

Curvas de Engel e transformação de Box-Cox: uma aplicação aos dispêndios em alimentação e educação na cidade de São Paulo *

JOSÉ ADELINO SOUZA MEDEIROS **

1 — Introdução

A técnica de transformação de variáveis de Box-Cox será utilizada na estimação de Curvas de Engel. Estamos interessados em obter a especificação e os parâmetros de funções representativas do dispêndio em alimentação e em educação. Foram selecionados esses dois itens para confrontar dois bens diferenciados (respectivamente, bem de primeira necessidade e bem superior).

Através da transformação Box-Cox podemos reduzir o grau de arbitrariedade quando necessitamos escolher a função apropriada. Os próprios dados, além de utilizados na obtenção das estimativas dos parâmetros, indicam a melhor especificação através da maximi-

* Uma versão preliminar deste trabalho foi apresentada nos seminários realizados pelo IPE-USP (Série "Trabalhos para Discussão" n.º 18/76). Agradecemos ao Dr. Denisard C. de Oliveira Alves pela ajuda na realização deste e pelo incentivo que nos foi dado. Fomos beneficiados também com as sugestões do Dr. Joanílio R. Teixeira e do Dr. Marcos J. A. P. Pacca. Somos gratos a Elcúterio F. da S. Prado e ao Dr. Derli C. Machado da Silva, que colaboraram na versão final do trabalho. As imperfeições remanescentes são de nossa responsabilidade.

** Do Instituto de Pesquisas Espaciais, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), São José dos Campos, São Paulo.

zação da verossimilhança. A partir de dados de orçamentos familiares para a cidade de São Paulo (período de 1971 a 1972), faremos aplicações numéricas mostrando as potencialidades da técnica.

Para os itens estudados (alimentação e educação), de acordo com a técnica utilizada, veremos que as formas especificativas resultantes e as elasticidades-renda terão comportamentos distintos. Para simplificar a exposição, descreveremos os resultados esperados para a função que representa o dispêndio em alimentação, mas o raciocínio pode ser facilmente transposto para a educação. Os dados de orçamentos familiares na cidade de São Paulo encontram-se na Tabela 9 (no Apêndice). A hipótese mantida nesta análise é que, na média, as diferenças nos padrões de consumo, entre famílias ricas e pobres, sejam provocadas por suas alterações na renda corrente e as outras diferenças sejam estocásticas. Em outras palavras, estamos supondo que as famílias sejam razoavelmente homogêneas em cada classe de renda.

Considerar-se-ão as recomendações de Brown e Deaton,¹ ignorando-se os efeitos da renda passada do consumidor, a história do consumo, o efeito das expectativas e o efeito da poupança. Assim, a variável renda será tratada como se fosse idêntica ao dispêndio total do consumidor. Um outro fator que contribui para a não utilização da variável renda é que, na análise de orçamentos familiares, esta usualmente contém mais erros de medida. Então, vários fatores contribuem para que a variável explicativa seja o dispêndio total, e embora dispuséssemos de dados sobre renda não os utilizamos. Apesar de eliminar os erros de medida, isso introduziu um viés porque não há independência entre a variável explicativa, que utilizaremos, e o termo aleatório. Essa dependência é provocada pela determinação simultânea do dispêndio total e do dispêndio no item considerado. Acreditamos, contudo, que o efeito líquido seja favorável à utilização do dispêndio total. Um argumento adicional é dado por Philips,² que defende a utilização dos dispêndios totais

¹ A. Brown e A. Deaton, "Models of Consumption Behavior: A Survey", in *The Economic Journal*, vol. LXXXII, n.º 328 (dezembro de 1972), pp. 1145-1236.

² L. Philips, *Applied Consumption Analysis* (Amsterdã: North Holland, 1974), p. 111.

que estão sob o controle da família, e que são, provavelmente, um melhor indicador do que a renda observada, a qual inclui toda espécie de componentes transitórios (positivos ou negativos).

A seguir, fornecemos as características da variável explicada (dispêndio no bem). Se o consumo fosse medido em quantidades físicas, as dificuldades encontradas seriam de difícil superação. Devido às diferenças de unidades de medida e à variedade de modos em que os bens são colocados no mercado, nem mesmo um esquema sofisticado de números índices daria completa solução ao problema. Dessa forma, a variável dependente é o dispêndio monetário no bem. Contudo, assim procedendo, também surgem problemas. Fendt³ mostra a conseqüência de utilizar esse dado, pois quando consideramos as classes de renda mais altas não é medido apenas um aumento na quantidade consumida, mas também a maneira como varia a qualidade do produto adquirido. Essa desvantagem, porém, não é de maior gravidade do que as deficiências citadas, se forem utilizadas as quantidades físicas. Sabemos, também, que o problema se reduz quando consideramos o dispêndio monetário em grupos de bens (ao invés de tomados isoladamente) e este é o caso que nos interessa, pois analisaremos a alimentação no domicílio e a educação.

Definimos as duas variáveis que constam na função-dispêndio: o dispêndio em alimentação e o dispêndio total. Como os preços são mantidos constantes, essa relação é a conhecida Curva de Engel, que pode ser encarada como uma curva de demanda quando os preços são fixados. O próximo passo é mencionar as propriedades que essa função deve ter. Como os preços são constantes, o problema é bastante simplificado. Desaparecem as restrições da teoria da demanda no que se relaciona com: a) homogeneidade de grau zero das funções nos preços e renda, isto é, ausência de ilusão monetária; b) simetria na substituição entre bens; e c) negatividade do efeito-substituição. A única restrição que permanece é a consistência do modelo ou o critério da adição (*adding up*), o qual exige que a soma dos dispêndios, nos vários bens (ou grupos de bens) que

3 R. Fendt Jr., "Padrões de Consumo na Guanabara: Uma Análise Econômetrica de Sistemas Alternativos de Curvas de Engel" (Rio de Janeiro: Ministério do Planejamento e Coordenação Geral/IPÊA, 1970), p. 41, mimeo.

compõem a cesta do consumidor, seja igual ao dispêndio total (renda). Uma outra forma de ver esta condição é obter a soma ponderada da elasticidade-renda. Este somatório deve ser igual a um, para cada classe de renda, e os vários pesos considerados são as razões-dispêndio no bem sobre o dispêndio total. É oportuno notar que seria mais apropriado utilizar o termo “elasticidade-dispêndio”, ao invés de “elasticidade-renda”. Contudo, com as observações feitas anteriormente, estes dois valores se igualam. Após o sumário das principais condições a atender, vejamos as leis de Ernst Engel e as evidências empíricas de interesse para este estudo:

- a) a alimentação é o item mais importante no orçamento familiar;
- b) a proporção do dispêndio alocada para a alimentação cresce quando a renda aumenta;
- c) a proporção do dispêndio utilizada em vestuário e habitação permanece aproximadamente constante; e
- d) a proporção do dispêndio utilizada em bens de luxo (superiores) cresce quando a renda aumenta.

Em adição aos aspectos citados, podemos esperar que as Curvas de Engel, que forem estimadas, também sejam compatíveis com as evidências empíricas usuais. Citamos a seguir algumas delas:

- a) o crescimento da renda implica a diversificação da cesta do consumidor;
- b) existem bens que só são consumidos após determinado nível de renda;
- c) o comportamento da Curva de Engel é diferenciado para os diversos tipos de bens (por exemplo, comparando bens de primeira necessidade com bens superiores, os primeiros apresentam pontos de saturação ou saciedade de forma mais evidente); e
- d) a elasticidade-renda deve variar à medida que alteramos o nível de renda do consumidor.

Das evidências empíricas acima, exploramos um pouco mais a variabilidade da elasticidade-renda. Fendt⁴ discute esse aspecto e

⁴ *Ibid.*

afirma que, supondo elasticidade constante para todos os consumidores (e para todos os estratos de renda, se os dados forem agrupados de alguma forma), não é possível a análise da variação da elasticidade de mercado com variações diferenciadas no crescimento da renda. Brown e Deaton⁵ também discutem o comportamento esperado para a elasticidade-renda e afirmam que existem numerosas evidências empíricas que justificam a proposição de que, para um amplo intervalo de bens, as elasticidades-renda são funções decrescentes da renda.

E a elasticidade variável (geralmente decrescente) é compatível com outras duas características já citadas: a) a de que novos bens passam a ser adquiridos; e b) a saturação do nível de demanda. É bom frisar que a elasticidade constante não é incompatível com as leis de Engel anteriormente citadas. Contudo, se for possível obter funções que apresentem um bom ajuste e variabilidade na elasticidade-renda, estaremos mais próximos da realidade, pois sabemos que, na prática, havendo uma variação na renda, o consumidor apresenta "sensibilidade" diferente no consumo de determinado bem, dependendo de seu nível de renda. Em resumo, a elasticidade-renda variável acentua os efeitos das leis de Engel e permite analisar o comportamento de mercado em resposta às variações no crescimento da renda.

