

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DE RIBEIRÃO PRETO
DEPARTAMENTO DE ECONOMIA

CELSO VILELA CHAVES CAMPOS

**Previsão da Arrecadação de Receitas Federais:
Aplicações de Modelos de Séries Temporais para o Estado de São Paulo**

Orientador: Prof. Dr. Rudinei Toneto Júnior

RIBEIRÃO PRETO
2009

Profa. Dra. Suely Vilela
Reitora da Universidade de São Paulo

Prof. Dr. Rudinei Toneto Júnior
Diretor da Faculdade de Economia, Administração e Contabilidade de Ribeirão
Preto

Prof. Dr. Walter Belluzzo Júnior
Chefe do Departamento de Economia

CELSO VILELA CHAVES CAMPOS

**Previsão da Arrecadação de Receitas Federais:
Aplicações de Modelos de Séries Temporais para o Estado de São Paulo**

Dissertação apresentada ao Programa de Pós-Graduação em Economia da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo como requisito para obtenção do título de Mestre em Economia.

Orientador: Prof. Dr. Rudinei Toneto Júnior

RIBEIRÃO PRETO
2009

AUTORIZO A DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

FICHA CATALOGRÁFICA

Campos, Celso Vilela Chaves

Previsão da Arrecadação de Receitas Federais: Aplicações de Modelos de Séries Temporais para o Estado de São Paulo. Ribeirão Preto, 2009.

133 p. : il. ; 30cm

Dissertação de Mestrado, apresentada à Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo.

Orientador: Toneto Júnior, Rudinei

1. Arrecadação. 2. Previsão. 3. Séries de tempo. 4. ARIMA. 5. Modelos dinâmicos univariados e multivariados. 6. Modelo Estrutural.

ERRATA

CAMPOS, C. V. C. **Previsão da Arrecadação de Receitas Federais: Aplicações de Modelos de Séries Temporais para o Estado de São Paulo.** 2009. 133 f. Dissertação (Mestrado) – Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, 2009.

página	linha	onde se lê	leia-se
57	13	(IVV, IPI e PIB)	(IVV, IPI, PIB e TCR)
75 (Tabela 20)	5	(0,61)	(0,061)
82	2 (nota rodapé 33)	(LII)	(LLR)
84	2	(DLII/DLIPI/DTCR)	(DLII/DLIPI/DLTCR)
84	17	(método dos indicadores, dos modelos ARIMA e equações simultâneas)	(método dos indicadores, dos modelos ARIMA e de equações simultâneas)
84	21	(Cofins)	(Contribuição para o Financiamento da Seguridade Social)
84	23	(Trasferência)	(Transferência)

FOLHA DE APROVAÇÃO

Celso Vilela Chaves Campos

Previsão da Arrecadação de Receitas Federais: Aplicações de Modelos de Séries Temporais para o Estado de São Paulo.

Dissertação apresentada à Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo para obtenção do título de Mestre.

Aprovado em:

Banca Examinadora

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

DEDICATÓRIA

À Jussara, minha esposa, pelo incentivo e grande apoio ao meu retorno aos estudos.

Aos meus filhos, Tiago e Amanda, pela paciência com que suportaram a minha ausência em vários momentos ao longo do período de elaboração deste trabalho.

“Pessoas que simulam prever o futuro serão consideradas desordeiras nos termos da subdivisão 3, seção 901 do código criminal e estarão sujeitas a uma multa de US\$250 e/ou seis meses de prisão.”

Seção 889, Código de Processo Criminal do Estado de Nova York

Pindyck e Rubinfeld (2004)

RESUMO

CAMPOS, C. V. C. **Previsão da Arrecadação de Receitas Federais: Aplicações de Modelos de Séries Temporais para o Estado de São Paulo**. 2009. 133 f. Dissertação (Mestrado) – Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, 2009.

O objetivo principal do presente trabalho é oferecer métodos alternativos de previsão da arrecadação tributária federal, baseados em metodologias de séries temporais, inclusive com a utilização de variáveis explicativas, que reflitam a influência do cenário macroeconômico na arrecadação tributária, com o intuito de melhorar a acurácia da previsão da arrecadação. Para tanto, foram aplicadas as metodologias de modelos dinâmicos univariados, multivariados, quais sejam, Função de Transferência, Auto-regressão Vetorial (VAR), VAR com correção de erro (VEC), Equações Simultâneas, e de modelos Estruturais. O trabalho tem abrangência regional e limita-se à análise de três séries mensais da arrecadação, relativas ao Imposto de Importação, Imposto Sobre a Renda das Pessoas Jurídicas e Contribuição para o Financiamento da Seguridade Social - Cofins, no âmbito da jurisdição do estado de São Paulo, no período de 2000 a 2007. Os resultados das previsões dos modelos acima citados são comparados entre si, com a modelagem ARIMA e com o método dos indicadores, atualmente utilizado pela Secretaria da Receita Federal do Brasil (RFB) para previsão anual da arrecadação tributária, por meio da raiz do erro médio quadrático de previsão (RMSE). A redução média do RMSE foi de 42% em relação ao erro cometido pelo método dos indicadores e de 35% em relação à modelagem ARIMA, além da drástica redução do erro anual de previsão. A utilização de metodologias de séries temporais para a previsão da arrecadação de receitas federais mostrou ser uma alternativa viável ao método dos indicadores, contribuindo para previsões mais precisas, tornando-se ferramenta segura de apoio para a tomada de decisões dos gestores.

Palavras-chave: Arrecadação. Previsão. Séries de tempo. ARIMA. Modelos dinâmicos univariados e multivariados. Modelo Estrutural.

ABSTRACT

CAMPOS, C. V. C. **Federal Revenue Collection Forecast: Application of Time Series Models at the State of Sao Paulo. 2009.** 133 f. Dissertation (Master's degree.) - Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, 2009.

The main objective of this work is to offer alternative methods for federal tax revenue forecasting, based on methodologies of time series, inclusively with the use of explanatory variables, which reflect the influence of the macroeconomic scenario in the tax collection, for the purpose of improving the accuracy of revenues forecasting. Therefore, there were applied the methodologies of univariate dynamic models, multivariate, namely, Transfer Function, Vector Autoregression (VAR), VAR with error correction (VEC), Simultaneous Equations, and Structural Models. The work has a regional scope and it is limited to the analysis of three series of monthly tax collection of the Import Duty, the Income Tax Law over Legal Entities Revenue and the Contribution for the Social Security Financing – Cofins, under the jurisdiction of the state of São Paulo in the period from 2000 to 2007. The results of the forecasts from the models above were compared with each other, with the ARIMA moulding and with the indicators method, currently used by the Secretaria da Receita Federal do Brasil (RFB) to annual foresee of the tax collection, through the root mean square error of approximation (RMSE). The average reduction of RMSE was 42% compared to the error committed by the method of indicators and 35% of the ARIMA model, besides the drastic reduction in the annual forecast error. The use of time-series methodologies to forecast the collection of federal revenues has proved to be a viable alternative to the method of indicators, contributing for more accurate predictions, becoming a safe support tool for the managers decision making process.

Key Words: Collection. Forecasting. Time Series. ARIMA. univariate and multivariate dynamic models. Structural Models.

LISTA DE ILUSTRAÇÕES

Gráficos 1 e 2 - Série do Imposto de Importação (em R\$ 10 ⁶ deflacionado pelo IER e em logaritmo).....	98
Gráficos 3 e 4 - Série do Imposto Sobre a Renda da Pessoa Jurídica – Demais Empresas Obrigadas a Apuração do Lucro Real (em R\$ 10 ⁶ deflacionado pelo IER e em logaritmo).....	98
Gráficos 5 e 6 - Série da Contribuição para o Financiamento da Seguridade Social (em R\$ 10 ⁶ deflacionado pelo IER e em logaritmo).....	99
Gráfico 7 - Estatísticas descritivas – Série LII.....	99
Gráfico 8 - Estatísticas descritivas – Série LLR.....	99
Gráfico 9 - Estatísticas descritivas – Série LCF.....	100
Gráfico 10 - Estatísticas descritivas – Série LPIB.....	100
Gráfico 11 - Estatísticas descritivas – Série LIVV.....	100
Gráfico 12 - Estatísticas descritivas – Série LIPI.....	101
Gráfico 13 - Estatísticas descritivas – Série LTCR.....	101
Gráfico 14 – <i>Boxplots</i> – Série LII.....	101
Gráfico 15 – <i>Boxplots</i> – Série LLR.....	102
Gráfico 16 – <i>Boxplots</i> – Série LCF.....	102
Gráfico 17 – Correlograma em nível da série LII.....	102
Gráfico 18 – Correlograma em nível da série LLR.....	103
Gráfico 19 – Correlograma em nível da série LCF.....	103
Gráfico 20 – Correlograma em primeira diferença da série LII.....	104
Gráfico 21 – Correlograma em primeira diferença da série LLR.....	104
Gráfico 22 – Correlograma em primeira diferença mais primeira diferença sazonal da série LLR.....	105
Gráfico 23 – Séries Índice de Vendas no Varejo – São Paulo, Índice de Produção Industrial – Geral – São Paulo e Produto Interno Bruto - Índice.....	105
Gráfico 24 – <i>Scatter plot</i> IVV x IPI, IVV x PIB, IPI x PIB_I.....	106
Gráfico 25 – Correlograma da série Lpib.....	106
Gráfico 26 – Correlograma da série LIVV.....	107
Gráfico 27 – Correlograma da série LIPI.....	107
Gráfico 28 – Série Taxa de Câmbio Real – Índice.....	108
Gráfico 29 – Correlograma da série LTCR.....	108

LISTA DE TABELAS

Tabela 1 - Método dos Indicadores - Erro percentual absoluto médio da previsão - MAPE...	14
Tabela 2: Modelos básicos de séries de tempo com tendência.....	33
Tabela 3 - Teste RU - DF-GLS (ERS)	51
Tabela 4 - Teste RU - Ng-Perron.....	52
Tabela 5 - Teste RU - DF-GLS (ERS)	52
Tabela 6 - Teste RU - Ng-Perron.....	52
Tabela 7 – Resultados dos testes de exogeneidade das variáveis explicativas em modelos multivariados	56
Tabela 8 – Modelagem ARIMA - Série: II – Intervalo dos dados: 04/2002 a 12/2006	63
Tabela 9 – Modelagem ARIMA - Série: IRPJ_DOLR – Intervalo dos dados: 10/2002 a 12/2006	64
Tabela 10 – Modelagem ARIMA - Série: COFINS_DE – Intervalo dos dados: 06/2002 a 12/2006	65
Tabela 11 – Modelos dinâmicos univariados e de Função de Transferência – II/IRPJ_DOLR/COFINS_DE	66
Tabela 12 – Modelos VAR/VEC/SEM – II - Intervalo dos dados: 07/2000 a 12/2006.....	69
Tabela 13 - Análise de cointegração I(1) - II.....	70
Tabela 14 - Modelo VEC/SEM DLII/DLIVV/DLTCR – modelo nº 3 – Tabela 12	70
Tabela 15 - Modelo VEC/SEM DLII/DLIVV/DLTCR – modelo nº 6 – Tabela 12	71
Tabela 16 - Modelo VEC/SEM DLII/DLIPI/DLTCR – modelo nº 9 – Tabela 12.....	72
Tabela 17 - Modelo VEC/SEM DLII/DLPIB/DLTCR – modelo nº 12 – Tabela 12	73
Tabela 18 – Modelos VAR/VEC/SEM – IRPJ_DOLR - Intervalo dos dados: 02/2001 a 12/2006	73
Tabela 19 - Análise de cointegração I(1) – IRPJ_DOLR.....	74
Tabela 20 - Modelo VEC/SEM DLLR/DLPIB – modelo nº 3 – Tabela 18.....	75
Tabela 21 - Modelo VEC/SEM DLLR/DLIVV – modelo nº 6 – Tabela 18.....	76
Tabela 22 - Modelo VEC/SEM DLLR/DLIPI – modelo nº 9 – Tabela 18	76
Tabela 23 – Modelos VAR/VEC/SEM – COFINS - Intervalo dos dados: 02/2001 a 12/2006	77
Tabela 24 - Análise de cointegração I(1) – COFINS	78
Tabela 25 - Modelo VEC/SEM DLCPF/DLPIB – modelo nº 3 – Tabela 23	78
Tabela 26 - Modelo VEC/SEM DLCPF/DLIVV – modelo nº 6 – Tabela 23	79
Tabela 27 - Modelo VEC/SEM DLCPF/DLIPI – modelo nº 9 – Tabela 23.....	80

Tabela 28 - Modelos Estruturais – EM e EMX – II/IRPJ_DOLR/COFINS_DE.....	81
Tabela 29 - Comparação do RMSE e do erro anual de previsão do melhor modelo de cada metodologia	83
Tabela 30 - Arrecadação Realizada 2007 – São Paulo - em R\$1.000.000,00.....	91
Tabela 31 – Testes de igualdade de variância	93
Tabela 32 - Valores dos logaritmos das séries/ 10^6 deflacionados pelo IER	94
Tabela 33 – Detecção inicial de <i>outliers</i> – variáveis principais	95
Tabela 34 – Detecção inicial de <i>outliers</i> – variáveis explicativas.....	96
Tabela 35 – Principais modelos e componentes estruturais de séries de tempo.....	97

LISTA DE SIGLAS

ACF	Função de autocorrelação.
ADF	Teste de raiz unitária Dickey-Fuller aumentado.
AIC	Critério de informação de Akaike.
AO	<i>Outlier</i> do tipo aditivo.
AR	Modelo ou componente auto-regressivo.
ARIMA	Modelo auto-regressivo integrado de médias móveis.
ARMA	Modelo auto-regressivo de médias móveis.
BIC/SBC	Critério Bayesiano de Schwartz.
COFINS_DE	Série da Contribuição para o Financiamento da Seguridade Social – Demais Empresas.
DF-GLS (ERS)	Teste de raiz unitária Elliott-Rothenberg-Stock.
DGP	Processo gerador de dados.
DR	Modelo dinâmico univariado.
DS	Tendência estocástica ou diferença estacionária.
E_{anual}	Erro anual de previsão.
EM	Modelo Estrutural.
EMX	Modelo Estrutural com inclusão de variável exógena.
IER	Índice Específico da Receita.
II	Série do Imposto de Importação.
IPI	Série do Índice de Produção Industrial – Indústria Geral – São Paulo.
IRPJ_DOLR	Série do Imposto sobre a Renda das Pessoas Jurídicas – Demais Obrigadas ao Lucro Real.
IVV	Série do Índice de Vendas no Varejo – Total – São Paulo.
LCF	Logaritmo da série COFINS_DE.
LII	Logaritmo da série II.
LLR	Logaritmo da série IRPJ_DOLR.
LS	<i>Outlier</i> do tipo mudança estrutural.
MA	Modelo ou componente média móvel.
MAD	Desvio absoluto médio.
MAPE	Erro percentual absoluto médio.
MI	Método dos indicadores.
NgP	Teste de raiz unitária de Ng-Perron.
NID $(0, \sigma^2)$	Variável aleatória, com média zero e variância σ^2 , normalmente distribuída e serialmente independente.
PACF	Função de autocorrelação parcial.
PDF	Função densidade de probabilidade.
PP	Teste de raiz unitária Phillips-Perron.
RW	Processo <i>random walk</i> ou passeio aleatório.

RFB	Secretaria da Receita Federal do Brasil.
RMSE	Raiz do erro quadrático médio.
SEM	Modelo de equações simultâneas.
SRRF/8ªRF	Superintendência da Receita Federal do Brasil da 8ª Região Fiscal.
TCR	Série da Taxa de Câmbio Real.
TFM	Modelo de Função de Transferência.
TS	Tendência estacionária.
URF	Forma reduzida irrestrita.
VAR	Modelo auto-regressivo vetorial.
VEC	Modelo VAR com correção de erro - Cointegrado.
WN	Processo <i>white noise</i> ou ruído branco.

SUMÁRIO

Introdução.....	13
1 Revisão bibliográfica.....	16
2 Marco teórico.....	19
2.1 Modelo teórico para as receitas dos impostos	19
2.2 Método dos indicadores.....	26
2.3 Modelagem Box-Jenkins – ARIMA.....	29
2.4 Modelo de Função de Transferência	36
2.5 Modelos dinâmicos de múltiplas equações	39
2.6 Modelo Estrutural	43
3 Procedimentos metodológicos.....	47
4 Análise das séries	49
4.1 Testes de igualdade de variâncias.....	49
4.2 Gráficos	50
4.3 Estatísticas descritivas	50
4.4 Correlogramas das séries	51
4.5 Testes de Raiz Unitária.....	51
4.6 Detecção inicial de <i>outliers</i>	53
4.7 Testes de exogeneidade das variáveis explicativas	55
5 Modelos e métodos de estimação	58
5.1 Métodos de comparação da previsão.....	60
6 Resultados.....	62
6.1 Modelagem Box-Jenkins – ARIMA.....	62
6.2 Modelos dinâmicos univariados e de Função de Transferência.....	65
6.3 Modelos dinâmicos de múltiplas equações	68
6.4 Modelos Estruturais.....	80
7 Conclusões.....	83
REFERÊNCIAS	88
APÊNDICE A - Tabelas.....	91
APÊNDICE B – Ilustrações	98
APÊNDICE C – Resultados dos principais modelos estimados	109

Introdução

Dentre as missões da Secretaria da Receita Federal do Brasil (RFB), destaca-se a de prover o Estado de recursos para garantir o bem-estar social. Uma das formas de se garantir o sucesso dessa missão se faz por meio de uma boa previsão da arrecadação das receitas federais. Entre as finalidades da RFB, conforme dispõe o inciso VIII do art. 1º do anexo de seu Regimento Interno, aprovado pela Portaria do Ministro de Estado da Fazenda – MF nº 95, de 30 de abril de 2007, publicada no D.O.U. de 2 de maio de 2007, consta: “realizar a previsão, o acompanhamento, a análise e o controle das receitas sob sua administração, bem como coordenar e consolidar as previsões das demais receitas federais, para subsidiar a elaboração da proposta orçamentária da União.” Pode-se considerar que subsidiar a elaboração da proposta orçamentária da União é um dos objetivos da previsão das receitas federais.

Porém, ainda restam as atividades do acompanhamento, da análise e do controle das receitas federais, em que se vislumbram outros objetivos, dentre os quais pode-se destacar:

- comparar o valor efetivamente arrecadado com o valor previsto, para cada tributo, propondo, se for o caso, ajustes nos valores previstos para o restante do ano;
- acompanhar as alterações do cenário macroeconômico e analisar sua influência sobre os valores arrecadados;
- identificar possíveis causas de desvios do valor arrecadado em relação ao previsto;
- auxiliar a atividade de fiscalização;
- verificar a eficácia do método de previsão adotado.

Na RFB as atividades de acompanhamento, de análise e de controle das receitas federais são efetuadas tanto em nível nacional, quanto regional e local. Ultimamente estas atividades têm adquirido grande importância no contexto institucional, juntamente com o acompanhamento dos grandes contribuintes (considerando que a arrecadação é bastante concentrada), em função da crescente necessidade de recursos demandados pelo Estado e a dificuldade de se elevar tributos, numa sociedade já saturada de tais aumentos.

A RFB utiliza, como se verá com mais detalhes em 2.2, um método de previsão das receitas federais, denominado pelo Órgão de “Método dos Indicadores”, em que, basicamente, a arrecadação prevista é igual à arrecadação do mesmo período do ano anterior multiplicada por índices que reflitam as variações de preço, de quantidade e da legislação no período. A

previsão é feita em nível regional¹, excluindo-se eventuais arrecadações atípicas, e o rateio entre as unidades da RFB é efetuado com base na participação de cada unidade na arrecadação realizada no ano anterior, levando-se em consideração, ainda, eventuais alterações de jurisdição de grandes empresas. A soma das previsões efetuadas por cada uma das Regiões Fiscais compõe a previsão nacional de arrecadação.

Este método, embora não seja complexo do ponto de vista teórico, é extremamente trabalhoso, em função, principalmente, do número de tributos a serem previstos (quatorze tributos que, com as subclassificações dão origem a cinquenta e cinco séries); da quantidade de índices de preço e quantidade envolvidos; da dificuldade de se obter um índice de alterações na legislação e da necessidade de se excluir as arrecadações atípicas ocorridas. Ademais, o método não tem levado a previsões precisas, conforme pode-se observar na tabela 1 abaixo, em que é apresentado o erro de previsão de alguns tributos, dentre os quais se encontram os que serão analisados neste trabalho, no período de 2001 a 2007, medido pelo erro percentual absoluto médio (MAPE²), que considera o erro relativo de cada previsão.

Tabela 1 - Método dos Indicadores - Erro percentual absoluto médio da previsão - MAPE

SÉRIE	2001	2002	2003	2004	2005	2006	2007	Média
II	19,2%	20,0%	16,1%	29,5%	28,4%	11,1%	5,7%	18,6%
IPI_OP	7,8%	8,5%	22,6%	10,9%	9,0%	5,3%	9,0%	10,5%
IRPJ_DOLR	27,2%	31,8%	21,4%	34,1%	20,8%	14,1%	19,3%	24,1%
IRPJ_NOLR	13,7%	4,8%	7,6%	12,2%	8,8%	4,4%	8,5%	8,6%
IRRF_T	6,8%	8,2%	5,5%	5,6%	7,5%	12,5%	3,9%	7,1%
COFINS_DE	4,0%	7,2%	7,8%	8,4%	4,1%	5,5%	5,6%	6,1%
PIS_DE	7,1%	9,0%	14,1%	5,3%	4,4%	5,1%	6,7%	7,4%
CSLL_DE	12,6%	21,5%	18,1%	23,1%	13,0%	10,2%	13,5%	16,0%
RA	4,8%	13,9%	4,0%	8,7%	5,7%	7,1%	10,3%	7,8%

Fonte: elaboração do autor.

Notas: II - Imposto de Importação.

IPI_OP - Imposto sobre Produtos Industrializados – Outros Produtos.

IRPJ_DOLR, IRPJ_NOLR - Imposto de Renda das Pessoas Jurídicas – Demais Obrigadas ao Lucro Real e Não Obrigadas ao Lucro Real.

IRRF_T - Imposto sobre a Renda Retido na Fonte – Rendimentos do Trabalho.

COFINS_DE - Contribuição para o Financiamento da Seguridade Social – Demais Empresas.

PIS_DE - Contribuição para o Programa de Integração Social - Demais Empresas.

CSLL_DE - Contribuição Social sobre o Lucro Líquido – Demais Empresas.

RA - Receita Administrada pela RFB.

¹ A RFB é estruturada, basicamente, em Órgãos Centrais e dez unidades descentralizadas, denominadas Superintendências Regionais da Receita Federal do Brasil (SRRF). As SRRF, por sua vez, jurisdicionam regiões fiscais, sendo o Estado de São Paulo a 8ª Região Fiscal. Estas se compõem, principalmente, de unidades locais denominadas Delegacias da Receita Federal do Brasil (DRF). A estrutura completa do órgão pode ser encontrada em: www.receita.fazenda.gov.br

² A definição do MAPE se encontra em 5.1.

Conforme se verifica, há uma tendência de diminuição dos erros de previsão, medidos pelo MAPE, ao longo dos anos, provavelmente devido a maior estabilidade econômica dos últimos anos. Deve-se ressaltar que o método dos indicadores funciona melhor quanto mais estável o ambiente econômico. Contudo, os erros de previsão são bastante elevados e variam muito de um ano para outro.

Este trabalho tem abrangência regional e limita-se à análise e previsão de três séries mensais da arrecadação, relativas ao Imposto de Importação, Imposto Sobre a Renda das Pessoas Jurídicas – Demais Obrigadas a Apuração do Lucro Real e Cofins – Demais Empresas, no âmbito da jurisdição do estado de São Paulo, no período de 2000 a 2007.

O problema de investigação pode ser expresso pelo seguinte questionamento: A utilização de metodologias de séries temporais pode melhorar a acurácia da previsão da arrecadação tributária federal, no âmbito da jurisdição da SRRF/8ªRF?

Para tanto, foram aplicadas as metodologias de Box-Jenkins (ARIMA), modelos dinâmicos univariados (DR), multivariados, quais sejam, Função de Transferência (TFM), Auto-regressão Vetorial (VAR), VAR com correção de erro (VEC), Equações Simultâneas (SEM), e de modelos Estruturais (EM). Nos modelos multivariados, além da variável principal de arrecadação, foram incluídas variáveis explicativas, que, acredita-se, tenham influência na arrecadação tributária e levem à diminuição dos erros de previsão.

A metodologia Box-Jenkins (ARIMA), aplicada à arrecadação tributária federal, por já ter sido desenvolvida em outros trabalhos, tais como, Melo (2001) e Siqueira (2002), não é o objetivo principal deste trabalho e foi desenvolvida com o objetivo de, juntamente com o método dos indicadores, servir de base de comparação para as demais metodologias acima citadas.

Portanto, o objetivo principal do trabalho é oferecer métodos alternativos de previsão da arrecadação tributária federal, baseados em metodologias de séries temporais, inclusive com a utilização de variáveis explicativas, que reflitam a influência do cenário macroeconômico na arrecadação tributária, com o intuito de se melhorar a acurácia da previsão da arrecadação.

1 Revisão bibliográfica

Quanto à revisão bibliográfica, diversos trabalhos, dentre os quais podem-se citar Melo (2001) e Siqueira (2002), propõem metodologias alternativas à atualmente utilizada pela RFB, para a atividade de previsão da arrecadação tributária. Melo (2001) propõe as metodologias de modelos auto-regressivos integrados de médias móveis (ARIMA, metodologia Box-Jenkins) e alisamento exponencial de Holt-Winters sazonal aditivo. Analisa séries do Imposto Sobre a Renda e suas subclassificações, no período de julho de 1994 a dezembro de 2000. Conclui que os dois modelos se mostraram superiores ao método dos indicadores utilizado pela RFB, gerando previsões mais acuradas. A metodologia proposta permitiu reduzir o erro de previsão médio de 10% para 0,17% aproximadamente, para o período de janeiro de 2000 a dezembro de 2000.

Siqueira (2002) também analisa a modelagem ARIMA, porém, para séries que apresentam ciclos estocásticos, apresenta um modelo dinâmico de regressão que incorpora variações de calendário (TDV – *Trading Day Variation*). Sua análise se baseia em séries de vários tributos federais (dez séries), no período de janeiro de 1989 a outubro de 2001. Conclui pela superioridade dos métodos escolhidos em relação ao método dos indicadores em oito das dez séries analisadas; que, apesar do método dos indicadores não fornecer previsões estatisticamente confiáveis, mostrou-se superior ao método em estudo para as séries da Cofins adicionada ao PIS e Outras Contribuições Sociais, devido, principalmente, à diversas alterações legais em variáveis significantes da arrecadação que não são devidamente capturadas pela metodologia ARIMA. Conclui, ainda, que as técnicas utilizadas dependem da natureza do tributo e não há padrão no tempo sobre o qual se podem fundar todas as previsões; que os modelos dinâmicos, em casos bem específicos, podem melhorar as previsões feitas pelos modelos ARIMA, e, por fim, que a abordagem de séries temporais normalmente apresenta melhor performance que a modelagem causal, pelo menos no curto prazo.

Santana (2004) utiliza modelos ARIMA, amortecimento exponencial de Holt-Winters e uma modelagem usando variáveis exógenas, para uma série do Imposto sobre Produtos Industrializados, no período de janeiro de 1995 a setembro de 2004. Comparam-se os três modelos entre si e com o método dos indicadores, concluindo pela superioridade deste último. Apresenta como possíveis explicações: primeiro o fato de ter sido utilizada a arrecadação bruta, sem eliminação de arrecadações atípicas; em segundo lugar, o fato do Imposto sobre Produtos Industrializados ter apresentado nos primeiros meses de 2004, desempenho superior

ao que vinha apresentando nos últimos anos devido ao alto crescimento econômico verificado no primeiro semestre de 2004, o que não é captado pelos modelos de séries temporais; e finalmente, os modelos estudados não levam em conta efeitos da alteração da legislação, muito presente no caso do Imposto sobre Produtos Industrializados, que tem sido usado como instrumento regulador de alguns setores econômicos, como o automobilístico, por exemplo, com variações freqüentes de alíquotas.

Finalmente, Peceguini (2001) utiliza modelos aritméticos (método convencional) e modelos econométricos de linha de tendência, séries de tempo – decomposição clássica e modelos ARIMA, aplicados à arrecadação do ICMS no Estado de São Paulo, no período de janeiro de 1995 a dezembro de 2000. Destaca-se, como o melhor método, aquele baseado na decomposição clássica de série temporal, que se constitui em uma mescla de modelo analítico e modelo de previsão. Sua desvantagem é exigir boas previsões para os valores das variáveis explicativas do comportamento da arrecadação. Os modelos de linha de tendência e o método convencional ficaram em segundo lugar e, por fim, a modelagem ARIMA. Todos os métodos têm suas vantagens e desvantagens, de maneira que, mesmo o modelo auto-regressivo, apesar de sua classificação, não deve ser descartado, visto que apresenta a vantagem de requerer exclusivamente dados (série histórica) da arrecadação do ICMS – SP.

Peceguini (2001) enfatiza, também, projeto executado na Secretaria de Estado dos Negócios da Fazenda do Governo do Estado de São Paulo, em que consultoria externa desenvolveu um modelo de previsão da arrecadação, baseado em Função de Transferência, apresentando bons resultados.

Esse modelo baseado em Função de Transferência, não foi objeto de exame neste trabalho. Uma de suas vantagens é o emprego de variáveis/indicadores antecedentes, o que permite uma avaliação do comportamento da arrecadação, antecipando reversão de tendência, porquanto não somente o valor mensal da arrecadação é importante, mas também a mudança de comportamento, de aumento real passando para diminuição real, ou vice-versa.

Outra vantagem do emprego de Função de Transferência advém do efeito defasado de uma determinada variável sobre a arrecadação. Suponhamos, por exemplo, que a taxa real de juros no mês "t-5" repercute na arrecadação do ICMS no mês "t". No caso deste exemplo, ao efetuarmos uma previsão com horizonte de 6 (seis) meses a partir de hoje, bastará estimar (projetar) um único valor futuro para a taxa real de juros. Isto não acontece no método da decomposição clássica de série de tempo, examinado neste trabalho, que requer, ainda neste exemplo, a previsão dos valores mensais da taxa real de juros para os seis meses vindouros. (PECEGUINI, 2001, p. 75)

Um ponto comum nas sugestões de trabalhos futuros propostas é a adição de variáveis explicativas (exógenas), tais como PIB, taxa de juros, consumo de energia elétrica, nível de emprego etc., nos modelos ARIMA, formando, assim, os modelos de Função de Transferência, que será um dos aplicados no presente trabalho.

Quanto à comparação de métodos de previsão aplicados a outros tipos de séries, Andrews (1994) compara o poder preditivo do modelo estrutural de séries de tempo com modelos ARIMA, Bayesian, Leawandowski (FORSYS) e auto-regressão de processos auto-regressivos média móvel (ARARMA) para 111 séries econômicas e de negócios, selecionadas em Makridakis *et alli* (1982). Conclui que, para dados sazonais e utilizando-se como medida de acurácia a média do MAPE, o modelo estrutural tem performance substancialmente melhor que os outros quatro métodos, especialmente para horizontes mais longos de previsão, embora a performance do modelo estrutural seja menos significativa quando se usa a mediana do MAPE.

2 Marco teórico

2.1 Modelo teórico para as receitas dos impostos

2.1.1 Imposto de Renda

De acordo com Barro (1993), um modelo simples para a receita tributária em um sistema com imposto de renda (pago tanto pelos produtores quanto pelos trabalhadores) assume que o imposto total arrecadado, em termos reais, $\mathbf{T}_t/\mathbf{P}_t$, é uma fração τ^3 da renda real sujeita ao imposto, ou seja:

$$\frac{\mathbf{T}_t}{\mathbf{P}_t} = \tau(\mathbf{Y}_t - \delta\mathbf{K}_{t-1} - \mathbf{E}_t),$$

em que, $\mathbf{Y}_t - \delta\mathbf{K}_{t-1}$ é o produto real líquido (com desconto da depreciação) e \mathbf{E}_t representa a renda total isenta do imposto. Assim, assumindo-se constante a taxa média do imposto, pode-se considerar que a receita tributária total é proporcional à renda sujeita à incidência do imposto.

Por outro lado, conforme asseveram Musgrave e Musgrave (1980), o modelo de determinação da renda e multiplicadores fiscais⁴, num sistema com o consumo \mathbf{C} , investimento \mathbf{I} , compras governamentais \mathbf{G} , e Imposto de Renda com alíquota \mathbf{t} , proporcional à renda \mathbf{Y} , conforme acima, pode ser descrito pelas seguintes equações:

$$\mathbf{Y} = \mathbf{C} + \mathbf{I} + \mathbf{G},$$

$$\mathbf{C} = \mathbf{a} + \mathbf{c}(1 - \mathbf{t})\mathbf{Y},$$

$$\mathbf{Y} = \frac{1}{1 - \mathbf{c}(1 - \mathbf{t})}(\mathbf{a} + \mathbf{I} + \mathbf{G}),$$

A constante \mathbf{a} se refere ao consumo autônomo e \mathbf{c} é a propensão marginal a consumir. O termo $1/(1 - \mathbf{c}(1 - \mathbf{t}))$ é o multiplicador fiscal. Para se saber a influência de alterações em \mathbf{G} , \mathbf{I} e em \mathbf{t} , mantidas as demais variáveis constantes, na renda de equilíbrio, aplica-se a derivada total na equação de \mathbf{Y} .

$$d\mathbf{Y}\Big|_{\substack{dt=0 \\ dI=0}} = \frac{1}{1 - \mathbf{c}(1 - \mathbf{t})} d\mathbf{G},$$

³ taxa média do imposto, assumida constante para toda faixa de renda por simplicidade.

⁴ No modelo considerado, por questão de simplicidade de exposição, o que não altera as conclusões básicas do modelo, o investimento, tratado como sendo fixo, e as compras governamentais são consideradas variáveis exógenas. Admite-se, ainda, que os salários nominais são rígidos para baixo, o que implica que os preços não podem cair, e que há um significativo desemprego de recursos na economia, de forma que um aumento na demanda agregada elevará a produção real sem afetar o nível de preços.

$$dY \Big|_{\substack{dt=0 \\ dG=0}} = \frac{1}{1-c(1-t)} dI,$$

$$dY \Big|_{\substack{dI=0 \\ dG=0}} = \frac{-cY}{[1-c(1-t)]} dt.$$

Neste modelo simples, fica claro que a existência de um sistema tributário dependente da renda reduz o tamanho do multiplicador fiscal e, portanto, amortece o impacto no produto de qualquer alteração exógena nos gastos privados. Assim, definindo-se o *deficit* fiscal (primário) como $G - T$, a consequência da existência do estabilizador automático é que o *deficit* fiscal aumenta quando a renda cai e vice-versa.

Segundo Carlin e Soskice (2006), para interpretar o significado do *deficit* fiscal registrado em qualquer tempo, é necessário saber se o produto está abaixo, acima ou no nível de equilíbrio. Para isso, é calculado o que se chama “*deficit* fiscal ajustado ciclicamente”, que nada mais é que o *deficit* que prevaleceria se a economia estivesse operando no produto de equilíbrio⁵, dados o sistema tributário vigente e os compromissos de gastos governamentais. O equilíbrio fiscal pode ser assim definido:

$$\begin{aligned} \text{deficit fiscal primário} &\equiv \text{deficit fiscal ajustado ciclicamente} + \text{impacto do estabilizador} \\ &\quad \text{automático} \\ &\equiv \text{estímulo fiscal discricionário} + \text{impacto do estabilizador} \\ &\quad \text{automático} \end{aligned}$$

$$G(Y_t) - T(Y_t) \equiv [G(Y_e) - T(Y_e)] + a(Y_e - Y_t),$$

em que a é uma constante e o termo $a(Y_e - Y_t)$ captura o impacto do estabilizador automático no *deficit* fiscal. Se o produto corrente, Y_t , está abaixo do produto ao nível de equilíbrio, Y_e , a economia está em recessão, o termo $a(Y_e - Y_t) > 0$ e a existência do estabilizador automático ajuda a retorná-la ao ponto de equilíbrio por meio do aumento dos gastos do governo com transferências, como por exemplo, para o seguro desemprego, e a diminuição das receitas dos impostos, aumentando, assim, o *deficit*. Se o *deficit* fiscal ajustado ciclicamente é zero, o que caracteriza ausência de estímulo fiscal discricionário, o *deficit* real simplesmente reflete a atuação do estabilizador automático e irá desaparecer assim que a economia retorne à posição de equilíbrio.

Ao se analisar as variações nos valores de G e T deve-se fazer distinção entre alterações provocadas por mudanças nos parâmetros fiscais, daquelas provenientes da

⁵ Geralmente, um problema importante é a determinação do produto de equilíbrio.

flexibilidade embutida no sistema fiscal, isto é, reações automáticas às alterações ocorridas no setor privado. No primeiro caso, do lado das receitas, uma modificação nos parâmetros fiscais pode ser caracterizada pela extinção de um imposto ou redução de alíquotas tributárias. Já no segundo caso, ajustamentos automáticos, pode-se ter alteração na receita tributária devido a uma mudança na base dos impostos, como a que pode ocorrer com a receita do IRPJ em resposta à alterações no lucro das empresas. Assim, a própria existência de um setor público configura um fator de estabilização, na medida em que as reações automáticas da receita tributária amortecem as repercussões das mudanças iniciais nos níveis dos gastos privados.

A medida do efeito amortecedor que o sistema fiscal exerce em relação às alterações originadas no setor privado, é dada pelo índice de estabilização automática, α , que mede a porcentagem de alteração na renda que não ocorre em função do efeito amortecedor gerado pela elevação das receitas tributárias, e é expresso por:

$$\alpha = \frac{ct}{1 - c(1 - t)}.$$

Conforme se verifica, a magnitude do estabilizador automático depende do tamanho do orçamento em relação à renda. Nos casos limite de $t = 0$ e $t = 1$, α terá o valor mínimo de zero e máximo de c . Na verdade o que importa é a variação da receita total T decorrente de uma modificação em Y , o que está relacionado com a elasticidade da receita em relação às alterações na renda. A elasticidade global depende, por sua vez, da estrutura tributária vigente, já que diverge consideravelmente entre os diversos impostos. Porém, Musgrave e Musgrave (1980) ressaltam que um imposto tem efeitos estabilizadores automáticos sempre que a sua elasticidade renda seja positiva, não sendo necessário que seja maior ou igual a um.

Pelo acima exposto conclui-se que, dado o sistema tributário vigente, um aumento da renda eleva as receitas tributárias que, por sua vez, no período seguinte diminui a renda disponível e conseqüentemente o produto. Assim, fica difícil estabelecer uma relação de causa e efeito entre as duas variáveis e surge o problema da exogeneidade da variável explicativa⁶ a ser utilizada nos modelos de Função de Transferência e Estrutural com variáveis explicativas. Sendo este o caso, a abordagem de auto-regressão vetorial (VAR), em que todas as variáveis são tratadas como endógenas, é a mais apropriada.

Portanto, na estimação de modelos para previsão da arrecadação do IRPJ, utiliza-se como variável explicativa o PIB ou alguma variável *proxy* do PIB, já que o IRPJ, no caso da série IRPJ_DOLR, é um tributo que depende da apuração do lucro real pelas empresas, que,

⁶ Uma explicação mais elaborada desta questão se encontra em 2.4.1

em última instância, está ligado ao nível da atividade econômica, embora esta relação possa não ser tão direta.

O desenvolvimento acima, embora se aplique ao sistema econômico como um todo, em que a estrutura tributária vigente é predominantemente influenciada pelo imposto de renda, entendo, se aplica, no caso concreto do presente trabalho, tanto às receitas advindas do IRPJ quanto às da Cofins. Neste último caso, trata-se de tributo que depende do faturamento das empresas, que, por sua vez, está diretamente ligado ao nível de atividade econômica. Já para o Imposto de Importação, devido a sua especificidade, pois depende do fluxo de transações externas, a abordagem deve ser diferente, como se verá a seguir.

2.1.2 Imposto de Importação

A receita do Imposto de Importação (II) é função da demanda de bens importados, da taxa real de câmbio⁷ e da alíquota média dos tributos devidos na importação. A demanda de bens importados, por sua vez, é função da renda doméstica, da taxa real de câmbio e dos tributos devidos na importação. Quanto maior a renda doméstica, maior a demanda por bens, sejam domésticos ou estrangeiros. No que diz respeito à taxa real de câmbio, quanto mais caros os bens domésticos relativamente aos bens estrangeiros, maior a demanda por bens importados. Assim, uma baixa taxa real de câmbio conduz ao aumento das importações. No caso dos tributos devidos na importação, quanto maior o nível de tributação, menor será a demanda por bens importados. Esta relação pode ser assim expressa, em que **M** é a demanda por bens importados, **Y** é a renda doméstica, **ε** a taxa real de câmbio e **t** a alíquota média dos tributos devidos na importação:

$$M = M(Y_{(+)}, \varepsilon_{(-)}, t_{(-)}). \quad (1)$$

Segundo Romer (2006), numa economia aberta e em equilíbrio, considerando que há barreiras ao ajuste instantâneo de preços e salários, o gasto real planejado, **E**, iguala o produto **Y**, o que pode ser expresso por:

⁷ A taxa nominal de câmbio pode ser definida como o preço da moeda estrangeira em termos da moeda doméstica, a qual denomina-se **e**. Para o Brasil, a taxa real de câmbio, **ε**, é construída pela multiplicação do nível de preços doméstico, **P**, pela taxa nominal de câmbio, **e**, dividido pelo nível de preços estrangeiro, **P***, ou seja **ε = eP/P***. Assim, uma diminuição (aumento) na taxa real de câmbio, isto é, um aumento (diminuição) no preço relativo dos bens domésticos em relação aos bens estrangeiros, é chamado de apreciação (depreciação) real. Na realidade, a receita do II é função da taxa nominal de câmbio. Porém, considerando que: i) a relação entre a taxa nominal e a real (**e/ε**) é igual a uma constante e ii) a demanda de bens importados, **M**, utilizada no decorrer desta sub-seção em um sistema de equações, é função da taxa real de câmbio, considerar a receita do II função da taxa real de câmbio simplifica a análise e não altera o conceito básico que se pretende mostrar.

$$Y = E(Y_{(+)}, r_{(-)}, T(Y)_{(-)}, G_{(+)}, \varepsilon_{(+)}), \quad (2)$$

com E crescente em ε , ou seja, um aumento de ε corresponde a uma depreciação da moeda doméstica em relação à moeda estrangeira; r é a taxa de juros real, $T(Y)$, a receita de impostos, considerada como função da renda, e G , o gasto governamental, considerado exógeno.

