

## ESTRUTURA PRODUTIVA OU ESCOLARIDADE? UMA ANÁLISE DOS FATORES EXPLICATIVOS DA DESIGUALDADE DE RENDA ENTRE AS REGIÕES SUDESTE E NORDESTE POR QUANTIL PARA O PERÍODO ENTRE OS ANOS DE 1970 E 2010

### Productive structure or schooling? An analysis of the explanatory factors of income inequality between the southeast regions and northeast regions per quantil for the period between 1970 and 2010

**Rodrigo C. Oliveira**

Economista. Mestrado em Economia pela Universidade Federal de Pernambuco - UFPE. Professor Assistente da Faculdade de Ciências Econômicas, Universidade Federal da Bahia - FCE/UFBA. rodrigo.coliveira13@gmail.com

**Raul da Mota Silveira Neto**

Economista. Doutor em Economia. Prof. do Departamento de Ciências Econômicas da Universidade Federal de Pernambuco - UFPE.

**Resumo:** Desde a primeira década do século XX as desigualdades regionais brasileiras foram preocupação das autoridades e dos estudiosos. Porém, ainda não foi possível identificar com precisão as causas explicativas destas disparidades. Ao longo das décadas de 1960 e 1980 o argumento desenvolvimentista de estímulo a industrialização e melhoria da estrutura produtiva no Nordeste predominou sobre a formulação de políticas públicas regionais, porém, os críticos deste sistema argumentam que deveria ter sido dada maior ênfase às políticas de valorização do capital humano individual, sobretudo da educação. Neste sentido, o presente artigo propõe-se a analisar, ao longo do tempo, quais são os fatores predominantes para a explicação da desigualdade regional no Brasil: Capital Humano ou Estrutura Produtiva? 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, tanto as características produtivas, quanto os retornos a estas características são fundamentais para compreender a evolução da desigualdade regional no Brasil. Ademais, a estrutura produtiva desempenha papel mais importante até a década de 1991 e para os quantis inferiores, enquanto o capital humano desempenha papel mais importante a partir da década de 2000 e para os quantis superiores.

**Palavras Chave:** Desigualdade Regional; Educação; Decomposição.

**Abstract:** Since the first decade of the twentieth century Brazilian regional inequalities were concern of the authorities and researchers. However, it was not possible to accurately identify the causes of these disparities. Between the 1960s and 1990s the predominant argument was that the industrialization and improvement of the productive structure in the Northeast would help to reduce the regional inequalities, however, critics of this view argue that it should have been given greater emphasis to policies that value the education. In this sense, this article aims to analyze, over time, which are the predominant factors in explaining regional inequality in Brazil: Human Capital or Productive Structure? It was used as a methodological strategy decomposition using *Recentered Influence Functions (RIF)* proposed by Firpo, Fortin and Lemieux (2007, 2009). The main results indicate that both the productive characteristics, and the returns to these characteristics are fundamental to understand the evolution of regional inequality in Brazil. Moreover, the productive structure plays important role until the decade of 1991 and for the lower quantiles, while the human capital plays important role from the 2000s and the upper quantiles.

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

Ferreira e Ellery Jr. (1996) mostraram que existiu 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 similar, 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 brasileiros. 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.

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.

Por fim, 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 regressão<sup>1</sup>.

Note-se que, estes trabalhos possuem uma série de limitações. Neste sentido, a presente pesquisa busca superar estas limitações e trazer elementos novos para a análise da desigualdade de renda regional brasileira por duas razões. 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 literatura<sup>2</sup>. A mais importante delas é o fato de permitir atribuir um Efeito Composição e um Efeito Estrutura Salarial para cada variável determinante do nível

1 Estes problemas são bem explicados em Firpo, Fortin e Lemieux (2010).

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

de renda das economias ao longo de diferentes pontos da distribuição de renda, nos moldes da tradicional decomposição de Oaxaca-Blinder, o que não é possível com os métodos até aqui aplicados para estudo das disparidades regionais no Brasil. A segunda contribuição refere-se a análise da desigualdade regional entre as regiões Nordeste e Sudeste para todos os Censos Demográficos desde a década de 1970. Isto é, para os anos nos quais foram realizados: 1970, 1980, 1991, 2000 e 2010. Os dados foram compatibilizados para permitir uma análise ao longo do tempo e, além disto, a escolha do Censo justifica-se pela disponibilidade de 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; AZZONI, 2006). Os enfoques gerais ao tratamento das desigualdades regionais na literatura, de forma geral, se diferenciam por enfatizar mais ou menos estes dois conjuntos 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).

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 ideia de Mincer pode ser simplificada utilizando uma equação que busca apreender como são determinados os salários dos indivíduos:

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

Onde  $\beta_0$  seria o salário de um indivíduo desqualificado,  $\beta_1$  o retorno a educação,  $\beta_2$  e os retornos associados a experiência e  $\beta_3$  o retorno associado às demais características dos indivíduos.

Atualmente, grande parte da literatura em economia da educação e economia do trabalho busca encontrar um valor causal para  $\beta_1$ . 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 gêmeos<sup>3</sup>.

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

<sup>3</sup> O survey de Card (1999) possui informações detalhadas sobre a evolução desta literatura.

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 de convergência entre grupos de países<sup>4</sup> 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 mo-

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

delos 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 ideia de que o crescimento econômico não aparece simultaneamente em toda a parte, mas sim em pontos ou polos 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 ideias 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 microeconômica, 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 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 fun-

cionando 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 está 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 infraestrutura e do mix de indústrias entre as regiões.

