

VIÉSES NO CÁLCULO DE NÚMEROS-ÍNDICES UTILIZANDO A FÓRMULA DE TÖRNQVIST EM SÉRIES COM AUSÊNCIA DE DADOS¹

José Roberto Vicente²
Lilian Cristina Anefalos³
Denise Viani Caser⁴

1 - INTRODUÇÃO

A necessidade de se agregar bens heterogêneos pode ser resolvida tecnicamente com o emprego de fórmulas de números-índices. Diversos estudos efetuados a partir da década de 70 demonstraram a inadequação do uso de fórmulas clássicas, como as de Laspeyres e Paasche e, concomitantemente, as vantagens do emprego de índices superlativos, sendo os mais conhecidos os de Fisher e Törnqvist⁵. A escolha entre fórmulas superlativas tem importância relativa, já que seus resultados aproximam-se até a segunda ordem (DIEWERT, 1978).

A fórmula de Törnqvist vem sendo bastante utilizada, uma vez que é exata para uma forma agregativa translog homogênea (e, portanto, superlativa), enquanto que a fórmula de Fisher é exata para uma função agregativa quadrática de ordem dois homogênea (e, portanto, também superlativa). Todavia, aquela primeira fórmula apresenta problemas quando existem valores nulos nas séries que se deseja agregar, exigindo interferência para que possa ser empregada, com procedimentos nem sempre explicitados.

Para averiguar os efeitos de diversas dessas interferências, o objetivo deste trabalho foi o de efetuar algumas simulações, com dados dos Censos Agropecuários, comparando os resultados obti-

dos com os provenientes de outras fórmulas de números-índices, que não apresentam essa limitação.

2 - METODOLOGIA

As fórmulas de números-índices utilizadas foram as de Laspeyres, Paasche, Fisher e Törnqvist, que podem ser representadas para a aplicação mais simples - base fixa de ponderação e base fixa de cálculo - para índices de quantidade entre dois períodos de tempo (SILVA e CARMO, 1986), como:

a) Laspeyres

$$LQ_{0,1} = \frac{\sum_{i=1}^n p_0^i \cdot q_1^i}{\sum_{i=1}^n p_0^i \cdot q_0^i} = \sum_{i=1}^n \frac{q_1^i}{q_0^i} \cdot w_0^i$$

b) Paasche

$$PQ_{0,1} = \frac{\sum_{i=1}^n p_1^i \cdot q_1^i}{\sum_{i=1}^n p_1^i \cdot q_0^i} = \frac{1}{\sum_{i=1}^n \frac{q_0^i}{q_1^i} \cdot w_1^i}$$

c) Fisher

$$FQ_{0,1} = \sqrt{LQ_{0,1} \cdot PQ_{0,1}}$$

d) Törnqvist

$$TQ_{0,1} = \prod_{i=1}^n \left(\frac{q_1^i}{q_0^i} \right)^{\frac{w_0^i + w_1^i}{2}}$$

Onde a letra inicial, L, P, F ou T identifica o tipo de índice, e a letra Q indica que se referem a quantidades, p_1^i é o preço do item i, no período 1, p_0^i o preço do item i, no período 0, q_1^i a quantidade do item i, no período 1, q_0^i a quantidade

¹Estudo elaborado com auxílio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). Os autores agradecem a colaboração do estagiário Hernani Pinheiro Fernandes.

²Engenheiro Agrônomo, Doutor, Pesquisador Científico do Instituto de Economia Agrícola e bolsista do CNPq.

³Engenheira Agrônoma, Mestre, Pesquisadora Científica do Instituto de Economia Agrícola.

⁴Estatístico, Pesquisadora Científica do Instituto de Economia Agrícola.

⁵Ver, por exemplo, DIEWERT (1976; 1978), STAR; HALL (1976) e ALLEN; DIEWERT (1981).

do item i , no período 0, $w_0^i = \frac{p_0^i \cdot q_0^i}{\sum_{i=1}^n p_0^i \cdot q_0^i}$ a partici-

pação do item i , na relação orçamentária do período 0 e $w_1^i = \frac{p_1^i \cdot q_1^i}{\sum_{i=1}^n p_1^i \cdot q_1^i}$ a participação do item i ,

na relação orçamentária do período 1.