De acordo com Romer (2006), a variável r , em um modelo em que a oferta monetária é tratada de forma endógena, é função da renda e do nível de inflação (π), ou seja:

$$r = r(Y_{(+)}, \pi_{(+)}). \quad (3)$$

O sistema formado pelas equações (1), (2) e (3) pode ser tratado como um sistema não linear de funções implícitas se existir uma divisão entre as variáveis exógenas e endógenas, de tal modo que, se substituirmos valores numéricos para as variáveis exógenas, o sistema resultante possa ser resolvido de maneira única para os valores endógenos correspondentes (SIMON E BLUME 2004, p. 361). Aplicando-se o teorema da função implícita ao sistema de equações, tem-se:

$$\begin{aligned} dM &= \frac{\partial M}{\partial Y} dY + \frac{\partial M}{\partial \varepsilon} d\varepsilon + \frac{\partial M}{\partial t} dt, \\ dY &= \frac{\partial E}{\partial Y} dY + \frac{\partial E}{\partial r} dr + \frac{\partial E}{\partial T} T'(Y) dY + \frac{\partial E}{\partial G} dG + \frac{\partial E}{\partial \varepsilon} d\varepsilon. \\ dr &= \frac{\partial r}{\partial Y} dY + \frac{\partial r}{\partial \pi} d\pi. \end{aligned}$$

Chamando:

$$\begin{aligned} M_Y &= \frac{\partial M}{\partial Y} > 0, M_\varepsilon = \frac{\partial M}{\partial \varepsilon} < 0, M_t = \frac{\partial M}{\partial t} < 0; \\ E_Y &= \frac{\partial E}{\partial Y} > 0, e < 1, E_r = \frac{\partial E}{\partial r} < 0, E_T = \frac{\partial E}{\partial T} < 0, \\ E_G &= \frac{\partial E}{\partial G} > 0, E_\varepsilon = \frac{\partial E}{\partial \varepsilon} > 0, T'(Y) = \frac{\partial T}{\partial Y} = T_Y > 0; \\ r_Y &= \frac{\partial r}{\partial Y} > 0, r_\pi = \frac{\partial r}{\partial \pi} > 0. \end{aligned}$$

Na forma matricial:

$$\begin{bmatrix} 1 & -M_Y & 0 \\ 0 & A & -E_r \\ 0 & -r_Y & 1 \end{bmatrix} \begin{bmatrix} dM \\ dY \\ dr \end{bmatrix} = \begin{bmatrix} M_\varepsilon d\varepsilon + M_t dt \\ E_\varepsilon d\varepsilon + E_G dG \\ r_\pi d\pi \end{bmatrix}.$$

Em que:

$$A = 1 - E_Y - E_T T_Y \text{ e } D = 1 - E_Y - E_T T_Y - r_Y E_r = A - r_Y E_r.$$

Resolvendo para dM , dY e dr :

$$\begin{bmatrix} dM \\ dY \\ dr \end{bmatrix} = \begin{bmatrix} 1 & -M_Y & 0 \\ 0 & A & -E_r \\ 0 & -r_Y & 1 \end{bmatrix}^{-1} \begin{bmatrix} M_\varepsilon d\varepsilon + M_t dt \\ E_\varepsilon d\varepsilon + E_G dG \\ r_\pi d\pi \end{bmatrix}.$$

Aplicando-se a regra de Cramer para inversão de matrizes:

$$dM = \frac{\begin{vmatrix} M_\varepsilon d\varepsilon + M_t dt & -M_Y & 0 \\ E_\varepsilon d\varepsilon + E_G dG & A & -E_r \\ r_\pi d\pi & -r_Y & 1 \end{vmatrix}}{D};$$

$$dY = \frac{\begin{vmatrix} 1 & M_\varepsilon d\varepsilon + M_t dt & 0 \\ 0 & E_\varepsilon d\varepsilon + E_G dG & -E_r \\ 0 & r_\pi d\pi & 1 \end{vmatrix}}{D};$$

$$dr = \frac{\begin{vmatrix} 1 & -M_Y & M_\varepsilon d\varepsilon + M_t dt \\ 0 & A & E_\varepsilon d\varepsilon + E_G dG \\ 0 & -r_Y & r_\pi d\pi \end{vmatrix}}{D}.$$

Após algumas manipulações algébricas:

$$dM = \frac{1}{D} \{ [M_\varepsilon D + M_Y E_\varepsilon] d\varepsilon + [M_t D] dt + [M_Y E_G] dG + [M_Y E_r r_\pi] d\pi \},$$

$$dY = \frac{1}{D} \{ E_\varepsilon d\varepsilon + E_G dG + E_r r_\pi d\pi \},$$

$$dr = \frac{1}{D} \{ [E_\varepsilon r_Y] d\varepsilon + [E_G r_Y] dG + A r_\pi d\pi \}.$$

Supondo uma economia com nível de preços estável, sem inflação, e a estrutura tributária na importação, bem como os gastos governamentais fixos, chega-se a seguinte solução em relação às alterações de M e Y , com relação a ε :

$$\left. \frac{dM}{d\varepsilon} \right|_{\substack{dt=0 \\ dG=0 \\ d\pi=0}} = M_\varepsilon + \frac{M_Y E_\varepsilon}{D},$$

$$\left. \frac{dY}{d\varepsilon} \right|_{\substack{dG=0 \\ d\pi=0}} = \frac{E_\varepsilon}{D}.$$

Como o objetivo é saber a influência de alterações em ε na receita dos tributos devidos na importação, R , que por sua vez, é função da demanda de importações M , da alíquota média dos tributos, t , e da taxa real de câmbio, ε , aplica-se a derivada total na equação de R abaixo:

$$R = R(M_{(+)}, \epsilon_{(+)}, t_{(+)}),$$

$$dR = \frac{\partial R}{\partial M} dM + \frac{\partial R}{\partial \epsilon} d\epsilon + \frac{\partial R}{\partial t} dt.$$

Chamando $\partial R/\partial M = R_M > 0$ e $\partial R/\partial \epsilon = R_\epsilon > 0$, tem-se:

$$\left. \frac{dR}{d\epsilon} \right|_{dt=0} = R_M \frac{dM}{d\epsilon} + R_\epsilon,$$

$$\left. \frac{dR}{d\epsilon} \right|_{\substack{dt=0 \\ dG=0 \\ d\pi=0}} = R_M \left[M_\epsilon + \frac{M_Y E_\epsilon}{D} \right] + R_\epsilon. \quad (4)$$

O único termo negativo na equação acima é M_ϵ . Assim, o resultado final sobre a receita de importação, decorrente de uma variação no câmbio, depende basicamente de M_ϵ e M_Y , as sensibilidades da demanda de importados em relação a taxa de câmbio real e em relação à renda, respectivamente. Se M_ϵ é grande, ou seja, M é muito sensível à variação de ϵ , a expressão entre colchetes será negativa se $M_Y E_\epsilon / D < M_\epsilon$ e o resultado final dependerá de R_ϵ , a sensibilidade de R em relação a ϵ . Caso R_ϵ supere o primeiro termo do lado direito da equação acima, um aumento em ϵ provocará um aumento em R . Caso contrário, um aumento em ϵ provocará uma diminuição em R . Por outro lado, se M_ϵ é pequeno, a expressão entre colchetes será positiva desde que $M_Y E_\epsilon / D > M_\epsilon$ e o resultado final positivo, o que significa que um aumento da taxa de câmbio real eleva a receita dos tributos devidos na importação.

A análise acima pretende estabelecer o resultado final do processo de alteração da receita tributária do II em função de uma alteração na taxa de câmbio. Entretanto, considerando-se os efeitos dinâmicos da variação da taxa de câmbio, num primeiro momento, dada uma demanda por bens importados, a diminuição da taxa de câmbio (apreciação) leva à diminuição do II em função da diminuição da base de incidência do imposto, o que pode ser chamado de efeito preço, que neste caso é negativo. Porém, à medida que o tempo passa a mudança nos preços relativos altera o comércio exterior. A menor taxa de câmbio leva ao aumento da demanda de bens importados, que ficam mais baratos, em relação aos nacionais, com o conseqüente aumento do II, tudo isso mantida inalterada a alíquota do imposto, o que pode ser chamado de efeito quantidade, neste caso positivo. Portanto, o efeito final sobre a receita do II vai depender de qual efeito irá prevalecer, o que está diretamente ligado à magnitude dos parâmetros M_ϵ , M_Y , E_ϵ , D e R_ϵ , da equação 4. Assim, em última instância, renda (PIB) e taxa de câmbio real parecem ser boas variáveis explicativas do comportamento do imposto de importação.

2.2 Método dos indicadores

Segundo Santana (2004), o método dos indicadores, atualmente utilizado pela RFB na elaboração da previsão dos tributos federais, consiste na multiplicação da arrecadação do período anterior (base) por:

- um índice de preço – Efeito-Preço, que representa a variação inflacionária a que está sujeito o fato econômico gerador da arrecadação;
- um índice de quantidade – Efeito-Quantidade, que representa a variação real desse fato gerador;
- um índice de legislação – Efeito-Legislação, que representa o efeito causado na arrecadação por modificações na legislação tributária;
- outros índices – Efeito-Residual, que representam quaisquer influências na arrecadação tributária.

Esse método pode ser resumido na seguinte fórmula:

$$A_t = A_{t-1}(1 + \Delta P)(1 + \Delta Q)(1 + \Delta L)(1 + \Delta U)$$

Em que:

A_t = arrecadação prevista para determinado período do ano t;

A_{t-1} = arrecadação efetiva do mesmo período do ano t-1;

ΔP = variação percentual do indicador de preços;

ΔQ = variação percentual do indicador de quantidades;

ΔL = variação percentual decorrente de alterações na legislação (geralmente variação de alíquotas);

ΔU = variação percentual de qualquer outro indicador que tenha influência na arrecadação e não possa ser enquadrado nos indicadores básicos acima.

Assim, a qualidade da previsão depende, principalmente, da obtenção de bons indicadores de preço e quantidade específicos para cada tributo.

Ademais, na utilização do método, as séries tributárias que servem de base são corrigidas para eliminação de arrecadações atípicas, que não tendem a se repetir no período seguinte.

2.2.1 Análise Econométrica

Partindo-se da equação acima, aplicando-se o logaritmo natural (ln) a ambos os lados da equação e lembrando-se, ainda, que para pequenas variações percentuais nos índices, $\ln(1 + \Delta I^{\%}) \approx \Delta I^{\%}$, tem-se:

$$\ln A_t = \ln A_{t-1} + \ln(1 + \Delta P_t) + \ln(1 + \Delta Q_t) + \ln(1 + \Delta L_t) + \ln(1 + \Delta U_t),$$

$$\ln A_t \approx \ln A_{t-1} + \Delta P_t + \Delta Q_t + \Delta L_t + \Delta U_t \approx \ln A_{t-1} + \Delta I_t,$$

em que $\Delta I_t = \Delta P_t + \Delta Q_t + \Delta L_t + \Delta U_t$. Assim:

$$\ln A_t \approx \ln A_{t-1} + \Delta I_t. \text{ Chamando } \ln A_t = a_t \text{ e } \Delta I_t = i_t,$$

$$a_t \approx a_{t-1} + i_t. \tag{5}$$

A equação acima se assemelha, conforme se verá mais detalhadamente em 2.3, a um processo auto-regressivo de ordem um, AR(1), que, de acordo com Brockwell e Davis (2002), pode ser representado por:

$$X_t = \phi X_{t-1} + Z_t, \tag{6}$$

em que $\{Z_t\} \sim \mathbf{WN}(\mathbf{0}, \sigma^2)$, $|\phi| < 1$, e Z_t é não correlacionado com X_s , para todo $s < t$. Ou seja no caso da equação (5), $\phi = 1$, e o choque Z_t é representado pela soma das variações percentuais nos diversos índices presentes no método, podendo assumir valores tanto positivos quanto negativos.

Na equação (6), desde que $|\phi| < 1$, aplicando-se o processo iterativo, verifica-se que existe uma única solução estacionária para a equação acima, dada por:

$$X_t = \sum_{j=0}^{\infty} \phi^j Z_{t-j}.$$

Caso $|\phi| > 1$, a série representada pela equação acima não converge. Entretanto, pode-se reescrever o processo AR(1) da seguinte forma:

$$X_t = -\phi^{-1} Z_{t+1} + \phi^{-1} X_{t+1}.$$

Aplicando-se novamente o mesmo procedimento iterativo, constata-se que a única solução estacionária será:

$$X_t = -\sum_{j=1}^{\infty} \phi^{-j} Z_{t+j}.$$

Esta solução é considerada não natural, pois X_t é correlacionado com valores futuros de Z_s , diferentemente do que ocorre quando $|\phi| < 1$, em que X_t é não correlacionado com Z_s para todo $s > t$. Assim, é usual, na modelagem de séries de tempo estacionárias, restringir a atenção a processos AR(1) com $|\phi| < 1$. Neste caso, $\{X_t\}$ é considerada uma função causal ou

independente do futuro de $\{Z_t\}$, chamado de processo auto-regressivo causal. Se $|\phi| > 1$, a solução é não causal, desde que X_t é função de Z_s para $s > t$. Portanto, não há perda em se desconsiderar processos AR(1) com $|\phi| > 1$. Ademais, se $|\phi| = 1$, não há solução estacionária para a equação (6) (BROCKWELL E DAVIS, 2002, p. 54).

Voltando-se à análise do método dos indicadores, a equação (5), em que $\phi = 1$, representa um processo não estacionário, com presença de raiz unitária, chamado processo *random walk* (RW), ou passeio aleatório, cuja solução, considerando-se que a_0 é a condição inicial, é dada por:

$$a_t = a_0 + \sum_{j=1}^t i_j.$$

Ou seja, trata-se de um processo totalmente aleatório, com média constante, dada por a_0 , já que a esperança de i_t , $E(i_t) = 0$, para todo t , $i_t \sim WN(0, \sigma^2)$. Ademais, o choque i_t , tem efeito permanente sobre a_t , o que se reflete diretamente nas previsões de a_{t+s} , $s > 0$. Sua variância não é constante, sendo função direta de t , e se aproxima do infinito à medida que t tende para infinito. Assim, um processo RW vagueia sem exibir qualquer tendência de crescimento ou decrescimento.

O erro de previsão a partir do período t , s períodos à frente, $e_t(s)$, para um processo AR(1), conforme o da equação (6), é definido por:

$$e_t(s) \equiv a_{t+s} - E(a_{t+s}) = Z_{t+s} + \phi Z_{t+s-1} + \phi^2 Z_{t+s-2} + \dots + \phi^{s-1} Z_{t+1}.$$

Desde que o valor esperado da equação acima é zero, as previsões são não viesadas, porém não precisas. A variância do erro de previsão será dada por:

$$\text{Var}[e_t(s)] = \sigma^2 [1 + \phi^2 + \phi^4 + \dots + \phi^{2(s-1)}].$$

Nota-se que a variância do erro de previsão é função crescente de s , o que faz com que previsões de curto prazo sejam mais confiáveis que previsões de longo prazo. No limite, à medida que s tende para infinito, a variância do erro de previsão converge para $\sigma^2 / (1 - \phi^2)$, que é a variância incondicional de X_t . No caso do método dos indicadores, com $\phi = 1$, a variância do erro de previsão cresce de forma muito mais acentuada e, no caso limite, à medida que s tende para infinito, a variância do erro de previsão também tende para infinito.

Portanto, conforme destaca Melo (2001), o método dos indicadores utilizado pela Secretaria da Receita Federal não está reproduzindo um processo auto-regressivo causal, estacionário, e suas previsões não são confiáveis, uma vez que as condições básicas de estacionariedade não são satisfeitas. A circunstância que atenua o erro de previsão obtido pelo método dos indicadores é o fato de que as previsões são feitas anualmente, para um curto

horizonte de tempo. Tal método, aplicado a previsões para longos períodos de tempo, certamente levaria a erros inaceitáveis.

2.3 Modelagem Box-Jenkins – ARIMA

Enders (2004) destaca que os modelos auto-regressivos de médias móveis - ARMA, podem ser vistos como uma classe especial de equações em diferenças estocásticas lineares. Trata-se de um modelo covariância estacionário na medida em que tem média e covariância finitas e independentes do tempo. O modelo ARMA(p,q), pode ser representado pela seguinte equação:

$$y_t = \phi_0 + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{i=0}^q \theta_i \varepsilon_{t-i},$$

em que ϕ_0 é uma constante, ϕ_1, \dots, ϕ_p são os coeficientes da parte auto-regressiva, $\theta_0 = 1$, $\theta_1, \dots, \theta_q$ são os coeficientes da parte média móvel, $\varepsilon_t, \dots, \varepsilon_{t-q}$ são as componentes aleatórias do modelo e y_t é o processo modelado.

Para um modelo ARMA ser estacionário, as raízes da equação característica⁸ inversa do modelo acima devem se situar dentro do círculo unitário. Entretanto, se uma ou mais raízes é igual a um, a seqüência $\{y_t\}$ é um processo não estacionário, integrado, chamado auto-regressivo integrado de médias móveis (ARIMA). Raízes maiores que um caracterizam um processo explosivo. Enquanto uma série não estacionária pode ser transformada em estacionária por meio da aplicação de uma ou mais diferenças, e assim ser modelada como um processo ARMA, o mesmo não é verdadeiro para uma série explosiva.

O modelo ARMA (p,q) acima pode ser escrito de forma mais compacta da seguinte maneira:

$$\phi(L)y_t = \theta(L)\varepsilon_t,$$

em que $\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p$, e $\theta(L) = 1 + \theta_1 L + \dots + \theta_q L^q$, L^9 é o operador defasagem.

Caso todas as raízes do polinômio $\phi(L)$ se situem fora do círculo unitário¹⁰, está satisfeita a condição de estacionariedade, que é uma das condições necessárias para se aplicar a metodologia ARIMA. Neste caso, o processo $\{y_t\}$ é dito causal, ou uma função causal de

⁸ Para uma equação de segunda ordem do tipo $y_t - a_1 y_{t-1} - a_2 y_{t-2} = 0$, a equação característica inversa será $\alpha^2 - a_1 \alpha - a_2 = 0$. As raízes α_1 e α_2 desta equação são chamadas raízes características inversas (ENDERS 2004, p. 22).

⁹ O operador L é definido como um operador linear, tal que para qualquer valor de y_t : $L^1 y_t = y_{t-1}$. As propriedades do operador L podem ser encontradas em Enders (2004), p. 38-39.

¹⁰ Que é o mesmo que dizer que as raízes da equação característica se situem dentro do círculo unitário.

$\{\epsilon_t\}$. Entretanto, outra condição é necessária, qual seja, de que o modelo seja invertível, o que significa que possa ser expresso por um processo auto-regressivo convergente ou de ordem finita. Para que o processo acima seja invertível, é necessário que todas as raízes do polinômio $\theta(L)$ se situem fora do círculo unitário.

Enders (2004) enfatiza que não há nada de inadequado com modelos não invertíveis, que, ainda assim, podem ser estacionários. O problema é que não há como estimar tais modelos e, portanto, deve-se restringir a atenção a modelos causais e invertíveis.

Ainda segundo Enders (2004), a metodologia Box-Jenkins utiliza um método de três estágios para selecionar o modelo apropriado para o propósito de estimar e prever séries de tempo univariadas.

No primeiro estágio - identificação, as séries são plotadas em um gráfico e examinadas por meio das funções de autocorrelação (ACF) e autocorrelação parcial (PACF). Assim, obtemos informações sobre possíveis valores fora do padrão (*outliers*), valores ausentes e quebras estruturais nos dados. No caso de séries não estacionárias, o primeiro passo é diferenciá-las até que se tornem estacionárias¹¹. A comparação das funções ACF e PACF¹² pode indicar vários modelos ARMA plausíveis.

No segundo estágio - estimação, cada um dos possíveis modelos identificados no primeiro estágio tem seus parâmetros estimados, levando-se em conta o princípio da parcimônia e as condições de estacionariedade e invertibilidade. Nesse estágio surge uma questão importante: o aumento das defasagens do modelo (p e/ou q) necessariamente reduzirá a soma dos quadrados dos resíduos. Entretanto, aumenta-se a quantidade de parâmetros a serem estimados e assim perdem-se graus de liberdade. Ademais, a inclusão de parâmetros irrelevantes reduz o poder de previsão do modelo. Entre os diversos critérios de seleção de modelos que consideram o *trade off* entre soma dos quadrados dos resíduos e parcimônia, os mais utilizados são o Critério de Informação de Akaike (AIC) e o Critério Bayesiano de Schwartz (SBC ou BIC). O critério AIC trabalha melhor em pequenas amostras enquanto o critério BIC tem desempenho superior em grandes amostras.

Por fim tem-se o terceiro estágio, que envolve checagem e diagnóstico. Um bom modelo se ajustará bem aos dados e os resíduos do modelo estimado devem ser não correlacionados serialmente. Pode-se estimar o modelo considerando-se somente uma porção

¹¹ Desde que a série seja tendência estacionária. Este aspecto será mais aprofundado em 2.3.2, que trata de raízes unitárias.

¹² As propriedades das funções ACF e PACF para um modelo ARMA(p,q) podem ser encontradas em Enders (2004), p. 66.

dos dados, deixando, por exemplo, os últimos dados para comparação com os valores previstos pelo modelo estimado. Assim, a soma dos quadrados dos resíduos é uma boa maneira de se comparar os diversos modelos estimados. Aqueles com baixo poder de previsão devem ser eliminados.

Resumindo, um modelo bem estimado deve:

- ser parcimonioso;
- ter coeficientes que implicam estacionariedade e invertibilidade;
- ajustar-se bem aos dados;
- ter resíduos que se aproximam de um ruído branco;
- ter coeficientes que não se alteram sobre o período amostral;
- ter bom poder preditivo.

2.3.1 Sazonalidade

Sazonalidade, definida como a propriedade de determinadas séries de apresentar certos comportamentos cíclicos ou periódicos, é uma característica de diversos processos econômicos e séries de arrecadação de tributos geralmente apresentam forte padrão sazonal, seja devido ao próprio processo econômico, que em última instância é o responsável pela arrecadação tributária, seja devido à forma de arrecadação do tributo analisado, como ocorre por exemplo, no caso do Imposto Sobre a Renda da Pessoa Jurídica (IRPJ) – Lucro Real Trimestral, em que se apura o tributo devido em função do lucro real em determinado trimestre, com o pagamento podendo ser efetuado em até três cotas mensais, com incidência de juros a partir da segunda cota, o que faz com que os recolhimentos se concentrem mais na primeira cota.

Enders (2004) ressalta que previsões que ignoram os padrões sazonais terão alta variância. Muitos ignoram o padrão sazonal sob o argumento de trabalhar com dados dessazonalizados ou ajustados sazonalmente. Porém esta parece não ser a melhor alternativa, pois mesmo dados ajustados podem reter ainda algum padrão sazonal. Assim, a estimação conjunta dos coeficientes ARMA e sazonais frequentemente leva a melhores resultados (ENDERS 2004, p. 93-94).

Na modelagem Box-Jenkins para dados sazonais de período s ¹³, os coeficientes sazonais da ACF e PACF aparecem nas defasagens $s, 2s, 3s, \dots$, ao invés de $1, 2, 3, \dots$. Os

¹³ $s = 12$ para dados mensais e $s = 4$ para dados trimestrais.

coeficientes do modelo sazonal podem ser tratados de forma aditiva ou multiplicativa. Por exemplo, um modelo aditivo, identificado como ARIMA ((1,4),0,1), tem a seguinte representação:

$$(1 - a_1L - a_4L^4) y_t = (1 + b_1L) \varepsilon_t ,$$

enquanto um modelo multiplicativo, identificado como SARIMA (1,0,1) \times (1,0,0)₄, é representado por:

$$(1 - a_1L) (1 - a_4L^4) y_t = (1 + b_1L) \varepsilon_t .$$

Não há uma forma de tratamento preferida e a checagem e diagnóstico do modelo estimado é a melhor maneira de se obter o modelo mais apropriado. Caso o modelo apresente uma raiz unitária sazonal, é necessária uma diferenciação sazonal, que, no caso de dados mensais terá a forma $(1 - L^{12}) = \Delta_{12}$. Neste caso, a ordem de diferenciação sazonal, representada por **D**, é igual a um. Isto posto, o modelo multiplicativo será representado por SARIMA (p,d,q) \times (P,D,Q)_s, em que:

- d = número de diferenças não sazonais;
- p e q = coeficientes não sazonais ARMA;
- P = número de coeficientes auto-regressivos multiplicativos;
- Q = número de coeficientes média móvel multiplicativos;
- D = número de diferenças sazonais;
- s = período do padrão sazonal.

Generalizando, Brockwell e Davis (2002) assim definem o modelo sazonal multiplicativo:

“Se **d** e **D** são inteiros não negativos, então $\{X_t\}$ é um processo SARIMA (p,d,q) \times (P,D,Q)_s com período **s**, se a série diferenciada $Y_t = (1 - L)^d (1 - L^s)^D X_t$ é um processo ARMA definido por $\phi(L)\Phi(L^s)Y_t = \theta(L)\Theta(L^s)Z_t$, $\{Z_t\} \sim WN(0, \sigma^2)$, $\phi(z) = 1 - \phi_1z - \dots - \phi_pz^p$, $\Phi(z) = 1 - \Phi_1z - \dots - \Phi_Pz^P$, $\theta(z) = 1 + \theta_1z + \dots + \theta_qz^q$ e $\Theta(z) = 1 + \Theta_1z + \dots + \Theta_Qz^Q$.” (BROCKWELL e DAVIS, 2002, p. 203).

Na prática **D** quase nunca é maior que um e especialmente para séries mensais, com **s** = 12, **P** e **Q** raramente são maiores que um. Isto é particularmente válido para séries com quantidade de dados não suficientes para garantir uma estimação consistente de modelos com **P** ou **Q** maiores que um (BOX, JENKINS e REINSEL, 2008, p. 378).

2.3.2 Raiz Unitária

Para aplicação da modelagem Box-Jenkins¹⁴ é necessário que a série temporal seja estacionária. Para tal, conforme já explicitado, pode ser necessário aplicar o operador diferença (Δ) à série. Para séries que apresentam componente de tendência, é importante diferenciar entre aquelas que apresentam tendência estacionária, daquelas cuja tendência é estocástica. No primeiro caso, os desvios em relação à tendência determinista são temporários, seguindo um processo estacionário ARMA, e o modelo é chamado tendência estacionária (TS). No segundo caso, os choques têm impacto permanente na média condicional da série, não seguindo um processo estacionário em relação à tendência determinista ou à média do processo e os modelos são chamados tendência estocástica (diferença estacionária - DS). Os modelos básicos que apresentam tendência, seja ela determinista ou estocástica, podem ser sintetizados na seguinte tabela.

Tabela 2: Modelos básicos de séries de tempo com tendência

Modelo	Equação básica	Solução geral	Componentes
Tendência estacionária	$y_t = y_{t-1} + a_0 + A(L)\Delta\varepsilon_t$	$y_t = y_0 + a_0t + A(L)\varepsilon_t$	tendência determinista linear + comp. estac.
<i>Random Walk (RW)</i>	$y_t = y_{t-1} + \varepsilon_t$	$y_t = y_0 + \sum_{i=1}^t \varepsilon_i$	tendência estocástica
<i>RW</i> mais tendência	$y_t = y_{t-1} + a_0 + \varepsilon_t$	$y_t = y_0 + \sum_{i=1}^t \varepsilon_i + a_0t$	tend. estocástica + tend. determinista linear
<i>RW</i> mais ruído	$\Delta y_t = \varepsilon_t + \Delta\eta_t$	$y_t = y_0 + \sum_{i=1}^t \varepsilon_i + \eta_t$	tend. estocástica + comp. ruído branco
Tendência mais ruído	$\Delta y_t = a_0 + \varepsilon_t + \Delta\eta_t$	$y_t = y_0 + a_0t + \sum_{i=1}^t \varepsilon_i + \eta_t$	tend. estocástica + tend. determinista linear + comp. ruído branco
Tendência mais componente irregular	$\Delta y_t = a_0 + \varepsilon_t + A(L)\Delta\eta_t$	$y_t = y_0 + a_0t + \sum_{i=1}^t \varepsilon_i + A(L)\eta_t$	tend. estocástica + tend. determinista linear + comp. estac.

Fonte: elaboração do autor – baseados em (ENDERS 2004).

Nota: Nos modelos acima ε_t e $\eta_t \sim WN$ e $E(\varepsilon_t\eta_{t-s}) = 0$ para todo t e s . $A(L)$ é o polinômio no operador defasagem, que aplicado em determinada série resulta em uma série estacionária.

A correta identificação da tendência presente na série, se determinista ou estocástica, é fundamental para aplicação do procedimento correto para se remover a tendência e obter uma série estacionária que poderá ser tratada por meio da modelagem Box-Jenkins. Há duas maneiras de se remover a tendência: para uma série TS a remoção se faz por meio da inclusão

¹⁴ A estacionariedade da série temporal também é requerida nas demais modelagens aplicadas neste trabalho, com exceção dos modelos estruturais.

de um componente dependente do tempo na equação de regressão; para uma série DS, ou seja, com raiz unitária, aplicar a diferença à série conduzirá à estacionariedade.

Aplicar o método errado de remoção de tendência gera um sério problema. Aplicar a diferença em um processo TS introduz raiz unitária não invertível no componente MA do modelo. Por outro lado, retirar uma tendência determinista de um processo DS não resulta em uma série estacionária (ENDERS, 2004, p. 167).

Considere a equação de regressão:

$$y_t = a_0 + a_1 z_t + e_t \quad (7)$$

Enders (2004) enfatiza que as hipóteses básicas requeridas pelo modelo acima são a estacionariedade de $\{y_t\}$ e $\{z_t\}$ e que os erros tenham média zero e variância finita. Na presença de seqüências não estacionárias, tem-se a chamada regressão espúria, que apresenta elevado R^2 e estatísticas t significantes, porém sem significado econômico.

Assim, deve-se ter cuidado ao se trabalhar com variáveis não estacionárias. Enders (2004) elenca quatro casos a se considerar, tendo por base a equação (7) acima:

- Caso 1: se $\{y_t\}$ e $\{z_t\}$ são estacionárias, o modelo é adequado;
- Caso 2: se $\{y_t\}$ e $\{z_t\}$ são integradas de ordem diferente, o modelo não faz sentido;
- Caso 3: se $\{y_t\}$ e $\{z_t\}$ são integradas de mesma ordem e os resíduos contém uma tendência estocástica, o modelo não faz sentido e teremos o caso da regressão espúria. Neste caso recomenda-se aplicar a diferença às séries, tornando-as estacionárias (caso as séries sejam DS);
- Caso 4: se $\{y_t\}$ e $\{z_t\}$ são integradas de mesma ordem e os resíduos são estacionários, $\{y_t\}$ e $\{z_t\}$ são cointegradas¹⁵ e possuem uma relação estável no longo prazo.

Isto posto, percebe-se a importância dos testes de raiz unitária como pré-requisito para a correta estimação do modelo ARIMA. Brockwell e Davis (2002) destacam que o problema de raiz unitária surge quando o polinômio AR ou MA do modelo ARMA tem raízes no círculo unitário ou próximas deste. Uma raiz próxima do círculo unitário no polinômio autorregressivo sugere que os dados devem ser diferenciados, enquanto que no polinômio média móvel indica que os dados foram sobrediferenciados.

Maddala e Kim (1998) destacam que além da importância da presença de tendência determinista em séries macroeconômicas, a especificação da tendência determinista

¹⁵ A definição de cointegração se encontra em 2.5.

desempenha papel essencial nos testes de raiz unitária e está intimamente relacionada com o poder e tamanho dos testes. Se uma variável de tendência, que está presente no verdadeiro processo gerador de dados (DGP), é omitida da regressão, o poder do teste tende a zero à medida que o tamanho da amostra aumenta.

Ao se estimar uma regressão de um processo TS usando a primeira diferença, teremos sobrediferenciação, enquanto que a estimação de um processo DS em nível, conduzirá a subdiferenciação (MADDALA e KIM, 1998, p. 88-89).

Dentre uma ampla gama de testes de raiz unitária disponíveis, destacam-se os testes ADF (Dickey-Fuller aumentado) e PP (Phillips-Perron). Porém, estes testes apresentam problemas de distorção de tamanho (PP) e baixo poder (ADF). Dois testes que procuram contornar estes problemas são Ng-Perron (NgP) e Elliott-Rothenberg-Stock DF-GLS (DF-GLS (ERS)). O teste NgP mantém um bom poder, ao mesmo tempo em que corrige problemas de distorção de tamanho na presença de resíduos MA negativos, que são comuns na maioria das séries macroeconômicas, além da vantagem de não necessitar da identificação de observações atípicas (*outliers*) antes da realização do teste (MADDALA e KIM, 1998, p. 109). Os testes NgP e DF-GLS (ERS) são, portanto, os testes de raiz unitária utilizados no presente trabalho.

Outra questão importante nos testes de raiz unitária diz respeito à relação entre a frequência das observações e o poder dos testes. Maddala e Kim (1998) apresentam resultados de estudos que concluem que o poder dos testes depende mais da extensão do intervalo dos dados, que do número de observações. Para dados de fluxo (caso do presente estudo), agregar os dados resulta em menor poder dos testes. Apresentam, ainda, resultados de estudo de Ng (1995) que analisou o poder dos testes de RU variando a frequência das observações, bem como a extensão do intervalo dos dados (S), cujas conclusões são: (i) aumentar a frequência das observações enquanto mantém-se S constante, aumenta o poder do teste, mas a taxa decrescentes; (ii) o poder decresce se a frequência das observações é aumentada, mas S é diminuída; (iii) o poder aumenta se S aumenta, mesmo se o número de observações é mantido constante (pela redução da frequência das observações).

O efeito do ajuste sazonal dos dados também é examinado em Maddala e Kim (1998). O questionamento que se faz é se os testes RU devem ser aplicados aos dados ajustados ou não pela sazonalidade. A conclusão é de que os testes ADF e PP são viesados na direção da não rejeição da raiz unitária quando se utilizam dados ajustados sazonalmente. Assim, espera-se que os testes RU tenham maior poder se aplicados aos dados sem ajuste, o que é utilizado neste trabalho.

2.4 Modelo de Função de Transferência

Segundo Enders (2004), o modelo de Função de Transferência (TFM) pode ser expresso da seguinte forma:

$$y_t = a_0 + A(L)y_{t-1} + C(L)z_t + B(L)\varepsilon_t, \quad (8)$$

em que $A(L)$, $B(L)$ e $C(L)$ são polinômios no operador defasagem L .

As seqüências $\{y_t\}$ e $\{z_t\}$ representam as séries das variáveis endógena e exógena, respectivamente. O objetivo é estimar o parâmetro a_0 e os parâmetros dos polinômios $A(L)$, $B(L)$ e $C(L)$. O polinômio $C(L)$ é chamado função de transferência, na medida em que revela como o movimento na variável exógena z_t afeta $\{y_t\}$. Os coeficientes de $C(L)$, são chamados de pesos da função de transferência.

Segundo Enders (2004), há duas técnicas para se estimar o modelo TFM. A primeira¹⁶ estima a equação

$$y_t = a_0 + \sum_{i=1}^p a_i y_{t-i} + \sum_{i=0}^n c_i z_{t-i} + \varepsilon_t \quad (9)$$

ao invés de se estimar a equação (8). Aqui, $\{y_t\}$ é representado por um processo $AR(p)$ que também é afetado por valores presentes e passados de $\{z_t\}$. Esta especificação, chamada de ADL (*autoregressive distributed lags*), se inicia com valores grandes para p e n , julgados exequíveis. Então, a partir de testes F e t , eliminam-se os coeficientes desnecessários. Podem ser usados, também, os critérios de ajustamento de Akaike (AIC) e Schwartz (BIC) para se encontrar os valores de p e n que resultam no melhor modelo. O inconveniente do método é a possibilidade de se obter modelos sobreparametrizados, já que pode haver um modelo ARMA para $\{y_t\}$ mais parcimonioso que a representação AR da equação (9). O segundo método estima a equação (8) diretamente. Embora a identificação de $B(L)$ possa ser difícil, pode levar a modelos mais parcimoniosos.

Para validade do modelo TFM, especialmente quando o objetivo é a previsão de valores futuros de y_t , requer-se que z_t seja uma variável fortemente exógena. A próxima subseção trata da questão da exogeneidade da variável explicativa.

¹⁶ Esta foi a técnica utilizada no presente trabalho, por meio do módulo *PcGive* do programa *GiveWin 2.02* (ver seção 5).

2.4.1 Exogeneidade

Enders (2004) destaca que o ponto fundamental na análise de função de transferência é que admite-se que $\{z_t\}$ é um processo exógeno que se desenvolve independentemente de $\{y_t\}$, ou seja, $\{y_t\}$ não tem efeito sobre $\{z_t\}$. Considerando $C(L) = c_0 + c_1L + c_2L^2 + \dots$, se $c_0 = 0$, o valor contemporâneo de z_t não afeta y_t . Assim, para a validade do modelo TFM¹⁷, é necessário que haja causalidade no sentido de $\{z_t\}$ para $\{y_t\}$, sem a possibilidade de haver realimentação entre a variável dependente e as variáveis explicativas.

Para o entendimento desta questão, três conceitos precisam ficar bem entendidos, quais sejam, exogeneidade fraca, causalidade de Granger e exogeneidade forte.

Segundo Harvey (1989), no caso de séries de tempo, com a variável dependente $\mathbf{y} = [y_1, \dots, y_T]'$ e a variável exógena designada pelo vetor $\mathbf{X} = [x_1, \dots, x_T]'$, a função densidade de probabilidade conjunta (pdf conjunta) é dada por:

$$L(\mathbf{y}, \mathbf{X}; \lambda) = \prod_{t=1}^T p(y_t, x_t | Y_{t-1}, X_{t-1}; \lambda),$$

em que λ é o conjunto completo de parâmetros sobre o qual a pdf conjunta de \mathbf{y} e \mathbf{X} dependem. Considere que o modelo possa ser reparametrizado em função de um novo conjunto de parâmetros θ^* , de tal forma que $\theta^* = [\theta', \theta_x]$, em que θ' e θ_x tem variação livre (*variation-free*), ou seja, para qualquer valor de θ_x , θ pode assumir qualquer valor no seu espaço de parâmetros e vice-versa. Suponha, ainda, que os parâmetros de interesse são funções apenas de θ . A variável x_t será fracamente exógena com relação a θ se a equação acima puder ser fatorada em:

$$L(\mathbf{y}, \mathbf{X}; \theta^*) = \left[\prod_{t=1}^T p(y_t | Y_{t-1}, X_t; \theta) \right] \left[\prod_{t=1}^T p(x_t | X_{t-1}, Y_{t-1}; \theta_x) \right]$$

Assim, o segundo termo acima, a distribuição marginal de x_t , pode ser ignorado já que não contém informação em θ (parâmetros de interesse). Portanto, inferência sobre os parâmetros de interesse θ pode ser efetuada por meio da distribuição condicional de y_t , o primeiro termo da equação acima. Isto equivale a considerar a variável explicativa como se fosse fixa em repetidas amostras, mesmo que ela seja gerada por um processo estocástico.

Granger e Newbold (1977) tratam da questão da causalidade a partir de duas regras fundamentais: (i) O futuro não pode causar o passado. Causalidade estrita somente pode ocorrer com o passado causando o presente ou o futuro; (ii) Somente faz sentido discutir

¹⁷ Os conceitos aqui apresentados também se aplicam ao modelo estrutural com inclusão de variáveis exógenas.

causalidade para um grupo de processos estocásticos. Não é possível detectar causalidade entre dois processos determinísticos.

A partir daí, definem causalidade da seguinte maneira: Seja $\mathbf{P}(\mathbf{A}|\mathbf{B})$ a função de distribuição condicional de \mathbf{A} dado \mathbf{B} e Ω_t toda a informação disponível no universo no tempo t . Pergunta-se: a série \mathbf{X}_t causa a série \mathbf{Y}_t ? Então, se $\mathbf{P}(\mathbf{Y}_{t+1}|\Omega_t) = \mathbf{P}(\mathbf{Y}_{t+1}|\Omega_t - \mathbf{X}_t)$, em que $\Omega_t - \mathbf{X}_t$ é toda a informação do universo exceto a informação proveniente de \mathbf{X}_t , \mathbf{X}_t não causa \mathbf{Y}_t . Caso contrário, \mathbf{X}_t causa \mathbf{Y}_t .

Ainda segundo Granger e Newbold (1977), em vez de se lidar com toda a informação do universo, considera-se o conjunto de informações \mathbf{I}_t e uma definição mais útil em função deste conjunto de dados disponível é a seguinte: Seja $\mathbf{f}_t(\mathbf{I})$ a previsão ótima de \mathbf{Y}_{t+1} usando a informação contida em \mathbf{I}_t , com o termo de erro $\mathbf{e}_t(\mathbf{I}) = \mathbf{Y}_{t+1} - \mathbf{f}_t(\mathbf{I})$ e a variância $\sigma_t^2(\mathbf{I}) = \text{var}(\mathbf{e}_t(\mathbf{I}))$. Se $\mathbf{I}_t - \mathbf{X}_t$ representa o conjunto de informações à exceção da informação proveniente de \mathbf{X}_t , então pode ser dito que \mathbf{X}_t causa \mathbf{Y}_t , com respeito a \mathbf{I}_t , se $\sigma_t^2(\mathbf{I} - \mathbf{X}) > \sigma_t^2(\mathbf{I})$.

Assim, a causalidade de Granger estabelece que uma variável \mathbf{x} Granger causa uma variável \mathbf{y} , se \mathbf{y}_t puder ser mais bem prevista com o uso da informação do passado de \mathbf{x} , do que sem a utilização desta informação, mantidas as demais informações idênticas. Dizer que \mathbf{X}_{t-1} não Granger causa \mathbf{y}_t equivale a dizer que a distribuição de \mathbf{y}_t é independente dos valores passados de \mathbf{x}_t , ou seja,

$$p(y_t | Y_{t-1}, X_{t-1}; \theta_y) = p(y_t | Y_{t-1}; \theta_y), \quad t = 1, \dots, T.$$

Ademais, conforme estabelece Covey e Bessler (1992), a causalidade de \mathbf{X} para \mathbf{Y} requer alguma teoria ou crença de que \mathbf{X} causa \mathbf{Y} e, ainda, evidências de que a adição de \mathbf{X} no conjunto de informações \mathbf{I}_t resulta em melhora da performance preditiva dos valores futuros de \mathbf{Y} . Melhora do poder preditivo sem uma teoria que sustente o modelo é insuficiente, pois é possível obter causalidade espúria entre duas variáveis porque uma terceira variável, que causa a ambas, foi deixada de fora do conjunto \mathbf{I}_t .

Quando o interesse está na previsão da variável \mathbf{y}_t , é necessário estabelecer o conceito de exogeneidade forte das variáveis explicativas. Um vetor de variáveis explicativas \mathbf{x}_t é fortemente exógena com relação a \mathbf{y}_t , para o conjunto de parâmetros θ , se \mathbf{x}_t é fracamente exógena com relação a θ e \mathbf{y}_t não causar \mathbf{x}_t no sentido de Granger. Ou seja, \mathbf{x}_t não pode ser influenciado pelos valores contemporâneos ou passados de \mathbf{y}_t . É importante notar que exogeneidade forte implica em exogeneidade fraca, mas não o inverso.

É importante notar que a afirmação “**X** Granger causa **Y**” não implica que **Y** é o efeito ou resultado de **X**. A causalidade de Granger mede a precedência de uma variável em relação à outra e não pode, por si só, indicar causalidade, no uso mais comum do termo. Note-se, também, que a causalidade nos dois sentidos é freqüente, ou seja, podemos ter “**X** Granger causa **Y**” e “**Y** Granger causa **X**” (EViews 5 USER’S GUIDE, p. 376).

2.5 Modelos dinâmicos de múltiplas equações

Modelos TFM pressupõem que a trajetória da variável dependente seja influenciada pela trajetória de uma variável independente ou exógena. Porém, muitos sistemas econômicos apresentam realimentação, ou seja, a trajetória da variável dita independente é afetada pela trajetória da variável dependente e a hipótese básica para a funcionalidade do modelo TFM é violada. Neste caso, uma alternativa é trabalhar com um sistema de equações que trata todas as variáveis simetricamente sem fazer distinção entre variável dependente e independente, o que é chamado de auto-regressão vetorial (VAR).

Para ilustrar, de acordo com Enders (2004), base desta sub-seção, considere um VAR de primeira ordem formado por duas variáveis, em que a trajetória de $\{y_t\}$ seja afetada pelo presente e passado de $\{z_t\}$ e que a trajetória de $\{z_t\}$ seja afetada pelo presente e passado de $\{y_t\}$, ou seja:

$$\begin{aligned} y_t &= b_{10} - b_{12}z_t + \gamma_{11}y_{t-1} + \gamma_{12}z_{t-1} + \varepsilon_{yt}, \\ z_t &= b_{20} - b_{21}y_t + \gamma_{21}y_{t-1} + \gamma_{22}z_{t-1} + \varepsilon_{zt}. \end{aligned} \quad (10)$$

No sistema acima assume-se que y_t e z_t são estacionários, ε_{yt} e ε_{zt} são resíduos ruído branco não correlacionados, com desvios padrão σ_y e σ_z , respectivamente. O sistema acima, chamado de VAR estrutural ou sistema primitivo possui realimentação, desde que y_t e z_t afetam um ao outro. Após um pouco de álgebra matricial, o sistema (10) pode ser escrito na forma VAR padrão, dada por:

$$x_t = A_0 + A_1x_{t-1} + e_t, \quad (11)$$

em que:

$$\begin{aligned} A_0 &= B^{-1}\Gamma_0, \quad A_1 = B^{-1}\Gamma_1, \quad e_t = B^{-1}\varepsilon_t, \\ x_t &= \begin{bmatrix} y_t \\ z_t \end{bmatrix}, \quad B = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}, \quad \Gamma_0 = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix}, \quad \Gamma_1 = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}, \quad \varepsilon_t = \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}. \end{aligned}$$

Definindo a_{i0} como o elemento i do vetor A_0 , a_{ij} o elemento da linha i , coluna j da matriz A_1 e e_{it} o elemento i do vetor e_t , o VAR na forma padrão pode ser escrito como:

$$\begin{aligned} y_t &= a_{10} + a_{11}y_{t-1} + a_{12}z_{t-1} + e_{1t}, \\ z_t &= a_{20} + a_{21}y_{t-1} + a_{22}z_{t-1} + e_{2t}. \end{aligned} \quad (12)$$

No sistema acima, e_{1t} e e_{2t} , que se compõem dos dois choques ε_{yt} e ε_{zt} , são individualmente não correlacionados serialmente, porém correlacionados entre si. Desde que $e_t = \mathbf{B}^{-1}\varepsilon_t$, são expressos por:

$$\begin{aligned} e_{1t} &= (\varepsilon_{yt} - b_{12}\varepsilon_{zt}) / (1 - b_{12}b_{21}), \\ e_{2t} &= (\varepsilon_{zt} - b_{21}\varepsilon_{yt}) / (1 - b_{12}b_{21}). \end{aligned} \quad (13)$$

A condição para estabilidade do sistema (12) requer que as raízes de $(\mathbf{1} - \mathbf{a}_{11}\mathbf{L})(\mathbf{1} - \mathbf{a}_{22}\mathbf{L}) - (\mathbf{a}_{12}\mathbf{a}_{21}\mathbf{L}^2)$ se situem fora do círculo unitário.

