

DETERMINANTES DOS DIFERENCIAIS DE RENDIMENTOS DO TRABALHO: UMA ABORDAGEM HIERÁRQUICA PARA OS ESTADOS BRASILEIROS*

Luzia Maria Cavalcante de Melo

Doutoranda em economia pelo CEDEPLAR-UFMG

E-mail: luziamcm@gmail.com

RESUMO Diferenças nos rendimentos podem ser determinadas tanto pelo lado da oferta quanto pelo lado da demanda por trabalhadores. Pelo lado da oferta, trabalhadores com características diferenciadas auferem rendimentos diferenciados. Pelo lado da demanda, as diferenças na estrutura de mercado de trabalho local podem se refletir em ganhos salariais diferenciados, mesmo para trabalhadores com características semelhantes. Esse trabalho tem como objetivo identificar variáveis determinantes do diferencial de rendimentos do trabalho no Brasil. Como os rendimentos são influenciados tanto por características individuais como por características do local onde o trabalhador está inserido, optou-se por utilizar a metodologia de modelos hierárquicos, que permite incorporar, no mesmo modelo, variáveis medidas em níveis diferentes. Apesar de dificuldades relacionadas ao tamanho da amostra e ao recorte espacial dado no trabalho, os resultados mostram que tanto as características individuais, como as economias de aglomeração possuem efeitos sobre o diferencial de rendimentos entre as Unidades da Federação do Brasil.

Código JEL: R10, R23

Palavras-chave: Economias de Aglomeração, Modelos Hierárquicos, Rendimentos

ABSTRACT Differences in income may be determined by both the supply side by side as demand for workers. On the supply side, workers with different characteristics earn different incomes. On the demand side, differences in the structure of local labor market may be reflected in earnings wage differential, even for workers with similar characteristics. This work aims to identify variables determining the differential of income from work in Brazil. As the income is influenced by both individual characteristics and by characteristics of the place where the employee is inserted, we opted to use the methods of hierarchical models, which allows to incorporate, in the same model, variables measured at different levels. In spite of difficulties related to sample size and the cut taken in the work space, the results show that both individual characteristics such as economies of agglomeration have effects on the gap in incomes between the units of the Federation of Brazil.

Key-words: Economies of Agglomeration, Hierarchical Models, Income

* Artigo recebido em março/2009 e aceito em maio/2009.

1. Introdução

O Brasil se caracteriza por fortes desigualdades de renda não só entre os indivíduos, mas também entre as Regiões e até mesmo entre as Unidades da Federação (UF's). A desigualdade de renda entre diversas localidades reflete, entre outros aspectos, a desigualdade na distribuição das atividades produtivas. Uma vez que a atividade produtiva é distribuída de maneira desigual entre as UF's, isso gera também desigualdades nos mercados de trabalho e, conseqüentemente, nos níveis salariais.

O processo de industrialização no país se caracterizou pela forte aglomeração da atividade produtiva em São Paulo. Apesar de, entre o final da década de 1960 e meados da década de 1980, terem surgido fatores desaglomerativos, como o aumento dos custos da renda fundiária urbana e o aumento dos congestionamentos urbanos; o movimento de desaglomeração se deu em direção a Região Metropolitana de São Paulo (RMSP) e as demais UF's das Regiões Sudeste e Sul, que ainda concentram grande parte da atividade econômica do país.

De fato, a região Nordeste apresentou altas taxas de crescimento nas últimas décadas, e as regiões Centro-Oeste e pequena parte da região Norte foram favorecidas positivamente pela expansão das fronteiras agrícolas e minerais, que proporcionou a modernização da agricultura e atraiu a moderna agroindústria para o local. Contudo, isso não foi o bastante para modificar de forma significativa a concentração da atividade produtiva no país, permanecendo a segregação econômica entre Norte-Nordeste e Sudeste-Sul.

A literatura sobre mercado de trabalho mostra que características pessoais influenciam na desigualdade dos rendimentos. Pesquisas mostram que em média os homens possuem rendimentos maiores que as mulheres, o nível salarial dos brancos é mais alto que o dos não brancos – isso pode estar correlacionado com o fato de que a qualificação dos não brancos é em média mais baixa que a dos brancos, mas ainda assim, existem evidências de que existe um componente de discriminação que diferencia o salário de uma pessoa branca em relação a uma não branca, mesmo sendo ambas com o mesmo nível de escolaridade –, pessoas mais velhas ganham mais que as mais jovens (influência da experiência) e o nível dos rendimentos é mais alto quanto mais alta é a escolaridade do indivíduo.

No entanto, não só as características individuais, mas também as características do mercado de trabalho e da estrutura produtiva do local onde o trabalhador está inserido influenciam na determinação dos rendimentos. Assim, diferenças no nível de rendimento dos trabalhadores podem também ser determinadas por diferenças nos atributos econômicos urbanos, como aglomerações urbanas, concentração da atividade produtiva e a estrutura de demanda por trabalho, bem como características individuais podem ter influência diferenciada de um local para outro de acordo com as diferenças entre os atributos locais.

O objetivo deste trabalho é analisar a influência de características pessoais e de atributos locais sobre a diferença nos rendimentos do trabalho entre as UF's do Brasil.

Considerando que as características individuais e locais são medidas em níveis diferentes, esta análise será feita utilizando a metodologia de modelos hierárquicos. O uso desses modelos permite resolver o problema da unidade de análise quando se tem variáveis independentes em distintos níveis de agregação. Neste trabalho será usado um modelo hierárquico de dois níveis, a saber: o modelo de nível 1 que contém variáveis explicativas ao nível do indivíduo e capta a diferença nos rendimentos devido a características individuais e, o modelo de nível 2, que contém variáveis ao nível da UF e capta a diferença nos rendimentos determinadas por características locais. Apesar das diversas características locais que podem ter influência sobre os rendimentos, optou-se nesse trabalho por considerar variáveis que capturem efeitos das economias de aglomeração. Com essa metodologia será possível verificar os efeitos diretos das características individuais sobre os rendimentos, bem como estimar o efeito direto dos atributos locais sobre os rendimentos dos indivíduos, controlando-se pelas características pessoais dos trabalhadores.

