

Afinal, Quão Importantes são as Desigualdades de Escolaridade para Explicar as Disparidades Regionais de Renda no Brasil?

Area 10: Economia Regional e Urbana
JEL: R120. R110. J310.

Rodrigo Carvalho Oliveira¹
Raul da Mota Silveira Neto²

Resumo

Trabalhos recentes tem evidenciado que apesar da queda recente da desigualdade regional de renda no Brasil, esta ainda permanece bastante elevada. No entanto, a maioria destes trabalhos utiliza métodos insuficientes e focam apenas na desigualdade Sudeste - Nordeste. Neste sentido, o presente trabalho busca entender a desigualdade regional de renda no Brasil e analisar o papel da educação. Utilizou-se como estratégia metodológica a decomposição utilizando Recentered Influence Functions (RIF), proposta por Firpo, Fortin e Lemieux (2007, 2009). Os principais resultados indicam que os diferenciais de características produtivas são mais importantes para a compreensão do diferencial de rendimentos no Brasil do que o diferencial de retorno a estas características.

Palavras Chave: Desigualdade Regional. Educação. Decomposição.

Abstract

Recent work has shown that despite the recent fall in regional income inequality in Brazil, this still remains quite high. However, most of these studies use inadequate methods and focus only on inequality between Southeast and Northeast. In this sense, this paper aims to understand the regional income inequality in Brazil and analyze the role of education. We apply a decomposition method that use the Recentered Influence Functions (RIF) regression, proposed by Firpo , Fortin and Lemieux (2007 , 2009). The main results indicate that the differences in productive characteristics are most important for understanding the earnings differential in Brazil than the return differential to these characteristics.

¹Doutorando em Economia pela Universidade Federal de Pernambuco (PIMES), Professor Substituto da UEFS. Contato: (71) 8861-1025. rodrigo.coliveira13@gmail.com

²Professor da Universidade Federal de Pernambuco, Doutor em Economia pela USP, Bolsista de Produtividade do CNPQ. Contato: rau.silveira@uol.com.br

Keywords: Regional Inequality. Education. Decomposition

1. Introdução

Há muito tempo que as desigualdades regionais brasileiras são objeto de estudo e preocupação por parte dos formuladores de políticas públicas no Brasil. Segundo Diniz (2009), as primeiras atenções das autoridades brasileiras remetem ao século XIX, com a preocupação em relação às secas no Nordeste. No entanto, foi a partir do diagnóstico GTDN, elaborado por Celso Furtado em 1959, que documentou-se a desigualdade regional brasileira e fez-se um esforço teórico para entender as causas deste fenômeno e propor soluções para a resolução do mesmo, principalmente com a elaboração de um planejamento regional e a criação da Sudene.

Por outro lado, ao longo das décadas de 1970 e 1980 o interesse dos economistas pelos problemas regionais foram arrefecidos, sobretudo em virtude do sucesso do estado de bem estar social na Europa e, no Brasil, por causa da ditadura militar. Segundo Azzoni (1997), o interesse dos economistas pelo estudos das desigualdades regionais ressurgiu no início da década de 1990, quando o tema foi novamente colocado em pauta pela literatura de crescimento econômico, com a discussão sobre convergência de renda³.

Ferreira e Ellery Jr (1996) mostraram que existia um processo de convergência do PIB per capita entre os estados brasileiros e que este processo de convergência era mais lento do que aquele verificado entre os estados dos EUA. Ferreira (2000) realiza um estudo similiar, onde demonstra que há um processo de convergência de renda per capita entre os estados brasileiros nos anos de 1970 e 1986. No entanto, o autor destaca que este processo de convergência foi suavizado a partir do ano de 1986 até o ano de 1995. Trabalhos mais recentes, como o de Mossi et al (2003) e o de Gondim et al (2007) utilizam a metodologia desenvolvida por Quah (1997) para analisar o processo de convergência entre os estados brasileiros⁴. De um modo geral, estes trabalhos concluem que no Brasil se verifica um processo de convergência dentro de cada região, mas não se verifica este processo de convergência entre todos os estados do país.

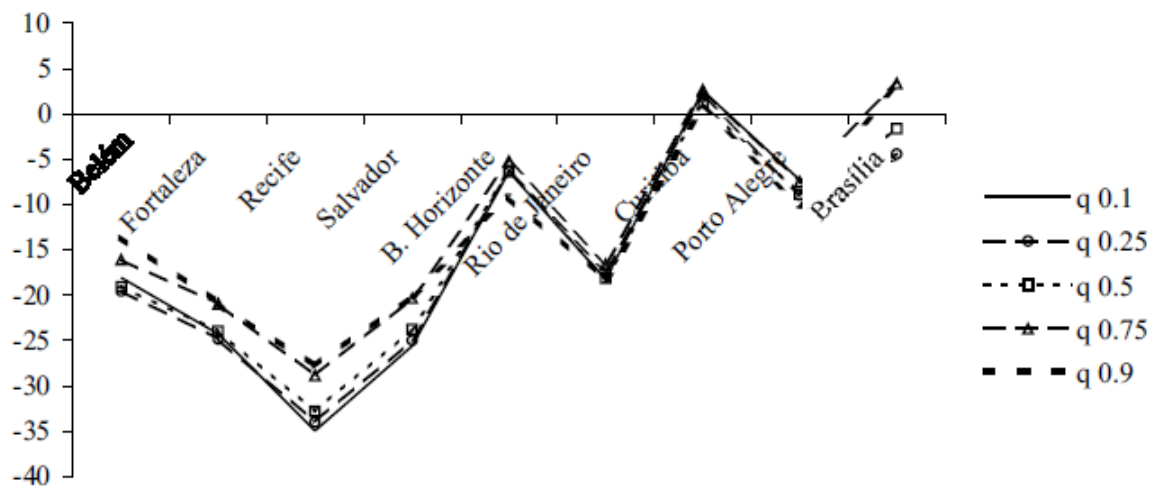
Mais recentemente, Silveira Neto e Azzoni (2012) e Oliveira e Silveira Neto (2013), mostraram que a redução recente das disparidades regionais de renda parece ser explicada tanto por movimentos associados à renda do trabalho, como por movimentos de renda associados aos programas de transferência de renda. Note-se que tais movimentos em favor de menores desigualdades regionais observadas para a renda do trabalho, a princípio, são consistentes, por um lado, com a conhecida importância da educação para entender os níveis de desigualdade regional de renda no Brasil (DUARTE, FERREIRA e SALVATO, 2003; SILVEIRA NETO e MENEZES, 2008) e, por outro, com o aumento do nível de escolaridade dos indivíduos presente em todas as unidades da federação (mudanças na oferta de trabalho).

Tais trabalhos contribuem para o entendimento da dinâmica da desigualdade, mas pouco elucidam a respeito dos condicionantes dos níveis de desigualdade. Neste sentido, há um conjunto de trabalhos que tratam especificamente deste ponto; pela relação direta com os resultados deste trabalho, quatro são destacados aqui. Silveira Neto e Menezes (2008) aplicaram o método de decomposição Oaxaca-Blinder para estudar as disparidades de renda entre o Nordeste e o Sudeste, concluindo que a educação dos indivíduos possui um papel fundamental para a explicação das

³sobretudo a partir dos trabalhos de Baumol (1986), Barro (1991), Barro e Sala-i-Martin (1992) e Mankiw et al (1992).

⁴Outros trabalhos aplicando estas metodologias de convergência para unidades geográficas brasileiras de convergência dentro de cada região, mas não se verifica este processo de convergência entre todos os estados do país.

Figura 1: Diferença de rendimentos entre São Paulo e as demais regiões metropolitanas brasileiras por quantil.



Fonte: Campelo e Silveira Neto (2003)

disparidades de rendimentos médios. Porém, ainda são escassas as evidências de trabalhos que buscam analisar o diferencial regional de renda para diferentes pontos da distribuição dos rendimentos além da média. Um primeiro esforço neste sentido foi o trabalho de Silveira Neto e Campelo (2003), o qual aplicou regressões quantílicas para analisar o diferencial de rendimentos entre as principais regiões metropolitanas brasileiras. As conclusões deste trabalho indicam que há grandes variações na desigualdade por quantis de renda, sendo estas maiores para os quantis inferiores - indivíduos mais pobres - e menores para os quantis superiores - indivíduos mais ricos. Além disto, as diferenças ainda são mais evidentes quando compara-se a região metropolitana de São Paulo com as regiões metropolitanas do Nordeste (Figura 1).