Quando o VAR contém variáveis não estacionárias, é possível haver uma combinação linear das variáveis que seja estacionária. Neste caso as variáveis são cointegradas e estimar um VAR com as variáveis em diferença conduz a resultados incorretos.

Engle e Granger (1987 apud Enders, 2004, p. 322) assim definem cointegração:

Os componentes do vetor $\mathbf{x}_t = (\mathbf{x}_{1t}, \mathbf{x}_{2t}, \dots, \mathbf{x}_{nt})'$ são cointegrados de ordem \mathbf{d}, \mathbf{b} , $\mathbf{x}_t \sim \mathbf{CI}(\mathbf{d}, \mathbf{b})$ se:

- 1 todos os componentes de \mathbf{x}_t são integrados de ordem \mathbf{d} ;
- 2 existe um vetor $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_n)$ tal que a combinação linear $\boldsymbol{\beta}\mathbf{x}_t = \boldsymbol{\beta}_1\mathbf{x}_{1t} + \boldsymbol{\beta}_2\mathbf{x}_{2t} + \dots + \boldsymbol{\beta}_n\mathbf{x}_{nt}$ é integrada de ordem $(\mathbf{d} - \mathbf{b})$, em que $\mathbf{b} > \mathbf{0}$.

O vetor $\boldsymbol{\beta}$ é chamado de vetor de cointegração.

A partir da definição acima, conclui-se que cointegração refere-se à combinação linear de variáveis não estacionárias e integradas de mesma ordem, além do que, o número máximo de vetores de cointegração é de $\mathbf{n} - \mathbf{1}$, que corresponde ao posto de cointegração de \mathbf{x}_t . Neste trabalho, todas as variáveis são integradas de ordem 1, $I(1)$, e, assim, as relações de cointegração, quando existentes, são $\mathbf{CI}(1,1)$, ou seja, existe uma ou mais combinações lineares que são $I(0)$.

A principal característica de variáveis cointegradas é que suas trajetórias no tempo são influenciadas pela extensão de qualquer desvio do equilíbrio de longo prazo, ou seja, a dinâmica da relação de curto prazo, chamada de representação de correção de erro, é influenciada pelo desvio da relação de longo prazo. Um vetor $\mathbf{x}_t = (\mathbf{x}_{1t}, \mathbf{x}_{2t}, \dots, \mathbf{x}_{nt})'$ tem uma representação de correção de erro (VEC) se puder ser expresso na forma:

$$\Delta \mathbf{x}_t = \boldsymbol{\pi}_0 + \boldsymbol{\pi}_1 \mathbf{x}_{t-1} + \boldsymbol{\pi}_1 \Delta \mathbf{x}_{t-1} + \dots + \boldsymbol{\pi}_p \Delta \mathbf{x}_{t-p} + \boldsymbol{\varepsilon}_t, \quad (14)$$

em que: $\boldsymbol{\pi}_0$ = vetor $(\mathbf{n} \times \mathbf{1})$ de interceptos com elementos $\boldsymbol{\pi}_{i0}$;

$\boldsymbol{\pi}_i$ = matriz $(\mathbf{n} \times \mathbf{n})$ de coeficientes com elementos $\boldsymbol{\pi}_{jk}(\mathbf{i})$;

π = matriz com elementos π_{jk} tal que um ou mais $\pi_{jk} \neq 0$;

ε_t = vetor ($n \times 1$) com elementos ε_{it} , tal que ε_{it} pode ser correlacionado com ε_{jt} .

Supondo que todas variáveis em \mathbf{x}_t sejam I(1), se há uma representação de correção de erro como em (14), necessariamente há uma combinação linear de variáveis I(1) que é estacionária. O termo $\pi\mathbf{x}_{t-1}$ é estacionário já que todos os outros termos da equação (14) também o são. Cada linha de π é um vetor de cointegração de \mathbf{x}_t . Se todos os elementos de π são iguais a zero, a equação (14) se transforma num VAR tradicional em primeira diferença e não há representação de correção de erro, desde que $\Delta\mathbf{x}_t$ não responde ao desvio do período anterior do equilíbrio de longo prazo. Entretanto, se um ou mais elementos π_{jk} diferem de zero, $\Delta\mathbf{x}_t$ responde ao desvio do período anterior do equilíbrio de longo prazo e, portanto, estimar \mathbf{x}_t como um VAR em primeira diferença não é apropriado se \mathbf{x}_t tem uma representação de correção de erro.

Para examinar a relação entre cointegração e correção de erro, considere o VAR simples da equação (12), que pode ser reescrito, ignorando-se os termos dos interceptos, da seguinte maneira:

$$\begin{bmatrix} \Delta y_t \\ \Delta z_t \end{bmatrix} = \begin{bmatrix} a_{11} - 1 & a_{12} \\ a_{21} & a_{22} - 1 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}. \quad (15)$$

Para que $\{y_t\}$ e $\{z_t\}$ sejam CI(1,1), é necessário que uma raiz característica¹⁸ inversa seja igual a um e a outra seja menor que um em valor absoluto. A condição de que a maior das duas raízes seja igual a um implica que:

$$a_{11} = [(1 - a_{22}) - a_{12}a_{21}] / (1 - a_{22}). \quad (16)$$

Desde que as variáveis sejam cointegradas, a_{12} e/ou a_{21} são diferentes de zero e a condição de que a segunda raiz seja menor que um em valor absoluto requer que:

$$\begin{aligned} a_{22} &> -1 \text{ e} \\ a_{12}a_{21} + a_{22}^2 &< 1. \end{aligned} \quad (17)$$

Assim, a equação (15), com a restrição da equação (16), implica que o modelo de correção de erro, com a_{12} e a_{21} diferentes de zero e o vetor de cointegração normalizado com relação à variável y_t , será:

$$\begin{aligned} \Delta y_t &= \alpha_y (y_{t-1} - \beta z_{t-1}) + e_{1t}, \\ \Delta z_t &= \alpha_z (y_{t-1} - \beta z_{t-1}) + e_{2t}. \end{aligned} \quad (18)$$

Em que:

¹⁸ A equação característica inversa do sistema (12) é dada por $\lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0$, $\lambda = 1/L$, L é o operador defasagem.

$$\begin{aligned}\alpha_y &= -a_{12}a_{21}/(1-a_{22}), \\ \alpha_z &= a_{21}, \\ \beta &= (1-a_{22})/a_{21}.\end{aligned}$$

Portanto, o modelo de correção de erro é o VAR em primeira diferença com a inclusão do termo de correção de erro, $\alpha_y(y_{t-1} - \beta z_{t-1})$ e $\alpha_z(y_{t-1} - \beta z_{t-1})$. α_y e α_z são chamados de velocidades de ajustamentos. Quanto maior α_y , maior a resposta de Δy ao desvio do período anterior do equilíbrio de longo prazo. No caso da equação (18), o vetor de cointegração normalizado é $[1, -\beta]$ e as restrições impostas para assegurar que as variáveis sejam CI(1,1), garantem a existência do modelo de correção de erro. Assim, um sistema cointegrado pode ser visto como uma forma restrita do VAR geral.

Caso as variáveis sejam cointegradas, as linhas de π , na equação (14), não são linearmente independentes e o posto de π pode ser usado para se determinar se as variáveis do sistema são cointegradas. Se o posto de π é igual a zero, não há cointegração e o VAR em primeira diferença das n variáveis é o modelo adequado; se o posto de π é completo, ou seja, igual a n , todas as variáveis em x_t são I(0) e o VAR em nível das n variáveis é o modelo adequado; finalmente, se o posto de π é r , com $0 < r < n$, há r relações de cointegração que são I(0)¹⁹ e o VEC é a especificação adequada para a estimação.

Em sistemas cointegrados, geralmente ambas as variáveis respondem aos desvios do equilíbrio de longo prazo. Porém, é possível que um parâmetro α (mas não ambos) seja zero. Caso $\alpha_z = 0$, $\{z_t\}$ não responde ao desvio do equilíbrio de longo prazo, todo o ajuste é feito por $\{y_t\}$, e $\{z_t\}$ é considerada fracamente exógena. Ademais, em sistemas cointegrados é necessário reinterpretar a causalidade de Granger: $\{y_t\}$ não Granger causa $\{z_t\}$ se os valores passados Δy_{t-i} não entram na equação de Δz_t e se $\alpha_z = 0$. Neste caso, $\{z_t\}$ será fortemente exógena.

De acordo com Hendry (1995), processos integrados podem ser expressos em termos de variáveis em diferenças e relações de cointegração, sem perda de informação, mas a redução na dimensão do vetor de parâmetros pode trazer ganho. Assim, para se obter um modelo mais parcimonioso, após o VEC, o modelo é reduzido para I(0), formando um modelo de equações simultâneas, em diferenças e com as relações de cointegração (identidades), a partir do qual restrições podem ser impostas em cada equação individualmente, bem como sobre as identidades, e os testes de adequação do modelo são efetuados. Após a redução, a análise estatística convencional pode ser implementada. Caso os parâmetros estimados para as

¹⁹ Neste caso não está se considerando a possibilidade das variáveis serem I(2), pois neste trabalho todas as variáveis são I(1).

identidades de determinada equação (variável dependente y_2 , por exemplo) sejam iguais a zero, a respectiva variável é considerada fracamente exógena. Se, adicionalmente, valores passados das outras variáveis (y_1, y_3, \dots) podem ser retirados da equação de y_2 , sem perda de informação, ou seja, y_1, y_3 etc. não causam y_2 no sentido de Granger, então y_2 é fortemente exógena. Neste caso, alternativamente, pode-se estimar um modelo TFM em que y_2 é inserida como variável exógena, tornando os modelos mais simples de serem estimados, devido à diminuição do número de parâmetros, e previsões dinâmicas podem ser efetuadas.

Ainda segundo Hendry (1995), quando o propósito é a previsão, manter as variáveis integradas pode ou não ser satisfatório. Engle e Yoo (1987, apud Hendry, 1995, p. 354) chegam à conclusão de que o URF VAR em nível tem performance pior que o VEC para longos horizontes de previsão, enquanto que Clements e Hendry (1994, apud Hendry, 1995, p. 354), concluem que, considerando a média quadrática do erro de previsão, em sistemas em diferenças e/ou cointegrados, ocorre o oposto, dependendo de como o VEC é estimado.

2.6 Modelo Estrutural

Modelos Estruturais de séries de tempo são modelos formulados diretamente em termos de seus componentes de interesse, quais sejam, tendência, ciclo, sazonalidade e componente irregular, os quais têm uma interpretação direta. Além de terem considerável apelo intuitivo, provêm uma clara ligação com modelos de regressão. Os modelos básicos descritos a seguir são baseados em Harvey (1989), onde podem ser vistos com mais detalhes.

A essência da abordagem estrutural em séries de tempo pode ser entendida por meio do modelo estrutural mais simples, denominado passeio aleatório mais ruído (*random walk plus noise*) ou modelo nível local (*local level*), que pode ser descrito como:

$$y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim \text{NID}(0, \sigma_\varepsilon^2), \quad (19)$$

$$\mu_t = \mu_{t-1} + \eta_t, \quad \eta_t \sim \text{NID}(0, \sigma_\eta^2), \quad (20)$$

em que $\text{NID}(0, \sigma^2)$, refere-se a uma variável aleatória, com média zero e variância σ^2 , normalmente distribuída e serialmente independente. A essência deste modelo se baseia no componente de nível, μ_t , neste caso um processo passeio aleatório (RW – *random walk*), composto do nível no período anterior mais um distúrbio ruído branco (WN – *white noise*), e no componente irregular, ε_t , que também é uma seqüência WN, ambos não diretamente observáveis. Entretanto, μ_t pode ser estimado e previsões futuras serão iguais à sua estimativa no tempo T . Esta estimativa será a média ponderada das observações e a extensão na qual

observações passadas são descontadas depende do valor relativo entre as variâncias dos dois distúrbios. No caso extremo de σ_ε^2 igual a zero, as observações seguem um processo RW e a previsão de valores futuros é simplesmente a última observação, y_T . Caso σ_η^2 seja zero, o nível é constante e a melhor previsão para o futuro é a média amostral. As duas grandes virtudes deste modelo são a simplicidade e a interpretação direta.

O modelo de regressão no qual as variáveis explicativas são função do tempo constitui o ponto de partida para o desenvolvimento de modelos de séries de tempo. Considere o modelo a seguir:

$$y_t = \alpha + \beta t + \sum_j \gamma_j z_{jt} + \varepsilon_t, \quad (21)$$

em que α e β são coeficientes associados à tendência e γ está associado ao componente sazonal. Saliente-se que outros componentes, tais como ciclo ou uma tendência quadrática no tempo podem ser adicionados.

A característica essencial do modelo acima está no fato de que o único componente estocástico do modelo é $\varepsilon_t \sim \mathbf{NID}(0, \sigma^2)$, o que o torna inapropriado para a maioria das aplicações, principalmente para se fazer previsões, já que os componentes de tendência e sazonalidade, por serem deterministas, fazem com que todas as observações tenham o mesmo peso no momento da previsão. Entretanto, ao se fazer previsões, faz mais sentido colocar relativamente mais peso nas observações mais recentes. Assim, a abordagem essencialmente estática da equação (21) deve ser alterada para permitir a inclusão de componentes estocásticos no modelo, tornando-o um modelo dinâmico, com parâmetros variantes no tempo.

Dessa forma, os diversos componentes da abordagem estrutural de séries de tempo, quando tratados de maneira estocástica, podem ser representados como se segue:

Tendência: representa o componente de longo prazo da série e em sua modelagem mais simples consiste de um componente de tendência, que pode ter diversas formas, mais um distúrbio aleatório. Na forma linear tem-se:

$$y_t = \mu_t + \varepsilon_t, \quad t = 1, \dots, T, \quad (22)$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t,$$

$$\beta_t = \beta_{t-1} + \zeta_t, \quad t = \dots, -1, 0, 1, \dots \quad \eta_t \sim \mathbf{NID}(0, \sigma_\eta^2), \zeta_t \sim \mathbf{NID}(0, \sigma_\zeta^2), \quad (23)$$

em que η_t , que afeta o nível da tendência, e ζ_t , que afeta a inclinação, são não correlacionados. Quanto maiores as variâncias, maiores serão os movimentos estocásticos na tendência. Se σ_η^2 e σ_ζ^2 forem iguais a zero, tem-se o caso limite de tendência determinista.

Ciclo: seja ψ_t uma função cíclica do tempo, com frequência λ_c e período $2\pi/\lambda_c$. O ciclo, expresso em termos das funções seno e co-seno, pode ser representado por:

$$\psi_t = \alpha \cos \lambda_c t + \beta \sin \lambda_c t, \quad (24)$$

em que $(\alpha^2 + \beta^2)^{1/2}$ é a amplitude e $\tan^{-1}(\beta/\alpha)$ é a fase. Acima, tem-se a forma determinista do ciclo. De forma estocástica, considerando $y_t = \psi_t + \varepsilon_t$, $t = 1, \dots, T$, tem-se:

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos \lambda_c & \sin \lambda_c \\ -\sin \lambda_c & \cos \lambda_c \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix}, \quad (25)$$

em que κ_t e κ_t^* são distúrbios WN, não correlacionados e com mesma variância para permitir a identificação do modelo; ρ é um fator de amortecimento, com valor entre zero e um. Ressalte-se que o modelo acima se transforma em um processo AR(1) quando $\lambda_c = 0$ ou π .

Sazonalidade: partindo-se do componente sazonal na equação (21) e introduzindo o componente estocástico, tem-se:

$$\gamma_t = -\sum_{j=0}^{s-1} \gamma_{t-j} + \omega_t, \quad \omega_t \sim \text{NID}(0, \sigma_\omega^2). \quad (26)$$

Quanto maior a variância σ_ω^2 em relação σ_ε^2 , maior peso as observações recentes têm na construção do padrão sazonal para a função de previsão. Uma maneira alternativa de se representar o componente sazonal pode ser feita por meio de termos trigonométricos, de forma similar ao componente cíclico. Assim, o componente sazonal no tempo t será:

$$\gamma_t = \sum_{j=1}^{\lfloor s/2 \rfloor} (\gamma_j \cos \lambda_j t + \gamma_j^* \sin \lambda_j t), \quad (27)$$

em que $\lambda_j = 2\pi j/s$, $j = 1, \dots, \lfloor s/2 \rfloor$, é a frequência sazonal. Neste modelo, devido à característica do componente sazonal de apresentar alterações relativamente suaves ao longo dos anos, muitas vezes podemos desprezar frequências mais altas, chamadas harmônicas, e considerar a frequência mais baixa, ou fundamental, que é a que realmente importa.

Até aqui os principais componentes do modelo estrutural foram tratados separadamente, de maneira sucinta, para facilidade de compreensão. Ocorre que estes componentes geralmente estão presentes nos modelos de forma agregada, formando diversas classes de modelos estruturais. Um resumo dos principais modelos estruturais e suas propriedades pode ser encontrado na tabela 35 no Apêndice A, retirada de Harvey (1989).

De acordo com Harvey (1989), o modelo estrutural básico (BSM, modelo F da tabela 35 - Apêndice A) pode ser ampliado pela inclusão de um conjunto de variáveis x_t capazes de

explicar algo do comportamento de \mathbf{y}_t . Neste caso, tem-se o modelo estrutural com inclusão de variáveis exógenas (EMX), que pode ser definido como:

$$y_t = \mu_t + \gamma_t + \mathbf{x}_t' \boldsymbol{\delta} + \varepsilon_t, \quad t = 1, \dots, T,$$

em que \mathbf{x}_t é um vetor ($\mathbf{k} \times \mathbf{1}$) de variáveis explicativas, que devem ser fracamente exógenas²⁰, e $\boldsymbol{\delta}$ é o vetor associado de parâmetros. Evidentemente, o modelo pode ser manipulado para acrescentar ou excluir componentes. Se valores defasados de \mathbf{x}_t entram na equação acima, a variável \mathbf{x}_t é chamada de indicador de direção (*leading indicator*).

²⁰ De maneira análoga ao modelo TFM, quando o propósito é a previsão, a variável explicativa deve satisfazer a exigência mais rigorosa da exogeneidade forte (ver 2.4.1).

3 Procedimentos metodológicos

No presente trabalho são analisadas as séries de três tributos, compreendendo o período de 2000 a 2007, no âmbito da jurisdição da SRRF/8ªRF, que compreende o estado de São Paulo. Os tributos a serem analisados, segundo a classificação adotada pela RFB, são:

- Imposto de Importação – II;
- Imposto sobre a Renda das Pessoas Jurídicas – Demais Obrigadas ao Lucro Real – IRPJ_DOLR;
- Contribuição para o Financiamento da Seguridade Social – Demais Empresas – COFINS_DE;

Quanto às séries do IRPJ e da Cofins, a designação acima exclui a arrecadação das empresas do setor financeiro, que, supõem-se, possuem uma dinâmica própria e diferenciada das demais empresas.

A escolha dos tributos a serem analisados levou em conta a participação de cada um na arrecadação total da SRRF/8ªRF para o ano de 2007, bem como os âmbitos de incidência dos tributos, que, no presente estudo são, o setor externo, o lucro e o faturamento das empresas, para o II, IRPJ_DOLR e COFINS_DE, respectivamente. Os tributos acima especificados foram responsáveis por aproximadamente 24% da arrecadação da SRRF/8ªRF, conforme se verifica na tabela 30 do Apêndice A, que contém a arrecadação tributária federal do estado de São Paulo para o ano de 2007, por tributo (com suas respectivas subclassificações). Importante ressaltar que a 8ª RF é responsável por aproximadamente 43% da arrecadação tributária total do país, justificando-se, assim, a escolha do estado de São Paulo para a realização do presente trabalho. Na tabela 30 do Apêndice A, os tributos que serão analisados estão marcados com um asterisco (*) à direita do nome.

Para cada uma das séries acima especificadas, são efetuadas previsões da arrecadação por meio das seguintes modelagens:

- Box-Jenkins (ARIMA);
- modelo dinâmico univariado - DR e modelo de Função de Transferência - TFM, em que variáveis exógenas são incluídas no modelo dinâmico univariado;
- modelos dinâmicos de múltiplas equações: Auto-regressão Vetorial - VAR, Modelo VAR com Correção de Erros - VEC e Modelos de Equações Simultâneas - SEM;
- modelo Estrutural (EM e EMX) sem e com inclusão de variáveis exógenas, respectivamente;

Para conclusão final, os métodos de previsão são comparados, entre si e com a previsão efetuada pela RFB para a SRRF/8ªRF, em que se utilizou o método dos indicadores. Os dados reais da arrecadação de 2007 serviram de base de comparação.

Os dados das séries de arrecadação dos tributos envolvidos na análise foram obtidos diretamente da Receita Federal do Brasil. Quanto às séries das variáveis explicativas, foram utilizadas as seguintes²¹:

- Taxa Real de Câmbio - TCR: IPEA (Instituto de Pesquisa Econômica Aplicada – www.ipea.gov.br) - Taxa de câmbio - efetiva real - INPC - exportações - índice (média 2000 = 100) – Ref: IPEA - GAC12_TCERXTINPC12.
- PIB mensal: fonte: Bacen (Banco Central do Brasil – www.bcb.gov.br) - Ref.: série nº 4380 - PIB mensal - Valores correntes (R\$ milhões) – deflacionado pelo IER²² e transformado em índice.
- Índice de Produção Industrial – IPI: IBGE (Instituto Brasileiro de Geografia e Estatística – www.ibge.gov.br) - Produção física industrial – Indústria Geral – São Paulo - Índice de base fixa mensal sem ajuste sazonal (Base: média de 2002 = 100).
- Índice de Vendas no Varejo – IVV: Bacen (Banco Central do Brasil – www.bcb.gov.br) - Ref.: série nº 1475 - Índice volume de vendas no varejo (2003=100) - Total - São Paulo.

Para cada uma das séries das variáveis explicativas são efetuadas previsões destas variáveis para o ano de 2007. Estes valores são incorporados aos modelos de Função de Transferência e Estrutural com variáveis explicativas das variáveis dependentes de arrecadação. Isto se mostrou necessário devido ao fato de que, numa situação real de aplicação do método à previsão da arrecadação, no momento da previsão, os valores reais das variáveis explicativas não são conhecidos. Portanto, nas tabelas de resultados dos modelos de Função de Transferência e Estrutural com variáveis exógenas, constam duas linhas com os erros de previsão, uma com a utilização dos valores reais das variáveis exógenas e outra com os valores previstos destas mesmas variáveis, para o ano de 2007. Para previsão das variáveis explicativas foram utilizadas as modelagens ARIMA, dinâmica univariada e Estrutural.

²¹ Todos os números-índice tiveram a base alterada para janeiro de 2000 = 100.

²² IER – Índice Específico da Receita (ponderado 55% IPCA – 45% IGP-DI) – índice de preço utilizado pela RFB na elaboração da previsão pelo método dos indicadores, que também será usado neste trabalho como índice de deflação das séries dos tributos analisados.

4 Análise das séries

4.1 Testes de igualdade de variâncias

Os testes de igualdade de variância têm o objetivo de se verificar a necessidade de transformação da série para estabilização da variância. Quando a série não apresenta variância constante, geralmente se aplica a transformação Box-Cox, definida como (BROCKWELL e DAVIS, 2002, p. 188):

$$f_{\lambda} = \begin{cases} \lambda^{-1}(U_t^{\lambda} - 1), & U_t \geq 0, \lambda > 0, \\ \ln(U_t), & U_t > 0, \lambda = 0, \end{cases}$$

Pankratz (1991) enfatiza que a transformação logarítmica, com $\lambda = 0$, é usada na maioria das vezes e que geralmente faz com que os dados fiquem mais próximos da distribuição normal.

Foram efetuados testes para se apurar a hipótese de constância da variância nas séries sem transformação e transformadas (transformação Box-Cox com $\lambda = 0$ - transformação logarítmica). A rejeição da hipótese nula indica a necessidade da transformação. A tabela 31 no Apêndice A apresenta o resultado dos testes, resumidos a seguir:

- II: O teste de Brown-Forsythe, que é superior em termos de robustez e poder, não rejeita H0 de variância constante para a série transformada, porém rejeita H0 para a série sem transformação, justificando-se, assim, a transformação. O logaritmo da série II é denominado LII.
- IRPJ_DOLR - LR: Nenhum dos três testes rejeita H0 de variância constante, para ambas as séries, sem e com transformação, porém vamos utilizar o logaritmo da série para padronização em relação às outras séries. O logaritmo da série IRPJ_DOLR é denominado LLR.
- COFINS_DE - CF: Os testes Levene e Brown-Forsythe rejeitam H0 de variância constante, para ambas as séries. Porém, vamos utilizar o logaritmo da série para padronização em relação às outras séries. O logaritmo da série COFINS_DE é denominado LCF.

A transformação logarítmica também foi aplicada às séries das variáveis explicativas.

4.2 Gráficos

Os gráficos das séries se encontram no Apêndice B (gráficos 1 a 6) e os valores das séries transformadas na tabela 32 do Apêndice A. Os valores das séries, em R\$ 10⁶, deflacionados pelo IER, e após transformação logarítmica, referem-se à arrecadação divulgada pela RFB, que é a base de comparação utilizada pelo órgão para se apurar a acurácia da previsão. Compõem-se das seguintes rubricas²³: arrecadação apropriada no dia + arrecadação apropriada em atraso no decêndio + saldo das retificações + Cide + Refis + retenção pessoa jurídica de direito privado + Paes + Estorno Paes + TPAI + RPEM + valor retenção de órgão público. Para as séries do IRPJ e Cofins, o início da série ocorre em agosto de 2000²⁴.

Os gráficos das séries das variáveis explicativas se encontram no Apêndice B (gráficos 23 e 28). Destacam-se na análise gráfica a alta variabilidade até 2004 e a tendência de queda a partir de 2003 da TCR, o forte componente sazonal presente nas séries IPI, PIB e, principalmente, IVV, além da clara tendência ascendente nestas séries a partir de 2003. Conforme era de se esperar, pois se tratam do próprio PIB ou *proxys* deste, o gráfico de dispersão (gráfico 24 - *scatter plot*) entre as séries PIB, IVV e IPI mostra forte correlação positiva entre elas, que só não é maior entre IVV/IPI e IVV/PIB devido ao elevado componente sazonal presente no mês de dezembro na série IVV. Os coeficientes de correlação, quando se consideram as séries dessazonalizadas, que foram utilizadas para detecção inicial de *outliers*, conforme se verá mais adiante, e após transformação logarítmica, são de 0,76, 0,79 e 0,87 entre LIPI/LIVV, LPIB/LIVV e LPIB/LIPI, respectivamente.

4.3 Estatísticas descritivas

As estatísticas descritivas das séries se encontram no Apêndice B (gráficos 7 a 9), merecendo destaque o fato de que a única série em que se rejeita a hipótese de distribuição normal, ainda assim, somente ao nível de 10%, é a série LII. Ao nível de 5% todas as séries apresentam distribuição normal, segundo o teste Jarque-Bera.

²³ Cide – Contribuição de Intervenção no Domínio Econômico – Combustíveis – Lei nº 10336/2001. Refis / Paes – programas de parcelamento instituídos pelas Leis nº 9964/2000 e 10864/2003, respectivamente. TPAI – tributação do patrimônio de afetação – Lei nº 10931/2004. RPEM – Retenção de CSLL, Cofins e Pis/Pasep, efetuada por órgão público dos Estados e Municípios nos pagamentos efetuados à pessoas jurídicas de direito privado pelo fornecimento de bens e serviços – Lei nº 10833/2003, art. 33.

²⁴ Disponíveis no sistema *Data Warehouse* – DW-Arrecadação da RFB.

Com relação às variáveis explicativas, somente a série LIVV não apresenta distribuição normal, conforme se verificam nos gráficos 10 a 13.

4.4 Correlogramas das séries

Os correlogramas das séries em nível, em primeira diferença e em primeira diferença sazonal se encontram no Apêndice B (gráficos 17 a 22), merecendo os seguintes comentários:

- LII: ACF indica possível presença de raiz unitária.
- LLR: ACF também indica presença de raiz unitária e forte presença de componente sazonal. A série em primeira diferença mantém o forte padrão sazonal, que somente é abrandado quando se aplica, além da primeira diferença, também uma primeira diferença sazonal com $s = 12$.
- LCF: ACF e PACF sugerem um processo AR(1).

Com relação às variáveis explicativas, os correlogramas das séries em nível se encontram nos gráficos 25 a 27 e 29, no Apêndice B.

4.5 Testes de Raiz Unitária

Para se testar a presença de raiz unitária, conforme já explicitado em 2.3.2, foram efetuados os testes DF-GLS (ERS) e Ng-Perron. Os testes foram efetuados a partir das definições padrão do software *Eviews*. Quanto à seleção automática do comprimento da defasagem foi utilizado o critério modificado de Akaike (MAIC), que é o critério recomendado pelo próprio manual do *software*, no caso do teste Ng-Perron (EVIEWES 5 USER'S GUIDE, p. 514). Todos os testes confirmam a raiz unitária para as três séries, para ambas as especificações (constante e constante mais tendência linear). Os resultados dos testes se encontram nas tabelas 3 e 4 abaixo.

Tabela 3 - Teste RU - DF-GLS (ERS)

Série	Especificação do teste	Estatística do Teste	Valores críticos			Conclusão
			1%	5%	10%	
LII	cte	-1,54	-2,59	-1,94	-1,61	RU
	cte + tend. linear	-2,55	-3,65	-3,09	-2,79	RU
LLR	cte	0,87	-2,60	-1,95	-1,61	RU
	cte + tend. linear	-0,86	-3,71	-3,14	-2,85	RU
LCF	cte	-1,54	-2,60	-1,95	-1,61	RU
	cte + tend. linear	-1,56	-3,71	-3,14	-2,85	RU

Fonte: elaboração do autor

Tabela 4 - Teste RU - Ng-Perron

Série	Especificação do teste	Estatística MZa	Valores críticos			Estatística MZt	Valores críticos			Conclusão
			1%	5%	10%		1%	5%	10%	
LII	cte	-4,83	-13,80	-8,10	-5,70	-1,50	-2,58	-1,98	-1,62	RU
	cte + tend. linear	-11,49	-23,80	-17,30	-14,20	-2,38	-3,42	-2,91	-2,62	RU
LLR	cte	0,34	-13,80	-8,10	-5,70	0,45	-2,58	-1,98	-1,62	RU
	cte + tend. linear	-0,76	-23,80	-17,30	-14,20	-0,58	-3,42	-2,91	-2,62	RU
LCF	cte	-3,68	-13,80	-8,10	-5,70	-1,30	-2,58	-1,98	-1,62	RU
	cte + tend. linear	-4,35	-23,80	-17,30	-14,20	-1,40	-3,42	-2,91	-2,62	RU

Fonte: elaboração do autor

A seguir apresentam-se os testes de raiz unitária das variáveis explicativas, em que também se confirma a presença de raiz unitária para as quatro séries, para ambas as especificações (constante e constante mais tendência linear).

Tabela 5 - Teste RU - DF-GLS (ERS)

Série	Especificação do teste	Estatística do Teste	Valores críticos			Conclusão
			1%	5%	10%	
LTCR	cte	-1,11	-2,60	-1,95	-1,61	RU
	cte + tend. linear	-1,24	-3,67	-3,11	-2,81	RU
LIPI	cte	1,14	-2,60	-1,95	-1,61	RU
	cte + tend. linear	-2,55	-3,69	-3,12	-2,82	RU
LPIB	cte	0,76	-2,60	-1,95	-1,61	RU
	cte + tend. linear	-1,28	-3,68	-3,12	-2,82	RU
LIVV	cte	-0,77	-2,60	-1,95	-1,61	RU
	cte + tend. linear	-1,68	-3,69	-3,12	-2,82	RU

Fonte: elaboração do autor

Tabela 6 - Teste RU - Ng-Perron

Série	Especificação do teste	Estatística MZa	Valores críticos			Estatística MZt	Valores críticos			Conclusão
			1%	5%	10%		1%	5%	10%	
LTCR	cte	-2,88	-13,80	-8,10	-5,70	-1,17	-2,58	-1,98	-1,62	RU
	cte + tend. linear	-3,49	-23,80	-17,30	-14,20	-1,25	-3,42	-2,91	-2,62	RU
LIPI	cte	0,38	-13,80	-8,10	-5,70	0,59	-2,58	-1,98	-1,62	RU
	cte + tend. linear	0,32	-23,80	-17,30	-14,20	0,45	-3,42	-2,91	-2,62	RU
LPIB	cte	1,85	-13,80	-8,10	-5,70	1,19	-2,58	-1,98	-1,62	RU
	cte + tend. linear	-1,12	-23,80	-17,30	-14,20	-0,60	-3,42	-2,91	-2,62	RU
LIVV	cte	4,99	-13,80	-8,10	-5,70	1,73	-2,58	-1,98	-1,62	RU
	cte + tend. linear	3,47	-23,80	-17,30	-14,20	7,62	-3,42	-2,91	-2,62	RU

Fonte: elaboração do autor

4.6 Detecção inicial de outliers

Uma primeira análise gráfica dos dados, além da inspeção visual da série, que permite tirar conclusões sobre a presença de *outliers* (dados fora do padrão), se faz por meio dos chamados *Boxplots*, em que se obtém, ainda, informações sobre a dispersão dos dados.

Os *boxplots* das séries se encontram no Apêndice B (gráficos 14 a 16), com as seguintes conclusões:

- LII: embora verifica-se a presença de um *near outlier* em fevereiro de 2004 e dois em 2005 (fevereiro e agosto), quando se considera a totalidade dos dados não há presença de *outliers*.
- LLR: embora verifica-se a presença de um *near outlier* em janeiro de 2001 e um *far outlier* em janeiro de 2005, quando se considera a totalidade dos dados não há presença de *outliers*.
- LCF: Não se verifica a presença de *outliers*.

Esta porém, não é a maneira adequada de se inferir sobre a presença de *outliers*, já que os dados estão segregados por ano. Assim, segundo Harvey e Koopman (1992), a presença de *outliers* e quebras estruturais podem ser identificados por meio da utilização dos resíduos auxiliares, que são obtidos da estimação de modelos estruturais.

Os dois principais tipos de *outliers* encontrados são o aditivo (AO), que afeta somente uma determinada observação e a mudança estrutural (IO ou LS), que afeta o valor da série a partir de uma determinada observação. Segundo Tsay (1986), o modelo geral com *outlier* pode ser expresso da seguinte maneira:

$$X_t = \sum_{j=1}^m \omega_j v_j(L) \xi_t^{(T_j)} + Z_t, \quad Z_t = [\theta(L)/\phi(L)] a_t,$$

em que X_t é a série observada, Z_t a série não observada, que segue um processo ARMA(p,q), livre de *outliers*, ω_j representa a magnitude do *outlier*, $\xi_t^{(T)} = \mathbf{1}$ se $t = T$ e zero caso contrário, $v_j(L) = \mathbf{1}$ para AO, $v_j(L) = [\theta(L)/\phi(L)]$ para o LS no tempo $t = T_j$ e m é o número de *outliers*. Destaque-se que a presença de *outliers* pode causar sérios problemas de viés na estimação de modelos ARMA.

A partir de uma determinada especificação inicial do modelo estrutural, a análise dos resíduos auxiliares do componente irregular é um bom indicador de *outliers* do tipo aditivo, enquanto que os resíduos auxiliares do componente de nível identifica *outliers* do tipo LS.

Harvey e Koopman (1992) destacam que em um modelo gaussiano, valores dos resíduos auxiliares padronizados maiores que dois em valor absoluto, indicam a presença de *outliers*.

No presente trabalho utilizou-se o modelo básico estrutural²⁵ como especificação inicial para detecção de *outliers*, por meio da análise dos resíduos auxiliares, levando-se em consideração, ainda, que os dados foram previamente ajustados sazonalmente, conforme recomendação de Koopman *et alli* (2000). A análise dos resíduos dos modelos estimados também pode indicar a presença de *outliers* e quebras estruturais, mas geralmente não dão uma clara indicação da fonte do problema. Em todas as séries detectou-se algum tipo de *outlier*, conforme pode ser verificado nas tabelas 33 e 34, no Apêndice A, relativas às variáveis principais e explicativas, respectivamente.

Deve-se enfatizar que estes modelos básicos pretendem fornecer um ponto de partida para a identificação de *outliers*, principalmente para utilização em modelagens outras que não a estrutural. Quando da estimação de determinado modelo, a identificação dos *outliers* se desenvolve conjuntamente à especificação do modelo, não sendo necessariamente utilizados todos os identificados nas tabelas acima citadas, bem como podendo surgir outros.

Visando obter um melhor conhecimento das séries de arrecadação analisadas, a seguir é apresentada uma análise gráfica juntamente com os resultados das quebras estruturais encontradas e a análise da arrecadação efetuada pela RFB.

A análise gráfica da série do II parece mostrar comportamento bem distinto entre os períodos 2000 – 2002/2003 e 2003/2004 – 2006. O método de detecção de *outliers* aponta mudanças mais pronunciadas no nível da série em novembro / dezembro de 2001, maio de 2003 e março de 2004. A primeira queda, que foi negativa, está provavelmente associada à crise provocada pelos atentados terroristas à Nova York em setembro de 2001. Quanto à segunda, a regulamentação do imposto de importação sofreu alteração por meio do Decreto nº 4543, de 26 de dezembro de 2002, em vigor a partir de jan/2003 (novo Regulamento Aduaneiro, revoga o Decreto nº 91030/85) e esta, juntamente com a diminuição da demanda por força do desaquecimento econômico e a forte valorização cambial, ocorridos no início de

²⁵ Geralmente o modelo básico estrutural com pequenas alterações na especificação apresentou resultados satisfatórios em termos da adequação do modelo. A exceção foi o modelo para a variável explicativa TRC, em que foi necessária a introdução de defasagens da variável dependente para que se alcançasse a normalidade dos resíduos.

2003²⁶, podem ter sido as causas desta mudança estrutural. Por fim, a alteração de março de 2004 é positiva e possivelmente é devida à recuperação da atividade econômica.

A análise gráfica da série IRPJ_DOLR parece mostrar comportamento bem distinto entre os períodos 2000 – 2003 (nível aproximadamente constante) e 2004 – 2006 (com tendência ascendente). O método de detecção de *outliers* aponta mudança mais pronunciada no nível da série em agosto de 2004. A causa desta alteração também parece ser devida à recuperação da atividade econômica. Apresenta, ainda, forte *outlier* do tipo aditivo em maio de 2003.

Quanto à série da COFINS_DE, a análise gráfica parece mostrar comportamento bem distinto entre os períodos 2000 – meados de 2004 (nível aproximadamente constante) e segundo semestre de 2004 – 2006 (com tendência ascendente). O método de detecção de *outliers* aponta mudança (negativa) mais pronunciada no nível da série em julho de 2004. No ano de 2004 a legislação da Cofins sofreu muitas alterações, principalmente em função da instituição da Cofins não cumulativa²⁷ e da incidência sobre importações²⁸. Assim, a Cofins sobre o faturamento apresentou tendência de declínio ao longo de 2004, em virtude da instituição da incidência não cumulativa e da incidência sobre importações. Já a Cofins não cumulativa apresentou tendência de estabilização (apenas oscilações sazonais) no 2º semestre, quando já se fizeram sentir os efeitos da incidência sobre importações (desembaraço de mercadorias por filiais) e de outras alterações de legislação relacionadas à desoneração de cadeias produtivas (Lei nº 10925/2004 e MP nº 183/2004) e ao estímulo aos investimentos (MP nº 219/2004). Assim, em função das profundas e diversas alterações legislativas, o imposto oscilou bastante em 2004, com tendência de queda na arrecadação, que passa a crescer novamente, com tendência ascendente, a partir de 2005, neste ponto devido, provavelmente, à recuperação da atividade econômica.

4.7 Testes de exogeneidade das variáveis explicativas

Os testes de exogeneidade das variáveis explicativas, tendo por base a causalidade de Granger em sistemas cointegrados, para identificação das variáveis mais apropriadas para as

²⁶ Lembre-se que no segundo semestre de 2002 houve espantosa desvalorização cambial face à instabilidade decorrente do processo eleitoral e, num primeiro momento, a valorização do câmbio, ocorrida no início de 2003, diminui a base de cálculo e, conseqüentemente, a arrecadação do II.

²⁷ Lei nº 10833, de 29 de dezembro de 2003, com vigência a partir de março de 2004, que foi projetada de sorte a manter os níveis de arrecadação. Dessa maneira, para compensar diversas deduções permitidas à base de cálculo, promoveu-se elevação da alíquota de 3% para 7,6%.

²⁸ Lei nº 10865, de 30 de abril de 2004, com vigência a partir de maio de 2004.

modelagens de Função de Transferência e Estrutural com variáveis exógenas, foram efetuados a partir da análise dos coeficientes e da velocidade de ajustamento (parâmetro α) dos modelos VEC/SEM²⁹ e são apresentados na tabela 7 a seguir.

Tabela 7 – Resultados dos testes de exogeneidade das variáveis explicativas em modelos multivariados

Modelo VEC/SEM	Posto da matriz Π (*1)	Variável explicativa	Velocidade de ajustamento (*2)		Var. principal --> Var. explicativa (*3)	Conclusão (*4)
			α_1	α_2		
DLII/DLIVV/DLTCR VEC/SEM Nº 3 – Tab. 12	2	DLIVV	-	-0,12 (0,004)	SIM	endógena
		DLTCR	-	-0,31 (0,000)	SIM	endógena
DLII/DLIVV/DLTCR VEC/SEM Nº 6 – Tab. 12	2	DLIVV	0,12 (0,000)	-0,054 (0,048)	SIM	endógena
		DLTCR	0,17 (0,000)	-0,028 (0,000)	SIM	endógena
DLII/DLIPI/DLTCR VEC/SEM Nº 9 – Tab. 12	2	DLIPI	0,29 (0,000)	-0,027 (0,000)	SIM	endógena
		DLTCR	-	-0,026 (0,000)	SIM	endógena
DLII/DLPIB/DLTCR VEC/SEM Nº 12 – Tab. 12	2	DLPIB	-0,066 (0,000)	-0,44 (0,000)	NÃO	endógena
		DLTCR	-0,17 (0,000)	0,80 (0,000)	SIM	endógena
DLLR/DLPIB VEC/SEM Nº 3 – Tab. 18	1	DLPIB	-	-	NÃO	exógena forte
DLLR/DLIVV VEC/SEM Nº 6 – Tab. 18	0	DLIVV	-	-	SIM	endógena
DLLR/DLIPI VEC/SEM Nº 9 – Tab. 18	1	DLIPI	0,054 (0,018)	-	SIM	endógena
DLCF/DLPIB VEC/SEM Nº 3 – Tab. 23	1	DLPIB	-	-	NÃO	exógena forte
DLCF/DLIVV VEC/SEM Nº 6 – Tab. 23	1	DLIVV	-	-	SIM	exógena fraca
DLCF/DLIPI VEC/SEM Nº 9 – Tab. 23	1	DLIPI	0,34 (0,000)	-	SIM	endógena

Fonte: elaboração do autor.

Notas: (*1) Posto da matriz Π do modelo URF VAR correspondente.

(*2) Coeficiente e **t-prob** da velocidade de ajustamento (α) presente na equação da variável dependente no modelo VEC/SEM.

(*3) Indica se algum *lag* da variável principal está presente na equação da variável explicativa.

(*4) Indica a condição da variável explicativa no modelo.

