
"Não fosse o futuro tão incerto, as cartomantes e os economistas iriam a falência"

John K. Galbraith

Este relatório principal reproduz os resultados de um sistema de previsão com indicadores antecedentes para o fluxo de passageiros de vôos domésticos e internacionais, e é parte de uma linha de pesquisa encomendada pelo Ministério do Turismo¹. O relatório apresenta a metodologia, os resultados detalhados dos indicadores antecedentes e os testes estatísticos sobre o desempenho do sistema. Um relatório adicional apresenta a lista-gem das variáveis utilizadas no sistema e detalhes estatísticos.

1 – Introdução

Estima-se que em 2005 entraram 5 milhões de turistas no Brasil, que apesar de um número alvissareiro em comparação com o passado, é modesto em comparação com o observado em outros países. Entre 1983 e 2004, a entrada de turistas estrangeiros no Brasil cresceu a taxa média anual de quase 6 %, contra um crescimento médio de apenas 2,7 % do PIB. E a partir de 1990, o ritmo de crescimento atingiu mais de 11,2 % ao ano, contra 2,5 % para o PIB. Como a base é pequena fica fácil ostentar altas taxas de crescimento, mas manter um crescimento acelerado por um horizonte maior vai exigir diversas medidas articuladas entre o setor privado e o governo, principalmente com suporte de ações complementares de melhoria de infraestrutura e de melhor divulgação.

A principal via de entrada de turistas é o transporte aéreo, com quase 70 % do total, seguido pelo transporte terrestre, através das fronteiras secas, com mais 24 %. Considerando a extensão do Brasil, o transporte aéreo deve manter a predominância na preferência dos turistas estrangeiros, e é este segmento que deve atrair a atenção dos investimentos e melhorias de infraestrutura no médio e longo prazo.

¹ Pesquisa “Descrição de perfis e dinâmica da oferta e demanda de serviços turísticos”, Ministério do Turismo, dezembro de 2006



Aéreo	69,7 %
Terrestre	24,0 %
Marítimo	1,1 %
Fluvial	0,4 %
Não especificado	4,8 %

Fonte : Embratur

Figura 1 – Vias de acesso de turistas, 2004

No curto prazo, a infra-estrutura e demais condições de hotelaria, transporte interno, apoio de agências e profissionais qualificados estabelecem um limite ao crescimento. É interessante notar que o fluxo de turistas apresenta forte flutuação, que independe das flutuações cíclicas da atividade geral. De fato, a correlação entre o crescimento do PIB real e o do fluxo anual de turistas estrangeiros é de apenas 4,7 %, para o período 1983-2004. Ou seja, as fases do ciclo econômico do Brasil não influenciam o ingresso de turistas. Seria esperado que as condições econômicas internas, com seus reflexos no desemprego, na violência etc atuassem como um dos fatores nas decisões dos visitantes de outros países. Assim, fases de expansão econômica ou de contração não parecem ser fatores importantes de atração ou de rejeição de turistas. Se alguma inferência estatística pode ser feita, cabe notar que a correlação é mais elevada (24,3 %) com a variação do fluxo de turistas antecedendo o crescimento do PIB em um ano. Esta relação não implica em qualquer efeito multiplicador ou de causalidade do turismo para atividade econômica, mas é interessante assinalar que existem evidências de outros países de que a entrada de turistas antecede as flutuações do PIB.² De qualquer forma, a correlação com dados anuais não é significativamente diferente de zero a 5 %.

O ingresso de turistas deve ser explicado por outros fatores, que podem ser divididos em fatores econômicos domésticos, como a taxa de câmbio e os custos de hotelaria e de alimentação; fatores econômicos externos, como os custos do turismo em outros países; fatores institucionais, como a violência e a qualidade de serviços; e fatores naturais, como a existência de atrações turísticas, recursos naturais etc. A ordenação da importância de cada fator deve ser objeto de pesquisas específicas.

² Ver por exemplo, National Statistical Coordination Board, Technical notes on the computation of the composite leading economic indicator, The Philippines Leading Economic Indicators System –LEIS, 2º trimestre de 2006

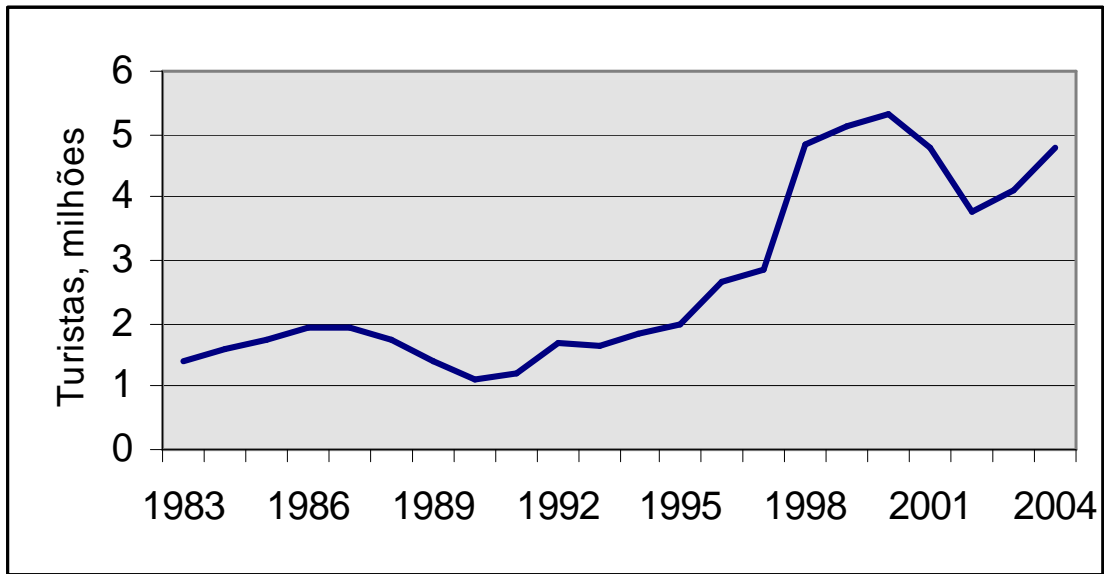


Figura 2 – Fluxo de turistas, em milhões de pessoas

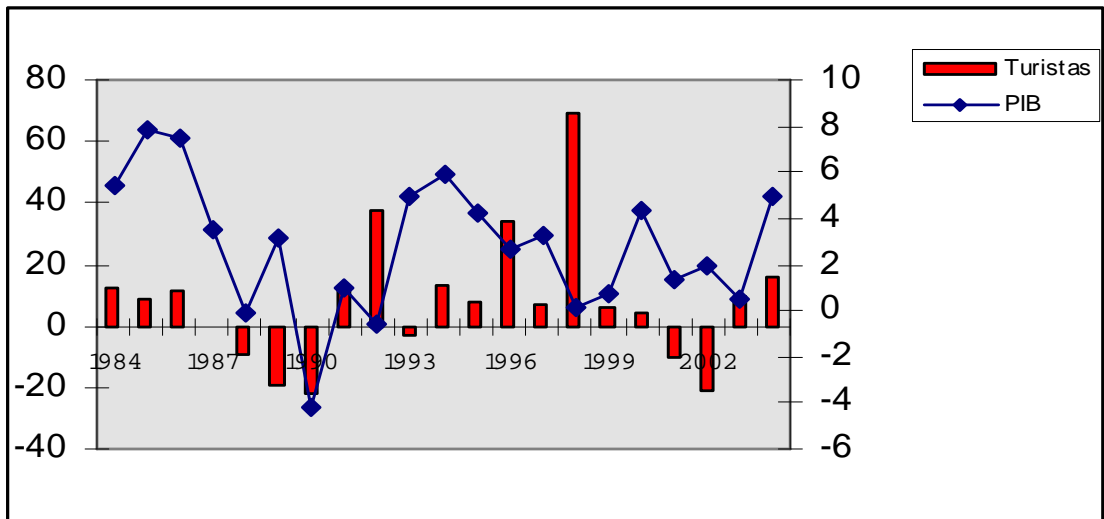
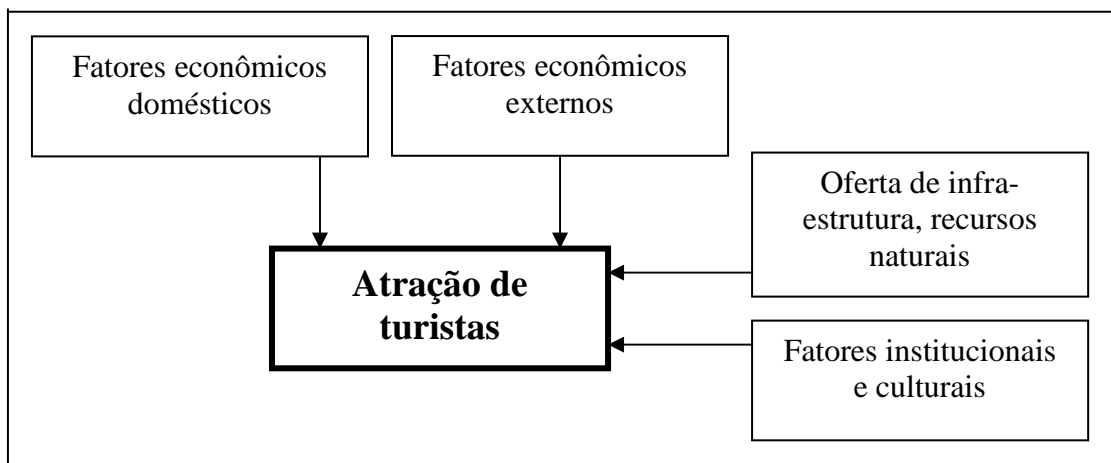


Figura 3 – Crescimento do PIB e do fluxo anual de turistas



Quadro 1 – Forças de atração para o turismo

Sendo o transporte aéreo a principal via de acesso do turismo no Brasil, o volume de passageiros transportados em vôos internacionais é uma boa *proxy* para representar as condições do turismo. Como o objetivo da pesquisa é o desenvolvimento de indicadores antecedentes de curto prazo para o turismo, a montagem empírica do sistema de previsão cíclica é favorecida pela existência de informações mensais do fluxo de passageiros/km.

2 – As flutuações no fluxo de passageiros

As estatísticas de passageiros/km são coletadas pelo DAC e EMBRATUR, e estão disponíveis a nível mensal a partir de 1974, para os vôos domésticos, e a partir de 1990, para os vôos internacionais. São estas as estatísticas utilizadas como *proxy* para as condições do turismo no Brasil. É impossível separar no fluxo de passageiros, as viagens de negócio, de turismo, e outros motivos, principalmente nos vôos domésticos, que tem o predomínio de viagens não-turísticas. Por outro lado, existe o predomínio do motivo turismo nas viagens internacionais, mas também não é possível separar o ingresso de turistas estrangeiros da saída de brasileiros para o turismo no exterior. Assim, as flutuações no fluxo de passageiros sofrem a influência de fatores domésticos, como a atividade econômica que afeta diretamente o fluxo de passageiros em viagens de negócio, e indiretamente a geração de renda (que por sua vez, influencia a demanda por turismo), e de fatores externos.

A Figura 4 reproduz a evolução do fluxo mensal de passageiros/km, e estas estatísticas serão utilizadas na montagem dos indicadores antecedentes. As séries mensais têm forte sazonalidade, concentrada nos meses outubro-janeiro no caso de passageiros em vôos domésticos, e em dezembro-janeiro, para os passageiros de vôos internacionais.

Num exercício simples, o crescimento do fluxo anual de passageiros domésticos sofre a influência da variação do PIB do Brasil e dos EUA (representando o crescimento econômico do resto do mundo), e da taxa real de câmbio R\$/US\$. Apesar da forte correlação serial nos resíduos, as elasticidades-renda assumiram valores iguais a 1,98 e 1,13, respectivamente, e a da taxa de câmbio, o valor de 0,1. O coeficiente de correlação R^2 ajustado foi de 0,81. A regressão poderia ser melhorada, e as evidências recomendam que o sistema de previsão incorpore variáveis relativas aos demais países.

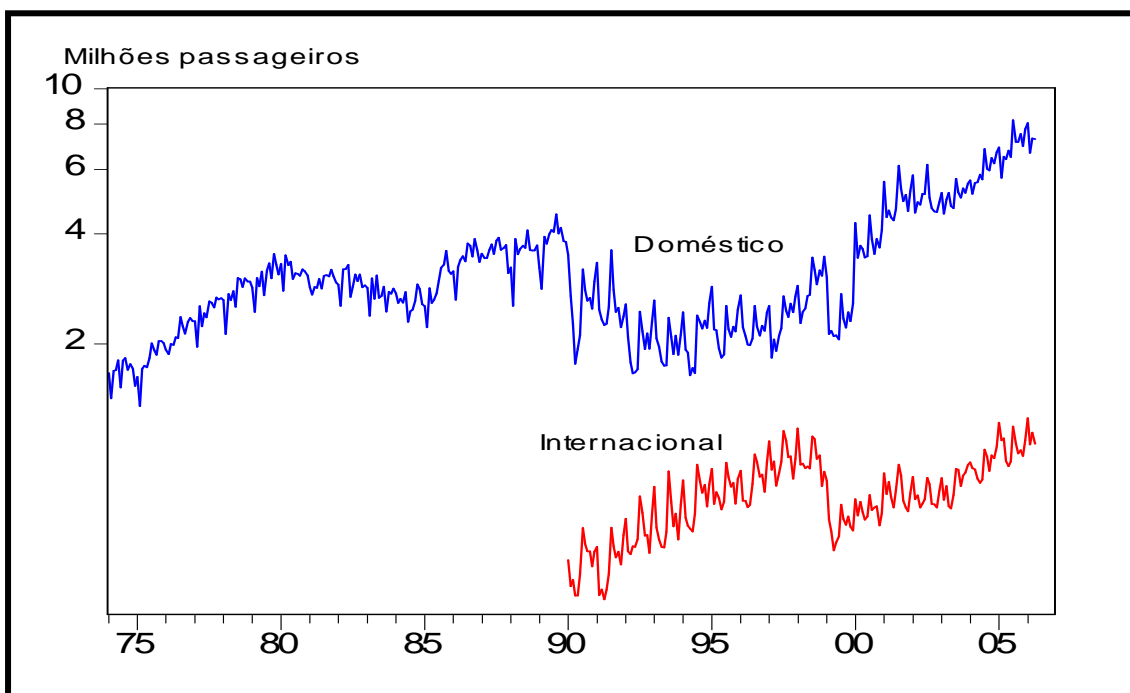


Figura 4 – Fluxo de passageiros transportados/km, em logs

Tabela 1 – Coeficientes de sazonalidade, fluxo de passageiros/km
Cálculo pela relação com a média-móvel

Meses	Doméstico	Internacional	Total
Janeiro	1,081	1,152	1,178
Fevereiro	0,892	0,958	0,939
Março	0,982	0,970	0,954
Abril	0,936	0,903	0,908
Mai	0,964	0,896	0,910
Junho	0,937	0,949	0,926
Julho	1,116	1,183	1,193
Agosto	1,051	1,085	1,047
Setembro	0,986	0,987	0,971
Outubro	1,048	1,010	1,020
Novembro	0,998	0,919	0,946
Dezembro	1,033	1,034	1,058

A decomposição por análise espectral das séries dessazonalizadas mostra que 78 % da variância total do fluxo de passageiros em vôo doméstico é explicado por ciclos maiores que cinco anos, enquanto para os de vôos internacionais, os ciclos desta frequência explicam 47 %, com importante participação das flutuações com menos de 12 meses.

Tabela 2 – Decomposição cíclica do fluxo de passageiros, livre de sazonalidade, em percentual da variância total, %

Ciclos com :	Doméstico ^a	Internacional ^b	Total ^b
Mais de 5 anos	78,4	47,3	49,0
De 3 a 5 anos	8,0	14,8	13,1
De 1 a 3 anos	6,3	10,2	11,2
Menos de 12 meses	7,4	27,7	26,7
Total	100,0	100,0	100,0

^a Período 1970-2006. ^b Período 1990-2006.

Mesmo com a eliminação da sazonalidade, as séries permanecem com flutuações menores que 12 meses, sem importância para os indicadores antecedentes. A Figura 5 compara a série de passageiros domésticos com e sem sazonalidade. Em seguida, a Figura 6 mostra os resultados de um filtro de média-móvel de 12 meses, com o qual as flutuações sazonais e as menores que 12 meses são eliminadas.

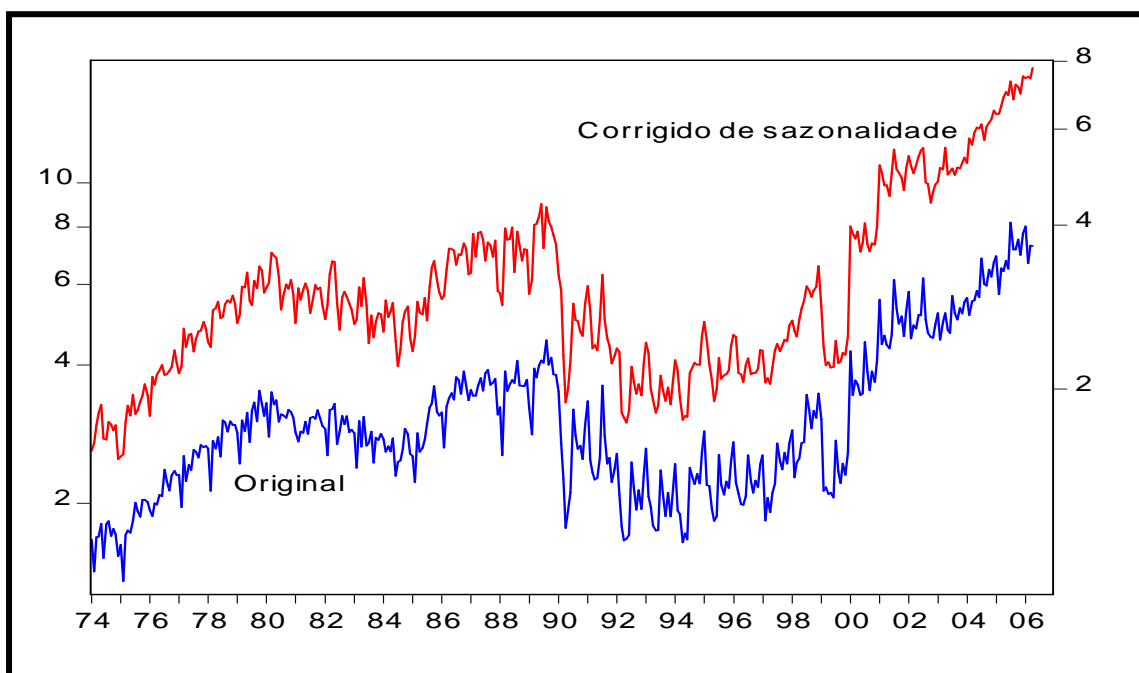


Figura 5 – Fluxo de passageiros domésticos, com e sem sazonalidade.

