

Identificação de Nichos de Mercado para Países Exportadores: uma Análise Multivariada para o Ano de 2011

Identifying Niche Markets for Exporting Countries: a Multivariate Analysis for 2011

Vinícius de Azevedo Couto Firme*

Cláudio Roberto Fóffano Vasconcelos**

Resumo: Este trabalho tem por objetivo a identificação de nichos de mercado com base nas exportações de 97 tipos de produtos realizadas por 68 países em 2011. Para tanto, técnicas de análise multivariada são utilizadas no intuito de determinar os grupos de países que tendem a exportar produtos semelhantes. Os resultados indicam que as exportações de produtos da indústria pesada (metal e cimento) e tecnológicos (máquinas, reatores e eletrônicos) são o principal fator de diferenciação entre os países. A China é o país com maior pontuação nesses critérios (*outlier*), seguida pelo grupo de Estados Unidos e Alemanha. O Brasil está inserido em um grupo intermediário e parece concorrer, de forma mais intensa, com países como Índia, México, Malásia, Turquia e outras nove economias. De modo geral, qualquer acordo comercial que vise economias inseridas em um mesmo grupo deveria considerar o fato de que elas apresentam baixa complementaridade entre si e disputam os mesmos nichos no mercado externo.

Palavras-chave: Comércio internacional. Nichos de mercado. Análise multivariada.

Abstract: This study aimed to identify market niches based on exports of 97 types of products made by 68 countries in 2011. Therefore, multivariate analysis techniques were used in order to determine which groups of countries are exporting similar products. The results indicate that exports of products of heavy industry (metal and concrete) and technology (machinery, reactors and electronic products) are the main differentiation factor between countries. China was the country with the highest score in these criteria (*outlier*) followed by Group USA and Germany. Brazil was included into an intermediate group and seems to compete more intensively with countries like India, Mexico, Malaysia, Turkey and nine other economies. In general, any trade agreement that focuses on economies inserted into the same group should consider the fact that they are not complementary between themselves and have been competing for the same niche in the foreign market.

Keywords: International trade. Market niches. Multivariate analysis.

JEL Classification: F01; F14; R12; C38.

* Doutor em Economia Aplicada pela Universidade Federal de Juiz de Fora (UFJF). Professor do Departamento de Economia da UFJF - campus Governador Valadares. E-mail: vinicius.firme@ufjf.edu.br

** Doutor em Economia pela Universidade Federal de Pernambuco (UFPE). Professor do Programa de Pós-Graduação em Economia Aplicada da Universidade Federal de Juiz de Fora (UFJF). E-mail: claudio.foffano@ufjf.edu.br

1 Introdução

Este trabalho analisou as exportações, referentes ao ano de 2011, realizadas por 68 países, desagregadas em 97 produtos, a fim de definir possíveis nichos de mercado. Para tanto, técnicas de análise multivariada foram empregadas no intuito de identificar agrupamentos de países que exportam produtos semelhantes.

Segundo a teoria da base exportadora, uma das primeiras a tentar explicar o crescimento regional, boa parte das disparidades poderia ser explicada pela capacidade de cada região em produzir bens exportáveis. North (1977) revela que o crescimento regional não implicaria apenas na exportação de produtos industrializados. Segundo ele, nada impediria que uma região agrícola-exportadora se desenvolvesse rapidamente. A ideia básica é a de que o aumento da produção da base exportadora exerce um efeito multiplicador sobre as atividades não básicas ou de mercado interno (BALASSA, 1989; SIRKIN, 1959; SOUZA, 1980; TIEBOUT, 1977). Desde então, estudos têm demonstrado que os setores exportadores tendem a ser mais produtivos e têm mais capacidade de induzir o crescimento que aqueles que se destinam apenas a atender o mercado interno (AW; HWANG, 1995; DELGADO *et al.*, 2002; EATON *et al.*, 2004; FERREIRA; ROSSI, 2003; GOMES; ELLERY, 2007; HAY, 2001; HIDALGO; MATA, 2009).

Uma vez que as exportações constituem um componente importante ao crescimento,¹ seria estratégico identificar quais países concorrem entre si e quais fatias de mercado externo eles disputam. Imagina-se que cada país é uma empresa que exporta uma cesta de produtos. Assim, conforme lembra Porter (1979), nem toda a empresa (ou país) está em concorrência com todas as demais ou na mesma intensidade. Na realidade, é mais provável que algumas estejam disputando um determinado segmento, no qual obtêm vantagem competitiva, enquanto outras estejam voltadas para setores em que são mais eficientes.

Dessa forma, seria possível agrupar os países de acordo com o tipo de produto que vendem (exportam) e a escala (tamanho) de produção. Assim, aqueles que estiverem no mesmo grupo (ou em grupos próximos) concorreriam entre si de modo mais intenso que os demais países incluídos em grupos distantes. A identificação do nicho de mercado ao qual cada país pertence, assim como seus concorrentes, pode ser um fator decisivo no mercado internacional. Ferraz *et al.* (1997) argumenta que os países inseridos em um nicho de mercado tendem a apresentar vantagem competitiva nesse ramo, dificultando a entrada de novos integrantes. Assim, a manutenção de um espaço privilegiado no mercado seria estratégica e

1 Cabe ressaltar que a base exportadora sozinha não explica integralmente o crescimento econômico global, principalmente quando a região se industrializa e aumenta de tamanho. Afinal, se se pensar no mundo como um país, se concluirá que não existem exportações, no entanto, a renda cresce (TIEBOUT, 1977).

garantiria a lucratividade dos membros. Para Mascarenhas e Aakker (1989), a dificuldade em se inserir nesses nichos funcionaria como uma barreira à concorrência e ajudaria a explicar o bom desempenho de empresas e países. Porter (1979, 1989) argumenta que, sem a devida proteção, qualquer estratégia bem-sucedida seria rapidamente imitada e a rentabilidade entre os grupos tenderia à igualdade.

Dado o exposto, este artigo utiliza técnicas de análise multivariada a fim de identificar os nichos de mercado ao qual cada país pertence, conforme o tipo de produto exportado. Nota-se que, quanto maior o grau de correlação entre os produtos comercializados por dois ou mais países, maior tenderá a ser a competição entre eles, dificultando, assim, o fechamento de acordos comerciais e políticas de preferência tarifária. Os resultados indicam que as exportações da indústria pesada (metal e cimento) e de produtos tecnológicos (máquinas, reatores e eletrônicos) são o principal fator de diferenciação entre os países. Em se tratando de Brasil, conclui-se que seria difícil fechar acordos bilaterais com países como Índia, México, Malásia, Turquia e os demais incluídos no agrupamento brasileiro. O fato é que essas economias não parecem ser complementares à brasileira. Ao invés disso, concorrem diretamente com os produtos brasileiros no mercado internacional.

Além desta introdução, este trabalho está organizado da seguinte forma: a metodologia e a base de dados utilizadas são apresentadas na seção 2. A seção 3 contém os resultados. Logo após encontra-se a conclusão, referências e anexo.

2 Metodologia e Base de Dados

Os dados referentes às exportações foram obtidos junto ao International Trade Centre (INTERNATIONAL TRADE CENTRE, 2015) e têm como base o ano de 2011 (expressos em milhares de dólares correntes). Inicialmente, foi possível obter informações para 221 países e 97 produtos. No entanto, como será visto, os métodos utilizados neste trabalho são sensíveis a *outliers* e informações pouco relevantes. Logo, optou-se por trabalhar apenas com países que exportavam uma parcela significativa de produtos.² Essa triagem inicial resultou em uma amostra contendo 68 países.

Uma vez definida a amostra, já seria possível criar grupos de países exportadores segundo o tipo de bem exportado. No entanto, a elevada quantidade de produtos dificultava a caracterização dos grupos. Além disso, alguns produtos apresentavam uma correlação elevada, indicando que poderiam fazer parte da mesma indústria ou cadeia produtiva. Assim, foi realizada uma agregação dos 97 produtos, utilizando a técnica multivariada de análise de *clusters* (AC), que indicou a existência de 25 cestas de produtos exportáveis. Feito isso, utilizou-se o método

2 Na realidade, todos os países que apresentavam missings em algum dos 97 produtos foram excluídos.

de análise de componentes principais (ACP) sobre os 25 produtos. Tal técnica revelou que não era necessário analisar todos esses produtos, uma vez que apenas três componentes explicavam, aproximadamente, 83% da sua variância total. Dessa forma, esses componentes resultantes da ACP foram utilizados para agregar os países de acordo com o método de AC.

2.1 Método de Análise de *Clusters* (AC)

A AC caracteriza-se por ser um método de análise multivariada, com o objetivo de verificar a existência de grupos dentro de um conjunto de observações. Para tanto, utiliza algum critério de forma a maximizar a homogeneidade entre os elementos intragrupos e maximizar a heterogeneidade dos elementos intergrupos (MINGOTI, 2007).³

De acordo com Vicini (2005), a estrutura básica da aplicação da AC pode ser representada em etapas, da seguinte forma: a) formulação do problema; b) definição de uma medida de similaridade ou dissimilaridade; c) escolha de um processo de aglomeração; e d) avaliação e interpretação dos resultados. Além disso, como essa técnica é sensível à escala das variáveis, recomenda-se que estas sejam normalizadas. Dado que as questões *a* e *d* dependem do propósito do pesquisador, são analisados apenas os pontos *b* e *c*. Sendo assim, o próximo passo consiste na definição de uma medida de similaridade ou dissimilaridade. Kageyama e Leone (1999, p. 20), argumentando sobre as medidas de distância, utilizadas como critérios de dissimilaridade, revelam que “[...] a distância mais usual, quando os valores das variáveis referentes aos indivíduos são valores reais, é a distância Euclidiana”.⁴ Essa afirmação é corroborada por Manly (1986 *apud* PEREIRA, 2002, p. 36). Logo, neste trabalho utilizou-se essa medida de distância.⁵

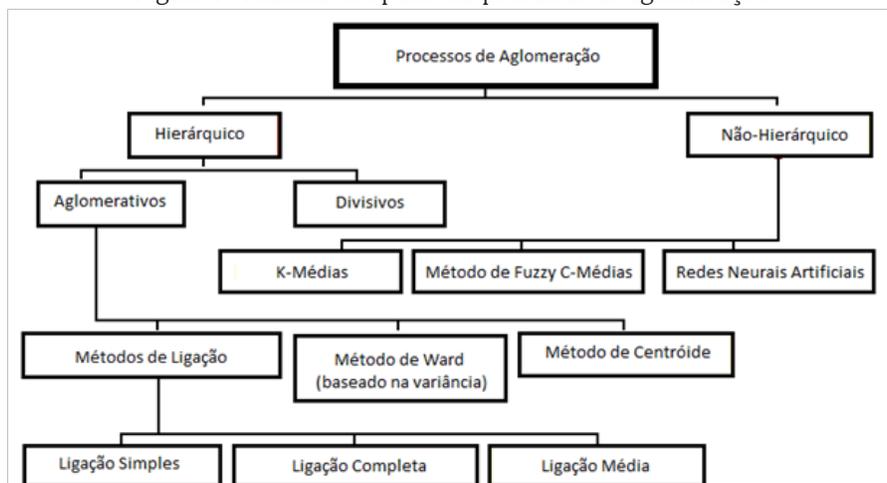
Uma vez determinado o critério de (dis)similaridade, deve-se selecionar a técnica de aglomeração (ou formação de *clusters*). A Figura 1 resume o processo de decisão quanto ao método de aglomeração adotado.

3 É importante lembrar que esse método avalia todo um conjunto de relações interdependentes sem estabelecer uma relação de causa e efeito entre as variáveis (não se trata de uma regressão). O mesmo vale para ACP e AF.

4 Considerando uma análise para o caso de dois vetores, X_i e X_k , com $i \neq k$, a distância euclidiana é: $d(X_i, X_k) = \sqrt{[(X_i - X_k)' I (X_i - X_k)]}$, em que I representa uma matriz identidade.

