

Rendimento e desigualdade nas regiões metropolitanas brasileiras¹

Resumo: A equação minceriana tem sido a base de vários estudos referentes aos determinantes do rendimento. Essa equação relaciona a variável dependente rendimento-hora com várias variáveis explicativas tais como escolaridade, experiência, setor de atividade, tipo de ocupação ou profissão, modalidade de contrato e características individuais de gênero e cor, dentre outras. Nesse artigo, o objetivo é investigar os determinantes do rendimento em cada região metropolitana brasileira e discutir suas diferenças. Para tanto, são utilizados dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) para os anos de 2013, 2014 e 2015. Primeiramente, são estimadas equações de rendimento para cada região metropolitana através do modelo de dois estágios de Heckman (1979) para corrigir viés de seleção, juntamente com a decomposição de Oaxaca-Blinder para mensurar o peso da segmentação entre as regiões metropolitanas. Em seguida, é feita uma análise de desigualdade para cada região baseada no coeficiente de Gini. Os resultados revelam que há diferenças consideráveis em termos de magnitude entre as regiões metropolitanas, por exemplo, analisando a variável cor, é observado que a região metropolitana que tem a maior magnitude de diferença salarial, cerca de 14%, é Salvador-BA, e a menor diferença salarial, 6%, encontra-se na região de Porto Alegre-RS. Com relação a decomposição de Oaxaca-Blinder, verificou-se que na maioria das regiões metropolitanas as diferenças de remuneração ocorreram devido a segmentação de mercado, indicando que os atributos pessoais e capital humano não são os únicos a influenciar na remuneração, o que sugere a necessidade de políticas públicas focalizadas. No tocante aos resultados dos coeficientes de Gini, as estimativas indicam que a região metropolitana de Brasília-DF é a mais desigual e Fortaleza-CE é a menos desigual.

Palavras-chave: Equação de rendimento minceriana; Modelo de seleção de Heckman; Decomposição de Oaxaca-Blinder; Coeficiente de Gini.

Abstract: The Mincerian equation has been the basis of several studies on earning determinants. This equation is related to the dependent variable hourly earnings, with explanatory variables such as schooling, experience, activity sector, type of person or job, contract modality, and individual characteristics of gender and skin color, among others. The aim of this paper is to investigate the earning determinants for each Brazilian metropolitan region and their differences. To this end, data from the Brazilian National Household Sample Survey (PNAD) for the years 2013, 2014 and 2015, are used. First, Mincerian equations for each metropolitan region are estimated using the two-stage model by Heckman (1979) to correct sample selection bias, along with the Oaxaca-Blinder decomposition to measure the weight of segmentation between metropolitan regions. Then, an inequality analysis is done for each region based on the Gini coefficient. The results show that there are considerable differences in magnitude between metropolitan regions, for example,

Danúbia R. Cunha²

Leticia A. Lavoratto³

Helton Saulo⁴

Daniel T. G. N. Maciel⁵

² Doutoranda em Economia - Programa de Pós-Graduação em Economia - Universidade Católica de Brasília
Email: danubiarodrigues@hotmail.com

³ Doutoranda em Economia - Programa de Pós-Graduação em Economia - Universidade Católica de Brasília
Email: leticialavoratto@hotmail.com

⁴ Professor Assistente - Departamento de Estatística, Universidade de Brasília – UNB.
Email: heltonsauro@gmail.com

⁵ Doutor em economia e professor da FATEC-MT.
Email: daniel_mt_br@hotmail.com

¹ Os autores agradecem aos pareceristas pelas contribuições e sugestões, aos quais contribuíram para uma versão melhorada do artigo. O presente trabalho foi realizado com apoio da CAPES.

by analyzing the variable skin color, it was observed that the metropolitan region that has the largest magnitude of earnings change, about 14%, is Salvador-BA, and the lowest earning difference, 6%, is Porto Alegre-RS. With respect to the Oaxaca-Blinder decomposition, it was found that in most metropolitan regions the earning differences occurred due to market segmentation, indicating that personal attributes and human capital are not the only factors influencing remuneration, which suggest the need for focused public policies. Regarding the results of the Gini coefficients, the estimates indicate that the metropolitan region of Brasília-DF is the most unequal and Fortaleza-CE is the least unequal.

Keywords: *Mincerian earnings equation; Heckman's selection model; Oaxaca-Blinder decomposition; Gini Coefficient.*

Classificação JEL: C13; J31; R10.

1. Introdução

A existência de diferenciais de salários afeta o nível da desigualdade de rendimento e pode ser explicada por uma série de fatores (Ramos e Vieira, 2001): i) A diferença de salários entre indivíduos igualmente produtivos pode se dar de forma compensatória quando existe uma compensação pelos riscos existentes entre os postos de trabalho, ou seja, caso haja insalubridade, riscos de acidentes, entre outros. Essa diferença nos rendimentos está diretamente ligada a particularidades do posto de trabalho; ii) Outro fator determinante da desigualdade de renda está relacionado ao capital humano, isso ocorre quando a remuneração está diretamente relacionada aos atributos produtivos de cada indivíduo, com destaque para os níveis de escolaridade e a experiência do trabalhador, porém isso faz com que haja uma diferença significativa de renda entre os indivíduos que não possuem acesso a uma educação de qualidade; iii) A existência de preconceito e discriminação no mercado de trabalho também é um determinante de desigualdade de renda, pois remunera de forma diferente indivíduos com as mesmas habilidades e atributos produtivos baseado em sexo, cor, local de residência, entre outros; e iv) Há ainda as desigualdades salariais derivadas da segmentação existente no mercado do trabalho, ou seja, remuneração diferente para indivíduos igualmente produtivos, mas que estão em localizações distintas. No geral, são todos os fatores que diferenciam duas localidades que não estão diretamente ligadas aos atributos das pessoas, como mercados melhores e mais desenvolvidos, políticas públicas focalizadas, segurança, economia, renda e investimento que o local recebe, fatores atrativos locais, etc.

Os diferenciais de rendimentos estão relacionados ao fato do mercado de trabalho apresentar muitas imperfeições, entre elas à discriminação e a segmentação. Nos trabalhos de Barros e Mendonça (1993, 1996), os autores destacam que a causa de trabalhadores com a mesma produtividade, porém remunerados de forma diferente é devido a essas imperfeições.

Um aspecto relevante nos estudos relacionados à renda no Brasil, é a análise regional sob a ótica das regiões metropolitanas brasileiras. Recentemente, alguns autores têm se dedicado ao tema. Ribeiro (2015), por exemplo, analisa as desigualdades de renda nas regiões metropolitanas brasileiras levando em conta aspectos de localização na metrópole. Komatsu e Menezes-Filho (2015) avaliam o peso do crescimento real do salário mínimo, da ampliação do setor formal, das condições de oferta e demanda, e das características da força de trabalho, sobre a distribuição salarial nas regiões

metropolitanas brasileiras. Por outro lado, Almeida e Azzoni (2016) estimam os níveis de custo de vida das regiões no período 1996-2014, observando que Brasília, São Paulo e Rio de Janeiro apresentam os maiores custos de vida. Outros estudos associados podem ser encontrados em Silva e Fonseca-Neto (2014), Carneiro et al. (2016), Fochezatto (2016) e Montali e Lessa (2016).

Nesse contexto, o objetivo primário desse artigo é investigar os determinantes do rendimento salarial bem como analisar suas diferenças, nas seguintes regiões metropolitanas brasileiras: Belém-PA, Fortaleza-CE, Recife-PE, Salvador-BA, Belo Horizonte-MG, Rio de Janeiro-RJ, Curitiba-PR, Porto Alegre-RS, Brasília-DF e São Paulo-SP. A decomposição de Oaxaca-Blinder (Oaxaca, 1973) será utilizada para avaliar se há diferenças nas remunerações entre as regiões metropolitanas, e caso haja, será analisado se essa diferença nas remunerações se dá pelo fato da diferença nos atributos dos próprios trabalhadores ou se é devido à existência de segmentação de mercado. Para tal, os dados da Pesquisa Nacional por Amostragem de Domicílio (PNAD) de 2013, 2014 e 2015, e a equação de rendimentos minceriana, serão utilizados. As estimativas são realizadas através do modelo de correção de vies de seleção proposto por Heckman (1979), uma vez que os dados da PNAD de rendimentos-hora são observados apenas para pessoas empregadas, ou seja, um subconjunto da população sob estudo, provocando um vies de seleção amostral. O objetivo secundário desse artigo é estudar o nível de desigualdade em renda nas regiões metropolitanas brasileiras, bem como responder quais regiões têm maior e menor desigualdade. Nesse sentido, o coeficiente de Gini, indicador que revela o grau da desigualdade de renda, será estimado no contexto paramétrico e não paramétrico. Em termos paramétricos, serão utilizadas duas distribuições de probabilidade usualmente utilizadas no contexto de renda, a saber, as distribuições Singh-Maddala e Dagum; ver Hajargasht et al. (2012). De modo geral, destaca-se que os estudos realizados nesse trabalho são de certa forma inéditos quando levados em consideração apenas o fato de que todas as regiões brasileiras são estudadas e a análise é feita individualmente. Os resultados do artigo são contrastados com resultados já apresentados na literatura e a intenção é justamente corroborar tais resultados. Destaca-se, no entanto, que muitos desses resultados são bastante informativos e podem ajudar na formulação de políticas públicas focadas na especificidade de cada região metropolitana.

