

Notas sobre uma nova decomposição do índice de Gini*

José W. Rossi**

Neste artigo, a decomposição setorial do índice de Gini, recentemente proposta por Braulke (1983), é aplicada a dados de renda do Brasil. Os resultados obtidos indicam que tal decomposição é suficientemente precisa. Como o procedimento pressupõe distribuições de rendas setoriais do tipo Pareto, que certamente não se aplicam ao caso do Brasil, sugerimos que a fórmula seria provavelmente resistente a desvios dessa forma funcional.

1 — Introdução

Sabe-se já de algum tempo [e. g., Bhattacharya e Mahalanobis (1967)] que o índice de Gini não pode ser decomposto nas desigualdades dentro dos grupos e entre os grupos; há um terceiro termo que indica o grau em que as rendas se superpõem nos vários grupos. Pyatt (1976), com a sua abordagem do jogo estatístico para a decomposição desse índice, contribuiu apenas para um melhor entendimento dessa dificuldade. Agora Braulke (1983) apresenta-nos uma outra decomposição do índice de Gini, mas que é derivada supondo-se distribuições de rendas setoriais do tipo Pareto.

Dado o fato de que a função de Pareto não é em geral uma boa especificação para a maioria das distribuições empíricas de renda (um bom ajustamento é geralmente obtido apenas para a cauda superior de tais distribuições), há a tentação em rejeitar-se a decomposição de Braulke sem quaisquer outras considerações. Assim, parece ser de algum interesse a revelação de que tal decomposição fornece resultados razoavelmente pre-

* Sou imensamente grato a um parecerista desta revista, tanto pelas sugestões apresentadas, como pelas várias correções feitas aos meus cálculos anteriores para as Tabelas 1 e 2. As correções relativas à Tabela 2 foram mais substanciais e conduziram-me, inclusive, a alterar algumas das minhas conclusões anteriores.

** Da Universidade Federal do Rio de Janeiro e do IBGE.

cisos, mesmo com dados de renda do Brasil, onde tem-se um apreciável segmento populacional com rendas baixas ou nulas.¹ A próxima seção preocupa-se justamente com tal aplicação e a Seção 3 apresenta comentários gerais.

2 — Aplicação para dados de renda do Brasil

A decomposição de Bralke para uma economia com dois setores é a seguinte:

$$G = \frac{c^2 b}{k} \frac{(1 - a_1)}{1 + a_1} + \frac{(1 - c)^2}{k} \frac{(1 - a_2)}{1 + a_2} + \\ + \frac{c(1 - c)(1 - b)}{k} + \frac{2c(1 - c)}{k} A$$

onde $k = cb + 1 - c$ e:

$$A = \frac{b(1 - a_1)^2}{1 - a_1 a_2} \left(\frac{a_1 b}{a_2} \right)^{\frac{a_1}{1 - a_1}} \quad \text{se } a_1 b \leq a_2 \\ = b - 1 + \frac{(1 - a_2)^2}{1 - a_1 a_2} \left(\frac{a_1 b}{a_2} \right)^{\frac{a_2}{a_2 - 1}} \quad \text{se } a_1 b \geq a_2$$

e onde b é a razão entre as rendas médias do primeiro e segundo setores, c é a fração populacional do primeiro setor e a_i é o parâmetro da função de Pareto do setor i , baseado na curva de Lorenz:

$$y_i = 1 - (1 - x_i)^{a_i}$$

que é relacionada ao índice de Gini através de:

$$G_i = (1 - a_i) / (1 + a_i)$$

Os dois primeiros termos da decomposição são uma média ponderada dos índices de Gini setoriais, onde os pesos são o produto das frações de

¹ Para discussões sobre a distribuição de renda no Brasil, ver, por exemplo, Bacha e Taylor (1978), Fishlow (1972), Langoni (1973), Denslow Jr. e Tyler (1983) e Hoffmann e Duarte (1972).

renda e população dos respectivos setores (note-se que cb/k e $(1-c)/k$ são, respectivamente, as frações de renda do primeiro e segundo setores), medindo então a desigualdade dentro do grupo. O terceiro termo (ou o seu valor negativo, se $b > 1$), por outro lado, é o índice de Gini na ausência de qualquer desigualdade dentro de cada setor (isto é, $a_1 = a_2 = 1$, já que $a_i = 0$ e $a_i = 1$ correspondem, respectivamente, à perfeita desigualdade e perfeita igualdade), podendo então ser considerada como a desigualdade entre os grupos. Esses três primeiros termos equivalem, portanto, aos dois primeiros da decomposição de Bhattacharya e Mahalanobis (1967), correspondendo às desigualdades dentro dos setores e entre os setores, respectivamente. Segue-se, pois, que o último termo indica a desigualdade devida à superposição da renda nos dois setores; evidentemente, isto daria uma medida exata da superposição da renda apenas se as distribuições de rendas setoriais forem exatamente Pareto, caso contrário tem-se apenas uma medida aproximada.

