

# Previsão da produção industrial: indicadores antecedentes e modelos de série temporal\*

RICARDO A. MARKWALD \*\*  
AJAX R. BELLO MOREIRA \*\*  
PEDRO L. VALLS PEREIRA \*\*

*O trabalho tem como objetivo a comparação da metodologia de indicador antecedente (leading indicator) com modelos estruturais de séries temporais visando à previsão das reversões cíclicas de uma série-alvo (no caso, o índice da produção da indústria geral). São explorados métodos alternativos de agregação das séries antecedentes, optando-se pela escolha de ponderadores obtidos a partir de uma regressão de mínimos quadrados ordinários entre a série-alvo e as séries antecedentes. Adicionalmente, é ajustada uma função de previsão que relaciona dinamicamente o indicador antecedente com a série-alvo e são calculadas também as probabilidades de ocorrência de reversão. Na linha sugerida por Neftçi, admite-se que a série-alvo possa ser representada por um processo estocástico que nos momentos (pontos) de reversão muda de regime. A partir desse pressuposto, o trabalho procede ao reconhecimento desses pontos com base nas séries-alvo observada e prevista, para uma dada probabilidade de falsa indicação.*

## 1 — Introdução

A previsão de variáveis econômicas admite o emprego de técnicas e métodos diversos. Dentre aqueles de uso mais difundido, cabe destacar os modelos econométricos que, baseados num modelo conceitual, ajustam funções que relacionam um conjunto restrito de variáveis econômicas, simulando, ainda, seus valores e trajetórias no contexto de cenários alternativos, nem sempre consistentes.

A técnica dos indicadores antecedentes (*leading indicators*) pressupõe a escolha de séries isoladas (indicadores parciais ou barômetros) ou a construção de índices compostos cuja trajetória antecipe o comportamento de uma série-alvo, servindo, portanto, como “preditor” dessa variável. Esta técnica não requer a especificação antecipada de um modelo teórico e não precisa definir de maneira estrita o sentido de causalidade entre as variáveis — como nos modelos econométricos —, podendo, portanto, aproveitar ao máximo as informações de um conjunto amplo de séries.

\* Os autores agradecem ao Dr. José Geraldo Hossanah Cordcero (Diace/Depec/Bacen), pela sua inestimável colaboração na montagem do banco de séries temporais.

\*\* Do Instituto de Pesquisas do IPEA.

Igualmente extenso é o uso de modelos estruturais de série temporal que, com base na história passada da própria variável, inferem seu valor futuro. São parcimoniosos quanto às hipóteses envolvidas e com potencial de captar mudanças no processo gerador das séries.

A capacidade de adaptação dos modelos estruturais de séries temporais parece especialmente interessante num contexto de instabilidade econômica, em face da hipótese básica da técnica de indicadores antecedentes que, basicamente, supõe a existência de um conjunto de séries que antecipem de forma constante a trajetória da série-alvo.

Um dos objetivos para o qual os indicadores antecedentes parecem se mostrar mais apropriados diz respeito à previsão das reversões (*turning points*) na trajetória de uma certa variável-alvo. Neste sentido, é preciso destacar que a previsão do nível e das reversões de uma série respondem a preocupações de caráter diferenciado e implicam problemas bastante distintos, ainda que relacionados.

A antecipação de pontos de reversão, no caso de variáveis representativas do nível da atividade econômica, é sem dúvida de vital importância para a adequada e oportuna implementação de políticas de caráter anticíclico. Esta preocupação não se encontra necessariamente presente nos exercícios rotineiros de previsão do nível de uma variável, os quais podem, no entanto, ser melhor desenvolvidos com base na correta antecipação das reversões, minimizando assim os erros de previsão, particularmente graves quando as reversões ocorrem em momentos próximos ao patamar inicial de projeção.

A previsão do nível da série implica um “preditor” que minimize os erros de previsão. A previsão das reversões requer, por sua vez, um “preditor” com capacidade para detectar os pontos de reversão sem incorrer em omissões e tornando mínima a indicação falsa de reversão. A despeito de sua maior exigência — decorrente da necessidade de prever tanto o nível como as reversões da série —, o primeiro dispõe de técnicas mais desenvolvidas. Já o tratamento estatístico rigoroso da previsão das reversões ainda está em fase de desenvolvimento, sendo que algumas das contribuições mais relevantes datam do início desta década.

Um procedimento alternativo à construção de um “preditor” das reversões — que exige uma função de perda assimétrica e que só é relevante no entorno dos momentos de reversão — resulta na formulação de Neftçi, que vê este problema como o do reconhecimento de rupturas no processo gerador das séries.

A previsão com base em indicadores antecedentes tem início com a obra de Mitchell (1913) e adquiriu verdadeira relevância com a publicação através do NBER de um conjunto de trabalhos editados por Moore (1961). Na década de 70, apareceram contribuições significativas devidas a Long (1970), Stekler (1972), Hymans (1973), Zarnovitz e Boschan (1975a e 1975b) e Wecker (1979). Dentre os trabalhos mais recentes, cabe destacar, ainda, Moore e Zarnovitz (1982), Neftçi (1982), Moore (1983), Zarnovitz (1985), Kling (1987) e Zellner, Hong e Gulati (1987).

No caso do Brasil, a técnica dos indicadores antecedentes tem sido utilizada por Contador (1977 e 1981), cabendo registrar, ainda, um trabalho recente desenvolvido no âmbito do Centro de Estudos Econômicos (1986), vinculado à empresa SERFINA.

Os modelos de indicadores antecedentes e os estruturais de série temporal têm qualidades complementares, e em princípio não é claro qual terá o melhor desempenho. Daí a proposta deste trabalho de construir uma função de previsão com base nos indicadores antecedentes e comparar os resultados quanto à aderência e à capacidade de identificar reversões, com as previsões de um modelo estrutural de série temporal.

Enfim, vamos avaliar se o ônus adicional de manter em operação a técnica de indicadores antecedentes é compensado por capacidade preditiva adicional.

Este trabalho utiliza como variável-alvo o índice da produção da indústria geral levantado pelo IBGE, que foi escolhido pelo fato de a produção industrial apresentar um comportamento de padrão semelhante ao do Produto Interno Bruto. De resto, a produção industrial é a série tradicionalmente usada como variável-alvo na maioria dos estudos [Nilsson (1987)].

Também foi implantada a metodologia de Neftçi que calcula as probabilidades de ocorrência de reversões e identifica as reversões das séries.

Além desta introdução, o trabalho inclui outras três seções:

- Indicadores antecedentes: hipóteses (Seção 2);
- Metodologia e principais resultados (Seção 3);
- Medida da probabilidade de reversão (Seção 4); e
- Conclusões (Seção 5).