Além desta seção introdutória, o restante desse artigo é organizado da seguinte forma. A Seção 2 apresenta o referencial teórico apontando trabalhos anteriores relativos aos determinantes de rendimentos. A Seção 3 descreve o modelo de Heckman e a decomposição de Oaxaca-Blinder, além do coeficiente de Gini no contexto paramétrico e não paramétrico. A Seção 4 apresenta a base de dados e discute os resultados dos modelos estimados. Por fim, na Seção 5 serão abordadas as considerações finais, sintetizando os principais resultados encontrados.

2. Determinantes do rendimento no Brasil

O trabalho de Langoni (1973) foi uma das primeiras contribuições sobre a desigualdade de renda do trabalho brasileira. O autor estuda profundamente como a desigualdade de renda é gerada e exposta no mercado de trabalho. Através das informações do Censo, o autor tenta descrever o perfil da distribuição da renda dos indivíduos, fazendo uma comparação do ano de 1970 com o ano de 1960, para analisar o processo das desigualdades através dos diferenciais de salários de acordo com o trabalhador e seu posto de trabalho. Além de verificar as relações dos diferenciais de rendimentos com características como educação, idade, sexo, região de residência, entre outros. O autor concluiu que a educação é uma das principais causas da desigualdade

de renda, além de que um dos principais fatores do elevado grau de desigualdade salarial no Brasil se dá em virtude do seu sistema educacional ser excludente (Barros e Mendonça, 1995).

Com o intuito de analisar os rendimentos dos trabalhadores no mercado de trabalho, é utilizada a teoria do capital humano, desenvolvida em meados da década de 60 devido a preocupações com o crescimento econômico e uma melhor distribuição de renda, tendo como precursores Schultz (1961), Becker (1975) e Mincer (1974). De acordo com a teoria do capital humano, o aumento da renda está relacionado positivamente com o crescimento do nível educacional do indivíduo, pois a educação afeta diretamente a produtividade. Dessa forma, a educação é uma variável de grande importância para explicar os rendimentos (Barros e Mendonça, 1995).

De acordo com Sachsida et al. (2004), utiliza-se a equação minceriana⁶ para medir os retornos da educação, entre outras variáveis que também são determinantes da renda do trabalhador. Segundo o autor, o sucesso no mercado de trabalho está diretamente relacionado com a educação do trabalhador, pois as remunerações são diferentes para cada nível educacional.

Tanto a educação quanto as habilidades do trabalhador, passaram a ter maior importância com o surgimento da teoria do capital humano, que afirma que investimento em educação pode melhorar as aptidões e habilidades do trabalhador, tornando-o mais produtivo, e esses aumentos no conhecimento, qualificação e nas habilidades dos trabalhadores levam a mudanças na economia, nas questões sociais e também políticas, pois estas são impulsionadas por novas descobertas decorrentes desses novos conhecimentos, ou seja, podem influenciar positivamente no crescimento do país (Bento, 2006).

O investimento em capital humano gera benefícios e custos. O aumento da capacidade produtiva e das habilidades dos indivíduos se dá na escola e em programas de treinamento dentro das empresas, porém, é necessário investir em livros, materiais, mensalidades, além, de num primeiro momento haver perda de produção, pois o treinamento exige tempo tanto do trabalhador quanto do instrutor. Os indivíduos esperam reembolsar esses custos através de rendimento maiores decorrente dessa capacitação (Kassouf, 1997).

A partir disso, de acordo com Schultz (1961) e Becker (1975), um círculo virtuoso é criado, pois, se as despesas com a capacitação aumentarem a produtividade, esses investimentos terão retornos positivos. Esse aumento da produtividade gera aumento de salário e serve ainda como incentivo para novos investimentos em treinamentos e capacitação.

Segundo Oliveira (2001) a desigualdade educacional apresenta relação negativa com a renda per capita, ou seja, os trabalhadores que estudaram em boas escolas e possuem maior nível educacional são os que exibem maiores rendimentos no trabalho. Essa desigualdade na escolaridade tem uma relação negativa com a renda, aumentando a desigualdade de renda, pois os mais pobres muitas vezes não tem acesso a uma educação de qualidade.

Senna (1976) desenvolveu um dos primeiros trabalhos no Brasil tratando do retorno da educação nos rendimentos. O autor comparou os retornos da educação e da experiência de trabalhadores do setor público e privado, onde concluiu que para o setor privado, os trabalhadores que possuem maior escolaridade são os que possuem mais habilidade no trabalho, enquanto que no setor público, acontece o contrário, quanto maior o nível de educação menor o nível de habilidade.

Já no trabalho de Cavalieri e Fernandes (1998) é comparado os diferenciais de salários por gênero e cor entre as regiões metropolitanas brasileiras, utilizando as variáveis de gênero, cor, idade, anos de estudo, salário e jornada de trabalho através dos dados

⁶: O modelo salarial proposto por Mincer, que é amplamente utilizado para estimar retornos da educação na renda do trabalho, inclui além de escolaridade a variável experiência, supondo que a experiência é contínua e iniciada após a conclusão da escola.

da PNAD de 1989. O trabalho não levou em conta a possibilidade de viés de seleção por considerar apenas os assalariados do setor privado. Como resultados o trabalho apresentou que as pessoas de cor branca apresentam maiores rendimentos que os não brancos e também que os homens apresentam rendimentos superiores às mulheres.

No que tange aos métodos de estimação da equação minceriana, Sachsida et al. (2004) avaliaram a possível ocorrência de viés na equação de Mincer estimada através de MQO - Mínimos Quadrados Ordinários. Para os autores, o viés pode ocorrer devido a existência de salário de mercado, endogeneidade ou ainda por "viés de habilidades". E também como grande parte dos adultos que compõem a amostra não participa do mercado de trabalho, ou seja, são zerados na amostra, estimar a equação minceriana por MQO pode gerar estimadores tendenciosos (Kassouf, 1994). Dessa forma, para solucionar o problema do viés de seleção, utilizaremos o modelo de dois estágios de Heckman (1979).

Heckman (1979) propôs um modelo de duas equações, cujo objetivo era solucionar o problema de viés de seleção. Neste modelo, são incluídas características pessoais dos indivíduos, onde são avaliados os fatores determinantes da participação deste indivíduo no mercado de trabalho, para em seguida, realizar a estimação de uma segunda equação referente ao salário desses indivíduos (Pereira e De Oliveira, 2016).

De acordo com Heckman (1979) pode ocorrer viés de seleção amostral da seguinte forma: por autoseleção dos indivíduos ou das unidades de dados investigadas ou por seleção de amostras realizadas pelas decisões dos analistas dos dados, o que na prática assemelha-se a autoseleção. Dessa forma, a maneira mais adequada para as estimações das equações mincerianas é através do modelo desenvolvido por Heckman (1979), pois este faz a correção do viés de seleção apresentado pelos modelos de MQO (Pereira e De Oliveira, 2016). Dessa forma, este trabalho utiliza o modelo de Heckman (1979) para a estimação da equação minceriana e obtém os resultados esperados eliminando o problema de viés de seleção.

Como há o interesse em comparar os determinantes do rendimento em cada região metropolitana brasileira, esse trabalho utiliza a decomposição de Oaxaca-Blinder para analisar se as regiões apresentam remuneração diferenciada por haver segmentação no mercado de trabalho ou se é devido à diferença das características dos próprios trabalhadores.

No trabalho de De Carvalho et al. (2006), os autores utilizam os dados da PNAD de 2003 e decompõem o diferencial de salário por sexo e cor, estimando a equação de salários minceriana e corrigindo viés de seleção através do modelo de Heckman, em seguida fazem uso da decomposição de Oaxaca-Blinder considerando características produtivas e discriminação. Os resultados apontaram para a importância da correção de viés de seleção e ainda que a diferença de remuneração entre homens de cor branca e mulheres de cor preta e parda se dá pela discriminação.

