

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

INSTITUTO DE ECONOMIA

MONOGRAFIA DE BACHARELADO

**MOBILIDADE INTRAGERACIONAL DE RENDA NO
BRASIL (2003-2014):**

Uma abordagem markoviana

FÁBIO MOREIRA ALVES ESPERANÇA

matrícula nº 116031052

ORIENTADOR: Prof. Getúlio Borges da Silveira Filho

JANEIRO 2021

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

INSTITUTO DE ECONOMIA

MONOGRAFIA DE BACHARELADO

**MOBILIDADE INTRAGERACIONAL DE RENDA NO
BRASIL (2003-2014):**

Uma abordagem markoviana

Fábio Esperança

FÁBIO MOREIRA ALVES ESPERANÇA

matrícula nº 116031052

ORIENTADOR: Prof. Getúlio Borges da Silveira Filho

JANEIRO 2021

As opiniões expressas neste trabalho são da exclusiva responsabilidade do autor

AGRADECIMENTOS

Agradeço especialmente à minha família, pelo contínuo apoio durante todos esses anos.

Agradeço também a todo o corpo docente do Instituto de Economia, pela imensurável contribuição que tiveram e continuam tendo na minha trajetória intelectual; em especial, agradeço ao meu orientador, Getúlio Borges, pela paciência e solicitude com que me guiou durante todo o processo de urdidura deste trabalho, apesar da distância física.

Agradeço ainda a todos ao corpo discente do Instituto de Economia, pelo agradável e contínuo convívio durante os últimos anos, e a todos os técnicos e servidores da UFRJ, que ajudaram e seguem ajudando a construir uma educação pública de qualidade.

RESUMO

Este trabalho visa a estudar o processo de mobilidade social entre os anos de 2003 e 2014 para o Brasil, bem como para alguns de seus subgrupos populacionais. Para isso, foi usado o edifício teórica das cadeias de Markov, bem como a abordagem axiomática dos índices de mobilidade proposta por SHORROCKS (1978) e o método de entropias relativas, apresentado por AEBI et. al. (1999). Este estudo observou que o Brasil no período pode ser caracterizado como um país com uma série de disparidades sociais, bem como uma baixa mobilidade social e uma baixa probabilidade de indivíduos mais pobres alcançarem as classes superiores; apesar disso, notou-se significativa melhora na distribuição de renda do país, com ascensão das classes mais pobres em direção às classes médias.

Palavras-chave: Cadeias de Markov. Entropia relativa. Mobilidade social.

ABSTRACT

The purpose of this study is to address the social mobility in Brazil (as well as in some of its populational subsets) between 2003 and 2014. To this end, the theoretic edifice of Markov chains was used, as well as the axiomatic approach of mobility indices proposed by SHORROCKS (1978) and the relative entropy method as presented by AEBI et. al. (1999). This study found that Brazil in this period can be characterized as a country with a series of social inequalities, as well as a low social mobility, with a low probability that individuals in the poorer classes reach the upper classes; despite that, it was possible to note a significant improvement in the income distribution, with a rise in the income of poorer individuals towards the middle class.

Keywords: Markov chains. Relative entropy. Social mobility.

LISTA DE ILUSTRAÇÕES

| | |
|---|----|
| FIGURA 1 - Média da renda familiar per capita (Brasil)..... | 18 |
| FIGURA 2 - Média da renda familiar per capita (subgrupos raciais)..... | 19 |
| FIGURA 3 - Média da renda familiar per capita (macrorregiões)..... | 21 |

LISTA DE TABELAS

| | |
|--|----|
| TABELA 1 - Critério de estratificação de renda..... | 17 |
| TABELA 2 - Percentual da população por classe de renda (Brasil)..... | 18 |
| TABELA 3 - Percentual da população por classe de renda (brancos)..... | 20 |
| TABELA 4 - Percentual da população por classe de renda (pretos e pardos)..... | 20 |
| TABELA 5 - Percentual da população por classe de renda (Nordeste)..... | 21 |
| TABELA 6 - Percentual da população por classe de renda (Norte)..... | 22 |
| TABELA 7 - Percentual da população por classe de renda (Centro - Oeste)..... | 22 |
| TABELA 8 - Percentual da população por classe de renda (Sudeste)..... | 22 |
| TABELA 9 - Percentual da população por classe de renda (Sul)..... | 22 |
| TABELA 10 - Matriz de transição estimada (Brasil)..... | 23 |
| TABELA 11 - Matriz de transição estimada (pretos e pardos)..... | 24 |
| TABELA 12 - Matriz de transição estimada (brancos)..... | 25 |
| TABELA 13 - Matriz de transição estimada (Norte)..... | 25 |
| TABELA 14 - Matriz de transição estimada (Nordeste)..... | 26 |
| TABELA 15 - Matriz de transição estimada (Sudeste)..... | 26 |
| TABELA 16 - Matriz de transição estimada (Centro - Oeste)..... | 26 |
| TABELA 17 - Matriz de transição estimada (Sul)..... | 27 |
| TABELA 18 - Índices de mobilidade (Brasil)..... | 28 |
| TABELA 19 - Índices de mobilidade (subgrupos raciais)..... | 28 |
| TABELA 20 - Índices de mobilidade (macrorregiões)..... | 29 |
| TABELA 21 - Distribuição-limite das classes de renda (Brasil)..... | 30 |
| TABELA 22 - Distribuição-limite das classes de renda (pretos e pardos)..... | 30 |

| | |
|---|----|
| TABELA 23 - Distribuição-limite das classes de renda (brancos)..... | 31 |
| TABELA 24 - Distribuição-limite das classes de renda (Nordeste)..... | 31 |
| TABELA 25 - Distribuição-limite das classes de renda (Norte)..... | 31 |
| TABELA 26 - Distribuição-limite das classes de renda (Sudeste)..... | 31 |
| TABELA 27 - Distribuição-limite das classes de renda (Sul)..... | 32 |
| TABELA 28 - Distribuição-limite das classes de renda (Centro - Oeste)..... | 32 |
| TABELA 29 - Matriz a priori estimada (Brasil)..... | 40 |
| TABELA 29 - Matriz a priori estimada (brancos)..... | 40 |
| TABELA 29 - Matriz a priori estimada (pretos e pardos)..... | 40 |
| TABELA 29 - Matriz a priori estimada (Centro - Oeste)..... | 41 |
| TABELA 29 - Matriz a priori estimada (Nordeste)..... | 41 |
| TABELA 29 - Matriz a priori estimada (Norte)..... | 41 |
| TABELA 29 - Matriz a priori estimada (Sudeste)..... | 41 |
| TABELA 29 - Matriz a priori estimada (Sul)..... | 42 |

SUMÁRIO

| | | |
|--------------|--|-----------|
| 1 | INTRODUÇÃO..... | 11 |
| 2 | CADEIAS DE MARKOV E A ABORDAGEM AXIOMÁTICA DOS ÍNDICES DE MOBILIDADE..... | 13 |
| 2.1 | Processos estocásticos..... | 13 |
| 2.2 | Cadeias de Markov..... | 13 |
| 2.3 | Matrizes de Markov e sua aplicabilidade à mobilidade de renda..... | 15 |
| 2.4 | Índices de mobilidade | 16 |
| 3 | O MÉTODO DE ENTROPIAS RELATIVAS..... | 18 |
| 3.1 | O problema da estimação com dados agregados..... | 18 |
| 3.2 | A entropia relativa..... | 18 |
| 3.3 | O IPFP..... | 19 |
| 4 | ANÁLISE EMPÍRICA..... | 21 |
| 4.1 | Estatísticas descritivas..... | 21 |
| 4.1.1 | Brasil..... | 22 |
| 4.1.2 | Subgrupos raciais..... | 23 |
| 4.1.3 | Macrorregiões..... | 24 |
| 4.2 | Matrizes de transição..... | 27 |
| 4.2.1 | Brasil..... | 27 |
| 4.2.2 | Subgrupos raciais..... | 28 |
| 4.2.3 | Macrorregiões..... | 30 |
| 4.3 | Índices de mobilidade..... | 32 |
| 4.3.1 | Brasil..... | 32 |

| | |
|--------------------------------------|-----------|
| 4.3.2 Subgrupos raciais..... | 32 |
| 4.3.3 Macrorregiões..... | 33 |
| 4.4 Distribuições-limite..... | 33 |
| 4.4.1 Brasil..... | 34 |
| 4.4.2 Subgrupos raciais..... | 35 |
| 4.4.3 Macrorregiões..... | 35 |
| 5 CONCLUSÕES..... | 37 |
| REFERÊNCIAS..... | 38 |
| APÊNDICE A..... | 40 |
| APÊNDICE B..... | 41 |

1 INTRODUÇÃO

Cada vez mais, a desigualdade social vem sendo percebida como uma questão candente no debate econômico internacional, com a publicação de uma grande quantidade de estudos dedicados a mensurá-la e analisá-la de diversos pontos de vista. A maioria dos estudos sobre desigualdade, entretanto, tratam-na de maneira estática, ou seja, buscam apresentar o estado da desigualdade em um ponto do tempo, o que provê uma análise incompleta do panorama, visto que desconsidera a dinâmica da distribuição de renda no tempo (FIELDS e OK, 1999).

Uma abordagem alternativa a essa é dada pelos estudos de mobilidade de renda, que visam a mostrar como a distribuição de renda de uma determinada sociedade evolui ao longo do tempo, o que pode dar um quadro mais completo da sua dinâmica social. De acordo com CORAK (2016, p. 2):

Snapshots leave more to the imagination than movies. This is certainly the case when it comes to measuring and understanding inequality. Whether the degree of inequality in a society is “too high” or “too low” is hard to say, and therefore it is hard to imagine what the public policy implications should be. How did it arise? What will happen to it in the future? These seem to be reasonable questions to ask in trying to interpret a picture taken at a particular point in time.¹

A literatura sobre mensuração da mobilidade social, entretanto, é bastante heterogênea, podendo variar tanto no período de análise quanto no método de análise. Em relação ao primeiro, as chamadas análises intrageracionais buscam mensurar a evolução da renda de um indivíduo ao longo do tempo, enquanto as análises ditas intergeracionais se concentram em diferenças de rendimentos entre pais e filhos; em relação aos métodos de análise, a mensuração pode se dar por métodos “welfaristas”, “markovianos”, entre outros (para uma introdução aos principais métodos de mensuração da mobilidade de renda, ver FIELDS e OK (1999)).

