

NOTA:

Estimação de Índices Trimestrais de Renda para o Brasil Uma Nova Técnica (*)

Celso L. Martone^()**

A ausência de dados trimestrais de renda real no Brasil tem criado sérias dificuldades aos trabalhos econométricos, particularmente na área de Teoria Monetária, onde a renda nacional é uma das medidas mais importantes para o teste de qualquer teoria. Um levantamento direto da renda trimestral, nos moldes do que é feito nas contas nacionais, pelo volume de recursos que envolveria, está fora de alcance de um pesquisador ou grupo de pesquisadores. Entretanto, existe um grande volume de dados ou indicadores individuais do comportamento trimestral e até mensal do nível de atividade que se encontram rapidamente disponíveis. Este conjunto de indicadores constitui uma riqueza estatística cujo potencial informativo não tem sido totalmente explorado até agora.

Esta nota procura utilizar esta informação estatística disponível, aplicando uma técnica que raramente é usada em Economia, qual seja a Análise de Componentes Principais. Na primeira parte, exponho sinteticamente o que é um componente

(*) Uma primeira versão desta nota foi apresentada em junho de 1972 no Seminário de Problemas Latino-Americanos da Universidade de Chicago, em forma mimeografada, contendo estimativas até 1970.

(**) O autor é professor do Instituto de Pesquisas Econômicas da Universidade de São Paulo.

principal e na segunda parte indico os principais resultados obtidos da análise com os dados brasileiros. Na terceira e última parte, alguns testes são feitos com a série estimada de renda real trimestral e indicadas suas limitações atuais.



Dadas n observações sobre um conjunto de k variáveis originais ou fatores X , onde X é uma matriz $n \times k$, o procedimento consiste em se determinar um único vetor \underline{z} , de ordem $n \times 1$, que descreva melhor (no sentido de Mínimos Quadrados) as colunas de X . Desta forma, deveremos minimizar a soma dos quadrados dos desvios $(X - \underline{z} \underline{a}')$ onde \underline{a}' é um vetor linha $1 \times k$ de "pesos" ou "cargas" dos fatores, um para cada variável original. Como não conhecemos a priori \underline{z} e \underline{a}' , um procedimento iterativo terá que ser usado para resolver o problema ⁽¹⁾.

Inicialmente, para obter unicidade de \underline{z} , impomos a restrição:

$$\underline{z}' \underline{z} = 1 \quad (1)$$

Nosso objetivo é, como ficou claro, minimizar a expressão:

$$\begin{aligned} W &= \text{tr} (X - \underline{z} \underline{a}')' (X - \underline{z} \underline{a}') \\ &= \text{tr } X' X - 2 \underline{z}' X \underline{a} + \underline{a}' \underline{a} \end{aligned} \quad (2)$$

Diferenciando parcialmente em relação a \underline{a} e igualando a zero, obtemos um primeiro resultado:

$$\underline{a} = X' \underline{z} \quad (3)$$

que, substituído novamente em (2), nos permite ficar apenas com a incógnita \underline{z} :

$$W = \text{tr } X' X - \underline{z}' X X' \underline{z} \quad (4)$$

(1) Esta exposição segue de perto a de Henri Theil. *Principles of Econometrics*, John Wiley & Sons, 1971, pp. 46-52, onde uma discussão mais completa pode ser encontrada.

Portanto, podemos agora minimizar W em relação a \underline{z} ou, o que é o mesmo, maximizar $(\underline{z}' \mathbf{X} \mathbf{X}' \underline{z})$ sujeito à restrição (1), isto é:

$$\max_{(\underline{z})} [\underline{z}' \mathbf{X} \mathbf{X}' \underline{z} - \Theta (\underline{z}' \underline{z} - 1)] \quad (5)$$

onde Θ é um multiplicador de Lagrange.

