

# Metodologia de estimação do PIB trimestral utilizando procedimentos de cointegração e filtros de Kalman\*

Luiz Fernando Cerqueira\*\*

Professor Adjunto do Departamento de  
Economia da Universidade Federal  
Fluminense (UFF)

## Resumo

*Neste artigo, apresenta-se uma metodologia para a estimação da série do PIB trimestral no Brasil, para o período 1960-96, a qual parte da estimação da equação estática do procedimento de cointegração de Engle-Granger (EG), envolvendo dados anuais do PIB e variáveis a ele relacionadas. Os coeficientes estimados da regressão são utilizados com os dados trimestrais dessas variáveis, para construir uma primeira estimativa da série do PIB trimestral. Em seguida, aprimoram-se esses resultados com procedimentos de filtro de Kalman. Estimam-se modelos univariados com valores omissos e processos de "benchmarking", bem como um modelo multivariado SURE com variáveis explicativas. Os modelos apresentam boas "performances" no que concerne à obediência das hipóteses do modelo gaussiano linear. A principal contribuição do artigo é apresentar uma metodologia consistente e original para a recuperação e/ou estimação e oferecer bases para o desenvolvimento de uma metodologia de previsão e/ou extrapolação do PIB brasileiro trimestral.*

## Palavras-chave

**Filtro de Kalman; construção e estimação de modelos; PIB.**

---

\* Artigo recebido em fev. 2007 e aceito para publicação em dez. 2007.

\*\* E-mail: lcer@uol.com.br.

O autor agradece os comentários e o apoio técnico de Adrian Pizzinga e Cristiano Fernandes, bem como as observações cuidadosas de dois pareceristas anônimos. Todos os erros remanescentes são de responsabilidade do autor. O autor é especialmente grato ao CNPq e à FAPERJ pelo apoio financeiro.

## ***Abstract***

*In this paper we present a methodology to estimate the Brazilian GDP's quarterly series from 1960 to 1996. Firstly, the suggested methodology applies Engle-Granger two-step cointegration procedure, involving annual GDP data and some GDP correlated variables. The estimated coefficient of the long-run relationship is used to create a quarterly equation with the GDP and its covariates. From this equation the first quarterly GDP estimate is obtained. The following step consists of improving this estimate with the state space framework, which use the Kalman filter. We estimated univariate models with missing values and benchmarking process, as well as a multivariate SURE with explanatory variables. The estimated model shows appropriate goodness of fit and the diagnostic tests for standardized innovations suggest the assumptions underlying the Gaussian linear model are valid for the data. The paper's major contribution is to present an approach for estimating the quarterly GDP series and developing a methodological framework for its forecasting.*

## ***Key words***

**Kalman filter; model construction and estimation; GDP.**

**Classificação JEL: C32, C51, E01.**

## **1 Introdução**

O objetivo deste artigo é descrever e aplicar a metodologia desenvolvida para a recuperação e a estimação do PIB trimestral no período 1960-96, utilizando procedimentos de cointegração combinados com métodos de espaço de estado em modelos estruturais de séries de tempo. Dentro desse esquema, estima-se uma relação de longo prazo, com frequência anual, entre o PIB e séries relacionadas, e aplica-se o vetor de cointegração obtido sobre os dados trimestrais dessas variáveis, para computar uma primeira estimativa do PIB. A série estimada é encadeada para trás, com a série oficial do PIB estimada pelo IBGE a partir de 1980. O passo seguinte é filtrar esses resultados com modelos estruturais em espaço de estado.

A motivação inicial era estimar a série do índice do PIB trimestral para o período 1960-79<sup>1</sup> e obter uma série completa para todo o período após 1960, com o intuito de utilizá-la em estudos posteriores, por exemplo, na estimação de uma demanda por moeda, para esse período ou para subperíodos, em que a renda possuía um papel central como variável determinante. Entretanto, na démarche empreendida no processo de estimação, apareceram dificuldades operacionais, relatadas na subseção 3.1, que conduziram à estimação para um período mais amplo e a uma metodologia "mais complexa", diferente da inicialmente pretendida, que se resumia aos procedimentos tradicionais de cointegração. Acabou-se chegando, assim, à escolha de modelos estruturais em estado de espaço.

Alguns estudos anteriores a este, utilizando enfoques diversos, procuraram sanar essa deficiência das estatísticas oficiais. Em geral, na ausência de estatísticas coletadas com o objetivo de computar o PIB, a montagem de estimativas trimestrais para o produto real agregado pode adotar três critérios: (a) um trabalho de interpolação dos dados anuais; (b) um levantamento próprio a partir de amostras de bens e serviços; e (c) uma combinação dos dois primeiros critérios (Contador; Santos Filho, 1987).

Existem diversas técnicas para a interpolação de séries, desde métodos simples, baseados em suposições de comportamento linear ou quadrático, até métodos mais elaborados, que lidam mais apropriadamente com valores omissos, a exemplo do ajuste aos dados de funções *spline* cúbicas e filtro de Hodrick-Prezcott, que utilizam suavizações exatas do filtro de Kalman. Há ainda procedimentos de controle ótimo de funcionais de custo quadráticas, que utilizam soluções recursivas, baseadas também no filtro de Kalman. Outros métodos adotam séries disponíveis em bases mensais ou trimestrais, que possuem forte relação com a variável objeto de interesse. No Brasil, a preferência dos pesquisadores recai sobre o uso da interpolação com séries relacionadas com o produto. Neste estudo, segue-se também este último caminho.

Na primeira etapa da metodologia, obteve-se uma estimativa do PIB trimestral a partir dos coeficientes de uma regressão, rodada com frequência anual, entre o PIB e a produção de veículos automotores e a de cimento, o consumo industrial de energia elétrica nas regiões do Rio de Janeiro e de São

---

<sup>1</sup> A partir de 1992, a série do índice do PIB trimestral não dessazonalizado passou a ser divulgada periodicamente pelo IBGE, tendo sido computada retroativamente desde 1980. Ela passou por diversas revisões. A utilizada neste artigo esteve disponível no *site* do IPEADATA e referia-se àquela calculada pelo IBGE com a "metodologia antiga". Atualmente, ela aparece, no *site*, encadeada com a série calculada pela metodologia nova, série esta que não está disponível no *site* do IBGE.

Paulo e a receita tributária real do Tesouro Nacional<sup>2</sup>, utilizando as séries em índices com base em 1980. Se as séries são todas integradas de mesma ordem, faz sentido realizar um teste de cointegração e utilizar o vetor de cointegração associado à equação do PIB como relação de longo prazo de interesse entre essas variáveis. Embora o tamanho da amostra seja um empecilho para essa estratégia, especialmente para a implementação do procedimento de Johansen, o teste de cointegração de Engle-Granger (EG), por ser menos exigente em termos de número de observações, é factível.<sup>3</sup> Assim, os resultados obtidos com a equação estática de Engle-Granger constituem uma informação *a priori* a ser utilizada. O índice do PIB trimestral é, então, computado com uma equação em que os coeficientes estimados da equação anual são multiplicados pelos índices trimestrais respectivos das séries relacionadas.

A natureza restritiva da já convencional metodologia raiz unitária/cointegração não se coaduna com o uso de modelos que admitem uma heterogeneidade temporal mais geral (Caporale; Pittis, 2002). No período em estudo, as mudanças foram endêmicas. Nesse caso, os processos econômicos são modelados mais adequadamente na forma de espaços de estado, que possuem uma maior flexibilidade para tratar de uma gama variada de problemas na análise de séries temporais. Nesta abordagem, assume-se que a evolução da série no tempo é determinada por uma série não observada de vetores  $\alpha_t$ 's aos quais está associada uma série de observações  $y_t$ 's. A relação entre os  $\alpha_t$ 's e os  $y_t$ 's é especificada por um modelo de espaço de estado. O propósito da análise de espaço de estado é inferir as propriedades relevantes dos  $\alpha_t$ 's a partir do conhecimento das observações. Nesses modelos, as observações são constituídas por tendência, sazonalidade, ciclos, componentes de regressão e mais um termo irregular. Os componentes são geralmente representados por formas de modelos de *random walk* e são moldados previamente, antes de serem colocados junto às observações no modelo de espaço de estado. Outra característica dos modelos de espaço de estado é que eles são flexíveis. Devido à natureza recursiva destes e das técnicas computacionais usadas para analisá-los, permitem, de imediato, mudanças conhecidas na estrutura do sistema ao longo do tempo (Durbin; Koopman, 2004).

---

<sup>2</sup> As informações sobre a obtenção e o tratamento dos dados utilizados neste artigo estão descritas detalhadamente em Cerqueira (2007).

<sup>3</sup> O que se deve ao fato de que defasagens — nesse caso, apenas da variável dependente — são exigidas somente no segundo estágio do procedimento, quando é aplicado um teste de raiz unitária sobre os resíduos da equação estática estimada no primeiro estágio. E, com frequência anual, o número de defasagens requerido é pequeno, embora os seus resultados devam ser tomados com cautela, porque é um teste, como o de Johansen, idealizado para grandes amostras.

Modelos de espaço de estado são bastante gerais, cobrindo uma ampla variedade de modelos. Observações multivariadas podem ser tratadas como extensões da teoria univariada. Pode-se permitir a presença de valores omissos. Variáveis explicativas podem ser incorporadas ao modelo. Além disso, se necessário, pode-se admitir que os coeficientes da regressão variem estocasticamente ao longo do tempo. Ajustes e variações de calendário também podem ser incluídos. Da mesma forma, observações mensais e trimestrais podem ser ajustadas, para terem consistência com observações anuais totais — esse processo de harmonização das estimativas trimestrais com o total anual recebe o nome de *benchmarking* e é, muitas vezes, utilizado por agências produtoras de dados oficiais.

A etapa seguinte da metodologia consiste, então, em aprimorar a estimativa inicial através de procedimentos de filtro de Kalman. Utilizam-se três modelos distintos para a decomposição em espaço de estado da série do PIB em seus componentes estruturais: um modelo univariado com valores omissos; um multivariado SURE, em que a obtenção do PIB filtrado por cada modelo, respectivamente, é feita sem ser adicionado o ruído gerado pelo termo irregular aos demais componentes; e um terceiro, univariado, com *benchmarking*, harmonizando os dados de frequência trimestral com os anuais, através da extração do erro de medida existente nos valores trimestrais. Até o momento da finalização da presente versão, não se tinha conhecimento de trabalhos semelhantes na literatura econômica brasileira.

O artigo está organizado da seguinte forma. Na seção 2, é feita uma revisão da literatura brasileira sobre estimações do PIB trimestral. Na seção 3, a metodologia é descrita em seus vários aspectos: na subseção 3.1, apresenta-se a primeira etapa da metodologia, que utiliza o procedimento de cointegração, na 3.2, é feita uma breve resenha sobre o assunto, junto com uma rápida apresentação dos modelos estruturais com *benchmarking*; e, na 3.3 são demonstrados os procedimentos da segunda etapa da metodologia. Na seção 4, são mostrados os resultados dos modelos estimados em espaço de estado. Na seção 5, oferece-se uma conclusão. No **Apêndice A**, apresentam-se os gráficos das séries empregadas, e, no **Apêndice B**, as matrizes desenhadas para se estimar o modelo de *benchmarking*.