5 Outras medidas poderiam ter sido utilizadas. Para Mingoti (2007), a distância de Minkowsky é menos afetada por *outliers*. Já a distância de Mahalanobis (1936) atenuaria as distorções geradas por variáveis com elevada variância.

Figura 1 - Resumo dos possíveis processos de aglomeração



Fonte: Elaboração própria a partir de Mingoti (2007).

O primeiro passo consiste em definir se a técnica utilizada será do tipo hierárquica ou não. A principal diferença entre as duas consiste no conhecimento prévio do número ótimo de grupos (g). Caso g seja conhecido, pode-se utilizar a abordagem não hierárquica, do contrário, os modelos hierárquicos são mais indicados.

Mingoti (2007) recomenda que a utilização do método hierárquico seja mais exploratória do que definitiva. A autora argumenta que esse método auxilia a determinação do número ótimo de grupos (g).⁶ Segundo Andrade (2009, p. 63-64), quando se trabalha com um número elevado de informações, os métodos não hierárquicos são preferíveis. Para o autor, “[...] os algoritmos computacionais utilizados nos métodos não hierárquicos são do tipo iterativo e, em comparação com os métodos hierárquicos, têm a maior capacidade de análise de um conjunto de dados com um grande número de observações”.⁷

Cabe ressaltar que, embora superiores (do ponto de vista computacional), os modelos não hierárquicos requerem que o pesquisador tenha conhecimento prévio do número de *clusters* (g). Além disso, deve-se definir quais serão as semen-

6 Como esse modelo pressupõe que existe uma hierarquia aglomerativa, é possível traçar uma “árvore histórica” do processo de agrupamento (desde o ponto em que cada indivíduo representava um grupo, até o ponto em que todos os indivíduos se encontram no mesmo grupo). Esse histórico recebe o nome de dendograma (TIMM, 2002) e permite visualizar quais indivíduos são mais homogêneos e quais são *outliers* (ANDRADE, 2009).

7 As técnicas não hierárquicas requerem a especificação prévia do número de grupos desejado (k). Assim, a cada estágio de agrupamento novos *clusters* podem se formar através da divisão ou junção de grupos já combinados em passos anteriores até que se atinjam os k grupos pré-determinados. Logo, se dois elementos tiverem sido colocados num mesmo grupo, não necessariamente eles permanecerão juntos na partição final, impossibilitando a construção de dendogramas (ANDRADE, 2009; MINGOTI, 2007).

tes iniciais (centroides) que iniciarão o processo de aglomeração via k -médias.⁸ Mingoti (2007, p. 193) sugere alguns procedimentos para a escolha das sementes. Dentre eles, o que parece mais eficiente (e menos arbitrário) consiste na utilização de uma técnica hierárquica para definição dos K grupos. Feito isso, calcula-se a média vetorial desses grupos e utiliza-se como semente no método não hierárquico.⁹ Logo, optou-se pelo método hierárquico com finalidade de encontrar g e seus respectivos elementos. Dessa forma, foi possível obter as sementes, necessárias ao modelo não hierárquico, através da média dos elementos de cada grupo criado. A classificação hierárquica ascendente, ou aglomerativa (mais comum), consiste em considerar cada indivíduo como um grupo. A partir disso, utiliza-se alguma medida de similaridade para agrupar os elementos continuamente até que exista apenas um grupo contendo todos os indivíduos (KAGEYAMA; LEONE, 1999). Obviamente, um agrupamento contendo todos os indivíduos seria muito heterogêneo. Logo, deve-se utilizar algum critério para definir o número adequado de grupos. Conforme demonstrado na Figura 1, existem cinco métodos de agrupamentos hierárquicos: a ligação média produz grupos com variância semelhante e tende a apresentar partições melhores que as ligações simples e completa (MINGOTI, 2007; VICINI, 2005). Já o método de centroides e o de Ward (1963) partem de uma estrutura de seleção (vetorial) semelhante, porém o método de Ward (1963) tende a produzir *clusters* com variância mínima e com um número semelhante de elementos intragrupos.¹⁰

Conforme ressalta Afonso e Melão (2007) e Mingoti (2007), não existe consenso sobre qual das técnicas de agrupamento hierárquico descritas deve ser utilizada. Como cada uma pode induzir a um agrupamento diferente, o pesquisador deve recorrer a alguns critérios de decisão que ajudam a definir o número de *clusters* g final. Nesse caso, dois critérios foram adotados: a) estatística pseudo-F;

8 Também poderiam ter sido utilizados o método *fuzzy* e a técnica de redes neurais. Quanto ao método *fuzzy* (c médias), desenvolvido por Bezdek (1981), também se trata de uma técnica iterativa em que o pesquisador deve definir o número de *clusters* (c) a priori. No entanto, diferentemente da k médias, esse método estima a probabilidade de que cada elemento pertença a um dos c grupos pré-determinados. Lima (2001 *apud* MINGOTI 2007), usando simulações de Monte Carlo, argumenta que a técnica *fuzzy* e a k médias apresentam resultados superiores às redes neurais. O autor compara diversas técnicas hierárquicas e não hierárquicas e, embora as não hierárquicas (*fuzzy* e k médias) tenham sido superiores, o método de Ward também apresentou bons resultados.

9 Os outros procedimentos consistem na escolha: a) aleatória das sementes (ineficiente); b) baseada na variável de maior variância; c) baseada em valores discrepantes dos dados; d) arbitrária (necessário justificar); e) escolha dos k primeiros elementos do banco de dados (interessante se os primeiros elementos forem discrepantes).

10 Afonso e Melão (2007) testam os cinco métodos descritos, e apenas a ligação média e o método de Ward produziram grupos com coerência econômica. Assim como Silva e Simões (2004), os autores optam pela técnica de Ward. A favor dos centróides, tem-se que: "[...] centroid method of measuring similarity is employed because this method is more robust to outliers than most other hierarchical methods" (FAN *et al*, 2007, p. 38).

e b) estatística pseudo- T^2 .¹¹ Mingoti (2007) ainda revela que, em qualquer um dos testes de definição do número g de *clusters*, pode haver trechos em que ocorrem diversos saltos nos critérios de decisão (aumentando e diminuindo diversas vezes de forma abrupta). Nesses casos, a autora sugere que se trata de um trecho de indecisão no qual a definição do número de grupos não é clara. Assim, caso o pesquisador acredite que g esteja nesse trecho, deve adotar outro método de definição.¹²

Feito isso, é possível obter o número de grupos e as sementes necessárias ao modelo não hierárquico. A classificação não hierárquica visa encontrar uma partição de n elementos em K grupos (*clusters*) de modo a garantir a maior coesão interna (similaridade intragrupo) e a maior heterogeneidade externa (intergrupos). Nesse método, cada elemento amostral é alocado àquele cluster cujo centroide ou semente (vetor de médias amostral) é o mais próximo do vetor de valores observados para o respectivo indivíduo. O método, originalmente, é composto por quatro passos: a) escolhe-se K centroides, reconhecidos como “sementes” ou “protótipos”, para se iniciar o processo de partição; b) compara-se, então, cada elemento do conjunto de dados a cada centroide inicial, por meio de uma medida de distância. Assim, o indivíduo é alocado ao grupo cuja distância é menor; c) aplicando-se o segundo passo a cada um dos n elementos amostrais, recalculam-se os valores dos centroides para cada novo grupo gerado e, então, repete-se a etapa *b*, considerando os centroides dessas novas classes; e d) os passos *b* e *c* devem ser repetidos até que todos os indivíduos amostrais estejam bem alocados em seus grupos. Isto é, até que nenhuma realocação de elementos seja necessária (MINGOTI, 2007).

11 O pseudo-F, proposto por Calinski e Harabasz (1974), é semelhante ao seu homônimo da econometria clássica e está baseado na variância entre os grupos, a cada nível de agregação. Assim, um valor alto do teste é desejável e implicaria na rejeição à hipótese de homogeneidade entre os grupos criados. Já o pseudo-T de Duda e Hart (1973) assemelha-se à estatística *T* de *student* tradicional elevada ao quadrado. Assim, um valor elevado do teste não seria desejável, pois implicaria na rejeição da hipótese de que os elementos intra-grupos são semelhantes.

12 A autora cita ainda o critério CCC (*cubic clustering criterium*) para a definição de g . No entanto, argumenta que esse método não é usual e está disponível em um número reduzido de pacotes estatísticos.

2.2 Análise de Componentes Principais (ACP)¹³

Segundo Kageyama e Leone (1999), este método permite indicar as proximidades entre os indivíduos ou os vínculos entre as variáveis, possibilitando resumir as informações em um conjunto inferior ao original, com a menor perda de generalidade possível.¹⁴ Basicamente, consiste em explicar a estrutura de variância e covariância da matriz de ρ variáveis através da construção de componentes que são combinações lineares entre as variáveis originais.¹⁵ Assim, se existem ρ variáveis, é possível obter ρ componentes principais. Contudo, geralmente busca-se uma redução do número de variáveis originais, e o foco recai sobre os k componentes resultantes (de modo que $k < \rho$) (MINGOTI, 2007). Percebe-se que esse método também possibilita a visualização de variáveis pouco relevantes (menos correlacionadas com as demais). Além disso, cabe ressaltar que a redução de variáveis só é possível quando as ρ variáveis iniciais não são independentes e possuem coeficientes de correlação não nulos (VICINI, 2005).

O resultado dessa combinação gera um número de componentes principais, ortogonais entre si (não correlacionados), exatamente igual ao número de variáveis originais. Uma vez que tais componentes foram obtidos, deve-se selecionar os mais importantes (que expressam boa parte da variância original dos dados). A qualidade da aproximação irá depender do número de componentes principais mantidos no sistema e da proporção da variância total explicada por esses componentes (MINGOTI, 2007).¹⁶ Feito isso, é importante interpretar as principais características de cada componente (o que cada um representa com relação aos dados selecionados) e quais variáveis foram mais importantes para construção de cada um.

- 13 Embora a análise fatorial (AF) seja uma versão mais elaborada da ACP (MINGOTI, 2007; TABACHNICK; FIDELL, 2007), e também pudesse ser utilizada para o mesmo fim, ela requer que alguns pressupostos sejam satisfeitos. Segundo Pereira (2001), a AF requer que as variáveis apresentem distribuição normal. Sendo assim, foi realizado um teste univariado sobre cada uma das variáveis [proposto por D'Agostino *et al.* (1990), com o ajustamento sugerido por Royston (1991)] e quatro testes multivariados sobre a amostra completa (teste abrangente de Doornik e Hansen (2008); teste consistente de Henze e Zirkler (1990); teste da curtose de Mardia (1970); e teste da obliquidade de Mardia (1970)). Em todos os casos, a normalidade foi rejeitada. Logo, optou-se pela ACP, que, segundo Mingoti (2007), não requer nenhuma pressuposição inicial. De qualquer modo, Hair *et al.* (2006) argumentam que a ACP e a AF costumam convergir quando o número de variáveis supera 30.
- 14 O método de componentes principais também é útil na identificação de *clusters*. Segundo Pereira (2001), esse método, quando restrito a duas dimensões, equivale à análise de agrupamentos. Contudo, deve-se estar ciente de que tal restrição no número de componentes pode causar uma perda relevante de informações.
- 15 Bouruche e Saporta (1982) argumentam que a construção dos componentes principais é semelhante à fatorial, pois a redução do número de variáveis não se faz por uma simples seleção de variáveis, mas sim pela construção de novas variáveis sintéticas, obtidas pela combinação linear das variáveis iniciais, por meio dos fatores (ou componentes).
- 16 O primeiro componente principal contém a maior parcela da variância total, enquanto o último apresenta a menor contribuição. Além disso, o somatório de todos os componentes precisa explicar 100% da variância total dos dados.