Duarte, Ferreira e Salvato (2003) mostram, a partir de um exercício contrafactual utilizando o método desenvolvido por DiNardo, Fortin e Lemieux (1996), que as disparidades regionais de renda no Brasil poderiam ser atenuadas caso o nível de educação do Nordeste fosse próximo àquele verificado no Sudeste do Brasil. Os autores realizam esta mesma análise para os estados de São Paulo e Ceará, mostrando que as disparidades de renda entre estes estados também poderiam ser atenuadas caso o nível de escolaridade dos indivíduos no Ceará fosse maior. No entanto, o método utilizado não permite calcular uma decomposição detalhada nos moldes da tradicional decomposição Oaxaca Blinder, o que significa que não é possível atribuir e separar os efeitos das variáveis entre efeitos composição e o efeito retorno das características. Deste modo, não é possível saber, por exemplo, se o papel da escolaridade para a análise das disparidades ocorre, pois o nível de escolaridade entre as regiões é diferente ou se o retorno a educação entre as regiões é diferente.

Guimarães, Cavalcanti e Silveira Neto (2006) realizam uma análise que supera parte das limitações do trabalho anterior, através da aplicação do método desenvolvido por Machado e Mata (2005). Com resultado diferente daquele de Duarte et al (2003), a principal conclusão deste trabalho é que os retornos a escolaridade possuem um papel muito mais importante na explicação das disparidades regionais do que as diferenças em termos de nível de escolaridade *per se*. No entanto, este método possui algumas limitações. Primeiramente, ele é bastante intensivo computacionalmente. Além disto ele não permite realizar uma decomposição detalhada robusta do componente

composição e, por fim, não utiliza uma estrutura de regressão⁵.

Note-se que, estes trabalhos se limitam a analisar a desigualdade entre as regiões Sudeste e Nordeste. Neste sentido, este artigo busca trazer elementos novos para a análise da desigualdade de renda regional brasileira por três motivos. A primeira e mais importante contribuição é a utilização da decomposição baseada na *Recentered Influence Function Regression (RIF)* (FIRPO, FORTIN, LEMIEUX, 2007, 2009), um método recente e ainda não utilizado para analisar a disparidade regional de renda no Brasil. Além de ser de fácil implementação computacional, este método possui diversas vantagens em relação aos outros desta literatura⁶. A mais importante delas é o fato de permitir atribuir um efeito composição e um efeito estrutura salarial para cada variável determinante do nível de renda das economia ao longo de diferentes pontos da distribuição de renda, nos moldes da tradicional decomposição de Oaxaca-Blinder, algo não possível com os métodos até aqui aplicados para estudo das disparidades regionais no Brasil. A segunda contribuição refere-se a análise da desigualdade regional das quatro regiões brasileiras em relação a Região Sudeste, trazendo novos elementos para a compreensão do problema. Os dados para esta análise foram obtidos utilizando o Censo Demográfico para o ano de 2010, o qual permite investigar uma série de fatores, individuais, regionais, e do ambiente de modo a buscar quais são as principais explicações das disparidades regionais brasileiras.

Além desta introdução, este artigo possui mais quatro seções. Na segunda seção apresentaremos as evidências teóricas que embasam nosso trabalho. Na terceira será apresentada o método de decomposição que utiliza a RIF e a base de dados. Na quarta são apresentados os resultados e na quinta seção a conclusão do trabalho.

2. Referencial Teórico

Do ponto de vista teórico, as desigualdades regionais são explicadas tanto por fatores relacionados a oferta de trabalho como por fatores relacionados a demanda por trabalho (MENEZES e AZZONI, 2006). Os enfoques gerais ao tratamento das desigualdades regionais na literatura, de forma geral, se diferenciam por enfatizar mais ou menos estes dois conjunto de fatores. Duas linhas gerais de tratamento são aqui destacadas: os enfoques baseados no capital humano e nos modelos de crescimento e os enfoques específicos da Economia Regional, que também enfatizam a importância de atributos locais.

2.1. Educação e Crescimento Econômico

O primeiro passo para compreender o papel da educação sobre a desigualdade regional de renda é entender quais os mecanismos pelos quais a educação afeta a renda de um indivíduo. Atualmente, esta é uma literatura bastante extensa em economia, mas os trabalhos pioneiros de maior relevância começaram a surgir no início da década de 1960, com Schultz (1960; 1961) e Becker (1962) que passaram a destacar a educação como uma forma de investimento. Este investimento, definido como capital humano, mesmo sem poder ser vendido é uma forma de capital, pois provê uma produtividade com valor econômico. Mais ainda, Schultz (1961) defendia que as diferenças salariais entre grupos, raças e estados americanos podem ser explicadas pelas diferenças de capital humano entre os agentes econômicos (FILHO, PESSOA, 2010).

⁵Estes problemas serão detalhados na seção metodológica

⁶DiNardo, Fortin e Lemieux(1996) e Machado e Mata (2005)

Em 1970, o trabalho de Mincer aparece como mais uma contribuição da educação para explicar as regularidades no mercado de trabalho, em que pese, os diferenciais de salários entre os indivíduos. A idéia de Mincer pode ser simplificada utilizando uma equação que busca apreender como são determinados os salários dos indivíduos:

$$\ln w = \beta_0 + \beta_1 educ + \beta_2 Exper + \beta_3 Exper^2 + X_i \beta_i \quad (1)$$

Onde β_0 seria o salário de um indivíduo desqualificado, β_1 o retorno a educação, β_2 e β_3 os retornos associados a experiência e β_i o retorno associado às demais i características dos indivíduos.

Atualmente, grande parte da literatura em economia da educação e economia do trabalho busca encontrar um valor causal para β_1 . A dificuldade maior reside no problema de viés de variável omitida, principalmente pela dificuldade de controlar o efeito da habilidade dos indivíduos. Dentre as tentativas de resolver este problema destacam-se os trabalhos de Angrist e Krueger (1991), que utilizam o trimestre de nascimento do indivíduo nos EUA como variável instrumental e Ashenfelter e Rouse (1998) e Rouse (1997), que realizam estimativas a partir de uma base de dados com indivíduos gêmeos⁷.

A importância desta linha de pesquisa para a compreensão das desigualdades regionais reside na seguinte explicação: indivíduos com maiores níveis educacionais tendem a ter filhos mais educados, de modo que seus filhos também possuirão salários mais elevados. Portanto, se estes indivíduos estiverem espacialmente concentrados, então a geração da riqueza deverá ocorrer de modo desigual entre diferentes regiões. Outro argumento é que, dada a possibilidade de migração, só pode haver diferença na renda per capita entre as pessoas de diferentes regiões caso as características dos trabalhadores sejam diferentes. Neste sentido, pode haver um problema regional apenas devido a um motivo social, isto é, uma região pode ser mais pobre em virtude das características embutidas nos moradores desta (PESSOA, 2001).

A importância da literatura de crescimento para a explicação das desigualdades regionais reside no fato de ter caminhado no sentido de explicar se existe uma tendência de convergência das rendas per capita entre os países. A justificativa para tal investigação é a conclusão do modelo de Solow de que economias com menor estoque de capital per capita tendem a crescer a uma taxa mais elevada do que economias que já se encontram mais próximas do nível de equilíbrio de longo prazo.

No entanto, os trabalhos de Baumol (1986), Barro (1991) e Barro e Sala-i-Martin (1992) indicam que a hipótese de convergência não pode ser verificada de maneira geral para todos os países ou regiões. Em outras palavras, quando se toma uma amostra de países com características bastante diferentes não é possível verificar convergência entre eles. Esta situação, na qual os países possuem níveis de preferências e tecnologia bastante diferentes, ficou conhecido na literatura como *convergência absoluta*. Já quando se analisam países com características econômicas parecidas, leia-se tecnologia e preferências, o nível de produto per capita em estado estacionário destes países tendem a se igualar. Este processo ficou conhecido na literatura como *convergência condicional*. Neste caso, nos dois últimos trabalhos citados, os autores encontraram evidências deste processo de convergência para regiões geográficas com características parecidas, como os estados dos EUA, regiões da França e Distritos do Japão.

Um problema destes trabalhos era o fato de apesar de o modelo de Solow conseguir descrever razoavelmente a trajetória de crescimento de uma economia, ele possuía pequena capacidade de descrever os diferenciais de renda entre duas economias. Mankiw Romer e Weil (1992) introduzem o capital humano explicitamente na função de produção e conseguem explicar de forma mais satisfatória as diferenças de renda entre diferentes economias. Os autores mostram que o processo

⁷O survey de Card (1999) possui informações detalhadas sobre a evolução desta literatura.

de convergência entre grupos de países⁸ se apresenta mais rápido com a introdução do capital humano no modelo. Portanto, esta é uma primeira evidência empírica da importância do capital humano, com destaque para a escolaridade das pessoas, sobre a distribuição de renda entre regiões específicas.

Como foi colocado na introdução, estes modelos também podem ser aplicados a diferentes unidades geográficas, tornando-se de grande importância para explicar os diferenciais regionais de renda dentro de países ou regiões, sobretudo daqueles com grande extensão territorial - como no caso do Brasil -, o que ocasiona uma grande diversidade cultural, econômica e social.