A obtenção de funções que atendam a todas as características citadas é tarefa difícil. Quando satisfazem o critério da adição (por exemplo, as especificações lineares), não apresentam as evidências empíricas desejáveis, e portanto o ajuste estatístico é bastante precário. Philips⁶ reforça essa posição afirmando: "... muito tem sido ganho em termos de poder descritivo, muito tem sido perdido em termos de plausibilidade teórica. Introduzindo alterações mais realísticas na elasticidade-renda, Prais e Houthakker perderam contato com a teoria da maximização da utilidade. Sem dúvida, não mais existe qualquer referência a uma função utilidade específica. E, muito mais, as especificações adotadas não são compatíveis com

⁵ Brown e Deaton, *op. cit.*, p. 1.173.

⁶ Philips, *op. cit.*, p. 113.

a maximização da utilidade, pois elas não satisfazem exatamente o critério da adição. A abordagem é inteiramente pragmática”.

Estão sendo feitas pesquisas que visam a obter funções que apresentem um bom ajuste estatístico e satisfaçam o critério da adição, como, por exemplo, Houthakker,⁷ que utiliza a função utilidade indireta, e Leser.⁸ Este autor afirma: “... o problema de encontrar a forma da Curva de Engel mais apropriada é antigo em econometria, e ainda não apareceu uma solução de aceite geral. Genericamente falando, talvez seja válido afirmar que a especificação das formas das relações tem merecido menos atenção que os métodos de estimar os parâmetros para equações especificadas”.

Em artigo que discute o estado atual da teoria do consumidor, Houthakker⁹ é de opinião semelhante, e afirma que é necessário obter funções de demanda apropriadas para poder colocar a análise empírica da demanda em bases mais firmes e, também, a fim de contribuir para os seus aspectos estatísticos.

2 — Transformação de variáveis (transformação Box-Cox)

Vejamos as principais funções não lineares que melhor descrevem o comportamento das Curvas de Engel, apesar de não satisfazerem o critério da adição. Embora as funções lineares sejam compatíveis com esse critério, fornecemos dois fatos que implicam a não utilização destas especificações: de um lado, verificamos que, para muitos bens, existe um nível de renda abaixo do qual não há consumo; em outros casos, é possível detectar um nível de saturação para o consumo, que atua como limite superior, qualquer que seja o nível

⁷ H. S. Houthakker, “Additive Preferences”, in *Econometrica*, vol. XXVIII, n.º 2 (abril de 1960), pp. 244-257.

⁸ C. E. V. Leser, “Forms of Engel Curves”, in *Econometrica*, vol. XXXI, n.º 4 (outubro de 1963), pp. 694-703.

⁹ H. S. Houthakker, “The Present Stage of Consumption Theory”, in *Econometrica*, vol. XXIX, n.º 4 (outubro de 1961), pp. 704-740.

de renda. Colocando de lado a função linear, surgem várias especificações que apresentam um bom ajustamento aos dados:

$$\lg X_A = a + b \lg X_T + u \quad (1)$$

$$X_A = a + b \lg X_T + u \quad (2)$$

$$\lg X_A = a - b \frac{1}{X_T} + u \quad (3)$$

$$\lg X_A = a - \frac{b}{X_T} - c \lg X_T + u \quad (4)$$

$$\frac{X_A^{\lambda_1} - 1}{\lambda_1} = a + b \frac{X_T^{\lambda_2} - 1}{\lambda_2} + c \frac{X_T^{\lambda_3} - 1}{\lambda_3} + u \quad (5)$$

onde X_A representa o dispêndio em alimentação, X_T o dispêndio total, u é a variável aleatória e λ_1 , λ_2 e λ_3 são parâmetros de transformação.

Quando o bem analisado é superior, isto é, com elasticidade-renda maior que um, usualmente escolhe-se a função duplo-logaritmo — expressão (1). Quando a elasticidade é próxima de um, é usada a especificação linear. Se o bem aparece como superior, quando o nível de renda é baixo, e como de primeira necessidade, em níveis de renda mais altos, utiliza-se a forma mono-logaritmo — expressão (2). Contudo, esta função não possui valor de saturação, e sua elasticidade decresce continuamente em direção a zero. Então, para bens cuja demanda está próxima da saturação, a função escolhida é a log-inversa — expressão (3). Além dessas especificações, Zarembka¹⁰ cita a especificação log-log-inversa — expressão (4). Utilizando esta função, em níveis baixos de renda o bem é superior, à medida que a renda cresce torna-se de primeira necessidade e, finalmente, o bem é inferior. As elasticidades-renda, neste caso, são maiores que um e, por último, negativas. Contudo, não é comum obter este tipo de especificação porque se trabalha com dados agregados.

¹⁰ P. Zarembka, "An Econometric Analysis of the Food Consumption Function", in P. Zarembka (ed.), *Toward a Theory of Economic Development* (São Francisco: Holden Day, 1972), pp. 194-215.

A seguir, discutiremos a técnica de transformação de variáveis, introduzida por Box-Cox¹¹ e, em aplicação à econometria, formalizada por Zarembka.¹² Daremos as idéias gerais, já considerando as aplicações a serem feitas. A idéia fundamental da técnica é a introdução de parâmetros adicionais para indicar a forma especificativa apropriada e possibilitar que a elasticidade-renda possa variar, quando os níveis de renda se alteram. Tal inovação é mostrada através de uma relação linear entre uma transformação potência do dispêndio em alimentação (variável dependente) e a mesma transformação potência do dispêndio total (variável independente). A expressão (5) reflete esse tipo de transformação, sendo que está sendo considerado um caso amplo, pois permitimos que a variável dispêndio total sofra duas transformações, através dos parâmetros λ_2 e λ_3 .

A partir da expressão (5) podemos obter todas as outras expressões. Se o coeficiente de regressão (5) for nulo e se $\lambda_1 = \lambda_2 = 0$, obtemos a equação (1). Basta lembrar que, quando λ_1 tende para zero, o primeiro membro da expressão (5) tende para $\lg X_1$. Para obtermos o segundo membro de (1) procedemos da mesma forma. A expressão (2) é obtida quando $c = 0$, $\lambda_1 = 1$ e $\lambda_2 = 0$. Neste caso é necessário utilizar o argumento discutido anteriormente (quando λ_2 tende para zero) e redefinir os parâmetros de regressão, pois, quando $\lambda_1 = 1$, o novo intercepto linear de (2) é $a + 1$ (para simplificar a notação, deixamos o valor original de a ; na aplicação numérica, devemos alterá-lo). Se $c = 0$, $\lambda_1 = 0$ e $\lambda_2 = -1$, a partir de (5) obtemos a expressão (3), utilizando os mesmos argumentos discutidos anteriormente e adaptando-os para o presente caso. A expressão (4) resulta de (5) se $\lambda_1 = \lambda_3 = 0$ e $\lambda_2 = -1$, utilizando o mesmo procedimento adotado para obtermos as expressões (1) a (3). A especificação linear também é obtida de (5), bastando fazer $c = 0$, $\lambda_1 = 1$ e $\lambda_2 = 1$ e seguir o raciocínio descrito nos casos anteriores.

¹¹ G. E. P. Box e D. R. Cox, "An Analysis of Transformations", in *Journal of Royal Statistical Society (B)*, vol. XXVI, n.º 2 (abril de 1964), pp. 211-252.

¹² P. Zarembka, "Transformation of Variables in Econometrics", in P. Zarembka (ed.), *Frontiers in Econometrics* (Nova York: Academic Press, 1974), pp. 81-104.

A título de ilustração antecipamos um dos resultados que obtivemos neste estudo e fazemos comparações com duas especificações usuais: a especificação linear (ver Gráfico 1) e a especificação duplo-logaritmo — expressão (2) — que está no Gráfico 2. Utilizamos dados dos dispêndios em alimentação (ver Apêndice), e o valor ótimo do parâmetro de transformação que obtivemos¹³ é $\lambda_1 = \lambda_2 = \lambda = -0,85$ — para simplificar fizemos $c = 0$ e consideramos $\lambda_1 = \lambda_2$ (ver Gráfico 3). A Tabela 1 resume as principais estimativas.

TABELA 1

Resultados de estimações linear, duplo-logaritmo, utilizando-se a técnica de transformação de variáveis

Tipo de Especificação	Intercepto	Coefficiente	t	R ²	d
Linear	61,82	0,11	7,49	83,60	0,53
Duplo-Logaritmo	1,66	0,50	17,49	96,53	0,87
$\lambda = -0,85$	-0,38	1,31	41,93	99,38	2,13

FONTE: Regressões do tipo mencionado, utilizando-se dados da Tabela 9 (no Apêndice) para a alimentação.

