíba	0.581	0.712	0.636	0.613	0.790	0.692	0.583	0.775	0.624
Pernambuco	0.579	0.719	0.636	0.571	0.696	0.578	0.520	0.615	0.499
Alagoas	0.562	0.658	0.569	0.638	0.859	0.747	0.587	0.991	0.639
Sergipe	0.543	0.559	0.570	0.575	0.662	0.591	0.512	0.562	0.480
Bahia	0.577	0.672	0.626	0.595	0.805	0.634	0.536	0.631	0.537
Minas Gerais	0.561	0.614	0.586	0.578	0.712	0.598	0.503	0.542	0.449
Espírito Santo	0.559	0.660	0.575	0.575	0.672	0.594	0.508	0.579	0.456
Rio de Janeiro	0.552	0.608	0.563	0.558	0.622	0.549	0.515	0.578	0.456
São Paulo	0.520	0.512	0.489	0.534	0.566	0.502	0.491	0.484	0.418
Paraná	0.530	0.566	0.512	0.569	0.647	0.584	0.511	0.534	0.463
Santa Catarina	0.517	0.515	0.473	0.522	0.536	0.483	0.460	0.429	0.361
Rio G. do Sul	0.524	0.549	0.496	0.557	0.639	0.556	0.492	0.500	0.425
Mato G. do Sul	0.529	0.566	0.501	0.573	0.671	0.584	0.557	0.706	0.557
Mato Grosso	0.499	0.494	0.444	0.545	0.590	0.517	0.495	0.494	0.424
Goiás	0.569	0.665	0.598	0.570	0.678	0.577	0.508	0.594	0.450
Distrito Federal	0.559	0.575	0.574	0.560	0.573	0.574	0.585	0.632	0.628

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1996 a 2007, IBGE.

A partir dos resultados acima se pode inferir que a distribuição de renda do rural e urbano não destoa de forma a impactar negativamente no cotejamento da análise de eficiência entre os espaços considerados.

3.4.2. Estudo da eficiência do trabalho

Nas Tabelas 3.4 e 3.5, são apresentadas as eficiências relativas dos estados, contemplando a área urbana e área rural²⁸, de acordo com as fronteiras de eficiência nos anos de 1981, 1990, 1992, 1999, 2001 e 2007. Pretendeu-se verificar, ao analisar estas áreas em conjunto, se há muitos estados “rurais” atingindo a eficiência em relação aos estados “urbanos”. Como se pode verificar nesta Tabela, em 1981 havia três unidades eficientes, sendo duas no rural e uma no urbano, porém comparativamente as unidades urbanas são mais eficientes. Em 1981, a área urbana do Rio de Janeiro, São Paulo e de Minas Gerais possui quase o dobro da eficiência da área rural. São poucos os estados nos quais a área rural é mais eficiente, como Maranhão e Pernambuco. Em 1990, esta situação se inverte e persiste em 1992, havendo maior número de unidades eficientes no meio urbano. Todavia, neste período a diferença de eficiência é menor que o dobro nos estados supramencionados. Para os demais estados verifica-se o arrefecimento da diferença entre o rural e urbano, sendo que ambos tiveram ganho de eficiência porém a do meio rural foi mais expressiva.

Analisando os demais anos, 1999 a 2007, constata-se que o urbano também apresenta mais estados com eficiência máxima, o que já era esperado uma vez que as cidades trazem benefícios como maior oferta e concentração de bens e serviços públicos e privados, que se traduzem em elementos de eficiência, ou seja, a cidade dispõe de uma logística que oferece competitividade econômica. Contudo, mesmo que a princípio o rural apresente relativamente menos eficiência que o urbano, ao longo dos anos esse cenário vem melhorando. Em 2007 a diferença de eficiência é pouco expressiva para os estados do Sul, Sudeste e Centro-Oeste. Em contrapartida, para os estados do Norte e Nordeste a diferença persiste. Pode-se inferir que devido ao fato destas regiões serem menos desenvolvidas, elas apresentam um diferencial de renda maior entre o urbano e rural que se trata de um indicador de desenvolvimento.

²⁸ Nas quais as assinaladas com asterisco referem-se ao meio rural.

Tabela 3.4. Distribuição da Eficiência do Brasil para os Anos 1981, 1990 e 1992.

Estados	1981	1990	1992
	ET (RC)	ET (RC)	ET (RC)
Maranhão	0.525	0.397	0.397
Piauí	0.657	0.421	0.421
Ceará	0.651	0.42	0.42
Rio G. Norte	0.546	0.423	0.423
Paraíba	0.455	0.432	0.432
Pernambuco	0.47	0.448	0.448
Alagoas	0.305	0.464	0.464
Sergipe	0.176	0.476	0.476
Bahia	0.39	0.543	0.543
Minas Gerais	0.506	0.569	0.569
Esp. Santo	0.402	0.675	0.675
Rio de Jan.	0.703	0.593	0.593
São Paulo	1	0.854	0.854
Paraná	0.419	0.667	0.667
San. Catarina	0.439	0.805	0.805
Rio G. do Sul	0.717	0.68	0.68
M. G. do Sul	0.545	0.675	0.675
Mato Grosso	0.352	0.704	0.704
Goiás	0.332	0.749	0.749
Dist. Federal	0.733	1	1
Maranhão*	0.752	0.345	0.345
Piauí*	0.057	0.199	0.199
Ceará*	0.16	0.304	0.304
Rio G. Norte*	0.077	0.263	0.263
Paraíba*	1	0.201	0.201
Pernambuco*	0.736	0.376	0.376
Alagoas*	0.138	0.379	0.379
Sergipe*	0.106	0.45	0.45
Bahia*	0.352	0.411	0.411
Minas Gerais*	0.219	0.377	0.377
Esp. Santo*	0.222	0.308	0.308
Rio de Jan.*	0.317	0.394	0.394
São Paulo*	0.732	0.582	0.582
Paraná*	0.377	0.352	0.352
San. Catarina*	0.164	0.428	0.428
Rio G. Sul*	0.941	0.398	0.398
Mato G. Sul*	0.708	0.494	0.494
Mato Grosso*	0.393	0.572	0.572
Goiás*	0.315	0.535	0.535
Dist. Federal*	1	0.493	0.493
Média	0.477	0.496	0.496

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1996 a 2007, IBGE.

Mais especificamente, o rural ainda continua menos eficiente, porém as medidas relativas estão aumentando corroboradas pelo índice de Malmquist apresentado na próxima seção. Para atingir a fronteira, os estados ineficientes podem estudar os valores absolutos das variáveis da DMU eficiente, para que possíveis mudanças sejam sinalizadas.

Nos últimos anos tem-se observado a emergência de um espaço rural multifuncional com a introdução de uma maior diversificação econômica. Verifica-se políticas que procuram convertê-lo num espaço de consumo, emergindo atividades não agrícolas e, deste modo, diminuindo a diferença entre a renda rural e a urbana, um indicador de equilíbrio do desenvolvimento social. A análise de eficiência apresentada acima se contrapõe a certas convenções sobre o meio rural. A modernização da agricultura e o surgimento de novas ocupações mostram que nem sempre rural é sinônimo de atraso, ou ineficiência, dada a inserção de alta tecnologia nos processos produtivos no meio rural e consolidação de novas ocupações. E, principalmente, que o processo do êxodo passa por um período de transformação. Segundo o modelo de Lewis, o êxodo do campo para as cidades será realizado até o momento em que os salários urbanos não mais excederem o produto médio do campo acrescido dos custos da transferência do campo para a cidade. As mudanças em curso têm levado vários autores, como os do Projeto RURBANO a discutir sobre o 'Novo Rural Brasileiro', que se mostra mais eficiente e próximo do urbano.

Ademais, dos diversos fatores que determinam a remuneração do trabalhador, a educação é uma das variáveis que mais tem sido alvo de pesquisa, conforme aponta Coelho e Corseuil (2002). Sua importância consiste no fato de que ao adquirir educação o indivíduo obtém mais conhecimento, o que leva a desenvolver sua capacidade de raciocínio e conseqüentemente permite a ele executar de forma mais eficiente as tarefas que lhe forem conferidas no trabalho. Devido ao fato das novas ocupações necessitarem de mais qualificação, a expansão das ocupações não agrícolas como alternativas de renda nas zonas rurais, portanto, estaria estreitamente vinculada ao aumento da produtividade no rural.

Tabela 3.5. Distribuição da Eficiência do Brasil para os anos 1999, 2001 e 2007.

Estados	1999	2001	2007
	ET (RC)	ET (RC)	ET (RC)
Tocantins	0.491	0.578	0.505
Maranhão	0.34	0.399	0.352
Piauí	0.381	0.371	0.395
Ceará	0.443	0.434	0.357
Rio G. Norte	0.432	0.45	0.415
Paraíba	0.541	0.419	0.408
Pernambuco	0.406	0.45	0.365
Alagoas	0.396	0.425	0.456
Sergipe	0.425	0.437	0.432
Bahia	0.422	0.45	0.403
Minas Gerais	0.549	0.556	0.533
Esp. Santo	0.589	0.593	0.536
Rio de Jan.	0.668	0.659	0.552
São Paulo	0.783	0.823	0.676
Paraná	0.665	0.695	0.675
San. Catarina	0.716	0.746	0.712
Rio G. Sul	0.674	0.696	0.602
Mato G. Sul	0.584	0.666	0.681
Mato Grosso	0.626	0.775	0.613
Goiás	0.575	0.636	0.61
Dist. Federal	1	1	1
Tocantins*	0.381	0.312	0.266
Maranhão*	0.398	0.394	0.232
Piauí*	0.247	0.208	0.182
Ceará*	0.198	0.197	0.165
Rio G. Norte*	0.277	0.288	0.297
Paraíba*	0.205	0.207	0.209
Pernambuco*	0.226	0.232	0.194
Alagoas*	0.313	0.275	0.209
Sergipe*	0.254	0.293	0.237
Bahia*	0.315	0.301	0.252
Minas Gerais*	0.349	0.402	0.382
Esp. Santo*	0.381	0.351	0.432
Rio de Jan.*	0.431	0.505	0.438
São Paulo*	0.611	0.554	0.491
Paraná*	0.394	0.487	0.517
San. Catarina*	0.435	0.526	0.538
Rio G. Sul*	0.467	0.469	0.501
Mato G. Sul*	0.569	0.576	0.414
Mato Grosso	0.499	0.534	0.517
Goiás*	0.518	0.509	0.458
Dist. Federal*	1	0.72	0.6
Média	0.48	0.49	0.448

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1996 a 2007, IBGE.

3.4.3. Malmquist

A estimação do índice de *Malmquist* pelo uso da análise envoltória de dados, segundo Fare *et al.* (1994), permite decompor a mudança da produtividade total dos fatores em dois componentes: mudanças na eficiência técnica (VET) e mudanças na tecnologia (VTC).

Essa possibilidade de desmembramento do índice de *Malmquist* é muito importante, pois permite entender a origem das alterações na produtividade. De acordo com Nishimizu e Page (1982), quando os ganhos de produtividade estão sendo provenientes, principalmente, de deslocamentos na fronteira, estão ocorrendo inovações que proporcionam estes deslocamentos devido a mudanças na tecnologia. Nesta pesquisa, essas inovações dizem respeito à automatização das atividades, aumento da qualificação (anos de estudo) dos trabalhadores. Por outro lado, ganhos de produtividade relacionada ao deslocamento das unidades em relação à fronteira são provenientes da difusão das tecnologias ou de fatores conjunturais. Neste segmento, encontram-se as ocupações antes consideradas urbanas emergindo cada vez mais no meio rural e, deste modo, difundindo tecnologia.

Para analisar as mudanças na fronteira e em relação à fronteira entre os anos que a dissertação abarca o método aplica o algoritmo de programação linear de DEA para construção da fronteira de produção em ambos os períodos analisados e considera a distância de cada estado ou DMU considerada, antes e depois, às duas fronteiras distintas.

As Tabelas 3.6, 3.7 e 3.8 desta seção apresentam o resultado do índice de Malmquist²⁹, subdividido em 1981 a 1990, 1992 a 1999 e 2001 a 2007 respectivamente, compreendendo os anos intra censo. Neste contexto, verifica-se pela Tabela 3.6 que em média no período de 1981 a 1990 não houve variação maior que um na produtividade dos fatores (PTF) nos estados (área rural e urbana). Embora se observe um aumento de cerca de 20% na eficiência técnica VET, essa não foi suficiente para cobrir a queda de 64% na variação tecnológica VTC. Esse resultado indica que mesmo com o aumento da tecnologia, representada no rural pela mecanização da agricultura, esta não foi suficiente para melhorar a eficiência do trabalho e gerar maior rendimento ao trabalhador. Cabe destacar que a análise é feita apenas do lado do trabalhador, ou seja, sobre a capacidade de se extrair o máximo de renda com o uso dos insumos horas trabalhadas e anos de estudo.

²⁹ Novamente as assinaladas com asterisco referem-se ao meio rural.

Pode-se inferir que a mecanização não trouxe benefícios quanto ao rendimento do trabalhador, uma vez que pela análise DEA a tecnologia não se mostrou positiva. Analisando a mudança na eficiência técnica, verifica-se que o rural obteve maior desempenho que o urbano. Em outras palavras, tem-se mais estados do rural elevando a eficiência.