2.2. Economia Regional

A literatura clássica em economia regional destaca que há uma tendência de o processo de desenvolvimento econômico não ocorrer de maneira uniforme entre diferentes regiões econômicas (Hirschman, 1977; Myrdal, 1957; Perroux, 1977). Segundo Hirschman (1977), o progresso econômico não ocorre em toda a parte, e existem forças que levam a uma concentração espacial do crescimento econômico. Já Perroux (1977), defende a idéia de que o crescimento econômico não aparece simultaneamente em toda a parte, mas sim em pontos ou pólos de crescimento. No entanto, estes autores sugerem que o problema da concentração pode ser superado, caso existam efeitos de fluência (usando a terminologia de Hirschman) suficientemente fortes. Isto é: é possível que se desenvolvam nas regiões deprimidas atividades complementares àquelas localizadas nas regiões desenvolvidas de modo que externalidades positivas possam fluir de uma região para a outra. Por fim, Myrdal (1957) argumenta que existe um processo de causalção circular cumulativa, onde as regiões deprimidas tendem a continuar deprimidas, enquanto as regiões ricas tendem a apresentar um maior crescimento econômico.

Já a Nova Geografia Econômica estuda como forças centrífugas e centrípetas determinam a decisão de localização das firmas e dos trabalhadores entre regiões. Segundo Ruiz (2003) os principais autores desta linha de pesquisa são Paul Krugman, Masahisa Fujita e Anthony Venables e suas idéias são sintetizadas no livro *The Spatial Economy* (1999). Uma contribuição importante desta literatura é a tentativa de explicar a organização da produção a partir de uma lógica micro-econômica, sobretudo com a incorporação das teorias de retorno crescente, tal como o modelo de competição monopolística de Dixit-Stiglitz (1997).

Nocco (2004) sintetiza quais seriam os fatores causadores das forças centrípetas (ou de aglomeração) e das forças centrífugas (ou de dispersão). As primeiras seriam: i) a mobilidade dos trabalhadores quando o setor final possui retornos crescentes; ii) encadeamentos para trás e para frente quando as firmas produzem bens intermediários e finais, considerando que os bens finais são produzidos com retornos crescentes; e iii) vantagens tecnológicas da produção em uma determinada região. Já no segundo caso: i) imobilidade dos fatores de produção; ii) forte competição por um fator produtivo limitado; e iii) spillovers de conhecimento tecnológico de uma região que possui um setor mais moderno para regiões menos desenvolvidas.

Krugman (1991), por exemplo, demonstra a partir de um modelo com apenas dois tipos de trabalhadores, trabalhadores na manufatura e camponeses, sendo os primeiros móveis e os segundos imóveis, e incorporando retornos crescentes ao nível da firma e custos de transportes, que pode haver um diferencial salarial entre duas regiões a depender da magnitude dos custos de transportes a da parcela da população empregada na manufatura em cada região. Krugman argumenta que baixos custos de transporte favorecem a concentração da atividade econômica em apenas uma

⁸E neste estudo se destaca o grupo apenas com os países da OCDE.

região (centro) e, neste caso, os trabalhadores (camponeses) da outra região (periferia) receberiam salários mais baixos que os trabalhadores (na manufatura e camponeses) no centro.

Fujita e Thisse (2002) apresentam um modelo teórico que auxilia na compreensão da relação entre educação e desigualdade regional de renda. Este modelo considera duas regiões, um produto e dois fatores de produção: trabalhadores qualificados e trabalhadores não qualificados. Considera-se apenas o primeiro fator como móvel e não se considera a existência de custos de transporte. A partir deste modelo, os autores mostram que, mesmo as regiões possuindo dotações iniciais dos dois fatores iguais, haverá um equilíbrio com concentração dos trabalhadores qualificados em uma das regiões e este equilíbrio equalizará os salários dos trabalhadores qualificados entre as duas regiões. Por fim, este equilíbrio implica que os salários dos trabalhadores não-qualificados da região com mais trabalhadores qualificados serão maiores, gerando disparidades de renda entre as duas regiões.

Rice e Venables (2003) desenvolvem um modelo e estudam a desigualdade regional no Reino Unido. Este trabalho é muito importante pois mostra que grandes disparidades regionais podem ocorrer mesmo em uma economia de mercado funcionando bem, sem barreiras à mobilidade ou externalidades espaciais. Neste modelo, a desigualdade regional pode ocorrer caso algumas regiões possuam melhor acesso a mercado do que outras. Isto é, as duas regiões produzem um bem que é transacionado dentro de um mesmo país. Porém, uma região tem menor custo de transporte do que a outra, seja por condições geográficas (proximidade de maiores mercados consumidores) ou estruturais (presença de melhores portos, aeroportos, etc). Neste caso, tanto a produção per capita quanto o salário da região com menor custo de transporte serão maiores. Em outra análise, os autores mostram que estas disparidades podem ser ainda mais elevadas caso algumas regiões sejam especializadas na produção de bens para o mercado externo, produção esta caracterizada pela presença de externalidades tecnológicas e pecuniárias, e pela presença de trabalhadores altamente qualificados. Os autores destacam, também, que parece improvável que diferenças de capital humano, apenas, sejam capazes de explicar a grande disparidade de renda entre regiões. Portanto, os mecanismos que geram disparidades regiões nesta análise são as diferenças na composição da força de trabalho, da infra estrutura e do mix de indústrias entre as regiões.

Sendo Assim, diferentemente da literatura em crescimento econômico, esta ultima abordagem enfatiza a questão do espaço para a geração de riqueza dentro de uma região. Isto é, cada região sofre a influência de processos socioeconômicos particulares, que estão diretamente relacionados a sua localização e a suas estruturas produtivas, e não apenas o grau de capital humano presente na mesma. Ou seja, amenidades ambientais locais, a localização em relação a regiões economicamente dinâmicas e a presença de infraestrutura adequada também são condições fundamentais para o processo de desenvolvimento econômico e, conseqüentemente, para o grau de desigualdade de renda regional.

3. Método

Este artigo utilizará o método de regressão que utiliza a *Recentered Influence Function* para decompor as diferenças nas distribuições de rendimentos entre dois grupos distintos, baseado nos trabalhos de Firpo, Fortin e Lemieux (2007, 2009). Este método se insere na literatura de decomposição em economia, especialmente na área de economia do trabalho, a partir dos papers seminais de Oaxaca (1973) e Blinder (1973).

Enquanto o método original Oaxaca-Blinder possibilita a decomposição do diferencial de rendimento médio entre dois grupos, os métodos mais recentes permitem a decomposição do diferencial de rendimento entre dois grupos para medidas de distribuição além da média, tais como quantis,

variância e Gini (DINARDO, FORTIN e LEMIEUX, 1996; MACHADO E MATA, 2005; FIRPO, FORTIN E LEMIEUX, 2007, 2009).

A partir do método proposto, será possível decompor a desigualdade de renda entre as regiões Nordeste e Sudeste, no efeito composição e o efeito retorno salarial para qualquer ponto da distribuição dos rendimentos. Normalmente, o Efeito Composição é dito como a parte explicada da decomposição, enquanto o Efeito Estrutura Salarial é conhecido como a parte não explicada, ou devida a “discriminação”. Estes dois efeitos possuem fáceis interpretações. Enquanto o primeiro reflete o efeito sobre os rendimentos pelo fato dos indivíduos de cada região possuírem características diferentes, o segundo reflete o efeito sobre os rendimentos decorrente das diferenças de retorno às características de indivíduos similares, mas em grupos diferentes. No caso deste estudo, suponha que deseja-se avaliar a diferença de rendimentos entre as regiões A e B. O efeito composição irá avaliar em cada quantil, por exemplo, a diferença de rendimentos entre as regiões dada por diferenças nas características observadas dos indivíduos, tais como idade e educação. Já o efeito salarial irá captar o efeito dos retornos às variáveis sobre a diferenças nos rendimentos médios como, por exemplo, se o retorno à educação é diferente entre as duas regiões.

3.1. Decompondo as diferenças de rendimentos utilizando Recentered Influence Function

O método *Recentered Influence Function (RIF)* proposto por Firpo, Fortin e Lemieux (2009), permite generalizar a decomposição Oaxaca-Blinder para qualquer medida de distribuição, tal como a média, mediana, quantis, variância e índice de Gini. Este método possui algumas vantagens em relação a outras metodologias de decomposição⁹. Primeiramente ele baseia-se numa estrutura de regressão, de modo que é fácil de ser interpretado na prática, tal como o método Oaxaca-Blinder. Em segundo lugar, ele permite a decomposição detalhada, tanto do efeito composição, quanto do efeito retorno, o que não é possível através dos métodos de decomposição para além da média existentes. Em terceiro, ele possui a propriedade de ser *path independent*¹⁰. (FIRPO, FORTIN E LEMIEUX, 2010).

A realização da decomposição é levada a efeito em 2 estágios. O primeiro estágio consiste em decompor a estatística de distribuição de interesse entre o efeito composição e o efeito retorno utilizando uma abordagem de reponderação, onde os pesos são calculados de forma paramétrica e não paramétrica. Já no segundo estágio utiliza-se a regressão RIF para obter a contribuição de cada variável sobre o efeito composição e sobre o efeito retorno.

