

# **Crescimento do produto real brasileiro — 1900/1947.\***

Claudio Haddad

1. Introdução; 2. Os métodos de agregação empregados; 3. Os resultados; 4. Um teste mediante componentes principais.

## **1. Introdução**

O presente trabalho apresenta um resumo de nossas investigações quanto ao crescimento da economia brasileira na primeira metade do século XX.<sup>1</sup>

A seguir discutimos, sumariamente, os métodos de agregação empregados e a composição dos índices de produto real. Os principais resultados são apresentados na terceira parte do trabalho e, na quarta parte, as estimativas são testadas mediante o uso de componentes principais.

## **2. Os métodos de agregação empregados**

Os índices de produto calculados neste estudo são do tipo Divisia, também chamados índices de base móvel. Estes índices são definidos pelas

\* Agradeço aos Profs. A. Arberger, M. Nerlove e D. McCloskey pelo apoio e sugestões recebidos. Naturalmente os erros por ventura ainda existentes são de minha inteira responsabilidade.

<sup>1</sup> Para maiores detalhes ver Haddad, C. *Growth of Brazilian Real Output, 1900 — 1947*, Universidade de Chicago, June 1974. Tese de doutoramento.

suas variações percentuais. Em particular, um índice  $Q$  das quantidades  $q_i$  ( $i = 1, \dots, n$ ) é definido como

$$\frac{dQ}{Q} = \frac{\sum p_i dq_i}{\sum p_i q_i}$$

$p_i$  ( $i = 1, \dots, n$ ) sendo o preço relevante do bem  $i$ .

Conseqüentemente, o valor do índice no período  $t$  seria dado por

$$Q_t = Q_0 \exp. \int_c \frac{\sum p_i dq_i}{\sum p_i q_i}$$

$Q_0$  sendo o valor inicial (arbitrário) do índice.

Como no presente trabalho tratamos de períodos discretos, a fórmula acima reduz-se a

$$Q_t = Q_0 \cdot \frac{Q_1}{Q_0} \dots \frac{Q_t}{Q_{t-1}}$$

onde cada

$$\frac{Q_j}{Q_{j-1}} = \sum_i s_{i,j-1} \frac{q_{i,j}}{q_{i,j-1}}$$

com os diversos  $s_i$  representando as participações em valor do bem  $i$  no total do produto que, para efeito de cálculo, foram tomadas como sendo as participações no final o período  $j - 1$ . Desta maneira  $Q_t$  seria dado por uma cadeia de índices de Laspeyres para anos consecutivos.

Naturalmente, qualquer índice apresenta vantagens e desvantagens. A nossa escolha dos índices Divisia pode ser justificada pelos seguintes motivos:

- a) Eles evitam o problema da seleção de uma única cesta representativa de bens em relação a qual se mediria o crescimento econômico, o que é especialmente importante no nosso estudo, pois, evidentemente, a estrutura da economia brasileira variou sensivelmente de 1900 a 1947, tornando qualquer padrão único de medida extremamente arbitrário e sem muito sentido. Embora o problema de medirmos crescimento a longo prazo não seja solucionado com o uso dos índices Divisia, eles fazem mais sentido para comparações anuais, pois o crescimento é medido segundo o padrão do ano anterior. Como uma das principais aplicações das nossas séries de produtos seria na análise das flutuações cíclicas da economia, os índices Divisia são mais razoáveis sob este aspecto.
- b) Sob um ponto de vista operacional, é muito mais fácil trabalhar com os índices Divisia. Em primeiro lugar, o seu uso torna a agregação de

novos produtos simples e direta, pois cada variação de quantidade é ponderada pela participação do produto em questão no ano anterior, independentemente do número de produtos envolvidos no cálculo. Em segundo lugar, eles fazem com que o cálculo de um deflator implícito seja fácil e consistente. Seja  $Y_t = y_t P_t$ , onde  $P_t$  é a renda total nominal,  $y_t$  é a renda real  $P_t$  o deflator implícito no período  $t$ . Usando \* para denominar variações percentuais teremos, para cada período,  $Y_t^* = y_t^* + P_t^*(1)$ . Se tivermos  $Y_t^*$  e  $y_t^*$  e se  $y_t^*$  for um índice Divisia das quantidades produzidas, o cálculo de  $P_t^*$  é imediato. Como a taxa de inflação seria dada também sob a forma de um índice Divisia, o método é internamente consistente. Em outras palavras, trabalhando-se com índices Divisia a relação fica identicamente verdadeira para cada período  $t$ . Isto não aconteceria se usássemos índices de base fixa. Para validarmos a relação (1) teríamos de calcular um índice de preços que não seria nem de base fixa nem Divisia. Ou seja, o deflator implícito seria definido *ad hoc*, desprovido de qualquer interpretação econômica.

Uma aparente desvantagem dos índices Divisia é a de que, como eles são integrais de linha, seus valores dependem da trajetória de integração.<sup>2</sup> Isto significa que, em alguns casos, seu uso poderia levar a resultados enganosos. Tais casos, porém, são de difícil ocorrência na prática.<sup>3</sup> Além disso, os possíveis erros estariam certamente dentro do intervalo de confiança obtido ao se trabalhar com dados econômicos, especialmente brasileiros, para o passado.

Em vista disso, procuramos, sempre que possível, construir índices Divisia a partir dos dados originais. Quando dispúnhamos de informações sobre preços e quantidades, o cálculo era imediato. Na falta de dados sobre quantidades, interpolações foram feitas. Quando não dispúnhamos de dados sobre preços, as participações no valor foram extrapoladas ou interpoladas, de acordo com os casos abaixo:

(1) Um conjunto inicial de observações nos preços até o período  $i$  não era disponível. Neste caso, procedemos da seguinte maneira:

a) Primeiramente colocávamos as participações a partir de  $i$  em um gráfico a fim de investigarmos possíveis tendências, principalmente na vizinhança de  $i$ .

<sup>2</sup> Wold, H. *Demand Analysis*. New York, John Wiley & Sons, 1953. p. 132-9.

<sup>3</sup> Erros substanciais poderiam ocorrer em casos onde as variações anuais das quantidades fossem grandes e inversamente correlacionadas. Essa situação dificilmente ocorreria em estudos como o nosso.

b) Se alguma tendência era claramente distinguível, esta era estimada, em forma exponencial, por mínimos quadrados. O coeficiente de tendência era então aplicado na predição dos valores passados, antes de  $i$ .<sup>4</sup>

c) Se nenhuma tendência era discernível do período  $i$  ao período  $j$ ,  $j$  afastado de  $i$ , nós consideramos as participações antes de  $i$  iguais à média aritmética das participações de  $i$  a  $j$ .

(2) Dispúnhamos de observações para os períodos  $i$  e  $i + j$ , mas não dispúnhamos de observações para os  $j$  períodos intermediários.

O procedimento geral seguido nestes casos foi a interpolação linear das participações entre  $i$  e  $i + j$ .<sup>5</sup>

Na tabela 1 apresentamos os índices de produto real para os cinco setores da economia que serviram de base para o cálculo do produto total brasileiro, que também é apresentado na mesma tabela. A seguir discutimos, sumariamente, a composição dos índices.

