

RANKING DE COMPLEXIDADE ECONÔMICA DOS ESTADOS BRASILEIROS

Leo Cavalcante¹

Gustavo Kaique Araujo Monea²

Fernando Fagundes Ferreira³

RESUMO

Neste trabalho apresentamos uma análise do impacto do tipo de dados disponíveis na RAIS (Relação Anual de Informações Sociais) para o cálculo da complexidade econômica dos estados brasileiros. A metodologia utilizada foi o índice de complexidade econômica para obtenção do ranking e posteriormente o uso da análise de componentes principais para unificar diferentes variáveis candidatas para o cálculo da complexidade econômica. O estudo da complexidade econômica permite determinar desigualdades regionais da estrutura produtiva entre países ou entre estados em um território nacional. Contudo, as variáveis utilizadas no cálculo do índice de complexidade não precisa ser a mesma. A pergunta que surge é saber em que medida diferentes base de dados ou variáveis podem influenciar o ranking da complexidade econômica. O resultado principal neste estudo é que o índice de complexidade é relativamente robusto mediante mudança de variáveis, mas uma solução parcimoniosa é usar uma combinação de diferentes variáveis obtida por meio da análise de componentes principais.

Palavras-chave: Complexidade econômica; Análise de componentes principais; Ranking de complexidade; Vantagem comparativa revelada; Complexidade econômica brasileira.

ABSTRACT

In this paper we present an analysis of the impact of different type of input data available in RAIS (Annual Report of Social Information) for the calculation of the economic complexity for Brazilian states. The methodology used here was the economic complexity index to obtain the economic complexity ranking as well as the principal component analysis to unify different candidate variables for the calculation of such complexity index. The study of economic complexity index makes it possible to determine regional inequalities in the productive structure between countries or between states in a national level. However, the variables used in calculating the complexity index need not be the same. The question that arises is to what extent different databases or variables can influence the ranking of economic complexity. The main result in this study is that the complexity index is relatively robust by changing variables,

¹Mestre em Modelagem de Sistemas Complexos-EACH/Universidade de São Paulo- USP.

²Mestre em Modelagem de Sistemas Complexos-EACH/Universidade de São Paulo- USP.

³Doutor e Professor da Universidade de São Paulo-USP.



but a parsimonious solution is to use a combination of different variables obtained through the principal components analysis.

Keywords: Economic complexity index; Principal component analysis; Complexity ranking; Revealed comparative advantage; Brazilian economic complexity.

1 INTRODUÇÃO

Se pensarmos na economia como um sistema evolutivo, a complexidade econômica mede o grau de sofisticação ou diversidade da estrutura produtiva de uma nação no âmbito internacional ou de uma região no caso nacional. Estudos recentes sobre complexidade econômica revelou que nações mais desenvolvidas possuem maior complexidade. O índice de complexidade econômica está positivamente correlacionado com produto interno bruto, renda, grau de escolaridade. Isso leva a inferir que a complexidade caminha na direção do desenvolvimento. Mais ainda, existe uma forte correlação entre complexidade e desigualdade social (Hartmann et al.,2010). Portanto, medir a complexidade econômica é de grande interesse tanto do ponto de vista de análise comparativa econômica quanto do ponto de vista de políticas públicas.

Considere uma rede de produtos produzidos por um país. Provavelmente os produtos mais sofisticados possuem mais conexões do que produtos menos sofisticados. Isso se deve ao fato que os produtos mais sofisticados demandam mais inputs. Em termos de medidas de centralidade de redes podemos dizer que os produtos mais sofisticados têm maior centralidade e ocupam uma posição central na rede e os menos complexos ficam na periferia da rede. O mesmo acontece se consideramos a produção global e identificarmos os países que produzem produtos mais sofisticados. Os países tendem a especializar-se em produtos próximos daqueles que já estão consolidados no seu espaço de produtos. Por isso, os países que estão na parte central da rede conseguem mudar sua estrutura produtiva mais rapidamente do que aqueles que estão nos clusters mais periféricos da rede. Os países que fazem parte do cluster central são mais desenvolvidos do que aqueles que estão nos clusters periféricos. Isto pode trazer uma explicação sobre o motivo pelo qual os países mais pobres não conseguem desenvolver suas exportações e atingir um determinado nível mais elevado de renda (Hidalgo; Hausman, 2009). Esta rede de produtos é o ponto de partida para se medir a complexidade econômica.

