Comunicação II

O emprego de deflatores inadequados e o problema de erro comum nas variáveis em estudos econométricos – um comentário

AFFONSO CELSO PASTORE *

Em estimulante trabalho publicado no último número desta Revista, Kenneth King [1] questiona a validade dos resultados de alguns modelos econométricos envolvendo relações nas quais as variáveis independentes e a dependente são "deflacionadas" por um mesmo índice de preços.

A presente nota tem o duplo propósito de apresentar algumas evidências adicionais sobre o comportamento da demanda de moeda no Brasil, que fundamentam a ausência da correlação espúria arguída por King, e de avaliar criticamente os resultados de suas simulações, mostrando que são inadequadas ao tratamento do problema da demanda de moeda no Brasil.

Sua análise é desenvolvida visando especificamente as discussões em torno de uma versão mais simples do modelo de demanda de moeda, no qual o estoque real de meios de pagamento em *t* aparece como uma função da renda real em *t*, e da taxa de inflação verificada no mesmo período. Argumenta que os erros no índice utilizado para deflacionar a renda e os meios de pagamento, e a partir do qual se constrói a taxa de inflação em *t*, provocam vieses nos coeficientes da relação estimada, na mesma direção especificada pela Teoria Econômica. O argumento é ilustrado por algumas simulações, que partem de séries hipotéticas estritamente aleatórias

* Do Instituto de Pesquisas Econômicas da Universidade de São Paulo.

1 K. King, "O Emprego de Deflatores Inadequados e o Problema de Erro Comum nas Variáveis em Estudos Econométricos"; *Pesquisa e Planejamento,* (Vol. I, n.º 2, dez. 1971).

Pesq. Plan. Econ. Rio de Janeiro, 2(1): 117-130 jun. 1972

(que seriam as "verdadeiras" séries de renda e meios de pagamento reais), e por isso mesmo guardam correlações nulas *nas populações dos valores possíveis dessas variáveis.* Ao serem as séries pseudo-reais inflacionadas pelo "verdadeiro índice de preços" e deflacionadas pelo "deflator errado", geravam correlações estreitas e resultados estatisticamente significativos, mas provocados apenas pelos erros do deflator.

Colocar à prova as hipóteses sobre as quais os economistas constroem as suas teorias é um dos principais objetivos da análise econométrica. Nesse contexto, a econometria deve ser entendida como uma técnica de "rejeitar" hipóteses, não podendo ser utilizada para "prová-las", pelo simples fato de que sempre que encontramos uma hipótese consistente com um conjunto de fatos, podemos formular uma ou algumas hipóteses alternativas, também consistentes com o mesmo conjunto de fatos.

Os resultados econométricos apresentados por Cagan [2] em sua análise de sete hiperinflações, e pelo autor [3] no caso brasileiro, são compatíveis com a hipótese de que as flutuações nas expectativas de inflação constituem um elemento dominante na explicação das flutuações da demanda de moeda. King tenta demonstrar que essa situação pode também derivar de um caso extremo em que as correlações entre os "verdadeiros valores" da caixa real, de um lado, e da renda e da taxa de inflação esperada, de outro, seriam nulas, ou quase nulas, e que os resultados derivariam, fundamentalmente, da correlação espúria proveniente dos erros no deflator.

A existência de erros nos deflatores obviamente conduz a correlações espúrias, mas por esse motivo não está ela presente em meu trabalho (nem no de Cagan) porque: a) não utilizei o mesmo índice para deflacionar o estoque de meios de pagamento e a renda (nem Cagan, pois em sua demanda de moeda não comparece a renda como variável explicativa); b) não utilizei a taxa de inflação em t como

2 P. Cagan, "The Monetary Dynamics of Hiperinflation", em *Studies in the Quantity Theory of Money* (Editado por Milton Frieman, Chicago University Press, 1956).

3 A. C. Pastore, "Inflação e Política Monetária no Brasil", *Revista Brasileira de Economia,* (janeiro-março 1969).

uma aproximação para a taxa de inflação esperada (tampouco Cagan).