## 2.1 Agricultura

Neste índice estão incluídos os seguintes itens:

a) Culturas: Arroz, feijão, milho, cana-de-açúcar, algodão (rama e caroço), fumo, café e cacau de 1900 a 1920. A partir de 1920 foram também incluídas bananas, batatas, mandioca, laranjas e trigo.

b) Produção animal: De 1900 a 1920 o índice é baseado no abate de gado (bovino, suíno, ovino e caprino). De 1920 em diante foram incluídos leite, lã e ovos.

c) Produção extrativa: Borracha foi o único produto incluído até 1920. Babaçu, castanhas, cera de carnaúba e erva-mate, foram adicionados a partir desta data.

Cada índice parcial foi agregado mediante a participação de cada subsetor no valor adicionado agrícola, igual à produção total menos consumo intermediário.

Levando-se em conta a participação de cada produto nos anos base, esperaríamos que o índice de produto agrícola fosse composto de cerca de 85% da produção total a partir de 1920, e 72% antes desta data.

<sup>4</sup> Aqui, uma observação de cautela. Como os valores das observações estão compreendidos entre 0 e 1, um modelo linear pode ser inadequado na predição de participações passadas. Felizmente, em todos os casos em que foi tentado o método não apresentou problemas.

<sup>5</sup> Talvez fosse melhor, sob o ponto de vista teórico, fazer interpolações geométricas. Mas sempre que ambas foram tentadas as diferenças foram muito reduzidas, não compensando o cálculo adicional.

**Tabela 1**  
**Índices setoriais de produto 1900 — 1947**  
**(1939 = 100,0)**

Ano	Agricultura	Indústria	Transportes e comunicações	Comércio	Governo	Produto total
1900	27,7	11,0		15,5		17,9
01	32,1	11,3		16,7		20,0
02	35,1	11,7		17,4		21,4
03	34,8	12,0		17,7		21,5
04	34,4	12,6		18,1		21,6
05	34,8	12,9		19,0		22,1
06	36,2	13,6		19,9		23,1
07	42,1	14,8		22,2		26,3
08	36,5	14,8	13,5	20,7	26,4	23,7
09	38,5	18,0	14,0	24,1	27,8	26,2
1910	40,3	18,8	16,7	25,9	31,1	28,1
11	37,3	20,5	18,3	27,0	32,0	28,2
12	41,4	22,7	20,8	30,2	34,2	31,2
13	40,5	22,9	22,2	30,1	38,7	31,7
14	43,2	20,9	23,0	27,3	43,3	32,1
15	42,8	23,6	23,0	28,3	35,3	31,7
16	44,6	26,3	25,7	30,8	29,4	33,1
17	46,2	28,6	29,3	31,8	30,6	34,9
18	47,6	28,3	32,3	31,9	30,6	35,6
19	47,0	32,5	32,3	36,5	30,9	37,7
1920	53,3	34,2	34,8	40,6	33,2	41,5
21	55,5	33,6	33,6	38,9	39,1	42,3
22	55,8	39,9	37,9	42,7	42,4	45,6
23	58,0	45,2	44,6	51,0	37,7	49,5
24	58,6	44,7	49,6	53,3	35,3	50,2
25	56,7	45,2	56,9	54,5	34,7	50,2
26	58,5	46,3	55,7	55,5	46,1	52,8
27	64,8	51,3	62,6	60,6	52,4	58,5
28	76,7	54,9	64,6	67,9	52,4	65,2
29	76,9	53,7	67,0	67,3	59,0	65,9
1930	77,8	50,1	57,4	61,5	67,4	64,5
31	72,9	50,7	60,2	58,4	67,0	62,4
32	77,3	51,4	55,1	59,2	78,5	65,1
33	86,6	57,4	59,7	68,2	70,4	70,9
34	92,0	63,8	62,1	74,5	82,5	77,4
35	89,7	71,4	70,9	78,9	78,6	79,7
36	98,2	83,7	78,0	89,2	86,0	89,3
37	98,3	88,2	84,8	94,3	95,1	93,4
38	102,4	91,5	92,7	96,2	102,6	97,6
39	100,0	100,0	100,0	100,0	100,0	100,0
1940	98,2	97,3	100,7	96,3	107,8	99,0
41	104,4	103,5	108,6	102,2	103,6	103,9
42	99,8	104,9	106,2	95,7	102,7	101,1
43	107,1	119,1	114,8	108,6	94,3	109,7
44	109,7	131,9	128,6	119,2	99,6	118,0
45	107,3	139,1	131,9	122,7	111,3	121,8
46	116,3	164,8	136,1	135,2	123,4	135,9
47	117,1	170,2	140,0	145,0	118,4	139,2

A mais séria omissão no cálculo do índice agrícola foi a da produção florestal (extração de madeiras, lenha e carvão vegetal). À medida que as atividades florestais foram proporcionalmente mais importantes no passado, o índice de produto total tenderia a subestimar a renda agrícola para o passado e, conseqüentemente, a superestimar o crescimento do produto.

## 2.2 Indústria

As atividades incluídas no índice industrial foram as seguintes:

- a) Manufaturas: De 1900 a 1907 a única manufatura incluída foi a de tecidos de algodão. De 1908 a 1911 incluímos tecidos de algodão e um índice de alimentos. De 1912 a 1921 foram adicionados outros tecidos, bebidas, produtos de fumo, sapatos, chapéus, fósforos, perfumaria e produtos farmacêuticos. Para o período 1921-25 foram adicionados siderurgia, móveis e gráficas. A partir de 1926, papel e couro foram incluídos; de 1927, cimento; de 1931, pneumáticos e, a partir de 1940, montagem de automóveis. Finalmente, no período 1945-47 foram incluídas todas as atividades acima, exceto produção de chapéus e móveis. Cada classe de manufaturas foi agregada mediante sua participação no valor adicionado total. Para o (valor adicionado / produção total), interpoladas através dos anos censitários.
- b) Mineração: De 1920 a 1940 o índice inclui produção de carvão, sal, manganês e ouro. Minério de ferro foi incluído a partir de 1941.
- c) Eletricidade e utilidades públicas: De 1902 a 1928 o índice foi baseado em uma média móvel de três períodos da energia total instalada. Para 1928 — 1939, o índice reflete o consumo total de energia elétrica no Rio de Janeiro e em São Paulo. De 1939 a 1947 o índice foi obtido de Loeb.<sup>6</sup>

Como no caso anterior, podemos estimar, a grosso modo, a participação dos produtos incluídos no índice no valor adicionado industrial total, levando em consideração os seus valores nos anos base. Para os períodos indicados estas participações seriam as seguintes: 1900-1907, 21%; 1907-1911, 42%; 1911-1921, 65%; 1921-1925, 68%; 1925-1939, 70%; 1939-1947, 69%.

<sup>6</sup> Loeb, Gustaaf. Números índice do desenvolvimento físico da produção industrial do Brasil, 1939-1949. *Revista Brasileira de Economia*, n. 7, p. 44-51, mar. 1953.

Deve-se enfatizar que as atividades que não entraram diretamente no cálculo do índice não foram negligenciadas. O valor adicionado das mesmas foi estimado mediante interpolação das suas participações no produto industrial, nos anos censitários.

### 2.3 Transportes e comunicações

As seguintes atividades foram diretamente incluídas no índice:

- a) Transportes: De 1908 a 1947 o índice inclui transporte ferroviário e marítimo.
- b) Comunicações: De 1908 a 1947 o índice reflete o volume de cartas e telegramas enviados através do Serviço de Correios e Telégrafos.