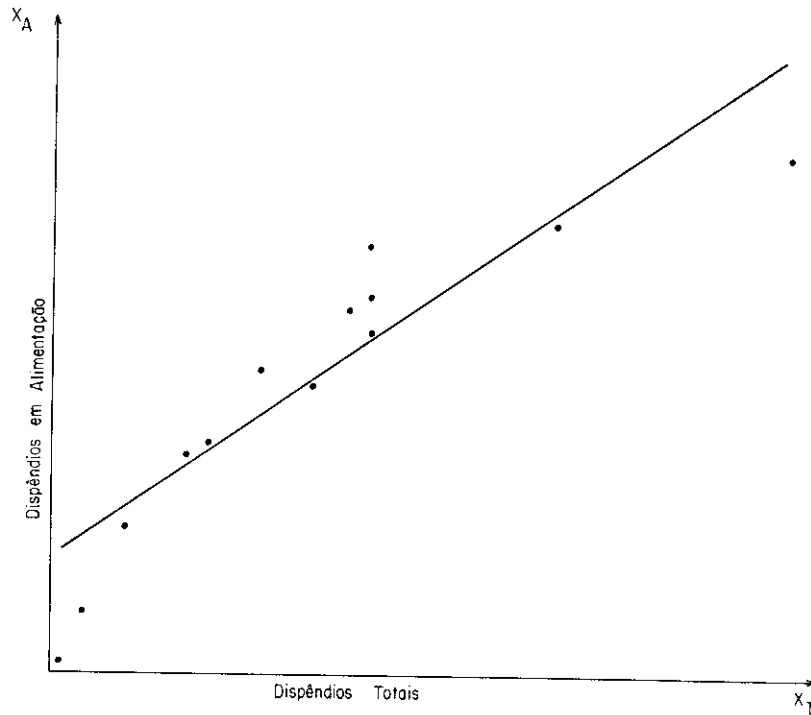
Faremos a seguir algumas observações relacionadas com a Tabela 1. Tomando a expressão (5) e considerando $c = 0$ e $\lambda_1 = \lambda_2 = 1$, obtemos os mesmos resultados da especificação linear, exceto o intercepto, que se altera (como já discutimos); o valor do intercepto é 60,931. Os resultados apresentados pela especificação duplo-logaritmo foram os mesmos obtidos pela transformação de variáveis, tomando (5) e considerando $c = 0$ e $\lambda_1 = \lambda_2 = \lambda = 0,00001$.

Analisando o coeficiente de determinação, chegamos à conclusão que o valor ótimo da transformação de variáveis ($\lambda_1 = \lambda_2 = \lambda = -0,85$) proporciona o R^2 maior, considerando as outras duas estimativas da Tabela 1. Contudo, é importante realçar que, na técnica de transformação de variáveis, o R^2 não é um indicador

¹³ Posteriormente descreveremos o procedimento adotado para obter o valor ótimo de λ .

Gráfico 1

DISPÊNDIOS EM ALIMENTAÇÃO – ESPECIFICAÇÃO LINEAR

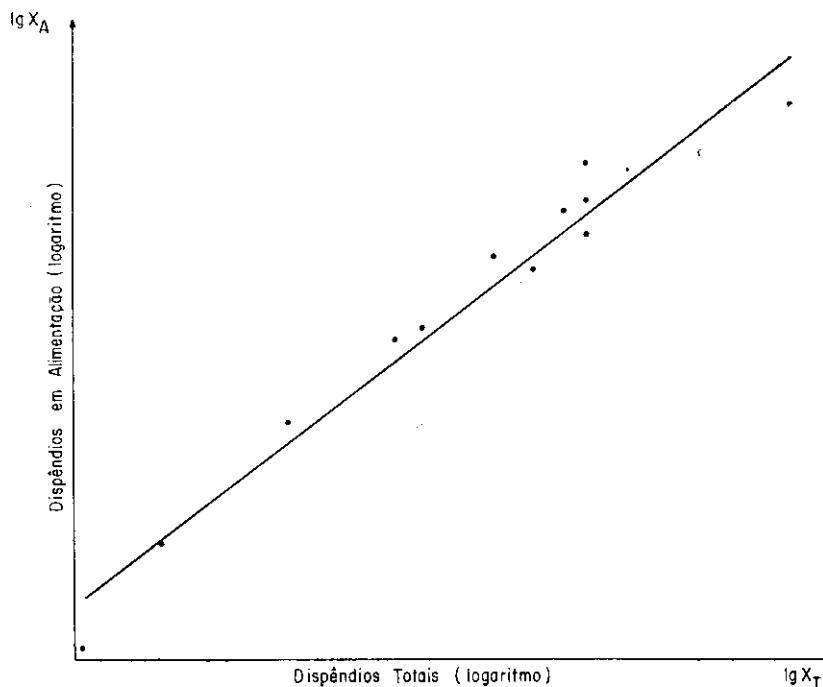


relevante para mostrar a forma especificativa apropriada. Conforme veremos posteriormente (por exemplo, na Tabela 3) o λ ótimo ($\lambda = -0,85$) não corresponde ao maior R^2 . Um outro aspecto relacionado com o coeficiente de determinação é o fato de ser excessivamente alto, devido ao nível de agregação dos dados. Praticamente todos os sistemas mostram um bom ajustamento dos dados. Devemos, então, analisar esse indicador com maior cuidado. Conforme salienta Johnston,¹⁴ "... a perda de eficiência quando nos movemos de dados originais para dados agrupados depende da maneira do agrupamento, e finalmente podemos notar que o R^2 computado a partir dos dados agrupados freqüentemente pode ser

¹⁴ J. Johnston, *Econometric Methods* (Tóquio: McGraw-Hill, 1972), p. 230.

Gráfico 2

DISPÊNDIOS EM ALIMENTAÇÃO – ESPECIFICAÇÃO DUPLO-LOGARITMO



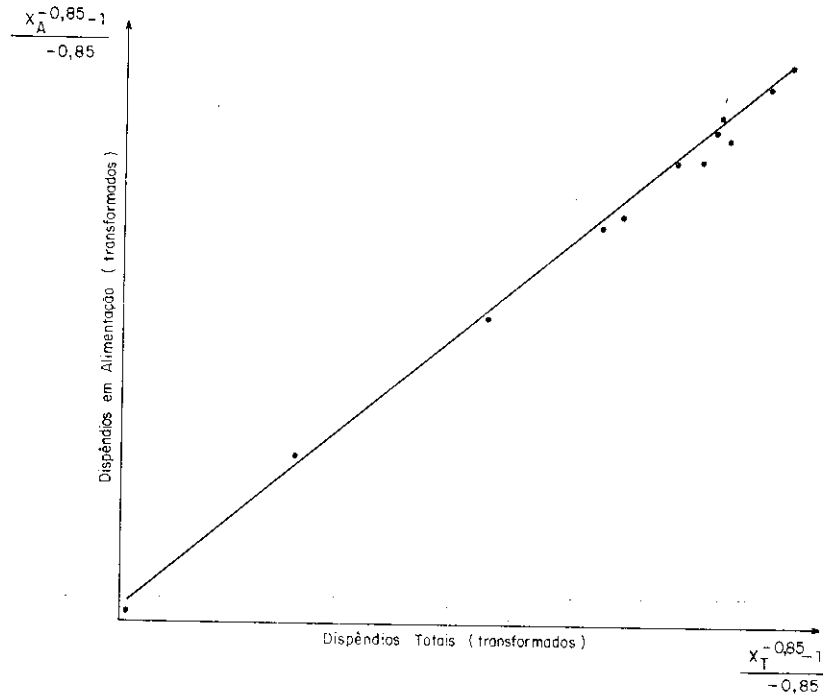
substancialmente maior que o R^2 para os dados individuais e pode ser um indicador enganoso”.

Nas análises econométricas que vêm sendo feitas, a opção tem sido por formas especificativas mais flexíveis e por maior ajustamento estatístico. Em comparações internacionais realizadas por Houthakker,¹⁵ foram adotadas funções duplo-logaritmo. Kirsten,¹⁶ a partir de dados da pesquisa-piloto de orçamentos familiares realizada de maio a junho de 1971 pelo IPE/USP para a cidade de

¹⁵ H. S. Houthakker, “An International Comparison of Household Expenditure Patterns, Commemorating the Centenary of Engel’s Law”, in *Econometrica*, vol. XXV, n.º 4 (outubro de 1957), p. 260.