Este trabalho está dividido em quatro seções além desta introdução. Na próxima seção será feita uma breve revisão da literatura que fornece evidências sobre os determinantes dos diferenciais salariais. A seção seguinte apresenta o modelo, as variáveis incluídas, os dados e o universo da amostra, bem como o recorte territorial considerado. Em seguida são apresentados os resultados e suas respectivas análises. Por fim são apresentadas as considerações finais.

2. Determinação dos Diferenciais Locacionais de Rendimentos: Uma breve revisão da literatura

Uma das características marcantes da economia brasileira é a aglomeração da atividade produtiva em determinados locais. Segundo Marshall (1890) a existência de muitas firmas de uma indústria em determinado local gera ganhos de aglomeração para as mesmas, uma vez que reduz custos de transporte, atrai mão-de-obra especializada etc. Assim, a aglomeração, como também a especialização local, em determinada atividade gera externalidades positivas para a indústria, denominadas economias de localização.

Já para Jacobs (1969), a principal fonte de externalidades positivas é a diversidade de atividades econômicas que existem nos centros urbanos. Mesmo reconhecendo a eficiência produtiva proporcionada pela especialização dos aglomerados produtivos, Jacobs defende que são as inovações que movem o crescimento econômico, e que estas são mais propensas de serem desenvolvidas e disseminadas em ambientes de alta escala urbana e diversificação econômica.

Hoover (1936) já classificava as economias de aglomeração. Segundo ele as economias de aglomeração podem ter duas classificações distintas, a saber:

Economias de Urbanização: do tipo proposta por Jacobs (1969), são externas às firmas, mas internas ao centro urbano onde estas estão inseridas. Dependem do nível global de atividades econômicas do local e agem sobre todas as firmas das várias indústrias.

Economias de Localização: do tipo proposto por Marshall (1890), onde a especialização em uma atividade econômica gera economias de escala externas às firmas, mas interna à indústria.

O alto nível de aglomeração incide em custos como poluição, congestionamentos urbanos e altos aluguéis, fatores esses que funcionam como forças de desaglomeração. Contudo, os ganhos de produtividade, ou as economias de aglomeração, decorrentes da proximidade geográfica, contrapõem-se às deseconomias de aglomeração. Entre os vários aspectos em que esses ganhos de produtividade podem se refletir, um deles é sobre os rendimentos do trabalho. As economias de aglomeração podem se refletir sobre os rendimentos sejam pelo aumento da produtividade do trabalhador, sejam para compensar os efeitos das deseconomias de aglomeração. Sendo assim, o nível de rendimentos pode ser diferente de uma localidade para outra dependendo do nível de aglomeração da atividade econômica.

Segundo Molho (1992), os diferenciais salariais compensatórios são os principais determinantes das diferenças salariais entre regiões, ou seja, regiões com custo de vida mais alto teriam um nível salarial também mais alto, de forma a equilibrar os salários reais e o padrão de vida dos trabalhadores.

No entanto, tanto diferenças no custo de vida como diferenças de salários podem ser explicadas por vantagens ou desvantagens relativas locais. Combes *et al* (2004) afirmam que a explicação para as diferenças de salários entre regiões é dada pelas diferenças na composição da força de trabalho – principalmente no que se refere à qualificação dessa força de trabalho – e pelas diferenças inter-regionais de produtividade decorrentes da aglomeração econômica e da estrutura produtiva local.

Vários trabalhos empíricos também tentam mostrar os efeitos das aglomerações sobre os ganhos e as diferenças de rendimentos entre localidade. Topel (1986) investiga as diferenças dos mercados locais e o impacto de mudanças nestes sobre emprego e salário. O autor verifica que locais que apresentaram crescimento econômico acima da média tiveram aumento relativo dos salários e diminuição do desemprego e conclui que distúrbios no mercado de trabalho, principalmente do lado da demanda, têm impacto sobre os salários.

Com o objetivo de investigar como a evolução das diferenças regionais na desigualdade salarial, entre as Regiões Metropolitanas do Brasil, pode ser explicada por movimentos regionais e temporais na distribuição da educação e nos perfis salariais por grupos, Reis e Barros (1990) concluem que os diferenciais salariais variam mais acentuadamente entre os níveis educacionais e que, o poder explicativo da educação varia muito entre as regiões; sendo mais significativos para explicar diferenças salariais na região Nordeste e que as diferenças regionais na desigualdade salarial são bem maiores que as diferenças temporais. Isso indica que as diferenças salariais tendem a ser maiores em regiões mais pobres e que, diferentes níveis de educação da população acentuam a desigualdade salarial entre regiões.

Em estudo sobre os diferenciais salariais entre as Regiões Metropolitanas de São Paulo e Salvador, Menezes *et al* (2002) verificam que atributos pessoais como

sexo e cor contribuem para o diferencial no salário – nesse estudo os diferenciais foram acentuados para mulheres e negro – no entanto, a concentração econômica da Região Metropolitana de SP em relação à Região Metropolitana de Salvador foi o principal determinante para o diferencial de salários.

Silveira Neto e Campelo (2003) estudaram as disparidades de rendimentos entre Regiões Metropolitanas, empregando uma variável *dummy* para cada Região a fim de captar os efeitos regionais sobre tais disparidades. Os autores concluem que as disparidades de rendimentos são maiores entre as Regiões Metropolitanas do Nordeste e a região Metropolitana de São Paulo, que os atributos pessoais são bastante significativos para explicar essas disparidades. No entanto, os diferenciais regionais não diminuem mesmo controlando pela influência desses atributos, o que sugere que fatores locais exercem importante influência sobre o diferencial nos rendimentos.