Em 1939 e em 1949 as atividades acima representaram cerca de 39% e 47% do total do setor, respectivamente. Como anteriormente, no cálculo do índice total, cada atividade foi ponderada pela sua participação no valor adicionado do setor.

### 2.4 Comércio

O índice de produto no setor comércio foi calculado como sendo uma média ponderada dos índices agrícola, industrial e de um índice de *quantum* de importações. Ainda hoje esta é a maneira pela qual o índice é estimado pelo Centro de Contas Nacionais da Fundação Getúlio Vargas.<sup>7</sup>

A justificativa desse procedimento é a de que deve-se esperar que a quantidade de bens comercializados no país seja uma função direta da sua produção interna, agrícola e industrial, mais os bens obtidos pela importação. Os pesos de ponderação foram obtidos via regressões *cross-section* entre os estados brasileiros, para os anos censitários de 1939 e 1949.

### 2.5 Governo

O índice de produto no setor público reflete os salários totais pagos nas três esferas de governo, federal, estadual e municipal.

Usando as nossas estimativas de valor adicionado conjuntamente com os dados da Fundação Getúlio Vargas, construímos a tabela 2, abaixo.

<sup>7</sup> Fundação Getúlio Vargas. *Contas nacionais do Brasil*. Rio de Janeiro, Fundação Getúlio Vargas, 1972. p. 22.

As participações da tabela 2 juntamente com a discussão acima sugerem que de 1939 a 1949 o índice de produto total seria composto de aproximadamente 68% da produção doméstica. Se a participação relativa dos setores (1) a (5), na tabela 2, tivesse permanecido em torno de 80% em períodos anteriores a 1939, em 1911 incluiríamos, no índice de produto, cerca de 60% das atividades produtivas.

**Tabela 2**

**Participações setoriais na renda nacional 1939 e 1949**  
(percentagens)

Setor	1939	1949
1. Agricultura	26,1	24,5
2. Indústria	21,9	26,5
3. Transportes e comunicações	6,2	7,0
4. Comércio	19,3	18,0
5. Governo	6,6	7,3
6. Aluguéis	3,8	3,4
7. Intermediários financeiros	4,9	4,4
8. Outros serviços	11,2	8,9
9. Total	100,0	100,0
10. Total de (1) a (5)	80,0	83,3

Fonte: Para o valor adicionado nos setores (1) a (5) ver Haddad. op. cit.

Para os outros setores ver Fundação Getúlio Vargas, *Contas nacionais do Brasil, novas estimativas*. Suplemento de *Conjuntura Econômica*, n. 24, jun. 1970.

### 3. Os resultados

Na figura 1 apresentamos, em escala aritmética, os índices de produto total e portanto total *per capita*, que se encontram reproduzidos na tabela 3.

O índice de produto real é uma média ponderada dos cinco índices setoriais acima discutidos: agricultura, indústria, transportes e comunicações, comércio e Governo, reproduzidos na tabela 1. Estes índices setoriais são reproduzidos na figura 2. O índice de produto total reflete a produção doméstica, não tendo sido feita nenhuma correção para variações nas relações de troca.

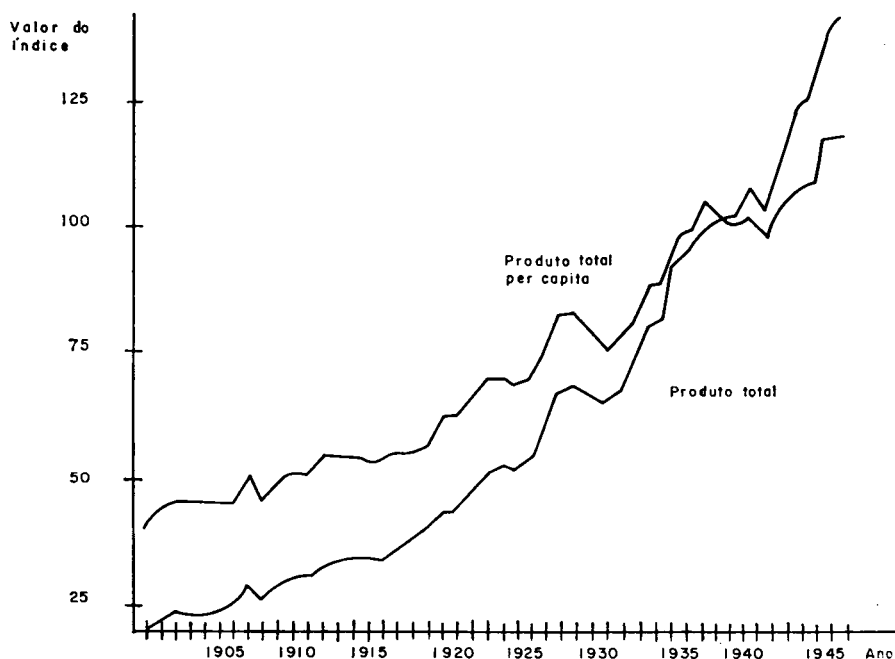


Na tabela 4 apresentamos taxas anuais de crescimento dos índices, para períodos selecionados. Os dados relativos a períodos de após guerra, foram obtidos da Fundação Getulio Vargas.<sup>8</sup>

Estimando-se tendências geométricas por mínimos quadrados para o período 1900-1947, obtém-se as seguintes taxas anuais de crescimento dos índices: agricultura, 3,03%; indústria, 5,7%; transportes e comunicações, 5,62%; comércio, 4,67%; Governo, 4,16%; produto total, 4,28%; população, 2,07%; produto total *per capita*, 2,20%.

Figura 1

### Produto total e *per capita*



<sup>8</sup> Fundação Getulio Vargas, 25 anos de economia brasileira. 1972. 2 p. Suplemento de *Conjunto Econômica*, n. 26, nov. 1972.

Tabela 3

Índice de produto real, população e produto real **per capita**,

1900-1947 (1939 = 100,0)

Ano	Produto real (1)	População (2)	Produto real per capita (3)	Ano	Produto real (1)	População (2)	Produto real per capita (3)
1900	17,9	44,6	40,1	1924	50,2	73,8	68,0
01	20,0	45,7	43,7	25	50,2	75,3	66,7
02	21,4	46,6	45,9	26	52,8	76,8	68,8
03	21,5	46,7	45,1	27	58,5	78,4	74,7
04	21,6	48,6	44,5	28	65,2	80,0	81,4
05	22,1	49,6	44,6	29	65,9	81,6	80,7
06	23,1	50,7	45,6				
07	26,3	51,8	50,9	1930	64,5	83,3	77,5
08	23,7	52,9	44,8	31	62,4	85,0	73,4
09	26,2	54,0	48,5	32	65,1	86,8	75,1
				33	70,9	88,5	80,0
1910	28,1	55,1	50,9	34	77,4	90,4	85,6
11	28,2	56,3	50,1	35	79,7	92,2	86,5
12	31,2	57,5	54,3	36	89,3	94,1	94,9
13	31,7	58,7	53,9	37	93,4	96,0	97,3
14	32,1	60,0	53,5	38	97,6	95,5	102,2
15	31,7	61,2	51,8	39	100,0	100,0	100,0
16	33,1	62,5	52,9				
17	34,9	63,9	54,7	1940	99,0	102,0	97,0
18	35,6	65,2	54,5	41	103,9	104,4	99,5
19	37,7	66,6	56,6	42	101,1	106,9	94,6
				43	109,7	109,4	100,2
1920	41,5	68,0	61,0	44	118,0	112,0	105,4
21	42,3	69,4	60,9	45	121,8	114,7	106,1
22	45,6	70,8	64,4	46	135,9	117,4	115,7
23	49,5	72,3	68,5	47	139,2	120,2	115,8

