

# DETERMINANTES DA DESIGUALDADE DE RENDA NO BRASIL: UMA ABORDAGEM BASEADA NA DECOMPOSIÇÃO DE SHAPLEY

**Vladimir Faria dos Santos**

Graduado em Ciências Econômicas, mestre e doutor em Economia Aplicada pela Universidade Federal de Viçosa. É professor associado da Universidade Federal Fluminense.

E-mail: vladimirsantos@id.uff.br



<http://orcid.org/0000-0002-7160-9986>

**Como citar este artigo:** Santos, V. F. dos. (2023). Determinantes da desigualdade de renda no Brasil: Uma abordagem baseada na decomposição de Shapley. *Revista de Economia Mackenzie*, 20(2), 13–37. doi:10.5935/1808-2785/rem.v20n2p.13–37

**Recebido em:** 26/7/2021

**Aprovado em:** 15/5/2023



Este artigo está licenciado com uma Licença Creative Commons - Atribuição-NãoComercial 4.0 Internacional

## Resumo

Este trabalho teve como objetivo avaliar os principais determinantes da desigualdade de renda no Brasil. Para isso, foi estimada uma equação de rendimentos que leva em conta, simultaneamente, a possibilidade de viés de seletividade e a endogeneidade. Em seguida, foi usado um método de decomposição baseado no valor de Shapley, um conceito de solução para jogos cooperativos. Os resultados mostram que a educação foi o principal determinante da desigualdade de rendimentos, independentemente do índice de desigualdade considerado.

**Palavras-chave:** Desigualdade de renda; decomposição de Shapley; educação; teoria dos jogos; regressão.

Classificação *JEL*: I24, O15.

## 1 INTRODUÇÃO

A economia brasileira tem passado por um processo de desconcentração de renda. Embora possam ser notadas reduções na desigualdade de renda a partir do ano de 1997, foi após 2001, como enfatizam Barros et al. (2007), que a queda se tornou contínua e mais acentuada.

Segundo Ferreira et al. (2006), há três fatores que podem ter contribuído para a diminuição na concentração de renda. Primeiro, a redução das diferenças de rendimentos entre grupos educacionais distintos, advindo, provavelmente, do declínio prolongado nos retornos da educação. Segundo, devido à convergência de renda das famílias localizadas em áreas urbanas e rurais. Por fim, o terceiro fator é a expansão dos programas governamentais de transferência de renda e a melhoria no seu grau de focalização. Além desses fatores, os autores destacam a importância da estabilidade macroeconômica alcançada pelo Plano Real, o que eliminou as altas taxas de inflação no período anterior ao plano.

Apesar da redução, a renda ainda é distribuída de uma forma bem desigual no Brasil. Conforme os dados da Pesquisa Nacional por Amostra de Domicílios (Pnad) de 2014, o coeficiente de Gini foi de aproximadamente 0,501. Já a proporção da renda apropriada pelos 10% mais ricos, segundo a renda

domiciliar *per capita*, foi de cerca de 40,3%, enquanto os 10% mais pobres apropriaram-se de aproximadamente 1,5%.

As possíveis causas da alta desigualdade de renda na economia brasileira já vem sendo estudadas há um longo tempo. Uma das principais causas, constatadas em diversos estudos, é a qualificação (educação). Para Langoni (1973), foi a partir do processo de desenvolvimento econômico – que se iniciou por volta de 1960 e que alterou a estrutura econômica do Brasil, passando de agrária para a industrial – que se agravou o problema da iniquidade distributiva no país. A justificativa, segundo o autor, está no mercado de trabalho, visto que houve expansão da demanda de mão de obra qualificada, motivada pelo crescimento do setor industrial; com isso, houve aumento dos salários relativos em favor dos mais qualificados.

Conforme Senna (1975 como citado em Castelo Branco, 1979), a concentração industrial e a escolaridade média explicaram, conjuntamente, cerca de 65% das diferenças de salários médios entre os componentes do setor industrial brasileiro no ano de 1970.

Trabalhos mais recentes também forneceram evidências com relação aos principais fatores que determinam a desigualdade de renda brasileira. A educação quase sempre é um dos fatores que têm uma contribuição importante para explicar a desigualdade (Souza et al., 2013; Cunha & Vasconcelos, 2012; Berni, 2007; Ramos & Vieira, 2001).

Embora a qualificação seja importante, há outros fatores relevantes para explicar a desigualdade. Salardi (2005), por exemplo, analisou, para o ano de 2002, os determinantes da desigualdade de renda para o Brasil utilizando métodos de decomposição – dentre eles, o método de Fields. O autor utilizou variáveis que captam características demográficas, nível educacional, área geográfica etc. Os resultados mostraram que a desigualdade brasileira é explicada, em sua maior parte, pelas diferenças regionais e étnicas.

Siqueira e Siqueira (2006), para a região Nordeste, decompueram a desigualdade de renda em desigualdade intergrupo e intragrupo, utilizando dados da PNAD. Os resultados mostraram que a grande maioria da iniquidade distributiva é encontrada dentro dos estados e dos grupos de gênero e raça.

Diante do exposto, este trabalho teve como objetivo quantificar os principais fatores determinantes da desigualdade de renda brasileira. Para tanto, foi usado um procedimento de decomposição baseado em uma regressão. A decomposição é realizada aplicando um conceito da teoria dos jogos cooperativos, a saber, o valor de Shapley. Dentre os trabalhos que já aplicaram o

método, pode-se mencionar Wan e Zhou (2005), Gunatilaka e Chotikapanich (2009) e Santos e Vieira (2015). A diferença metodológica em relação aos artigos mencionados é que se levou em consideração, simultaneamente, o problema da endogeneidade e do viés de seletividade.

## 2

## DESIGUALDADE DE RENDA: ASPECTOS TEÓRICOS

A literatura que trata da desigualdade de renda pode ser dividida em duas grandes áreas: aquela que analisa os efeitos macroeconômicos sobre a desigualdade e aquela que foca o relacionamento entre fatores microeconômicos e desigualdade (Hammill, 2005).

Segundo Hammill (2005), os trabalhos que abordam a análise macroeconômica da desigualdade possuem algumas características em comum. Eles normalmente utilizam os agregados macroeconômicos, tais como PIB, taxa de juros, crescimento econômico, sistema tributário, entre outros. A renda, na maioria das vezes, é definida como PIB ou PIB *per capita*.

Kuznets (1955) foi um dos pioneiros a abordar a desigualdade num contexto macroeconômico. O autor introduziu a famosa curva do U-invertido, que estabelece que, inicialmente, a distribuição de renda torna-se mais desigual à medida que a renda cresce; porém, com o desenvolvimento do país, essa tendência se reverte e a distribuição torna-se mais igualitária.