Outro aspecto importante é o fato de que o espaço de produtos mostra quais as possibilidades que um país possui em termos tecnológicos e sua capacidade para mudar sua estrutura produtiva. Produtos com alta complexidade não podem ser feitos por economias mais simples, pois exigem conhecimentos e recursos que estão presentes nos clusters centrais da rede

e são escassos na periferia da rede. A complexidade econômica é expressa pela composição das estruturas produtivas que emergem para manter e combinar diversos tipos de conhecimento (Hidalgo; Hausman, 2009),

Por meio do espaço de produtos, (Hidalgo; Hausman, 2009), criaram o índice de complexidade econômica (ECI). Este índice mede a quão complexa e diversificada é uma economia, ou seja, como é a sua rede de parceiros comerciais e quão ubíquos são os produtos. Produtos poucos ubíquos tendem a ser mais sofisticados e não podem ser feitos por quaisquer estruturas produtivas. Para que um país seja considerado complexo deve ser diverso nas relações comerciais e os seus produtos devem ser poucos ubíquos e exigir especialização maior para serem produzidos.

Entende-se por complexidade econômica a multiplicidade de conhecimentos incorporados em uma economia. Segundo (Hidalgo; Hausman, 2009), como há um limite naquilo que os indivíduos sabem, a única maneira para que as sociedades possam expandir suas bases de conhecimento é através da interação entre indivíduos. Essas interações formam redes cada vez mais complexas de cadeias produtivas.

Neste artigo serão feitas duas adaptações no índice de complexidade. A primeira diz respeito ao espaço geográfico aplicado, neste caso a aplicação não será feita para países, mas sim para estados brasileiros entre o 2010-2019. A questão a ser endereçada é a equivalência dos rankings quando se usa vínculos, estabelecimentos e massa salarial (Mearly et al., 2019). A segunda adaptação diz respeito aos inputs usados para computar do ECI. Aqui será usado uma variação do que foi proposto por (Mearly et al., 2019), ou seja, ao invés de usar apenas o número de vínculos por indústrias, será usado uma combinação de três inputs, a saber: vínculos na indústria, massa salarial e quantidade de estabelecimentos para calcular o ECI. Essa combinação é obtida pelo emprego da análise de componentes principais.

2 REFERENCIAL TEÓRICO

A partir do surgimento de instrumentos de análise de sistemas adaptativos complexos, novos métodos de obtenção da resposta sobre a complexidade foram propostos, em especial para a economia. Existem várias propostas para o cálculo da complexidade econômica. Nossa escolha foi por uma formulação que relaciona a complexidade com redes de relacionamentos entre produtos e países e é devido a (Hidalgo et al., 2007). Essa proposta é facilmente adaptável para o caso em que os países são mapeados em estados, províncias ou cidades. Capacitando-as quanto possível. Tal estudo abriu a possibilidade de analisar uma evolução da estrutura produtiva de um país, com base na Vantagem Comparativa Revelada (VCR) de seus produtos, bem como a criação de uma rede que represente tudo o que é produzido naquele país. (Hidalgo; Hausmann, 2009), elaboraram o chamado índice de complexidade econômica apresentado na segunda seção que trata da metodologia.