Fontes (2006) analisou os determinantes dos diferenciais salariais em cidades médias e regiões metropolitanas brasileiras no período entre 1991 e 2000, usando um modelo hierárquico de dois níveis; o modelo de nível 1 com características do indivíduo e o modelo de nível 2 com características do centro urbano onde o indivíduo está inserido. Os resultados mostram que os atributos pessoais produtivos, como educação e experiência, foram significantes para explicar os diferenciais salariais entre os centros urbanos, contudo, uma grande parte da variabilidade permaneceu não explicada. A inclusão do setor do trabalho do indivíduo também foi significativa para explicar as diferenças salariais entre os centros urbanos, indicando que essas diferenças salariais se devem também às diferenças na especialização produtiva local. Após a inclusão das variáveis de nível 2, o autor observa que tanto a escala urbana – medida pela população total – como a estrutura produtiva dos centros – medida pela densidade local da oferta de serviços produtivos modernos – foram significantes para a determinação dos diferenciais de salário entre os centros urbanos, indicando a importância das economias de aglomeração e urbanização sobre a produtividade e, conseqüentemente, sobre o nível salarial.

Galinari *et al* (2007), através do método de análise de componente principal, estima os efeitos das economias de aglomeração sobre os salários industriais para as cidades médias no Brasil. Seus resultados mostram que tanto características de aglomeração quanto de diversificação desses locais possuem efeitos significantes sobre os salários.

É neste contexto de disparidades nos rendimentos entre indivíduos e entre localidades, que este trabalho se propõe a contribuir para esta discussão, apresentando evidências dos determinantes dos diferenciais de rendimentos entre as UF's do Brasil. Este trabalho se diferencia de alguns aqui citados pelo uso e algumas variáveis e pelo uso do método de modelos hierárquicos. O uso desse método possibilita melhores estimativas para os parâmetros uma vez que incorpora dois níveis de agregação diferentes de dados e permite a partição da variância observada nos salários entre as fontes de variação envolvidas, ou seja, o indivíduo e a Unidade da Federação.

3. Aspectos Metodológicos

Nesta seção serão apresentadas a fonte dos dados utilizados, a composição da amostra, o recorte espacial, a definição das variáveis e o modelo utilizado.

3.1. *Fonte de dados, Recorte Espacial e Composição da Amostra*

Esse trabalho usa a metodologia de modelos hierárquicos em dois níveis. Aqui, o modelo de nível 1 capta os efeitos das características individuais do trabalhador sobre as diferenças nos rendimentos. Os dados utilizados na composição da amostra de indivíduos foram obtidos da Pesquisa Nacional por Amostra de Domicílios (PNAD) para o ano de 2006.

O modelo de nível de 2 capta não só os efeitos das características individuais do trabalhador sobre as diferenças de rendimento como também características do local onde o trabalhador está inserido. O recorte espacial dado nesse trabalho abrange as 27 Unidades da Federação do Brasil. As variáveis do modelo de nível 2 foram obtidas no *site* do IBGE e calculadas a partir da Pesquisa Industrial Anual (PIA). As variáveis explicativas de nível 2 também são referentes ao ano de 2006.

A amostra obtida da PNAD 2006 selecionou os indivíduos com as seguintes características:

- Residentes em áreas urbanas das 27 Unidades da Federação;
- Maiores de 10 anos de idade;
- Declararam ter trabalho principal no setor formal na semana de referência;
- Possuem rendimento em dinheiro no trabalho principal;
- Declararam sexo, cor ou raça, idade, escolaridade, setor do trabalho principal, rendimento e jornada de trabalho.

A amostra exclui os residentes em áreas não urbanas uma vez que a análise no nível 2 (nível da UF) leva em consideração variáveis relacionadas às economias de aglomeração, que é um aspecto tipicamente da estrutura urbana. Consideram-se apenas os trabalhadores do setor formal afim de não captar distorções salariais inerentes às diferenças entre os setores formal e informal e até mesmo dentro do próprio setor informal, diferenças essas que fogem do escopo de análise desse trabalho.

3.2. *Definição das variáveis*

A variável explicada no modelo é o logaritmo do rendimento-hora mensal do trabalho principal (*lnrendhr*). Essa variável foi calculada a partir de informações contidas na PNAD do ano de 2006.

As variáveis explicativas de nível 1, que captam as características do indivíduo, incluídas nesse modelo são as seguintes:

- *Sexo*: Variável *dummy* que assume valor 1 se o indivíduo for homem, e 0 se for mulher. A variável *sexo* é incluída devido a evidências na literatura de que os rendimentos das mulheres são, em média, menores que os dos homens (SAVEDOFF, 1990).
- *Cor*: Variável *dummy* que assume valor 1 se o indivíduo declarou ser branco, e 0 se o indivíduo declarou ser não branco. A variável *cor* é incluída no modelo para captar diferenciais salariais existentes entre pessoas de cor diferente.
- *Idade2*: É a idade do indivíduo elevada à segunda potência. Essa variável é comumente usada na literatura para captar a experiência, uma vez que os rendimentos do trabalho não são uma função linear da idade.
- *Nível de Escolaridade*: Foram construídas variáveis *dummy* para captar o diferencial de rendimentos entre os salários devido a diferentes níveis de escolaridade. A variável assume três classificações: *fund* (nível fundamental completo), *medio* (nível médio ou segundo grau completo) e *sup* (nível superior completo). Assim, a classificação *fund* assume valor 1 se o trabalhador possui nível fundamental completo, e 0 se possui outro nível de escolaridade; a classificação *medio* assume valor 1 se o trabalhador possui nível médio completo, e 0 se possui outro nível de escolaridade; e a classificação *sup* assume valor 1 se o trabalhador possui nível superior completo, e 0 se possui outro nível de escolaridade¹⁰.
- *Setor de Atividade do Trabalho Principal*: Essa variável não expressa propriamente uma característica do trabalhador. Ela reflete características do mercado de trabalho onde o trabalhador está inserido e, sendo assim, permite captar diferenciais de rendimentos entre os setores de atividade (Fontes, 2006). A variável é uma *dummy* que assume três classificações: *ind*, que assume valor 1 se o trabalhador está empregado na indústria e 0 se está