Diferenciando em relação a \underline{z} e igualando a zero, obtemos:

$$(\mathbf{X} \mathbf{X}' - \Theta \mathbf{I}) \underline{z} = 0 \quad (6)$$

que mostra que \underline{z} é um vetor característico da matriz $(\mathbf{X} \mathbf{X}')$ correspondente à raiz característica Θ .

Pré-multiplicando agora (6) por \underline{z}' , vem:

$$\underline{z}' \mathbf{X} \mathbf{X}' \underline{z} = \Theta \quad (7)$$

que implica que, a fim de maximizar (5), devemos tomar Θ como a maior raiz característica associada à matriz $(\mathbf{X} \mathbf{X}')$. O vetor \underline{z} , neste caso, é denominado primeiro componente principal das k variáveis originais representadas em \mathbf{X} .

Além disso, pré-multiplicando (6) por \mathbf{X}' , obtemos:

$$(\mathbf{X}' \mathbf{X} - \Theta \mathbf{I}) \mathbf{X}' \underline{z} = (\mathbf{X}' \mathbf{X} - \Theta \mathbf{I}) \underline{a} = 0 \quad (8)$$

que mostra o vetor \underline{a} também como um vetor característico da matriz $(\mathbf{X}' \mathbf{X})$ correspondente à raiz característica Θ .

Finalmente, usando (3) e (6), conseguimos uma relação associando todas as variáveis relevantes:

$$\underline{z} = \frac{1}{\Theta} \mathbf{X} \underline{a} \quad (9)$$

Utilizando o mesmo procedimento, o segundo componente principal pode ser calculado, correspondente à segunda maior raiz de $(\mathbf{X} \mathbf{X}')$ e assim sucessivamente.

II

Um problema fundamental da utilização desta técnica em Economia é o de se identificar os componentes principais estimados, o que em geral não é possível. No caso presente, entretanto, a identificação do componente principal é imposta como condição inicial, através de uma conveniente seleção do conjunto de variáveis originais. Como queremos obter um índice de renda real, selecionamos as variáveis originais na base do conhecimento a priori de sua alta correlação com a renda real, eliminando desta forma o problema da identificação. O primeiro componente principal é então caracterizado como um índice de renda real trimestral.

Apesar disso, uma seleção cuidadosa das variáveis originais é necessária. A escolha ideal teria que recair sobre aquelas variáveis que estivessem altamente correlacionadas com a renda real sem serem, ao mesmo tempo, partes da renda ou que, pelo menos, representassem o comportamento de um segmento substancial da economia. A razão disto é que, embora um componente da renda, digamos a produção de eletrodomésticos, pudesse estar correlacionada com ela, suas flutuações trimestrais poderiam não representar bem as flutuações trimestrais do nível de atividade. Em outras palavras, a série de produção de eletrodomésticos estaria captando não somente os movimentos do nível agregado de renda mas também os movimentos específicos deste setor de atividade. Neste caso estariamos embutindo no índice estimado as variações estacionais e variabilidade geral das variáveis escolhidas, sem que isto representasse o comportamento da renda real.

Foram então escolhidas oito variáveis originais, divididas em dois grupos ⁽²⁾:

Insumos Básicos à Produção

1. Consumo Industrial de Energia Elétrica (0,992)
2. Produção Nacional de Cimento (0,983)

(2) Grande parte dos dados utilizados pode ser encontrada na revista **Análise e Perspectiva Econômica**, APEC Editora, Rio de Janeiro (varios números).

3. Produção Nacional de Aço (0,945)
4. Produção Nacional de Borracha Sintética (0,905)
5. Produção Nacional de Veículos (0,976)

Indicadores do Nível de Atividade

6. Valor Real dos Cheques Compensados no País (0,979)
7. Valor Real da Receita do ICM no Estado de São Paulo (0,724)
8. Índice de Emprego Industrial na Cidade de São Paulo (0,966)

O período coberto pela análise foi 1961-72 trimestralmente. Devido à não uniformidade de algumas séries, este período foi dividido em dois, 1961-64 e 1965-72, tendo sido feita uma análise separada para cada um. O resultado da computação é uma variável padronizada z_t , com média nula e desvio-padrão unitário. Para transformá-la de volta num índice do PNB real, fizemos:

$$y_t = \bar{y} + s_y z_t \quad (10)$$

onde y é o índice estimado do PNB real trimestral, \bar{y} é a média do índice oficial anual do PNB real, como aparece nas contas nacionais do Brasil, e s_y é o desvio-padrão deste índice anual.