Sendo assim, diferentemente da literatura em crescimento econômico, esta última 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 Metodologia

Este trabalho utilizará o método de regressão proposto por Firpo, Fortin e Lemieux (2007), que se vale da *Recentered Influence Function* para decompor as diferenças nas distribuições de rendimentos entre dois grupos distintos. Este método se insere na literatura de decomposição em economia, especialmente na área de economia do traba-

lho, 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; LEMIEUX, 1996; MACHADO; MATA, 2005; FIRPO; FORTIN; 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 Estrutura 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 diferencial de rendimentos pelo fato dos indivíduos de cada região possuírem características diferentes, o segundo reflete o diferencial de rendimentos decorrente das diferenças de retorno às características de indivíduos similares, mas em grupos diferentes. No caso deste estudo, deseja-se avaliar a diferença de rendimentos entre as regiões Sudeste e a região Nordeste.

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 Estrutura Salarial irá captar o efeito dos retornos às variáveis sobre a diferenças nos rendimentos médios. Por exemplo, se compararmos uma pessoa com nível superior no Nordeste com uma pessoa com nível superior no Sudeste, o retorno ao nível superior será o mesmo em cada região? Ou na região Sudeste o retorno ao nível superior é maior que o retorno na região Nordeste?

### 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 decomposição<sup>5</sup>. 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* (FIRPO; FORTIN; 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 Estrutura Salarial 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.

O primeiro estágio pode ser obtido de modo bastante similar à reponderação proposta em Dinardo, Fortin e Lemieux (1996). Este primeiro estágio é necessário pois, como discutido em Barsky et al. (2002), quando a média condicional não é uma função linear, a decomposição Oaxaca-Blinder pode não proporcionar estimativas consistentes dos Efeitos Estrutura Salarial e do Efeito Composição. E, como debatido no mesmo trabalho, em geral o rendimento médio condicional é não linear. Deste modo, serão criados pesos, para reponderar a distribuição dos rendimentos da região de interesse de modo a obter uma distribuição contrafactual de interesse  $Fy_a^c$ . Os pesos de reponderação são computados como:

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

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

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

<sup>5</sup> Oaxaca (1973), Blinder (1973), Machado e Mata (2005) e Dinardo, Fortin e Lemieux (1996).

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

Onde  $\widehat{p}$  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:

$$\widehat{w}_1^*(T_1) = \frac{\widehat{w}_1(T_1)}{\sum_{j=1}^N \widehat{w}_1(j)} = \frac{T_i}{N \cdot \widehat{p}}$$

$$\widehat{w}_0^*(T_i) = \frac{\widehat{w}_0(T_i)}{\sum_{j=1}^N \widehat{w}_0(j)} = \frac{1 - T_i}{N \cdot (1 - \widehat{p})}$$

$$\widehat{w}_c^*(T_i X_i) = \frac{\widehat{w}_c(T_i)}{\sum_{j=1}^N \widehat{w}_c(j)} = \frac{1 - T_i \cdot \frac{\widehat{p}(X_i)}{1 - \widehat{p}(X_i)}}{\sum_{j=1}^N \frac{\widehat{p}(X_j)}{1 - \widehat{p}(X_j)}}$$

Em termos práticos, o cálculo do peso de ponderação é bem simples. Basta estimar via um modelo Probit, utilizando as características observáveis dos indivíduos como regressores, qual a probabilidade de um indivíduo morar no Sudeste, dado que ele mora no Nordeste. Note que este procedimento é bastante similar a ideia de pareamento via Propensity Score Matching, o que mostra uma primeira evidência da relação das literaturas de decomposição e avaliação de programas. A partir dos pesos estimados, obtém-se a distribuição contrafactual dos rendimentos de interesse, como:

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

e

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

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 deseja-se estimar

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

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

$$\widehat{\mu}_t = N^{-1} \sum_{i=1}^N \widehat{w}_t(T_1) Y_i$$

e

$$\widehat{\mu}_c = N^{-1} \sum_{i=1}^N \widehat{w}_c(T_1, X_i) Y_i$$

Em seguida seria possível estimar as diferenças entre os grupos:

$$\widehat{\Delta}_o^v = \widehat{\mu}_1 - \widehat{\mu}_0 = \widehat{\Delta}_s^v + \widehat{\Delta}_x^v$$

com  $\widehat{\Delta}_s^v = \widehat{\mu}_1 - \widehat{\mu}_c$  representando o Efeito Estrutura Salarial e  $\widehat{\Delta}_x^v = \widehat{\mu}_c - \widehat{\mu}_0$  representando o Efeito Composição.

Importante ressaltar que, tal como explicado em Firpo, Fortin e Lemieux (2010), o efeito Estrutura Salarial possui uma interpretação muito importante. Ele representa o diferencial de rendimentos que é explicado pelo retorno às características dos indivíduos depois de controlado por todas as características dos mesmos. Isto significa que o diferencial de rendimentos é explicado pelo fato de o indivíduo situar-se no Nordeste e não no Sudeste. Em outras palavras, o componente estrutura salarial representa a valorização diferente das características produtivas nas diferentes regiões. Este componente é interpretado, na literatura de economia do trabalho, como um efeito da discriminação. No caso do presente estudo significa que pessoas com as mesmas características são remuneradas de modo diferente apenas pelo fato de estarem em regiões diferentes.

Firpo, Fortin e Lemieux (2010) pontuam que, o Efeito Estrutura Salarial agregado possui uma interpretação causal, muito similar a literatura de avaliação de programas, pois:

This selection based on observables assumption allows for selection biases as long they are the same for the two groups. For example, if unobservable ability and education are correlated, a linear regression of  $Y$  on  $X$  will not yield consistent estimates of structural parameters (i.e. the return to education). But the aggregate decomposition remain valid as long as the dependence structure between ability and education is the same in group A and B (FIRPO; FORTIN; LEMIEUX, p. 9, 2010).