¹⁰ A variável *Nível de Escolaridade* também inclui os analfabetos (ou pessoas com até três anos de estudo). Porém, de acordo com as características dos modelos de regressão com variáveis *dummy*, se uma variável *dummy* possui *m* categorias, a introdução da variável no modelo deve conter *m-1* variáveis *dummy* (GUAJARATI, 2006). Por isso, o modelo possui *dummy* apenas para os trabalhadores com nível fundamental, médio e superior. Os trabalhadores de nível de escolaridade abaixo destes também estão incluídos na amostra, contudo, são a categoria de controle e não possuem *dummy* que os representem no modelo.

empregado em outro setor; *serv*, que assume valor 1 se o trabalhador está empregado no setor serviços e 0 se está empregado em outro setor; e *Adm-Pub*, que assume valor 1 se o trabalhador está empregado na administração pública e 0 se está empregado em outro setor¹¹.

As tabelas abaixo mostram uma descrição de cada uma dessas variáveis na amostra considerada.

Tabela 1
Composição amostral das variáveis de nível 1

Variable	Obs	Mean	Std. Dev.	Min	Max
Sexo	77.801	0,5847483	0,4927686	0	1
Cor	77.801	0,5450701	0,4979677	0	1
Idade2	77.801	1434,265	906,8001	144	8.281
Fund	77.801	0,1668359	0,3728317	00	1
Medio	77.801	0,4310099	0,4952207	00	1
Sup	77.801	0,1492783	0,3563649	0	1
Ind	77.801	0,2432617	0,4290545	0	1
Serv	77.801	0,517654	0,4996914	0	1
AdmPub	77.801	0,536839	0,4855621	0	1

Fonte: Elaboração própria a partir dos dados da PNAD 2006

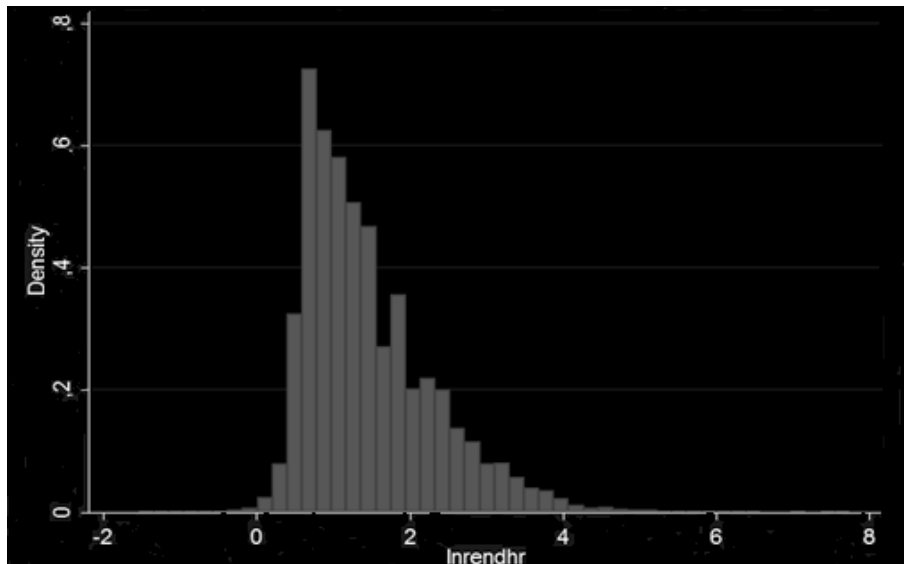
Tabela 2
Descrição da Variável ln do Rendimento-hora mensal por UF

UF	Mean	Std. Dev.	Frequência
Rondônia	1,5316912	0,87406465	957
Acre	1,6485094	0,90710214	557
Amazonas	1,4528354	0,75634648	1.462
Roraima	1,5933305	0,91685662	353
Pará	1,3464029	0,78369131	2.911
Amapá	1,5654773	0,7535877	464
Tocantins	1,3438565	0,81459553	771
Maranhão	1,2954893	0,88873448	585
Piauí	1,3588634	0,85384788	519
Ceará	1,2241231	0,79471882	3.544
Rio Grande do Norte	1,2778619	0,80207826	892
Paraíba	1,3431119	0,84826365	874
Pernambuco	1,2397911	0,8131243	4.039
Alagoas	1,2802072	0,87167169	647
Sergipe	1,3218748	0,81571888	879
Bahia	1,3423015	0,81703109	5.426
Minas Gerais	1,3401472	0,78769963	8.369
Espírito Santo	1,3286649	0,77702597	1.638
Rio de Janeiro	1,5642776	0,85868453	6.249
São Paulo	1,595563	0,78554156	12.357
Paraná	1,5065679	0,79015603	4.753
Santa Catarina	1,5355824	0,7503823	2.732
Rio Grande do Sul	1,5240702	0,78528525	7.287
Mato Grosso do Sul	1,4252356	0,83002313	1.547
Mato Grosso	1,5080481	0,8377754	1.520
Goiás	1,3443785	0,75204491	3.317
Distrito Federal	1,9668976	1,0413828	3.152
Brasil	1,460779	0,82824585	77.801

Fonte: Elaboração própria a partir dos dados da PNAD 2006

¹¹ A variável *Setor de Atividade do Trabalho Principal* também inclui trabalhadores do setor de “comércio de reparação”. Porém, como este setor foi escolhido como categoria de controle, o modelo não possui uma *dummy* que o represente; pela mesma justificativa da nota acima, para a variável *Nível de Escolaridade*.