Tal comportamento, salvo algumas particularidades, também é observado para o período de 1992 a 1999. Não obstante, em 2001 a 2007, conforme Tabela 3.8, em média houve variação positiva de 57% na PTF, sendo que os ganhos em produtividade total dos estados brasileiros se devem mais às variações tecnológicas que aos outros componentes. Desta forma, a inovação tecnológica vem sendo, predominantemente, a razão dos ganhos de produtividade dos estados brasileiros. Observa-se também que os ganhos de eficiência técnica (VET) são os que menos contribuem para os ganhos em produtividade. Neste sentido, o efeito da aproximação dos estados na direção da fronteira tecnológica do país definido como VET (efeito *catching up*) é pouco expressiva. Intuitivamente pode-se afirmar que esse efeito revela a capacidade de absorção da tecnologia disponível que se difunde no território nacional.

Entretanto, como revelam os resultados desta seção, a absorção desta capacidade tecnológica é menos intensa quando comparada com a velocidade de difusão apenas no período de 2001 a 2007 analisado. Neste último período tem-se o efeito expressivo do VTC, que trata do deslocamento da fronteira entre dois períodos (progresso técnico ou tecnológico).

Tabela 3.6. Cálculo do índice de Malmquist para os estados brasileiros, 1981 a 1990.

Estados	VET	VTC	PTF
Maranhão	0.756	0.376	0.284
Piauí	0.64	0.429	0.275
Ceará	0.645	0.395	0.255
Rio G. do Norte	0.774	0.44	0.34
Paraíba	0.951	0.475	0.452
Pernambuco	0.953	0.439	0.418
Alagoas	1.519	0.413	0.628
Sergipe	2.705	0.425	1.151
Bahia	1.394	0.44	0.614
Minas Gerais	1.124	0.439	0.493
Espírito Santo	1.679	0.475	0.797
Rio de Jan.	0.844	0.524	0.442
São Paulo	0.854	0.442	0.378
Paraná	1.591	0.415	0.66
San. Catarina	1.834	0.447	0.82
Rio G. do Sul	0.948	0.455	0.431
Mato G. do Sul	1.238	0.394	0.488
Mato Grosso	2.001	0.392	0.785
Goiás	2.258	0.401	0.906
Dist. Federal	1.365	0.523	0.714
Maranhão*	0.459	0.253	0.116
Piauí*	3.524	0.253	0.892
Ceará*	1.902	0.253	0.481
Rio G. do Norte*	3.428	0.273	0.935
Paraíba*	0.201	0.253	0.051
Pernambuco*	0.511	0.254	0.13
Alagoas*	2.755	0.253	0.697
Sergipe*	4.257	0.257	1.093
Bahia*	1.168	0.253	0.296
Minas Gerais*	1.724	0.259	0.447
Espírito Santo*	1.387	0.258	0.358
Rio de Jan.*	1.241	0.294	0.365
São Paulo*	0.796	0.279	0.222
Paraná*	0.934	0.262	0.245
San. Catarina*	2.613	0.274	0.717
Rio G. do Sul*	0.423	0.265	0.112
Mato G. do Sul*	0.697	0.253	0.176
Mato Grosso*	1.455	0.261	0.38
Goiás*	1.697	0.256	0.434
Dist. Federal*	0.493	0.32	0.158
Média	1.192	0.339	0.404

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1981 a 2007, IBGE.

Tabela 3.7. Cálculo do índice de Malmquist para os estados brasileiros, 1992 a 1999.

Estados	VET	VTC	PTF
Maranhão	0.857	0.016	0.013
Piauí	0.906	0.016	0.014
Ceará	1.055	0.016	0.016
Rio G. do Norte	1.022	0.016	0.016
Paraíba	1.25	0.017	0.021
Pernambuco	0.905	0.016	0.015
Alagoas	0.854	0.016	0.014
Sergipe	0.892	0.016	0.014
Bahia	0.777	0.016	0.012
Minas Gerais	0.965	0.016	0.015
Espírito Santo	0.872	0.016	0.014
Rio de Janeiro	1.125	0.017	0.019
São Paulo	0.917	0.017	0.015
Paraná	0.997	0.016	0.016
Santa Catarina	0.889	0.016	0.014
Rio G. Sul	0.992	0.016	0.016
Mato G. do Sul	0.865	0.016	0.013
Mato Grosso	0.889	0.016	0.014
Goiás	0.767	0.016	0.012
Dist. Federal	1	0.017	0.017
Maranhão	1.153	0.016	0.018
Piauí	1.241	0.016	0.019
Ceará	0.652	0.016	0.01
Rio G. do Norte	1.055	0.016	0.016
Paraíba	1.018	0.016	0.016
Pernambuco	0.6	0.016	0.009
Alagoas	0.824	0.016	0.013
Sergipe	0.563	0.016	0.009
Bahia	0.767	0.016	0.012
Minas Gerais	0.926	0.016	0.014
Espírito Santo	1.239	0.016	0.019
Rio de Janeiro	1.095	0.016	0.017
São Paulo	1.048	0.016	0.016
Paraná	1.118	0.016	0.017
Santa Catarina	1.016	0.016	0.016
Rio G. do Sul	1.175	0.016	0.018
Mato G. do Sul	1.153	0.016	0.018
Mato Grosso	0.873	0.016	0.014
Goiás	0.968	0.016	0.015
Dist. Federal	2.027	0.016	0.033
Média	0.959	0.016	0.015

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1981 a 2007, IBGE.

Tabela 3.8. Cálculo do índice de Malmquist para os estados brasileiros, 2001 a 2007.

Estados	VET	VTC	PTF
Tocantis	0.874	1.738	1.519
Maranhão	0.883	1.738	1.534
Piauí	1.063	1.738	1.848
Ceará	0.821	1.738	1.428
Rio G. do Norte	0.923	1.738	1.604
Paraíba	0.973	1.738	1.692
Pernambuco	0.813	1.738	1.413
Alagoas	1.075	1.738	1.868
Sergipe	0.99	1.738	1.72
Bahia	0.894	1.738	1.554
Minas Gerais	0.958	1.738	1.664
Espírito Santo	0.903	1.738	1.569
Rio de Janeiro	0.837	1.793	1.501
São Paulo	0.821	1.738	1.427
Paraná	0.97	1.738	1.687
Santa Catarina	0.955	1.738	1.659
Rio G. Sul	0.865	1.738	1.504
Mato G. do Sul	1.023	1.738	1.778
Mato Grosso	0.791	1.738	1.374
Goiás	0.959	1.738	1.666
Dist. Federal	1	1.793	1.793
Tocantis	0.852	1.738	1.481
Maranhão	0.588	1.738	1.022
Piauí	0.876	1.738	1.522
Ceará	0.838	1.738	1.456
Rio G. do Norte	1.032	1.738	1.793
Paraíba	1.009	1.738	1.754
Pernambuco	0.838	1.738	1.457
Alagoas	0.76	1.738	1.321
Sergipe	0.809	1.738	1.405
Bahia	0.838	1.738	1.456
Minas Gerais	0.949	1.738	1.649
Espírito Santo	1.229	1.738	2.136
Rio de Janeiro	0.868	1.738	1.508
São Paulo	0.886	1.738	1.54
Paraná	1.063	1.738	1.847
Santa Catarina	1.023	1.738	1.777
Rio G. do Sul	1.067	1.738	1.854
Mato G. do Sul	0.719	1.738	1.25
Mato Grosso	0.969	1.738	1.684
Goiás	0.9	1.738	1.565
Dist. Federal	0.833	1.738	1.448
Média	0.906	1.74	1.576

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1981 a 2007, IBGE.

Segundo Laurenti e Del Grossi (1999) verifica-se uma contínua perda de capacidade de absorção de mão-de-obra das atividades agrícolas, em função da mecanização e, indiretamente, na maior atratividade relativa das atividades não-agrícolas quanto a menor pilosidade do trabalho, maiores segurança e remuneração. Alves e Valente (2003) ressalta que as atividades não-agrícolas no rural são responsáveis pela maior oferta de ocupações e postos de trabalho, cujas remunerações são mais elevadas quando comparadas com as obtidas nas atividades rurais ligadas à agropecuária tradicional. Destarte, o meio rural brasileiro apresenta crescente diversificação de atividades agrícolas e não-agrícolas, onde se espera uma dissipação da tecnologia antes concentrada apenas no urbano.

3.4.4. Análise de eficiência, 1981 a 2007

Esta seção se destina a comparar os resultados obtidos em 2007 com os de 1981. Foram consideradas como DMUs os estados brasileiros, dividindo-os segundo a situação censitária. Excluíram-se os estados do Norte, Rondônia, Acre, Amazonas, Roraima, Pará, Amapá, uma vez que a PNAD não cobria a área rural desses estados em 1981. Ademais, Tocantins também não é analisado, pois ainda pertencia a Goiás. No problema de eficiência desta seção³⁰, também não se considera o Distrito Federal, em consonância com o estudo de Marinho e Benegas (2002), pela relação amorfa com os demais estados. A pesquisa, portanto, abrange 19 estados e, conseqüentemente, 38 DMUs.

É interessante destacar que as análises realizadas neste ensaio referem-se somente à comparação das unidades urbanas e rurais. Ou seja, busca-se verificar se diante do aumento das ocupações não-agrícolas, a área rural aproximou-se, em termos de eficiência do trabalho, da área urbana. Sendo assim, não se preocupou em verificar o porquê de um estado ser mais eficiente do que o outro, importando somente as discrepâncias quanto ao setor censitário (rural e urbano).

As Figuras 3.2 e 3.3 demonstram a eficiência técnica dos estados brasileiros, com o coorte para o setor censitário (urbano e rural), dos anos de 1981 e 2007.

³⁰ Dado que o objetivo primordial do presente ensaio refere-se ao cotejamento do rural com o urbano, o Distrito Federal será considerado nas análises da próxima seção. Para esta comparação, o método DEA não é influenciado por outilier, uma vez que as comparações são relativas.

Como se pode verificar, somente uma unidade urbana, São Paulo, e uma rural, Paraíba, que apresentaram eficiência máxima em 1981. Próximos da eficiência máxima no urbano, apenas Rio Grande do Sul e Rio de Janeiro, que mesmo assim, não ultrapassaram 80%. No rural, Rio Grande do Sul chega bem próximo da máxima eficiência, cerca de 95%. Excetuando Mato Grosso do Sul, Maranhão, Pernambuco e São Paulo, nenhuma DMU rural apresentou eficiência superior a 50%. Os estados do nordeste e sudeste apresentaram o urbano bem mais eficiente que o rural. Em contrapartida, no centro-oeste e sul, o rural tem o desempenho um pouco melhor que a área urbana dos estados brasileiros.

Em 2007, como se pode observar na figura 3.2, o rural não obteve nenhuma DMU mais eficiente que o urbano. Não obstante, é preciso salientar, pois se trata de uma medida relativa, que o urbano aumentou sua eficiência em relação a 1981, bem como o rural, que será corroborado na próxima seção pelo índice de malmquist. Ademais, mais da metade dos estados do rural, em 2007, obtiveram eficiência superior a 50%. Neste sentido, numa análise superficial, pode-se inferir que o rural de 2007 apresenta maior eficiência em relação a 1981³¹. Cotejando os dados de eficiência de 1981 e 2007, verifica-se que tanto na área urbana quanto na rural a maioria dos estados chegou mais próxima de seus respectivos benchmarks. No urbano, destacam-se Paraná, Mato Grosso do Sul e Rio de Janeiro, todos com mais de 90%. No rural, embora tenha ocorrido aproximação, esta não foi suficiente para colocar nenhuma DMU rural acima de 80% da máxima eficiência encontrada.

Diante do exposto acima, verifica-se maior homogeneidade das eficiências entre as situações censitárias dos estados brasileiros em 2007. Neste momento, pode-se corroborar a visão de Graziano (1999) e Kageyama (2003). Segundo estes autores, é perceptível a expansão da interseção entre o rural e o urbano. Em outras palavras, as diferenças entre o rural e urbano estão se atenuando. Pode-se inferir que uma das causas desta aproximação seja devido à maior diversificação das ocupações. As atividades não-agrícolas, geralmente demandam maior qualificação e conseqüentemente são responsáveis por maiores salários. Ademais, o simples fato de apresentar ao trabalhador que reside no rural maior possibilidade de emprego, pela lei da oferta e demanda, também aumenta sua remuneração em qualquer ocupação. Neste sentido, as ocupações não-agrícolas contribuíram para o aumento da eficiência na área rural dos estados brasileiros.

³¹ Esta assertiva será corroborada pelo índice de produtividade (Malmquist) da próxima subseção.