A expectativa de inflação é aproximada por uma média móvel de pesos geometricamente declinantes das taxas passadas de inflação, sendo que a primeira taxa de inflação a comparecer no cômputo da expectativa é a taxa efetivamente verificada em *t-1*; inexistindo, conseqüentemente, o comparecimento simultâneo do nível geral de preços em *t* como deflator dos meios de pagamento em *t* e no numerador da taxa de inflação em *t*.[4]

Submeto-me a uma outra objeção, de que interpolei a renda real mensal seguindo um critério já exposto anteriormente[5], e reconheço que esse procedimento introduz um *viés para baixo* no coeficiente da renda, *mas esse viés é no sentido contrário ao especificado pela Teoria Econômica, devendo piorar os resultados, e não melhorá-los.*

Acredito, conseqüentemente, que a qualidade dos resultados apresentados (por mim e por Cagan) não derivem fundamentalmente do fenômeno da correlação espúria, e que sejam, ao contrário, uma indicação de que não temos razões para rejeitar a hipótese de que a demanda de moeda possa ser adequadamente representada na forma como foi discutida naqueles trabalhos.

Uma apreciação dos resultados das simulações

Toda a análise de King consiste em estimar coeficientes de "pseudo-modelos" de demanda de moeda, em que denomina por R_i cada uma das séries reais aleatórias, e nos quais comparecem sempre dois índices de preços: p^e, o índice errado, e p^v, o índice correto. São sempre apresentadas duas especificações. A primeira coloca $y = R_1 \cdot \frac{p^v}{p^e}$

4 Ver A. C. Pastore, *op. cit.*, p. 105.

5 Ver A. C. Pastore, *op. cit.*, pp. 101 a 103, e em particular o Apêndice. O processo de interpolação consiste basicamente em distribuir a renda anual entre os meses, proporcionalmente a um índice que contém os ciclos mensais da arrecadação do IVC (imposto sobre vendas e consignações), e a variação estacional da arrecadação daquele imposto. O processo de interpolação impõe que a soma das rendas mensais seja igual à renda do ano e, conseqüentemente, não existem erros entre anos, somente entre meses. Se a arrecadação do IVC for sensível às variações cíclicas e estacionais da renda, então estará garantida uma elevada correlação entre as flutuações mensais da renda *proxy* e da renda verdadeira.

em função de $x = R_2 \frac{p^v}{p^e}$, e a segunda inclui na relação acima a variável $z = p_t^e/p_{t-1}^e$. Claramente, a primeira variável será a "pseudo-caixa real", a segunda a "pseudo-renda real" e a terceira a "pseudo-taxa de inflação". É claro que se o mesmo deflator errado for utilizado para a obtenção de y e x, o coeficiente de regressão por mínimos quadrados sofrerá um viés na direção da unidade.

A expressão para o viés assintótico é dada por: [6]

$$\varepsilon(b) = \frac{\beta\sigma_x^2}{\sigma_x^2 + \sigma_v^2} + \frac{\sigma_v^2}{\sigma_x^2 + \sigma_v^2} \tag{1}$$

onde ε (b) indica a esperança assintótica do estimador de mínimos quadrados, σ_x^2 é a variância da "verdadeira componente" de X (isto é, a variância da verdadeira variável real) e σ_v^2 é a variância derivada do erro de deflacionamento. É claro que ε (b) é uma média do verdadeiro coeficiente β, ponderado pela participação da variância da "verdadeira" variável no total da variância de x (a componente verdadeira mais o erro), e da unidade, ponderada esta pela participação da variância do erro de deflacionamento no total da variância de x.

6 Supomos que a relação entre os "verdadeiros valores" das variáveis reais é linear, ou seja,

$$Y = \beta X + u$$

e que os valores reais observados sejam, respectivamente,

$$y = Y + v$$
$$x = X + v$$

o estimador de mínimos quadrados de β, utilizando y e x, será:

$$b = \frac{M_{YX} + M_{Yv} + M_{Xv} + M_{vv}}{M_{XX} + M_{vv} + 2M_{Xv}}$$

onde M_{ij} está designando o momento direto ou cruzado das respectivas variáveis. Pelo fato de a variável aleatória comparecer no numerador e denominador da expressão acima não podemos calcular a média de sua distribuição por amostragem para pequenas amostras. Podemos, entretanto, calcular a esperança assintótica de b, que é dada pelo limite da probabilidade da distribuição de b, cuja expressão é dada, na hipótese de que os erros de deflacionamento sejam independentes das verdadeiras variáveis reais, pela expressão (1) no texto.