Para cenários regionais, o ECI foi adaptado de várias maneiras para capturar a complexidade econômica das regiões nacionais. No cenário brasileiro, uma iniciativa começou com o projeto DataViva em 2013 e foi aperfeiçoado em (Freitas; Paiva ,2015), utilizando a sofisticação da pauta de exportação de diferentes localidades nacionais para medir uma complexidade econômica de estados e municípios. De forma semelhante, com base na pauta de exportação, (Hausmann et al., 2016) estudaram regiões do Panamá, enquanto (Reynolds et al,2018) estudaram as estruturas produtivas das províncias australianas. Recentemente, (Gao ; Zhou, 2018) adaptaram a medida ECI para estudo da complexidade regional da China. A novidade trabalho de Gao e Zhou foi o uso de variáveis de entrada extraídas da bolsa de valores de Xangai e Shenzhen. Eles usaram valor de mercado das empresas para medir a vantagem comparativa revelada dos setores atividades definidas, criando um proxy para a produção regional. Os resultados do estudo da complexidade regional chinesa são muito semelhantes ao caso brasileiro. Ou seja, as regiões mais complexas estão ao longo da costa chinesa, enquanto estava dispersa no interior do país, explicando o padrão de crescimento chinês marcado pela grande desigualdade em todo o país. No caso brasileiro, os estados mais complexos estão na costa sul e sudeste.

Vale ressaltar que, Mealy et al. (2018) usaram diferentes bancos de dados para análise da complexidade regional para estudar a complexidade econômica regional para Estados Unidos e para o Reino Unido. Eles mostram que também há pouca variação em relação à variável utilizada, semelhante ao que encontramos aqui. Isso reforça a robustez do método para diferentes bancos de dados. Claro que os resultados dos rankings são altamente correlacionados, mas não idênticos. Por isso, no presente trabalho o foco foi mostrar como escolher o melhor índice.

Outros estudos que analisam a complexidade envolvendo o Brasil e suas possíveis consequências no desenvolvimento econômico são os seguintes (Gala et al., 2017), (Sbardella et al., 2017) e (Alencar et al., 2018) .

3 METODOLOGIA

3.1 Índice de Complexidade Econômica

Para calcular o índice de complexidade econômica usaremos a metodologia criada por (Hidalgo;Hausmann,2009) e adaptada por (Mearly et al,2019).. O ponto de partida para obter este índice é calcular a vantagem comparativa revelada (VCR), conforme consta na equação (1).

$$RCA_{ip} = \frac{\frac{X_{ip}}{\sum_p X_{ip}}}{\frac{\sum_p X_{ip}}{\sum_i \sum_p X_{ip}}} \quad (1)$$

X_{ip} = Quantidade de vínculos i no estado p

i = número considerado de estados

P = número considerado de produtos

Essas medidas são usadas para construir a matriz $M \in R^{pxi}$ que liga cada estado aos produtos que produz, com as entradas. Um $RCA_{ip} \geq 1$ Significa que o estado e tem vantagem comparativa revelada no setor i , ou seja, o setor i é proporcionalmente mais maior do que em todos as outras localidades da economia.

$$M_{ip} = \begin{cases} 1 & \text{se } RCA_{ip} \geq 1 \\ 0 & \text{se } RCA_{ip} < 1 \end{cases}$$

Para medir a diversidade e ubiquidade simplesmente somam-se as linhas ou colunas da Matriz M_{cp} , definido formalmente conforme abaixo.

$$Diversidade = K_{p,0} \sum_i M_{i,p}. \quad (2)$$

$$Ubiquidade = K_{i,0} \sum_p M_{ip}. \quad (3)$$

O índice de complexidade econômica (ou Economic Complex Index- ECI) é definido como o autovetor (K) associado ao segundo maior autovalor de \tilde{M}_{pp} :

$$\tilde{M}_{pp} = \frac{1}{k_{p,0}} \sum_i \frac{M_{ip} M_{p',i}}{k_{i,0}}. \quad (4)$$

Este autovetor (K) é o que captura a maior quantidade de variação do sistema e que define o índice de complexidade econômica e sua formulação é dada pela expressão:

$$ECI = \frac{\bar{K} - \langle \bar{K} \rangle}{stdv(\bar{K})} = \frac{m^2 k_p - m \sum_p k_p}{\sqrt{m \sum_p (m k_p - \sum_p k_p)^2}}. \quad (5)$$

Aqui a notação $\langle \rangle$ representa a média sobre as componentes do vetor K e stdv é o desvio-padrão sobre o mesmo vetor. Assim, o ECI é uma quantidade padronizada.