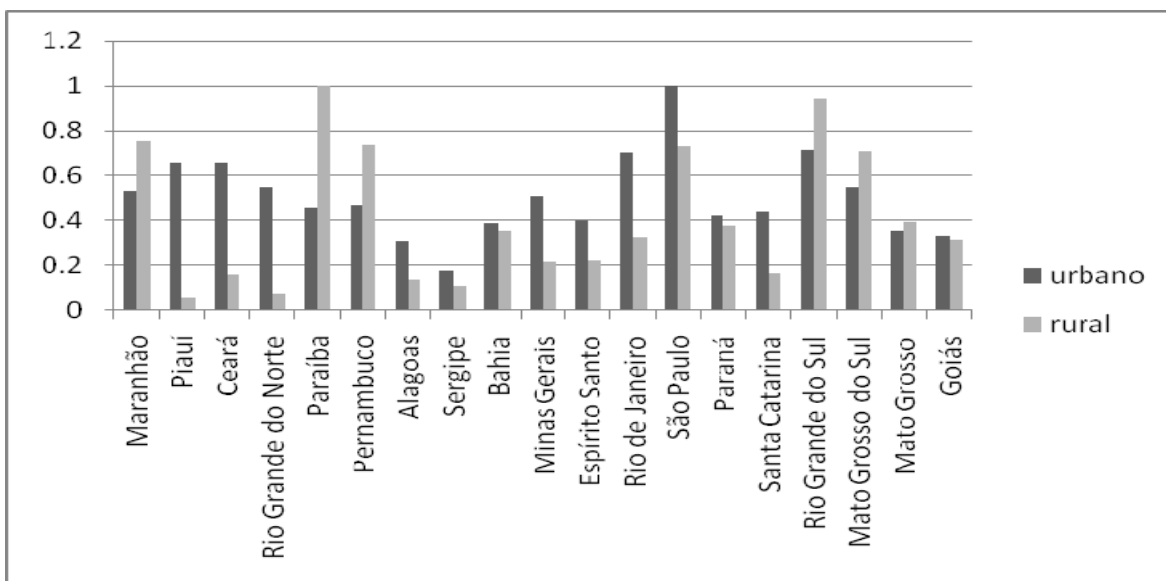


Figura 3.2. Eficiência Técnica dos estados brasileiros, 1981.

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1981, IBGE.

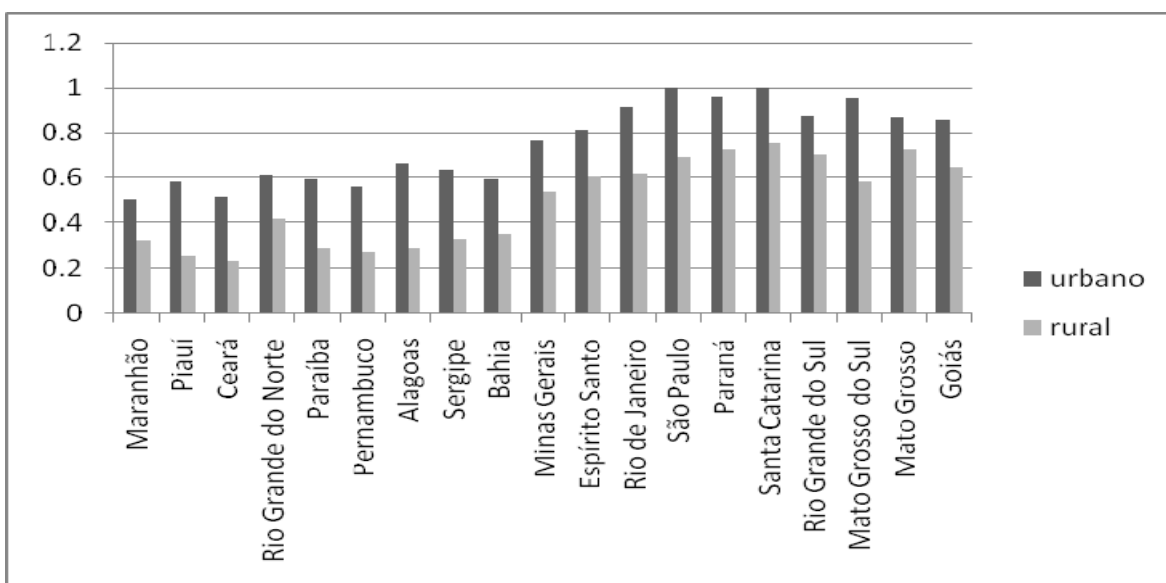


Figura 3.3. Eficiência Técnica dos estados brasileiros, 2007.

Fonte: Elaborado pela autora a partir dos microdados da PNAD-2007, IBGE.

Índice de Malmquist

Conforme apresentado na metodologia, o índice de Malmquist permite decompor a variação da PTF em mudança de eficiência técnica e mudança tecnológica. As figuras 3.4, 3.5 e 3.6 representam o índice de Malmquist PTF e suas decomposições VTC e VET.

A partir da figura 3.4 é possível verificar que nos estados do sudeste, excetuando São Paulo, o desempenho da eficiência técnica da parte rural foi superior que a urbana. É

importante salientar, que esse aumento não significa que o sudeste rural possui maior eficiência técnica que o urbano. O que se depreende é que a área rural do sudeste obteve ganhos de eficiência técnica superiores aos do urbano na comparação do que possuíam em 1981 em relação a 2007. Esse tipo de deslocamento das unidades em relação à fronteira, VET, são provenientes da difusão das tecnologias ou de fatores conjunturais. Sendo assim, o deslocamento em pauta indica que a tecnologia, antes mais centralizada no urbano, penetra cada vez mais no rural. Quanto aos fatores conjunturais, pode-se inferir que concerne às mudanças na composição das ocupações no rural, antes composta em grande parte por ocupações agrícolas.

Considerando os dois setores censitários, em média, os ganhos de eficiência técnica foram cerca de 53% para o conjunto de estados analisado.

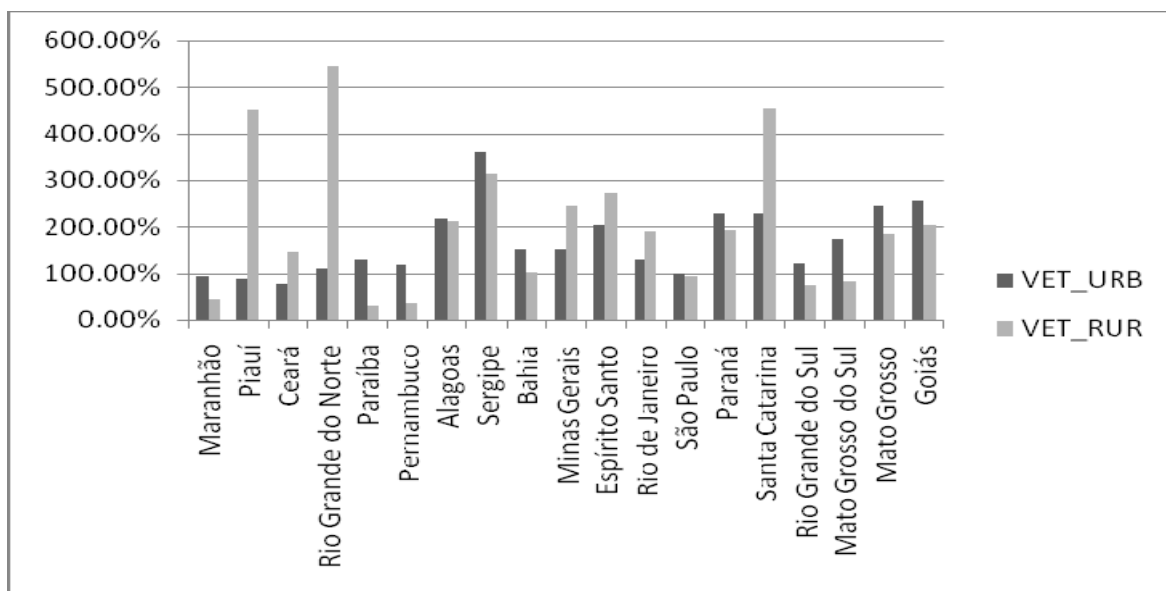


Figura 3.4: Variação da Eficiência Técnica (VET) dos estados brasileiros, 1981-2007.
 Fonte: Elaborado pela autora a partir dos microdados da PNAD-1981 e 2007, IBGE.

No que diz respeito ao VTC, verifica-se que a variação urbana superou a variação rural em todos os estados. Isso era esperado, uma vez que o VTC é o deslocamento da fronteira devido a inovações, que são em sua grande parte, criada nas cidades.

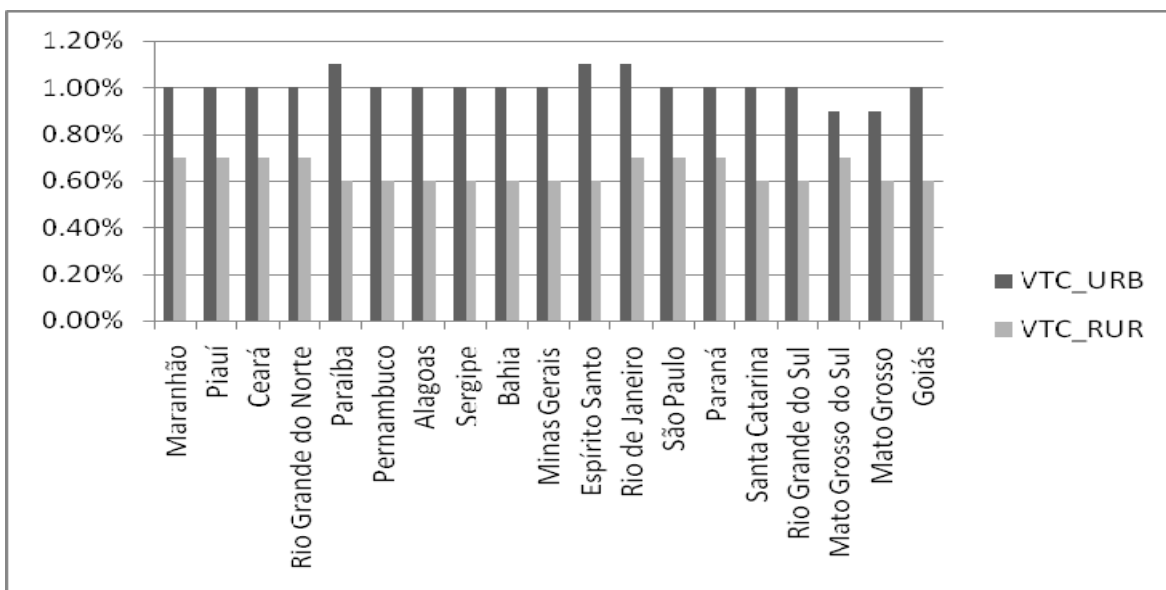


Figura 3.5: Variação Tecnológica (VTC) dos estados brasileiros, 1981-2007.

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1981 e 2007, IBGE.

Em média, a variação na produtividade total dos fatores do trabalho variou 1,2% para o conjunto dos estados analisados. Quanto ao PTF, temos um equilíbrio entre o meio rural e o meio urbano, onde as variações no urbano não se mostraram expressivamente maiores que o urbano. Como demonstrado acima, esse comportamento foi devido ao grande aumento da eficiência técnica do rural. Caso contrário, o gráfico do PTF seria semelhante ao VTC, no qual a variação urbana superou a rural em todos os estados.

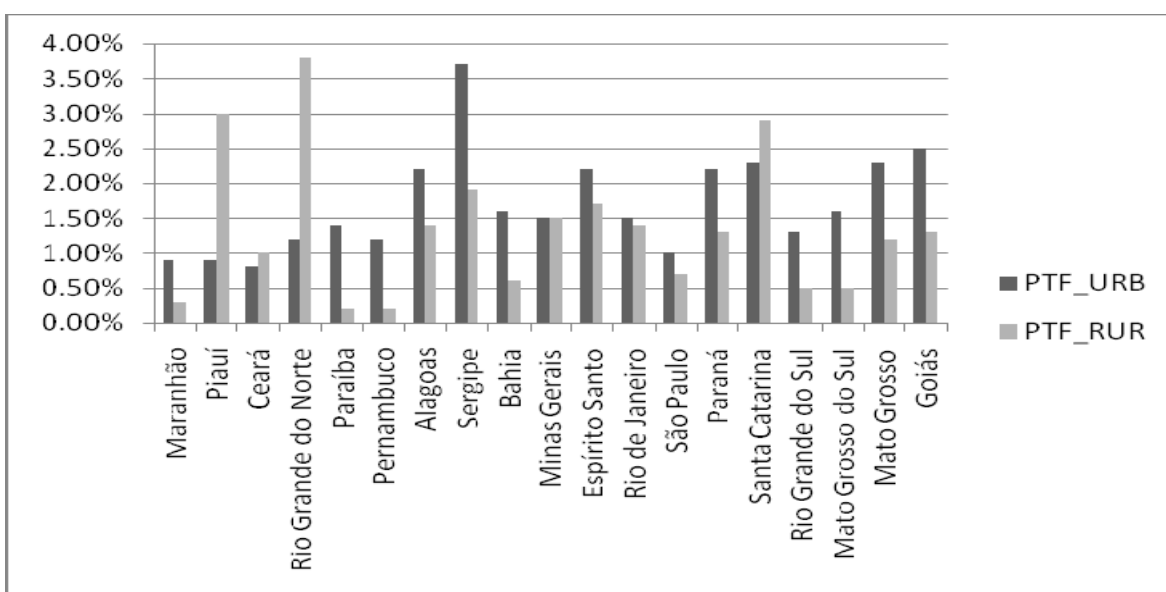


Figura 3.6: Variação da Produtividade Total dos Fatores dos estados brasileiros, 1981-2007.

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1981 e 2007, IBGE.

É importante ressaltar que no índice de Malmquist não há a comparação intertemporal da mesma unidade, apenas a comparação com outras unidades. Ou seja, mesmo que haja ganhos de eficiência técnica alto para o rural, este aumento é em relação às mesmas unidades em outro ano. Sendo assim, se não houver ganhos tecnológicos altos no rural, a PTF não será tão alta, que foi o que se encontrou no presente ensaio. Ademais, a melhoria da eficiência técnica em algumas unidades não significa, necessariamente, que tais unidades estão mais próximas da nova fronteira do que as que não obtiveram ganhos de eficiência. Pode ser que as unidades que não tiveram ganhos de eficiência já estavam próximas da fronteira, ou mesmo nela, no início do período.