É fácil verificar que se fizermos $\beta = 0$, impondo uma correlação nula entre os "verdadeiros" valores das variáveis reais, ainda assim poderemos polarizar a distribuição por amostragem do estimador de β em torno de valores plausíveis, desde que tomemos $\sigma_v^2/\sigma_{\hat{X}} + \sigma_v^2$ igual a esse valor plausível.

Sabemos, por outro lado, que os estimadores de Mínimos Quadrados gozam da propriedade de ter a variância reduzida à medida que aumentamos o tamanho da amostra. A regra para que as simulações mostrem maior número de resultados significativos é, portanto, aumentar o número de observações. Conseqüentemente, é possível produzir quaisquer resultados mediante uma escolha adequada da variância do erro e do número de observações.

O que é dramático nesse tipo de exercício, contudo, é que nunca poderemos saber exatamente se esses resultados guardam alguma semelhança com a "verdadeira" demanda, e se têm condições de insinuar que a magnitude do viés nas estimativas de Mínimos Quadrados da demanda é elevado. Isso porque eles partem de hipóteses arbitrárias sobre os erros, escolhidas de forma a tornar plausíveis os resultados, e não indicam necessariamente o viés que está efetivamente presente nos modelos estimados com os dados existentes.

Basta fixar uma variância do erro que atinja entre 70% e 90% da variância total da variável independente que, ainda que a correlação entre as séries hipotéticas seja nula, em experimentos repetidos a média da distribuição do estimador de Mínimos Quadrados tenderá a se fixar em torno de 0,7 a 0,9. Em experimentos individuais podemos não obter exatamente esses valores, existindo igual probabilidade de obtermos valores maiores e menores do que aqueles em torno dos quais a distribuição está polarizada. É por isso que simulações dessa natureza devem ser conduzidas dentro dos padrões de experimentos Monte-Carlo, em que são realizadas experiências repetidas, e o viés de pequenas amostras somente pode ser aproximado pela diferença entre a média da distribuição por amostragem tabulada nos experimentos repetidos, e o "verdadeiro valor" do coeficiente. A validade de resultados individuais é, em simulações estatísticas, altamente questionável.

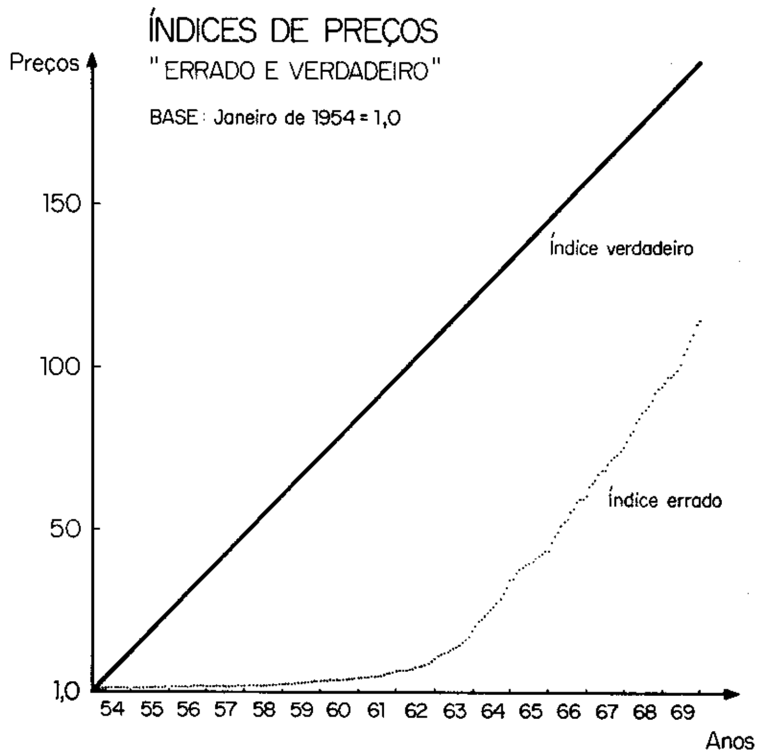
Na especificação I, as séries pseudo-reais são inflacionadas por um índice estritamente linear, que assume o valor 1 em janeiro de 1954, e atinge o valor 228 em dezembro de 1970. São posteriormente deflacionados pelo "índice errado", que é o índice 2 da Conjuntura Econômica, que assume o valor 1,3 no início do período chegando a 208 ao final. Aparentemente, os desvios entre os dois índices são pequenos, chegando-se a uma diferença de apenas 10% ao final do período que, se espalhada em todos os 228 meses, efetivamente provocaria pequenos desvios entre os valores verdadeiros e errados dos índices mensais. Essa observação, aliada à significação da "pseudo-renda", indicaria que as estimativas são extremamente sensíveis a esse erro.