3.2 Dados Empíricos

Para este artigo os dados foram extraídos de 561 classes conforme a nomenclatura da CNAE 95 oriundos da RAIS – Relações anuais de informações sociais, fornecido pelo ministério do trabalho e emprego, entre os anos de 2010 e 2019 e para os 26 estados da federação mais o distrito federal.

ECI Vínculos: É o ECI calculado a partir dos vínculos empregatícios das classes da CNAE 95, para todos os estados e entre o período de 2010 a 2019. Esta variável leva em conta o número de empregos formais em cada estabelecimento na base de dados RAIS, para todos os estados da federação.

ECI Massa Salarial: É o ECI calculado a partir da massa salarial das classes da CNAE 95, para todos os estados e entre o período de 2010 a 2019. Neste caso, a variável leva em conta os salários dos trabalhadores de cada estabelecimento na base RAIS, para todos os estados da federação.

ECI Estabelecimentos: É o ECI calculado a partir do total de estabelecimentos das classes pertencentes a CNAE 95, para todos os estados e entre o período de 2010 a 2019.

4 ANÁLISE DE RESULTADOS

Nesta seção iremos apresentar os resultados das análises do ranking de complexidade dos estados brasileiros e distrito federal entre 2010 e 2019 calculados a partir de três variáveis de entradas distintas. Em seguida iremos apresentar uma combinação linear entre elas, obtida pelo uso da PCA, considerada aqui uma proxy parcimoniosa para resolver eventuais mudanças de posição no ranking.

A tabela 1 mostra a correlação entre os três índices de complexidade. Vemos que todos eles se correlacionam acima de 95%, sendo a maior correlação encontrada foi entre o ECI vínculos e ECI de estabelecimentos com 99% de correlação, a menor correlação é entre ECI massa salarial e ECI estabelecimentos que é 96%.

Tabela 1: Matriz de Correlação com inputs distintos para o ECI

	ECI Vínculos	ECI Estabelecimento	ECI Massa Salarial
ECI Vínculos	1	0.993647	0.968937
ECI Estabelecimentos	0.993647	1	0.963553
ECI Massa Salarial	0.968937	0.963553	1

Fonte: Elaborado pelos autores, com a base de dados da RAIS.

A tabela 2 mostra os 5 estados mais complexos para cada um dos índices calculados em 2010. O estado de São Paulo ficou entre o mais complexo do período para os três índices calculados. O ECI de estabelecimentos e vínculos mantém as mesmas posições dos estados, já no caso de ECI o Rio de Janeiro substituiu o estado do Amazonas e figura em segundo lugar.

Tabela 2: Top 5 complexidades para os três índices em 2010

ECI estabelecimentos	ECI massa	ECI vínculos
SP	SP	SP
RS	RJ	RS
PR	RS	PR
SC	SC	SC
AM	PR	AM

Fonte: Elaborado pelos autores, com a base de dados da RAIS.

Para os estados menos complexos o ECI dos estabelecimentos e dos vínculos possuem os estados nas mesmas posições, sendo Tocantins o menos complexo conforme tabela 3. Já o

ECI da massa salarial tem uma mudança nas posições com os mesmos estados dos outros dois índices.

Tabela 3: Top 5 menores complexidades para os três índices em 2010

ECI estabelecimentos	ECI massa	ECI vínculos
TO	AP	TO
MA	AC	MA
AP	RR	AP
AC	TO	AC
RR	MA	RR

Fonte: Elaborado pelos autores, com a base de dados da RAIS.