Após o trabalho de Kuznets (1955), diversos outros (teóricos e empíricos) surgiram com o intuito de verificar a relação entre crescimento econômico e desigualdade de renda (Kaldor, 1957; Alesina & Rodrik, 1991; Persson & Tabellini, 1994; Li & Zou, 1998; Barro, 2000; Forbes, 2000; García-Peñalosa & Turnovsky, 2006; Jin, 2009).

As evidências empíricas e teóricas não são conclusivas. Há trabalhos que encontraram uma relação negativa entre crescimento e desigualdade. Alesina e Rodrik (1994), por exemplo, desenvolveram um modelo de crescimento econômico endógeno, utilizando trabalho e capital como fatores de produção para analisar o relacionamento entre política e crescimento econômico, considerando os conflitos distributivos existente entre os agentes. Por meio do teorema do votante mediano, os autores verificaram que a desigualdade é um fator que reduz o crescimento econômico, visto que a má distribuição de

renda tende a influenciar o votante mediano a preferir maior tributação sobre o capital, que, por sua vez, reduz o crescimento da renda.

Seguindo a mesma linha de Alesina e Rodrik (1991), Persson e Tabellini (1994) desenvolveram um modelo teórico para relacionar crescimento econômico, desigualdade de renda e instituições políticas. De forma geral, o resultado dos autores evidenciou uma relação negativa entre desigualdade e crescimento. A explicação para isso estaria no fato de que a má distribuição da renda tende a gerar políticas que não protegem os direitos de propriedade e não permitem a completa apropriação privada dos retornos do investimento. Nesse cenário, portanto, os incentivos ao investimento são baixos e, por consequência, a taxa de crescimento econômico fica aquém do que poderia alcançar.

Alesina e Perotti (1996) argumentam que quanto mais desigual é a distribuição de riqueza, maior tende a ser a instabilidade política, econômica e social do país, o que incentiva os indivíduos a se envolverem em atividades ilegais. Em um ambiente de insegurança e incerteza, a acumulação de capital e, por consequência, os investimentos, são menores. Assim, a desigualdade levaria a uma taxa de crescimento futuro menor.

Por outro lado, há estudos que não encontraram relação negativa entre a má distribuição de riqueza e o crescimento da economia. Como exemplo, podem-se citar García-Peñalosa e Turnovsky (2006). Os autores analisaram a relação entre crescimento e desigualdade de riqueza com base na pressuposição de que a taxa de crescimento da economia e a distribuição de renda são ambas endógenas e influenciadas por mudanças estruturais e políticas macroeconômicas. Seus resultados evidenciaram que taxas de crescimento maiores estão associadas à distribuição mais desigual da riqueza.

Li e Zou (1998) também investigaram, teórica e empiricamente, o relacionamento entre desigualdade e crescimento. Assim como García-Peñalosa e Turnovsky (2006), os autores observaram, tanto teórica como empiricamente, que, ao incorporar o consumo público na função de utilidade, a má distribuição de renda pode proporcionar taxas de crescimento maiores. Esse mesmo resultado foi obtido por Forbes (2000), que encontrou uma relação positiva no curto e médio prazo entre crescimento e desigualdade.

É importante enfatizar que há também trabalhos que verificaram relação ambígua em seus resultados. Entre esses, podem-se citar Barro (2000) e Jin (2009). Especificamente, Barro (2000) encontrou uma relação negativa entre

crescimento e desigualdade em países pobres, mas uma relação positiva quando foram considerados os países mais ricos.

Todos esses trabalhos supracitados utilizam um enfoque macroeconômico da desigualdade que, de acordo com Hammill (2005), omite algumas considerações que existem a respeito dos efeitos microeconômicos. Segundo esse autor, dada a probabilidade de que a combinação dos efeitos micro e macroeconômico afetem a desigualdade, conclusões a respeito das suas causas – com base somente nos fatores macroeconômicos – podem ser questionáveis.

As características microeconômicas, na maioria das vezes, são aquelas relacionadas à demografia, ao mercado de trabalho e à educação (Hammill, 2005). Segundo Garner e Terrell (2001), entre os mecanismos de transmissão, a composição da família (característica demográfica) pode ter uma função importante na determinação da desigualdade. Nesse contexto, Garner e Terrell (2001) analisaram a mudança na desigualdade de renda na Eslováquia no período de 1988 a 1996, e seus resultados sugerem que as alterações nas características demográficas têm contribuído para aumentar a desigualdade no país.

Da mesma forma, o trabalho de Cowell e Jenkins (1995), que analisaram a economia norte-americana, sugere que entre 20% e 30% (dependendo do modelo utilizado) da desigualdade de renda é explicada por características demográficas (sexo, raça e idade).

Leibbrandt e Woolard (2001) analisaram a relação entre mercado de trabalho e desigualdade de renda familiar na África do Sul. Os resultados demonstram que a desigualdade de renda é extremamente relacionada ao acesso ao mercado de trabalho e às variações nos salários de mercado. Por meio da decomposição do índice de Gini, esses autores mostraram que a renda do trabalho (salário) é a principal causa da desigualdade de renda na África do Sul. Fatores relacionados à demografia também se mostraram importantes para explicar a desigualdade.

Fields et al. (2003) estudaram a dinâmica da renda *per capita* das famílias usando dados da Indonésia, África do Sul, Espanha e Venezuela. Os autores mostraram que mudanças no mercado de trabalho que afetem o chefe da família é o fator mais importante para mudanças na renda familiar.

Outro fator importante, e que é amplamente reconhecido na literatura, para explicar a desigualdade é a educação. Diversos trabalhos mostram que a educação e a experiência estão intimamente relacionadas com a desigualdade

de renda<sup>1</sup>. De uma perspectiva teórica, quanto maior a disparidade educacional, maior é a desigualdade de renda, visto que o primeiro tende a ampliar a diferença de rendas (Duman, 2008). De acordo com Langoni (1973), o crescimento da desigualdade de renda brasileira, a partir da década de 1960, é decorrente, basicamente, do diferencial de salários entre os trabalhadores mais qualificados e menos qualificados. Em outras palavras, para Langoni, o crescimento da taxa de retorno da escolaridade foi o principal fator para o agravamento da distribuição de renda.

Cabe mencionar, segundo Alejos (2003), que na literatura é possível encontrar várias teorias que têm como objetivo esclarecer o fenômeno da desigualdade de renda, sendo possível dividi-las em dois grupos: aquelas que focam as características dos trabalhadores e aquelas que enfatizam a estrutura do mercado de trabalho. Podem-se classificar, dentro do primeiro grupo, as teorias do capital humano e da discriminação no mercado de trabalho. No segundo grupo, estão as teorias dos diferenciais compensatórios e salário-eficiência e a teoria do mercado de trabalho dual.