No gráfico a seguir reproduzem-se os valores dos índices errado e verdadeiro, evidenciando-se que os erros são enormes, e praticamente impossíveis de acontecer na realidade. O índice "verdadeiro" é 10 vezes maior que o errado, no início do período, 30 vezes maior em 1957/58, e apenas uma vez e meia maior em 1970. As taxas de variação do índice verdadeiro começam em 100% ao ano, no início do período, declinando continuamente até 6% ao ano, no final, enquanto que as taxas do índice errado situam-se em torno de 10 a 20% no início, passam por um máximo de 90% (quando as taxas do índice "verdadeiro" estão em 10%), e declinam, estabilizando-se em 20%, quando a taxa do índice errado cai em torno de 6 a 7%.

Essa verdadeira "barriga", gerada pelos desvios entre os índices "verdadeiro" e "errado", estará presente nas duas séries reais erradas, e ainda que a correlação entre as originais fosse nula, as duas novas séries forçosamente estariam correlacionadas. Mas será que "barriga" semelhante existiria quando da comparação do índice teórico de preços, sem erros, para o Brasil, com o índice 2 da Conjuntura?

Na especificação II, King supõe o erro da *proxy* como uma proporção ε_t do verdadeiro índice de preços, sendo ε_t um número aleatório. Os resultados deste experimento são bem piores do que os anteriores, mas nem por isso suas hipóteses são mais plausíveis. Exprimindo o índice "errado" como uma proporção ε_t do índice verdadeiro,

$$P_t^* = (1 + \varepsilon_t) P_t$$



podemos extrair uma relação simples entre a "taxa de inflação errada" e a "taxa de inflação verdadeira", dada por

$$\frac{P_t^*}{P_{t-1}^*} = \frac{1 + \varepsilon_t}{1 + \varepsilon_{t-1}} \cdot \frac{P_t}{P_{t-1}}$$

As simulações de King mostram que resultados significativos começam a aparecer somente quando o erro está contido entre -30% e $+30\%$ do índice verdadeiro. Não se conhece qual seja a forma da distribuição de erros[7], mas, diante dos extremos em que os erros

[7] Se os erros foram extraídos ao acaso de uma distribuição normal de média zero, então a probabilidade de erros extremos, em torno de $\pm 30\%$, será relativamente pequena, mas serão freqüentes os erros entre $\pm 5\%$. Se eles foram extraídos ao acaso de uma distribuição retangular, então a probabilidade de erros extremos é substancialmente mais elevada.

podem flutuar, devem ser relativamente prováveis os que passam de -5% para $+5\%$. Se a "verdadeira taxa de inflação" fosse de 2%, a "taxa errada" seria, nessa hipótese de erros, de

$$1{,}02 \times \frac{1{,}05}{0{,}95} = 1{,}13$$

ou seja, estimar-se-ia uma taxa de 13%, quando a verdadeira seria de 2%, o que parece um resultado altamente implausível.