## 2 — Indicadores antecedentes: hipóteses

A economia apresenta fases mais ou menos persistentes e pronunciadas de crescimento e queda da produção. O objetivo deste trabalho é antecipar o esgotamento de cada fase e prever o nível da produção utilizando modelos baseados em determinadas hipóteses, as quais pressupõem que:

a) os eventos econômicos medidos por séries temporais se relacionem de tal forma no tempo que o valor futuro da série-alvo seja determinado pelo valor presente de fatores não observáveis, e que estes fatores sejam representados por séries temporais relacionadas com a série-alvo (este pressuposto não implica necessariamente uma relação causal — no sentido de um modelo conceitual — entre as séries antecedentes e a série-alvo);

b) as séries temporais possam ser decompostas em componentes de tendência, sazonalidade, ciclo e irregular, e que a relação mencionada em *a* diga exclusivamente respeito às componentes de tendência e ciclo;

c) a relação de dependência no tempo das séries antecedentes com a série-alvo seja a mesma para todas as séries — no sentido de terem a mesma função de transferência —, uma vez que cada série tenha sido convenientemente defasada; e

d) a série-alvo possa ser representada por um processo estocástico  $\{Y_t\}$  que nos pontos de reversão muda de regime:

$$Y_t = Y_{t-1} + e_t$$

$$e_t \sim N(w) \quad \text{se } t \text{ é um período normal; e}$$

$$e_t \sim C(w) \quad \text{se } t \text{ é um período de crise.}$$

A hipótese *a* implica que as componentes de ciclo e tendência das diversas séries representem os diferentes fatores que deveriam esgotar os possíveis efeitos sobre a série-alvo. Para isto, as séries antecedentes devem cobrir os diversos aspectos da economia, o que é uma forma de compensar a falta do modelo conceitual.

As séries devem ter um número suficiente de observações, de forma que a relação estabelecida tenha validade estatística e sua estabilidade ao longo do tempo possa ser testada, tornando ainda possível a consideração de diferentes momentos da trajetória da série.

A hipótese *b* implica que as componentes de sazonalidade e irregular sejam específicas da série e não estejam relacionadas àqueles fatores comuns.

A hipótese *c* é necessária para tratar da relação dinâmica das séries antecedentes com a série-alvo. De fato, cuidar isoladamente de cada série antecedente levaria a um modelo com número excessivo de variáveis explicativas.

A hipótese *d* é necessária ao cálculo das probabilidades de reversão e implica supor que na economia tudo se passa como se houvesse dois estados — o normal e o de crise —, com leis próprias de funcionamento representadas pelas distribuições de probabilidade  $N(w)$  e  $C(w)$ .

Neste caso, o problema de reconhecer o momento de reversão é o mesmo de reconhecer o momento de mudança da lei de probabilidade que rege o funcionamento da economia.

### 3 — Metodologia e principais resultados

#### 3.1 — Elaboração do banco de séries temporais (banco Lead)

O objetivo da criação de um banco de séries temporais é dispor de um conjunto amplo de séries do qual possam ser selecionadas, por procedimento estatístico, aquelas que de fato antecedem a série-alvo.

O banco deve possuir alguns atributos genéricos como o de cobrir os diversos aspectos da economia e contar com séries que tenham um número

suficiente de observações, compatível com a dimensão temporal do fenômeno cíclico, objeto de estudo. A literatura que trata da construção de indicadores antecedentes destaca, ainda, alguns atributos específicos a serem preenchidos pelas séries antecedentes e que devem ser levados em consideração quando da montagem do banco, uma vez que aquelas são apenas um subconjunto deste. Neste sentido, ainda condicionado à base estatística disponível, é recomendável que:

- a) a periodicidade das séries seja mensal;
- b) as séries não estejam sujeitas a revisão freqüente, nem excessiva;
- c) a divulgação das séries seja imediata e de fácil obtenção; e
- d) as séries não apresentem descontinuidade nem encadeamentos.

Estas considerações nortearam a montagem do banco Lead, constituído de aproximadamente 240 séries agrupadas em nove blocos. As séries mais longas iniciam-se em janeiro de 1971 (mais de 200 observações), enquanto as mais curtas iniciam-se em janeiro de 1981 (mínimo de 85 observações). A estrutura e composição do banco Lead é a seguinte:

Bloco	Número de séries
Produção	61
Consumo de vendas	59
Moeda e crédito	15
Salário e emprego	36
Finanças públicas	6
Preços	3
Setor externo	29
Negócios e outros	23
Economia internacional	5
Total	238

O banco é pobre em série como estoques, encomendas, pedidos em carteira, novos pedidos, sondagens e enquetes, etc. No entanto, inexistem no Brasil levantamentos confiáveis, rotineiros e contínuos que cubram estes aspectos da atividade econômica. Neste sentido, o banco reproduz apenas as limitações de nossa base estatística.

### 3.2 — Identificação das séries antecedentes

Após a montagem do banco Lead, procedeu-se à seleção daquelas séries efetivamente relacionadas com a série-alvo. Dado que as séries coletadas tinham dimensão diversa (séries de fluxo, de estoque, em valores nominais ou em valores reais, etc.), foi preciso estabelecer alguns critérios para homogeneizar as séries. Estes critérios, chamados de filtros, foram os seguintes:

a) Se a série  $A_t$  mede valores acumulados dentro do ano, ela é transformada em fluxo:

$$F_t = A_t - A_{t-1} \quad \text{para } t \neq \text{janeiro; e}$$
$$F_t = A_t \quad \text{para } t = \text{janeiro.}$$

b) Se a série está medida em cruzados correntes, é deflacionada pelo IGP-DI.

c) Se a série é um fluxo, isto é, mede os acontecimentos relativos a um mês isolado, é tomada a média móvel de 12 meses para suavizá-la, ou seja, para diminuir a variância da componente irregular:

$$y_t = (F_t + F_{t-1} + \dots + F_{t-11})/12$$

d) Se a série registra a evolução de um estoque, ou seja, a acumulação de valores ao longo do tempo, ela já está suavizada e não necessita da transformação sugerida em c.

Vale mencionar que a série de estoque deveria receber um tratamento diferente do mencionado para que fosse homogeneizada da mesma forma. Neste caso:

$$F_t = E_t - E_{t-1}$$
$$y_t = (F_t + F_{t-1} + \dots + F_{t-11})/12 = (E_t - E_{t-12})/12$$

Entretanto, este filtro, quando aplicado às séries de estoque do banco, resultou em séries com variância excessiva, mesmo após a aplicação de c. Optou-se, então, por não utilizar este procedimento aparentemente mais correto.