Nota-se pelas figuras acima que, em média, os ganhos em produtividade total dos estados brasileiros se devem muito mais às variações técnicas que aos outros componentes. Em outras palavras, a inovação tecnológica não vem sendo predominantemente a razão dos ganhos de produtividade dos estados brasileiros e sim, a difusão da tecnologia já existente.

3.5. Conclusão

Nesta dissertação, partiu-se da hipótese de que o aumento das ocupações não agrícolas e o aumento dos anos de estudos implicariam uma evolução positiva da Eficiência Técnica e da Produtividade Total dos Fatores do meio rural, aproximando-o do meio urbano. Essa hipótese foi investigada e confirmada.

Na comparação de 2007 em relação a 2000, a componente VTC da PTF foi a principal responsável pelo valor positivo. Mais especificamente, o que se destacou foi o deslocamento da fronteira de produção, onde ambos os espaços, rural e urbano, expandiram a sua eficiência. Não obstante, no meio rural a expansão foi mais expressiva.

Não obstante, foi a variação na eficiência técnica (VET) que se destacou no meio rural e no meio urbano. Isso demonstra que houve difusão da tecnologia já existente. A partir dos resultados positivos para o rural, pode-se inferir que o aumento da eficiência do trabalho no rural esteja correlacionado, em parte, com o arrefecimento do êxodo rural. A escolha de viver no meio rural ou no meio urbano envolve uma análise de custo-benefício baseada principalmente no rendimento do trabalho. Na decisão de migrar para a cidade, o diferencial de salário se sobressai perante outras variáveis, como ajuste ao novo estilo de vida. Neste sentido percebe-se que o rural tem se mostrado mais atrativo comparativamente às duas últimas décadas. Ou seja, o trabalhador que reside no rural

consegue, em média, a partir das horas trabalhadas e anos de estudo extrair maior rendimento do trabalho do que antes.

Não obstante, verifica-se que o diferencial de eficiência do rural e urbano é maior no Norte e Nordeste brasileiro. Tal fato reforça a teoria de que o diferencial de renda entre campo e cidade é um indicador de desenvolvimento social. Sendo assim, mais uma vez essas regiões são ponto de reflexão, devendo-se analisar a composição de insumos e produto das demais regiões (benchmarks) a fim de nortear políticas para melhorar este quadro.

5. Considerações finais

Nos últimos anos tem-se observado a emergência de um espaço rural multifuncional com a introdução de uma maior diversificação econômica. Verifica-se principalmente a consolidação das ocupações não agrícolas, diminuindo a diferença entre a renda rural e a urbana. A modernização da agricultura e o surgimento de novas ocupações mostram que nem sempre rural é sinônimo de atraso, ou ineficiência, dada a inserção de alta tecnologia nos processos produtivos no meio rural e surgimento de novas ocupações. As mudanças em curso têm levado vários autores, como os do Projeto RURBANO a discutir sobre o ‘Novo Rural Brasileiro’, que se mostra mais eficiente e próximo do urbano.

A presente pesquisa mostrou a evolução das características quanto à ocupação dos indivíduos residentes no rural e no urbano, no período 1981-2007, propiciando subsídios para a construção de políticas públicas mais eficazes, capazes de atender às especificidades de cada setor censitário. Ademais, primou pela busca de estatísticas mais consistentes, optando por incorporar o plano amostral, a determinação do limiar de educação, a correção do viés de seletividade pelo método de Heckman, análise de eficiência pelo método DEA e Malmquist-DEA, nos dois artigos apresentados.

Nos dois ensaios apresentados, partiu-se da hipótese de que o aumento das ocupações não agrícolas e o aumento dos anos de estudos no campo implicariam em maior similaridade dos Determinantes da renda no meio rural e numa evolução positiva da Eficiência Técnica e da Produtividade Total dos Fatores no rural, respectivamente.

Pode-se considerar que tais hipóteses foram confirmadas. Por exemplo, no princípio do período analisado um ano a mais de estudo impactava mais na renda do trabalhador urbano que na do rural. O mesmo vale para a posição na ocupação, onde o fato de ser empregador representava um diferencial positivo na renda do trabalhador rural bem mais expressivo

que no meio urbano. Quanto à eficiência, ao longo dos anos os estados rurais foram aumentando sua produtividade, sendo a difusão da tecnologia a responsável por essa melhora. As unidades rurais ainda estão aquém das unidades urbanas, mais essa diferença se arrefeceu.

Analisando conjuntamente os dois ensaios, pode-se inferir que a estabilidade da população rural revelada no primeiro ensaio pode ser explicada pela melhora da eficiência do trabalho. Ou seja, o trabalhador não precisa migrar para o meio urbano para obter o retorno do seu investimento em educação. Adicionalmente, apresentou-se no primeiro ensaio que o retorno na educação para cada ano de estudo no rural igualou-se ao do urbano. Outros fatores positivos foram o aumento do rendimento familiar rural, da proporção de pessoas ocupadas no meio rural. Isso tudo foi possível sem aumento da desigualdade de renda no rural, pelo contrário, houve queda.

Diante do exposto, percebe-se que apesar do rural apresentar um rendimento médio inferior ao urbano, verificou-se que em relação a outros aspectos houve aproximação destes dois espaços. Antes bem distintos, estes espaços ficaram mais iguais quanto às características de ocupação. Aumentou a similaridade da quantidade de horas trabalhadas, do limiar da educação e da queda do impacto da posição na ocupação no retorno da renda do trabalhador residente no rural.

As análises supramencionadas reforçam o pensamento de que na virada do século o rural ganhou mais dinamização no que se refere as ocupações que se reverteram em maior eficiência, embora ainda não capaz de suprir o diferencial de renda existente. Mais também há de considerar que o rural apresenta ocupações ainda menos qualificadas onde se pode inferir que isso impacta negativamente na média do salário recebido pelos moradores rurais.

Ou seja, no balanço geral o rural se apresenta mais desenvolvido do que há três décadas, colocando o Brasil na posição de destaque neste cenário já que o diferencial de renda entre campo e cidade é um indicador de desenvolvimento social.

6. Referências Bibliográficas

Aigner, D. J.; Chu, S. F. Estimating the industry production function. *American Economic Review*, v. 58, n. 4, p. 826-839, September 1968.

AIGNER , D. , LOVELL, C.A .K. and SCHMIDT. Formulation and Estimation of Estochastic Frontier Production Function Models , *Journal of Econometrics* , 1977.

ALVES, M. O.; VALENTE Jr., A. S. Perfil das atividades não-agrícolas no rural do pólo de desenvolvimento de agronegócios baixo jaguaribe: resultados preliminares. Juiz de Fora: SOBER, 2003.

AMADEO, E., CAMARGO, G., BARROS R. P., MENDONÇA, R. A natureza e o funcionamento do mercado de trabalho brasileiro desde de 1980. Rio de Janeiro: IPEA, 1994.

ASHENFELTER, O., HARMON, C., OOSTERBEEK, H. A review of estimates of the schooling/earnings relationship with tests of publication bias, *Labour Economics*, 1999.

BAPTISTA, A.J.M.S. Progresso tecnológico, mudanças na eficiência e produtividade na pesca artesanal em Cabo Verde, na década de 90. Viçosa: UFV, Dissertação de mestrado, 2002.

BAPTISTA, A.J.M.S. GOMES, A.P.; BAPTISTA, A.J.M.dos S., Análise Envoltória de Dados: Conceitos e Modelos Básicos. In Métodos Quantitativos em Economia. Cap. 5, Viçosa, Ed. 1. Editora Universidade Federal de Viçosa, 2004

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

BARROS, R. P. et al. Acesso ao trabalho e produtividade no Brasil: Implicações para crescimento, pobreza e desigualdade. Rio de Janeiro: IPEA, 2004.

BARROS, R. P. e MENDONÇA, R. S. P. Os determinantes da desigualdade no Brasil. Rio de Janeiro, IPEA, 1995.

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

BARROS, R. P., MENDONÇA, R., *Welfare, inequality, poverty, social indicators and social programmes in Brazil in the 180s*. Rio de Janeiro: IPEA, 1993.

BECKER, Gary S. *Human Capital*, 1st ed. New York: Columbia University Press for the National Bureau of Economic Research, 1964.

BECKER, G.S., Investment in Human Capital: A Theoretical Analysis, in *Journal of Political economy*, Vol. 70, Suplemento, Parte 2, 1962

BERNI, H. A. A. Evolução dos Determinantes da Desigualdade de Renda Salarial no Nordeste. 2007. 48f. Dissertação (Mestrado em Economia) –CAEN, Universidade Federal do Ceará, Fortaleza, 2007.

BLUME, R. Território e ruralidade: A desmistificação do fim do rural. Porto Alegre: FCE, 2004.

BRITO, Fausto. O deslocamento da população brasileira para as metrópoles. *Estudos Avançados*, 57, USP, 2006.

CARVALHO, A. P. Decomposição do Diferencial de Salários no Brasil em 2003: Uma Aplicação dos Procedimentos de Oaxaca e Heckman em Pesquisas Amostrais Complexas, Dissertação de Mestrado, Escola Nacional de Ciências Estatísticas, Rio de Janeiro, 2006.

CAVES, D.; CHRISTENSEN, L. ; DIEWERT, E. The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica*, 1982.

CHAPUIS, R. La géographie agraire et la géographie rurale. In. BAILLY, A.S. (coord.). *Les concepts de la géographie humaine*. 2.ed. Paris, Masson, 1991.

CHARNES A., COOPER, W.W.; RHODES, E.. Measuring The Efficiency of Decision Making Units . *European Journal of Operational Research*, 2, 1978.

CHROMY, J. R. & ABEYASEKARA, S. (2005). Statistical analysis of survey data. In of Economic, D. & of United Nations, S. A., editors, *Household Sample Surveys in Developing and Transition Countries*, chapter 19, pages 389–417. New York: United Nations.

COELHO, A. M.; CORSEUIL, C.H. Diferenciais salariais no Brasil: um breve panorama. In: CORSEUIL, G. H. (Ed.). *Estrutura salarial: aspectos conceituais e novos resultados para o Brasil*. Rio de Janeiro: IPEA, 2002.

COLE, W.E.; SANDERS, R.D. Internal migration and urban employment in the Third World. *American Economic Review*, n.75, p.481-494, 1985. FALLON, P.R. Education and the duration of job search and unemployment

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

COURA, R. M. Produtividade total dos fatores (PTF) na agricultura paulista: 1985-2001. 2004. 61 f. Dissertação (Mestrado) – Universidade Federal de Viçosa, Viçosa, 2004.

DEBREU, G. (1951). The Coefficient of Resource Utilisation. *Econometrica*, 19, 273-292.
Ehrlich. Pesquisa Operacional: curso introdutório. Editora Atlas, 1991.

DEL GROSSI, M. & GRAZIANO DA SILVA, J. Mudanças recentes no mercado de trabalho rural. *Parcerias Estratégias*, Brasília, CGEE, n. 22, junho de 2006.

DIEESE, Departamento Intersindical de Estatística e Estudos Socioeconômicos . Salário Mínimo. 2007 Disponível em: <http://www.dieese.org.br/rel/rac/salminmai09.xml>

DINIZ, B. C. ; SILVEIRA, F. G. O mercado de trabalho rural: evolução recente, composição da renda e a dimensão regional. In: BUAINAIN, A. M.; DEDECCA, C. ; MIRANDA, C. (Org.). Emprego e trabalho na agricultura brasileira. Brasília: IICA, 2009.

EFRON, B. Bootstrap methods: another look at the Jackknife. *Annals of Statistics*, v.7, n.1,1979.

FÄRE, R.; Grosskopf, S. & ROOS, P. (1998). Malmquist productivity indexes: a survey of theory and practice. In: *Index Numbers: Essays in Honour of Sten Malmquist* [editado por R. Färe, S. Grosskopf e R. Russell], Kluwer Academic Publishers, 127-190.

FISHLOW, A. "Brazilian Size Distribution of Income," *American Economic Review*, 1972. Versão em português: TOLIPAN e TINELLI (1975).

FREEMAN, R. B. Demand for education. In Ashenfelter, O. & Layard, R., editors, *Handbook of Labor Economics*, volume 1 of Handbook of Labor Economics, Elsevier,1987.

FRIED, H.; LOVELL, C. A. K.; SCHMIDT, S. Measurement of productive efficiency: techniques and applications. New York: Oxford University, 1993.

FRIEDMANN, J. “A Spatial Framework for Urban Policy: New Directions, New Challenges”, *OECD International Conference: What policies for globalising cities? Rethinking the urban policy agenda*, Madrid, 2007.

GANDRA, Rodrigo M. “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, 2002.

GOLANY, B.; ROLL Y. "An Application Procedure For DEA". *Omega (Uk)*, v. 17, n. 3, 1989.

GOMES, A. P., BAPTISTA, A. J. MEDINA, S. Análise Envoltória de Dados: Conceitos e modelos básicos. In: *Métodos Quantitativos em Economia*. Ed UFV, 2004.

GRAZIANO DA SILVA, J. O novo rural brasileiro. Campinas, Unicamp/IE,1999 (Coleção Pesquisas, n.1).

GRAZIANO DA SILVA, J; DEL GROSSI, E. O novo rural brasileiro: uma atualização para 1992-98. IE/Unicamp. 2001 (texto para discussão).