Gráfico 1
Distribuição da Variável Explicada (*lnrendhr*)



Fonte: Elaboração própria a partir dos dados da PNAD 2006

A tabela 1 mostra, respectivamente, a média, o desvio padrão e os valores de máximo e mínimo para cada variável de nível 1. A tabela 2 mostra, respectivamente, a média e o desvio padrão da variável dependente do modelo, bem como a frequência de distribuição das variáveis entre as UF's. Pela tabela 1 vemos também que a amostra para o modelo de nível 1 contém 77.801 observações, distribuídas entre as 27 UF's de acordo com a última coluna da tabela 2. Contudo, como o recorte espacial o trabalho considera as UF's, o modelo de nível 2 possui apenas 27 observações. O gráfico 1 mostra a distribuição da variável explicada, logaritmo do rendimento-hora mensal (*lnrendhr*).

A fim de captar efeitos de economias de aglomeração sobre os diferenciais de rendimento entre as UF's foram consideradas as seguintes variáveis de nível 2¹²:

- *Taxa de Urbanização (urban)*: A taxa de urbanização da UF é a variável utilizada no modelo para captar efeitos de economias de urbanização.
- *Concentração Industrial (ConcInd)*: Essa variável representa a densidade e escala industrial da UF. Segundo a literatura (Marshall, 1890), a concentração industrial gera economias de aglomeração internas e externas à firma,

¹² A variável Taxa de Urbanização foi obtida no *site* do IBGE e é referente ao ano de 2006. A Concentração Industrial e o Índice de *Hirschman-Herfindal (HH)* foram calculados a partir de dados da Pesquisa Industrial Anual (PIA) do ano de 2006.

e/ou o setor, que podem ter efeitos positivos sobre os salários. A variável é calculada pela razão entre o emprego industrial local e o emprego industrial nacional.

- *Índice de Hirschman-Herfindal (HH)*¹³: Essa variável capta efeitos da variabilidade produtiva entre as UF's sobre os diferenciais de rendimentos. Assim, permite verificar se é a diversidade ou a especialização produtiva das UF's que influenciam de forma positiva os rendimentos locais. O *HH* varia entre 0 e 2; se o *HH*=0 tem-se o caso de um setor industrial perfeitamente diverso, quando o *HH* tende a 2 o setor industrial é completamente concentrado.

3.3. O modelo

Os modelos hierárquicos se diferenciam dos modelos de regressão simples pela sua capacidade e incorporar variáveis explicativas medidas em diferentes níveis. No caso deste trabalho, como discutido anteriormente, sabe-se que os diferenciais nos rendimentos são determinados tanto por características individuais do trabalhador, como por características do local onde o trabalhador está inserido. Assim, esse tipo de problema exige uma abordagem hierárquica, que possa captar tanto as características do trabalhador – unidade de nível 1 –, como as de seu local de trabalho – unidade de nível 2.

Além de poder incorporar variáveis medidas em níveis distintos, a metodologia de modelos hierárquicos permite a partição da variância residual entre os dois níveis, a utilização mais eficiente dos dados e a estimação mais consistente dos parâmetros, uma vez que incorpora variáveis que não poderiam ser incorporadas em análises de regressão medida em um único nível.

O modelo hierárquico utilizado nesse trabalho possui a seguinte estrutura:

- Indivíduo empregado = unidade de nível 1 (subscrito $i = 1, 2, \dots, n_i$);
- Unidade da Federação (UF) = unidade de nível 2 (subscrito $j = 27$);
- $\ln \text{rendhr}_{ij}$ é a variável resposta do modelo, e representa o rendimento/hora mensal do trabalho principal do indivíduo i na UF j .

— O modelo hierárquico nulo ou não condicional (não inclui variáveis explicati-

¹³ Cálculo do *HH*: $HH_i = \sum_{j=1}^n [(E_{ij} / E_i) - (E_j / E_p)]^2$,

Onde: E_j = emprego na UF i no setor j ;

E_i = emprego industrial total na UF i ;

E_j = emprego nacional no setor j ;

E_p = emprego industrial total no Brasil;

n = número de setores considerados no cálculo ($n=25$).

vas e possui apenas o intercepto) para a variável $\ln \text{rendhr}_{ij}$ é especificado da seguinte forma:

Onde:

$$\ln \text{rendhr}_{ij} = \beta_{0j} + \varepsilon_{ij} \quad \varepsilon_{ij} \sim N(0, \sigma_{\varepsilon}^2) \quad (1)$$

$$\beta_{0j} = \beta_{00} + u_{0j} \quad u_{0j} \sim N(0, \sigma_{u0}^2) \quad (2)$$

A partição da variância permite calcular o efeito das características da UF sobre a variação na variável resposta. Esse efeito pode ser medido pelo coeficiente de β_{00} : representa a média global do rendimento/hora mensal do trabalho principal do indivíduo.

u_{0j} : representa o afastamento do rendimento/hora mensal do trabalho principal da UF j em relação à média global (efeito aleatório associado ao nível 2).

ε_{ij} : representa o efeito aleatório associado ao nível 1.

σ_{u0}^2 : variância do rendimento/hora mensal do trabalho principal entre as UF's.