Os critérios a a d acima mencionados obedeceram à necessidade de homogeneizar as séries e, adicionalmente, diminuir a variância da componente irregular. O passo seguinte foi a retirada das componentes sazonal e tendencial. Para tanto:

e) Foi calculada a taxa de variação de 12 meses de cada série objetivando retirar a componente sazonal suposta multiplicativa.

No que se refere à retirada da componente tendencial, trata-se de uma opção aparentemente não considerada nos trabalhos precedentes. Contudo, conforme já destacado na literatura [ver Granger e Nelbold (1974)], verificou-se que a identificação das séries antecedentes torna-se problemática quando é mantida a componente tendencial, uma vez que os resultados ficam mais ambíguos.

De fato, as funções de autocorrelação e correlação cruzada defasada calculadas sobre séries com tendência de crescimento apresentam um decaimento suave, o que torna o reconhecimento da defasagem característica mais ambíguo e pode apresentar correlações significativas entre séries não relacionadas. Com base nesta constatação foi introduzido mais um filtro:

f) Para aquelas séries que, após  $e$ , ainda apresentaram tendência foi tomada, recorrentemente, a diferença de primeira ordem até obter estacionaridade.

O critério de testar a não-estacionaridade consiste em verificar se a série co-integra com a constante [ver Engle e Granger (1987)]. Seja:

$$dw = \sum_{t=2}^n (y_t - E(y)) * (y_{t-1} - E(y)) / \sum_{t=2}^n (y_t - E(y))^2$$

onde  $n$  é o número de observações da amostra.

No caso em que  $dw < 0,335$ , então a série ainda tem tendência, sendo considerada estacionária no caso contrário. Com base neste teste, todas as séries ficaram estacionárias após a primeira diferença.

Em suma, a aplicação destes filtros objetivou o isolamento da componente cíclica, possibilitando, assim, testar a relação dinâmica entre esta série e a série-alvo.

O critério para o reconhecimento da dependência baseou-se na função de correlação cruzada entre a série-alvo e as séries antecedentes previamente filtradas. Procedeu-se, assim, ao cálculo dessa correlação entre a série-alvo ( $y$ ) e cada uma das séries candidatas ( $x$ ) para diversas defasagens da série candidata, ou seja:

$$\rho_j = E\{ (x_{t-j} - E(x_{t-j})) (y_t - E(y_t)) \} \quad j = -5, \dots, 0, \dots, 12$$

Foram selecionadas aquelas que antecipavam a série-alvo, mas não eram antecipadas por esta. Este é um critério aproximado de selecionar as séries causadoras (do ponto de vista estatístico) da série-alvo, mas não causadas por ela. Adicionalmente, foi adotado o critério de que as séries selecionadas antecipassem de forma efetiva – já considerando o atraso na publicação – em quatro meses à série-alvo.

Em síntese, foram escolhidas séries tais que:

$$\rho_{\tau} = \max_{\tau} \{ \rho_{\tau} \}$$

$$\tau \geq 4 \quad \text{e} \quad \tau \geq n^{-1/2}$$

O parâmetro  $\tau$ , como definido, é chamado de defasagem característica da série e foi utilizado como referência para colocar todas as séries em fase com a série-alvo.

O procedimento de filtrar e identificar as séries relevantes foi implementado com todas as séries do banco de dados e resultou na seleção do conjunto de 66 séries.

O cálculo da correlação foi realizado com tamanhos de amostras diferentes, dependendo do comprimento da série candidata. A maior parte das séries permitia o cálculo da correlação a partir de janeiro de 1981. Tendo em vista avaliar a estabilidade da correlação estimada, o cálculo da mesma foi desenvolvido para o período 1981/88 e repetido para o período mais curto (1981/88).

O teste de estabilidade<sup>1</sup> da relação de dependência permitiu, juntamente com outros critérios — inexistência de correlação significativa de defasagens negativas, sinal correto da correlação e antecipação efetiva suficiente —, a seleção das seguintes séries antecedentes para compor o indicador (a relação das séries selecionadas é apresentada juntamente com suas fontes, defasagem característica, atraso na publicação e mnemônico):

Fonte	$\epsilon$	Atraso	Nome
ASCP	12	2	Falências na indústria metalúrgica em São Paulo
ACSP	12	2	Falências no comércio de tecidos & armarinho em São Paulo
ACSP	11	3	Faturamento de bens de consumo imediato
Anfavea	9	3	Venda de veículos nos mercados externo e interno
Acesp	7	1	Consultas de venda a prazo (SPC) — total São Paulo
Cacex	9	3	Exportação de produtos básicos
Cacex	9	3	Exportação de semi-manufaturados (US\$)
Bacen	12	3	Dívida mobiliária federal fora Bacen

Vale notar que só foram consideradas as séries que antecipassem efetivamente — defasagem característica menos atraso da publicação — em pelo menos seis meses a série-alvo.

### 3.3 — Construção de indicadores antecedentes

Selecionadas as séries, o indicador antecedente fica determinado dentro da classe das funções lineares por ponderadores e pela defasagem característica — defasagem associada à maior correlação — com que cada série antecedente é considerada. Assim:

$$I_t = \sum_{j \in A} a_j X_{j,t-\epsilon_j}$$

onde  $A$  é o conjunto das séries antecedentes.

<sup>1</sup> O cálculo das correlações foi realizado utilizando a amostra toda e apenas os últimos cinco anos. A dependência foi considerada estável se a correlação se manteve significativa nas duas amostras.



A consideração de séries antecedentes com e sem tendência, bem como de esquemas de ponderação alternativos, resultou na construção de três indicadores diferenciados:

– Indicador *IC*:

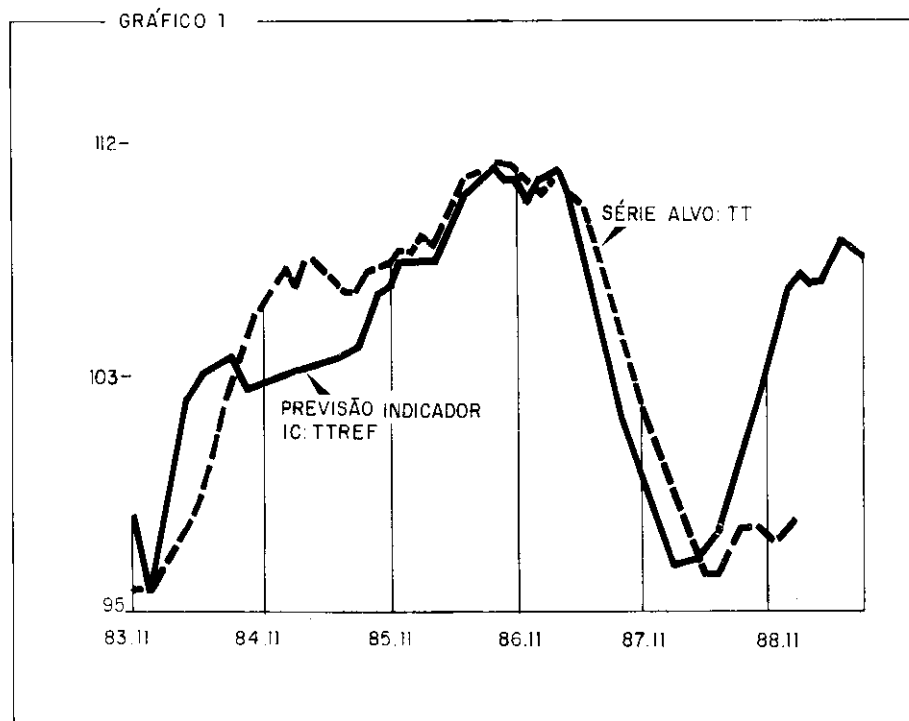
Foi construído de acordo com a metodologia usual, ou seja, é uma soma das séries antecedentes devidamente padronizadas e defasadas, ponderada pela sua correlação na defasagem característica com a série-alvo, ou seja:

$$\alpha_j = \rho_j / \sigma_j$$

onde  $\rho_j$  é a correlação com a série-alvo na defasagem característica e  $\sigma_j$  é o desvio-padrão da série antecedente.

O Gráfico 1 apresenta este indicador numa escala conveniente, juntamente com a série-alvo, podendo-se observar a instabilidade da relação de precedência do indicador construído, o que sem dúvida condiciona sua capacidade preditiva. Não foi possível identificar qual ou quais séries seriam responsáveis por esta indesejada instabilidade.

– Indicador *ID*:



Foi construído de forma semelhante ao anterior, mas com séries transformadas em estacionárias. Os resultados obtidos foram mais uma vez pouco satisfatórios.

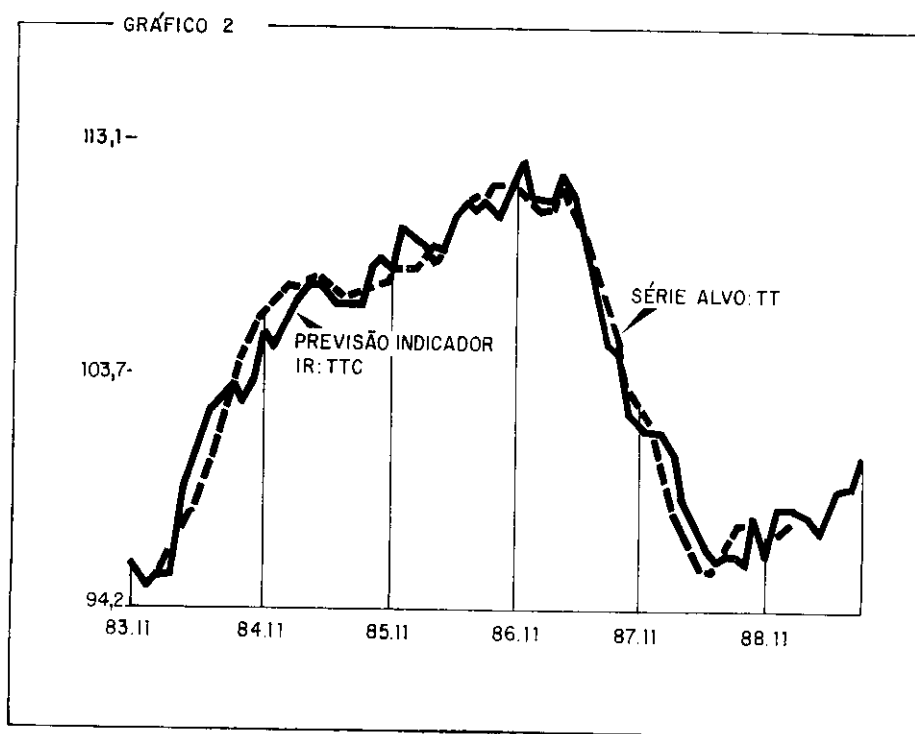
– Indicador *IR*:

Os ponderadores, neste caso, foram obtidos a partir do ajustamento de uma regressão de mínimos quadrados ordinários entre a série-alvo e as séries antecedentes, cabendo aqui algumas considerações.

A hipótese *a* da Seção 2 estabeleceu a premissa da existência de fatores comuns não observáveis condicionantes da série-alvo no futuro e representados pelas séries antecedentes no presente. Portanto, estas séries são interdependentes e, possivelmente, a inversa da matriz de covariância entre estas séries seja mal condicionada, o que implica não-unicidade dos estimadores de mínimos quadrados ordinários.

Entretanto, combinações lineares destes parâmetros são estimáveis, isto é, satisfazem o teorema de Gauss-Markov, ou seja, o “preditor” da série-alvo (*Y*) obtido é estimável de forma eficiente e consistente, apesar de os ponderadores não serem estimáveis isoladamente, o que impede a interpretação dos seus valores.

Então, o indicador *IR* definido acima é ótimo no sentido de ser o melhor “preditor” da série-alvo, dado o conjunto de informações. O Gráfico 2 apresenta este indicador, bem como a série-alvo.



### 3.4 — Função de previsão

Os indicadores construídos são compostos de séries que têm uma relação dinâmica com a série-alvo. Portanto, em princípio, a função de previsão deve ser dinâmica, ficando definida pela ordem das defasagens distribuídas da variável independente, pela ordem do polinômio auto-regressivo da série-alvo e pela ordem de diferença, se os polinômios tiverem raízes unitárias, utilizadas para se transformar em estacionárias as séries consideradas.

Em princípio, não existe qualquer motivo que indique a necessidade de tomar a diferença das séries e retirar a tendência de crescimento das variáveis para estimar a função de previsão. Por este motivo, foram considerados tanto modelos no nível, utilizando-se os indicadores *IC* e *IR*, como modelos nas primeiras diferenças, utilizando-se o indicador *ID*.

É apresentada, a seguir, a função de autocorrelação da série-alvo e de correlação cruzada com os indicadores construídos, confirmando a existência de relação dinâmica e sugerindo, ainda, que o polinômio auto-regressivo seja de ordem 2 nas primeiras diferenças das séries (o polinômio de defasagens distribuídas do indicador foi definido arbitrariamente com a mesma ordem do polinômio auto-regressivo):

*Correlação cruzada e autocorrelação da série-alvo*

Defasagem	-5	-4	-3	-2	-1	0	1	2	3	4	5
<i>IC, Y</i>	.48	.57	.64	.70	.74	.76	.78	.78	.76	.73	.68
<i>ID, ΔY</i>	.24	.37	.44	.52	.60	.63	.54	.48	.39	.35	.28
<i>IR, Y</i>	.73	.79	.86	.91	.95	.97	.95	.92	.87	.81	.74
$\Delta Y, \Delta Y$							.79	.75	.62	.49	.38
$\Delta Y, \Delta Y^*$							.79	.35	-.13	-.19	-.08

A linha  $\Delta Y, \Delta Y^*$  refere-se à função de autocorrelação parcial da série-alvo.