¹⁶ J. T. Kirsten, *Metodologia de Construção de Índices de Preços ao Consumidor – Custo de Vida*, Série Monografias (São Paulo: IPE/USP, 1975), p. 70.

DISPÊNDIOS EM ALIMENTAÇÃO –
 ESPECIFICAÇÃO USANDO TRANSFORMAÇÃO DE VARIÁVEIS
 ($\hat{\lambda} = -0,85$)



São Paulo, efetuou as estimações para as funções linear, mono-logarítmica e duplo-logarítmica e comenta: "... analisando os resultados obtidos, verificou-se que o poder descritivo das regressões é, para alguns grupos de componentes, bem próximo nas três especificações adotadas e, em outros grupos, uma particular especificação se destaca em relação às demais".

Para os itens que nos interessam (alimentação e educação), em ambos os casos a forma especificativa que Kirsten escolheu foi a duplo-logarítmica, que apresentava melhores R^2 , t e d em relação às outras duas funções. Uma outra aplicação foi realizada por

Fendt,¹⁷ que, analisando o consumo para o antigo Estado da Guanabara, utilizou as mesmas formas especificativas já citadas (duplo-logaritmo e mono-logaritmo). Mais adiante faremos as comparações com esses dois autores.

Utilizando a transformação Box-Cox, nenhuma das formas especificativas é rejeitada *a priori*. Como vimos, as formas especificadas em (1), (2), (3) e (4), assim como a aritmética, são casos particulares de uma forma funcional mais geral dada pela expressão (5). Os próprios dados são utilizados para indicar a função apropriada e, posteriormente, proporcionam as estimativas dos parâmetros. Além das expressões (1) a (4) e linear, podemos obter outras (dependendo dos valores λ) não ficando restrito aos casos tradicionais.

A nossa preocupação fundamental reside na análise estatística da escolha da forma especificativa e na satisfação das evidências empíricas anteriormente citadas. Do ponto de vista da teoria do consumidor não houve contribuição. No entanto, achamos oportuno mencionar o artigo sempre atual de Haavelmo,¹⁸ que analisa as "hipóteses *a priori* sobre um fenômeno real". No caso em discussão, embora difícil, não é impossível que essas hipóteses (critério da adição no sentido em que é utilizado) sejam modificadas a partir do crescimento das evidências empíricas, indicando formas especificativas inconsistentes com essas hipóteses. Assim, devemos considerar a possibilidade de a formulação teórica tornar-se inadequada (pelo menos em alguns aspectos), quando testada contra os dados escolhidos como verdadeiras variáveis.

No nosso estudo empírico restringiremos a análise ao caso particular em que $c = 0$ e $\lambda_1 = \lambda_2 = \lambda$. Isto implica considerarmos uma única transformação para a variável explicativa e eliminarmos os casos em que as transformações para as variáveis explicada e explicativa sejam diferentes, isto é, $\lambda_1 \neq \lambda_2$. Desta forma, considerando essas simplificações, a partir de (5) obtemos a expressão (6), que

¹⁷ Fendt Jr., *op. cit.*

¹⁸ T. Haavelmo, "The Probability Approach in Econometrics", in *Econometrica*, vol. XII, n.º 1 (janeiro de 1944), pp. 9 e 144.

servirá de base para a análise (a expressão (7) é mera repetição da expressão (6), em notação menos trabalhosa):

$$\frac{X_A^\lambda - 1}{\lambda} = a + b \frac{X_T^\lambda - 1}{\lambda} + u \quad (6)$$

$$X_A^{(\lambda)} = a + b X_T^{(\lambda)} + u \quad (7)$$

Embora no presente estudo façamos as simplificações mencionadas, poderíamos, com trabalho adicional, utilizar a expressão (5) e obter famílias de funções tais que as do tipo mono-logaritmo e log-inversa seriam casos particulares — expressões (2) e (3), respectivamente. Zarembka¹⁹ fornece a metodologia necessária para efetuar essas extensões. No caso da educação acreditamos que, se utilizássemos transformações diferentes, as melhorias não seriam significativas, pois esse bem apresenta características de bem superior. O mesmo não ocorrerá com a alimentação (item de primeira necessidade). Neste caso, à medida que a renda cresce, as características de saturação são mais evidentes, pelo que uma função tipo mono-logaritmo é mais representativa, implicando transformações diferentes para as variáveis dispêndio em alimentação e dispêndio total. Apesar dessas limitações, as aplicações numéricas, que efetuamos nas Seções 4 e 5, já conseguimos mostrar uma sensível melhora em relação aos procedimentos usuais.

A idéia fundamental da transformação Box-Cox é a introdução de um parâmetro adicional, λ , para indicar a forma especificativa apropriada e possibilitar que a elasticidade-renda possa variar, quando os níveis de renda se alteram. Tal inovação é mostrada através de uma relação linear entre uma transformação potência do dispêndio em alimentação (variável dependente) e a mesma transformação potência do dispêndio total (variável independente) como está mostrado na expressão (6). A variável aleatória, u , é normalmente distribuída, com média zero e variância constante, independente de X_T . No entanto, é bom lembrar que a hipótese de normalidade é uma

¹⁹ Zarembka, "Transformations of . . .", *op. cit.*, p. 83, e "An Econometric . . .", *op. cit.*, p. 208.

aproximação, pois, como realça Zarembka,²⁰ “mesmo que a suposição de normalidade de u exija que a variável dependente varie em todo o campo real, a aproximação à normalidade é boa, se a probabilidade dos valores da variável dependente, na região excluída, for pequena”.

Como dissemos anteriormente, os próprios dados são utilizados para indicar o valor ideal de λ , não sendo necessário impor nenhuma forma especificativa em particular. No caso da alimentação, é razoável esperar que a elasticidade-renda se situe entre 0 e 1 e que decresça quando o dispêndio total aumentar. Para isso ocorrer é necessário que o valor de λ seja, necessariamente, inferior a zero, o que é confirmado nas aplicações que fizemos. Construiremos, também, intervalos de confiança para λ que reforçam os resultados obtidos.

Tínhamos a preocupação de não poder voltar às variáveis originais, e se isso acontecesse haveria dificuldade em obter a elasticidade-renda, que é a principal estimativa de que necessitamos. Contudo, como será mostrado abaixo, transformações algébricas simples indicam o valor da elasticidade que interessa (η):

$$\eta = \frac{\partial X_A}{\partial X_T} \cdot \frac{X_T}{X_A} = \hat{b} \left(\frac{X_T}{X_A} \right)^{\hat{\lambda}} \quad (8)$$

Para obter esse resultado, basta derivar a expressão (6) conforme indicado em (8). Estamos utilizando derivada parcial porque, mesmo quando possuímos mais de uma variável explicativa (como, por exemplo, a introdução de economias de escala que comentaremos na Seção 4), o resultado se mantém. As influências das variáveis explicativas (ou variável explicativa) são captadas através de \hat{b} (estimativa do coeficiente de regressão de X_T). $\hat{\lambda}$ representa o valor ótimo de λ . É importante notar que, à medida que a renda aumenta, isto é, que passamos para classes de renda mais altas ou, o que é equivalente, quando cresce o dispêndio total, a elasticidade-renda diminui (lembrando que estamos efetuando a análise para a ali-

²⁰ P. Zarembka, “Functional Forms in the Demand for Money” in *Journal of American Statistical Association*, vol. LXIII (junho de 1968), p. 504.

mentação; para outros itens o comportamento pode ser diferente). Um aumento de renda implica um acréscimo em X_T . E η diminui porque $d\eta/dX_T < 0$, pois $\hat{\lambda} < 0$ e os outros componentes da expressão (9) são positivos:

$$\frac{d\eta}{dX_T} = \hat{\lambda} \hat{b} \frac{1}{X_A} \left(\frac{X_T}{X_A} \right)^{\lambda-1} < 0 \quad (9)$$

Quando utilizamos a especificação duplo-logaritmo, o valor de λ tende para zero e a expressão acima se anula. Portanto, verificamos que a elasticidade é constante para todos os níveis de renda.