No segundo estágio, calcula-se a contribuição de cada variável sobre o Efeito Composição ( $\widehat{\Delta}_x^v$ ) e sobre o Efeito Estrutura Salarial ( $\widehat{\Delta}_s^v$ ) tal como na tradicional decomposição Oaxaca-Blinder. Nes-

te 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

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

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:

$$IF(y; q_T) = \frac{T - 1\{y \leq q_T\}}{fY(q_T)}$$

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

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

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

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

Para caso dos quantis, o qual é o maior interesse deste estudo, primeiramente estima-se o quantil amostral  $\widehat{Q}_T$  tal como em Koenker e Basset (1978) e a densidade  $f(\widehat{Q}_T)$  neste ponto utilizando funções

Kernel9, posteriormente substitui-se estas estimativas na equação (2). Por fim, é estimado o efeito da mudança na distribuição de uma variável sobre o sobre o quantil marginal t de Y, numa especificação que toma a forma:

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

E o parâmetro de interesse pode ser calculado via OLS:

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

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

$$\hat{\gamma}_t^v = \left( \sum_{i=1}^N \hat{w}_1^*(T_i).X_i.X_i^T \right)^{-1} \cdot \sum_{i=1}^N \hat{w}_t^*(T_i).R\hat{I}F(Y_i; v_t).X_i$$

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

$$\text{Com } R\hat{I}F(y; v_t) = \widehat{v}_t + \widehat{I}F(y; v_t) \text{ e}$$

$$R\hat{I}F(y; v_c) = \widehat{v}_c + \widehat{I}F(y; v_c).$$

Por fim, é possível decompor os Efeito Composição ( $\widehat{\Delta}_X^\varphi$ ) e Efeito Estrutura Salarial ( $\widehat{\Delta}_S^\varphi$ ) como:

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

$$\widehat{\Delta}_X^\varphi = \sum_{i=1}^N \hat{w}_1^*(T_i) - \widehat{w}_0^*(T_i). X_i. \widehat{\gamma}_c^v + R^\mu$$

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

Um problema comum destes métodos de decomposição é a invariância do grupo base (Oaxaca e Ransom, 1999). Isto é, ao utilizar variáveis explicativas categóricas o resultado da estimação varia a depender do grupo base escolhido. Neste caso, de modo a superar este problema, implementamos a correção proposta por Yunn (2005). Esta correção consiste na estimação da regressão diversas

vezes alterando o grupo base escolhido. Posteriormente toma-se a média dos coeficientes estimados como o parâmetro correto.

### 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 são os Censos Demográficos realizados a partir do ano de 1970 até o ano 2010. A utilização do Censo Demográfico possui diversas vantagens. Primeiramente permitirá analisar a desigualdade regional brasileira numa perspectiva temporal mais ampla, a partir de uma base de dados compatibilizada. Não há precedente de análises similares para o Brasil em período tão longo de tempo. Segundo, 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. Por fim, 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. A variável de rendimento foi deflacionada utilizando 2010 como ano base. Neste caso, compatibilizou-se os cinco censos disponíveis neste período de modo que foi possível obter as seguintes variáveis para análise (TABELA 1).

Tabela 1 – Variáveis Seleccionadas para Estudo

Variáveis	
Dependente	
LnRenda	Logaritmo do Rendimento de Todos os Trabalhos
Explicativas	
Médio	Dummy = 1 caso o indivíduo possui ensino médio
Superior	Dummy = 1 caso o indivíduo possui ensino superior
Idade	Idade em anos
Tfamília	Número de pessoas na família
Gênero	Dummy igual a 1 caso seja homem
Agropecuária	Dummy = 1 para o setor no qual o indivíduo trabalha
Ind. Transformação	
Ind. Extrativa	
Serviços	
Governo	
Comércio	
Construção Civil	
Outros	

Fonte: Elaboração Própria

Antes de iniciar a análise de algumas evidências que podem ser observadas a partir dos dados cabe ressaltar que no período entre 1970 e 2010 o Brasil vivenciou diferentes ambientes macroeconômicos e condições sociais, neste sentido, analisaremos estas estatísticas à luz destes acontecimentos. Em 1970 consolidava-se o governo militar e o País estava no meio do chamado Milagre Econômico com forte investimento em infraestrutura e estímulo a industrialização. A partir de 1974 enfrentou diversos momentos conturbados da economia mundial e, de modo a contorná-los, lançou mão do II PND. Neste período o Brasil não registrou em nenhum ano crescimento inferior a 5% do PIB, chegando a uma taxa média de crescimento nesta década de, aproximadamente, 8%.

Em seguida veio um período de forte crise na década de 1980, iniciada com as crises do petróleo e consolidada com o grande problema de endividamento externo do País. Este período ficou conhecido como a década perdida. A taxa média de crescimento do PIB no período foi de 1%, sendo que em alguns anos o país vivenciou uma retração forte do PIB. Ao final da década de 1980 houve o processo de redemocratização do país, juntamente com uma agenda político-social que introduziu diversos direitos aos cidadãos, como o acesso à educação e saúde.

A década de 1990 também foi marcada pela estabilização econômica promovida a partir de 1994 com a introdução do Plano Real. O controle da inflação ocorreu concomitantemente a um período de relativo crescimento do salário mínimo, o que beneficiou os indivíduos mais pobres em todas as regiões do País. Apesar da estabilização, ocorreram algumas crises internacionais que prejudicaram o crescimento da economia brasileira no período, como a crise do México em 1994, a crise dos países asiáticos em 1997 e a crise da Rússia em 1998.

Por fim, a partir dos anos 2000 o cenário macroeconômico muda bastante. Primeiramente pois até 2007 verificou-se um período de crescimento da economia mundial. Segundo pois dada a estabilidade macroeconômica do período anterior, foi possível o Governo Federal introduzir diversas políticas de redução da pobreza, em especial o programa Bolsa Família, e manter a política de valorização do salário mínimo.