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

GROSSKOPF, S. Efficiency and productivity. The measurement of productive efficiency. New York: Oxford University, 1993.

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

HARRIS, J., TODARO, M. MIGRATION, unemployment and development: two-sector analysis, *American Economic Review*, 1970.

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

HOFFMANN, R. and KASSOUF, A.L. “Deriving conditional and unconditional marginal effects in log earnings equations estimated by Heckman’s procedure”. *Applied Economics*, Londres, v. 37, n. 11, 2005.

HOFFMANN, R.; LEONE E. T. Participação da mulher no mercado de trabalho e desigualdade da renda domiciliar per capita no Brasil: 1981-2002. *Nova Economia*. Belo Horizonte. maio-agosto de 2004

HOFFMANN, R. Mensuração da desigualdade e da pobreza no Brasil. In: HENRIQUES, R. (eds) *Desigualdade e pobreza no Brasil*. Rio de Janeiro: IPEA, 2000.

HOFFMANN, Rodolfo. Distribuição da renda e desenvolvimento econômico do Brasil. Rio de Janeiro: Expressão e Cultura, 1973.

HOFFMANN, Rodolfo; DUARTE, João Carlos. A Distribuição de Renda no Brasil. *Revista de Administração de Empresas*, v.12, n.2, p. 46-66, 1972.

HOFFMANN, R; SIMÃO, R. C. S. 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, maio/ago. 2005

KAGEYAMA, A. Os rurais e os agrícola de São Paulo no Censo de 2003. IE/UNICAMP. 2003 (texto para discussão)

KISH, L. (1965). *Survey Sampling*. New York: Wiley.

KNEIPP, S.M.; Yarandi H.N. Complex sampling designs and statistical issues in secondary analysis. *West J Nurs Res* 2002.

LANGONI, C. G. Distribuição de renda e desenvolvimento Rio de Janeiro: Expressão e Cultura, 1973.

LAURENTI, A. C ; GROSSI, Mauro Eduardo Del . A evolução das pessoas ocupadas em atividades agrícolas e não-agrícolas nas áreas rurais do Brasil. In: XXXVII Congresso Brasileiro de Economia e Sociologia Rural, 1999, Foz do Iguaçu. O agronegócio do Mecosul e a sua inserção na economia mundial. Brasília : Sociedade Brasileira de Economia e Sociologia Rural-SOBER, 1999. p. 269-269.

LEWIS, W. A. **Economic development with unlimited supplies of labour**. Manchester school of economics and social studies, p. 139-192, may. 1954.

LEWIS, A. “The Evolution of the International Economic Order”. Princeton University Press, 1977.

LIMA, Chaves Ricardo. Economics of rural labor migration: a survey in the literature. In: LIMA, João Policarpo Rodrigues. Economia e Região. Nordeste: ensaios. Recife: PIMES, Editora Universitária da UFPE, 1999.

LUMLEY, T. (2004). Analysis of complex survey samples. *Journal of Statistical Software*, 9(8):1–19. Disponível em: <http://ideas.repec.org/a/jss/jstsof/09i08.html>.

MALMQUIST, S. “Index Numbers and Indifference Curves”, *Trabajos de Estadística* 4(1), 1953.

MALTHUS, T.R. Ensaio sobre a população. Coleção Os Economistas. São Paulo: Nova Cultural, 1996.

MANSO, C. A.; BARRETO, F. A.; TEBALDI, E. O desequilíbrio regional brasileiro: novas perspectivas a partir das fontes de crescimento “pró-pobre”. *Revista Econômica do Nordeste*, v.37, nº37, 2006.

MARINHO, E.; ATALIBA, F., & LIMA, F. (2001). Produtividade, variação tecnológica e variação da eficiência das regiões e estados brasileiros. ANPEC. XXIX Encontro Nacional de Economia, Salvador.

MARINHO, E.; BENEGAS, M. (2002). Avaliação Inter/Intra-regional de produtividade total dos estados e regiões brasileiras: Uma abordagem não-paramétrica utilizando o conceito de meta-fronteira de produção. Latin American Meeting of the Econometric Society. 24 a 27 de julho de 2002. Anais em CD.

MATOS, R.; SATHLER, D.; UMBELINO G. Urbano influente e rural não agrícola em Minas Gerais. Anais do XI Seminário sobre a Economia Mineira, 2004.

MARSHALL, A. *Principles of Economics*. 8th ed. London: Macmillan, 1920.

MARTINS, N.; FIGEIREDO, C.; AZEVEDO, F.; RIBEIRO, F.; PROENÇA, M.; ABREU, M. J. PROJECTO “CIDADES INTELIGENTES”. Apresentação do Projecto INTELLIGENT CITIES no Seminário “Cidades e Regiões de Inovação: O Papel dos Parques de C&T” FCT-UNL - 14 de Dezembro de 2005.

MARX, K. *O Capital* (Rio de Janeiro: Bertrand) Vol. 1, Livro 1, 1994.

MENEZES-FILHO, N.A.; MENDES, M.; DE ALMEIDA, S.E. "O diferencial de salários formal-informal no Brasil: segmentação ou viés de seleção?", *Revista Brasileira de Economia*, v. 58, n. 2, 2004.

MENEZES, W. F., et al. Diferenciações regionais de rendimentos do trabalho: uma análise das regiões metropolitanas de São Paulo e de Salvador. In: ENCONTRO NACIONAL DE ECONOMIA, 30, 2002, Nova Friburgo. Anais. Nova Friburgo: ANPEC, 2002. Disponível em CD-ROM.

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)*.

MINCER, J. Investment in human capital and personal income distribution. *The journal of Political Economy*, v. 46, n.4. August, 1958.

MORLEY, S. “The Income Distribution Problem in Latin America and the Caribbean”, ECLAC, Santiago do Chile, 2001.

MOURA, R. L. Testando as Hipóteses do Modelo de Mincer para o Brasil. *Revista Brasileira de Economia*. v. 62 n. 4 / p. 407–449. Out-Dez 2008.

MURPHY, K. M. & Welch, F. The structure of wages. *Quarterly Journal of Economics*, 1992.

NEDER, H. D. The effects of non-agricultural activities on income distribution in rural Brazil. *Revista de Economia e Sociologia Rural*. Brasília, v. 41, n.2, abr./jun. 2003.

NEY, M. G.; HOFFMANN, R. Desigualdade de renda na agricultura: o efeito da posse da terra. *Economia*, v. 4, n. 1, p. 113-152. ANPEC, Niterói, jan./jun. 2003.

NIEDERAUER, C. A. P. *Ethos*: Um modelo para medir a produtividade de NOGUEIRA, N. V., MUSTEFAGA, P. S. Informativo técnico da revista Gleba, Edição ior. Florianópolis, 1999. 132p. Dissertação (Mestrado em Engenharia de Produção) Programa de Pós-dos 144 Engenharia de Produção, Univ aduação em Engenharia de Produção, na Universidade Federal de Santa Catarina - UFSC.

NISHIMIZU, M.; PAGE, J. M. (1982) “Total factor productivity growth, technical progress and technical efficiency change: Dimensions of productivity change in Yugoslavia, 1965.

PESQUISA NACIONAL POR AMOSTRA DE DOMICÍLIOS 1981. Brasil. Rio de Janeiro: IBGE, 1981 a 1990.

PESQUISA NACIONAL POR AMOSTRA DE DOMICÍLIOS 1981. Brasil. Rio de Janeiro: IBGE, 1992 a 1999.

PESQUISA NACIONAL POR AMOSTRA DE DOMICÍLIOS 2006. Brasil. Rio de Janeiro: IBGE, 2001 a 2007.

PNUD - Programa das Nações Unidas para o Desenvolvimento. Atlas do desenvolvimento humano no Brasil. Brasília: PNUD, 1999.

PSACHARAPOULOS, G. Returns to education: an updated international comparison, *Comparative Education*, 1981.

PSACHAROPOULOS, G.; PATRINOS, H. A. (2004). Returns to investment in education: a further update. *Education Economics*, 12(2):111–134. Disponível em: <http://ideas.repec.org/a/taf/edecon/v12y2004i2p111-134.html>.

PINDYCK, R. S.; RUBINFELD, D. L. *Microeconomia*. São Paulo. Makron Books, 1994.
REDFIELD, R. The Folk Society. *American Journal of Sociology*, 1947.

REIS, D. S. O rural e o urbano no Brasil. *Caderno de Geografia (PUCMG)*, v. 15, p. 77-92, 2005.

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

ROCHA, L.E.V; BASTOS, P.M.A, SANTOS, G. C. Evolução da distribuição da renda das pessoas ocupadas residentes no meio rural do Estado de Minas Gerais , 1981 a 2003. In: ANAIS DO XLIII CONGRESSO BRASILEIRO DE ECONOMIA E SOCIOLOGIA RURAL, 2005. CD ROM.

ROCHA, S. Estimação de linhas de indigência e de pobreza: opções metodológicas no Brasil. In: HENRIQUES, R. (org). *Desigualdade e pobreza no Brasil*. Rio de Janeiro: IPEA, 2000.

ROCHA, L. E. V.; BASTOS, P. M. A.; SANTOS, G. C. Evolução da distribuição da renda e da pobreza das famílias ocupadas e residentes no meio rural do Estado de Minas Gerais, de 1981 a 2003. In: ANAIS DO XLIII CONGRESSO BRASILEIRO DE ECONOMIA E SOCIOLOGIA RURAL, 2005. CD ROM.

SARACENO, E. "Recent trends in rural development and their conceptualisation". *Journal of Rural Studies*, v.10, n.4, 1994.

SCHNEIDER, S. Rurbanização e pluriatividade: o mercado de trabalho não-agrícola e a pluriatividade das famílias em áreas rurais (um estudo de caso). In: CARVALHO, F. M. A.; GOMES, M. F. M.; LIRIO, V.S.(orgs). *Desigualdades sociais: pobreza, desemprego e questão agrária*. Viçosa: UFV, 2003.

SCHULTZ, T. W. Investment in Human Capital. *The American Economic Review*, 1961.

SEPLAG, Secretaria do Planejamento e Gestão. 2009. Available in: <http://www.seplag.ce.gov.br/>

SILVA, Pedro Luis do Nascimento, PESSOA, Djalma Galvão Carneiro and LILA, Maurício Franca. Análise estatística de dados da PNAD: incorporando a estrutura do plano amostral. *Ciênc. saúde coletiva*, 2002, vol.7, no.4, ISSN 1413-8123.

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

SMITH, Adam (1999 [1776]) "Da Divisão do Trabalho" in *A Riqueza das Nações*, Vol. 1, Lisboa, Edições Calouste Gulbenkian, 1999.

SOROKIN, P. A.; ZIMMERMAN, C.C. e GALPIN, C. J. Diferenças Fundamentais entre o mundo rural e urbano. (1929) In: MARTINS, J. S. *Introdução Crítica à Sociologia Rural*. São Paulo: Hucitec, 1986.

TODARO, M.P. A model of labor migration and urban unemployment in less developed countries. *American Economic Review*, n.59, p.138-148, 1969.

THUROW, L. *Generating Inequality*, New York, Basic Books, 1975.

WATSON, Donald S. & HOLMAN, Mary A. Microeconomia. São Paulo, Edições Saraiva, 1979. Bradley, S.P.; Hax, A.C.; Magnanti, T.L. Applied Mathematical Programming. Menlo Park, California, Addison-Wesley, 1977.

APÊNDICE 1

Resultados complementares

Tabela A.1. Análise descritiva das principais variáveis utilizadas na pesquisa.

Variáveis	TODOS		URBANO		RURAL	
	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.
Economicamente ativos	0.71	0.002	0.68	0.002	0.80	0.005
Renda do trabalho principal	561.15	13.447	640.53	14.152	229.39	5.849
Log da razão de renda por hora trabalhada	1.7165	0.005	1.7549	0.004	1.53	0.006
Rendimento Familiar	1909.514	43.7243	2162.40	46.983	851.67	20.573
Número de horas trabalho principal (semana)	27.34	0.09	26.66	0.084	30.20	0.238
Número de Componentes na Fam.	4.19	0.015	4.07	0.012	4.70	0.036
Gênero (Masculino)	0.4841	0.001	0.4751	0.0012	0.52	0.00
Idade (em dezenas de anos)	3.53	0.006	3.53	0.006	3.54	0.011
idade^2 (em dezenas de anos)	14.60	0.047	14.54	0.048	14.83	0.086
Cor preta ou indígena	0.05	0.002	0.05	0.002	0.05	0.004
Cor amarela ou pardas	0.39	0.005	0.36	0.006	0.49	0.012
Cor branca	0.56	0.005	0.58	0.007	0.46	0.013
Educação em anos de estudo	6.65	0.071	7.28	0.056	4.05	0.063
Empregados	0.61	0.006	0.68	0.003	0.35	0.008
Conta-própria	0.23	0.002	0.22	0.002	0.28	0.004
Empregador	0.04	0.001	0.05	0.001	0.02	0.001
N_remunerado, Const Prop. Uso, Próprio consumo	0.12	0.005	0.05	0.002	0.34	0.007
Agrícola	0.2384	0.010	0.0860	0.005	0.75	0.009
Não-agrícola	0.76	0.010	0.91	0.005	0.25	0.009

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1995

Tabela A.2. Análise descritiva das principais variáveis utilizadas na pesquisa.