Um dos métodos mais profícuos de avaliação da evolução da mobilidade de renda de uma sociedade se dá através do uso de matrizes de transição de Markov. Essas matrizes nos mostram qual a probabilidade de que pessoas que estão em certa classe de renda em um período migrem para outra classe determinada no período seguinte; a partir dessas matrizes,

¹ " Retratos deixam mais à imaginação do que filmes. Este certamente é o caso quando se trata de medir e entender a desigualdade. Se o grau de desigualdade de uma sociedade é 'muito alto' ou 'muito baixo' é difícil de dizer e, portanto, é difícil imaginar quais deveriam ser as implicações em termos de políticas públicas. Como ela [a desigualdade] emergiu? O que irá ocorrer com ela no futuro? Essas parecem questões razoáveis de se perguntar ao tentar interpretar um retrato tirado em um momento específico do tempo." (tradução nossa)

ainda, é possível obter índices de mobilidade social, o que permite a comparação entre sociedades.

Entretanto, uma série de dificuldades emerge na estimação dessas matrizes, em especial a ausência de disponibilidade de dados individuais, o que torna impossível o acompanhamento da evolução de renda de cada indivíduo ao longo do tempo. Apesar desse obstáculo, alguns métodos podem ser utilizados para se estimar consistentemente matrizes de transição a partir de dados agregados.

Essa monografia divide-se em 3 capítulos, além desta Introdução e da Conclusão: o capítulo 1 visa a introduzir ao leitor o conceito de processos estocásticos, em especial as cadeias de Markov, bem como suas hipóteses subjacentes, sua aplicabilidade para a análise econômica e os índices de mobilidade que podem ser obtidos a partir delas; o capítulo 2 tem como objetivo apresentar o método de entropia relativa, utilizado para estimar consistentemente matrizes de transição com dados agregados. O capítulo 3, por fim, busca aplicar os métodos apresentados para estimar matrizes de transição para a economia brasileira como um todo e para subgrupos populacionais (divididos por raça e região) no período entre 2003 e 2014, de modo a responder às seguintes questões: qual o nível de mobilidade social da economia brasileira no período? Quais subgrupos tiveram maior mobilidade de renda?

2 CADEIAS DE MARKOV E A ABORDAGEM AXIOMÁTICA DOS ÍNDICES DE MOBILIDADE

2.1 Processos estocásticos

Um processo estocástico pode ser definido simplesmente (DOBROW, 2016) como um conjunto de variáveis aleatórias $\{X_t, t \in I\}$. O conjunto I , indexador do processo, é normalmente identificado com o tempo; outrossim, o conjunto (notemo-lo S) onde as variáveis X_t tomam valores é denominado "espaço de estados". Um processo estocástico pode ser classificado pelas características do seu conjunto indexador, de seu "espaço de estados", ou pelas relações que as variáveis do processo têm entre si: quanto ao conjunto indexador, o processo pode ser de tempo discreto (quando o indexador é um subconjunto dos números naturais) ou de tempo contínuo (quando o indexador é um intervalo real²); em relação ao espaço de estados, podemos ter um processo de estado finito (em que o espaço de estado é um conjunto finito) ou infinito.

2.2 Cadeias de Markov

As cadeias de Markov são um tipo específico de processo estocástico, em que a relação de dependência entre as variáveis X_0, X_1, \dots (no caso de tempo contínuo) tomando valores no conjunto (finito) S , é dada por:

$$P(X_{t+1} = j | X_0 = x_0, \dots, X_t = i) = P(X_{t+1} = j | X_t = i)$$

Portanto, a idiosincrasia das cadeias de Markov é que, nelas, o valor que a variável toma em um determinado período depende explicitamente apenas do seu valor no instante mais recente, estando sua trajetória pregressa expressa nele. Mais formalmente, essa propriedade é denominada "markovianidade de primeira ordem"³. Ademais, se as probabilidades condicionais independem do tempo t , i.e., $P(X_{t+1} = j | X_t = i) = P(X_1 = j | X_0 = i)$, dizemos que a cadeia é "temporalmente homogênea", e denotamos:

$$p_{ij} := P(X_{t+1} = j | X_t = i), \quad \forall i, j \in S$$

² Mais geralmente, um retângulo em \mathbb{R}^d .

³ De modo mais geral, um processo de Markov é dito de i -ésima ordem quando o valor no período atual depende explicitamente apenas de seu valor nos i períodos anteriores.

A probabilidade condicional p_{ij} de uma cadeia de Markov homogênea é denominada "probabilidade de transição" do estado i para o j , e a matriz quadrada P definida por:

$$P := (p_{i,j}), i, j \in S$$

É chamada "matriz de transição" do processo.

Uma matriz de transição, então, reúne em suas células as probabilidades de transição entre todos os estados do processo; como corolário de sua definição, ela possui ainda duas características que a distinguem das demais matrizes quadradas. Essas características, tão cruciais quanto triviais, são:

$$(i) p_{ij} \geq 0, \forall i, j \in S$$

$$(ii) \sum_{j \in S} P_{ij} = 1, \forall i \in S$$

Pelas duas condições acima, então, uma matriz de transição qualquer deve ter todas as suas células não-negativas (o que advém diretamente da definição das células como uma probabilidade), e cada uma de suas linhas deve somar a unidade.

De posse da matriz de transição de um período P , é possível, então, estender seus resultados para intervalos maiores: a matriz de transição de t períodos pode ser obtida pela simples elevação à t -ésima potência de P . De acordo com DOBROW (2016, p. 53):

$$P_{ij}^t = P(X_t = j | X_0 = i), \forall i, j \quad (1)$$

Assim, a função de probabilidade $P(X_t = i)$ da variável X_t pode ser dada pela igualdade (Dobrow, 2016, p.56):

$$P(X_t = j) = \sum_{i \in S} P(X_t = j | X_0 = i) P(X_0 = i) = \sum_{i \in S} n_i P_{ij}^t \quad (2)$$

O vetor-linha n , que indica, em sua i -ésima célula, a proporção da população que se encontrava no estado i ao início do processo, é denominada "distribuição inicial". De posse desta e da matriz de transição homogênea P que rege o processo, temos, pelas equações:

$$P(X_t = j) = (n P^t)_j, \forall j \in S \quad (3)$$

Desse modo, munidos da distribuição inicial n e da matriz P , é possível mapear a distribuição de probabilidade das variáveis X_0, \dots, X_t , em qualquer etapa do processo.

2.3 Matrizes de Markov e sua aplicabilidade à mobilidade de renda

As cadeias de Markov vistas na seção anterior podem muito convenientemente retratar a evolução da distribuição de renda de uma sociedade ao longo de um período. Nesse caso, definimos o "espaço de estados" S como as diferentes classes de renda de uma economia; assim, a célula p_{ij} da matriz de transição nos indica a proporção de indivíduos que se encontravam na classe i e migraram para a classe j .

Para que utilizemos matrizes de Markov, entretanto, uma série de hipóteses deve ser adicionada. São elas, segundo SHORROCKS (1976):

- a) Homogeneidade populacional, i.e., a probabilidade de transição de indivíduos que estão na mesma classe de renda é igual;
- b) Markovianidade de primeira ordem⁴;
- c) Homogeneidade temporal, i.e., a matriz de transição é homogênea, mantendo-se constante ao longo do tempo.

Desse modo, a evolução da distribuição de renda de uma sociedade pode ser descrita pelo seguinte processo:

$$n_{t+1} = n_t P \quad (4)$$

Em que n_t representa a distribuição da população entre as classes em cada período t .

Se for possível ter acesso a cada um dos indivíduos da amostra, a estimação da matriz pode ser facilmente obtida por meio do método de máxima verossimilhança, e as suas células serão dadas pela proporção:

$$\hat{p}_{ij} = \frac{k_{ij}}{\sum_{j \in S} k_{ij}}$$

Onde k_{ij} representa o número total de indivíduos que migraram da classe i para a classe j , e o denominador representa todos os indivíduos que estavam na classe j . Se, entretanto, não for possível obter dados de cada indivíduo ao longo do tempo, o processo de estimação torna-se menos direto.

⁴ Essa hipótese, entretanto, pode ser relaxada de modo a permitir cadeias de Markov de ordens superiores.

2.4 Índices de mobilidade

De posse da matriz de transição estimada, é possível obter uma série de índices, que buscam resumir a magnitude da mobilidade social de determinada economia, permitindo, portanto, a comparação mais direta entre diferentes sociedades.

Com esse intuito, é possível definir axiomáticamente uma função real contínua M (denominada "índice de mobilidade") sobre o conjunto Π das matrizes de transição, que atende às seguintes condições (FIGUEIREDO, 2006):

a) Normalização, i.e., $0 \leq M(P) \leq 1, \forall P \in \Pi$

b) Monotonicidade: como as transições entre classes são dadas pelos elementos fora da diagonal principal, se há o aumento de um desses componentes em detrimento das células da diagonal, o índice deve aumentar. Assim, escrevendo $P > P'$ quando $p_{ij} \geq p'_{ij}, \forall i \neq j$ e $p_{ij} > p'_{ij}$ para algum $i \neq j$, temos que $P > P'$ implica $M(P) > M(P')$ (SHORROCKS, 1978).

c) Imobilidade, i.e., $M(I) = 0$

d) Mobilidade perfeita: $M(P) = 1$, se $P = u * x'$, onde $u = [1 \ 1 \ \dots \ 1]'$ e $x' u = 1$.

Um índice de mobilidade, portanto, é uma função que deve tomar valores entre 0 e 1, sendo 0 reservado para matrizes-identidade (em que qualquer indivíduo permaneceria eternamente na mesma classe) e 1 para matrizes com linhas iguais (nas quais a classe de um indivíduo no período presente é completamente independente de sua posição no período anterior). Ademais, FIGUEIREDO (2006, p. 438) indica um modo de ordenar socialmente as matrizes ao afirmar que "se uma matriz P apresentar maior mobilidade do que P' , ela será socialmente preferível". Entretanto, o índice de mobilidade não nos indica a direção da mobilidade, de modo que uma matriz com alta mobilidade pode ser compatível com uma piora significativa na distribuição de renda, como será visto posteriormente; por esse motivo, furtaremos-nos de fazer estas hierarquizações entre índices.

Um dos principais índices que atendem às condições acima é o índice de Prais (Figueiredo, 2010):

$$M_p(P) = \frac{r - \text{tr}(P)}{r - 1}$$

Onde r representa a ordem da matriz e $tr(P)$, seu traço.