Figura 2

Comportamento dos índices de produtos setoriais

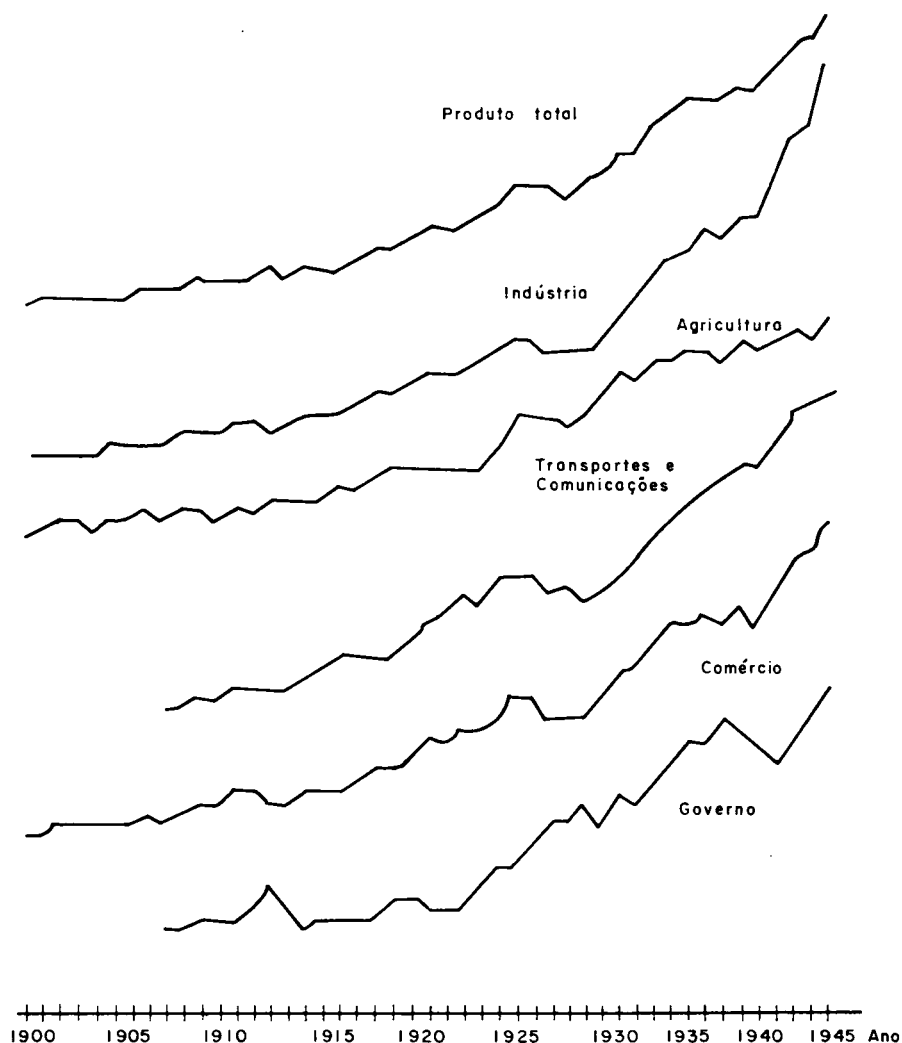


Tabela 4

## Taxa acumulada de crescimento (%) dos índices de produto

Período	Agri- cultura	Indústria	Transp. comun.	Comércio	Governo	Prod. total	População	Produto total <i>per capita</i>
Pré-guerra:								
1900-02 — 1910-12	2,31	6,25	—	5,37	—	3,96	2,14	1,74
1910-12 — 1920-22	3,29	5,65	6,66	3,92	1,66	3,97	2,12	1,75
1920-22 — 1930-32	3,30	3,51	5,00	3,89	6,40	4,04	2,05	1,95
1930-32 — 1940-42	2,86	7,20	6,20	5,09	3,96	4,71	2,08	2,56
1940-42 — 1945-47	2,42	9,18	5,27	6,49	2,36	5,49	2,38	3,02
1900-02 — 1945-47	2,88	6,03	5,85 <sup>a</sup>	4,74	3,75 <sup>a</sup>	4,30	2,12	2,15
Pós-guerra:								
1945-47 — 1969-71	4,39	8,31	8,05	5,90	7,13 <sup>b</sup>	6,45	3,32	3,03
Período total:								
1900-02 — 1969-71	3,56	6,82	6,74 <sup>c</sup>	5,17	6,23 <sup>d</sup>	5,05	2,54	2,46

<sup>a</sup> 1910-12 — 1945-47<sup>c</sup> 1910-12 — 1969-71<sup>b</sup> 1945-47 — 1965-67<sup>d</sup> 1910-12 — 1965-67

Fonte: As taxas de pré-guerra foram calculadas das tabelas 1 e 3, e as do pós-guerra a partir da Fundação Getúlio Vargas, *25 anos de economia brasileira*, suplemento especial. As taxas são centradas nos médios dos três anos base.

As nossas estimativas em conjunto com os valores dados pela Fundação Getúlio Vargas indicam, então, que a renda *per capita* brasileira cresceu a aproximadamente 2,2% na primeira metade deste século o que, juntamente com uma taxa de 3% ao ano para o período de pós-guerra até 1970, resultou em um crescimento de cerca de 2,5% ao ano no período 1900-1970.

O primeiro fato marcante a respeito das taxas apresentadas na tabela 4 é que elas são bem elevadas para padrões internacionais, como demonstramos na tabela 5.

Tabela 5

**Taxas anuais de crescimento da renda  
per capita de países seleccionados**

País	Período	Taxas acumuladas anuais de crescimento da renda per capita (%)
Estados Unidos	1840-1960	1,56
	1880-1920	1,61
	1920-1960	1,81
Grã-Bretanha (U.K.)	1881-1921	0,86
	1841-1957-59	1,20
Alemanha	1851-55-1958-60	1,43
Suécia	1881-85-1921-25	2,08
	1861-65-1958-60	2,47
Japão	1878-82-1918-22	3,05
	1878-82-1958-60	2,81
França	1913-1958-60	1,28
URSS	1913-1958	3,71

Fonte: Kuznets, Simon. Notes on the pattern of U.S. economic growth. In: *The reinterpretation of American Economic History*. Ed. by R.W. Fogel and S.L. Engerman. New York, Harper & Row, 1971. p. 18-9.

Como podemos observar, a taxa média de crescimento brasileiro é mais elevada que a americana, inglesa, francesa e alemã. É da mesma ordem de grandeza que a taxa de crescimento da Suécia, sendo inferior somente às taxas de crescimento do Japão e da URSS.