A última simulação apresenta uma especificação ainda diversa do erro, supondo que este é proporcional à taxa de inflação efetivamente verificada, flutuando os limites de erro entre -50% e $+ 50\%$ da taxa efetivamente verificada. Os resultados são melhores do que os da especificação II (apenas no que diz respeito ao coeficiente da "pseudo-renda", e não quanto ao coeficiente da "pseudo-inflação", que é, apenas, mediocremente significativo), mas não comparáveis aos anteriores, porque foram tomadas novas séries aleatórias. Como estamos diante de um resultado isolado, desconhecemos se esse coeficiente de correlação relativamente elevado entre a "pseudo-renda" e a "pseudo-caixa real" deriva de uma correlação alta entre as séries pseudo-reais corretas, ou de uma flutuação por amostragem. A menos que um experimento Monte-Carlo confirme os resultados, as evidências apresentadas são de pouca validade. Fica-se, por outro lado, em posição extremamente desconfortável, pois apesar da afirmativa de que os erros especificados nas simulações são plausíveis, não existe qualquer critério objetivo para julgar-lhe a plausibilidade. Erros de medida no índice 2 da Conjuntura, não podem ser observados, e conseqüentemente não se pode saber a sua magnitude.

Exercícios de simulação podem gerar dúvidas em entusiastas de análise empírica somente quando ajudam a compreender os erros que podem ocorrer em análises concretas. Do contrário, são totalmente destituídos de interesse. No caso presente, a impossibilidade de saber se os erros de deflacionamento arguídos são verdadeiros ou não, impede sua aplicabilidade prática, e a existência de correlação espúria somente pode ser investigada analisando diretamente a demanda de moeda no Brasil.

A demanda de moeda no Brasil

O modelo de demanda de moeda anteriormente estimado para o Brasil pode ser especificado na forma

$$(2) \qquad (M/P)_t = \alpha_0 + \alpha_1 \, y_t + \alpha_2 \, \Pi^e_t + \alpha_3 \, (M/P)_{t-1} + u_t$$

onde y_t é a renda real corrente, Π^e_t a taxa de inflação esperada, construída de acordo com a hipótese de expectativas adaptadas, proposta por Cagan, e na qual admite-se que a taxa esperada em *t* seja igual à esperada em *t-1*, mais uma correção que é uma proporção γ da diferença entre a taxa esperada e a ocorrida naquele período. Essa suposição implica dizer que a taxa esperada pode ser expressa como uma média móvel de pesos geometricamente declinantes das taxas passadas, ou seja,

$$(3) \qquad \Pi^e_t = \sum_{j=0}^{\infty} \gamma(1 - \gamma)^j \, \Pi_{t-j-1}$$

O modelo utilizado por Cagan é um caso particular de (2), no qual se supõe $\alpha_1 = \alpha_3 = 0$, sendo a ausência da renda explicada pelo fato de que as variações de renda, no período das hiper-inflações analisadas, eram muito pequenas, e admitindo que o "ajustamento Nerloviano" da demanda de moeda era instantâneo.[8] O mo-

[8] Admite-se que a demanda de moeda de longo prazo, isto é, o volume real de meios de pagamento que a coletividade desejará reter depois de decorrido um período suficientemente longo para que ocorra um completo reajuste de seu *portfólio*, é dada por:

$$(M/P)^*_t = f(y_t, \Pi^e_t)$$

Supõe-se, além disso, que a caixa real de curto prazo em *t* seja igual à de curto prazo em t-1, mais um ajustamento que é uma proporção β do hiato que se deseja cobrir a longo prazo, ou seja:

$$(M/P)_t - (M/P)_{t-1} = \beta[(M/P)^*_t - (M/P)_{t-1}]$$

Reconhece-se, na segunda equação acima, a equação de ajustamento proposta por Nerlove. Ver, a esse respeito, M. Nerlove, *The Dynamics of Supply: Estimation of Farmer's Response to Price* (The Johns Hopkins Univ. Press, 1958.) Admitindo-se a linearidade de f, e substituindo a primeira na segunda equação, somos conduzidos à expressão do texto.

delo utilizado por King é também um caso particular de (2), onde se supõe $\alpha_3 = 0$ e $\gamma = 1$. Esta última restrição conduziria, por uma aplicação imediata de (3), à relação $\Pi^e_t = \Pi_{t-1}$, mas nas simulações a taxa de inflação não foi tomada defasada de um período, e sim no mesmo momento da caixa real, ou seja $\Pi^e_t = \Pi_t$. Para analisar em maior profundidade a consistência dos resultados, estimou-se aqui a relação (2) utilizando-se agora dados da renda ligeiramente diversos dos anteriores, devido aos resultados da revisão das Contas Nacionais, realizada pela Fundação Getulio Vargas.