Formalmente, o método consiste em encontrar as combinações lineares ($b_{i1}, b_{i2}, \dots, b_{ip}$), que acompanharão as variáveis originais (X_1, X_2, \dots, X_p), de forma a produzir um novo conjunto de variáveis não correlacionadas entre si, denominadas componentes principais (CP_1, CP_2, \dots, CP_p). Cada componente gerado é um vetor, disposto de modo a captar o máximo de variância possível do conjunto de variáveis originais (JOHNSON; WICHERN, 2002). Assim:

$$CP_i = b_{i1} X_1 + b_{i2} X_2 + b_{ip} X_p, \text{ sujeito a: } b_{i1}^2 + b_{i2}^2 + \dots + b_{ip}^2 = 1 \quad (1)$$

em que $b_{i1}, b_{i2}, \dots, b_{ip}$ são os loadings associados à cada variável X_i . Em outras palavras, são os elementos do autovetor (b_i), obtido com base no autovalor (λ_i) da matriz de variância-covariância associada ao conjunto de variáveis originais padronizadas (Z_i). Segundo Johnson e Wichern (2002), isso garante que a variância máxima ($\lambda_i = \text{Var}[Z_i]$) é obtida apenas quando $b_i = \lambda_i$. Ou seja, o autovetor normalizado (b_i) corresponde ao autovalor da matriz de covariância amostral (λ_i).

Manly (1986) sintetiza os passos do método de componentes principais da seguinte forma:

1. Inicia-se com a padronização das variáveis originais X_1, X_2, \dots, X_p , para que tenham média zero e desvio padrão um, isto é, $Z_i = [(X_i - \bar{X})/\sigma_i]$;
2. Calcula-se a matriz de correlação;
3. Encontram-se os autovalores ($\lambda_1, \lambda_2, \dots, \lambda_p$) e seus respectivos autovetores (b_1, b_2, \dots, b_p), lembrando que os coeficientes do i -ésimo componente principal são obtidos de b_i , sendo λ_i sua variância; e
4. Descarta-se um número de componentes que permita uma análise dos resultados de modo satisfatório.¹⁷

Segundo Mingoti (2007), deve-se estar atento ao fato de que a ACP é afetada pelas variáveis de maior variância. Assim, a técnica é pouco útil em dados em que há variâncias muito distintas. Por fim, Reis (1997) ressalta que a aplicação da ACP deve conter: a) a descrição das variáveis incluídas na análise; b) o percentual de variância total e individual explicada pelos componentes principais retidos; c) a contribuição de cada variável em cada componente (*factor loadings*), antes e depois de ser aplicado um método de rotação de fatores;¹⁸ e d) a interpretação de cada componente principal retido.

17 Não há uma regra única para definir o número de componentes retidos. Alguns autores se valem da variância explicada por estes componentes. Outros usam o método de Kaiser (1958) (no qual se mantém apenas os autovalores ≥ 1) ou se baseiam em análises gráficas do tipo *screeplot*, idealizada por Cattell (1966). O mesmo vale para a AF.

18 A rotação é controversa. Em particular, o componente rotacionado deixa de apresentar variância máxima (STATA CORP LP, 2009, p. 576). Para Mingoti (2007), se o resultado não rotacionado já apresenta boa qualidade, a rotação poderia piorá-lo. Caso contrário, recomenda-se o uso da rotação ortogonal, pois apenas redistribui a variância explicada pelos *loadings* entre os componentes rotacionados, sem afetar a parcela da variância explicada por esses componentes (MINGOTI,

3 Resultados

Conforme mencionado, as exportações dos 68 países considerados neste artigo apresentaram desagregação para 97 tipos de produtos. Dessa forma, a simples agregação dos países em grupos, conforme os 97 produtos exportados, seria de difícil interpretação. Além disso, o fato de alguns produtos apresentarem correlação próxima à unidade prejudicava o modelo. Logo, o primeiro passo consistiu da agregação dos produtos. Como a AC apresenta maior robustez no caso dos modelos não hierárquicos, devido à interatividade envolvida no processo de estimação (ANDRADE, 2009; LIMA, 2001), optou-se por utilizar o método de K médias para se determinar como os elementos serão agrupados. No entanto, essa técnica requer o conhecimento prévio do número de *clusters* e da definição das respectivas sementes iniciais que serão o ponto de partida para a agregação dos grupos. Logo, realizou-se o procedimento proposto por Mingoti (2007), que consiste em usar o resultado de um agrupamento hierárquico para definir o número ótimo de grupos. Feito isso, utilizou-se a média dos elementos de cada grupo, oriundo da agregação hierárquica, como semente inicial do modelo não hierárquico.

Visando evitar uma agregação arbitrária, cinco técnicas de agrupamento hierárquico foram consideradas: método de Ward, método de centroides, ligação completa, ligação média e ligação simples. Dessa forma, foi possível utilizar o resultado dessas agregações para definir um número ótimo de *clusters* com base em critérios estatísticos (ver Tabela 1).

Os resultados apresentados na Tabela 1 indicam que a agregação da amostra em oito grupos seria uma boa opção, ou seja, pouca informação seria perdida após o agrupamento. Nota-se que outros agrupamentos também poderiam ser escolhidos. No entanto, no caso de oito *clusters*, a média de variação do teste de Calinski/Harabasz (Pseudo-F), considerando os cinco agrupamentos, foi alta e positiva, enquanto a do teste de Duda (Pseudo-T) foi alta e negativa. Isso indica que muita informação seria perdida caso o conjunto de dados sofresse mais uma agregação (se ficasse com sete grupos). Além disso, o efeito negativo de mais uma agregação foi captado por três dos cinco tipos de agrupamento em ambos os testes, fato que confere certa robustez ao resultado.

Uma vez definido o número de *clusters* ($g = 8$), deve-se calcular as sementes que serão utilizadas no modelo não hierárquico. No entanto, os elementos incluídos em cada grupo não são necessariamente os mesmos quando se observam os

2007; SCHNEIDER; WAQUIL, 2001). Já as rotações oblíquas afetam a ortogonalidade entre os componentes e, segundo Johnson (1998), não deveriam ser utilizadas. Para Rencher (2002), serviriam apenas para testar a ortogonalidade dos componentes.

cinco tipos de agrupamentos hierárquicos utilizados. Logo, optou-se por comparar os elementos dos grupos gerados pelo método de Ward e de ligação média.¹⁹

Tabela 1 - Testes para a definição do número ótimo de *clusters* segundo o resultado dos métodos de agrupamentos hierárquicos para os 97 produtos

Número de <i>clusters</i>	Teste de Calinski/Harabasz (Pseudo-F)					Teste de Duda (Pseudo-T)								
	Métodos de agrupamento					Δ Média (%)	Número de Δ (+)	Métodos de agrupamento					Δ Média (%)	Número de Δ (-)
	a	b	c	d	e			a	b	c	d	e		
1	-	-	-	-	-	-	-	55.49	3.70	47.95	31.63	22.66	-	-
2	55.49	3.70	47.95	31.63	22.66	-	-	40.27	6.82	48.66	96.31	25.61	17.62	1
3	46.82	5.38	65.03	87.76	27.08	23.85	4	17.22	2.12	43.62	35.31	48.67	-98.50	4
4	46.28	4.34	79.95	83.71	43.42	5.27	2	16.45	1.45	10.18	29.58	.	-99.69	4
5	45.73	3.64	67.99	92.34	36.01	-9.85	1	12.83	38.70	16.52	35.03	11.47	30.49	1
6	40.78	11.95	60.88	95.44	34.38	8.85	2	10.17	4.39	4.04	22.68	8.24	-242.05	5
7	38.11	10.20	53.35	96.57	32.31	-8.70	1	7.86	60.55	11.08	32.32	13.24	38.90	1
8	36.61	24.75	51.63	109.34	33.35	13.23	3	4.12	.	23.03	9.49	12.69	-70.95	3
9	35.82	21.65	55.60	101.29	34.60	-2.74	2	5.12	5.38	7.25	18.05	5.77	-67.66	2
10	33.36	21.35	54.71	105.18	33.06	-2.27	1	6.26	2.05	8.70	13.70	6.65	-29.22	2
11	33.3	19.86	51.97	102.74	32.36	-3.50	0	3.98	1.86	10.36	22.56	8.51	1.93	2
12	31.81	18.58	52.34	114.5	32.76	0.13	3	5.76	2.01	9.00	19.26	5.65	-8.90	3
13	30.86	17.58	53.84	113.24	32.14	-1.80	1	5.18	1.87	9.08	10.85	4.22	-25.84	4
14	30.01	16.70	53.01	116.8	31.13	-1.97	1	4.61	.	19.13	14.43	4.97	20.02	1
15	29.41	15.44	56.20	117.73	30.64	-1.07	2	5.88	1.30	8.25	15.54	3.79	-33.57	2
16	28.49	14.64	55.69	121.93	29.82	-1.78	1	3.57	6.11	11.64	6.83	9.88	-4.55	2
17	28.23	15.38	59.26	120.22	31.64	2.85	3	4.78	2.70	9.71	26.33	5.06	-28.41	3
18	28.27	14.71	58.52	123.52	31.59	-0.63	2	5.37	1.29	2.48	19.35	12.99	-72.97	3
19	27.95	14.11	57.84	129.51	35.08	1.60	2	8.06	6.76	2.82	5.47	6.34	-46.46	2
20	27.49	13.91	56.54	129.59	35.82	-0.66	2	5.94	2.02	7.96	8.25	15.26	-22.72	2

Fonte: Elaboração própria a partir dos resultados obtidos com o software Stata 11.

Notas: 1) Método de Ward (a), método de centroides (b), ligação completa (c), ligação média (d) e ligação simples (e); 2) Δ Média (%) representa a variação média (percentual), considerando os cinco métodos de agrupamento, entre o valor obtido nos testes para o *cluster* n e $n-1$, $\{[(C_n - C_{(n-1)})/C_n] * 100\}$; 3) Número de Δ (+) é o número de casos, considerando o teste de Pseudo-F sobre os cinco métodos de agrupamento, em que a variação do teste foi positiva; 4) Número de Δ (-) é o número de casos, considerando o teste de Pseudo-T sobre os cinco métodos de agrupamento, em que a variação do teste foi negativa.

Conforme observado na Tabela 6 (ver Anexo), os métodos de Ward (a) e de ligação média (b) produziram grupos com elementos distintos, de modo que, quando são considerados apenas os elementos que permaneceram no mesmo grupo em ambos os agrupamentos (a e b), é possível obter um novo número ótimo de *clusters*. Esse novo agrupamento (c) revelou que existem 18 grupos de produtos que se mantiveram agregados em ambos os métodos (a e b) e 12 produtos que migraram de grupo. Assim, como a intenção é minimizar a perda de informações devido à agregação, os 18 grupos mais os 12 produtos foram utilizados como sementes para a estimação do método K médias (d). Os resultados dos agrupamentos via Ward (a), ligação média (b) e K médias (d) foram combinados com a agregação

19 Segundo Mingoti (2007), o método de ligação média produz partições melhores que a simples e a completa. Afonso e Melão (2007), após testar os cinco agrupamentos, concluíram que apenas os métodos de Ward e de ligação média faziam sentido econômico. Por fim, Lima (2001), realizando testes de Monte Carlo sobre diversas técnicas, percebeu que os melhores resultados (hierárquicos) foram obtidos via método de Ward.

proposta pelo sistema harmonizado - nomenclatura comum do Mercosul (NCM),²⁰ de forma a produzir uma agregação com grupos de produtos homogêneos e coerentes com a NCM. Como resultado, foram criadas 25 cestas de produtos (e).