A tabela indica o modelo VEC/SEM resultado da redução para I(0) dos modelos URF VAR mais adequados, o posto da matriz Π , que corresponde ao número de relações de

²⁹ A descrição dos modelos está apresentada em 6.3 e Apêndice C.

cointegração presentes no modelo, a velocidade de ajustamento da variável explicativa e se algum *lag* da variável principal está presente na equação da variável explicativa. Caso o(s) parâmetro(s) α seja(m) igual(is) a zero, a variável é dita fracamente exógena. Se, adicionalmente, nenhum *lag* da variável principal está presente na equação da respectiva variável explicativa, a variável é dita fortemente exógena e um modelo de Função de Transferência ou Estrutural, mais simples e com menos parâmetros, pode ser estimado com a inclusão da variável explicativa como variável exógena. Caso contrário, ou seja, o(s) parâmetro(s) α seja(m) diferente(s) de zero, independentemente da condição do *lag* da variável principal estar presente na equação da variável explicativa, a variável é endógena e o modelo VEC/SEM é o mais adequado para se efetuar previsões da variável principal, que é a finalidade deste trabalho.

A conclusão é de que para os modelos que envolvem o Imposto de Importação, todas as variáveis explicativas (IVV, IPI e PIB) devem ser tratadas como endógenas e o modelo VEC/SEM é o mais adequado. Já para os modelos do Imposto de Renda e Cofins, a variável PIB pode ser considerada fortemente exógena e modelos de Função de Transferência e Estrutural com variáveis exógenas também foram estimados. A variável IVV se mostrou fracamente exógena para o modelo da Cofins, já que, embora $\alpha = 0$, o *lag* 3 de DLDCF está presente na equação de DLIVV, com **t-prob** de 1,2%. A retirada de DLDCF_3 da equação de DLIVV é aceita pelo teste LR de sobre-identificação da restrição, que considera a redução somente entre um modelo inicial e um final. Porém, reduções a partir de modelos intermediários são rejeitadas e, assim, DLDCF_3 foi mantida na equação de DLIVV. Isto posto e, ademais, considerando que se trata de somente um *lag* de DLDCF e com probabilidade superior a 1%, também foi estimado um modelo de Função de Transferência para DLDCF, com DLIVV como variável exógena. A variável IPI, para ambos os modelos, IRPJ e Cofins, e IVV, para o modelo do IRPJ, devem ser tratadas como endógenas.

5 Modelos e métodos de estimação

Nos modelos multivariados para as séries do IRPJ_DOLR e da COFINS_DE, além da série do próprio tributo, foram utilizadas as séries do PIB ou as *proxys* do PIB, IPI ou IVV, individualmente para cada especificação. Para a série do II, além destas, também utilizadas individualmente, foi utilizada conjuntamente a série da TCR, em cada especificação.

Para estimação dos modelos foram utilizados os seguintes softwares e métodos de estimação:

- ARIMA: E-views (versão 5.0) – OLS – *Ordinary Least Square*;
- Modelos univariados dinâmicos: *PcGive* (versão 10.0b) - OLS – *Ordinary Least Square*;
- Modelos multivariados dinâmicos – VAR/VEC/SEM: *PcGive* (versão 10.0b) – URF VAR: OLS – *Ordinary Least Square*; VEC: *Reduced Rank Regression*; SEM: FIML – *Full Information Maximum Likelihood*;
- Modelos de Função de Transferência - TFM: *PcGive* (versão 10.0b) - OLS – *Ordinary Least Square*;
- Modelos Estruturais – EM e EMX: *Stamp (Structural Time Series Analyser, Modeller and Predictor* – versão 6.20) – *Maximum Likelihood*.

PcGive e *Stamp* são módulos do software *GiveWin* (versão 2.02).

Na especificação ARIMA, para permitir a comparação entre os diversos modelos por meio dos critérios de ajustamento AIC e BIC, as estimações foram efetuadas sobre o mesmo intervalo de dados. Ocorre que, quando o modelo estimado inclui variáveis em diferença, especialmente diferença sazonal, isso leva à perda de muitas observações no início da série (para série mensal a primeira diferença sazonal leva à perda de 12 observações). Assim, o intervalo de estimação da série é determinado em função do modelo com maior número de parâmetros a serem estimados. Na especificação dos demais modelos (DR, TFM, VAR/VEC/SEM, EM) também foi mantido o mesmo intervalo de dados, porém, diferente do intervalo de estimação da especificação ARIMA, já que não há perda de informação devido à diferença sazonal presente neste último.

A modelagem dinâmica multivariada é usada, também, para se testar a exogeneidade das variáveis explicativas. O roteiro básico, tanto para estimação e previsão, quanto para identificação das variáveis exógenas, envolve os seguintes passos no *PcGive*:

- 1º- O procedimento tem início com um modelo bem geral – VAR irrestrito (URF VAR) com pelo menos 6 defasagens das variáveis dependentes, constante

(irrestrita), tendência (restrita), sazonalidade centrada (CSEAS - média zero, irrestrita), *outliers* das variáveis dependentes, geralmente conhecidas das outras modelagens (principalmente modelos estruturais), ou identificadas por meio dos resíduos do próprio modelo. As variáveis *dummies* que representam os *outliers* são tratadas também como irrestritas.

- 2º- A partir da especificação acima, reduz-se o modelo gradativamente, retirando-se as variáveis não significativas no teste **F** (*retained regressors*), até se chegar a um modelo mais parcimonioso. No caso das variáveis dependentes mantém-se sempre o mesmo número de *lags* para todas elas. A redução do modelo é testada por meio do comando “*Progress*” (teste-**F** – baseado em Rao’s *F-approximation*). São verificados, também, os testes de constância de parâmetros, autocorrelação, normalidade e heteroscedasticidade, com ênfase no sistema como um todo, sem deixar de considerar, quando possível, estes mesmos testes em relação às variáveis dependentes isoladamente.
- 3º- A etapa anterior finaliza o VAR irrestrito. Procede-se, então, ao teste de cointegração para descobrir se há cointegração entre as variáveis, por meio do posto (**r**) da matriz **II**. Também é feita análise de cointegração I(2).
- 4º- A partir do posto de **II**, estima-se novamente o modelo, agora VAR cointegrado, com imposição do posto de **II**, sem restrições impostas sobre a velocidade de ajustamento (**α**) e/ou o vetor de cointegração (**β**).
- 5º- Impõem-se restrições sobre os parâmetros **α** e/ou **β**, para verificar a exogeneidade fraca das variáveis que supostamente foram utilizadas no modelo para ajudar a prever a arrecadação.
- 6º- A etapa anterior finaliza o VAR cointegrado. Para se obter um modelo mais parcimonioso, após o passo 4 ou 5, o modelo é reduzido para I(0), formando um modelo de equações simultâneas, em diferenças e com as relações de cointegração (identidades), a partir do qual restrições podem ser impostas em cada equação individualmente, bem como sobre as identidades, e os testes de adequação do modelo são efetuados. A redução do modelo é testada por meio do comando “*Progress*” (neste caso um teste χ^2 – *over-identifying restrictions*). Caso o parâmetro estimado para as identidades de determinada equação (variável dependente **y**₂, por exemplo) seja igual zero, a respectiva variável é considerada fracamente exógena. Se, adicionalmente, valores passados das outras variáveis (**y**₁, **y**₃...) podem ser retirados da equação de **y**₂, sem perda de informação, ou seja,

y_1 , y_3 etc. não causam y_2 no sentido de Granger, então y_2 é fortemente exógena. Neste caso, pode-se estimar um modelo de Função de Transferência ou Estrutural em que y_2 é inserida como variável exógena, tornando os modelos mais simples de serem estimados, devido à diminuição do número de parâmetros, e previsões dinâmicas podem ser efetuadas.

- 7º- Quando o posto de Π (3º passo) é zero, estima-se ao VAR em 1ª diferença, já que as variáveis são I(1) e não há relação de cointegração entre elas. No caso de posto completo, o VAR irrestrito já é o modelo a ser utilizado, o que indica que as variáveis são I(0).

As modelagens dinâmica univariada e de Função de Transferência seguem basicamente os dois primeiros passos da modelagem multivariada. Quanto aos modelos Estruturais, inicia-se com o modelo básico estrutural (BSM – modelo **F** – Tabela 35), a partir do qual, em função dos resultados obtidos em cada etapa, são feitas alterações nos componentes do modelo, sempre verificando os testes de autocorrelação, normalidade e heteroscedasticidade dos resíduos, até se chegar ao modelo mais adequado.

5.1 Métodos de comparação da previsão

Os modelos estimados utilizam dados até dezembro de 2006. Os valores de 2007 são deixados de fora da amostra para permitir a comparação com os valores previstos pelos modelos. Segundo Melo (2001), os métodos mais utilizados para se medir a acurácia da previsão são o desvio absoluto médio (MAD), o erro quadrático médio (MSE) e o erro percentual absoluto médio (MAPE). Todos utilizam os resíduos em seus cálculos, definido como a diferença entre o valor previsto pelo modelo e o valor real, ou seja: $e_t = P_t - R_t$. Assim, os erros de previsão são definidos pelas seguintes fórmulas, atentando-se para o fato de que no presente trabalho utiliza-se a raiz do erro quadrático médio (RMSE), ao invés do MSE:

$$\text{MAD} = \frac{\sum |e_t|}{n}, \quad \text{RMSE} = \sqrt{\frac{\sum e_t^2}{n}}, \quad \text{MAPE} = \frac{\sum \left| \frac{e_t}{R_t} \right|}{n} * 100,$$

em que n é o número de valores previstos obtidos dos dados passados³⁰. O MAD é definido como a média dos valores absolutos de cada resíduo; o RMSE é a raiz quadrada da média dos

³⁰ Neste trabalho $n = 12$.

valores quadráticos de cada resíduo; o MAPE considera o erro relativo de cada previsão. Nas tabelas de resultados dos modelos estimados também é apresentado o erro de previsão anual, que pretende mensurar qual o erro percentual anual da previsão, bem como se o valor previsto foi superior ou inferior ao valor real. Esta medida é importante pois, quando se trata de estimar a receita para fins de orçamento, e mesmo para a avaliação do desempenho anual da Secretaria da Receita Federal do Brasil, o que realmente importa é uma previsão anual bem feita. Esta medida é definida por:

$$E_{\text{anual}} = \frac{\sum_1^{12} P_t - \sum_1^{12} R_t}{\sum_1^{12} R_t} * 100$$

Melo (2001) destaca que se erros elevados de previsão são inaceitáveis, então o uso do RMSE faz-se necessário, enquanto que se é possível ignorar alguns erros elevados, o MAD funciona melhor. O MAPE, por se tratar de uma medida percentual é utilizado para comparar a acurácia de duas séries temporais diferentes. Neste trabalho, embora constem nos quadros comparativos as quatro medidas descritas, o RMSE será utilizado como critério de acurácia para as comparações dos métodos de previsão apresentados. Para que os erros de previsão dos modelos em nível pudessem ser comparados com os erros dos modelos em diferenças, no caso destes últimos, foi efetuada a transformação da previsão da variável em diferença para previsão em nível.

6 Resultados

A seguir são apresentados os resultados das diversas modelagens tratadas no presente trabalho. Procurou-se dar ênfase nos resultados das previsões, com a apresentação dos quatro erros de previsão descritos em 5.1, bem como nos principais testes de adequação dos modelos. As especificações são resumidas e os resultados dos coeficientes estimados dos principais modelos se encontram no Apêndice C. Os erros de previsão obtidos pelo método dos indicadores também constam de cada tabela, para efeito de comparação.

6.1 Modelagem Box-Jenkins – ARIMA

Os resultados da modelagem ARIMA se encontram nas tabelas 8, 9 e 10. A escolha dos melhores modelos se baseou nos critérios de ajustamento (AIC e BIC), significância dos coeficientes estimados, normalidade dos resíduos, não correlação dos resíduos, estacionariedade e invertibilidade do modelo, e, finalmente, nos menores erros de previsão RMSE. As observações do ano de 2007 foram deixadas de fora da amostra para possibilitar a comparação com os valores previstos pelos modelos.

Na modelagem ARIMA os *outliers* não foram tratados. Deve-se enfatizar que o objetivo do presente trabalho é o de se desenvolver metodologias alternativas ao método dos indicadores da RFB para previsão da arrecadação federal. A modelagem ARIMA para previsão de receitas federais, que é uma das metodologias alternativas, já foi elaborada por dois outros trabalhos, conforme citado na seção 1, e, portanto, não é o objetivo mais importante aqui. Assim, a procura do modelo ARIMA ideal não foi aprofundada como deveria, por exemplo, com o tratamento dos *outliers* e quebras estruturais. O foco foi encontrar modelos adequados do ponto de vista econométrico e que servissem de base de comparação com o método dos indicadores e com os demais modelos elaborados neste trabalho. A seguir são apresentados os resultados dos modelos para as três séries tratadas.

6.1.1 Imposto de Inportação – II

O correlograma de DLII sugere AR(1) ou MA(1) com provável componente sazonal. Testaram-se diversos modelos ARIMA $(p,d,q) \times (P,D,Q)_{12}$, com $p \leq 2$, $d = 1$, $q \leq 1$, $P \leq 1$, $D = 0$, $Q \leq 1$. Os modelos foram estimados sobre o mesmo número de observações, compreendendo o período de abril de 2002 a dezembro de 2006, para viabilizar o efeito

comparativo. A previsão foi efetuada e comparada com os valores da arrecadação para o ano de 2007. Os resultados apontam para o modelo 5 – SARIMA (0,1,1)x(0,0,1)₁₂ - como o melhor modelo, com RMSE 15% inferior ao do método dos indicadores. O teste Ljung-Box (estatística Q) aponta a presença de correlação serial nos resíduos da terceira defasagem. Porém, um segundo teste de correlação serial foi efetuado (teste Breusch-Godfrey LM), não indicando a presença de autocorrelação dos resíduos.

Tabela 8 – Modelagem ARIMA - Série: II – Intervalo dos dados: 04/2002 a 12/2006

Modelos	SARIMA (p,d,q) x (P,D,Q) s=12 c = cte ; t = tend. determinista	Normal. Resíduos (*2)	Correlação Serial dos Resíduos			Estac. e Invert. (*5)	Erros de previsão			
			p (*3)	Estat. Q (Ljung-Box) (prob) (*4)	Estat. Breusch- Godfrey LM (prob)		MAD	RMSE	MAPE (%)	Anual (%)
1	(1,1,0) x (1,0,1)	0,87	4	4,77 (0,03)	4,36 (0,35)	sn/sn	30,8	38,9	10,4	-9,3
2	(1,1,0) x (0,0,1)	0,89	3	2,63 (0,11)	5,58 (0,13)	s/sn	18,2	23,0	6,6	4,7
3	(1,1,0) x (0,0,0)	0,11	12	19,79 (0,05)	15,68 (0,21)	s/-	27,8	35,7	9,4	-8,2
4	(0,1,1) x (1,0,1)	0,92	4	3,32 (0,07)	3,09 (0,54)	sn/sn	31,2	39,1	10,5	-9,6
5 (*1)	(0,1,1) x (0,0,1)	0,67	3	3,12 (0,08)	5,87 (0,12)	-/sn	15,0	19,3	5,4	2,4
6	(0,1,1) x (0,0,0)	0,28	15	22,42 (0,07)	24,67 (0,05)	-/s	27,7	35,4	9,3	-8,0
7	(0,1,0) x (0,0,1)	0,17	2	11,62 (0,00)	13,57 (0,00)	-/sn	14,9	19,6	5,4	2,4
8	(2,1,1) x (1,0,1)	0,61	6	2,94 (0,09)	3,00 (0,81)	sn/sn	29,4	37,7	9,9	-8,9
9	(2,1,1) x (0,0,1)	0,96	5	1,67 (0,20)	2,65 (0,75)	s/sn	19,9	26,0	7,3	6,6
10	(2,1,0) x (1,0,1)	0,75	10	11,72 (0,07)	15,22 (0,12)	sn/sn	31,0	39,0	10,5	-9,4
RFB							15,6	22,6	5,7	2,5

Fonte: elaboração do autor.

Notas: (*1) O resultado completo deste modelo se encontra no Apêndice C.

(*2) Probabilidade do teste Jarque-Bera para normalidade dos resíduos.

(*3) Ordem do *lag*. Na estatística Q refere-se à ordem que apresenta probabilidade inferior a 10%. Caso todas as probabilidades sejam superiores a 10%, refere-se à ordem da primeira autocorrelação computada. Este mesmo valor de p é usado na estatística Breusch-Godfrey LM.

(*4) Estatística ajustada pelos números de coeficientes ARMA estimados.

(*5) s = sim ; n = não ; sn = sim mas com módulo das raízes da eq. característica inversa superior a 0,9.

6.1.2 Imposto sobre a renda das Pessoas Jurídicas – IRPJ_DOLR

O correlograma de DLLR sugere forte presença sazonal, como era de se esperar tendo em vista a sistemática de recolhimentos do IRPJ em três cotas subsequentes ao período de apuração e concentração de recolhimentos na primeira cota. Assim, foram estimados também modelos depois de aplicada uma diferença sazonal com $s = 12$. O correlograma depois de aplicada a primeira diferença mais uma diferença sazonal, sugere AR(1) ou MA(1), com SMA(12) ou SAR(12). Os modelos foram estimados sobre o mesmo número de observações,

compreendendo o período de outubro de 2002 a dezembro de 2006. Os resultados apontam para o modelo 6, SARIMA (1,1,0)x(0,1,1)₁₂, como o melhor modelo, com RMSE aproximadamente 20% inferior ao do método dos indicadores. Aqui, também o teste Ljung-Box (estatística Q), aponta a presença de correlação serial nos resíduos da terceira defasagem. Porém, o segundo teste de correlação serial não indica a presença de autocorrelação dos resíduos. Destaque-se que o modelo 9, que é um modelo sem componente estocástico, somente com constante e tendência determinista, apresentou ótimo resultado do RMSE, 16% inferior ao modelo RFB, porém apresentou critérios de ajustamento AIC, BIC bem superiores aos demais.

Tabela 9 – Modelagem ARIMA - Série: IRPJ_DOLR – Intervalo dos dados: 10/2002 a 12/2006

Modelos	SARIMA (p,d,q) x (P,D,Q) s=12 c = cte ; t = tend. determinista	Normal. Resíduos (*2)	Correlação Serial dos Resíduos				Estac. e Invert. (*5)	Erros de previsão			
			p (*3)	Estat. Q (Ljung-Box) (prob) (*4)	Estat. Breusch- Godfrey LM (prob)			MAD	RMSE	MAPE (%)	Anual (%)
1	(0,1,1) x (0,0,1)	0,85	3	1,25 (0,26)	0,00 (1,00)	-/sn	144,9	177,0	20,3	-19,1	
2	(1,1,0) x (0,0,1)	0,74	3	0,97 (0,32)	0,05 (1,00)	s/sn	155,3	189,9	21,6	-22,1	
3	(0,1,1) x (1,0,0)	0,57	3	2,38 (0,12)	1,92 (0,59)	sn/s	133,1	159,9	18,7	-17,5	
4	(1,1,0) x (1,0,0)	0,86	3	2,76 (0,10)	3,65 (0,30)	sn/-	149,8	177,7	21,1	-21,4	
5	(0,1,1) x (0,1,1) + c	0,63	3	2,51 (0,11)	1,65 (0,65)	-/sn	186,9	215,0	27,3	-28,2	
6 (*1)	(1,1,0) x (0,1,1)	0,16	3	3,28 (0,07)	3,63 (0,30)	s/sn	97,9	123,0	14,0	-13,6	
7	(1,1,0) x (1,1,0)	0,16	6	12,62 (0,01)	11,38 (0,08)	sn/-	133,6	156,2	19,5	-18,2	
8	(1,0,0) x (0,0,1) + c + t	0,92	3	2,46 (0,12)	2,05 (0,56)	s/sn	101,0	144,5	14,0	-9,1	
9	(0,0,0) x (0,0,0) + c + t	0,27	1	1,23 (0,27)	1,19 (0,27)	-/-	95,8	127,9	14,1	-1,7	
RFB							132,6	152,2	19,3%	-17,3%	

Fonte: elaboração do autor.

Notas: (*1) O resultado completo deste modelo se encontra no Apêndice C.

(*2) Probabilidade do teste Jarque-Bera para normalidade dos resíduos.

(*3) Ordem do *lag*. Na estatística Q refere-se à ordem que apresenta probabilidade inferior a 10%. Caso todas as probabilidades sejam superiores a 10%, refere-se à ordem da primeira autocorrelação computada. Este mesmo valor de p é usado na estatística Breusch-Godfrey LM.

(*4) Estatística ajustada pelos números de coeficientes ARMA estimados.

(*5) s = sim ; n = não ; sn = sim mas com módulo das raízes da eq. característica inversa superior a 0,9.

6.1.3 Cofins – COFINS DE

O correlograma da série LCF sugere um AR(1), com possível componente sazonal AR. Os modelos foram estimados sobre o mesmo número de observações, compreendendo o período de junho de 2002 a dezembro de 2006. Os resultados apontam para o modelo 3,

SARIMA ((1,10),0,(12))x(0,1,0)₁₂ com constante e tendência determinista, como o melhor modelo, porém com o RMSE pouco superior ao do método dos indicadores. Os modelos para a série LCF ficaram mais complexos que os demais, devido à dificuldade de se encontrar normalidade nos resíduos das regressões.

Tabela 10 – Modelagem ARIMA - Série: COFINS_DE – Intervalo dos dados: 06/2002 a 12/2006

Modelos	SARIMA (p,d,q) x (P,D,Q) s=12 c = cte ; t = tend. determinista	Normal. Resíduos (*2)	Correlação Serial dos Resíduos			Estac. e Invert. (*5)	Erros de previsão			
			p (*3)	Estat. Q (Ljung-Box) (prob) (*4)	Estat. Breusch- Godfrey LM (prob)		MAD	RMSE	MAPE (%)	Anual (%)
1	(0,1,(10)) x (0,1,1)	0,82	3	3,14 (0,08)	3,21 (0,36)	-/sn	88,9	111,6	7,5	-6,1
2	((1,10),0,(12)) x (0,1,0)	0,98	4	1,54 (0,22)	1,14 (0,89)	sn/sn	103,4	123,7	8,6	-8,2
3 (*1)	((1,10),0,(12)) x (0,1,0) + c + t	0,83	4	1,54 (0,21)	1,85 (0,76)	sn/sn	73,7	89,4	6,2	-5,5
4	((10),1,1) x (1,0,1)	0,19	5	4,32 (0,04)	4,35 (0,50)	sn/sn	78,6	98,7	6,7	-4,3
5	((1,10),0,(12)) x (1,0,0)	0,83	5	3,76 (0,05)	3,68 (0,60)	sn/sn	84,0	99,5	7,0	-6,1
RFB							66,3	85,0	5,6	-4,5

Fonte: elaboração do autor

Notas: (*1) O resultado completo deste modelo se encontra no Apêndice C.

(*2) Probabilidade do teste Jarque-Bera para normalidade dos resíduos.

(*3) Ordem do *lag*. Na estatística Q refere-se à ordem que apresenta probabilidade inferior a 10%. Caso todas as probabilidades sejam superiores a 10%, refere-se à ordem da primeira autocorrelação computada. Este mesmo valor de p é usado na estatística Breusch-Godfrey LM.

(*4) Estatística ajustada pelos números de coeficientes ARMA estimados.

(*5) s = sim ; n = não ; sn = sim mas com módulo das raízes da eq. característica inversa superior a 0,9.

6.2 Modelos dinâmicos univariados e de Função de Transferência³¹

Os resultados dos modelos dinâmicos univariados (DR) e de Função de Transferência (TFM) são apresentados na tabela 11 a seguir. Todos os modelos estão em diferenças, devido à presença de raiz unitária e levam em consideração o resultado dos testes de exogeneidade das variáveis explicativas. Assim, somente foram estimados modelos TFM quando a variável explicativa se mostrou fortemente exógena, conforme resultado apresentado na tabela 7.

³¹ Os modelos de Função de Transferência são modelos dinâmicos de uma única equação, embora envolvam mais de uma série. Assim, por sua semelhança aos modelos dinâmicos univariados e por terem sido estimados pelos mesmos métodos e programas deste, são tratados no mesmo tópico.

Modelo		Testes					Erros de Previsão			
Nº	Especificação (*1)	CP (*2)	AR (*3)	N (*4)	H (*5)	RESET (*6)	MAD	RMSE	MAPE (%)	Anual (%)
1	DLII	0,52 (0,89)	1,27 (0,29)	0,95 (0,62)	1,13 (0,36)	3,05 (0,09)	27,1	31,2	9,3	-9,4
	Modelo RFB						15,6	22,6	5,7	2,5
2	DLLR	1,51 (0,15)	1,03 (0,41)	0,89 (0,64)	0,70 (0,79)	1,09 (0,30)	72,8	82,7	11,1	-5,0
	Modelo RFB						132,6	152,2	19,3	-17,3
3	DLCF	0,85 (0,60)	0,23 (0,95)	1,17 (0,56)	0,46 (0,91)	0,43 (0,51)	71,9	90,6	6,0	-5,0
4 (*7)	DLCF	0,86 (0,59)	1,28 (0,28)	1,96 (0,38)	0,77 (0,63)	1,48 (0,23)	55,3	72,5	4,6	-3,2
	(DLPIB_1)						62,6	80,4	5,2	-3,8
5 (*7)	DLCF	1,04 (0,42)	0,71 (0,62)	1,35 (0,51)	0,32 (0,97)	0,18 (0,67)	47,2	67,0	4,1	-1,0
	(DLIVV , DLIVV_1)						53,5	73,7	4,6	-2,3
	Modelo RFB						66,3	85,0	5,6	-4,5

Fonte: elaboração do autor.

Notas: (*1) A especificação completa do modelo se encontra no Apêndice C. A variável entre parênteses indica a variável, com respectiva defasagem, utilizada como exógena nos modelos TFM.

(*2) Teste Chow de constância de parâmetros. Segue uma distribuição $F(\mathbf{h}, \mathbf{T} - \mathbf{k})$, em que \mathbf{h} é o horizonte de previsão (12), \mathbf{T} o número de observações e \mathbf{k} o número de parâmetros.

(*3) Teste de autocorrelação dos resíduos, com cinco defasagens segundo distribuição $F(\mathbf{s}, \mathbf{T} - \mathbf{k} - \mathbf{s})$, em que \mathbf{s} é o número de defasagens utilizadas.

(*4) Teste de normalidade dos resíduos. Segue distribuição $\chi^2(2)$.

(*5) Teste de heteroscedasticidade. Teste F - baseado em White (1980).

(*6) *RESET* teste (*Ramsey Regression Equation Specification Error Test*). Testa se o modelo está bem especificado. Segue distribuição $F(\mathbf{1}, \mathbf{T} - \mathbf{k} - \mathbf{1})$.

(*7) Tratam-se de modelos TFM, em que o primeiro resultado dos erros de previsão utiliza os dados reais das variáveis exógenas para o ano de 2007, enquanto que no segundo resultado são utilizados os dados previstos destas mesmas variáveis, conforme detalhado na seção 3.

O modelo nº 1, para o II, estimado para o período agosto de 2000 a dezembro de 2006, não consegue aproximar a previsão obtida pelo método dos indicadores, além de apresentar significância ao nível de 10% no *Reset* teste, o que indica má especificação do modelo. Não se conseguiu chegar a resultado melhor com a modelagem dinâmica univariada.

O modelo nº 2, para o IRPJ, estimado para o período fevereiro de 2001 a dezembro de 2006, apresenta ótimos resultados, tanto com relação à congruência do modelo quanto com relação às ótimas previsões para 2007, com RMSE e erro anual de previsão 46% e 71% inferiores, respectivamente, ao obtido pelo método dos indicadores. Adite-se que este foi o melhor resultado obtido para o IRPJ, superando todas as metodologias desenvolvidas neste trabalho. Tendo em vista os resultados dos testes de exogeneidade das variáveis explicativas, estimou-se modelo de Função de Transferência com DLPIB como variável exógena, que,

porém, não apresentou bons resultados, já que nenhuma defasagem de DLPIB se mostrou significativa.

O modelo nº 3, para a Cofins, estimado para o período março de 2001 a dezembro de 2006, apresenta resultado próximo ao apresentado pelo método dos indicadores. Quando se acrescenta DLPIB como variável exógena (modelo TFM nº 4), o coeficiente da primeira defasagem, DLPIB_1, se mostra bastante significativo e há sensível melhora nos erros de previsão, com RMSE 15% inferior ao obtido pelo método dos indicadores. O modelo nº 5, Função de Transferência com DLIVV como variável exógena, melhora ainda mais os resultados, com RMSE 21% inferior ao obtido pelo método dos indicadores e erro anual de previsão de apenas 1%. Neste modelo, o coeficiente de DLIVV, contemporâneo a DLCF, é significativo ao nível de 5%, enquanto que o coeficiente da primeira defasagem, DLIVV_1, é altamente significativo (**t-prob = 0,000**).

Entretanto, o primeiro resultado dos erros de previsão para os modelos de Função de Transferência nº 4 e nº 5 utilizam os dados reais das variáveis exógenas, DLPIB e DLIVV, respectivamente, para o ano de 2007. Neste caso, constata-se a utilidade destas variáveis, pois os erros de previsão diminuem. Assim, num segundo momento, foi efetuada a previsão destas variáveis explicativas para o ano de 2007, por meio das diversas técnicas utilizadas neste trabalho, conforme explicitado na seção 3. A incorporação destes valores previstos aos modelos de Função de Transferência levou ao segundo resultado apresentado para os erros de previsão.

A melhor previsão para o PIB em 2007, obtida por meio de modelagem dinâmica univariada, prevê variação positiva de 4,7% em relação a 2006, enquanto que o PIB real de 2007 foi 5,2% superior ao de 2006. Esta previsão 10% inferior ao valor real fez com que o erro anual de previsão aumentasse de 3,2% para 3,8%, ou seja um aumento de 0,6 pontos percentuais, porém, ainda mantendo-se inferior aos obtidos pelo método dos indicadores e pelo modelo dinâmico univariado nº 3, que foram de 4,5% e 5,0%, respectivamente. Importante salientar que, no quarto trimestre de 2006, a mediana da expectativa do PIB anual para 2007, obtida no sítio do Banco Central do Brasil, era de 3,5%, que, se utilizada no modelo de Função de Transferência nº 4, obviamente após transformação em índice mensal, levaria a erro anual de previsão maior que o obtido com a utilização da previsão do PIB por meio de séries temporais.

Com relação ao IVV, o melhor resultado foi obtido pela modelagem ARIMA e prevê variação positiva para o ano de 2007 de 8,7% em relação a 2006. A variação real foi de 12,5%, ou seja 44% superior, o que elevou o erro anual de previsão de 1,0% para 2,3%,

mantendo-se, também, inferior ao obtido pelo método dos indicadores e pelo modelo dinâmico univariado nº 3.

Merece destaque o fato de que as previsões das variáveis exógenas foram feitas rapidamente e, se efetuadas com maior rigor, certamente levariam a resultados mais precisos. O maior objetivo neste ponto foi mostrar a utilidade destas variáveis na previsão da arrecadação e não obter previsões mais acuradas para as variáveis exógenas.

6.3 Modelos dinâmicos de múltiplas equações

Os resultados dos modelos dinâmicos de múltiplas equações são apresentados nas tabelas 12, 18 e 23, para o II, IRPJ_DOLR e COFINS_DE, respectivamente. As tabelas apresentam, para cada especificação das variáveis envolvidas, o resultado do modelo VAR irrestrito (URF VAR), do modelo VAR cointegrado ou do modelo URF VAR em diferenças (quando não há presença de relações de cointegração entre as variáveis) e do modelo VEC/SEM reduzido para $I(0)$, que é o modelo final apropriado para se efetuar previsões dinâmicas. São apresentadas, também, tabelas com os coeficientes estimados e **t-prob** dos modelos VEC/SEM³², que embasam as conclusões apresentadas em 4.7, bem como tabelas com os resultados dos testes que determinam se há e qual o número de relações de cointegração entre as variáveis do sistema.

6.3.1 Imposto de Importação – II

Todas as especificações para o II envolvem as variáveis taxa de câmbio real e o PIB ou alguma *proxy* do PIB, Índice de Vendas no Varejo ou Índice de Produção Industrial e são apresentadas na tabela 12 a seguir. A tabela 13 mostra o resultado dos testes do traço, para se verificar a presença e a quantidade de relações de cointegração $I(1)$, por meio da aplicação da metodologia de Johansen. Não foram encontradas relações de cointegração $I(2)$ entre as variáveis.

³² Os coeficientes das variáveis irrestritas não são mostrados nestas tabelas e se encontram no Apêndice C, que apresenta os resultados completos destes modelos.

Tabela 12 – Modelos VAR/VEC/SEM – II - Intervalo dos dados: 07/2000 a 12/2006

Modelo			Testes				Erros de Previsão			
Nº	Especificação (*1)	Modelo Final (*2)	CP (*3)	AR (*4)	N (*5)	H (*6)	MAD	RMSE	MAPE (%)	Anual (%)
1	LII/LIVV/LTCR	URF_VAR(3)	27,76 (0,84)	1,62 (0,02)	2,53 (0,87)	0,59 (1,00)	14,5	17,9	5,0	-4,3
2	LII/LIVV/LTCR	VAR_C(3,2)	-	-	2,13 (0,91)	0,30 (1,00)	13,3	14,8	4,7	-1,6
3	DLII/DLIVV/DLTCR	VEC_SEM	29,38 (0,77)	1,57 (0,03)	3,06 (0,80)	0,69 (0,98)	13,0	14,5	4,6	-1,2
4	LII/LIVV/LTCR	URF_VAR(6)	22,87 (0,96)	1,42 (0,08)	3,43 (0,75)	0,08 (1,00)	16,8	19,1	5,9	-5,2
5	LII/LIVV/LTCR	VAR_C(6,2)	-	-	2,69 (0,85)	249,02 (0,16)	10,3	12,4	3,7	-1,4
6	DLII/DLIVV/DLTCR	VEC_SEM	36,75 (0,43)	1,42 (0,07)	4,22 (0,65)	0,51 (1,00)	12,8	15,5	4,6	-3,0
7	LII/LIPI/LTCR	URF_VAR(4)	24,19 (0,93)	1,16 (0,27)	1,59 (0,95)	0,29 (1,00)	29,7	34,1	10,2	-10,5
8	LII/LIPI/LTCR	VAR_C(4,2)	-	-	0,76 (0,99)	0,01 (1,00)	14,8	16,8	5,2	-3,7
9	DLII/DLIPI/DLTCR	VEC_SEM	25,13 (0,91)	0,96 (0,54)	0,91 (0,99)	0,56 (1,00)	11,0	14,2	4,0	0,7
10	LII/LPIB/LTCR	URF_VAR(3)	30,58 (0,72)	0,87 (0,70)	0,87 (0,99)	0,47 (1,00)	14,0	17,1	4,9	-1,2
11	LII/LPIB/LTCR	VAR_C(3,2)	-	-	1,63 (0,95)	0,20 (1,00)	13,2	16,6	4,7	0,2
12	DLII/DLPIB/DLTCR	VEC_SEM	33,73 (0,58)	0,83 (0,75)	1,58 (0,95)	0,55 (1,00)	13,1	16,5	4,7	0,4
Modelo RFB							15,6	22,6	5,7	2,5

Fonte: elaboração do autor.

Notas: (*1) Var.1/Var.2/Var.3. Os resultados dos parâmetros estimados dos modelos URF_VAR e VEC_SEM se encontram no Apêndice C.

(*2) URF_VAR(P) = forma reduzida irrestrita com **P** defasagens. VAR_C (P,R) = VAR cointegrado com **P** defasagens; **R** é posto da matriz **II**. VEC_SEM = modelo de representação de correção de erro – modelo de equações simultâneas.

(*3) Teste de constância de parâmetros usando V[E] (URF_VAR), definida como a matriz de variância dos erros de previsão, a qual leva em conta a incerteza dos parâmetros e a intercorrelação entre os erros de previsão ou V[e], que leva em conta somente a incerteza dos parâmetros (VEC_SEM). Segue distribuição $\chi^2(nh)$, em que **n** é o número de variáveis dependentes e **h** o horizonte de previsão.

(*4) Teste vetorial de autocorrelação dos resíduos, com cinco defasagens segundo distribuição **F(np, Nr-q)**, em que **n** é a dimensão do sistema; **p = ns**; **s** é o número de defasagens utilizadas (5); **N = T - k - p - (n - p + 1)/2**; **r = [(n²p² - 4) / (n² + p² - 5)]^{1/2}**; **q = (np)/2 - 1**; **T** é o número de observações; **k** é o número de regressores no sistema original. Para maiores detalhes ver DOORNIK (1996).

(*5) Teste vetorial de normalidade dos resíduos. Segue distribuição $\chi^2(2n)$, em que **n** é o número de variáveis dependentes.

(*6) Teste vetorial de heteroscedasticidade. Segue distribuição **F(gh, Ns-q)**. Para maiores detalhes ver DOORNIK (1996).

Tabela 13 - Análise de cointegração I(1) - II

Modelo URF_VAR (*1)		H0: posto \leq	Estat. traço (p-valor) (*2)
Nº	Especificação		
1	LII/LIVV/LTCR	0	126,18 (0,000)
		1	51,66 (0,000)
		2	4,40 (0,686)
4	LII/LIVV/LTCR	0	144,52 (0,000)
		1	58,07 (0,000)
		2	4,58 (0,661)
7	LII/LIPI/LTCR	0	141,22 (0,000)
		1	40,28 (0,000)
		2	6,79 (0,377)
10	LII/LPIB/LTCR	0	120,26 (0,000)
		1	28,20 (0,000)
		2	1,99 (0,159)

Fonte: elaboração do autor.

Notas: (*1) Referem-se aos modelos da Tabela 12.

(*2) Segundo metodologia de Johansen.

Os modelos nº 1 a 3 envolvem LII/LIVV/LTCR, com o URF VAR contendo três *lags*, constante, tendência, sazonalidade e *outliers* do tipo AO em 10/2002, 11/2002, 02/2003 e 02/2004, e do tipo LS em 11/2001, 05/2002, 01/2003 e 04/2005. Com exceção de 02/2004, relativo ao LIVV, todos os demais foram incluídos em função da análise dos resíduos de LTCR. O sistema apresenta duas relações de cointegração, sem que se possa considerar LIVV ou LTCR como variáveis exógenas (ver tabelas 13 e 14). O modelo nº 3, redução final para I(0), cujo resultado se encontra na tabela 14, apresenta ótimos resultados de previsão, com RMSE 36% inferior ao método dos indicadores e erro anual de previsão de apenas 1,2%. O único problema do modelo é apresentar autocorrelação dos resíduos. No modelo nº 3 são aceitas as restrições impostas de que as velocidades de ajustamentos do primeiro vetor de cointegração das variáveis DLIVV e DLTCR são iguais a zero, ou seja $\alpha_{1_DLIVV} = \alpha_{1_DLTCR} = 0$.

Tabela 14 - Modelo VEC/SEM DLII/DLIVV/DLTCR – modelo nº 3 – Tabela 12

Regressor	Lag / α (*1)	Equação (*2)		
		DLII	DLIVV	DLTCR
DLII	1	-	0,07 (0,008)	0,16 (0,000)
	2	-0,14 (0,037)	-	0,10 (0,000)
	α_1	-1,79 (0,000)	-	-
	α_2	1,83 (0,000)	-	-
DLIVV	1	-1,29 (0,000)	-0,46 (0,000)	0,28 (0,000)
	2	-0,31 (0,000)	-0,11 (0,000)	-
	α_1	-	-	-
	α_2	-	-0,12 (0,004)	-

Tabela 14 - Modelo VEC/SEM DLII/DLIVV/DLTCR – modelo nº 3 – Tabela 12

Regressor	Lag / α (*1)	Equação (*2)		
		DLII	DLIVV	DLTCR
DLTCR	1	0,43 (0,039)	-	-
	2	-	-	-0,25 (0,001)
	α_1	-	-	-
	α_2	-	-	-0,31 (0,000)

Fonte: elaboração do autor.

Notas: (*1) *Lag* do regressor presente na equação da variável dependente juntamente com a velocidade de ajustamento α da relação de cointegração no modelo VEC/SEM.

(*2) Coeficiente e **t-prob** do respectivo *lag* do regressor, ou da velocidade de ajustamento (α), presente na equação da variável dependente no modelo VEC/SEM.

Uma segunda especificação, modelos nº 4 a 6 da tabela 12, para tentar resolver o problema de autocorrelação dos resíduos, aumenta o número de *lags* do URF VAR para seis e acrescenta mais um *outlier* em 10/2001 para o LIVV, com resultados de previsão muito semelhantes. O problema agora é que o número de parâmetros a serem estimados aumenta muito, o que é indesejável, para pouco ganho, já que o modelo ainda continua com autocorrelação dos resíduos ao nível de 10%. O critério de informação de Schwarz, BIC, aumenta de -9,56 para -9,02 no URF VAR. Além do mais, há um pequeno aumento nos erros de previsão, o que faz do modelo nº 3 o mais adequado. Os resultados do modelo nº 6, redução final para I(0) se encontram na tabela 15 abaixo.

Tabela 15 - Modelo VEC/SEM DLII/DLIVV/DLTCR – modelo nº 6 – Tabela 12

Regressor	Lag / α (*1)	Equação (*2)		
		DLII	DLIVV	DLTCR
DLII	1	-0,37 (0,001)	-	-
	2	-0,37 (0,000)	-	-
	3	-	-	-
	4	-	0,051 (0,011)	-
	5	-	-	-
	α_1	-0,42 (0,000)	-	-
	α_2	-0,057 (0,000)	-	-
	DLIVV	1	-0,86 (0,001)	-0,26 (0,000)
2		-	-0,11 (0,000)	0,22 (0,000)
3		-	-	0,096 (0,000)
4		-	-	-
5		0,21 (0,000)	0,098 (0,000)	-
α_1		-	0,12 (0,000)	-
α_2		-	-0,0054 (0,048)	-
DLTCR		1	0,99 (0,000)	-
	2	-	-	-0,15 (0,023)
	3	-	-	0,14 (0,055)
	4	-	-	-
	5	-	-	0,28 (0,000)
	α_1	-	-	0,17 (0,000)
	α_2	-	-	-0,028 (0,000)

Fonte: elaboração do autor.

Notas: (*1) *Lag* do regressor presente na equação da variável dependente juntamente com a velocidade de ajustamento α da relação de cointegração no modelo VEC/SEM.

(*2) Coeficiente e **t-prob** do respectivo *lag* do regressor, ou da velocidade de ajustamento (α), presente na equação da variável dependente no modelo VEC/SEM.

Nos modelos n° 7 a 9 da tabela 12, a *proxy* do PIB passa a ser o Índice de Produção Industrial. O URF VAR contém quatro *lags*, constante, tendência, sazonalidade, os mesmos *outliers* para LTCR, do tipo AO em 12/2001 e LS em 09/2003 para o LIPI e do tipo AO em 04/2004 para o LII. O sistema continua com duas relações de cointegração, sem que se possa considerar LIPI ou LTCR como variáveis exógenas (ver tabelas 13 e 16). O modelo n° 9, redução final para I(0), cujo resultado se encontra na tabela 16, também apresenta ótimos resultados de previsão, um pouco superiores ao modelo com LIVV, com RMSE 37% inferior ao do método dos indicadores e erro anual de previsão de apenas 0,7%, sem problemas nos testes de adequação do modelo. No modelo n° 9, $\alpha_{1_DLTCR} = 0$.

Tabela 16 - Modelo VEC/SEM DLII/DLIPI/DLTCR – modelo n° 9 – Tabela 12

Regressor	Lag / α (*1)	Equação (*2)		
		DLII	DLIPI	DLTCR
DLII	1	-0,60 (0,000)	-0,09 (0,007)	0,20 (0,000)
	2	-0,24 (0,002)	-	0,15 (0,000)
	3	-	-	-
	α_1	0,20 (0,009)	-	-
	α_2	-0,05 (0,000)	-	-
	DLIPI	1	-	-
DLIPI	2	-	-	-0,15 (0,032)
	3	0,55 (0,001)	0,31 (0,000)	-
	α_1	-	0,29 (0,000)	-
	α_2	-	-0,027 (0,00)	-
	DLTCR	1	0,93 (0,000)	-
DLTCR	2	-	-	-0,27 (0,001)
	3	-	-	-
	α_1	-	-	-
	α_2	-	-	-0,026 (0,000)

Fonte: elaboração do autor.

Notas: (*1) *Lag* do regressor presente na equação da variável dependente juntamente com a velocidade de ajustamento α da relação de cointegração no modelo VEC/SEM.

(*2) Coeficiente e **t-prob** do respectivo *lag* do regressor, ou da velocidade de ajustamento (α), presente na equação da variável dependente no modelo VEC/SEM.