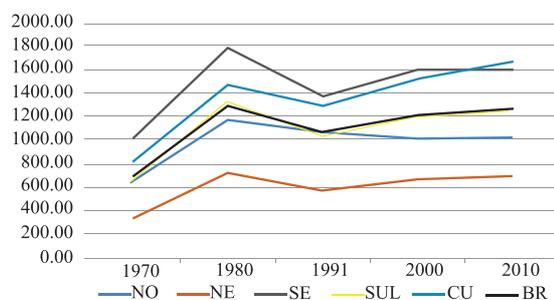
A partir desta contextualização, podemos analisar os resultados. Na figura 1 podem-se ser visualizadas a evolução das rendas médias de cada uma das cinco macrorregiões brasileiras. Duas evidências se destacam, a primeira é o fato de as rendas de cada região moverem-se da mesma forma ao longo do período, acompanhando os movimentos da economia brasileira como um todo. A segunda é o fato de região Sudeste ser a mais rica desde 1970, enquanto a região Nordeste é a mais pobre.

Em seguida apresentamos na figura 2 a evolução do índice de Gini estadual e do índice de municipal no Brasil. Primeiramente observa-se que o Gini municipal é bem superior ao estadual, o que reflete o fato de os municípios no Brasil serem muitos mais desiguais do que os estados brasileiros, um fato estilizado na literatura nacional que utiliza os modelos de crescimento para analisar convergência de renda. Em seguida, observa-se que no período da década de 1970 (milagre econômico e II PND) verifica-se uma grande redução do índice de Gini regional enquanto durante a década perdida este índice aumenta consideravelmente. Isso significa que durante a fase de crescimento da economia brasileira os estados e municípios mais pobres cresceram a taxas mais elevadas do que os estados e municípios mais ricos, enquanto no período de crise os estados e municípios mais pobres também sofreram mais que os estados e municípios mais ricos.

Esta evidência é muito importante pois mostra que as economias mais atrasadas estão mais dependentes dos movimentos da economia nacional, refletindo uma relativa falta de estrutura para contrapor os momentos macroeconômicos mais difíceis. Por seu turno, a partir de 1991 há uma tendência comum em termos de desigualdade regional de renda, onde os dois índices calculados refletem a estabilização econômica, os períodos de relativo crescimento, a valorização do salário mínimo e a introdução dos programas sociais.

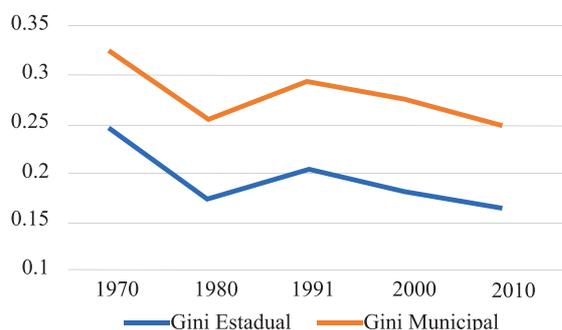
Esta evidência é muito importante pois mostra que as economias mais atrasadas estão mais dependentes dos movimentos da economia nacional, refletindo uma relativa falta de estrutura para contrapor os momentos macroeconômicos mais difíceis. Por seu turno, a partir de 1991 há uma tendência comum em termos de desigualdade regional de renda, onde os dois índices calculados refletem a estabilização econômica, os períodos de relativo crescimento, a valorização do salário mínimo e a introdução dos programas sociais.

Figura 1 – Evolução das Rendas Médias Per Capita Regional



Fonte: Elaboração própria

Figura 2 – Evolução dos Índices de Gini Estadual e Municipal



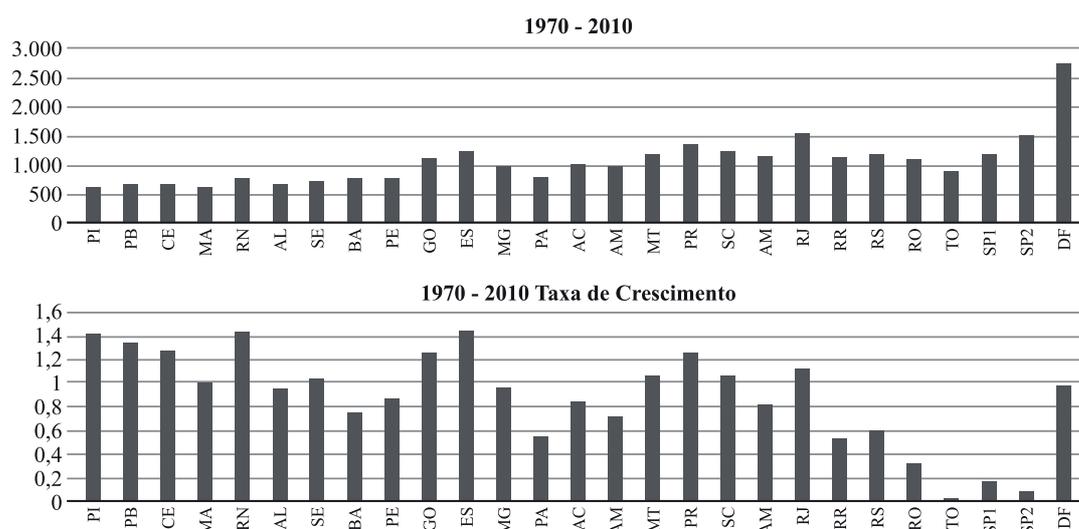
Fonte: Elaboração própria

Outra evidência interessante é a figura 3. Nesta ordena-se os estados do mais pobre ao mais rico em 1970 e reporta-se na parte superior da figura a renda destes estados em 2010. Já na parte inferior reporta-se a taxa de crescimento da renda

destes estados entre 1970 e 2010. Pode-se observar nesta figura duas evidências interessantes. A primeira é o fato de que, com exceção dos estados de Goiás, Espírito Santo, Minas Gerais e Tocantins, os estados mais ricos em 1970 continuam sendo os estados mais ricos em 2010. A segunda é a tendência de convergência das rendas *per capita* estaduais, principalmente quando analisa-se a partir de uma perspectiva de clubes de convergência, onde os estados mais pobres em 1970 apresentam taxas de crescimento mais elevadas do que os estados mais ricos, tal como preconiza a literatura de crescimento econômico.