Logo, a técnica AC permitiu agregar 97 produtos em 25 cestas sem perda significativa de informação e mantendo-se coerente com a NCM. Nota-se que seria possível agrupar os países com base nessas cestas, no entanto analisar o perfil de cada grupo de países de acordo com a distribuição de 25 cestas de produtos pode ser uma tarefa complicada. Assim, utilizando a técnica ACP percebeu-se que uma parcela significativa da variância dessas cestas poderia ser explicada por um número pequeno de componentes. Tanto o critério de Kaiser (1958) (Tabela 2), quanto o de *screeplot* (Gráfico 1) sugerem que apenas três componentes devam ser utilizados. Juntos, estes componentes explicariam mais de 83% da variância das cestas (Tabela 2). Isso não apenas facilitou a análise, uma vez que diminuiu a quantidade de variáveis de 25 para 3, como organizou os elementos (que, na realidade, são as 25 cestas de produtos) de acordo com sua importância nos respectivos componentes (*loadings*).

Tabela 2 - Percentual da variância (σ^2) explicada pelos componentes

Componentes	Autovalor	Proporção	Acumulada
Comp1	15.55	0.62	0.62
Comp2	3.82	0.15	0.78
Comp3	1.30	0.05	0.83
Comp4	0.95	0.04	0.87
Comp5	0.71	0.03	0.89
Comp6	0.55	0.02	0.92
Comp7	0.40	0.02	0.93
Comp8	0.34	0.01	0.94
Comp9	0.27	0.01	0.96
Comp10	0.21	0.01	0.96
Comp11	0.18	0.01	0.97
Comp12	0.15	0.01	0.98
Comp13	0.13	0.01	0.98
Comp14	0.11	0.00	0.99
Comp15	0.08	0.00	0.99

Observações: 68; Variáveis: 25; Traço: 25.

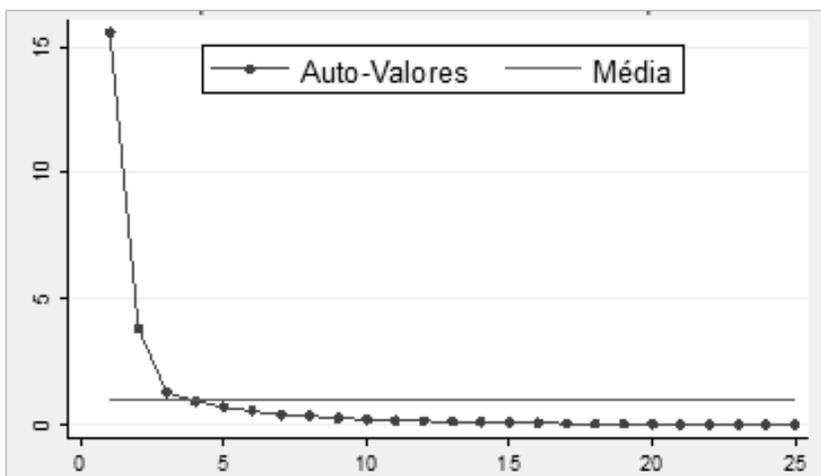
Fonte: Elaboração própria do autor com base nos resultados obtidos no software STATA 11.

Nota: Optou-se por apresentar apenas os 15 componentes mais relevantes gerados pela ACP.

20 Ver Brasil (2012).

Vale ressaltar que, no caso da ACP, dois testes de qualidade do modelo são largamente empregados. O teste de Kaiser, Meyer e Olkin (KMO)²¹ indica a proporção da variância dos dados, que pode ser considerada comum a todas as variáveis (fator comum). Dessa forma, o valor de 0.83 revela que um elevado grau de adequabilidade da amostra ao método. Além disso, o teste de esfericidade de Bartlett rejeita a hipótese de que a matriz de dados seja não correlacionada (fato que inviabilizaria a utilização da ACP) (ver notas do Gráfico 1).

Gráfico 1. Screeplot - Resultado dos Autovalores após a ACP



Fonte: Elaboração própria do autor com base nos resultados obtidos no *software* STATA 11.

Notas: Outros testes de adequação da amostra são:

1. Kaiser-Meyer-Olkin (KMO): 0.83; 2. Esfericidade de Bartlett: $\chi^2=3179.25$; Prob.=0.00.

Na Tabela 3 é apresentada a importância de cada componente na explicação da variância dos dados. É possível perceber que somente o primeiro componente é responsável por 62% da variância das 25 cestas de produtos entre os 68 países. O segundo e o terceiro componentes explicam, respectivamente, 15% e 5%. Portanto, embora todos os componentes busquem identificar os elementos que diferenciam os países, o primeiro é o que apresenta maior peso nessa diferenciação. Dessa forma, o primeiro componente reflete o desempenho global dos países em termos das 25 cestas exportadas, ponderando de modo mais intenso as cestas que têm maior capacidade de diferenciar países. Como os maiores elementos desse componente correspondem à exportação de cimento, metais, máquinas, reatores e equipamentos eletrônicos, pode-se associar esse componente à indústria pesada de produtos siderúrgicos (metal) e minerais (cimento, amianto, etc.) e a produtos intensivos em tecnologia (máquinas, reatores e eletrônicos). O segundo compo-

21 Maiores detalhes sobre esta medida de adequação amostral podem ser obtidos em Kaiser (1974).

nente opõe produtos do setor têxtil (peles e couro, artigos de vestuário e algodão e têxteis) aos produtos alimentícios (frutas, animais vivos, produtos de origem vegetal e animal, bebidas, tabaco, etc.) e farmacêuticos (farmacêuticos e fertilizantes). O terceiro componente divide os países em exportadores de produtos farmacêuticos e exportadores de produtos e combustíveis minerais (ver Tabela 3).

Como cada uma das 25 cestas possui um coeficiente (*loading*) que mensura sua importância nos componentes descritos na Tabela 3, é possível hierarquizar ou classificar os países, em cada componente, através do cálculo de seus scores. Assim, caso o objetivo seja classificar os países de acordo com o componente 1, deve-se ponderar e somar as 25 cestas exportadas por um país j . Os *loadings* obtidos na ACP para o componente 1 devem ser utilizados nas ponderações. Isso produzirá um número (score) que pode ser utilizado para agrupar os países.

Tabela 3 - *Loadings* dos três componentes principais retidos para as 25 cestas de produtos

Número	Cestas de produtos	Cimento, metais, máquinas, reatores e eletrônicos		Setor têxtil vs. produtos alimentícios e farmacêuticos		Produtos e combustíveis minerais vs. farmacêuticos	
		Comp1 (62% de σ^2)	Comp2 (15% de σ^2)	Comp2 (15% de σ^2)	Comp3 (5% de σ^2)	Comp3 (5% de σ^2)	Comp3 (5% de σ^2)
1	Obras de metal e cimento	↑	0.25	↔	-0.02	↘	-0.05
2	Cimento, amianto, cerâmica e vidros	↑	0.24	↘	-0.08	↘	-0.03
3	Máquinas, reatores e eletrônicos	↑	0.24	↔	-0.03	↘	0.00
4	Tecidos	↑	0.23	↓	-0.17	↘	0.03
5	Obras de arte e diversas	↑	0.23	↔	0.02	↘	-0.04
6	Siderúrgicos e pedras preciosas	↑	0.23	↗	0.13	↘	0.02
7	Mobiliário, equipamentos musicais e de relojoaria	↑	0.22	↓	-0.17	↘	-0.04
8	Castaria, filamentos e fibras sintéticas	↑	0.22	↓	-0.19	↘	0.06
9	Brinquedos, jogos e artigos de desporto	↑	0.22	↓	-0.20	↘	0.06
10	Plástico, borracha, celulose, papel e madeira	↑	0.22	↑	0.21	↘	-0.09
11	Algodão e têxteis	↑	0.22	↓	-0.20	↔	0.11
12	Peles e couro	↑	0.21	↓	-0.21	↘	0.05
13	Freios e suas partes	↑	0.21	↓	-0.25	↘	0.08
14	Químicos diversos	↑	0.21	↑	0.23	↓	-0.10
15	Artigos de vestuários	↑	0.21	↓	-0.27	↘	0.07
16	Minerais metálicos	↗	0.20	↑	0.19	↔	0.19
17	Cortiça, tapetes e seda	↗	0.20	↘	-0.12	↘	0.01
18	Material de transporte, armas e equipamentos médicos	↗	0.20	↑	0.23	↓	-0.15
19	Livros, jomais, fotos, etc.	↗	0.19	↑	0.23	↓	-0.20
20	Produtos de origem animal e vegetal, açúcares, café, oleaginosas, cacau, tabaco e bebidas	↗	0.18	↑	0.24	↓	-0.12
21	Farmacêuticos e fertilizantes	↔	0.14	↑	0.28	↓	-0.29
22	Frutas, hortícolas, plantas, cereais, animais vivos e gorduras	↘	0.12	↑	0.29	↘	0.02
23	Cobre e suas obras	↘	0.11	↑	0.18	↘	-0.06
24	<i>Commodities</i> não especificadas	↓	0.07	↑	0.27	↑	0.52
25	Produtos e combustíveis minerais	↓	0.05	↑	0.22	↑	0.68
	Média		0.19		0.03		0.03

Fonte: Elaboração própria a partir dos resultados obtidos com o *software* Stata 11.

Nota: As setas indicam a orientação de cada produto nos componentes. Assim, os maiores *loadings* receberam o sinal ↑, seguido de ↗. Os *loadings* que ficaram próximo da média do componente ou que não se destacaram são indicados por ↔. Os menores coeficientes são representados por ↓, seguidos por ↘. Os elementos hachurados são os que apresentaram coeficiente de maior magnitude entre os componentes.

Vale lembrar que o agrupamento dos países realizado exclusivamente com base nos *scores* obtidos através do componente 1 poderia induzir a resultados distorcidos, uma vez que 38% de todas a variância dos dados estaria sendo ignorada. Para minimizar esse problema, os três componentes devem ser considerados. Assim, o próximo passo consistiu em agrupar os países com base nos *scores* obtidos através dos três componentes. O agrupamento seguiu o mesmo procedimento adotado para os produtos (início da seção 3), ou seja, utilizaram-se os métodos hierárquicos para definir o número ótimo de *clusters* e as sementes iniciais necessárias à estimação do método de *K* médias (não hierárquico).

A análise da média de variação dos coeficientes obtidos nos testes de Calinski/Harabasz (Pseudo-F) e Duda (Pseudo-T), considerando cinco métodos de agrupamentos hierárquicos, sugere que as partições 5, 8, 11 e 16 poderiam ser escolhidas (ver Tabela 4).