Por fim, nos modelos n° 10 a 12 da tabela 12, utiliza-se o próprio PIB. O URF VAR contém três *lags*, constante, sazonalidade, os mesmos *outliers* para LTCR e do tipo AO em 09/2000, 06/2001, 06/2005 e 04/2006 para o LPIB. O sistema possui duas relações de cointegração, sem que se possa considerar LPIB ou LTCR como variáveis exógenas (ver tabelas 13 e 17). O modelo n° 12, redução final para I(0), cujo resultado se encontra na tabela 17, apresenta bons resultados de previsão, um pouco superiores ao modelo com LIVV e LIPI, RMSE 27% inferior ao do método dos indicadores e erro anual de previsão de apenas 0,4%, sem problemas nos testes de adequação do modelo. No modelo n° 12, $\alpha_{1_DLII} = \alpha_{2_DLPIB} = 0$.

Tabela 17 - Modelo VEC/SEM DLII/DLPIB/DLTCR – modelo nº 12 – Tabela 12

Regressor	Lag / α (*1)	Equação (*2)		
		DLII	DLPIB	DLTCR
DLII	1	-0,46 (0,000)	-0,035 (0,018)	0,16 (0,000)
	2	-0,31 (0,004)	-	0,12 (0,003)
	α_1	-	-	-
	α_2	0,75 (0,000)	-	-
DLPIB	1	-1,24 (0,000)	-0,24 (0,000)	-0,31 (0,015)
	2	-1,01 (0,001)	-0,44 (0,000)	-0,36 (0,002)
	α_1	-	-0,020 (0,000)	-
	α_2	-	-	-
DLTCR	1	1,05 (0,000)	-	-
	2	0,41 (0,093)	-	-0,28 (0,005)
	α_1	-	-	0,047 (0,009)
	α_2	-	-	0,46 (0,000)

Fonte: elaboração do autor.

Notas: (*1) *Lag* do regressor presente na equação da variável dependente juntamente com a velocidade de ajustamento α da relação de cointegração no modelo VEC/SEM.

(*2) Coeficiente e **t-prob** do respectivo *lag* do regressor, ou da velocidade de ajustamento (α), presente na equação da variável dependente no modelo VEC/SEM.

Para o II, o melhor sistema, tanto em termos de congruência do modelo, quanto com relação aos erros de previsão é o modelo nº 9, com LIPI como variável explicativa, que, além de superar o método dos indicadores, apresenta RMSE 26%, 54% e 50%, inferior ao dos melhores modelos ARIMA, dinâmico univariado e Estrutural, respectivamente.

6.3.2 Imposto sobre a Renda das Pessoas Jurídicas – IRPJ DOLR

As especificações para o IRPJ envolvem o PIB ou alguma *proxy* do PIB e são apresentadas na tabela 18 a seguir. A tabela 19 mostra o resultado dos testes do traço para se verificar a presença e a quantidade de relações de cointegração I(1), por meio da aplicação da metodologia de Johansen e não foram encontradas relações de cointegração I(2) entre as variáveis.

Tabela 18 – Modelos VAR/VEC/SEM – IRPJ_DOLR - Intervalo dos dados: 02/2001 a 12/2006

Nº	Modelo		Testes				Erros de Previsão			
	Especificação (*1)	Modelo Final (*2)	CP (*3)	AR (*4)	N (*5)	H (*6)	MAD	RMSE	MAPE (%)	Anual (%)
1	LLR/LPIB	URF_VAR(4)	27,47 (0,28)	0,91 (0,57)	0,53 (0,97)	0,58 (0,98)	102,4	115,2	15,0	-7,0
2	LLR/LPIB	VAR_C(4,1)	-	-	0,59 (0,96)	0,33 (1,00)	96,3	110,1	14,1	-4,9
3	DLLR/DLPIB	VEC_SEM	35,10 (0,07)	1,02 (0,45)	1,15 (0,89)	0,96 (0,55)	97,7	111,2	14,4	-5,4

Tabela 18 – Modelos VAR/VEC/SEM – IRPJ_DOLR - Intervalo dos dados: 02/2001 a 12/2006

Modelo			Testes				Erros de Previsão			
Nº	Especificação (*1)	Modelo Final (*2)	CP (*3)	AR (*4)	N (*5)	H (*6)	MAD	RMSE	MAPE (%)	Anual (%)
4	LLR/LIVV	URF_VAR(6)	20,87 (0,65)	0,85 (0,65)	2,63 (0,62)	0,67 (0,96)	69,6	80,3	11,4	5,7
5	DLLR/DLIVV	URF_VAR(5)	27,67 (0,27)	1,06 (0,41)	3,00 (0,56)	0,81 (0,82)	123,1	148,4	18,1	-17,9
6	DLLR/DLIVV	SEM	30,60 (0,17)	1,08 (0,38)	2,27 (0,69)	0,83 (0,78)	122,5	148,2	18,0	-17,4
7	LLR/LIPI	URF_VAR(4)	13,85 (0,95)	1,00 (0,46)	0,66 (0,96)	0,37 (1,00)	74,6	85,6	11,5	-0,2
8	LLR/LIPI	VAR_C(4,1)	-	-	0,78 (0,94)	0,34 (1,00)	90,3	102,3	13,4	-6,7
9	DLLR/DLIPI	VEC_SEM	17,98 (0,80)	1,17 (0,29)	1,58 (0,81)	0,63 (0,96)	85,5	91,9	13,0	-3,3
Modelo RFB							132,6	152,2	19,3	-17,3

Fonte: elaboração do autor.

Notas: (*1) Var.1/Var.2/Var.3. Os resultados dos parâmetros estimados dos modelos URF_VAR e VEC_SEM se encontram no Apêndice C.

(*2) URF_VAR(P) = forma reduzida irrestrita com **P** defasagens. VAR_C (P,R) = VAR cointegrado com **P** defasagens; **R** é posto da matriz **Π**. VEC_SEM = modelo de representação de correção de erro – modelo de equações simultâneas.

(*3) Teste de constância de parâmetros usando V[E] (URF_VAR), definida como a matriz de variância dos erros de previsão, a qual leva em conta a incerteza dos parâmetros e a intercorrelação entre os erros de previsão ou V[e], que leva em conta somente a incerteza dos parâmetros (VEC_SEM). Segue distribuição $\chi^2(nh)$, em que **n** é o número de variáveis dependentes e **h** o horizonte de previsão.

(*4) Teste vetorial de autocorrelação dos resíduos, com cinco defasagens segundo distribuição **F(np, Nr-q)**, em que **n** é a dimensão do sistema; **p = ns**; **s** é o número de defasagens utilizadas (5); **N = T - k - p - (n - p + 1)/2**; **r = [(n²p² - 4) / (n² + p² - 5)]^{1/2}**; **q = (np)/2 - 1**; **T** é o número de observações; **k** é o número de regressores no sistema original. Para maiores detalhes ver DOORNIK (1996).

(*5) Teste vetorial de normalidade dos resíduos. Segue distribuição $\chi^2(2n)$, em que **n** é o número de variáveis dependentes.

(*6) Teste vetorial de heteroscedasticidade. Segue distribuição **F(gh, Ns-q)**. Para maiores detalhes ver DOORNIK (1996).

Tabela 19 - Análise de cointegração I(1) – IRPJ_DOLR

Modelo URF_VAR (*1)		H0: posto ≤	Estat. traço (p-valor) (*2)
Nº	Especificação		
1	LLR/LPIB	0	58,44 (0,000)
		1	4,98 (0,606)
4	LLR/LIVV	0	17,30 (0,401)
		1	3,72 (0,778)
7	LLR/LIPI	0	63,95 (0,000)
		1	5,79 (0,498)

Fonte: elaboração do autor.

Notas: (*1) Referem-se aos modelos da Tabela 18.

(*2) Segundo metodologia de Johansen.

Nos modelos n° 1 a 3 da tabela 18, PIB é utilizada no VAR. O modelo n° 1, URF VAR, contém quatro *lags*, constante, tendência, sazonalidade e *outliers* do tipo AO em 05/2003 e 01/2006, e do tipo LS em 08/2004 para o LLR, e do tipo AO em 01/2001 para o LPIB. O sistema apresenta uma relação de cointegração e $\alpha_1_{DLPIB} = \mathbf{0}$ (ver tabelas 19 e 20). O modelo n° 3, redução final para I(0), cujo resultado se encontra na tabela 20, apresenta bons resultados de previsão, com RMSE 27% inferior ao do método dos indicadores e erro anual de previsão de apenas 5,4%, ou seja, quase 70% inferior ao do método dos indicadores. No modelo n° 3, o fato de $\alpha_1_{DLPIB} = \mathbf{0}$, juntamente com o fato de que defasagens de DLLR não entram na equação de DLPIB, faz desta última uma variável exógena forte.

Tabela 20 - Modelo VEC/SEM DLLR/DLPIB – modelo n° 3 – Tabela 18

Regressor	Lag / α (*1)	Equação (*2)	
		DLLR	DLPIB
DLLR	1	0,31 (0,029)	-
	2	0,27 (0,019)	-
	3	0,17 (0,61)	-
	α_1	-1,35 (0,000)	-
DLPIB	1	-	-0,19 (0,038)
	2	-	-
	3	-	0,45 (0,000)
	α_1	-	-

Fonte: elaboração do autor.

Notas: (*1) *Lag* do regressor presente na equação da variável dependente juntamente com a velocidade de ajustamento α da relação de cointegração no modelo VEC/SEM.

(*2) Coeficiente e **t-prob** do respectivo *lag* do regressor, ou da velocidade de ajustamento (α), presente na equação da variável dependente no modelo VEC/SEM.

Nos modelos n° 4 a 6 da tabela 18, a *proxy* do PIB passa a ser o Índice de Vendas no Varejo. O modelo n° 4, URF VAR, contém seis *lags*, tendência, sazonalidade e um *outlier* do tipo AO em 05/2003 para o LLR. O sistema apresenta posto da matriz Π igual a zero, ou seja, não há relação de cointegração entre LLR e LIVV (ver tabela 19). Neste caso deve-se estimar um modelo VAR em diferenças, que é o que foi feito no modelo n° 5, com cinco *lags* das variáveis em diferenças. O resultado não foi satisfatório em termos de previsões. Mesmo o modelo n° 6, cujo resultado se encontra na tabela 21, redução final do modelo n° 5, com algumas restrições impostas, não melhora em nada a previsão para 2007. Ainda assim, o RMSE é ligeiramente inferior ao do método dos indicadores. Neste modelo, conforme se verifica na tabela 21, DLLR causa DLIVV no sentido de Granger.

Tabela 21 - Modelo VEC/SEM DLLR/DLIVV – modelo nº 6 – Tabela 18

Regressor	Lag / α (*1)	Equação (*2)	
		DLLR	DLIVV
DLLR	1	-0,70 (0,000)	-
	2	-0,40 (0,000)	-0,035 (0,017)
	3	-0,24 (0,013)	-
	4	-0,34 (0,001)	-
	5	-0,35 (0,000)	-0,040 (0,003)
	α_1	-	-
DLIVV	1	-	-0,17 (0,000)
	2	-	-0,32 (0,000)
	3	0,39 (0,020)	-
	4	0,60 (0,002)	-
	5	0,48 (0,002)	0,093 (0,001)
	α_1	-	-

Fonte: elaboração do autor.

Notas: (*1) *Lag* do regressor presente na equação da variável dependente juntamente com a velocidade de ajustamento α da relação de cointegração no modelo VEC/SEM.

(*2) Coeficiente e **t-prob** do respectivo *lag* do regressor, ou da velocidade de ajustamento (α), presente na equação da variável dependente no modelo VEC/SEM.

Nos modelos nº 7 a 9 da tabela 18, a variável utilizada como *proxy* do PIB é o Índice de Produção Industrial. O modelo nº 7, URF VAR, contém quatro *lags*, constante, tendência, sazonalidade e *outlier* do tipo AO em 05/2003 para o LLR. O sistema apresenta uma relação de cointegração e DLIPI não pode ser considerada exógena (ver tabelas 19 e 22). O modelo nº 9, redução final para I(0), cujo resultado se encontra na tabela 22, apresenta os melhores resultados de previsão, com RMSE 40% inferior ao do método dos indicadores e erro anual de previsão de apenas 3,3%, 81% inferior ao do método dos indicadores.

Tabela 22 - Modelo VEC/SEM DLLR/DLIPI – modelo nº 9 – Tabela 18

Regressor	Lag / α (*1)	Equação (*2)	
		DLLR	DLIPI
DLLR	1	-	-
	2	-	-0,050 (0,016)
	3	-	-0,055 (0,008)
	α_1	-0,89 (0,000)	-
DLIPI	1	-1,87 (0,000)	-0,39 (0,000)
	2	-	-
	3	-	0,25 (0,003)
	α_1	-	0,054 (0,018)

Fonte: elaboração do autor.

Notas: (*1) *Lag* do regressor presente na equação da variável dependente juntamente com a velocidade de ajustamento α da relação de cointegração no modelo VEC/SEM.

(*2) Coeficiente e **t-prob** do respectivo *lag* do regressor, ou da velocidade de ajustamento (α), presente na equação da variável dependente no modelo VEC/SEM.

Para o IRPJ, o melhor sistema, tanto em termos de congruência do modelo, quanto com relação aos erros de previsão é o modelo nº 9, com LIPI como variável explicativa, que, além de superar o método dos indicadores, apresenta RMSE 25% inferior ao do modelo ARIMA. Entretanto, o modelo dinâmico univariado em diferenças apresentou o melhor resultado para este tributo.

6.3.3 Cofins – COFINS DE

Da mesma maneira que para o IRPJ, as especificações para a Cofins envolvem o PIB ou alguma *proxy* do PIB e são apresentadas na tabela 23 a seguir. A tabela 24 mostra o resultado dos testes do traço para se verificar a presença e a quantidade de relações de cointegração I(1), por meio da aplicação da metodologia de Johansen e, também, não foram encontradas relações de cointegração I(2) entre as variáveis.

Tabela 23 – Modelos VAR/VEC/SEM – COFINS - Intervalo dos dados: 02/2001 a 12/2006

Nº	Modelo		Testes				Erros de Previsão			
	Especificação (*1)	Modelo Final (*2)	CP (*3)	AR (*4)	N (*5)	H (*6)	MAD	RMSE	MAPE (%)	Anual (%)
1	LCF/LPIB	URF_VAR(4)	16,67 (0,86)	1,15 (0,32)	0,45 (0,98)	0,89 (0,67)	52,5	65,7	4,9	3,8
2	LCF/LPIB	VAR_C(4,1)	-	-	0,43 (0,98)	0,62 (0,96)	53,0	66,2	4,9	3,9
3	DLCF/DLPIB	VEC_SEM	18,04 (0,80)	0,42 (0,98)	0,75 (0,95)	0,77 (0,83)	54,7	67,1	5,0	4,1
4	LCF/LIVV	URF_VAR(5)	39,90 (0,02)	0,72 (0,79)	3,24 (0,52)	0,56 (0,99)	89,6	101,4	8,0	7,7
5	LCF/LIVV	VAR_C(5,1)	-	-	2,24 (0,69)	0,43 (1,00)	38,4	52,3	3,5	1,3
6	DLCF/DLIVV	VEC_SEM	58,04 (0,00)	0,67 (0,85)	1,40 (0,84)	0,89 (0,68)	34,6	48,5	3,1	0,4
7	LCF/LIPI	URF_VAR(4)	15,84 (0,89)	1,19 (0,28)	1,20 (0,88)	0,38 (1,00)	75,2	87,6	6,3	-5,4
8	LCF/LIPI	VAR_C(4,1)	-	-	2,35 (0,67)	0,30 (1,00)	78,1	91,1	6,6	-5,5
9	DLCF/DLIPI	VEC_SEM	17,06 (0,85)	0,82 (0,68)	4,43 (0,35)	0,61 (0,97)	77,5	90,1	6,5	-5,4
	Modelo RFB						66,3	85,0	5,6	-4,5

Fonte: elaboração do autor.

- Notas: (*1) Var.1/Var.2/Var.3. Os resultados dos parâmetros estimados dos modelos URF_VAR e VEC_SEM se encontram no Apêndice C.
- (*2) URF_VAR(P) = forma reduzida irrestrita com **P** defasagens. VAR_C (P,R) = VAR cointegrado com **P** defasagens; **R** é posto da matriz **Π**. VEC_SEM = modelo de representação de correção de erro – modelo de equações simultâneas.
- (*3) Teste de constância de parâmetros usando V[E] (URF_VAR), definida como a matriz de variância dos erros de previsão, a qual leva em conta a incerteza dos parâmetros e a intercorrelação entre os erros de previsão ou V[e], que leva em conta somente a incerteza dos parâmetros (VEC_SEM). Segue distribuição $\chi^2(nh)$, em que **n** é o número de variáveis dependentes e **h** o horizonte de previsão.
- (*4) Teste vetorial de autocorrelação dos resíduos, com cinco defasagens segundo distribuição **F(np, Nr-q)**, em que **n** é a dimensão do sistema; **p = ns**; **s** é o número de defasagens utilizadas (5); **N = T - k - p - (n - p + 1)/2**; **r = [(n²p² - 4) / (n² + p² - 5)]^{1/2}**; **q = (np)/2 - 1**; **T** é o número de observações; **k** é o número de regressores no sistema original. Para maiores detalhes ver DOORNIK (1996).
- (*5) Teste vetorial de normalidade dos resíduos. Segue distribuição $\chi^2(2n)$, em que **n** é o número de variáveis dependentes.
- (*6) Teste vetorial de heteroscedasticidade. Segue distribuição **F(gh, Ns-q)**. Para maiores detalhes ver DOORNIK (1996).

Tabela 24 - Análise de cointegração I(1) – COFINS

Modelo URF_VAR (*1)		H0: posto ≤	Estat. traço (p-valor) (*2)
Nº	Especificação		
1	LCF/LPIB	0	33,37 (0,000)
		1	1,57 (0,246)
4	LCF/LIVV	0	45,72 (0,000)
		1	7,44 (0,310)
7	LCF/LIPI	0	83,56 (0,000)
		1	9,62 (0,149)

Fonte: elaboração do autor.

Notas: (*1) Referem-se aos modelos da Tabela 23.

(*2) Segundo metodologia de Johansen.

Nos modelos nº 1 a 3 da tabela 23, PIB é utilizada no VAR. O modelo nº 1, URF VAR, contém quatro *lags*, sazonalidade e um *outlier* do tipo AO em 06/2001 para LPIB, e do tipo LS em 07/2004 para o LCF. O sistema apresenta uma relação de cointegração e $\alpha_{1_DLPIB} = \mathbf{0}$ (ver tabelas 24 e 25). O modelo nº 3, redução final para I(0), cujo resultado se encontra na tabela 25, apresenta bons resultados de previsão, com RMSE 21% inferior ao do método dos indicadores. No modelo nº 3, o fato de $\alpha_{1_DLPIB} = \mathbf{0}$, juntamente com o fato de que defasagens de DLCF não entram na equação de DLPIB, faz desta última uma variável exógena forte.

Tabela 25 - Modelo VEC/SEM DLDCF/DLPIB – modelo nº 3 – Tabela 23

Regressor	Lag / α (*1)	Equação (*2)	
		DLCF	DLPIB
DLCF	1	-	-
	2	-	-
	3	-	-
	α_1	-0,60 (0,000)	-

Tabela 25 - Modelo VEC/SEM DLCPF/DLPIB – modelo nº 3 – Tabela 23

Regressor	Lag / α (*1)	Equação (*2)	
		DLCF	DLPIB
DLPIB	1	-	-0,20 (0,029)
	2	-0,44 (0,090)	-
	3	-	0,47 (0,000)
	α_1	-	-

Fonte: elaboração do autor.

Notas: (*1) *Lag* do regressor presente na equação da variável dependente juntamente com a velocidade de ajustamento α da relação de cointegração no modelo VEC/SEM.

(*2) Coeficiente e **t-prob** do respectivo *lag* do regressor, ou da velocidade de ajustamento (α), presente na equação da variável dependente no modelo VEC/SEM.

Nos modelos nº 4 a 6 da tabela 23, a *proxy* do PIB passa a ser o Índice de Vendas no Varejo. O modelo nº 4, URF VAR, contém cinco *lags*, tendência, sazonalidade e um *outlier* do tipo LS em 07/2004 para o LCF. O sistema apresenta uma relação de cointegração (ver tabela 24). O modelo nº 6, redução final para I(0), cujo resultado se encontra na tabela 26, apresenta ótimos resultados de previsão, com RMSE e erro anual de previsão 43% e 91% inferiores ao do método dos indicadores, respectivamente. No modelo nº 3, o fato de $\alpha_1_{DLPIB} = 0$, faz de DLIVV uma variável fracamente exógena, já que o *lag* 3 de DLCF é significativo na equação de DLIVV. Conforme já comentado em 4.7, a retirada de DLCF_3 da equação de DLIVV é aceita pelo teste LR de sobre-identificação da restrição, que considera a redução somente entre um modelo inicial e um final. Porém, reduções a partir de modelos intermediários são rejeitadas e, assim, DLCF_3 foi mantida na equação de DLIVV. O problema deste modelo está no teste de constância de parâmetros, que se mostrou altamente significativo desde o início da modelagem do URF VAR.

Tabela 26 - Modelo VEC/SEM DLCPF/DLIVV – modelo nº 6 – Tabela 23

Regressor	Lag / α (*1)	Equação (*2)	
		DLCF	DLIVV
DLCF	1	-	-
	2	-	-
	3	-	0,11 (0,012)
	4	-	-
	α_1	-0,74 (0,000)	-
DLIVV	1	-	-0,31 (0,000)
	2	-0,24 (0,017)	-0,53 (0,000)
	3	-	-0,15 (0,000)
	4	-	-0,27 (0,000)
	α_1	-	-

Fonte: elaboração do autor.

Notas: (*1) *Lag* do regressor presente na equação da variável dependente juntamente com a velocidade de ajustamento α da relação de cointegração no modelo VEC/SEM.

(*2) Coeficiente e **t-prob** do respectivo *lag* do regressor, ou da velocidade de ajustamento (α), presente na equação da variável dependente no modelo VEC/SEM.

Nos modelos nº 7 a 9 da tabela 23, a variável utilizada como *proxy* do PIB é o Índice de Produção Industrial. O modelo nº 7, URF VAR, contém quatro *lags*, constante, tendência, sazonalidade e um *outlier* do tipo LS em 07/2004 para o LCF. O sistema apresenta uma relação de cointegração e DLIPI não pode ser considerada exógena (ver tabelas 24 e 27). O modelo nº 9, redução final para I(0), cujo resultado se encontra na tabela 27, apresenta resultados de previsão próximos aos alcançados pelo método dos indicadores.

Tabela 27 - Modelo VEC/SEM DLDCF/DLIPI – modelo nº 9 – Tabela 23

Regressor	Lag / α (*1)	Equação (*2)	
		DLDCF	DLIPI
DLDCF	1	-	-0,10 (0,017)
	2	-	-
	3	-0,22 (0,004)	-
	α_1	-0,31 (0,000)	-
DLIPI	1	-	-
	2	-0,57 (0,000)	-
	3	-0,37 (0,006)	0,23 (0,004)
	α_1	-	0,34 (0,000)

Fonte: elaboração do autor.

Notas: (*1) *Lag* do regressor presente na equação da variável dependente juntamente com a velocidade de ajustamento α da relação de cointegração no modelo VEC/SEM.

(*2) Coeficiente e **t-prob** do respectivo *lag* do regressor, ou da velocidade de ajustamento (α), presente na equação da variável dependente no modelo VEC/SEM.

Para a COFINS_DE, o melhor sistema, com relação aos erros de previsão é o modelo nº 6, com LIVV como variável explicativa, que, além de superar o método dos indicadores, apresenta RMSE 46%, 28% e 12%, inferior ao dos modelos ARIMA, dinâmico univariado e Estrutural, respectivamente. Deve-se enfatizar, também, o baixíssimo erro de previsão anual do modelo, de apenas 0,4%.

6.4 Modelos Estruturais

Os modelos Estruturais foram estimados por meio do software *Stamp (Structural Time Series Analyser, Modeller and Predictor)*, que é um módulo do *GiveWin*. A partir do modelo básico estrutural (BSM – Modelo F - Tabela 35 no Apêndice A), foi implementada uma rotina para detecção de *outliers*, a partir da série ajustada sazonalmente. Identificados os *outliers*, tanto do tipo AO quanto LS, somente os mais significativos foram inseridos no modelo estrutural da série original (sem ajuste sazonal) e efetuada a estimação. Os resultados dos

modelos Estruturais, sem e com a inclusão de variáveis exógenas, respectivamente, se encontram na tabela 28.

Tabela 28 - Modelos Estruturais – EM e EMX – II/IRPJ_DOLR/COFINS_DE

Mod. n°	Var. Dep.	Componentes do modelo estrutural (*1)						Testes Estat. (t-prob) (*2)				Erros de previsão			
		N	T	S	I	Outlier	Variável X_N	N	Estat. Q(P,d)	H	MAD	RMSE	MAPE (%)	E_anual (%)	
1	LII	SL	NS	FS	S	-	-	0,18 (0,91)	Q (7,6)	10,72 (0,10)	0,40 (0,99)	27,5	30,1	9,5	-9,7
2	LII	SL	NS	FS	S	-	LII_3	0,35 (0,84)	Q (7,3)	3,61 (0,31)	0,52 (0,95)	25,5	28,3	8,8	-9,0
RFB												15,6	22,6	5,7	2,5
3	LLR	SL	FS	TS	S	I200305 L200408	-	0,07 (0,96)	Q (8,6)	10,35 (0,11)	1,83 (0,10)	104,5	123,8	15,4	-11,6
4(*3)	LLR	SL	FS	TS	S	I200305	LIVV_2	1,33 (0,51)	Q (8,4)	3,92 (0,42)	1,84 (0,10)	67,9 78,3	83,5 92,4	10,5 11,9	-1,6 -4,8
RFB												132,6	152,2	19,3	-17,3
5	LCF	SL	FS	FS	S	-	-	0,22 (0,90)	Q (7,6)	3,42 (0,75)	1,25 (0,30)	54,3	69,7	4,6	-3,2
6	LCF	SL	SL	FS	S	L200407	-	0,23 (0,89)	Q (8,6)	6,09 (0,41)	1,20 (0,34)	44,1	57,1	3,8	-1,6
7	LCF	SL	SL	FS	S	L200407	LCF_1	0,25 (0,88)	Q (8,5)	5,20 (0,39)	1,16 (0,36)	43,8	55,2	3,8	-1,2
RFB												66,3	85,0	5,6	-4,5

Fonte: elaboração do autor.

Notas (*1)N = Nível: SL (estocástico); FL (fixo); NL (ausente).

T = tendência ou inclinação: SS (estocástica); FS (fixa); NS (ausente).

S = componente sazonal: DS (*dummy*); TS (trigonométrico); FS (fixo); NS (ausente).

I = componente irregular: S (sim); N(não).

Outlier: identificação do tipo do outlier, I para tipo aditivo (AO) e L para alteração no nível (LS), seguido do ano e mês de ocorrência.

Variável X_N: identificação da variável explicativa com respectiva defasagem.

(*2)N: teste de normalidade dos resíduos de Doornik e Hansen (1994). Segue distribuição $\chi^2(2)$.

Q(P,d): estatística Q de Box-Ljung com base nas primeiras P autocorrelações dos resíduos, distribuída aproximadamente como $\chi^2(d)$, em que $d = P + 1$ - número de parâmetros do modelo. Estatística ajustada pelo número de parâmetros do modelo.

H: teste de heteroscedasticidade não paramétrico com distribuição $F(h,h)$, em que h é o inteiro mais próximo de T/3.

(*3) Trata-se de modelo EMX, em que o primeiro resultado dos erros de previsão utiliza os dados reais da variável exógena LIVV_2 para o ano de 2007, enquanto que no segundo resultado são utilizados os dados previstos desta mesma variável, conforme detalhado na seção 3.

6.4.1 Imposto de Importação – II

Para a série do LII os resultados dos modelos estruturais não foram satisfatórios. Os modelos foram estimados para o período julho de 2000 a dezembro de 2006. O modelo n° 1 apresenta erros de previsão bastante superiores aos dos demais modelos tratados neste trabalho. Foram detectados outliers mais significativos, do tipo LS, nos meses de dezembro

de 2001, maio de 2003 e março de 2004, que quando incorporados aos modelos, resultaram em previsões menos acuradas. O modelo nº 2, com a incorporação da terceira defasagem de LII como variável explicativa, praticamente não melhora o poder preditivo em relação ao modelo nº 1, com alguma melhora nos critérios de ajustamento AIC e BIC e diminuição da autocorrelação dos resíduos.

6.4.2 Imposto sobre a Renda da Pessoa Jurídica – IRPJ_DOLR

Quanto à série IRPJ_DOLR, cujos modelos foram estimados para o período fevereiro de 2001 a dezembro de 2006, os *outliers* mais significativos, já incorporados ao modelo nº 3, em função da não normalidade dos resíduos do modelo básico sem eles, são do tipo AO no mês de maio de 2003 e LS no mês de agosto de 2004. Este modelo já apresenta resultado superior ao do método dos indicadores e similar ao ARIMA. O modelo nº 4, com a incorporação da segunda defasagem de LIVV como variável explicativa³³ e somente o *outlier* do tipo AO de maio de 2003, melhora bastante o resultado, com RMSE 45% inferior ao do método dos indicadores e da mesma ordem de grandeza do RMSE do modelo dinâmico univariado, que é o melhor resultado para esta série. A inclusão de LPIB como variável exógena não mostrou bons resultados.

6.4.3 Cofins – COFINS_DE

O modelo básico, nº 5, sem *outliers* ou variáveis exógenas, já apresenta resultado do erro de previsão RMSE 18% inferior ao do método dos indicadores. O modelo nº 6, com a inclusão do *outlier* do tipo LS no mês de julho de 2004, melhora sensivelmente o poder preditivo em relação ao modelo básico nº 5. O modelo nº 7, em que se acrescenta a primeira defasagem de LCF, introduz pequena melhora ao modelo anterior, com RMSE 35% inferior ao do método dos indicadores e pouco superior ao do modelo de equações simultâneas nº 6 (DLCF/DLIVV) da tabela 23, que apresenta o melhor poder preditivo entre todas as modelagens. A inclusão das variáveis exógenas, LPIB ou LIVV, não conduz a resultados satisfatórios.

³³ Embora o resultado do teste de exogeneidade das variáveis explicativas, apresentado em 4.7, indique que IVV deve ser tratada como endógena, pelo fato de não haver relação de cointegração entre LII e LIVV (ver tabela 19), foi estimado este modelo EMX com LIVV como variável exógena.

7 Conclusões

O objetivo geral do presente trabalho foi apresentar métodos alternativos ao método dos indicadores utilizado pela RFB, para a previsão da arrecadação tributária federal, por meio de metodologias de séries temporais. Foram utilizadas as metodologias ARIMA, modelos dinâmicos univariados e de Função de Transferência, modelos dinâmicos de múltiplas equações e modelos Estruturais. As séries de arrecadação analisadas se referem ao Imposto de Importação (II), Imposto sobre a Renda da Pessoa Jurídica – Demais Empresas Obrigadas à Apuração do Lucro Real (IRPJ_DOLR) e Contribuição para o Financiamento da Seguridade Social – Demais Empresas (COFINS_DE). A escolha dos tributos analisados levou em conta a participação de cada um na arrecadação total da SRRF/8ªRF para o ano de 2007, bem como o âmbito de incidência do tributo, que, no presente estudo é, respectivamente, o setor externo, o lucro e o faturamento das empresas. Importante ressaltar que a 8ª RF é responsável por aproximadamente 43% da arrecadação tributária federal total do país, justificando-se, assim, a escolha do estado de São Paulo para a realização do presente trabalho.

Foram estimados diversos modelos em cada uma das metodologias propostas. A tabela 29 a seguir apresenta o resultado da comparação dos erros de previsão entre si e com o método dos indicadores, utilizado pela Secretaria da Receita Federal do Brasil para previsão anual da arrecadação tributária. A medida de acurácia da previsão se faz por meio da utilização da raiz do erro quadrático médio da previsão (RMSE) do melhor modelo de cada metodologia, para o ano de 2007. São indicados, também, os erros anuais de previsão, devido à importância em relação ao orçamento anual, bem como em relação às metas de arrecadação a serem cumpridas pela RFB.

Tabela 29 - Comparação do RMSE e do erro anual de previsão do melhor modelo de cada metodologia

Série	MI/RFB		ARIMA		DR/TFM		VEC/SEM		EM/EMX	
	RMSE E_anual	Mod. (tab)	RMSE E_anual	Mod. (tab)	RMSE E_anual	Mod. (tab)	RMSE E_anual	Mod. (tab)	RMSE E_anual	Mod. (tab)
II	22,6 +2,5%	5 (8)	19,3 +2,4%	1 (11)	31,2 -9,4%	9 (12)	14,2 +0,7%	2 (28)	28,3 -9,0%	
IRPJ_DOLR	152,2 -17,3%	6 (9)	123,0 -13,6%	2 (11)	82,7 -5,0%	9 (18)	91,9 -3,3%	4 (28)	83,5 -1,6%	
COFINS_DE	85,0 -4,5%	3 (10)	89,4 -5,5%	5 (11)	67,0 -1,0%	6 (23)	48,5 +0,4%	7 (28)	55,2 -1,2%	

Fonte: elaboração do autor.

Para a série do Imposto de Importação, o modelo nº 9 da tabela 12, modelo dinâmico multivariado VEC/SEM, formado por DLII/DLIPI/DTCR, apresentou o melhor resultado, com RMSE 37% inferior ao do método dos indicadores e 26% inferior ao do ARIMA, além do erro de previsão anual ser de apenas 0,7%, frente a 2,5% apresentado pelo método dos indicadores. Em seguida aparece o modelo nº 5 da tabela 8, ARIMA, com RMSE cerca de 15% inferior ao do método dos indicadores. As modelagens dinâmica univariada e Estrutural não apresentaram resultados satisfatórios, sem conseguir alcançar os erros de previsão obtidos pela RFB.

Para o Imposto sobre a Renda da Pessoa Jurídica todas as metodologias se mostraram superiores ao método dos indicadores. O modelo ARIMA nº 6 da tabela 9 já apresenta melhora em relação ao modelo da RFB, com RMSE quase 20% inferior. As demais metodologias apresentam resultados semelhantes, com sensível melhora tanto no RMSE quanto no erro anual de previsão. O modelo de equações simultâneas nº 9 da tabela 18, formado por DLLR/DLIPI, apresenta RMSE 40% inferior ao do método dos indicadores. Já os modelos nº 4 da tabela 28, Estrutural com LIVV como variável exógena, e nº 2 da tabela 11, dinâmico univariado, exibem resultados semelhantes, com RMSE 46%, 33% e 10% inferiores aos do método dos indicadores, dos modelos ARIMA e equações simultâneas, respectivamente. Ressalte-se que o modelo Estrutural apresenta um erro de previsão anual de apenas 1,6%, frente aos 17,3% apresentado pelo método dos indicadores, ou seja, redução de 91%.

Finalmente, para a série da Cofins, o modelo ARIMA nº 3 da tabela 10 apresenta RMSE aproximadamente igual ao do método dos indicadores, porém sem superá-lo. Em seguida aparece o modelo nº 5 da tabela 11, Função de Transferência com DLIVV como variável exógena, com RMSE 21% inferior. O modelo estrutural nº 7 da tabela 28 melhora ainda mais o resultado, com o RMSE 35% inferior ao do método dos indicadores. Como melhor modelo aparece o de nº 6 da tabela 23, modelo de equações simultâneas com DLCF/DLIVV, com RMSE 43%, 46%, 28% e 12% inferior ao do método dos indicadores e dos modelos ARIMA, Função de Transferência e Estrutural, respectivamente. Ademais, o erro anual de previsão é de apenas 0,4%, ou seja, 91% inferior ao erro obtido pelo método dos indicadores.

Quanto à utilização das variáveis explicativas, tanto na modelagem de múltiplas equações, quanto nos modelos de Função de Transferência, chega-se à conclusão que, para o Imposto de Importação, nos modelos VAR/VEC/SEM, com as variáveis explicativas Índice de Vendas no Varejo ou Índice de Produção Industrial e Taxa de Câmbio Real, os resultados

foram muito bons e semelhantes. A utilização do PIB nos modelos também oferece bons resultados, embora um pouco inferiores aos demais. Os resultados dos testes de exogeneidade das variáveis explicativas não permitem considerar a Taxa de Câmbio Real, o Produto Interno Bruto, o Índice de Vendas no Varejo ou o Índice de Produção Industrial Geral como variáveis exógenas fortes e, assim, estas variáveis não são adequadas às metodologias de Função de Transferência e Estrutural com variáveis exógenas. O pré-requisito de validade destes modelos é o de que as variáveis explicativas possuam exogeneidade forte, ou seja exogeneidade fraca juntamente com não causalidade de Granger no sentido da variável dependente para a variável explicativa. Todas as combinações de variáveis utilizadas, LII e LTCCR com LPIB, LIVV ou LIPI, apresentaram 2 relações de cointegração, o que indica relações de longo prazo entre estas variáveis.

Ainda com relação ao Imposto de Importação, de maneira geral, a modelagem de múltiplas equações (VAR/VEC/SEM) apresentou desempenho superior às modelagens ARIMA, dinâmica univariada e Estrutural. Importante destacar a sensível melhora nas previsões dos modelos com a inclusão das variáveis explicativas, principalmente em relação aos modelos dinâmico univariado e Estrutural.

No caso do IRPJ, o Índice de Produção Industrial parece ser a melhor variável a ser utilizada na modelagem multivariada, seguido do PIB. IRPJ/IVV foi a única combinação de variáveis que não apresentou relação de cointegração e o VAR em diferenças estimado para esta relação não apresentou bons resultados. As demais combinações, IRPJ/PIB e IRPJ/IPI, mostram relação de longo prazo entre as variáveis. O modelo Estrutural, com a utilização do Índice de Vendas no Varejo como variável exógena apresentou resultado satisfatório em relação ao método dos indicadores e à modelagem ARIMA. Porém, a modelagem univariada superou todas as demais. Isto pode ter acontecido devido ao fato de que o IRPJ depende do lucro real apurado pelas empresas, que pode não ter relação direta com o nível da atividade econômica, representado pelas variáveis PIB, IVV e IPI.

Para a Cofins, Índice de Vendas no Varejo e PIB parecem ser as melhores alternativas tanto para os modelos de múltiplas equações, quanto para modelos de Função de Transferência. Como a Cofins depende do faturamento das empresas, parece intuitivo que estas duas variáveis tenham apresentado os melhores resultados. O Índice de Produção Industrial, por sua vez, mede a produção industrial, que pode não ter relação direta com o faturamento das empresas. Todas as combinações de variáveis, Cofins/PIB, Cofins/IVV e Cofins/IPI, mostram relação de longo prazo.

No que diz respeito aos erros de previsão, verifica-se na tabela 1 que há uma tendência de diminuição dos erros de previsão cometidos pelo método dos indicadores, medidos pelo MAPE, ao longo dos anos, provavelmente devido a maior estabilidade econômica dos últimos anos. Assim, merece atenção o fato de que os erros médios do II, IRPJ_DE e COFINS_DE são maiores que os erros cometidos no ano de 2007, base de comparação, especialmente no caso do II, em que a média do MAPE no período 2001 a 2007 é de 18,6%, enquanto que o erro de 2007 é de apenas 5,7%. Ou seja, deve-se levar em consideração que os valores previstos pelos métodos desenvolvidos neste trabalho estão sendo comparados com ótimas previsões efetuadas pelo método dos indicadores para o ano de 2007, o que reforça ainda mais os resultados alcançados.

Ainda com relação ao desempenho do método dos indicadores, quanto mais estável o ambiente econômico, melhor será sua performance, que é totalmente dependente da arrecadação do mesmo período do ano anterior, além de uma boa previsão dos indicadores a serem utilizados. Por outro lado, metodologias de séries temporais geralmente trabalham com uma história maior dos dados e, quanto mais informação é acrescentada à série, com o decorrer dos anos, e desde que adequadamente tratada, mais robustos e melhores se tornam os resultados.

Outro ponto que merece consideração diz respeito ao tratamento e identificação das arrecadações atípicas (*outliers*). Conforme enfatizado por Tsay (1986), a existência de *outliers* pode causar sérios problemas de viés nas estimativas dos parâmetros auto-regressivos e de médias móveis. Uma identificação inicial dos *outliers* e quebras estruturais foi efetuada por meio da utilização dos resíduos auxiliares obtidos da estimação de modelos estruturais. Porém, os *outliers* identificados não foram incorporados aos modelos ARIMA e esta pode ter sido uma das causas da não obtenção de um melhor resultado destes modelos. Importante salientar que, mesmo sem um tratamento mais profundo, os modelos ARIMA superam ou igualam o método dos indicadores.

Por fim, merece destaque o fato de que as séries de arrecadação utilizadas no trabalho referem-se à arrecadação divulgada pela RFB, que é a base de comparação utilizada pelo órgão para se apurar a acurácia da previsão e este foi o motivo da escolha desta série, ou seja, ter uma base de comparação com a previsão efetuada pela RFB. Entretanto, esta arrecadação se compõe de muitas rubricas, tais como retificações, restituições, compensações, Refis, Paes etc. Muitas destas rubricas podem ser consideradas atípicas, com é o caso, por exemplo, do Refis e Paes, que são programas de refinanciamento de dívidas antigas, e restituições e

compensações³⁴. O fato de se usar a arrecadação divulgada pode ter contribuído para o surgimento de *outliers* nas séries. Na série do II, este fato parece não ter muita relevância, com as duas arrecadações sendo bastante próximas. Contudo, para as séries IRPJ_DOLR e COFINS_DE, há meses em que a diferença chega à 10%, o que é significativo e pode alterar o resultado. Assim, uma opção de continuação do trabalho, que me parece apropriada, é a utilização da série da arrecadação bruta classificada. Esta arrecadação é o somatório da arrecadação oriunda dos Documentos de Arrecadação das Receitas Federais (DARF), após a classificação por um sistema interno, que, no caso de pagamentos unificados, como é o caso, por exemplo, do Simples, retenção de órgãos públicos e dívida ativa da União, “explode” o código de recolhimento em seus diversos componentes³⁵. Portanto, esta série está livre de várias fontes de *outliers*. Ademais, para as séries IRPJ_DE e COFINS_DE, o início da série da arrecadação divulgada, constante do sistema *Data Warehouse* Arrecadação, ocorre em agosto de 2000, enquanto que para a arrecadação bruta classificada, o início se dá em janeiro do mesmo ano, o que proporciona um maior número de observações.

Como conclusão final, a aplicação dos métodos desenvolvidos neste trabalho levou à redução média do RMSE de 42% em relação ao erro cometido pelo método dos indicadores e de 35% em relação à modelagem ARIMA, além da drástica redução do erro anual de previsão. A técnica foi aplicada a três tributos que são representativos do sistema tributário e, a partir de agora, seria interessante aplicá-la aos demais tributos administrados pela RFB. Como toda técnica, deve ser aprimorada constantemente, em especial, com relação à tentativa de escolha de outras variáveis explicativas a serem incluídas nos sistemas multivariados.

Esta dissertação mostrou que a utilização de metodologias de séries temporais para a previsão da arrecadação de receitas federais é uma alternativa ao método dos indicadores atualmente utilizado pela Receita Federal do Brasil, contribuindo para previsões mais precisas, tornando-se ferramenta segura de apoio para a tomada de decisões dos gestores.

³⁴ Restituições são consideradas arrecadações negativas, enquanto que a compensação entre tributos de diferentes espécies implica na arrecadação positiva para o tributo devido e arrecadação negativa para o tributo com crédito.

³⁵ Por exemplo, o Simples recolhido no código 6106, é explodido nos códigos 7104 – IRPJ, 7606 – Cofins, 7200 – Pis e 7307 – CSLL.