Pereira e De Oliveira (2016) utilizaram do modelo de Heckman para correção de viés amostral da decomposição de Oaxaca-Blinder para analisar os diferenciais de rendimento entre homens e mulheres no estado do Rio Grande do Sul, através dados do Censo de 2010. O trabalho evidenciou a importância da educação na determinação dos rendimentos e por meio da decomposição de Oaxaca-Blinder constatou a existência de discriminação contra as mulheres, que possuem em média um ano a mais de estudo, porém esse diferencial não é repassado aos seus rendimentos.

Cirino e Lima (2011) analisaram os determinantes dos rendimentos no mercado de trabalho nacional, comparando entre homens e mulheres e as regiões metropolitanas de Belo Horizonte e Salvador.

Utilizaram o modelo de seleção amostral de Heckman e obtiveram como resultados que existe discriminação contra as mulheres no que se refere à idade e escolaridade,

e ainda que a região metropolitana de Belo Horizonte apresentou maiores impactos de escolaridade e idade comparada a região metropolitana de Salvador. Porém, em 2012, os autores acrescentaram ao trabalho a decomposição de Oaxaca-Blinder para analisar as diferenças de rendimentos entre as regiões e obtiveram como resultado que devido a uma maior concentração e aglomeração econômica, os rendimentos da região metropolitana de Belo Horizonte são maiores que os da região metropolitana de Salvador.

3. Aspectos metodológicos

Esta seção destina-se a apresentar o modelo de dois estágios proposto por Heckman (1979), a decomposição de Oaxaca-Blinder (Oaxaca, 1973) e a obtenção do coeficiente de Gini.

3.1 Modelo de Heckman

Seja $Y_1, \dots, Y_n$ variáveis aleatórias independentes seguindo uma distribuição normal com média μ e desvio-padrão σ , ou seja, $Y_i \sim N(\mu, \sigma)$. Considere que há k variáveis explicativas, $x_i = (1, x_{i1}, x_{i2}, \dots, x_{ik})^T$, associadas com Y_i , respectivamente, tal que $E[Y_i | X_i = x_i] = \mu_i = x_i^T \beta$, $i = 1, \dots, n$, em que $\beta = (\beta_0, \beta_1, \beta_2, \dots, \beta_k)^T$ é um vetor de $k + 1$ parâmetros desconhecidos, e x_i é a i -ésima linha da matriz X , cuja dimensão é $n \times k$. Assim, temos o seguinte modelo de regressão linear normal:

$$Y_i = x_i^T \beta + \varepsilon_i, \quad i = 1, \dots, n \quad (1)$$

em que $\varepsilon \sim N(0, \sigma)$. A equação minceriana de rendimento é tipicamente descrita pela Equação (1) em que Y_i é o logaritmo do rendimento-hora recebido pelo trabalhador, ou seja, rendimento-hora = rendimento mensal / (horas trabalhadas * 4,33), e X é uma matriz de variáveis explicativas tais como escolaridade, experiência, estado civil, dentre outras. O uso do modelo (1) no contexto da equação de rendimento não é adequada, uma vez que os rendimentos-hora são observados apenas para pessoas empregadas e não para aqueles que decidem ficar em casa ou não podem, por algum motivo, trabalhar. Ou seja, a variável rendimento-hora é observada somente em um subconjunto da população sob estudo, e desse modo, dizemos que tal variável está sujeita a um truncamento oculto e que possui viés de seleção amostral. Heckman (1976) propôs um modelo de seleção amostral baseado no seguinte sistema de equações

$$\begin{aligned} D_i^* &= z_i^T \gamma + v_i, \\ Y_i^* &= x_i^T \beta + \varepsilon_i, \end{aligned} \quad (2)$$

em que z_i é a i -ésima linha da matriz de variáveis Z , x_i é a i -ésima linha da matriz de variáveis X , γ e β são vetores de parâmetros desconhecidos, D_i^* é não observada, mas é maior ou menor que um $\xi \in \mathbb{R}$ fixo, Y_i^* é observada somente se $D_i^* > \xi$, e $(v_i, \varepsilon_i) \sim N_2(0, 0, \sigma_v = 1, \sigma_\varepsilon, \rho)$, ou seja, os termos de erros seguem uma distribuição normal bivariada com coeficiente de correlação ρ . Desse modo, a equação de participação é dada por $D_i = 1$ se $D_i^* > 0$ e 0 caso contrário, e a equação de intensidade é dada por $Y_i = Y_i^*$ se $D_i = 1$ e *n.a.* caso contrário. Tal modelo é conhecido como modelo de Heckman ou de seleção amostral, também denominado modelo *tobit* tipo 2. Os parâmetros do modelo podem ser estimados por máxima verossimilhança, no entanto Heckman (1979) propôs um procedimento denominado método de dois estágios que é utilizado nesse artigo para estimar a equação de rendimentos. Os passos são dados por:

Passo 1. Estimar via máxima verossimilhança o vetor γ de parâmetros no modelo *probit* padrão:

$$D_i^* = z_i^T \gamma + v_i, \text{ com } D_i = 1 \text{ se } D_i^* > 0 \text{ e } 0 \text{ caso contrário;}$$

Passo 2. Obter a estimativa do inverso da razão de Mills, dado por $\hat{\lambda}_i = \phi(z_i^T \hat{\gamma}) / \Phi(z_i^T \hat{\gamma})$, a partir dos resultados da estimação do modelo *probit* dado no Passo 1, em que ϕ e Φ denotam as funções densidade de probabilidade e de distribuição acumulada da normal padrão;

Passo 3. Estimar um modelo de regressão conforme (1), via mínimos quadrados, com a estimativa do inverso da razão de Mills ($\hat{\lambda}_i$ do Passo 2) como regressor adicional, ou seja, $Y_i = x_i^T \beta + \beta_\lambda \hat{\lambda}_i + \varepsilon_i$.

A vantagem do método de dois estágios encontra-se na sua simplicidade, uma vez que o ajuste tradicional via máxima verossimilhança requer algoritmos complicados. Como o interesse principal é na equação de intensidade, deve-se ter uma boa noção sobre as influências x_i . Porém, isto não é necessário para a equação de participação e para z_i . Por isso, faz sentido supor que $x_i \neq z_i$. Decisões de participação no mercado de trabalho não precisam estar interligadas.

3.2 Decomposição de Oaxaca-Blinder

A segmentação regional pode ser mensurada pela diferença nas remunerações entre duas regiões que não é explicada pelo modelo de Mincer, dessa forma segue-se a decomposição de Oaxaca-Blinder, juntamente com a extensão proposta por Fairlie (2005)⁷ para mensurar o valor dessa segmentação.

Definindo a região A e outra região qualquer B , pode-se comparar a diferença nas remunerações da seguinte forma⁸:

$$\Delta Y_g = \overline{Y_g^A} - \overline{Y_g^B} = \left[\frac{\sum_{i=1}^{n^A} \Theta(\sum_{j=0}^k \beta_j^A x_i^A)}{N^A} - \frac{\sum_{i=1}^{n^B} \Theta(\sum_{j=0}^k \beta_j^B x_i^B)}{N^B} \right].$$

Pode-se somar e subtrair sem perda:

$$\left[\frac{\sum_{i=1}^{n^A} \Theta(\sum_{j=0}^k \beta_j^B x_i^A)}{N^A} - \frac{\sum_{i=1}^{n^B} \Theta(\sum_{j=0}^k \beta_j^B x_i^B)}{N^B} \right].$$

Logo, temos:

$$\overline{Y_g^A} - \overline{Y_g^B} = \left[\frac{\sum_{i=1}^{n^A} \Theta(\sum_{j=0}^k \beta_j^A x_i^A)}{N^A} - \frac{\sum_{i=1}^{n^B} \Theta(\sum_{j=0}^k \beta_j^B x_i^B)}{N^B} \right] + \left[\frac{\sum_{i=1}^{n^A} \Theta(\sum_{j=0}^k \beta_j^B x_i^A)}{N^A} - \frac{\sum_{i=1}^{n^B} \Theta(\sum_{j=0}^k \beta_j^B x_i^B)}{N^B} \right]$$

Reorganizando:

$$\overline{Y_g^A} - \overline{Y_g^B} = \left[\frac{\sum_{i=1}^{n^A} \Theta(\sum_{j=0}^k \beta_j^A x_i^A)}{N^A} - \frac{\sum_{i=1}^{n^B} \Theta(\sum_{j=0}^k \beta_j^B x_i^B)}{N^B} \right] + \left[\frac{\sum_{i=1}^{n^A} \Theta(\sum_{j=0}^k \beta_j^B x_i^A)}{N^A} - \frac{\sum_{i=1}^{n^B} \Theta(\sum_{j=0}^k \beta_j^B x_i^B)}{N^B} \right]$$

Seja o efeito de parâmetros dado por

$$D = \left[\frac{\sum_{i=1}^{n^A} \Theta(\sum_{j=0}^k \beta_j^A x_i^A)}{N^A} - \frac{\sum_{i=1}^{n^B} \Theta(\sum_{j=0}^k \beta_j^B x_i^B)}{N^B} \right]$$

e o efeito das médias dado por

⁷: A decomposição de Oaxaca-Blinder inicialmente é feita para uma regressão linear comum, estimada por MQO, Fairlie (2005) estende essa abordagem permitindo que seja estimada para qualquer função. No caso do presente artigo será utilizado o modelo proposto por Fairlie (2005), pois as estimações seguem o modelo de Heckman.