Para obter o λ ótimo, partimos da expressão (6) e utilizamos a abordagem de máxima verossimilhança. O logaritmo neperiano da função de verossimilhança, em termos das N observações originais, é expresso por:

$$\lg V = - \frac{N}{2} \lg \hat{\sigma}_u^2(\lambda) + (\lambda - 1) \sum \lg X_A \quad (10)$$

onde $\hat{\sigma}_u^2(\lambda)$ representa a variância estimada da disturbância da regressão $X_A^{(\lambda)}$ sobre $X_T^{(\lambda)}$. Nas tabelas das Seções 4 e 5 utilizaremos o termo "erro-padrão da estimativa", fornecido pelo programa de computação usado²¹ — esse valor é a raiz quadrada de $\hat{\sigma}_u^2(\lambda)$. O somatório se refere a todas as observações da variável considerada (varia de 1 a N). Todos os pormenores, inclusive a obtenção do jacobiano da transformação da variável dependente, podem ser vistos em Zarembka.²²

A partir dos diversos valores de λ , pertencentes ao intervalo de confiança estabelecido (que discutiremos no final desta seção) e utilizando o valor correspondente de $\hat{\sigma}_u^2(\lambda)$, obtemos diversos valores de $\lg V$ e, possivelmente, convergimos para o máximo $\lg V$, que indica o λ ótimo e, portanto, a forma especificativa apropriada. Nos vários resultados apresentados neste trabalho (Tabela 3 e seguintes), sempre convergimos para um λ que maximizasse o valor de verossimilhança. Todavia, não há garantia de que isso sempre

²¹ Bourroughs, *Programa BASIS* (Detroit: Bourroughs, 1971).

²² Zarembka, "Transformations of...", *op. cit.*, pp. 85-86.

ocorra, e pode acontecer, inclusive, que os valores encontrados sejam máximos locais. Mas, considerando que os resultados numéricos obtidos são satisfatórios, quando comparados com estudos semelhantes, podemos temporariamente descuidar desses aspectos. Uma indicação de pesquisas futuras é investigar o comportamento da função de verossimilhança e verificar a possibilidade de obter o λ ótimo sem ter que estimar várias regressões.

Em relação à não-normalidade da variável aleatória, o procedimento descrito é robusto, e Zarembka²³ realça que o efeito da forma distribucional é secundário, quando modelos complexos se ajustam satisfatoriamente. Contudo, o mesmo não ocorre quando existe heterocedasticidade. Existe um viés na estimação de λ , na direção que estabiliza a variância do erro. Neste caso, os esforços visam a obter estimativas de λ consistentes, sob hipóteses relativas ao nível de heterocedasticidade. Entretanto, é possível obter informações sobre a magnitude do viés. Na seção seguinte analisaremos em detalhe o problema da heterocedasticidade.

Devemos, ainda, investigar as implicações decorrentes do procedimento de transformar as variáveis originais. Ao verificar a disposição dos pontos nos Gráficos 2 e 3, e tomando como base o Gráfico 1, notamos que os pontos se “deslocam” para a direita. É provável que isso ocorra porque estamos utilizando a mesma transformação para ambas as variáveis (explicada e explicativa), e os dados considerados (alimentação) seriam melhor representados se as transformações fossem diferentes, pois este item apresenta uma tendência de saturação para os níveis de renda mais altos. Em pesquisas posteriores pretendemos estudar o caso das transformações diferentes para ambas as variáveis. Contudo, apesar da limitação citada, a parte empírica deste estudo apresentou resultados satisfatórios.

Mostramos, a seguir, a forma de obter o intervalo de confiança para λ . Utilizando a expressão (10) e variando os valores de λ , obtemos uma série de valores para $lg V(\lambda)$.²⁴ Aplicando o método da razão de verossimilhanças,²⁵ sabemos que o logaritmo dessa ra-

²³ *Ibid.*, pp. 88-95.

²⁴ Nesta parte utilizaremos a notação $lg V(\lambda)$ ao invés de $lg V$.

²⁵ Ver Box e Cox, *op. cit.*, p. 216, e Zarembka, “An Econometric...”, *op. cit.*, p. 210.

zão é assintoticamente distribuído com $\frac{1}{2} \chi^2$ (meio qui-quadrado), com um grau de liberdade (que corresponde ao número dos parâmetros de transformação; no caso em estudo temos um parâmetro). Esse resultado é dado pela expressão (11), onde $\hat{\lambda}$ é o valor ótimo de λ , isto é, aquele valor que maximiza $lg V(\lambda)$:

$$lg V(\hat{\lambda}) - lg V(\lambda) \leq \frac{1}{2} \chi^2(0,05) \quad (11)$$

Para o nível de confiança escolhido de 0,05, o valor de $\frac{1}{2} \chi^2(0,05)$ com um grau de liberdade é 1,92. Com esses ingredientes é imediata a obtenção do intervalo de confiança para λ , ao nível de 5%. Os limites desse intervalo (λ_a e λ_b) são obtidos nos pontos em que a expressão (11) se torna uma igualdade. Considerando que a função $lg V(\lambda)$ tem o formato indicado no Gráfico 4, a igualdade mencionada ocorre nos pontos A e B.

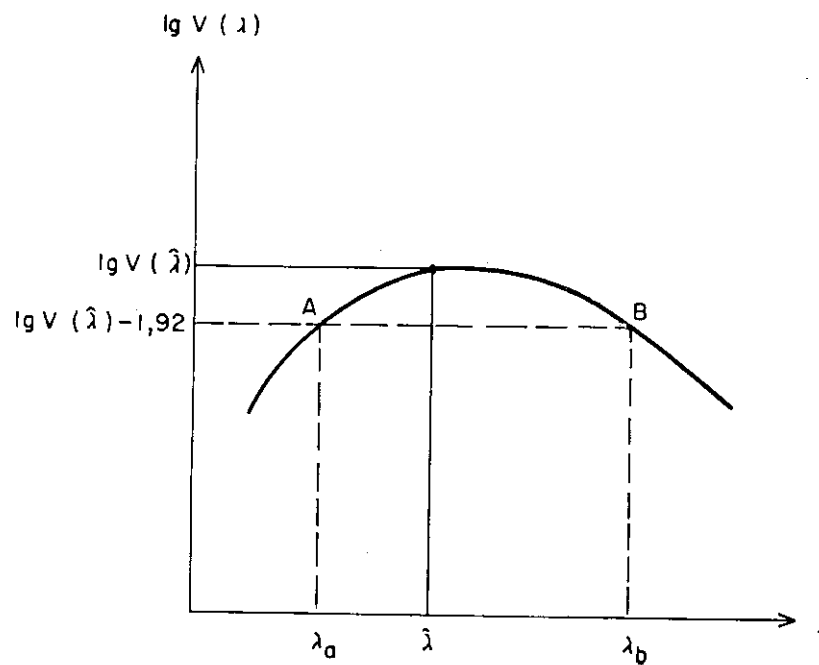
A partir dos pontos A e B obtemos, respectivamente, λ_a e λ_b , que definem o intervalo de interesse. Notar que a função $lg V(\lambda)$ usualmente não é simétrica e, portanto, obtemos um intervalo para λ não simétrico. Na parte empírica deste estudo fornecemos vários intervalos de confiança para λ , utilizando o procedimento ora discutido. Os resultados obtidos permitem mostrar as potencialidades da técnica de transformação de variáveis.

Como veremos nas aplicações numéricas, entre as várias estatísticas apresentadas, fornecemos também a estatística de Durbin-Watson. No contexto que nos interessa ela é útil porque, à medida que convergimos para a transformação ótima ($\lambda = \hat{\lambda}$), o valor de Durbin-Watson (d) aproxima-se de 2. Portanto, a hipótese $d = 2$ não é rejeitada, e há conseqüentemente fortes razões para acreditar que os resíduos não estão autocorrelacionados. Este fato é um dos indicadores de que a forma especificativa é apropriada, pois sabemos que a especificação inadequada é uma das fontes de autocorrelação. Kirsten²⁶ justifica a utilização do teste de Durbin-Watson para o caso que nos interessa e argumenta: "... o cálculo de auto-

²⁶ Kirsten, *op. cit.*, p. 70.

Gráfico 4

OBTENÇÃO DO INTERVALO DE CONFIANÇA PARA λ



correlação da componente residual tem validade para séries temporais. Entretanto, como a variável dispêndio total (tomada como variável explicativa) está ordenada, utilizou-se o teste de Durbin-Watson em *cross-section*.

3 — Agrupamento dos dados

Ao trabalhar com dados agrupados, devemos ter o cuidado de efetuar algumas ponderações, de forma a manter as características dos dados originais. Sabemos que, com o agrupamento, a hipótese de homo-

cedasticidade usualmente não é atendida. À medida que consideramos classes de renda mais alta, a variância dos desvios cresce com a renda, porque varia o padrão de comportamento do consumidor (uma diversificação maior) e, principalmente, porque acreditamos que o número de observações seja menor nas classes de renda mais alta. As estimativas continuam não viesadas, mas se tornam ineficientes.

Para contornar o problema, adotamos as sugestões de Zarembka²⁷ e, antes de qualquer transformação de variáveis, efetuamos a ponderação de todas as variáveis através da multiplicação pela raiz quadrada do número de indivíduos na observação agrupada — variável I da Tabela 9 (no Apêndice). Esta é a consequência prática da aplicação do método de mínimos quadrados generalizados para os dados agrupados, a fim de ser corrigida a heterocedasticidade do termo errático, devida ao agrupamento.