A tabela 4 mostra os estados mais complexos em 2019 segundo os três índices calculados, os cinco estados mais complexos segundo o ECI estabelecimentos e ECI vínculos são os mesmos e seguem a mesma ordem. Já para o ECI da massa salarial, o estado do Rio de Janeiro entra no top cinco, mas fica em quinto lugar.

Tabela 4: Top 5 complexidades para os três índices em 2019

ECI estabelecimentos	ECI massa	ECI vínculos
SP	SP	SP
SC	SC	SC
RS	RS	RS
PR	AM	PR
AM	RJ	AM

Fonte: Elaborado pelos autores, com a base de dados da RAIS.

A tabela 5 mostra os estados menos complexos em 2019, novamente as posições do ECI dos estabelecimentos e vínculos são as mesmas, já para o ECI da massa as posições mudaram, mas mantiveram os mesmos estados que compõe os outros dois índices.

Tabela 5: Top 5 menores complexidades para os três índices em 2019

ECI estabelecimentos	ECI massa	ECI vínculos
AL	PI	AL
MA	AC	MA
AC	TO	AC
PI	MA	PI
TO	AL	TO

Fonte: Elaborado pelos autores, com a base de dados da RAIS.

A questão que se coloca é qual variável usar para o cálculo do índice de complexidade econômica? Uma solução proposta aqui é criar um índice a partir da análise de componentes principais (PCA) (Mauricio et al.,2015) e tomar a primeira, caso tenha alto poder de explicação. Foi justamente o caso, pois a primeira componente da PCA explica 99% da variação do conjunto de dados. Neste caso a contribuição marginal de se adicionar mais um componente é mínimo, conforme a figura 1.

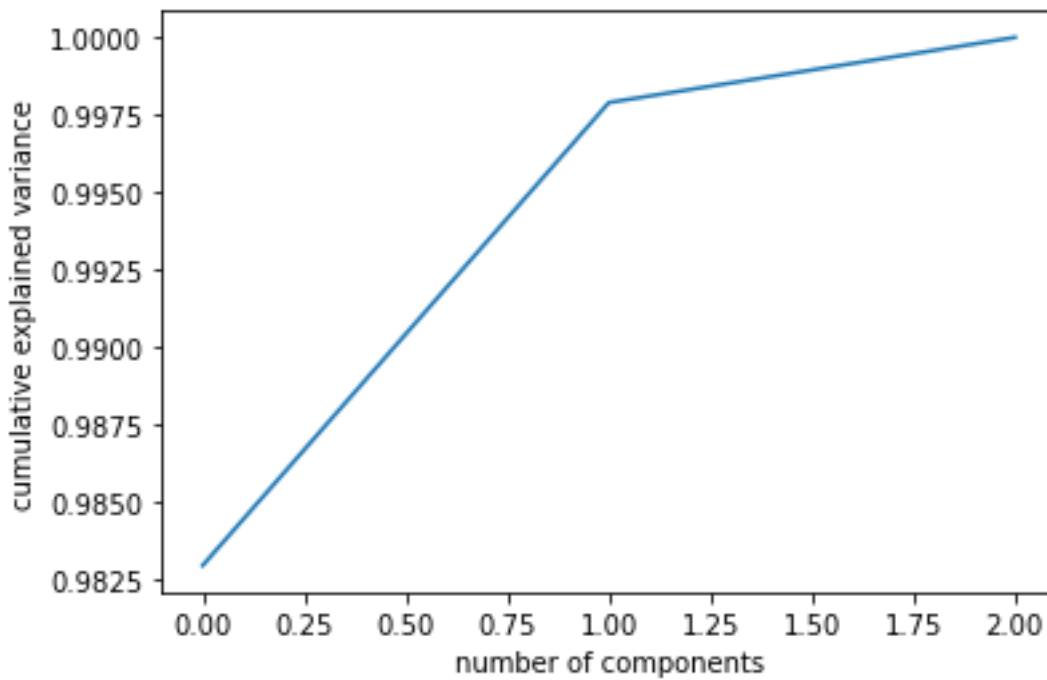


Figura 1. Variância explicada pela PCA em função do número de componentes principais utilizando dados da RAIS entre 2010 e 2019. Elaboração próprios autores.