No primeiro estágio é necessário obter a distribuição contrafactual de interesse $F_{Y_A^c}$, a qual possibilita a realização da decomposição agregada, pois permite estimar diretamente os efeitos composição e estrutura salarial. A abordagem proposta por Firpo, Fortin e Lemieux (2007) é bastante similar ao processo de reponderação proposto por Dinardo, Fortin e Lemieux (1996). Os pesos de reponderação são computados como:

$$\hat{w}_1(T) = \frac{T}{\hat{p}}$$

e

$$\hat{w}_0(T) = \frac{1 - T}{1 - \hat{p}}$$

⁹Oaxaca (1973), Blinder (1973), Machado e Mata (2005) e Dinardo, Fortin e Lemieux (1996).

¹⁰Diz-se que uma decomposição é *path independent* quando a ordem na qual os diferentes elementos da decomposição detalhada são calculados não afeta os resultados da decomposição.

com $\hat{p} = N^{-1} \sum_{i=1}^N T_i$, e T indicando se o indivíduo participa do grupo 0 ou 1. No caso deste trabalho, o grupo zero será representado, sempre, pela região Sudeste, enquanto o grupo 1 será uma das outras quatro regiões brasileiras. Posteriormente a função de reponderação $w_c(T, X)$ é estimada:

$$\hat{w}_c(T, X) = \frac{(1 - T)}{\hat{p}} \cdot \frac{\hat{p}(X)}{1 - \hat{p}(X)}$$

Onde \hat{p} é um estimador da verdadeira probabilidade de estar no grupo 1, dado o vetor de características observáveis X . E, para obter os pesos somando 1, são utilizados os seguintes procedimentos de normalização:

$$\begin{aligned}\hat{w}_1^*(T_i) &= \frac{\hat{w}_1(T_i)}{\sum_{j=1}^N \hat{w}_1(j)} = \frac{T_i}{N \cdot \hat{p}} \\ \hat{w}_0^*(T_i) &= \frac{\hat{w}_0(T_i)}{\sum_{j=1}^N \hat{w}_0(j)} = \frac{1 - T_i}{N \cdot (1 - \hat{p})} \\ \hat{w}_c^*(T_i, X_i) &= \frac{\hat{w}_c(T_i)}{\sum_{j=1}^N \hat{w}_c(j)} = \frac{(1 - T_i) \cdot \frac{\hat{p}(X_i)}{1 - \hat{p}(X_i)}}{\sum_{j=1}^N \frac{\hat{p}(X_j)}{1 - \hat{p}(X_j)}}\end{aligned}$$

A partir destes pesos, é possível computar as distribuições acumuladas dos rendimentos como:

$$\hat{F}_t(y) = \sum_{i=1}^N \hat{w}_t^*(T_i) \cdot \mathbb{1}\{Y_i \leq y\}$$

e

$$\hat{F}_c(y) = \sum_{i=1}^N \hat{w}_c^*(T_i, X_i) \cdot \mathbb{1}\{Y_i \leq y\}$$

Para exemplificar a utilização destes pesos podemos demonstrar o caso da estimação de diferenças médias entre dois grupos $t = 0$ e $t = 1$. Suponha que deseje-se estimar

$$Y_{ti} = X_i \beta_i + \epsilon_{ti}$$

No primeiro estágio utilizaríamos os pesos para computar $\hat{\mu}_t$ e $\hat{\mu}_c$:

$$\hat{\mu}_t = N^{-1} \sum_{i=1}^N \hat{w}_t(T_i) Y_i$$

e

$$\hat{\mu}_C = N^{-1} \sum_{i=1}^N \hat{w}_C(T_i, X_i) Y_i$$

Em seguida seria possível estimar as diferenças entre os grupos:

$$\hat{\Delta}_o^\varphi = \hat{\mu}_1 - \hat{\mu}_0 = \hat{\Delta}_s^\varphi + \hat{\Delta}_x^\varphi$$

com $\hat{\Delta}_s^\varphi = \hat{\mu}_1 - \hat{\mu}_C$ representando o efeito estrutura salarial e $\hat{\Delta}_x^\varphi = \hat{\mu}_C - \hat{\mu}_0$ representando o efeito composição.

Já no segundo estágio, calcula-se a contribuição de cada variável sobre o efeito efeito composição ($\hat{\Delta}_x^u$) e sobre o efeito estrutura salarial ($\hat{\Delta}_s^u$), tal como na tradicional decomposição Oaxaca-Blinder. Neste estágio é necessário compreender o conceito de *Influence Function* (IF).

Seja uma estatística de distribuição $v(F_y)$, então $IF(Y; v, F_y)$ representa a influência de uma observação individual sobre esta estatística de distribuição. Posteriormente, adicionando-se a estatística $v(F_y)$ de volta na função de influência produz o que os autores chamam de *Recentered Influence Function* (RIF), a qual é a maior contribuição de Firpo, Fortin e Lemieux (2007) e o que diferencia o trabalho destes autores dos outros antecedentes.

Especificamente,

$$RIF(y; v) = v(F) + IF(y; v)$$

Uma propriedade bastante importante desta, é o fato de sua expectativa condicional ser igual a $v(F)$:

$$\int RIF(y; v).dF(y) = \int (v(F) + IF(y; v)).dF(y) = v(F)$$

Este segundo estágio é realizado em duas etapas. A primeira é calcular a *recentered influence function* para a estatística de distribuição de interesse. Firpo, Fortin e Lemieux (2007) mostram como calcular esta medida para a média, quantis, variância e o índice de Gini. Vejamos, por exemplo, para o caso da média e dos quantis:

- Média:

$$IF(y; \mu) = \lim_{\epsilon \rightarrow 0} \frac{[(1 - \epsilon).\mu + \epsilon.y - \mu]}{\epsilon} = y - \mu$$

$$RIF(y; \mu) = IF(y; \mu) + \mu$$

- Quantis:

$$IF(y; q_\tau) = \frac{\tau - \mathbb{1}\{y \leq q_\tau\}}{f_Y(q_\tau)}$$

$$RIF(y; q_\tau) = q_\tau + IF(y; q_\tau) \tag{2}$$

O caso da mediana (me) exemplifica de forma simples:

$$IF(y; me) = \frac{1/2 - \mathbb{1}\{y \leq q_{me}\}}{f_Y(q_{me})}$$

$$RIF(y; me) = me + \frac{1/2 - \mathbb{1}\{y \leq q_{me}\}}{f_Y(q_{me})}$$

Para caso dos quantis, o qual é o maior interesse deste estudo, primeiramente estima-se o quantil amostral \hat{Q}_τ tal como em Koenker e Basset (1978) e a densidade $f(\hat{Q}_\tau)$ neste ponto utilizando funções Kernel¹¹, posteriormente substitui-se estas estimativas na equação (2). Por fim, é estimado

¹¹Ver FIRPO, FORTIN E LEMIEUX (2007)

o efeito da mudança na distribuição de uma variável sobre o sobre o quantil marginal t de Y, numa especificação que toma a forma

$$E[RIF(Y_i; Q_\tau)|X_i] = X_i\gamma_i$$

E o parâmetro de interesse pode ser calculado via OLS:

$$R\hat{I}F(Y_i; Q_\tau)|X_i] = X_i\hat{\gamma}_i$$

Onde $\hat{\gamma}$ captura o “efeito parcial do quantil não condicional” de X. Sejam γ_t^v , com $t \in [1, 0]$ os parâmetros estimados da regressão e γ_c^v o parâmetro da regressão contrafactual estimada, então utilizando os pesos de reponderação e a RIF encontrada, pode-se estimar estes parâmetros como:

$$\hat{\gamma}_t^v = \left(\sum_{i=1}^N \hat{w}_t^*(T_i) \cdot X_i \cdot X_i^T \right)^{-1} \cdot \sum_{i=1}^N \hat{w}_t^*(T_i) \cdot R\hat{I}F(Y_i; v_t) \cdot X_i$$

$$\hat{\gamma}_c^v = \left(\sum_{i=1}^N \hat{w}_c^*(T_i, X_i) \cdot X_i \cdot X_i^T \right)^{-1} \cdot \sum_{i=1}^N \hat{w}_c^*(T_i, X_i) \cdot R\hat{I}F(Y_i; v_t) \cdot X_i$$

Com $R\hat{I}F(y; v_t) = \hat{v}_t + \hat{I}F(y; v_t)$ e $R\hat{I}F(y; v_c) = \hat{v}_c + \hat{I}F(y; v_c)$.