Considerando os correlogramas acima, foi julgado necessário incluir componentes da variável independente até a defasagem de ordem 3 (para as variáveis no nível) e de ordem 2 (para as variáveis medidas na primeira diferença). Portanto, os modelos considerados foram:  $AD(2,2)$ <sup>2</sup> sobre as diferenças das séries,  $AD(3,3)$  no nível. Observe-se que o primeiro modelo é um caso particular do segundo.

<sup>2</sup> A notação  $AD(n,m)$  significa um modelo com defasagens distribuídas de ordem  $m$  na variável independente e de ordem  $n$  no polinômio auto-regressivo da variável dependente. Para uma descrição deste tipo de modelo, ver Hendry, Pagan e Sargan (1984).

#### A – Modelo com mecanismo de correção de erro (MCE)

O modelo MCE tem sido amplamente usado desde sua formulação inicial devida a Sargan (1964). Este tipo de modelo tem a vantagem de reter informações sobre o nível das séries, de modo que as relações de longo prazo entre as variáveis do modelo estejam presentes. Mais recentemente, Engle e Granger (1987) introduziram o conceito de série integrada de ordem "d" e o conceito de co-integração<sup>3</sup> entre séries e mostraram que os modelos MCE são equivalentes a séries co-integradas.

As séries  $Y$  e as componentes de  $IR$  têm tendência de crescimento e tornam-se estacionárias na primeira diferença, ou seja, são ditas integradas de ordem 1. Adicionalmente, a hipótese de que  $Y$  e as componentes de  $IR$  são co-integradas de ordem 1 não é rejeitada.

Então, o índice  $IR$  pode ser interpretado como o valor de equilíbrio de longo prazo para  $Y$ , e o modelo MCE, juntamente com a relação dinâmica acima sugerida, define a relação de curto prazo:

$$\begin{aligned} \Delta Y_t = & \alpha \Delta Y_{t-1} + \beta \Delta Y_{t-2} + \gamma \Delta IR_t + \varepsilon \Delta IR_{t-1} \\ & + \phi \Delta IR_{t-2} + \delta (Y_{t-1} - IR_{t-1}) \end{aligned}$$

Vale mencionar que a relação acima foi especificada lembrando a hipótese  $c$  da Seção 2, que determina que o padrão dinâmico de resposta das séries componentes é o mesmo para todas as séries.

#### B – Modelo no nível

Considerando a sugestão de relação dinâmica acima mencionada, e dado que o indicador  $IC$  não co-integra com  $Y$ , temos:

$$\begin{aligned} Y_t = & \alpha Y_{t-1} + \beta Y_{t-2} + \gamma Y_{t-3} + \delta IC_t + \varepsilon IC_{t-1} + \\ & + \phi IC_{t-2} + \eta IC_{t-3} \end{aligned}$$

#### C – Modelo com séries estacionárias, primeira diferença

Considerando a relação dinâmica mencionada e que não cabe o uso do mecanismo de correção de erros aplicados sobre séries já estacionárias, temos:

$$\Delta Y_t = \alpha \Delta Y_{t-1} + \beta \Delta Y_{t-2} + \gamma ID_t + \varepsilon ID_{t-1} + \phi ID_{t-2}$$

<sup>3</sup> Para uma resenha da teoria de co-integração, ver Pereira (1988).

### 3.5 — Modelo estrutural de série temporal

O modelo de série temporal utilizado como referência é o descrito em Harvey e Pereira (1989), que decompõe a série nas componentes de tendência, ciclo, sazonalidade e efeito do número de dias trabalhados em cada mês. Previsões para o nível da produção industrial são, então, obtidas por este modelo e servirão de comparação com a função de previsão do indicador antecedente.

O modelo de série temporal utilizado como referência é o descrito em Pereira (1988), que decompõe a série nas componentes de tendência ( $\mu$ ), ciclo ( $\psi$ ), sazonalidade ( $\xi$ ) e efeito do número de dias trabalhados em cada mês ( $\gamma$ ) e realiza previsões para o nível da produção industrial.

A discussão e detalhamento deste modelo está descrita na referência. Serão apresentadas, a seguir, apenas as equações de definição do modelo, que foi estimado utilizando o programa STAMP (Structural Time Series Analyser Modeller and Predictor) de S. Peters:

$$y_t = \mu_t + \psi_t + \xi_t + \gamma_t d_t + e_t$$

onde:

$$\begin{aligned} \mu_t &= \mu_{t-1} + \beta_{t-1} + \xi_1 \\ \beta_t &= \beta_{t-1} + \xi_2 \\ \psi_t &= \rho_{t-1} \psi_{t-1} \cos(w) + \rho_{t-1} \psi_{t-1}^* \sin(w) + \xi_3 \\ \psi_t^* &= -\rho_{t-1} \psi_{t-1} \sin(w) + \rho_{t-1} \psi_{t-1}^* \cos(w) + \xi_4 \\ \xi_t &= \xi_{t-1} \phi + \xi_5 \\ \gamma_t &= \gamma_{t-1} + \xi_6 \end{aligned}$$

As variáveis  $\xi$  referem-se a choques aleatórios associados a cada variável de estado, e  $d_t$  refere-se ao número de dias trabalhados padronizados de forma a ter média zero.

O sistema STAMP estima as variâncias dos choques aleatórios e os parâmetros  $\rho$  e  $\omega$  que definem o decaimento e o ciclo da produção industrial. Com base nos parâmetros do modelo estimados para cada período, foram realizadas previsões a  $h$  passos dentro da amostra e construída a taxa de crescimento a 12 meses do valor acumulado nos últimos 12 meses, que é a variável de previsão do modelo dos indicadores antecedentes.

Como a variável de previsão do modelo dos indicadores antecedentes foi a taxa de crescimento a 12 meses, construiu-se, com base nos componentes suavizados do modelo estrutural de série temporal, previsões a  $h$  passos dentro da amostra e, a seguir, a taxa de crescimento a 12 meses do valor acumulado nos últimos 12 meses foi obtida utilizando-se estas previsões.