Entretanto, um problema emerge na formulação dos índices: a comparação entre matrizes que tratam de períodos com durações distintas. Se não for levado em conta o efeito do tempo, é possível chegar à conclusão de que uma sociedade tem maior mobilidade social do que outra pelo simples fato de que o período medido foi maior. Levemos em conta, por exemplo, as seguintes matrizes de transição, retiradas de SHORROCKS (1978):

$$P = \begin{bmatrix} 0,9 & 0,1 & 0 \\ 0,3 & 0,4 & 0,3 \\ 0,3 & 0,3 & 0,4 \end{bmatrix} \quad Q = P^2 = \begin{bmatrix} 0,44 & 0,28 & 0,28 \\ 0,28 & 0,44 & 0,28 \\ 0,28 & 0,28 & 0,44 \end{bmatrix}$$

Nesse caso, a matriz Q é apenas a matriz P elevada ao quadrado (e, logo, definida para o dobro do tempo). Desse modo, embora ambas as matrizes descrevam o mesmo processo, a diferença no período compreendido faz com que Q pareça mais móvel. Utilizando o índice de Prais, por exemplo, temos:

$$M_p(P) = 0,64 < 0,84 = M_p(Q)$$

A fim de debelar esse problema, SHORROCKS (1978) adiciona às quatro condições anteriores uma outra:

(v) Invariância temporal, i.e., matrizes que descrevem processos iguais devem ter mesmo índice, independentemente do tamanho do período que compreendam. Formalmente, sendo T o número de períodos compreendidos pela matriz, devemos ter::

$$M(P, T) = M(P^k, kT), \quad k \geq 1$$

É possível, entretanto, construir índices que satisfaçam o critério de invariância temporal e que podem ser demonstrados compatíveis com as quatro condições, desde que os autovalores de P sejam reais e não-negativos (Figueiredo, 2006). Entre eles:

$$M_L(P) = 1 - \lambda_2, \text{ sendo } \lambda_2 \text{ o segundo maior autovalor da matriz de transição.}$$

Assim, portanto, as cadeias de Markov nos provêm uma poderosa ferramenta de análise dos processos de mobilidade social das economias, nos mostrando as probabilidades de transição de indivíduos entre diversas classes de renda, além de permitir, por meio da utilização dos índices de mobilidade, a comparação entre sociedades e períodos distintos.

3 O MÉTODO DE ENTROPIA RELATIVA

3.1 O problema da estimação com dados agregados

Como vimos, a estimação de matrizes de transição quando há disponibilidade de dados individuais é simples e direta. Entretanto, muitas vezes essas informações não estão acessíveis, sendo possível apenas obtermos dados relativos à proporção de indivíduos em cada classe a cada período de tempo; desse modo, o problema que emerge é o de estimar as matrizes de transição tendo como informação apenas as seguintes condições (denominadas "restrições de continuidade" ou "condições inicial e terminal"):

$$n_{t+1} = P n_t \quad (5)$$

$$\sum_{i \in S} P_{ij} = 1, \forall i \in S \quad (6)$$

Entretanto, para quaisquer processos com $S > 2$, as equações acima não nos dão uma solução única para o problema, visto que temos S^2 incógnitas (dadas pelas células da matriz P) e apenas $2S$ equações. É preciso, portanto, buscar em outro local a teoria que nos permita estimar P .

3.2 A entropia relativa

Segundo COVER e THOMAS (1991, p. 18), o conceito de entropia relativa pode ser aproximadamente⁵ descrito como "a measure of distance between two distributions"⁶. Ainda segundo COVER e THOMAS (1991, p.18): "the relative entropy $D(p||q)$ is the measure of the inefficiency of assuming that the distribution is q when the true distribution is p ."⁷ Numericamente, a entropia relativa (também denominada "distância Kullback-Leibler") entre as funções de probabilidade $p(x)$ e $q(x)$ é dada por:

$$D(p||q) = \sum p(x) * \log\left(\frac{p(x)}{q(x)}\right) = E_p \log\left(\frac{p(x)}{q(x)}\right) \quad ^8$$

⁵ "Aproximadamente" porque, ao contrário da distância, a entropia relativa não é simétrica.

⁶ "Uma medida de distância entre duas distribuições." (tradução nossa)

⁷ "A entropia relativa $D(p||q)$ é a medida da ineficiência de assumir que a distribuição é q quando a verdadeira distribuição é p ." (tradução nossa)

⁸ Caso a variável seja contínua, a entropia relativa é dada pela integral $\int_{-\infty}^{\infty} p(x) \log\left(\frac{p(x)}{q(x)}\right) dx$.

É possível demonstrar, ademais, que $D(p||q)$ é sempre não-negativa, e igual a zero se e somente se as duas distribuições coincidem, i.e., $p = q$.

Assim, o problema que se põe é o seguinte: dadas uma matriz de transição a priori P_0 e as "restrições de continuidade" (equações 5 e 6), é preciso encontrar a matriz de transição P que minimize a entropia relativa $D(P||P_0)$; a solução, então, é dada pelo Lagrangeano (GOLAN, 2006):

$$L = D(P||P_0) + \sum_{i \in S} \lambda_i (n_{t+1_i} - \sum_{i \in S} p_{ij} n_{t_i}) + \sum_{j \in S} \mu_j (1 - \sum_{i \in S} p_{ij})$$

De modo intuitivo, o que buscamos é aquela matriz de transição que, dadas as restrições de continuidade, esteja mais próxima da matriz hipotética P_0 . O que nos resta obter, portanto, é a matriz a priori P_0 . AEBI et. al. (2008) sugerem utilizar como matriz hipotética uma matriz tridiagonal, i.e., que tem como suposição o fato de que um indivíduo só pode migrar em um período de tempo para classes adjacentes àquela em que ele se encontra atualmente. Dada essa hipótese, estimamos por meio de mínimos quadrados ordinários⁹ a matriz P_0 e ajustamo-la por meio do método de entropias relativas.

3.3 O IPFP

O método de entropia relativa acima pode ser convenientemente implementado por meio do algoritmo denominado *Iterative Proportional Fitting Procedure* (IPFP).

Originalmente descrito por DEMING e STEPHAN (1940), o algoritmo de "*Iterative proportional fitting procedure*" (IPFP) pode ser compreendido como "a procedure for assigning values to internal cells based on known marginal totals in a multidimensional matrix"¹⁰ (LOVELACE et. al., 2015, p.2). No nosso caso, o algoritmo tem como função ajustar os valores das células da matriz de transição a priori de modo a coincidir com as restrições de continuidade (5) e (6). Ademais, é possível demonstrar que a aplicação do algoritmo faz com que a matriz convirja à matriz "verdadeira", gerando estimativas "que são

⁹ Para uma exposição mais detalhada do método, ver apêndice A.

¹⁰ "Um procedimento para atribuir valores às células internas baseado nas marginais totais conhecidas em uma matriz multidimensional." (tradução nossa)

equivalentes às estimativas de máxima verossimilhança e essas estimativas minimizam a entropia relativa" ¹¹ (AEBI et. al., 1999, p. 132, tradução nossa).

¹¹ "that are equivalent to maximum likelihood estimates and that these estimates minimize the relative entropy"

ANÁLISE EMPÍRICA

4.1 Base de dados

O critério de renda utilizado neste trabalho foi o de "renda familiar per capita", oriundo da Pesquisa Nacional de Amostra por Domicílio (PNAD), recolhido pelo Instituto Brasileiro de Geografia e Estatística (IBGE) no mês de setembro de cada ano junto a famílias; a renda foi deflacionada pelo Índice de Preços ao Consumidor Amplo (IPCA), calculado pelo mesmo IBGE. Devido à pesquisa do Censo, a PNAD não foi realizada no ano de 2010; os dados deste ano, portanto, foram aproximados por meio de uma interpolação linear utilizando os dados dos demais anos.

A divisão em classes foi feita de acordo com o critério do Centro de Políticas Sociais da Fundação Getúlio Vargas, reproduzido abaixo, que usa como indicador a renda familiar per capita, tendo como referência os preços de janeiro de 2014.

Tabela 1: Critério de estratificação de renda

| Classes econômicas | Limite inferior | Limite superior |
|--------------------|-----------------|-----------------|
| Classe E | 0 | R\$ 1.254,00 |
| Classe D | R\$ 1.255,00 | R\$ 2.004,00 |
| Classe C | R\$ 2.005,00 | R\$ 8.640,00 |
| Classe B | R\$ 8.641,00 | R\$ 11.261,00 |
| Classe A | R\$ 11.262,00 | - |

Fonte: NERI (2014)

Por fim, os subgrupos foram estabelecidos da seguinte forma: além de uma análise geral para o Brasil como um todo, foram feitas duas subdivisões: a primeira, uma subdivisão racial, em que comparamos a evolução de indivíduos brancos com aquela de indivíduos negros (estes definidos como aquelas pessoas que apareciam, na variável "raça ou cor de pele" da PNAD, sob a denominação "pretos" ou "pardos"); a segunda, uma divisão em que buscamos comparar a mobilidade nas cinco diferentes macrorregiões do país (Centro-Oeste, Nordeste, Norte, Sudeste, Sul).

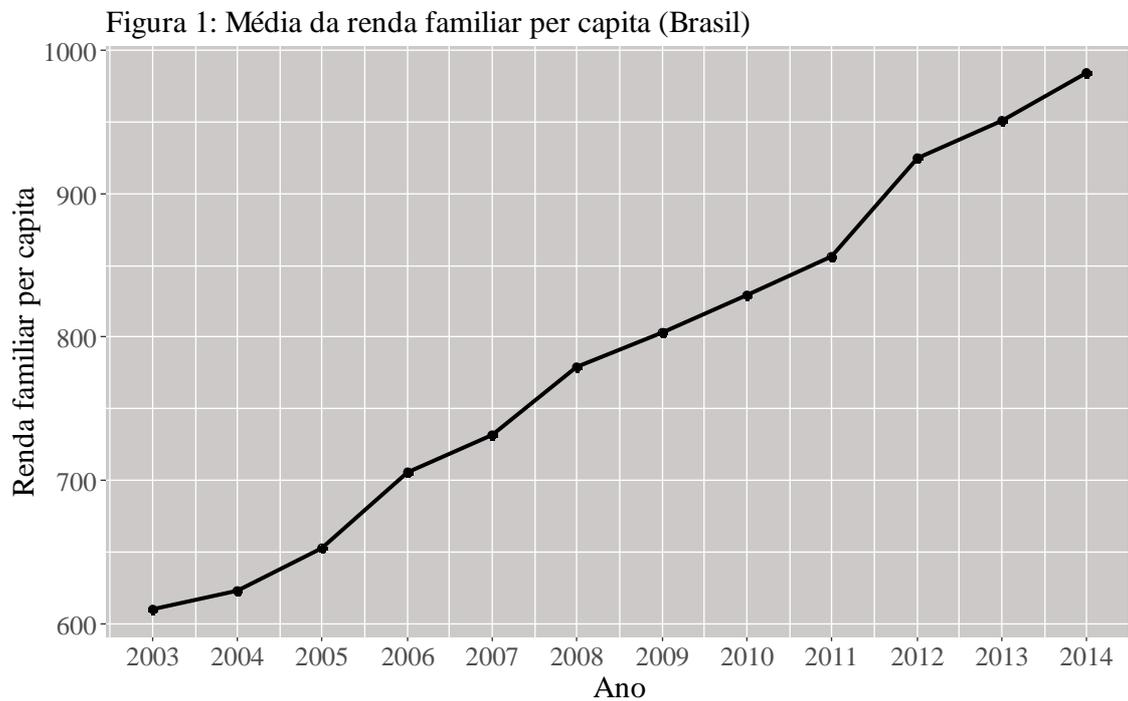
4.2 Estatísticas descritivas

Embora, como vimos, as matrizes de Markov nos provenham uma potente ferramenta de análise da mobilidade social de uma economia, há uma série de aspectos da evolução de uma sociedade que pode ser apreendida pela simples observação de algumas estatísticas

descritivas, como a média do rendimento familiar per capita e a evolução das proporções da população em cada classe de renda.