⁸: Note que $Y_g^A = \ln(\text{rendimento-hora})$.

$$E = \left[\frac{\sum_{i=1}^{n^A} \Theta(\sum_{i=0}^k \beta_i^B x_i^A)}{N^A} - \frac{\sum_{i=1}^{n^B} \Theta(\sum_{i=0}^k \beta_i^B x_i^B)}{N^B} \right]$$

Desse modo

$$\Delta Y_g = \overline{Y_g^A} - \overline{Y_g^B} = D + E$$

Assim, a diferença nas remunerações entre as regiões A e B (ΔY_g) é decomposta em Efeito de Médias (E) e Efeito de Parâmetros (D). Sendo que o Efeito de Médias mensura qual seria a diferença nas remunerações se a região A e B tivessem a mesma valoração dos atributos pessoais, ou seja, qual seria a diferença de remuneração entre as regiões se não existisse segmentação. Já o Efeito de Parâmetros mensura a diferença nas remunerações entre as regiões devido à diferença na valoração dos atributos pessoais entre as duas regiões, com isso o D apresenta uma estimativa para o peso da segmentação regional entre duas regiões.

3.3 Coeficiente de Gini

O coeficiente de Gini é uma medida de desigualdade de renda que toma valores entre zero e um, em que zero corresponde à completa igualdade e um corresponde à completa desigualdade. Nesse trabalho, duas distribuições de probabilidade serão utilizadas para estimar o coeficiente de Gini, a saber, as distribuições Singh-Maddala e Dagum (Hajargasht et al., 2012). Estimativa não paramétrica também será utilizada para estimar o coeficiente de Gini.

A distribuição Singh-Maddala é um dos modelos mais utilizados para descrever rendimentos pessoais. Em particular, uma variável aleatória Z segue uma distribuição Singh-Maddala se sua função densidade de probabilidade e função de distribuição acumulada são dadas, respectivamente, por

$$f(z) = \frac{aqz^{a-1}}{b^a \left(1 + \left(\frac{z}{b}\right)^a\right)^{1+q}}, \quad F(z) = 1 - \left(1 + \left(\frac{z}{b}\right)^a\right)^{-q}, \quad z > 0,$$

em que $a > 0$ e $q > 0$ são parâmetros de forma e b é um parâmetro de escala. O coeficiente de Gini da distribuição Singh-Maddala é dado por

$$\text{Gini-SM: } 1 - \frac{\Gamma(q)\Gamma(2q-1/a)}{\Gamma(q-1/a)\Gamma(2q)},$$

em que Γ é a função matemática gama; ver Hajargasht et al. (2012).

A distribuição Dagum é menos conhecida que a Singh-Maddala, mas vem sendo também largamente utilizada para descrever a renda pessoal. Dado que Z segue uma distribuição Dagum, então a função densidade de probabilidade e função de distribuição acumulada associadas são dadas, respectivamente, por:

$$f(z) = \frac{aqz^{ap-1}}{b^{ap} \left(1 + \left(\frac{z}{b}\right)^a\right)^{p+1}}, \quad F(z) = 1 - \left(1 + \left(\frac{z}{b}\right)^a\right)^{-p}, \quad z > 0, \quad (5)$$

em que $a > 0$ e $p > 0$ são parâmetros de forma e b é um parâmetro de escala. O coeficiente de Gini baseado na distribuição Dagum é dado por:

$$\text{Gini-D} = \frac{\Gamma(p)\Gamma(2p-1/a)}{\Gamma(p-1/a)\Gamma(2q)} - 1, \quad (6)$$

em que Γ é a função matemática gama.

No contexto não paramétrico, isto é, sem assumir uma distribuição paramétrica para os dados, o coeficiente de Gini pode ser definido por:

$$\text{Gini-NP} = 1 - \sum_{i=0}^{n-1} (Y_{i+1} - Y_i)(Z_{i+1} - Z_i),$$

em que Y_i é a proporção acumulada da característica da população na amostra e Z_i é a proporção acumulada da renda.

4. Abordagem empírica

Nessa seção, é feita uma descrição da base de dados. Em seguida, as estimativas do modelo de Heckman para cada região metropolitana são discutidas juntamente com a decomposição de Oaxaca-Blinder. Finalmente, os resultados referentes ao coeficiente de Gini são apresentados. Todos os resultados foram obtidos por meio do software estatístico R [<https://www.r-project.org/>].

4.1 Base de dados

Nesse trabalho, a base de dados da PNAD, elaborada pelo Instituto Brasileiro de Geografia e Estatística (IBGE), é utilizada para os anos de 2013 a 2015⁹. Trabalha-se com uma amostra composta por indivíduos ocupados com idade entre 18 e 65 anos, com informações completas de seus rendimentos, e no total a amostra é composta por 172.750 observações. Os dados coletados são das dez principais regiões metropolitanas (RMs) do Brasil, a saber, Belém-PA, Fortaleza-CE, Recife-PE, Salvador-BA, Belo Horizonte-MG, Rio de Janeiro-RJ, Curitiba-PR, Porto Alegre-RS, Brasília-DF e São Paulo-SP.

Os valores nominais do rendimento são deflacionados pelo Índice Nacional de Preços ao Consumidor (INPC) para o ano de 2015. A seguir, é feita a divisão das variáveis utilizadas no modelo. Para tal, é feita a separação das variáveis que afetam a equação de participação, àquelas que afetam a equação de intensidade, e as variáveis que afetam as duas equações. Tal escolha é feita através da literatura pertinente; ver Borjas (2012).

As covariáveis comuns, que afetam a equação de participação e equação de intensidade, são dadas na Tabela 1. Já as covariáveis que afetam somente a equação de intensidade são apresentadas na Tabela 2. Por fim, as covariáveis que afetam apenas a equação de participação são apresentadas na Tabela 3.

4.2 Modelos de Heckman estimados

A Tabela 4 reporta os resultados das estimativas segundo o modelo de Heckman apresentado na Seção 2 para cada região metropolitana, com os respectivos valores-p. O teste de Wald é usado para testar as hipóteses de que $H_0: \theta = 0$ versus $H_1: \theta \neq 0$; e é definida por

$$W = \frac{[\hat{\theta} - \theta_0]}{\text{Erro Padrão}(\hat{\theta})} \sim N(0,1),$$

em que $\hat{\theta}$ e θ_0 são o estimador e o seu respectivo valor proposto sob H_0 , respectivamente.

⁹: A base de dados é organizada em um Modelo Empilhado - Polled, ou seja, é feito o empilhamento de todos os anos e não há variável de controle incluída.

A partir dos resultados apresentados na Tabela 4, é possível perceber que grande parte dos coeficientes estimados apresentaram significância estatística (5%). Nota-se que as variáveis sexo, idade, cor e educação, por exemplo, afetam positivamente o rendimento-hora do trabalhador em todas as regiões metropolitanas. Tais resultados são corroborados, por exemplo, por estudos que procuram investigar discriminação por gênero, em que homens têm rendimentos superiores quando comparado com as mulheres; ver Giuberti e Menezes-Filho (2005). É importante ressaltar que a inserção da mulher no mercado de trabalho vem sendo acompanhada, ao longo dos anos, por um elevado grau de discriminação, não só no que tange à qualidade das ocupações que têm sido criadas, mesmo com um aumento expressivo da taxa de atividade feminina nas últimas décadas, mas principalmente no que se refere à desigualdade da distribuição de renda entre homens e mulheres. Nos resultados da Tabela 4, pode-se verificar que, de fato, os homens ainda possuem maiores rendimentos em todas as regiões metropolitanas brasileiras. Note que, na região metropolitana de Belo Horizonte-MG, homens possuem rendimento-hora médio 25% maior que mulheres dessa região, porcentagem mais alta entre todas as regiões consideradas. Em Belém-PA, tem-se a menor taxa de diferenciação salarial, no entanto, os homens têm rendimento-hora médio 18% maior em relação as mulheres.

Tabela 1: Covariáveis que afetam a equação de participação e equação de intensidade.