Procedemos da forma descrita e efetuamos as ponderações recomendadas. Na Tabela 2 verificamos que o λ que maximiza a função de verossimilhança se situa entre $-1,10$ e $-0,80$. Utilizando o método de mínimos quadrados ordinários (ver Tabela 3), o λ ótimo é $-0,85$, que pertence ao intervalo obtido com o método de mínimos quadrados generalizados.

TABELA 2

Regressões utilizando o método de mínimos quadrados generalizados

λ	Verossimilhança ^a
-0,60	567
-0,80	821
-1,00	883
-1,10	321

FONTE: Regressões $(X_A \sqrt{I})^{(\lambda)} = \hat{a} + \hat{b} (X_T \sqrt{I})\lambda$, utilizando-se dados da Tabela 9 (no Apêndice).

^a Multiplicado por 10^{-38}

²⁷ Zarembka, "An Econometric...", *op. cit.*

Embora possa ocorrer que o método generalizado melhore os resultados apresentados nas Seções 4 e 5, acreditamos que, considerando a forma como foram divididas as classes de renda, que constam em nossa amostra, as melhorias não são significativas. Os números de indivíduos que constam em cada uma das classes de renda consideradas (que são os pesos utilizados) não diferem de modo relevante. Assim, não existe uma tendência decrescente acentuada, para a variável que serve como peso, e o processo de correção da heterocedasticidade fica prejudicado. Então, para a aplicação numérica apresentada neste trabalho, optamos pela utilização do método de mínimos quadrados ordinários. Contudo, é recomendável que as classes de renda sejam subdivididas em maior número e de forma conveniente, para introduzirmos o método generalizado e melhorarmos os resultados. Não efetuamos essas alterações para não incorrer em custos adicionais de computação. Por outro lado, nossos resultados, quando comparados com aqueles obtidos pelos métodos tradicionais, já conseguem satisfazer a um número maior de evidências empíricas.

4 — Resultados para a alimentação

Através da maximização da função de verossimilhança obtivemos $\hat{\lambda} = -0,85$, compatível com o comportamento esperado para a elasticidade-renda, no caso da alimentação. O intervalo para λ , com coeficiente de confiança de 95%, é definido pelos limites $-1,20$ e $-0,50$. Para obter estes valores utilizamos o procedimento fornecido no final da Seção 2 e os dados para a função de verossimilhança da Tabela 3. Nesta tabela mostramos as estimativas usuais em regressão e fornecemos as elasticidades-renda para cada um dos valores de λ indicados. É possível notar a convergência para o λ ótimo. Indicamos as elasticidades para as classes de menor e de maior renda. A Tabela 4 detalha as elasticidades-renda para as 13 classes de renda consideradas, para o λ ótimo e para os dois imediatamente próximos.

Acreditamos desnecessários maiores comentários, visto que na Seção 2 já discutimos as contribuições que a transformação de variáveis

TABELA 3
 Resultados para os dispêndios em alimentação (I): $X_A = f(X_T)$

λ	Intercepto	Coefficiente	Erro- Padrão da Estimativa	R^2	t^a	d	Verossimilhança ^d	Elasticidade- Renda
-1,20	-0,68	1,81	0,00017	99,38	41,89	1,70	12,08	0,81 a 0,16
-1,10	-0,60	1,66	0,00025	99,40	42,69	1,87	34,48	0,80 a 0,18
-1,05	-0,56	1,58	0,00031	99,41	42,93	1,95	43,64	0,78 a 0,19
-1,00	-0,52	1,51	0,00039	99,41	43,00	2,02	45,69	0,76 a 0,20
-0,95	-0,47	1,44	0,00048	99,41	42,88	2,08	63,67	0,76 a 0,21
-0,85 ^b	-0,38	1,31	0,00076	99,38	41,93	2,13	69,55 ^c	0,74 a 0,24
-0,80	-0,33	1,25	0,00097	99,45	41,08	2,12	60,43	0,73 a 0,23
-0,60	-0,08	1,02	0,00268	99,13	35,48	1,85	20,38	0,68 a 0,31
-0,40	-0,25	0,82	0,00806	98,66	28,44	1,44	2,28	0,63 a 0,37
-0,20	0,78	0,65	0,02491	97,82	22,22	1,10	0,22	0,57 a 0,43
-0,01	1,60	0,51	0,07254	96,61	17,70	0,88	0,02	0,51 a 0,50

FONTE: Regressões $X_A^{(\lambda)} = \hat{a} + \hat{b} X_T^{(\lambda)}$, utilizando-se dados da Tabela 9 (no Apêndice).

^a Variável explicativa significativa ao nível de 1%.

^b λ ótimo.

^c Valor máximo da verossimilhança.

^d Multiplicado por 10^{-10} .

TABELA 4

Elasticidades-renda para os dispêndios em alimentação

Classes de Renda	λ		
	--0,80	--0,85	--0,95
1	0,73	0,74	0,77
2	0,70	0,71	0,73
3	0,61	0,61	0,62
4	0,52	0,52	0,51
5	0,49	0,46	0,48
6	0,46	0,46	0,44
7	0,39	0,39	0,37
8	0,39	0,38	0,37
9	0,41	0,40	0,39
10	0,41	0,40	0,38
11	0,36	0,35	0,33
12	0,30	0,29	0,37
13	0,25	0,21	0,21

FONTES: Utilização da expressão (8), estimativas da Tabela 3 e dados da Tabela 9 (no Apêndice).

pode dar, no que se relaciona à utilização de formas funcionais não-tradicionais (e, talvez, mais representativas) e à não imposição de elasticidade constante para os diversos níveis de renda. De resto, lembramos que o R^2 deve ser encarado com ressalvas (inclusive devido à alta agregação dos dados) e podemos ver que a função ótima não corresponde ao maior valor do coeficiente de determinação. Em relação ao teste de Student, todos são aceitos ao nível de confiança de 1%, o mesmo ocorrendo para o teste de Durbin-Watson (pelo menos para os valores de λ próximos a $\hat{\lambda}$).

Como a alimentação é enquadrada dentro das características de bens de primeira necessidade, podemos esperar que sua elasticidade-renda se situe entre um e zero e decresça à medida que passamos para classes de renda mais altas. Nossos resultados atendem a essas previsões e, conforme podemos ver na Tabela 4, encontramos valores

entre 0,74 e 0,24 (partindo da classe de renda mais baixa para a mais alta). Kirsten²⁸ obteve 0,68, valor constante para todas as classes, pois utilizou a especificação duplo-logaritmo (dados para a cidade de São Paulo, período maio/junho de 1971). Fendt²⁹ obteve 1,88 a 0,32 (e 1,49 a 0,28 usando outro sistema) com especificações mono-logaritmo, que permitem variabilidade da elasticidade-renda, mas prejudicam a interpretação dos valores para as classes de renda iniciais, em virtude de apresentarem nessas classes elasticidades superiores a um (Fendt utilizou dados do antigo Estado da Guanabara, período 1967/68).

Além das duas variáveis utilizadas, isto é, o dispêndio *per capita* na alimentação em função do dispêndio total, podem ocorrer economias de escala, que seriam captadas através da introdução de uma nova variável explicativa: o número médio de habitantes por família. Famílias maiores estariam habilitadas a economizar na compra, estocagem e preparação de alimentação.³⁰ Contudo, uma simples visualização dos dados sugere que essa variável, pelo menos no nível de agregação em 13 classes de renda, não é significativa. *Mutatis mutandis*, aplicaremos o mesmo raciocínio para os dispêndios em educação, mas julgamos conveniente utilizar o número de membros por família.

Para a alimentação, a Tabela 5 confirma a não significância da nova variável introduzida. Apresentamos, na coluna correspondente a "Habitantes por Família", o valor de *t* e o coeficiente de confiança. Mesmo a altos níveis de erro, essa variável não é significativa. Contudo, a variável dispêndio total é significativa. Neste caso obtivemos $\hat{\lambda} = -1,00$, e o intervalo para λ , com coeficiente de confiança de 95%, é definido pelos limites $-1,30$ e $-0,70$. A Tabela 5 mostra o comportamento da transformação Box-Cox para este caso e serve principalmente como ilustração adicional. Como a nova variável explicativa não foi aceita (ao nível de significância de 5%), devemos basear as conclusões na Tabela 4, anteriormente discutida.