Como os índices de Gini setoriais (e rendas médias) encontram-se prontamente disponíveis para o Brasil, não é necessário ajustar qualquer função de Pareto aos dados de renda; ² o parâmetro da função pode ser obtido do correspondente índice de Gini. A adoção deste procedimento cria, aliás, uma situação interessante, vale dizer, isto garante que os três primeiros termos da decomposição de Braulke (isto é, as desigualdades entre os grupos e dentro dos grupos) sejam iguais àqueles do esquema Bhattacharya-Mahalanobis-Pyatt. Tais termos seriam certamente diferentes caso os parâmetros a_i fossem obtidos diretamente da estimação da função de Pareto. Assim, em certo sentido os nossos resultados fornecem uma estimativa da superposição da renda apenas. Os resultados da Tabela 1 devem, então, ser vistos nessa perspectiva e, como eles bem indicam, os erros de aproximação obtidos pela fórmula proposta por Braulke são sempre muito reduzidos (com a exceção do caso "Fronteira" em 1980). ³ Naturalmente, se as distribuições de renda setoriais fossem exatamente do tipo Pareto, tais

² As informações relativas a 1960 são de Jain (1975) e aquelas para 1970 e 1980 vêm de Denslow Jr. e Tyler (1983). A unidade recipiente da renda é a população economicamente ativa.

³ A possibilidade do cálculo incorreto do índice de Gini não deve naturalmente ser excluída neste caso, e não se tem como verificar tal ponto, já que esses índices são dados para a análise. Tomou-se apenas o cuidado para que os cálculos deste estudo fossem corretamente obtidos.

TABELA I

Índice de Gini atual e aproximado

Regiões	Anos	Parâmetros*			Índice de Gini agregado		Erro em %		
		$a_1(G_1)$	$a_2(G_2)$	b	c	Atual (1)		$G(a_1, a_2, b, c)$ (2)	
Sudeste	1970	0,3750(0,4545)	0,3010(0,5373)	0,410	0,3150	0,5450	0,5527	0,0077	1,41
	1980	0,2840(0,5576)	0,3050(0,5326)	0,5280	0,3330	0,5610	0,5621	0,0011	0,20
Nordeste	1970	0,4240(0,4045)	0,2590(0,5886)	0,3770	0,5920	0,5570	0,5537	-0,0033	-0,60
	1980	0,3610(0,4695)	0,2580(0,5898)	0,4340	0,4580	0,5860	0,5801	-0,0059	-1,01
Fronteira**	1970	0,4940(0,3387)	0,3100(0,5267)	0,4930	0,5520	0,4800	0,4724	-0,0076	-1,58
	1980	0,3310(0,5026)	0,2630(0,5835)	0,6350	0,5170	0,5330	0,5368	-0,0262	-4,49
Brasil	1960	0,3980(0,4306)	0,3470(0,4848)	0,4430	0,4410	0,5050	0,5098	0,0048	0,96
	1970	0,3890(0,4399)	0,2890(0,5516)	0,3680	0,4020	0,5650	0,5670	0,0020	0,35
1980	0,2950(0,5444)	0,2790(0,5637)	0,4790	0,2720	0,5900	0,5808	-0,0092	-1,56	

*Os subíndices 1 e 2 representam, respectivamente, os setores rural e urbano; b é a razão entre as rendas médias dos dois setores; c é a fração populacional do setor rural; e, finalmente, os números entre parênteses são os correspondentes índices de Gini setoriais.

**Inclui os estados do Amazonas, Pará, Goiás, Mato Grosso do Sul, Rondônia e Acre e os territórios de Amapá e Roraima.

erros seriam nulos. Como a distribuição de renda no Brasil certamente não é do tipo Pareto (pelas razões já apontadas), isso parece sugerir que a fórmula de decomposição proposta por Braulke seria provavelmente resistente a desvios de tal forma funcional.

Para medir mudanças na desigualdade de renda entre dois períodos quaisquer, Braulke propõe:

$$\begin{aligned} G(a^1, b^1, c^1) - G(a^0, b^0, c^0) = & \{G(a^1, b^1, c^1) - G(a^0, b^1, c^1)\} \\ & + \{G(a^0, b^1, c^1) - G(a^0, b^0, c^1)\} \\ & + \{G(a^0, b^0, c^1) - G(a^0, b^0, c^0)\} \end{aligned}$$