A figura 2 mostra a variação das posições no Ranking de complexidade usando o ECI após a aplicação da PCA. Os cinco estados mais complexos no início do ranking (2010) foram os seguintes, SP, RS, SC, PR e AM. Já os cinco estados menos complexos no início foram RR, MA, AC, AP e TO. No ano de 2019 os cinco mais complexos foram SP, SC, RS, PR e AM. Os menos complexos foram TO, AL, PI, AC e MA.

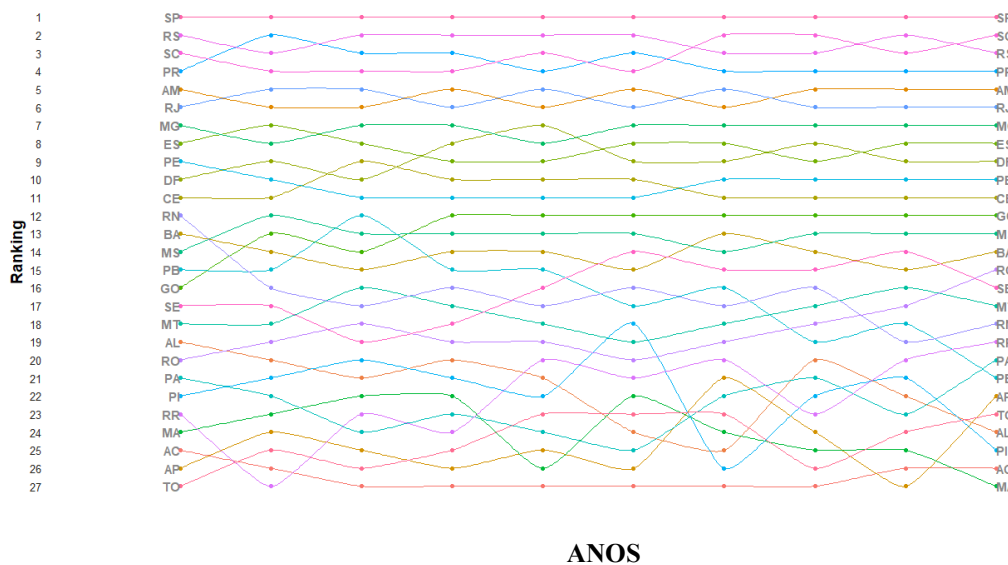


Figura 2: Ranking de complexidade usando a primeira componente da PCA (2010-2019). Elaboração próprios autores.

Para reduzir a grande variabilidade que na alternância das posições, foi aplicada uma média móvel para os ECI PCA, a alternância entre os dados mudou de forma significativa. Os estados mais complexos em 2012 eram SP, RS, PR, SC e RJ. Em 2019 eram os estados de SP, SC, RS, PR e AM. Os menos complexos em 2012 eram os estados do MA, RR, AP, AC e TO. Em 2019 os menos complexos eram PI, AP, TO, MA e AC. Conforme a figura 3 abaixo.