### 3

## METODOLOGIA: DECOMPOSIÇÃO DE SHAPLEY

Diferentemente dos métodos tradicionais de decomposição por subgrupo populacional ou por fontes de renda na análise da desigualdade de renda, a abordagem baseada em regressão tem a vantagem de permitir a inclusão de quaisquer variáveis explicativas, tais como variáveis econômicas, sociais e demográficas (Gunatilaka & Chotikapanich, 2009). Conforme Wan e Zhou (2005), todas as abordagens baseadas em regressão se iniciam com uma função de rendimentos (*income generating function*), que pode ser escrita como:

$$\ln(Y_i) = \beta_0 + \sum_{j=1}^K \beta_j X_{ji} + \varepsilon_i \quad (1)$$

1 Ver Morley (2001).

em que  $Y_i$ , nesta pesquisa, refere-se aos rendimentos de todos os trabalhos do indivíduo  $i$  sobre o número de horas mensais trabalhadas,  $X_{ji}$  são os fatores que determinam a renda e  $\varepsilon_i$  é um termo de erro aleatório com as propriedades usuais. O uso de uma especificação semi-log é justificado, conforme Wan e Zhou (2005), pelo fato de a variável renda ter distribuição aproximadamente log-normal.

Foram consideradas as seguintes variáveis explicativas, escolhidas de acordo com a literatura (Santos & Vieira, 2013; Ney & Hoffmann, 2009; Kassouf, 1994):

- a) variável binária para gênero, que assume valor 1 para homens e 0, caso contrário;
- b) variável binária para cor, que assume valor 1 para cor branca e 0, caso contrário;
- c) idade, medida em dezenas de anos. Foi considerada também idade ao quadrado, visto que o rendimento não varia linearmente com a idade;
- d) escolaridade, em anos de estudo;
- e) cinco variáveis binárias para distinguir a região geográfica: Sudeste, Sul, Centro-Oeste, Norte e Nordeste (categoria de referência);
- f) três variáveis binárias para distinguir a posição na ocupação: empregador, conta própria e empregado (categoria de referência).

De posse da equação de rendimentos, é possível computar a contribuição de cada variável para a desigualdade de renda. O procedimento de decomposição proposto por Shorrocks (1999) segue a mesma lógica do valor de Shapley. Por isso, então, ela é denominada decomposição de Shapley.

A vantagem da decomposição de Shapley é que ela pode ser aplicada a quaisquer tipos de especificação econométrica adotados para a equação de rendimentos (Wan, 2004). Dado que a forma funcional da equação (1) é do tipo semi-log, a equação de rendimentos, em termos da variável original rendimento, é não linear. Desse modo, não seria possível utilizar (ao menos em termos da variável rendimento), por exemplo, a decomposição de Fields, utilizada por Salardi (2005), mas sim a decomposição de Shapley<sup>2</sup>. Outra vantagem desse método é que ele pode ser utilizado para decompor qualquer índice de desigualdade (Índice de Gini, Índice de Theil [L e T], Coeficiente de

---

2 Salardi (2005) utilizou o logaritmo do rendimento, e não a variável rendimento, para decompor a equação de rendimentos.



Variação etc.), diferentemente, por exemplo, do método de Fields, que considera somente o Coeficiente de Variação.

O procedimento da decomposição de Shapley pode ser entendido, conforme Wan e Zhou (2005), utilizando a seguinte função de rendimentos:

$$Y = f(X_1, \dots, X_k) \quad (2)$$

Normalmente, o vetor  $X$  é diferente para cada indivíduo (observações). Assim, substitui-se  $X_k$  por sua média amostral, eliminando qualquer diferença, em termos de  $X_k$ , entre os indivíduos. A renda resultante, chamada de  $Y_k$ , ainda difere entre os indivíduos, porém não mais em razão de  $X_k$ . Em outras palavras, a desigualdade em  $Y_k$ , denotada por  $I(Y_k)$  e que foi obtida, no caso deste trabalho, por meio do índice de Gini, de Theil-T e Theil-L, é causada por diferenças em  $X$ , excluída a variável  $X_k$ . Dessa forma, a contribuição da variável  $X_k$  na desigualdade total ( $C_k$ ) pode ser conseguida da seguinte forma:  $C_k = I(Y) - I(Y_k)$ , para  $k = 1, 2, \dots, K$ . O próximo passo é substituir, além de  $X_k$ ,  $X_j$  pela sua média amostral. Da mesma maneira que o passo anterior, é necessário obter a renda resultante,  $Y_{kj}$ . A contribuição do segundo passo, ou “segundo round”, pode ser obtida por  $C_k = I(Y_j) - I(Y_{jk})$  para  $k, j = 1, 2, \dots, K (k \neq j)$ . Por meio do mesmo processo, é obtida a contribuição do “terceiro round”,  $C_k = I(Y_{ij}) - I(Y_{ijk})$  para  $k, j, i = 1, 2, \dots, K (k \neq j \neq i)$ . O procedimento continua até que todo o vetor  $X$  seja substituído pela média amostral.

Na equação (2), a variável dependente é, geralmente, do tipo semi-log, o que torna necessário transformar a equação, depois de estimada, em um modelo linear para que a decomposição de Shapley seja aplicada em  $Y$  e não em  $\log(Y)$ .

### ■ 3.1 Viés de seletividade e endogeneidade

Um dos requisitos para que a decomposição de Shapley seja aplicada é, como já foi enfatizado, que haja uma equação de rendimentos. Assim sendo, há a necessidade de estimá-la. Porém, no processo de estimação, deve-se levar em conta a possibilidade de que apareçam alguns problemas econométricos, a saber: o viés de seletividade e a endogeneidade.

De acordo com Kassouf (1994), a estimação das equações de rendimentos, por meio do método dos Mínimos Quadrados Ordinários (MQO), pode gerar coeficientes viesados. A explicação para isso está na provável seletividade amostral, uma vez que é comum utilizar, no processo de estimação, somente indivíduos que estão empregados, isto é, que possuem algum rendimento proveniente do trabalho. Aqueles trabalhadores que não recebem renda do trabalho não seriam considerados nos cálculos. Para levar em conta essa possibilidade, Heckman (1979) desenvolveu um método que permite estimativas consistentes.

Assim como o viés de seletividade, o problema da endogeneidade pode gerar estimativas viesadas e inconsistentes. De acordo com Cameron e Trivedi (2009), a pressuposição fundamental para que os estimadores de MQO sejam consistentes é a não correlação entre o termo erro e os regressores. Caso isso não ocorra, não é possível fornecer uma interpretação causal no modelo estimado.