## 2 Revisão da literatura

Na ausência de estatísticas oficiais coletadas com o objetivo de computar o PIB, a montagem de estimativas trimestrais para o produto real agregado pode adotar três critérios: (a) interpolação dos dados anuais, (b) levantamento

próprio a partir de amostras de bens e serviços e (c) combinação dos dois primeiros critérios (Contador; Santos Filho, 1987). No Brasil, a preferência dos pesquisadores recaiu sobre o uso da interpolação com séries relacionadas ao produto.<sup>4</sup>

Dentre os estudos que utilizam variáveis relacionadas com o PIB como forma de se obterem valores interpolados, cabe destacar-se o de Cardoso (1981), o qual serviu de base para análises posteriores na mesma linha metodológica, como em Rossi (1988), Pastore (1991), Pinotti (1991) e Cerqueira (1993). Nesse artigo, o autor obtém estimativas trimestrais para o PIB, no período 1966-79, a partir dos coeficientes de uma regressão do PIB anual contra a produção anual de cimento e de veículos automotores, o consumo industrial de energia elétrica no eixo Rio de Janeiro—São Paulo e a arrecadação tributária real do Tesouro Nacional. No entender de Contador e Santos Filho (1987), o trabalho de Cardoso (1981) possui deficiências no que concerne à excessiva participação da produção industrial e à imposição de que a sazonalidade do produto real estimado seja a mesma apresentada pela combinação das variáveis. Esses autores, na construção de sua série de PIB trimestral para o período 1973-85, combinam amostragem com interpolação e utilizam séries desagregadas para o produto dos setores agropecuária, indústria, comércio e transportes. Contudo a metodologia empregada é diferente daquela utilizada pelo IBGE. As séries desagregadas usadas diferem, assim como os pesos associados a cada uma, de modo a compor a série agregada.<sup>5</sup>

Além dos problemas apontados, os dados trimestrais, obtidos através dessa metodologia, apresentam outros dignos de nota. Há descontinuidade dos dados utilizados para os setores agropecuária, comércio e para o produto do Governo, assim como dificuldade em obtê-los. Há também um certo empirismo na escolha das *proxies* das variáveis utilizadas, o que torna o procedimento árduo de se replicar e estender para períodos maiores e anteriores a 1974. Por essas razões, esse procedimento não será empregado.

Embora, à primeira vista, as estimativas que utilizam variáveis relacionadas ao PIB possam falhar, por não empregarem um número de variáveis suficientemente representativo de setores produtivos da economia, uma combinação linear dessas variáveis disponíveis pode constituir uma boa *proxy* do PIB. Assim sendo, é lícito que se utilize como ponto de partida da metodologia

---

<sup>4</sup> Para uma análise detalhada, ver Contador e Santos Filho (1987).

<sup>5</sup> Nakane (1994) procura corrigir essas deficiências, utilizando as mesmas ponderações que o IBGE, de modo a tornar as séries comparáveis. Porém, como sugerido pelo próprio autor, a série estimada apresenta distorções, devido às diferenças do padrão de sazonalidade entre

de recuperação e estimação do PIB uma estimativa inicial, obtida por essa técnica baseada na estimação do PIB trimestral através de variáveis relacionadas a ele.

## 3 Metodologia

### 3.1 Primeira etapa: recuperação e estimação do PIB trimestral através de técnicas de cointegração

A idéia inicial era fazer a estimativa do PIB trimestral apenas para o período anterior a 1980.<sup>6</sup> A metodologia de estimação partiria do teste de cointegração entre o PIB e as variáveis a ele relacionadas, com frequência anual, utilizando o procedimento de Johansen. Os coeficientes do termo de correção de erros estimado seriam, então, aplicados sobre as séries em frequência trimestral, gerando a série do PIB. As variáveis incluídas seriam: PIB, produção de veículos automotores e de cimento, consumo industrial de energia elétrica nas regiões do Rio de Janeiro e de São Paulo, receita tributária real, produção de petróleo e gás e de aço em lingotes. As séries estão expressas em índices com base em 1980.

No entanto, o pequeno tamanho da amostra (16 observações) revelou-se, desde logo, um entrave. O reduzido número de graus de liberdade impediu a colocação de um número maior de defasagens para a correção da correlação serial nos resíduos, o que invalidava o teste de cointegração, o qual, ademais, requer "grandes amostras". Além disso, a possível presença de séries  $I(2)$  no período 1965-80 poderia estar causando o excesso de correlação serial. A solução encontrada para dar continuidade aos experimentos foi reduzir o número de variáveis e aumentar o tamanho da amostra. Assim, alongou-se a amostra até o limite da disponibilidade de dados, a qual passou, então, a cobrir o período 1960-96.<sup>7</sup> O número de variáveis foi reduzido para cinco. A produção de aço foi retirada,

---

<sup>6</sup> O método de relacionar séries indicadoras à série anual do PIB, o *benchmark*, não representa um processo de *benchmarking* estrito senso, embora resolva um problema de valores omissos existente nas estatísticas oficiais do período 1960-79. O procedimento descrito nesta seção inspira-se em Chow e Lin (1971).

<sup>7</sup> As séries de energia elétrica e receita tributária só estão disponíveis a partir de 1960. O truncamento em 1996 deveu-se ao fato de a série do índice do PIB trimestral, pela metodologia antiga do IBGE, ter sido interrompida no terceiro trimestre de 1997. Para não se ter que trabalhar com fração de ano, a amostra foi truncada no final de 1996.



por ser intensiva em eletricidade, o que estava acarretando inversão de sinais nos coeficientes do mecanismo de correção de erros (ECM). Além disso, sua presença aumentava a correlação serial dos resíduos. A produção de petróleo, por sua vez, foi retirada em função do seu crescimento muito acelerado após 1980, o que levou a série a se afastar das demais.

Os estudos realizados com o procedimento de Johansen para o período 1960-96 não apresentaram resultados aceitáveis. Embora os testes de raiz unitária tenham indicado que as cinco séries utilizadas são  $I(1)$ , os resíduos do vetor de correção de erros (VECM) apresentaram elevada correlação serial, que não foi possível corrigir com um polinômio de lags de ordem cinco, número máximo imposto pelo tamanho da amostra ( $T = 35$ ). Além disso, o VECM não mostrou boas características de estabilidade, apresentando, mesmo após a restrição sobre o espaço dos vetores de cointegração, raízes muito próximas a um. Assim, essa estratégia teve de ser abandonada.<sup>8</sup>

Um procedimento alternativo é o teste de dois estágios de Engle-Granger. Nessa hipótese, a equação estática é usada para fornecer os coeficientes de longo prazo, e o teste de raiz unitária aplicado sobre os resíduos da equação indica a existência de uma relação estável de longo prazo entre as variáveis.

Os testes realizados estão expostos no Quadro 1. O índice do PIB é a variável dependente. Do lado direito da equação aparece um termo constante (C), uma tendência (TT) e os índices de consumo de energia elétrica (IEES), de produção de veículos (IAUTO), da receita tributária real (IRTNRS) e da produção de cimento (ICIM). Na parte inferior do Quadro 1, estão as estatísticas  $R^2$ , DW e SER para a equação estática e o resultado do teste de raiz unitária dos resíduos junto com os *p-values* dos testes *portmanteau* para a presença de correlação serial e o valor do critério de informação bayesiano.

No nível de 10%, rejeita-se a hipótese nula de não cointegração entre as variáveis. O teste apresenta resíduos sem correlação serial. Como esperado, o  $R^2$  está bem próximo de **um**, caracterizando um ajuste perfeito e eliminando a necessidade de correção para o viés de equações simultâneas ou correlação serial (Engle; Granger, 1991).<sup>9</sup> Ademais, a multicolinearidade existente entre as

---

<sup>8</sup> Cabe mencionar-se que, nos experimentos realizados, algumas variáveis apareceram com sinais negativos, o que se deveu provavelmente à colinearidade existente entre certas séries, como energia elétrica com aço e cimento. Entretanto as séries do PIB estimadas com os coeficientes do ECM apresentaram-se semelhantes às obtidas com a estática de Engle-Granger.

<sup>9</sup> Contudo sempre se pode argumentar sobre a presença de regressores endógenos, que tornariam as estimativas linearmente viesadas, com distribuição limite assimétrica e contendo parâmetros de *nuisance*. A solução dessas deficiências, proposta por Phillips e Hansen



variáveis colaborou no sentido de se obter esse ajuste. Por outro lado, os estudos sobre os componentes sazonais estruturais através de modelos em espaço de estado e de Holt-Winters indicam que a série estimada e a do IBGE possuem padrões de sazonalidade muito semelhantes, cujas amplitudes crescem progressivamente, com a série estimada apresentando uma amplitude inferior à do IBGE após 1980. Isso não tem implicações para este estudo, dado que a série que se irá usar assume, entre 1980 e 1996, os valores da própria série do IBGE.

#### Quadro 1

##### Teste de cointegração de Engle-Granger da variável dependente PIB

VARIÁVEIS	COEFICIENTES
C (1)	0,112423
TT (2)	0,007836
IEES (3)	0,283306
IAUTO (4)	0,108377
IRTNRS (5)	0,098286
ICIM (6)	0,231170
$R^2 = 0,9984$   $DW = 1,5793$   $SER = 0,0164$   $t = -4,6341 \uparrow(7)$   $Lags = 0$	
$Q(1) = 0,931$   $Q(4) = 0,228$   $Q(8) = 0,479$   $Q(12) = 0,126$   $SIC = -5,4545$	

(1) Termo constante. (2) Tendência. (3) Índice de consumo de energia elétrica. (4) Índice de produção de veículos. (5) Índice da receita tributária. (6) Índice da produção de cimento. (7) O símbolo  $\uparrow$  representa rejeição da hipótese nula de uma raiz unitária no nível de significância de 10%.