Tabela 4 - Testes para a definição do número ótimo de *clusters* segundo o resultado dos métodos de agrupamentos hierárquicos para os 68 países

Número de <i>clusters</i>	Teste de Calinski/Harabasz (Pseudo-F)					Teste de Duda (Pseudo-T)								
	Métodos de agrupamento					Δ Média (%)	Número de Δ(+)	Métodos de agrupamento					Δ Média (%)	Número de Δ(-)
	a	b	c	d	e			a	b	c	d	e		
1	-	-	-	-	-	-	-	46.39	61.72	61.72	61.72	61.72	-	-
2	46.39	61.72	61.72	61.72	61.72	-	-	21.82	68.33	68.33	68.33	68.33	-0.15	1
3	97.00	96.51	96.51	96.51	96.51	0.49	5	26.55	11.49	11.49	11.49	11.49	-3.92	4
4	78.16	78.42	78.42	78.42	78.42	-0.29	0	29.38	.	71.39	62.32	.	0.58	0
5	133.69	59.72	138.41	128.43	59.72	0.15	3	17.00	71.39	7.32	.	6.34	-4.74	2
6	153.03	115.44	127.42	106.75	53.09	0.05	2	17.88	4.24	14.26	9.96	3.56	-4.02	2
7	175.14	106.90	128.07	100.71	46.66	-0.04	2	7.32	3.28	5.02	21.46	60.17	-0.42	3
8	198.53	97.05	116.16	118.02	87.39	0.13	3	7.10	5.02	46.84	4.48	9.86	-1.54	3
9	193.96	89.25	167.57	108.67	89.25	0.03	2	20.93	14.26	.	.	7.77	0.35	1
10	177.95	98.89	167.78	98.89	81.17	-0.05	2	5.07	9.67	30.23	65.14	5.33	-1.35	3
11	183.64	90.09	193.25	183.91	79.32	0.13	3	8.23	7.77	.	7.77	9.82	-1.70	2
12	187.41	84.65	188.63	181.47	84.65	0.00	2	6.07	65.14	6.38	9.67	23.36	0.33	1
13	194.18	174.39	200.78	174.39	112.04	0.20	4	.	.	7.77	7.10	.	-0.09	1
14	227.99	160.81	208.38	195.05	102.70	0.03	3	8.05	34.22	9.67	8.23	34.22	0.17	0
15	239.25	150.38	208.12	216.39	95.19	0.00	2	6.43	32.63	32.63	32.63	4.01	-1.28	3
16	266.87	168.71	254.40	267.95	94.48	0.14	4	.	4.33	16.99	5.78	5.78	-2.95	3
17	301.72	186.07	301.64	274.10	90.24	0.09	4	17.06	4.17	5.78	4.35	14.04	-0.43	3
18	312.41	205.64	317.47	294.22	109.29	0.11	5	8.75	8.23	8.75	14.52	.	0.15	1
19	339.09	257.97	345.53	345.53	102.09	0.11	4	9.67	.	14.46	34.22	.	0.36	0
20	399.44	244.98	386.98	348.75	95.59	0.04	3	5.78	4.35	34.22	.	22.01	-0.05	1

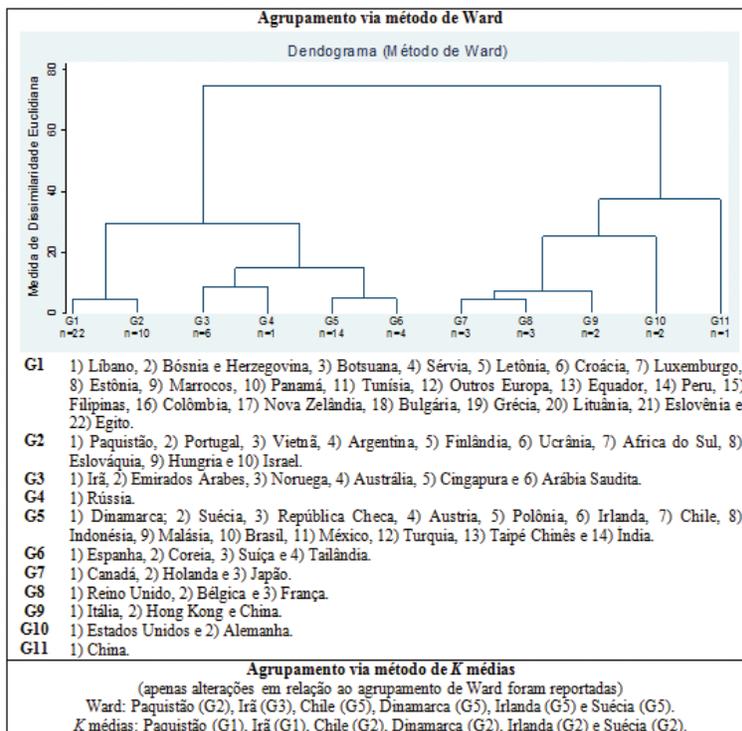
Fonte: Elaboração própria a partir dos resultados obtidos com o *software* Stata 11.

Notas: 1) Análise análoga a da Tabela 1; 2) como o objetivo era obter um número (g) reduzido de grupos, foram reportados apenas os casos em que $g \leq 20$.

Logo, visando evitar a perda de informações devido ao excesso de agregações, excluíram-se os agrupamentos com cinco e oito partições. O agrupamento com onze partições foi escolhido com base no critério da parcimônia (seria difícil visualizar diferenças significativas entre um número elevado de grupos). Conforme realizado anteriormente, optou-se pela utilização dos métodos de Ward e de ligação média para definir os elementos que irão servir de base à construção das sementes necessárias ao método não hierárquico. Contudo, o método de ligação média produziu partições de pouco valor analítico (seis dos onze grupos continham apenas um país) e o método de Ward acabou sendo escolhido para definição das sementes.

A análise do dendograma apresentado no Quadro 1 revela que, dos onze grupos gerados via método de Ward, dois são compostos por apenas um país.

Quadro 1 - Agrupamento dos 68 países através das técnicas de Ward (hierárquica) e de K médias (não hierárquica)



Fonte: Elaboração própria a partir dos resultados obtidos com o *software* Stata 11.

Notas: As médias dos 11 grupos gerados via agrupamento hierárquico de Ward foram utilizadas como sementes para o método não hierárquico de K médias.

No caso mais extremo tem-se a China (G11), país que havia apresentado o maior score referente ao primeiro componente da ACP. Embora a Rússia (G4) também esteja sozinha em seu agrupamento, não deve ser considerada um *outlier*, uma vez que seu perfil exportador está classificado entre o G3 e o G5. Essa é uma característica interessante dos dendogramas, pois possibilitam visualizar a proximidade entre os grupos. Assim, é possível perceber que as exportações brasileiras (G5) estão mais próximas do G6 (Coreia, Tailândia, Espanha e Suíça) do que da Rússia (G4) e estão distantes da China (G11), Estados Unidos e Alemanha (pertencem ao G10). A utilização das médias desses grupos como semente para o método de K médias corroborou a robustez do agrupamento. Dos 68 países, apenas seis foram realocados em grupos diferentes após a utilização do método não hierárquico. São eles: Paquistão, Irã, Chile, Dinamarca, Irlanda e Suécia. No caso de Paquistão e Irã, o método de K médias indicou que ambos deveriam ser agrupados com

o G1, juntamente com Líbano, Bósnia, Egito, etc. Já o Chile, Dinamarca, Irlanda e Suécia deveriam pertencer ao grupo de Portugal, Argentina, África do Sul, etc. (G2) (ver Quadro 1).

A análise das características do agrupamento gerado via K médias (ver Tabela 5) revela que o componente 1 foi responsável pela maior parte do agrupamento. Para se verificar isso, basta observar as médias dos scores obtidos por cada grupo. É possível perceber que, no caso do componente 1, são raras as ocasiões em que o ordenamento é alterado entre os grupos. Em outras palavras, a média de $G1 < G2... < G11$ em quase todos os casos. O G4, representado pela Rússia, obteve uma média expressiva no componente 2 e a maior das médias no componente 3. Isso ocorre porque esse país foi o maior exportador de produtos e combustíveis minerais (Cesta H) e outras *commodities* (Cesta I) entre os 68 países. No caso da cesta H, pode-se destacar os produtos combustíveis minerais, petróleo, produtos de destilação, etc. (P27) (ver Tabela 6, no Anexo), uma vez que a Rússia foi a maior exportadora desses produtos no período. No caso do G11 (China), a análise das médias normalizadas das cestas exportadas sugere que esse país é um *outlier*. Segundo a Tabela 5, das nove cestas mais importantes para definição dos grupos, a China foi a maior exportadora em cinco.

A análise do dendograma revelou certa proximidade entre a Rússia (G4) e os países do G3. O fato é que, em média, tanto a Rússia quanto os países do G3 são exportadores intensivos da cesta H e apresentam baixo desempenho na maioria das outras cestas. Nesse caso, os produtos exportados em comum são petróleo e derivados (note de que o G3 é composto por dois grandes exportadores desse produto: Arábia Saudita, e Emirados Árabes).²² As pequenas médias obtidas nas cestas exportadas pelo G1 parecem indicar que esse grupo é composto por países pouco representativos em termos de fluxo comercial para o mercado internacional. Além disso, nota-se que o resultado do componente 2 foi o grande responsável pela separação do G11 e do G10, pois, enquanto a China (G11) obteve o menor valor, os Estados Unidos e a Alemanha (G10) apresentaram o maior. Isso indica que a China é intensiva em produtos têxteis (Cesta F e G), enquanto a Alemanha e os Estados Unidos são mais voltados para a exportação de produtos alimentícios e farmacêuticos (Cestas D e E) (ver Tabela 5).

A Tabela 5 ainda revela que o grupo em que se encontra o Brasil (G5) poderia ser classificado como um exportador de médio porte. As exportações desse grupo estão longe do montante exportado pelo G11, G10, G9, G8 e, até mesmo, pelo G7. O G5 parece só ter capacidade de ameaçar esses grupos no que se refere às exportações de frutas, hortícolas, plantas, cereais, animais vivos e gorduras (cesta D) e produtos e combustíveis minerais (cesta H). Embora o G4 e o G3 estejam próximos do grupo brasileiro, eles parecem concentrar boa parte de suas expor-

22 Na técnica de Ward, o Irã também faria parte desse grupo. No entanto, a análise da Tabela 5 está baseada no resultado via K médias, no qual o Irã faz parte do G1.

tações, enquanto as exportações do G5 e do G6 são mais diversificadas. Assim, é mais provável que a competição por espaço no mercado internacional seja mais intensa entre o G4 e o G3 e entre o G5 e o G6. Outra característica interessante sobre o grupo brasileiro (G5) refere-se à elevada variância verificada para os scores dos países desse grupo em todos os três componentes. Esse poderia ser um indício de que o G5 é o menos homogêneo entre os grupos selecionados. De qualquer modo, pode ser difícil para o Brasil fechar acordos bilaterais com países incluídos no seu agrupamento (e mesmo em agrupamentos próximos). O fato é que esses países concorrem diretamente com o Brasil. Portanto, seria mais interessante buscar acordos comerciais com países que pudessem complementar a produção brasileira (agrupamentos distantes) e não ameaçá-la.

Com objetivo de entender melhor os grupos criados, algumas características macroeconômicas são apresentadas na Tabela 5. É possível perceber que o grupo de maior crescimento do PIB *per capita*, entre 2005 e 2011, foi o G11 (China), seguido do G4 (Rússia) e do G5 (Brasil e outros). Ao que parece, os grupos iniciais (G1, G2, G3, G4, G5, G6) obtiveram um crescimento maior que os finais (G7, G8, G9, G10).

Tabela 5 - Características dos grupos gerados pela técnica de análise de *clusters* (K médias)

Grupos	Componentes (médias)			Média das exportações normalizadas (máx.: 1 e mín.: 0): apenas as mais relevantes nos três componentes								
	comp1	comp2	comp3	Cesta A	Cesta B	Cesta C	Cesta D	Cesta E	Cesta F	Cesta G	Cesta H	Cesta I
G1	-2.05	-0.76	-0.17	0.00	0.01	0.00	0.02	0.01	0.01	0.00	0.03	0.01
G2	-1.36	-0.32	-0.39	0.02	0.02	0.02	0.07	0.08	0.02	0.01	0.03	0.01
G3	-1.12	0.57	1.80	0.02	0.01	0.05	0.05	0.04	0.00	0.00	0.51	0.14
G4	0.05	3.42	6.29	0.06	0.02	0.01	0.12	0.14	0.00	0.00	1.00	1.00
G5	-0.14	-0.16	-0.02	0.08	0.07	0.07	0.16	0.04	0.03	0.01	0.12	0.01
G6	0.76	0.58	-0.79	0.07	0.08	0.10	0.13	0.24	0.03	0.02	0.09	0.01
G7	2.83	2.51	0.84	0.14	0.15	0.22	0.29	0.19	0.03	0.04	0.31	0.27
G8	3.08	2.17	-1.23	0.12	0.14	0.11	0.22	0.54	0.06	0.06	0.17	0.05
G9	3.43	-1.16	-0.25	0.18	0.20	0.25	0.11	0.15	0.16	0.10	0.05	0.07
G10	11.24	6.60	-1.34	0.56	0.39	0.49	0.64	0.81	0.08	0.09	0.31	0.40
G11	23.82	-9.49	1.45	1.00	1.00	1.00	0.14	0.19	1.00	1.00	0.13	0.04
Componentes (variância)				Características macroeconômicas de cada grupo (média)								
Grupos	comp1	comp2	comp3	PIB <i>per capita</i>	Δ PIB <i>per capita</i>	Pop.	Δ Pop.	Conta Corrente.	Δ Conta Corrente.	(Conta Corr.) PIB	Δ (Conta Corr.) PIB	
G1	-0.07	-0.23	-1.46	16.68	31.29	26.86	5.18	-0.74	5.19	-2.78	-36.99	
G2	-0.20	-1.74	-0.79	23.61	28.11	23.26	4.68	1.24	-95.39	-0.24	-90.35	
G3	-0.69	1.06	0.34	45.16	17.94	13.30	18.12	53.81	48.18	13.60	-21.54	
G4	-	-	-	16.74	40.85	142.41	-0.27	101.14	19.77	5.47	-50.57	
G5	-2.85	-4.05	-22.97	17.12	37.67	212.91	6.01	-19.46	74.25	-1.10	13.16	
G6	0.69	0.44	-0.85	28.78	25.82	41.77	4.27	18.03	-44.13	4.03	-55.37	
G7	0.44	0.19	0.81	39.15	16.63	59.65	3.08	44.79	-107.61	2.24	-97.13	
G8	0.18	0.09	-0.26	36.33	14.96	45.56	3.98	-36.37	122.44	-1.43	76.00	
G9	0.40	-1.02	-0.15	39.80	24.03	33.89	3.60	-30.00	163.80	0.48	112.64	
G10	0.01	0.19	-1.77	43.14	19.45	196.86	2.26	-134.00	4.98	1.30	-16.71	
G11	-	-	-	8.38	104.32	1348.12	3.10	2.75	-53.65	201.00	49.89	
Média total (apenas macro)				24.29	29.46%	80.87	5.72%	0.33	-33.51%	2.49	0.96%	