Os números entre parêntesis defronte de cada uma das variáveis originais na relação dada acima são as cargas de cada variável na formação do primeiro componente principal e são também os coeficientes de correlação simples entre cada uma das variáveis e o componente principal.

Na tabela 1, são apresentados os resultados para o período 1961-72, com todas as séries baseadas no ano de 1965, tanto na forma de índice como na forma de cruzeiros reais daquele ano. Apresentamos também o índice estacionalmente ajustado do PNB real na coluna (3) da tabela, onde os fatores estacionais são os seguintes:

TRIMESTRES	FATORES
I	0,9616
II	1,0021
III	1,0167
IV	1,0207

Os fatores estacionais foram calculados pelo método das médias móveis e foram normalizados por uma média geométrica. O comportamento estacional da economia brasileira, segundo os índices, é em forma de escada, com um degrau mais alto do primeiro para o segundo trimestre e quase que um patamar do terceiro para o quarto trimestre. Assim, uma queda do nível de atividade de 5,8% no primeiro trimestre, uma elevação de 4,2% no segundo trimestre, de 1,5% no terceiro trimestre e de 0,4% no quarto trimestre podem ser consideradas como normais, resultantes da simples estrutura da economia.

Para o período 1961-64, a fração total da variância explicada pelo primeiro componente principal foi de 58%, enquanto que, para o período 1965-72, foi de 88%. Esta diferença é explicável pelo fato de que, para 1961-64, os indicadores de números 3, 7 e 8 não estão disponíveis.

III

O índice trimestral estimado é, nessa altura, muito mais um índice de produto industrial do que propriamente um índice de renda real global da economia. A razão disto é que a maioria dos indicadores utilizados como variáveis originais está intimamente associada com o setor industrial. Não há disponibilidade de indicadores trimestrais do setor agrícola no Brasil, o que se explica em grande parte pelo fato de que, embora a renda do setor seja gerada continuamente durante o ano, sua realização, só se dá em certos meses, que coincidem com a época das colheitas.

TABELA 1
ÍNDICES TRIMESTRAIS ESTIMADOS DO PNB REAL
DO BRASIL: 1961-72

Anos Trimestres	Índice do PNB Real (1965 = 100)	PNB Real a Preços de 1965 (Cr\$ Bilhões)	Índice do
			PNB Real Ajustado (1965 = 100)
	(1)	(2)	(3)
1961 I	91,5	31,8	92,6
II	94,8	32,9	96,1
III	95,0	33,0	93,8
IV	96,7	33,6	95,5
1962 I	94,6	32,9	95,8
II	98,6	34,2	99,9
III	102,7	35,7	101,5
IV	104,1	36,2	102,7
1963 I	100,8	35,0	102,1
II	98,0	34,0	99,3
III	100,9	35,0	99,6
IV	99,7	34,6	98,4
1964 I	99,1	34,4	100,4
II	99,8	34,7	101,2
III	103,7	36,0	102,4
IV	105,9	36,8	104,6
1965 I	100,2	34,8	104,2
II	92,3	32,1	92,2
III	99,2	34,5	97,6
IV	108,3	37,6	106,1
1966 I	106,6	37,0	110,9
II	113,8	39,5	113,6
III	117,3	40,7	115,4
IV	115,1	40,0	112,8
1967 I	106,9	37,1	111,2
II	112,6	39,1	112,4
III	118,7	41,2	116,8
IV	119,8	41,6	117,4
1968 I	116,6	40,5	121,2
II	130,7	45,4	130,4
III	134,3	46,6	132,1
IV	141,8	49,2	138,9
1969 I	136,7	47,5	142,2
II	145,4	50,5	145,1
III	151,6	52,6	149,1
IV	149,5	51,9	146,5
1970 I	144,4	50,2	150,2
II	151,6	52,7	151,3
III	156,5	54,4	154,0
IV	160,9	55,9	157,6
1971 I	154,0	53,5	160,1
II	162,3	56,4	162,0
III	171,6	59,6	168,8
IV	178,4	62,0	174,8
1972 I	175,7	61,0	182,8
II	185,9	64,6	185,5
III	192,7	66,9	189,6
IV	200,1	69,5	196,1