Talvez a representatividade das variáveis utilizadas seja questionável, porém um pouco de reflexão indica que elas representam, de forma adequada, os principais setores produtivos da economia. De fato, o setor automobilístico é, desde os anos 60, o carro-chefe da economia. O consumo de cimento, por sua vez, representa bem o nível de atividade da construção civil, responsável pela geração de grande parte do emprego na economia. Da mesma forma, o consumo industrial de energia elétrica na principal região econômica do País é um indicador relevante do nível de atividade da economia. Por fim, é evidente que a receita

---

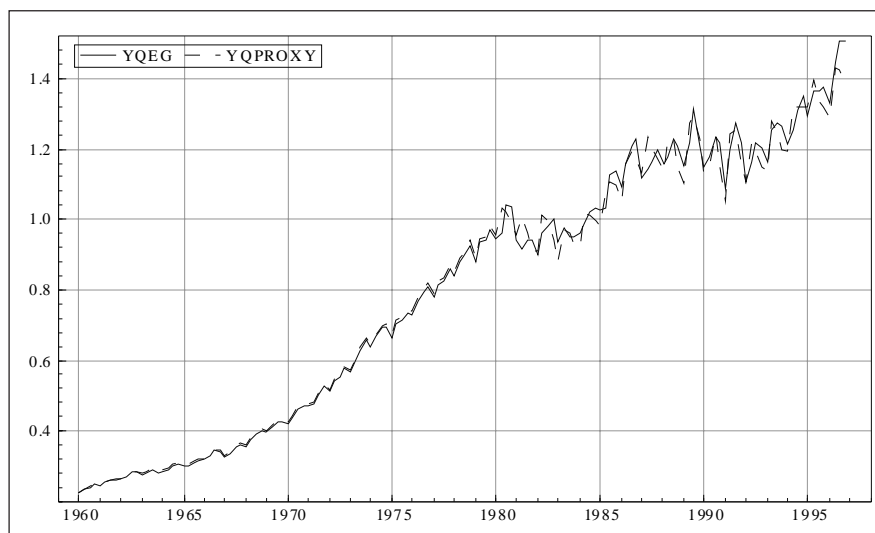
(1990), é utilizar os estimadores *fully modified* LS. Os experimentos realizados com esse procedimento e estatisticamente válidos não apresentaram resultados diferentes dos obtidos com as estimativas de EG. Assim sendo, não há por que se preocupar com endogeneidade.

tributária federal representa adequadamente as atividades do setor governamental, especialmente quando se tem em conta que parte dela é repassada a estados e municípios e que ela é inteiramente gasta, pelo menos em grande parte do período de interesse.<sup>10</sup>

Sendo assim, a série obtida como uma combinação linear dessas variáveis, em que os escalares são os coeficientes da equação acima, fornece uma *proxy* inicial bastante razoável para o índice do PIB trimestral, especialmente porque ela recupera os gaps existentes nas estatísticas do período 1960-79, onde as observações diretas não estão disponíveis. Na Figura 1, aparece a série estimada com os coeficientes da equação de Engle-Granger para o período (YQEG), juntamente com essa série encadeada (para trás), a partir do primeiro trimestre de 1980 com a série do IBGE (YQPROXY). É esta última que será filtrada na segunda etapa da metodologia proposta e será denominada proxy do PIB — seja em índice, seja em valor.

Figura 1

Índice do PIB trimestral — 1960-96



NOTA: 1. YQEG é a série estimada com os coeficientes da equação de Engle-Gayer.  
2. YQPROXY é a série do IBGE.

<sup>10</sup> Os dados sobre o setor de transportes não foram incluídos, por não se ter obtido uma série de tamanho adequado.

## 3.2 Modelos estruturais em espaço de estado<sup>11</sup> com *benchmarking*

Um problema comum em estatísticas oficiais é o ajustamento de observações mensais e trimestrais, obtidas através de levantamentos ou amostragem e sujeitas a erros, aos dados anuais totais, obtidos através de censos ou levantamentos mais detalhados e que são, por hipótese, assumidos livres dos erros de amostragem. O total anual é chamado de *benchmark*; e o processo de harmonização das estimativas com o total anual, de *benchmarking*. Mais especificamente, *benchmarking* — "calendarização" ou harmonização — é um processo que consiste em combinar de forma "ótima" duas fontes de medição de uma mesma série temporal, geralmente obtidas a partir de frequências distintas, onde a série de frequência mais baixa, a série de *benchmark*, é assumida como mais confiável estatisticamente. No processo de *benchmarking*, procura-se ajustar a série de frequência mais alta à série *benchmark*.

Existem duas metodologias principais de aplicação do *benchmarking* a uma série temporal: um *approach* puramente numérico e outro de modelagem estatística. O *approach* numérico difere do segundo, por não especificar um modelo estatístico que se suponha que a série obedeça. O *approach* numérico abrange a família dos métodos baseados na minimização de uma soma quadrática, propostos por Denton (1971)<sup>12</sup> — que segue o princípio da preservação de movimentos —, Bassi (1958) e Ginsburgh (1973). O método de modelagem estatística, por sua vez, envolve os modelos baseados em processos Autoregressive Integrated Moving Average (ARIMA), propostos por Hillmer e Trabelsi (1987), os modelos de espaço de estado, propostos por Durbin e Queenneville (1997), e os modelos que usam um conjunto de regressões, tais como os de Cholette e Dagum (1994) e Mian e Laniel (1993), assim como as referências citadas por esses autores.<sup>13</sup>

Neste artigo, faz-se uso de um procedimento de *benchmarking* que utiliza modelos estruturais, onde uma série temporal é decomposta em componentes

---

<sup>11</sup> Para o leitor interessado em modelos de espaço de Estado e filtro de Kalman, ver Durbin e Koopman (2004) e Harvey (1990).

<sup>12</sup> Para uma aplicação desse procedimento, ver Di Fonzo e Marini (2003).

<sup>13</sup> Em adição, Chow e Lin (1971) propuseram um modelo multivariado de mínimos quadrados generalizados para interpolação, distribuição e extrapolação de séries temporais. Embora não sendo um método de *benchmarking* em um sentido estrito, o método proposto está relacionado ao de modelagem estatística, particularmente ao último grupo.

de tendência, sazonalidade e termo irregular.<sup>14</sup> A seguir, alinhavam-se os passos seguidos — em ordem — por Durbin e Koopman (2004) e Durbin e Queenneville (1997).

Define-se um **modelo em espaço de estado (EE) linear gaussiano** pelas seguintes equações:

$$\begin{aligned} y_t &= Z_t \alpha_t + \varepsilon_t, & \varepsilon_t &\sim N(0, H_t), \\ \alpha_{t+1} &= T_t \alpha_t + R_t \eta_t, & \eta_t &\sim N(0, Q_t), \\ \alpha_1 &\sim N(a_1, P_1) \end{aligned} \quad (1)$$

onde  $y_t$  é um vetor de observações  $p \times 1$ ;  $\alpha_t$  é chamado de **vetor de estado**, é não observável e de dimensão  $m \times 1$ ;  $\varepsilon_t$  e  $\eta_t$  são termos de erro independentes; as matrizes do sistema  $Z_t, T_t, R_t, H_t$  e  $Q_t$ , em geral, contêm parâmetros desconhecidos, os quais se aglutinam em um **vetor de parâmetros  $\psi$** . A estimação do modelo em (1) é efetuada pelo **filtro de Kalman** — vetor de estado — combinado à **máxima verossimilhança** — vetor de parâmetros. O FK é formado por um conjunto de equações que estimam, recursivamente no tempo, a média e a variância condicional do vetor de estado. Detalhes sobre essas equações, sobre suas deduções respectivas e sobre a combinação destas com a estimação de máxima verossimilhança encontram-se em Harvey (1990, cap. 3) e em Durbin e Koopman (2004, cap. 4 e 7).

Por questões de facilidade de exposição, suponha-se que as observações sejam mensais e obtidas em *surveys* realizados também mensalmente ( $s = 12$ ). Elas são denotadas por  $y_t$  e os verdadeiros valores que se deseja estimar, por  $y_t^*$ , para  $t = 12(i-1) + j$ ,  $i = 1, \dots, l$  e  $j = 1, \dots, 12$ , onde  $l$  é o número de anos. Assim, o erro dos *surveys* é  $e_t = y_t - y_t^*$ , denotado por  $\sigma_t^s \xi_t^2$ , onde  $\sigma_t^s$  é o desvio-padrão do erro de *survey* no tempo  $t$ . O erro é modelado como um processo AR(1) com variância unitária. Em princípio, processos Autoregressive Moving Average (ARMA) de ordem mais alta podem ser usados. Assuma-se que os valores de  $\sigma_t^s$  estejam disponíveis a partir da prática dos *surveys* e que os erros sejam livres de viés.

Os valores de *benchmark* são dados por  $x_t = \sum_{j=1}^{12} y_{12(i-1)+j}^*$  para  $i = 1, \dots, l$ . Suponha-se que os valores anuais estejam disponíveis para to-

<sup>14</sup> Caminhando em direção oposta, Moauro e Savio (2005) apresentam um procedimento para desagregação de séries temporais observadas em uma dada frequência para uma outra de frequência maior. O método sugerido é um modelo multivariado SURE, estimado com filtro de Kalman.

dos os anos de estudo; assim, na prática, os valores censitários terão, em geral, uma defasagem de um ano ou dois em relação às observações dos *surveys*. Como modelo da série de observações, usa-se

$$y_t = \mu_t + \gamma_t + \varepsilon_t + \sigma_t^s + \xi_t^s \quad t = 1, \dots, 12l \quad (2)$$

onde  $\mu_t$  é a tendência e  $\gamma_t$  é o termo sazonal.

A série é organizada da seguinte forma:

$$y_1, \dots, y_{12}, x_1, y_{13}, \dots, y_{24}, x_2, y_{25}, \dots, y_{12l}, x_l$$

Considere-se que o instante na série em que o *benchmark*  $x_t$  ocorre é  $t = (12i)'$ ; assim, o instante  $t = (12i)'$  ocorre na série entre  $t = 12i$  e  $t = 12i + 1$ . Tomando o modelo de *random walk* integrado para o componente de tendência e o de *quasi-random walk* para a componente sazonal, tem-se

$$\begin{aligned} \Delta^2 \mu_t &= \zeta_t, & \zeta_t &\sim NID(0, \sigma_\zeta^2), \\ \gamma_t &= -\sum_{j=1}^{11} \gamma_{t-j} + \omega_t, & \omega_t &\sim NID(0, \sigma_\omega^2) \end{aligned} \quad (3)$$

É conveniente colocarem-se os erros das observações — de medida — dentro do vetor de estado. Assim tem-se

$$\alpha_t = (\mu_t, \dots, \mu_{t-11}, \gamma_t, \dots, \gamma_{t-11}, \varepsilon_t, \dots, \varepsilon_{t-11}, \xi_t^s)'$$

Dessa forma,  $y_t = Z_t \alpha_t$ , onde,

$$Z_t = (1, 0, \dots, 0, 1, 0, \dots, 0, 1, 0, \dots, 0, \sigma_t^s) \quad t = 1, \dots, n,$$

e  $x_i = Z_t \alpha_t$ , onde,

$$Z_t = (1, \dots, 1, 0, \dots, 0, 1, \dots, 1, 0) \quad t = (12i)'$$

para  $i = 1, \dots, l$ . Pode-se, então, escrever a transição de  $\alpha_t$  para  $\alpha_{t+1}$  entre  $t = 12i - 11$  e  $t = 12i - 1$ . De  $t = 12i$  para  $t = (12i)'$ , a transição é a identidade. De  $t = (12i)'$  para  $t = 12i + 1$ , a transição é a mesma que de  $t = 12i - 11$  para  $t = 12i - 1$ .

A matriz de transição  $T_t$  é uma matriz bloco-diagonal com elementos

$$\begin{bmatrix} 2 & -1 & 0_{1,10} \\ \mathbf{I}_{11} & 0_{11,1} \end{bmatrix}, \begin{bmatrix} \mathbf{I}_{11}' & 0_{1,1} \\ \mathbf{I}_{11} & 0_{11,1} \end{bmatrix}, I_k, \begin{bmatrix} 0_{1,12} \\ \mathbf{I}_{11} & 0_{11,1} \end{bmatrix}, \phi \quad \text{para } t = 1, \dots, n$$

e é igual a  $I_{37}$ , quando  $t = (12i)'$ ,  $i = 1, \dots, l$ . O vetor de erros de estado é  $v_t = [\xi_t, \omega_t, \varepsilon_t, \kappa_t]$ , se  $t = 12i + 2, \dots, 12(i + 1)$  e  $v_t = 0$  para  $t = (12i)'$ ,  $i = 1, \dots, l$ . Os demais termos do modelo estão descritos acima.

Assim, emprega-se uma forma modificada do modelo de espaço de estado introduzido no começo da seção, no qual a variância do erro da equação de observações é zero para  $t = 1, \dots, n$  e é igual à variância do *benchmark* para  $t = (12i)'$ ,  $i = 1, \dots, l$ . Supõe-se também que a matriz de erros de *benchmark* seja uma matriz diagonal, porque, de outra forma, o vetor de estado se tornaria muito grande. A equação de transição é, então, dada por

$$\alpha_t = T_t \alpha_{t-1} + R_t v_t, \quad t = 1, \dots, 12, (12)', 13, \dots \quad (4)$$

onde  $R_t$  é a matriz de seleção formada apropriadamente pelas colunas da matriz identidade  $I_{37}$ .

### 3.3 Segunda etapa: procedimentos de estimação em espaço de estado<sup>15</sup>

A partir dos resultados apresentados na subseção 3.1, os componentes estruturais do PIB trimestral foram extraídos, utilizando-se modelos de espaço de estado univariados e multivariados. Foram estimados três modelos estruturais univariados: um com valores omissos, um com processo de *benchmarking* e um terceiro com *missings* e *benchmarking*. Além deles, estimou-se também um modelo estrutural multivariado — SURE — uma generalização do modelo estrutural básico.

As estimações foram feitas para o período 1960-96, com as séries em índices nos modelos univariados e com valores omissos e multivariados. Neste último, usaram-se as mesmas variáveis do modelo da subseção 3.1. Nos modelos com processo de *benchmarking*, utilizaram-se as séries em quantidades, e o PIB foi estimado em valores reais de 1980.<sup>16</sup> Neste último, a soma dos trimestres de cada ano é diferente da do PIB anual computado oficialmente; para as demais séries, o total anual é igual à soma das quantidades e/ou valores trimestrais.

<sup>15</sup> Uma síntese da metodologia é apresentada em Cerqueira (2007).

<sup>16</sup> A série do PIB trimestral em valores constantes de 1980 é calculada multiplicando-se a série de índice pelo valor do PIB médio trimestral de 1980.

Os experimentos preliminares com modelos estruturais básicos univariados e multivariados em frequência trimestral indicaram a presença de uma forte heteroscedasticidade, que, se não tratada, poderia comprometer os resultados estimados. A heteroscedasticidade no período 1960-96 tem relação com o comportamento da taxa de crescimento da economia: um longo período de crescimento, de 1960 a 1980, seguido de um período de taxas baixas de crescimento, oscilantes e sem tendência definida. Ela também guarda relação com o comportamento da taxa de inflação<sup>17</sup>. A partir de 1973, a taxa de inflação entrou numa rota ascendente, intercalada com períodos de quedas violentas, derivadas dos congelamentos de preços. A partir de meados de 1994, com o Plano Real, ela caiu abruptamente, seguindo uma rota declinante.

A solução para o problema passa pelos mínimos quadrados ponderados. O procedimento é tratar a heteroscedasticidade através da normalização das séries, com uma série que seja proporcional aos recíprocos dos desvios-padrão dos erros. Uma série adequada seria, então, aquela que guardasse relação com o PIB e com a inflação. A receita tributária, pela teoria, deveria guardar uma relação de endogeneidade com o PIB e a inflação. No entanto, no caso da economia brasileira, diversas modificações na tributação foram efetuadas durante o período, o que conduziu a um aumento progressivo e exógeno da receita tributária real. A opção como fator de ponderação recaiu sobre o índice<sup>18</sup> da receita tributária real.<sup>19</sup> Como a série está disponível mensalmente, calculou-se a receita do trimestre e, a partir dela, a série de **índice**. Com isso, a receita tributária real deixou de ser uma variável isolada no modelo e passou a compor combinações com as demais. No que concerne ao PIB, essa série — em valor ou em índice —, dividida pelo índice da receita tributária, passa a ser a variável a ser estimada — PIB ponderado. O PIB estimado é recuperado, multiplicando-se o PIB ponderado pelo índice da receita.

Experimentos com a estimação de modelos estruturais básicos com a série ponderada do PIB indicaram uma falha no processo de estimação dos componentes estruturais, decorrente do surgimento de *spikes* antes inexistentes

---

<sup>17</sup> A taxa de inflação é definida como a variação percentual do IGP-DI da FGV.

<sup>18</sup> A razão do uso da variável em índice é que a receita real expressa em reais de 1980 assume valores muito baixos, devido às conversões monetárias causadas pela inflação, o que faz os valores divididos pela série explodirem, tornando o procedimento estatístico e computacional pouco manejável.

<sup>19</sup> A taxa de inflação seria uma ótima candidata, não fossem os congelamentos de preços ocorridos durante o período, os quais levaram a taxa de inflação próxima a zero. Assim, a divisão de uma série por ela tem como resultado a geração de *outliers* de difícil tratamento.



na série. A solução foi estimar-se um modelo estrutural<sup>20</sup> para a série de receita real e extrair-se os *outliers* identificados na série de resíduos auxiliares da equação de medida e a quebra observada nos resíduos da equação de estado.<sup>21</sup> Dessa forma, para corrigir a heteroscedasticidade identificada, utiliza-se a série de receita tributária ajustada, devidamente tratada com *dummies* de política e de intervenção. As novas séries ajustadas são plotadas no **Apêndice A**, na Figura A.2.

Para efeito das estimações, os parâmetros foram tomados como tendo densidade, a *priori*, difusa. Assim, não foram supostos valores a *priori* da matriz de covariância, embora *priors* para os coeficientes das variáveis explicativas tenham sido utilizados nas estimações feitas com valores omissos.

O primeiro modelo estimado (PIB1) considera a presença de valores omissos na série do PIB trimestral, no período 1º trim./60-4º trim./79. O modelo utilizado foi o estrutural básico com tendência, mas sem inclinação, com *dummies* sazonais estocásticas, termo irregular e variáveis explicativas.<sup>22</sup> A variável do lado esquerdo da equação de medida é a série de índice do PIB ponderado ajustado com valores observados de 1980 a 1996 e com valores omissos de 1960 a 1979, os quais se deseja estimar. Se o período anterior a 1980 for inteiramente coberto com valores omissos, o pacote não realiza adequadamente as recursões. Então, da série estimada do PIB, selecionaram-se aleatoriamente 50% da amostra para assumir valores omissos, para os trimestres anteriores a 1980.<sup>23</sup> Por construção, esses valores omissos foram substituídos, através do algoritmo de suavização, por valores interpolados, consistentes com a estrutura de componentes do modelo. As variáveis explicativas — veículos, cimento e energia elétrica —, devidamente ajustadas, entram no lado direito da equação de medida,

---

<sup>20</sup> O modelo utilizado inclui um componente estrutural de tendência sem inclinação, um ciclo de curta duração, *dummies* sazonais trigonométricas e um termo irregular.

<sup>21</sup> O critério para extração foi o resíduo normalizado assumir um valor maior do que 3.5. Foram feitas intervenções no 3º trim./80, no 1º trim./86, no 4º trim./86, no 3º trim./89, no 2º trim./90, no 4º trim./91 e no 1º trim./92, no termo irregular, e uma no nível da série, no 4º trim./89.

<sup>22</sup> O modelo foi estimado em Eviews 5.1, que usa o procedimento descrito em Koopman, Shephard e Doornik (1999) para estimações com *missings*. Vale a pena ter em conta que o pacote funciona “mais” apropriadamente, quando são fornecidos *priors* dos valores iniciais, matriz de covariância e valores dos coeficientes do sistema.

<sup>23</sup> A seleção foi realizada utilizando 61 pseudonúmeros aleatórios extraídos de uma normal (0,1). Desses números, as duas primeiras casas decimais em módulo eram usadas para escolher uma posição entre 1 e 80, à qual se atribuía um *missing*. Foram necessários 61 números para conseguir selecionar 40, já que alguns eram repetidos; e outros, superiores a 80. Uma *random seed* foi utilizada.

apropriadamente estacionarizadas,<sup>24</sup> conforme recomendado por Durbin e Koopman (2004). O PIB reportado não inclui o componente irregular, por ser este uma estimativa do distúrbio aleatório. Assim, sem esse termo, o PIB estimado fornece uma informação sobre a capacidade de os componentes de tendência e sazonalidade estocástica explicarem o seu comportamento.

O segundo modelo univariado (PIB2) é o de *benchmarking*, que torna o PIB trimestral "consistente" com o PIB anual calculado pelo IBGE, ao corrigir os erros de medida. Supõe-se que o erro de medida siga um processo AR(1). O modelo foi estimado com o PIB em valores ponderados pelo índice médio do ano da receita tributária.<sup>25</sup> O modelo de estimação foi o gaussiano linear geral,<sup>26</sup> com tendência local, *dummies* sazonais e sem variáveis explicativas. O modelo obedece a estrutura do modelo com processo *benchmarking*, tal como descrito na subseção 3.2. As matrizes  $Z$ ,  $T$ , e  $R$ , construídas para operacionalizar as estimações com *benchmark*, estão apresentadas no **Apêndice B**. O primeiro grupo de matrizes refere-se à estimação do modelo univariado reportado na seção 4. Essas matrizes, escritas para os dados trimestrais e totais anuais, também constituem o primeiro passo para a implementação do modelo multivariado com quatro variáveis dependentes, que não chegou a ser estimado. O segundo grupo de matrizes refere-se apenas ao modelo multivariado.

Procedeu-se também a um exercício de modelo *benchmarking* contendo valores omissos (PIB3) — dado que o modelo trabalha naturalmente com a presença de observações omissas, embora não em grandes proporções da amostra. Esses valores foram concentrados no período anterior a 1980 e escolhidos pela sua atipicidade.<sup>27</sup> As demais características do modelo são as descritas acima. Em ambos os modelos, o PIB reportado não inclui o erro de medida.

---

<sup>24</sup> O teste clássico para a presença de raízes unitárias sazonais, HEGY, em geral rejeita a hipótese nula e, tal como apontado em Maddala e Kim (2002), apresenta várias deficiências. Por esse teste, concluir-se-ia que não existem raízes unitárias sazonais em nenhuma das variáveis explicativas utilizadas. A solução para detectar e tratar um componente sazonal estocástico foi pesquisar o comportamento da série através de modelos de espaço de estado. Coincidentemente, o mesmo modelo foi ajustado para todas as séries. Assim, foi estimado um modelo com tendência linear local, sazonalidade estocástica e termo AR(1). Nas séries de veículos e energia elétrica, o termo AR(1) estacionário dominou o componente irregular; com isso, este último foi retirado. Nesses casos, as séries estacionarizadas são representadas pelo componente AR(1). No caso da série cimento, a série estacionarizada é representada pelo termo irregular.

<sup>25</sup> O uso da média anual tem como objetivo preservar a relação de *benchmarking* das séries, à exceção do PIB.

<sup>26</sup> A estimação foi feita em SsfPack3.

<sup>27</sup> Foram retirados da amostra e substituídos por valores omissos os PIBs do período 1º trim.-4º trim./66, 4º trim./73, 1º trim./75, 1º trim./76, 1º trim./79 e 1º trim.-2º trim./94. Os dois últimos foram removidos por terem exatamente o mesmo valor.

Cabe observar que a solução encontrada para tratamento da heteroscedasticidade não implicou não-linearidades nos moldes descritos por Durbin e Queenneville (1997). O que se fez foi transformar variáveis linearmente e estimar o modelo, de tal maneira que os valores trimestrais fossem corrigidos para que a soma dos trimestres fosse igual ao total anual. O procedimento de tratamento da heteroscedasticidade aqui assumido é uma alternativa ao modelo multiplicativo proposto por Durbin e Queenneville (1997), o qual incorre em não-linearidades. Ademais, a metodologia proposta é computacionalmente mais eficiente, além de permitir que se averigüe, através das inovações padronizadas, se a heteroscedasticidade foi tratada corretamente.

O último modelo estimado foi um multivariado, em que as quatro variáveis — PIB, veículos, cimento e energia elétrica — estão endogeneizadas. As séries são usadas em índices, divididas previamente pelo índice da receita tributária ajustada de *outliers* (Figura A.2). A especificação envolve uma tendência local para cada variável, *dummies* sazonais na forma trigonométrica — porque, no modelo, lidaram melhor com a correlação serial presente — e variáveis explicativas estacionarizadas, também para tratamento da correlação serial (Durbin; Koopman, 2004). No caso, a escolha recaiu sobre a série de automóveis com seis defasagens. O interesse deste último modelo é exatamente apresentar a extração dos componentes estruturais em um ambiente multivariado e comparar com os resultados dos modelos univariados. O PIB estimado reportado também não inclui o componente irregular, permitindo que os componentes de tendência e sazonalidade estocástica de cada variável expliquem o seu comportamento.

## 4 Resultados

Nas Tabelas 1 e 2, são apresentados os testes de diagnósticos dos modelos estimados descritos na seção anterior.<sup>28</sup> Todos os testes foram feitos com a série de resíduos padronizados do modelo, as chamadas inovações (Durbin; Koopman, 2004). O modelo PIB1 refere-se ao modelo com *missings*; o PIB2, ao modelo com *benchmarking*; e o PIB3, ao modelo com *missings* e *benchmarking* (Tabela 2). O modelo PIB4 é o multivariado estrutural (Tabela 3).

---

<sup>28</sup> Sobre a notação, cabem algumas informações. P. E. V. é a predição da variância do erro. DH é o teste de normalidade de Doornik-Hansen com distribuição aproximada  $\chi^2$ . H(h) é o teste para presença de heteroscedasticidade de Goldfeld-Quandt, com distribuição aproximada F(h,h). Os coeficientes de determinação são reportados para as séries em níveis, para diferenças e para diferenças sazonais (RS2). Por fim, são reportados os critérios de informação de Akaike e Schwartz (Bayes).

Tabela 1

Testes de diagnósticos dos modelos com *missing* (PIB1), com *benchmarking* (PIB2) e com *missing e benchmarking* (PIB3)

ESTATÍSTICAS	PIB1(1)	PIB2	PIB3
LogL .....	110 236	229 262	203 655
Iterações .....	(2)16	(2)64	(2)61
P. E. V. ....	0,0030	...	...
Erro-padrão .....	0,0549	...	...
DH .....	(3) 0,0008 [0,9996]	21,895 [0,0000]	(3)67,648 [0,0000]
H(h) .....	(4) 1,2269 [0,2430]	(5) 0,7382 [0,8747]	(6) 1,3209 [0,1548]
DW .....	1,9540 [0,7576]	1,8561 [0,3278]	1,7346 [0,0863]
Q(12) .....	16,505 [0,1234]	12,964 [0,3717]	16,522 [0,168]
CumPer (K-S) (7) ....	0,0877	0,1544	0,1134
R2 .....	0,9377	...	...
RD2 .....	0,7307	...	...
AIC .....	-1,9483	-5,3337	-5,2086
SIC (Bayes) .....	-1,7985	-5,1729	-5,0406

NOTA: Os números entre colchetes são os p-valores correspondentes das estatísticas.

(1) Neste modelo, há 107 observações válidas. (2) Interação forte. (3) O teste utilizado foi o de Bera-Jarque, que também tem distribuição aproximada  $\chi^2$ . (4) h = 47. (5) h = 58. (6) h = 54. (7) K – S (0,10) = 0,1078.

Tabela 2

## Testes de diagnósticos do modelo multivariado PIB4

ESTATÍSTICAS	MODELO MULTIVARIADO PIB4			
	PIB	CIM	EE	AUTO
LogL .....	1 457,73	1 457,73	1 457,73	1 457,73
Iterações .....	(1) 38	38	38	38
P. E. V. ....	0,0088	0,0068	0,0103	0,0117
Erro-padrão .....	0,0938	0,0825	0,1017	0,1079
DH .....	2,8403	4,2242	0,6834	0,2926
	[0,2417]	[0,1210]	[0,7105]	[0,8639]
H(46) .....	0,5234	0,6467	1,0551	1,3326
	[0,9848]	[0,9285]	[0,4282]	[0,1668]
DW .....	1,8763	1,9279	1,8781	1,8263
	[0,4611]	[0,6675]	[0,4677]	[0,3007]
Q(10) .....	13,667	10,423	13,986	9,2378
	[0,0909]	[0,2366]	[0,0821]	[0,3226]
CumPer (K-S) ...	0,1991	0,1181	0,2270	0,1537
R2 .....	0,8112	0,6041	0,7272	0,6424
RD2 .....	0,4631	0,4855	0,4515	0,5275
RS2 .....	0,3384	0,3584	0,3269	0,4242
AIC .....	-3,7610	-4,0187	-3,6007	-3,4809
SIC (Bayes) .....	-2,3247	-2,5824	-2,1644	-2,0446

NOTA: Os números entre colchetes são os p-valores correspondentes das estatísticas.

(1) Interação muito forte.

Tabela 3

## Avaliação das predições dos modelos PIB1, PIB2, PIB3, PIB4 e série do IBGE — 1980-96

ESTATÍSTICAS	PIB1	PIB2	PIB3	PIB4	Série do IBGE YQ6096
Raiz quadrada do EQM.....	0,0089	0,0112	0,0158	0,0245	0,0427
EQM (%).....	0,0069	0,0078	0,0121	0,0196	0,0372
Erro absoluto médio.....	0,0070	0,0073	0,0091	0,0199	0,0356
Erro absoluto médio (%).....	0,0054	0,0052	0,0066	0,0158	0,0312
<b>Coefficiente Desigualdade de Theil .....</b>	<b>0,0034</b>	<b>0,0043</b>	<b>0,0060</b>	<b>0,0094</b>	<b>0,0185</b>
Proporção do viés .....	0,0008	0,2847	0,1428	0,0021	0,0398
Proporção da variância.....	0,0008	0,2685	0,0947	0,0050	0,0263
Proporção da covariância...	0,9984	0,4469	0,7625	0,9929	0,9339
Pseudo R2.....	0,9969	0,9979	0,9928	0,9770	0,9198

O modelo PIB1, sem *benchmarking*, apresenta diagnósticos adequados. É importante frisar a inexistência de quaisquer indicações de quebras no termo de nível da série, nem da presença de *outliers* em 1980, um ano crítico, dado que nele ocorreu o encadeamento da série estimada com a série publicada pelo IBGE.<sup>29</sup> Pode-se concluir que o modelo estimado capturou adequadamente os componentes estruturais da série, o que, provavelmente, decorreu de a sazonalidade ter sido especificada estocasticamente.

As avaliações das previsões um passo à frente, feitas para o período 1980-1996, estão reportadas na Tabela 3. Esse é o período para o qual estão disponíveis as estimativas oficiais do IBGE, obtidas no *site* do IPEADATA. Apenas se ressalta que as estatísticas são bem claras quanto à acurácia das estimativas realizadas pelo modelo PIB1. O gráfico do PIB1 — estimado pela soma dos componentes de tendência e sazonais — convertido para valores reais de 1980 aparece na Figura 3. Na quinta coluna da Tabela 3, as estatísticas referem-se ao PIB em valores estimados, com os coeficientes da equação de longo prazo do procedimento de Engle-Granger sem encadeamento com a série do IBGE, denotada por YQ6096 (ver subseção 3.1).

O modelo PIB2 corrige os valores trimestrais do PIB, de tal maneira que a soma dos PIBs trimestrais estimados se iguala à do PIB oficial do IBGE. A heteroscedasticidade foi inteiramente tratada. Não há presença de correlação serial nas inovações padronizadas, mas o teste K-S é rejeitado a 1%. A normalidade das inovações é rejeitada, o que não chega a ser um problema, dado que não se pretende fazer inferência estatística sobre parâmetros desconhecidos, mas, sim, extrair componentes estruturais. Contudo é uma indicação de que o comportamento dos valores observados não está sendo representado apropriadamente pelo modelo em consideração. Frisa-se também que não foram encontrados quaisquer indícios de quebras no termo de nível da série resultante, nem de presença de *outliers* na vizinhança de 1980.

Entre as limitações dos resultados desse modelo, ressalta-se a ausência de *dummies* de intervenção, o que foi detectado através de experimentos realizados com o modelo com estrutura de *benchmarking* — rodados em ambiente de modelo estrutural com a especificação incluindo variáveis explicativas. Por outro lado, o erro de medida estimado possui um coeficiente de correlação serial

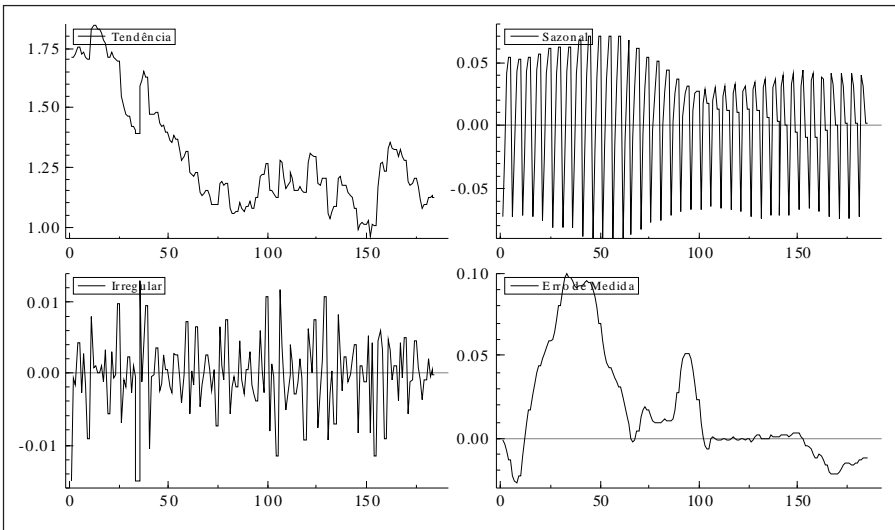
---

<sup>29</sup> Para corroborar essa afirmativa, fez-se uma regressão da série PIB1, em níveis e em diferenças, contra termos determinísticos (constante, tendência e *dummies* sazonais), estimada por mínimos quadrados recursivos. Os coeficientes, assim como seus intervalos de confiança, não apresentam mudanças de comportamento em 1980, apesar de serem bastante instáveis. Os resultados estão disponíveis aos interessados.

de primeira ordem muito próximo de um (0,9727),<sup>30</sup> o que significa que esse termo absorveu ou refletiu a correlação serial, que, de outra forma, estaria contida nas inovações. Interpreta-se essa forte persistência como sendo "herdada" de estruturas que, potencialmente, deveriam estar embutidas nos demais componentes do vetor de estado,<sup>31</sup> conforme se visualiza no quarto painel da Figura 2. Julga-se esse problema superável com a mudança da especificação do modelo.<sup>32</sup>

Figura 2

Componentes extraídos do modelo univariado PIB2 — 1960-96



<sup>30</sup> O erro de medida estimado  $\xi_t$ , sem os totais anuais tem ordem de integração próxima de um. Os testes de raiz unitária com polinômio de defasagens apropriadamente especificado aceitam, em conjunto, a hipótese nula de raiz unitária;  $\xi_t$  segue, assim, um processo ARIMA (4,1,0). A hipótese do modelo é que  $\xi_t$  é  $AR(1) \sim (0,1)$ .

<sup>31</sup> Também é importante ter em conta que, devido às inúmeras dificuldades computacionais para estimação desse modelo, esta análise é incompleta, posto não ter sido possível estimar os resíduos auxiliares.

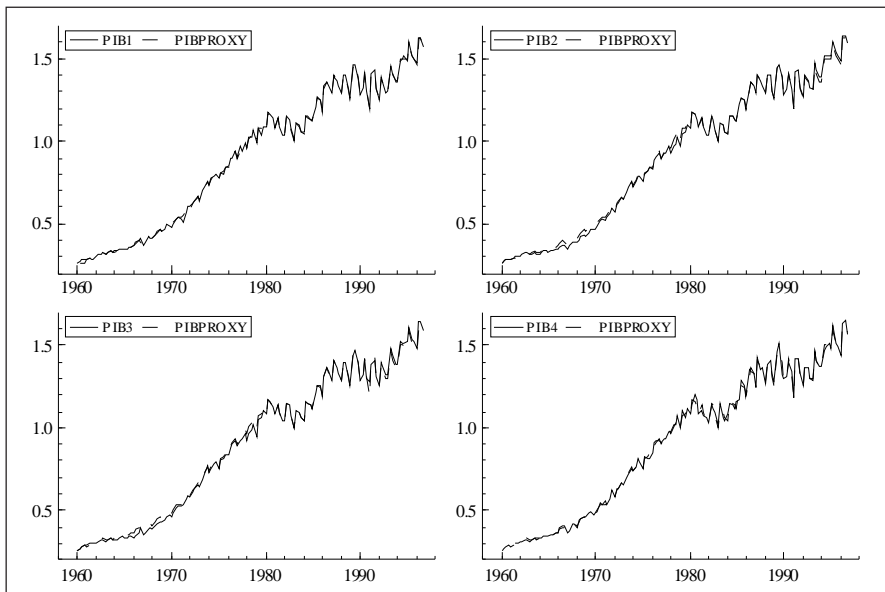
<sup>32</sup> Para testar a validade do procedimento de *benchmarking* empregado, pode-se adotar o seguinte procedimento: tomar uma série observada ou simulada com frequência trimestral e cujos totais anuais sejam iguais à soma dos trimestres e introduzir um erro seguindo um processo AR(1) nos dados reais; em seguida, aplicar o procedimento de *benchmarking* e comparar os resultados para observar se a série foi corrigida do erro de medida, tal como proposto pelo método.



Todas as observações feitas para o modelo PIB2 aplicam-se ao modelo PIB3 estimado contendo alguns valores omissos (ver seção anterior). Em especial, destacam-se três resultados dos modelos de *benchmarking*. O primeiro — não reportado nesta versão do artigo — é que o modelo PIB2 estimado sem tratamento de heteroscedasticidade produz praticamente a mesma série, não havendo diferença material entre elas. Assim, para esse modelo de estado de espaço — *benchmarking* —, o filtro de Kalman é robusto à heteroscedasticidade. O segundo é que a presença de dados omissos, por sua vez, afeta as estimativas, mas não de modo expressivo. O terceiro é um resultado evidenciado pela análise da avaliação das previsões para o período 1980-96, apresentada na Tabela 3. O aparente viés das estimativas vem do fato de que, a partir do 4º trimestre de 1990, o PIB estimado sobreestima sistematicamente as estimativas do PIB feitas pelo IBGE (Figura 3). Em especial, o PIB3 não acompanha os movimentos bruscos do PIB estimado pelo IBGE ocorridos em 1991, subestimando-os, o que sugere que o procedimento adotado está efetivamente corrigindo as estimativas trimestrais oficiais, de modo que a soma dos trimestres seja consistente com ou igual ao PIB anual. Nesse caso, o viés detectado é uma característica do PIB trimestral oficial estimado, e, portanto, são as estimativas do IBGE, obtidas no IPEADATA, que não apresentam consistência.

Figura 3

## PIBs estimados — 1960-96



O modelo multivariado associado ao PIB4 estimado envolveu as séries de índice do PIB, de produção de cimento, de veículos automotores e de consumo de energia elétrica no Rio de Janeiro e em São Paulo. Todas elas foram divididas pelo índice da receita tributária real, extraindo-se previamente os *outliers*. A série de produção de automóveis em defasagens apropriadamente estacionarizada entrou como variável explicativa, com o intuito de controlar a correlação serial. Os testes dos resíduos padronizados indicam que os pressupostos básicos do modelo estão sendo respeitados. Por outro lado, não há, nos resíduos auxiliares das variáveis, indicações da presença de *outliers* ou de quebras, o que revela que o modelo captura adequadamente os componentes estruturais da variável.<sup>33</sup> O PIB4 é estimado como a soma dos componentes de tendência e sazonal a menos do termo irregular. A sua aderência aos dados do IBGE é evidente.

Vale ressaltar que os componentes sazonais extraídos dos modelos estimados (ver Figura 2 para o modelo PIB2; os demais não foram reportados) não apresentam quebras significativas em torno do ano de 1980 (observação 29). O que se verifica é a sazonalidade oscilando ao longo de todo o período. Isso permite deduzir que o encadeamento, *per se*, da série estimada na primeira etapa com a do IBGE não implicou alterações materiais do padrão sazonal.

Na Figura 3 estão plotados os quatro PIBs estimados pela metodologia de espaço de estado e o PIBPROXY — estimado pela estática de Engle-Granger e encadeado com a série do IBGE. Olhados individualmente, cada um dos resultados estimados apresenta boas propriedades, e todos representam um avanço em relação às estimativas do PIB com os coeficientes da estática de Engle-Granger. Por sua vez, os procedimentos de estimação utilizando filtro de Kalman com valores omissos e com processo de *benchmarking* são inovadores na literatura brasileira, ao que se juntam o procedimento de tratamento da heteroscedasticidade e o emprego de variáveis explicativas em modelos univariados e multivariados.

As estatísticas dos testes de diagnóstico e da avaliação da capacidade preditiva dos modelos não indicam claramente qual modelo escolher, entre o univariado com valores omissos (PIB1) e o multivariado (PIB4), para ser usado na recuperação do PIB trimestral, no período 1960-79, e como substituto do PIB do IBGE nos anos de 1980 a 1996. Os modelos com *benchmarking* (PIB2 e

---

<sup>33</sup> Os componentes extraídos da equação do PIB — não reportados — indicam que o componente de tendência se assemelha ao do modelo PIB1, embora a sazonalidade apresente um padrão mais instável e de maior amplitude, provavelmente devido à presença de quatro variáveis dentro do modelo. O termo irregular, tal como no modelo do PIB1, é um termo auto-regressivo estável, no caso, um AR(8).

PIB3), por corrigirem o erro de medida existente nos dados trimestrais, harmonizam os dados recuperados com os valores anuais conhecidos e constituem um aperfeiçoamento em relação aos próprios dados trimestrais da fonte oficial — o IBGE — no período 1980-96.

## 5 Conclusão

Neste artigo, apresentou-se uma metodologia para a estimação do PIB trimestral para o período 1960-96 e foram reportadas as estimativas realizadas. Embora o objetivo inicial fosse estimá-lo apenas para o período 1960-79, a necessidade de uma amostra maior levou a se fazerem estimativas para um período mais longo. A necessidade de uma *proxy* judiciosa para o PIB, para períodos anteriores a 1980, relaciona-se diretamente à realização de estudos empíricos que cubram um período mais longo da história econômica brasileira.

Inicialmente, realizou-se um experimento com um modelo simples, porém rigoroso — dado seus estimadores serem superconsistentes —, que foi o de estimar os coeficientes da equação estática do teste de cointegração de dois estágios de Engle-Granger, realizado com frequência anual. As séries utilizadas foram o PIB, as produções de cimento e de veículos automotores, o consumo industrial de energia elétrica nas regiões do Rio de Janeiro e de São Paulo e a receita tributária do Tesouro Nacional. O vetor de cointegração estimado foi, então, aplicado sobre os dados trimestrais das séries, para gerar a estimativa do PIB trimestral. Os resultados indicaram que a melhor conduta seria estender os estudos com uma metodologia mais robusta, o que conduziu a se estimarem modelos que utilizam métodos de espaço de estado.

Continuou-se a pesquisa com diferentes modelos estruturais, dadas as estimativas realizadas com o modelo acima. Foram estimados um modelo univariado, com valores omissos no período 1960-79, utilizando-se variáveis explicativas, e um modelo com processo de *benchmarking*, que corrige o erro de medida dos valores dos trimestres de modo que a soma dos PIBs trimestrais totalize o PIB anual respectivo. Realizou-se também um experimento com modelo de *benchmarking* e valores omissos. Finalmente, estimou-se um modelo estrutural multivariado a quatro séries e com variáveis explicativas. Todos eles foram estimados com correção de heteroscedasticidade. Estimaram-se, portanto, quatro modelos concorrentes e quatro séries de PIB trimestral. À exceção do modelo de *benchmarking*, nenhum dos PIBs recuperados inclui o componente irregular. A lógica do *benchmarking* considera o componente irregular, porque o componente de erro de medida é retirado da sua composição.

Os resultados foram bastante satisfatórios no que concerne aos diagnósticos dos resíduos, os quais respeitam as hipóteses do modelo gaussiano linear. Da mesma forma, os PIBs estimados apresentaram boa aderência aos dados e previram adequadamente o PIB oficial no período após 1980. O PIB estimado com *benchmarking* corrige, ao máximo, os erros de medida e inequivocamente mantém consistência com o PIB anual, aperfeiçoando os próprios dados oficiais. Em nível global, o modelo estimado com *missing values* e o modelo SURE apresentaram as melhores *performances*. Em conjunto, as estatísticas reportadas não indicam claramente qual a melhor estimativa. Certamente, todas representaram uma melhora em relação ao PIB estimado com a estática de Engle-Granger. A escolha para prever o PIB do período 1960-96 deverá recair sobre uma delas.

Como recomendação de estudos futuros, indicam-se três direções. A primeira é utilizando os modelos discutidos neste artigo para pesquisar outros períodos, preferencialmente que se estendam até o presente e que permitam a inclusão de outras variáveis relacionadas com o PIB. Como consequência natural, o caminho para a realização de extrapolações fica aberto. Uma segunda, complementar, é a de estender a metodologia para modelos multivariados de *benchmarking*, o que traria maior precisão e confiabilidade às séries finais do PIB trimestral. Finalmente, sugere-se estender o procedimento para a recuperação e a correção de erros de medida de dados mensais a partir de dados trimestrais previamente corrigidos.

A principal contribuição deste artigo é apresentar uma metodologia consistente e original para a recuperação e a estimação do PIB trimestral do período 1960-96 e que, ademais, é capaz de corrigir os dados da pesquisa trimestral do IBGE. Por fim, essa metodologia oferece bases concretas para o desenvolvimento de um procedimento de previsão e/ou extrapolação do PIB trimestral.

## Apêndice

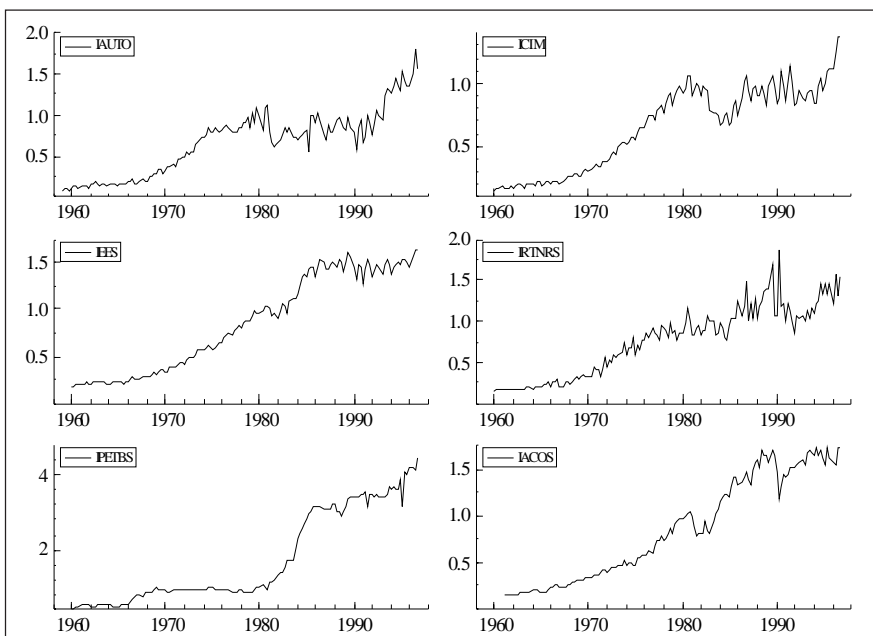
### Apêndice A: dados utilizados

Na Figura A.1, estão os dados em índices trimestrais encadeados com as observações interpoladas ou projetadas, respectivamente, da produção de automóveis (IAUTO), de cimento (ICIM), do consumo industrial de energia elétrica (IEES), da receita tributária real (IRTNRS), da produção de petróleo (LGN) e gás

liquefeito natural (IPETBS) e de aço em lingotes (IACOS).<sup>1</sup> Na Figura A.2, aparecem as séries do PIB estimado, YRTN, as séries observadas da produção de automóveis, de cimento e consumo industrial de energia elétrica, todas divididas pela série da receita tributária real ajustada e expressa em índices.

Figura A.1

Índices trimestrais das variáveis intervenientes no PIB — 1960-96

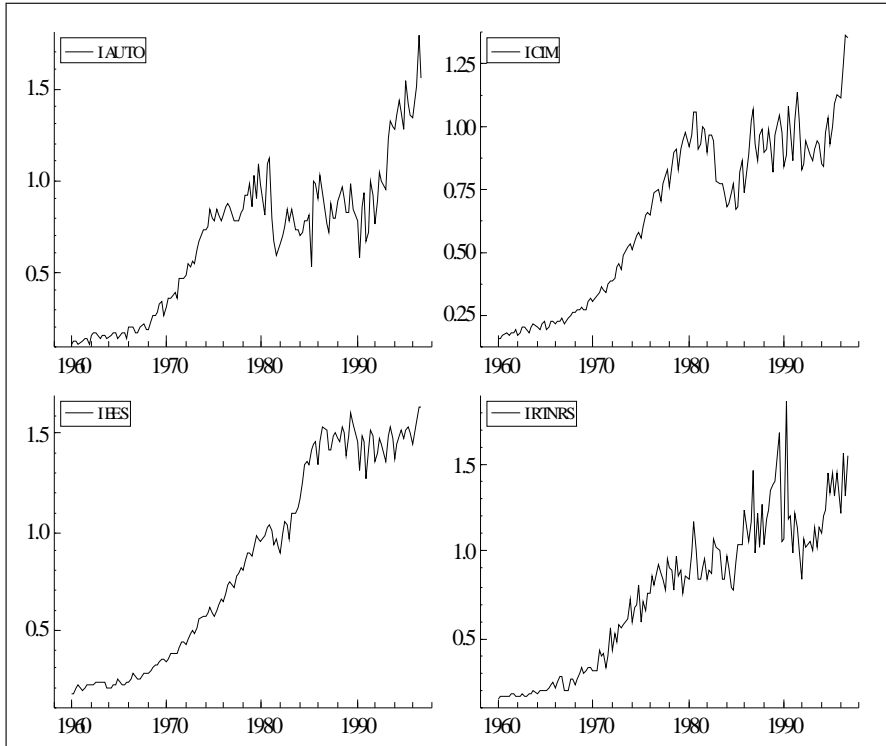


NOTA: IAUTO é a produção de automóveis; ICIM, a de cimento; IEES é o consumo industrial de energia elétrica; IRTNRS, a receita tributária real; IPETBS é a produção de petróleo e LGN; e IACOS, a produção de aço em lingotes.

<sup>1</sup> Para as séries que possuem apenas frequência anual para parte do período — produção de petróleo, aço e receita tributária nacional em valores correntes —, construiu-se um teste ADF de raiz unitária para um período “específico”. Com a especificação do teste, deduz-se o processo ARIMA da série, isto é, um *random walk* puro, um *random walk* com um processo auto-regressivo de ordem 1 e o coeficiente de auto-regressão associado, ou um processo integrado de segunda ordem. Com essas informações, realiza-se a interpolação descrita em Bertsekas (1976) e encadeia-se a série interpolada com a série trimestral observada a partir do período em que os dados não estão disponíveis. Para a série de produção de petróleo,

Figura A.2

Índices trimestrais das variáveis ajustadas — 1960-96



NOTA: IAUTO é a produção de automóveis; ICM, a de cimento; IEES é o consumo industrial de energia elétrica; e IRTNRS, a receita tributária real.

escolheu-se interpolar no período 1955-69, em função da homogeneidade do crescimento nesse período, ainda não muito pronunciado. A maior parte dos testes indicou a série como sendo  $I(2)$ , o que levou à escolha desta especificação. A interpolação da produção de aço foi feita para o período 1961-97, a partir da constatação de que ela segue um processo do tipo *random walk* puro. A série de receita tributária nacional foi interpolada no período 1952-73, supondo que a série no período era  $I(2)$ . A série foi convertida para reais e encadeada com a série observada no 3º trim./65.

## Apêndice B: matrizes do modelo de benchmarking

A seguir, apresentam-se o grupo de matrizes referente a um modelo de *benchmarking* univariado com nível, *dummies* sazonais e termo de erro de medida, juntamente com os vetores de estados  $\alpha_t$  e de choques de estado  $\tilde{\eta}_t$ , associados às matrizes. Por conveniência, o vetor de observações está inserido dentro do vetor de estado. Lembre-se de que o modelo espaço de estado é dado por

$$y_t = Z_t \alpha_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, H_t),$$

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t, \quad \eta_t \sim N(0, Q_t) \quad t = 1, \dots, n, \quad \alpha_1 \sim N(a_1, P_1),$$

onde no modelo univariado  $y_t$  tem uma variável; e o modelo multivariado, quatro variáveis empilhadas.

Para os  $Y$ 's (trimestres), as matrizes são:

$$Z_t = [1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 1],$$

$$T_t = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \phi \end{bmatrix},$$

$$R_t = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



Para os  $X$ 's (totais anuais), as matrizes são  $Z_t = [1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1]$ ,  
 $T_t =$  Identidade  $12 \times 12$

$$R_t = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Os vetores de estados e de choques de estados são, respectivamente, os seguintes:

$$\alpha_t' = [\mu_t, \mu_{t-1}, \mu_{t-2}, \mu_{t-3}, \gamma_t, \gamma_{t-1}, \gamma_{t-2}, \gamma_{t-3}, \varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \varepsilon_{t-3}, \xi_t^s]$$

e  $\tilde{\eta}_t' = [\eta_t, \omega_t, \varepsilon_t^*, \kappa_t]$

As matrizes que seguem se referem a um modelo de *benchmarking* com quatro variáveis dependentes (vetor de medidas com dimensão quatro), uma delas sendo o PIB. A especificação de cada variável é a mesma do modelo univariado. As matrizes  $Z_t$ ,  $T_t$  e  $R_t$  possuem uma dimensão a menos que as matrizes do modelo univariado, porque delas foi retirado o erro de medida  $\xi_t^s$ , o qual é recolocado posteriormente.

O procedimento de construção das matrizes segue os seguintes passos: (a) reconstruir as matrizes do modelo univariado sem considerar o erro de medida, para os  $Y$ 's, e para os  $X$ 's, as quais não são rerepresentadas; (b) construir as matrizes,  $\tilde{Z}_t = Z_t \otimes I_4$ ,  $\tilde{T}_t = T_t \otimes I_4$ , e  $\tilde{R}_t = R_t \otimes I_4$ , nas quais  $\otimes$  é o produto de Kronecker; (c) construir as matrizes finais, que aparecem a seguir, montadas a partir dos componentes descritos, seguidas das matrizes bloco-diagonal de covariância  $Q_t$  e  $H_t$ .