O segundo fato importante a notar na tabela 5 é que o crescimento brasileiro tem se acelerado no decorrer do século XX. De uma taxa média de crescimento *per capita* ligeiramente abaixo de 2%, de 1900 até os anos 30, a economia evoluiu para uma taxa de, aproximadamente 2,5% dos anos 30 ao começo da década de 50 e daí para 3% até o final dos anos 60. Como a taxa de crescimento do produto *per capita* tem se situado em torno de 6,5% de 1968 para cá, é possível que o desenvolvimento brasileiro tenha-se acelerado ainda mais, desta vez substancialmente.

Em vista dessas surpreendentes taxas de crescimento, a pergunta que imediatamente nos ocorre é: como é possível que o Brasil ainda hoje

seja um país subdesenvolvido, com uma renda *per capita* em torno de somente US\$ 550,00?

Só existem duas razões possíveis:

- a) que o crescimento tenha se iniciado de uma base muito pequena; e
- b) que o crescimento da economia antes de 1900 tenha se processado a taxas muito mais baixas.

Entretanto, um raciocínio simples mostra imediatamente que a primeira hipótese não é suficiente para explicar o fenômeno. Com uma taxa de crescimento anual de 2% um país dobra a sua renda *per capita* em aproximadamente 35 anos. Como as nossas estimativas indicariam uma renda *per capita* brasileira de aproximadamente US\$ 40,00 (a preços de 1947) em 1900-2, se a taxa de crescimento de 2% tivesse se mantido através do século XIX, encontraríamos um valor de US\$ 10,00 para a renda *per capita* em 1830 e um valor de cerca de US\$ 5,00 em 1800. Como esses valores são claramente abaixo do nível de subsistência, podemos concluir que as taxas de crescimento encontradas no século XX não poderiam ter sido sustentadas através do século XIX.

#### 4. Um teste mediante componentes principais

Infelizmente não dispomos de dados sobre pagamentos de fatores de produção em volume suficiente para testarmos as nossas estimativas sob um ponto de vista da renda. Entretanto, tentamos contornar este problema de duas maneiras e aqui apresentamos um teste baseado na técnica dos componentes principais, mediante o uso de funções de demanda estimadas.<sup>9</sup> A idéia é a seguinte:

Suponhamos que a função de demanda pelo bem  $i$  para o passado seja dada, em forma matricial, por

$$\underline{q}_i = P_i \eta_i + \epsilon_i \underline{y} + \underline{u}_i \quad (I)$$

onde,

$\underline{q}_i$  = vetor ( $n \times 1$ ) de quantidades demandadas do bem  $i$  nos diversos períodos.

$p_i$  = matriz ( $n \times k$ ) de preços relativos (preço do bem  $i$  e preços de bens complementares e substitutos).

<sup>9</sup> Um outro teste relativo ao produto agrícola, é apresentado em Haddad, op. cit. p. 24-6.

$\underline{y}$  = vetor ( $n \times 1$ ) de renda real em diferentes períodos.

$\underline{u}_i$  = vetor ( $n \times 1$ ) de distúrbios aleatórios.

$\underline{\eta}_i, \epsilon_i$  = parâmetros.

Considerando-se as variáveis sob forma logaritmica e admitindo-se as conhecidas propriedades de Gauss-Markov relativas às variáveis independentes e ao termo aleatório,  $\underline{\eta}_i$  e  $\epsilon_i$  representariam o vetor de elasticidade preço e a elasticidade renda, respectivamente. Admitindo-se que as quantidades demandadas sejam as consumidas do bem  $i$ , se dispusermos de um estimador  $\hat{\eta}_i$  para  $\underline{\eta}_i$ , poderemos definir resíduos  $\underline{r}_i$  como

$$\underline{r}_i = q_i - P_i \underline{\eta}_i = P_i (\underline{\eta} - \hat{\underline{\eta}}_i) + \epsilon_i \underline{y} + \underline{u}_i \text{ se } \hat{\underline{\eta}}_i$$

é não-tendencioso,

$$E\underline{r}_i = \epsilon_i \underline{y}, \text{ pois } E\underline{u}_i = 0 \text{ por hipótese.}$$

Se repetíssemos o mesmo procedimento para outros bens, definiríamos um conjunto de resíduos  $\underline{r}_i$  que poderiam ser arranjados numa matriz  $R$ , tal que

$$R = (\underline{r}_1, \underline{r}_2, \dots, \underline{r}_l)$$

e sendo o número de resíduos obtidos. Todas as colunas de  $R$  teriam um componente comum, renda real, multiplicados pelas respectivas elasticidades renda. Além disso, se os estimadores dos  $\eta_i$ 's forem próximos dos seus valores reais e se os distúrbios aleatórios  $\underline{u}_i$  forem não correlacionados, o efeito renda seria o único fator comum às colunas de  $R$ .

Antes de prosseguirmos, vamos discutir como poderíamos obter estimadores não-tendenciosos para  $\underline{\eta}_i$ .

A primeira idéia seria estimar  $\underline{\eta}_i$  aplicando-se mínimos quadrados à equação:

$$\underline{q}_i = P_i \underline{\eta}_i + \underline{u}_i \quad (\text{II})$$

Entretanto, como não estamos incluindo a renda real na estimativa, as elasticidades preço estimadas seriam tendenciosas. A tendência seria igual à elasticidade renda multiplicada pelo vetor de coeficientes obtidos pela regressão, em mínimos quadrados, da renda real no vetor de preços

relativos.<sup>10</sup> A fim de diminuirmos a tendenciosidade deveríamos incluir na equação (II) uma variável instrumental positivamente correlacionada com a renda real. Como esta última cresce através do tempo, uma variável tempo seria uma tal variável instrumental. Então, estimaríamos, ao invés da equação (II), a equação (III) a seguir:

$$\underline{q}_i = P_i \underline{y}_i + \gamma \underline{t} + \underline{u}_i \quad (\text{III})$$

onde,

$t' = (0, 1, \dots, n)$ . Estimando-se esta equação por mínimos quadrados, encontraríamos um estimador para  $\underline{\eta}_i$  que seria menos tendencioso que no caso acima.

Voltemos à matriz  $R$ , agora formada por resíduos  $\underline{r}_i$  obtidos pelas elasticidades preço estimadas como na equação (III). Se as colunas de  $R$  fossem estritamente proporcionais entre si, teríamos:

$$R = \underline{p} \cdot \underline{a}'$$

onde,

$\underline{a}'$  é um vetor linha de múltiplos escalares e  
 $\underline{p}$  é o componente comum.<sup>11</sup>

Como não temos uma proporcionalidade exata, obteríamos uma matriz de discrepâncias igual a  $R - \underline{p} \cdot \underline{a}'$ .

Para determinarmos  $\underline{p}$  necessitamos de um critério e de uma restrição. Seguindo Theil, e escolhendo como critério selecionar  $\underline{p}$  e  $\underline{a}$  tal que a soma dos quadrados das  $n \times 1$  discrepâncias é minimizado e impondo-se a restrição de que  $\underline{p}$  tenha variância unitária,<sup>12</sup>  $\underline{p}$  seria então o primeiro componente principal obtido da matriz  $R$ , o que explicaria uma maior proporção da variância de suas colunas.<sup>13</sup>

<sup>10</sup> O argumento é baseado em Theil, H. Specification errors and the estimation of economic relationships. *Review of the International Institute of Statistics*, n. 25 p. 41-51. 1957.