Entretanto, a aproximação da renda real agregada por um índice do produto industrial parece funcionar bem. Num artigo recente, M. Keran usou um índice de produto industrial como "proxy" para o PNB real para vários países europeus que não dispõem de dados trimestrais de PNB com resultados satisfatórios. Ele fez também um teste com vários países onde existem séries de ambos os índices e concluiu que, exceto para um único caso, os resultados foram quase os mesmos, independentemente do índice utilizado⁽³⁾. Isto significa que, a curto prazo, as flutuações, do setor industrial aproximam satisfatoriamente as flutuações gerais da economia.

Uma forma de testar a qualidade do índice estimado consiste em se computar as taxas anuais de crescimento implicadas pelo índice e compará-las com as taxas anuais oficiais de crescimento das contas nacionais⁽⁴⁾. Na tabela 2, é feita esta comparação, para o período 1962-72. Como se pode observar, em termos de taxas anuais, a semelhança das séries de taxas é grande. O coeficiente de correlação entre as taxas anuais do índice estimado com as taxas do produto total é de 65%, enquanto que este mesmo coeficiente entre as taxas estimadas e as do produto industrial é de 88%, o que reafirma o comentário feito acima de que o índice é, presentemente, mais associado ao comportamento do setor industrial.

Outro teste, de natureza não quantitativa, que se pode imaginar é verificar, com base na experiência factual da economia brasileira, se o índice estimado capta as chamadas "crises conjunturais" que ocorreram no Brasil durante a última década. Assim, por exemplo, podemos verificar que as crises de 1963 e início de 1964, do primeiro semestre de 1965, do fim de 1966 e início de 1967 e de fins de 1969, para citar as que despertaram maior atenção na época, correspondem a quedas no índice de produto estimado. Naturalmente, nada se pode dizer com respeito à magnitude das flutuações da série estimada e até que ponto elas refletem com exatidão a amplitude das flutuações da economia. As flutuações trimestrais da série estimada se

(3) Michael Keran, "Monetary and Fiscal Influences on Economic Activity: The Foreign Experience", *Review, Federal Reserve Bank of St. Louis*, fevereiro, 1970.

(4) Observe-se que nenhum tipo de ajustamento foi feito no índice estimado para compatibilizá-lo com a série anual oficial.

TABELA 2

COMPARAÇÃO DAS TAXAS ANUAIS DE CRESCIMENTO

Anos	Índice Estimado (1)	Contas Nacionais	
		Produto Industrial (2)	Produto Total (3)
1962	5,8	7,8	5,2
1963	-0,2	0,2	1,6
1964	2,3	5,1	2,9
1965	-2,1	-4,7	2,8
1966	13,2	11,7	5,1
1967	1,1	3,0	4,8
1968	14,3	15,5	8,4
1969	11,4	10,8	9,0
1970	5,2	11,2	9,5
1971	8,6	10,4	12,2
1972	13,2	—	—

situam entre um máximo positivo de 12,1%, ocorrido no segundo trimestre de 1968, e um máximo negativo de -7,8%, ocorrido no segundo trimestre de 1965.

A série de renda real trimestral apresentada nesta nota poderá ser de grande utilidade, na medida em que ela venha a ser permanentemente atualizada, como um indicador para a política econômica, como uma variável independente em uma variedade de estudos econométricos e talvez mesmo como variável dependente, obedecidas as devidas cautelas.