σ_{ε}^2 : variância do rendimento/hora mensal do trabalho principal intra-UF.

correlação intra-UF, o qual mede a proporção da variância entre as UF's em face da variância total, isto é, o quanto da variação do rendimento/hora do trabalho entre os indivíduos empregados é explicado por diferenças existentes entre as UF's onde trabalham. O coeficiente é calculado da seguinte maneira:

O coeficiente varia entre 0 e 1. Se seu valor for nulo, significa que as UF's são homogêneas entre si e que não existe diferencial salarial entre elas. Se seu valor

$$\rho = \frac{\sigma_{u0}^2}{\sigma_{\varepsilon}^2 + \sigma_{u0}^2} \quad (3)$$

for 1, toda variabilidade nos rendimentos deve-se as diferenças entre as UF's. Nesta situação, as características individuais do trabalhador em nada afetam os rendimentos salariais.

O modelo de nível 1 inclui as variáveis no nível do indivíduo que buscam explicar as variações nos rendimentos. A equação a seguir representa o modelo de nível 1, com o intercepto variando aleatoriamente entre as UF's:

$$\ln rendhr = \beta_{0j} + \beta_1 \text{sexo} + \beta_2 \text{cor} + \beta_3 \text{idade2} + \beta_4 \text{fund} + \beta_5 \text{medio} + \beta_6 \text{Sup} + \beta_7 \text{ind} + \beta_8 \text{serv} + \beta_9 \text{AdmPub} + \varepsilon_{ij} \quad (4)$$

$$\beta_{0j} = \beta_{00} + u_{0j}$$

$$\varepsilon_{ij} \sim N(0, \sigma^2_\varepsilon)$$

$$u_{0j} \sim N(0, \sigma^2_{u0})$$

O intercepto β_{0j} tem o índice j para indicar a existência de um parâmetro para cada UF. O intercepto é decomposto no valor médio global, β_{00} , e no efeito individual da UF, u_{0j} , que é o componente aleatório de nível 2 associado ao intercepto.

A equação a seguir representa o modelo de nível 2, que inclui, além das variáveis no nível do indivíduo, as variáveis no nível da UF. Como a unidade de nível 2 são as UF's brasileiras, possui portanto somente 27 observações, o que compromete a poder explicativo do modelo. Para não haver perda de graus de liberdade em um modelo já com tão poucas observações, optou-se por estimar um modelo para cada uma das variáveis de nível 2. Assim temos, respectivamente, o modelo com a variável taxa de urbanização (*urban*), o modelo com o nível de concentração industrial da UF (*ConcInd*), e o modelo com o índice de diversificação (*HH*):

$$\ln rendhr = \beta_{0j} + \beta_1 \text{sexo} + \beta_2 \text{cor} + \beta_3 \text{idade2} + \beta_4 \text{fund} + \beta_5 \text{medio} + \beta_6 \text{Sup} + \beta_7 \text{ind} + \beta_8 \text{serv} + \beta_9 \text{AdmPub} + \varepsilon_{ij} \quad (5)$$

$$\beta_{0j} = \beta_{00} + \beta_{10} \text{Urban} + u_{0j} \quad (6)$$

$$\beta_{0j} = \beta_{00} + \beta_{10} \text{ConcInd} + u_{0j} \quad (7)$$

$$\beta_{0j} = \beta_{00} + \beta_{10} \text{HH} + u_{0j} \quad (8)$$

$$\varepsilon_{ij} \sim N(0, \sigma^2_\varepsilon)$$

$$u_{0j} \sim N(0, \sigma^2_{u0})$$

4. Análise dos Resultados

Os resultados estimados para o modelo nulo, o modelo de nível 1 e o modelo de nível 2 estão na tabela 3. Esses resultados foram obtidos através do método de estimação de máxima verossimilhança restrita. Todos os resultados – com exceção dos dois últimos coeficientes do modelo de nível 2 – são significantes a 5%.

O modelo nulo mostra a partição da variância entre os níveis de análise; sem influência de variáveis explicativas. A variação observada nos salários é de 0,6862361; sendo que 3,8% dessa variação é dada pela variação entre a média dos salários nas UF's e, 96,2% da variação é dada pela variação dentro das UF's, ou seja, entre os trabalhadores. Esse resultado indica que as diferenças entre características pessoais dos trabalhadores têm um peso muito maior sobre a variação nos salários que as características da UF onde estes estão inseridos.

Resultados dos Modelos					
Variável	Modelo Nulo	Modelo de Nível 1	Modelo de Nível 2	Modelo de Nível 2	Modelo de Nível 2
Intercepto	1,43718 (0,000)	0,3699449 (0,000)	-0,2268331 (0,000)	0,3745026 (0,000)	0,3595147 (0,000)
Sexo		0,2642722 (0,000)	0,2642645 (0,000)	0,2642732 (0,000)	0,2642768 (0,000)
Cor		0,1478973 (0,000)	0,1476971 (0,000)	0,1478892 (0,000)	0,1478292 (0,000)
Idade2		0,000237 (0,000)	0,000237 (0,000)	0,000237 (0,000)	0,000237 (0,000)
Fund		0,2143674 (0,000)	0,2143629 (0,000)	0,2143634 (0,000)	0,2143634 (0,000)
Médio		0,5526023 (0,000)	0,5526165 (0,000)	0,5525961 (0,000)	0,5526045 (0,000)
Sup		1,446496 (0,000)	1,446524 (0,000)	1,446489 (0,000)	1,446503 (0,000)
Ind		0,0380169 (0,000)	0,0380123 (0,000)	0,0380199 (0,000)	0,038002 (0,000)
Serv		-0,0437101 (0,000)	-0,0436788 (0,000)	-0,0437069 (0,000)	-0,0436996 (0,000)
AdmPub		0,2354933 (0,000)	0,2356675 (0,000)	0,235494 (0,000)	0,2355348 (0,000)
Urban			0,0075304 (0,003)		
HH				-0,0404139 (0,772)	
ConcInd					0,2788524 (0,461)
Partição da Variância					
Trabalhadores	0,6603542	0,3718082	0,3718082	0,3718081	0,3718082
UF	0,0258819	0,0184823	0,014028	0,019177	0,0188169
ρ	3,8%	4,7%	3,6%	4,9%	4,8%

Fonte: Elaboração própria.
***P-valor entre parêntese

Ao introduzir as variáveis no nível do indivíduo (modelo de nível 1), a variação nos salários se reduz para 0,3902905, uma redução de 43% em relação o modelo nulo. Embora só tenham sido introduzidas variáveis no nível do indivíduo, a variabilidade entre as UF's diminuiu 28,6% e a variação entre os indivíduos diminuiu 44%. Um fato curioso, observado na partição da variância do modelo de nível 1, é que apesar da variância entre as UF's ter diminuído em termos absolutos, a participação da variação entre as UF's em relação à variação total aumentou, passando de 3,8% para 4,7%. A explicação para isto pode estar na alta variação das características individuais entre as UF's.