Por fim, é possível decompor os Efeitos Dotação ($\hat{\Delta}_X^\varphi$) e Efeito Estrutura dos Salários ($\hat{\Delta}_S^\varphi$) como:

$$\hat{\Delta}_S^v = \left(\sum_{i=1}^N \hat{w}_1^*(T_i) \cdot X_i \right) \cdot (\hat{\gamma}_1^v - \hat{\gamma}_C^v)$$

$$\hat{\Delta}_X^v = \left(\sum_{i=1}^N (\hat{w}_1^*(T_i) - \hat{w}_0^*(T_i)) \cdot X_i \right) \cdot \hat{\gamma}_C^v + R^\mu$$

Onde R^μ é um erro de aproximação. E sob as hipóteses de linearidade e média condicional igual a zero, $\gamma_\mu^0 = \gamma_\mu^c$ e $R^\mu = 0$, de modo que esta decomposição se torna idêntica à decomposição Oaxaca-Blinder.

Como debatido em Barsky et al (2002), o método Oaxaca Blinder proporciona estimativas consistentes dos efeitos composição e retorno apenas quando a suposição de expectativa condicional é linear. Barsky et al (2002) e Dinardo, Fortin e Lemieux (1996) propõem realizar uma abordagem de reponderação para resolver o problema. No entanto este procedimento inviabiliza a realização da decomposição detalhada. O método proposto por Firpo, Fortin e Lemieux (2007, 2009) resolve este problema através da estimação de γ_c^v ao invés de γ_0^v , como na estimação tradicional Oaxaca-Blinder.

3.2. Análises empíricas utilizando RIF

Firpo, Fortin e Lemieux (2011), analisam a evolução da desigualdade de renda entre as décadas de 1980, 1990 e 2000 nos Estados Unidos. Os autores concluem que a mudança tecnológica foi responsável por grande parte das mudanças na distribuição dos salários nas últimas três décadas. Deste modo, argumentam que deve-se introduzir variáveis que capturem as tarefas desenvolvidas pelos indivíduos, bem como as ocupações que eles exercem, para compreender melhor a evolução das distribuições de salários.

Ndoye (2013), para estudar o papel da educação sobre a desigualdade entre as regiões rurais e urbanas em Senegal. A principal conclusão deste autor é que o efeito do retorno a educação é significativo para explicar as disparidades e que este efeito aumenta à medida que caminhamos para a parte superior da distribuição. Medina (2013) analisa a diferença de salários entre gêneros nas cidades da Nicarágua nos anos de 2005 e 2009. Os resultados indicam que o efeito estrutura dos salários é o maior responsável pelas diferenças salariais. Heywood e Parente (2012) utilizam o método para analisar o diferencial de salários entre brancos e negros nos EUA.

O trabalho empírico mais importante para embasar o artigo proposto foi escrito por Galego e Pereira (2013), os quais analisam a desigualdade de rendimentos entre as cinco macro regiões portuguesas. Os autores encontram que a educação, o tipo de ocupação e o tamanho da firma são os principais fatores explicativos das disparidades regionais no país. Além disto, os autores concluem que estas variáveis possuem diferentes impactos sobre o efeito composição e o efeito estrutura dos salários.

Este método também já possui algumas aplicações ao Brasil. Salardi (2012) investiga as diferenças salariais de raça e gênero no Brasil nas últimas duas décadas. Brito, Machado e Kerstenetzky (2013) analisam o papel do salário mínimo sobre a evolução da desigualdade de renda no Brasil entre 2001 e 2010.

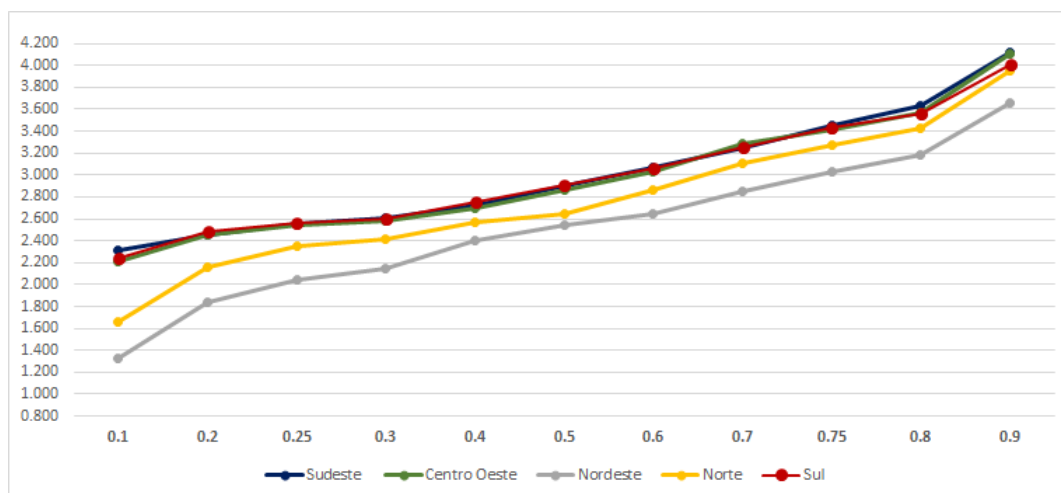
3.3. Base de dados e evidências iniciais

A base de dados utilizada neste estudo é o Censo Demográfico realizado para o ano de 2010 e foi escolhida por diversas razões. O censo abrange todo o território nacional, sendo realizado em quase todos os municípios do país, o que permitirá analisar a desigualdade dentre as diferentes regiões. O censo contém vasta gama de informações sobre as características socioeconômicas dos indivíduos, o que possibilita avaliar as características da oferta e da demanda por trabalho, vasta gama de informações sobre as atividades produtivas em que os indivíduos estão empregados e quais cargos eles exercem.

Nossa amostra consta de 3.089.083 milhões de pessoas entre Homens e Mulheres em todas as regiões brasileiras, maiores de 10 anos e menores de 65 anos, que declararam trabalhar entre 20 e 80 horas semanais, com ou sem carteira assinada.

A partir da utilização do Censo demográfico, estimou-se um modelo na qual a variável dependente representa o Total dos Rendimentos obtidos em todos os trabalhos por Hora. Optou-se por esta especificação da variável dependente com o objetivo de não comparar indivíduos com jornadas diárias de trabalhos diferentes, o que por si só já acarretaria diferenças salariais. Na figura 2, podem ser observados os salários-hora estimados dos indivíduos em cada região Brasileira por Quantil. Pode-se observar que, para o ano de 2010, os salários das regiões Sudeste, Sul e Centro Oeste são bastante parecidos em toda a distribuição dos rendimentos, enquanto a uma diferença significativa destes em relação às regiões Nordeste e Norte.

Figura 2: Salários-Hora Individuais de cada região brasileira por quantil



Fonte: Elaboração Própria

As variáveis explicativas de interesse deste estudo são duas dummies para medir a escolaridade. A variável “médio” refere-se aos indivíduos com ensino médio completo e a variável “superior” refere-se apenas aos indivíduos com ensino superior completo. As demais variáveis utilizadas como controle são: Idade, dummies que indicam se o indivíduo é homem, casado, possui cor branca, se retorna diariamente do trabalho para casa, se passa mais de uma hora no trânsito entre a residência e o trabalho, se possui carteira assinada, 6 dummies que indicam o número de filhos. Também são utilizadas 8 variáveis dummy de setores em que o indivíduo trabalha e 9 variáveis dummy que indicam o cargo ocupado pelo indivíduo no trabalho principal.¹²

4. Resultados

A partir da estimação da “Recentered Influence Function” foi possível decompor o diferencial salarial de cada região brasileira em relação a região Sudeste para o ano de 2010. A equação estimada assume a forma:

$$RIF(Y_i; Q_\tau)|X_i = \beta_0 + \beta_1 medio + \beta_2 superior + \beta_j X_j$$

, onde τ é o quantil estimado de interesse e β_1 e β_2 são os parâmetros estimados de interesse.

Os resultados da decomposição agregada e da decomposição detalhada para as variáveis médio e superior são apresentados nas Tabelas 1 a 4¹³. Todos os resultados foram estatisticamente significantes a 1%.

Vejamos o caso da disparidade salarial entre Nordeste e Sudeste. A diferença, em logarítimos do salário hora, no quantil inferior (.10) da distribuição é de 0.982, o que significa um diferencial de salários-hora de, aproximadamente, 166% a favor da região Sudeste. Já no quantil superior (.90), o diferencial de rendimentos a favor da região Sudeste é de, aproximadamente, 57%.

¹²Infelizmente, por limitação do número máximo de páginas, não foi possível inserir as Tabelas com os resultados dos parâmetros estimados de todas as variáveis explicativas.