Foram calculados índices de produção para o setor de lavouras (quantidade produzida), utilizando-se dados do CENSO AGROPECUÁRIO (1970, 1975, 1980, 1985 e 1998), referentes às classes/grupos de atividade econômica agricultura e agropecuária (1970 a 1985) ou lavoura temporária, lavoura permanente e produção mista (1995), e preços recebidos pelos produtores publicados pela Fundação Getúlio Vargas (FGV). A disponibilidade de preços limitou o número de lavouras consideradas, que foram: algodão (arbóreo e herbáceo), amendoim, arroz, banana, batata, cacau, café, caju, cana, cebola, coco, feijão, fumo, juta, laranja, malva, mamona, mandioca, milho, pimenta do reino, sisal, soja, tomate, trigo e uva⁶.

As simulações efetuadas, para tornar possível o cálculo de índices pela fórmula de Törnqvist, foram:

- substituir as informações inexistentes por um número muito pequeno (0,0000000001);
- substituir as informações inexistentes por um;
- substituir as informações inexistentes pela informação anterior⁷;

- considerar o resultado $\left(\frac{q_1^i}{q_0^i}\right)^{\frac{w_0^i + w_1^i}{2}}$ igual a 1,

quando q_0^i ou q_1^i eram iguais a zero. Esse último procedimento, aplicado apenas para o cálculo pela fórmula de Törnqvist, implica desconsiderar as variações de casos com a ocorrência de zeros.

⁶GASQUES; CONCEIÇÃO (2000) estimaram índices com um número maior de culturas e com explorações animais. Apesar do ganho em abrangência, a inexistência de fontes de preços confiáveis, como os da FGV, para a maioria desses produtos, deve introduzir erros consideráveis nas medidas obtidas. Os problemas de vieses na fórmula de Törnqvist com um número maior de itens, cujos dados nem sempre existem, tendem a ser exacerbados.

⁷Nos casos em que as informações apareciam a partir de determinado ano, essas foram repetidas em todos os pontos anteriores; produtos inexistentes em todos os anos foram considerados com valor igual a 1.

Os resultados dessas simulações foram comparados aos obtidos sem efetuá-las, propiciados pelas fórmulas de Laspeyres, Paasche e Fisher.

Os índices foram estimados para o Brasil e para o Estado de São Paulo, para verificar o comportamento dos vieses a diferentes níveis de agregação.

3 - RESULTADOS E DISCUSSÃO

O cálculo dos índices foi efetuado ponto a ponto, e cada medida refere-se à variação em relação ao ponto anterior; em seguida, foi efetuado o encadeamento, que representa a variação total do período 1970-1995.

Trabalhando-se com os dados originais dos Censos Agropecuários, nota-se que as medidas fornecidas pelas fórmulas de Laspeyres, Paasche e Fisher são muito aproximadas; em nível de Brasil, a maior diferença ocorre em 1985 (base: 1980=100), em que o índice Laspeyres atinge 132,69 e o Paasche 135,67⁸. Com dados do Estado de São Paulo, a maior distância entre esses índices também aparece em 1985, com o Laspeyres resultando em 134,14, e o Paasche, em 138,95. Nos índices encadeados, percebe-se diferença razoável entre as duas fórmulas com dados do Brasil, o que é consistente com as evidências registradas na literatura em favor das fórmulas superlativas (Tabela 1).

Na primeira simulação - em que os dados iguais a zero foram substituídos por um valor muito pequeno - praticamente não houve alteração nos resultados das três fórmulas. Já os resultados fornecidos pela fórmula de Törnqvist apresentam enorme superestimação, que ocorre principalmente no ano de 1980 (índice igual a 190,01 para o Brasil e 223,12 para São Paulo, ambos com base em 1975=100) e se mantém no encadeamento, indicando variações 59,6% e 81,6% maiores do que as do índice Fisher para o Brasil e São Paulo, respectivamente. Esse problema ocorre devido à inexistência de dados de cebola, coco, juta, malva, mamona, pimenta e sisal, nos anos de 1970 e 1975. Quando esses produtos aparecem, em 1980, tendem a apresentar participações elevadíssimas no índice de Törnqvist; a cebola, por exemplo, responde por 27,8% das variações do índice em nível nacional, e por 67%

⁸O índice Fisher é a média geométrica do dois anteriores.

das variações do Estado de São Paulo.