Dado que a endogeneidade e o viés de seletividade são problemas que podem causar estimativas viesadas e inconsistentes, Wooldridge (2002) propôs um método que busca corrigir, simultaneamente, esses dois possíveis casos. Esse método é uma extensão do procedimento de Heckman que leva em consideração o potencial endógeno de algumas variáveis, tais como a educação. O modelo pode ser especificado, segundo Wooldridge (2002), a partir de três equações (populacionais), isto é:

$$y_1 = Z_1\delta_1 + \alpha_1 y_2 + u_1 \quad (3)$$

$$y_2 = Z\delta_2 + v_2 \quad (4)$$

$$y_3 = 1(Z\delta_3 + v_3 > 0) \quad (5)$$

em que  $\delta_1$ ,  $\delta_2$ ,  $\delta_3$ , e  $\alpha_1$  são parâmetros a serem estimados; e  $u_1$ ,  $v_2$ , e  $v_3$ , são os termos de erros estocásticos. Na primeira equação,  $y_1$  é a variável de interesse – que, neste trabalho, representa o logaritmo natural dos rendimentos de todos os trabalhos do indivíduo  $i$  sobre o número de horas mensais trabalhadas – e  $y_2$  é a potencial variável endógena – educação, neste estudo. A equação (4) seria o primeiro estágio no método dos Mínimos Quadrados de Dois Estágios,

que tem como objetivo obter o instrumento a ser inserido na equação de rendimentos. A terceira equação, da mesma forma que no procedimento de Heckman, é a equação de seleção. Essa equação, que normalmente é obtida por meio de um modelo *probit*, busca levar em conta o possível viés de seleção.

Dentre os casos nos quais esse modelo se aplica, cita-se aquele em que  $y_2$  é sempre observado, mas é uma variável explicativa endógena na equação (3). Essa situação se ajusta perfeitamente ao que se pretende fazer neste trabalho, visto que  $y_1$  é o rendimento e  $y_2$  é a variável “anos de estudos”.<sup>3</sup> Se  $y_1$  e  $y_2$  fossem observados para cada variável do vetor  $Z$ , seria necessário estimar somente a equação (3) por Mínimos Quadrados de Dois Estágios (MQ2E), considerando  $y_2$  como endógeno.

Segundo Wooldridge (2002), esse método apresenta as seguintes pressuposições:

- a) o vetor  $Z$  e a variável  $y_3$  são sempre observáveis, porém  $y_1$  e  $y_2$  são observados somente quando  $y_3 = 1$  (o indivíduo participa do mercado de trabalho);
- b)  $u_1$  e  $v_3$  não são correlacionados com  $Z$ ;
- c)  $v_3$  possui distribuição normal, com média zero e variância constante;
- d)  $E(u_1|v_3) = \gamma_1 v_3$ ; e
- e)  $E(Z' | v_2) = 0$ .

Assumindo as pressuposições (a)–(e), a equação de rendimentos pode ser estimada, segundo Wooldridge (2002), por meio dos seguintes passos:

- a) Estime  $\delta_3$  por meio da regressão (*probit*) de  $y_3$  sobre  $Z$  usando a amostra completa. Em seguida, calcule a razão inversa de Mills.
- b) Utilizando a amostra cujas observações contêm somente indivíduos que trabalham ( $y_3 = 1$ ), estime a seguinte equação:

$$y_{1i} = Z_{i1}\delta_1 + \alpha_1 y_{i2} + \gamma_1 \hat{\lambda}_{i3} + erro_i \quad (6)$$

por MQ2E utilizando os instrumentos  $(Z_{i1}, \hat{\lambda}_{i3})$ .

3 Anos de estudos é sempre observado, independente se o indivíduo está inserido no mercado de trabalho ou não.

Em virtude de se utilizar, como regressor, uma variável estimada, os erros-padrão são inválidos (Wooldridge, 2002). Dessa forma, optou-se por utilizar as técnicas de bootstrap para gerar os erros-padrão.

Dada a falta de boas variáveis que podem ser utilizadas como instrumentos na PNAD, serão usadas variáveis que não estão nesse banco de dados. Assim, as variáveis que foram usadas como instrumentos, com base em Teixeira (2006) e Costa (2008), são as seguintes: a Lei nº 5.692, de 1971, e a oferta de professores.

O primeiro instrumento foi utilizado para refletir a mudança ocorrida após a Lei nº 5.692, de 1971. De acordo com Machado e Gonzaga (2007), o sistema educacional do Brasil era dividido, por volta da década de 1960, em duas fases: o ensino primário e o médio. O primeiro durava quatro anos, com início aos 7 anos de idade. O segundo tinha como objetivo a formação dos adolescentes e era dividido em duas fases: o ginásio, com duração de quatro anos; e o colegial, que tinha pelo menos três séries. Em 1971, a Lei nº 5.692 incorporou os quatro anos do ensino ginásial à escola primária, o que aumentou de quatro para oito anos o tempo mínimo de permanência na escola. Então, crianças entre 7 e 14 anos eram obrigadas a frequentar a escola. Para refletir essa alteração no sistema educacional, foi usada uma variável *dummy* que assume o valor 1 se o indivíduo nasceu em 1957 ou após isso; e 0, caso contrário.

O segundo instrumento refere-se ao número de professores do ensino fundamental em cada estado quando o indivíduo completa 7 anos. A escolha se efetivou a partir da ideia de que o aumento da oferta de professores está associado a maiores oportunidades de estudos e, por conseguinte, a mais anos de instrução.

Diante disso, a equação de seleção (modelo probit) foi estimada da seguinte forma<sup>4</sup>:

$$Z_i = \beta_0 = \beta_1 \text{Sexo}_i + \beta_2 \text{Cor}_i + \beta_3 \text{Tam} + \beta_4 \text{Id} + \beta_5 \text{Id}_i^2 + \beta_6 \text{Sudeste} + \beta_7 \text{Sul} + \beta_8 \text{Centro\_Oeste} + \beta_9 \text{Norte} + \beta_{10} \text{prof} + \beta_{11} \text{lei} + \varepsilon_i \quad (7)$$

em que:

$Z_i$  = variável binária que reflete a condição de atividade do  $i$ -ésimo indivíduo, isto é, 1 se faz parte da força de trabalho e 0, caso contrário;

---

4 É importante frisar que a variável educação não deve ser inserida como regressor no modelo *Probit*, visto que ela é a variável a ser instrumentalizada.