Analisando estes dois gráficos, pode-se concluir que, apesar de os estados mais pobres terem crescido a taxas mais elevadas que os estados mais ricos, este crescimento não foi suficiente para alterar o padrão da distribuição regional de renda para uma perspectiva mais igualitária.

Figura 3 – Rendas médias Estaduais e suas Taxas de Crescimento entre 1970 e 2010



Fonte: Elaboração própria

## 4 Resultados

Antes de iniciar a análise, explica-se alguns procedimentos adotados para simplificar a interpretação dos resultados, pois foram estimados um total de 825 parâmetros. Primeiramente consolida-se os resultados em gráficos, de modo que ficou mais fácil de observá-los. Segundo, reporta-se no apêndice os parâmetros estimados apenas para as variáveis principais de interesse. Não foram introduzidas todas as variáveis devido à limitação de espaço do

texto. Por fim, de modo a simplificar o texto, toda vez que estiver analisando “indústria” refere-se à “indústria de transformação”.

Na figura 4 apresentam-se os resultados estimados para o logaritmo do rendimento de todos os trabalhos para os quantis 10, 25, 50, 75 e 90 da distribuição dos rendimentos, entre os anos de 1970 e 2010. A análise desta figura é bastante interessante pois pode-se observar, de um modo geral, para todos os quantis, três movimentos: i) aumento das rendas entre 1970 e 1980; ii) redução das ren-

das entre 1980 e 1991; e iii) evolução das rendas a partir de 1991. Isto significa que quando se analisa as rendas por quantis, os movimentos são bastante similares àqueles observados para a renda média como um todo, seguindo os movimentos ocorridos em cada década no Brasil.

Em seguida apresentamos na figura 5 os diferenciais em termos de logaritmos dos rendimentos e a contribuição dos Efeito Composição e do Efeito Estrutura Salarial para cada quantil ao longo do tempo. Com exceção do quantil 10, pode-se observar uma forte queda da desigualdade a partir do ano 1991. Aqui cabe ressaltar que as pessoas no quantil 10 são representadas por indivíduos que recebem menos de R\$ 100,00 reais por mês em todos os anos analisados, tanto para o Nordeste, quanto para o Sudeste. Enquanto os indivíduos que estão a partir do quantil 50 são aqueles próximos a faixa de recebimento do salário mínimo vigente em cada ano.

Esta tendência pode ser explicada, portanto, por dois movimentos. Para os quantis inferiores, refletiria uma elevação dos rendimentos provenientes das políticas sociais implementadas a partir do plano real e reforçada a partir de 2003 com o governo Lula. E, para os quantis superiores, pode estar refletindo tanto a política de valorização do salário mínimo, quanto a expansão do ensino médio e superior, que teve fortalecimento a partir de constituição de 1988, fortaleceu-se com a Lei de Diretrizes e Bases de 1996 e ampliou o seu efeito sobre a região Nordeste com a política de expansão da educação superior a partir do século XXI.

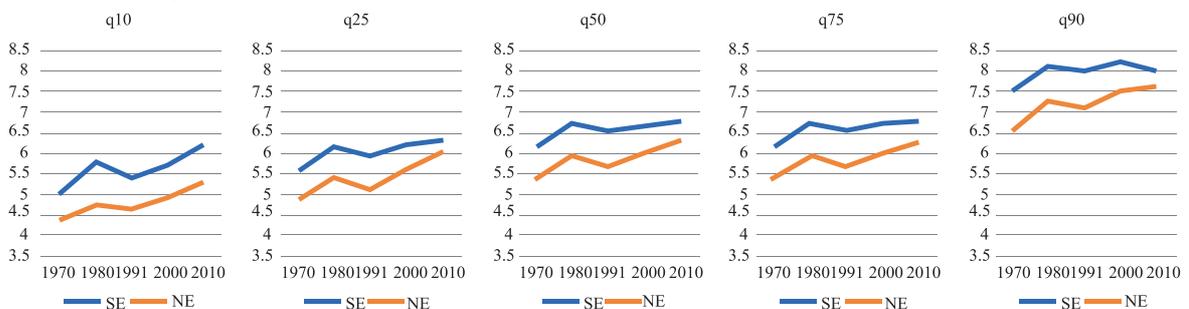
Quando se analisa a decomposição agregada, temos a primeira contribuição inédita deste trabalho: o Efeito Estrutura Salarial é mais importante para a explicação da desigualdade regional do que o Efeito Composição. Como o Efeito Estrutura Salarial agregado possui uma interpretação

causal, podemos afirmar que há um diferencial de rendimentos entre as pessoas das regiões Sudeste e Nordeste devido a uma discriminação regional. Ou, dito de outra forma, como resultado da remuneração diferente dos mesmos atributos produtivos devido ao fato dos indivíduos se localizarem em diferentes regiões.

Entretanto, há de se notar que a importância do Efeito Estrutura Salarial vai diminuindo a medida que avançamos para os quantis superiores, chegando muito próximo ao Efeito Composição para o quantil 90. Isto significa que o elemento "discriminatório" tende a ser reduzido à medida que caminhamos para quantis mais elevados da distribuição dos rendimentos. Uma possível explicação para este resultado é o fato de haver uma maior possibilidade de arbitrar dentre os indivíduos de rendimentos mais elevados. Isto é, se o indivíduo possui atributos produtivos que deveriam lhe conferir uma elevada remuneração, ele buscará residir em localidades que ofereçam tal benefício. Ou seja, um indivíduo com educação superior e outras características que o tornam produtivo buscará residir em algum município ou região que remunere de forma satisfatória estes atributos. Acredito que posso encontrar evidências na literatura que corroborem com este ponto.

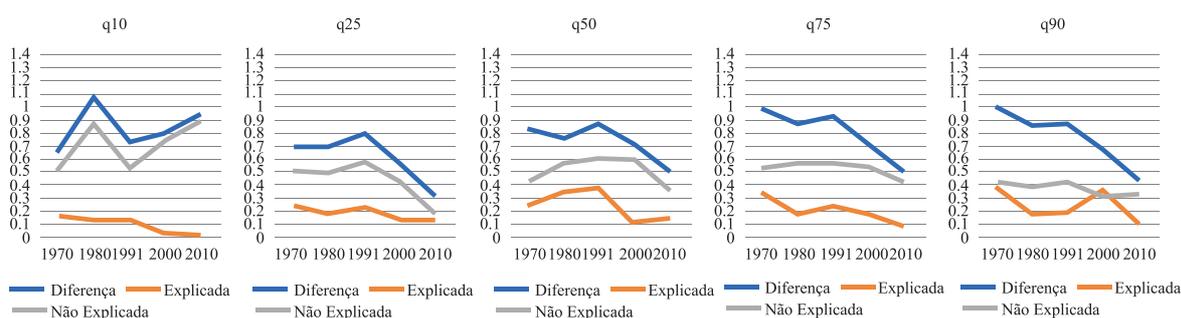
Já no caso da decomposição agregada do Efeito Composição (Figura 6) podemos observar que, apesar de sua contribuição para a desigualdade total ser inferior a do componente estrutura salarial, ela permanece muito importante. Cabe ressaltar que o efeito composição representa as diferenças de rendimentos devidos às diferenças de características em cada grupo. Caso não houvesse diferentes retornos às características ("discriminação entre as regiões"), as diferenças de rendimentos seriam remuneradas segundo este componente.

Figura 4 – Logaritmo rendimentos estimados para as regiões Nordeste e Sudeste por quantil e ao longo do tempo



Fonte: Elaboração própria

Figura 5 – Diferença no Logarítmo Rendimentos Estimados, Efeitos Composição e Estrutura Salarial Agregado



Fonte: Elaboração própria

Na decomposição detalhada do Efeito Composição está a segunda contribuição empírica e, provavelmente, a mais importante deste trabalho. Na literatura econômica brasileira verifica-se duas explicações, e consequentes proposições de resolução, para o problema da desigualdade regional. Enquanto a tradição desenvolvimentista baseada principalmente nos escritos de Celso Furtado defendeu que para reduzir as desigualdades regionais o Brasil deveria investir numa estrutura produtiva no Nordeste complementar àquela instalada no Sudeste, os pensadores neoclássicos, sobretudo aqueles que se baseiam na literatura de crescimento econômico, tem defendido que o investimento em educação seria o essencial, e que este foi negligenciado até a década de 1990. Os resultados deste estudo sugerem que as duas linhas de pesquisa estão corretas, porém cada argumento tem maior aderência a determinados momentos do tempo e para determinados quantis.

Observa-se na figura 6 que, para todos os quantis, a agropecuária possui um papel essencial para a explicação do Efeito Composição, enquanto a partir de 1991 a variável superior passa a desempenhar papel mais relevante. Ressalta-se ainda que a agropecuária tem maior peso nos quantis mais baixos enquanto o ensino superior tem papel mais importante a partir do quantil 75.

Este resultado sugere que, para os quantis inferiores e para anos anteriores a 1991, trabalhar na agropecuária no Sudeste ao invés de trabalhar em outro setor no Sudeste proporciona uma renda bem mais elevada do que a renda que os indivíduos que trabalham no setor agropecuário no Nordeste recebem em relação a trabalhar nos demais setores do Nordeste. Por outro lado, possuir ensino superior ao invés de outro tipo de qualificação (ensino médio ou fundamental) no Sudeste confere uma

renda mais elevada do que possuir ensino superior no Nordeste relativamente a outro tipo de qualificação no Nordeste.

Mais difícil de ser verificado nos gráficos, porém não menos importante, está o caso da contribuição da indústria de transformação (Ver Tabela 2 no apêndice). Até o ano 1991 e para os quantis 10, 25 e 50, a indústria é o segundo maior fator explicativo do Efeito Composição. Isto é, com exceção da agropecuária, trabalhar na indústria no Sudeste ao invés de outro setor, confere uma renda superior do que trabalhar na indústria no Nordeste, em comparação a indivíduos que trabalham em outros setores no Nordeste. Ao observar os anos 2000 e 2010 e os quantis 75 e 90, a indústria continua possuindo acentuada importância, porém o Ensino Superior toma o posto de segundo colocado na explicação do Efeito Composição agregado.

No caso da decomposição detalhada do Efeito Estrutura Salarial analisa-se o diferencial de retorno para cada característica produtiva entre as regiões. Por exemplo, se tomarmos duas pessoas com nível superior, uma no Sudeste e uma no Nordeste, o retorno por possuir nível superior é maior na primeira região? Se sim, isto significa que há um maior retorno a educação no Sudeste do que no Nordeste, mesmo os indivíduos possuindo as mesmas características produtivas.

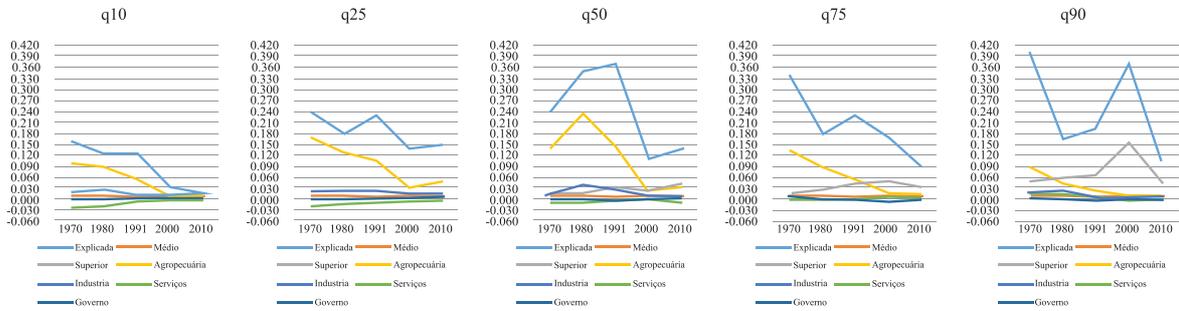
Os dados permitem observar que, dentre as variáveis que consideramos de maior relevância para este estudo, a educação é a variável que mais se destaca para a explicação do componente estrutura salarial. Esta observação se torna ainda mais evidente quando se analisa a partir do quantil 50 e, para o caso dos quantis 75 e 90, no qual o componente nível superior é importante em todo o período analisado.