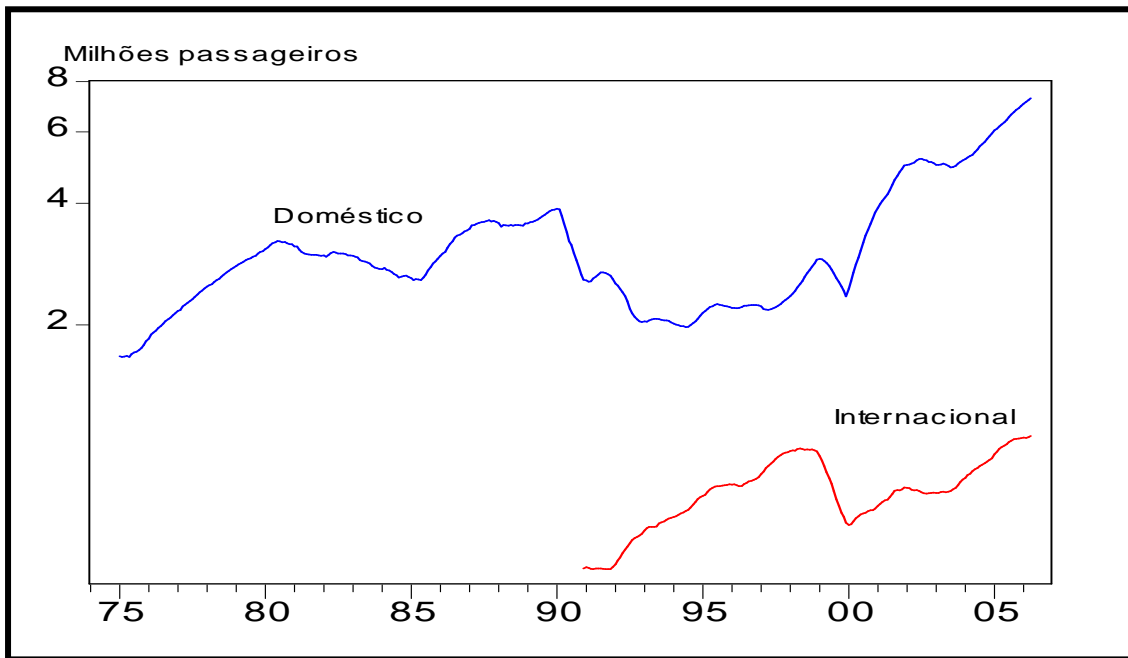


Figura 6 – Fluxo de passageiros/km, média-móvel de 12 meses.

Finalmente, as flutuações cíclicas ficam mais visíveis na Figura 7 com a taxa de crescimento em 12 meses das séries aplainadas com o filtro de média-móvel. A Figura 8 mostra as mesmas séries com as áreas sombreadas identificando as fases de desaceleração do PIB. Como esperado, as flutuações dos fluxos de passageiros em vôos domésticos estão mais associadas com as fases cíclicas do PIB do que os fluxos de passageiros internacionais.

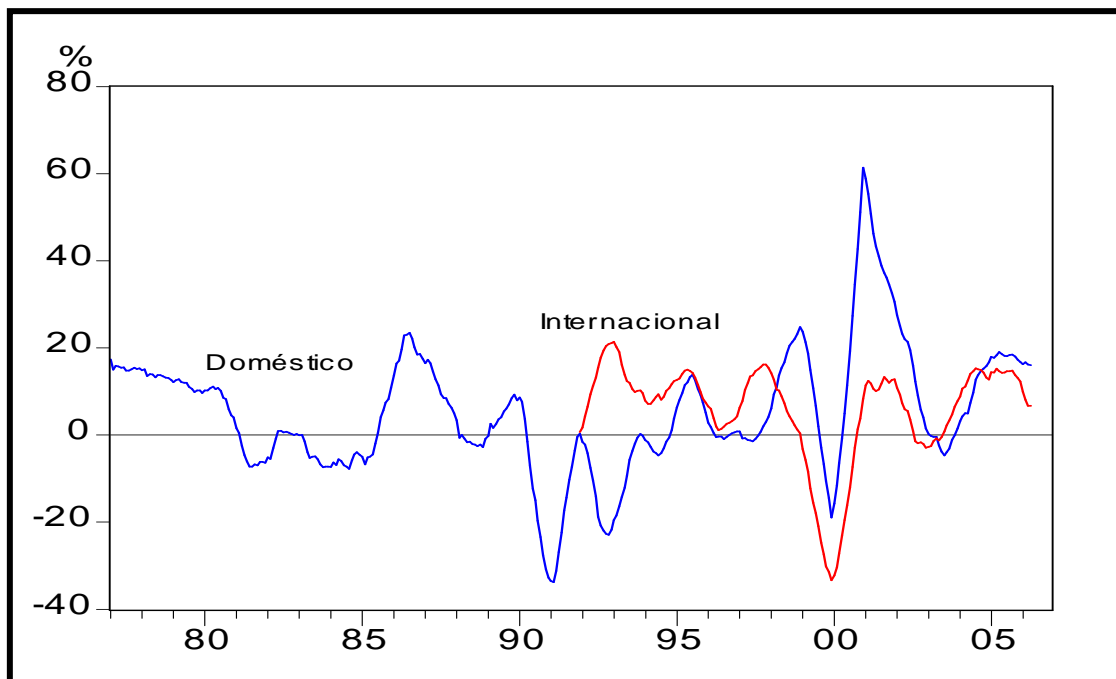


Figura 7 – Taxas de crescimento em 12 meses, fluxo de passageiros.

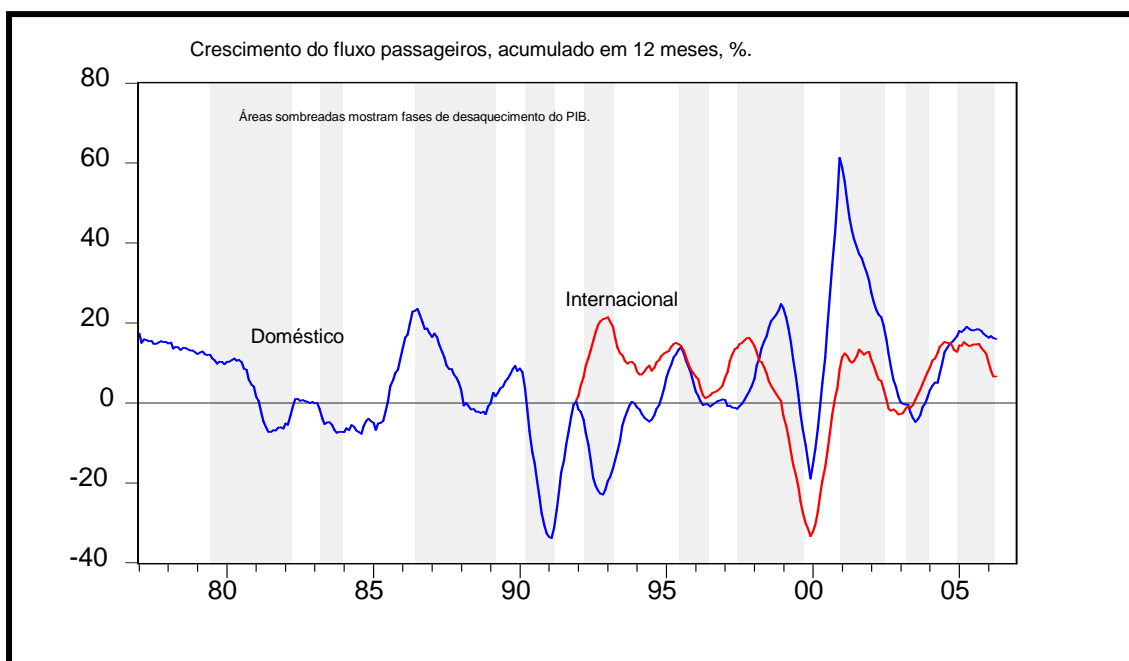


Figura 8 – Taxas de crescimento em 12 meses, fluxo de passageiros. Áreas sombreadas mostram as fases de desaceleração do PIB

3 – A metodologia

A análise e a previsão do ambiente futuro atendem a duas atividades básicas : (a) ao planejamento estratégico, geralmente voltado para o médio e longo prazo, e (b) a tática operacional, para o curto prazo. No planejamento estratégico, a preocupação é como escolher e adotar linhas mestras de ação para moldar o futuro. Na tática operacional, o futuro já é determinado pelas condições passadas e presentes, e o interesse é voltado para a solução de problemas e de obstáculos que podem causar desvios do curso traçado no planejamento estratégico.

As duas metodologias mais usadas na previsão de ciclos econômicos são os modelos estruturais e a técnica dos indicadores antecedentes. Os modelos estruturais são mais robustos teoricamente, mas não são adequados para a previsão da cronologia das reversões cíclicas, principalmente durante períodos de mudanças estruturais. Os indicadores antecedentes, por outro lado, são teoricamente menos formalizados, mas tem a vantagem da flexibilidade.

Os indicadores antecedentes aceitam o conteúdo informacional contido num grupo de variáveis, sem se preocupar com a formalização de um "modelo" propriamente dito. A exigência é de que a informação contida nas variáveis deve refletir uma antecipação do futuro. A agregação de inúmeras variáveis com tal característica permite que sejam montados índices antecedentes compostos - IAC. A grande vantagem da técnica de indicadores antecedentes, além de sua simplicidade e a mensagem intuitiva, sobre as metodologias

alternativas, é que prescinde do conhecimento prévio sobre o sentido da causalidade entre variáveis, o que não ocorre com os modelos econométricos. O objetivo último da técnica dos indicadores antecedentes não é desvendar e eliminar a incerteza ou identificar os fatores que determinam o comportamento dos ciclos econômicos, mas sim aceitar o conteúdo de informação contido num conjunto de variáveis. Foge da discussão acadêmica interminável sobre o "melhor" modelo e parte direto para o "teste do melhor pudim" de Friedman.

A técnica dos indicadores antecedentes compostos surgiu em 1919 nos Estados Unidos, com os esforços pioneiros de Burns e Mitchell³ no National Bureau of Economic Research (NBER). Atualmente os indicadores antecedentes são divulgados pelo Departamento de Comércio.⁴ Na Europa, a OECD desenvolve sistemas similares para os seus países membros.⁵

No Brasil até o final da década de 60, pouca atenção era dada à previsão de ciclos econômicos, ou mesmo da evolução da conjuntura.⁶ Como um marco pioneiro, desde 1968, a Fundação Getúlio Vargas vem realizando pesquisas junto a empresas, e os resultados de agregação das respostas são conhecidos como "Sondagens Conjunturais". Por outro lado, diversos economistas e estatísticos vem apresentando previsões com base na técnica de indicadores compostos⁷. É importante salientar que a informação fornecida pela técnica de

³ Burns, A F. & W.C. Mitchell, "Statistical indicators of cyclical revivals", NBER Bulletin, no. 69, Cambridge, Ma., 1938; e "Measuring business cycles", Studies in Business Cycles, no.2, Cambridge, Ma., NBER, 1946.

⁴ Uma descrição bem humorada dos primórdios dos indicadores antecedentes nos EUA é encontrada em Lempert, Leonard H., "Leading indicator sour grapes", Business Economics, vol.14, no.1, janeiro 1979, pp.83-86. O "estado de arte" até os anos 80 está em Ratti, Ronald A., "A descriptive analysis of economic indicators", Federal Reserve Bank of St. Louis, Review, vol.67, no.1, janeiro de 1985, pp.14-24.

⁵ OECD, "OECD leading indicators and business cycles in member countries 1960-1985", Sources and methods, vol.39, 1987, Paris, França. Outras descrições estão disponíveis em Ebanks, Walter, "The growth cycle in the industrialized world", Business Economics, vol.14, no.1, janeiro de 1979, pp.67-71; Berk, J.M. & J.A. Bikker, "International interdependence of business cycles in the manufacturing industry: the use of the leading indicators for forecasting and analysis", Journal of Forecasting, vol.14, 1995, pp.1-23; e Klein, Phillip A., "Analyzing growth cycles and leading indicators in Pacific Basin countries", Columbia Journal of World Business, vol.18, no.3, outono de 1983, pp.3-15.

⁶ Uma honrosa exceção é encontrada num artigo publicado pela Fundação Getúlio Vargas em 1948, "A conjuntura no Brasil desde 1822", Conjuntura Econômica, abril de 1948, pp.19-27.

⁷ Contador, C.R., Ciclos Econômicos e Indicadores de Atividade no Brasil, (Rio, IPEA, 1977); e outros textos do mesmo autor: "Leading Indicators for the Industrial Sector", Brazilian Economic Review, no.5, 1979, pp.1-32; "O desempenho dos indicadores antecedentes na cronologia das reversões", Relatório de Pesquisa, no.99, COPPEAD/UFRJ, agosto de 1990. Outros trabalhos importantes são o de Markwald, R.A., Ajax R.B. Moreira e Pedro L. Valls Pereira, "Previsão da Produção Industrial : Indicadores Antecedentes e Modelos de Série Temporal", Pesquisa e Planejamento Econômico, vol.19, no.2, agosto de 1989, pp.233-254; Melo Souza, Mary & Moyses

indicadores antecedentes não se preocupa primordialmente em prever o nível das variáveis ou mesmo a sua taxa de crescimento. Para a previsão do nível absoluto, a técnica é menos acurada, embora possa fornecer estimativas com dimensão próxima a da variável sendo prevista. Se o objetivo é prever a cronologia das reversões cíclicas, e para esta finalidade, o desempenho da técnica é considerado excepcional. De importante para os objetivos maiores da pesquisa, a previsão das flutuações cíclicas e das suas intensidades deve permitir identificar os períodos de maiores e menores pressões sobre a capacidade instalada do complexo do turismo.

A análise econômica de um mercado pressupõe que os seus ciclos são formados, em parte, pela influência das flutuações macroeconômicas - que, por princípio, afetam todos os setores - e, em parte, pela influência de fatores próprios da atividade do setor. Se os fatores macroeconômicos são predominantes, os ciclos do setor em estudo têm fases similares aos dos ciclos agregados, diferindo pelo período de defasagem e pela amplitude das flutuações.⁸ Se os fatores específicos predominam, os ciclos do setor tendem a independer dos ciclos macroeconômicos.

Os ciclos têm diferentes causas, efeitos, duração e amplitude. Em termos de formato, os ciclos são convencionalmente identificados pelas letras V, U e L, no caso de fases em torno de vales, ou de V, U e L invertidos, no caso de picos. Os formatos mudam com o conceito do ciclo. Um ciclo em V significa geralmente fases rápidas e agudas, tanto das fases de contração

Tenenblat, "Indicadores antecedentes para as exportações e importações totais brasileiras", Sociedade Brasileira de Econometria, Anais, III Encontro Brasileira de Econometria, Curitiba, dezembro de 1991, pp.551-558; Oliveira, A.X. & F.A. Pino, "Indicador antecedente para a industria de transformação : uma proposta alternativa", III Escola de Séries Temporais e Econometria, EPGE/FGV, Rio de Janeiro, julho de 1989, e mais recentemente Chauvet, Marcelle, "Turning point analysis of the leading inflation indicators – activity report", Banco Central, abril-maio de 2000; Chauvet, Marcelle, "Leading inflation indicators for Brazil: Final Report", Banco Central, dezembro de 1999; Chauvet, Marcelle; Solange Gouvea; Marta Baltar Moreira e Jose Ricardo da Costa e Silva, "Two methodologies to build inflation leading indicators for Brazil", Seminario One Year of Inflation Targeting in Brazil, Banco Central, 10-11 julho 2000; Contador, C.R., "Economic activity in 2001 :what the leading indicators forecast", Seminario DIMAC no.44, março de 2001; Contador, C.R., "Inflation targeting and leading indicators : some notes", Seminário One year of Inflation Targeting", Banco Central, 10-11 julho 2000; Contador, C.R., "Economic activity in 2001: what the leading indicators forecast", Seminario sobre Indicadores Antecedentes, IPEA/CEPAL/OECD, Rio, 4-5 dez.2000; Gouvea, Solange; Marta Baltar e Pedro Albuquerque, "Indicadores antecedentes para a inflacao", Seminario sobre Indicadores Antecedentes, IPEA/CEPAL/OECD, Rio, 4-5 dez.2000;

⁸ Assim, setores com forte elasticidade ao ciclo macroeconômico apresentam ciclos com flutuações mais intensas, e setores com elasticidade modesta (menor que um), menores flutuações. A elasticidade-renda do fluxo de passageiros em voos domésticos é 2,1 no Brasil.

como de expansão. O formato em V é encontrado com mais frequência nos ciclos de curta duração. O exemplo no Brasil é encontrado em 1981-82, para o crescimento do PIB real, e em 1980 nos EUA. O ciclo em U tem fases de contração e expansão mais longas e suaves, como no Brasil em 1997-2001 e nos EUA em 1981-83 nos EUA. E finalmente, o ciclo em L é caracterizado por longas durações, como no Brasil em 1963-68, e no Japão, em 1990-2001.

A tipologia dos ciclos tem implicações importantes tanto para a condução da política macroeconômica, como para o planejamento e operações de instituições e empresas. Por exemplo, considerando os efeitos retardados das medidas de política econômica, não tem sentido insistir na sintonia fina (alias, em nenhum caso também!) perante os ciclos em V. A duração do ciclo é curta, e as medidas anticíclicas podem na verdade intensificar, ao invés de amortecer, a amplitude das flutuações. Numa empresa, um ciclo em V que ocorre num mercado, pega de surpresa a administração de estoques, mas não derruba o desempenho da organização no médio e longo prazo.