Fonte: Elaboração própria a partir dos resultados obtidos com o *software* Stata 11 e de dados do International Monetary Fund (2015).

Notas: O símbolo Δ representa a variação entre o período de 2005 e 2011. As cestas de produtos são: obras de metal e cimento (A); cimento, amianto, cerâmica e vidros (B); máquinas, reatores e eletrônicos (C); frutas, hortícolas, plantas, cereais, animais vivos e gorduras (D); farmacêuticos e fertilizantes (E); artigos de vestuários (F); freios e suas partes (G); produtos e combustíveis minerais (H); e *commodities* não especificados (I).

Embora a China esteja mais próxima dos grupos de menor crescimento, ela apresenta características de *outlier* e não deve ser comparada aos mesmos. No caso dos grupos de maior crescimento, o G3 (Austrália, Arábia Saudita, Emirados Árabes, entre outros) se destaca pelo fraco desempenho econômico. Uma das possíveis causas pode estar associada ao elevado crescimento populacional verificado no período (Tabela 5). Conforme relembra Mankiw, Romer e Weil (1992), altas taxas de crescimento populacional costumam prejudicar o crescimento do PIB *per capita*.²³

4 Considerações Finais

Este trabalho teve como objetivo definir os principais nichos de mercado em termos de comércio internacional. Para tanto, as técnicas AC e ACP foram utilizadas no intuito de organizar os países de acordo com seu perfil exportador. Inicialmente, os dados apresentavam desagregação para 68 países e 97 produtos. Assim, a agregação dos produtos através da AC e, posteriormente, da ACP revelou que o principal componente de diferenciação entre as cestas exportadas pelos países está associado à exportação de produtos siderúrgicos (metal) e minerais (cimento, amianto, etc.) e a produtos intensivos em tecnologia (máquinas, reatores e eletrônicos). Somente o componente mencionado explicaria mais de 62% da variância total dos dados. Além desse, foram considerados dois outros componentes, que juntos explicam, aproximadamente, 20% da variância dos dados. Um deles opõe produtos do setor têxtil (peles e couro, artigos de vestuário e algodão e têxteis) aos produtos alimentícios (frutas, animais vivos, produtos de origem vegetal e animal, bebidas, tabaco, etc.) e farmacêuticos (farmacêuticos e fertilizantes). O outro divide os países em exportadores de produtos farmacêuticos e exportadores de produtos e combustíveis minerais.

Os scores gerados através dos três componentes selecionados foram utilizados para agrupar os países através do método de AC. Os primeiros resultados indicaram que 11 grupos podiam ser destacados. No caso mais extremo, tem-se a China (G11), país que havia apresentado o maior score referente ao primeiro componente da ACP. Outro país que ficou sozinho em seu grupo foi a Rússia (G4). Todavia, não deve ser considerada um *outlier*, uma vez que seu perfil exportador está classificado entre o G3 e o G5. As exportações brasileiras (G5) estão mais próximas do G6 (Coreia, Tailândia, Espanha e Suíça) do que da Rússia (G4) e estão distantes da China (G11), Estados Unidos e Alemanha (sendo os dois últimos pertencentes ao G10).

23 Cabe lembrar que os métodos utilizados neste artigo não estabelecem uma relação de causalidade (são apenas descritivos). Dessa forma, não é possível afirmar que o crescimento superior ou o melhor desempenho em transações correntes de um determinado grupo seja explicado pelo tipo de produto exportado por esse grupo.

Com relação ao tipo de produto exportado pelos grupos, tem-se que o G4, representado pela Rússia, foi o maior exportador de produtos e combustíveis minerais e outras *commodities* entre os 68 países (com destaque para a exportação de combustíveis minerais, petróleo, produtos de destilação, etc.). No caso do G11, composto unicamente pela China, das nove cestas mais importantes para definição dos grupos, a China foi a maior exportadora em cinco. O G3 (Arábia Saudita, Emirados Árabes, entre outros) apresentou certa proximidade com a Rússia (G4). Um dos fatores que contribui para esse resultado é o de que ambos (G3 e G4) são exportadores intensivos em petróleo e derivados e apresentam baixo desempenho na maioria das outras cestas. As pequenas médias obtidas nas cestas exportadas pelo G1 (Líbano, Marrocos, Tunísia, entre outros) parecem indicar que esse grupo é composto de países pouco representativos em termos de fluxo comercial para o mercado internacional. Além disso, o resultado do componente 2 parece ter sido o grande responsável pela separação do G11 e G10. Enquanto a China (G11) é intensiva em produtos têxteis, a Alemanha e os Estados Unidos (G10) são mais voltados para a exportação de produtos alimentícios e farmacêuticos.

O grupo em que se encontra o Brasil (G5) poderia ser classificado como um exportador de médio porte, sendo responsável por uma parcela representativa da exportação de frutas, hortícolas, plantas, cereais, animais vivos e gorduras e produtos e combustíveis minerais. Embora o G4 e o G3 estejam próximos do grupo brasileiro, parecem concentrar boa parte de suas exportações, enquanto as exportações do G5 e do G6 (Espanha, Coreia, Suíça e Tailândia) são mais diversificadas. Assim, é mais provável que a competição por espaço no mercado internacional seja mais intensa entre o G4 e o G3 e entre o G5 e o G6. Outra característica interessante sobre o grupo brasileiro (G5) refere-se à elevada variância verificada para os *scores* dos países desse grupo em todos os três componentes. Isso poderia ser um indício de que o G5 é o menos homogêneo entre os grupos selecionados. Apesar de ser o mais heterogêneo dos grupos, os resultados indicam que seria difícil para o Brasil fechar acordos bilaterais com países incluídos no seu agrupamento (e mesmo em agrupamentos próximos). O fato é que países como Índia, México, Malásia, Turquia e os demais incluídos no agrupamento brasileiro parecem concorrer diretamente com o Brasil. Logo, seria mais interessante buscar acordos comerciais com países que pudessem complementar a produção brasileira (agrupamentos distantes) sem ameaçá-la.

Os novos poços de petróleo encontrados no território brasileiro, denominados pré-sal, podem fazer com que o país se especialize ainda mais na produção de petróleo e derivados. Tal fato poderia afetar a pauta de exportações brasileira tornando-a menos diversificada no futuro. Nesse caso, o Brasil caminharia em direção à Rússia (G4). Caso queira ir em direção à China (G11), Estados Unidos e Alemanha (G10), o país precisaria incentivar o setor de alta tecnologia a fim de

ganhar espaço no mercado externo de máquinas, reatores e eletrônicos (cesta C). Precisaria, ainda, investir na área de bioquímica, criando laboratórios e formando mão de obra capaz de inovar e produzir novas patentes. Caso contrário, dificilmente conseguiria adentrar no mercado de produtos farmacêuticos e fertilizantes (cesta E). Além disso, os pedidos de proteção comercial realizados por empresas brasileiras fornecem alguns indícios interessantes. Os dados de Vasconcelos e Firme (2011, p. 172) indicam que mais de 50% dos casos *antidumping* (principal medida de proteção comercial) iniciados no país tiveram como alvo os setores de metais comuns e suas obras (cesta A) e a indústria química (na qual a cesta E está incluída). O setor de máquinas, aparelhos e material elétrico (incluído na cesta C) também se destaca pelo excesso de pedidos de proteção e aparece como o quarto setor mais demandante. Uma vez que Araújo Jr. *et al.* (2001) *apud* Firme e Vasconcelos (2012, p. 262) revelam que essas medidas costumam “desviar o foco da falta de competitividade interna”, esse seria um indício da fragilidade desses setores ao enfrentar a concorrência externa. Na outra ponta encontram-se os produtos do reino animal e vegetal e os demais produtos alimentícios, bebidas e fumo (equivalente à cesta D), em que os pedidos de proteção não chegam a 10% do total, apesar do elevado número de empresas atuando nesses segmentos. O fato é que, conforme demonstrado neste artigo, as exportações brasileiras desse segmento parecem se destacar. Logo, não haveria necessidade de pedidos de proteção.

As características macroeconômicas dos grupos revelam que os grupos iniciais (G1, G2, G3, G4, G5, G6) obtiveram um crescimento do PIB *per capita* maior que os finais (G7, G8, G9, G10) entre 2005 e 2011. Embora a China (G11) esteja mais próxima dos grupos de menor crescimento, apresenta características de *outlier* e não deve ser comparada com eles. No entanto, como os métodos utilizados neste artigo não estabelecem uma relação de causalidade (são apenas descritivos), não é possível afirmar que o crescimento superior de um determinado grupo seja explicado pelo tipo de produto exportado por esse grupo.

Para estudos futuros sugere-se estimar algum modelo causal que verifique a influência desses grupos no desempenho econômico dos países e realizar o mesmo procedimento de agrupamento em um período diferente para verificar se os grupos se mantêm. Neste caso, haveria um indício de barreira à entrada e saída de novos membros nos grupos.

Referências

AFONSO, M. A. D.; MELÃO, N. F. R. Para uma tipologia socioeconômica da área metropolitana do Porto: uma análise estatística multivariada. *Tékhnē - Polytechnical Studies Review*, v. 5, n. 8, p. 215-242, 2007.

ANDRADE, C. M. C. *Crédito e crescimento econômico: uma análise da relevância dos tipos de crédito no Brasil*. 2009. 139 f. Dissertação (Mestrado em Economia) - Centro

de Desenvolvimento e Planejamento Regional, Faculdade de Ciências Econômicas, Universidade Federal de Minas Gerais, Belo Horizonte, 2009.

AW, B. Y.; HWANG, A. Productivity and the export market: a firm-level analysis. *Journal of Development Economics*, v. 47, n. 2, p. 313-332, Aug. 1995.

BALASSA, B. Outward orientation. In: CHENERY, H.; SRINIVASAN, T. N. (Ed.). *Handbook of Development Economics*. Amsterdam: North Holland, 1989. v. 2. cap. 31. p. 1645-1689.

BEZDEK, J. C. *Pattern recognition with fuzzy objective function algorithms*. New York: Plenum Press, 1981.

BOUROCHE, J. M.; SAPORTA, G. *Análise de dados*. Rio de Janeiro: Zahar, 1982.