onde os sobreíndices indicam o tempo de referência e a é escrito por (a_1, a_2) . Os três termos entre colchetes são, respectivamente, o componente “dentro dos setores”, o componente “entre os setores” e um puro efeito migratório. A aplicação desta fórmula aos dados do Brasil produziu os resultados da Tabela 2. Agora, as mudanças observadas são previstas com pouca precisão. De qualquer modo, os principais resultados indicam que: ⁴ *a*) os componentes “dentro dos setores” são responsáveis pela maior parte das mudanças totais do índice de Gini, tanto do “Sudeste” como do “Brasil” em 1960/70 e 1970/80; *b*) os componentes “entre setores” são negativos em todos os casos (exceto para o “Brasil” em 1960/70), indicando pois a sua contribuição para a redução da desigualdade; e, finalmente, *c*) o componente puramente migratório só é de alguma importância no caso do Nordeste, onde, aliás, foi maior o processo de urbanização. Como última observação sobre esses resultados, ressalte-se que as mudanças encontradas para o “Brasil” em 1970/80 não diferem muito daquelas calculadas por Denslow Jr. e Tyler (1983) usando o índice de Theil e uma economia com três setores; note-se, entretanto, que é grande, nesse caso, o erro de aproximação do nosso procedimento, conforme indica, aliás, a Tabela 2.

⁴ Os resultados para o Brasil no período 1960/70 diferem um pouco daqueles obtidos por Braulke (1983, Tab. 2), porque a distribuição de renda para 1970 utilizada por aquele autor provém de Jain (1975), enquanto os nossos dados são os de Denslow Jr. e Tyler (1983); esses resultados diferem um pouco entre as duas fontes.

TABELA 2

Decomposição da variação na desigualdade agregada

Regiões	Períodos	Variação no índice de Gini agregado segundo as várias fontes				Erro em %	
		Atual (1)	Componente "dentro dos setores" (2)	Componente "entre os setores" (3)	Efeito migração (4)		Erro (5) = (1)-(2)-(3)-(4)
Sudeste	1970/80	0,0160	0,0272	-0,0181	-0,0003	-0,0066	-41,14
Nordeste	1970/80	0,0290	0,0230	-0,0144	0,0178	-0,0026	-8,98
Fronteira	1970/80	0,1030	0,1026	-0,0241	0,0659	0,0186	-18,09
Brasil	1960/70	0,0600	0,0402	0,0174	-0,0004	-0,0029	-4,76
	1970/80	0,0250	0,0316	-0,0167	-0,0010	-0,0111	-44,59
	1960/80	0,0850	0,0809	-0,0051	-0,0048	-0,0140	-16,48

3 — Observações finais

A decomposição do índice de Gini proposta por Braulke parece produzir resultados suficientemente precisos no caso de uma economia com dois setores. Além disso, a sua implementação requer apenas dados básicos. Tais qualidades fazem-na uma ferramenta analítica potencialmente útil em trabalhos aplicados.

Apesar dessas considerações, parece existir também alguma limitação ao uso da técnica. Embora Braulke sugira que o seu método poderia provavelmente ser generalizado para uma economia com n setores, não fica muito clara qual a complexidade da sua fórmula. Mesmo na eventualidade de tal generalização, não é também claro se os resultados seriam ainda tão precisos quanto no caso da economia com dois setores. Adicionalmente, embora a decomposição da desigualdade por uma dimensão e duas classificações dentro de tal dimensão pudesse ser um útil exercício, ela provavelmente revelará tanto quanto ocultará. A consideração de mais dimensões e/ou mais que duas classificações traria, naturalmente, esclarecimentos adicionais à questão da desigualdade. Mas, então, apenas medidas de desigualdade tais como o índice de Theil e a variância do logaritmo da renda poderiam ser usados, não o índice de Gini.

Bibliografia

- BACHA, E. L., e TAYLOR, L. Brazilian income distribution in the 60's: "facts", model results and the controversy. *Journal of Development Studies*, 14 (2) :271-97, 1978.
- BHATTACHARYA, N., e MAHALANOBIS, M. Regional disparities in household consumption in India. *American Statistical Association Journal*, 62: 143-61, 1967.
- BRAULKE, M. An approximation to the Gini coefficient for a population based on sparse information for sub-groups. *Journal of Development Economics*, 12:75-81, 1983.
- DENSLow JR., D., e TYLER, W. G. Perspective on poverty and income inequality in Brazil: an analysis of the changes during the 1970's. *World Bank Staff Working Paper*, 601, 1983.

- FISHLOW, A. Brazilian size distribution of income. *American Economic Review*, 62:391-401, maio 1972.
- HOFFMANN, R., e DUARTE, J. C. A distribuição de renda no Brasil. *Revista de Administração de Empresas*, 12 (2) :46-66, mar. 1972.
- JAIN, S. *Size distribution of income: a compilation of data*. Washington, The World Bank, 1975.
- LANGONI, C. G. *Distribuição da renda e crescimento econômico do Brasil*. Rio de Janeiro, Expressão e Cultura, 1973.
- PYATT, G. On the interpretation and disaggregation of Gini coefficients. *Economic Journal*, 86:243-55, 1976.

(Originais recebidos em maio de 1984. Revisos em setembro de 1984.)