*Gênero* = variável binária para sexo, que assume valor 1 para homens e 0, mulheres;

*Cor* = variável binária para cor, que assume valor 1 para cor branca e 0, caso contrário;

*Tam* = número de componentes (pessoas) no domicílio;

*Id* = idade, medida em dezenas de anos;

*Id<sup>2</sup>* = idade ao quadrado, em dezenas de anos;

*Sudeste* = variável binária para região Sudeste, que assume valor 1 para Sudeste e 0, caso contrário;

*Prof* = oferta de professores em cada estado no ano em que o indivíduo faz 7 anos;

*Sul* = variável binária para região Sul, que assume valor 1 para Sul e 0, caso contrário;

*Centro\_oeste* = variável binária para região Centro-Oeste, que assume valor 1 para Centro-Oeste e 0, caso contrário;

*Norte* = variável binária para região Norte, que assume valor 1 para Norte e 0, caso contrário (Nordeste é a categoria de referência);

*Lei* = variável *dummy* que reflete a Lei nº 5.692, de 1971.

A partir da equação (7), encontra-se a razão inversa de Mills e a incorpora na equação de rendimentos (1), que já foi especificada anteriormente.

## ■ 3.2 Fonte de dados

A maior parte dos dados foi oriunda da Pesquisa Nacional por Amostra de Domicílios (PNAD), referente ao ano de 2014. O número de professores do ensino fundamental de cada estado foi coletado no *site* do Instituto Brasileiro de Geografia e Estatística (IBGE) [www.ibge.gov.br], especificamente no item estatísticas do século XX.

A decomposição baseada numa regressão (Shapley) envolve, dependendo do número de variáveis que são inseridas na equação de rendimentos, uma grande quantidade de cálculos, o que se torna extremamente trabalhoso caso seja feito manualmente. Dessa forma, a decomposição foi implementada por um programa Java desenvolvido pelo World Institute for Development Economics Research of the United Nations University (UNU-Wider).

## 4

# RESULTADOS E DISCUSSÃO

Levando em consideração a possibilidade de viés de seletividade e endogeneidade, estimou-se a equação de rendimentos para o Brasil. Os resultados podem ser vistos na Tabela 1<sup>5</sup>.

**Tabela 1**

**Equação de rendimentos para o Brasil, metodologia de Wooldridge, 2014**

Variáveis	Coefficientes	Erros-padrão
Constante	-1,0736	0,2181***
Anos de estudos	0,1400	0,0106***
Idade	0,5280	0,0448***
Idade ao quadrado	-0,0437	0,0056***
Gênero	0,4193	0,0356***
Cor	0,0596	0,0179***
Sudeste	0,1440	0,0197***
Sul	0,1968	0,0219***
Centro-Oeste	0,2386	0,0182***
Norte	0,1307	0,0185***
Empregador	0,5284	0,0266***
Conta própria	0,1297	0,0169***
Lambda	0,2407	0,0708***
Estatística Sargan	2,304 <sup>NS</sup>	
Estatística Durbin-Wu-Hausman	28,2060***	
Estatística Wu-Hausman	28,1934***	

Fonte: Elaborada pelo autor com base na PNAD de 2014.

Nota 1: Os erros-padrão foram calculados por meio de *bootstrap* (100 replicações).

Nota 2: \*\*\* significativo a 1%.

Nota 3: NS - Não significativo.

5 Os coeficientes da equação de participação estão na Tabela 1A do Apêndice.

No intuito de verificar a possível existência de endogeneidade na variável anos de estudos, foram executados os testes de exogeneidade de Durbin-Wu-Hausman e de Wu-Hausman. Em ambos os testes, como pode ser visto na Tabela 1, rejeita-se, ao nível de 1% de significância, a hipótese de exogeneidade. Portanto, há evidências de que a variável educação seja correlacionada com o termo de erro. Então, o método dos MQO não é apropriado para a estimação do modelo.

Como já mencionado, a equação de rendimentos foi estimada considerando dois instrumentos: número de professores e a Lei nº 5.692, de 1971. Cabe ressaltar que os dois instrumentos foram válidos, de acordo com a estatística de Sargan. Além disso, a razão inversa de Mills (lambda) mostrou-se significativa, indicando que sua inclusão era necessária para evitar o viés de seletividade.

Com relação aos coeficientes estimados, todos foram estatisticamente significativos ao nível de 1% e apresentaram os sinais de acordo com o que era esperado.

Conforme os resultados da Tabela 1, nota-se que o rendimento do trabalho é influenciado positivamente, como já era esperado, pela variável anos de estudos. A estimativa do coeficiente da variável mencionada também pode ser interpretada como a taxa de retorno da educação. Assim, para cada ano adicional de estudo, tem-se, no Brasil, um aumento de 15% nos rendimentos do trabalho (por hora)<sup>6</sup>.

No que se refere à idade, *proxy* para experiência, verifica-se que o coeficiente associado à variável idade é positivo, porém, o que é relacionado com a idade ao quadrado é negativo. Isso mostra que a relação entre experiência e rendimentos tem o comportamento do U-invertido, ou seja, à medida que o indivíduo ganha mais experiência em seu ambiente de trabalho, os rendimentos tendem a crescer; entretanto, ao atingir certo ponto, a renda passa a cair.

As estimativas dos coeficientes anos de estudos, idade e idade ao quadrado estão de acordo com a teoria clássica do capital humano, que estabelece uma relação positiva entre educação e rendimentos, e uma forma parabólica para a variável experiência. Os resultados confirmam o que já é praticamente unânime na literatura, isto é, o efeito positivo do capital humano sobre os rendimentos.

---

6 Para evitar qualquer tipo de imprecisão, Wooldridge (2006) sugeriu utilizar a fórmula  $100[\exp(x)-1]$  para calcular o acréscimo percentual nos rendimentos gerado por cada ano a mais de estudo, em que  $x$  é o coeficiente estimado da variável educação.

Quanto à variável gênero, observa-se que o diferencial de rendimentos, mesmo considerando os rendimentos por hora, ainda é grande. De acordo com as estimativas, homens recebiam, em média, 52,1% a mais do que as mulheres, descontados os efeitos oriundos dos outros regressores. Embora não seja possível afirmar com exatidão, dado o modelo utilizado, uma possível explicação é a discriminação no mercado de trabalho. Ueda e Hoffmann (2002) também encontraram, para o Brasil, estimativas elevadas para o coeficiente gênero. Os valores obtidos por esses autores, referentes ao ano de 2002, variaram de 0,4415 a 0,4506, dependendo do modelo. No meio rural do Brasil, Ney e Hoffmann (2009) verificaram que os homens recebiam uma renda aproximadamente 54% superior à renda das mulheres. Se comparado a esses trabalhos, houve uma redução no diferencial, porém, deve-se destacar que os procedimentos metodológicos não são iguais.