4.2.1. Brasil

No período em questão, observou-se no Brasil como um todo um aumento significativo e persistente na renda média familiar per capita (Figura 1), que passou de R\$609,66 em setembro de 2003 para R\$983,94 no mesmo mês de 2014 (ambos em preços correntes de janeiro de 2014), um incremento de aproximadamente 61%.



Fonte: IBGE. Elaborado pelo autor (2021)

Ademais, é possível observar no período um acentuado incremento na proporção de indivíduos nas classes mais altas e médias, o que ocorreu em detrimento da classe mais pobre (E), a única classe que teve retração no seu tamanho, apesar de seguir compreendendo a maior parte da população brasileira. Tal processo pode ser observado na Tabela 2 abaixo.

Tabela 2: Percentual da população por classe de renda (Brasil)

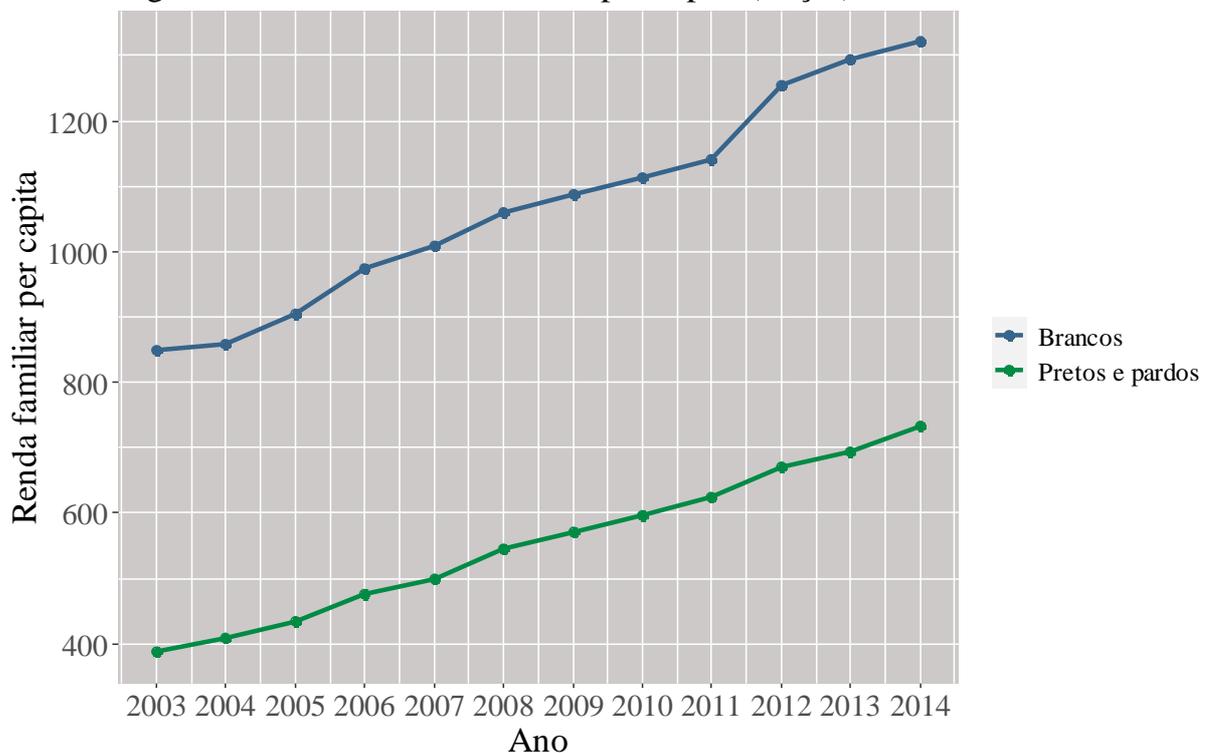
| Classe | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | Varição total |
|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|---------------|
| A | 0.11 | 0.14 | 0.15 | 0.17 | 0.20 | 0.20 | 0.21 | 0.23 | 0.25 | 0.26 | 0.31 | 0.27 | 137.80 |
| B | 0.17 | 0.13 | 0.15 | 0.19 | 0.18 | 0.20 | 0.21 | 0.23 | 0.24 | 0.28 | 0.33 | 0.25 | 47.76 |
| C | 5.08 | 5.14 | 5.48 | 6.02 | 6.23 | 6.61 | 6.96 | 7.13 | 7.33 | 8.15 | 8.65 | 8.53 | 67.83 |
| D | 5.25 | 5.29 | 5.48 | 5.92 | 6.41 | 7.21 | 7.37 | 7.62 | 7.90 | 9.17 | 9.38 | 10.20 | 94.32 |
| E | 89.39 | 89.30 | 88.73 | 87.71 | 86.98 | 85.79 | 85.25 | 84.80 | 84.28 | 82.14 | 81.33 | 80.75 | -9.66 |

Fonte: IBGE.

4.2.2. Subgrupos raciais

No caso dos subgrupos raciais, é possível observar uma grande heterogeneidade, com larga diferença entre a renda familiar média per capita da população branca, de um lado, e de pretos e pardos, de outro. Apesar disso, ambos os subgrupos experimentaram um incremento persistente de renda no período, o que pode ser visto na Figura 2; para pretos e pardos, o incremento foi de aproximadamente 89,5% (indo de R\$386,00 para R\$732,87), enquanto que, para brancos, foi de 55,8%, indo de R\$848,04 em 2003 para R\$1321,11 em 2014.

Figura 2: Média da renda ramiliar per capita (Raças)



Fonte: IBGE. Elaborado pelo autor (2021)

O mesmo pode ser observado tendo como variável de análise a proporção de indivíduos em cada classe de renda (Tabelas 3 e 4): apesar da grande heterogeneidade entre os subgrupos, tendo os brancos presença muito maior nas classes mais altas, e os pretos e pardos tendo participação consideravelmente mais alta na classe mais baixa, pode ser observado para ambos os grupos um aumento significativo da presença nas classes mais altas e médias, com queda na proporção de indivíduos na classe E.

Tabela 3: Percentual da população por classe de renda (brancos)

| Classe | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | Varição |
|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|---------|
| A | 0.21 | 0.24 | 0.28 | 0.33 | 0.37 | 0.38 | 0.39 | 0.42 | 0.45 | 0.50 | 0.60 | 0.61 | 184.71 |
| B | 0.31 | 0.23 | 0.28 | 0.35 | 0.35 | 0.36 | 0.39 | 0.41 | 0.44 | 0.53 | 0.64 | 0.49 | 61.47 |
| C | 8.54 | 8.53 | 9.17 | 9.91 | 10.40 | 10.86 | 11.31 | 11.46 | 11.64 | 13.12 | 14.06 | 14.34 | 67.93 |
| D | 7.97 | 8.05 | 8.31 | 8.97 | 9.60 | 10.45 | 10.69 | 10.88 | 11.09 | 12.81 | 12.91 | 13.98 | 75.38 |
| E | 82.97 | 82.94 | 81.96 | 80.45 | 79.29 | 77.96 | 77.23 | 76.83 | 76.38 | 73.04 | 71.80 | 70.58 | -14.93 |

Fonte: IBGE.

Tabela 4: Percentual da população por classe de renda (pretos e pardos)

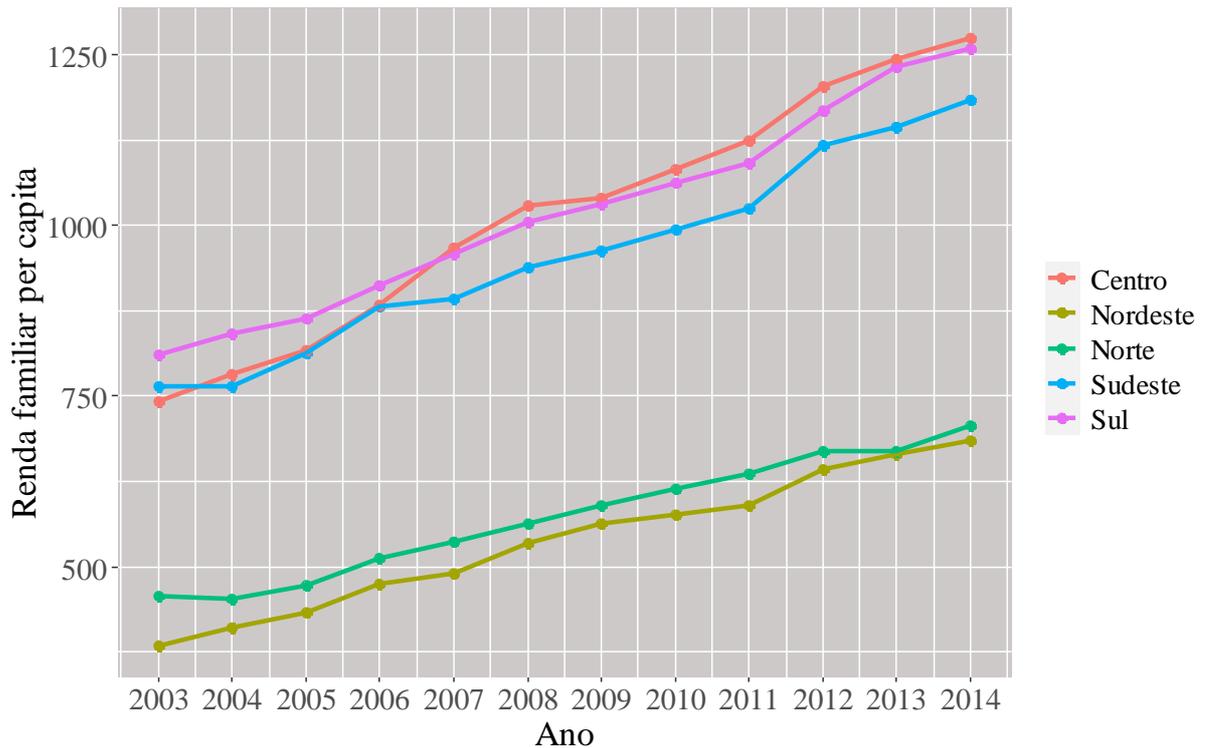
| Classe | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | Varição |
|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|---------|
| A | 0.02 | 0.04 | 0.04 | 0.04 | 0.05 | 0.05 | 0.07 | 0.07 | 0.08 | 0.08 | 0.09 | 0.09 | 350.29 |
| B | 0.04 | 0.04 | 0.04 | 0.04 | 0.04 | 0.08 | 0.07 | 0.08 | 0.09 | 0.09 | 0.10 | 0.11 | 183.70 |
| C | 1.85 | 2.02 | 2.26 | 2.67 | 2.73 | 3.06 | 3.39 | 3.57 | 3.79 | 4.32 | 4.57 | 4.78 | 159.11 |
| D | 2.69 | 2.78 | 3.03 | 3.33 | 3.74 | 4.52 | 4.67 | 4.95 | 5.28 | 6.40 | 6.74 | 7.43 | 176.39 |
| E | 95.41 | 95.13 | 94.63 | 93.92 | 93.44 | 92.29 | 91.81 | 91.33 | 90.77 | 89.12 | 88.51 | 87.59 | -8.20 |

Fonte: IBGE.