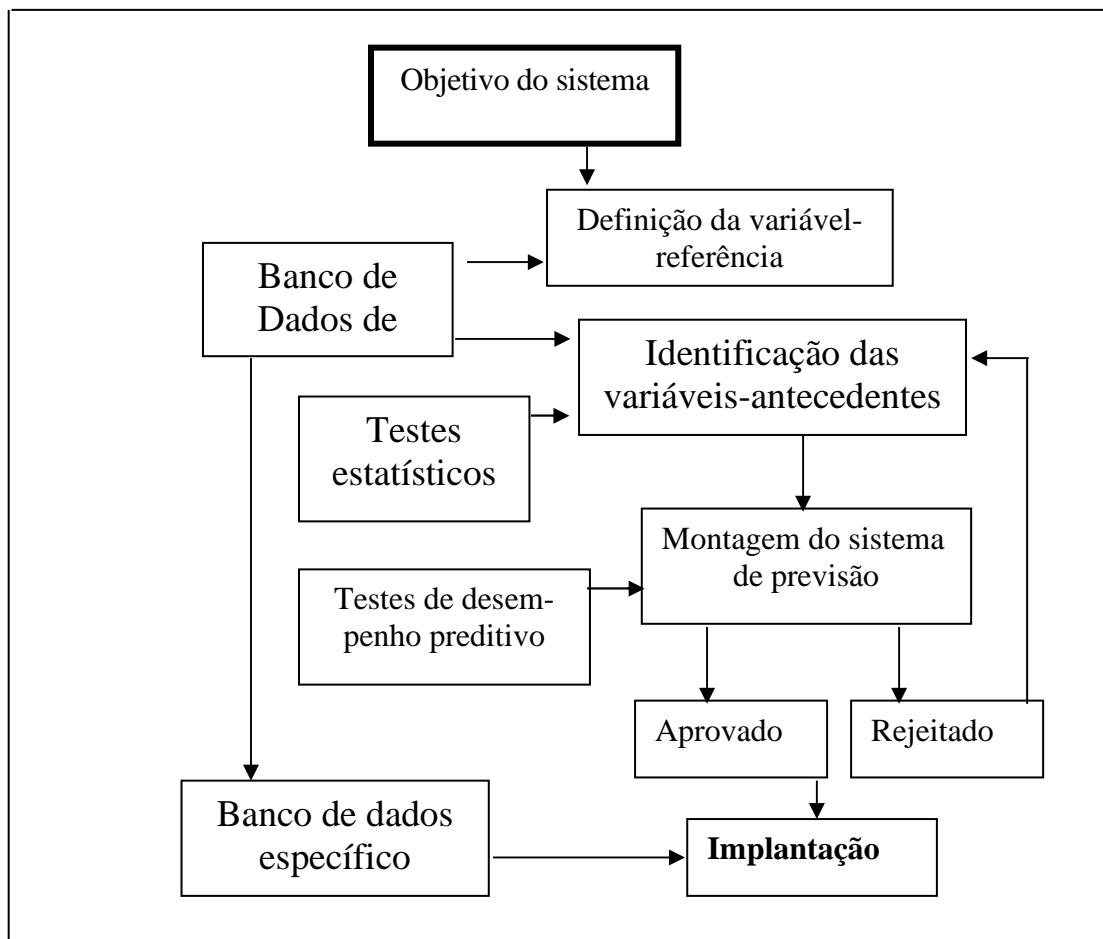
No outro extremo, os ciclos em L exigem medidas mais drásticas. Para um país, região, ou setor, uma longa fase de estagnação pode significar a necessidade de uma profunda reorganização da estrutura produtiva, reforma fiscal, investimentos maciços em setores críticos, legislação mais moderna, menor regulação etc. Para uma empresa, um ciclo em L pode ser um indício de que o seu mercado já atingiu a fase de declínio na conceituação do modelo BCG - *Boston Consulting Group* e não há muito que fazer, salvo a mudança de ramo de atividade, buscar a diversificação de negócios e de mercados regionais, ou mesmo se desfazer da empresa e mudar de emprego.

Naturalmente, para que a previsão seja viável na prática, é crucial manter operacional e simples o tratamento estatístico. Na impossibilidade de considerar cada ciclo como específico e diferente de qualquer outro ocorrido no passado ou a ocorrer no futuro – o que tornaria irrelevante a idéia de previsão - vamos aproveitar o fato de que os ciclos compartilham de algumas características comuns, independentes das suas causas, duração e amplitude.

Para prever um fenômeno é necessário que ele possa ser representado por conceitos empíricos. Se o objetivo é prever o longo prazo de um fenômeno, uma análise de tendência é suficiente na maioria das vezes. Se, por um lado, este tipo de previsão não causa maiores problemas, por outro, tem o inconveniente de fornecer informações singelas e geralmente erradas, quando ocorre uma quebra no estado do fenômeno, ou seja, uma mudança cíclica. Mas qual ciclo causa problemas? Certamente, não os ciclos sazonais, nem as oscilações do curtíssimo prazo. Em cada caso existe um ciclo mais importante ou mais relevante que os demais. As dificuldades aparecem na identificação do que seja o "ciclo relevante". Como ciclo relevante, a análise utilizará o conceito de taxa de crescimento em 12 meses do fluxo acumulado exposto nas Figuras 7 e 8.

4 – Etapas da montagem do sistema de indicadores antecedentes

Os conceitos básicos do sistema de indicadores antecedentes e o fluxo de informação entre as etapas estão reproduzidos no esquema a seguir, com detalhes que serão explicados ao longo do texto.



Quadro 2 – Fluxo de informação de sistemas de indicadores antecedentes

Definido o objetivo do sistema (as variáveis-referência ou objeto) e as características desejadas (horizonte da previsão e o período do ciclo), os ciclos relevantes são ampliados por processos estatísticos. Em seguida, procedimentos semelhantes são realizados no Banco de Dados, contendo as variáveis que compõem um subconjunto para a agregação num índice composto. Uma bateria de testes estatísticos de desempenho examina a validade preditiva do indicador. Se o indicador antecedente composto passa nos testes de desempenho fora do período utilizado na estimação dos seus parâmetros, a sua composição é utilizada na montagem do sistema de previsão, com um banco de dados específico. Se for rejeitado, reinicia-se o processo de escolha das variáveis-insumo é reiniciado, submetido a novos testes de desempenho e assim por diante.

5 - Definição do horizonte e o objetivo da previsão

De um modo geral, a natureza dos objetivos estabelece as duas características básicas para o sistema de previsão; as dimensões do tempo e o fenômeno a ser previsto. Qualquer sistema de previsão envolve três dimensões temporais : (a) o período ou intervalo de previsão; (b) a antecedência ou horizonte "a frente" da previsão; e (c) a frequência das previsões e das revisões. "Período" corresponde à unidade de tempo para a qual as previsões são feitas, como, por exemplo, o mês ou o trimestre. Em geral, a escolha da unidade de tempo envolve compromissos, de um lado, entre a desagregação no tempo e, de outro, a disponibilidade de informações, o custo de montagem do sistema, a distribuição dos erros do sistema e o risco/custo das previsões. Em geral, quanto menor a unidade de tempo, mais difícil se torna a coleta de dados estatísticos e maior o risco das previsões devido ao aumento das oscilações erráticas. Assim, previsões do fluxo mensal de vendas estão mais sujeitas a erros do que previsões para um período mais longo, como, por exemplo, um ano.

A "antecedência da previsão" ou "períodos a frente" mostra o horizonte ou número de períodos no futuro cobertos pela previsão. Quanto maior o horizonte, maior o risco de erros nas previsões. Por exemplo, a previsão da taxa de crescimento do fluxo de passageiros, realizada com três trimestres de antecedência, envolve um risco muito maior do que quando realizada com antecedência de apenas um trimestre. É claro que, a cabo de três trimestres, dispomos de muito mais informações e conhecimento sobre o crescimento provável no quarto trimestre do que há dois trimestres atrás.

Por outro lado, quanto maior o horizonte da previsão, menor o papel das flutuações de curto prazo, e maior a probabilidade de acerto da previsão, desde que bem fundamentada. Por isso, o horizonte da previsão não deve ser menor do que o tempo necessário para a implantação e os primeiros efeitos das medidas. Novamente, há um conflito entre a magnitude provável do erro e o horizonte da previsão.

Finalmente, a frequência com que as previsões e as revisões são preparadas confunde-se muitas vezes com a própria unidade de tempo. Novas previsões são feitas, incorporando as informações mais recentes. Quando possível, as previsões devem ser revistas com frequência, fazendo-se até várias revisões por período. É de se esperar que cada nova revisão implique uma redução no erro da previsão. Contudo, nem sempre isto ocorre.

Ao determinar o que deve ser previsto - por exemplo, o fluxo de passageiros de transporte aéreo – já definimos parte o campo da pesquisa. Porém, apenas restringir o campo não garante que a solução seja viável operacionalmente. O número de variáveis que afetam ou antecipam o comportamento das vendas pode ser imenso. O detalhamento excessivo, longe de simplificar, pode complicar o sistema de previsão, e a experiência mostra

que o sacrifício de algumas variáveis pouco ou nada afeta a qualidade da previsão. Este é o fundamento do “princípio da parcimônia”. O grau de detalhe exigido no sistema de previsão é afetado por muitos fatores : a disponibilidade e qualidade das informações estatísticas; o custo de análise, considerações teóricas, etc. Dentre os citados, os primeiros são fatores limitantes em qualquer sistema de previsão.

Como regra básica, só é possível fornecer previsões quantitativas para fenômenos também expressos de forma quantitativa. Variáveis qualitativas - como qualidade da vida; bem-estar; estabilidade de regras e de contratos etc. - são importantes, mas por não serem representáveis quantitativamente não podem ser objeto de previsão com esta metodologia. Da mesma forma, mudanças qualitativas têm efeitos nos fenômeno que se deseja prever, e, se não quantificadas, não podem ser incorporadas diretamente no sistema de previsão. Exemplos destas mudanças qualitativas são os choques heterodoxos que marcaram a economia brasileira na década de 80 e início dos 90..

Algumas vezes, a variável que se pretende prever ou acompanhar é quantitativa, mas simplesmente não é coletada a intervalos menores que um ano. Isto ocorre, por exemplo, com as informações típicas de censos. Outras vezes, a informação disponível é de má qualidade. A existência e a importância desses problemas forçam o investigador a armar-se com doses extras de imaginação e de conhecimento mais detalhado do fenômeno em estudo e da economia brasileira, em geral. Com frequência ele é forçado a empregar variáveis *proxy* para substituir a variável-referência, ou simplesmente assumir que as estatísticas disponíveis retratam satisfatoriamente o fenômeno a ser previsto.

6 - Selecionando os ciclos

Sabemos que qualquer série histórica é formada pela superposição de um grande número de processos, tais como tendência, vários componentes cíclicos - com durações diversas inclusive sazonalidade - e por resíduos puramente aleatórios. A tendência histórica do fenômeno em estudo é, na maioria das vezes, pouco importante. As oscilações aleatórias, no outro extremo, são impossíveis de serem previstas (por definição), e normalmente eliminadas no processo de agregação. Mesmo dentre os componentes cíclicos existem alguns que, dada a sua regularidade, podem ser ignorados, como os ciclos sazonais com duração de doze meses. O interesse central do sistema de previsão está direcionado para os demais ciclos.

A forma mais simples de eliminar a tendência e a sazonalidade, e ainda amortecer as oscilações aleatórias de variáveis é, no caso de uma série que retrate um fluxo, primeiro, "aplinar" os seus valores através da média-móvel em doze termos. Em seguida, extrai-se a taxa de crescimento em doze meses da série acumulada. Se a variável corresponde a um estoque, geralmente basta

calcular a sua taxa de crescimento em doze meses. Este é o procedimento adotado nas variáveis usadas neste estudo.

Infelizmente, não existe uma regra rígida que forneça um filtro único e geral para todas as variáveis envolvidas e cada série deveria receber um tratamento específico. Como este procedimento torna-se oneroso, existe uma sugestão ditada pelo bom senso e pela economia de recursos e de tempo, de que seja utilizado um filtro que melhor atenda a maioria das variáveis. Embora, a primeira vista esta atitude pareça grosseira, uma análise exaustiva de experimentação comprovou que um filtro bastante satisfatório para eliminar a tendência e a sazonalidade e reduzir as oscilações espúrias tem o formato,

$$z(t) = (1 - L^{12}) Z(t) \quad (1)$$

onde t é a dimensão de tempo; $Z(t)$ é uma série histórica genérica expressa em logaritmos (previamente deflacionada e acumulada, se for o caso); L , o operador-retardo definido como

$$L^i Z(t) = Z(t-i) \quad (2)$$

e $z(t)$, a série "filtrada".

O emprego deste filtro oferece várias vantagens. Suponhamos uma série qualquer $Z(t)$ formada por uma tendência, diversos ciclos, inclusive um sazonal, e variações aleatórias. Dependendo da série, alguns destes componentes são mais importantes que os outros. Por exemplo, em algumas variáveis, a tendência predomina sobre os outros componentes; em outras séries, as flutuações de curto prazo são os mais importantes, e assim por diante. Representando cada um destes processos por $Z_{\omega_i}(t)$ onde ω_i corresponde a duração do ciclo deste período, desde $i = 1$, que corresponde às flutuações aleatórias, até infinito, que corresponde a tendência (um ciclo de duração infinita), teríamos

$$Z(t) = \sum_{i=1} Z_{\omega_i}(t) \quad (3)$$

Por definição, cada um destes componentes cíclicos Z_w independe dos demais (covariância nula) e a variância da série $Z(t)$ é a soma das variâncias de cada ciclo. No caso do fluxo de passageiros de transporte aéreo doméstico foi constatado que as flutuações de médio e longo prazo, com duração acima de um ano são componentes importantes, explicando pouco mais de 90 % da variância da série, enquanto os ciclos de curto prazo, com duração até um ano, explicam uns 7 %.

O filtro descrito pelo processo (1) permite eliminar a tendência e a sazonalidade; amortecer os ciclos de curtíssimo prazo; e ampliar os de médio

prazo. Com este cuidado, as séries utilizadas nos indicadores antecedentes são compostas pelos ciclos de maior interesse para a previsão.

7 - Identificação das variáveis antecedentes

A montagem de indicadores antecedentes exige uma análise prévia dos retardos e avanços entre a variável-referência a ser prevista (fluxo de passageiros de transporte aéreo) e as demais. A partir daí, é assumida a hipótese básica de que a mesma estrutura de retardos e avanços, estimada com as informações passadas, permanece válida para o futuro próximo. Os retardos e avanços permitem classificar as variáveis em antecedentes, coincidentes e retardadas. É possível que ocorram mudanças nas estruturas de retardos e avanços entre séries, o que recomenda uma crítica periódica (digamos, a cada oito ou doze meses) da classificação das variáveis. Uma variável identificada, por exemplo, no período 1980-95 como antecedente com oito meses de avanço sobre uma outra, pode apresentar um avanço diferente no período 1995-05 ou mesmo tornar-se coincidente ou retardada (fato mais improvável).⁹

Uma simples análise visual revela informações interessantes. Por exemplo, a Figura 9 reproduz as flutuações do fluxo de passageiros em vôos domésticos e do Índice Bovespa (deflacionado pelo IGP-M). As fases cíclicas seguem em geral o mesmo sentido, com as flutuações do Bovespa antecedendo em 13 meses o fluxo de passageiros. Como a Bolsa de Valores antecipa as condições futuras, este resultado sugere que esta variável pode ser uma candidata a compor o indicador antecedente. Um outro exemplo na Figura 10, onde agora a comparação é feita com a taxa de câmbio real (deflacionada pelo IGP-M), onde os movimentos são opostos: a valorização do real (ou seja, queda no câmbio real) desestimula as viagens aéreas, e a antecedência é de oito meses. É um resultado esperado, uma vez que a desvalorização da moeda nacional eleva o custo relativo do transporte aéreo. Mais um exemplo: as flutuações na taxa de desemprego antecedem as do fluxo de passageiros de vôos domésticos¹⁰.

A forma mais rigorosa de identificar os avanços e retardos entre variáveis utiliza o correlograma entre variações passadas e futuras de duas séries devidamente filtradas. Por este critério, diz-se que uma variável-insumo X antecede a variável-meta Y, se as maiores correlações significantes são encontradas entre o valor atual de Y e os valores passados de X. As variáveis X e Y são ditas coincidentes, se a maior correlação é encontrada para valores não defasados (retardo nulo), e X é dita retardada em relação a Y, se as

⁹ A mudança do avanço entre variáveis ao longo do tempo foi examinada por Edlung, Per-Olov & Henning T. Sogaard, "Fixed versus time-varying transfer functions for modelling business cycles", *Journal of Forecasting*, vol.12, nos.3 e 4, abril de 1993, pp.345-364.

¹⁰ A correlação é de -37,3 %, significante a 5 %, e o avanço é de 11 meses.

correlações mais elevadas são encontradas entre valores passados de Y e os correntes de X.

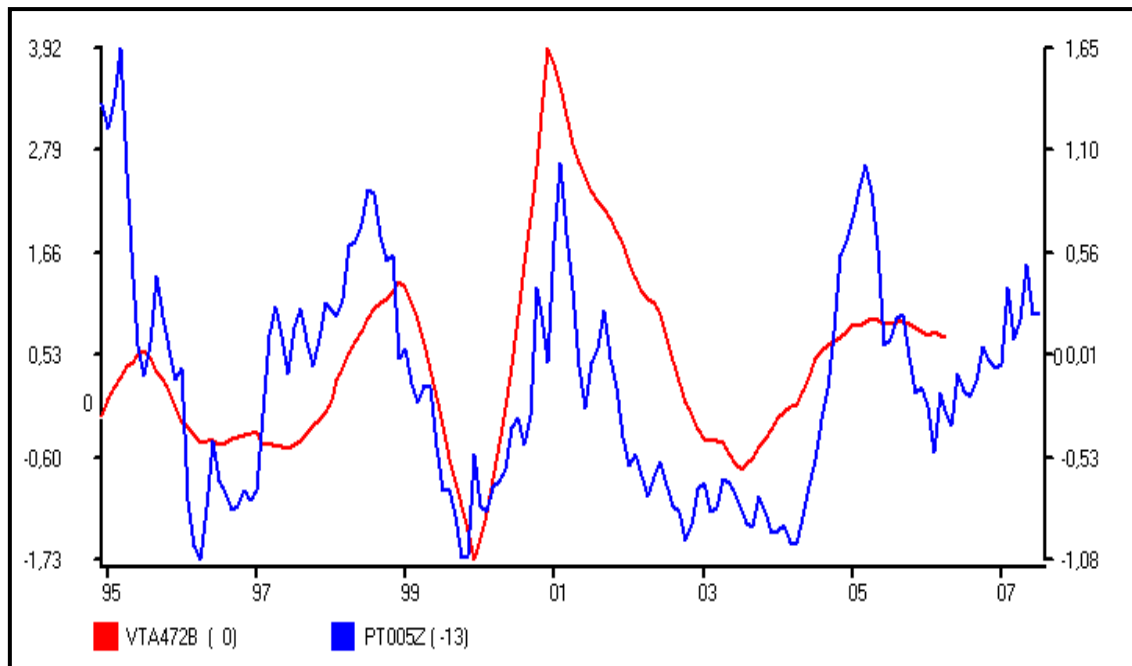


Figura 11 – Flutuações no fluxo de passageiros (linha vermelha) e Bolsa de Valores (azul)

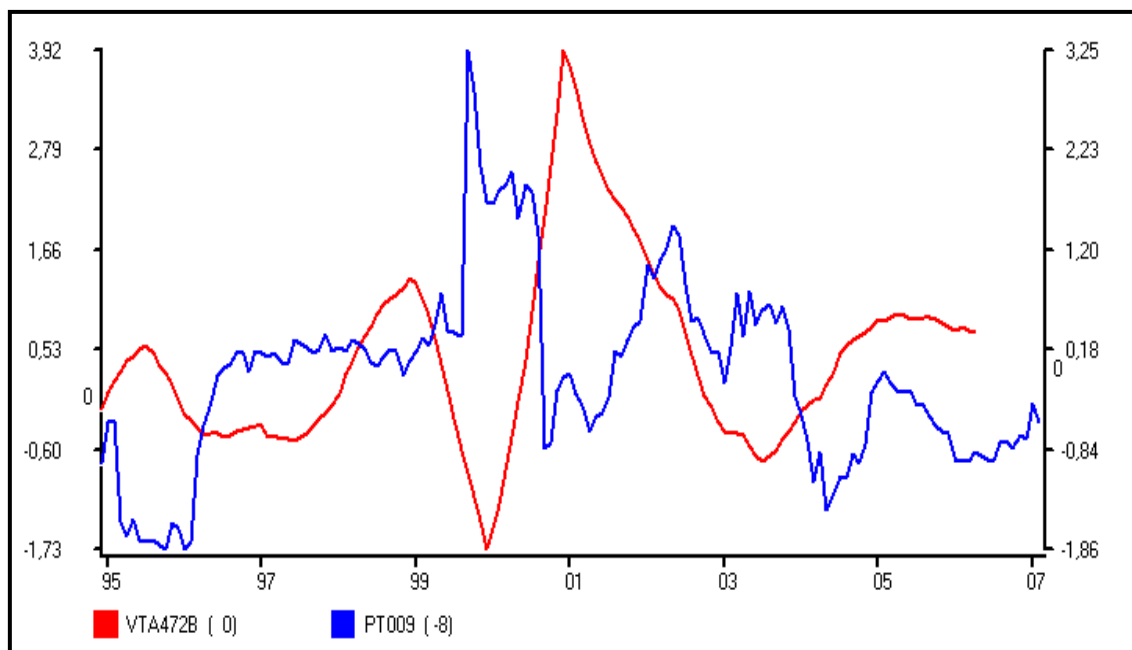


Figura 12 – Flutuações no fluxo de passageiros (linha vermelha) e taxa de câmbio azul)

A análise da antecedência entre variáveis deve ser feita com as séries expressas em três escalas diferentes:

- Original : em valores originais,
- Dicotomia de fases, em que as séries recebem valor um durante as fases de expansão, e zero, nas contrações; e
- Dicotomia das reversões, com valores dicotômicos para as datas de reversão (valor um nas datas de vale e menos um nos picos).

A qualificação de uma variável-insumo X como boa candidata a compor um indicador antecedente exige que o seu avanço em relação à série referência Y seja aproximadamente o mesmo nos três critérios de antecedência. Usualmente, a análise da antecedência aborda apenas a escala original. O resultado pode ser enganoso. Vamos usar um exemplo com séries construídas artificialmente.

Por construção, três variáveis: Y (a referência); X1 e X2 (duas variáveis-insumo) tem a mesma cronologia de reversões. Na Figura 13, o gráfico (a) reproduz as séries expressas em valor original; em (b), as fases de expansão e contração foram convertidas com filtros em valores zero e um; e em (c), as datas de picos e vales em valores -1 e 1. Nos dois gráficos (b) e (c), as variáveis X1 e X2 tem o mesmo formato, e por isso apenas uma é representada.

Não obstante a mesma cronologia cíclica – com correlação igual a 100 % com os critérios (b) e (c) – a correlação (linear) entre os valores não é tão visível. De fato, a correlação entre Y e X1 (e X2) é igual a 55,9 %. E seria possível construir um exemplo em que a correlação entre valores fosse elevada, e muito pequena (não significativa) com os outros critérios. Ou seja, os resultados da análise de antecedência podem ser enganosos (proveniente de correlações espúrias) se utilizamos apenas um dos critérios. Mas em conjunto, os três critérios fornecem resultados robustos e confiáveis.

A Figura 14 mostra como funciona a análise dos correlogramas com os três critérios. A variável-referência é o crescimento do acumulado em 12 meses do fluxo de passageiros de vôos domésticos e a variável-insumo, o crescimento do Índice Bovespa. As variáveis examinadas foram previamente "filtradas" para eliminar os componentes (tendência e sazonalidade) sem interesse. A figura está dividida em duas partes. Na esquerda, a tabela (que no sistema SIA[®] é visualizada em detalhes com a barra de rolamento) mostra a correlação e a defasagem em cada critério. Os retardos e avanços estão expressos em meses, com valores negativos e positivos, respectivamente. Na direita, a figura reproduz o correlograma com as defasagens no eixo horizontal e as correlações no vertical.

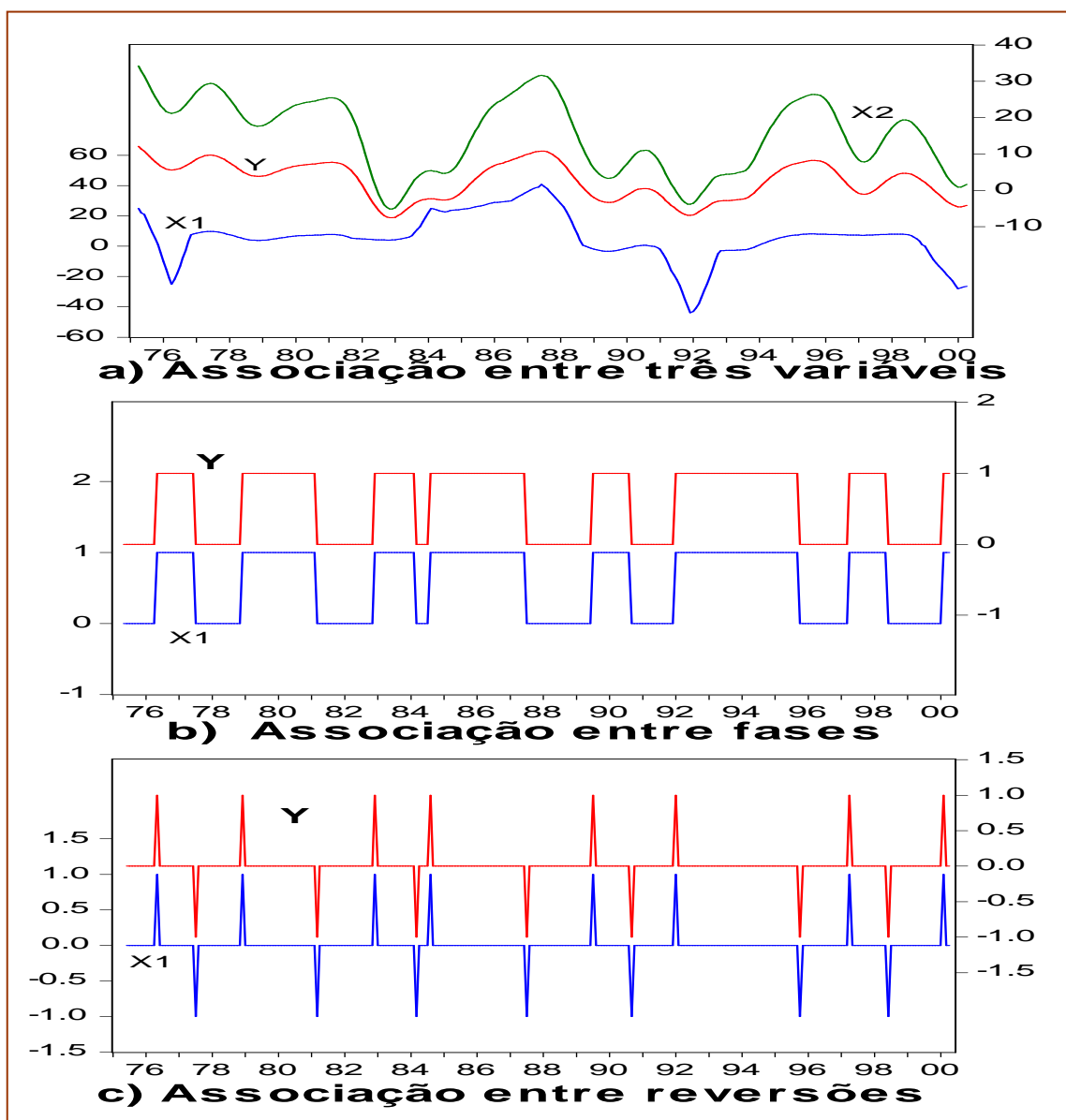


Figura 13 – Análise visual de antecedência entre variáveis

As duas linhas tracejadas horizontais assinalam os valores críticos para que a correlação (em valor absoluto) seja significativamente diferente de zero ao nível de 5%. Valores negativos (positivos) para a defasagem mostram a faixa em que a variável-insumo X antecipa (retarda-se) a variável-referência Y. Para a montagem de indicadores antecedentes é necessário – ainda que não suficiente como veremos mais adiante – que a variável-insumo X seja significativamente correlacionada com Y no segmento negativo das defasagens. A figura apresenta o exemplo interessante de uma variável que, embora não influencie (ou seja, não causa) o faturamento real da Indústria,

tem condições de antecipar os seus movimentos.¹¹ Os mercados especulativos incorporam as expectativas dos agentes sobre o futuro, e, portanto servem com candidatos a compor indicadores antecedentes. Na figura, o índice de Bolsa de Valores (IBOVESPA) antecipa o fluxo de passageiros em treze meses, com correlação positiva. A figura direita reproduz os correlogramas com os três critérios. A barra azul mostra a correlação simples entre os níveis das séries; na linha vermelha a correlação é entre as fases (0 e 1); e a verde, entre as datas de reversão (-1 e 1).

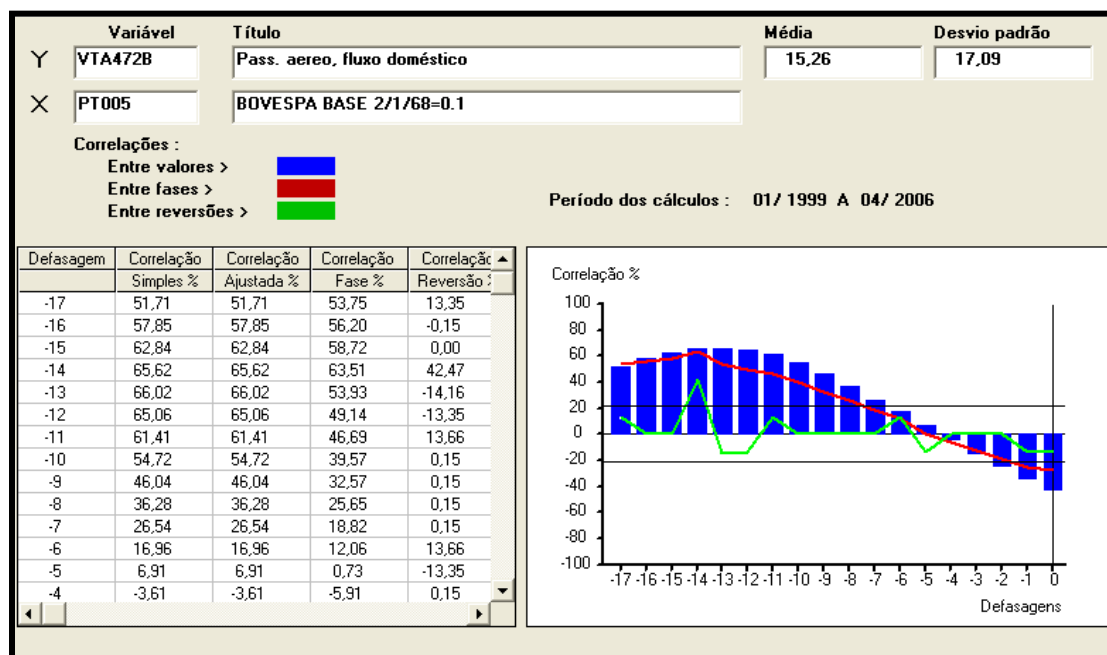


Figura 14 – Correlogramas com os três critérios

As figuras seguintes ilustram outros exemplos, onde apenas a parte do correlograma é apresentada. A Figura 15 mostra que o crédito às pessoas físicas antecede o fluxo de passageiros em aproximadamente quatro meses. É um avanço pequeno, considerando a demora na divulgação das estatísticas, mas o resultado indica que as flutuações na oferta de crédito (deflacionado pelo IGP-

¹¹ Mills, Leonard, "Can stock prices reliably predict recessions?", *Business Review*, Federal Reserve Bank of Philadelphia, setembro/outubro de 1988, pp.3-14; Peek, Joe & Eric S. Rosengren, "The stock market and economic activity", *New England Economic Review*, Federal Reserve bank of Boston, maio/junho de 1988, pp.39-50; Chauvet, Marcelle; "Stock market fluctuations and the business cycle", *Journal of Economic and Social Measurement*, a sair, 2000; Fama, Eugene F. & Kenneth R. French; "Business conditions and expected returns on stocks and bonds", *Journal of Financial Economics*, vol. 25, no.1, novembro de 1989, pp.23-49; Lee, Bong-Soo, "Causal relations among stock returns, interest rates, real activity and inflation", *Journal of Finance*, vol. 47, no.4, setembro de 1992, pp.1591-1603; Loungani, Prakash, Mark Rush & William Tave, "Stock market dispersion and business cycles", *Economic Perspective, Federal Reserve Bank of Chicago*, vol.15, no.1, jan/fev. 1991, pp.2-8; Pearce, Douglas K., "Stock prices and the economy", *Economic Review, Federal Reserve Bank of Kansas City*, novembro de 1983, pp.7-22; Saied, Mahdavi, "The link between the rate of growth of stock prices and the economy", *American Economist*, vol.35, outono de 1991, pp.41-48; Tang, Gordon Y.N.; S.C. Mak & Daniel F.S. Choi; "The causal relationship between stock index futures and cash index prices in Hong Kong", *Applied Financial Economics*, vol.2, dezembro de 1992, pp.187-190.

DI) com retardo de quatro meses e o fluxo de passageiros tem uma correlação positiva acima de 60 %. A Figura 16 reproduz o correlograma com as consultas ao SPC, que reflete o ânimo de consumo das famílias, e a antecedência é de cinco meses, com correlação acima de 70%. A Figura 17 completa os exemplos, mostrando os correlogramas entre o fluxo de passageiros e a taxa real de juros cobrada no desconto de duplicatas, onde a correlação é negativa e os juros antecipam o fluxo de passageiros em oito meses.

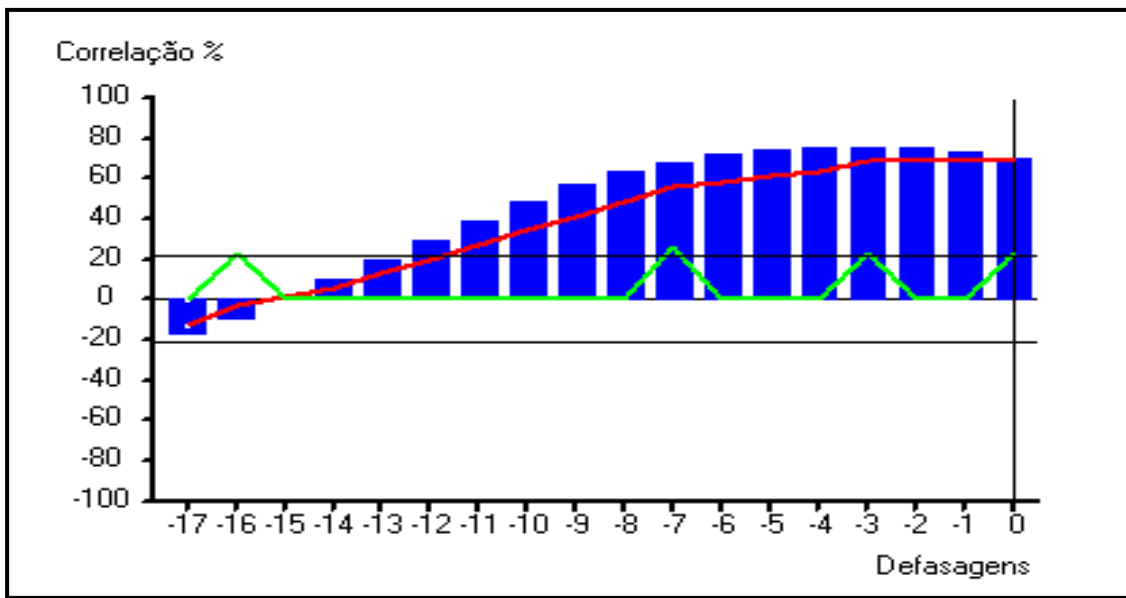


Figura 15 – Correlogramas entre fluxo de passageiros domésticos e oferta real de crédito às pessoas físicas.

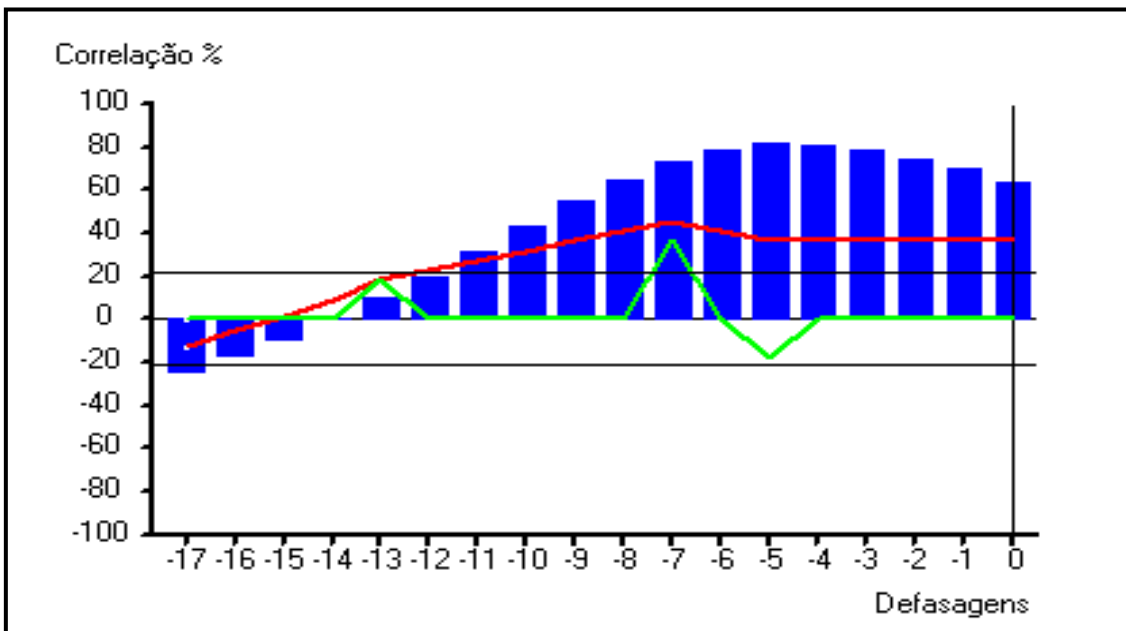


Figura 16 – Correlogramas entre fluxo de passageiros domésticos e as consultas ao SPC

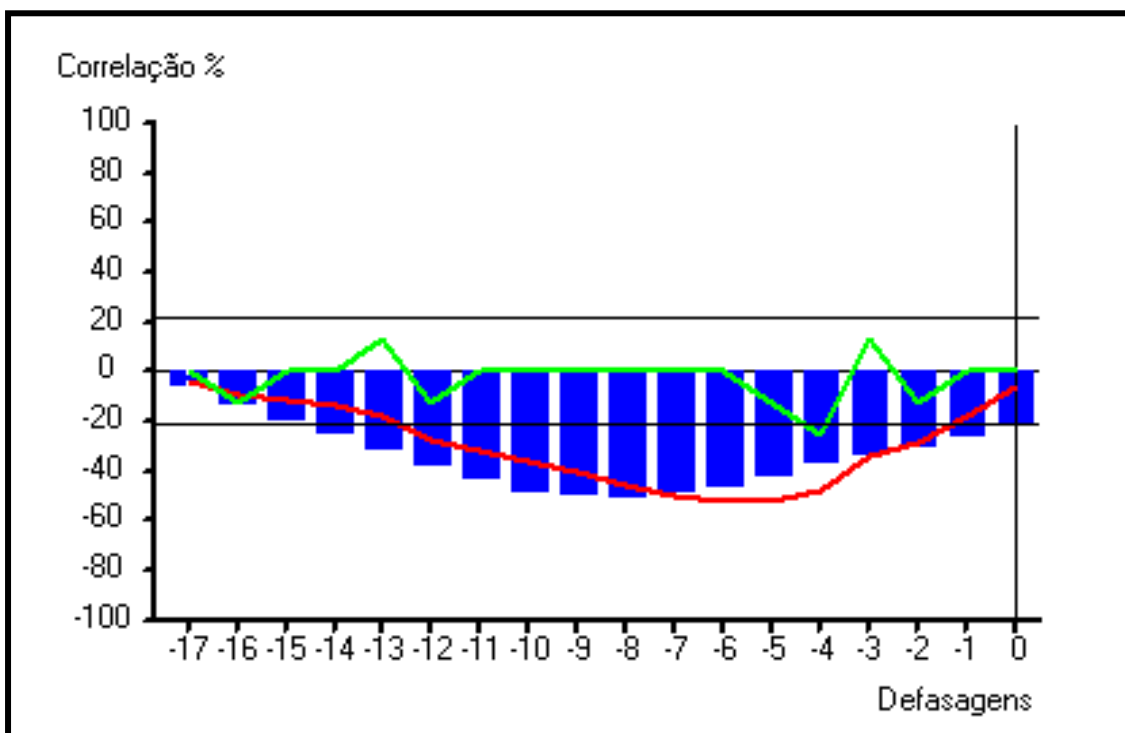


Figura 17 – Correlogramas entre fluxo de passageiros domésticos e a taxa real de juros em duplicatas

Como as amostras para o cálculo dos correlogramas podem diferir, é necessário ajustar a correlação para os diferentes números de observações. A correlação simples $R_{YX}(\lambda)$ considera nos cálculos o período comum disponível para as duas variáveis X e Y, e a ajustada $R'_{YX}(\lambda)$ pondera a correlação simples pela razão entre o número de observações comuns N' e o do período solicitado N.

$$R'_{YX}(\lambda) = R_{YX}(\lambda) \cdot N'/N \quad (4)$$

Ainda que com correlação significativa e com o sentido esperado, não é garantido que a variável-insumo seja incorporada na composição do indicador antecedente final, pois necessita ainda passar por outros testes. É possível que a mesma informação (no sentido de "inovação") esteja contida em outras variáveis, e não teria sentido duplicar a informação fornecida no indicador antecedente.

Em resumo, os avanços satisfatórios e significantes, sob o ponto de vista estatístico e com lógica econômica, recomendam que variáveis, como as examinadas nas figuras sejam selecionadas preliminarmente como candidatas à composição do indicador antecedente. Por outro lado, outras variáveis são irrelevantes para a composição dos indicadores, quer por apresentarem um avanço insuficiente ou inexistente, ou com correlações não-significantes, dentro das linhas tracejadas.

8 - Eliminando a informação redundante

Dependendo do tamanho do banco de dados, o número de variáveis que antecedem significativamente o faturamento real da Indústria pode ser elevado. Um bom número de variáveis é classificado como antecedente simplesmente devido a correlação espúria. Os candidatos diminuem ainda mais com a crítica de bom senso sobre o sinal da correlação e com a eliminação daquelas com avanço operacional considerado insuficiente ou estatisticamente instável. Mas, nem todas as variáveis aprovadas pelo bom senso devem compor o indicador agregado, pois é provável que muitas forneçam o mesmo tipo de informação. Em termos mais simples, isto significa que muitas variáveis que seriam componentes ou insumo do indicador antecedente estão fortemente associadas entre si, ou seja, algumas delas são redundantes e podem ser eliminadas.

Existem duas questões interligadas: (1) o diagnóstico da severidade da informação redundante, e (2) como identificar a sua origem para a correção (eliminação da informação desnecessária).

8.1 Diagnóstico da severidade

O teste adotado no sistema de indicadores antecedentes SIA consiste em examinar o valor do determinante da matriz formada pelas correlações simples entre as variáveis-insumo. No caso de singularidade, o determinante é zero; no caso de independência (ortogonalidade) perfeita - isto é, todas as correlações entre variáveis distintas são nulas - o determinante é igual a um. A questão é como identificar quais as variáveis que mais contribuem para a multicolinearidade, ou seja, onde se localiza a redundância de informação.

8.2 Identificação e localização da redundância

A identificação da fonte da redundância é semelhante ao enfrentado normalmente no caso de multicolinearidade elevada nos modelos de regressão múltipla,

$$Y = X B + u \quad (5)$$

onde Y corresponde ao vetor com N observações da variável explicada; X , a matriz com as M variáveis independentes com N observações X ; B , o vetor de parâmetros; e u , os resíduos aleatórios. Por hipótese, as variáveis Y e X são distribuídas $N(0,1)$ e os resíduos u atendem as propriedades distributivas convencionais. O estimador por mínimos quadrados é obtido pré-multiplicando todos os termos da expressão (5) pela transposta de X , e em seguida pela inversa de $X'X$, ou

$$B = (X'X)^{-1}X'Y \quad (6)$$

com matriz de variância-covariância,

$$V(B) = V^2 (X'X)^{-1} \quad (7)$$

A multicolinearidade corresponde a uma condição de interdependência entre as variáveis-insumo, que independe da sua forma de associação com a variável a ser prevista. Se a interdependência entre os insumos - ou no caso de regressão múltipla, entre as variáveis explicativas X - é elevada, a matriz $X'X$ é singular, e o determinante de sua inversa $(X'X)^{-1}$, infinito. O vetor de parâmetros da resposta dos insumos na explicada torna-se indeterminado. No caso da técnica de indicadores antecedentes, diz-se que existe duplicidade ou redundância de informação no conjunto de variáveis-insumo.

A metodologia dos indicadores antecedentes leva vantagens sobre a regressão múltipla, pois a eliminação de variáveis independentes importantes num modelo de regressão pode acarretar erros de especificação, o que não ocorre com os indicadores antecedentes, Mas a seleção da informação relevante para os indicadores ante-cedentes tem traços comuns com o problema da multicolinearidade nas regressões. Aqui como lá, as variáveis não despertam igual interesse; algumas são mais importantes que as outras. O critério da parcimônia nos recomenda isolar um pequeno grupo de variáveis, com conteúdo informacional relevante e variado, com estatísticas de fácil aceso e com o mínimo de atraso.

Para isolar este grupo de variáveis existem três métodos: (1) a “regra de bolso”, sugerida por Klein¹²; (2) o enfoque de Farrar & Glauber¹³; e (3) o teste F ¹⁴. A “regra de bolso” de Klein parte do princípio de que toda multicolinearidade é relativa e a sua existência depende da correlação simples entre duas variáveis independentes ser maior que a correlação múltipla da variável dependente com todas as independentes. Esta regra serve inclusive para ordenar as variáveis que tornam mais séria a multicolinearidade.

Entretanto, apesar do bom senso e de atender a um teste mais imediato, esta regra tem o inconveniente de desviar a atenção do fato de que a multicolinearidade é um problema de interdependência e não de dependência. Além disso, a multicolinearidade completa, no caso de singularidade perfeita dentro de um conjunto de variáveis independentes, é consistente com baixas correlações entre pares de variáveis. Basta lembrar que cada vetor de um

¹² Klein, Lawrence R., An Introduction to Econometrics, (Englewood Cliffs, N.J. Prentice Hall, 1962), p.101

¹³ Farrar, D.E. & R.R. Glauber, "Multicolineaty in regression analysis : the problem revisited", Review of Economics and Statistics, vol. 39, fevereiro de 1967, pp.92-107 1975, pp.366-368.

¹⁴ Rossi, José W., “A matriz de correlação revisitada”, Revista Brasileira de Estatística, vol.38, outubro/dezembro de 1977, pp.379-384.

conjunto de variáveis *dummy*, cuja combinação linear iguala a uma constante, é perfeitamente colinear com a constante de regressão, e ainda assim a correlação simples entre os pares de variáveis pode ser nula.

Um teste mais severo com a regra de Klein consiste em examinar o valor do determinante da matriz $X'X$ formada pelas correlações entre as variáveis-insumo. No caso de singularidade, o determinante é zero; no caso de independência (ortogonalidade) perfeita - isto é, todas as correlações entre variáveis distintas são nulas - o determinante é igual a um. Infelizmente, o valor crítico para o determinante não é definido, e o teste de Klein em si não indica como resolver o problema da multicolinearidade, quando detectada.

Farrar e Glauber¹⁵ desenvolveram um teste para identificar a existência de uma multicolinearidade severa e identificar a sua causa.¹⁶ O teste consiste em comparar a estimativa do determinante da matriz de correlações simples das variáveis-insumo com o nível crítico. Se a estimativa for superior ao valor crítico, diz-se que a multicolinearidade não chega a afetar substancialmente o modelo e as variáveis são significativamente independentes entre si.

Se r^{ii} é o elemento diagonal da matriz inversa $(X'X)^{-1}$ correspondente a variável X_i , sabemos que,

$$\text{Det } \{(X'X)_{ii}\} = r^{ii} \text{ Det } \{X'X\} \quad (8)$$

onde Det representa o determinante e $(X'X)_{ii}$ corresponde a matriz $X'X$, sem a variável X_i .

Assim, se X_i for independente (ortogonal) das demais variáveis,

$$\text{Det } \{(X'X)_{ii}\} = \text{Det } \{X'X\} \quad (9)$$

e portanto $r^{ii} = 1$. Se X_i for perfeitamente dependente (colinear),

$$\text{Det } (X'X) = 0 \quad (10)$$

e portanto r^{ii} é infinito. O simples conhecimento dos valores assumidos por r^{ii} por si só permite identificar a origem da multicolinearidade, mas é possível melhorar o teste.

Observamos que a transformação de Det $(X'X)$ atende a distribuição Qui-quadrado χ^2 ,

¹⁵ Klein, *op.cit.*

¹⁶ Existem críticas ao teste de Farrar e Glauber, como as de Kumar, Krishna T., "Multicollinearity in regression analysis", *Review of Economics and Statistics*, vol. 45, março de 1975, pp.365-369; Wichers, Robert C., "The detection of multicollinearity: a comment", *Review of Economics and Statistics*, vol.45, março de 1975, pp.366-368; O'Hagan, John & Brendan McCabe, "Tests for the severity of multicollinearity in regression analysis: a comment", *Review of Economics and Statistics*, vol.45, março de 1975, pp.368-370. As críticas ao teste de Farrar-Glauber são pertinentes no caso de regressão múltipla com variáveis *dummy*, o que não é o nosso caso.

$$\chi^2(gl) = -\frac{1}{6} \{ N - 1 - \frac{1}{6} (2 M + 5) \} \text{Log Det} (X'X) \quad (11)$$

onde $gl = 0,5 M (M - 1)$ corresponde aos graus de liberdade; N é o número de observações e M , o de variáveis. O valor encontrado em (11) pode ser testado com a tabela da distribuição Qui-quadrado. Observe que $X'X$ corresponde à matriz quadrada simétrica formada pelos coeficientes de correlação entre as variáveis.

O teste sobre a existência de uma multicolinearidade severa consiste em comparar a estimativa do determinante $|X'X|$ com o nível crítico $|X'X|^*$, conforme apresentado por Contador e Ferraz¹⁷. Se a estimativa for superior ao valor crítico, diz-se que a multicolinearidade não chega a afetar substancialmente o modelo. No nosso caso, o mesmo tipo de teste é aplicado ao conjunto de variáveis aceitas preliminarmente para compor um indicador antecedente. Isto permite examinar o grau de duplicidade da informação contida no conjunto de variáveis.

A identificação da variável ou variáveis que duplicam a informação contida nas demais pode ser feita com o enfoque desenvolvido anteriormente por Wilks¹⁸, e reproduzido por Farrar e Glauber¹⁹. Lembrando que podemos escrever,

$$r_{ii} = (1 - R_{xi}^2)^{-1} \quad (12)$$

onde r_{ii} é o elemento i da diagonal da matriz inversa $(X'X)^{-1}$ e R_{xi}^2 é o coeficiente de correlação múltipla (ao quadrado) entre x_i e os outros elementos de X . Quando x_i for colinear (perfeitamente dependente) dos demais X , $R_{xi}^2 = 1$, e $r_{ii} = \infty$. Se x_i é ortogonal, $R_{xi}^2 = 0$ e $r_{ii} = 1$

Após algum algebrismo encontramos,

$$R_{ii} - 1 = \frac{R_{xi}^2}{1 - R_{xi}^2} \quad (13)$$

e notamos que a expressão (13) é a relação entre a variância explicada e não explicada, ou seja a conhecida distribuição F . Isto significa que transformações lineares de (13) terão o mesmo tipo de distribuição. Ainda Wilks havia observado que

¹⁷ Contador, C.R. & C.B. Ferraz, "Parcimônia, informação redundante e multi-colinearidade", COPPEAD/UFRJ, dezembro de 1996.

¹⁸ Wilks, S., "Certain generalization in the analysis of variance", *Biometrika*, vol.24, 1932

¹⁹ *Op.cit.*, p.102

$$\omega = (r^{ii} - 1) \left(\frac{m-n}{n-1} \right) \quad (14)$$

pode ser visto como uma distribuição F com $\underline{m-n}$ e $\underline{n-1}$ graus de liberdade. Este teste de "informação duplicada" – tem a vantagem de não requerer normalidade conjunta de X.

Portanto, um valor estimado de ω menor que o valor crítico ω^* indica ortogonalidade da variável x_i com as demais variáveis do conjunto X. Em caso contrário, isto é $\omega > \omega^*$, a informação de x_i já está contida nas demais variáveis e portanto x_i pode ser eliminada da composição do indicador antecedente.

Portanto, um valor estimado de ω menor que o valor crítico ω^* indica ortogonalidade da variável x_i com as demais variáveis do conjunto X. Em caso contrário, isto é $\omega > \omega^*$, a informação de x_i já está contida nas demais variáveis e portanto x_i pode ser eliminada da composição do indicador antecedente.

Em seguida, a matriz $X'X$ é recomposta “passo-a-passo”, sem a variável eliminada no “passo” anterior. Novos valores r^{ii} são obtidos e testados sucessivamente, até que as variáveis restantes sejam livres de colinearidade elevada. Este processo fornece resultados semelhantes aos da técnica de regressão múltipla *step-wise*, onde o critério de eliminação sucessiva de variáveis é baseado no teste “t”.

O teste da estatística F, apresentado por Rossi²⁰, identifica a localização da colinearidade severa entre as variáveis-insumo através de regressões múltiplas sucessivas entre as variáveis-insumo. Seja EY o indicador antecedente da variável-referência Y e X, o conjunto de k variáveis-insumo, comprovadamente com sério problema de colinearidade. Fazemos numa primeira etapa, um conjunto de k regressões múltiplas onde a variável “dependente” muda sucessivamente e as demais k-1 são as variáveis “independentes”. Cada variável “dependente” X é explicada sucessivamente pelas demais.

$$\left. \begin{array}{l} X_1 = g_1 (X_2, X_3, \dots, X_k) \\ X_2 = g_2 (X_1, X_3, \dots, X_k) \\ \dots \\ X_k = g_k (X_1, X_2, \dots, X_k) \end{array} \right\} \quad (15)$$

Se a informação contida na variável-insumo X_j for uma combinação linear das informações parciais contidas em qualquer conjunto de X, a estatística F é significativamente diferente de zero, então a variável X_j pode ser descartada. Na segunda etapa, as k-1 variáveis-insumo restantes são

²⁰ Op.cit.

submetidas a k-1 regressões, no mesmo critério acima, e retirada aquela variável mais explicada pelas demais. No final do processo, temos um conjunto de variáveis não correlacionadas (ou com colinearidade pouco severa) entre si. O problema deste enfoque é o grau de liberdade no caso de um grupo numeroso de variáveis iniciais. Em casos limites, o número de variáveis-insumo candidatas podem ser maior do que o número de observações.

A eliminação da informação redundante é feita em duas etapas. Na primeira, as variáveis identificadas como antecedentes pelos correlogramas (e após a eliminação daquelas que agridem o bom senso) são distribuídas por classes de fenômenos. Na segunda, as variáveis são submetidas aos demais testes. O grau de severidade do teste pode ser constatado pelo seguinte resultado. De um banco de dados contendo mais de 2900 variáveis mensais, cerca de 458 mostraram ser antecedentes às flutuações no fluxo de passageiros. Após a análise crítica de bom senso, este número diminui para 72 variáveis. Com o teste de redundância de informação, o número final de variáveis-insumo do indicador antecedente é entre dez e vinte séries.

9 - Composto os indicadores antecedentes

Do exposto sobre a defasagem dos movimentos cíclicos entre variáveis e sobre a redundância de informações, podemos resumir um conjunto de regras básicas para nortear a escolha das variáveis-insumo. Para que seja escolhida como componente final de um indicador composto, uma variável deve atender a seis critérios básicos, nesta ordem de prioridade:

- a) Homogeneidade no número de períodos de avanço: o avanço da variável-insumo deve ser (aproximadamente) o mesmo nos três critérios.
- b) estabilidade, significância estatística e lógica econômica do seu avanço em relação ao fenômeno a ser previsto. O correlograma entre a variável-insumo e a variável-referência, embora possa sofrer mudanças no seu formato, não deve indicar avanços muito distintos em diferentes períodos amostrais. Por exemplo, se o crédito à pessoa física antecipa em quatro meses as flutuações no faturamento real da Indústria no período 1990 a 2005, dizemos que esta variável é considerada confiável se em outros períodos - ou em sub-períodos - ela mantiver um avanço não muito distinto dos quatro meses. Este cuidado evita que o avanço estimado seja espúrio, principalmente naqueles casos em que a teoria e o bom senso não são suficientes para reconhecer *a priori* a variável como antecedente.
- c) a disponibilidade de informações atualizadas e de boa qualidade, com o mínimo de atraso. Não basta que o avanço estimado através de correlogramas seja elevado, estável, e aceito pelo bom senso. Devido ao atraso na divulgação, o avanço estimado estatisticamente pode encurtar-se bastante

na prática. Por isso, é importante distinguir entre o "avanço estatístico", obtido mecanicamente com o correlograma, e o "avanço operacional", que corresponde a diferença entre o avanço estatístico e o período médio necessário para obter dados atualizados da variável.²¹ O avanço estatístico é utilizado na montagem do indicador composto, mas em última instância será o avanço operacional que apontará a utilidade da previsão.

- d) quanto mais estável e uniforme for o avanço da variável-insumo nos ciclos de periodicidade e com características distintas. O ideal seria construir indicadores compostos de tal forma que cada variável antecedesse o fenômeno a ser previsto por um determinado período constante, em qualquer ciclo, por exemplo, oito meses para as concordatas, nove meses para um agregado monetário, etc. Estas condições são satisfeitas com o "retardo puro" (*pure delay*), mas a possibilidade de sua ocorrência é mínima.
- e) quanto mais estável e uniforme a relação entre o indicador e a atividade a ser prevista. Por exemplo, se fosse realizada uma regressão entre as flutuações do indicador e do faturamento real da Indústria, decompostas em ciclos com periodicidade distinta, seria desejável que o coeficiente da regressão, no caso denominado de "ganho" ou "resposta", fosse idêntico para os diversos ciclos. O "ganho" é dito "amortecido", "idêntico" ou "ampliado" numa determinada frequência, se, para cada movimento de um por cento no indicador ocorre, na média, uma variação inferior, igual, ou superior a 1% no nível de atividade, naquele mesmo ciclo. Uniformidade no ganho significa que ele assume valor único em todas as frequências, não importando a periodicidade do ciclo;
- f) originalidade do conteúdo informacional da série em relação às demais variáveis componentes do indicador. Em termos práticos, esta regra diz que deve se evitar incluir variáveis que refletem o mesmo tipo de informação, muito embora elas isoladamente possam estar fortemente correlacionadas com o fenômeno a ser previsto. Por exemplo, o crédito e a oferta de moeda tendem a ser fortemente correlacionados e ambos antecedem o fluxo de passageiros. Estão também fortemente correlacionados entre si (correlação de 75%). Assim, não devem ser incluídos em bloco na composição final do indicador, pois isto apenas duplicaria o tipo de informação (liquidez da economia).

Estas seis condições permitem montar indicadores antecedentes compostos, cuja qualidade preditiva depende da hipótese básica de que a estrutura de avanço das variáveis escolhidas (componentes) permaneça válida para o futuro próximo. Esta hipótese não é restrita à metodologia dos

²¹ Por exemplo, o avanço estatístico do crédito sobre o fluxo de passageiros é de quatro meses, mas o avanço operacional diminui para dois meses, devido ao atraso na divulgação das estatísticas pelo Bacen. Em alguns casos, o atraso na divulgação produz avanços operacionais tão pequenos ou negativos (o atraso na divulgação é maior do que avanço estatístico), que elimina a variável-insumo.

indicadores antecedentes como parece a primeira vista. Na verdade, ela é crucial em qualquer sistema de previsão.

Para evitar que as séries com as maiores flutuações dominem as demais variáveis no processo de agregação, as variáveis finais (inclusive a variável-referência), já deflacionadas e/ou acumuladas em doze meses, e depuradas pelo filtro (2), são normalizadas.

$$z'(t) = \frac{z(t) - a(z)}{\sigma_z} \quad (16)$$

onde $a(z)$ retrata a média; e σ_z , desvio-padrão da série no período amostral escolhido.

Em seguida, as M variáveis-insumo finais escolhidas são ponderadas e agregadas num índice composto,

$$y^*(t) = \sum_{i=1}^M w_i z'_i(t - i) \quad (17)$$

onde w_i é o peso da variável antecedente normalizada z'_i defasada em i períodos (meses). O peso w_i é baseado na correlação ajustada entre a variável X_i e a variável-referência Y que se pretende prever.²² Ou seja,

$$w_i = r'_i / \{ \sum r'_i \} \quad (18)$$

onde r'_i é a correlação ajustada (máxima) entre as taxas de variação da variável X_i e da variável-referência Y (faturamento real da Indústria). O valor do avanço em meses é obtido diretamente do correlograma, e corresponde ao avanço estatístico que conduz a maior correlação ajustada (valor absoluto) entre as variáveis.

Outro conceito importante é o avanço médio $av.$ do indicador, definido pela média ponderada dos avanços de cada variável X_i ,

$$av. = \sum_{i=1}^M i [r'_i] / \{ \sum_{i=1}^M [r'_i] \} \quad (19)$$

Como a fração de um período não tem sentido prático, os avanços médios devem ser expressos no maior número inteiro. Ou seja, embora a estimativa com a expressão (18) possa fornecer um avanço médio, digamos de

²² Outras alternativas são utilizar pesos iguais, ou baseados em critérios subjetivos, ou ainda através de coeficientes de regressão, como feito por Maher, John E., "Forecasting industrial production", *Journal of Political Economy*, vol.65, abril de 1957, pp.158-165.

8,7 meses, para fins práticos dizemos que o indicador tem o avanço médio de 8 meses.

Para que o indicador $y^*(t)$ tenha a mesma dimensão da variável-referência torna-se necessário "desnormalizá-lo", o que é feito pela expressão;

$$E \{ y(t) / \Theta \} = a(y) + \sigma_y y^*(t) \quad (20)$$

onde Θ representa o conjunto de informações disponíveis nas variáveis componentes, e $a(\cdot)$, o operador-média. A expressão (19) tem a vantagem de apresentar um indicador expresso na mesma unidade de Y . Embora não seja o preditor mais eficiente para os níveis ou das taxas de crescimento de Y , o indicador $E\{y(t)\}$ fornece a melhor previsão possível para a cronologia cíclicas futuras.

Atendendo as regras acima, montamos indicadores antecedentes para o fluxo de passageiros em vôos domésticos e internacionais, expresso em taxa de crescimento do acumulado em 12 meses. A Figura II.8 abaixo reproduz o ajuste do indicador para o fluxo doméstico no período após 1994, com previsões cíclicas até o início de 2007. O indicador antecedente é formado por 19 variáveis-insumo. Pela composição das variáveis-insumo, o avanço estatístico médio é de nove meses, reduzido para sete meses (avanço operacional) devido ao atraso médio na divulgação das variáveis-insumo. A taxa de crescimento observada está em cor azul, e a previsão com o IA, em vermelho.

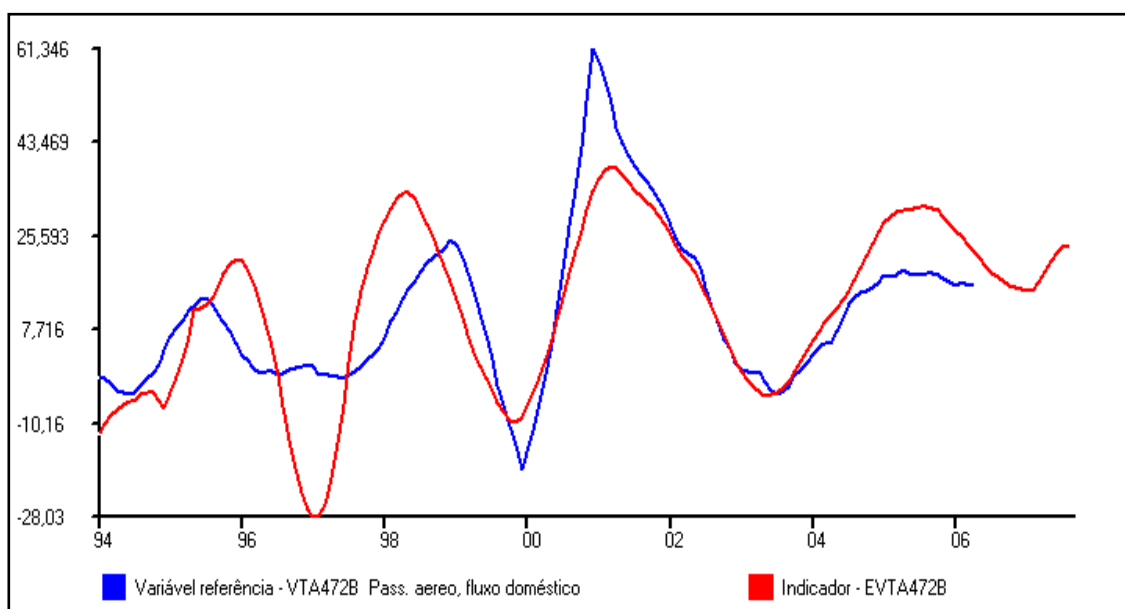


Figura 17 – Desempenho do indicador antecedente para o fluxo de passageiros domésticos

A Figura 18 mostra o indicador para o fluxo de passageiros em vôos internacionais, sendo o indicador formado por 16 variáveis-insumo e com

avanço estatístico de sete meses, reduzido para cinco meses com o atraso das estatísticas.

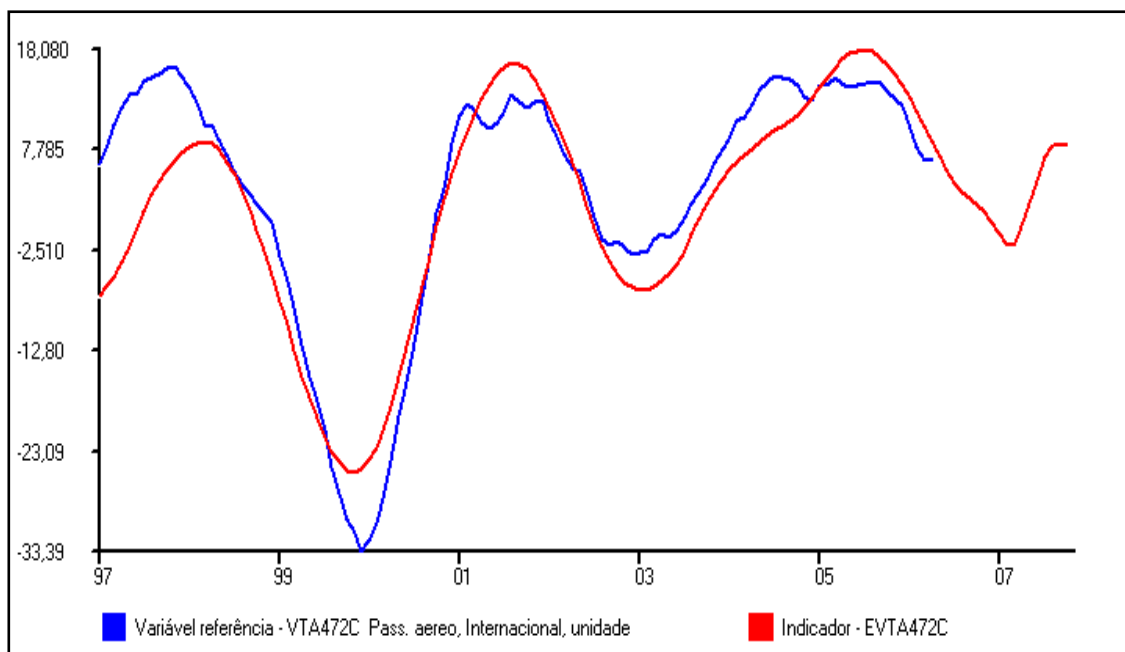


Figura 18 – Desempenho do indicador antecedente para o fluxo de passageiros internacionais

Ambos indicadores retratam com fidelidade as fases cíclicas, com uma composição de variáveis domésticas e externas. As correlações dos indicadores antecedentes com as flutuações observadas são de 87 % para o fluxo doméstico e 96 %, para o internacional. No aspecto visual, o ajuste dos indicadores antecedentes é satisfatório, captando as reversões cíclicas ocorridas desde a década de 90. Os indicadores prevêm que a fase de contração do fluxo de passageiros se mantém até o final de 2006, e a fase de recuperação deve iniciar em 2007.

Apesar do ajuste cíclico satisfatório, o sistema de indicadores antecedentes não é infalível na cronologia das reversões. O passado de acerto das datas de reversão recomenda confiança nas suas previsões. Porém, isto não significa que o erro seja nulo. Embora a esperança do erro cronológico (diferença entre a data efetiva e a prevista) de uma reversão cíclica seja nula, na prática, os indicadores antecedentes não estão livres deste problema. Para amostras com um número elevado de ciclos, a análise da cronologia das reversões mostra a distribuição de frequência dos erros cronológicos.

10 - Desempenho nas reversões

Uma forma de avaliar o desempenho do indicador antecedente é comparar a sua tendência com o da variável-referência. O indicador ideal retrata cronologicamente as mesmas fases cíclicas da variável-referência. Se

o indicador aponta uma fase de expansão (contração), enquanto a variável-referência, uma fase de contração (expansão), ele se torna suspeito. O teste para esta análise é feito através de uma Tabela de Contingência, onde as primeiras diferenças do indicador e da variável-referência são distribuídas em quatro células.²³

Se as mudanças (primeiras diferenças) do indicador (ΔI) e da variável-referência (ΔY) tiverem distribuição independente e aleatória entre si, o número total de observações correspondentes à fase de expansão da variável-referência (N^+) estaria igualmente distribuído entre as células A e B do Quadro 3, e os da fase de contração (N^-), entre as células C e D.

Um indicador ideal, que acerta todas as mudanças de tendência (ou seja, sempre em fase com a variável-referência) tem as células A (na observações) e D (nd observações) nulas, e as células C (nc) e B (nb) com número de observações respectivamente iguais a N^- e N^+ . Comparando o número observado em cada célula com o esperado (entre parênteses), aplicamos o teste Qui-quadrado, com um grau de liberdade,

$$\chi^2 (1) = \sum_i \frac{(n_i - en_i)^2}{en_i} \quad (21)$$

Quadro 3
O teste de desempenho de tendência

		ΔI		
		-	+	Σ
ΔY	+	A, n1 (en1= $N^+/2$)	B, n2 (en2= $N^+/2$)	N^+
	-	C, n3 (en3= $N^-/2$)	D, n4 (en4 = $N^-/2$)	N^-

Se $\chi^2 (1)$ calculado for maior do que o valor da tabela Qui-quadrado, a hipótese de independência entre as tendências do indicador e da variável-referência é rejeitada.

²³ Conforme sugerido por Schnader, M.H. & H.O.Stekler, "Evaluating prediction of change", *Journal of Business*, vol.63, no.1, parte 1, janeiro de 1990, pp.99-107. Ver também Henriksson, Roy D. & Robert Merton, "On market timing and investment performance 2 : statistical procedures for evaluating forecasting skills", *Journal of Business*, vol.54, outubro de 1981, pp.513-533.

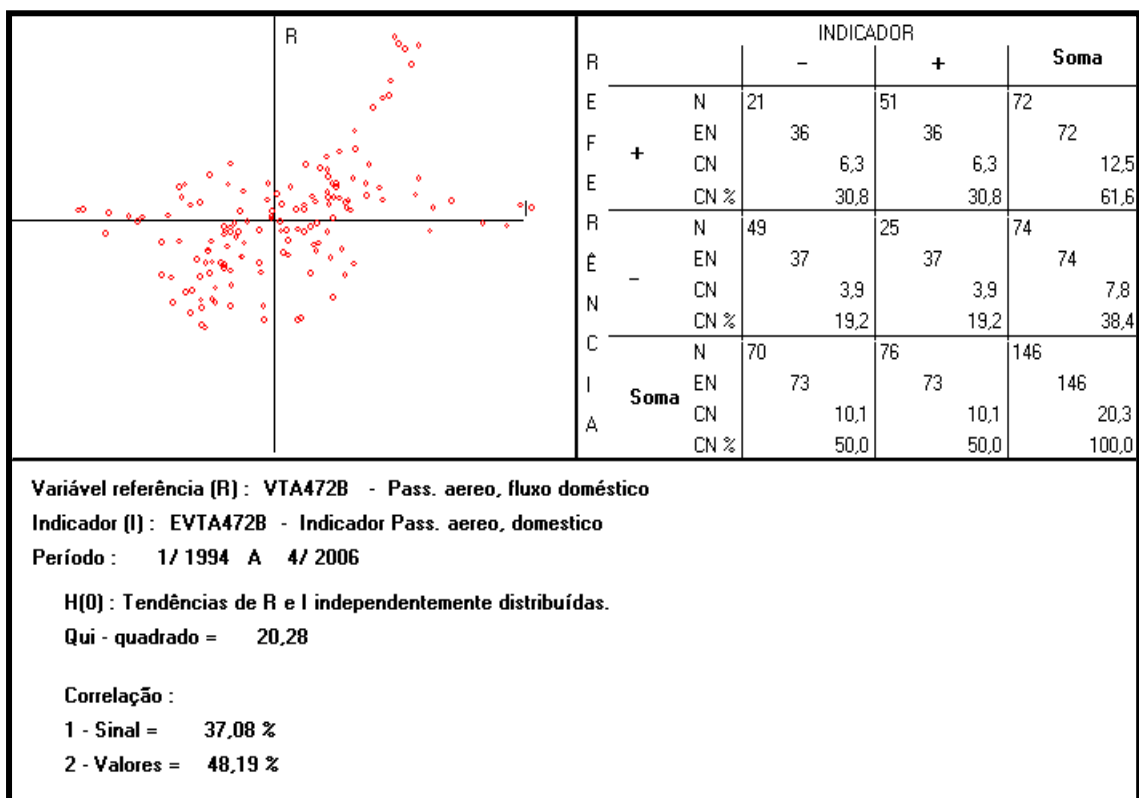


Figura 19 – O desempenho do indicador antecedente dos passageiros domésticos nas fases.

Na Figura 19, N representa o número de observações na célula; EN, o número esperado; CN, a contribuição da célula para o valor do Qui-quadrado; e CN%, a contribuição percentual. Aplicando o teste no indicador antecedente para as flutuações do fluxo de passageiros domésticos, a hipótese de independência entre as fases é rejeitada ao nível de 1 %, com Qui-quadrado igual a 20,3. Observe que a dispersão dos pontos está concentrada ao longo da diagonal no sentido crescente.²⁴

²⁴ O teste é fornecido pelo sistema SIA® onde N, na Figura 4.2, representa o número de observações na célula; EN, o número esperado; CN, a contribuição da célula para o valor do Qui-quadrado; e CN%, a contribuição percentual.

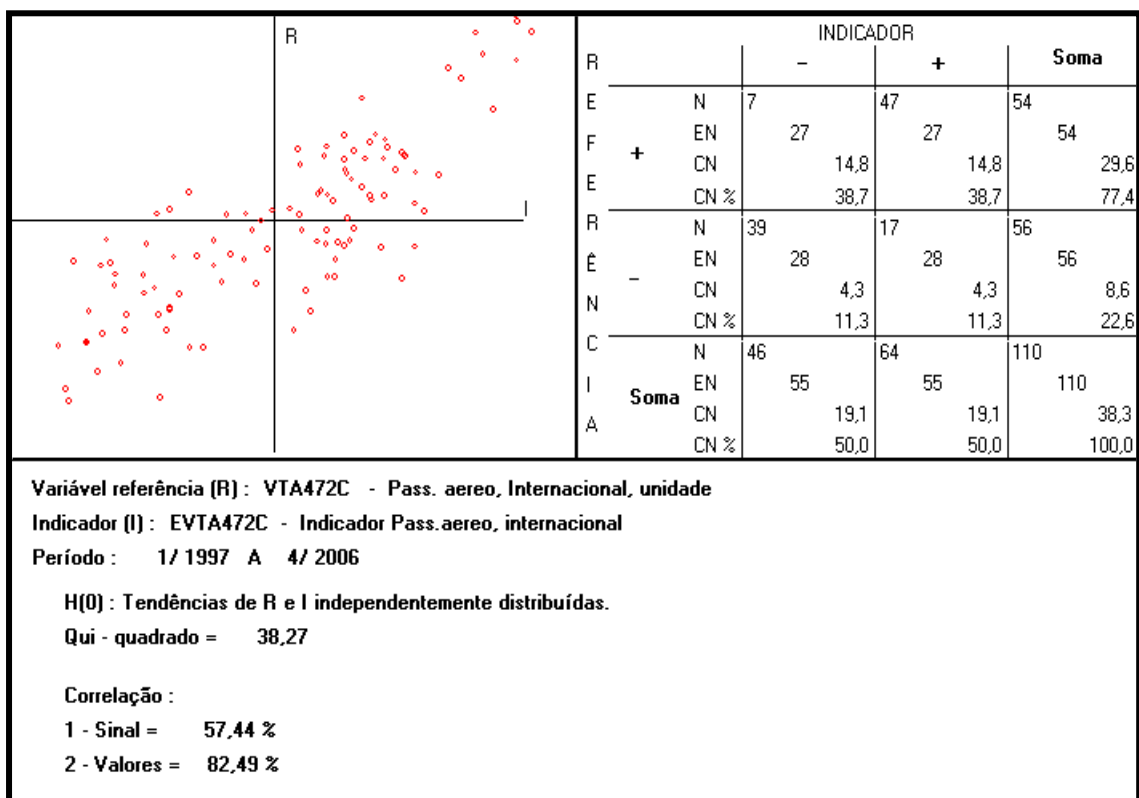


Figura 20 – O desempenho do indicador antecedente dos passageiros internacionais.

Para o indicador do fluxo de passageiros internacionais, o teste Qui-quadrado é de 38,3, igualmente satisfatório. Outros testes de desempenho *ex-post* podem ser extraídos do Quadro 3. Vamos adicionar dois outros. A relação entre a soma do número de pontos nos quadrantes A e D, em comparação com a de B e C fornece uma medida do erro relativo *der*,

$$der = (n1 + n4)/(n2+n3) \quad (21)$$

que, no caso da Figura 18 para os vôos domésticos corresponde a,

$$(21 + 25)/(49 + 51) = 46/100 = 0,46$$

Outro teste examina a gravidade do erro. O número (relativo) de pontos no quadrante D fornece a medida da gravidade. Note que estes pontos indicam o número de vezes em que o indicador antecedente sinalizou uma expansão rejeitada pela variável-referência, que apontou uma retração. É um erro sério, mais grave do que os pontos em A, em que o indicador sinalizou uma contração desmentida pela expansão na variável-referência. Este teste pode ser denominado de gravidade relativa do erro – *gre*,

$$gre = n4/n1 \quad (22)$$

e quanto menor a relação *gre*, menor a gravidade do erro. Os dados produzem um *gre* igual a,

$$gre = 25/22 = 1,19$$

11 - Probabilidade das reversões

Ainda que o desempenho passado do indicador antecedente em prever as reversões ocorridas sirva para qualificar o seu ajuste, este desempenho é de pouca ajuda, quando se trata de inferir a sua qualidade nas previsões de futuras reversões. Como qualquer método de previsão, os indicadores antecedentes podem errar, e o usuário deve estar consciente desta possibilidade. Ainda assim, interessa saber a probabilidade de que a reversão prevista deve ocorrer. A avaliação do desempenho do indicador antecedente em prever *ex-ante* as reversões cíclicas utiliza diversos critérios, alguns populares, e outros apoiados em métodos estatísticos mais rigorosos.

Por construção, os indicadores antecedentes avançam no tempo, para períodos em que a variável-referência ainda não é conhecida. Se for necessário aguardar a divulgação das estatísticas da variável-referência para apontar a ocorrência de uma reversão, o objetivo principal do indicador está perdido. Por esta razão, foram desenvolvidas algumas regras para sinalizar a reversão antes que ela ocorra efetivamente.

O critério mais popular é qualificar uma nova fase se o indicador mantém o mesmo sentido por três meses consecutivos.²⁵ Uma variante – utilizado com preços de ações – é baseada na acumulação das variações do indicador até que o valor atinja um certo percentual estabelecido pelo analista.²⁶ Um terceiro critério utiliza a regra baseada no número de meses em que o indicador antecedente é menor do que o último pico.²⁷ Outras regras simples combinam critérios tradicionais e estatísticos.²⁸ Porém, antes de rejeitar estas regras simplistas, é importante que sejam submetidas a testes de desempenho. Se rejeitadas devido ao acúmulo de erros cometidos, então o analista deve decidir se avança para outras metodologias mais complexas estatisticamente.

²⁵ Vaccara, B. & V. Zarnowitz, “How good are the leading indicators?”, Proceedings of the Business and Statistics Section, American Statistical Association, 1977, pp.41-50.

²⁶ Para exemplo da aplicação deste critério para previsão dos índices de Bolsas de Valores veja Higgins, Bryon, “Is a recession inevitable this year?”, Economic Review, Federal Reserve Bank of Kansas City, janeiro de 1988, pp.3-16; Malabre, Alfred, “As economy goes, so goes stock market”, Wall Street Journal, 9 de fevereiro de 1987.

²⁷ Alexander, S. & H. Stekler, “Forecasting industrial production – leading series versus autoregression”, Journal of Political Economy, vol.67, 1959, pp.402-409.

²⁸ Hymans, S., “On the use of leading indicators to predict cyclical turning points”, Brookings Papers on Economic Activity, vol.2, 1973, pp.339-384; Zarnowitz, V. & G. Moore, “Sequential signals of recession and recovery”, Journal of Business, vol.55, 1982, pp.57-85; Keen, H., “Leading economic indicators can be misleading, study shows”, Journal of Business Forecasting, vol.2, 1983, pp.13-14.

A metodologia de Neftçi foi apresentada em 1982²⁹ e tornou-se o padrão para avaliação do desempenho de indicadores antecedentes, através de probabilidade. O princípio básico da regra de Neftçi é semelhante aos critérios tradicionais: uma reversão seguida de uma contração (expansão) no indicador antecedente sinaliza uma recessão (recuperação) no futuro. Cada observação adicional de queda (aumento) no indicador reforça a probabilidade de recessão (recuperação), até que o valor acumulado da probabilidade atinge o nível crítico, predeterminado e imposto pelo analista. A regra de Neftçi utiliza três componentes:

a) O nível crítico da probabilidade

O valor da probabilidade crítica – de escolha do analista – é uma solução de compromisso entre a possibilidade de erros de falso sinal e a definição tardia da reversão (recuperação). A escolha de uma probabilidade crítica pequena – por exemplo, 60 % - gera o alarme de reversão, que pode vir a ser desmentida por um falso sinal. O nível crítico de 60 % significa que existe uma probabilidade de 40 % de que a reversão não ocorra, ou seja, a probabilidade de 40 % de falso sinal. Por outro lado, um nível crítico de probabilidade mais elevado – por exemplo, 90 % - reduz a probabilidade de um falso sinal, mas exige meses adicionais de observação (para acumulação das probabilidades) e estes meses adicionais podem ser excessivos e preciosos, considerando o tempo necessário para as medidas preventivas.

Assim, o dilema é entre escolher um nível crítico mais baixo para a probabilidade e correr o risco de um falso sinal, mas ter mais tempo de antecedência, ou escolher uma probabilidade mais elevada, com menor chance de falso sinal, porém com pouco tempo para a prevenção. Quanto maior o valor crítico, menor o número (risco) de falsos sinais e maior o número de reversões não antecipadas/omitidas, lembrando que a probabilidade de reversão e mudança para uma fase de recessão (expansão) é calculada enquanto a variável-referência está passando por uma fase de expansão (recessão). A Figura 21 descreve o dilema. A linha representa a probabilidade de reversão, efetivamente ocorrida (no futuro) em t^* . As linhas horizontais mostram os níveis críticos de probabilidade. Se escolhida a probabilidade de 60 %, o alarme de reversão é disparado em t_0 , e o analista apressado corre o risco (e o emprego!) de sinalizar um falso sinal. Tem, porém mais tempo para as medidas de precaução.

²⁹ Neftçi, N.S., "Optimal Prediction of Cyclical Downturns", *Journal of Economic Dynamics and Control*, Vol.4, 1982, pp.225-241. Para aplicações, ver Niemira, Michael P., "An International Application of Neftçi's Probability Approach for Signalling Growth Recessions and Recoveries using Turning Point Indicators", em Lahiri, Kalal & Geoffrey H. Moore (eds), *Leading Economic Indicators : New Approaches and Forecasting Records*, (Cambridge, Cambridge University Press, 1991).

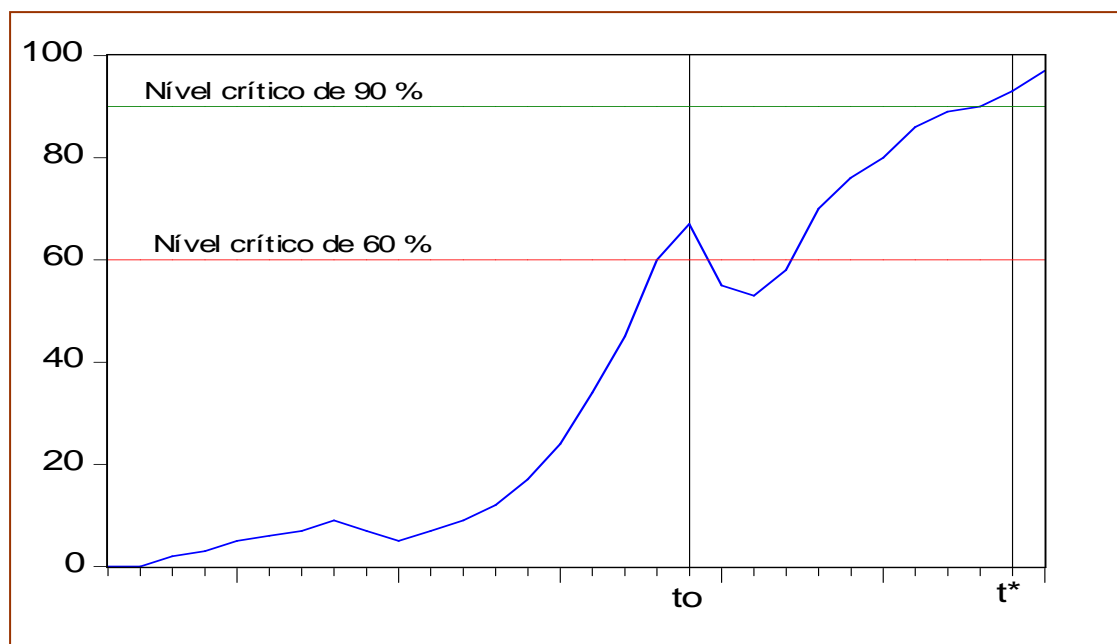


Figura 21 – Probabilidade e o nível crítico para decisão

Por outro lado, se o nível crítico for 90 %, o alarme dispara nas vésperas na reversão. A probabilidade de falso sinal é de apenas 10 %. É um consolo, mas pode ter um preço alto. Enfim, o nível crítico da probabilidade é decidido pelo usuário, que deve balancear os benefícios e custos do alarme prematuro com risco de falso sinal e tempo para as precauções, ou o alarme tardio, com quase certeza, mas sem tempo para a reação.

b) A probabilidade da reversão $Prob_i$

As reversões apontadas num indicador antecedente assinalam (exclusive os falsos sinais) reversões na variável-referência. Esta informação é transformada numa distribuição de probabilidades, com os parâmetros estimados com o modelo *probit*,

$$P(\Delta E_y) = F(\alpha_0 + \alpha_1 \Delta E_y) \quad (23)$$

onde $P(\Delta E_y)$ é a probabilidade de mudança na variável-referência y dada a mudança no indicador antecedente E_y , e $F(.)$ a função cumulativa de probabilidade de distribuição Normal, que transforma (através da *probit*) ΔE_y em valores entre zero e um.

Na Figura 22, os pontos zero e um no eixo vertical representam as variações negativas (zero) e positivas (um) para a variável-referência. $F(.)$ é a função *probit*, estimada com os pontos de dispersão em zero e um.³⁰

A estimação por máxima verossimilhança fornece a probabilidade de que uma reversão ocorrerá no futuro, onde Prob1 refere-se as observações do indicador antecedente pertencentes à fase de expansão, e Prob2, à fase de contração.

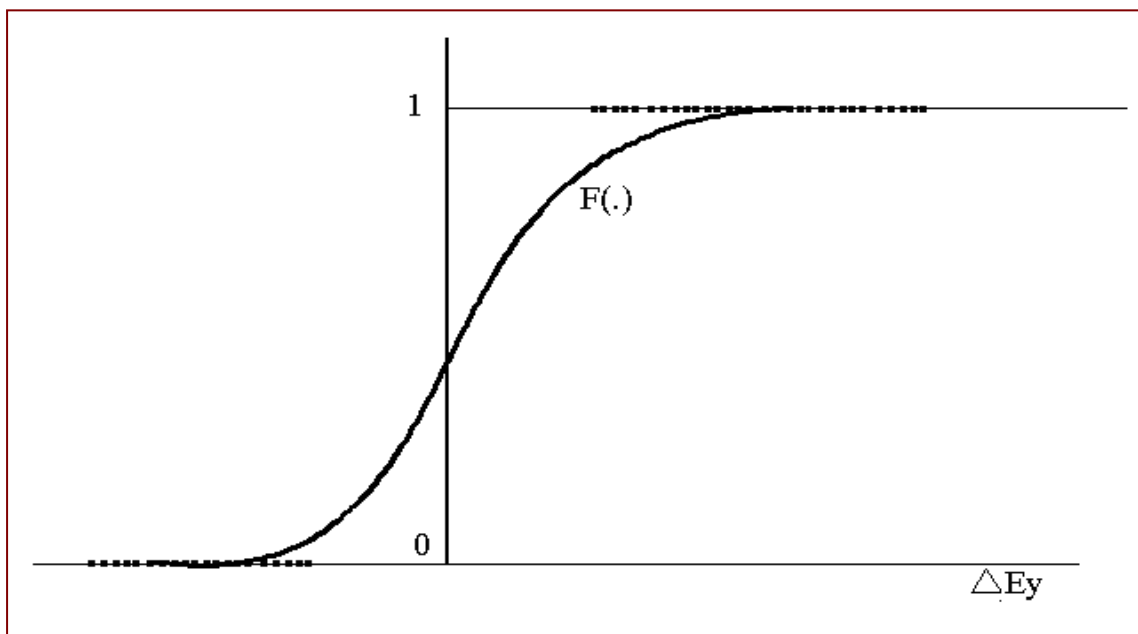


Figura 22 – A distribuição *probit*

c) A distribuição prior

A probabilidade da reversão aumenta com a duração da fase corrente em comparação com a sua duração média histórica. Se as fases de expansão têm uma média histórica de 14 meses, e estamos no décimo mês de expansão ininterrupta, a probabilidade de que a fase esteja terminando é alta. No próximo mês (ou seja, no décimo-primeiro mês), a probabilidade aumenta, e assim por diante, até que a reversão ocorre, e a probabilidade é zerada. A distribuição incondicional de probabilidade *prior* Pr (incondicional, pois independe dos valores assumidos pelo indicador antecedente ou pela variável-referência) é calculada distribuindo a duração média da fase por períodos, iniciando por zero no mês seguinte à última reversão até um valor máximo a partir de uma data anterior à da duração histórica (ou seja, em algum número de meses menor do que 14, no nosso exemplo).

Na metodologia de Neftçi, a probabilidade de uma reversão – independente do tipo de fase que se inicia – tem como elemento a imposição

³⁰ O sistema SIA utiliza o método de estimação com o algoritmo Levenberg-Marquardt. A subrotina está descrita em Press, William e outros, *Numerical Recipes : the art of scientific computing*, (New York, Cambridge University Press, 1987).

de um determinado formato para a distribuição da função cumulativa, escolhido geralmente pelo usuário. A nossa abordagem estabelece um número fixo de períodos para início da contagem cumulativa da probabilidade *prior*, independente da estabilidade da duração da fase cíclica. E se as fases tiverem duração variável ? Certamente, a distribuição *prior* deve considerar esta possibilidade e incorporar outros elementos ao critério. Seja por exemplo a duração média de 14 meses para a fase de expansão dos ciclo de duas variáveis Y_1 e Y_2 . Pelo princípio mais simples, o valor ns para início da contagem da distribuição cumulativa da probabilidade *prior* de uma reversão seria o mesmo.

Porém suponhamos que a variável Y_1 tenha as fases de expansão sempre com a duração de 14 meses, enquanto para a variável Y_2 a duração varia entre 6 e 20 meses. Ora, se o mesmo valor inicial ns for aplicado nos dois casos, a probabilidade *prior* para Y_2 pode não crescer tão rápido como deveria.

Tabela 3
Parâmetros para o cálculo da probabilidade de reversão

Parâmetro	Picos	Vales
Fluxo de passageiros domésticos:		
Duração média da fase	13 meses	19 meses
Desvio-padrão	5 meses	15 meses
Valor inicial nulo até	3 meses	0
Incremento por período	20 %	6 %
Fluxo de passageiros internacionais:		
Duração média da fase	30 meses	12 meses
Desvio-padrão	14 meses	8 meses
Valor inicial nulo até	6 meses	0
Incremento por período	16 %	12 %

d) O critério de decisão

A probabilidade $Prob_t$ de reversão com a regra de Neftçi em t é obtida combinando os três elementos acima,

$$Prob_t = \{ \psi Prob_{t-1} + [Pr (1 - Prob_{t-1}) Prob1] \} / \{ Prob_{t-1} + [Pr (1 - Prob_{t-1}) Prob1 + (1 - Prob_{t-1}) Prob2 (1 - Pr)] \} \quad (24)$$

onde Pr é a distribuição *prior*; $Prob1$, a probabilidade de que uma nova observação do indicador pertença a uma fase de expansão; e $Prob2$, de que pertença a uma recessão. $Prob1$ e $Prob2$ são obtidas com a função *probit*. O parâmetro ψ é o fator de amortecimento para evitar que a probabilidade

composta $Prob_t$ cresça muito rápido. É sugerido que ψ esteja no intervalo 0,2 a 0,6, conforme recomendado por Bikker e Kennedy.³¹

Se

$$Prob_t \geq Prob^* \quad (25)$$

é dado o alarme de uma reversão eminente, onde $Prob^*$ corresponde ao nível crítico da probabilidade.

Os indicadores antecedentes para as flutuações do crescimento do fluxo dos passageiros têm as probabilidades calculadas com os parâmetros da Tabela II.2. As probabilidades estão nas figuras seguintes.

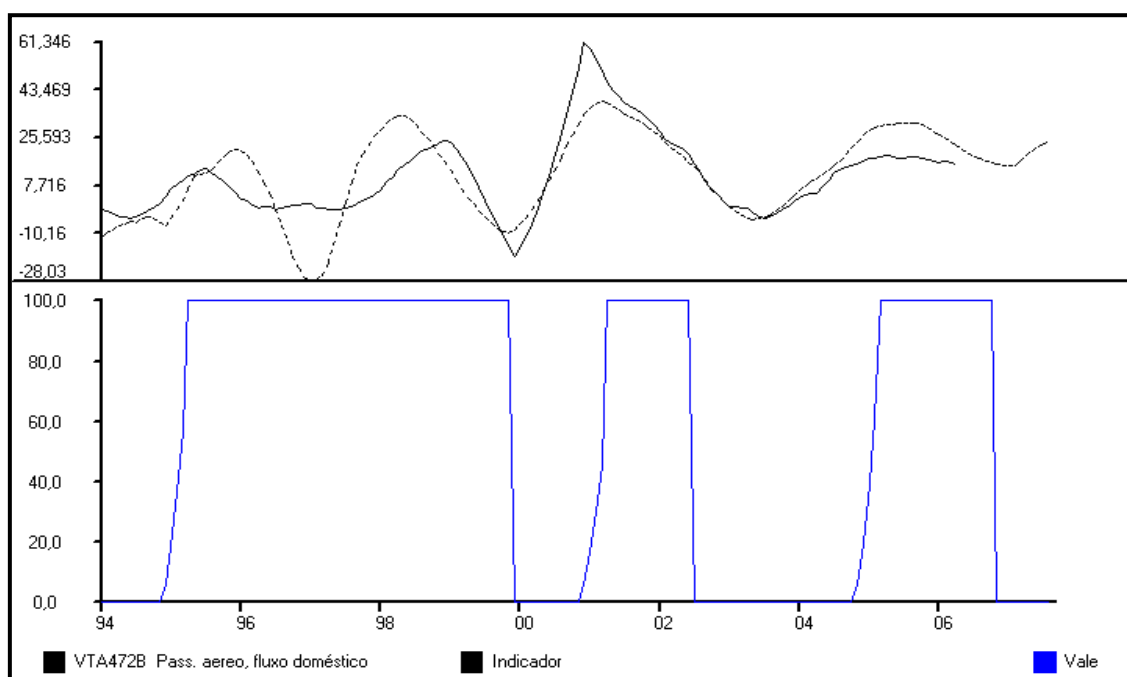


Figura 23 – Probabilidade de ocorrência de uma fase de expansão, fluxo doméstico

³¹ Bikker, J. A. & N.O Kennedy, “Composite leading indicators of underlying inflation for seven EU countries”, *Journal of Forecasting*, vol.18, julho de 1999, pp.225-258.

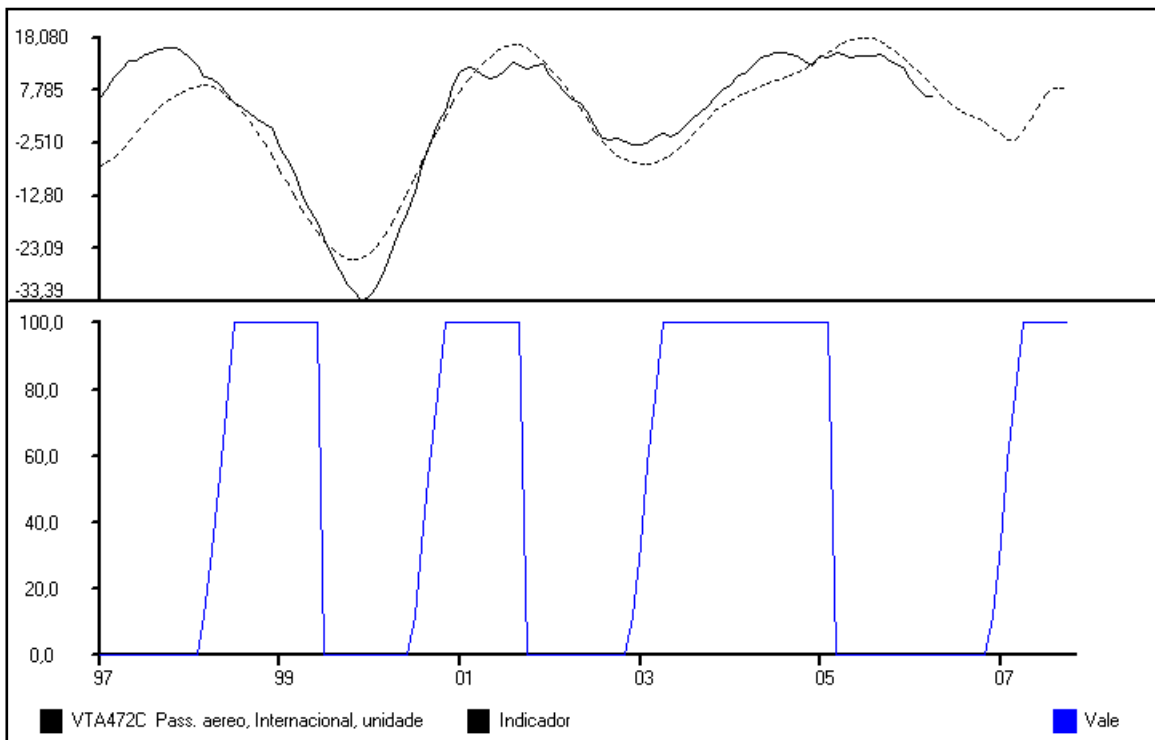


Figura 24 – Probabilidade de ocorrência de uma fase de expansão, fluxo internacional

Com os parâmetros adotados, a probabilidade de ocorrência da fase de expansão prevista pelo indicador antecedente para o fluxo doméstico ainda não foi sinalizada. Para o fluxo internacional, por outro lado, já existe uma probabilidade crescente de ocorrência a partir do início de 2007.

Referências :

- Auerbach, A.J., "The index of leading indicators : `measurement without theory' thirty-five years later", Review of Economics and Statistics, vol. 64, 1982, pp.589-595
- Azzoni, C.R. & Z.A. Latif, "Indicador de movimentação econômica IMEC-FIPE", XVII Encontro Brasileiro de Econometria, Anais, vol.1, 1995, pp.53-69
- Barnes, Leo, "Long-lead vs. medium-lead cycle indicators as business and stock market forecasters", Business Economics, vol.12, no.2, março de 1977, pp.1-11
- Barros, Alexandre Rands, "A periodization of the business cycles in the Brazilian economy, 1856-1985", Revista Brasileira de Economia, vol.47, janeiro/março de 1993, pp.53-82
- Bernanke, Ben, "On the predictive power of interest rates and interest rate spreads", New England Economic Review, novembro- dezembro de 1990, pp.51-68
- Boldin, Michael D., "Dating turning points in the business cycle", Journal of Business, vol.67, no.1, janeiro de 1994, pp.97-132
- Boschen, John & Leonard Mills, "Tests of the relation between money and output in the real business cycle model", Journal of Monetary Economics, vol.22, novembro de 1988, pp.355-374
- Burkholder, Alex A., "In defense of leading indicators", Business Economics, vol.14, no.1, janeiro de 1979, pp.87-89
- Canova, Fabio & Gianni de Nicolo, "Stock returns and real activity: a structural approach", European Economic Review, vol.39, no.5
- Conference Board Inc., "Business Cycle Indicators", Boletim Mensal, vários números
- Contador, Claudio R., "Indicadores da atividade econômica no Brasil", Pesquisa e Planejamento Econômico, vol.6, no.1, abril de 1976, pp. 1-60
- Contador, Claudio R., "Indicadores de atividade no Brasil : uma revisão", Documentos de Política Econômica, no.30, julho de 1976, INPES/IPEA
- Contador, Claudio R., Ciclos econômicos e indicadores de atividade, (Rio, INPES/IPEA, 1977), 237 páginas.
- Contador, Claudio R., "Leading indicators for the industrial sector", Brazilian Economic Review, no.5, 1979, pp.1-32
- Contador, Claudio R., "Inflation targeting and leading indicators for Brazil: some notes", junho de 2000, apresentado no Seminário "One Year of Inflation Targeting", Banco Central do Brasil, Rio, 10-11 de julho de 2000, e como "Metas inflacionárias e política econômica: o emprego de indicadores antecedentes", Relatório SILCON 46, abril 2000
- Chauvet, Marcelle; Solange Gouvea; Marta Baltar Moreira e Jose Ricardo da Costa e Silva, "Two methodologies to build inflation leading

- indicators for Brazil”, Seminário One Year of Inflation Targeting in Brazil, Banco Central, 10-11 julho 2000
- Contador, Claudio R., “Economic activity in 2001 : what the leading indicators forecast”, apresentado no Seminário sobre Indicadores Antecedentes, IPEA/CEPAL/OECD, Rio, 4-5 de dezembro de 2000.
- Contador, Claudio R. & Clarisse B. Ferraz, “Parcimônia, informação redundante e multicolinearidade”, Relatório COPPEAD 312, março de 1997
- Cribari Neto, Francisco, "The cyclical component in Brazilian GDP", Revista de Econometria, Ano 13, no.1, abril de 1993, pp.1-22
- Diebold, Francis X., "Are long expansions followed by short contractions ?", Federal Reserve Bank of Philadelphia, Business Review, julho/agosto de 1993, pp.3-11
- Diebold, Francis X. & Glenn D. Rudebusch, "A nonparametric investigation of duration dependence in the American business cycle", Journal of Political Economy, vol. 98, junho de 1990, pp. 596-616
- Diebold, Francis X. & Glenn D. Rudebusch, "Turning point prediction with the composite leading index: an ex ante analysis", em Kajal Lahiri e Geoffrey H. Moore (eds.), Leading economic indicators : new approaches and forecasting records, (Cambridge, Cambridge Univ. Press, 1991), pp.231-256
- Dotsey, Michael, "The predictive content of the interest rate term spread for future economic growth", Federal Reserve Bank of Richmond, Economic Quarterly, vol. 84, no.3, Verão de 1998, pp.31-52
- Dueker, Michael J., “Strengthening the case for the yield curve as a predictor of US recessions”, Federal Reserve Bank of St. Louis, Review, vol.79, no.2, março-abril de 1997, pp.41-51
- Emerson, Rebecca & David F. Hendry, “An evaluation of forecasting using leading indicators”, Journal of Forecasting, vol.15, no.4, abril de 1996, pp.271-291
- Estrella, Arturo & Frederic S. Mishkin, “Prediction of U.S. recessions: financial variables as leading indicators”, Federal Reserve Bank of New York, Research Paper no. 9609, maio de 1996
- Estrella, Arturo & Frederic S. Mishkin, “The yield curve as a predictor of U.S. recessions”, Current Issues in Economics and Finance, Federal Reserve Bank of New York, vol.2, no.7, junho de 1996, 6 p.
- Falk, Barry, "Further evidence on the asymmetric behavior of economic time series over the business cycle", Journal of Political Economy, vol.94, no.5, setembro de 1986, pp.1096-1109
- Fava, Vera L. & Denisard C. Alves, "Indicador de movimentação econômica, Plano Real e análise de intervenção", Revista Brasileira de Economia, vol.51, no.1, janeiro/março de 1997, pp. 133-143

- Fundação Getúlio Vargas, "A conjuntura no Brasil desde 1822", Conjuntura Econômica, vol.3, abril de 1948, pp.19-27.
- Gabisch, Gunter & Hans-Walter Lorenz, Business cycle theory : a survey of methods and concepts, (Berlin, Springer-Verlag, 1989)
- Ghysels, Eric, "On scoring asymmetric periodic probability models of turning-point forecasts", Journal of Forecasting, vol.12, no.3 & 4, abril de 1993, pp.227-238
- Goldstein, Morris, "Presumptive indicators/early warning signals of vulnerability to financial crises in emerging market economies", mimeo., Washington, Institute for Economics, 1996
- Harvey, Campbell R., "Forecasts of economic growth from the bond and stock markets", Financial Analysts Journal, vol.45, no.5, setembro/outubro de 1989, pp.38-45
- Harvey, Campbell R., "The yield curve, stock returns and the prediction of Canadian economic growth", Working Paper, Duke University, 1993.
- Jun, Duk Bin & Young Jin Joo, "Predicting turning points in business cycles by detection of slope changes in leading composite index", Journal of Forecasting, vol.12, no. 3 & 4, abril de 1993, pp.197-213
- Lahiri, K. & G. H. Moore (eds.), Leading economic of indicators new approaches and forecasting records, (Cambridge, Cambridge University Press, 1991)
- Mills, Leonard, "Can stock prices reliably predict recessions?", Federal Reserve Bank of Philadelphia, Business Review, setembro/outubro de 1988, pp.3-14
- Neftci, Salih N., "Optimal prediction of cyclical downturns", Journal of Economic Dynamics and Control, vol.4, novembro de 1982, pp. 225-241
- Niemira, M., "An international application of Neftci's probability approach for signaling growth recessions and recoveries using turning point indicators", em Kajal Lahiri e Geoffrey H. Moore (eds), op.cit., pp.91-108
- OECD, "OECD leading indicators and business cycles in member countries 1960-1985", Sources and Methods, vol.39, 1987, Paris
- Palash, C. J. & L.J. Radecki, "Using monetary and financial variables to predict cyclical downturns", New York Federal Reserve Bank, Quarterly Review, 1985, pp.36-45
- Pearce, Douglas K., "Stock prices and the economy", Federal Reserve Bank of Kansas City, Economic Review, novembro de 1983, pp.7-22
- Petersen, Bruce & Steven Strongin, "Why are some industries more cyclical than others?", Journal of Business & Economics Statistics, vol.14, no.2, abril de 1996, pp.189-198
- Remolona, Eli M., "Global stock markets and links in real output", Federal Reserve Bank of New York, Research Paper no. 9109, março de 1991

- Schwert, G. William, "Stock returns and real activity : a century of evidence", Journal of Finance, vol.45, setembro de 1990, pp.1237-1257
- Stock, James H. & Mark W. Watson, "A probability model of the coincident economic indicators", em K.Lahiri & G.H. Moore (eds.), Leading economic indicators : new approaches and forecasting records, (Cambridge, Cambridge University Press, 1991)
- Stock, James H. & Mark W. Watson, "A procedure for predicting recessions with leading indicators", Econometric Issues and Recent Performance, Federal Reserve Bank of Chicago, Working Paper 92-7, 1992
- Temin, Peter, "The causes of american business cycles: an essay in economic historiography", em Jeffrey Fuhrer & Scott Schuh (eds.), Beyond shocks : what causes business cycles?, Federal Reserve bank of Boston, 1998
- Vasko, T.; R. Ayres & L.Fontvieille (eds), Life cycles and long waves, (Berlin, Springer-Verlag, 1990), Lecture notes in economics and mathematical systems no. 340
- Westlund, Anders H. & S. Ohlen, "On testing for symmetry in business cycles", Empirical Economics, vol.16, abril de 1991, pp.479-502
- Zarnowitz, Victor, Business cycles : theory, history, indicators and forecasting, (Chicago, University of Chicago Press, 1991)
- Zarnowitz, Victor & P. Braun, "Major macroeconomic variables and leading indicators: some estimates of their interrelations 1886-1982", Working Paper 2812, NBER, 1992