<sup>11</sup> A discussão que se segue é baseada em Theil, H. *Principles of Econometrics*. New York, John Wiley & Sons, 1971. p. 46-55.

<sup>12</sup> Theil. *Principles* ... impõe  $\underline{p}' \cdot \underline{p} = 1$ , ou seja, comprimento unitário. Variância unitária é mais de acordo com os programas de computadores existentes.

<sup>13</sup> Theil. *Principles*.



Conseqüentemente, deveríamos esperar que o primeiro componente principal obtido da matriz  $R$  refletisse, de perto, o comportamento da renda real estandarizada.<sup>14</sup> E, se as nossas estimativas representam a renda real “verdadeira” esperaríamos que, após estandarizá-las, elas se movessem no tempo de acordo com o primeiro componente principal. Em outras palavras, o primeiro componente principal obtido através dos resíduos de demandas, como explicado acima, serviria de um teste poderoso dos nossos resultados.

Foram então obtidas informações sobre quantidades consumidas, preços próprios e preços de complementos ou substitutos de 10 produtos para o período 1911-1939.<sup>15</sup> Reduzindo-se as quantidades consumidas a níveis *per capita*, deflacionando-se os preços pelo deflator implícito obtido pelas nossas estimativas,<sup>16</sup> tirando-se os logaritmos naturais de todas as variáveis e aplicando-se mínimos quadrados comuns à equação (III), acima indicada, estimamos o vetor de elasticidades preço. Os produtos e as variáveis independentes usadas encontram-se listadas na tabela 8.

Possuindo-se as elasticidades preço estimadas, computamos os diversos resíduos  $\underline{r}_i$  como,

$$\underline{r} = \underline{q}_i - P \hat{\underline{y}}$$

formando a matriz  $R$ . A seguir, extraímos o primeiro componente principal de  $R$ .<sup>17</sup>

<sup>14</sup> Como as quantidades dos bens estão expressas em unidades diferentes, o procedimento seguido pelo computador é primeiramente estandarizar todas as colunas de  $R$  e então retirar os componentes principais da nova matriz. Cada coluna de  $R$  seria então transformada para

$$(\epsilon_i y_i + \phi_i) \cdot \frac{1}{(\epsilon_i^2 S^2 + d_i^2)^{1/2}},$$

onde,

$\underline{y}_i$  é renda real estandarizada,

$S$  é o desvio padrão da renda,

$\phi_i$  é igual a  $P (\underline{y} - \hat{\underline{y}}) + \underline{u}_i$  em forma estandarizada

$d_i$  é o desvio padrão desta expressão.

<sup>15</sup> A nossa escolha dos produtos e do período foi ditada pela existência de dados sobre produção, importação e preços. As quantidades consumidas foram admitidas como iguais às quantidades produzidas mais a importação líquida.

<sup>16</sup> O deflator implícito é apresentado em Haddad, op. cit. tabela 76, p. 191. Restaria saber se o uso do nosso deflator implícito poderia tornar o teste tendencioso. Em princípio não vemos razão para tal, pois encontramos uma correlação muito elevada entre o nosso deflator e outros índices de preço existentes para o mesmo período como o apresentado por Villela e Suzigan, p. 424.

<sup>17</sup> O primeiro componente principal explicou 56% da variância dos resíduos.

Tabela 8

Variáveis usadas nas estimativas das funções de demanda

Produtos	Variáveis independentes
Arroz	Preço próprio, preço de feijão e de milho, tempo
Feijão	Preço próprio, preço de arroz e de milho, tempo
Milho	Preço próprio, preço de arroz e de feijão, tempo
Azeite de oliva	Preço próprio, preço de banha, tempo
Açúcar	Preço próprio, tempo
Bacalhau	Preço próprio, tempo
Farinha de trigo	Preço próprio, tempo
Calçados	Preço próprio, tempo
Derivados do fumo	Preço próprio, tempo
Tecidos de algodão	Preço próprio, preço geral dos tecidos, tempo

Tirando-se logaritmos naturais da renda *per capita* estandarizada e regressando o resultado no componente principal extraído da matriz  $R$ , obtivemos o resultado

$$\begin{aligned}\underline{y}_s &= 0,935 \underline{p} \\ &\quad (13,9) \\ R^2 &= 0,87 \quad n = 29 \\ DW &= 1,20\end{aligned}$$

$\underline{y}_s$  sendo o produto transformado e  $\underline{p}$  o componente principal. O número entre parênteses é a estatística  $t$  do coeficiente de correlação 0,935.  $\underline{y}_s$  e  $\underline{p}$  encontram-se plotados na figura 3. Podemos observar que, exceto para 1917, 1921 e 1939, os seus movimentos são extremamente semelhantes. Como 1917 e 1939 foram anos de guerra, durante os quais o consumo era restringido, o mau resultado para estes anos não é surpreendente.

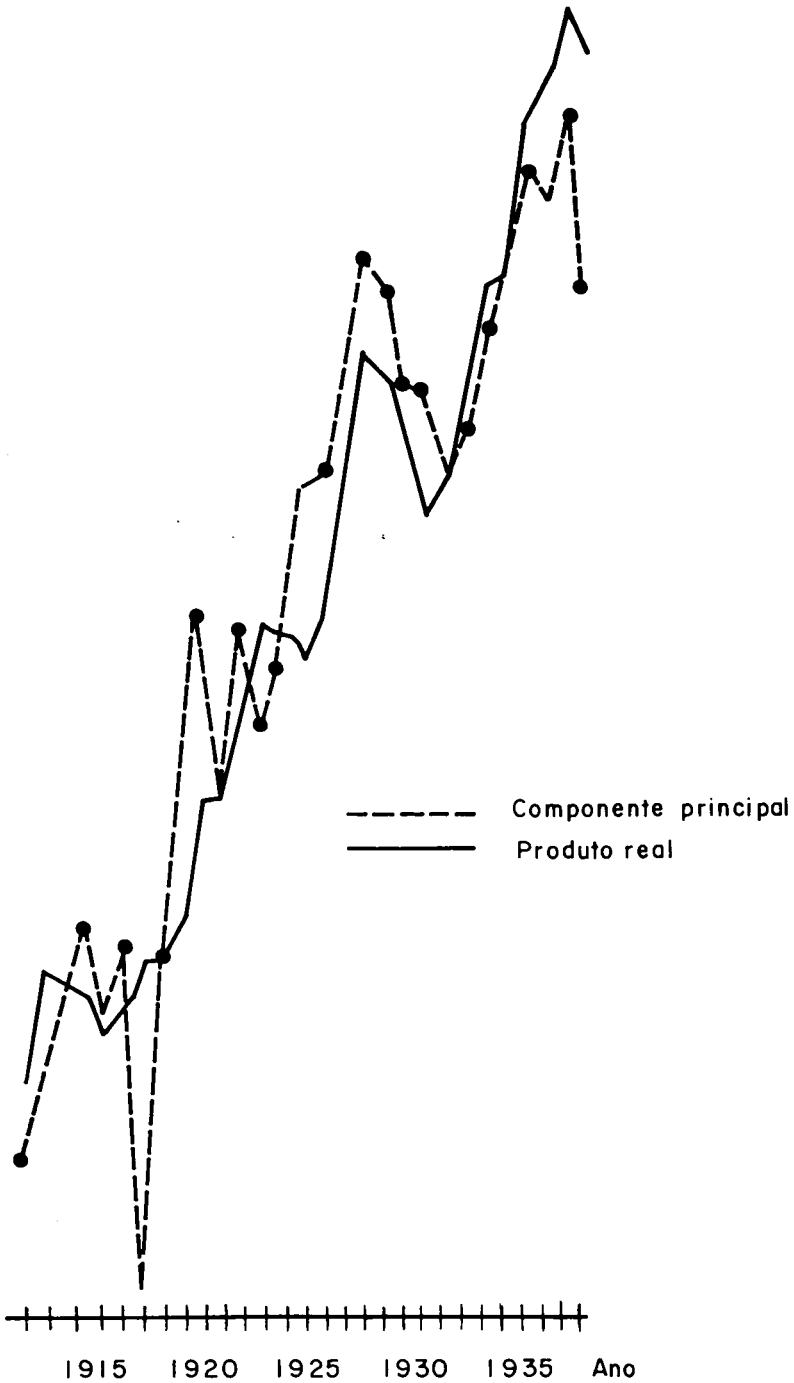
No geral, nossos resultados podem ser considerados muito bons. Mesmo se a nossa série de produto correspondesse à série “verdadeira” não obteríamos uma correlação perfeita entre  $\underline{y}_s$  e  $\underline{p}$  devido a erros de especificação das variáveis, problemas de identificação,<sup>18</sup> etc. Conseqüentemente, um coeficiente de correlação de 0,935, não diferente da unidade a um nível de significância de 10%, juntamente com um elevado  $R^2$ , podem ser considerados bons resultados.