Sexo	1 para homem e 0 para mulher
Cor	1 para brancos e 0 para não-brancos
Estado Civil	1 para indivíduos casados e 0 para não-casados
Idade e Idade ²	incluídas na regressão as variáveis "idade" e seu valor ao quadrado
Educação	1 tem como proxy os anos formais de estudo, que variam entre 0 a 16 anos
Categoria (alta, média-alta, média, média-baixa, baixa)	conjunto de variáveis binárias usadas para captar a categoria de ocupação, segmentadas segundo critérios sócioeconômicos e tendo a categoria Baixa como referência ¹⁰
Contrato de trabalho (com-carteira, sem-carteira, conta-própria, funcionário-público)	as variáveis "sem-carteira", "conta própria" e "funcionário público" são dummies que buscam captar o tipo de ocupação do indivíduo no mercado de trabalho, tendo como referência os indivíduos com carteira-referência os indivíduos com carteira
Ano (2013, 2014 e 2015)	são as dummies temporais para os anos analisados, tendo como referência o ano de 2013
Setor (agrícola, atividades industriais, construção, comércio, alimentação e outros, educação, saúde e serviços sociais)	são variáveis dummies que buscam captar efeitos de agrupamento por atividade dos indivíduos, tendo como referência os indivíduos que trabalham no setor público

Fonte: Elaboração própria.

Tabela 2: Covariáveis que afetam somente a equação de intensidade.

Sindicato	1 para indivíduos que eram associados a algum sindicato no mês de referência e 0 para aqueles que não eram associados
Previdência	1 para indivíduos que eram contribuintes de algum tipo de previdência no mês de referência e 0 para aqueles que não eram contribuintes
Experiência	número de anos no trabalho principal, que pode variar entre 0 a 56 anos

Fonte: Elaboração própria.

¹⁰: Para mais detalhes sobre o procedimento, ver Greene (2003).

Tabela 3: Covariáveis que afetam somente a equação de participação.

Chefe	tem o intuito de captar a condição do indivíduo na unidade domiciliar, assumindo valor 1 se o indivíduo de referência do domicílio é chefe da família e valor 0 caso contrário (não-chefe)
Menores	variável dummy usada para captar se há crianças menores de 10 anos no domicílio
Inatividade	1 para domicílios que haviam desempregados no mês de referência e 0 para aqueles domicílios em que todos estavam empregado

Fonte: Elaboração própria.

Tabela 4: Estimativas do modelo de Heckman do rendimento-hora para as regiões metropolitanas indicadas.

indicadas.	Belém	Fortaleza	Recife	Salvador	B, Horizonte	Rio Janeiro	São Paulo	Curitiba	Porto Alegre	Brasília
(Intercept)	0,793	0,652	0,759	0,813	0,594	0,968	0,736	0,826	0,824	0,463
Sexo	0,180	0,192	0,194	0,231	0,255	0,198	0,206	0,234	0,207	0,205
Idade	0,024	0,031	0,027	0,030	0,046	0,035	0,046	0,042	0,037	0,056
Idade ²	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001	-0,001
Cor	*0,027	0,095	0,088	0,146	0,104	0,131	0,092	0,075	0,063	0,106
Estado Civil	0,082	*-0,028	0,012	*-0,008	*-0,025	*0,034	*0,001	*0,075	*-0,019	*0,020
Educação	0,038	0,044	0,049	0,046	0,050	0,049	0,051	0,034	0,050	0,054
Alta	0,806	0,814	0,733	0,765	0,703	0,716	0,826	0,600	0,730	0,914
Média-alta	0,365	0,369	0,256	0,257	0,236	0,169	0,290	0,243	0,267	0,438
Média	0,066	0,153	*0,023	*0,025	*0,007	*-0,045	*0,031	0,102	0,053	0,115
Média-baixa	*-0,025	*0,017	-0,128	-0,116	-0,127	-0,196	-0,156	*-0,050	-0,086	*-0,054
Agrícola	-0,322	-0,673	-0,538	-0,632	-0,531	-0,464	-0,344	-0,711	-0,466	-0,418
Atividades industriais	-0,204	-0,314	-0,146	-0,075	-0,201	-0,139	-0,125	-0,115	-0,243	-0,411
Construção	-0,135	-0,070	*-0,037	-0,079	-0,100	*-0,043	*-0,045	*-0,077	-0,139	-0,281
Atividades terciárias	-0,201	-0,281	-0,207	-0,224	-0,271	-0,242	-0,154	-0,149	-0,273	-0,404
Educação	-0,129	-0,166	-0,150	-0,098	-0,182	-0,162	-0,091	-0,077	-0,164	-0,296
Outros serviços	-0,135	-0,191	-0,159	-0,169	-0,195	-0,134	-0,109	-0,108	-0,194	-0,296
Sem carteira	-0,063	-0,061	-0,078	-0,067	*0,022	-0,041	-0,037	*0,028	*0,005	*-0,006
Conta própria	-0,035	*0,016	*-0,004	-0,062	0,154	*0,012	0,101	0,075	0,101	0,140
Funcionário Público	0,422	0,284	0,272	0,286	0,251	0,273	*0,047	0,281	0,234	0,523
Sindicato	0,130	0,095	0,149	0,102	0,093	0,123	0,104	0,130	0,178	0,090
Previdência	0,282	0,258	0,350	0,418	0,391	0,219	0,413	0,215	0,288	0,368
Experiência	0,011	0,008	0,009	0,009	0,011	0,010	0,012	0,010	0,011	0,014
2014	*0,020	*0,029	*0,013	-0,074	*-0,003	*-0,019	*-0,010	0,059	*-0,006	*-0,034
2015	*0,045	*-0,002	*-0,051	-0,100	-0,047	*0,029	-0,046	*-0,042	*-0,005	*0,058
IMills	0,794	*-0,021	*0,128	*0,435	-1062,00	*0,151	-1017,00	1330,00	-0,871	*0,034

Fonte: Elaboração própria. Obs.: i) * denota que não é significativo a 5%; ii) Atividades terciárias: Comércio e reparação; Alojamento e alimentação; Transporte, armazenagem e comunicação; iii) Educação: Educação, saúde e serviços sociais; iv) Outros: outros serviços coletivos, sociais e pessoais; atividades mal definidas.

O mesmo diferencial salarial ocorre na variável "cor", ou seja, o rendimento-hora médio dos indivíduos de cor branca é maior do que dos de cor não-branca, mesmo após a inclusão de variáveis de controle para idade e anos de estudo. Um ponto a ser destacado é sobre a região metropolitana de Salvador-BA, onde tem-se que indivíduos brancos têm rendimento-hora médio cerca de 14% maior do que os não-brancos, maior taxa entre as regiões metropolitanas consideradas. Historicamente o estado da Bahia é composto em sua maioria por indivíduos negros, o que leva a crer que os baianos brancos são maioria em empregos de elite e os negros ocupam vagas com menor qualificação. Para Henrique (2001)[p. 20], considerando que a maioria da população negra concentra-se nos segmentos de menor renda per capita e, por outro lado, que os negros ricos são menos ricos que os brancos ricos, é possível desenhar dois Brasis: um "Brasil branco", mais rico e mais desigual, e um "Brasil negro", mais igual. O autor ainda acrescenta que o "Brasil branco" é cerca de 2,5 vezes mais rico que o "Brasil negro".

Para a variável "educação" é observado na Tabela 4 que um ano a mais de estudo afeta o rendimento-hora médio positivamente em todas as regiões metropolitanas. Tal resultado é corroborado pela literatura, em especial, Lau et al. (1993) encontram uma relação positiva entre rendimento e anos de estudo para o caso brasileiro, resultado que corrobora com a teoria do capital humano, a qual relaciona o nível educacional dos indivíduos como variável explicativa para aumentos na renda.

No que se refere às categorias de ocupação, os resultados da Tabela 4 mostram que

os ocupantes das categorias Alta, Média-alta e Média, têm rendimento-hora médio superior se comparado com aqueles que ocupam a categoria Baixa, porcentagem que vai diminuindo conforme deslocam para as categorias mais baixas. Por exemplo, trabalhadores da região metropolitana de São Paulo-SP, que ocupam a categoria Alta, têm um rendimento-hora médio em torno de 82% maior quando comparados com a categoria Baixa, mas os indivíduos que ocupam a categoria Média, têm apenas cerca de 3% acrescidos no rendimento-hora médio quando comparados com a categoria Baixa.

Outra análise importante a ser destacada são os setores de atividade: agrícola, atividades industriais, construção, comércio, alimentação e outros; educação, saúde e serviços sociais que têm como referência o setor público. O setor em que os indivíduos trabalham também influi no salário recebido.

A partir dos resultados da Tabela 4, tem-se que todos os setores citados, quando comparados com o setor público, possuem um menor nível de rendimento-hora médio em todas as regiões metropolitanas consideradas. A ocupação do indivíduo, obviamente, também influi no rendimento-hora recebido, que é corroborado na variável “tipo de contrato de trabalho”, que está separada em: com carteira, sem-carteira, conta-própria, funcionário-público, e tendo como referência os indivíduos com carteira. Aqui nota-se que, funcionários públicos têm rendimento-hora médio maior que todas as demais em todas as regiões metropolitanas, com destaque para a região metropolitana de Brasília-DF, onde é encontrado o maior percentual, cerca de 52%, ou seja, funcionários públicos desta região têm rendimento-hora médio 52% maior do que trabalhadores com carteira assinada, verifica-se com esse resultado que os salários do setor público dessa região são mais elevados se comparado com outras regiões metropolitanas.

Ainda nesta análise, tem-se que trabalhadores por conta própria têm rendimento-hora médio cerca de 14% maior do que trabalhadores com carteira. Trabalhadores por conta-própria têm rendimentos mais elevados em todas as regiões exceto em Belém, Recife e Salvador, onde trabalhadores por conta-própria têm rendimentos menores do que os trabalhadores com carteira. A análise destas duas variáveis, setor de atividade e tipo de contrato de trabalho, traz uma similaridade, pois, os resultados mostram que o setor que tem os maiores rendimentos é o público, e o tipo de contrato em que os trabalhadores percebem maiores proventos são os funcionários públicos. Por fim, é observado que trabalhadores sindicalizados, que contribuem para a previdência possuem rendimentos mais elevados em todas as regiões metropolitanas.

4.3 Resultados da decomposição de Oaxaca

A seguir são apresentadas as tabelas relacionadas à decomposição de Oaxaca-Blinder realizadas, que permitem separar a diferença nas remunerações¹¹ entre as regiões e verificar se essa diferença acontece devido à segmentação de mercado ou aos atributos individuais dos trabalhadores. Na Tabela 5 temos o Efeito de Parâmetro (D), que apresenta uma estimativa para a existência de segmentação entre os setores, onde as características produtivas são valorizadas de forma diferente.

O Efeito de Parâmetros (D) mensura a diferença das remunerações devido à diferença na valoração dos atributos produtivos, dessa forma esse efeito apresenta uma estimativa para o peso da segmentação entre as regiões metropolitanas comparadas duas a duas. As maiores diferenças nas remunerações devido à segmentação foram encontradas na comparação entre as regiões metropolitanas de Fortaleza com São Paulo (-0,59) e de Fortaleza com Brasília (-0,57). Esse resultado é

¹¹: As tabelas apresentarão os valores da decomposição, sem realizar a transformação para salário/hora, pois a variável dependente do nosso modelo de Mincer é $\ln(\text{rendimento-hora})$, logo quando referirmos à diferença nas remunerações estaremos avaliando em logaritmo, da seguinte forma:

$\ln(\text{rendimento por hora}^a) - \ln(\text{rendimento por hora}^b)$

esperado, visto que São Paulo e Brasília são economicamente mais desenvolvidas que Fortaleza e isso é demonstrado pelo peso negativo dessa segmentação.

Tabela 5: Efeito de Parâmetros - D.

Efeito D	Belém	Fortaleza	Recife	Salvador	B. Horizonte	Rio Janeiro	São Paulo	Curitiba	Porto Alegre	Brasília
Belém	0,00	0,13	0,05	-0,13	-0,21	-0,16	-0,44	-0,17	-0,18	-0,42
Fortaleza	-0,12	0,00	-0,08	-0,24	-0,34	-0,29	-0,59	-0,34	-0,35	-0,57
Recife	-0,05	0,09	0,00	-0,16	-0,25	-0,22	-0,50	-0,24	-0,26	-0,49
Salvador	0,10	0,25	0,15	0,00	-0,10	-0,07	-0,35	-0,11	-0,13	-0,34
B. Horizonte	0,24	0,36	0,26	0,11	0,00	0,04	-0,25	0,00	-0,02	-0,23
Rio Janeiro	0,18	0,31	0,22	0,07	-0,03	0,00	-0,28	-0,04	-0,06	-0,26
São Paulo	0,45	0,58	0,49	0,33	0,23	0,28	0,00	0,25	0,23	-0,01
Curitiba	0,22	0,36	0,26	0,09	-0,00	0,03	-0,26	0,00	-0,01	-0,27
Porto Alegre	0,23	0,35	0,26	0,09	-0,00	0,05	-0,24	0,02	0,00	-0,23
Brasília	0,40	0,50	0,43	0,27	0,16	0,22	-0,08	0,17	0,15	0,00

Fonte: Elaboração própria.

Na Tabela 6 abaixo temos o Efeito de Médias (E), que representa a diferença na remuneração média dos trabalhadores devido a diferença de capital humano, tais como a experiência, a educação, pois alteram a produtividade do trabalhador.

O efeito de médias (E) mensura quanto da diferença dos rendimentos é explicada pela diferença nas características produtivas, dessa forma esse efeito apresenta uma estimativa da diferença na remuneração entre as regiões metropolitanas que é devido a diferença de capital social, comparando duas a duas.

As maiores diferenças nas remunerações foram encontradas entre as regiões metropolitanas de Brasília com Fortaleza (0,33) e Brasília com Belo Horizonte (0,23). Isso é devido em grande parte a diferença no nível educacional médio e na maior parcela de servidores públicos encontrado em Brasília.

Tabela 6: Efeito de Médias - E.

Efeito E	Belém	Fortaleza	Recife	Salvador	B. Horizonte	Rio Janeiro	São Paulo	Curitiba	Porto Alegre	Brasília
Belém	0,00	0,07	-0,03	-0,02	-0,03	-0,09	-0,08	-0,06	-0,05	-0,20
Fortaleza	-0,08	0,00	-0,10	-0,10	-0,09	-0,15	-0,13	-0,09	-0,07	-0,25
Recife	0,02	0,08	0,00	-0,01	-0,00	-0,05	-0,04	-0,01	0,01	-0,15
Salvador	0,04	0,09	0,01	0,00	0,01	-0,03	-0,02	0,03	0,04	-0,14
B. Horizonte	-0,00	0,07	-0,01	-0,02	0,00	-0,05	-0,04	0,00	0,02	-0,16
Rio Janeiro	0,06	0,13	0,05	0,03	0,04	0,00	0,01	0,05	0,07	-0,12
Sao Paulo	0,06	0,14	0,05	0,04	0,05	-0,00	0,00	0,04	0,07	-0,10
Curitiba	0,00	0,06	-0,00	-0,01	-0,00	-0,04	-0,03	0,00	0,01	-0,13
Porto Alegre	-0,00	0,08	-0,01	-0,01	-0,00	-0,07	-0,05	-0,02	0,00	-0,16
Brasília	0,22	0,33	0,21	0,21	0,23	0,16	0,18	0,22	0,24	0,00

Fonte: Elaboração própria.

Na Tabela 7 é apresentado o Efeito Total, que mede a diferença na remuneração entre duas regiões metropolitanas como sendo a soma do Efeito de Parâmetros (D) e do Efeito de Médias (E), conforme demonstrado na metodologia.

Essa diferença representa a diferença total nas remunerações ($\overline{Y_g^A} - \overline{Y_g^B}$), conforme já era esperado as maiores diferenças foram encontradas para as regiões metropolitanas de Brasília com Fortaleza (0,82) e de São Paulo com Fortaleza (0,71), pois existe segmentação e maior capital humano para Brasília e São Paulo do que para Fortaleza.

Fato interessante ocorre quando o Efeito de Parâmetros (D) que mensura a segmentação supera o Efeito de Médias (E), pois significa que apesar deles terem capital social semelhante ou superior acaba ocorrendo que a segmentação é mais forte e isso faz com que a região de menor capital social acabe tendo maiores remunerações. Isso ocorreu entre as regiões metropolitas de Belém com Recife, Curitiba com Salvador, Porto

Alegre com Belém, Porto Alegre com Recife e Porto Alegre com Salvador.

A Tabela 8 demonstra o peso da segmentação na diferença de remuneração entre duas regiões (Efeito Total). Valores superiores a 100% ou inferiores a -100% indicam que o Efeito de Parâmetros (segmentação) é maior que o Efeito Total. Na maioria dos casos, observamos que o peso da segmentação é superior a 50%, isso é um indicativo de que os atributos pessoais e capital humano não são os únicos a influenciar na remuneração, o que sugere a necessidade de políticas públicas focalizadas.

Tabela 7: Efeito Total.

Efeito Total	Belém	Fortaleza	Recife	Salvador	B. Horizonte	Rio Janeiro	São Paulo	Curitiba	Porto Alegre	Brasília
Belém	0,00	0,20	0,02	-0,15	-0,23	-0,24	-0,52	-0,23	-0,23	-0,62
Fortaleza	-0,20	0,00	-0,17	-0,34	-0,43	-0,44	-0,71	-0,43	-0,42	-0,82
Recife	-0,02	0,17	0,00	-0,17	-0,26	-0,27	-0,54	-0,25	-0,25	-0,65
Salvador	0,15	0,34	0,17	0,00	-0,09	-0,10	-0,37	-0,08	-0,08	-0,48
B. Horizonte	0,23	0,43	0,26	0,09	0,00	-0,01	-0,28	0,00	0,01	-0,39
Rio Janeiro	0,24	0,44	0,27	0,10	0,01	0,00	-0,28	0,01	0,01	-0,38
Sao Paulo	0,52	0,71	0,54	0,37	0,28	0,28	0,00	0,29	0,29	-0,11
Curitiba	0,23	0,43	0,25	0,08	-0,00	-0,01	-0,29	0,00	0,00	-0,40
Porto Alegre	0,23	0,42	0,25	0,08	-0,01	-0,01	-0,29	-0,00	0,00	-0,40
Brasília	0,62	0,82	0,65	0,48	0,39	0,38	0,11	0,40	0,40	0,00

Fonte: Elaboração própria.

Tabela 8: Percentual de Segmentação no Efeito Total.

Segmentação	Belém	Fortaleza	Recife	Salvador	B. Horizonte	Rio Janeiro	São Paulo	Curitiba	Porto Alegre	Brasília
Belém	X	0,66	2,23	0,90	0,89	0,65	0,85	0,76	0,80	0,68
Fortaleza	0,60	X	0,45	0,70	0,79	0,67	0,82	0,79	0,83	0,69
Recife	1,96	0,52	X	0,97	0,99	0,82	0,92	0,96	1,05	0,76
Salvador	0,71	0,73	0,91	X	1,12	0,72	0,94	1,30	1,55	0,71
B. Horizonte	1,01	0,83	1,02	1,25	X	-5,11	0,87	0,42	-2,73	0,60
Rio Janeiro	0,75	0,70	0,83	0,70	-3,74	X	1,02	-2,74	-3,98	0,69
São Paulo	0,87	0,81	0,91	0,90	0,82	1,02	X	0,86	0,78	0,10
Curitiba	0,99	0,85	1,02	1,13	0,75	-2,11	0,89	X	-8,68	0,68
Porto Alegre	1,00	0,82	1,04	1,13	0,30	-3,39	0,82	-13,07	X	0,59
Brasília	0,65	0,60	0,67	0,56	0,42	0,58	-0,74	0,44	0,38	X

Fonte: Elaboração própria.

4.4 Coeficientes de Gini estimados

As Tabelas 9 e 10 apresentam as estimativas obtidas pelo método da máxima verossimilhança dos parâmetros das distribuições Singh-Maddala e Dagum, bem como os respectivos valores dos critérios de informação Akaike (AIC) e Bayesiano (BIC) para cada região metropolitana brasileira. Nessas tabelas são apresentadas ainda os coeficientes de Gini estimados baseados nas distribuições Singh-Maddala e Dagum e na abordagem não paramétrica.

Os resultados das Tabelas 9 e 10 mostram que os ajustes referentes à distribuição Dagum são melhores do que os baseados distribuição na Singh-Maddala, uma vez que ela possui valores associados de AIC e BIC menores. Por outro lado, os valores das estimativas dos coeficientes de Gini baseados nessas duas distribuições apresentam valores próximos, e por conseguinte próximos também aos valores estimados baseados na abordagem não paramétrica. Os resultados obtidos mostram que as regiões metropolitanas com maior desigualdade em renda foram Brasília [0,65 (Singh-Maddala), 0,65 (Dagum), 0,73 (Não paramétrico)], São Paulo [0,67 (Singh-Maddala), 0,63 (Dagum), 0,66 (Não paramétrico)] e Rio de Janeiro [0,65 (Singh-Maddala), 0,63 (Dagum), 0,72 (Não paramétrico)]. As menores desigualdades foram observadas em Fortaleza [0,49 (Singh-Maddala), 0,47 (Dagum), 0,51 (Não paramétrico)], Belo Horizonte [0,49 (Singh-Maddala), 0,48 (Dagum), 0,50 (Não paramétrico)] e Curitiba [0,50 (Singh-Maddala), 0,49 (Dagum), 0,53 (Não paramétrico)].

Tabela 9: Estimativas dos parâmetros do modelo Singh-Maddala (SM), AIC, BIC e coeficientes de Gini estimados baseados nos modelos SM e não paramétrico (NP).

	$\hat{a}$	$\hat{b}$	$\hat{q}$	AIC	BIC	Gini-SM	Gini-NP
Belém	0,9	68,74	6,68	82029,99	82052,27	0,58	0,56
Fortaleza	1,43	11,64	2,09	91153,57	91176,28	0,49	0,51
Recife	0,82	121,98	8,86	107105	107128,1	0,6	0,56
Salvador	1,39	11,5	1,68	111989,8	112012,9	0,55	0,59
Belo Horizonte	1,63	12,51	1,54	125565,7	125589	0,49	0,5
Rio Janeiro	0,96	29,3	2,39	183483	183507,2	0,65	0,72
São Paulo	0,7	316,49	9,21	203493,7	203518,3	0,67	0,66
Curitiba	1,46	19	1,87	76978,96	77000,63	0,5	0,53
Porto Alegre	1,16	27,78	2,86	168795,5	168819,6	0,54	0,56
Brasília	1	34,22	2,18	105004,7	105027	0,65	0,73

Fonte: Elaboração própria.

Tabela 10: Estimativas dos parâmetros do modelo Dagum (D), AIC, BIC e coeficientes de Gini estimados baseados nos modelos D e não paramétrico (NP).

	$\hat{a}$	$\hat{b}$	$\hat{q}$	AIC	BIC	Gini-D	Gini-NP
Belém	2,61	14,53	0,27	78982,32	79004,59	0,55	0,56
Fortaleza	2,6	9,9	0,43	89267,09	89289,79	0,47	0,51
Recife	2,53	15,33	0,26	102950	102973,1	0,57	0,56
Salvador	2,23	11,44	0,49	110303,6	110326,7	0,52	0,59
Belo Horizonte	2,4	12,8	0,54	124306,9	124330,2	0,48	0,5
Rio Janeiro	1,89	19,89	0,4	179655,4	179679,6	0,63	0,72
São Paulo	2,21	26,7	0,25	196379,3	196403,9	0,63	0,66
Curitiba	2,45	17,48	0,47	75945,5	75967,17	0,49	0,53
Porto Alegre	2,39	17,13	0,39	165300,8	165324,9	0,53	0,56
Brasília	1,74	25,58	0,46	103932,9	103955,2	0,65	0,73

Fonte: Elaboração própria.

5 Considerações finais

Nesse artigo, a equação minceriana de rendimentos foi estudada para as seguintes regiões metropolitanas brasileiras: Belém-PA, Fortaleza-CE, Recife-PE, Salvador-BA, Belo Horizonte-MG, Rio de Janeiro-RJ, Curitiba-PR, Porto Alegre-RS, Brasília-DF e São Paulo-SP. As estimações foram realizadas utilizando o modelo de dois estágios de Heckman considerando os dados da PNAD para os anos de 2013, 2014 e 2015 juntamente com a decomposição de Oaxaca-Blinder para mensurar o peso da segmentação entre as regiões metropolitanas e ainda uma análise de desigualdade baseada no coeficiente de Gini. Os resultados encontrados corroboram com outros estudos que investigaram as diferenças nos rendimentos, como o trabalho de Giuberti e Menezes-Filho (2005) que investiga a discriminação por gênero, onde os homens apresentam rendimentos superiores às mulheres. Henrique (2001) que compara a diferenças nos rendimentos devido à cor, onde segundo o autor existem dois Brasis, um "Brasil branco", mais rico e mais desigual, e um "Brasil negro", onde a população negra se concentra em setores de menor remuneração. E ainda Lau et al. (1993) que encontram uma relação positiva entre rendimento e anos de estudo para o caso brasileiro.

No que se refere à diferença nos rendimentos entre homens e mulheres, o trabalho verificou que os homens apresentam as maiores remunerações se comparados às mulheres em todas as regiões metropolitanas analisadas, em especial na região metropolitana de Belo Horizonte, onde o rendimento-hora médio dos homens é 25% superior ao das mulheres. Já para a variável cor, o rendimento-hora médio encontrados dos indivíduos de cor branca é superior ao dos de cor não branca, com destaque para a região metropolitana de Salvador-BA onde o rendimento-hora médio dos brancos é 14% a mais do que os não brancos, apesar do estado ser composto em sua maioria por

negros. Quanto a variável educação observou-se que um ano a mais de estudo afeta positivamente o rendimento-hora de todas as regiões metropolitanas brasileiras. A decomposição de Oaxaca-Blinder confirma os resultados apresentados por Azzoni e Servo (2001) onde apontam que as regiões metropolitanas de Brasília e São Paulo são as que possuem maiores rendimentos, enquanto que as regiões metropolitanas de Recife e Fortaleza apresentam os menores rendimentos. Os resultados evidenciaram que as regiões metropolitanas de Fortaleza com São Paulo (-0,59) e de Fortaleza com Brasília (-0,57) são as regiões com maiores diferenças nas remunerações devido à segmentação, onde as características produtivas são valorizadas de forma diferente. Resultado esperado, visto que São Paulo e Brasília são economicamente mais desenvolvidos que Fortaleza.