Os resultados obtidos para a equação (1) foram

$$(4) \quad \left(\frac{M}{P}\right)_t = \underset{\substack{(5,887)\\ [0.398]}}{0,173\, y_t} - \underset{\substack{(4,360)\\ [-0,306]}}{1.083\ \Pi^e_t} + \underset{\substack{(26,416)\\ [0,890]}}{0,826} \left(\frac{M}{P}\right)_{t-1} + 0,273$$

$$R^2 = 0,975 \quad DW = 1,859 \quad n = 188$$

onde todas as variáveis estão expressas em logaritmos. Os números entre parênteses logo abaixo dos coeficientes são os valores de T de Student, e os números entre colchetes são os coeficientes de correlação parcial entre a respectiva variável independente e a dependente. O valor do coeficiente de expectativas utilizado na simulação foi $\gamma = 0,1$.[9]

Suponhamos que, efetivamente, o índice geral de preços possua substanciais erros de medida, e que mesmo que não tenhamos deflacionado a renda por esse índice exista, por qualquer efeito fora de nosso controle, uma elevada correlação inversa entre os erros do índice e a renda interpolada. Neste caso, a correlação entre (M/P) e y seria positiva e possivelmente acentuada, podendo, inclusive, a significância da demanda estar substancialmente alterada pela presença de correlação espúria.

9 O método utilizado para estimar γ foi o processo iterativo que consiste em dar valores sucessivos a γ, entre 0 e 1, gerando as respectivas séries de espectativas, e tomando como estimativa de γ, aquele valor que maximizou o coeficiente de determinação da relação estimada. Prova-se que, sob certas hipóteses, essa é uma estimativa de máxima verossimilhança de γ.

Ver, a esse respeito, P. Cagan, *op. cit.*, Apêndice A, pp. 92 a 96.

Essa dificuldade pode ser facilmente evitada colocando-se o volume de meios de pagamento em termos nominais em t como uma função de todas as variáveis presentes no segundo membro de (2), e adicionando como variável explicativa o nível geral de preços.

Se não existir qualquer correlação espúria derivada de erros no deflator, os resultados devem ser próximos dos anteriores, e desde que todas as variáveis estão expressas em logaritmos, o coeficiente do nível geral de preços não deve diferir significativamente da unidade.

Os resultados obtidos, livres agora do viés do deflacionamento, e sem impor previamente a restrição de que a demanda de moeda é homogênea de grau um no nível geral de preços, são:

$$(5)\quad M_t = \underset{\substack{(5{,}684)\\ [0{,}388]}}{0{,}156\, y_t} - \underset{\substack{(4{,}166)\\ [-0{,}296]}}{0{,}970\, \Pi_t^e} + \underset{\substack{(28{,}931)\\ [0{,}906]}}{0{,}841 \left(\frac{M}{P}\right)_{t-1}} + \underset{\substack{(289{,}876)\\ [0{,}999]}}{0{,}995\, P_t} + 0{,}189$$

$$R^2 = 0{,}999 \qquad DW = 1{,}878 \qquad n = 188$$

O coeficiente de P_t não difere significativamente da unidade, e o que é mais importante, a simples comparação entre os coeficientes desvios-padrão e correlações parciais, nas estimativas (4) e (5), mostra que são praticamente os mesmos.

Estamos diante de um ilustrativo teste de consistência, evidenciando de forma bastante clara a ausência de um viés significativo derivado do deflacionamento com índices errados.

Na tabela abaixo, apresentamos um resumo das elasticidades de curto e de longo prazos da demanda de moeda com relação à renda e à taxa de inflação esperada, bem como estimativas dos coeficientes de expectativas (γ) e de ajustamento (β). É fato sabido que as elasticidades de longo prazo são extremamente sensíveis ao viés de especificação, mas, apesar disso, as estimativas apresentadas mostram-se extremamente próximas entre si. Ora, se na primeira das formulações poderia existir viés derivado dos erros no índice de preços, na seguinte essa correlação espúria foi eliminada. Mesmo assim, os resultados não se alteram. Conseqüentemente, esse viés, se existente, deve ser desprezível, e não serve para explicar a qualidade dos resultados.