Destaque maior para os quantis 75 e 90, onde

pode-se observar que o retorno a educação favorece a redução da desigualdade de renda. Isto é, o retorno a educação no quantil superior da distribuição é maior no Nordeste do que no Sudeste. A explicação para este resultado vem da própria di-

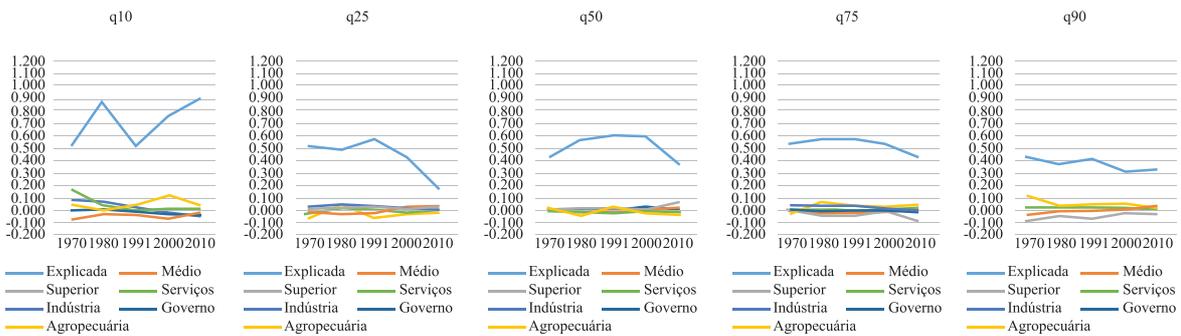
nâmica do mercado de trabalho, pois como há menos trabalhadores qualificados no Nordeste que no Sudeste, os movimentos de oferta e demanda elevam o prêmio salarial relativo à qualificação. Esta é a terceira contribuição inédita deste trabalho.

Figura 6 – Efeito composição detalhado



Fonte: Elaboração própria

Figura 7 – Efeito estrutura salarial detalhado



Fonte: Elaboração própria

## 5 Considerações Finais

Este artigo teve como objetivo analisar a desigualdade de rendimentos entre as regiões Sudeste e Nordeste do Brasil entre os anos de 1970 e 2010. Para a realização deste objetivo utilizou-se os censos demográficos realizados nos anos de 1970, 1980, 1991, 2000 e 2010 e estimou-se o logaritmo dos rendimentos médios por quantil para cada região através da Recentered Influence Function e depois a desigualdade foi decomposta nos moldes da tradicional decomposição Oaxaca - Blinder para diferentes quantis da distribuição dos rendimentos.

A análise de todo este período é de grande importância, pois em cada década analisada o país passou por cenários macroeconômicos bastante diversos, configurando dinâmicas das economias regionais muito diferentes. Como foi apresentado

ao longo do texto, nos períodos de crescimento da economia nacional as regiões mais pobres cresceram a taxas mais elevadas, enquanto nos períodos de recessão, estas regiões também foram as mais impactadas.

Os resultados encontrados contribuem com a literatura que busca analisar a desigualdade regional brasileira. Em primeiro lugar, encontrou-se que existe um diferencial de rendimentos que é explicado simplesmente pelo fato de as pessoas residirem em regiões diferentes, resultado este captado pelo Efeito Estrutura Salarial. Em outros termos, foi observado que, depois de controlar pelas características produtivas dos indivíduos, há um diferencial de rendimentos explicado por retornos desiguais a estas características entre as regiões. Em segundo, também pode-se observar que o Efeito Composição é relevante. Isto significa que as diferenças de características entre os indivíduos das regiões também

explicam a desigualdade regional de renda.

Ao realizar a decomposição detalhada três evidências se destacaram. A Primeira foi o papel da estrutura produtiva (indústria e, principalmente, agropecuária) explicando grande parte do Efeito Composição para os quantis inferiores até o ano de 1991. A segunda foi o fato de o nível superior passar a ser a característica mais importante para a explicação da desigualdade de renda entre as regiões a partir dos anos 2000 e para os quantis superiores. E, por fim, tem-se o fato de que ao comparar duas pessoas com as mesmas características em diferentes regiões, observou-se que o retorno a escolaridade no Nordeste é bastante superior ao retorno a Escolaridade no Sudeste para os quantis 75 e 90 e, sobretudo, a partir do ano 2000.

Outra evidência interessante é que a indústria de transformação, mesmo não assumindo papel protagonista em nenhum período, foi bastante importante para explicar a desigualdade em todos os anos e em todos os períodos.

Estes resultados possuem grande relevância para o necessário debate acerca da construção das políticas públicas regionais no Brasil. Não há precedentes na literatura de trabalho que tenha analisado a desigualdade regional em período tão longo a partir de uma base de dados compatibilizada e que utiliza o método de decomposição que possibilita realizar a decomposição Oaxaca-Blinder para todos os quantis da distribuição dos rendimentos.

## Referências

- ANGRIST, J. D.; KRUEGER, A. B. Does compulsory school attendance affect schooling and earnings? **Quarterly Journal of Economics**, v. 106, p. 979–1014, 1991.
- ASHENFELTER, O.; ROUSE, C. E. Income, schooling and ability: Evidence from a newsample of identical twins. **Quarterly Journal of Economics**, v. 113, p. 253–284. 1998.
- 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, 1997.
- BARRO, R. Economic growth in a cross section of countries. **Quarterly Journal of Economics**, v. 106, p. 407-43, 1991.
- BARRO, R.; SALA-I-MARTIN, X. Convergence. **Journal of Political Economy**, v. 100, n. 2, p. 223-251, 1992.
- BARSKY, R.; BOUND, J.; CHARLES, K. C.; LUPTON, J. P. Accounting for the Black-White Wealth Gap: A Nonparametric Approach, **Journal of the American Statistical Association**, v. 97, n. 459, p. 663-673, 2002.
- BAUMOL, W. J. Productivity growth, convergence and welfare. **American Economic Review**, v. 76, p. 1072-1085, December, 1986.
- BECKER, G. S. Human capital. Columbia Press University. 1964.
- DINARDO, J.; FORTIN, N. M.; LEMIEUX, T. Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach, **Econometrica**, v. 64, n. 5, p. 1001-1044, 1996.
- DINIZ, C. C. Celso Furtado e o Desenvolvimento Regional. **Nova Economia**, n 19, p. 227-249, Maio - Agosto, 2009.
- 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**, v. 32, p. 479-489, 2000.
- FERREIRA, P. C.; ELLERY Jr. R. G. 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, P. 38-56, 1995.
- FIRPO, S.; FORTIN, N.; LEMIEUX, T. Decomposing Wage Distributions using Recentered Influence Functions Regressions, mimeo, University of British Columbia. 2007. Disponível em: [http://economie.esg.uqam.ca/upload/files/seminaires/Lemieux\\_Thomas.pdf](http://economie.esg.uqam.ca/upload/files/seminaires/Lemieux_Thomas.pdf)

FIRPO, S.; FORTIN, N. M.; LEMIEUX, T. Unconditional Quantile Regressions, **Econometrica**, v. 77, n. 3, p. 953-973, 2009.

FORTIN, N.; LEMIEUX, T.; 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.

GUIMARÃES, J. F.; CAVALCANTI, T.; SILVEIRA-NETO, R. Accounting For Labor Income Differences, In: Brazil: The Role Of Human Capital. Encontro Nacional de Economia 35, 2006, **Anais...** ANPEC - Associação Nacional dos Centros de Pós-graduação em Economia.

HEYWOOD, J. S.; PARENT, D. **The White-Black Wage Gap and the Method of Pay**. McGill University. 2009. 46 p. Disponível em: <http://digitalcommons.ilr.cornell.edu/cr/17/>

HIRSCHMAN, A. Transmissão inter-regional e internacional do crescimento econômico. In: SCHWARTZMAN (Org), **Economia Regional – Textos Selecionados**. Cedeplar-MINTER, p. 291-313, 1977.

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

KOENKER, R.; BASSETT, G. Regression Quantiles, **Econometrica**, v. 46, n. 33-50, 1978

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

MACHADO, J. F.; MATA, J. Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression, **Journal of Applied Econometrics**, n. 20, p. 445-465, 2005.

MANKIW, G.; ROMER, D.; WELL, D. A Contribution to the Empirics of Economic Growth, *QJE*, may, P. 407- 437, 1992.

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, 2003.

MYRDAL, G. **Teoria econômica e regiões subdesenvolvidas**. Fundo de Cultura, Rio de Janeiro, 1957.

PERROUX, F. O conceito de polo de crescimento. In: SCHWARTZMAN (Org), **Economia Regional – Textos Selecionados**. CEDEPLAR MINTER, p. 145-156, 1977.

OAXACA, R. Male-Female Wage Differentials in Urban Labor Markets. **International Economic Review**, v. 14, p. 693-709, 1973

OAXACA, R. L.; RANSOM, M. R. On discrimination and the decomposition of wage differentials. **Journal of Econometrics**, v. 61, n. 5, p. 21-23, 1994.

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

PESSOA, S. Existe um problema de desigualdade regional no Brasil? In: ENCONTRO NACIONAL DE ECONOMIA, 29, 2000. **Anais...** Salvador. ANPEC. 2000.

PESSÔA, S. A.; BARBOSA FILHO, Fernando de Holanda . Retorno da Educação no Brasil. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v. 38, p. 97-125, 2008.

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

ROUSE, C. E. **Further estimates of the economic return to schooling firm a new sample of twins**. Unpublished Discussion Paper (Industrial Relations Section, Princeton University, 1997.

SILVEIRA NETO, R. M.; CAMPELO, A. K.  
O perfil das disparidades regionais de renda no Brasil: evidências a partir de regressões quantílicas para os anos de 1992 e 2001. In: ENCONTRO NACIONAL DE ECONOMIA, 31., 2003, Porto Seguro. **Anais...** Porto Seguro, 2003.

SILVEIRA NETO, R. M.; MENEZES, T. A.  
**Disparidades Regionais de Renda no Brasil: Analisando a Importância do Capital Humano.**  
In: Fundação Konrad-Adeneur. (Org.). Fórum Brasil-Europa. Fundação Konrad-Adeneur. 2008.

SILVEIRA NETO, R. M.; AZZONI, C. R.  
Disparidades regionais de renda no Brasil: qual a importância das amenidades regionais. In: ENCONTRO DE ECONOMIA REGIONAL - ANPEC NORDESTE, 21, 2004. **Anais...** Fortaleza: Banco do Nordeste/ANPEC. 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.

SCHULTZ, T. W. **Investment in Human Capital.** The Free Press, New York. 1971. 272 p.

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

YUN, M. A Simple Solution to the Identification Problem in Detailed wage Decomposition. **Economic Inquiry**, v. 43, p. 766-772. with "Erratum," **Economic Inquiry** (2006), v. 44, p. 198, 2005.