²⁸ Kirsten, *op. cit.*

²⁹ Fendt Jr., *op. cit.*

³⁰ Ver Zarembka, "An Econometric...". *op. cit.*

TABELA 5
 Resultados para os dispendios em alimentação (II) : $X_A = f(X_T, H)$

λ	Intercepto	Dispendio Total Per Capita		Habitantes por Família		Erro-Padrão da Estimativa	R ² Ajustado	d	Verosimilhança ^e	Elasticidade-Renda
		Coefficiente	t ^a	Coefficiente	t ^b					
-1,20	-0,86	2,04	17,17	-0,008	-2,06	0,00015	99,48	1,82	61,49	0,92 a 0,18
-1,10	-0,76	1,84	17,61	-0,009	-1,87	0,00023	99,47	1,95	101,94	0,88 a 0,20
-1,05	-0,70	1,74	17,67	-0,010	-1,73	0,00029	99,45	2,00	103,77	0,87 a 0,21
-1,00 ^c	-0,65	1,65	17,63	-0,010	-1,58	0,00036	99,43	2,03	129,33 ^d	0,85 a 0,22
-0,95	-0,59	1,56	17,43	-0,011	-1,40	0,00046	99,40	2,04	110,71	0,83 a 0,23
-0,85	-0,46	1,39	16,93	-0,010	-1,01	0,00076	99,32	2,02	69,55	0,79 a 0,25
-0,80	-0,39	1,31	16,55	-0,010	-0,80	0,00098	99,27	2,00	52,89	0,77 a 0,26
-0,60	-0,08	1,02	14,70	0,001	0,01	0,00282	98,96	1,85	10,51	0,68 a 0,31
-0,40	0,20	0,78	12,77	0,036	0,73	0,00824	98,47	1,71	1,71	0,60 a 0,35
-0,20	-0,82	0,59	12,09	0,130	1,31	0,02415	97,77	1,59	0,27	0,51 a 0,40
-0,01	1,48	0,45	9,75	0,336	1,73	0,06672	96,87	1,52	0,05	0,44 a 0,44

FONTE: Regressões $X_A(\lambda) = \hat{a} + \hat{b} X_T(\lambda) + \hat{c} H(\lambda)$, utilizando dados da Tabela 9 (no Apêndice).

^a Variável explicativa significativa ao nível de 1%.

^b Variável explicativa não significativa ao nível de 5%.

^c λ ótimo.

^d Valor máximo da verosimilhança.

^e Multiplicado por 10 - 10.

5 — Resultados para a educação

A mesma sistemática utilizada para a alimentação deve ser transposta para a educação. Todos os comentários feitos na Seção 4 para os valores de t , d , R^2 e verossimilhança são válidos para este caso.

A diferença fundamental reside nos valores que devem ser atribuídos a λ , em decorrência do comportamento da elasticidade-renda no caso da educação. Optamos por analisar a alimentação e a educação exatamente para realçar as diferenças. Essa divergência não ocorreria se analisássemos, por exemplo, os dispêndios em vestuário e habitação. Nestes casos podemos esperar que as Curvas de Engel resultantes (e, portanto, as elasticidades-renda) tenham comportamentos semelhantes.

Enquanto no caso da alimentação esperávamos que λ fosse negativo, conforme mostra a expressão (9), no caso da educação as previsões são mais difíceis. O ponto de partida para a análise da educação é verificar o comportamento esperado para sua elasticidade-renda. Se a educação for encarada como bem superior, possui elasticidade-renda superior a 1.

Além de a elasticidade-renda para a educação possivelmente apresentar valores maiores que 1, é necessário investigar se, à medida que passamos para as classes de rendas mais altas, seu valor cresce ou decresce. Uma possível interpretação é considerar a educação como uma alternativa que possibilita às classes de renda mais baixas auferirem rendas maiores. Desse modo, a elasticidade seria decrescente. Estamos implicitamente supondo que o dispêndio em educação, registrado na pesquisa de orçamentos familiares, mede principalmente a parcela de educação encarada como “investimento”. A educação “consumo” para as classes de renda mais baixas é um bem de luxo e, neste contexto, a elasticidade-renda (além de ser maior que 1) cresceria à medida que a renda aumentasse. Na Tabela 6 obtivemos o valor ótimo de λ ($\hat{\lambda} = 0,05$), que justificaria a tentativa de interpretação dada anteriormente. O intervalo para λ (usando o procedimento apresentado na Seção 2), com coeficiente de confiança de 95%, é definido pelos limites $-0,20$, e $+0,40$. A tabela fornece, ainda, os resultados para a função de verossimilhança (onde

TABELA 6
 Resultados para os dispendios em educação (I): $Y_E = f(Y_T)$

λ	Intercepto	Coefficiente	Erro-Padrão da Estimativa	R^2	t^a	d	Verossimilhança ^a	Elasticidade-Renda
-0,30	-11,48	4,66	0,11	96,13	16,52	1,91	6,31	1,34 a 1,98
-0,10	-8,40	2,26	0,22	96,03	16,30	1,64	71,24	1,51 a 1,70
-0,05	-7,93	1,90	0,26	95,89	16,02	1,60	96,35	1,55 a 1,64
-0,03	-7,77	1,77	0,29	95,83	15,89	1,59	105,29	1,57 a 1,62
0,03	-7,37	1,44	0,36	95,50	15,44	1,55	124,30	1,62 a 1,56
0,05 ^b	-7,26	1,34	0,40	95,50	15,28	1,55	127,31 ^c	1,64 a 1,54
0,15	-6,91	0,95	0,61	94,96	14,39	1,52	117,27	1,74 a 1,46
0,20	-6,84	0,81	0,77	94,64	13,93	1,52	101,15	1,80 a 1,43
0,25	-6,83	0,68	0,97	94,29	13,48	1,53	82,39	1,87 a 1,40
0,50	-7,83	0,31	3,23	92,33	11,55	1,57	16,98	2,28 a 1,27

FONTE: Regressões $Y_E^{(\lambda)} = \hat{a} + \hat{b} Y_T^{(\lambda)}$, utilizando-se dados da Tabela 9 (no Apêndice).

^a Variável explicativa significativa ao nível de 1%.

^b λ ótimo.

^c Valor máximo da verossimilhança.

^d Multiplicado por 10^{-10} .

notamos a convergência para o valor máximo) e os demais resultados da regressão estimada.

A exemplo do que foi feito para a alimentação, a Tabela 7 mostra a elasticidade-renda da educação para o λ ótimo e para os dois imediatamente próximos. E, de acordo com as afirmações anteriormente feitas, a tendência da elasticidade é decrescente. Em trabalhos futuros pretendemos voltar a esse ponto, inclusive verificando a influência da educação gratuita e financiada. Como os dispêndios em educação, computados na Tabela 9 (no Apêndice), medem os gastos efetivamente realizados, adiamos por enquanto essa discussão.

Os resultados obtidos para a elasticidade (ver Tabela 7) situam-se entre 1,64 e 1,54 (partindo da classe de renda mais baixa para a mais alta). Citamos, a seguir, os trabalhos realizados por Kirsten³¹ e Fendt,³² em pesquisas já citadas, quando tratamos da alimentação.

TABELA 7
Elasticidades-renda para os dispêndios em educação

Classes de Renda	λ		
	0,03	0,05	0,14
1	1,62	1,64	1,74
2	1,64	1,68	1,86
3	1,61	1,62	1,67
4	1,60	1,60	1,62
5	1,58	1,57	1,52
6	1,57	1,56	1,49
7	1,56	1,53	1,42
8	1,55	1,53	1,41
9	1,57	1,56	1,51
10	1,54	1,51	1,37
11	1,55	1,52	1,39
12	1,56	1,54	1,45
13	1,56	1,55	1,46

FONTES: Utilização da expressão (8), estimativas da Tabela 6 e dos dados da Tabela 9 (no Apêndice).

³¹ Kirsten, *op. cit.*

³² Fendt Jr., *op. cit.*

Para a educação, o primeiro obteve 1,83 e o segundo 1,60 (e 1,78 usando outro sistema). Ambos os autores utilizaram a especificação duplo-logaritmo, e todas as classes de renda exibem conseqüentemente a mesma elasticidade. Esta restrição é superada pela técnica de transformação de variáveis.

Nas estimações feitas para a educação utilizamos os "Dispêndios por Família", ao invés de "Dispêndios *Per Capita*", usados na alimentação. Quando se trata do item educação julgamos que os valores utilizados são mais representativos.

A Tabela 8 mostra a introdução de uma nova variável explicativa para os dispêndios em educação. Além do dispêndio total, consideramos o número médio de membros por família. Neste caso não estamos interessados em economias de escala, como no caso da alimentação, mas sim em medir um possível aumento nos dispêndios, devido a um maior número de membros por família. Também aqui, essa nova variável não é relevante, ao nível de significância de 5%. Contudo, a variável dispêndio total por família é significativa. Neste caso, $\hat{\lambda} = 0,03$, e o intervalo para λ , com coeficiente de confiança de 95%, é definido pelos limites $-0,20$ e $+0,40$. Fornecemos esta tabela para mostrar mais um caso do comportamento da técnica de transformação de variáveis. Como a variável explicativa adicional não foi aceita, a análise da educação toma como base as Tabelas 6 e 7.