Em relação aos resultados dos coeficientes, verifica-se que indivíduos do sexo masculino, de cor branca, com mais experiência e nível educacional mais alto auferem rendimentos maiores. No que diz respeito à contribuição do nível educacional sobre o diferencial de rendimentos, os resultados mostram que os salários de indivíduos com nível médio são, em média, 158% maiores que os salários de indivíduos que possuem apenas nível fundamental. Esse fosso aumenta quando comparamos os salários dos níveis fundamental e médio com os salários dos indivíduos com nível superior. Em média, os indivíduos com nível superior auferem rendimentos maiores que indivíduos que possuem nível médio na ordem de 162% e, em relação aos indivíduos que possuem nível fundamental a diferença aumenta, em média, para 575%.

As variáveis *ind*, *serv* e *AdmPub*, apesar de não representarem uma característica inerente ao indivíduo, permitem captar informações relacionadas à estrutura de demanda por trabalho e ao diferencial salarial entre o setores. Pelos resultados, observa-se que, em média, indivíduos empregados no setor público auferem maiores rendimentos que indivíduos empregados na indústria que, por sua vez, auferem rendimentos maiores que indivíduos empregados no setor de serviços. Esses resultados podem refletir não só o padrão de distribuição salarial entre os setores produtivos, mas também, o padrão de altos salários do serviço público em relação ao setor privado.

Através do modelo de nível 2 buscou-se captar os efeitos de economias de aglomeração, existentes nas UF's, sobre as diferenças nos rendimentos. As estimativas para as variáveis de nível 1 não foram afetadas pela introdução das variáveis de nível 2. Devido ao fato da unidade de nível 2 possuir poucas observações – 27 Unidades da Federação – optou-se por estimar uma regressão para cada variável considerada afim de não perder graus de liberdade, o poderia comprometer a significância das estimativas. Das três variáveis de nível 2 que tentam captar efeitos de economias de aglomeração sobre os rendimentos, apenas a variável *taxa de urbanização* foi significativa. O sinal positivo do parâmetro estimado dessa variável corrobora com a literatura, indicando que, localidades com maior taxa de urbanização possuem também níveis de rendimento maiores, como colocado por Jacobs (1969).

As variáveis *concentração industrial* e *HH* apresentam parâmetros com valores e sinais que corroboram com a teoria, contudo os testes mostram que são insignificantes. O valor positivo do parâmetro da variável *ConcInd* indica que, em média, os salários são mais altos em UF's que apresentam nível de concentração industrial mais

alto em relação ao país. Isso está de acordo com a teoria que diz que a concentração industrial, na medida em que requer mão-de-obra especializada e gera economias de localização, gera também efeitos positivos sobre os salários. O *HH* é um índice que busca captar se é a diversidade ou a especialização produtiva das UF's que influenciam de forma positiva os rendimentos. Uma vez que esse índice varia entre 0, quando a indústria é perfeitamente diversa, e 2, quando a indústria é perfeitamente especializada, o valor negativo do parâmetro indica que quanto mais especializada em determinada atividade produtiva é a UF, em média, menores são os rendimentos obtidos pelos trabalhadores dessa UF. Esse resultado vai de encontro com as teorias das economias de localização, mas corrobora com Jacobs (1969), que defende que é a diversidade, e não a especialização do local, que induz inovação e aumento de produtividade, gerando assim, efeitos positivos sobre os salários. Apesar dos parâmetros para essas variáveis serem insignificantes, seus valores e sinais estão de acordo com os resultados obtidos por Galinari *et al* (2007) para as mesmas variáveis, embora usem recorte espacial e método de estimação diferentes.

Uma explicação para os parâmetros dessas duas variáveis terem dado insignificantes pode estar no recorte espacial que foi dado no estudo. Primeiro, o fato do recorte ser por UF faz com que a unidade de nível 2 do modelo tenha apenas 27 observações, que é considerada uma amostra pequena e pode comprometer a significância da estimativa. Segundo, índices como o de concentração industrial e de diversificação/especialização, como os utilizados aqui, são índices que captam melhor os efeitos de diversificação/especialização em recortes espaciais menores, como o foi o caso de Galinari *et al* (2007) que encontram resultados significativos para os parâmetros estimados dessas variáveis usando como recorte espacial as cidades médias. Ou seja, o recorte espacial amplo dado nesse estudo pode ter sido deficiente na hora de captar a variabilidade entre as localidades, o que compromete as estimativas.

5. Considerações Finais

Este trabalho buscou identificar variáveis que explicam os diferenciais de rendimentos no Brasil. Com base na literatura, sabe-se que os rendimentos do trabalho são influenciados tanto por características individuais dos trabalhadores, como por características do local onde o trabalhador está inserido. Diante disso, tornou-se conveniente para o propósito desse trabalho o uso da metodologia de modelos hierárquicos. O recorte espacial aqui utilizado consiste nas 27 Unidades da Federação (UF), buscando captar os efeitos das características de cada uma delas sobre os diferenciais de rendimento entre as mesmas.