Variáveis	TODOS		URBANO		RURAL	
	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.
Economicamente ativos	0.73	0.002	0.72	0.002	0.79	0.005
Renda do trabalho principal	566.07	13.800	638.17	14.544	239.45	8.536
Log da razão de renda por hora trabalhada	1.2416	0.004	1.2671	0.003	1.11	0.004
Rendimento Familiar	1949.86	46.018	2188.48	48.792	865.20	28.045
Número de horas trabalho principal (semana)	26.76	0.10	26.12	0.092	29.63	0.289
Número de Componentes na Fam.	4.19	0.016	4.08	0.014	4.71	0.033
Gênero (Masculino)	0.4823	0.001	0.4742	0.0011	0.52	0.00
Idade (em dezenas de anos)	3.53	0.006	3.52	0.006	3.54	0.013
idade^2 (em dezenas de anos)	14.58	0.050	14.53	0.049	14.80	0.103
Cor preta ou indígena	0.06	0.002	0.06	0.002	0.07	0.003
Cor amarela ou pardas	0.38	0.006	0.36	0.006	0.47	0.013
Cor branca	0.56	0.007	0.58	0.007	0.46	0.013
Educação em anos de estudo	6.89	0.075	7.48	0.063	4.25	0.074
Empregados	0.63	0.007	0.70	0.003	0.38	0.012
Conta-própria	0.23	0.003	0.21	0.002	0.27	0.005
Empregador	0.04	0.001	0.04	0.001	0.02	0.001
N_remunerado, Const Prop. Uso, Próprio consumo	0.11	0.005	0.05	0.002	0.33	0.009
Agrícola	0.216	0.010	0.0789	0.004	0.73	0.011
Não-agrícola	0.78	0.010	0.92	0.004	0.27	0.011

Fonte: Elaboração dos autores a partir dos microdados da PNAD 1996.

Tabela A.3. Análise descritiva das principais variáveis utilizadas na pesquisa.

Variáveis	TODOS		URBANO		RURAL	
	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.
Economicamente ativos	0.69	0.002	0.67	0.002	0.78	0.006
Renda do trabalho principal	556.15	11.968	630.68	12.951	232.74	8.392
Log da razão de renda por hora trabalhada	1.2371	0.003	1.2660	0.003	1.09	0.005
Rendimento Familiar	1895.84	40.048	2143.79	42.124	823.72	24.781
Número de horas trabalho principal (semana)	26.56	0.09	25.92	0.086	29.36	0.254
Número de Componentes na Fam.	4.11	0.015	4.00	0.012	4.56	0.035
Gênero (Masculino)	0.4843	0.001	0.4759	0.0009	0.52	0.00
Idade (em dezenas de anos)	3.54	0.006	3.54	0.006	3.55	0.010
idade^2 (em dezenas de anos)	14.68	0.045	14.63	0.050	14.90	0.075
Cor preta ou indígena	0.06	0.001	0.06	0.002	0.05	0.003
Cor amarela ou pardas	0.39	0.007	0.36	0.007	0.50	0.013
Cor branca	0.56	0.007	0.58	0.007	0.45	0.014
Educação em anos de estudo	6.92	0.062	7.54	0.049	4.24	0.065
Empregados	0.62	0.005	0.69	0.003	0.37	0.009
Conta-própria	0.23	0.002	0.22	0.002	0.28	0.006
Empregador	0.04	0.001	0.05	0.001	0.02	0.001
N_remunerado, Const Prop. Uso, Próprio consumo	0.11	0.004	0.05	0.001	0.33	0.006
Agrícola	0.223	0.009	0.0765	0.003	0.73	0.012
Não-agrícola	0.78	0.009	0.92	0.003	0.27	0.012

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1997

Tabela A.4. Análise descritiva das principais variáveis utilizadas na pesquisa.

Variáveis	TODOS		URBANO		RURAL	
	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.
Economicamente ativos	0.69	0.003	0.67	0.002	0.77	0.005
Renda do trabalho principal	543.77	12.002	615.61	12.719	231.62	7.121
Log da razão de renda por hora trabalhada	1.7286	0.005	1.7662	0.004	1.54	0.006
Rendimento Familiar	1891.64	40.818	2133.78	42.742	841.42	23.860
Número de horas trabalho principal (semana)	26.17	0.12	25.62	0.115	28.54	0.265
Número de Componentes na Fam.	4.05	0.016	3.94	0.014	4.52	0.028
Gênero (Masculino)	0.4846	0.001	0.4760	0.0010	0.52	0.00
Idade (em dezenas de anos)	3.55	0.006	3.55	0.007	3.57	0.011
idade^2 (em dezenas de anos)	14.79	0.044	14.74	0.059	15.03	0.090
Cor preta ou indígena	0.06	0.002	0.06	0.002	0.06	0.003
Cor amarela ou pardas	0.38	0.007	0.36	0.007	0.50	0.012
Cor branca	0.55	0.007	0.58	0.007	0.45	0.013
Educação em anos de estudo	7.11	0.049	7.73	0.050	4.41	0.077
Empregados	0.62	0.006	0.68	0.003	0.39	0.012
Conta-própria	0.23	0.003	0.22	0.002	0.28	0.006
Empregador	0.04	0.001	0.05	0.001	0.02	0.001
N_remunerado, Const Prop. Uso, Próprio consumo	0.11	0.004	0.05	0.001	0.31	0.007
Agrícola	0.213	0.009	0.0731	0.004	0.70	0.012
Não-agrícola	0.79	0.009	0.93	0.004	0.30	0.012

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1998.

Tabela A.5. Análise descritiva das principais variáveis utilizadas na pesquisa.

Variáveis	TODOS		URBANO		RURAL	
	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.
Economicamente ativos	0.70	0.003	0.68	0.002	0.79	0.005
Renda do trabalho principal	507.09	10.539	568.01	11.805	230.70	7.953
Log da razão de renda por hora trabalhada	1.7162	0.004	1.7503	0.004	1.54	0.006
Rendimento Familiar	1779.77	35.110	1984.02	39.543	850.56	29.837
Número de horas trabalho principal (semana)	25.91	0.10	25.36	0.097	28.44	0.251
Número de Componentes na Fam.	4.01	0.012	3.89	0.013	4.51	0.032
Gênero (Masculino)	0.4832	0.001	0.4751	0.0011	0.52	0.002
Idade (em dezenas de anos)	3.55	0.006	3.55	0.006	3.56	0.011
idade^2 (em dezenas de anos)	14.79	0.046	14.75	0.048	14.96	0.091
Cor preta ou indígena	0.06	0.002	0.06	0.002	0.06	0.004
Cor amarela ou pardas	0.40	0.005	0.38	0.005	0.50	0.012
Cor branca	0.54	0.006	0.56	0.006	0.44	0.013
Educação em anos de estudo	7.29	0.063	7.89	0.059	4.59	0.070
Empregados	0.61	0.005	0.68	0.003	0.38	0.010
Conta-própria	0.24	0.003	0.22	0.002	0.27	0.007
Empregador	0.04	0.001	0.05	0.001	0.02	0.001
N_remunerado, Const Prop. Uso, Próprio consumo	0.11	0.004	0.05	0.002	0.33	0.007
Agrícola	0.212	0.009	0.0746	0.004	0.71	0.009
Não-agrícola	0.79	0.009	0.93	0.004	0.29	0.009

Fonte: Elaborado pela autora a partir dos microdados da PNAD-1999

Tabela A.6. Análise descritiva das principais variáveis utilizadas na pesquisa.

Variáveis	TODOS		URBANO		RURAL	
	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.
Economicamente ativos	0.70	0.002	0.68	0.001	0.78	0.005
Renda do trabalho principal	503.56	5.713	555.73	6.622	205.62	5.853
Log da razão de renda por hora trabalhada	1.7154	0.002	1.7444	0.002	1.52	0.006
Rendimento Familiar	1744.54	19.553	1913.52	22.656	780.05	20.726
Número de horas trabalho principal (semana)	26.11	0.08	25.66	0.076	28.70	0.263
Número de Componentes na Fam.	3.89	0.009	3.80	0.008	4.37	0.031
Gênero (Masculino)	0.4823	0.001	0.4748	0.0009	0.53	0.002
Idade (em dezenas de anos)	3.56	0.004	3.56	0.004	3.60	0.011
idade^2 (em dezenas de anos)	14.84	0.033	14.76	0.035	15.30	0.091
Cor preta ou indígena	0.06	0.001	0.06	0.001	0.05	0.003
Cor amarela ou pardas	0.39	0.003	0.37	0.003	0.51	0.010
Cor branca	0.54	0.003	0.56	0.003	0.44	0.010
Educação em anos de estudo	7.56	0.028	8.08	0.028	4.57	0.056
Empregados	0.64	0.003	0.70	0.002	0.36	0.010
Conta-própria	0.22	0.002	0.21	0.002	0.28	0.007
Empregador	0.04	0.001	0.05	0.001	0.03	0.002
N_remunerado, Const Prop. Uso, Próprio consumo	0.10	0.003	0.05	0.001	0.33	0.008
Agrícola	0.190	0.005	0.0683	0.002	0.75	0.010
Não-agrícola	0.81	0.005	0.93	0.002	0.25	0.010

Fonte: Elaborado pela autora a partir dos microdados da PNAD-2001

Tabela A.7. Análise descritiva das principais variáveis utilizadas na pesquisa.

Variáveis	TODOS		URBANO		RURAL	
	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.
Economicamente ativos	0.71	0.002	0.69	0.001	0.78	0.005
Renda do trabalho principal	781.55	8.455	892.03	9.526	264.44	7.167
Log da razão de renda por hora trabalhada	1.2262	0.001	1.2464	0.001	1.09	0.005
Rendimento Familiar	1738.52	18.632	1905.60	21.409	777.33	15.293
Número de horas trabalho principal (semana)	26.21	0.08	25.80	0.075	28.54	0.254
Número de Componentes na Fam.	3.85	0.008	3.77	0.007	4.37	0.032
Gênero (Masculino)	0.4828	0.001	0.4750	0.0008	0.53	0.002
Idade (em dezenas de anos)	3.58	0.004	3.58	0.004	3.59	0.011
idade^2 (em dezenas de anos)	14.98	0.032	14.94	0.035	15.22	0.088
Cor preta ou indígena	0.06	0.001	0.06	0.001	0.05	0.003
Cor amarela ou pardas	0.40	0.003	0.38	0.003	0.51	0.011
Cor branca	0.54	0.003	0.56	0.003	0.44	0.010
Educação em anos de estudo	7.75	0.028	8.27	0.027	4.78	0.058
Empregados	0.64	0.003	0.69	0.002	0.36	0.010
Conta-própria	0.22	0.002	0.21	0.002	0.28	0.005
Empregador	0.04	0.001	0.05	0.001	0.02	0.001
N_remunerado, Const Prop. Uso, Próprio consumo	0.10	0.003	0.05	0.001	0.34	0.008
Agrícola	0.191	0.005	0.0718	0.002	0.75	0.010
Não-agrícola	0.81	0.005	0.93	0.002	0.25	0.010

Fonte: Elaborado pela autora a partir dos microdados da PNAD-2002

Tabela A.8. Análise descritiva das principais variáveis utilizadas na pesquisa.

Variáveis	TODOS		URBANO		RURAL	
	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.
Economicamente ativos	0.71	0.002	0.69	0.001	0.78	0.005
Renda do trabalho principal	723.21	7.773	818.72	8.887	272.95	8.620
Log da razão de renda por hora trabalhada	1.2170	0.001	1.2358	0.001	1.09	0.004
Rendimento Familiar	1599.87	16.706	1740.03	19.140	785.68	18.108
Número de horas trabalho principal (semana)	25.91	0.08	25.52	0.071	28.21	0.247
Número de Componentes na Fam.	3.81	0.008	3.73	0.008	4.28	0.028
Gênero (Masculino)	0.4829	0.001	0.4757	0.0008	0.53	0.002
Idade (em dezenas de anos)	3.60	0.004	3.60	0.004	3.61	0.011
idade^2 (em dezenas de anos)	15.11	0.032	15.06	0.035	15.38	0.089
Cor preta ou indígena	0.06	0.001	0.07	0.001	0.05	0.003
Cor amarela ou pardas	0.40	0.003	0.38	0.003	0.52	0.010
Cor branca	0.53	0.003	0.55	0.003	0.43	0.010
Educação em anos de estudo	7.94	0.029	8.44	0.027	5.00	0.058
Empregados	0.64	0.003	0.70	0.002	0.36	0.010
Conta-própria	0.22	0.002	0.21	0.002	0.28	0.005
Empregador	0.04	0.001	0.05	0.001	0.02	0.001
N_remunerado, Const Prop. Uso, Próprio consumo	0.10	0.003	0.05	0.001	0.34	0.008
Agrícola	0.192	0.005	0.0741	0.002	0.75	0.010
Não-agrícola	0.81	0.005	0.93	0.002	0.25	0.010

Fonte: Elaborado pela autora a partir dos microdados da PNAD-2003

Tabela A.9. Análise descritiva das principais variáveis utilizadas na pesquisa.