REFERÊNCIAS

- ANDREWS, R.L. Forecasting Performance of Structural Time Series Models. **Journal of Business & Economic Statistics**, Alexandria (USA), v. 12, n. 1, p. 129-33, Jan. 1994.
- BLANCHARD, O. **Macroeconomics**. 4. ed. New Jersey: Prentice Hall, 2005.
- BARRO, R.J. **Macroeconomics**. 4. ed. New York: John Wiley & Sons. Inc., 1993.
- BOX, G.E.P.; JENKINS, G.M.; REINSEL, G.C. **Time Series Analysis: Forecasting and Control**. 4. ed. New Jersey: Wiley, 2008.
- BROCKWELL, P.J.; DAVIS, R.A. **Time Series: Theory and Methods**. 2. ed. New York: Springer, 1991.
- BROCKWELL, P.J.; DAVIS, R.A. **Introduction to Time Series and Forecasting**. 2. ed. New York: Springer, 2002.
- CARLIN, W.; SOSKICE, D. **Macroeconomics. Imperfections, Institutions et Policies**. Oxford: Oxford University Press, 2006.
- COVEY, T.; BESSLER, D.A. Testing for Granger's Full Causality. **The Review of Economics and Statistics**, Cambridge (USA), v. 74, n. 1, p. 146-153, Feb. 1992.
- DOORNIK, J.A. Testing Vector Error Autocorrelation and Heteroscedasticity in Dynamic Models. **Nuffield College**, Oxford OX1 1NF, UK, Sept. 1996.
- DOORNIK, J.A.; HENDRY, D.F. **Modelling Dynamic Systems Using PcGive 10 – Volume II**. London: Timberlake Consultants Ltd, 2001.
- ENDERS, W. **Applied Econometric Time Series**. 2. ed. USA: Wiley, 2004.
- GRANGER, C.W.J.; NEWBOLD, P. **Forecasting Economic Time Series**. Orlando: Academic Press, 1977.
- GUARAGNA, P.R.S.; MELLO, M.R. **Um Modelo de Previsão para a Arrecadação do ICMS**. Brasília: ESAF, 2002. 46 p. Monografia premiada em 2º lugar no VII Prêmio Tesouro Nacional - 2002, Tributação, Orçamentos e Sistemas de Informação sobre Administração Financeira Pública. out. 2002. Disponível em: <http://www.tesouro.fazenda.gov.br/Premio_TN/VIIIPremio/conteudo_mono_pr7_tema3.html>. Acesso em: 15 set. 2007.
- HAMILTON, J.D. **Time Series Analysis**. Princeton: Princeton University Press, 1994.
- HARVEY, A.C. **Forecasting, Structural Time Series Models and the Kalman Filter**. Cambridge: Cambridge University Press, 1989.
- HARVEY, A.C. Diagnostic Checking of Unobserved-Components Time Series Models. **Journal of Business & Economic Statistics**, Alexandria (USA), v. 10, n. 4, p. 377-89, Oct. 1992.
- HENDRY, D.F. **Dynamic Econometrics**. New York: Oxford University Press, 1995.
- KOOPMAN, S.J.; HARVEY, A.C.; DOORNIK, J.A.; SHEPHARD, N. **Stamp – Structural Time Series Analyser, Modeller and Predictor**. 3. ed. London: Timberlake Consultants Ltd, 2000.
- MADDALA, G.S.; KIM, I.M. **Unit Roots, Cointegration, and Structural Change**. 1. ed. Cambridge: Cambridge University Press, 1998.

MAKRIDAKIS, S. *et alli*. The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition. **Journal of Forecasting**. Hoboken (USA), v. 1, n. 2, p. 111-53, Apr./June 1982.

MELO, B.S.V. **Modelo de Previsão para Arrecadação Tributária**. Brasília: ESAF, 2001 97 p. Monografia vencedora em 1º lugar no VI Prêmio Tesouro Nacional – 2001. Orçamentos e Sistemas de Informação sobre a Administração Financeira Pública, Brasília (DF). Disponível em: <http://www.tesouro.fazenda.gov.br/Premio_TN/conteudo_mono_pr6tema3.html>. Acesso em: 15 set. 2007.

NG, S. Testing for Unit Roots in Flow Data Sampled at Different Frequencies. **Economic Letters**. Orlando, v. 47, n. 3/4, p. 237-42, Mar. 1995.

MUSGRAVE, R.A.; MUSGRAVE, P.B. **Finanças Públicas. Teoria e Prática. Public Finance in Theory and Practice**. Tradução de Carlos Alberto Primo Braga. Rio de Janeiro: Campus; São Paulo: Ed. da Universidade de São Paulo, 1980.

PANKRATZ, A. **Forecasting with Dynamic Regression Models**. New York: John Wiley & Sons Inc., 1991.

PECEGUINI, E.E. **Análise Comparativa de Métodos de Previsão Aplicados à Arrecadação do ICMS - Estado de São Paulo**. Brasília : ESAF, 2001. 47 p. Monografia premiada em 2º lugar no VI Prêmio Tesouro Nacional – 2001. Orçamentos e Sistemas de Informação sobre a Administração Financeira Pública, São Paulo (SP). Disponível em <http://www.tesouro.fazenda.gov.br/Premio_TN/conteudo_mono_pr6_tema3.html>. Acesso em: 15 set. 2007.

PINDYCK, R.S.; RUBINFELD, D.L. **Econometria - Modelos & Previsões. Econometric Models and Economic Forecasts**. Tradução da 4. ed. Revisão técnica: Hamilton Massataki Kai. Rio de Janeiro: Campus, 2004.

PINTO, T.R. **Aplicação de Modelos Não Lineares em Negociação Automática no Mercado Acionário Brasileiro**. Rio de Janeiro. 2006. Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica da PUC-Rio. Disponível em <http://www.maxwell.lambda.ele.puc-rio.br/cgi-bin/db2www/PRG_0490.D2W/INPUT?CdLinPrg=pt>. Acesso em: 27 set. 2007.

ROMER, D. **Advanced Macroeconomics**. 3.ed. Boston: Mc Graw-Hill Irwin, 2006.

SANTANA, I.C. **Modelo Econométrico para Previsão do IPI**. Brasília: Esaf, 2004. 23 p. Trabalho final do curso de Introdução à Estatística e Econometria Aplicada ministrado pela ESAF no período de 14/06 a 15/10 de 2004.

SIQUEIRA, M.L. **Melhorando a Previsão de Arrecadação Tributária Federal Através da Utilização de Modelos de Séries Temporais**. Brasília: ESAF, 2002. 84 p. Monografia premiada em 1º lugar no VII Prêmio Tesouro Nacional - 2002, Tributação, Orçamentos e Sistemas de Informação sobre a Administração Financeira Pública. out. 2002. Disponível em <http://www.tesouro.fazenda.gov.br/Premio_TN/VIIIPremio/conteudo_catalogos_pr7.html>. Acesso em: 15 set. 2007.

SIMON, C.P.; BLUME, L. **Matemática Para Economistas**. Tradução: Dr. Claus Ivo Doering. São Paulo: Bookman, 2004.

TSAY, R.S. Time Series Model Specification in the Presence of Outliers. **Journal of the American Statistical Association**, Alexandria (USA), v. 81, n. 393, p. 132-141, Mar. 1986.

TSAY, R.S. Time Series and Forecasting: Brief History and Future Research. **Journal of the American Statistical Association**, Alexandria (USA), v. 95, n. 450, p. 638-643, June 2000.

WHITE, H. A Heteroskedastic-consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. **Econometrica**, Princeton, v. 48, n. 4, p. 817-838, May 1980.

APÊNDICE A - Tabelas

Tabela 30 - Arrecadação Realizada 2007 – São Paulo - em R\$1.000.000,00

Código de Agregação - CA	Tributo		Arrecadação	Participação na Arrec. Total da SRRF (%)
010	I.Importação	*	6.114	3,23%
020	I.Exportação		4	0,00%
060	I.P.I		15.136	7,99%
070	Fumo		824	0,43%
080	Bebidas		799	0,42%
090	Automóveis		2.236	1,18%
110	Outros Produtos		7.378	3,89%
120	Vinculado		3.314	1,75%
125	Demais Receitas		355	0,19%
126	A. Legais		230	0,12%
140	I.R.Total		70.918	37,42%
150	I.R.P.F.		6.093	3,21%
160	Carnê-Leão		3.653	1,93%
170	Quotas		1.682	0,89%
175	Demais Receitas		457	0,24%
176	A.Legais		301	0,16%
190	I.R.P.J.		30.929	16,32%
195	Obrigadas ao Lucro Real		22.522	11,88%
200	Financeiras		8.252	4,35%
210	Demais Obrigadas	*	14.270	7,53%
230	Não Obrig L.Real		6.029	3,18%
240	Incentivos Fiscais		22	0,01%
245	Demais Receitas		1.359	0,72%
246	A.Legais		996	0,53%
260	I.R.R.F.		33.896	17,88%
270	Trabalho		14.804	7,81%
280	Capital		12.499	6,59%
290	Fundos R.Fixa		5.192	2,74%
300	Fundos R.Variável		481	0,25%
310	Títulos R.Fixa		4.237	2,24%
320	Demais Rend Capital		2.589	1,37%
330	Residentes Exterior		4.376	2,31%
340	Outros Rendimentos		1.801	0,95%
345	Demais Receitas		176	0,09%
346	A.Legais		240	0,13%
360	I.O.F.		5.390	2,84%
365	Operações de Crédito		3.666	1,93%
375	Operações de Seguro		1.511	0,80%
385	Outras Operações		180	0,09%

Tabela 30 - Arrecadação Realizada 2007 – São Paulo - em R\$1.000.000,00

Código de Agregação - CA	Tributo	Arrecadação	Participação na Arrec. Total da SRRF (%)
395	Demais Receitas	18	0,01%
396	A.Legais	15	0,01%
420	I.T.R.	113	0,06%
495	C.P.M.F.	23.007	12,14%
500	COFINS	41.203	21,74%
510	Financeiras	2.265	1,20%
520	Demais Empresas	* 25.113	13,25%
521	Importação	7.700	4,06%
525	Demais Receitas	5.102	2,69%
526	A Legais	1.022	0,54%
540	PIS / PASEP	10.617	5,60%
550	PIS	7.880	4,16%
553	Financeiras	508	0,27%
554	Demais Empresas	5.702	3,01%
555	Importação	1.670	0,88%
560	PASEP	1.279	0,67%
565	Demais Receitas	1.049	0,55%
566	A. Legais	409	0,22%
580	C.S.L.L	14.984	7,91%
590	Financeiras	3.521	1,86%
600	Demais Empresas	9.005	4,75%
605	Demais Receitas	1.842	0,97%
606	A. Legais	616	0,32%
612	CIDE	529	0,28%
613	Cide-Combustíveis	150	0,08%
614	Cide-Remessas	379	0,20%
620	FUNDAF	118	0,06%
622	C.Previdenciária	0	0,00%
644	Pagamento Unificado	-46	-0,02%
649	Outras Rec. Adm.	1.445	0,76%
650	REC. ADM. PELA SRF	189.532	100,00%

Fonte: SRRF 8ª RF.

Tabela 31 – Testes de igualdade de variância

Série	Método	df	Valor	Probabilidade
II	Bartlett	6	24,77720	0,0004
	Levene	(6, 77)	4,705234	0,0004
	Brown-Forsythe	(6, 77)	3,567087	0,0036
LII	Bartlett	6	12,04440	0,0610
	Levene	(6, 77)	2,196637	0,0522
	Brown-Forsythe	(6, 77)	1,513009	0,1851
LR	Bartlett	6	6,662193	0,3532
	Levene	(6, 70)	0,752695	0,6094
	Brown-Forsythe	(6, 70)	0,524272	0,7880
LLR	Bartlett	6	5,472622	0,4848
	Levene	(6, 70)	1,289514	0,2734
	Brown-Forsythe	(6, 70)	0,844644	0,5399
CF	Bartlett	6	7,998352	0,2382
	Levene	(6, 70)	2,763809	0,0181
	Brown-Forsythe	(6, 70)	2,552309	0,0271
LCF	Bartlett	6	9,830185	0,1320
	Levene	(6, 70)	3,052032	0,0104
	Brown-Forsythe	(6, 70)	2,924480	0,0133

Fonte: elaboração do autor.

Tabela 32 - Valores dos logaritmos das séries/10⁶ deflacionados pelo IER

Série	jan	fev	mar	abr	mai	jun	jul	ago	set	out	nov	dez	Média	Mediana
2000														
LII	5,80	5,80	5,85	5,80	5,99	5,97	5,96	6,05	5,99	6,09	6,13	6,03	5,95	5,98
LLR	NA	NA	NA	NA	NA	NA	NA	5,52	5,62	5,71	5,36	5,50	5,54	5,52
LCF	NA	NA	NA	NA	NA	NA	NA	6,98	7,05	7,00	7,04	7,09	7,03	7,04
2001														
LII	5,95	5,75	6,11	5,92	6,11	6,01	6,05	6,08	5,86	6,01	5,83	5,58	5,94	5,98
LLR	6,01	5,42	5,60	5,53	5,34	5,29	5,58	5,21	5,28	5,48	5,25	5,43	5,45	5,42
LCF	7,12	6,97	6,93	7,07	7,03	7,10	7,00	7,06	7,06	6,98	7,04	7,11	7,04	7,05
2002														
LII	5,59	5,50	5,55	5,59	5,60	5,61	5,87	5,77	5,78	5,93	5,77	5,66	5,69	5,64
LLR	6,01	5,28	5,81	5,72	5,28	5,25	5,45	5,17	5,36	5,71	5,42	5,36	5,48	5,39
LCF	7,10	6,91	6,90	6,95	6,98	7,00	6,99	6,98	7,08	6,98	7,08	7,02	7,00	6,99
2003														
LII	5,65	5,61	5,62	5,55	5,44	5,37	5,48	5,38	5,50	5,55	5,51	5,52	5,52	5,52
LLR	6,09	5,46	5,77	5,80	6,06	5,32	5,66	5,36	5,47	5,69	5,81	5,69	5,68	5,69
LCF	7,04	6,86	6,90	6,87	6,85	6,90	6,92	6,97	7,03	7,05	7,08	7,02	6,96	6,94
2004														
LII	5,43	5,22	5,65	5,60	5,51	5,68	5,66	5,68	5,68	5,61	5,71	5,58	5,58	5,63
LLR	6,29	5,71	5,85	6,03	5,60	5,67	5,91	6,06	6,12	6,22	6,03	6,13	5,97	6,03
LCF	7,09	7,02	7,05	7,14	7,08	7,02	6,79	6,85	6,87	6,86	6,89	6,90	6,96	6,96
2005														
LII	5,54	5,39	5,58	5,46	5,49	5,51	5,50	5,65	5,47	5,50	5,51	5,50	5,51	5,50
LLR	6,66	6,23	6,26	6,20	6,01	6,07	6,09	6,12	6,08	6,16	5,97	6,11	6,16	6,12
LCF	7,05	6,72	6,75	6,88	6,84	6,90	6,91	6,92	7,02	6,96	6,96	7,03	6,91	6,91
2006														
LII	5,52	5,31	5,46	5,35	5,57	5,52	5,46	5,62	5,53	5,58	5,56	5,55	5,50	5,53
LLR	6,37	6,37	6,54	6,30	6,14	5,98	6,23	6,21	6,42	6,25	6,20	6,11	6,26	6,24
LCF	7,15	6,87	6,90	6,99	6,94	6,97	6,87	7,07	7,09	6,98	7,02	7,06	6,99	6,98
2007														
LII	5,57	5,45	5,65	5,55	5,66	5,55	5,63	5,78	5,67	5,79	5,74	5,61	5,64	5,64
LLR	6,72	6,18	6,81	6,40	6,22	6,36	6,45	6,44	6,36	6,65	6,57	6,51	6,47	6,44
LCF	7,15	6,89	6,85	7,05	7,00	7,06	7,02	7,03	7,11	7,04	7,16	7,22	7,05	7,05

Fonte: elaboração do autor.

Tabela 33 – Detecção inicial de *outliers* – variáveis principais

Série (*1)	Mês/Ano	Tipo de <i>outlier</i> (*2)	Valor do Resíduo Auxiliar (*3)	Estatística (*4)
SeasAdj_Lii = Level (SL+NS) + I	09/2001	LS	-2,2219	0,0145
	11/2001	LS	-2,6268	0,0051
	12/2001	LS	-2,6573	0,0047
	07/2002	LS	2,0910	0,0198
	07/2002	AO	2,1632	0,0167
	03/2003	LS	-2,1223	0,0184
	05/2003	LS	-2,7202	0,0040
	08/2003	AO	-2,1477	0,0173
	02/2004	AO	-2,2805	0,0126
	03/2004	LS	2,5984	0,0055
SeasAdj_Llr = Trend (SL+SS) + I	05/2003	AO	4,2271	0,0000
	06/2003	LS	-2,3028	0,0120
	06/2004	LS	2,1475	0,0175
	07/2004	LS	2,3597	0,0104
	08/2004	LS	2,8123	0,0031
	01/2006	AO	-2,4250	0,0088
	02/2006	LS	2,0115	0,0239
SeasAdj_Lcf = Level (SL+NS) + I	04/2003	LS	-2,0440	0,0222
	02/2004	LS	2,5264	0,0068
	06/2004	LS	-2,5290	0,0067
	07/2004	LS	-3,4958	0,0004
	07/2004	AO	-2,0777	0,0205
	01/2005	AO	2,5842	0,0058
	02/2005	LS	-2,0771	0,0206
	08/2006	LS	2,4313	0,0087
08/2006	AO	2,1896	0,0158	

Fonte: elaboração do autor.

Notas: (*1) Para séries com forte componente sazonal, trata-se da série ajustada sazonalmente.

SL: nível estocástico ; SS: inclinação estocástica ; NS: inclinação ausente ; I: componente irregular.

(*2) AO: *additive outlier* – LS: *level shift*.

(*3) Valor do resíduo auxiliar do componente irregular (AO) ou do componente de nível (LS). Somente estão mostrados os períodos em que o valor (em módulo) foi superior a 2, conforme Harvey and Koopman (1992).

(*4) Estatística de distribuição padrão (não pode ser usada como teste padrão devido à correlação serial presente nos resíduos auxiliares).

Tabela 34 – Detecção inicial de *outliers* – variáveis explicativas

Série (*1)	Mês/Ano	Tipo de <i>outlier</i> (*2)	Valor do Resíduo Auxiliar (*3)	Estatística (*4)
Ltcr = Level (SL+NS) + Expl vars (Ltcr_1,Ltcr_6) + I	11/2001	LS	-2,6094	0,0055
	10/2002	AO	3,1343	0,0012
	11/2002	LS	-2,4420	0,0085
	11/2002	AO	-3,2917	0,0008
	02/2003	AO	2,5205	0,0069
	03/2005	AO	2,4074	0,0092
	04/2005	LS	-2,3789	0,0099
SeasAdj_Lpib = Trend (SL+FS) + I	03/2000	LS	-2,3234	0,0113
	05/2000	LS	2,4861	0,0074
	06/2000	LS	2,5217	0,0068
	06/2000	AO	2,6259	0,0051
	09/2000	AO	-2,0996	0,0194
	06/2001	LS	-3,0621	0,0015
	06/2001	AO	-3,0034	0,0018
	12/2002	LS	-2,1288	0,0181
SeasAdj_Livv = Trend (SL+SS) + I	09/2003	LS	2,1546	0,0170
	02/2000	AO	2,1376	0,0177
	03/2002	AO	2,5930	0,0056
	11/2002	LS	-2,0998	0,0194
	12/2002	LS	-2,9287	0,0022
	01/2004	LS	2,6425	0,0049
	02/2004	LS	2,5052	0,0071
	08/2004	LS	-2,0401	0,0222
SeasAdj_Lipi = Trend (SL+FS) + I	02/2005	AO	-2,1748	0,0162
	01/2000	AO	-2,1465	0,0174
	02/2000	LS	2,1415	0,0176
	02/2000	AO	2,1680	0,0165
	10/2001	LS	-2,0684	0,0208
	12/2001	AO	-2,6386	0,0050
	04/2002	AO	2,4491	0,0082
	10/2002	LS	2,1358	0,0178
	10/2002	AO	2,2615	0,0132
	09/2003	LS	2,5351	0,0065

Fonte: elaboração do autor.

Notas: (*1) Para séries com forte componente sazonal, trata-se da série ajustada sazonalmente.

SL: nível estocástico ; SS: inclinação estocástica ; FS: inclinação fixa ; NS: inclinação ausente ; I: componente irregular.

(*2) AO: *additive outlier* – LS: *level shift*.

(*3) Valor do resíduo auxiliar do componente irregular (AO) ou do componente de nível (LS). Somente estão mostrados os períodos em que o valor (em módulo) foi superior a 2, conforme Harvey and Koopman (1992).

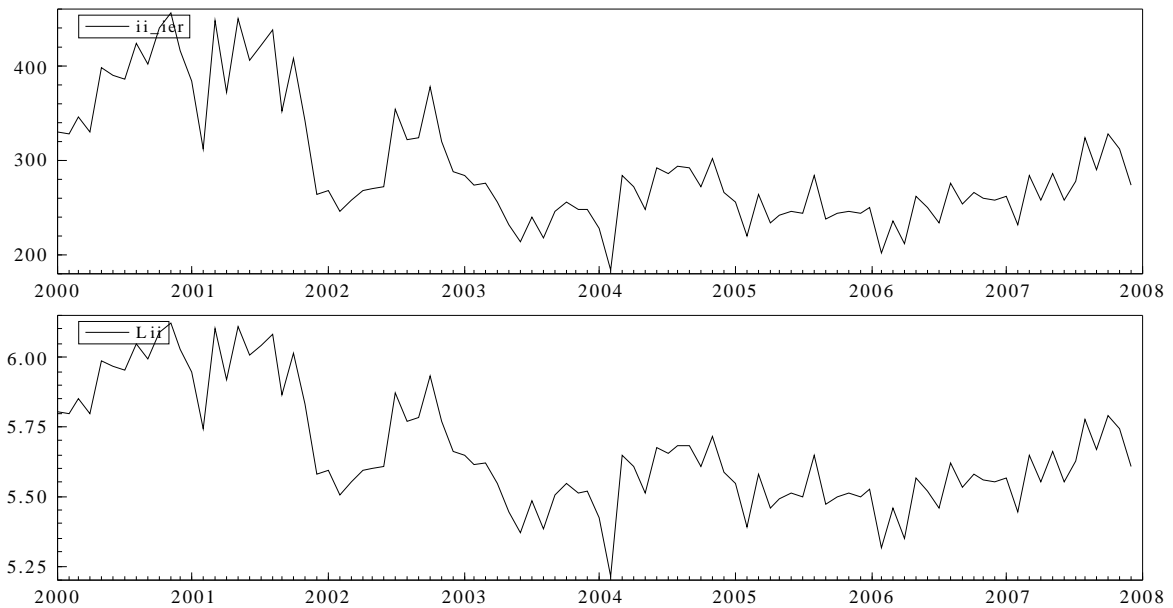
(*4) Estatística de distribuição padrão (não pode ser usada como teste padrão devido à correlação serial presente nos resíduos auxiliares).

Tabela 35 – Principais modelos e componentes estruturais de séries de tempo

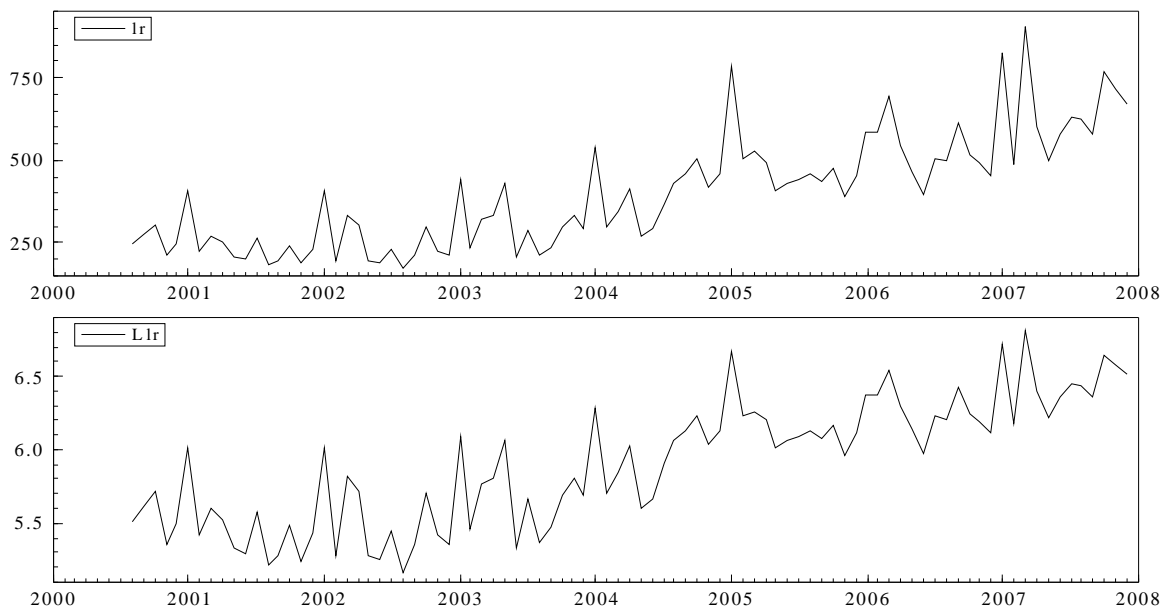
Modelo	Componentes	Especificação	Op. estac. $\Delta(L)$	Forma reduzida	Comentários
	1a <i>Random Walk (RW)</i>	$\mu_t = \mu_{t-1} + \eta_t$	$\Delta = 1 - L$	-	-
	1b <i>RW com intercepto</i>	$\mu_t = \mu_{t-1} + \beta + \eta_t$	Δ	-	-
A Nível local / RW mais ruído	2 tendência estocástica	$y_t = \mu_t + \varepsilon_t$, c/ μ_t como em (1a) $\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t$ $\beta_t = \beta_{t-1} + \zeta_t$	Δ Δ^2	ARIMA(0,1,1) -	Função de previsão é EWMA RW c/ intercepto se $\sigma_\zeta^2 = 0$. RW duplamente integrado se $\sigma_\eta^2 = 0$.
B Tendência linear local	3 ciclo estocástico	$y_t = \mu_t + \varepsilon_t$, c/ μ_t como em (2) $\begin{bmatrix} \Psi_t \\ \Psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos \lambda_c & \sin \lambda_c \\ -\sin \lambda_c & \cos \lambda_c \end{bmatrix} \begin{bmatrix} \Psi_{t-1} \\ \Psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix}$ ψ é o ciclo, $0 \leq \rho < 1$ e $0 \leq \lambda_c < \pi$	1	ARIMA(0,2,2) ARMA(2,1)	Função de previsão Holt-Winters não sazonal Se transforma em AR(1) se $\lambda_c = 0$ ou π
C Ciclo mais ruído		$y_t = \mu + \psi_t + \varepsilon_t$, $0 \leq \rho < 1$	1	cte + ARMA(2,2)	-
D Tendência mais ciclo		$y_t = \mu_t + \psi_t + \varepsilon_t$, c/ μ_t como em (2)	Δ^2	ARIMA(2,2,4)	-
E Tendência cíclica	4 ciclo não estacionário 5a sazonalidade variável <i>dummy</i> 5b sazonalidade trigonométrica	$y_t = \mu_t + \varepsilon_t$, $\mu_t = \mu_{t-1} + \psi_{t-1} + \beta_{t-1} + \eta_t$, c/ β_t como em (2) idem (3) c/ $\rho = 1$ $\gamma_t = \sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t$ $\gamma_t = \sum_{j=1}^{[s/2]} \gamma_{t,j}^*$, $\gamma_{t,j}^*$ é um ciclo não estacionário (4), c/ $\lambda_c = \lambda_j = 2\pi j/s$, $j = 1, 2, \dots, [s/2]$	Δ^2 $1 - 2 \cos \lambda_c L + L^2$ $S(L) = 1 + L + L^2 + \dots + L^{s-1}$ $S(L)$	ARIMA(2,2,4) $(1 - 2 \cos \lambda_c L + L^2)\psi_t \sim \text{MA}(1)$ $S(L)\gamma_t \sim \text{WN}$ $S(L)\gamma_t \sim \text{MA}(s-2)$	Modelo de tend. amortecida se $\lambda_c = 0$ e β_t é removido - - se desenvolve mais suavemente que 5a
F Modelo estrutural básico (BSM)		$y_t = \mu_t + \gamma_t + \varepsilon_t$, c/ μ_t como em (2) e γ_t como em (5a) ou (5b)	$\Delta_s = (1 - L)(1 - L^s)$	$\Delta \Delta_s y_t \sim \text{MA}(s+1)$	(a) Previsões a partir de Holt-Winters são similares (b) Próximo do modelo <i>airline</i> p/ algumas séries

Fonte: Harvey (1989), p.510-11.

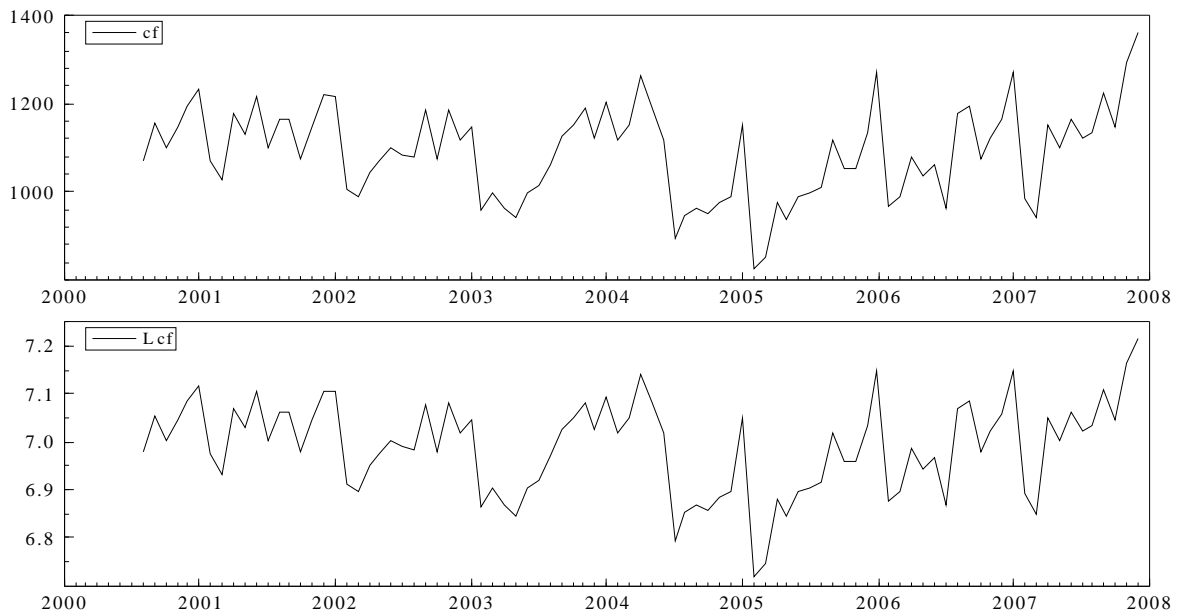
APÊNDICE B – Ilustrações



Gráficos 1 e 2 - Série do Imposto de Importação (em R\$ 10⁶ deflacionado pelo IER e em logaritmo)



Gráficos 3 e 4 - Série do Imposto Sobre a Renda da Pessoa Jurídica – Demais Empresas Obrigadas a Apuração do Lucro Real (em R\$ 10⁶ deflacionado pelo IER e em logaritmo)



Gráficos 5 e 6 - Série da Contribuição para o Financiamento da Seguridade Social (em R\$ 10⁶ deflacionado pelo IER e em logaritmo)

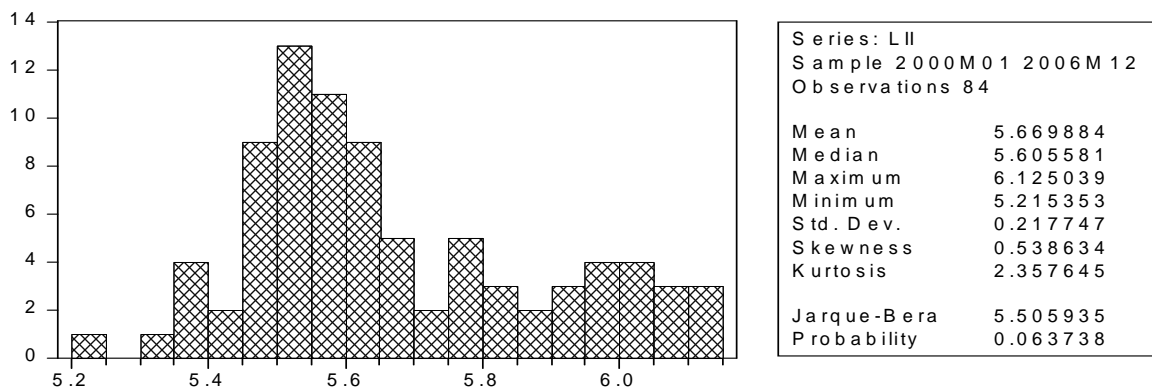


Gráfico 7 - Estatísticas descritivas – Série LII

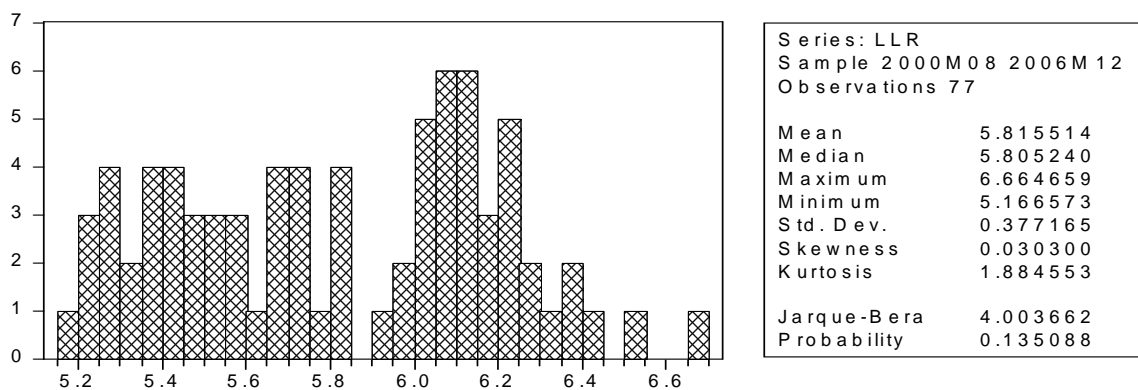


Gráfico 8 - Estatísticas descritivas – Série LLR

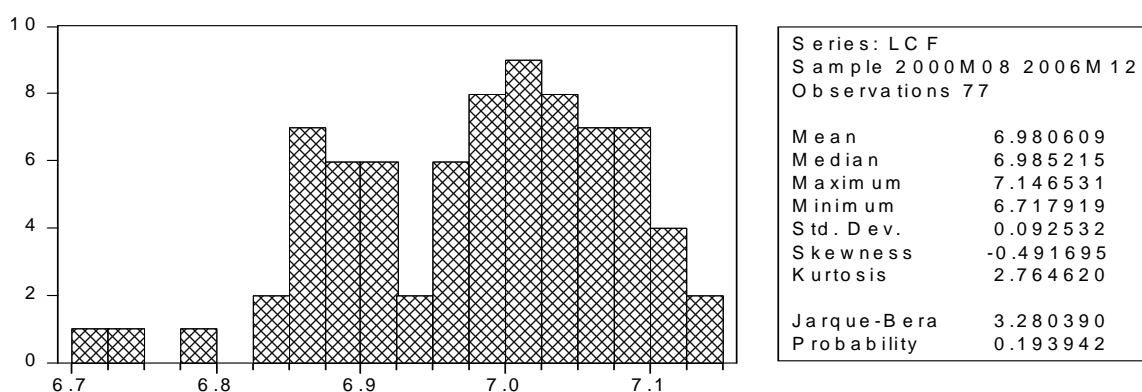


Gráfico 9 - Estatísticas descritivas – Série LCF

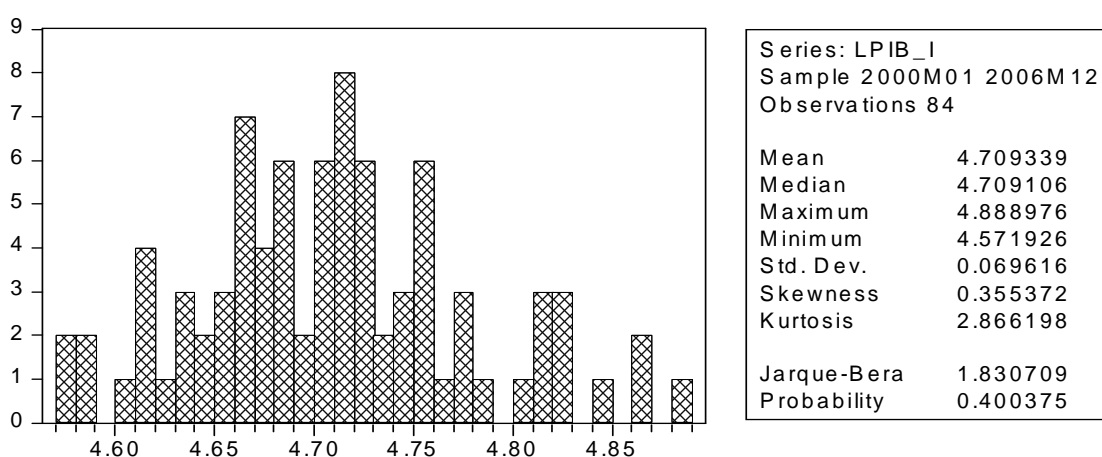


Gráfico 10 - Estatísticas descritivas – Série LPIB

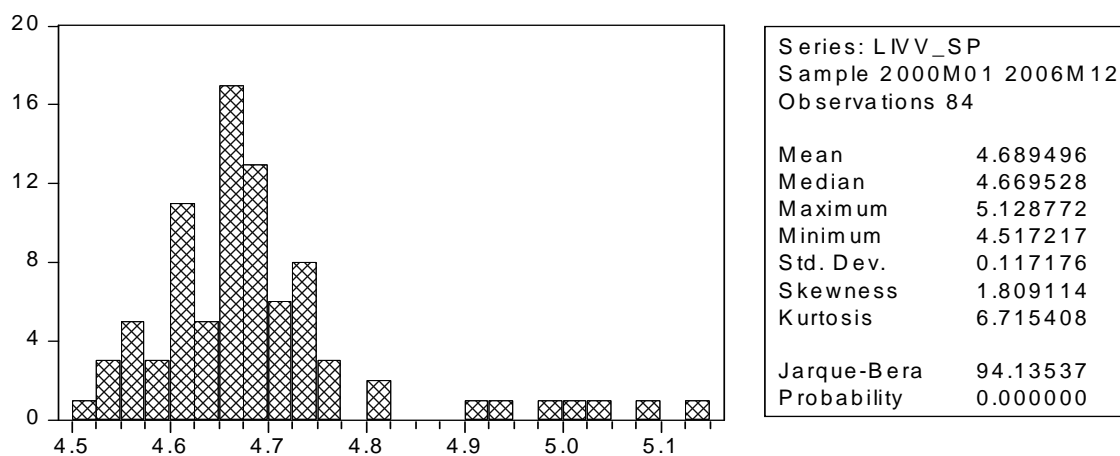


Gráfico 11 - Estatísticas descritivas – Série LIVV

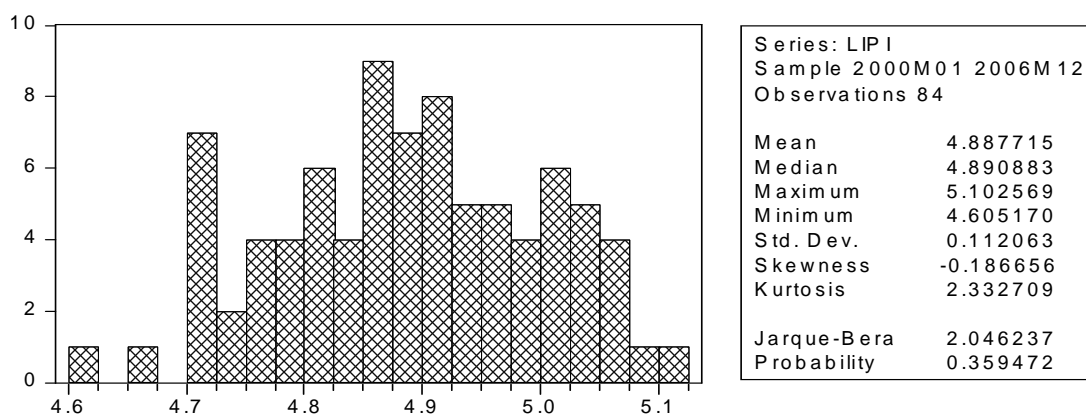


Gráfico 12 - Estatísticas descritivas – Série LIPI

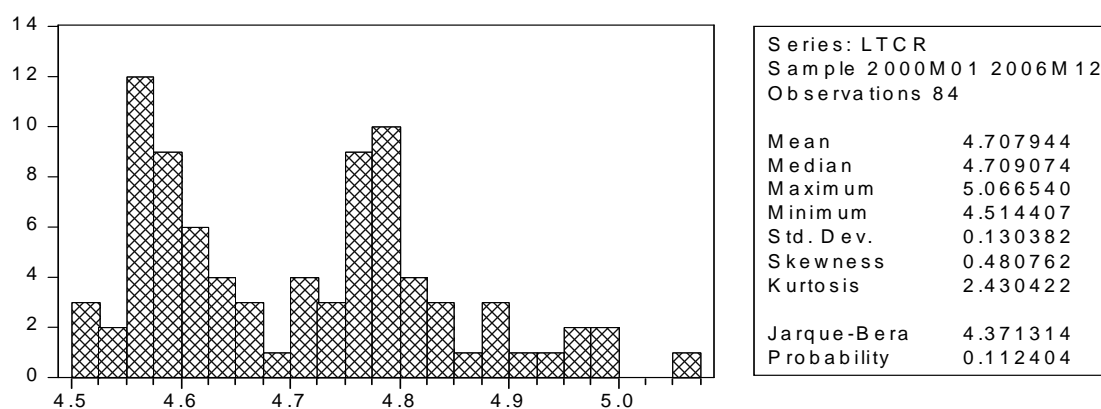
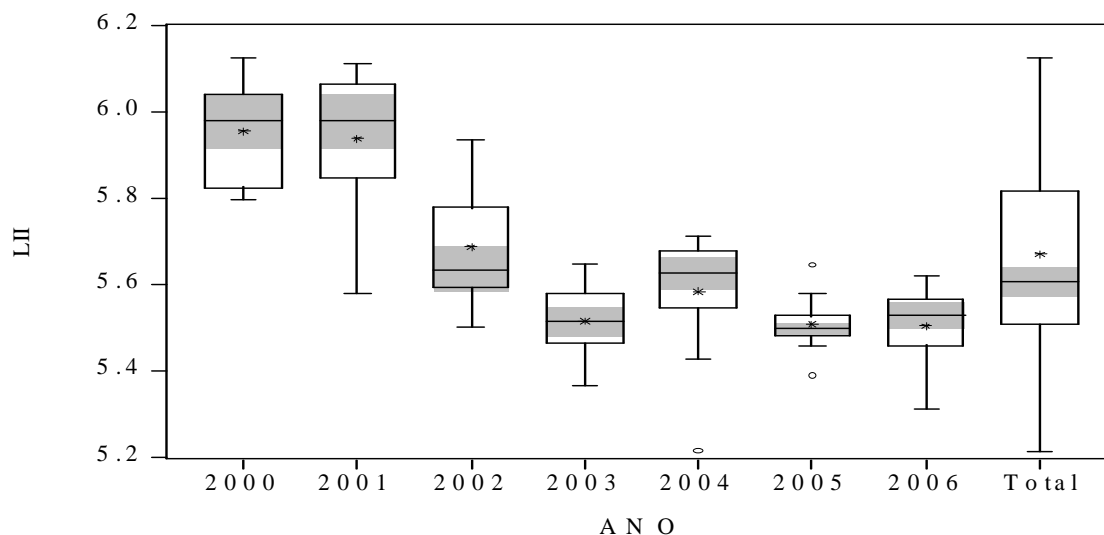


Gráfico 13 - Estatísticas descritivas – Série LTCR

Gráfico 14 – *Boxplots*³⁶ – Série LII

³⁶ *Boxplots* – a caixa representa a faixa entre o primeiro (1Q) e o terceiro (3Q) quartis, cuja diferença é denominada IQR (faixa de interquartil). Observações entre $(1Q - 1,5 IQR)$ e $(1Q - 3 IQR)$ ou entre $(3Q + 1,5 IQR)$ e $(3Q + 3 IQR)$ são consideradas *near outliers* (representadas por um pequeno círculo no gráfico), enquanto que observações abaixo de $(1Q - 3 IQR)$ ou acima de $(3Q + 3 IQR)$ são consideradas *far outliers* (representadas por um pequeno quadrado).