Apesar destes resultados encorajadores, porém, é possível que, submetida ao mesmo teste, uma série “verdadeira” saísse bem melhor, caso

<sup>18</sup> A maioria dos produtos da tabela 8 eram muito comercializados internacionalmente, o que, por aumentar o grau de exogeneidade da variável preço, reduziria o problema de identificação.

Figura 3

Produto real estandartizado e o primeiro componente principal



em que os nossos resultados, sob um ponto de vista relativo, não seriam tão bons quanto aparentam ser.

Em virtude disso, resolvemos testar, sob condições semelhantes, uma tal série "verdadeira". Devido à facilidade de obtenção dos dados e à coincidência de período de análise, decidimos aplicar o mesmo teste às estimativas de renda real de Richard Stone para o Reino Unido, de 1920 a 1938.<sup>19</sup>

Novamente selecionamos dez produtos, com aproximadamente as mesmas características que os escolhidos previamente, e repetimos o procedimento anterior. Os produtos escolhidos e as variáveis independentes são listados na tabela 9

Tabela 9  
Variáveis empregadas nas estimativas das  
funções de demanda para o Reino Unido,  
1920 — 1938

Produtos	Variáveis independentes
Pão	Preço próprio, tempo
Ovos	Preço próprio, tempo
Carne de porco	Preço próprio, preço da carne de carneiro, tempo
Carne de galinha	Preço próprio, preço de carne de carneiro, tempo
Açúcar	Preço próprio, tempo
Chá	Preço próprio, preço de café, tempo
Mobiliário	Preço próprio, tempo
Calçados	Preço próprio, preço geral de roupas, tempo
Tecidos para roupas	Preço próprio, preço geral de roupas, tempo
Derivados do fumo	Preço próprio, tempo

Estandartizando-se a renda real *per capita* obtida por Stone<sup>20</sup> e regressando-a no primeiro componente principal obtido dos resíduos de demanda estimadas a partir da tabela 10, da mesma maneira que anteriormente, encontramos o seguinte resultado:

$$\underline{y}_t^* = 0,947 \underline{p}_t^* \\ (12,1)$$

$$R^2 = 0,90 \quad n = 19$$

$$DW = 1,24$$

<sup>19</sup> Stone, Richard. *The measurement of consumer's expenditure and behavior in the United Kingdom, 1920-1938*. Cambridge, University Press, 1954. 2 v.

<sup>20</sup> Reproduzida em Mitchell, B. R. *Abstract of British historical statistics*. Cambridge, University Press, 1962. p. 368.

onde,

$y_s^*$  é a renda *per capita* estandarizada de Stone e

$p^*$  é o primeiro componente principal da matriz de resíduos. Ambas as variáveis encontram-se plotadas na figura 4.

Repetindo-se o mesmo procedimento para a nossa série de produto, desta vez para o período 1920-1939, obtemos:

$$\underline{y_s} = 0,947 \underline{p} \\ (12,5)$$

$$R^2 = 0,90 \quad n = 20$$

$$DW = 1,63$$

Comparando-se as duas relações estimadas, podemos ver que elas são quase idênticas. O comportamento da renda estandarizada de Stone e o primeiro componente principal, apresentados na figura 4, é também semelhante ao encontrado no nosso caso. A renda e o componente principal movem-se quase coincidentemente, à exceção dos anos anormais de 1925 e 1930-1931.

Em conclusão, aplicando-se o mesmo teste para uma série reputada como “verdadeira” como a de Stone, encontramos resultados semelhantes aos nossos. Naturalmente, isso aumenta, de modo substancial, a confiança nas nossas estimativas.

Para finalizar, é mister fazermos uma comparação das nossas estimativas com estudos anteriores.