Variáveis	TODOS		URBANO		RURAL	
	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.
Economicamente ativos	0.71	0.002	0.70	0.001	0.79	0.004
Renda do trabalho principal	721.78	7.298	821.96	8.355	287.64	8.163
Log da razão de renda por hora trabalhada	1.2201	0.001	1.2383	0.001	1.11	0.004
Rendimento Familiar	1617.02	16.339	1765.98	18.839	832.91	18.609
Número de horas trabalho principal (semana)	26.30	0.07	25.96	0.070	28.11	0.228
Número de Componentes na Fam.	3.76	0.009	3.67	0.008	4.25	0.028
Gênero (Masculino)	0.4821	0.001	0.4738	0.0008	0.53	0.002
Idade (em dezenas de anos)	3.62	0.004	3.62	0.004	3.60	0.012
idade^2 (em dezenas de anos)	15.24	0.034	15.22	0.034	15.30	0.096
Cor preta ou indígena	0.06	0.001	0.07	0.001	0.05	0.003
Cor amarela ou pardas	0.41	0.003	0.39	0.003	0.54	0.010
Cor branca	0.52	0.003	0.54	0.003	0.41	0.010
Educação em anos de estudo	8.05	0.030	8.59	0.027	5.15	0.057
Empregados	0.64	0.004	0.71	0.002	0.37	0.010
Conta-própria	0.22	0.002	0.21	0.002	0.28	0.005
Empregador	0.04	0.001	0.05	0.001	0.02	0.002
N_remunerado, Const Prop. Uso, Próprio consumo	0.09	0.003	0.04	0.001	0.33	0.007
Agrícola	0.196	0.005	0.0725	0.002	0.73	0.010
Não-agrícola	0.80	0.005	0.93	0.002	0.27	0.010

Fonte: Elaborado pela autora a partir dos microdados da PNAD-2004

Tabela A.10. Análise descritiva das principais variáveis utilizadas na pesquisa.

Variáveis	TODOS		URBANO		RURAL	
	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.
Economicamente ativos	0.72	0.002	0.71	0.001	0.79	0.004
Renda do trabalho principal	751.16	8.194	858.39	9.407	289.62	8.424
Log da razão de renda por hora trabalhada	1.2323	0.001	1.2515	0.001	1.11	0.004
Rendimento Familiar	1699.56	17.107	1860.95	19.918	856.28	19.037
Número de horas trabalho principal (semana)	26.24	0.07	25.95	0.072	27.75	0.227
Número de Componentes na Fam.	3.73	0.009	3.63	0.007	4.22	0.029
Gênero (Masculino)	0.4833	0.001	0.4753	0.0008	0.53	0.002
Idade (em dezenas de anos)	3.63	0.004	3.63	0.004	3.60	0.012
idade^2 (em dezenas de anos)	15.33	0.035	15.34	0.035	15.30	0.102
Cor preta ou indígena	0.07	0.001	0.07	0.001	0.05	0.003
Cor amarela ou pardas	0.42	0.003	0.40	0.003	0.55	0.010
Cor branca	0.51	0.003	0.53	0.003	0.40	0.009
Educação em anos de estudo	8.19	0.030	8.73	0.027	5.33	0.065
Empregados	0.64	0.003	0.71	0.002	0.38	0.009
Conta-própria	0.22	0.002	0.21	0.002	0.27	0.005
Empregador	0.04	0.001	0.05	0.001	0.02	0.001
N_remunerado, Const Prop. Uso, Próprio consumo	0.10	0.003	0.04	0.001	0.33	0.007
Agrícola	0.190	0.005	0.0700	0.002	0.71	0.012
Não-agrícola	0.81	0.005	0.93	0.002	0.29	0.012

Fonte: Elaborado pela autora a partir dos microdados da PNAD-2005

Tabela A.11. Análise descritiva das principais variáveis utilizadas na pesquisa.

Variáveis	TODOS		URBANO		RURAL	
	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.
Economicamente ativos	0.72	0.001	0.71	0.001	0.78	0.005
Renda do trabalho principal	811.27	8.708	920.50	9.801	314.40	8.795
Log da razão de renda por hora trabalhada	1.2484	0.001	1.2664	0.001	1.13	0.004
Rendimento Familiar	1844.15	20.041	2018.99	23.124	906.64	17.007
Número de horas trabalho principal (semana)	26.31	0.07	26.18	0.068	26.99	0.221
Número de Componentes na Fam.	3.67	0.008	3.58	0.007	4.13	0.030
Gênero (Masculino)	0.4820	0.001	0.4738	0.0008	0.53	0.002
Idade (em dezenas de anos)	3.66	0.004	3.66	0.004	3.65	0.011
idade^2 (em dezenas de anos)	15.56	0.034	15.54	0.036	15.64	0.089
Cor preta ou indígena	0.08	0.001	0.08	0.001	0.06	0.004
Cor amarela ou pardas	0.42	0.003	0.39	0.003	0.54	0.009
Cor branca	0.51	0.003	0.53	0.003	0.39	0.009
Educação em anos de estudo	8.39	0.028	8.93	0.026	5.50	0.062
Empregados	0.65	0.003	0.71	0.002	0.38	0.009
Conta-própria	0.21	0.002	0.20	0.001	0.27	0.005
Empregador	0.04	0.001	0.05	0.001	0.02	0.001
N_remunerado, Const Prop. Uso, Próprio consumo	0.09	0.002	0.04	0.001	0.32	0.007
Agrícola	0.179	0.004	0.0657	0.002	0.70	0.012
Não-agrícola	0.82	0.004	0.93	0.002	0.30	0.012

Fonte: Elaborado pela autora a partir dos microdados da PNAD-2006

Tabela A.12: Análise descritiva das principais variáveis utilizadas na pesquisa.

Variáveis	TODOS		URBANO		RURAL	
	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.
Economicamente ativos	0.72	0.002	0.71	0.001	0.77	0.004
Renda do trabalho principal	839.10	8.766	944.98	9.853	343.00	10.217
Log da razão de renda por hora trabalhada	1.2632	0.002	1.2798	0.002	1.15	0.004
Rendimento Familiar	1881.64	19.821	2045.43	22.626	994.61	24.650
Número de horas trabalho principal (semana)	26.31	0.07	26.27	0.071	26.56	0.221
Número de Componentes na Fam.	3.66	0.008	3.58	0.007	4.08	0.027
Gênero (Masculino)	0.4834	0.001	0.4755	0.0008	0.53	0.002
Idade (em dezenas de anos)	3.69	0.004	3.69	0.004	3.68	0.012
idade^2 (em dezenas de anos)	15.78	0.035	15.76	0.037	15.91	0.099
Cor preta ou indígena	0.08	0.001	0.09	0.001	0.07	0.004
Cor amarela ou pardas	0.41	0.003	0.39	0.003	0.54	0.010
Cor branca	0.50	0.003	0.52	0.003	0.39	0.010
Educação em anos de estudo	8.52	0.029	9.04	0.027	5.68	0.063
Empregados	0.66	0.003	0.72	0.002	0.39	0.010
Conta-própria	0.21	0.002	0.20	0.002	0.26	0.005
Empregador	0.04	0.001	0.04	0.001	0.02	0.001
N_remunerado, Const Prop. Uso, Próprio consumo	0.09	0.003	0.04	0.001	0.33	0.010
Agrícola	0.170	0.004	0.0598	0.002	0.69	0.012
Não-agrícola	0.83	0.004	0.94	0.002	0.31	0.012

Fonte: Elaborado pela autora a partir dos microdados da PNAD-2007

APÊNDICE 2

EXEMPLO DO ALGORITMO DESENVOLVIDO PARA O ANO DE 2007

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

cd "D:\PATRICIA\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:\PATRICIA\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:\PATRICIA\PNADS\2007\domicilio07.dta" , replace

*****
*****
*****

#delimit;
clear;
use "D:\PATRICIA\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:\PATRICIA\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_fal 738-749 rend_p 690-701
fam_comp 752-753 areacen 754-754
sitcen 755-755 peso 756-760 pesofam 761-765
using "D:\PATRICIA\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:\PATRICIA\PNADS\2007\domicilio07.dta";
#delimit cr

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

sum

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

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

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

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

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

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

```

```

svydes, single

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

use "D:\PATRICIA\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:\PATRICIA\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:\PATRICIA\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 DO.FILE idonepsu - ANO DE 2007

use "D:\PATRICIA\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:\PATRICIA\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:\PATRICIA\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)

drop horas_total

#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:\PATRICIA\PNADS\2007\pess07.dta" , replace

*****

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

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

recode rend_tod (1.00e+12/2.29e+12=.) (-1=.)
recode rend_fal (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 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)

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

gen ln_rend_p = ln(rend_p)

gen ln_horas_p = ln(horas_p)

drop if ln_rend_p <1

gen ln_rend_p_ln_horas = (ln_rend_p/ln_horas_p)

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

*****

set more off

*****OUTRA FORMA DE CRIAR DUMMY'S E ATRAVES DO COMANDO
*****tab varname, gen(varname)

***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=.)
replace edu = . if idade <15

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

*****

*** 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 regiao =.;
replace regiao = 1 if uf >=50 & uf < 88;
replace regiao = 2 if uf >=31 & uf <= 35;
replace regiao = 4 if uf >=41 & uf <=43;
replace regiao = 3 if uf <=17;
replace regiao = 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 regiao==2;
replace sudeste = 0 if regiao~2;
gen sul =.;
replace sul = 1 if regiao ==4;
replace sul = 0 if regiao ~4;
gen centro =.;
replace centro = 1 if regiao==1;
replace centro = 0 if regiao~1;
gen norte=.;
replace norte = 1 if regiao ==3;
replace norte = 0 if regiao ~3;
tab sudeste;
tab sul;
tab centro;
tab norte;
#delimit cr

save "D:\PATRICIA\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

```

```

#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:\PATRICIA\PNADS\2007\pess07.dta" , replace

*****
*****

***seperacao da cor***
**brancas=2
**preta=4
**amarela=6
**parda=8
**indigena=0

*****
*****definindo rótulos para cor ou raça*****

label define definicao 2 branca 4 parda 8 negro 6 amarela 0 indigena

label values cor definicao

tab cor

label define gender 1 homem 0 mulher

label values sexo gender

tab sexo
*****

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:\PATRICIA\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:\PATRICIA\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:\PATRICIA\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 (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:\PATRICIA\PNADS\2007\pess07.dta" , replace

sum
*****
*****TESTE DAS 15 VAR. DUMMY'S PARA EDU*****
*****

*****todos*****
#delimit;
svy: heckman ln_rend_p_ln_horas sexo idade10 idade10q amarela_pard
brancas edu2 edu3 edu4 edu5 edu6 edu7 edu8 edu9
edu10 edu11 edu12 edu13 edu14 edu15 edu16 cont_propria empregador
n_agricola if idade>=15 & idade<71
, select(cond_ativid = sexo fam_comp rend_fal idade10 idade10q
amarela_pard edu2 edu3 edu4 edu5 edu6 edu7 edu8 edu9
edu10 edu11 edu12 edu13 edu14 edu15 edu16 brancas);
#delimit cr;

*****urbano*****
#delimit;
svy: heckman ln_rend_p_ln_horas sexo idade10 idade10q amarela_pard
brancas edu2 edu3 edu4 edu5 edu6 edu7 edu8 edu9
edu10 edu11 edu12 edu13 edu14 edu15 edu16 cont_propria empregador
n_agricola if idade>=15 & idade<71 & urbano==1
, select(cond_ativid = sexo fam_comp rend_fal idade10 idade10q
amarela_pard edu2 edu3 edu4 edu5 edu6 edu7 edu8 edu9
edu10 edu11 edu12 edu13 edu14 edu15 edu16 brancas);
#delimit cr;

*****rural*****
#delimit;
svy: heckman ln_rend_p_ln_horas sexo idade10 idade10q amarela_pard
brancas edu2 edu3 edu4 edu5 edu6 edu7 edu8 edu9
edu10 edu11 edu12 edu13 edu14 edu15 edu16 cont_propria empregador
n_agricola if idade>=15 & idade<71 & rural==1
, select(cond_ativid = sexo fam_comp rend_fal idade10 idade10q
amarela_pard edu2 edu3 edu4 edu5 edu6 edu7 edu8 edu9
edu10 edu11 edu12 edu13 edu14 edu15 edu16 brancas);
#delimit cr;

*****CRIANDO EFEITO LIMIAR DA EDU TODOS*****

#delimit;
gen zt=.;
replace zt = 0 if edu <=12;
replace zt = 1 if edu >12;
gen st= edu - 12;
gen eledu_t = zt*st;
#delimit cr;

*****CRIANDO EFEITO LIMIAR DA EDU URBANO*****

#delimit;
gen zu=.;
replace zu = 0 if edu <=12;

```

```

replace zu = 1 if edu >12;
gen su= edu - 12;
gen eledu_u = zu*su;
#delimit cr;

*****CRIANDO EFEITO LIMIAR DA EDU RURAL*****

#delimit;
gen zr=.;
replace zr = 0 if edu <=12;
replace zr = 1 if edu >12;
gen sr= edu - 12;
gen eledu_r = zr*sr;
#delimit cr;

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

#delimit;
svy linearized : mean rend_p if idade>=15 & idade<71 ;
svy linearized : total rend_p if idade>=15 & idade<71;
svy linearized : mean ln_rend_p_ln_horas if idade>=15 & idade<71 ;
svy linearized : mean idade10 if idade>=15 & idade<71 ;
svy linearized : mean idade10q if idade>=15 & idade<71;
svy linearized : mean edu if idade>=15 & idade<71;
svy linearized : mean rend_fal if idade>=15 & idade<71;
svy linearized : mean fam_comp if idade>=15 & idade<71;
svy linearized : mean horas_p if idade>=15 & idade<71;
svy linearized : proportion cond_ativid if idade>=15 & idade<71;
svy linearized : proportion preta_ind if idade>=15 & idade<71;
svy linearized : proportion amarela_pard if idade>=15 & idade<71;
svy linearized : proportion cor if idade>=15 & idade<71;
svy linearized : proportion sexo if idade>=15 & idade<71;
svy linearized, subpop(if idade >=15 & idade <71) : tabulate pos_ocup,
se;
svy linearized, subpop(if idade >=15 & idade <71) : tabulate atividade,
se;
#delimit cr