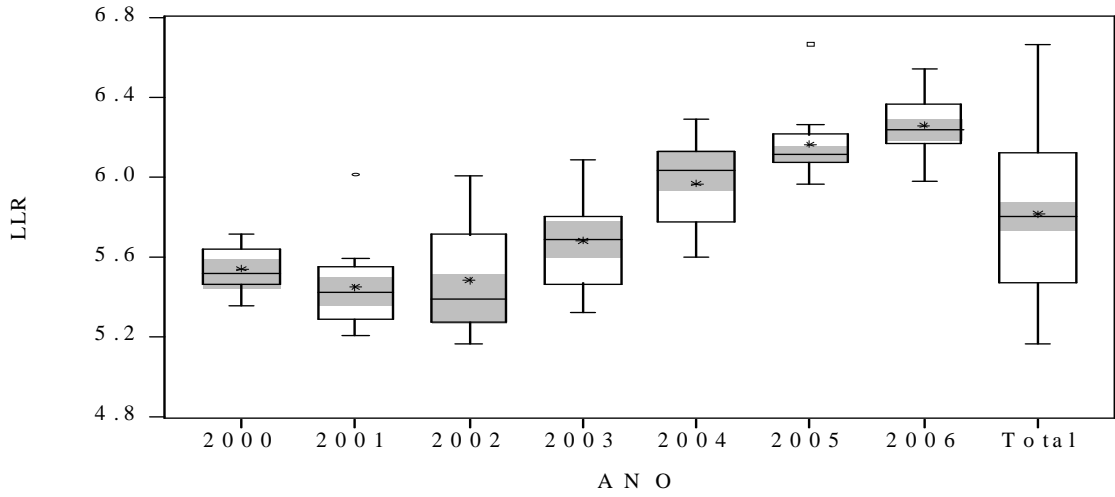


Gráfico 15 – Boxplots – Série LLR

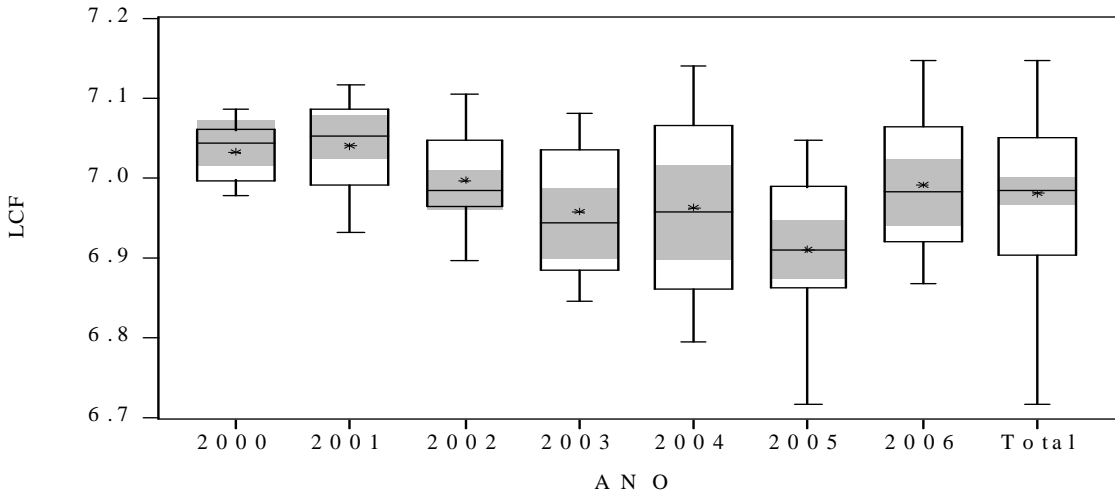


Gráfico 16 – Boxplots – Série LCF

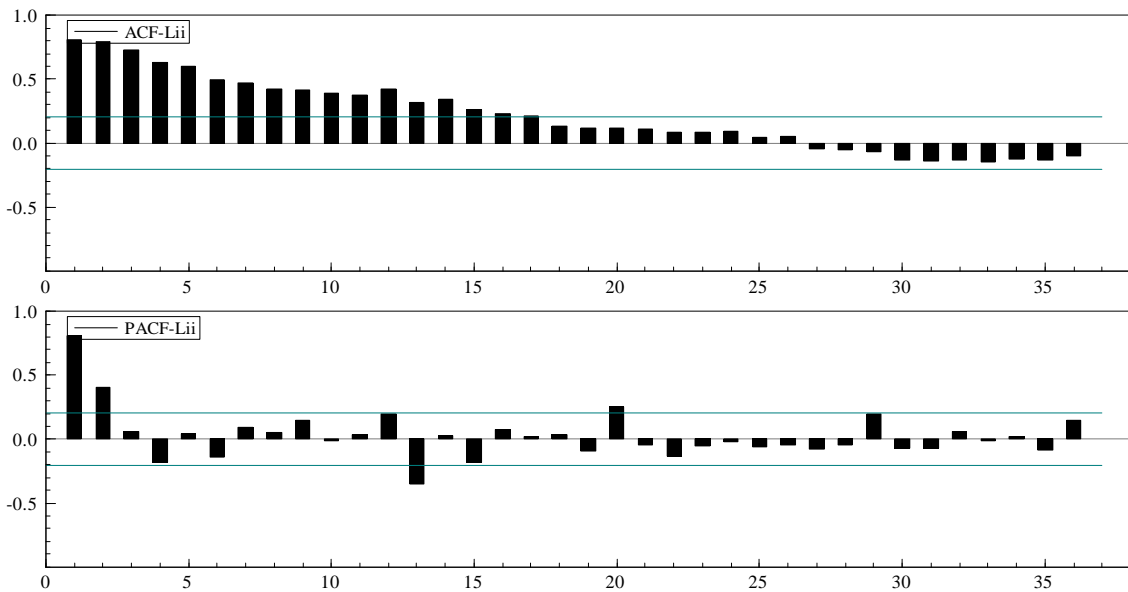


Gráfico 17 – Correlograma em nível da série LII

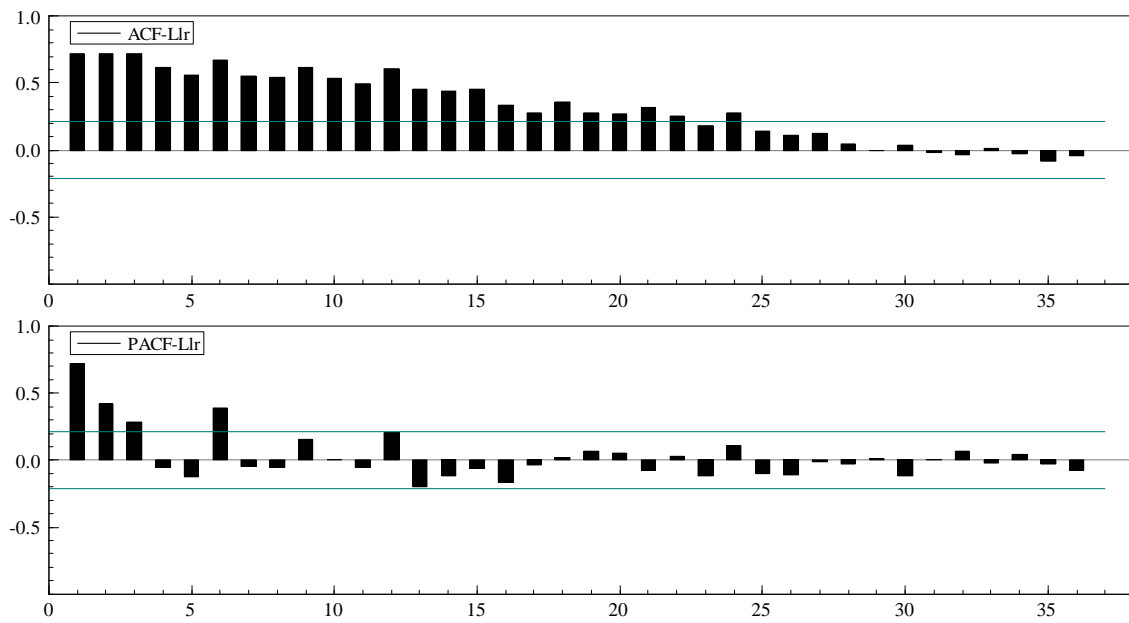


Gráfico 18 – Correlograma em nível da série LLR

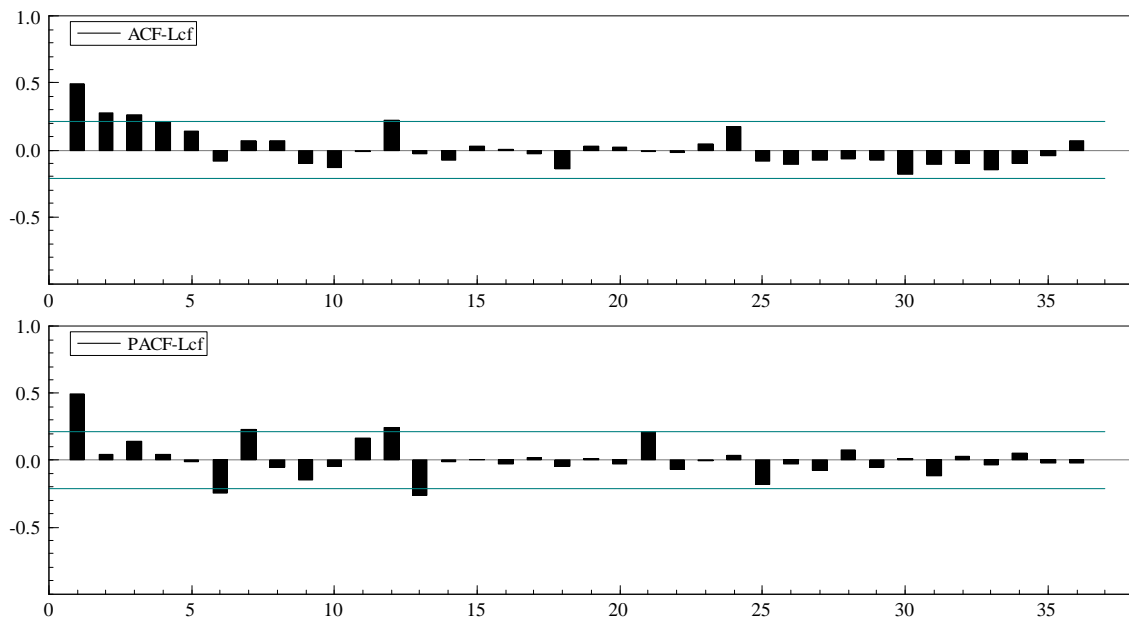


Gráfico 19 – Correlograma em nível da série LCF

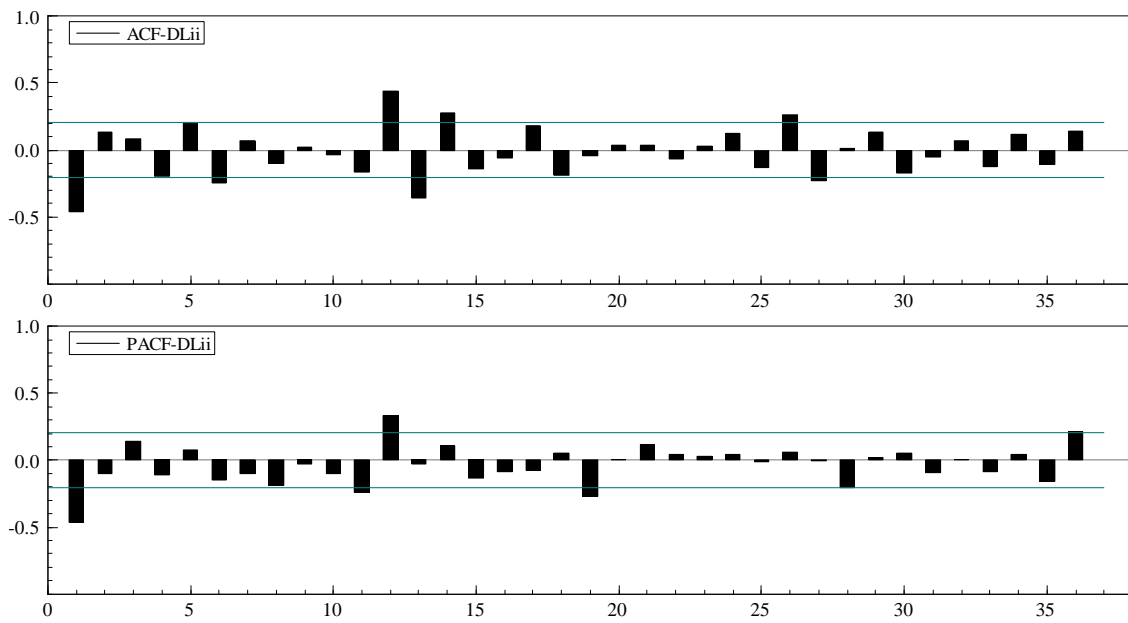


Gráfico 20 – Correlograma em primeira diferença da série LII

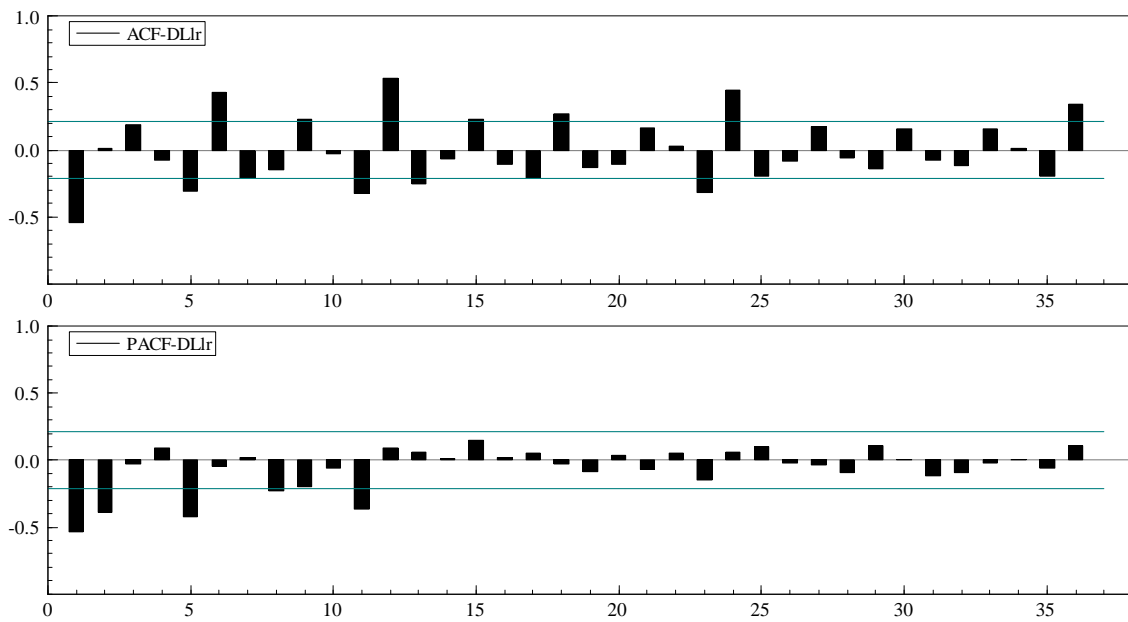


Gráfico 21 – Correlograma em primeira diferença da série LLR

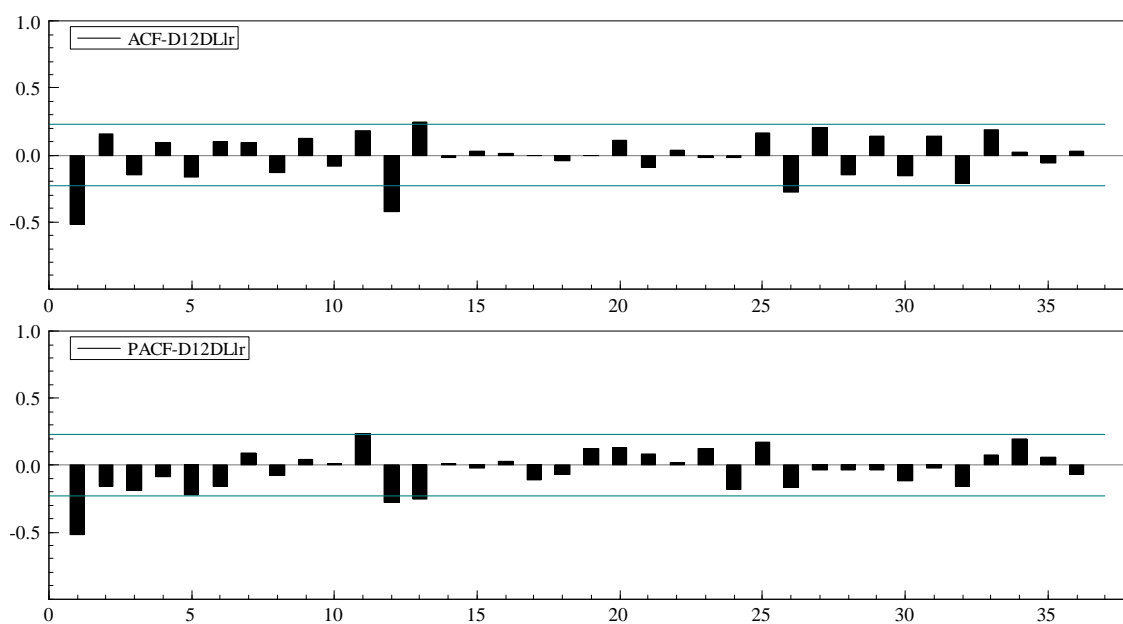


Gráfico 22 – Correlograma em primeira diferença mais primeira diferença sazonal da série LLR

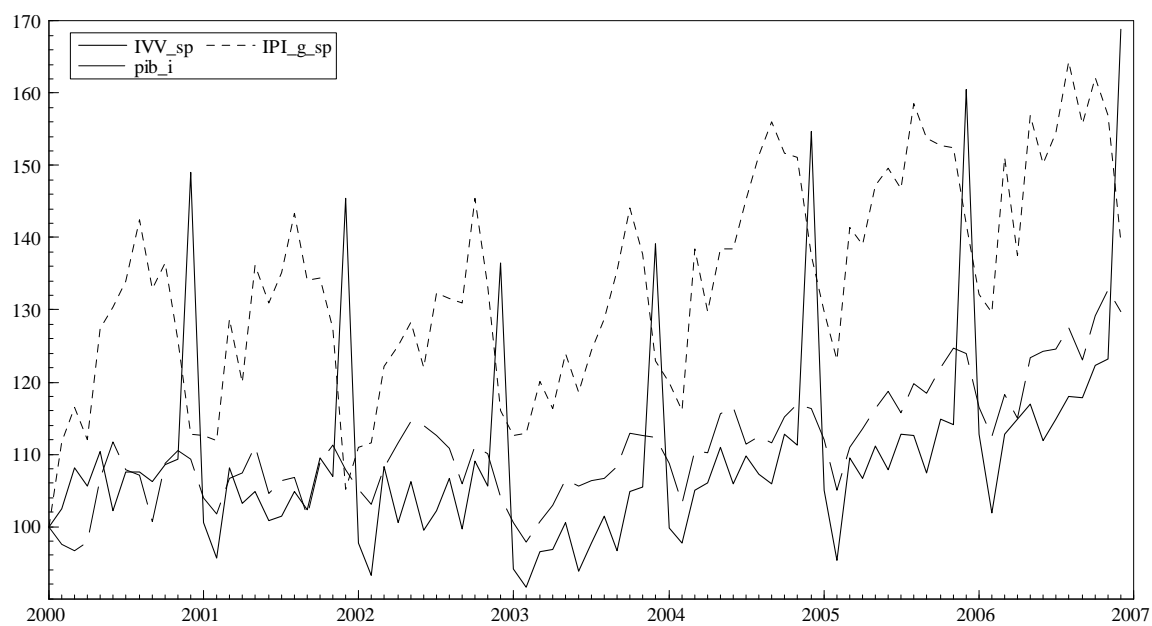


Gráfico 23 – Séries Índice de Vendas no Varejo – São Paulo, Índice de Produção Industrial – Geral – São Paulo e Produto Interno Bruto - Índice

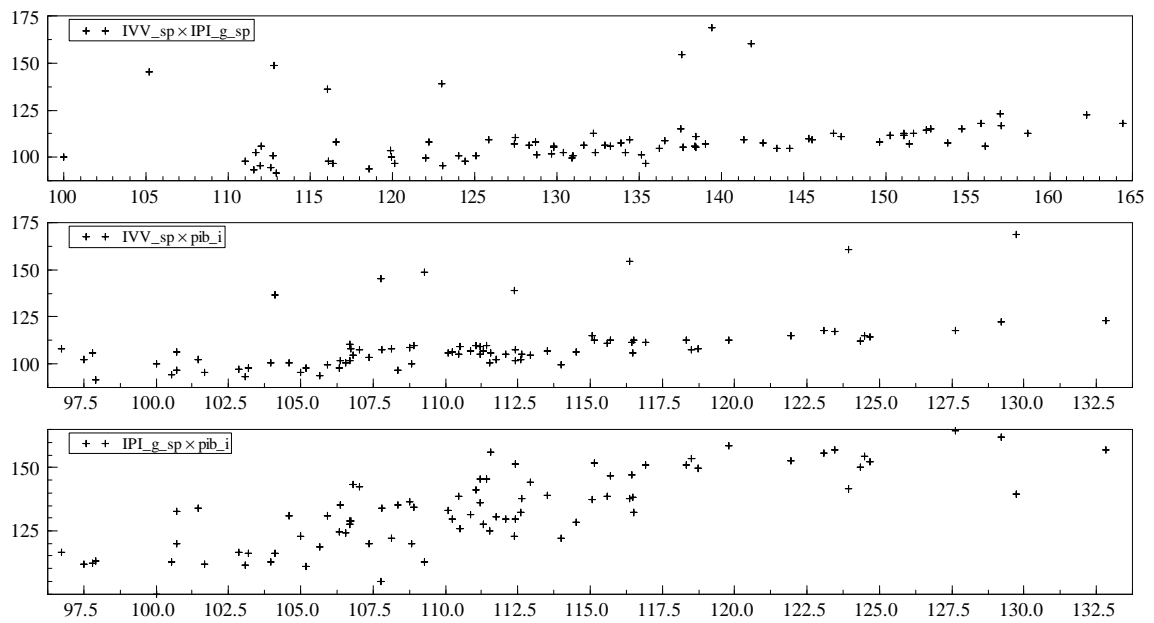


Gráfico 24 – Scatter plot IVV x IPI, IVV x PIB, IPI x PIB_I

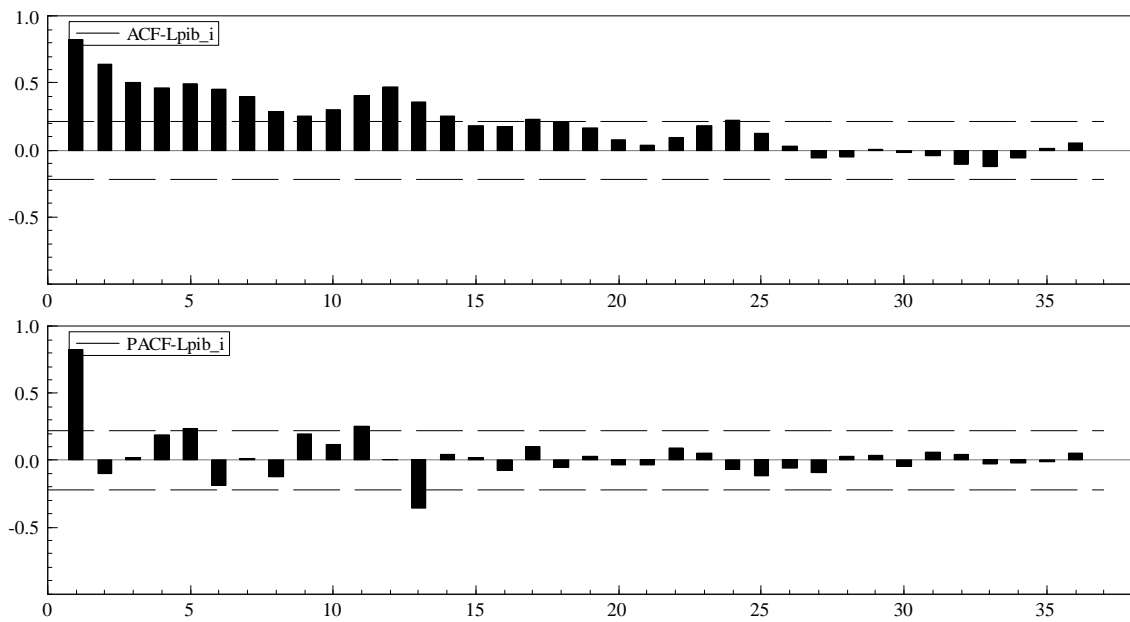


Gráfico 25 – Correlograma da série Lpib

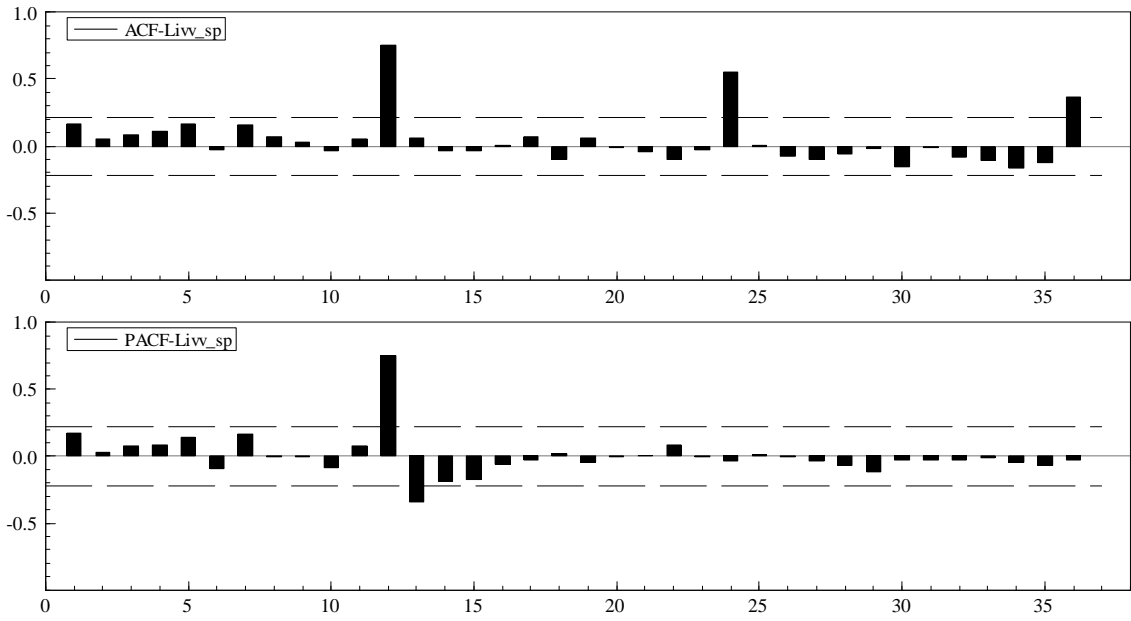


Gráfico 26 – Correlograma da série LIVV

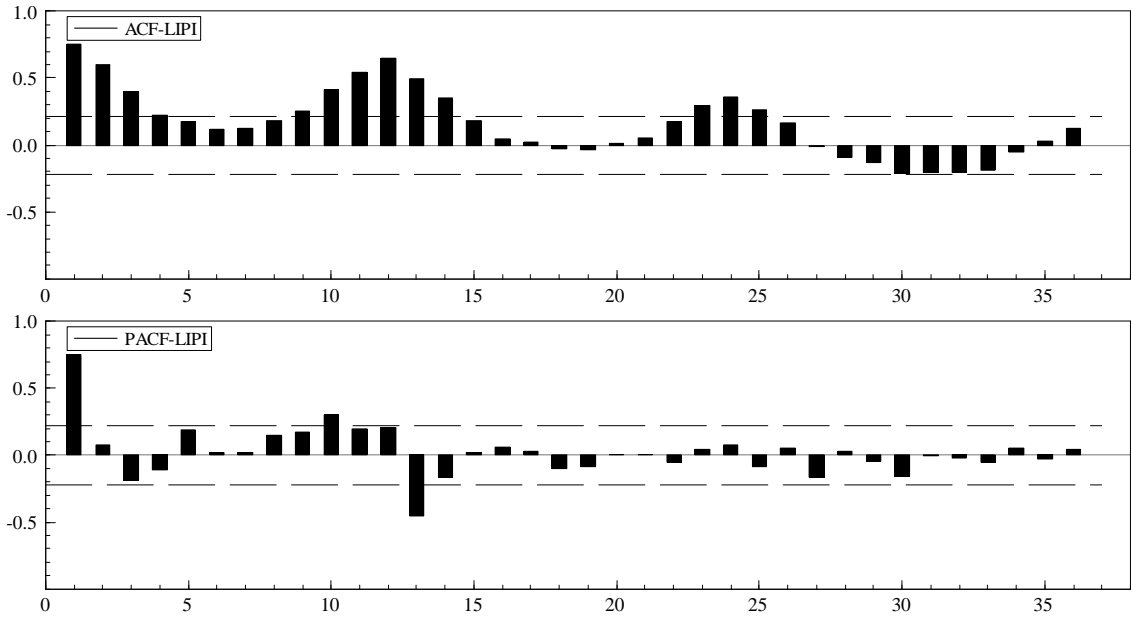


Gráfico 27 – Correlograma da série LIPI

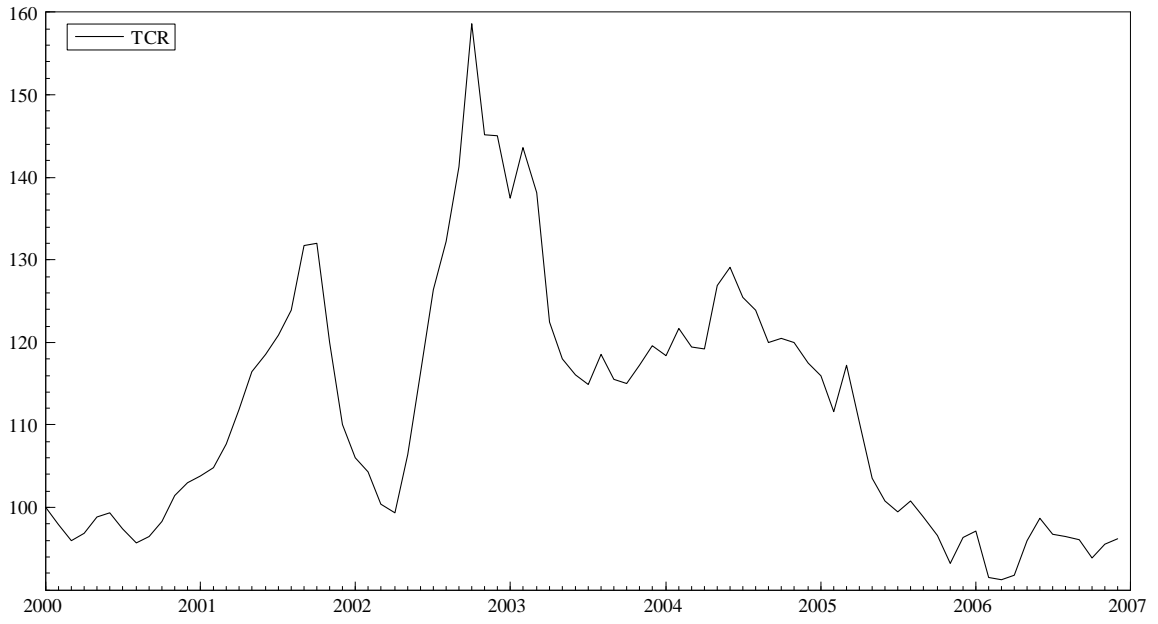


Gráfico 28 – Série Taxa de Câmbio Real – Índice

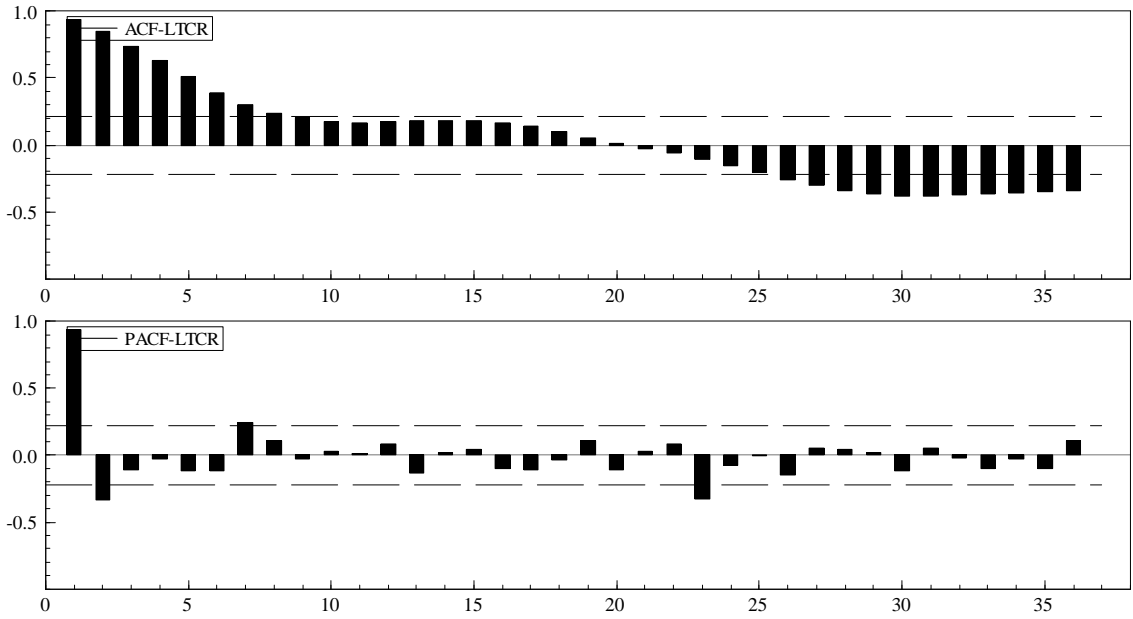


Gráfico 29 – Correlograma da série LTCR

APÊNDICE C – Resultados dos principais modelos estimados

Especificação de modelos ARIMA

II - Tabela 8 - Mod. nº 5

Dependent Variable: DLOG(II)
 Method: Least Squares
 Date: 04/23/08 Time: 22:10
 Sample: 2002M04 2006M12
 Included observations: 57
 Convergence achieved after 12 iterations
 Backcast: 2001M03 2002M03

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MA(1)	-0.296603	0.107744	-2.752854	0.0080
SMA(12)	0.873503	0.030167	28.95573	0.0000
R-squared	0.524268	Mean dependent var		8.24E-06
Adjusted R-squared	0.515619	S.D. dependent var		0.120967
S.E. of regression	0.084190	Akaike info criterion		-2.077024
Sum squared resid	0.389838	Schwarz criterion		-2.005338
Log likelihood	61.19518	Durbin-Watson stat		2.314598
Inverted MA Roots	.96-.26i	.96+.26i	.70+.70i	.70-.70i
	.30	.26-.96i	.26+.96i	-.26-.96i
	-.26+.96i	-.70-.70i	-.70-.70i	-.96+.26i
	-.96-.26i			

IRPJ DOLR – Tabela 9 - Mod. nº 6

Dependent Variable: DLOG(LR,1,12)
 Method: Least Squares
 Date: 05/05/08 Time: 10:07
 Sample: 2002M10 2006M12
 Included observations: 51
 Convergence achieved after 9 iterations
 Backcast: 2001M10 2002M09

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.478440	0.127479	-3.753080	0.0005
MA(12)	-0.854118	0.041348	-20.65657	0.0000
R-squared	0.581666	Mean dependent var		-0.001588
Adjusted R-squared	0.573128	S.D. dependent var		0.276638
S.E. of regression	0.180742	Akaike info criterion		-0.545063
Sum squared resid	1.600721	Schwarz criterion		-0.469305
Log likelihood	15.89911	Durbin-Watson stat		2.053279
Inverted AR Roots	-.48			
Inverted MA Roots	.99	.85-.49i	.85+.49i	.49-.85i
	.49+.85i	.00+.99i	-.00-.99i	-.49-.85i
	-.49+.85i	-.85+.49i	-.85-.49i	-.99

COFINS DE – Tabela 10 - Mod. nº 3

Dependent Variable: DLOG(CF,0,12)
 Method: Least Squares
 Date: 12/07/08 Time: 21:16
 Sample: 2002M06 2006M12
 Included observations: 55
 Convergence achieved after 11 iterations
 Backcast: 2001M06 2002M05

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.110772	0.048964	-2.262314	0.0281
@TREND	0.001709	0.000820	2.083504	0.0423
AR(1)	0.584897	0.103558	5.648012	0.0000
AR(10)	-0.273527	0.103411	-2.645059	0.0109
MA(12)	-0.868470	0.047088	-18.44343	0.0000
R-squared	0.824395	Mean dependent var		-0.005452
Adjusted R-squared	0.810346	S.D. dependent var		0.131341
S.E. of regression	0.057198	Akaike info criterion		-2.798087
Sum squared resid	0.163581	Schwarz criterion		-2.615603
Log likelihood	81.94740	F-statistic		58.68238
Durbin-Watson stat	1.954577	Prob(F-statistic)		0.000000
Inverted AR Roots	.92+.26i .05-.86i -.79+.27i	.92-.26i .05+.86i -.79-.27i	.58+.69i -.47-.70i	.58-.69i -.47+.70i
Inverted MA Roots	.99 .49+.86i -.49+.86i	.86+.49i .00-.99i -.86-.49i	.86-.49i -.00+.99i -.86+.49i	.49-.86i -.49-.86i -.99

Especificação modelos univariados e de Função de Transferência – II / IRPJ DOLR / COFINS DE – Tabela 11

Mod. nº 1: DLII

EQ(6) Modelling DLii by OLS (using II.in7)
 The estimation sample is: 2000 (8) to 2006 (12)

	Coefficient	Std. Error	t-value	t-prob	Part.R^2
DLii_1	-0.457706	0.09142	-5.01	0.000	0.2609
DLii_3	0.182504	0.08502	2.15	0.035	0.0609
CSeasonal_3	-0.104516	0.03843	-2.72	0.008	0.0944
CSeasonal_4	-0.100145	0.04219	-2.37	0.020	0.0735
CSeasonal_5	-0.171113	0.04076	-4.20	0.000	0.1989
CSeasonal_6	0.134109	0.04424	3.03	0.003	0.1146

Mod. nº 2: DLLR

EQ(8) Modelling DLlr by OLS (using IRPJ.in7)
 The estimation sample is: 2001 (2) to 2006 (12)

	Coefficient	Std. Error	t-value	t-prob	Part.R^2
DLlr_1	-0.605234	0.06272	-9.65	0.000	0.6162
DLlr_2	-0.454470	0.08924	-5.09	0.000	0.3090
DLlr_3	-0.296906	0.08829	-3.36	0.001	0.1632
DLlr_4	-0.421693	0.08033	-5.25	0.000	0.3221
DLlr_5	-0.502107	0.06522	-7.70	0.000	0.5054
CSeasonal_1	0.114334	0.05740	1.99	0.051	0.0640
CSeasonal_3	0.160396	0.05789	2.77	0.008	0.1169
CSeasonal_4	0.655589	0.06655	9.85	0.000	0.6259
CSeasonal_6	0.280096	0.07432	3.77	0.000	0.1967
I200305_LR	0.610456	0.1302	4.69	0.000	0.2750
I200601_LR	-0.328488	0.1406	-2.34	0.023	0.0860
L200408_LR	0.0521620	0.02443	2.14	0.037	0.0729

I200408_LR 0.396640 0.1275 3.11 0.003 0.1430

Mod. nº 3: DLCF

EQ(16) Modelling DLcf by OLS (using Cofins.in7)
The estimation sample is: 2001 (3) to 2006 (12)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
DLcf_1	-0.376197	0.09775	-3.85	0.000	0.1928
DLcf_2	-0.216598	0.07648	-2.83	0.006	0.1146
CSeasonal_1	-0.0561329	0.02475	-2.27	0.027	0.0766
CSeasonal_4	0.0514105	0.02566	2.00	0.049	0.0608
CSeasonal_5	-0.214140	0.02619	-8.18	0.000	0.5188
CSeasonal_6	-0.0821194	0.03008	-2.73	0.008	0.1073
CSeasonal_10	-0.0577256	0.02544	-2.27	0.027	0.0767
I200407_CF	-0.233387	0.05897	-3.96	0.000	0.2017

Mod. nº 4: DLCF(DLPIB)

EQ(17) Modelling DLcf by OLS (using Cofins.in7)
The estimation sample is: 2001 (2) to 2006 (12)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
DLcf_1	-0.248179	0.06594	-3.76	0.000	0.1789
DLpib_i_1	1.08892	0.1903	5.72	0.000	0.3349
CSeasonal_4	0.0932121	0.02366	3.94	0.000	0.1928
CSeasonal_5	-0.141411	0.02392	-5.91	0.000	0.3496
I200407_CF	-0.252794	0.04945	-5.11	0.000	0.2868
I200608_CF	0.171018	0.04967	3.44	0.001	0.1543

Mod. nº 5: DLCF(DLIVV)

EQ(14) Modelling DLcf by OLS (using Cofins.in7)
The estimation sample is: 2001 (2) to 2006 (12)

	Coefficient	Std.Error	t-value	t-prob	Part.R^2
DLcf_1	-0.316852	0.07176	-4.42	0.000	0.2307
DLcf_2	-0.157515	0.07284	-2.16	0.034	0.0671
DLIVV_sp	0.106120	0.04834	2.20	0.032	0.0690
DLIVV_sp_1	0.457326	0.04736	9.66	0.000	0.5893
CSeasonal	0.0532544	0.02299	2.32	0.024	0.0763
I200407_CF	-0.232012	0.05223	-4.44	0.000	0.2328

Especificação dos modelos VAR/VEC/SEM

II - Tabela 12

Mod. nº 1: LII/LIVV/LTCR

SYS(1) Estimating the system by OLS (using II.in7)
The estimation sample is: 2000 (7) to 2006 (12)

URF equation for: Lii

		Coefficient	Std.Error	t-value	t-prob
Lii_1		-0.160998	0.1041	-1.55	0.128
Lii_2		-0.102002	0.1170	-0.871	0.387
Lii_3		0.187261	0.09232	2.03	0.048
Livv_sp_1		1.10816	0.3144	3.52	0.001
Livv_sp_2		0.911479	0.3048	2.99	0.004
Livv_sp_3		0.256031	0.1017	2.52	0.015
LTCR_1		0.959793	0.2431	3.95	0.000
LTCR_2		-0.383884	0.3531	-1.09	0.282
LTCR_3		-0.157847	0.2557	-0.617	0.540
Trend		-0.00570011	0.002096	-2.72	0.009
Constant	U	-6.07665	2.276	-2.67	0.010
Cseasonal	U	-0.0553942	0.02686	-2.06	0.044
Cseasonal_3	U	-0.117022	0.02966	-3.95	0.000
Cseasonal_4	U	-0.509983	0.1116	-4.57	0.000
Cseasonal_5	U	-0.547466	0.1144	-4.79	0.000
Cseasonal_7	U	-0.0189067	0.04179	-0.452	0.653
Cseasonal_9	U	-0.102751	0.03138	-3.27	0.002
I200402_IVV	U	-0.0949005	0.07307	-1.30	0.200
I200210_TCR	U	0.0881837	0.07260	1.21	0.230