6 — Conclusões

Este trabalho optou por uma forma especificativa mais geral, onde casos particulares podem não satisfazer certas restrições teóricas, mas, por outro lado, possibilitam bom ajustamento. Verificamos, por exemplo, que o crescimento da renda ocasiona a substituição de bens de primeira necessidade por bens superiores e, então, podemos empregar curvas diferentes, que mostram os padrões do comportamento da elasticidade dos diversos bens.

Ao estimar uma forma especificativa mais geral, a partir de testes de hipóteses conseguimos escolher aquela função que melhor reflete

TABELA 8
 Resultados para os dispendios em educação (II): $Y_B = f(Y_T, M)$

λ	Intercepto	Dispendio Total por Família		Membros por Família		Erro-Padrão da Estimativa	R^2 Ajustado	d	Verossimilhança	Elasticidade-Renda
		Coefficiente	t^a	Coefficiente	t^b					
-0,30	-10,64	4,12	8,88	0,69	1,41	0,11	96,12	1,87	11,03	1,24 a 1,75
-0,10	- 8,40	2,06	9,64	0,85	1,20	0,21	95,83	1,57	91,57	1,38 a 1,55
-0,05	- 8,11	1,74	9,68	0,90	1,14	0,26	95,63	1,53	114,00	1,42 a 1,51
-0,03	- 8,02	1,63	9,68	0,93	1,11	0,28	95,54	1,51	120,63	1,44 a 1,49
0,03 ^c	- 7,83	1,33	9,65	1,01	1,04	0,36	95,22	1,48	130,06 ^d	1,50 a 1,45
0,05	- 7,80	1,24	9,62	1,04	1,02	0,40	95,10	1,47	128,54	1,52 a 1,44
0,15	- 7,86	0,89	9,44	1,24	0,91	0,62	94,41	1,46	105,97	1,63 a 1,37
0,20	- 8,04	0,76	9,32	1,38	0,87	0,78	94,02	1,46	87,30	1,69 a 1,34
0,25	- 8,33	0,64	9,19	1,54	0,83	0,99	93,60	1,46	68,51	1,76 a 1,32
0,50	-12,31	0,29	8,59	3,14	0,75	3,30	91,34	1,52	12,58	2,16 a 1,21

FONTE: Regressões $Y_E^{\lambda} = \hat{a} + \hat{\delta} Y_T^{\lambda} + \epsilon M^{\lambda}$, utilizando dados da Tabela 9 (no Apêndice).

^a Variável explicativa significativa ao nível de 1%.

^b Variável explicativa não significativa ao nível de 5%.

^c λ ótimo.

^d Valor máximo da verossimilhança.

^e Multiplicado por 10^{-10} .

o comportamento dos dispêndios na população. Foi possível atender às evidências empíricas, citadas na Seção 1, e evitar alguns resultados restritivos, como por exemplo aqueles apresentados pela função duplo-logaritmo, que impõe a mesma elasticidade para todos os estratos de renda. Evitamos, ainda, a obtenção de valores de elasticidade que “contrariam o bom senso”, para o caso das classes de renda menores, o que ocorre quando são utilizadas especificações mono-logaritmo. Finalmente, não nos restringimos às funções tradicionais, pois o parâmetro que discrimina entre as várias formas especificativas pode assumir qualquer valor. Em outras aplicações de transformação de variáveis que pesquisamos constatamos, inclusive, que a imposição da especificação duplo-logaritmo ou linear pode levar à rejeição *a priori* de teorias alternativas, tanto no que se refere à especificação, como também com relação à inclusão ou não de variáveis explicativas adicionais.³³

A não satisfação do critério da adição, pela forma especificativa escolhida pelo teste de hipóteses, já ocorre quando utilizamos as abordagens usuais (ver Seção 2). Então acreditamos que pelo menos do ponto de vista estatístico estejamos contribuindo para melhorar os resultados que vêm sendo obtidos. Por outro lado, existe validade em explorar os aspectos econométricos, porque sabemos que a teoria acerca do critério da adição ainda não chegou a conclusões definitivas, pelo menos do ponto de vista operacional.³⁴

Apêndice -- dados utilizados

Os dados brutos mostrados na Tabela 9 a seguir foram obtidos na Pesquisa de Orçamentos Familiares (POF) da cidade de São Paulo, realizada pelo Instituto de Pesquisas Econômicas no período 1971/72. Foram obtidos em quatro etapas, correspondentes a quatro trimestres, numa amostragem de cerca de 800 famílias em cada vez. A unidade base de levantamento fixada foi o mês e os dados que

³³ Ver Zarembka “Functional Forms...”, *op. cit.*

³⁴ Esta posição está explícita em Houthakker, “The Present State...”, *op. cit.*, p. 722, Leser, *op. cit.*, p. 694, e Philips, *op. cit.*, p. 113.

TABELA 9

Dados brutos (unidades: Cr\$ de 1971/72, média de quatro amostras)

Classes de Renda (Cr\$)	A	B	T	I	H	M	F
	Dispêndio em Alimentação (Média Mensal)	Dispêndio em Educação (Média Mensal)	Dispêndio Total (Média Mensal)	Número de Indivíduos na Classe	Número Médio de Habitantes por Família	Número Médio de Membros por Família	Número de Questionários (Número de Famílias na Classe)
1 até 225,60..	4.401,61	154,20	8.574,85	99	2,91	2,61	34
2 225,61 a 451,20..	32.754,19	778,75	67.096,57	572	3,32	2,79	172
3 451,21 a 1.353,60..	365.020,93	21.247,56	895.013,52	4.731	4,23	3,77	1.117
4 1.353,61 a 1.804,80..	132.298,10	11.590,07	385.217,55	1.397	4,62	3,89	302
5 1.804,81 a 2.256,00..	93.533,68	13.196,25	297.664,62	938	4,96	4,17	189
6 2.256,01 a 2.707,20..	71.993,11	12.559,36	249.738,80	613	4,75	3,94	129
7 2.707,21 a 3.158,40..	55.848,08	16.298,50	235.805,43	486	4,71	4,10	103
8 3.158,41 a 3.609,60..	31.106,77	9.597,27	132.105,72	229	4,32	3,41	53
9 3.609,61 a 4.060,80..	38.333,52	7.117,90	153.754,60	262	4,67	3,73	56
10 4.060,81 a 4.512,00..	27.367,39	10.255,55	111.756,08	206	5,02	4,29	41
11 4.512,01 a 4.963,20..	17.918,76	6.814,85	84.687,29	144	5,33	4,29	27
12 4.963,21 a 7.444,80..	64.706,16	23.083,55	377.952,35	426	4,43	3,61	96
13 acima de 7.444,81..	53.679,42	23.000,31	399.411,97	315	5,16	3,91	61

FONTE: J. T. Kirsten, A. Berndt, C. N. P. Moraes, E. L. G. Alves, J. A. Carletti, M. N. S. Rubião e S. W. Song, "Orçamentos Familiares na Cidade de São Paulo" (São Paulo: IPE/USP, 1973), pp. 160 a 245. Período: agosto de 1971 a julho de 1972.

apresentamos representam médias mensais do período agosto de 1971 a julho de 1972. Optamos por utilizar os dados agrupados em 13 classes de renda a fim de minimizar o custo de computação e pelo fato de ser possível, com esse nível de agregação, chegar aos principais objetivos do trabalho. No entanto, cumpre realçar que nossos resultados numéricos seriam sensivelmente melhorados se pudéssemos subdividir as classes de renda de forma que o número de indivíduos, em cada uma, apresentasse variabilidade maior.

Na primeira linha da tabela mencionada relacionamos a notação das variáveis utilizadas ao longo deste trabalho. A partir desses elementos, fizemos algumas modificações, a fim de trabalhar em termos *per capita* ou em valores por família. Descreveremos, a seguir, os valores efetivamente utilizados nas várias regressões, sendo que a amostra consiste em 13 classes de renda:

- I — número de indivíduos;
- F — número de famílias;
- H — número médio de habitantes por família;
- M — número médio de membros por família;
- $X_A = A/I$ — dispêndio em alimentação *per capita* (média mensal);
- $X_T = T/I$ — dispêndio total *per capita* (média mensal);
- $Y_E = E/F$ — dispêndio em educação por família (média mensal); e
- $Y_B = T/F$ — dispêndio total por família (média mensal).