Quando a variável cor é considerada, nota-se que o diferencial de rendimentos entre os brancos e não brancos é bem menor do que o diferencial entre homens e mulheres. Conforme a estimativa, pessoas da cor branca recebem, em média, mantendo tudo o mais constante, 6,14% a mais do que os indivíduos da categoria base. O resultado também sugere que pode haver discriminação no mercado de trabalho com relação à cor da pessoa.

Em se tratando das regiões, observa-se que os indivíduos das regiões Sudeste, Sul, Centro-Oeste e Norte têm rendimentos, em média, superiores aos indivíduos da região Nordeste (categoria base). O resultado já era esperado, visto que o Nordeste é a região, no Brasil, com os piores indicadores sociais. Ainda que tenha melhorado nos últimos anos, o Nordeste ainda apresenta a maior proporção de pobres, a maior desigualdade de renda e a menor renda *per capita* do Brasil.

Deve-se destacar o resultado para o Centro-Oeste. A região apresentou o maior coeficiente, o que significa que os rendimentos dos indivíduos que vivem no Centro-Oeste são, em média, cerca de 27% maiores que os rendimentos dos indivíduos da região Nordeste. Uma possível explicação para isso pode estar na importância do agronegócio na região, principalmente na produção de grãos. Uma das principais características do agronegócio do Centro-Oeste é o seu caráter patronal, baseado em grandes propriedades e na *plantation*, destinada ao mercado externo (Guilhoto et al., 2007). Ademais, é um estado com alta renda domiciliar *per capita*, mas com elevada desigualdade de renda. Brasília, por exemplo, tem a maior renda domiciliar *per capita* do Brasil.



Observa-se também, na Tabela 1, que o indivíduo empregador possui rendimentos, em média, bem acima daquele que é empregado (categoria de referência). De acordo com a estimativa, a diferença é de cerca de 70%. Resultado similar, porém, em uma proporção menor, ocorre quando se consideram as pessoas que trabalham por conta própria.

Embora seja importante identificar os principais determinantes dos rendimentos das pessoas, saber quais fatores explicam a desigualdade de renda também é de grande relevância. Assim, por meio da equação de rendimentos (Tabela 1) foram calculadas as contribuições de cada uma das variáveis para a desigualdade. Para computar os níveis de iniquidade, foram usadas as três principais medidas, a saber, índice de Gini e os dois índices da classe de entropia generalizada: Theil-L e Theil-T. Da mesma forma que Gunatilaka e Ghotikapanich (2009), optou-se por trabalhar com algumas variáveis agregadas. Assim, a desigualdade foi decomposta em seis componentes: Educação, Idade (idade e idade ao quadrado), Gênero, Cor, Região (Sudeste, Sul, Centro-Oeste e Norte) e Posição (empregador e conta própria). Os resultados podem ser vistos na Tabela 2.

**Tabela 2**

**Decomposição (Shapley) baseada na equação de rendimentos estimada pelo método de Wooldridge (%), 2014**

Variáveis	Coefficiente de Gini	Coefficiente de Theil-L	Coefficiente de Theil-T
Educação	32,41	33,10	18,41
Idade	16,17	2,85	2,41
Gênero	6,00	3,43	2,32
Cor	0,77	0,50	0,30
Região	3,26	0,44	0,26
Posição	3,48	2,45	1,78
Resíduo	37,91	57,23	74,52
Total	100	100	100

Fonte: Elaborada pelo autor com base na PNAD de 2014.

Conforme os resultados, nota-se – quando o coeficiente de Gini é utilizado – que o modelo explica cerca de 62% do total da desigualdade de rendimentos, o que demonstra o bom poder explicativo do método de decomposição. Segundo Wan e Zhou (2005), modelos que explicam somente 30% a 40% da desigualdade de renda, deixando o restante para o resíduo, não geram estimativas seguras (confiáveis).

Dentre as variáveis em análise e levando em consideração o índice de Gini, educação é o principal fator para explicar a desigualdade de rendimentos no Brasil. De acordo com a PNAD de 2014, o coeficiente de Gini – aplicado sobre o rendimento total por hora<sup>7</sup> – foi de 0,57. Desse total, a educação contribuiu com 32,41%. Acrescentando a contribuição da idade (*proxy* para experiência), tem-se uma proporção de 48,58%, isto é, as variáveis relacionadas ao capital humano explicam quase 50% do índice de Gini.

O sinal positivo da variável educação já era esperado. Segundo Wan (2004), em razão do efeito marginal positivo da educação sobre os rendimentos (ver Tabela 1) e dada a correlação positiva entre anos de estudos e rendimentos, espera-se que a variável educação contribua positivamente para a desigualdade de rendimentos.

Ainda que tenha havido, nos últimos anos, melhorias em alguns indicadores da educação, o Brasil ainda enfrenta sérios problemas educacionais. Como exemplo, pode-se citar a taxa de analfabetismo. Conforme o Instituto de Pesquisa Econômica Aplicada (Ipea, 2016), cerca de 8,3% das pessoas com 15 anos ou mais, em 2014, são analfabetas, ou seja, não sabem ler, nem escrever um simples bilhete. Quando se considera a região Nordeste, essa proporção é ainda maior, aproximadamente 16,61%. A redução das desigualdades educacionais deve, portanto, ser encarada de um modo muito sério se a intenção for a redução, no longo prazo, da desigualdade de renda.

Por meio da variável gênero, observa-se que o diferencial de rendimentos existentes entre os sexos é um fator relevante para explicar a iniquidade distributiva. De acordo com a equação de rendimentos (Tabela 1), verificou-se que os rendimentos dos homens são, em média, 52% superiores aos das mulheres. Essa diferença impacta a desigualdade de renda, acentuando-a. A variável gênero é responsável por 6% da desigualdade, mensurada pelo coeficiente de Gini. Com relação à variável cor, a parcela do índice de Gini

---

7 Rendimentos de todos os trabalhos do indivíduo *i* sobre o número de horas mensais trabalhadas.

atribuída aos diferenciais de cor é de 0,77%, valor bem menor se comparado com a variável gênero.

As diferenças regionais, que no Brasil são bem acentuadas, têm uma parcela na formação da desigualdade de rendimentos. É notório que, no processo de desenvolvimento brasileiro, há regiões que se beneficiaram mais e outras, menos. Embora haja evidências de que em alguns períodos houve alguma convergência de renda entre as regiões (Ferreira, 2000), a disparidade regional ainda é grande, o que contribui, assim, para a desigualdade de renda. Como pode ser visto na Tabela 2, cerca de 3,3% do índice de Gini pode ser explicado pelas diferenças regionais. Valor similar foi encontrado para a variável posição na ocupação (3,48%).