¹³Optamos por apresentar apenas os resultados da decomposição detalhada para as variáveis de educação em função da grande quantidade de variáveis utilizadas no estudo, o que ultrapassaria o limite máximo de páginas permitido

Tabela 1: Logarítimos dos salários por hora estimados por quantil e decomposição da diferença salarial entre Sudeste e Nordeste

	0.1	0.2	0.25	0.3	0.4	0.5	0.6	0.7	0.75	0.8	0.9
Sudeste	2.307	2.453	2.555	2.611	2.720	2.898	3.065	3.242	3.449	3.626	4.116
Nordeste	1.325	1.843	2.039	2.151	2.398	2.549	2.639	2.850	3.023	3.182	3.663
Diferença	0.982	0.610	0.516	0.460	0.322	0.348	0.426	0.392	0.426	0.444	0.453
Efeito Composição	0.361	0.305	0.254	0.266	0.214	0.176	0.186	0.218	0.255	0.239	0.294
Médio	0.008	0.008	0.007	0.008	0.009	0.009	0.009	0.012	0.016	0.145	0.158
Superior	0.013	0.013	0.012	0.013	0.023	0.027	0.033	0.048	0.070	0.079	0.127
Efeito Estrutura	0.621	0.305	0.263	0.194	0.108	0.173	0.239	0.174	0.170	0.205	0.160
Médio	-0.040	-0.057	-0.050	-0.068	0.009	0.034	0.040	-0.028	-0.004	0.003	-0.027
Superior	-0.009	-0.016	-0.015	-0.025	0.015	0.029	0.042	-0.018	0.003	0.013	-0.051

Fonte: Elaboração Própria

Tabela 2: Logarítimos dos salários por hora estimados por quantil e decomposição da diferença salarial entre Sudeste e Norte

	0.1	0.2	0.25	0.3	0.4	0.5	0.6	0.7	0.75	0.8	0.9
Norte	2.306	2.453	2.555	2.610	2.719	2.897	3.064	3.241	3.448	3.626	4.116
Sudeste	1.654	2.163	2.347	2.411	2.565	2.644	2.868	3.110	3.275	3.424	3.958
Diferença	0.652	0.290	0.207	0.199	0.154	0.232	0.196	0.131	0.173	0.201	0.157
Efeito Composição	0.25	0.18	0.13	0.13	0.16	0.17	0.16	0.15	0.18	0.19	0.20
Médio	0.005	0.005	0.004	0.004	0.007	0.008	0.008	0.009	0.011	0.011	0.011
Superior	0.008	0.009	0.008	0.008	0.019	0.024	0.031	0.038	0.057	0.067	0.098
Efeito Estrutura	0.397	0.111	0.078	0.071	-0.009	0.062	0.033	-0.015	-0.009	0.014	-0.043
Médio	-0.039	-0.053	-0.024	-0.023	0.26	-0.006	-0.008	-0.033	0.27	-0.019	0.012
Superior	-0.007	-0.013	-0.004	-0.005	0.021	0.002	0.000	-0.028	0.034	-0.024	0.02

Fonte: Elaboração Própria

Tabela 3: Logarítimos dos salários por hora estimados por quantil e decomposição da diferença salarial entre Sudeste e Centro Oeste

	0.1	0.2	0.25	0.3	0.4	0.5	0.6	0.7	0.75	0.8	0.9
CO	2.307	2.453	2.555	2.611	2.720	2.898	3.065	3.284	3.449	3.626	4.116
Sudeste	2.210	2.453	2.539	2.585	2.694	2.864	3.024	3.242	3.410	3.563	4.110
Diferença	0.097	0.000	0.016	0.026	0.026	0.034	0.041	0.042	0.039	0.063	0.006
Efeito Composição	0.086	-0.055	0.044	0.044	0.072	0.067	0.066	-0.058	0.082	0.078	0.086
Médio	0.006	-0.005	0.004	0.004	0.009	0.009	0.010	-0.010	0.014	0.013	0.014
Superior	0.004	0.004	0.003	0.003	0.008	0.010	0.013	-0.015	0.024	0.277	0.041
Efeito Estrutura	0.011	0.055	-0.027	-0.019	-0.046	-0.034	-0.025	0.100	-0.043	-0.015	-0.079
Médio	-0.042	0.003	0.000	-0.001	-0.003	-0.001	0.001	0.004	-0.009	-0.002	-0.026
Superior	-0.022	0.002	0.000	-0.002	-0.005	-0.003	-0.001	0.000	-0.016	-0.021	-0.031

Fonte: Elaboração Própria

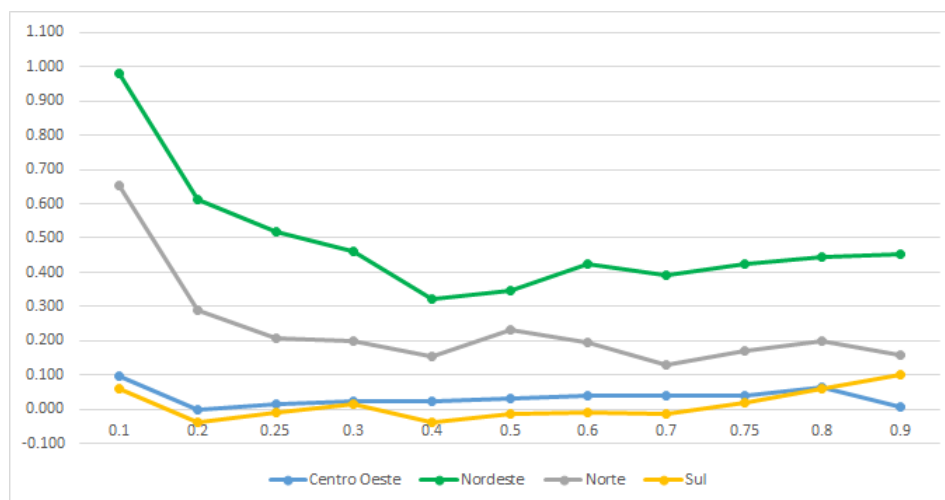
Tabela 4: Logarítimos dos salários por hora estimados por quantil e decomposição da diferença salarial entre Sudeste e Sul

Sul	0.1	0.2	0.25	0.3	0.4	0.5	0.6	0.7	0.75	0.8	0.9
Sudeste	2.307	2.453	2.555	2.611	2.720	2.898	3.075	3.242	3.449	3.626	4.116
Sul	2.244	2.490	2.566	2.597	2.758	2.912	3.065	3.255	3.430	3.565	4.014
Diferença	0.062	-0.037	-0.010	0.014	-0.038	-0.014	-0.010	-0.013	0.019	0.061	0.102
Efeito Composição	0.059	-0.020	-0.018	0.021	-0.018	-0.019	-0.018	-0.016	0.024	0.026	0.036
Médio	0.005	-0.004	-0.003	0.003	-0.006	-0.007	-0.007	-0.007	0.010	0.010	0.010
Superior	0.004	-0.004	-0.004	0.004	-0.009	-0.011	-0.015	-0.017	0.027	0.030	0.036
Efeito Estrutura	0.004	0.057	0.029	-0.008	0.056	0.033	0.028	0.029	-0.006	0.036	0.066
Médio	-0.027	-0.004	0.001	-0.011	-0.005	-0.001	0.003	-0.004	0.007	0.009	-0.003
Superior	-0.013	-0.003	0.000	-0.010	-0.005	-0.004	0.001	-0.009	0.012	0.021	0.022

Fonte: Elaboração Própria

Nas Figuras 3 a 6 apresentamos de forma gráfica algumas informações importantes extraídas das tabelas 1 a 4. A primeira evidência são as disparidades de salários das regiões Nordeste e Norte em relação à região Sudeste para todos os quantis da distribuição (Figura 3), destacando que a maior diferença de rendimentos é verificada nos quantis inferiores da distribuição. Esta evidência corrobora com os resultados obtidos em Campelo e Silveira Neto (2005) e Guimarães, Cavalcanti e Silveira Neto (2006). Além disto, o trabalho contribui com a literatura citada, pois esta apenas analisava os diferenciais de renda entre as regiões Sudeste e Nordeste. Já as regiões Centro Oeste e Sul não apresentam diferenciais de salários significativos em relação à região Sudeste, sendo que, em boa parte dos pontos da distribuição, estas regiões apresentam salários mais elevados do que a região Sudeste.

Figura 3: Diferença de rendimentos entre Sudeste e as demais regiões metropolitanas brasileiras por quantil.