4.2.3. Macrorregiões

Assim como no caso dos subgrupos raciais, a presença de heterogeneidades inter-regionais fica bastante evidente ao olharmos para a evolução das médias dos rendimentos (Figura 3), com uma distinção bastante clara entre um grupo de estados mais ricos (Sul, Sudeste e Centro-Oeste) e outro de estados mais pobres (Norte e Nordeste), embora tenha havido crescimento da renda média para todas as regiões no período abordado.

Figura 3: Média da renda familiar per capita (macrorregiões)



Fonte: IBGE. Elaborado pelo autor (2021)

Um padrão similar pode ser observado nas proporções da população de cada região nas classes de renda: embora haja grande heterogeneidade entre as proporções de cada região nas classes de renda, com proporções muito superiores nas classes médias e superiores nas macrorregiões mais ricas, durante todo o período é possível observar, para todas elas, um grande crescimento nos estratos médios e superiores, em detrimento da classe mais baixa, como pode ser observado nas tabelas abaixo.

Tabela 5: Percentual da população por classe de renda (Nordeste)

| Classe | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | Varição |
|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|---------|
| A | 0.04 | 0.08 | 0.08 | 0.08 | 0.09 | 0.10 | 0.12 | 0.13 | 0.14 | 0.14 | 0.18 | 0.15 | 304.50 |
| B | 0.07 | 0.06 | 0.06 | 0.11 | 0.09 | 0.12 | 0.14 | 0.13 | 0.11 | 0.16 | 0.19 | 0.14 | 101.30 |
| C | 2.57 | 2.87 | 2.96 | 3.33 | 3.36 | 3.79 | 4.05 | 4.00 | 3.94 | 4.52 | 4.46 | 4.77 | 85.55 |
| D | 2.51 | 2.79 | 2.99 | 3.03 | 3.33 | 3.72 | 4.11 | 4.20 | 4.30 | 4.90 | 5.15 | 5.52 | 119.52 |
| E | 94.81 | 94.21 | 93.92 | 93.45 | 93.13 | 92.26 | 91.59 | 91.55 | 91.51 | 90.27 | 90.02 | 89.43 | -5.68 |

Fonte: IBGE.

Tabela 6: Percentual da população por classe de renda (Norte)

| Classe | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | Variação |
|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------|
| A | 0.01 | 0.05 | 0.06 | 0.06 | 0.06 | 0.07 | 0.09 | 0.10 | 0.11 | 0.10 | 0.08 | 0.10 | 762.58 |
| B | 0.08 | 0.06 | 0.04 | 0.06 | 0.05 | 0.08 | 0.09 | 0.11 | 0.13 | 0.14 | 0.13 | 0.14 | 83.43 |
| C | 3.04 | 2.81 | 3.00 | 3.46 | 3.65 | 3.57 | 4.05 | 4.30 | 4.57 | 4.87 | 4.96 | 5.40 | 77.57 |
| D | 3.78 | 3.51 | 3.72 | 3.86 | 4.55 | 5.10 | 5.06 | 5.23 | 5.42 | 6.30 | 6.30 | 6.80 | 79.99 |
| E | 93.09 | 93.58 | 93.18 | 92.55 | 91.69 | 91.18 | 90.71 | 90.25 | 89.78 | 88.60 | 88.53 | 87.55 | -5.95 |

Fonte: IBGE.

Tabela 7: Percentual da população por classe de renda (Centro-Oeste)

| Classe | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | Variação |
|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------|
| A | 0.21 | 0.26 | 0.25 | 0.32 | 0.49 | 0.47 | 0.45 | 0.47 | 0.49 | 0.49 | 0.62 | 0.69 | 225.31 |
| B | 0.27 | 0.24 | 0.22 | 0.35 | 0.45 | 0.46 | 0.46 | 0.49 | 0.52 | 0.65 | 0.61 | 0.48 | 77.98 |
| C | 7.13 | 7.49 | 8.09 | 8.42 | 9.32 | 9.75 | 9.97 | 10.26 | 10.60 | 11.88 | 12.93 | 12.22 | 71.50 |
| D | 6.43 | 6.03 | 6.29 | 6.81 | 7.55 | 8.50 | 8.16 | 8.70 | 9.33 | 10.94 | 11.41 | 12.90 | 100.68 |
| E | 85.96 | 85.98 | 85.14 | 84.09 | 82.19 | 80.82 | 80.96 | 80.08 | 79.06 | 76.03 | 74.42 | 73.71 | -14.26 |

Fonte: IBGE.

Tabela 8: Percentual da população por classe de renda (Sudeste)

| Classe | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | Variação |
|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------|
| A | 0.17 | 0.17 | 0.22 | 0.23 | 0.24 | 0.23 | 0.28 | 0.30 | 0.32 | 0.34 | 0.40 | 0.43 | 152.72 |
| B | 0.22 | 0.17 | 0.23 | 0.25 | 0.23 | 0.24 | 0.24 | 0.27 | 0.30 | 0.32 | 0.44 | 0.38 | 70.15 |
| C | 6.86 | 6.71 | 7.22 | 8.17 | 8.18 | 8.56 | 8.87 | 9.08 | 9.33 | 10.64 | 11.35 | 11.36 | 65.62 |
| D | 7.01 | 7.18 | 7.27 | 8.07 | 8.57 | 9.78 | 9.82 | 9.98 | 10.16 | 11.90 | 11.84 | 12.87 | 83.53 |
| E | 85.74 | 85.77 | 85.05 | 83.28 | 82.78 | 81.20 | 80.79 | 80.37 | 79.89 | 76.80 | 75.98 | 74.96 | -12.57 |

Fonte: IBGE

Tabela 9: Percentual da população por classe de renda (Sul)

| Classe | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | Variação |
|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------|
| A | 0.17 | 0.20 | 0.17 | 0.25 | 0.26 | 0.27 | 0.24 | 0.26 | 0.27 | 0.35 | 0.39 | 0.35 | 98.44 |
| B | 0.26 | 0.17 | 0.25 | 0.21 | 0.19 | 0.21 | 0.25 | 0.28 | 0.31 | 0.34 | 0.43 | 0.35 | 31.72 |
| C | 7.05 | 7.31 | 7.97 | 8.19 | 8.90 | 9.34 | 10.03 | 10.27 | 10.53 | 11.18 | 12.61 | 13.17 | 86.73 |
| D | 7.88 | 8.05 | 8.46 | 9.16 | 9.89 | 10.76 | 11.22 | 11.60 | 12.01 | 13.91 | 14.67 | 15.70 | 99.38 |
| E | 84.63 | 84.26 | 83.15 | 82.18 | 80.76 | 79.42 | 78.27 | 77.60 | 76.87 | 74.23 | 71.90 | 70.43 | -16.78 |

Fonte: IBGE.

4.3 Matrizes de transição

As estimações das matrizes de transição foram feitas da seguinte maneira: primeiramente, por meio do método de mínimos quadrados (exposto em detalhes no apêndice A), estimamos as matrizes de um período (tridiagonais) *a priori* P_0 para todos os subgrupos.¹²

Uma função de probabilidades bidimensional para cada período, denominada D_{0t} , pode ser obtida pela multiplicação (AEBI et. al., 1999) $D_{0t} = \text{diag}(n_t) P_0$, em que o operador diag transforma o vetor n_t em uma matriz quadrada diagonal. A função de distribuição bidimensional verdadeira (denominemo-la D_t), entretanto, deve atender às seguintes condições (AEBI et. al., 1999), equivalentes às "restrições de continuidade" (5) e (6):

$$D_t \tau = n_t$$

$$D'_t \tau = n_{t+1}$$

Onde τ é um vetor de dimensão 5×1 com todas as células iguais a 1, e a matriz D'_t é a transposta da matriz D_t . É fácil notar, ademais, que as condições acima são equivalentes à igualdade entre as distribuições marginais de D e as proporções n_t e n_{t+1} . Assim, podemos ajustar, com o uso do IPFP, a distribuição D_{0t} usando suas distribuições marginais, obtendo a distribuição ajustada D_t ¹³. A matriz de transição ajustada para cada um dos períodos, P_t , é obtida pela equação:

$$P_t = \text{diag}(n_t)^{-1} D_t$$

Por fim, a matriz de transição para todo o período é obtida pela multiplicação das matrizes P_t de cada período.

4.3.1 Brasil

A interpretação de qualquer matriz de transição, como vimos, deve ser feita da seguinte maneira: a célula p_{ij} da matriz nos indica a probabilidade de que um indivíduo se encontre na classe j em um período (no nosso caso, o ano de 2014), dado que ele estava na classe i no período inicial (2003). Desse modo, a Figura 4 abaixo nos provê de algumas informações valiosas acerca da mobilidade de renda brasileira no período.

¹² As matrizes *a priori* estimadas se encontram no apêndice B.

¹³ Neste caso, o IPFP foi implementado por meio do pacote "mipfp" (BARTHÉLEMY e SUESSE, 2018) do software R (R CORE TEAM, 2020).