A despeito dos inúmeros fatores associados a uma localidade, e que podem influenciar os diferenciais de rendimento, neste trabalho optou-se por considerar as variáveis associadas às economias de aglomeração. Segundo a literatura brevemente revisada nesse trabalho, as economias de aglomeração, divididas em economias de localização e de urbanização, geram efeitos positivos sobre os salários. Três variáveis

foram usadas como *proxies* para captar os efeitos das economias de aglomeração sobre os rendimentos: taxa de urbanização, concentração industrial e o índice *Hirschman-Herfindal*.

Os resultados para o modelo nulo mostraram que, apesar de pequena, existe variação no rendimento dada pela variação entre características das UF's. Os parâmetros do modelo no nível do indivíduo (nível 1) são todos significantes. Pelos resultados estimados da equação de nível 1, pode-se concluir que indivíduos do sexo masculino, de cor branca com níveis de escolaridade mais altos e empregados nos setores de indústria ou na administração pública auferem rendimentos maiores. Esses resultados não foram afetados pela introdução das variáveis de nível 2.

Em relação às variáveis de nível 2, apesar dos sinais dos parâmetros serem condizentes com a teoria, seus valores foram insignificantes, com exceção do parâmetro para a variável taxa de urbanização. Consideram-se aqui três possíveis explicações para esse resultado. Primeiro, os parâmetros podem não ter captado a influência das variáveis de nível 2 devido o número pequeno de observações (27 UF's). Segundo, entre os indivíduos, a variabilidade dos rendimentos é alta, mas entre os rendimentos médios das UF's a variabilidade é pequena, o que pode ser visto pela partição da variância. Terceiro, o problema pode estar no recorte espacial que foi dado no trabalho. Em geral, as variáveis de nível 2 aqui consideradas, são abordadas em trabalhos empíricos através de recortes espaciais menores, como municípios ou microrregiões. Assim, considerar um recorte espacial grande, como as UF's, pode não ser eficiente para captar efeitos de economias de aglomeração, uma vez que esses efeitos podem ser melhor medidos em recortes espaciais pequenos.

Sendo assim, não é pertinente concluir que o diferencial nos rendimentos é determinado somente por características ligadas ao indivíduo. Apesar das variáveis de nível 2 (com exceção da taxa de urbanização) terem sido insignificantes na explicação do diferencial de rendimentos, a literatura – e alguns trabalhos empíricos – defendem e mostram que as economias de aglomeração exercem efeitos significantes sobre os rendimentos.

Portanto, o exercício feito neste trabalho mostra a importância da abordagem de modelos hierárquicos para captar, em um único modelo, efeitos de variáveis medidas em níveis distintos. E sugere o uso da metodologia de modelos hierárquicos com recortes espaciais menores, como microrregiões, por exemplo, para captar os efeitos das economias de aglomeração sobre os diferenciais salariais.

Referências Bibliográficas

- Combes, P-P., Duranton, G., Gobillon, L. 2004. "Spatial wage disparities: sorting matters!" *Discussion paper*. n. 4240. London. Centre for Economic Policy Research (CEPR).
- Fontes, G. G. 2006. *Atributos Urbanos e Diferencias Regionais de Salário no Brasil, 1991 e 2000*. Dissertação de Mestrado. CEDEPLAR. Faculdade de Ciências Econômicas, Universidade Federal de Minas Gerais, Belo Horizonte.
- Galinari, R., Crocco, M. A., Lemos, M. B., Basques, M. F. D. 2007. "Efeitos das Economias de Aglomeração sobre os Salários Industriais: uma aplicação ao caso brasileiro." *Economia Contemporânea*. Rio de Janeiro. v.11, n. 3, dezembro.
- Galinari, R., Lemos, M. B. 2007. "Economias de Aglomeração no Brasil: evidências a partir da concentração industrial paulista." In: *Anais do XXXV Encontro Nacional de Economia - ANPEC*. Recife.
- Gujarati, D. N. 2006. *Econometria Básica*. Rio de Janeiro: Elsevier.
- HOOVER. E. M. 1936. *Location theory and the shoe and leather industries*. Cambridge, MA: Harvard University.
- IBGE. 2006. *Pesquisa Industrial Anual*. Rio de Janeiro: IBGE.
- Jacobs, J. 1969. *The Economy of Cities*. Nova York: Random House.
- Marshall, A. 1985 [1890]. *Princípios de Economia*. São Paulo: Nova Cultural.
- Menezes, W. F., Carrera-Fernandez, J., Dedecca, C. 2005. "Diferenciações Regionais de Rendimentos do Trabalho: Uma Análise das Regiões Metropolitanas de São Paulo e Salvador." *Estudos Econômicos*. São Paulo. v. 05, n. 2, abril/junho.
- Molho, I. 1992. "Local Pay Determination." *Journal of Economic Surveys*. v. 6, n. 2.
- Raudenbush, S. W. 2001. *Hierarchical Linear Models*. SAGE-USA.
- Reis, J. G. A., Barros, R. P. 1990. "Desigualdade salarial e distribuição de educação: a evolução das diferenças regionais no Brasil." *Pesquisa e Planejamento Econômico*. Rio de Janeiro. v. 20, n. 3, dezembro.
- Savedoff, W. 1990. "Os diferenciais de salários no Brasil: segmentação versus dinamismo da demanda." *Pesquisa e Planejamento Econômico*. Rio de Janeiro, v20, n. 3.
- Silveira Neto, R. M., Campelo, A. K. 2003. "O Perfil das Disparidades de Renda no Brasil: evidências a partir de regressões quantílicas para os anos de 1992 e 2001." *Anais do XXXI Encontro Nacional de Economia - ANPEC*. Porto Seguro.
- Topel, R. 1986. "Local Labor Markets." *Journal of Political Economy*, vol. 94, no. 3.