Para os  $Y$ 's (trimestres), as matrizes são:

$$\bar{Z}_t = \begin{pmatrix} \left( \begin{array}{c} 1 \\ 0 \\ 0 \\ 0 \end{array} \right) \\ \left( \begin{array}{c} \tilde{Z}_t \end{array} \right) \end{pmatrix}, \quad \bar{T}_t = \begin{pmatrix} \left( \begin{array}{c} \tilde{T}_t \end{array} \right) & \left( \begin{array}{c} 0 \\ 48 \times 1 \end{array} \right) \\ \left( \begin{array}{c} 0 \\ 1 \times 48 \end{array} \right) & (\phi) \end{pmatrix}, \quad \bar{R}_t = \begin{pmatrix} \left( \begin{array}{c} \tilde{R}_t \end{array} \right) \left( \begin{array}{c} 0 \\ 48 \times 1 \end{array} \right) \\ \left( \begin{array}{c} 0 \\ 1 \times 12 \end{array} \right) & \left( \begin{array}{c} 1 \end{array} \right) \end{pmatrix}$$

$$Q = \text{diagcat}(\Sigma_\eta, \Sigma_\omega, \Sigma_\varepsilon, \sigma_\xi)$$

Para os  $X$ 's (totais anuais), as matrizes são:

$$\bar{Z}_t = \begin{pmatrix} \left( \begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \end{array} \right) \\ \left( \begin{array}{c} \tilde{Z}_t \end{array} \right) \end{pmatrix}, \quad \bar{T}_t = \begin{pmatrix} \left( \begin{array}{c} \tilde{T}_t \end{array} \right) & \left( \begin{array}{c} 0 \\ 48 \times 1 \end{array} \right) \\ \left( \begin{array}{c} 0 \\ 1 \times 48 \end{array} \right) & (1) \end{pmatrix}, \quad \bar{R}_t = \begin{pmatrix} \left( \begin{array}{c} \tilde{R}_t \end{array} \right) \left( \begin{array}{c} 0 \\ 48 \times 1 \end{array} \right) \\ \left( \begin{array}{c} 0 \\ 1 \times 12 \end{array} \right) & \left( \begin{array}{c} 0 \\ 1 \times 12 \end{array} \right) \end{pmatrix}$$

$$Q = \text{diagcat}(\Sigma_\eta, \Sigma_\omega, \Sigma_\varepsilon, \sigma_\xi) \quad Y's,$$

Observe-se que as matrizes do passo (b) devem ter as dimensões  $48 \times 4$  ( $\tilde{Z}_t$ ),  $48 \times 48$  ( $\tilde{T}_t$ ) e  $8 \times 12$  ( $\tilde{R}_t$ ), tanto para os  $Y$ 's (trimestres) quanto para os  $X$ 's (totais anuais), e que os blocos de  $Q_t$  são "cheios", porque as componentes de mesma interpretação são correlacionadas. A matriz  $\sigma_\xi$  ( $4 \times 1$ ) é univariada, pois existe um erro de medida, e este é associado ao PIB. O vetor de espaço do modelo multivariado não é apresentado, em função da sua dimensão ( $37 \times 1$ ). Porém ele nada mais é do que o vetor de estado do modelo univariado com cada grupo de componentes repetido quatro vezes, devidamente indexados pelo número da variável dependente (quatro) correspondente. O vetor de choques de estado é dado por

$$\bar{\eta}'_t = \left[ \eta_{t1}, \eta_{t2}, \eta_{t3}, \eta_{t4}, \omega_{t1}, \omega_{t2}, \omega_{t3}, \omega_{t4}, \varepsilon_{t1}^*, \varepsilon_{t2}^*, \varepsilon_{t3}^*, \varepsilon_{t4}^*, \kappa_t \right]$$

Os  $\eta$ 's estão associados aos níveis; os  $\omega$ 's às sazonalidades; os  $\varepsilon^*$ 's, às componentes irregulares; e  $\kappa_t$  é o choque do erro de medida do PIB, lembrando que esse erro de medida segue um AR(1), como evidenciado pela matriz  $\bar{T}$  relativa aos (trimestres).

## Referências

- ANUÁRIO ESTATÍSTICO DO BRASIL. Rio de Janeiro: IBGE, v. 50, 1990.
- ANUÁRIO ESTATÍSTICO DO BRASIL. Rio de Janeiro: IBGE, v. 54, 1994.
- BANCO CENTRAL DO BRASIL. Disponível em: <<http://bcb.gov.br/>>. Acesso em: jun. 2004.
- BASSI, V. L. Appendix A. In: \_\_\_\_ (Ed.). **Economic Forecasting**. New York: McGraw-Hill, 1958.
- BERTSEKAS, D. P. **Dynamic programming and stochastic control**. New York: Academic Press, 1976.
- BOLETIM MENSAL. Brasília: Banco Central do Brasil, maio 1994.
- CAPORALE, G. M.; PITTIS, N. Unit root versus other types of time heterogeneity, parameter time dependence and superexogeneity. **Journal of Forecasting**, v. 21, p. 207-223, 2002.
- CARDOSO, E. Uma equação para a demanda de moeda no Brasil. **Pesquisa e Planejamento Econômico**, v. 11, n. 3, p. 617-655, 1981.
- CERQUEIRA, L. F. **Demanda por moeda, regressões espúrias e níveis de instabilidade hiperinflacionários**. Dissertação (Mestrado)-Pontifícia Universidade Católica, Departamento de Economia, Rio de Janeiro, 1993.
- CERQUEIRA, L. F. **Metodologia de estimação do PIB trimestral utilizando procedimentos de cointegração e filtros de Kalman**. Niterói: UFF. Departamento de Economia, 2007. (Textos para discussão; n. 216).
- CHOLETTE, P. A.; DAGUM, E. B. Benchmarking time series with autocorrelated survey errors. **International Statistical Review**, v. 62, n. 3, p. 365-377, 1994.
- CHOW, G. C.; LIN, A. Best linear unbiased interpolation, distribution and extrapolation of time series by related series. **Review of Economics and Statistics**, v. 53, p. 372-375, 1971.
- CONJUNTURA ECONÔMICA. Rio de Janeiro: FGV, v. 31, n. 4, abr.1977.
- CONJUNTURA ECONÔMICA. Rio de Janeiro: FGV, v. 33, n. 8, ago.1979.
- CONTADOR, C. R.; SANTOS FILHO, W. A. C. Produto Interno Bruto trimestral: bases metodológicas e estimativas. **Pesquisa e Planejamento Econômico**, v. 17, n. 3, p. 711-742, 1987.

DENTON, F. T. Adjustment of monthly or quarterly series to annual totals: an approach based on quadratic minimization. **Journal of the American Statistical Association**, v. 66, p. 99-102, 1971.

DI FONZO, T.; MARINI, M. **Benchmarking systems of seasonally adjusted time series according to Denton's Movement Preservation Principle**. Padova: Università di Padova, Dipartimento di Scienze Statistiche, 2003. (Working paper; n. 2003-09).

DURBIN, J.; KOOPMAN, S. J. **Time series analysis by state space methods**. Oxford: Oxford University, 2004.

DURBIN, J.; QUEENNEVILLE, B. Benchmarking by state space models. **International Statistical Review**, v. 65, n. 1, p. 21-48, 1997.

ELLIOT, G.; ROTHENBERG, T. J.; STOCK, J. H. Efficient tests for an autoregressive unit root. **Econometrica**, v. 64, p. 813-836, 1996.

ENGLE, R. F.; GRANGER, C. Cointegration and error correction: representation, estimation, and testing. In: \_\_\_\_ (Ed.). **Long-run economic relationships: readings in cointegration**. Oxford: Oxford Univ., 1991a.

ENGLE, R. F.; GRANGER, C. W. J. Introduction. In: \_\_\_\_ (Ed.). **Long-run economic relationships: readings in cointegration**. Oxford: Oxford Univ., 1991.

GINSBURGH, V. A. A further note on the derivation of quarterly figures consistent with annual data. **Applied Statistics**, v. 22, n. 3, p. 368-374, 1973.

HARVEY, A. C. **Forecasting, structural time series and the Kalman Filter**. Cambridge: Cambridge University, 1990.

HARVEY, A. C.; KOOPMAN, S. J. Diagnostic checking of unobserved time series models. **Journal Business and Economic Statistics**, v. 10, p. 377-389, 1992.

HARVEY, A. C.; KOOPMAN, S. J. Multivariate structural time series models. In: HEJI, C. et al. (Ed). **Systematic dynamics in economics and financial models**. Chichester: Wiley & Sons, 1997. p. 269-298.

HARVEY, A. C.; KOOPMAN, S. J. Signal extraction and the formulation of unobserved components models. **Econometrics Journal**, n. 3, p. 84-107, 2000.

HILLMER, S. C.; TRABELSI, A. Benchmarking of economic time series. **Journal of the American Statistical Association**, v. 82, n. 400, p. 1064-1071, 1987.

IBGE. Sistema IBGE de Recuperação Automática-SIDRA. Disponível em: <<http://www.sidra.ibge.gov.br/>>. Acesso em: 21 jun. 2004.

IPEADATA. Disponível em: <<http://www.ipeadata.gov.br/>>. Acesso em: 21 jun. 2004.

JOHANSEN, S. **Likelihood-based inference in cointegrated vector auto-regressive models**. New York: Oxford University, 1995.

KOOPMAN, S. J.; SHEPHARD, H. Exact score for time series models in state space form. **Biometrika**, v. 79, p. 823-826, 1992.

KOOPMAN, S. J. et al. **Stamp 6.3: structural time series analyzer, modeller, and predictor**. London: Timberlake Consultants, 2000.

KOOPMAN, S. J.; SHEPHARD, H.; DOORNIK, J. A. Statistical algorithms for models in state space using SsfPack 2.2. **Econometrics Journal**, v. 2, p. 107-160, 1999.

MADDALA, G. S.; KIM, In-Moo. **Unit roots, cointegration and structural changes**. Cambridge: Cambridge University, 2002.

MIAN, I. U. H.; LANIEL, N. Maximum likelihood estimation of constant multiplicative bias benchmark model with application. **Survey Methodology**, v. 19, p. 165-172, Dec 1993.

MOAURO, F; SAVIO, G. Temporal disaggregating using multivariate structural time series models. **Econometrics Journal**, v. 8, n. 2, p. 214-234, 2005.

NAKANE, M. I. **Testes de exogeneidade fraca e de superexogeneidade para a demanda por moeda no Brasil**. Rio de Janeiro: BNDES, 1994.

NG, S.; PERRON, P. Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. **Review of Economic Studies**, v. 63, n. 3, p. 435-465, July 1997.

PASTORE, A. C. Déficit, dívida pública, moeda e inflação: uma resenha. (não publicado, 1991).

PHILLIPS, P. C. B.; HANSEN, B. E. Statistical inference in instrumental variables regression with I(1) processes. **Review of Economic Studies**, v. 58, p. 407-436, 1990.

PINOTTI, M. C. **Inovações financeiras e imposto inflacionário**. Tese (Doutorado)-Universidade de São Paulo; FEA, São Paulo, 1991.

QUANTITATIVE MICRO SOFTWARE. **Eviews 5**. Irvine, CA: Quantitative Micro Software, 2004.

ROSSI, J. M. A demanda de moeda no Brasil: o que ocorreu a partir de 1980? **Pesquisa e Planejamento Econômico**, v. 18, n. 1, p. 37-53, 1988.