Tabela 10: Matriz de transição estimada (Brasil)

| | A | B | C | D | E |
|---|---------|---------|---------|---------|---------|
| A | 0.40307 | 0.38732 | 0.20597 | 0.00360 | 0.00004 |
| B | 0.36176 | 0.34746 | 0.27918 | 0.01135 | 0.00025 |
| C | 0.02349 | 0.02106 | 0.87058 | 0.08190 | 0.00297 |
| D | 0.00612 | 0.00519 | 0.42475 | 0.47684 | 0.08710 |
| E | 0.00015 | 0.00012 | 0.02023 | 0.08139 | 0.89811 |

Fonte: IBGE. Elaboração do autor (2021)

Primeiramente, a alta probabilidade (89,8%) de que os indivíduos da classe E permaneçam nela nos mostra que, apesar de haver ocorrido uma retração dessa classe em favor das classes médias e altas (como visto na tabela 2), esse processo se dá muito lentamente, sendo os onze anos compreendidos pelo estudo longe de serem suficientes para sua concretização.

Uma outra característica marcante é a alta probabilidade de permanência na classe C (87%); assim, embora haja uma probabilidade relativamente alta de que pessoas que se encontravam na classe D migrem para a classe C (42,4%), é bem pouco provável que indivíduos da classe C progridam em direção às classes A e B (cerca de 2% para ambas). Esse dado sugere uma espécie de truncamento no processo de ascensão social brasileiro: embora relativamente frequente a ascensão das classes mais pobres em direção à classe média, a progressão em direção às classes mais altas se encontra praticamente vedada.

Por fim, as probabilidades baixíssimas de que indivíduos migrem de um extremo da distribuição de renda para o outro indicam, corroborando FIGUEIREDO (2010)¹⁴, que o fato de um indivíduo pertencer à classe mais pobre no início do período é determinante para que ele não alcance o topo da distribuição ao final (e a recíproca também é observada).

4.3.2 Subgrupos raciais

Pudemos observar, na análise das estatísticas descritivas, uma série de heterogeneidades entre pretos e pardos, de um lado, e brancos, de outro. Tal tendência é corroborada pela análise das matrizes de transição (Figuras 5 e 6): enquanto pretos e pardos

¹⁴ Figueiredo (2010), entretanto, utilizou como critério de estratificação de classe social quintis de renda, não a divisão por classes (A a E) aqui feita; ademais, o período compreendido pelo seu estudo foi de 1995 a 2005. A persistência dessa característica apesar das distinções metodológicas e temporais apenas corrobora sua validade como característica sistêmica da economia brasileira.

têm probabilidade de apenas 14,4% de permanecerem na classe A, esse número é de 47,9% para brancos. Quanto à probabilidade de permanecer na classe mais baixa, a situação se inverte, sendo de 84% para brancos e 91,5% para pretos e pardos. Tais dados nos indicam, portanto, que a população negra, além de ser, em média, mais pobre (como visto na seção 4.2.2), tem também menor probabilidade de sair da pobreza e de manter-se nas classes mais altas.

Tabela 11: Matriz de transição estimada (pretos e pardos)

| | A | B | C | D | E |
|---|---------|---------|---------|---------|---------|
| A | 0.14486 | 0.14562 | 0.68347 | 0.02546 | 0.00059 |
| B | 0.13321 | 0.13277 | 0.70505 | 0.02823 | 0.00073 |
| C | 0.02314 | 0.02691 | 0.89489 | 0.05291 | 0.00216 |
| D | 0.00862 | 0.01080 | 0.50335 | 0.40406 | 0.07316 |
| E | 0.00017 | 0.00025 | 0.01818 | 0.06548 | 0.91592 |

Fonte: IBGE. Elaboração do autor (2021)

Para pretos e pardos, ademais, a alta probabilidade de permanência na classe C, observada no Brasil como um todo, se confirma. Somado a isso, ainda, está também as altas probabilidades de migração das classes A, B e D em direção à classe C (68,3%, 70,5% e 50,3%, respectivamente); esse fato, entretanto, não é observado na população branca, em que indivíduos das classes mais altas (A e B) têm menor probabilidade de regredirem à classe C. Um dado que de certo modo contradiz essa tendência de perpetuação de disparidade social, entretanto, pode ser observado na probabilidade de indivíduos que estão na classe C regredirem à classe D, sendo esta bastante inferior para negros (aproximadamente 5, 29%) do que para brancos (42,8%).

Tabela 12: Matriz de transição estimada (brancos)

| | A | B | C | D | E |
|---|---------|---------|---------|---------|---------|
| A | 0.47910 | 0.42246 | 0.09536 | 0.00297 | 0.00011 |
| B | 0.41142 | 0.36254 | 0.15506 | 0.06584 | 0.00514 |
| C | 0.01902 | 0.01511 | 0.49913 | 0.42868 | 0.03806 |
| D | 0.01731 | 0.01366 | 0.48372 | 0.41640 | 0.06890 |
| E | 0.00093 | 0.00067 | 0.07419 | 0.08407 | 0.84013 |

Fonte: IBGE. Elaboração do autor (2021)

4.3.3 Macrorregiões

Para as macrorregiões, também é possível corroborar algumas das heterogeneidades observadas nas estatísticas descritivas: o grupo das regiões mais pobres que identificamos (Norte e Nordeste, Figuras 7 e 8 abaixo) é também o grupo em que se vê a maior probabilidade de um indivíduo permanecer na classe mais baixa, sendo de 94% para o Nordeste e 93% para o Norte, enquanto que, no grupo dos estados mais ricos, esse percentual não passa dos 86,4%.

Tabela 13: Matriz de transição estimada (Norte)

| | A | B | C | D | E |
|---|---------|---------|---------|---------|---------|
| A | 0.30842 | 0.29293 | 0.19982 | 0.19102 | 0.00781 |
| B | 0.31792 | 0.30183 | 0.19170 | 0.18159 | 0.00696 |
| C | 0.01037 | 0.01388 | 0.45202 | 0.48438 | 0.03936 |
| D | 0.00704 | 0.01018 | 0.39113 | 0.42086 | 0.17080 |
| E | 0.00017 | 0.00038 | 0.02722 | 0.03993 | 0.93229 |

Fonte: IBGE. Elaboração do autor (2021)

Tabela 14: Matriz de transição estimada (Nordeste)

| | A | B | C | D | E |
|---|---------|---------|---------|---------|---------|
| A | 0.38804 | 0.38121 | 0.22643 | 0.00418 | 0.00014 |
| B | 0.38841 | 0.38157 | 0.22579 | 0.00410 | 0.00014 |
| C | 0.02855 | 0.02686 | 0.84748 | 0.09191 | 0.00519 |
| D | 0.00874 | 0.00790 | 0.50350 | 0.39925 | 0.08061 |
| E | 0.00012 | 0.00010 | 0.01371 | 0.04514 | 0.94093 |

Fonte: IBGE. Elaboração do autor (2021)

Quanto à permanência na classe mais alta, embora ela seja mais alta em duas das regiões mais ricas - Sudeste e Centro-Oeste (Figuras 9 e 10, abaixo) com aproximadamente 49% cada, em contraste com 38,8% no Nordeste e 30,8% no Norte - ela também alcança seu menor patamar no Sul (Figura 11) uma região de renda per capita mais alta, não sendo possível estabelecer uma relação unívoca entre riqueza relativa e permanência nas classes superiores.

Tabela 15: Matriz de transição estimada (Sudeste)

| | A | B | C | D | E |
|---|---------|---------|---------|---------|---------|
| A | 0.49214 | 0.44319 | 0.04391 | 0.01979 | 0.00098 |
| B | 0.38012 | 0.34214 | 0.14634 | 0.12203 | 0.00937 |
| C | 0.01727 | 0.01486 | 0.47573 | 0.45109 | 0.04105 |
| D | 0.01305 | 0.01108 | 0.45919 | 0.43678 | 0.07991 |
| E | 0.00065 | 0.00053 | 0.05640 | 0.07793 | 0.86449 |

Fonte: IBGE. Elaboração do autor (2021)

Tabela 16: Matriz de transição estimada (Centro-Oeste)

| | A | B | C | D | E |
|---|---------|---------|---------|---------|---------|
| A | 0.49524 | 0.32320 | 0.16992 | 0.01152 | 0.00011 |
| B | 0.46625 | 0.30442 | 0.21252 | 0.01659 | 0.00022 |
| C | 0.04760 | 0.03334 | 0.82433 | 0.09244 | 0.00229 |
| D | 0.01394 | 0.01041 | 0.50167 | 0.37952 | 0.09446 |
| E | 0.00034 | 0.00030 | 0.03524 | 0.11394 | 0.85018 |

Fonte: IBGE. Elaboração do autor (2021)

Tabela 17: Matriz de transição estimada (Sul)

| | A | B | C | D | E |
|---|---------|---------|---------|---------|---------|
| A | 0.25841 | 0.23611 | 0.48389 | 0.02111 | 0.00048 |
| B | 0.18121 | 0.16629 | 0.61372 | 0.03756 | 0.00122 |
| C | 0.02458 | 0.02463 | 0.86986 | 0.07730 | 0.00364 |
| D | 0.00707 | 0.00757 | 0.44766 | 0.46110 | 0.07661 |
| E | 0.00029 | 0.00035 | 0.03854 | 0.13604 | 0.82478 |

Fonte: IBGE. Elaboração do autor (2021)

Quanto ao comportamento da classe C, ele pode ser dividido em dois grupos: aquelas regiões em que (acompanhando a tendência média do país) a probabilidade de permanência é superior a 80% (Sul, Nordeste e Centro-Oeste) e aquelas em que é mais baixa (entre 45 e 50%), tendo probabilidade mais alta de regressão rumo à classe D. Entretanto, dois aspectos comuns a todas as regiões (como foram também para os demais subgrupos) são: (i) a baixíssima probabilidade de que pessoas nas classes C, D e E acessem as classes A e B, o que corrobora a noção apontada de que haja um "truncamento" na ascensão social; (ii) a também baixíssima probabilidade de que indivíduos migrem de uma extremidade a outra da distribuição de renda.

4.4 Índices de mobilidade

4.4.1 Brasil

Os índices de mobilidade encontrados para o Brasil do período são dados pela tabela abaixo; esses números foram todos inferiores àqueles achados para o país nos períodos entre 1995 e 2005 (FIGUEIREDO, 2010) e entre 1987 e 2005 (FIGUEIREDO e ZYGELLMAN, 2012), e colocaria o Brasil como país com baixíssima mobilidade. Entretanto, as diferenças metodológicas entre este trabalho (utilizando critérios pecuniários de classes sociais) e os demais (que utilizam como critério de estratificação quintis de renda) tornam essas comparações imperfeitas, sendo mais profícua a comparação interna entre a mobilidade de todos os subgrupos.