Na segunda simulação - dados iguais a

zero substituídos por um - percebe-se uma maior

distância entre os resultados das fórmulas de

TABELA 1 - Variações da Produção Agrícola, Setor de Lavouras, Segundo Diferentes Fórmulas de Números-Índices, Brasil e Estado de São Paulo, 1970-1995

Fórmula	Brasil					São Paulo				
	1975	1980	1985	1995	Encadeado	1975	1980	1985	1995	Encadeado
Dados originais										
Laspeyres	140,92	117,11	132,69	104,54	228,92	125,40	133,72	134,14	90,87	204,39
Paasche	141,16	117,85	135,67	104,35	235,51	121,59	135,69	138,95	88,13	202,04
Fisher	141,04	117,48	134,17	104,45	232,19	123,48	134,70	136,52	89,49	203,21
Zero=0,0000000001										
Laspeyres	140,92	117,46	132,69	104,54	229,59	125,40	133,72	134,14	90,87	204,39
Paasche	140,81	117,85	135,67	104,35	234,92	121,59	135,69	138,95	88,13	202,04
Fisher	140,86	117,65	134,17	104,45	232,24	123,48	134,70	136,52	89,49	203,21
Törnqvist	142,87	190,01	134,02	101,88	370,65	124,18	223,12	151,74	87,79	369,08
Zero = 1										
Laspeyres	140,92	117,24	131,83	104,54	227,70	125,40	134,70	134,14	90,87	205,91
Paasche	140,98	119,09	135,67	104,35	237,69	121,17	135,69	138,95	88,13	201,35
Fisher	140,95	118,16	133,74	104,45	232,64	123,27	135,20	136,52	89,49	203,61
Törnqvist	142,77	143,96	133,36	102,98	282,29	124,35	156,31	140,72	87,79	240,12
Zero = anterior										
Laspeyres	138,43	116,75	132,69	104,62	224,36	123,33	133,24	133,90	90,88	199,96
Paasche	139,20	114,31	135,67	104,34	225,26	119,47	130,31	137,62	88,14	188,85
Fisher	138,82	115,52	134,17	104,48	224,81	121,38	131,77	135,75	89,50	194,33
Törnqvist	140,72	115,48	134,02	103,44	225,28	122,78	131,11	135,39	87,79	191,34
Zero: peso = 1										
Törnqvist	142,77	115,71	134,02	103,45	229,04	124,35	132,04	135,57	87,79	195,42

Fonte: Elaborada pelos autores a partir de dados básicos do IBGE e da FGV.

Laspeyres e de Paasche, mas os resultados do índice Fisher encadeados são ainda muito próximos, com diferenças inferiores a 0,5 ponto percentual. A fórmula de Törnqvist continua apresentando considerável superestimação, ainda que bem menor do que a da simulação anterior (21,3% para o Brasil e 17,9% para São Paulo, em relação ao resultado da fórmula de Fisher). Os pesos da cebola, nesse caso, são de 8,6% no índice nacional e de 17,9% no de São Paulo.

Na terceira simulação - informações inexistentes substituídas pela informação anterior - os índices encadeados tendem a convergir, com diferenças de 0,2% e 1,5% entre as fórmulas de Törnqvist e Fisher, para Brasil e São Paulo, respectivamente. Tomando-se como base o índice Fisher calculado sem interferência, nota-se uma subestimação nos resultados do índice Törnqvist, de cerca de 3% para o Brasil e 6% para o Estado de São Paulo.

A última simulação - que equivale a ignorar variações nos produtos com ocorrência de zeros - permite a melhor aproximação dos resultados da fórmula de Törnqvist com os outros índices calculados sem interferência; ainda permanece a tendência de subestimação verificada na simulação anterior, embora menos acentuada.