*****
*****ESTATÍSTICA DESCRITIVA URBANO*****
*****

#delimit;
svy linearized : mean rend_p if idade>=15 & idade<71 & urbano==1;
svy linearized : total rend_p if idade>=15 & idade<71 & urbano==1;
svy linearized : mean ln_rend_p_ln_horas if idade>=15 & idade<71 &
urbano==1;
svy linearized : mean idade10 if idade>=15 & idade<71 & urbano==1;
svy linearized : mean idade10q if idade>=15 & idade<71 & urbano==1;
svy linearized : mean edu if idade>=15 & idade<71 & urbano==1;
svy linearized : mean rend_fal if idade>=15 & idade<71 & urbano==1;
svy linearized : mean fam_comp if idade>=15 & idade<71 & urbano==1;
svy linearized : mean horas_p if idade>=15 & idade<71 & urbano==1;

```

```

svy linearized : proportion cond_ativid if idade>=15 & idade<71 &
urbano==1;
svy linearized : proportion preta_ind if idade>=15 & idade<71 &
urbano==1;
svy linearized : proportion amarela_pard if idade>=15 & idade<71 &
urbano==1;
svy linearized : proportion cor if idade>=15 & idade<71 & urbano==1;
svy linearized : proportion sexo if idade>=15 & idade<71 & urbano==1;
svy linearized, subpop(if idade >=15 & idade <71 & urbano==1) : tabulate
pos_ocup, se;
svy linearized, subpop(if idade >=15 & idade <71 & urbano==1) : tabulate
atividade,se;
#delimit cr

#delimit
svy linearized : mean rend_p if idade>=15 & idade<71 & urbano==1 &
agricola==1;
svy linearized : mean rend_p if idade>=15 & idade<71 & urbano==1 &
n_agricola==1;
svy linearized : total rend_p if idade>=15 & idade<71 & urbano==1 &
agricola==1;
svy linearized : mean rend_fal if idade>=15 & idade<71 & urbano==1 &
agricola==1;
svy linearized : total rend_p if idade>=15 & idade<71 & urbano==1 &
n_agricola==1;
svy linearized : mean rend_fal if idade>=15 & idade<71 & urbano==1 &
n_agricola==1;
#delimit cr

*****
*****
*****ESTATÍSTICA DESCRITIVA RURAL*****
*****
*****

#delimit;
svy linearized : mean rend_p if idade>=15 & idade<71 & rural==1;
svy linearized : total rend_p if idade>=15 & idade<71 & rural==1;
svy linearized : mean ln_rend_p_ln_horas if idade>=15 & idade<71 &
rural==1;
svy linearized : mean idade10 if idade>=15 & idade<71 & rural==1;
svy linearized : mean idade10q if idade>=15 & idade<71 & rural==1;
svy linearized : mean edu if idade>=15 & idade<71 & rural==1;
svy linearized : mean rend_fal if idade>=15 & idade<71 & rural==1;
svy linearized : mean fam_comp if idade>=15 & idade<71 & rural==1;
svy linearized : mean horas_p if idade>=15 & idade<71 & rural==1;
svy linearized : proportion cond_ativid if idade>=15 & idade<71 &
rural==1;
svy linearized : proportion preta_ind if idade>=15 & idade<71 & rural==1;
svy linearized : proportion amarela_pard if idade>=15 & idade<71 &
rural==1;
svy linearized : proportion cor if idade>=15 & idade<71 & rural==1;
svy linearized : proportion sexo if idade>=15 & idade<71 & rural==1;
svy linearized, subpop(if idade >=15 & idade <71 & rural==1) : tabulate
pos_ocup, se;
svy linearized, subpop(if idade >=15 & idade <71 & rural==1) : tabulate
atividade, se;
#delimit cr

#delimit

```

```

svy linearized : mean rend_p if idade>=15 & idade<71 & rural==1 &
agricola==1;
svy linearized : mean rend_p if idade>=15 & idade<71 & rural==1 &
n_agricola==1;
svy linearized : total rend_p if idade>=15 & idade<71 & rural==1 &
agricola==1;
svy linearized : mean rend_fal if idade>=15 & idade<71 & rural==1 &
agricola==1;
svy linearized : total rend_p if idade>=15 & idade<71 & rural==1 &
n_agricola==1;
svy linearized : mean rend_fal if idade>=15 & idade<71 & rural==1 &
n_agricola==1;
#delimit cr;

*****
***** Análise de Regressão total *****
*****

#delimit;
svy: heckman ln_rend_p_ln_horas urbano sexo idade10 idade10q amarela_pard
brancas edu eledu_t
cont_propria empregador n_agricola if idade>=15 & idade<71
, select(cond_ativid = urbano sexo fam_comp rend_fal idade10 idade10q
amarela_pard brancas edu eledu_t);
#delimit cr;

*****
***** Análise de Regressão URBANO*****
*****

#delimit;
svy: heckman ln_rend_p_ln_horas sexo idade10 idade10q amarela_pard
brancas edu eledu_u
cont_propria empregador n_agricola if idade>=15 & idade<71 & urbano==1
, select(cond_ativid = sexo fam_comp rend_fal idade10 idade10q
amarela_pard brancas edu eledu_u);
#delimit cr;

*****
***** Análise de Regressão RURAL*****
*****

#delimit;
svy: heckman ln_rend_p_ln_horas sexo idade10 idade10q amarela_pard
brancas edu eledu_r
cont_propria empregador n_agricola if idade>=15 & idade<71 & rural==1
, select(cond_ativid = sexo fam_comp rend_fal idade10 idade10q
amarela_pard brancas edu eledu_r);
#delimit cr;

*****
*****EFEITO MARGINAL*****
*****

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

*****
*****MULTICOLINEARIDADE*****
*****

```

```

*****

#delimit
corr ref pm_solteiro casal cond_ativid rend_fal 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 //

*****

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

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

***INDICES DE DESIGUALDADE POR ESTADOS TOTAL***

inequal2 rend_tod if uf==11 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==12 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==13 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==14 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==15 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==16 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==17 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==21 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==22 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==23 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==24 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==25 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==26 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==27 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==28 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==29 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==31 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==32 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==33 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==35 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==41 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==42 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==43 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==50 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==51 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==52 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==53 & rend_tod > 0 [fw=peso]

***INDICES DE DESIGUALDADE POR ESTADOS URBANO***

inequal2 rend_tod if uf==11 & urbano==1 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==12 & urbano==1 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==13 & urbano==1 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==14 & urbano==1 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==15 & urbano==1 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==16 & urbano==1 & rend_tod > 0 [fw=peso]
inequal2 rend_tod if uf==17 & urbano==1 & rend_tod > 0 [fw=peso]

```


ANÁLISE MALMQUIST 2007

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

cd "D:\PATRICIA\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:\PATRICIA\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:\PATRICIA\PNADS\2007\domicilio07.dta" , replace

*****
*****
*****

#delimit;
clear;
use "D:\PATRICIA\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:\PATRICIA\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
rend_tod 702-713 rend_fon 714-725 rend_fal 738-749 rend_p 690-701
fam_comp 752-753 areacen 754-754
sitcen 755-755 peso 756-760 pesofam 761-765
using "D:\PATRICIA\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:\PATRICIA\PNADS\2007\domicilio07.dta";
#delimit cr

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

sum

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

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

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

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

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

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

svydes, single

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

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

capture program drop progl
program define progl
use "D:\PATRICIA\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:\PATRICIA\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'
progl
}

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 DO.FILE idonepsu - ANO DE 2007

use "D:\PATRICIA\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:\PATRICIA\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:\PATRICIA\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)

drop horas_total

#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:\PATRICIA\PNADS\2007\pess07.dta" , replace

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

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

recode rend_tod (1.00e+12=.)
recode rend_fal (1.00e+12=.)
recode rend_p (1.00e+12=.)

***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_horas_t = ln(rend_tod/horas_totalmes)

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

*** 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 regiao =.;
replace regiao = 1 if uf >=50 & uf < 88;
replace regiao = 2 if uf >=31 & uf <= 35;
replace regiao = 4 if uf >=41 & uf <=43;
replace regiao = 3 if uf <=17;
replace regiao = 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 regiao==2;
replace sudeste = 0 if regiao~2;
gen sul =.;
replace sul = 1 if regiao ==4;
replace sul = 0 if regiao ~4;
gen centro =.;
replace centro = 1 if regiao==1;
replace centro = 0 if regiao~1;
gen norte=.;
replace norte = 1 if regiao ==3;
replace norte = 0 if regiao ~3;
tab sudeste;
tab sul;
tab centro;
tab norte;
#delimit cr

save "D:\PATRICIA\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 *****

recode pos_ocup (1/8=1) (14=.) (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

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

*****

***seperacao da cor****

recode cor (9=.)

#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:\PATRICIA\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;

****CRIANDO VARIAVEIS URBANO E RURAL EM APENAS UMA****
***urbano = 1 rural = 2****

#delimit;
gen urban_rural = .;
replace urban_rural=sitcen;
recode urban_rural (1/3=1) (4/8=2);
#delimit cr;

sum rural urbano

save "D:\PATRICIA\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 > 1;
replace n_agricola = 0 if atividade == 1;
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:\PATRICIA\PNADS\2007\pess07.dta" , replace

```

```
*****
*****CRIANDO DUMMY'S P/ EDUCACAO (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:\PATRICIA\PNADS\2007\pess07.dta" , replace
```

```
*****
*****TESTE DAS 15 VAR. DUMMY'S PARA EDU*****
*****
```

```
#delimit;
```

```
svy: heckman ln_rend_horas_p sexo idade10 idade10q brancas pardas
amarelas indigenas edu2 edu3 edu4 edu5 edu6 edu7 edu8 edu9
edu10 edu11 edu12 edu13 edu14 edu15 edu16 sudeste sul centro norte
n_agricola empregador cont_propria
select(cond_ativid = rend_fal fam_comp sexo idade10 idade10q brancas
pardas amarelas indigenas edu2 edu3
edu4 edu5 edu6 edu7 edu8 edu9 edu10 edu11 edu12 edu13 edu14 edu15 edu16
sudeste sul centro norte n_agricola empregador cont_propria);
#delimit cr;
```

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

```
#delimit;
sum sudeste sul centro norte cond_ativid ln_rend_tod rend_tod rend_fal
fam_comp sexo idade10 idade10q preta pardas amarelas brancas
edu eledu urbano n_agricola agricola empregado empregador cont_propria
[fw=peso];
#delimit cr;
```

```
*****
***** Análise de Regressão*****
*****
```

```
#delimit;
svy: heckman ln_rend_horas_p sexo idade10 idade10q brancas pardas
amarelas indigenas edu eledu
sudeste sul centro norte n_agricola empregador cont_propria
select(cond_ativid = rend_fal fam_comp sexo idade10 idade10q brancas
pardas amarelas indigenas edu eledu
sudeste sul centro norte n_agricola empregador cont_propria);
#delimit cr;
```

```
*****
*****
*****EFEITO MARGINAL*****
*****
```

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

```
*****
*****MULTICOLINEARIDADE*****
*****
```

```
#delimit
corr cond_ativid rend_fal ln_rend_horas_p fam_comp sexo idade10 idade10q
brancas pardas amarelas indigenas
edu eledu n_agricola sudeste sul centro norte n_agricola empregador
cont_propria;
#delimit cr;
```

```
*****
*****EFEITO DO PLANO AMOSTRAL*****
*****
```

```

*****

estat effects, meff deff // para analisar o efeito do plano amostral //

*****
*****CHEGANDO A EXISTENCIA DE HETEROCEDASTICIDADE*****
*****

#delimit
anova cond_ativid [fw=peso];
hettest;
anova ln_rend_horas_p [fw=peso];
hettest;
anova ln_rend_horas_t [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 brancas [fw=peso];
hettest;
anova pardas [fw=peso];
hettest;
anova amarelas [fw=peso];
hettest;
anova indigenas [fw=peso];
hettest;
anova edu [fw=peso];
hettest;
anova edu2 [fw=peso];
hettest;
anova urbano [fw=peso];
hettest;
anova n_agricola [fw=peso];
hettest;
anova sudeste [fw=peso];
hettest;
anova sul [fw=peso];
hettest;
anova centro [fw=peso];
hettest;
anova norte [fw=peso];
hettest norte;
anova empregador [fw=peso];
hettest norte;
anova cont_propria [fw=peso];
hettest norte;
#delimit cr;

*****análise Malmkish*****

**total
tabstat horas_p edu rend_p [fweight = peso], statistics( sum ) by(uf)
columns(variables) format(%15.0g)

```

```
**urbano
tabstat horas_p edu rend_p [fweight = peso] if urban_rural==1,
statistics( sum ) by(uf) columns(variables) format(%15.0g)

**rural
tabstat horas_p edu rend_p [fweight = peso] if urban_rural==2,
statistics( sum ) by(uf) columns(variables) format(%15.0g)
```

** FIM DO ALGORITMO **