Tabela 18: Índices de mobilidade (Brasil)

| | M_P | M_L |
|---------------|-----------|------------|
| Brasil | 0.5009885 | 0.10134514 |

Fonte: IBGE. Elaboração do autor (2021)

4.4.2 Subgrupos raciais

Na tabela 10 abaixo, podemos observar os índices de mobilidade para ambos os subgrupos raciais estudados. Parece, entretanto, haver uma contradição: embora pretos e pardos pareçam ter maior mobilidade social pelo índice de Prais (M_P), o índice M_L indica maior mobilidade entre os brancos. Isso pode ser explicado pelo método utilizado para obtenção do índice de Prais: como é utilizado para sua obtenção o traço da matriz, esse índice dá especial importância para os elementos da diagonal principal da matriz; a maior mobilidade entre pretos e pardos, portanto, pode ser explicada pela menor probabilidade de permanência destes nas classes superiores (como visto na seção 4.3.2), o que indica não uma característica positiva, mas negativa, da evolução de renda desse subgrupo no período.

Tabela 19: Índices de mobilidade (raças)

| | M_P | M_L |
|------------------------|-----------|------------|
| Branco | 0.6006757 | 0.15996161 |
| Pretos e pardos | 0.6268747 | 0.08430817 |

Fonte: IBGE. Elaboração do autor (2021)

4.4.3 Macrorregiões

Quanto às macrorregiões, os índices mostram maior consistência: embora haja distinções entre os dois, é possível distinguir entre um grupo de mobilidade mais alta (Norte, Sudeste e Sul) e outro menos móvel (Centro e Nordeste). Não é possível, entretanto, estabelecer uma relação unívoca entre a renda média per capita e o índice de mobilidade, havendo tanto regiões mais ricas com nível de mobilidade mais alto (Sudeste e Sul) quanto regiões mais pobres com mobilidade baixa (Nordeste), como pode ser observado na Tabela 20 abaixo.

Tabela 20: Índices de mobilidade (macrorregiões)

| | M_p | M_L |
|-----------------|-----------|------------|
| Centro | 0.5365759 | 0.13954399 |
| Nordeste | 0.5106804 | 0.06607747 |
| Norte | 0.6461479 | 0.16591346 |
| Sudeste | 0.5971768 | 0.15548315 |
| Sul | 0.6048914 | 0.16805177 |

Fonte: IBGE. Elaboração do autor (2021)

Outrossim, os índices de mobilidade não podem nos dar uma resposta quanto à direção da mobilidade: é possível obter um índice alto tanto quando há baixa probabilidade de permanência entre os estratos superiores de renda (o que indica uma piora nas condições de vida) como quando há uma melhora nas condições de vida, com alta probabilidade de ascensão das classes baixas em direção às superiores. Logo, os índices de mobilidade devem ser interpretados sempre em conjunção com a análise mais detalhada das matrizes de transição.

4.5 Distribuição-limite

Vimos pela equação (1) que uma matriz de transição para t períodos pode ser obtida pela elevação da matriz de transição P à t -ésima potência. Em muitos casos, a matriz possui um comportamento de longo prazo, convergindo para uma distribuição de equilíbrio independente de sua distribuição inicial (DOBROW, 2016). Nesses casos, o comportamento de longo prazo da matriz é dado por:

$$\lim_{t \rightarrow \infty} P^t = \Lambda$$

Onde Λ é uma matriz com todas as linhas iguais a λ , sendo λ a distribuição-limite da matriz P . Cada célula λ_i da distribuição λ indica a probabilidade de longo prazo de que o processo se encontre no estado i (DOBROW, 2016).

Ademais, se a matriz de transição P for regular¹⁵ (o que é o caso de todas as matrizes aqui estimadas), a distribuição-limite existe e é a única distribuição estacionária da cadeia. Assim, a distribuição-limite da matriz nos indica qual seria a distribuição de equilíbrio que teria lugar caso o processo regido pela matriz P se prolongasse indefinidamente. Obviamente, essa distribuição é meramente hipotética e potencial, e só apareceria de fato caso a estrutura social que rege o processo de mobilidade social se mantivesse constante indefinidamente, uma hipótese que não se verifica. Entretanto, a importância de olharmos para a distribuição-limite está em ver se há em moção um processo de convergência (mesmo que no longuíssimo prazo) entre os subgrupos populacionais.

4.5.1 Brasil

A distribuição-limite para o país como um todo nos mostra a continuação do processo observado anteriormente de uma redução da classe E, com um aumento das demais. Entretanto, a grande proporção da população que, em equilíbrio, estaria na classe C parece expor um limite à ascensão social no país: embora boa parte da população consiga alcançar os estratos intermediários, apenas uma pequena parcela consegue acessar as camadas superiores, o que pode ser visto na Tabela 21 abaixo.

Tabela 21: Distribuição-limite das classes de renda (Brasil)

| A | B | C | D | E |
|-------|-------|--------|--------|--------|
| 6,16% | 5,81% | 63,79% | 12,05% | 12,17% |

Fonte: IBGE. Elaboração do autor (2021)

¹⁵ Uma matriz de transição é dita regular se, para alguma potência qualquer dela, todas as células são positivas.

4.5.2 Subgrupos raciais

Para ambos os subgrupos raciais, a tendência vista de diminuição da classe E em favor das demais parece se confirmar, bem como o maior fortalecimento relativo das classes intermediárias, especialmente para a população negra (como visto na Tabela abaixo).

Tabela 22: Distribuição-limite das classes de renda (pretos e pardos)

| A | B | C | D | E |
|-------|-------|--------|-------|-------|
| 2,63% | 2,94% | 77,24% | 8,11% | 9,08% |

Fonte: IBGE. Elaboração do autor (2021)

Se, por um lado, a proporção de brancos (Tabela 23) entre as classes A e B segue sendo bastante superior àquela de negros, a situação se inverte para as classes inferiores, e a proporção de negros nas classes mais baixas torna-se menor, no estado estacionário, do que aquela de brancos, o que contraria as nossas expectativas.

Tabela 23: Distribuição-limite das classes de renda (brancos)

| A | B | C | D | E |
|-------|-------|--------|--------|--------|
| 7,33% | 6,35% | 35,20% | 29,71% | 21,40% |

Fonte: IBGE. Elaboração do autor (2021)

4.5.3 Macrorregiões

Para todas as macrorregiões também se confirma a diminuição na proporção da população entre a classe mais pobre, em direção principalmente à classe C. Entretanto, a região Norte ainda aparece, mesmo no estado estacionário, como tendo uma maioria da população na classe E, além de ter uma proporção bem menor do que as demais macrorregiões nas duas classes superiores. Logo, apesar da região Norte ter tido alta mobilidade na análise dos índices, essa mobilidade não se deu em direção às classes mais altas.

Tabela 24: Distribuição-limite das classes de renda (Nordeste)

| A | B | C | D | E |
|-------|-------|--------|--------|--------|
| 7,28% | 7,08% | 56,57% | 10,17% | 18,89% |

Fonte: IBGE. Elaboração do autor (2021)

Tabela 25: Distribuição-limite das classes de renda (Norte)

| A | B | C | D | E |
|-------|-------|--------|--------|--------|
| 0,98% | 1,09% | 17,82% | 19,73% | 60,37% |

Fonte: IBGE. Elaboração do autor (2021)

Tabela 26: Distribuição-limite das classes de renda (Sudeste)

| A | B | C | D | E |
|-------|-------|--------|--------|--------|
| 5,66% | 5,05% | 31,34% | 30,23% | 27,71% |

Fonte: IBGE. Elaboração do autor (2021)

Tabela 27: Distribuição-limite das classes de renda (Sul)

| A | B | C | D | E |
|-------|-------|--------|--------|-------|
| 3,35% | 3,24% | 73,58% | 12,70% | 7,11% |

Fonte: IBGE. Elaboração do autor (2021)

Tabela 28: Distribuição-limite das classes de renda (Centro-Oeste)

| A | B | C | D | E |
|--------|-------|--------|--------|-------|
| 14,74% | 9,77% | 57,51% | 10,48% | 7,51% |

Fonte: IBGE. Elaboração do autor (2021)

Outra cautela necessária ao analisar as distribuições-limite é a seguinte: como a distribuição-limite é única, independente da condição inicial, sua análise pode ser demasiada a-histórica, apagando as diversas desigualdades que se perpetuaram na sociedade brasileira. Além disso, como já foi apontado, ela é apenas hipotética, tendo lugar somente no caso em que o processo social descrito pela matriz se perpetue indefinidamente.

5 CONCLUSÃO

Este trabalho buscou estudar a mobilidade social do Brasil como um todo e de dois de seus subgrupos populacionais (raciais e regionais) entre os anos de 2003 e 2014. Para isso, fez uso de cadeias de Markov, da abordagem axiomática dos índices de mobilidade e do método de entropias relativas, bem como uma série de estatísticas descritivas.

Analisando a mobilidade social do Brasil (bem como de subgrupos raciais e regionais) no período entre 2003 e 2014, é inescapável a conclusão de que houve uma melhora substantiva na renda média per capita e na proporção em cada classe de renda de todas as parcelas estudadas da população brasileira. Entretanto, uma análise mais minuciosa das matrizes de transição desvela uma série de características negativas da economia brasileira no período: a alta probabilidade de permanência nas classes inferiores sugere que o processo de ascensão social observado se dá de maneira bastante lenta; a baixíssima probabilidade de que indivíduos nas três classes inferiores atinjam as classes A e B indica que há o que chamamos de "truncamento" do processo de melhoria na distribuição de renda. Nesse caso, embora o acesso das classes mais pobres em direção à classe média seja relativamente comum, a evolução desta em relação às classes mais ricas é praticamente vedada; esta análise é corroborada, ainda, pela análise das distribuições-limite.

Em relação aos subgrupos estudados, foi possível detectar uma série de heterogeneidades entre eles, tanto nas estatísticas descritivas quanto na estimação das matrizes de transição. A análise das distribuições-limite, ademais, nos permitiu observar uma relativa tendência à convergência entre os subgrupos raciais e entre boa parte das macrorregiões, com exceção da região Norte; entretanto, a análise das matrizes de transição nos demonstrou que esse movimento de convergência se dá de maneira demasiadamente lenta.

A partir da análise deste trabalho, alguns desenvolvimentos futuros podem ser feitos. Primeiramente, a extensão desta metodologia para os anos de 2015 a 2020 pode nos mostrar os efeitos da crise econômica sobre a mobilidade social da economia brasileira. Outra extensão possível com o uso dos dados da PNAD é a divisão entre rendimentos do trabalho e rendimentos assistenciais, o que nos permitiria analisar qual a parcela da mobilidade pode ser atribuída à assistência social e aquela que pode ser atribuída à variação salarial.