### 3.6 — Comparação dos resultados

Os modelos com base nos indicadores antecedentes e de série temporal foram estimados, construindo-se, então, para cada um dos modelos, a previsão a  $h$

passos no período abril de 1982/janeiro de 1989. Estimou-se o desvio-padrão do erro de previsão a partir do erro médio quadrático, para cada caso (a minimização deste erro constitui o critério de escolha dos modelos) :

*Desvio-padrão do erro de previsão*

Horizonte	1	2	3	4	6	12
MCE (A)	0,504	0,72	0,923	1,038	1,078	—
Nível (B)	0,53	0,871	1,276	1,649	—	—
Prim. dif. (C)	0,562	0,959	1,445	1,991	—	—
Indicador antecedente	1,077	1,077	1,077	1,077	1,077	—
Série temporal	0,157	0,432	0,562	0,761	1,118	1,762

Os resultados acima mostram que:

a) o modelo de série temporal é melhor que os modelos de indicadores antecedentes para previsões até quatro passos, sendo equivalente ao modelo MCE para previsões a seis passos;

b) o modelo MCE é superior em todos os horizontes aos demais modelos que utilizam indicadores antecedentes;

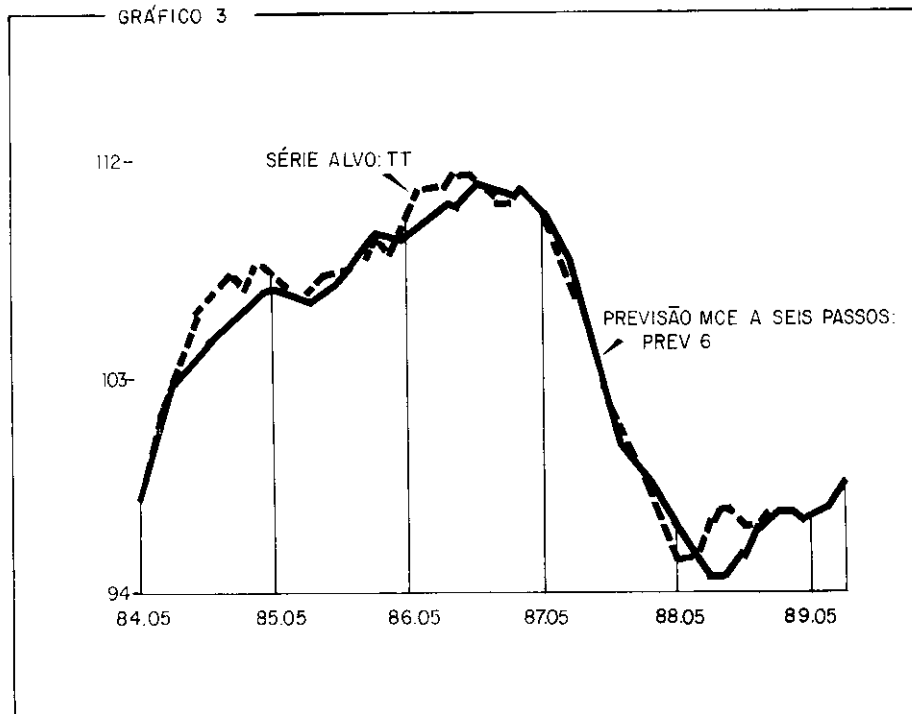
c) os resultados do modelo MCE aproximam-se do indicador antecedente — que é por construção seu ajustamento de longo prazo — à medida que o horizonte de previsão cresce; e

d) os demais modelos, embora sejam aproximadamente iguais para a previsão a um passo, perdem rapidamente a sua qualidade preditiva.

Em síntese, o modelo MCE é o melhor dentre os que utilizam indicadores antecedentes e equivalente ao modelo de série temporal para previsões no limite de sua capacidade de antecipação. Resta avaliar o comportamento destes modelos quanto à capacidade de identificar reversões (os Gráficos 3 e 4 mostram o desempenho dos modelos MCE e de série temporal na previsão a seis passos dentro da amostra).

#### 4 — Medida da probabilidade de reversão

A hipótese *d* da Seção 2 implica a existência de momentos  $t^*$  onde há a troca de regimes. Reconhecer estes momentos é o problema do tempo ótimo de parada. Ele consiste em determinar o momento  $\tau$  em que será reconhecida a reversão ocorrida em  $t^*$ , com um atraso médio mínimo e para uma dada probabilidade de falsa indicação  $\alpha$ .



Para isto temos o teorema mencionado em Neftçi (1982) e devido a Shirayev (1973), que determina:

$$\text{Se } \pi_t = \text{Prob} (t \geq t^*/F_t), \text{ então } \tau = \inf_t \{t : \pi_t \geq A(\alpha)\}$$

ou seja, dada a probabilidade  $\pi_t$  de ruptura e dado o conjunto de informação  $F_t$ , o momento ótimo de parada  $\tau$  é o primeiro momento tal que  $\pi_t$  seja maior do que uma função da probabilidade de falso sinal.

Em Neftçi, existe ainda um lema adicional que calcula  $\pi_t$  recursivamente:

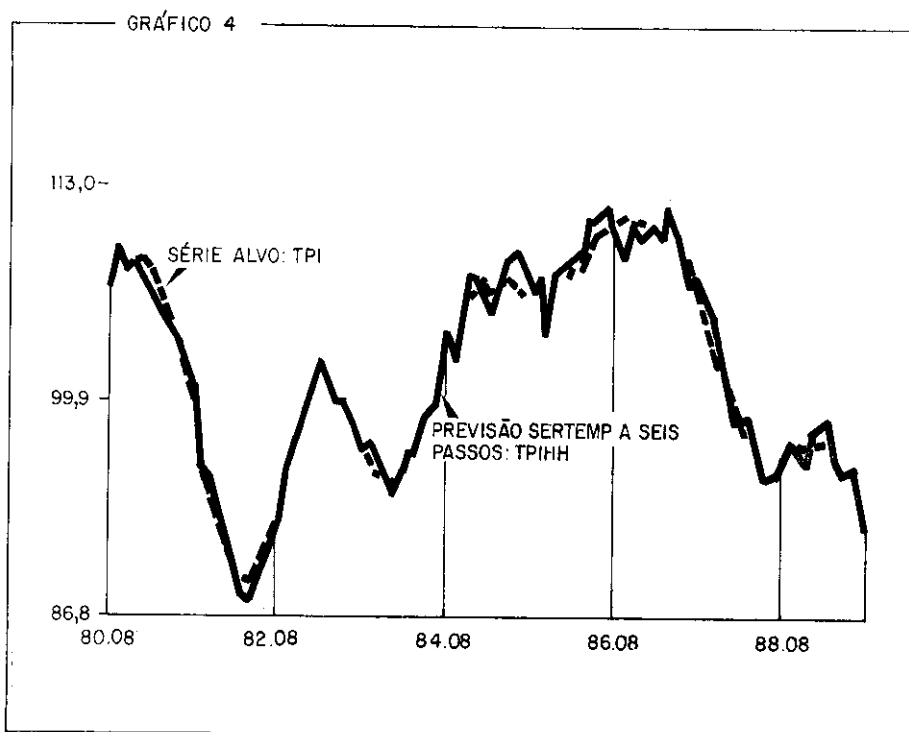
$$\pi_t = (\pi_{t-1} + \xi_t(1 - \pi_{t-1})) \mu_t / ((\pi_{t-1} + \xi_t(1 - \pi_{t-1})) \mu_t + (1 - \pi_{t-1})(1 - \xi_t) \nu_t)$$

onde:

$\xi_t$ : probabilidade *a priori* de ocorrência de reversão;

$\mu_t$ : Prob ( $\Delta Y \in [\Delta Y - \epsilon, \Delta Y + \epsilon]$ /regime de crise); e

$\nu_t$ : Prob ( $\Delta Y \in [\Delta Y - \epsilon, \Delta Y + \epsilon]$ /regime normal).