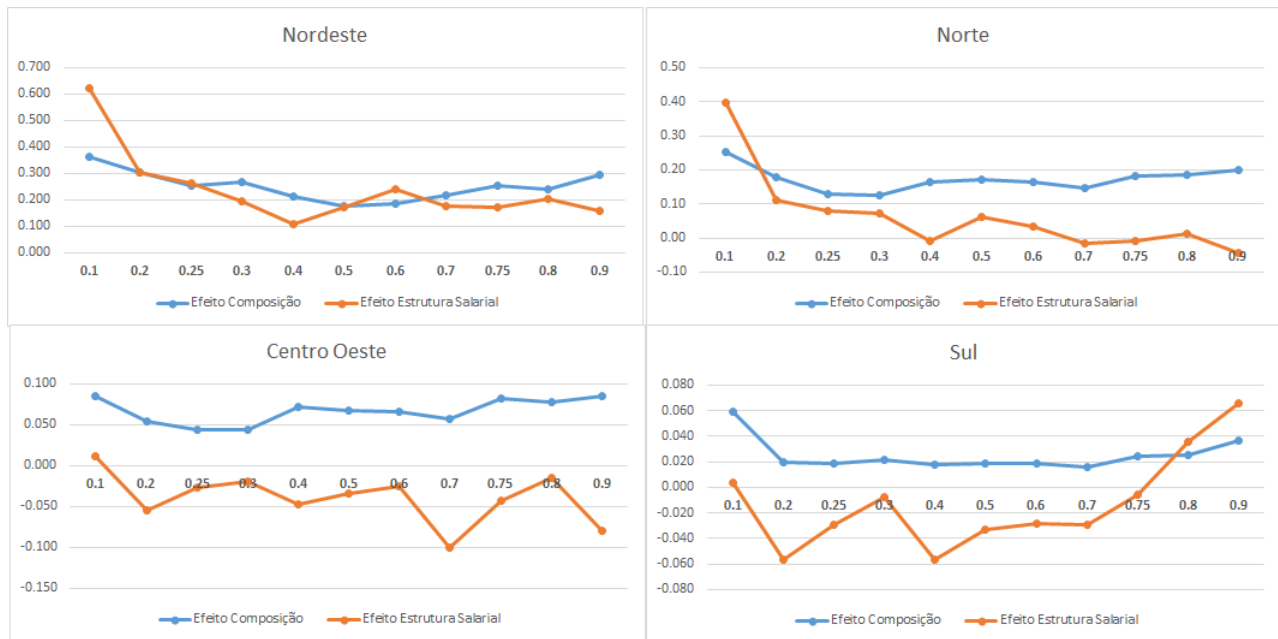


Fonte: Elaboração Própria

Já na figura 4 podem ser observados os resultados da decomposição agregada para cada região. O resultado que mais se destaca é o da região Nordeste, a qual apresenta os maiores valores, tanto para o efeito composição, quanto para o efeito estrutura salarial. Seguida da região Norte, Centro Oeste e Sul, respectivamente.

Além disto, de um modo geral, o efeito composição é mais importante para a explicação dos diferenciais de rendimento das regiões em relação à região Sudeste. Isto é, de um modo geral, o diferencial de rendimentos das regiões em relação a Sudeste é explicada principalmente pelas diferenças de atributos produtivos dos indivíduos. A única exceção é a região Nordeste, onde tanto o diferencial de atributos produtivos quanto o retorno a estes atributos é de fundamental importância para a explicação das diferenças salariais.

Figura 4: Decomposição agregada da desigualdade entre os efeitos composição e estrutura salarial para cada região



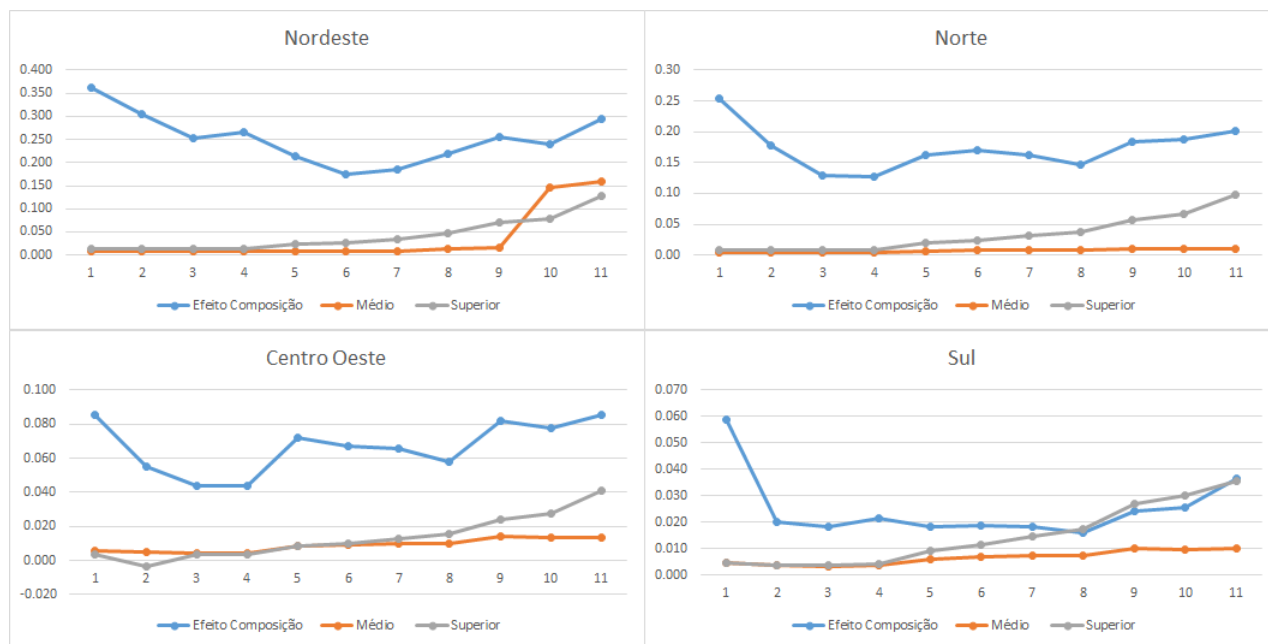
Fonte: Elaboração Própria

Na figura 5 apresentamos os resultados da decomposição detalhada do Efeito Composição. Pode-se observar que nos quantis inferiores da distribuição os valores do ensino médio e superior sobre o Efeito Composição são bastante similares e muito próximos de zero. Isto significa que, para os indivíduos mais pobres, possuir ensino médio ou superior não explica os diferenciais de rendimentos em relação a região Sudeste.

No entanto, acima do quantil .50, os valores começam a aumentar. Isto significa que, as diferenças do nível de educação são mais importantes para a explicação dos diferenciais regionais de salários nos quantis superiores da distribuição. Por fim, possuir ensino médio só explica os diferenciais de rendimentos entre Nordeste e Sudeste, enquanto que possuir ensino superior explica os diferenciais de rendimentos de todas as regiões em relação a região Sudeste.

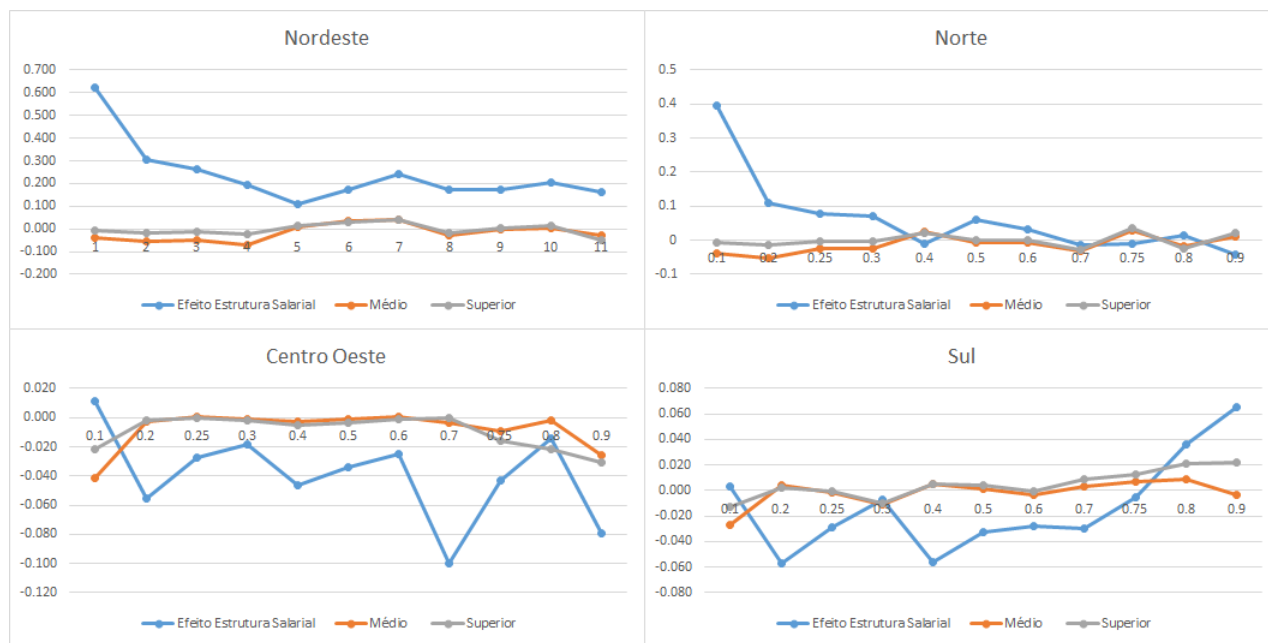
Na Figura 6 podem ser observados a contribuição das variáveis de educação sobre o Efeito Retorno Salarial para todos os quantis da distribuição. Neste caso, a análise permite observar que a contribuição da educação para o Efeito Estrutura Salarial é muito próxima de zero para todos os quantis da distribuição. Ou seja, o retorno a educação, não parece ser uma variável relevante para explicação dos diferenciais regionais de salários no Brasil. Por exemplo, os diferenciais de renda no Nordeste são explicados pois as pessoas possuem níveis menores de escolaridade, mas quando comparamos pessoas com níveis de escolaridades similares entre Nordeste e Sudeste, o retorno a escolaridade é similar.