BRASIL. Ministério da Fazenda. Secretaria da Receita Federal. *Tabela de incidência do imposto sobre produtos industrializados (TIPI)*. 2012. Disponível em: <www.receita.fazenda.gov.br/publico/tipi/tipi.doc>. Acesso em: 02 jun. 2015.

CALINSKI, T.; HARABASZ, J. A dendrite method for cluster analysis. *Communications in Statistics*, v. 3, n. 1, p. 1-27, 1974.

CATTEL, R. B. The screen test for the number of factors. *Multivariate Behavior Research*, v. 1, n. 5, p. 140-161, Jan. 1966.

D'AGOSTINO, R. B.; BELANGER, A. J.; D'AGOSTINO JR., R. B. A suggestion for using powerful and informative tests of normality. *American Statistician*, v. 44, n. 4, p. 316-321, 1990.

DELGADO, M. A.; FARIÑAS, J. C.; RUANO, S. Firm productivity and export markets: a non-parametric approach. *Journal of International Economics*, v. 57, n. 2, p. 397-422, 2002.

DOORNIK, J. A.; HANSEN, H. An omnibus test for univariate and multivariate normality. *Oxford Bulletin of Economics and Statistics*, v. 70, n. 1, p. 927-939, 2008.

DUDA, R. O.; HART, P. E. *Pattern classification and scene analysis*. New York: John Wiley & Sons, 1973.

EATON, J.; KORTUM, S.; KRAMARZ, F. Dissecting trade: firms, industries, and export destinations. *American Economic Review*, v. 94, n. 2, p. 150-154, 2004.

FAN, J. X. *et al.* Household food expenditure patterns: a cluster analysis. *Monthly Labor Review*, v. 130, n. 4, p. 38-51, Apr. 2007.

FERRAZ, J.; KUPFER, D.; HAGUENAUER, L. *Made in Brazil: desafios competitivos para a indústria*. Rio de Janeiro: Campus, 1997.

FERREIRA, P. C.; ROSSI, J. L. New evidence from Brazil on trade liberalization and productivity growth. *International Economic Review*, v. 44, n. 4, p. 1383-1405, Nov. 2003.

FIRME, V. A. C.; VASCONCELOS, C. R. F. Impactos de medidas "antidumping" adotadas pelos EUA sobre o setor siderúrgico de Minas Gerais e o restante do Brasil. *Revista Nova Economia*, Belo Horizonte, v. 22, n. 2, p. 261-302, mar./ago. 2012.

GOMES, V.; ELLERY JR., R. G. Perfil das exportações, produtividade e tamanho das firmas no Brasil. *Revista Brasileira de Economia*, v. 61, n. 1, p. 33-48, jan./mar. 2007.

HAIR, J. H. et al. *Multivariate data analysis*. 6. ed. New Jersey: Prentice Hall, 2006.

HAY, D. A. The post-1990 Brazilian trade liberalisation and the performance of large manufacturing firms: productivity, market share and profits. *The Economic Journal*, v. 111, n. 473, p. 620-641, July. 2001.

HENZE, N.; ZIRKLER, B. A class of invariant consistent tests for multivariate normality. *Communications in Statistics - Theory and Methods*, v. 19, n. 10, p. 3595-3617, 1990.

HIDALGO, A. B.; MATA, D. Produtividade e desempenho do exportador das firmas na indústria de transformação brasileira. *Estudos Econômicos*, São Paulo, v. 39, n. 4, p. 709-735, out./dez. 2009.

INTERNATIONAL MONETARY FUND. *International monetary fund: data and statistics*, 2015. Disponível em: <<http://www.imf.org/external/data.htm>>. Acesso em: 10 de dezembro de 2013.

INTERNATIONAL TRADE CENTRE. *International trade centre*, 2015. Disponível em: <<http://www.intracen.org>>. Acesso em: 10 de dezembro de 2013.

JOHNSON, D. E. *Applied multivariate methods for the data analysis*. New York: Duxbury Press, 1998.

JOHNSON, R. A.; WICHERN, D. W. *Applied multivariate statistical analysis*. New Jersey: Prentice-Hall, 2002.

KAGEYAMA A.; LEONE, E. T. *Uma tipologia dos municípios paulistas com base em indicadores sociodemográficos*. Campinas: IE/UNICAMP, jan. 1999. (Texto para discussão, n. 66).

KAISER, H. F. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, v. 23, p. 187-200, 1958.

KAISER, H. F. An index of factor simplicity. *Psychometrika*, v. 39, p. 31-36, 1974.

LIMA, J. O. *Uma comparação do método Fuzzy e redes neurais artificiais com os procedimentos de agrupamentos hierárquicos e não hierárquicos tradicionais*. 2001. 60 f. Dissertação (Mestrado em Estatística) - Programa de Pós-Graduação em Estatística, Faculdade de Estatística, Universidade Federal de Minas Gerais, Belo Horizonte, 2001.

MAHALANOBIS, P. C. On the generalized distance in statistics. *Proceedings of the National Institute of Science of India*, v. 12, n. 1. p. 49-55, Apr. 1936.

MANKIW, N. G.; ROMER, D.; WEIL, D. A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, v. 107, n. 2, p. 407-437, May. 1992.

MANLY, B. J. F. *Multivariate statistical methods: a primer*. London: Chapman and Hall, 1986.

MARDIA, K. V. Measures of multivariate skewness and kurtosis with applications. *Biometrika*, v. 57, n. 3, p. 519-530, 1970.

MASCARENHAS, B.; AKER, D. A. Mobility barriers and strategic groups. *Strategic Management Journal*, v. 10, p. 475-485, 1989.

MINGOTI, S. A. *Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada*. Belo Horizonte: UFMG, 2007. (Coleção Didática).

NORTH, D. Teoria de localização e crescimento econômico regional. In: SCHWARTZMAN, J. (Org.). *Economia regional: textos selecionados*. Belo Horizonte: Cedeplar, 1977.

PEREIRA, J. C. R. *Análise de dados qualitativos: estratégias metodológicas para as ciências da saúde, humanas e sociais*. 3. ed. São Paulo: Universidade de São Paulo, 2001.

PORTER, M. E. *Vantagem competitiva*. Rio de Janeiro: Campus, 1989.

PORTER, M. The structure within industries and companies performance. *The Review of Economics and Statistics*, v. 61, n. 2, p. 214-217, May 1979.

REIS, E. *Estatística multivariada aplicada*. Lisboa: Edições Sílabo, 1997.

RENCHER, A. C. *Methods of Multivariate Analysis*. 2. ed. New York: Wiley, 2002.

ROYSTON, P. An improved D'Agostino test. *Stata Technical Bulletin* 3: 23-24, v. 1, p. 110-112. Texas: Stata Press, 1991. (Reprinted in Stata Technical Bulletin Reprints).

SCHNEIDER, S.; WAQUIL, P. D. Caracterização sócio-econômica dos municípios gaúchos e desigualdades regionais. *Revista de Economia e Sociologia Rural*, Brasília, DF, v. 39, n. 3, p. 117-142, 2001.

SILVA, L.; SIMÕES, R. Oportunidades tecnológicas e produção científica: uma análise microrregional para o Brasil. *Revista EURE*, Santiago de Chile, v. 30, n. 90, p. 85-102, 2004.

SIRKIN, G. The theory of the regional economic base. *The Review of Economics and Statistics*, v. 41, n. 4, p. 426-429, Nov. 1959.

SOUZA, N. J. Conceito e aplicação da teoria da base econômica. *Perspectiva Econômica*, v. 10, n. 25, p. 117-130, mar. 1980.

STATACORP LP. *STATA multivariate statistics reference manual: release 11*. Stata Press: Texas, 2009.

TABACHNICK, B.; FIDELL, L. S. *Using multivariate analysis*. Boston: Allyn & Bacon, 2007.

TIEBOUT, C. As exportações e o crescimento econômico regional. In: SCHWARTZMAN, J. *Economia regional: textos escolhidos*. Belo Horizonte: Cedeplar, 1977. p. 315-323.

TIMM, N. H. *Applied multivariate analysis*. New York: Springer, 2002.

VASCONCELOS, C. R. F.; FIRME, V. A. C. Efetividade do instrumento antidumping no Brasil entre 1990 e 2007. *Revista Economia*, Brasília, DF, v. 12, p. 165-184, 2011.

VICINI, L. *Análise multivariada da teoria à prática*. Santa Maria: Departamento de Estatística, 2005.

Anexo

Tabela 6 - Agregação dos 97 produtos em 25 cestas de acordo com os métodos de Ward, ligação média e K médias

Produtos (97)	a	b	c	d	e	Agregação por grupos (25)
P08 Frutas frescas, nozes, cascas de cítricos, melão	1	1	1	19	1	Frutas, hortícolas, plantas, cereais, animais vivos e gorduras.
P01 Animais vivos	1	1	1	21	1	
P15 Animais, gorduras e óleos vegetais, produtos de clivagem, etc.	1	1	1	24	1	
P06 Árvores vivas, plantas, raízes, bulbos, flores de corte, etc.	1	1	1	26	1	
P14 Matérias para entrançar, produtos hortícolas e outros.	1	1	1	26	1	
P10 Cereais	2	1	2	2	1	
P11 Produtos da moagem, malte, amidos, insulina, glúten de trigo	2	1	2	21	1	
P13 Gomas, resinas, sucos, extratos vegetais e outros.	2	1	2	26	1	
P27 Combustíveis minerais, petróleo, produtos de destilação, etc.	3	2	3	3	2	Produtos e combustíveis minerais
P25 Sal, enxofre, terras, pedras, gesso, cal e cimento	3	2	3	5	2	
P26 Minérios, escórias e cinzas	3	2	3	17	2	
P28 Produtos químicos inorgânicos, componentes de metais preciosos, isótopos.	3	2	3	5	3	Químicos diversos
P32 Curtimento, tingimento, extratos taninos, pigmentos e derivados	3	2	3	11	3	
P34 Pastas de sabões, lubrificantes, ceras, velas, modelagem	3	2	3	11	3	
P35 Albuminoides, amidos modificados, colas, enzimas	3	2	3	21	3	
P36 Explosivos, pirotecnia, fósforos, pirofóricos, etc.	3	2	3	26	3	

continua...

continuação...

Produtos (97)	a	b	c	d	e	Agregação por grupos (25)
P29 Produtos químicos orgânicos	3	2	3	27	3	
P33 Óleos essenciais, perfumes, cosméticos, artigos de higiene pessoal	3	2	3	28	3	
P38 Produtos químicos diversos	3	2	3	28	3	
P37 Produtos para fotografia e cinematografia	3	4		26	3	
P39 Plásticos e suas obras	3	2	3	14	4	Plástico, borracha, celulose, papel e madeira
P47 Pastas de madeira, material celulósico fibroso, resíduos, etc.	3	2	3	21	4	
P44 Madeira e obras de madeira, carvão vegetal de madeira	3	2	3	23	4	
P40 Borracha e suas obras	3	2	3	30	4	
P48 Papel e cartão, artigos de celulose, papel e cartão	6	2	4	4	4	
P41 Couros e peles (exceto peles com pelo)	3	2	3	21	5	Peles e couro
P42 Artigos de couro, tripa, arreios, artigos de viagem	3	2	3	21	5	
P43 Peles com pêlo, peles artificiais e manufaturadas.	3	2	3	26	5	
P56 Estofa, feltros, falsos tecidos, fios, cordéis, etc.	4	3	5	21	6	Tecidos
P59 Tecidos impregnados, revestidos ou tecido laminado	4	3	5	21	6	
P58 Tecido especial ou tecido acolchoado, rendas, tapeçaria, etc.	4	3	5	26	6	
P60 Tecidos de malha	4	3	5	26	6	
P51 Lã, pêlos, fios de crina e seu tecido	6	2	4	26	6	
P68 Produtos de pedra, gesso, cimento, amianto, mica, artigos, etc.	6	3	6	11	7	Cimento, amianto, cerâmica e vidros
P69 Produtos cerâmicos	6	6	12	11	7	
P70 Vidro e suas obras	6	6	12	11	7	

continua...

continuação...