Vale mencionar que a expressão acima calcula a probabilidade de ocorrer uma crise dado que se está no regime normal, devendo ser derivada uma expressão análoga para calcular a probabilidade de passarmos de um regime de crise a um regime normal.

As probabilidades  $\mu_t$  e  $\nu_t$  foram estimadas a partir dos dados da amostra (1977/88), através da frequência relativa em diversas faixas de variação no nível da série-alvo. A identificação dos períodos de mudança de regime da série-alvo foi feita antes de se obter a estimativa das probabilidades.

A probabilidade *a priori*  $\xi_t$  é usualmente considerada como sendo função do comprimento esperado do ciclo econômico. No caso brasileiro, a instabilidade recente dos regimes dificulta esta forma de uso, não sendo, portanto, considerada no modelo.

Na tabela abaixo, estão indicados: o tipo de troca de regime — (C/N) crise para normal e (N/C) de normal para crise —, os momentos  $t^*$  de troca de regime e os momentos de reversão  $\tau$  identificados por este modelo, utilizando como nível de corte a probabilidade de reversão de 99%, para a série observada (Obs.) e para as previsões realizadas pelo modelo de série temporal (ST) a quatro, seis e 12 passos e o modelo MCE a quatro e seis passos:



*Data de mudança de regime*

Regime	$t^*$	Obs.	ST-4	ST-6	ST-12	MCE-4	MCE-6
C/N	12/77	—	—	—	—	—	—
N/C	12/80	—	—	—	—	—	—
C/N	2/82	3/82	3/82	4/82	6/82	—	—
N/C	2/83	5/83	8/83	7/83	6/87	—	—
C/N	11/83	1/84	3/84	1/84	1/84	3/84	—
N/C	9/86	4/87	2/87	4/87	1/87	2/87	12/86
C/N	6/88	7/88	8/88	2/88	7/88	10/88	10/88
N/C	—	—	2/89	N	4/89	N	N
Falso sinal		87/12	—	—	6/86 2/88	9/85	9/85

Na tabela acima, o sinal (—) indica que não foi possível identificar a reversão por falta de dados e (N) que a reversão mencionada não foi identificada pelo método. Conclui-se que:

a) para a série observada, o reconhecimento da reversão foi feito com atraso de dois a três meses, o que pode ser considerado satisfatório; a reversão de 1986 foi reconhecida com um atraso de sete meses, em virtude da dificuldade de identificar o momento da reversão, o que faz com que a determinação deste atraso fique prejudicada; e

b) os resultados para os modelos de série de tempo e indicadores antecedentes são equivalentes e razoáveis e não são significativamente afetados pela antecipação com que as previsões são realizadas.

## 5 — Conclusões

Em síntese, o trabalho estabeleceu:

a) a previsão com seis meses de antecedência do nível da série-alvo com erro de  $\pm 2\%$ , considerando o atraso na publicação das séries antecedentes, bem como da série-alvo (o avanço operacional é de quatro meses em relação ao período corrente);

b) uma medida da probabilidade de reversão calculada com base na série observada, bem como nas previstas com seis ou 12 meses de antecedência, confirmando a ocorrência — ou indicando a iminência — de reversão, com um atraso da ordem de dois meses na identificação dos momentos das reversões; e

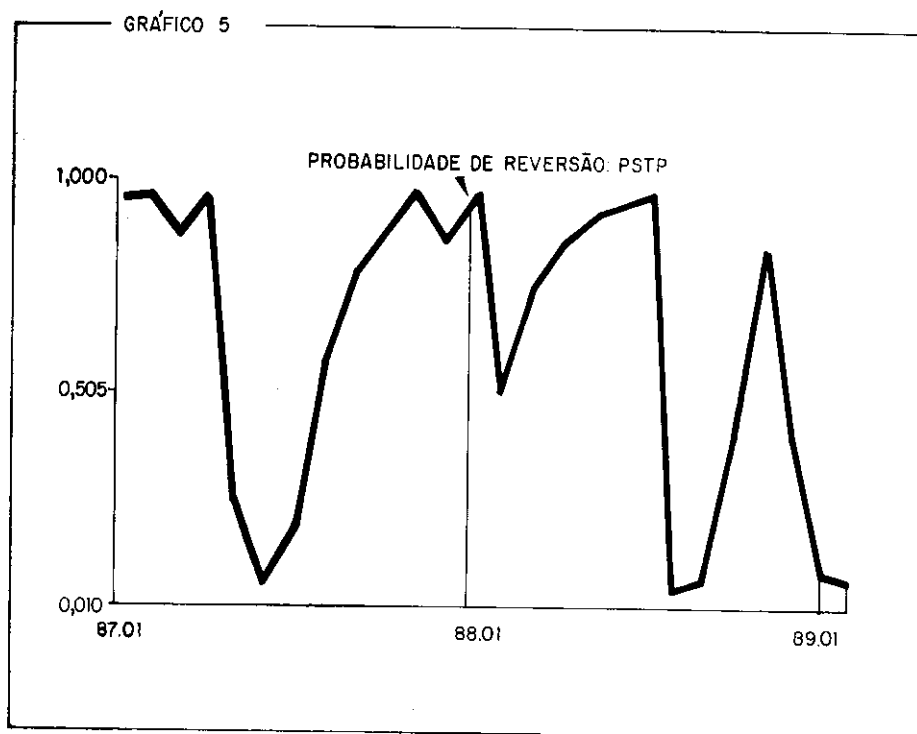
c) que os modelos de série temporal para a previsão do nível e da identificação das reversões são tão bons ou melhores que os modelos que utilizam

os indicadores antecedentes. A maior simplicidade de operação do modelo de série temporal e a maior fundamentação teórica tornam preferível o seu uso.