Figura 5: Importância da Escolaridade para o efeito composição para cada região (Sudeste é a referência)



Fonte: Elaboração Própria

Figura 6: Importância da Escolaridade para o efeito estrutura salarial para cada região (Sudeste é a referência)



Fonte: Elaboração Própria

5. Considerações Finais

Este estudo teve como objetivo entender a desigualdade regional de renda no Brasil e analisar o papel da educação. Utilizou-se como estratégia metodológica a decomposição utilizando Recentered Influence Functions (RIF). Primeiramente, a partir da estimação da RIF para cada quantil da distribuição dos salários, observou-se que a diferença de salários de cada região brasileira em relação a região Sudeste varia bastante. O Nordeste apresenta a maior disparidade de rendimentos em todos os quantis, seguido das regiões Norte, Centro Oeste e Sul. No entanto, para alguns pontos da distribuição, as regiões Centro Oeste e Sul apresentam salários maiores do que a região Sul.

No caso da região Nordeste e Norte, a desigualdade em relação ao Sudeste se mostra ainda maior nos quantis inferiores. Isto é, a desigualdade é maior dentre os mais pobres. Já para as regiões Centro Oeste e Sul, o padrão da desigualdade é praticamente o mesmo. Quando realizamos a Decomposição Agregada, uma variedade de situações foi apresentada. De um modo geral, o Efeito Composição é mais importante para a explicação da desigualdade em relação a região Sudeste do que o Efeito Estrutura Salarial. A única exceção ficou por conta da região Nordeste, onde os efeitos Composição são bem maiores para quantis mais altos da distribuição. Mas mesmo aqui, há importantes qualificações a fazer. Em relação às disparidades de renda entre o Sudeste e as regiões mais pobres (NO e NE), as disparidades de escolaridade só são importantes para os quantis mais altos da distribuição. Além disto, os efeitos Estrutura Salarial, em geral, contribuem para atenuar tais disparidades regionais.

Já os resultados da decomposição detalhada permitiram observar que, enquanto os diferenciais de educação entre as regiões são um importante fator contribuidor para a desigualdade regional de renda no Brasil, os retornos a educação não possuem forte efeito sobre a desigualdade. Isto significa, por exemplo, que as pessoas no Nordeste possuem um nível menor de escolaridade, mas quando comparamos pessoas com níveis de escolaridades similares, o retorno a escolaridade é similar.

É importante ressaltar que os resultados aqui encontrados, em larga medida, diferem daqueles encontrados por Duarte et al. (2003) e Guimarães et al. (2006). Especificamente, diferentemente do encontrado no primeiro trabalho, as disparidades de dotação de escolaridade entre o NE e o SE, por exemplo, nas evidências aqui encontradas são de forma substantiva relevante apenas para os quantis mais elevados da distribuição (por exemplo, para o quantil 0,9 explicam $100 \times (0,285/0,453)$ das disparidades de renda entre as duas regiões; já na mediana explicam apenas $100 \times (0,036/0,348)$ destas disparidades). Ao contrário de Guimarães et al. (2006), as evidências obtidas também apontam para um papel bem menos relevante para os retornos à escolaridade.

Assim, como as disparidades regionais entre as regiões mais pobres e mais rica do país são mais substantivos nos quantis mais baixos da distribuição, onde os disparidades de escolaridade tem papel menos relevante, as evidências encontradas neste trabalho não suportam o argumento de que as disparidades de escolaridade por si é a única e principal responsável pelas disparidades regionais brasileiras. Na verdade, os resultados indicam que outras características dos indivíduos e, sobretudo, dos mercados de trabalho das regiões mais pobres tem também papel relevante no entendimento destas disparidades.

6. Bibliografia

Azzoni, C. R. Concentração regional e dispersão das rendas per capita estaduais: análise a partir de séries históricas estaduais de Pib, 1939-1995. *Estudos Econômicos*. São Paulo, v.27, n. 3, p-341-393. Set-Dez, 1997.

Barro, R. Economic growth in a cross section of countries. *Quartely jornal of Economics*, 106, p. 407-43, May 1991.

Barro, R.; Sala-i-Martin, X. Convergence. *journal of Political Economy*, v. 100, n. 2, 1992.

Baumol, W. J. Productivity growth, convergence and welfare. *american Economic review*, 76, p. 1072-1085, December 1986.

Becker, G. S. Human capital. Columbia Press University. 1964.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux (1996), "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica* 64: 1001-1044.

Duarte, A.; Ferreira, P. C.; Salvato, M. Disparidades regionais ou educacionais? Um exercício com contrafactuais. Mimeografado. 2003.

Ferreira, A. Convergence in Brazil: recent trends and long-run prospects. *applied Economics*, 32, p. 479-489, 2000.

Ferreira, P. Ellery, Jr.; Ferreira, P. Convergência entre a renda per capitados estados brasileiros. *Revista de Econometria*, v. 16, n. 1, p. 83-103, 1996.

Ferreira, A.; Diniz, C. Convergência entre as rendas per capita estaduais no Brasil. *Revista de Economia Política*, v. 15, n. 4 (60), 1995

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux (2007), "Decomposing Wage Distributions using Recentered Influence Functions Regressions", mimeo, University of British Columbia.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux (2009) "Unconditional Quantile Regressions," *Econometrica* 77(3): 953-973.

FORTIN, Nicole ; LEMIEUX, Thomas ; FIRPO, S. . Decomposition Methods in Economics. In: David Card; Orley Ashenfelter. (Org.). *Handbook of Labor Economics*. 1ed.: Elsevier, 2011, v. 4, p. 1-102.

Gondim, J. L. B.; Barreto, F. A. F. D.; Carvalho, J. R. Condicionantes de Clubes de Convergência no Brasil. *Estudos Econômicos*. Instituto de Pesquisas Econômicas, v. 37, p. 71-100, 2007.

Kakwani, Nanaki. *Income Inequality and Poverty: Methods of Estimation and Policy Applications*. New York: Oxford University Press. 1980.

Koenker, Roger and G. Bassett (1978), "Regression Quantiles," *Econometrica*, 46, 33-50.

Leman, R.; Yitzhaki, S. Income Inequality Effects by Income Source: A New Approach and Applications to the United States. *Review of Economics and Statistics*, 67, 151–156. 1985.

Machado, José F. and José Mata (2005), "Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression", *Journal of Applied Econometrics* 20: 445-465.

Menezes, T. A.; Azzoni, C. R. . Convergência de Salários Entre as Regiões Metropolitanas Brasileiras: Custo de Vida e Aspectos da Demanda e Oferta de Trabalho. *Pesquisa e Planejamento Econômico*. Rio de Janeiro, v. 36, p. 449-470, 2006.

Mossi, M. B.; Aroca, P.; Fernández, I. J.; Azzoni, C. R. Growth dynamics and space in Brazil. *international regional science review*, v. 26, n. 3, p. 393-418, July 2003.

Oaxaca, Ronald (1973), "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review* 14: 693-709.

Oaxaca, Ronald L. and Michael R. Ransom(1994), "On discrimination and the decomposition of wage differentials". *Journal of Econometrics* 61: 5-21.

Oaxaca, Ronald L. and Michael R. Ransom (1999), "Identification in Detailed Wage Decompositions," *Review of Economics and Statistics* 81: 154-157.

Pessoa, S. Existe um problema de desigualdade regional no Brasil? In *Anais do Encontro Nacional de Economia*, Salvador. ANPEC. 2000.

Quah, D. Empirics for growth and distribution: stratification, polarization and convergence clubs. *Journal of Economic Growth*, v. 2, n. 1, p. 27-59, March 1997.

Silveira Neto, R. M. ; Menezes, T. A. Disparidades Regionais de Renda no Brasil: Analisando a Importância do Capital Humano. In: *Fundação Konrad-Adenauer*. (Org.). Fórum Brasil-Europa. Fundação Konrad-Adenauer. 2008.

Silveira Neto, R. M. ; Azzoni, C. R. Disparidades regionais de renda no Brasil: qual a importância das amenidades regionais. *Encontro de economia regional - Anpec Nordeste*. Fortaleza. 2004.

Silveira Neto, R. M. ; Azzoni, C. R. Social policy as regional policy: market and non-market factors determining regional inequality. *Journal of Regional Science*, v. 51, p. 1-18, 2011.

Zini Jr., A. A. Regional income convergence in Brazil and its socio-economic determinants. *Economia aplicada*, v. 2, n. 2, p. 383- 411, abr./jun. 1998.