Elasticidade da demanda de moeda

EQUAÇÃO	Com Relação à Renda		Com Relação à Inflação Esperada		Coeficiente de Ajustamento	Coeficiente de Espectativas
	CP	LP	CP	LP		
(4)	0,17	1,00	— 1,08	— 6,35	0,17	0,10
(5)	0,16	1,00	— 0,97	— 6,06	0,16	0,10

O coeficiente de determinação

O fato perturbador nesse modelo é a presença repetida de elevados coeficientes de determinação. Nos resultados anteriores, constatamos significâncias sempre aproximadamente iguais, se bem que os coeficientes de determinação tenham variado de forma visível. Na opinião do autor esse resultado tem muito pouca importância, pois a qualidade do modelo se infere basicamente da significância de seus coeficientes, e não do R^2 da função.

É fato sabido que modelos de defasagens distribuídas tendem, em geral, a apresentar elevados coeficientes de determinação. Suponhamos que a demanda de moeda pudesse ser expressa na forma

$$(6) \qquad y_t = \alpha x_t + \beta y_{t-1} + u_t$$

com todas as variáveis reduzidas às próprias médias, e onde y_t é o volume real de meios de pagamento em t, e x_t é a taxa de inflação em t-1. Se subtrairmos y_{t-1} de ambos os membros de (1) obtemos:

$$(7) \qquad y_t - y_{t-1} = \alpha x_t + (\beta - 1) y_{t-1} + u_t$$

É fácil verificar que apesar dessa transformação de variáveis, as estimativas de Mínimos Quadrados *a* e *b*, de α e β, serão numericamente iguais nas duas versões acima, pois,

$$\begin{bmatrix} \Sigma x_t^2 & \Sigma x_t y_{t-1} \\ \Sigma x_t y_{t-1} & \Sigma y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \Sigma y_t x_t \\ \Sigma y_t y_{t-1} \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \Sigma x_t^2 & \Sigma x_t y_{t-1} \\ \Sigma x_t y_{t-1} & \Sigma y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \Sigma (y_t - y_{t-1}) x_t \\ \Sigma (y_t - y_{t-1}) y_{t-1} \end{bmatrix}$$

O mesmo ocorre com relação aos desvios-padrão das estimativas, porque: a) a matriz de produtos diretos e cruzados é a mesma nas duas formulações; b) a soma de quadrados residuais também é a mesma, apesar da transformação de variáveis, pois,

$$S^2 = \Sigma (y_t - a x_t - b y_{t-1})^2 = \Sigma [y_t - y_{t-1} - a x_t - (b - 1) y_{t-1}]^2$$

Comprova-se, conseqüentemente, que os coeficientes estimados e o T de Student são exatamente os mesmos, mas somente por acaso é que o coeficiente de determinação será igual, pois a *base* do R^2 é diversa nas duas regressões. De fato, temos:

$$R_2^2 = 1 - \frac{S^2}{\Sigma (y_t - y_{t-1})^2} \cong 1 - \frac{S^2}{2 \Sigma y_t^2 (1 - r_{y_t y_{t-1}})}$$

o que nos permite escrever:

$$(9) \qquad (1 - R_1^2) \cong 2 (1 - r_{y_t y_{t-1}}) (1 - R_2^2)$$

sendo evidente que os coeficientes de determinação nas duas funções somente serão iguais quando o coeficiente de correlação simples entre y_t e y_{t-1} for igual a 0,5. Quanto maior a intercorrelação entre y_t e y_{t-1}, maior o valor de R_1^2 relativamente ao R_2^2, sem que se alterem os coeficientes da regressão, e permanecendo a mesma significância.

Como pela própria natureza dos modelos de defasagens distribuídas essa intercorrelação deve ser alta, devemos sempre esperar coeficientes de determinação elevados.[10] Os resultados anteriores devem, conseqüentemente, ser julgados pela sua significância, e não pelo coeficiente de determinação apresentado.

Esta é a explicação para o elevado R^2 de meu modelo, e não a presença de viés de deflacionamento comum, como pretende King.

10 O coeficiente de correlação simples entre as caixas reais em *t* e t-1 situa-se em torno de 0,97, o que explica o elevado coeficiente de determinação do modelo.