Para o Brasil, o índice Törnqvist encadeado apresenta-se 1,4% inferior ao Fisher original, e para o Estado de São Paulo, 3,8% inferior.

4 - CONCLUSÕES E CONSIDERAÇÕES FINAIS

As diversas simulações experimentadas com dados inexistentes de produção de lavouras presentes nos Censos Agropecuários pouco afetaram os resultados dos índices calculados pelas fórmulas de Laspeyres, Paasche e Fisher. Entretanto, alteraram significativamente os valores do índice Törnqvist, que variaram de superestimação elevada a subestimação discreta.

Esses problemas tendem a se agravar em níveis menores de agregação, em que um número maior de itens pode apresentar valores iguais a zero; o mesmo comportamento é esperado quando se aumenta o número de explorações consideradas.

ALLEN e DIEWERT (1981), comparando o desempenho de diversas fórmulas de números-índices em dados censitários, indicaram a de Fisher como a mais adequada, o que é consistente com os resultados encontrados neste estudo.

O comportamento irregular exibido pela

fórmula de Törnqvist reforça a necessidade de explicitar os procedimentos adotados em análises

nela baseadas, para permitir uma visão crítica dos resultados obtidos.

LITERATURA CITADA

ALLEN, R. C.; DIEWERT, W. E. Direct versus implicit superlative index number formulae. *Review of Economics and Statistics*, v. 63, n. 3, p. 430-435, Aug. 1981.

CENSO AGROPECUÁRIO. Rio de Janeiro: IBGE, 1970, 1975, 1980, 1985, 1998.

DIEWERT, W. E. Exact and superlative index numbers. *Journal of Econometrics*, North-Holland, v. 4, n. 2, p.115-145, May 1976.

_____. Superlative index numbers and consistency in aggregation. *Econometrica*, v. 46, n. 4, p. 883-900, Jul. 1978.

GASQUES, J. G.; CONCEIÇÃO, J. C. P. R. **Transformações estruturais da agricultura e produtividade total dos fatores**. Brasília: IPEA, nov. 2000. (Textos para Discussão, n. 768).

SILVA, G. L. S. P.; CARMO, H. C. E. Como medir a produtividade agrícola: conceitos, métodos e aplicações no caso de São Paulo. *Agricultura em São Paulo*, São Paulo, v. 33, t. 1/2, p. 139-170, 1986.

STAR, S.; HALL, R. E. An approximate Divisia index of total factor productivity. *Econometrica*, v. 44, n. 2, p. 257-263, Mar. 1976.

VIÉSES NO CÁLCULO DE NÚMEROS-ÍNDICES UTILIZANDO A FÓRMULA DE TÖRNQVIST EM SÉRIES COM AUSÊNCIA DE DADOS

RESUMO: O objetivo deste artigo é o de analisar as conseqüências do uso da fórmula de Törnqvist em séries que apresentam zeros. Alguns procedimentos usuais para contornar esse problema foram experimentados, e comparados os viéses deles decorrentes. As fontes dos dados foram os Censos Agropecuários do IBGE e a Fundação Getúlio Vargas. Os resultados mostraram que, dependendo do procedimento adotado, os índices calculados pela fórmula de Törnqvist variaram de grande superestimação a alguma subestimação.

Palavras-chave: números-índices, fórmula de Törnqvist, fórmula de Fisher.

BIASES IN THE CALCULATION OF INDEX-NUMBERS USING TÖRNQVIST FORMULA IN SERIES WITH ABSENCE OF DATA

ABSTRACT: The objective of this paper is to analyze the consequences of using Törnqvist formula in series containing zeros. To overcome this problem a few usual procedures were carried out, being their resulting biases subsequently compared. Data were drawn from the IBGE Agricultural Censuses and the Fundação Getúlio Vargas. The results showed that, depending upon the adopted procedure, the indexes derived from Törnqvist formula ranged from a great overestimation to a little underestimation.

Key-words: index numbers, Törnqvist formula, Fisher formula.

Recebido em 22/05/2001. Liberado para publicação em 06/06/2001.