A decomposição também foi conduzida por meio dos índices Theil-L e Theil-T. Como pode ser visto, a parcela da desigualdade que não foi explicada pelas variáveis explicativas – ou seja, o resíduo – foi bem maior do que aquela observada quando se utilizou o coeficiente de Gini. Para o coeficiente Theil-T, o resíduo é responsável por mais de 74% da variação na desigualdade, mostrando o baixo poder explicativo. Essa diferença nos resíduos também foi encontrada em Wan e Zhou (2005). Os autores usaram os coeficientes de Gini e de Theil-L no processo de decomposição. Em todas as suas estimativas, o resíduo gerado quando o coeficiente Theil-L foi utilizado ficou bem acima daquele verificado quando o índice de Gini foi a medida de desigualdade usada.

Ainda que os resíduos sejam grandes quando os coeficientes de Theil-L e Theil-T foram utilizados, as conclusões a respeito de quais variáveis contribuíram mais para explicar a desigualdade não se alteraram, ou seja, as variáveis relacionadas ao capital humano continuam sendo as principais causadoras da desigualdade de rendimentos.

As diferenças observadas, em termos de valores, quando distintas medidas são usadas, podem ser justificadas, conforme Wan e Zhou (2005), pela sensibilidade que cada índice possui nos diferentes segmentos da curva de Lorenz e pela utilização de distintas funções de bem-estar. Por exemplo, conforme Hoffmann (1998), o coeficiente Theil-T é mais sensível a mudanças nas rendas situadas no extremo direito da distribuição, ou seja, os relativamente mais ricos. Por outro lado, o índice Theil-L é mais sensível a alterações no extremo esquerdo da distribuição (os mais pobres). Por fim, o coeficiente de Gini é mais sensível a modificações nas rendas situadas no intervalo com maior densidade de frequência, isto é, em volta da moda ou mediana da distribuição.

## 5

# CONSIDERAÇÕES FINAIS

Historicamente, a renda no Brasil é distribuída de uma forma bem desigual. A partir do ano de 2001, a desigualdade de renda, mensurada com base nos dados da PNAD, entrou em uma trajetória decrescente e contínua. Todavia, o país ainda apresenta uma alta concentração de renda. Assim sendo, este trabalho buscou mensurar os principais determinantes da desigualdade de rendimentos no Brasil no ano de 2014.

Conforme os resultados, observou-se que as variáveis relacionadas ao capital humano são fundamentais para explicar os rendimentos do trabalho. A taxa de retorno da educação foi de aproximadamente 15%, o que mostra o alto impacto da escolaridade sobre os rendimentos. Além disso, anos de estudos foi a principal variável para explicar a desigualdade de rendimentos no Brasil. Ainda que tenha havido melhorias, em alguns aspectos, na educação, o país ainda enfrenta grandes desafios quando se trata de escolaridade. O número de pessoas que não sabem ler e escrever, por exemplo, permanece alto, o que contribui, conseqüentemente, com a concentração de renda. Possíveis soluções para este problema já vêm sendo propostas há tempos. Uma delas é o investimento em treinamento e qualificação dos mais pobres, o que tenderia a reduzir a desigualdade educacional, possibilitando, portanto, uma queda sustentável na desigualdade de renda. Além disso, dado o alto índice de analfabetismo no Brasil, deve haver políticas cuja prioridade máxima seja a alfabetização dos jovens e adultos.

Além das variáveis associadas ao capital humano, a variável gênero, embora em uma proporção menor, também foi importante para explicar a desigualdade. Ou seja, os diferenciais de rendimentos entre homens e mulheres contribuem, de uma forma considerável, para o coeficiente de Gini. Mesmo não sendo possível afirmar com exatidão com base na metodologia adotada neste trabalho, uma possível explicação é a discriminação no mercado de trabalho.

## DETERMINANTS OF INCOME INEQUALITY IN BRAZIL: AN APPROACH BASED ON SHAPLEY VALUE DECOMPOSITION

### Abstract

This study aimed to analyze the main determinants of individual income inequality in Brazil. For this, an earnings equation was estimated. We used a two-stage approach in order to control both for sample selection bias and endogeneity problems. After that, we used decompositions based on the Shapley value, a solution concept for cooperative games. The results show that education emerges as a most significant determinant of income inequality, no matter what inequality index was taken.

**Keywords:** Income inequality; Shapley decomposition; education; game theory; regression.

### Referências

- Alejos, L. (2003). *Contribution of the determinants of income inequality in Guatemala*. Cidade da Guatemala: Instituto de Investigaciones Económicas y Sociales, Universidad Rafael Ángel Landívar.
- Alesina, A., & Rodrik, D. (1991). Distributive politics and economic growth. *NBER Working Paper*, 3668.
- Alesina, A., & Perotti, R. (1996). Income distribution, political instability, and investment. *European Economic Review*, 4, 1203–1228.
- Angrist, J. D., & Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics*, 106(4), 979–1014.
- Barro, R. J. (2000). Inequality and growth in a panel of countries. *Journal of Economic Growth*, 5(1), 5–32.
- Barros, R. P., Carvalho, M., Franco, S., & Mendonça, R. (2007). *A queda recente da desigualdade de renda no Brasil*. (Texto para Discussão, nº 1258). São Paulo: IPEA.
- Berni, H. A. (2007). *Evolução dos determinantes da desigualdade de renda salarial no Nordeste*. Dissertação de mestrado, Coordenação do Programa de Pós-Graduação em Economia, Universidade Federal do Ceará, Fortaleza, Brasil.