Produtos (97)	a	b	c	d	e	Agregação por grupos (25)
P84 Máquinas, reatores nucleares, caldeiras, etc.	7	3		9	8	Máquinas, reatores e eletrônicos
P85 Produtos elétricos, equipamento eletrônico	7	5	10	10	8	
P52 Algodão	5	4	7	7	9	Algodão e têxteis
P63 Outros artigos têxteis manufaturados, roupas usadas, etc.	5	4	7	7	9	
P54 Filamentos sintéticos	5	4	7	21	10	Cestaria, filamentos e fibras sintéticas
P46 Manufaturas de material para entrançar, cestaria, etc.	5	4	7	26	10	
P55 Man-made fibras descontínuas	5	6		21	10	
P64 Calçados, polainas e artefatos semelhantes e suas partes	6	4	8	8	11	Artigos de vestuários
P61 Artigos de vestuário, acessórios, tricô ou crochê	6	4	8	12	11	
P67 Pele de pássaros, penas, flores artificiais, cabelo humano	6	4	8	26	11	
P62 Artigos de vestuário, acessórios (tricô e crochê exclusivos)	6	6	12	12	11	
P53 Produtos vegetais, fibras têxteis, fios de papel, tecido e outros.	6	6	12	26	11	
P66 Guarda-chuvas, bengalas, bengalas-assentos, chicotes, etc.	6	3	6	26	11	
P76 Alumínio e suas obras	5	5	9	19	12	
P79 Zinco e suas obras	5	5	9	26	12	
P80 Estanho e suas obras	5	7		26	12	
P75 Níquel e suas obras	6	7	13	26	12	
P78 Chumbo e suas obras	6	7	13	26	12	
P83 Obras diversas de metais comuns	7	5	10	11	13	Obras de metal e cimento

continua...

continuação...

Produtos (97)	a	b	c	d	e	Agregação por grupos (25)
P81 Outros metais comuns, cimentos e suas obras	7	5	10	26	13	
P96 Obras diversas		5		26	14	Obras de arte e diversas
P97 Obras de arte, de coleção e antiguidades		5		26	14	
P45 Cortiça e artigos de cortiça	3	6	11	26	15	Cortiça, tapetes e seda
P50 Seda	3	6	11	26	15	
P57 Tapetes e outros revestimentos para pavimentos	4	6		26	15	
P71 Pérolas, pedras preciosas, metais, moedas, etc.	6	7	13	13	16	Siderúrgicos e pedras preciosas
P73 Obras de ferro ou aço	6	7	13	14	16	
P72 Ferro e aço	6	7	13	16	16	
P87 Veículos diferentes, elétricos, ferroviários	7	7	14	1	17	Material de transporte, armas e equipamentos médicos
P90 Óptico, foto, técnicos, médicos, aparelhos, etc.	7	7	14	14	17	
P82 Ferramentas, implementos, talheres e utensílios de metal comum	7	7	14	21	17	
P86 Trilhos, locomotivas elétricas, material circulante, equipamentos	7	7	14	21	17	
P89 Embarcações e outras estruturas flutuantes	7	7	14	25	17	
P93 Armas e munições, suas partes e acessórios	7	7	14	26	17	
P88 Aeronaves e aparelhos espaciais e suas partes	7	7	14	29	17	
P94 Mobiliário, iluminação, sinalização, construções pré-fabricadas	8	7	15	15	18	
P91 Aparelhos de relojoaria e suas partes	8	7	15	21	18	

continua...

continuação...

Produtos (97)	a	b	c	d	e	Agregação por grupos (25)
P92 Instrumentos musicais, peças e acessórios	8	3		26	18	
P20 Vegetais, frutas, nozes, preparações alimentícias, etc.	1	8	16	11	19	Produtos de origem animal e vegetal, açúcares, café, oleaginosas, cacau, tabaco e bebidas
P21 Preparações alimentícias diversas	1	8	16	11	19	
P19 Cereais, farinha, amido, preparações de leite e de produtos	1	2		11	19	
P07 Produtos hortícolas, plantas, raízes e tubérculos	2	8	17	5	19	
P09 Café, chá, mate e especiarias	2	8	17	21	19	
P18 Cacau e suas preparações	2	8	17	21	19	
P16 Preparações de carne, peixes e frutos do mar	2	8	17	11	19	
P17 Açúcares e produtos de confeitaria	2	8	17	11	19	
P24 Tabaco e seus sucedâneos manufaturados	2	8	17	11	19	
P02 Carnes e miudezas, comestíveis	2	8	17	19	19	
P04 Produtos comestíveis de origem animal, lácteos, ovos, mel	2	8	17	19	19	
P12 Oleaginosas, grãos, sementes, frutos, etc.	2	8	17	20	19	
P23 Resíduos, resíduos da indústria de alimentos, ração animal	5	8		20	19	
P05 Produtos de origem animal, não especificados	2	8	17	26	19	
P03 Peixes, crustáceos, moluscos, invertebrados aquáticos, etc.	2	8	17	28	19	
P22 Bebidas, Bebidas alcoólicas e vinagres	2	8	17	28	19	

continua...

conclusão.

Produtos (97)	a	b	c	d	e	Agregação por grupos (25)
P31 Fertilizantes	3	8	18	5	20	Farmacêuticos e fertilizantes
P30 Produtos farmacêuticos	3	8	18	18	20	
P74 Cobre e suas obras	6	3	6	6	21	Cobre e suas obras
P49 Livros, jornais, fotos, etc.	6	6	12	11	22	Livros, jornais, fotos, etc.
P65 Freios e suas partes	6	4	8	26	23	Freios e suas partes
P95 Brinquedos, jogos, artigos de desporto		5		21	24	Brinquedos, jogos, artigos de desporto
P99 <i>Commodities</i> não especificadas				22	25	<i>Commodities</i> não especificadas

Fonte: Elaboração própria a partir de resultados obtidos no *software* Stata 11.

Nota: a) Método de Ward (8 *clusters*); b) ligação média (8 *clusters*); c) Ward + média (18 *clusters* + 12 *outliers*); d) método de *K* médias ($k = 30$); e) agregação final.

Recebido em: 17/02/2014.

Aceito em: 20/06/2014.

Revista Análise Econômica

Missão

Disseminar o conhecimento por meio da publicação de artigos, de natureza teórica ou empírica, e de resenhas bibliográficas de livros nacionais e estrangeiros da área de Ciências Econômicas.

Política Editorial

A *Análise Econômica* é uma publicação semestral que traz em todos os seus números contribuições originais de pesquisadores predominantemente de fora do Rio Grande do Sul. A revista busca disseminar o conhecimento, oferecendo aos leitores trabalhos de qualidade que avancem o conhecimento na área de Economia, tendo como público-alvo professores, pesquisadores e alunos de graduação e de pós-graduação, e como objetivo a integração desse público e a troca de informações atuais referentes à Ciência Econômica.

Desde a sua criação, em 1983, a *Análise Econômica* tem publicado em média dez artigos por edição. Aceitam-se para publicação trabalhos originais. Aceitam-se, também, notas e comentários sucintos de artigos já publicados, bem como resenhas de livros e teses. A revista aceita trabalhos em português, inglês e espanhol.

A revista utiliza o SEER, com submissões, pareceres e acompanhamento do processo editorial nesse sistema. Todas as edições estão disponíveis *on-line* com os artigos completos, que podem ser lidos na íntegra.

Política de Avaliação

Os textos submetidos são, primeiramente, examinados pelos Editores Responsáveis, que avaliam se o texto tem relação com o foco da revista e se segue as normas para publicação. Nessa fase, o artigo poderá ser recusado. Após exame prévio dos editores, o trabalho é encaminhado a avaliadores, no sistema *peer review*. Nessa etapa, a revista utiliza o sistema *blind review*, ou seja, os autores não são identificados em nenhuma fase da avaliação. A decisão dos avaliadores é registrada em pareceres, que são enviados aos autores, mantendo-se em sigilo os nomes desses avaliadores.

Normas para Publicação

A revista *Análise Econômica* é um periódico de publicação semestral na área de Economia. O autor que deseje publicar seu artigo deve submetê-lo à apreciação crítica e à revisão, conforme as normas da revista.

1. Aceitam-se para publicação trabalhos originais. Aceitam-se, também, notas e comentários sucintos de artigos já publicados, bem como resenhas de livros e teses. A revista *Análise Econômica* aceita trabalhos em português, inglês e espanhol.

2. Os textos submetidos são, primeiramente, examinados pelos editores responsáveis, que avaliam se o texto tem relação com o foco da revista e se segue as normas para publicação. Nessa fase, o artigo poderá ser recusado.

3. Após exame prévio dos editores, o trabalho é encaminhado a avaliadores, no sistema *peer review*. Nessa etapa, a revista utiliza o sistema *blind review*, ou seja, os autores não são identificados em nenhuma fase da avaliação. Por isso, antes de submeter o arquivo, exclua o(s) nome(s) do(s) autor(es) do trabalho submetido. A decisão dos avaliadores é registrada em pareceres, que são enviados aos autores, mantendo-se em sigilo os nomes desses avaliadores.

4. O texto deve ser digitado em tamanho de papel A4, formatado em espaço simples, corpo 12, em Word, limitando-se a 20 laudas, incluindo notas-de-rodapé, referências e anexos.

5. Os artigos devem apresentar, em português e em inglês: título; resumo (*Abstract*), com cerca de 10 linhas; três palavras-chave (*Keywords*) e Classificação JEL.

6. A apresentação dos trabalhos segue as normas da ABNT. As citações bibliográficas devem ser indicadas no próprio texto, contendo entre parênteses o sobrenome do autor, ano da publicação e, se for o caso, o número da página citada. Por exemplo: “Conforme Keynes (1936, p. 39), [...]” ou “(KEYNES, 1936, p. 39)”. As notas-de-rodapé devem ser destinadas exclusivamente a observações adicionais ao texto.

7. As referências completas deverão ser reunidas no fim do texto, em ordem alfabética, e elaboradas segundo a NBR 6023/2002. Exemplo: SAMUELSON, P. A. *Fundamentos da análise econômica*. São Paulo: Abril Cultural, 1983. 380p. (Coleção Os Economistas).

8. As fórmulas matemáticas e as notas-de-rodapé devem ser numeradas sequencialmente.

9. As ilustrações — tabelas, quadros, gráficos — devem ser numeradas e trazer legendas. Não usar cores além de preto e branco. Sempre indicar a fonte das ilustrações. Caso sejam elaboradas pelo autor, escrever: “Fonte: Elaboração própria”.

10. É de responsabilidade do autor providenciar a revisão gramatical e textual dos artigos em português e em língua estrangeira, bem como a formatação ABNT das referências. Isso deve ser feito antes da submissão do trabalho à avaliação.

11. As submissões devem ser feitas *on-line*. Para iniciar o processo de submissão de artigos basta se cadastrar na página principal da revista, clicando em CADASTRO. Nesse cadastro, marque a opção “Autor: Pode submeter à revista” (última opção do formulário). Ao finalizar o cadastro, acesse o *site* com seu *login* e senha, e clique em AUTOR. A seguir, inicie a submissão do artigo seguindo os cinco passos indicados pelo sistema. Por este *site*, é possível acompanhar o progresso do seu artigo dentro do processo editorial.

12. Ao submeter o artigo no *site*, o autor deverá preencher os metadados de submissão, para cada autor, com nome(s) completo(s), *e-mail*, instituição/afiliação, resumo da biografia, título, resumo e palavras-chave do trabalho em português.

13. As colaborações não serão remuneradas. Cada autor receberá três exemplares do número em que seu trabalho for publicado, e, caso possua coautor(es), cada um receberá dois exemplares.

Impressão:

