

**Área: Teoria Aplicada.**

**DIFERENCIAIS DE RENDIMENTOS ENTRE TRABALHADORES  
FORMAIS E INFORMAIS: UMA ANÁLISE PARA AS REGIÕES  
NORDESTE E SUDESTE**

**Daniel Lopes Ribeiro**

Graduando em Ciências Econômicas pela Universidade Federal do Ceará. Atualmente é assistente de pesquisa do Laboratório de Estudos da Pobreza – LEP-UFC.

Rua Coronel Estanislau Frota, s/n, Centro CEP 62.010-560, Sobral/CE.

Fone: (88) 3695-5703. Email: [danribeiro.ufc@gmail.com](mailto:danribeiro.ufc@gmail.com)

**Vitor Hugo Miro Couto Silva**

Professor dos cursos de Economia e Finanças da Universidade Federal do Ceará (UFC) – Campus de Sobral. Doutorando em Economia pelo CAEN/UFC. Pesquisador do Laboratório de Estudos da Pobreza – LEP-UFC.

Rua Coronel Estanislau Frota, s/n, Centro CEP 62.010-560, Sobral/CE.

Fone: (88) 3695-5703. Email: [vitormiro@gmail.br](mailto:vitormiro@gmail.br)

# DIFERENCIAIS DE RENDIMENTOS ENTRE TRABALHADORES FORMAIS E INFORMAIS: UMA ANÁLISE PARA AS REGIÕES NORDESTE E SUDESTE

## RESUMO

O presente estudo buscou analisar, controlando o viés de seletividade amostral, o diferencial de rendimentos entre trabalhadores formais e informais nas regiões Nordeste e Sudeste. Além disso, buscou-se captar como as características dos trabalhadores influenciam na escolha entre participar do mercado de trabalho, bem como na seleção setorial entre os setores formal e informal. Para tanto, a correção do viés de seleção foi feita baseada no método de Heckman em dois estágios (1979), assumindo a hipótese de correlação nula entre ambas as equações de seleção. Já para análise do diferencial se optou pela decomposição de Oaxaca e Blinder (1973) que detalha o diferencial de rendimentos em dois termos, *efeito preço* e *efeito dotação*. A base de dados utilizada para realização das estimativas foi a PNAD (2013), na qual os resultados obtidos sustentam a hipótese de que não há segmentação no mercado de trabalho das regiões selecionadas, haja vista que o principal determinante da diferença salarial em favor dos trabalhadores com carteira de trabalho assinada é explicada pela diferença nas características produtivas dos indivíduos.

**Palavras-Chaves:** Viés de seleção, Informalidade, Diferencial de rendimentos, Segmentação.

## ABSTRACT

This study investigates by controlling the bias of sample selectivity, the difference in income between formal and informal workers in the Northeast and Southeast. In addition, it sought to capture how the characteristics of workers influence the choice between participating in the labor market as well as the sectorial selection between the formal and informal sectors. Therefore, the correction of selection bias was made based on the Heckman method in two stages (1979), assuming the null hypothesis of correlation between both selection equations. As for differential analysis was decided by the decomposition of Oaxaca and Blinder (1973) detailing the income differential in two terms, price effect and effect allocation. The database used to perform the estimates was the National Household Survey (2013), in which the results support the hypothesis that there is segmentation in the work of selected regions market, given that the main determinant of the wage gap in favor of workers with a formal contract is explained by the difference in the productive characteristics of individuals.

**Key words:** Informality, Earnings gap, Segmentation, Sample selection bias.

**JEL codes:** J01, J3, D3.

# 1 INTRODUÇÃO

Apesar de amplamente conhecido como um dos países mais desiguais do mundo, o Brasil conseguiu diminuir vertiginosamente os níveis de desigualdades na última década, chegando em 2005 ao menor nível em 30 anos, segundo IPEA (2006)<sup>1</sup>. Contudo, não obstante aos resultados positivos alcançados, o país ainda possui uma das piores distribuições de renda. Portanto, entender que fatores influenciam essa disparidade é uma das questões mais importantes para a pauta de políticas públicas.

Segundo Ulyssea (2007), grande parte da desigualdade de rendimentos pode ser explicada pelas diferenças entre grupos no mercado de trabalho. Por exemplo, o autor mostra que os diferenciais de rendimentos entre trabalhadores formais e informais explicam cerca de 20% a desigualdade de renda observada ao longo da última década, não havendo muita variabilidade.

Ademais, podemos observar que os estudos acerca da informalidade passaram a se intensificar a partir do final da década de oitenta e início da década de noventa, motivados por mudanças intensas que ocorreram na economia e no mercado de trabalho e que segundo Ramos (2006), influenciaram para o aumento da informalidade no Brasil nesse período. Desde então, os trabalhos passaram a se dedicar a entender como a informalidade afeta a economia e o bem-estar dos trabalhadores, por exemplo, Barros et al. (1990) investigam como se comporta a mobilidade dos trabalhadores entre mercado de trabalho formal e informal, verificando a duração dos trabalhadores no emprego informal. Fernandes (1996) procura captar quais fatores influenciam a escolha ocupacional do trabalhador, a partir de um modelo de probabilidade não linear.

Embora se tenha assistido um aumento expressivo da informalidade a partir da década de noventa, nota-se que nos últimos anos houve uma reversão dessa tendência, aonde o número de trabalhadores informais vem diminuindo de forma abrupta. Conforme aponta Moura e Barbosa Filho (2014), a taxa de informalidade no Brasil caiu aproximadamente 11,7% entre 2002 e 2012. Esta queda ocorreu de forma sistêmica quando o percentual de trabalhadores assalariados sem carteira de trabalho assinada passou de 43,7% em 2002 para 39,3% em 2007 e posteriormente para 32,5% em 2012 [MOURA E BARBOSA FILHO, 2014].

Dado isso, o objetivo deste trabalho será o de investigar como se comporta os diferenciais de rendimentos entre trabalhadores formalizados e os que não possuem carteira de trabalho assinada, em duas regiões selecionadas (Nordeste e Sudeste) e, para tanto, lança mão do método de decomposição de Oaxaca-Blinder (BLINDER, 1973 e OAXACA, 1973), que detalha o diferencial de rendimentos em dois termos, *efeito preço* e *efeito dotação*. A estimação de equações de rendimentos foi realizada controlando o viés de seletividade amostral, que também permite avaliar como as características dos trabalhadores influenciam na escolha entre participar do mercado de trabalho, bem como na seleção setorial entre os setores formal e informal. A correção do viés de seleção foi feita baseada no método de Heckman em dois estágios (1979), assumindo a hipótese de correlação nula entre ambas as equações de seleção.

No mais, o artigo foi dividido em 6 seções incluindo esta introdução, onde na seção 2, logo abaixo, se têm uma revisão da literatura sobre informalidade e segmentação; na seção 3 se tem o desenvolvimento do modelo econométrico, tanto da escolha ocupacional e correção de viés de seleção, como da decomposição do diferencial de rendimentos; na seção 4 se apresenta uma breve exposição acerca da base de dados e estatística descritiva; na seção 5

---

<sup>1</sup> Conforme a nota técnica o coeficiente de Gini sofreu uma queda de 7% entre 2001 e 2004, passando de 0,593 para 0,569.

temos a apresentação dos resultados; e, por último, a seção 6 com algumas considerações finais.

## 2 REVISÃO DA LITERATURA

Um dos temas mais discutidos, na literatura relacionada ao mercado de trabalho, tem sido o diferencial de rendimentos entre grupos de trabalhadores. Obviamente, a definição desses grupos é particular e segue os objetivos específicos de cada análise. Um primeiro segmento, mais tradicional dessa literatura, considera características inatas dos trabalhadores como gênero e cor da pele. Outro considera segmentações baseadas na escolha e nas circunstâncias nas quais os trabalhadores são submetidos no mercado de trabalho.

Dentro desse segundo segmento, a avaliação do diferencial de rendimentos entre trabalhadores formais e informais recebeu grande atenção nos últimos anos. Este interesse vem do fato de se relacionar a segmentação no mercado de trabalho e estratégia ocupacional dos trabalhadores ao diferencial de rendimentos entre os dois setores, de acordo com o conceito formulado originalmente na obra de Dickens e Lang (1985), que explica o diferencial de rendimentos entre trabalhadores com os mesmos atributos produtivos por meio das características de ambos os setores<sup>2</sup>. Ou seja, trabalhadores com os mesmos atributos produtivos e com as mesmas preferências, obtêm diferentes níveis de utilidade por ocuparem distintos postos de trabalho. E conforme Barros et al. (1990), esta evidência pode ser interpretada como caracterização da existência de segmentação no mercado de trabalho.

Carneiro e Henley (2001) e Menezes-Filho *et al.* (2004), procuram evidências de segmentação no mercado de trabalho brasileiro e mostram que para o Brasil a mesma não é corroborada, pois os diferenciais salariais são explicados, em grande parte, pelos níveis de escolaridade e características observadas<sup>3</sup>.

Contudo, Tannuri-Pianto e Pianto (2002), utilizando o método de regressões quantílicas e o processo para correção de viés de seleção com base em Heckman (1979) e Buchinsky (1998), apontam que o mercado de trabalho brasileiro pode estar parcialmente segmentado. Os autores mostram que para os quantis mais elevados não há indícios de segmentação, pois o diferencial é inteiramente explicado pelas características produtivas dos trabalhadores. Porém, nos quantis inferiores e medianos, uma parte do diferencial é explicada pelos retornos nestes atributos o que sugere uma segmentação parcial no mercado de trabalho brasileiro, que desfavorece os trabalhadores informais situados nos quantis mais baixos. Soares (2004) argumenta que para trabalhadores do sexo feminino, analfabetos, sem experiência e cujo último trabalho era informal, o mercado de trabalho brasileiro pode estar segmentado.

Em relação ao sinal do diferencial entre grupos, um dos fatos estilizados mais bem consolidados na literatura sobre informalidade era a existência de significativos diferenciais de salários em favor dos trabalhadores formais, mesmo controlando por todas as características observáveis e pelo viés de seleção [ULYSSEA, 2005]. Fernandes (1996) a partir de dados da PNAD de 1989 mostra que neste ano, em média, os trabalhadores com carteira de trabalho assinada ganhavam 115% a mais que trabalhadores sem carteira<sup>4</sup>.

---

<sup>2</sup> Para maiores detalhes sobre o arcabouço teórico de segmentação no mercado de trabalho ver Barros e Reis (1990) e Barros (1993).

<sup>3</sup> Evidências de segmentação no mercado de trabalho são apresentadas na literatura internacional. Para os países de a América Central ver Funkhouser (1996); para o México ver Marcouiller *et al.* (1997).

<sup>4</sup> É importante destacar que o autor parte da hipótese de que existe uma segmentação no mercado de trabalho entre postos de trabalhos oferecidos por grandes empresas e postos de trabalhos oferecidos por pequenas empresas.

Entretanto, utilizando uma abordagem baseada no agrupamento dos dados por corte-tempo-escolaridade e no método de variáveis instrumentais, Menezes-Filho *et al.* (2004), estimam o diferencial de salários entre trabalhadores formais e informais. Os autores contra argumentam a ideia de que os trabalhadores formais recebem maiores salários que trabalhadores informais e chegam ao resultado que, condicional ao nível de escolaridade, a remuneração dos trabalhadores do setor informal supera a dos trabalhadores do setor formal, indicando que a proteção legal gozada pelos que se encontram na formalidade não lhes garante um maior nível salarial [MENEZES-FILHO *et al.*, (2004)].

Ainda assim, vários estudos têm mostrado, de forma sistemática, que os salários dos trabalhadores com registro em carteira são significativamente mais elevados, mesmo quando controlado por uma série de características individuais (em especial educação, idade e gênero), região de residência e ocupação.

Em um contexto regionalizado, Oliveira (2009) analisou o diferencial de salários entre trabalhadores formais e informais ao longo da distribuição de rendimentos no mercado de trabalho cearense. O autor observa que os trabalhadores formais se concentram nas áreas urbanas e na região metropolitana, e que indivíduos que pertencem a famílias pobres apresentam maiores chances de estarem ocupados em atividades sem carteira assinada.

Utilizando um modelo *logit* para verificar a probabilidade de se situar no mercado de trabalho formal, Fernandes (1996) mostra que a probabilidade de no mercado de trabalho informal cai com o aumento da escolaridade. Este autor também conclui que a probabilidade de um trabalhador estar no mercado de trabalho não regulamentado não depende apenas de seus atributos, mas de como esses atributos estão distribuídos na população.

### **3 MODELO ECONOMETRICO**

#### **3.1 Equação de salários e escolha ocupacional**

O modelo mais difundido na literatura, com o objetivo de captar os fatores que influenciam no rendimento dos indivíduos, é o proposto por Mincer (1974). O modelo, conforme afirma Menezes-Filho (2002), consiste em relacionar os salários auferidos com as características do trabalhador, sejam elas individuais (educação, experiência, sexo, raça) ou de seu trabalho (região, setor de atividade, ocupação entre outros).

Grande parte da literatura aplica a estimação dos modelos de equações de salários sob o arcabouço tradicional da teoria de capital humano e as diferenças salariais entre dois setores quaisquer são captadas por meio do uso de uma variável binária (*dummy*) em um modelo de equação única. Apesar de bastante difundido e aplicado, esse tipo de abordagem pode impor dois tipos importantes de limitações à análise.

O primeiro deles é que o coeficiente da variável binária, para o setor informal, por exemplo, representa uma diferença salarial sem qualquer dependência das características dos trabalhadores empregados nesse setor. Esse problema também inclui os coeficientes das demais variáveis explicativas, que são assumidos iguais para trabalhadores nos dois setores. Isso implica que trabalhadores formais e informais receberiam os mesmos retornos sobre suas características.

O segundo problema, remete a uma possível endogeneidade da variável que define o setor no qual o indivíduo está empregado, ou mesmo se ele decide por estar empregado, pois a decisão de oferta de trabalho e a alocação dos trabalhadores entre os setores não seria completamente aleatória, mas determinada por fatores observados e não observados. Dado isso, ao se desprezar informações relacionadas à estratégia ocupacional dos trabalhadores, pode se incorrer em um problema de seletividade amostral que é conhecido na literatura como

*viés de seleção*. Pois determinadas pessoas, de fato, podem levar em consideração os custos de oportunidade de estar ocupadas em determinado setor.

Uma forma de corrigir o primeiro problema consiste em estimar duas regressões separadas, uma para cada setor, em que as diferenças de perfil entre os dois tipos de trabalhadores são consideradas. Esse procedimento de estimar as equações separadamente é equivalente à estimação de uma única equação em que se adota a interação da variável binária indicativa do setor com cada uma das covariadas. No entanto, ao aplicar as interações também se pode incorrer em problemas na medida em que o hiato salarial passa a depender das distribuições das covariadas. A estimação de uma equação minceriana para cada setor também atende ao objetivo de estimar o diferencial de salários empregando um método de decomposição como o de Oaxaca-Blinder (1973).

No entanto, iremos considerar que o trabalhador faz pelo menos uma escolha antes de se alocar em determinado setor – formal ou informal. O indivíduo primeiramente determina se irá participar ou não do mercado de trabalho para em seguida, escolher em qual setor irá trabalhar. Desconsiderar ambas as etapas pode gerar problemas nas estimações, prejudicando a análise em questão. Tal problema é conhecido como viés de seleção.

Para tanto, iremos corrigir a partir de um método de em dois estágios proposto em Heckman (1979) que basicamente consiste em estimar a equação de seleção – neste caso serão duas – e a partir delas obter as respectivas razões inversas de Mills e por fim, utilizá-las como covariadas na equação minceriana. Para isso, recorreremos a uma hipótese que, por hora, tornará possível as estimações. Suporemos que não existe correlação entre ambas as equações de seleção, seus erros são independentes, ou seja,  $\rho_{\varepsilon v} = 0$ .

$$\begin{aligned} P_i^* &= \delta_i' \theta_i + \varepsilon_i \\ S_i^* &= \gamma_i' \pi_i + v_i \end{aligned} \quad (1)$$

Na qual  $P_i^*$  e  $S_i^*$  são variáveis latentes que indicam a participação do indivíduo em no mercado de trabalho e seleção do trabalhador para o setor formal e informal, respectivamente. Note, porém, que as variáveis  $P_i^*$  e  $S_i^*$  são não observadas e, portanto, o valor que a variável binária irá assumir dependerá do seu sinal. Como a seguir:

$$P_i = \begin{cases} 1 & \text{se } P_i^* > 0 \\ 0 & \text{se } P_i^* \leq 0 \end{cases} \quad (2)$$

$$S_i = \begin{cases} 1 & \text{se } S_i^* > 0 \\ 0 & \text{se } S_i^* \leq 0 \end{cases} \quad (3)$$

Tradicionalmente, se emprega um modelo *probit* para a escolha ocupacional do trabalhador, no qual a variável dependente assume valor igual a 1 se o indivíduo está ocupado e 0 caso contrário. Para o caso de escolha setorial a variável binária assume valor 1 se o trabalhador pertence ao setor informal e 0 se pertence ao setor formal.

Considere  $P_i$  como sendo a variável binária de escolha de participação que será explicada por um vetor de variáveis de controle  $\delta_i$ , um vetor de parâmetros  $\theta_i$  e um vetor de erros aleatórios  $\varepsilon_i$ . Também seja  $S_i$  uma variável de escolha setorial, a ser explicada por um vetor de características  $\gamma_i$  e um vetor de parâmetros  $\pi_i$ . Os respectivos modelos de probabilidade serão dados como a seguir:

$$Pr(P_i = 1) = Pr(\delta_i' \theta_i + \varepsilon_i > 0) = Pr(\varepsilon_i > -\delta_i' \theta_i) = Pr(\varepsilon_i < \delta_i' \theta_i) = \Phi(\delta_i' \theta_i) \quad (4)$$

$$Pr(S_i = 1) = Pr(\gamma_i' \pi_i + v_i > 0) = Pr(v_i > -\gamma_i' \pi_i) = Pr(v_i < \gamma_i' \pi_i) = \Phi(\gamma_i' \pi_i) \quad (5)$$

Onde  $\Phi(\delta_i'\theta_i)$  e  $\Phi(\gamma_i'\pi_i)$  são as funções de distribuição acumulada normal padrão. Por sua vez, a segunda etapa consiste em estimar a equação de salários levando em consideração o viés de seleção:

$$W_i = Z_i'\alpha_i + u_i \quad (6)$$

Onde  $W$  é o vetor de salários,  $Z$  o vetor de características individuais,  $\alpha$  o vetor de parâmetros e  $u$  um vetor de erros aleatórios. Em seguida, obtém-se a razão inversa de Mills da seguinte maneira:

$$E(W_i|P_i = 1, S_i = 1) = Z_i^I\alpha_i^I + \frac{cov(u_i, \varepsilon_i)}{\sigma_\varepsilon}\lambda_i^O + \frac{cov(u_i, v_i)}{\sigma_v}\lambda_i^{SI} \quad (7)$$

$$E(W_i|P_i = 1, S_i = 0) = Z_i^I\alpha_i^I + \frac{cov(u_i, \varepsilon_i)}{\sigma_\varepsilon}\lambda_i^O + \frac{cov(u_i, v_i)}{\sigma_v}\lambda_i^{SF} \quad (8)$$

No qual  $\lambda_i^O = \frac{-\phi(\delta_i'\theta_i)}{\Phi(\delta_i'\theta_i)}$ ,  $\lambda_i^{SI} = \frac{-\phi(\delta_i'\theta_i)}{\Phi(\delta_i'\theta_i)}$  e  $\lambda_i^{SF} = \frac{\phi(\delta_i'\theta_i)}{1-\Phi(\delta_i'\theta_i)}$  são as razões inversas de Mills. É importante observar que:

$$\lambda_i = \begin{cases} \lambda_i^{SI} & \text{se } S_i = 1 \\ \lambda_i^{SF} & \text{se } S_i = 0 \end{cases} \quad (9)$$

Conforme afirma Heckman (1979), utilizamos as razões inversas de Mills como covariadas na equação de salários e estimamos via MQO, por tanto, as equações mincerianas passarão a ter a seguinte estrutura:

$$\ln W_i^I = Z_i^I\alpha_i^I + \gamma_i^I \frac{-\phi(\delta_i'\theta_i)}{\Phi(\delta_i'\theta_i)} + u_i^I = Z_i^I\alpha_i^I + \gamma_i^I\lambda_i^{SI} + \varphi^I\lambda_i^O + u_i^I \quad (10)$$

$$\ln W_i^F = Z_i^F\alpha_i^F + \gamma_i^F \frac{\phi(\delta_i'\theta_i)}{1-\Phi(\delta_i'\theta_i)} + u_i^F = Z_i^F\alpha_i^F + \gamma_i^F\lambda_i^{SF} + \varphi^F\lambda_i^O + u_i^F \quad (11)$$

Caso o coeficiente estimado de  $\lambda_i$  seja estatisticamente significativo, então fica comprovado o viés de seletividade amostral.

### 3.2 Decompondo o diferencial de rendimentos entre trabalhadores formais e informais

Existem diversas formas de se estimar o diferencial de rendimentos entre trabalhadores pertencentes a dois grupos distintos. O método mais usual no campo da economia do trabalho é a decomposição de Oaxaca e Blinder (OB), que apesar da simplicidade, permite inferências interessantes sobre os determinantes dos diferenciais de rendimentos. O método foi elaborado de forma independente nos trabalhos de Oaxaca (1973) e Blinder (1973) e passou a ser a principal referência nos estudos sobre desigualdades e estruturas salariais nas décadas que se seguiram. Oaxaca (1973) estudou o diferencial de rendimentos entre homens e mulheres. Já Blinder (1973), faz o mesmo estudo, porém entre homens brancos e negros.

O método de decomposição de OB<sup>5</sup> tem como base equações de rendimentos minceriana, estimadas por métodos de regressão, e possui como ponto forte a decomposição detalhada da contribuição de cada variável explicativa em termos de dois efeitos. O primeiro deles é denominado *efeito dotação* ou parte explicada, que atribui parte do diferencial a

<sup>5</sup> Para mais detalhes em relação ao processo de decomposição ver Jann (2008) e Fortin, N.M., T. Lemieux, S. Firpo. (2011).

diferenças nas características dos trabalhadores. O segundo corresponde ao *efeito preço* ou parte não explicada, e correspondem as diferenças em termos dos coeficientes, ou seja, dos retornos das variáveis empregadas no modelo.

A primeira etapa da decomposição consiste em estimar uma equação de salários para cada grupo de interesse. No presente estudo, são estimadas regressões para os salários de trabalhadores empregados no setor formal, indexados por F, e para os empregados no setor informal, indexados por I.

$$\ln W_F = \beta_0^F + \sum_{i=1}^n \alpha_i^F Z_i^F + \gamma_i^F \lambda_i^{SF} + \varphi^F \lambda_i^O + u_i^F \quad (12)$$

$$\ln W_I = \beta_0^I + \sum_{i=1}^n \alpha_i^I Z_i^I + \gamma_i^I \lambda_i^{SI} + \varphi^I \lambda_i^O + u_i^I \quad (13)$$

Para fins didáticos, reescreveremos as equações (12) e (13) na forma matricial, conforme abaixo:

$$\ln W_F = X'_F \beta_F + U_F \quad (14)$$

$$\ln W_I = X'_I \beta_I + U_I \quad (15)$$

Tal que  $(\ln W_j)$  seja logaritmo natural do salário-hora que será explicado por um vetor de variáveis de controle  $(X'_j)$  e um termo de erro aleatório  $(U_j)$ , onde  $j \in \{F, I\}$ . As variáveis de controle são: educação, experiência, sexo, cor/raça, RM, Urb., setores de atividade e as razões inversas de Mills  $(\lambda)$  obtidas separadamente para cada região.

Em seguida, pode-se expressar o diferencial total como:

$$\Delta \widehat{W} = E(W_F|X_F) - E(W_I|X_I) = \ln \bar{W}_F - \ln \bar{W}_I \quad (16)$$

Note que o diferencial total  $(\Delta \widehat{W})$ , pode ser explicado a partir das variáveis independentes do modelo.

$$\Delta \widehat{W} = \ln \bar{W}_F - \ln \bar{W}_I = \bar{X}'_F \hat{\beta}_F - \bar{X}'_I \hat{\beta}_I \quad (17)$$

Um método para se construir a decomposição de Oaxaca e Blinder (1973) é o *twofold*, ou seja, em dois termos. Somando e subtraindo  $\bar{X}'_F \hat{\beta}_I$  em (17) temos:

$$\Delta \widehat{W} = \ln \bar{W}_F - \ln \bar{W}_I = \bar{X}'_F \hat{\beta}_F - \bar{X}'_I \hat{\beta}_I + \bar{X}'_F \hat{\beta}_I - \bar{X}'_F \hat{\beta}_I \quad (18)$$

Rearranjando os termos obtêm-se a decomposição:

$$\Delta \widehat{W} = \ln \bar{W}_F - \ln \bar{W}_I = \underbrace{(\bar{X}'_F - \bar{X}'_I) \hat{\beta}_I}_{\text{Efeito dotação}} + \underbrace{\bar{X}'_F (\hat{\beta}_F - \hat{\beta}_I)}_{\text{Efeito preço}} \quad (19)$$

Similarmente, podemos obter:

$$\Delta \widehat{W} = \ln \bar{W}_F - \ln \bar{W}_I = \underbrace{(\bar{X}'_F - \bar{X}'_I) \hat{\beta}_F}_{\text{Efeito dotação}} + \underbrace{\bar{X}'_I (\hat{\beta}_F - \hat{\beta}_I)}_{\text{Efeito preço}} \quad (20)$$

O primeiro termo das equações (19) e (20) representa a parte do diferencial devido às diferenças nas características produtivas como educação e experiência, por exemplo. Se, por hipótese, os trabalhadores informais tivessem o mesmo nível de características observáveis que os trabalhadores formais  $(\bar{X}'_F = \bar{X}'_I)$ , então o diferencial seria totalmente explicado por características não observadas. Já o segundo termo é a parte não explicada ou *efeito preço*, na qual se têm a influência de características não observadas e geralmente é considerado como resultado de discriminação no mercado de trabalho.



#### 4 ESTATÍSTICA DESCRITIVA

Antes de apresentar alguns resultados e tecer alguns comentários acerca das estatísticas descritivas das variáveis explicativas, cabe esclarecer como fora feita a classificação entre setor formal e informal. Para tanto, deve-se conceituar o que significa informalidade.

Nota-se, que nas últimas décadas uma parcela não desprezível da literatura em economia do trabalho tem se dedicado a discussão do conceito de setor informal. No entanto, não há um consenso em torno da questão, o que implica em uma variedade de definições de informalidade, utilizadas tanto em trabalhos teóricos quanto empíricos [ULYSSEA, 2005]. Em nível internacional, a OIT (2002) afirma que o conceito de informalidade está associado não somente a proteção legal, mas também a vulnerabilidade a qual o trabalhador está sujeito.

Segundo Ulyssea (2005), no Brasil, o problema da definição foi em grande parte minimizado pelo fato de a legislação brasileira exigir que todos os trabalhadores assalariados possuam carteira de trabalho assinada, o que fez com que a informalidade ficasse amplamente associada à posse ou não da mesma. Diversos trabalhos se aproveitam desse fato e consideram apenas empregados sem carteira assinada como integrantes do setor informal tal como Carneiro e Henley (2001); Tannuri-Pianto e Pianto (2002); e Menezes-Filho *et al.* (2004). Outros definem o setor informal como a soma dos trabalhadores sem carteira e por conta própria, ou então, o conjunto de trabalhadores que não contribuem com a previdência social, por exemplo, Kassouf (1998) e Machado *et al.* (2012).

Para fins de análise, optou-se como trabalhador informal aquele que não possui carteira de trabalho assinada, ou seja, a posse da carteira de trabalho assinada é o que caracteriza a alocação dos indivíduos entre ambos os setores. Vale ressaltar, que os trabalhadores por conta própria não foram considerados.

Para tanto, foram utilizados microdados brasileiros que tem como fonte a base de dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) para o ano de 2013, na qual é elaborada pelo Instituto Brasileiro de Geografia e Estatística (IBGE). As variáveis foram construídas de acordo com a literatura de mercado de trabalho e capital humano e podem ser vistas detalhadamente a seguir, no Quadro 1.

**Quadro 1: Descrição das variáveis explicativas das regressões de salários**

Variável	Descrição
<b>Características individuais</b>	
Educação	Para a variável educação foram utilizadas <i>dummies</i> que indicam a categoria de escolaridade, na qual o indivíduo se encontra. As categorias foram divididas em quatro: pessoas com até 4 anos de estudo; pessoas com 5 a 8 anos de estudo; pessoas com 9 a 11 anos de estudo; e por ultimo, uma <i>dummy</i> para pessoas com 13 anos ou mais de estudo. É importante ressaltar, que para o modelo <i>probit</i> a variável <b>educação</b> foi considerada na sua forma contínua e um termo quadrático também foi considerado na estimação.
Experiência	Para experiência foi usada a idade do indivíduo como proxy. Um termo quadrático da idade também foi levado em consideração.
Gênero	Para o gênero foi criada uma <i>dummy</i> denominada <b>mulher</b> que assume 1 caso seja mulher e 0 quando o indivíduo for homem.
Cor/ Raça	A variável que indica raça está representada por uma <i>dummy</i> que indica 1 caso o individuo seja “não branco” e 0 caso contrário.
<b>Características do trabalho</b>	

Setor	Para indicar o setor de atividade que o indivíduo está alocado foram criadas <i>dummies</i> que representam o setor da <b>Agricultura, Indústria, Administração pública e Serviço e Comércio.</b>
<b>Características demográficas</b>	
Urbano	Indica se o indivíduo pertence à zona urbana e assume valor igual a 1 caso se situe na zona urbana e 0 caso contrário.
Metropolitano	A variável RM indica se o indivíduo se situa ou não em uma região metropolitana e assume valor igual a 1 caso se situe na região metropolitana e 0 caso contrário.

Fonte: Elaboração própria.

No recorte amostral foram consideradas tanto pessoas ocupadas como desocupadas. No que diz respeito à idade, apenas indivíduos que possuem entre 16 e 65 anos de vida foram levados em consideração e no que tange a categoria ocupacional, funcionários públicos e militares foram excluídos da amostra. A tabela 1 a seguir mostra a estatística descritiva das variáveis explicativas, que pode fornecer algumas informações sobre amostra.

**Tabela 1: Estatística descritiva das variáveis explicativas**

Variáveis explicativas	NORDESTE				SUDESTE			
	Formal		Informal		Formal		Informal	
<b>Características ind.</b>								
	Média	DP	Média	DP	Média	DP	Média	DP
Estudo (0-4 anos)	<b>0,1336</b>	0,3402	<b>0,3155</b>	0,4647	<b>0,1176</b>	0,3221	<b>0,2427</b>	0,428
Estudo (5-8 anos)	0,1885	0,3911	0,2584	0,4378	0,1989	0,3992	0,2615	0,439
Estudo (9-12 anos)	0,5171	0,4997	0,3321	0,4710	0,4880	0,4998	0,3500	0,477
Estudo ( $\geq$ 13 anos)	<b>0,1606</b>	0,3671	<b>0,0937</b>	0,2915	<b>0,1953</b>	0,3964	<b>0,1457</b>	0,352
Idade	35,094	10,650	34,804	11,756	35,929	11,325	37,73	12,75
Idade <sup>2</sup>	1345,1	816,5	1349,5	898,66	1419,2	878,29	1586	1011
Mulher	0,3772	0,4847	<b>0,4736</b>	0,4993	0,4267	0,4946	<b>0,5386</b>	0,498
Não Branco	0,7256	0,4462	0,7547	0,4302	0,4961	0,4999	0,5498	0,497
<b>Características do trabalho</b>								
Agricultura	0,0403	0,1967	0,1578	0,3645	0,0382	0,1918	0,1047	0,306
Indústria	0,3206	0,4667	0,1673	0,3732	0,3319	0,4709	0,1716	0,377
Administração Pública	0,0314	0,1745	0,0726	0,2596	0,0210	0,1436	0,0595	0,236
Serviço e Comércio	0,6076	0,4882	0,6021	0,4894	0,6087	0,4880	<b>0,6641</b>	0,472
<b>Características demográficas</b>								
Zona Urbana	0,9304	0,2544	0,7810	0,4135	0,9631	0,1883	0,8944	0,307
Metropolitano	0,5574	0,4967	0,3077	0,4615	0,5245	0,4994	0,4516	0,497
<b>Num. de Observações</b>	<b>14.519</b>		<b>102.26</b>		<b>25.543</b>		<b>7.816</b>	

Fonte: Elaboração própria. PNAD/ IBGE (2013).

Nota-se a forte presença das mulheres no mercado de trabalho de ambas as regiões e em ambos os setores. No entanto, percebe-se que embora ainda constitua-se de menos da metade dos trabalhadores na região Nordeste, na região Sudeste a maioria dos trabalhadores informais é do sexo feminino, aproximadamente 53,86%. Ademais, no que tange a escolaridade, é necessário destacar alguns fatos importantes, como por exemplo, a distribuição dos níveis educacionais dos trabalhadores por setor. Percebe-se então, que em ambas as regiões a maioria dos trabalhadores informais possui baixos níveis de escolaridade,

na região Nordeste 31,55% dos trabalhadores informais possuem apenas até 4 anos de estudo, enquanto o setor formal é constituído apenas por 13,36% desses trabalhadores. Já na região Sudeste o fato é semelhante, são 24,27% dos trabalhadores sem carteira assinada com até 4 anos de estudos, contra 11,76% de trabalhadores formais com a mesma escolaridade.

## 5 RESULTADOS

Nesta seção serão apresentados alguns resultados acerca das estimações econométricas. A tabela 2 apresenta os resultados da estimação dos modelos *probit* para escolha ocupacional, assim como os seus respectivos efeitos marginais, onde os modelos buscam explicar quais fatores influenciam na decisão de participação no mercado de trabalho, tanto para região Nordeste quanto Sudeste. O modelo estimado foi conjuntamente estatisticamente significativo e todas as variáveis apresentaram o sinal esperado. As variáveis que indicam o setor agrícola e a que representa trabalhadores com até 4 anos de estudos foram omitidas na hora da estimação para que não se incorresse na “armadilha da *dummy*” e, portanto, são tidas como base.

**Tabela 2: Modelo probit para escolha ocupacional: Participação no mercado de trabalho.**

<b>Variável dependente: Ocupado</b>				
	<b>Nordeste</b>		<b>Sudeste</b>	
Variáveis explicativas	Coeficiente	Efeito Marginal	Coeficiente	Efeito Marginal
<i>Características individuais</i>				
Estudos (5-8 anos)	-0,0219 (0,0293)	-0,0033	-0,0592*** (0,0332)	-0,0073***
Estudos (9-12 anos)	-0,0384 (0,0273)	-0,0059	-0,0288*** (0,0317)	-0,0034***
Estudos (13 ou mais anos)	0,1436*** (0,0374)	0,0204***	0,1598*** (0,0379)	0,0177***
Idade	0,0809*** (0,0049)	0,0124***	0,0722*** (0,0048)	0,0087***
Idade <sup>2</sup>	-0,0007*** (0,0000)	-0,0001***	-0,0006*** (0,0000)	-0,0000***
Mulher	-0,3685*** (0,0185)	-0,0593***	-0,3045*** (0,0185)	-0,0379***
Não Branco	-0,0684*** (0,0213)	-0,0102***	-0,0934*** (0,0189)	-0,0112***
<i>Características demográficas</i>				
Metropolitano	-0,0343* (0,0192)	-0,0053*	-0,0420*** (0,0187)	-0,0050**
Zona Urbana	-0,2187*** (0,0297)	-0,0302***	-0,3070** (0,0503)	-0,0299***
Intercepto	-0,0925 (0,0928)		0,4053 (0,1008)	
Num. de obs.	38.745		45.463	
LR chi2(15)	1874,38		1235,79	
Prob > chi2	0,00000		0,00000	
Pseudo R <sup>2</sup>	0,0749		0,0534	

**Fonte:** Elaboração própria. PNAD/IBGE. Entre parênteses os erros-padrão. (\*\*\*) Significante a 1%, (\*\*) Significante a 5%, (\*) Significante a 10%.

Como esperado, a variável educação apresentou sinal positivo para a região Nordeste, indicando que a probabilidade do indivíduo participar do mercado de trabalho aumenta com o aumento dos anos de estudo. No entanto, para o Sudeste a mesma lógica só ocorre para indivíduos com 13 anos ou mais de estudos, enquanto que pessoas com escolaridade inferior a essa estão mais propensas a estarem situados fora do mercado de trabalho. Ademais, a idade<sup>6</sup> também afeta na probabilidade de participar ou não do mercado de trabalho, porém de forma não linear, formando um “U” invertido, para as duas regiões selecionadas.

Além das variáveis de capital humano apresentadas acima, podemos notar que as mulheres possuem maior probabilidade de estarem fora do mercado de trabalho do que homens, sendo que as chances de estarem desocupadas são maiores para indivíduos nordestinos. O mesmo acontece em relação à cor da pele, trabalhadores “não brancos” possuem maior probabilidade de estarem desempregados, onde o sinal do impacto foi o mesmo para ambas as regiões, contudo, para o Sudeste a probabilidade de estar desocupado é maior para não brancos, se comparado ao Nordeste.

Dado a análise dos fatores que influenciam a participação no mercado de trabalho, também se faz necessário investigar que fatores determinam a alocação entre os setores formais e informais. Para isso, a tabela 3 traz os resultados das estimações do modelo *probit* para trabalhadores formais e informais de cada região selecionada.

**Tabela 3: Modelo probit para escolha ocupacional: Formal-Informal**

<b>Variável dependente: Trabalhador Informal</b>				
	<b>Nordeste</b>		<b>Sudeste</b>	
Variáveis explicativas	Coeficiente	Efeito Marginal	Coeficiente	Efeito Marginal
<i>Características individuais</i>				
Estudos (5-8 anos)	-0,2626*** (0,0276)	-0,1015***	-0,1830*** (0,0263)	-0,0542***
Estudos (9-12 anos)	-0,8050*** (0,0266)	-0,3054***	-0,5664*** (0,0260)	-0,1702***
Estudos (13 ou mais anos)	-0,9157*** (0,0355)	-0,3116***	-0,5399*** (0,0317)	-0,1425***
Idade	-0,0785*** (0,0051)	-0,0308***	-0,0612*** (0,0044)	-0,0188***
Idade <sup>2</sup>	0,0009*** (0,0000)	0,0003***	0,0007*** (0,0000)	0,0002***
Mulher	0,4428*** (0,0200)	0,1737***	0,3088*** (0,0179)	0,0958***
Não Branco	0,0535** (0,0207)	0,0210**	0,0731*** (0,0170)	0,0224***
<i>Características do trabalho</i>				
Indústria	-0,7069*** (0,0391)	-0,2612***	-0,6392*** (0,0404)	-0,1763***

<sup>6</sup> A idade do trabalhador é usada como *proxy* para experiência.

Administração Pública	0,4089*** (0,0545)	0,1620***	0,4330*** (0,0568)	0,1502***
Serviço e Comércio	-0,2884*** (0,0390)	-0,1136***	-0,2934*** (0,0394)	-0,0924***
<i>Características demográficas</i>				
Zona Urbana	-0,3711*** (0,0299)	-0,1471***	-0,3697*** (0,0365)	-0,1258**
Metropolitano	-0,4158*** (0,0192)	-0,1616***	-0,0381** (0,0173)	-0,0117***
Intercepto	2,4691*** (0,0994)	-	1,2943*** (0,0923)	-
Num. de obs.	22.392		28.997	
LR chi2(15)	4448,53		2595,15	
Prob > chi2	0,00000		0,00000	
Pseudo R <sup>2</sup>	0,1447		0,0792	

**Fonte:** Elaboração própria. PNAD/IBGE. Entre parênteses os erros-padrão. (\*\*\*) Significante a 1%, (\*\*) Significante a 5%, (\*) Significante a 10%.

O modelo para escolha setorial foi conjuntamente estatisticamente significante e todas as covariadas foram individualmente significantes a nível de, no máximo, 5%. No mais, em relação às variáveis de capital humano, nota-se que a educação apresentou sinal negativo, indicando que a probabilidade do indivíduo estar ocupado no setor informal diminui com o aumento dos anos de estudo para ambas as regiões. Entretanto, acresce que, a magnitude do impacto da educação na probabilidade do indivíduo se situar no mercado de trabalho informal é maior para o Nordeste, em relação ao Sudeste. Além disso, assim como a educação, a idade do trabalhador também afeta a probabilidade de estar empregado sem carteira assinada, porém de forma não linear, para as duas regiões, assumindo a forma de “U” indo de encontro com os trabalhos de Fernandes (1996) e Oliveira (2009).

No que se relaciona ao sexo e a cor da pele, notamos que ambas as características influenciam positivamente na escolha setorial, ou seja, o fato de ser mulher impacta em 55,7% as chances de se situar no mercado de trabalho informal para o Nordeste e em 36,17% para o Sudeste. Já o fato de ser não branco aumenta as chances de estar no setor informal em 5,5% e 7,6% para as regiões Nordeste e Sudeste, respectivamente<sup>7</sup>.

Entretanto, não são somente as variáveis de capital humano e características naturais que influenciam a alocação setorial, as variáveis demográficas, assim como o setor de atividade (Indústria, Comércio e Serviços, Adm. Pública e etc.) exercem também seu papel. Com isso, note que o simples fato do indivíduo residir em Zona urbana diminui as chances de estar no setor informal em aproximadamente 31%, tanto para o Nordeste quanto para o Sudeste. Já no que tange os setores de atividade, para as duas regiões, Indústria e o setor de comércio e serviços apresentaram sinal negativo, sugerindo que os seus respectivos trabalhadores estão menos propensos a se situarem na informalidade. Em contrapartida, a variável que indica a administração pública apresentou sinal negativo, indicando que trabalhadores desse setor estão mais propensos a não possuir carteira de trabalho assinada.

<sup>7</sup> Efeito relativo de cada variável na razão de chances pode ser expresso da seguinte forma:  $[\exp(\hat{\beta}) - 1] * 100$ .

**Tabela 4: Resultados das regressões de salários com correção de viés de seletividade: Formal-Informal**

<b>Variável dependente:</b> Logaritmo do Salário-Hora				
Variáveis explicativas	<b>Nordeste</b>		<b>Sudeste</b>	
<i>Características individuais</i>				
	Formal	Informal	Formal	Informal
Estudos (5-8 anos)	0,1686*** (0,0295)	0,1122*** (0,0335)	0,1661*** (0,0216)	0,0941 (0,0659)
Estudos (9-12 anos)	0,4431*** (0,0631)	0,2370*** (0,0900)	0,5297*** (0,0454)	0,1075 (0,2014)
Estudos (13 ou mais)	1,1896 *** (0,0727)	0,9037*** (0,1040)	1,3002*** (0,0458)	0,7275 *** (0,1918)
Idade	0,0717*** (0,0110)	0,0151*** (0,0125)	0,1166*** (0,0077)	0,0161 (0,0232)
Idade <sup>2</sup>	-0,0006*** (0,0001)	-0,0001*** (0,0001)	-0,0012*** (0,0000)	-0,0000 (0,0000)
Mulher	-0,4025*** (0,0474)	-0,1668*** (0,0589)	-0,5530*** (0,0317)	-0,1072 (0,1138)
Não Branco	-0,0698*** (0,0147)	0,0445*** (0,0191)	-0,2169 *** (0,0113)	0,0849 *** (0,0321)
<i>Características do trabalho</i>				
Indústria	0,3089*** (0,0690)	0,1511 *** (0,0791)	0,4662*** (0,0560)	0,0502 (0,2283)
Administração Pública	0,3068 *** (0,0516)	0,8140 *** (0,0592)	0,0283 (0,0500)	0,7066 *** (0,1577)
Serviço e Comércio	0,1527 *** (0,0461)	0,3070*** (0,0372)	0,2020*** (0,0363)	0,1550 (0,1028)
<i>Características demográficas</i>				
Zona Urbana	0,0331 (0,0425)	0,0179 *** (0,0466)	-0,1580*** (0,0419)	0,0076 (0,1309)
Metropolitano	0,1227** (0,0319)	0,1309*** (0,0513)	0,1368*** (0,0089)	0,1844*** (0,0230)
$\lambda^o$	-1,3627 *** (0,3179)	-0,2635 (0,3368)	-2,4565*** (0,2912)	0,1650 (0,5492)
$\lambda^S$	0,3019** (0,1359)	-0,2657 (0,1788)	0,9426*** (0,1670)	-0,5244 (0,4856)
Intercepto	-0,8171** (0,3799)	0,4332 (0,2591)	-1,8246*** (0, 1091)	0,4606 (0,2751)
Num. de obs.	12098	9679	20844	7088
F(17, 9617)	229,07	222,58	682,63	161,84
Prob > F	0,0000	0,0000	0,0000	0,0000
R <sup>2</sup> Ajustado	0,2088	0,2427	0,3141	0,2411

**Fonte:** Elaboração própria. PNAD/IBGE (2013). Entre parênteses os erros-padrão. (\*\*\*) Significante a 1%, (\*\*) Significante a 5%, (\*) Significante a 10%.

Em relação às discrepâncias entre as regiões, percebe-se que os retornos da educação dos trabalhadores formais são maiores para a região Nordeste, exceto para trabalhadores com

13 anos ou mais de estudo. Já para os informais os retornos são maiores para o Nordeste em todos os níveis de escolaridade. Com respeito ao impacto da idade/experiência, esta se dá de forma não linear e está de acordo com a literatura relacionada ao mercado de trabalho e capital humano. Essa não linearidade significa que os salários aumentam à medida que se aumenta a experiência, porém a taxas cada vez menores. Isso nos diz que ao atingir determinado nível de idade/experiência, os trabalhadores podem usufruir de menores salários, ou seja, a relação salário-experiência passará a ser negativa.

No que se relacionam as variáveis de discriminação (*mulher e não branco*), estas também obtiveram o sinal esperado e mostram que tanto para o Nordeste quanto para o Sudeste brancos e homens possuem vantagens salariais, sendo que para as duas características o impacto negativo nos salários é maior para a região Sudeste, tanto para formal quanto para informal, com exceção de da variável mulher, na qual o impacto no setor formal é maior para Nordeste se comparado com o Sudeste. Ou seja, o fato de ser mulher no mercado de trabalho informal nordestino faz com que seus rendimentos caiam em média 15,36%<sup>8</sup>. Já para a trabalhadora informal situada na região Sudeste a perda é em média de 10,16% em relação aos homens. No tocante ao impacto do setor de atividade nos níveis salariais, nota-se que o fato de se situar no setor de comércio e serviços faz com que haja uma elevação salarial, tanto para trabalhadores formais quanto informais para as duas regiões selecionadas. O mesmo sinal do impacto se dá para os trabalhadores da indústria em ambas as regiões.

No que diz respeito às características demográficas, tanto para os trabalhadores informais do Nordeste quanto para os do Sudeste, o fato de pertencer à zona urbana impacta positivamente nos salários. Já para o setor formal, o sinal do impacto é positivo apenas para o Nordeste, pois para o Sudeste residir em zona urbana sugere uma diminuição salarial, em média. Os trabalhadores que residem em regiões metropolitanas gozam de maiores salários em comparação àqueles que não residem, tanto para trabalhadores formais quanto informais. Por exemplo, o trabalhador informal que reside em uma região metropolitana nordestina tem em média uma vantagem salarial de 14%.

Em relação às razões inversas de Mills  $\lambda^O$  e  $\lambda^S$ , lembre-se que a correção fora feita tanta para decisão de participação, como para seleção setorial formal/informal. Para a decisão de participação no mercado de trabalho fora criada uma inversa de Mills,  $\lambda^O$ , que se mostrou significativa apenas para o setor formal em ambas as regiões, indicando que a omissão dessa variável ocasionaria estimativas enviesadas. Já para equação de seleção setorial, a razão inversa de Mills,  $\lambda^S$ , também fora estatisticamente significantes, apenas para equação minceriana estimada para trabalhadores formais. Com efeito, a não inclusão desta variável pode fornecer estimativas enviesadas, porém, apenas para o setor formal.

Contudo, a proposta principal deste trabalho é de investigar o comportamento do diferencial de rendimentos entre grupos em cada região, e para tanto, iremos a seguir analisar os resultados obtidos na decomposição de Oaxaca e Blinder, que se encontram na tabela 5.

**Tabela 5: Resultados da decomposição: Formal-Informal**

NORDESTE		SUDESTE	
Formal	1,8055	Formal	2,0160
Informal	1,3156	Informal	1,7908
Diferencial Total	0,4900	Diferencial Total	0,2252
Explicada	0,6620	Explicada	1,7282
Não explicada	-0,1720#	Não explicada	-1.5030
<b>Nordeste</b>		<b>Sudeste</b>	

<sup>8</sup> Efeito relativo de cada variável binária no modelo do tipo *log-lin*, pode ser expresso da seguinte forma:  $[\exp(\hat{\beta})-1]*100$ .

Variáveis explicativas	<i>Efeito Dotação</i>	<i>Efeito Preço</i>	<i>Efeito Dotação</i>	<i>Efeito Preço</i>
<i>Características individuais</i>				
Estudos (5-8 anos)	-0,0103*** (0,0020)	0,0147 (0,0117)	-0,0098*** (0,0016)	0,0196 (0,0189)
Estudos (9-12 anos)	0,0810*** (0,0119)	0,0681* (0,0363)	0,0792*** (0,0076)	0,1462*** (0,0715)
Estudos (13 ou mais anos)	0,0652*** (0,0064)	0,0239** (0,0106)	0,0494*** (0,0062)	0,0702*** (0,0243)
Idade	0,0102 (0,0111)	1,9748*** (0,5808)	-0,2292 *** (0,0249)	3,815*** (0,9300)
Idade <sup>2</sup>	0,0085 (0,0076)	-0,6785*** (0,2437)	0,2188 *** (0,0227)	-1,8265*** (0,4790)
Mulher	0,0400*** (0,0054)	-0,1120*** (0,0360)	0,0689 *** (0,0054)	-0,2433*** (0,0645)
Não Branco	0,0018*** (0,0005)	-0,0191 (0,0183)	0,0121*** (0,0016)	-0,0749 *** (0,0193)
<i>Características do trabalho</i>				
Indústria	0,0475*** (0,0107)	0,0262 (0,0174)	0,0746*** (0,0093)	0,0712** (0,0403)
Administração Pública	-0,0128*** (0,0023)	-0,0370*** (0,0059)	-0,0011 (0,0019)	-0,0407 (0,0101)
Serviço e Comércio	0,0010 (0,0010)	-0,0926*** (0,0356)	-0,0107 *** (0,0023)	0,0310 (0,0721)
<i>Características demográficas</i>				
Zona Urbana	0,0049 (0,0063)	0,0117 (0,0487)	0,0112 *** (0,0030)	0,1467 (0,1218)
Metropolitano	0,0280*** (0,0074)	-0,0024 (0,0177)	0,0078 *** (0,0010)	-0,0205** (0,0106)
$\lambda^o$	-0,0122*** (0,0035)	0,3269*** (0,0931)	-0,0047** (0,2589)	0,3844*** (0,0911)
$\lambda^l$	0,4086** (0,1839)	-0,4264** (0,1687)	1,4616*** (0,0025)	-1,6971*** (0,5940)
Intercepto	-	-1.2504*** (0,4599)	-	-2,2852*** (0,3891)
Num. de obs.	21.777		27932	
Num. de obs. (Formal)	12098		20844	
Num. de obs. (Informal)	9679		7088	

**Fonte:** Elaboração própria. PNAD/IBGE (2013). Entre parênteses os erros-padrão. (\*\*\*) Significante a 1%, (\*\*) Significante a 5%, (\*) Significante a 10%. (#) Estatisticamente insignificante.

A decomposição de Oaxaca e Blinder (1973), que foi descrita na seção 3.2, traz alguns resultados interessantes acerca do diferencial de salários entre trabalhadores formais e informais. O diferencial total de salários teve sinal positivo e foi estatisticamente significativo para as duas regiões e foram de 0,4900 para a região Nordeste e 0,2252 para a região Sudeste. Estes resultados podem ser interpretados como um diferencial salarial a favor dos trabalhadores formais de 63,23% para a região Nordeste e de 25,25% para a região Sudeste,



corroborando os resultados obtidos por Fernandes (1996) a nível nacional e para Oliveira (2009) para o estado do Ceará, na qual apontam um diferencial positivo para aos trabalhadores formais<sup>9</sup>.

No que se espera da contribuição de cada efeito (preço e quantidade) no diferencial total, tem-se que para o Sudeste o diferencial devido a características não observadas (*efeito preço*), favorece o setor informal, enquanto que para o Nordeste a contribuição se mostrou estatisticamente insignificante, sugerindo um diferencial total determinado apenas pelo *efeito dotação*, ou seja, apenas pelas diferenças nas características dos trabalhadores. Contudo, o *efeito preço* para o Sudeste explica percentualmente o diferencial total salarial em -77,75%. A explicação para isso seria que os retornos das características favorecem os trabalhadores informais, novamente indicando que o diferencial total é devido a diferenças nas características produtivas dos trabalhadores. Dado isso, pode-se dizer que os resultados sugerem que não há segmentação no mercado trabalho das regiões Nordeste e Sudeste, indo de acordo com as constatações de Carneiro e Henley (2001) e Menezes-Filho et al. (2004), obtidos a nível nacional.

No que diz respeito à contribuição marginal da educação, o que se observa é que ela impacta positivamente no diferencial explicado da decomposição para trabalhadores com escolaridade acima de 4 anos de estudos, e negativamente para os com baixa escolaridade favorecendo os trabalhadores informais. Já para variável experiência o *efeito quantidade* é negativo para o Sudeste, sugerindo que esta influencia de forma a diminuir o diferencial de salários, enquanto os seu *efeito preço* favorece os trabalhadores formais elevando a desigualdade de rendimentos.

## 6 CONSIDERAÇÕES FINAIS

A proposta do presente trabalho é de investigar, controlando a seletividade amostral, o diferencial de rendimentos entre trabalhadores formais e informais para as regiões Nordeste e Sudeste. Além do diferencial de rendimentos, pode se notar como as características dos trabalhadores influenciam na escolha entre o mercado de trabalho formal e informal, assim como os determinantes da participação no mercado de trabalho.

No modelo probit se observa que a educação impacta negativamente a probabilidade de se situar no mercado de trabalho informal e que a experiência se comporta no formato de “U” indo de acordo com resultados apresentados por Fernandes (1996) e Oliveira (2009). Os resultados obtidos pela decomposição sustentam a hipótese de que não há segmentação no mercado de trabalho para as duas regiões, dado que o determinante do diferencial total de rendimentos é a diferença nas características produtivas dos trabalhadores, haja vista que o *efeito preço* fora negativo e favorecera os trabalhadores informais “puxando para baixo” o diferencial total, enquanto para o Nordeste o mesmo se mostrou insignificante.

Dado isso, o próximo passo deste trabalho será investigar esse diferencial ao longo da distribuição de salários, podendo então, obter resultados mais precisos acerca do comportamento do mercado de trabalho regional, no que tange os setores formais e informais. Outro ponto a ser explorado será uma nova forma de se corrigir viés de seleção, na qual se negará a hipótese de correlação nula entre as duas equações de escolha, assumindo que a escolha de estar no mercado de trabalho e de se situar ou não no setor informal estão correlacionados.

---

<sup>9</sup> Percentual calculado a partir da seguinte expressão:  $[\exp(\text{Diferencial total})-1]*100$ . Para mais detalhes ver Jann (2008).

## REFERÊNCIAS

- BARROS, R. P. de, REIS, J. G. A., RODRIGUEZ, J. “Segmentação no mercado de trabalho: a carteira de trabalho na construção civil”. **Revista de Econometria**, v. 10, n. 2, p. 313-35, 1990.
- BARROS, R. P.; SEDLACEK, G. L.; VARANDAS, S. “Segmentação e mobilidade no mercado de trabalho: a carteira de trabalho em São Paulo”. **Pesquisa e Planejamento Econômico**, v. 20, n. 3, pp.87 - 103, 1990.
- BARROS, R. P. **The informal labor market in Brazil**. Mimeo. 1993
- BLINDER, A. S. Wage discrimination: reduced form and structural estimates. **The Journal of Human Resources**, v.8(3), p.436-455. 1973.
- CACCIAMALI, M.C. “As economias informal e submersa: conceitos e distribuição de renda”. In: CAMARGO, J.M. e GIAMBIAGI, F. (org.). **Distribuição de renda no Brasil**. Paz e Terra, São Paulo, 1991.
- CACCIAMALI, M.C.; FERNANDES, R. “Distribuição dos trabalhadores e diferenciais de salários entre mercados de trabalho regulamentado e não regulamentado”, **Pesquisa e Planejamento Econômico**, v. 23, n. 1, pp.135-56, 1993.
- CARNEIRO, F.G.; HENLEY, A. “**Modelling Formal vs. Informal Employment and Earnings: microeconomic evidence for Brazil**”, Anais do XXIX Encontro Nacional da Anpec, 2001.
- DICKENS, W.T.; LANG, K. "A test of dual labour market theory", **American Economic Review**, 75(4):1-22.1985.
- FERNANDES, R. Mercado de trabalho não regulamentado: participação relativa e diferenciais de salários. **Pesquisa e Planejamento Econômico**, São Paulo, SP, v. 26, p. 417-442, 1996.
- FORTIN, N.M., T. LEMIEUX, S. FIRPO. Decomposition Methods in Economics, in O. Ashenfelter and D. Card, eds., **Handbook of Economics**, Amsterdam: North-Holland, Vol. IV.A: 1-102. 2011
- FUNKHOUSER, E. “The urban informal sector in Central America: household survey evidence”. **World Development**, v. 24, n. 11, pp.1737-51, 1996.
- GUIMARÃES, D. B. **Avaliações de abordagens metodológicas da equação de rendimentos aplicadas em cenários selecionados no mercado de trabalho**. 117f. Tese (Doutorado) – Universidade Federal do Ceará. Programa de Pós Graduação em Economia, CAEN, Fortaleza, 2013.
- IPEA. **Sobre a recente queda da desigualdade de renda no Brasil**. Brasília: Ipea, ago.2006 (Nota Técnica). Disponível em: <http://www.ipea.gov.br>. Acessado em: 02 de setembro de 15.
- JANN, B. A Stata implementation of the Blinder-Oaxaca decomposition. **Stata Journal**, v. 8, n. 4, p. 453-479, May 2008.
- KASSOUF, A. L. Wage gender discrimination and segmentation in the Brazilian labor market. **Economia Aplicada**, v.2, 1998.
- LEE, L.F. “Unionism and wage rates: a simultaneous equation model with qualitative and limited dependent variables”, **International Economic Review**, 19: 415-33. 1978.

MARCOULLIER, D.; RUIZ DE CASILLA, V.; WOODRU, C. "Formal Measures of the informal sector wage gap in Mexico, El Salvador and Peru", **Economic Development and Cultural Change**, v. 45, n. 2, pp. 367-92, 1997.

MENEZES-FILHO, N.; MENDES, M.; ALMEIDA, E. S. O diferencial de salários formal-informal: segmentação ou viés de seleção? **Revista Brasileira de Economia**, v. 58, n.2, 2004.

MINCER, J. **Schooling, experience and earnings**. Columbia University Press: New York. 1974.

MOURA, L.R.; BARBOSA FILHO, F; H. "**Determinantes da Redução da Informalidade do Emprego no Brasil entre 2002 e 2012: Características Socioeconômicas, Setoriais e Institucionais**", Anais do XLII Encontro Nacional da Anpec, 2014.

OAXACA, R.L. "Male-Female Wage Differentials in Urban Labor Markets". **International Economic Review**, 14(3), 693-709. 1973.

OAXACA, R.L.; RANSOM, M. R. (1999). "Identification in Detailed Wage Decompositions". **Review of Economics and Statistics**, 81(1), 154-157.

OIT – ORGANIZAÇÃO INTERNACIONAL DO TRABALHO. **Decent work and the informal economy**. Geneva: ILO, 2002.

OLIVEIRA, V. H.; **A informalidade e o diferencial de salários no mercado de trabalho cearense**. Texto para Discussão n° 58 – IPECE, Ceará, 2009.

RAMOS, L.; FERREIRA, V. Padrões espacial e setorial da evolução da informalidade no período de 1991-2005. **Pesquisa e Planejamento Econômico**. V.36 (3), p.471-488. 2006.

SOARES, F. V. **Some stylized facts of the informal sector in Brazil**. IPEA, 2004a. (Texto para Discussão, n.1020).

SOARES, F.V. "**Do Informal Workers Queue for Formal Jobs in Brazil?**". IPEA, texto para discussão, n. 1021, 2004b.

TANNURI-PIANTO, M. E.; PIANTO, D. Informal employment in Brazil – a Choice at the top and segmentation at the bottom: a quantile regression approach. Encontro Brasileiro de Econometria, **Anais...SBE**. 2002.

ULYSSEA, G. Informalidade no mercado de trabalho brasileiro: uma resenha da literatura. **Revista de Economia Política**, v.26 (4) p.596-618. 2006.

ULYSSEA, Gabriel (2007) "Segmentação no mercado de trabalho e desigualdade de rendimentos no Brasil: Uma análise empírica" **Texto para Discussão IPEA**, no. 1261. Disponível em: <http://www.ipea.gov.br>. Acessado em: 02 de setembro de 15.