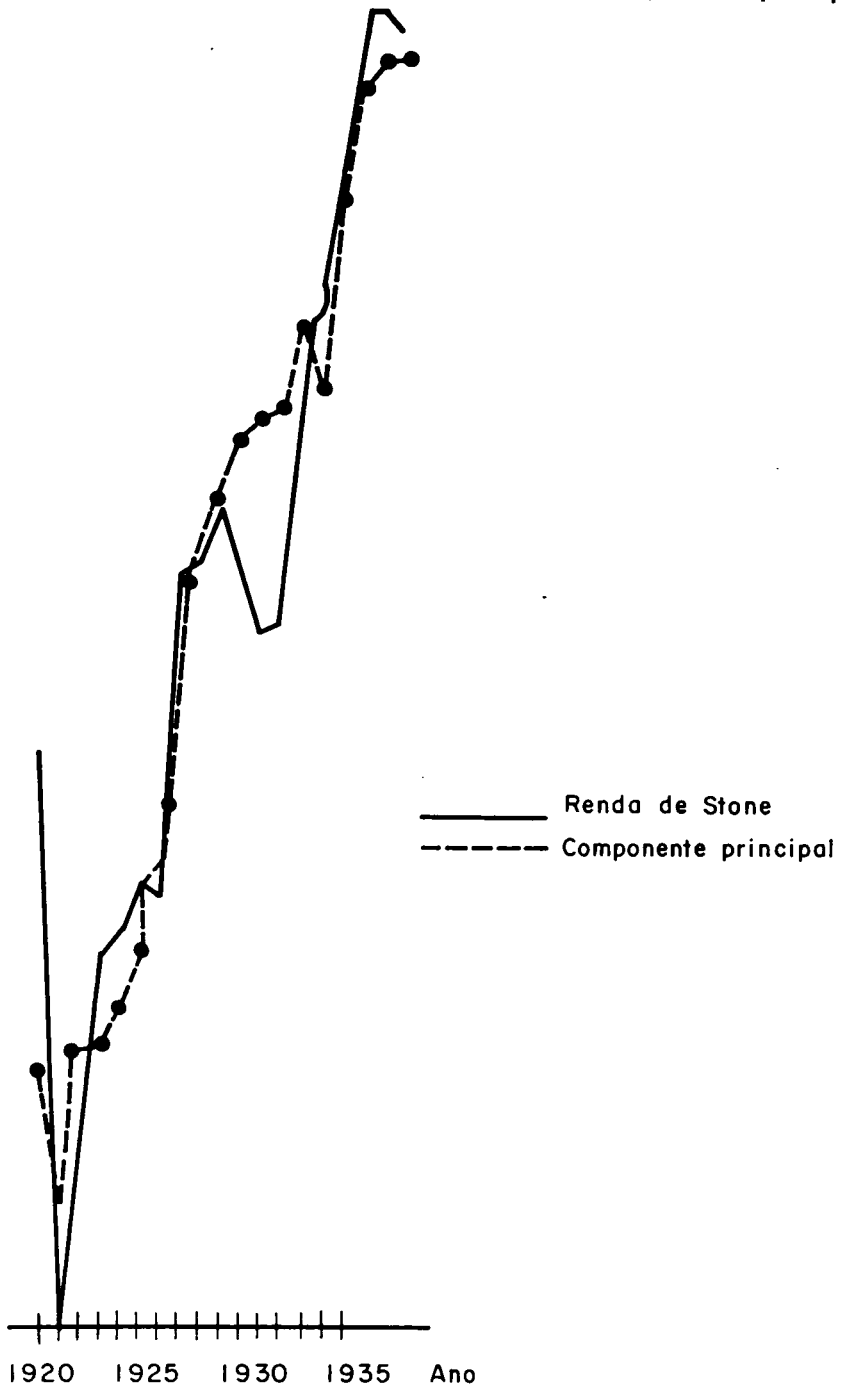
Segundo nosso conhecimento, já houve quatro tentativas sérias de estimação da renda brasileira para o passado. Estimativas de renda real foram obtidas por Loeb e Meiren para 1939-1953, pela Comissão Econômica para a América Latina (CEPAL) para 1920-1939, recentemente por Fishlow também para 1920-1939,<sup>21</sup> e por Villela e Suzigan para 1920-1945. As quatro estimativas são apresentadas na tabela 10, juntamente com a nossa série de produto real.

As estimativas feitas por Loeb e Meiren, Fishlow e Villela e Suzigan também foram feitas setorialmente. As da CEPAL foram baseadas em in-

<sup>21</sup> Loeb, Gustaaf & Van der Meiren, Pierre. O desenvolvimento da produção real e dos recursos disponíveis no Brasil, 1939-1953. *Revista Brasileira de Economia*, p. 141-58. dez. 1954. Para as estimativas da CEPAL ver Dias Carneiro, Octavio. *Past trends of structural relationships in the economic evolution of Brazil, 1920-1965*. Harvard, Harvard University, May 1966, mimeogr. Apresentado no Center for International Affairs. Fishlow, Albert. Origens e consequências da política de substituição de importações no Brasil. *Estudos Econômicos*, n. 2, dez. 1962; Villela, Annibal V. & Suzigan, Wilson. *Política do Governo e crescimento da economia brasileira 1889-1945*. Rio de Janeiro, IPEA/INPES, 1973.

Figura 4

Renda real estandartizada de Stone o primeiro componente principal



**Tabela 10**  
**Comparação com estimativas anteriores**  
**1920 — 1947**

Ano	Índices de produto real				
	CEPAL (1)	Loeb-Meiren (2)	Villela-Suzigan (3)	Fishlow (4)	Haddad (5)
1920	58,2		47,4	48	41,5
21	50,8		49,0	48	42,3
22	49,8		50,5	53	45,6
23	64,3		53,3	54	49,5
24	61,9		52,9	55	50,2
25	62,3		51,7	55	50,2
26	54,8		53,1	56	52,8
27	64,0		57,6	61	58,5
28	85,4		67,0	67	65,2
29	83,1		66,9	67	65,9
1930	72,5		66,4	66	64,5
31	68,3		64,4	65	62,4
32	68,5		67,9	68	65,1
33	80,5		74,9	74	70,9
34	84,2		80,0	80	77,4
35	92,7		81,4	83	79,7
36	94,3		90,4	92	89,3
37	99,6		92,5	95	93,4
38	97,8		96,9	98	97,6
39	100,0	100	100,0	100	100,0
1940		101			99,0
41		106			103,9
42		103			101,1
43		109			109,7
44		114			118,0
45		115			121,8
46		124			135,9
47		127			139,2

Fonte: Coluna (1) Carneiro, tabela I-A. Coluna (2) Loeb e Meiren, p. 142. Coluna (3) Villela e Suzigan, p. 431. Coluna (4) Fishlow, p. 64.

Índice de produto físico. Todos os índices são iguais a 100 em 1939.

dicadores de comportamento da economia como consumo de cimento, eletricidade, etc. À exceção de Loeb e Meiren e, em menor escala, por Villela e Suzigan, porém, os autores apenas indicam, em linhas gerais, como as estimativas foram obtidas, o que é um comportamento natural, pois em geral a estimativa da renda não era o principal objetivo de seus trabalhos. Em virtude disso uma cuidadosa comparação das nossas estimativas com as anteriores é uma tarefa extremamente difícil, se não impossível.

Entretanto, podemos abordar o problema de outra maneira, deixando de lado os detalhes da estimação e concentrando-nos nos resultados finais. Dentro da mesma linha de raciocínio, podemos aplicar o teste dos componentes principais às diversas estimativas para o período 1920-1939.<sup>22</sup>

Transformando as séries da tabela 10 em produto *per capita*, estandarizando-se os resultados e reintegrando-os no componente principal obtido anteriormente para o mesmo período, encontramos os resultados sumariados na tabela 11.

Tabela 11

Resultados das regressões das estimativas de produto  
no componente principal — 1920 — 1939

Estimativas	Coefficiente de correlação e valores <i>t</i>	<i>R</i> <sup>2</sup>	<i>D W</i>
Haddad	0,947 (12,5)	0,90	1,63
ECLA	0,775 ( 5,2)	0,60	1,61
Fishlow	0,920 ( 9,9)	0,85	0,88
Villela-Suzigan	0,884 ( 8,0)	0,78	1,02

Podemos imediatamente observar que as nossas estimativas se comportam melhor no teste, em qualquer padrão de comparação escolhido. Isto é pelo menos uma indicação de que as nossas séries corresponderiam mais de perto à série de renda “verdadeira” e, conseqüentemente, deveriam ter maior poder explicativo em análises macroeconômicas da economia brasileira.

<sup>22</sup> Como foi difícil a obtenção de dados para as funções de demanda no período 1939-1947, foi impossível a comparação das nossas estimativas com as de Loeb e Meiren por componentes principais. Só podemos dizer que, em geral, incluímos um maior número de produtos nas nossas estimativas setoriais comuns. Eles incluem dois setores adicionais: Serviços e Aluguéis. O produto no primeiro setor foi estimado considerando-se constante a produtividade média no setor. Desta maneira o índice refletiria o número de pessoas empregadas no setor, interpolado através dos anos censitários. Aluguéis rurais também refletem o volume da população rural. É nossa opinião que a inclusão dos dois setores adicionais faria com ue as estimativas tendessem a subestimar o crescimento do produto real.