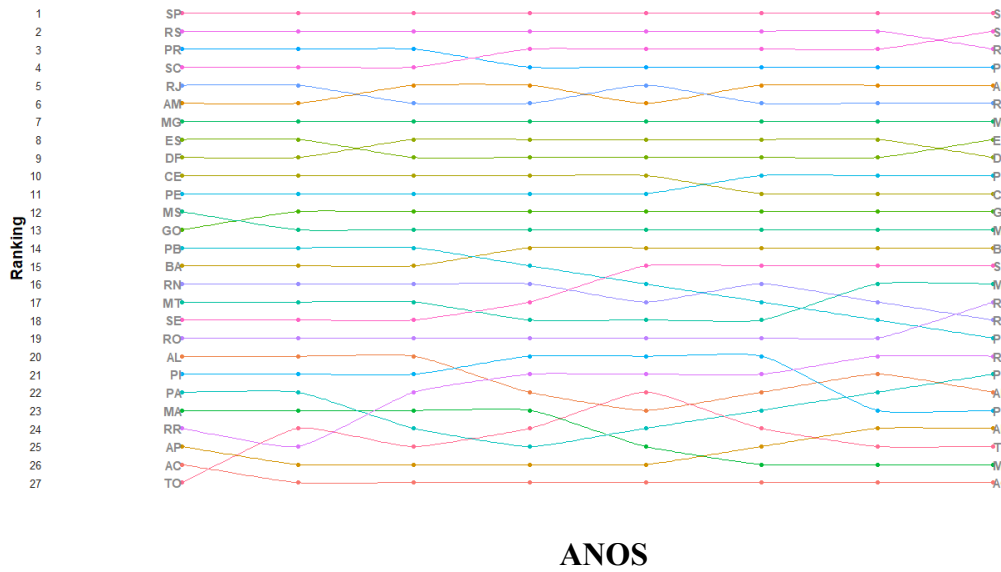


Figura 3: Ranking usando a média móvel do índice de complexidade via PCA (2012-2019). Elaboração próprios autores.

O estado de São Paulo é o mais complexo nas três variáveis utilizadas conforme esperado. Os três estados da região sul disputam as próximas colocações (segundo, terceiro e quarto lugares) no ranking. Este ranking não reflete uma competição direta e sim reflete o grau de sofisticação que estas economias possuem quando comparada entre si. Em outras palavras, podemos dizer que as estruturas produtivas são similares e por isso o valor da complexidade é tão próximo ponto de gerar esta aparente competição no ranking. Mesmo a média móvel não ainda manteve este padrão. Talvez aumentar a janela da média móvel possa diminuir a flutuação e permitir estabelecer um ordenamento estável. Em seguida, no ranking aparecem os demais estados do sudeste, Amazonas e Distrito Federal. Também é esperado que estes estados apareçam bem colocado, pois concentram maior população e maior parte do parque industrial brasileiro. No caso do Distrito Federal, a explicação é que a arrecadação dos tributos no Brasil no âmbito federal é centralizada na capital (Brasília). Com isso, a administração pública e o comércio têm grande vantagem comparativa em relação a maioria dos estados brasileiros. No caso do estado do Amazonas, que está na região norte, a explicação é a presença da zona franca de Manaus. Foi uma decisão política de desenvolvimento regional que levou a formação de um complexo regional cujo impacto se vê no índice de complexidade econômica.

Na parte inferior do ranking encontram-se estados do norte e nordeste e centro-oeste. Também é um resultado coerente e esperado. Estas regiões se destacam pela atividade agrícola e extrativista. Possuem piores índices de Gini, IDH, PIB per capita por exemplo. Basicamente, o ECI corrobora claramente uma diferença regional conhecida: sudeste e sul mais rico e desenvolvido e norte, nordeste centro-oeste mais pobre e menos desenvolvido. Este resultado é importante para validarmos a medida de complexidade regionalizada.

5 CONSIDERAÇÕES FINAIS

Como continuação deste trabalho podemos destacar três problemas a serem endereçados. O primeiro seria o estudo da relação entre desigualdade de renda e complexidade econômica. Segundo Hartmann et al. (2017), quanto maior a complexidade de uma economia, menor a desigualdade. Isto se deve ao fato de uma complexidade econômica mais elevada exigir mais qualificação da mão de obra. Assim, trabalhadores mais qualificados tendem a ganhar mais, tendem a se organizar melhor para reivindicar maiores salários. Não sabemos como seria esta relação em uma economia como a brasileiras. Talvez diferenças regionais de complexidade não tenha forte relação com a desigualdade de renda. Precisamos estudar esta relação para saber se a complexidade econômica de fato tem relação estatística significativa com a desigualdade. Outro ponto interessante seria estudar a relação entre desigualdade setorial e complexidade econômica. Neste caso espera-se que haja uma correlação significativa entre elas. Por fim, estudar a rede de estabelecimentos à semelhança à rede de produtos no trabalho seminal de Hidalgo et al. (2007), para o caso brasileiro e determinar a complexidade da estrutura produtiva tanto para estados quanto para microrregiões.

Este trabalho contribui para a literatura da complexidade econômica regional de duas maneiras complementares. A primeira foi mostrar que diferentes variáveis disponíveis na *RAIS* levaram à resultados muito parecidos, o que indica robustez da medida do índice de complexidade econômica. A segunda foi que o uso da primeira componente principal é uma alternativa promissora e parcimoniosa para decidir sobre eventuais impasses no ordenamento da complexidade mediante diferentes variáveis de entrada. A validação dos resultados empíricos se dá pela coerência nos rankings de complexidade, ou seja, a ordem dos estados do primeiro ao último colocado reflete o que se sabe em termos de desenvolvimento regional por Freitas e Paiva (2015). Consideramos que a adaptação do cálculo do *ECI* para o âmbito regional foi bem sucedida. Esperamos que o estudo da complexidade econômica possa sugerir no futuro estratégias de desenvolvimento regional baseada no conceito de redes e de uma economia como um sistema complexo evolutivo e adaptativo.

REFERÊNCIAS

ALENCAR, Júlia F. L.; FREIAS, Elton; ROMERO, João P.; BRITTO, Gustavo. Economic Complexity and Development: An Analysis of the Latin American Case. *Novos estud. CEBRAP* vol.37 no.2 São Paulo May/Aug. 2018

FREITAS, Elton Eduardo; PAIVA, Emilia Andrade. Diversificação e sofisticação das exportações: uma aplicação do product space aos dados do Brasil. *Revista Econômica do Nordeste*, v. 46, n. 3, p. 79-98, 2015.

GALA, Paulo; CAMARGO, Jhean; FREIAS, Elton. The Economic Commission for Latin America and the Caribbean (ECLAC) was right: scale-free complex networks and core-periphery patterns in world trade. *Cambridge Journal of Economics*, Volume 42, Issue 3, May 2018, Pages 633–651

GAO, Jian; ZHOU, Tao. Quantifying China's regional economic complexity. *Physica A* 492 (2018) 1591–1603.

HARTMANN, Dominik; GUEBARA, Miguel R.; FIGUEROA, Cristian Jara; ARISTÁN, Manuel; HIDALGO, Cesar. Linking Economic Complexity, Institutions and Income Inequality. *World Development* Volume 93, Pages 75-93. 2017

HAUSMANN, Ricardo; MORALES-ARILLA, Jose; SANTOS, Miguel. Economic Complexity in Panama: Assessing opportunities for productive diversification. *Development Economics: Microeconomic Issues in Developing Economies eJournal* 2016.

HIDALGO, César A. et al. The product space conditions the development of nations. *Science*, v. 317, n. 5837, p. 482-487, 2007.

HIDALGO, César A.; HAUSMANN, Ricardo. The building blocks of economic complexity. *Proceedings of the national academy of sciences*, v. 106, n. 26, p. 10570-10575, 2009.

MEALY, Penny; FARMER, J. Doyné; TEYTELBOYM, Alexander. A new interpretation of the economic complexity index. Alexander, A New Interpretation of the Economic Complexity Index (February 4, 2018), 2018.

REYNOLDS, Christian et al. A sub-national economic complexity analysis of Australia's states and territories. *Regional Studies*, v. 52, n. 5, p. 715-726, 2018.

SBARDELLA, Angelica; PUGLIESE, Emanuele; PIETRONERO, Luciano. Economic development and wage inequality: A complex system analysis. **Plos One** 19, 2017.

SILVA, Mauricio Corrêa da; Silva, GOMES, José Dionísio; BORGES, Erivan Ferreira. **Revista Brasileira de Biometria**. São Paulo, v.33, n.3, p.291-309, 2015