Parece surpreendente que os modelos de indicadores antecedentes que utilizam um maior volume de informações sejam superados pelos modelos estruturais de séries temporais. Entretanto, vale mencionar que a hipótese de estabilidade da defasagem característica não se verifica para as séries antecedentes em alguns períodos, talvez devido à instabilidade da economia brasileira nos períodos recentes. De outro lado, os modelos estruturais de série temporal, que por sua natureza são adaptativos, devem funcionar em contextos de instabilidade.

Os resultados que indicam a preferência pelos modelos estruturais de série temporal estão restritos, obviamente, para esta série-alvo, utilizando as metodologias adotadas e as séries disponíveis no banco, e não podem, a rigor, ser generalizados. Entretanto, considerando a amplitude do banco, os cuidados metodológicos e os comentários sobre as limitações dos indicadores antecedentes, fica implícita uma preferência pelos modelos de série temporal em geral.

O Gráfico 5 ilustra a probabilidade de reversão calculada com base na série-alvo, mostrando, por exemplo, que no mês de maio de 1989, e dadas



as previsões referentes ao mês de julho (avanço operacional de dois meses), a probabilidade de ocorrência de reversão teria sido julgada iminente. Já em julho, consideradas válidas as previsões do modelo para setembro, a fórmula de Neftçi confirmaria, com atraso mínimo, a ocorrência de uma mudança de regime antes de setembro. Os dados posteriores referentes à produção industrial de julho e agosto – não disponíveis no momento do cálculo da probabilidade de reversão – confirmariam que, de fato, ocorreu uma reversão em julho de 1988.

Ainda com base nas previsões do modelo de séries temporais em março de 1989, a fórmula de Neftçi não reconhece a possibilidade de nova reversão.

### Abstract

*The main purpose of this paper is to compare the leading indicator methodology with basic structural time series models, in order to predict the turning points of the general industrial production index, which is the target series chosen. Several methods of aggregation are used in order to obtain the weights of the series which are used to construct the leading indicator. The best weights are given by least square method. In order to improve the forecast property of the model, an error correction model was estimated. Another result, following Neftçi (1982), obtained in this paper is the probability of turning points. In order to derive this probability, the target series is assumed to be a stochastic process which switch regimes in a neighborhood of the turning points. These probabilities are computed, given a probability of false indication of turning point, for the observed and predicted target series.*

### Bibliografia

- CENTRO DE ESTUDOS ECONÔMICOS. *Indicador antecedente da indústria de transformação*. Mimeo, 1986.
- CONTADOR, C. R. *Ciclos econômicos e indicadores de atividade no Brasil*. Rio de Janeiro, IPEA/INPES, 1977 (Relatório de Pesquisa, 35).
- . *A previsão de ciclos econômicos com indicadores antecedentes*. Rio de Janeiro, Coppead/UFRJ, 1981 (Relatório Técnico, 50).
- ENGLE, R. F., e GRANGER, C. W. J. Co-integration and error correction: representation, estimation and testing. *Econometrica*, 55:251-76, 1987.
- GRANGER, C. W. J., e NELBOLD, P. Spurious regressions in econometrics. *Journal of Econometrics*, 2:111-20, 1974.
- HARRISON, P. J., e AMEEN, J. R. M. *Discount a regression method with non stationary parameters*. University of Warwick (Working Paper, 35).

- HARRISON, P. J., e JOHNSTON, F. R. *A regression method with non stationary parameters*. University of Warwick (Working Paper, 35).
- HARRISON, P. J. e STEVENS, C. F. *Baysean forecast*. University of Warwick (Working Paper, 13).
- HARVEY, A. C. *Time series models*. Philip Allan Publishers Limited, 1981.
- . *The econometric analysis of time series*. Philip Allan Publishers Limited, 1984.
- HARVEY, A. C. e PEREIRA, P. L. V. Trend seasonality and seasonal adjustment. In: MENTZ, R. P., ed. *Proceeding of the first seminar in applied statistics*. 1989.
- HENDRY, D. F., PAGAN, A. R., e SARGAN, J. D. Dynamic specification. In: GRILICHES, Z. e INTRILIGATOR, M., eds. *Handbook of econometrics*. Amsterdam, North Holland, 1984, v. II.
- HYMANS, S. H. On the use of leading indicators to predict cyclical turning points. *Brookings Papers on Economics Activity*, 2:339-81, 1973.
- KLING, J. L. Predicting the turning points of business and economic time series. *Journal of Business*, 60:201-38, 1987.
- LONG, R. Forecasting specific turning points. *Journal of American Statistical Association*, 65, 1970.
- MITCHELL, W. C. *Business cycles*. Berkeley, University of California Press, 1913.
- MOORE, G. H. *Business cycle indicators*. Princeton, Princeton University Press, 1961, v. 1.
- . *Business cycles, inflation and forecasting*. Cambridge, Mass. Ballinger, 1983.
- MOORE, G. H., e ZARNOVITZ, V. Sequential signals of recession and recovery. *Journal of Business*, 55:57-85, 1982.
- NEFTÇI, N. S. Optimal prediction of cyclical downturns. *Journal of Economic Dynamics and Control*, 4:225-41, 1982.
- NILSSON, R. OECD leading indicators. *OECD Studies*, (9), 1987.
- PEREIRA, P. L.V. Co-integração: uma resenha com implicações a séries brasileiras. *Revista de Econometria*, VIII(2), 1988.
- SARGAN, J. D. Wage and prices in the United Kingdom: a study in econometric methodology. In: HART, P. E., MILLS, G., e WITAKER, J. K., eds. *Econometric analysis for national economic planning*. London, Butterworths, 1964.

- SHIRYAYEV, A. N. *Statistical sequential analysis: on optimal stopping rules*. Providence, American Mathematical Society, 1973. (Translations of Mathematical Monographs).
- STEKLER, H. O. An analysis of turning points forecasts. *American Economic Review*, 62:274-8, 1972.
- WECKER, W. E. Predicting the turning points of a time series. *Journal of Business*, 72:35-50, 1979.
- ZARNOVITZ, V. Recent work on business cycles in historical perspective. *Journal of Economic Literature*, 23:523-80, 1985.
- ZARNOVITZ, V., e BOSCHAN, C. Cyclical indicators: an evaluation and new leading indexes. *Business Conditions Digest*, pp. v-xxii, 1975a.
- . New composite indexes of coincident and lagging indicators. *Business Conditions Digest*, pp. v-xxiv, 1975b.
- ZELNER, A., HONG, C., e GULATI, G. H. *Turning points in economic time series, loss structures and Bayesian forecasting*. Graduate School of Business University of Chicago, 1987 (Discussion Paper).

(Originais recebidos em maio de 1989. Revisitos em julho de 1989.)