REFERÊNCIAS

- AEBI, R., NEUSSER, K., STEINER, P. Improving models of income dynamics using cross-section information. **Swiss Journal of Economics and Statistics**, v. 144, p.117-151, 2008.
- BARTHÉLEMY, J., SUESSE T. mipfp: An R Package for Multidimensional Array Fitting and Simulating Multivariate Bernoulli Distributions. **Journal of Statistical Software**, v. 86 Code Snippets (2),p. 1-20, 2018.
- CORAK, M. **Inequality from generation to generation: the United States in comparison**. IZA Discussion Paper, n. 9929, 2016.
- COVER, T., THOMAS, J. **Elements of information theory**. New Jersey, NJ: John Wiley and Sons, 1991.
- DEMING, W.E., STEPHAN, F.F. On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. **Annals of Mathematical Statistics**, 11, p. 427-444, 1940.
- DOBROW, R. **Introduction to stochastic processes with R**. New Jersey, NJ: John Wiley and Sons, 2016.
- FIELDS, G. S.; OK, E. A. The measurement of income mobility: An introduction to the literature. In: SILBER, J. (Ed.). **Handbook on income inequality measurement** (p. 557-596). Norwell, MA: Kluwer Academic Publishers, 1999.
- FIGUEIREDO, E. A. Mobilidade intrageracional de renda no Brasil. **Nova econ.**, Belo Horizonte, v.20, n.3, p.427-455, set-nov. 2010.
- FGV SOCIAL. Qual a faixa de renda familiar das classes?. Disponível em: <https://cps.fgv.br/qual-faixa-de-renda-familiar-das-classes>. Acesso em: 27 de novembro de 2020.
- GOLAN, A. **Information and entropy econometrics: a review and synthesis**. Hanover, MA: Publishers Inc., 2006.
- IBGE. **Pesquisa Nacional de Amostra por Domicílios (PNAD)**. Disponível em: <https://www.ibge.gov.br/estatisticas/sociais/trabalho/19897-sintese-de-indicadores-pnad2.html?=&t=microdados>. Acesso em: 2020.
- JONES, M.T. Estimating Markov transition matrices using proportions data: an application to credit risk. **IMF Discussion Paper**, n. 219, 2005.
- LOVELACE, R. et. al. Evaluating the Performance of Iterative Proportional Fitting for Spatial Microsimulation: New Tests for an Established Technique. **Journal of Artificial Societies and Social Simulation**, v. 18, n.2, p.1-15, 2015.
- NERI, M. C. Brazil's middle classes. **Ensaio Econômicos**, n.759, 2014.
- R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2020.

SHORROCKS, A.F. Income mobility and the Markov assumption. **The Economic Journal**, v.86, n. 343, p. 566-578, set. 1976.

SHORROCKS, A.F. The measurement of mobility. **Econometrica**, v.46, n.5, p.1013-1024, set. 1978.

TIMOFEEVA, G., TIMOFEEV, N. State estimation problem for Markov chain model. In: PhysCon, 5., 2011, Léon, Espanha. **Anais eletrônicos**, Léon, 2011.

APÊNDICE A- ESTIMAÇÃO POR MÍNIMOS QUADRADOS

Como vimos, a evolução de um processo de Markov homogêneo pode ser resumida pela equação (4). Ademais, admitindo a existência de desvios, conquanto estejamos trabalhando com uma amostra da população, temos:

$$(n_t)_j = \sum_{i \in S} p_{ij} (n_{t-1})_i + (u_t)_j$$

Em termos matriciais, essa relação pode ser escrita (JONES, 2005) como:

$$y = X p + u$$

Onde $y = [n_A, n_B, \dots, n_E]'$

$$= [(n_1)_A, (n_2)_A, \dots, (n_t)_A \quad (n_1)_B, (n_2)_B, \dots, (n_t)_B \quad \dots \quad (n_1)_E, (n_2)_E, \dots, (n_t)_E]'$$

$$X_j = \begin{bmatrix} (n_1)_A & \dots & (n_1)_E \\ \vdots & \ddots & \vdots \\ (n_{t-1})_A & \dots & (n_{t-1})_E \end{bmatrix}$$

De modo que:

$$X = \begin{bmatrix} X_A & 0 & \dots & 0 \\ 0 & X_B & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & X_E \end{bmatrix}$$

$$p = [p_A \dots p_E]' = [p_{AA}, p_{BA}, \dots, p_{EA} \quad p_{AB}, p_{BB}, \dots, p_{EB} \quad \dots \quad p_{AE}, p_{BE}, \dots, p_{EE}]'$$

$$u = [u_A \quad u_B \quad \dots \quad u_E]' = [(u_1)_A, (u_2)_A, \dots, (u_t)_A \quad \dots \quad (u_1)_E, (u_2)_E, \dots, (u_t)_E]'$$

A solução é dada (TIMOFEEVA e TIMOFEEV (2011)) pela minimização da soma dos quadrados dos resíduos, condicionadas pelas restrições das matrizes de Markov (soma das linhas iguais à unidade e não-negatividade das células); no nosso caso, uma restrição adicional é dada pela tridiagonalidade da matriz. Formalmente:

$$\min. (y - Xp)'(y - Xp)$$

$$s. a. \sum_{j \in S} p_{ij} = 1$$

$$s. a. p_{ij} \geq 0, \quad \forall i, j \in S$$

$$s. a. p_{ij} = 0, \quad \forall i, j \in S; |i - j| > 2$$

APÊNDICE B- MATRIZES DE TRANSIÇÃO A PRIORI ESTIMADAS

Tabela 29: Matriz de transição a priori estimada (Brasil)

| | A | B | C | D | E |
|---|---------|---------|---------|---------|---------|
| A | 0.38569 | 0.61430 | 0.00000 | 0.00000 | 0.00000 |
| B | 0.60858 | 0.39141 | 0.00000 | 0.00000 | 0.00000 |
| C | 0.00000 | 0.00000 | 1.00000 | 0.00000 | 0.00000 |
| D | 0.00000 | 0.00000 | 0.04038 | 0.95961 | 0.00000 |
| E | 0.00000 | 0.00000 | 0.00000 | 0.00876 | 0.99123 |

Fonte: IBGE. Elaboração do autor (2021)

Tabela 30: Matriz de transição a priori estimada (brancos)

| | A | B | C | D | E |
|---|---------|---------|---------|---------|---------|
| A | 0.37173 | 0.62826 | 0.00000 | 0.00000 | 0.00000 |
| B | 0.64887 | 0.35112 | 0.00000 | 0.00000 | 0.00000 |
| C | 0.00000 | 0.00084 | 0.62667 | 0.37248 | 0.00000 |
| D | 0.00000 | 0.00000 | 0.44672 | 0.55327 | 0.00000 |
| E | 0.00000 | 0.00000 | 0.00000 | 0.01377 | 0.98622 |

Fonte: IBGE. Elaboração do autor (2021)

Tabela 31: Matriz de transição a priori estimada (pretos e pardos)

| | A | B | C | D | E |
|---|---------|---------|---------|---------|---------|
| A | 0.17093 | 0.82906 | 0.00000 | 0.00000 | 0.00000 |
| B | 0.51206 | 0.01774 | 0.47019 | 0.00000 | 0.00000 |
| C | 0.00000 | 0.00000 | 1.00000 | 0.00000 | 0.00000 |
| D | 0.00000 | 0.00000 | 0.04515 | 0.95484 | 0.00000 |
| E | 0.00000 | 0.00000 | 0.00000 | 0.00720 | 0.99279 |

Fonte: IBGE. Elaboração do autor (2021)

Tabela 32: Matriz de transição a priori estimada (Centro - Oeste)

| | A | B | C | D | E |
|---|---------|---------|---------|---------|---------|
| A | 0.60763 | 0.39236 | 0.00000 | 0.00000 | 0.00000 |
| B | 0.39569 | 0.53095 | 0.07335 | 0.00000 | 0.00000 |
| C | 0.00000 | 0.00221 | 0.99778 | 0.00000 | 0.00000 |
| D | 0.00000 | 0.00000 | 0.04657 | 0.95342 | 0.00000 |
| E | 0.00000 | 0.00000 | 0.00000 | 0.01262 | 0.98737 |

Fonte: IBGE. Elaboração do autor.

Tabela 33: Matriz de transição a priori estimada (Nordeste)

| | A | B | C | D | E |
|---|---------|---------|---------|---------|---------|
| A | 0.44006 | 0.55993 | 0.00000 | 0.00000 | 0.00000 |
| B | 0.57442 | 0.42557 | 0.00000 | 0.00000 | 0.00000 |
| C | 0.00000 | 0.00180 | 0.99819 | 0.00000 | 0.00000 |
| D | 0.00000 | 0.00000 | 0.04993 | 0.95006 | 0.00000 |
| E | 0.00000 | 0.00000 | 0.00000 | 0.00512 | 0.99487 |

Fonte: IBGE. Elaboração do autor (2021)

Tabela 34: Matriz de transição a priori estimada (Norte)

| | A | B | C | D | E |
|---|---------|---------|---------|---------|---------|
| A | 0.60832 | 0.39167 | 0.00000 | 0.00000 | 0.00000 |
| B | 0.28454 | 0.66544 | 0.05000 | 0.00000 | 0.00000 |
| C | 0.00000 | 0.00000 | 0.48570 | 0.51429 | 0.00000 |
| D | 0.00000 | 0.00000 | 0.45240 | 0.54759 | 0.00000 |
| E | 0.00000 | 0.00000 | 0.00000 | 0.00530 | 0.99469 |

Fonte: IBGE. Elaboração do autor (2021)

Tabela 35: Matriz de transição a priori estimada (Sudeste)

| | A | B | C | D | E |
|---|---------|---------|---------|---------|---------|
| A | 0.47532 | 0.52467 | 0.00000 | 0.00000 | 0.00000 |
| B | 0.55885 | 0.44114 | 0.00000 | 0.00000 | 0.00000 |
| C | 0.00000 | 0.00081 | 0.71299 | 0.28620 | 0.00000 |
| D | 0.00000 | 0.00000 | 0.30938 | 0.69061 | 0.00000 |
| E | 0.00000 | 0.00000 | 0.00000 | 0.01153 | 0.98847 |

Fonte: IBGE. Elaboração do autor (2021)

Tabela 36: Matriz de transição a priori estimada (Sul)

| | A | B | C | D | E |
|---|---------|---------|---------|---------|---------|
| A | 0.49562 | 0.50437 | 0.00000 | 0.00000 | 0.00000 |
| B | 0.43013 | 0.42107 | 0.14879 | 0.00000 | 0.00000 |
| C | 0.00000 | 0.00000 | 0.98886 | 0.01113 | 0.00000 |
| D | 0.00000 | 0.00000 | 0.05491 | 0.94508 | 0.00000 |
| E | 0.00000 | 0.00000 | 0.00000 | 0.01541 | 0.98458 |

Fonte: IBGE. Elaboração do autor (2021)