- Cameron, A. C., & Trivedi, P. K. (2009). *Microeconometrics using stata*. College.
- Castelo Branco, R. C. (1979). *Crescimento acelerado e o mercado de trabalho: A experiência brasileira*. Rio de Janeiro: Fundação Getúlio Vargas.
- Costa, R. R. F. (2008). *O efeito da educação sobre estado de saúde individual no Brasil*. Dissertação de mestrado, Centro de Desenvolvimento e Planejamento Regional da Faculdade de Ciências Econômicas, Universidade Federal de Minas Gerais, 2008.
- Cowell, F. A., & Jenkins, S. P. (1995). How much inequality can we explain? A methodology and an application to the United States. *The Economic Journal*, 105(4290), 421–430.
- Cunha, M. S., & Vasconcelos, M. R. (2012). Evolução da desigualdade na distribuição dos salários no Brasil. *Economia Aplicada*, 16(1), 105–136.
- Duman, A. (2008). Education and income inequality in turkey: does schooling matter? *Financial Theory and Practice*, 32(3), 369–385.
- Ferreira, A. (2000). Convergence in Brazil: recent trends and long-run prospects. *Applied Economics*, 479–489.
- Ferreira, F., Leite, P., Litchfield, J., & Ulyssea, G. (2006). Ascensão e queda da desigualdade de renda no Brasil. *Econômica*, 8(1), 147–169.
- Fields, G. S., Cichello, P. L., Freije, S., Menéndez, M., & Newhouse, D. (2003). Household income dynamics: a four-country story. *The Journal of Development Studies*, 40(2), 30–54.
- Forbes, K. J. (2000). A reassessment of the relationship between inequality and growth. *The American Economic Review*, 90(4), 869–887.
- García-Peñalosa, C., & Turnovsky, S. J. (2006). Growth and income inequality: a canonical model. *Economic Theory*, 28(1), 25–49.
- Garner, T. I., & Terrell, K. (2001). *Some explanations for changes in the distribution of household income in Slovakia: 1988 and 1996*. Working Paper, 345. Washington, D.C.: U.S. Department of Labour, Bureau of Labour Statistics, Office of Prices and Living Conditions.
- Guilhoto, J. J. M., Azzoni, C. R., Silveira, F. G., Ichihara, S. M., Diniz, B. P. C., & Moreira, G. R. (2007). *PIB da agricultura familiar: Brasil-Estados*. Brasília: Nead Estudos.
- Gunatilaka, R., & Chotikapanich, D. (2009). Accounting for Sri Lanka's expenditure inequality 1980-2002: regression-based decomposition approaches. *Review of Income and Wealth*, 55(4), 882–905.
- Hammill, M. (2005). Income inequality in Central America, Dominican Republic, and Mexico: assessing the importance of individual and household characteristics. *Serie Estudios y Perspectivas*, 43, Mexico.
- Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1), 153–161.

Hoffmann, R. (1998). *Distribuição de renda, medidas de desigualdade e pobreza*. São Paulo: Edusp.

Instituto de Pesquisa Econômica Aplicada [IPEA]. (2016). *Banco de Dados* <http://www.ipeadata.gov.br>

Jin, Y. (2009). A note on inflation, economic growth, and income inequality. *Macroeconomic Dynamics*, 13(1), 138–147.

Kaldor, N. (1957). A model of economic growth. *Economic Journal*, 67(268), 591–624.

Kassouf, A. L. (1994). The wage rate estimation using the Heckman procedure. *Revista de Econometria*, 14(1), 89–107.

Kuznets, S. (1955). Economic growth and income inequality. *American Economic Review*, 45(1), 1–28

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

Lei, J. J. (2005). *Parametric and semiparametric estimations of the return to schooling of wage workers in Canada*. Dissertação de mestrado, Departamento de Economia da Universidade do Colorado, Estados Unidos da América.

Leibbrandt, M., & Woolard, I. (2001). The labour market and household income inequality in South Africa: Existing evidence and new panel data. *Journal of International Development*, 13(6), 671–689.

Leigh, A.; & Ryan, C. (2008). Estimating returns to education using different natural experiment techniques. *Economics of Education Review*, 27, 149–160.

Li, H. Y., & Zou, H. F. (1998). Income inequality is not harmful to growth, theory and evidence. *Review of Development Economics*, 2(3), 318–334.

Machado, D. C., & Gonzaga, G. (2007). O impacto dos fatores familiares sobre a defasagem idade-série de crianças no Brasil. *Revista Brasileira de Economia*, 61(4), 449–476.

Morley, S. (2001). *The income distribution problem in Latin America and the Caribbean*. Economic Commission of Latin America and the Caribbean (ECLAC). Libros de Cepal, Santiago.

NEY, M. G.; HOFFMANN, R. (2009). Educação, concentração fundiária e desigualdade de rendimentos no meio rural brasileiro. *Revista de Economia e Sociologia Rural*, 47(1), 147–181.

Persson, T.; & Tabellini, G. (1994). Is inequality harmful for growth? *American Economic Review*, 84(3), 600–621.

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

Salardi, P. (2005). *How much of Brazilian inequality can explain? An attempt of income differentials using PNAD 2002*. 2005. <http://www-3.unipv.it/webdept/prin/workp05.pdf>

Santos, V. F., & Vieira, W. C. (2013). Effects of growth and reduction of income inequality on poverty in Northeastern Brazil, 2003-2008. *Revista de Economia Aplicada*, 17(4), 647–666.

Santos, V. F., & Vieira, W. C. (2015). Income inequality in the urban and rural sectors of the northeast region of Brazil. *Review of Urban and Regional Development Studies*, 27, 134–147.

Shorrocks, A. F. (1999). *Decomposition procedures for distributional analysis: a unified framework based on the Shapley value*. [Mimeo]. University of Essex, Reino Unido.

Siqueira, M. L., & Siqueira, M. L. (2006, maio). Desigualdade de renda no Nordeste brasileiro: uma análise de decomposição. *Anais do Encontro Regional de Economia*, Anpec/Banco do Nordeste, 11.

Souza, M. I. A., Taques, F. H., Oliveira, J. C. T., & Alencar, D. A. (2013). Relação entre a desigualdade e educação no Brasil: uma estimativa de dados em painel (1995-2009). *Textos de economia*, 16(2), 111–142.

Teixeira, W. M. (2006). *Equações de rendimentos e a utilização de instrumentos para o problema de endogeneidade da educação*. Tese de doutorado, Departamento de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, Brasil.

Ueda, E. M., & Hoffmann, R. (2002). Estimando o retorno da educação no Brasil. *Revista de Economia Aplicada*, 6(2), 209–238.

Wan, G. (2004). Accounting for income inequality in rural China: a regression-based approach. *Journal of Comparative Economics*, 32(2), 348–363.

Wan, G., & Zhou, Z. (2005). Income inequality in rural China: regression-based decomposition using household data. *Review of Development Economics*, 9(1), 107–120.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, Massachusetts: MIT Press.

WOOLDRIDGE, J. M. (2006). *Introdução à econometria: Uma abordagem moderna*. São Paulo: Thompson Learning.



## APÊNDICE

**Tabela 1A**

**Equações de participação (modelo *Probit*) no mercado de trabalho brasileiro, 2014**

Variáveis	Coefficientes	Erros-padrão
Constante	-1,5835	0,0725***
Gênero	0,9611	0,0138***
Cor	0,0482	0,0138***
Idade	1,2194	0,0346***
Idade ao quadrado	-0,1680	0,0041***
Sudeste	0,0314	0,0244 <sup>NS</sup>
Sul	0,1352	0,0309***
Norte	0,0977	0,0236***
Centro-Oeste	0,1114	0,0252***
Tam	-0,0511	0,0042***
Lei	0,0899	0,0295***
Número de professores	3,23e <sup>-07</sup>	1,58e <sup>-07**</sup>

Fonte: Elaborada pelo autor com base na Pnad de 2014.

Nota 1: \*\*\* significativo a 1%.

Nota 2: \*\* significativo a 5%.

Nota 3: NS - Não significativo.