Com relação à remuneração média dos trabalhadores devido à diferença de capital humano, tais como a experiência, a educação, as regiões metropolitanas de Brasília com Fortaleza (0,33) e Brasília com Belo Horizonte (0,23) apresentaram as maiores diferenças nas remunerações. Grande parte dessa diferença é devido à diferença no nível educacional médio e também por Brasília possuir a maior parcela de servidores públicos. De acordo com o Efeito Total e conforme esperado, as maiores diferenças foram encontradas para as regiões metropolitanas de Brasília com Fortaleza (0,82) e de São Paulo com Fortaleza (0,71), pois Brasília e São Paulo apresentam maiores índices de capital humano e segmentação do Fortaleza. As regiões metropolitanas de Belém com Recife, Curitiba com Salvador, Porto Alegre com Belém, Porto Alegre com Recife e Porto Alegre com Salvador apresentam um fato interessante, pois apesar de terem capital social semelhante ou superior, ocorre que a segmentação é mais forte e isso faz com que a região de menor capital social tenha maiores remunerações.

Com relação ao coeficiente de Gini os resultados obtidos mostram que as regiões metropolitanas com maior desigualdade em renda foram Brasília, São Paulo e Rio de Janeiro, enquanto que as menores desigualdades foram observadas em Fortaleza, Belo Horizonte e Curitiba. Os resultados evidenciaram que na maioria das regiões metropolitanas analisadas o peso da segmentação é superior a 50%, indicando que os atributos pessoais e o capital humano não são os únicos a influenciar a remuneração,

Referências

Almeida, A. N. e Azzoni, C. R. (2016). Custo de vida comparativo das regiões metropolitanas brasileiras: 1996-2014. *Estudos Econômicos*, 46:253–276.

Azzoni, C. e Servo, L. (2001). Education, cost of living and regional wage inequality in Brazil. *Papers in Regional Science*.

Barbosa-Filho, F. H., Pessoa, S. A., e Veloso, F. A. (2010). Evolução da produtividade total dos fatores na economia brasileira com ênfase no capital humano-1992-2007. *Revista Brasileira de Economia*, 64:91–113.

Barros, R. P. e Mendonça, R. S. P. (1993). Geração e reprodução da desigualdade de renda no Brasil. *perspectivas da economia brasileira 1994*. IPEA (Texto para discussão), 454.

Barros, R. P. e Mendonça, R. S. P. (1995). Os determinantes da desigualdade no Brasil. IPEA (Texto para discussão), 377.

Barros, R. P. e Mendonça, R. S. P. (1996). Os determinantes da desigualdade no Brasil.

perspectivas da economia brasileira. IPEA, pages 421–474.

Becker, G. (1975). *Human Capital: A Theoretical and Empirical Analysis*. Columbia University Press, New York.

Bento, S. D. (2006). A importância do capital humano dentro das organizações. In *Anais do XIII Simpósio de Engenharia de Produção*, Bauru, SP. SIMPEP.

Borjas, G. J. (2005). *Labor Economics*. McGraw Hill, United States.

Borjas, G. J. (2012). *Economia do trabalho*. Artmed, São Paulo, 5 edition.

Carneiro, D. M., Bagolin, I. P., e Tai, S. H. T. (2016). Determinantes da pobreza nas regiões metropolitanas do Brasil no período de 1995 a 2009. *Nova Economia*, 26:69–96.

Cavaleri, C. e Fernandes, R. (1998). Diferenciais de salários por gênero e cor: uma comparação entre as regiões metropolitanas brasileiras. *Revista de Economia Política*, 18:69.

Cirino, J. F. e Lima, J. F. (2011). Determinantes dos rendimentos no mercado de trabalho nacional: Uma comparação entre os gêneros e entre as regiões metropolitanas de Belo Horizonte e Salvador. *Revista Nexos Econômicos*, 5:107–136.

De Carvalho, A. P., Néri, M., e Do Nascimento Silva, D. B. (2006). Diferenciais de salários por raça e gênero no Brasil: aplicação dos procedimentos de Oaxaca e Heckman em pesquisas amostrais complexas. Instituto Brasileiro de Geografia e Estatística, Mimeographed document.

Fairlie, R.W. (2005). An extension of the Blinder-Oaxaca decomposition technique to logit and probit models. *Journal of Economic and Social Measurement*, 30:305–316.

Fochezatto, A. (2016). Estrutura setorial e distribuição de ocupações nas regiões brasileiras, 2006-2011. *EURE*, 42:237–262.

Giuberti, A. C. e Menezes-Filho, N. (2005). Discriminação de rendimentos por gênero: uma comparação entre o Brasil e os Estados Unidos. *Economia Aplicada*, 9:369–384.

Hajargasht, G., Griffiths, W. E., Brice, J., Rao, D. P., e Chotikapanich, D. (2012). Inference for income distributions using grouped data. *Journal of Business & Economic Statistics*, 30(4):563–575.

Heckman, J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement*, 5:475–492.

Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica*,

47:153–161.

Henrique, R. (2001). Desigualdade racial no Brasil: evolução das condições de vida na década de 90. Texto para Discussão IPEA, 807.

Jannuzzi, P. d. M. (2001). Status socioeconômico das ocupações brasileiras: medidas aproximativas para 1980, 1991 e anos 90. *Revista Brasileira de Estatística*, 2:47–74.

Kassouf, A. (1994). The wage rate estimation using the Heckman procedure. *Revista de Econometria*, 14:89–107.

Kassouf, A. (1997). Retornos à escolaridade e ao treinamento nos setores urbano e rural do Brasil. *Revista de Economia e Sociologia Rural*, 35:59–76.

Komatsu, B. K. e Menezes-Filho (2015). Salário mínimo e desigualdade salarial: um estudo com densidades contrafactuais nas regiões metropolitanas brasileiras. *Pesquisa e Planejamento Econômico*, 45:365–406.

Langoni, C. (1973). Distribuição da renda e desenvolvimento econômico do Brasil. *Expressão e Cultura*, Rio de Janeiro.

Lau, L., Jamison, D., Liu, S., e Riukin, S. (1993). Education and economic growth some crosssectional evidence from Brazil. *Journal of Development Economics*, 41:45–70.

Mincer, J. (1974). Schooling, Experience, and Earnings. National Bureau of Economic Research, Inc.

Montali, L. e Lessa, L. H. (2016). Pobreza e mobilidade de renda nas regiões metropolitanas brasileiras. *Cadernos Metrópole*, 36:503–533.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14:693–709.

Oliveira, J. B. (2001). Renda per capita, desigualdades de renda e educacional, e participação política no Brasil. IPEA (Texto para discussão), 827.

Pereira, R. M. e De Oliveira, C. A. (2016). Os diferenciais de rendimento por gênero no Rio Grande do Sul: uma aplicação do modelo de Heckman e da decomposição de Oaxaca-Blinder. *Redes*, 21:148–173.

Pereira, R. S. e Zavala, A. (2012). Educação e rendimentos do trabalho no Brasil: desafios de um cenário de transformações. *Si Somos Americanos, Revista de Estudios Transfronterizos*, 12:203–227.

Ramos, L. e Vieira, M. L. (2001). Desigualdade de rendimentos no Brasil nas décadas de 80 e 90: evolução e principais determinantes. IPEA (Texto para Discussão), 803.

Ribeiro, M. G. (2015). Território e desigualdades de renda em regiões metropolitanas do Brasil. *DADOS*, 58:913–949.

Sachsida, A., Loureiro, P. R. A., e Mendonça, M. (2004). Um estudo sobre retorno em escolaridade no Brasil. *Revista Brasileira de Economia*, 58:249–265.

Schultz, T. W. (1961). Investment in human capital. *American Economic Review*, 51:1–17.

Senna, J. (1976). Escolaridade, experiência no trabalho e salários no Brasil. *Revista Brasileira de Economia*, 30:163–193.

Silva, F. J. F. e Fonseca-Neto, F. A. (2014). Efeitos da crise financeira de 2008 sobre o desemprego nas regiões metropolitanas brasileiras. *Nova Economia*, 24:265–278.