I200211_TCR	U	-0.154924	0.07802	-1.99	0.052
I200302_TCR	U	0.283282	0.07507	3.77	0.000
L200111_TCR	U	-0.287879	0.06552	-4.39	0.000
L200205_TCR	U	0.0605231	0.04581	1.32	0.192
L200301_TCR	U	-0.0486687	0.05259	-0.925	0.359
L200504_TCR	U	-0.0292156	0.04078	-0.716	0.477

sigma = 0.0611487 RSS = 0.1981758773

URF equation for: Livv_sp

		Coefficient	Std.Error	t-value	t-prob
Lii_1		0.0636589	0.03578	1.78	0.081
Lii_2		-0.0457701	0.04024	-1.14	0.260
Lii_3		-0.0180440	0.03174	-0.569	0.572
Livv_sp_1		0.287694	0.1081	2.66	0.010
Livv_sp_2		0.223067	0.1048	2.13	0.038
Livv_sp_3		0.0790337	0.03497	2.26	0.028
LTCR_1		-0.0518745	0.08357	-0.621	0.537
LTCR_2		-0.0766108	0.1214	-0.631	0.531
LTCR_3		0.0582666	0.08789	0.663	0.510
Trend		0.00188136	0.0007205	2.61	0.012
Constant	U	2.21937	0.7824	2.84	0.006
Cseasonal	U	-0.0455853	0.009232	-4.94	0.000
Cseasonal_3	U	0.296931	0.01020	29.1	0.000
Cseasonal_4	U	-0.163299	0.03837	-4.26	0.000
Cseasonal_5	U	-0.206926	0.03932	-5.26	0.000
Cseasonal_7	U	-0.0264015	0.01437	-1.84	0.072
Cseasonal_9	U	-0.0606954	0.01079	-5.63	0.000
I200402_IVV	U	0.0717608	0.02512	2.86	0.006
I200210_TCR	U	0.0577150	0.02496	2.31	0.025
I200211_TCR	U	0.00782266	0.02682	0.292	0.772
I200302_TCR	U	0.0581644	0.02580	2.25	0.028
L200111_TCR	U	-0.0201623	0.02252	-0.895	0.375
L200205_TCR	U	-0.0267225	0.01575	-1.70	0.096
L200301_TCR	U	-0.0171873	0.01808	-0.951	0.346
L200504_TCR	U	-0.0165636	0.01402	-1.18	0.243

sigma = 0.0210202 RSS = 0.02341792828

URF equation for: LTCR

		Coefficient	Std.Error	t-value	t-prob
Lii_1		0.0313137	0.03556	0.881	0.383
Lii_2		-0.0470615	0.03999	-1.18	0.244
Lii_3		-0.0732618	0.03154	-2.32	0.024
Livv_sp_1		-0.0604903	0.1074	-0.563	0.576
Livv_sp_2		-0.217970	0.1041	-2.09	0.041
Livv_sp_3		-0.0370950	0.03475	-1.07	0.291
LTCR_1		1.02640	0.08305	12.4	0.000
LTCR_2		-0.382652	0.1206	-3.17	0.003
LTCR_3		0.256554	0.08735	2.94	0.005
Trend		0.00348645	0.0007160	4.87	0.000
Constant	U	2.44505	0.7776	3.14	0.003
Cseasonal	U	-0.00607481	0.009175	-0.662	0.511
Cseasonal_3	U	0.0120347	0.01013	1.19	0.240
Cseasonal_4	U	0.0348499	0.03813	0.914	0.365
Cseasonal_5	U	0.0521012	0.03908	1.33	0.188
Cseasonal_7	U	-0.0442662	0.01428	-3.10	0.003
Cseasonal_9	U	0.00278215	0.01072	0.260	0.796
I200402_IVV	U	0.0531529	0.02496	2.13	0.038
I200210_TCR	U	0.0853552	0.02480	3.44	0.001
I200211_TCR	U	-0.131119	0.02666	-4.92	0.000
I200302_TCR	U	0.137898	0.02564	5.38	0.000
L200111_TCR	U	-0.120202	0.02238	-5.37	0.000
L200205_TCR	U	0.0768331	0.01565	4.91	0.000
L200301_TCR	U	-0.147291	0.01797	-8.20	0.000
L200504_TCR	U	-0.0797462	0.01393	-5.72	0.000

sigma = 0.0208899 RSS = 0.02312858304

Mod. n° 3: DLII/DLIVV/DLTCR

Identity for Cia

Dlii	1.0000
Dlivv_sp	-0.37050
DLTCR	0.036601
Cia_1	1.0000

Constant -0.0070420
R^2 = 1 over 2000 (7) to 2006 (12)

Identity for Cib

Dlii 0.36002
Dlivv_sp 1.0000
DLTCR 0.28429
Cib_1 1.0000

Constant -0.010376

R^2 = 1 over 2000 (7) to 2006 (12)

MaxBFGS(): initial Hessian dimension error, using I

MOD(9) Estimating the model by FIML (using II.in7)

The estimation sample is: 2000 (7) to 2006 (12)

Equation for: Dlii

		Coefficient	Std.Error	t-value	t-prob
Dlivv_sp_1		-1.28922	0.2519	-5.12	0.000
DLTCR_1		0.433782	0.2046	2.12	0.039
Dlii_2		-0.140899	0.06586	-2.14	0.037
Dlivv_sp_2		-0.312180	0.07801	-4.00	0.000
Cia_1		-1.78533	0.1796	-9.94	0.000
Cib_1		1.82634	0.2781	6.57	0.000
Constant		-6.91601	1.540	-4.49	0.000
I200402_IVV	U	-0.0892335	0.06892	-1.29	0.201
I200210_TCR	U	0.0992119	0.06762	1.47	0.148
I200211_TCR	U	-0.153885	0.07069	-2.18	0.034
I200302_TCR	U	0.284167	0.07123	3.99	0.000
L200111_TCR	U	-0.302877	0.04281	-7.08	0.000
L200205_TCR	U	0.0755751	0.04189	1.80	0.077
L200301_TCR	U	-0.0548466	0.04060	-1.35	0.182
L200504_TCR	U	-0.0268590	0.03382	-0.794	0.431
Cseasonal	U	-0.0571735	0.02543	-2.25	0.029
Cseasonal_3	U	-0.118827	0.02647	-4.49	0.000
Cseasonal_4	U	-0.540034	0.09527	-5.67	0.000
Cseasonal_5	U	-0.565107	0.09734	-5.81	0.000
Cseasonal_7	U	-0.0195164	0.03657	-0.534	0.596
Cseasonal_9	U	-0.106812	0.02929	-3.65	0.001

sigma = 0.0590596

Equation for: Dlivv_sp

		Coefficient	Std.Error	t-value	t-prob
Dlii_1		0.0688853	0.02484	2.77	0.008
Dlivv_sp_1		-0.463122	0.07362	-6.29	0.000
Dlivv_sp_2		-0.106436	0.02559	-4.16	0.000
Cib_1		-0.120200	0.03961	-3.03	0.004
Constant		0.963089	0.3181	3.03	0.004
I200402_IVV	U	0.0760738	0.02456	3.10	0.003
I200210_TCR	U	0.0541142	0.02348	2.30	0.025
I200211_TCR	U	0.00447220	0.02490	0.180	0.858
I200302_TCR	U	0.0672840	0.02529	2.66	0.010
L200111_TCR	U	-0.0238423	0.01376	-1.73	0.089
L200205_TCR	U	-0.0250970	0.01346	-1.87	0.068
L200301_TCR	U	-0.0103026	0.01352	-0.762	0.449
L200504_TCR	U	-0.0217579	0.01137	-1.91	0.061
Cseasonal	U	-0.0451467	0.009099	-4.96	0.000
Cseasonal_3	U	0.288085	0.009330	30.9	0.000
Cseasonal_4	U	-0.213880	0.02442	-8.76	0.000
Cseasonal_5	U	-0.254920	0.03081	-8.27	0.000
Cseasonal_7	U	-0.0165827	0.01291	-1.28	0.204
Cseasonal_9	U	-0.0629155	0.009888	-6.36	0.000

sigma = 0.0211654

Equation for: DLTCR

		Coefficient	Std.Error	t-value	t-prob
Dlii_1		0.160010	0.02798	5.72	0.000
Dlivv_sp_1		0.277150	0.07231	3.83	0.000
Dlii_2		0.0990039	0.02532	3.91	0.000
DLTCR_2		-0.247310	0.07142	-3.46	0.001
Cib_1		-0.305916	0.03638	-8.41	0.000
Constant		2.47744	0.2921	8.48	0.000
I200402_IVV	U	0.0465109	0.02428	1.92	0.061
I200210_TCR	U	0.0981186	0.02352	4.17	0.000
I200211_TCR	U	-0.126987	0.02402	-5.29	0.000
I200302_TCR	U	0.125377	0.02442	5.13	0.000

L200111_TCR	U	-0.129812	0.01352	-9.60	0.000
L200205_TCR	U	0.0849187	0.01439	5.90	0.000
L200301_TCR	U	-0.148382	0.01349	-11.0	0.000
L200504_TCR	U	-0.0713989	0.01056	-6.76	0.000
Cseasonal	U	-0.00409461	0.008939	-0.458	0.649
Cseasonal_3	U	0.0151518	0.009009	1.68	0.098
Cseasonal_4	U	0.0309930	0.02415	1.28	0.205
Cseasonal_5	U	0.0789640	0.02892	2.73	0.008
Cseasonal_7	U	-0.0508010	0.01273	-3.99	0.000
Cseasonal_9	U	0.00506923	0.01021	0.496	0.622

sigma = 0.0207468

Mod. n° 4: LII/LIVV/LTCR

SYS(17) Estimating the system by OLS (using II.in7)
The estimation sample is: 2000 (7) to 2006 (12)

URF equation for: Lii

		Coefficient	Std.Error	t-value	t-prob
Lii_1		-0.170349	0.1254	-1.36	0.181
Lii_2		-0.0766933	0.1263	-0.607	0.547
Lii_3		0.228778	0.1099	2.08	0.043
Lii_4		-0.0477049	0.1095	-0.436	0.665
Lii_5		0.0719129	0.1068	0.674	0.504
Lii_6		-0.0211723	0.1104	-0.192	0.849
Livv_sp_1		1.01486	0.3770	2.69	0.010
Livv_sp_2		0.968455	0.3326	2.91	0.006
Livv_sp_3		0.225457	0.1356	1.66	0.103
Livv_sp_4		0.0233525	0.1091	0.214	0.832
Livv_sp_5		0.137864	0.1031	1.34	0.188
Livv_sp_6		-0.253137	0.1068	-2.37	0.022
LTCR_1		1.05379	0.2760	3.82	0.000
LTCR_2		-0.434620	0.3792	-1.15	0.258
LTCR_3		-0.357194	0.3998	-0.893	0.376
LTCR_4		0.299261	0.4407	0.679	0.501
LTCR_5		-0.0379644	0.3960	-0.0959	0.924
LTCR_6		-0.179758	0.2665	-0.674	0.503
Trend		-0.00543898	0.002205	-2.47	0.018
Constant	U	-5.35160	2.736	-1.96	0.057
Cseasonal	U	-0.0385637	0.02873	-1.34	0.186
Cseasonal_3	U	-0.120934	0.03323	-3.64	0.001
Cseasonal_4	U	-0.484486	0.1316	-3.68	0.001
Cseasonal_5	U	-0.562716	0.1273	-4.42	0.000
I200402_IVV	U	-0.0833250	0.07721	-1.08	0.286
I200210_TCR	U	0.0528194	0.07657	0.690	0.494
I200211_TCR	U	-0.144934	0.08416	-1.72	0.092
I200302_TCR	U	0.265369	0.09128	2.91	0.006
L200111_TCR	U	-0.264052	0.07722	-3.42	0.001
L200205_TCR	U	0.0463041	0.05936	0.780	0.439
L200301_TCR	U	-0.0322175	0.06277	-0.513	0.610
L200504_TCR	U	-0.0334164	0.04354	-0.767	0.447
I200110_IVV	U	-0.115909	0.07985	-1.45	0.154

sigma = 0.0612789 RSS = 0.1689797599

URF equation for: Livv_sp

		Coefficient	Std.Error	t-value	t-prob
Lii_1		0.117446	0.03978	2.95	0.005
Lii_2		-0.0135325	0.04009	-0.338	0.737
Lii_3		0.00330116	0.03488	0.0946	0.925
Lii_4		0.0353043	0.03473	1.02	0.315
Lii_5		-0.0517930	0.03387	-1.53	0.133
Lii_6		-0.00954806	0.03501	-0.273	0.786
Livv_sp_1		0.311227	0.1196	2.60	0.013
Livv_sp_2		0.0822255	0.1055	0.779	0.440
Livv_sp_3		0.145119	0.04304	3.37	0.002
Livv_sp_4		-0.0443625	0.03463	-1.28	0.207
Livv_sp_5		0.118703	0.03272	3.63	0.001
Livv_sp_6		-0.111381	0.03389	-3.29	0.002
LTCR_1		-0.144019	0.08755	-1.64	0.107
LTCR_2		0.0357564	0.1203	0.297	0.768
LTCR_3		-0.160106	0.1268	-1.26	0.213
LTCR_4		0.229856	0.1398	1.64	0.107
LTCR_5		-0.135814	0.1256	-1.08	0.285
LTCR_6		-0.00232454	0.08456	-0.0275	0.978

Trend		0.00132262	0.0006996	1.89	0.065
Constant	U	2.64274	0.8679	3.04	0.004
Cseasonal	U	-0.0330692	0.009114	-3.63	0.001
Cseasonal_3	U	0.303667	0.01054	28.8	0.000
Cseasonal_4	U	-0.158580	0.04175	-3.80	0.000
Cseasonal_5	U	-0.144119	0.04038	-3.57	0.001
I200402_IVV	U	0.0686496	0.02450	2.80	0.007
I200210_TCR	U	0.0645732	0.02429	2.66	0.011
I200211_TCR	U	0.00700199	0.02670	0.262	0.794
I200302_TCR	U	0.0255720	0.02896	0.883	0.382
L200111_TCR	U	0.0309957	0.02450	1.27	0.212
L200205_TCR	U	-0.0290061	0.01883	-1.54	0.131
L200301_TCR	U	0.0118574	0.01992	0.595	0.555
L200504_TCR	U	-0.0145961	0.01381	-1.06	0.296
I200110_IVV	U	0.0894227	0.02533	3.53	0.001

sigma = 0.0194417 RSS = 0.0170090349

URF equation for: LTCR

	Coefficient	Std.Error	t-value	t-prob	
Lii_1	-0.00521307	0.03950	-0.132	0.896	
Lii_2	0.00666667	0.03981	0.167	0.868	
Lii_3	-0.0599270	0.03464	-1.73	0.090	
Lii_4	-0.0217083	0.03448	-0.630	0.532	
Lii_5	-0.0336679	0.03363	-1.00	0.322	
Lii_6	0.0466067	0.03477	1.34	0.187	
Livv_sp_1	-0.0665538	0.1188	-0.560	0.578	
Livv_sp_2	-0.274784	0.1048	-2.62	0.012	
Livv_sp_3	-0.0532409	0.04273	-1.25	0.219	
Livv_sp_4	-0.0902034	0.03438	-2.62	0.012	
Livv_sp_5	-0.0120610	0.03249	-0.371	0.712	
Livv_sp_6	-0.0219348	0.03366	-0.652	0.518	
LTCR_1	1.01832	0.08694	11.7	0.000	
LTCR_2	-0.437190	0.1195	-3.66	0.001	
LTCR_3	0.323394	0.1259	2.57	0.014	
LTCR_4	-0.106894	0.1388	-0.770	0.445	
LTCR_5	0.304802	0.1248	2.44	0.019	
LTCR_6	-0.312226	0.08397	-3.72	0.001	
Trend	0.00370441	0.0006947	5.33	0.000	
Constant	U	3.77225	0.8618	4.38	0.000
Cseasonal	U	-0.000290315	0.009050	-0.0321	0.975
Cseasonal_3	U	0.0170043	0.01047	1.62	0.111
Cseasonal_4	U	0.0280934	0.04145	0.678	0.501
Cseasonal_5	U	0.0711749	0.04010	1.78	0.083
I200402_IVV	U	0.0579646	0.02433	2.38	0.021
I200210_TCR	U	0.0599580	0.02412	2.49	0.017
I200211_TCR	U	-0.140219	0.02651	-5.29	0.000
I200302_TCR	U	0.149908	0.02876	5.21	0.000
L200111_TCR	U	-0.108185	0.02433	-4.45	0.000
L200205_TCR	U	0.0781382	0.01870	4.18	0.000
L200301_TCR	U	-0.142760	0.01978	-7.22	0.000
L200504_TCR	U	-0.0880106	0.01372	-6.42	0.000
I200110_IVV	U	-0.0226601	0.02516	-0.901	0.373

sigma = 0.0193054 RSS = 0.01677148352

Mod. n° 6: DLII/DLIVV/DLTCR

Identity for Cia

Dlii	1.0000
Dlivv_sp	-3.8034
DLTCR	-0.99863
Cia_1	1.0000
Constant	0.014089

R² = 1 over 2000 (7) to 2006 (12)

Identity for Cib

Dlii	7.0252
Dlivv_sp	1.0000
DLTCR	3.2255
Cib_1	1.0000
Constant	-0.056093

R² = 1 over 2000 (7) to 2006 (12)

MaxBFGS(): initial Hessian dimension error, using I

MOD(38) Estimating the model by FIML (using II.in7)

The estimation sample is: 2000 (7) to 2006 (12)

Equation for: Dlii

		Coefficient	Std.Error	t-value	t-prob
Dlii_1		-0.371606	0.1000	-3.71	0.001
Dlivv_sp_1		-0.857950	0.2462	-3.48	0.001
DLTCR_1		0.988733	0.2292	4.31	0.000
Dlii_2		-0.369010	0.07461	-4.95	0.000
Dlivv_sp_5		0.209154	0.05513	3.79	0.000
Cia_1		-0.416799	0.09599	-4.34	0.000
Cib_1		-0.0574359	0.009165	-6.27	0.000
Constant		-3.27979	1.550	-2.12	0.040
I200402_IVV	U	-0.0557199	0.07060	-0.789	0.434
I200210_TCR	U	0.0364718	0.07034	0.519	0.607
I200211_TCR	U	-0.114061	0.07186	-1.59	0.119
I200302_TCR	U	0.326264	0.07276	4.48	0.000
L200111_TCR	U	-0.243627	0.04478	-5.44	0.000
L200205_TCR	U	0.0385140	0.04158	0.926	0.359
L200301_TCR	U	-0.0621677	0.03568	-1.74	0.088
L200504_TCR	U	-0.0618624	0.02932	-2.11	0.040
I200110_IVV	U	-0.144853	0.06760	-2.14	0.037
Cseasonal	U	-0.0462502	0.02530	-1.83	0.074
Cseasonal_3	U	-0.110177	0.02752	-4.00	0.000
Cseasonal_4	U	-0.378603	0.09183	-4.12	0.000
Cseasonal_5	U	-0.551394	0.09622	-5.73	0.000

sigma = 0.0607107

Equation for: Dlivv_sp

		Coefficient	Std.Error	t-value	t-prob
Dlivv_sp_1		-0.264223	0.06150	-4.30	0.000
Dlivv_sp_2		-0.105104	0.02445	-4.30	0.000
Dlii_4		0.0512243	0.01944	2.63	0.011
Dlivv_sp_5		0.0975918	0.01590	6.14	0.000
Cia_1		0.116430	0.02756	4.23	0.000
Cib_1		-0.00538694	0.002648	-2.03	0.048
Constant		2.21831	0.5032	4.41	0.000
I200402_IVV	U	0.0738237	0.02160	3.42	0.001
I200210_TCR	U	0.0656976	0.02084	3.15	0.003
I200211_TCR	U	0.0150169	0.02348	0.640	0.526
I200302_TCR	U	0.0513569	0.02222	2.31	0.025
L200111_TCR	U	0.0101467	0.01216	0.835	0.408
L200205_TCR	U	-0.0236321	0.01156	-2.04	0.047
L200301_TCR	U	-0.00615133	0.01082	-0.569	0.572
L200504_TCR	U	-0.0238849	0.009854	-2.42	0.019
I200110_IVV	U	0.0698620	0.02025	3.45	0.001
Cseasonal	U	-0.0346956	0.007933	-4.37	0.000
Cseasonal_3	U	0.298643	0.008444	35.4	0.000
Cseasonal_4	U	-0.154825	0.03140	-4.93	0.000
Cseasonal_5	U	-0.176151	0.02531	-6.96	0.000

sigma = 0.018713

Equation for: DLTCR

		Coefficient	Std.Error	t-value	t-prob
Dlivv_sp_1		0.486308	0.08070	6.03	0.000
DLTCR_1		0.331130	0.07041	4.70	0.000
Dlivv_sp_2		0.219023	0.03124	7.01	0.000
DLTCR_2		-0.150638	0.06391	-2.36	0.023
Dlivv_sp_3		0.0961278	0.02435	3.95	0.000
DLTCR_3		0.137923	0.06998	1.97	0.055
DLTCR_5		0.277854	0.06017	4.62	0.000
Cia_1		0.174412	0.02879	6.06	0.000
Cib_1		-0.0281268	0.003072	-9.16	0.000
Constant		4.56969	0.5478	8.34	0.000
I200402_IVV	U	0.0509910	0.02158	2.36	0.022
I200210_TCR	U	0.0508549	0.02166	2.35	0.023
I200211_TCR	U	-0.129467	0.02248	-5.76	0.000
I200302_TCR	U	0.148394	0.02378	6.24	0.000
L200111_TCR	U	-0.0907248	0.01283	-7.07	0.000
L200205_TCR	U	0.0688673	0.01415	4.87	0.000
L200301_TCR	U	-0.134633	0.01142	-11.8	0.000
L200504_TCR	U	-0.0904132	0.01080	-8.37	0.000
I200110_IVV	U	-0.0365221	0.02047	-1.78	0.081
Cseasonal	U	0.00405155	0.007868	0.515	0.609
Cseasonal_3	U	0.0183553	0.008319	2.21	0.032
Cseasonal_4	U	0.0702961	0.02922	2.41	0.020

Cseasonal_5 U 0.0624095 0.02804 2.23 0.031

sigma = 0.0183972

Mod. n° 7: LII/LIPI/LTCR

SYS(6) Estimating the system by OLS (using II.in7)
The estimation sample is: 2000 (7) to 2006 (12)

URF equation for: Lii

		Coefficient	Std.Error	t-value	t-prob
Lii_1		0.0824898	0.1355	0.609	0.546
Lii_2		0.382202	0.1576	2.43	0.019
Lii_3		0.398505	0.1442	2.76	0.008
Lii_4		0.196944	0.1440	1.37	0.178
LTCR_1		0.995356	0.2685	3.71	0.001
LTCR_2		-1.14537	0.3872	-2.96	0.005
LTCR_3		-0.208230	0.3736	-0.557	0.580
LTCR_4		0.0539126	0.2740	0.197	0.845
LIPI_1		-0.438889	0.2964	-1.48	0.145
LIPI_2		0.00247425	0.2931	0.00844	0.993
LIPI_3		0.423734	0.2809	1.51	0.138
LIPI_4		-0.778881	0.2521	-3.09	0.003
Trend		0.00127580	0.002049	0.623	0.537
Constant	U	4.84464	1.096	4.42	0.000
Cseasonal_2	U	-0.0231816	0.03412	-0.679	0.500
Cseasonal_3	U	-0.0450502	0.03513	-1.28	0.206
Cseasonal_4	U	-0.171341	0.04967	-3.45	0.001
Cseasonal_5	U	-0.308310	0.04531	-6.80	0.000
Cseasonal_7	U	-0.0678818	0.04428	-1.53	0.132
Cseasonal_9	U	-0.120541	0.04431	-2.72	0.009
I200210_TCR	U	0.0344543	0.07514	0.459	0.649
I200302_TCR	U	0.355325	0.08087	4.39	0.000
L200111_TCR	U	-0.0324734	0.08563	-0.379	0.706
L200205_TCR	U	0.0732036	0.05673	1.29	0.203
L200301_TCR	U	-0.0981272	0.06541	-1.50	0.140
L200504_TCR	U	-0.0236667	0.04280	-0.553	0.583
I200211_TCR	U	-0.0614145	0.08605	-0.714	0.479
I200112_IPI	U	-0.104205	0.08466	-1.23	0.224
L200309_IPI	U	0.158636	0.04932	3.22	0.002
I200404_II	U	0.237313	0.07882	3.01	0.004

sigma = 0.0626869 RSS = 0.18862301

URF equation for: LTCR

		Coefficient	Std.Error	t-value	t-prob
Lii_1		0.0152829	0.04785	0.319	0.751
Lii_2		-0.0488884	0.05566	-0.878	0.384
Lii_3		-0.184179	0.05091	-3.62	0.001
Lii_4		0.0293567	0.05086	0.577	0.566
LTCR_1		1.05543	0.09481	11.1	0.000
LTCR_2		-0.364556	0.1368	-2.67	0.010
LTCR_3		0.342994	0.1319	2.60	0.012
LTCR_4		-0.0362915	0.09675	-0.375	0.709
LIPI_1		-0.112434	0.1047	-1.07	0.288
LIPI_2		-0.0984742	0.1035	-0.951	0.346
LIPI_3		0.250572	0.09922	2.53	0.015
LIPI_4		-0.0973327	0.08903	-1.09	0.280
Trend		0.00265339	0.0007237	3.67	0.001
Constant	U	1.40495	0.3869	3.63	0.001
Cseasonal_2	U	0.00770305	0.01205	0.639	0.526
Cseasonal_3	U	0.00909318	0.01241	0.733	0.467
Cseasonal_4	U	-0.00902428	0.01754	-0.514	0.609
Cseasonal_5	U	-0.0269947	0.01600	-1.69	0.098
Cseasonal_7	U	-0.0212237	0.01564	-1.36	0.181
Cseasonal_9	U	-0.00237000	0.01565	-0.151	0.880
I200210_TCR	U	0.0849046	0.02653	3.20	0.002
I200302_TCR	U	0.130029	0.02856	4.55	0.000
L200111_TCR	U	-0.143087	0.03024	-4.73	0.000
L200205_TCR	U	0.0771433	0.02004	3.85	0.000
L200301_TCR	U	-0.147877	0.02310	-6.40	0.000
L200504_TCR	U	-0.0682845	0.01511	-4.52	0.000
I200211_TCR	U	-0.131593	0.03039	-4.33	0.000
I200112_IPI	U	-0.0194941	0.02990	-0.652	0.518
L200309_IPI	U	0.00496666	0.01742	0.285	0.777
I200404_II	U	0.000369045	0.02784	0.0133	0.989

sigma = 0.0221378 RSS = 0.02352397996

URF equation for: LIPI

		Coefficient	Std.Error	t-value	t-prob
Lii_1		-0.0429968	0.05006	-0.859	0.395
Lii_2		0.0802046	0.05823	1.38	0.175
Lii_3		0.0685664	0.05326	1.29	0.204
Lii_4		0.0625341	0.05321	1.18	0.246
LTCR_1		0.0167458	0.09919	0.169	0.867
LTCR_2		-0.0629272	0.1431	-0.440	0.662
LTCR_3		-0.00377993	0.1380	-0.0274	0.978
LTCR_4		-0.0425277	0.1012	-0.420	0.676
LIPI_1		0.222473	0.1095	2.03	0.048
LIPI_2		-0.0345203	0.1083	-0.319	0.751
LIPI_3		0.256761	0.1038	2.47	0.017
LIPI_4		-0.384555	0.09314	-4.13	0.000
Trend		0.00250174	0.0007571	3.30	0.002
Constant	U	3.93279	0.4048	9.72	0.000
Cseasonal_2	U	-0.0542597	0.01261	-4.30	0.000
Cseasonal_3	U	-0.124858	0.01298	-9.62	0.000
Cseasonal_4	U	-0.150715	0.01835	-8.21	0.000
Cseasonal_5	U	-0.157844	0.01674	-9.43	0.000
Cseasonal_7	U	-0.0868798	0.01636	-5.31	0.000
Cseasonal_9	U	-0.0902773	0.01637	-5.51	0.000
I200210_TCR	U	0.0656036	0.02776	2.36	0.022
I200302_TCR	U	0.0946353	0.02988	3.17	0.003
L200111_TCR	U	0.0136030	0.03164	0.430	0.669
L200205_TCR	U	-0.0174745	0.02096	-0.834	0.409
L200301_TCR	U	-0.0304112	0.02417	-1.26	0.214
L200504_TCR	U	0.00555707	0.01581	0.351	0.727
I200211_TCR	U	0.0398945	0.03179	1.25	0.216
I200112_IPI	U	-0.102478	0.03128	-3.28	0.002
L200309_IPI	U	0.101487	0.01822	5.57	0.000
I200404_II	U	0.00710146	0.02912	0.244	0.808

sigma = 0.0231613 RSS = 0.02574936157

Mod. n° 9: DLII/DLIPI/DLTCR

Identity for Cia

Dlii	1.0000
DLTCR	-0.17602
DLIPI	-2.8150
Cia_1	1.0000
Constant	0.0021463

R² = 1 over 2000 (7) to 2006 (12)

Identity for Cib

Dlii	6.7079
DLTCR	1.0000
DLIPI	2.3203
Cib_1	1.0000
Constant	-0.088529

R² = 1 over 2000 (7) to 2006 (12)

MaxBFGS(): initial Hessian dimension error, using I

MOD(13) Estimating the model by FIML (using II.in7)

The estimation sample is: 2000 (7) to 2006 (12)

Equation for: Dlii

		Coefficient	Std.Error	t-value	t-prob
Dlii_1		-0.603618	0.1077	-5.61	0.000
DLTCR_1		0.925126	0.2117	4.37	0.000
Dlii_2		-0.238599	0.07391	-3.23	0.002
DLIPI_3		0.545974	0.1560	3.50	0.001
Cia_1		0.199051	0.07347	2.71	0.009
Cib_1		-0.0534123	0.01194	-4.47	0.000
Constant		4.58953	0.9289	4.94	0.000
I200210_TCR	U	0.0243357	0.07375	0.330	0.743
I200302_TCR	U	0.357638	0.07774	4.60	0.000
L200111_TCR	U	-0.129432	0.05481	-2.36	0.022
L200205_TCR	U	0.00462810	0.04703	0.0984	0.922
L200301_TCR	U	-0.156069	0.04397	-3.55	0.001
L200504_TCR	U	-0.0784718	0.03490	-2.25	0.029
I200211_TCR	U	-0.0631621	0.07892	-0.800	0.427

I200112_IPI	U	-0.114095	0.08082	-1.41	0.164
L200309_IPI	U	0.0918919	0.04196	2.19	0.033
I200404_II	U	0.217773	0.07776	2.80	0.007
Cseasonal_2	U	-0.0154958	0.03206	-0.483	0.631
Cseasonal_3	U	-0.0621204	0.02932	-2.12	0.039
Cseasonal_4	U	-0.195496	0.03749	-5.21	0.000
Cseasonal_5	U	-0.327245	0.03833	-8.54	0.000
Cseasonal_7	U	-0.0505065	0.03567	-1.42	0.163
Cseasonal_9	U	-0.111804	0.03537	-3.16	0.003

sigma = 0.0644417

Equation for: DLTCR

		Coefficient	Std.Error	t-value	t-prob
Dlii_1		0.194592	0.03549	5.48	0.000
Dlii_2		0.153471	0.03661	4.19	0.000
DLTCR_2		-0.272894	0.07749	-3.52	0.001
DLIPI_2		-0.148839	0.06750	-2.20	0.032
Cib_1		-0.0259543	0.003943	-6.58	0.000
Constant		1.45035	0.2170	6.68	0.000
I200210_TCR	U	0.0925219	0.02415	3.83	0.000
I200302_TCR	U	0.123162	0.02559	4.81	0.000
L200111_TCR	U	-0.138339	0.01736	-7.97	0.000
L200205_TCR	U	0.0862246	0.01599	5.39	0.000
L200301_TCR	U	-0.142900	0.01502	-9.51	0.000
L200504_TCR	U	-0.0653663	0.01072	-6.10	0.000
I200211_TCR	U	-0.127361	0.02509	-5.08	0.000
I200112_IPI	U	-0.0191659	0.02726	-0.703	0.485
L200309_IPI	U	0.00800266	0.01177	0.680	0.500
I200404_II	U	-0.00298190	0.02538	-0.118	0.907
Cseasonal_2	U	0.0106810	0.01060	1.01	0.319
Cseasonal_3	U	0.00976955	0.009669	1.01	0.317
Cseasonal_4	U	-0.000801336	0.01119	-0.0716	0.943
Cseasonal_5	U	-0.0288993	0.01322	-2.19	0.034
Cseasonal_7	U	-0.0273685	0.01174	-2.33	0.024
Cseasonal_9	U	0.00229023	0.01053	0.217	0.829

sigma = 0.0213622

Equation for: DLIPI

		Coefficient	Std.Error	t-value	t-prob
Dlii_1		-0.0934751	0.03298	-2.83	0.007
DLIPI_3		0.307460	0.05410	5.68	0.000
Cia_1		0.287564	0.02594	11.1	0.000
Cib_1		-0.0268050	0.004007	-6.69	0.000
Constant		3.89486	0.3201	12.2	0.000
I200210_TCR	U	0.0773051	0.02530	3.06	0.004
I200302_TCR	U	0.0909750	0.02751	3.31	0.002
L200111_TCR	U	-0.00290383	0.01875	-0.155	0.878
L200205_TCR	U	-0.0324115	0.01512	-2.14	0.037
L200301_TCR	U	-0.0355270	0.01482	-2.40	0.020
L200504_TCR	U	-0.0104872	0.01202	-0.872	0.387
I200211_TCR	U	0.0579602	0.02747	2.11	0.040
I200112_IPI	U	-0.104147	0.02855	-3.65	0.001
L200309_IPI	U	0.0825461	0.01421	5.81	0.000
I200404_II	U	0.00236160	0.02716	0.0869	0.931
Cseasonal_2	U	-0.0528422	0.01113	-4.75	0.000
Cseasonal_3	U	-0.132755	0.01031	-12.9	0.000
Cseasonal_4	U	-0.160662	0.01297	-12.4	0.000
Cseasonal_5	U	-0.162564	0.01322	-12.3	0.000
Cseasonal_7	U	-0.0735887	0.01255	-5.86	0.000
Cseasonal_9	U	-0.0834144	0.01225	-6.81	0.000

sigma = 0.0228637

Mod. n° 10: LII/LPIB/LTCR

SYS(15) Estimating the system by OLS (using II.in7)
The estimation sample is: 2000 (7) to 2006 (12)

URF equation for: Lii

		Coefficient	Std.Error	t-value	t-prob
Lii_1		0.0929244	0.1043	0.891	0.377
Lii_2		0.104580	0.1163	0.899	0.373
Lii_3		0.261276	0.1173	2.23	0.030
Lpib_I_1		-0.381384	0.3792	-1.01	0.319

Lpib_I_2		0.210329	0.4425	0.475	0.637
Lpib_I_3		1.02105	0.3343	3.05	0.004
LTCR_1		0.818607	0.2618	3.13	0.003
LTCR_2		-0.584772	0.3488	-1.68	0.100
LTCR_3		-0.333223	0.2741	-1.22	0.230
Cseasonal	U	-0.0509166	0.03112	-1.64	0.108
Cseasonal_1	U	-0.00988460	0.03084	-0.320	0.750
Cseasonal_4	U	-0.114501	0.03300	-3.47	0.001
Cseasonal_5	U	-0.314419	0.03694	-8.51	0.000
Cseasonal_8	U	0.104517	0.03745	2.79	0.007
L200111_TCR	U	-0.179387	0.06786	-2.64	0.011
I200210_TCR	U	0.0892001	0.08045	1.11	0.273
I200211_TCR	U	-0.0722827	0.08086	-0.894	0.375
I200302_TCR	U	0.295325	0.07694	3.84	0.000
L200205_TCR	U	-0.0223167	0.04927	-0.453	0.652
L200301_TCR	U	-0.0348212	0.04339	-0.802	0.426
L200504_TCR	U	-0.111583	0.03530	-3.16	0.003
Constant	U	-0.269609	1.795	-0.150	0.881
I200604_PIB	U	-0.129032	0.07006	-1.84	0.071
I200009_PIB	U	-0.00282386	0.07644	-0.0369	0.971
I200106_PIB	U	-0.0699072	0.07204	-0.970	0.336
I200506_PIB	U	0.104306	0.07235	1.44	0.155

sigma = 0.0652468 RSS = 0.2213718078

URF equation for: Lpib_I

		Coefficient	Std.Error	t-value	t-prob
Lii_1		-0.0483192	0.01789	-2.70	0.009
Lii_2		0.0398012	0.01995	2.00	0.051
Lii_3		0.0233723	0.02012	1.16	0.251
Lpib_I_1		0.650140	0.06502	10.0	0.000
Lpib_I_2		-0.178608	0.07586	-2.35	0.022
Lpib_I_3		0.377641	0.05731	6.59	0.000
LTCR_1		0.0872365	0.04488	1.94	0.057
LTCR_2		-0.150089	0.05980	-2.51	0.015
LTCR_3		-0.0620781	0.04700	-1.32	0.192
Cseasonal	U	-0.0273487	0.005335	-5.13	0.000
Cseasonal_1	U	0.0322220	0.005288	6.09	0.000
Cseasonal_4	U	-0.0562324	0.005657	-9.94	0.000
Cseasonal_5	U	-0.0752986	0.006333	-11.9	0.000
Cseasonal_8	U	0.0523600	0.006420	8.16	0.000
L200111_TCR	U	0.0361562	0.01163	3.11	0.003
I200210_TCR	U	0.0138667	0.01379	1.01	0.319
I200211_TCR	U	0.0137111	0.01386	0.989	0.327
I200302_TCR	U	0.0163326	0.01319	1.24	0.221
L200205_TCR	U	-0.0364746	0.008446	-4.32	0.000
L200301_TCR	U	0.0338064	0.007439	4.54	0.000
L200504_TCR	U	-0.00888087	0.006052	-1.47	0.148
Constant	U	1.19934	0.3078	3.90	0.000
I200604_PIB	U	-0.0498205	0.01201	-4.15	0.000
I200009_PIB	U	-0.0585115	0.01310	-4.47	0.000
I200106_PIB	U	-0.0463764	0.01235	-3.75	0.000
I200506_PIB	U	0.0348202	0.01240	2.81	0.007

sigma = 0.0111859 RSS = 0.006506437922

URF equation for: LTCR

		Coefficient	Std.Error	t-value	t-prob
Lii_1		-0.0365513	0.03875	-0.943	0.350
Lii_2		-0.0283316	0.04321	-0.656	0.515
Lii_3		-0.104943	0.04358	-2.41	0.020
Lpib_I_1		0.214498	0.1409	1.52	0.134
Lpib_I_2		-0.0371178	0.1643	-0.226	0.822
Lpib_I_3		0.364510	0.1242	2.94	0.005
LTCR_1		1.06612	0.09722	11.0	0.000
LTCR_2		-0.329339	0.1296	-2.54	0.014
LTCR_3		0.270182	0.1018	2.65	0.011
Cseasonal	U	-0.00495812	0.01156	-0.429	0.670
Cseasonal_1	U	0.000234903	0.01146	0.0205	0.984
Cseasonal_4	U	-0.0102686	0.01226	-0.838	0.406
Cseasonal_5	U	-0.0156382	0.01372	-1.14	0.260
Cseasonal_8	U	0.0264307	0.01391	1.90	0.063
L200111_TCR	U	-0.112445	0.02520	-4.46	0.000
I200210_TCR	U	0.0906884	0.02988	3.03	0.004
I200211_TCR	U	-0.122526	0.03003	-4.08	0.000
I200302_TCR	U	0.113465	0.02858	3.97	0.000
L200205_TCR	U	0.0750064	0.01830	4.10	0.000

L200301_TCR	U	-0.0861831	0.01612	-5.35	0.000
L200504_TCR	U	-0.0577825	0.01311	-4.41	0.000
Constant	U	-1.52017	0.6669	-2.28	0.027
I200604_PIB	U	0.00554886	0.02602	0.213	0.832
I200009_PIB	U	-0.0346386	0.02839	-1.22	0.228
I200106_PIB	U	0.00487994	0.02676	0.182	0.856
I200506_PIB	U	0.00838953	0.02687	0.312	0.756

sigma = 0.0242335 RSS = 0.03053752826

Mod. n° 12: DLII/DLPIB/DLTCR

Identity for Cia

Dlii	1.0000
Dlpib_I	1.8601
DLTCR	4.0556
Cia_1	1.0000

R^2 = 1 over 2000 (7) to 2006 (12)

Identity for Cib

Dlii	-0.55065
Dlpib_I	1.0000
DLTCR	-0.32288
Cib_1	1.0000

R^2 = 1 over 2000 (7) to 2006 (12)
MaxBFGS(): initial Hessian dimension error, using I

MOD(23) Estimating the model by FIML (using II.in7)
The estimation sample is: 2000 (7) to 2006 (12)

Equation for: Dlii

	Coefficient	Std.Error	t-value	t-prob	
Dlii_1	-0.458070	0.09413	-4.87	0.000	
Dlpib_I_1	-1.24460	0.3155	-3.94	0.000	
DLTCR_1	1.05062	0.2415	4.35	0.000	
Dlii_2	-0.311865	0.1043	-2.99	0.004	
Dlpib_I_2	-1.00792	0.2912	-3.46	0.001	
DLTCR_2	0.407871	0.2384	1.71	0.093	
Cib_1	0.750376	0.1111	6.75	0.000	
L200111_TCR	U	-0.128622	0.04076	-3.16	0.003
I200210_TCR	U	0.0886177	0.07876	1.13	0.266
I200211_TCR	U	-0.0630974	0.07753	-0.814	0.419
I200302_TCR	U	0.297866	0.07531	3.96	0.000
L200205_TCR	U	-0.0162127	0.04679	-0.346	0.730
L200301_TCR	U	-0.00946563	0.03435	-0.276	0.784
L200504_TCR	U	-0.127113	0.02898	-4.39	0.000
I200604_PIB	U	-0.128276	0.06853	-1.87	0.067
I200009_PIB	U	-0.0123375	0.07424	-0.166	0.869
I200106_PIB	U	-0.0654939	0.06973	-0.939	0.352
I200506_PIB	U	0.118485	0.06971	1.70	0.095
Constant	U	0.0924870	0.02491	3.71	0.000
Cseasonal	U	-0.0494056	0.02995	-1.65	0.105
Cseasonal_1	U	-0.0141260	0.02993	-0.472	0.639
Cseasonal_4	U	-0.121844	0.03038	-4.01	0.000
Cseasonal_5	U	-0.321738	0.03539	-9.09	0.000
Cseasonal_8	U	0.111826	0.03370	3.32	0.002

sigma = 0.0640475

Equation for: Dlpib_I

	Coefficient	Std.Error	t-value	t-prob	
Dlii_1	-0.0351610	0.01436	-2.45	0.018	
Dlpib_I_1	-0.244101	0.05524	-4.42	0.000	
DLTCR_1	0.166263	0.04223	3.94	0.000	
Dlpib_I_2	-0.443402	0.04199	-10.6	0.000	
Cia_1	-0.0199596	0.003727	-5.36	0.000	
L200111_TCR	U	0.0204088	0.006419	3.18	0.002
I200210_TCR	U	0.0155183	0.01366	1.14	0.261
I200211_TCR	U	0.0104576	0.01387	0.754	0.454
I200302_TCR	U	0.0188144	0.01310	1.44	0.157
L200205_TCR	U	-0.0396294	0.007845	-5.05	0.000
L200301_TCR	U	0.0262633	0.005597	4.69	0.000
L200504_TCR	U	-0.0116757	0.004426	-2.64	0.011
I200604_PIB	U	-0.0491740	0.01199	-4.10	0.000
I200009_PIB	U	-0.0587189	0.01314	-4.47	0.000
I200106_PIB	U	-0.0477259	0.01209	-3.95	0.000

I200506_PIB	U	0.0337268	0.01200	2.81	0.007
Constant	U	0.673619	0.1255	5.37	0.000
Cseasonal	U	-0.0298906	0.005097	-5.86	0.000
Cseasonal_1	U	0.0333314	0.005267	6.33	0.000
Cseasonal_4	U	-0.0570017	0.005232	-10.9	0.000
Cseasonal_5	U	-0.0746891	0.006121	-12.2	0.000
Cseasonal_8	U	0.0532679	0.005510	9.67	0.000

sigma = 0.011328

Equation for: DLTCR

		Coefficient	Std.Error	t-value	t-prob
Dlii_1		0.162040	0.03601	4.50	0.000
Dlpib_I_1		-0.314005	0.1249	-2.51	0.015
Dlii_2		0.120971	0.03819	3.17	0.003
Dlpib_I_2		-0.361954	0.1132	-3.20	0.002
DLTCR_2		-0.284936	0.09685	-2.94	0.005
Cia_1		0.0476490	0.01751	2.72	0.009
Cib_1		0.462839	0.08440	5.48	0.000
L200111_TCR	U	-0.127960	0.01707	-7.49	0.000
I200210_TCR	U	0.0921709	0.02907	3.17	0.003
I200211_TCR	U	-0.122748	0.02906	-4.22	0.000
I200302_TCR	U	0.111249	0.02780	4.00	0.000
L200205_TCR	U	0.0753783	0.01740	4.33	0.000
L200301_TCR	U	-0.0942970	0.01273	-7.41	0.000
L200504_TCR	U	-0.0524138	0.01074	-4.88	0.000
I200604_PIB	U	0.00483489	0.02546	0.190	0.850
I200009_PIB	U	-0.0331463	0.02723	-1.22	0.229
I200106_PIB	U	0.00462230	0.02620	0.176	0.861
I200506_PIB	U	0.00280221	0.02551	0.110	0.913
Constant	U	-1.51169	0.5779	-2.62	0.012
Cseasonal	U	-0.00492178	0.01129	-0.436	0.665
Cseasonal_1	U	0.00131541	0.01116	0.118	0.907
Cseasonal_4	U	-0.00721782	0.01142	-0.632	0.530
Cseasonal_5	U	-0.0132133	0.01312	-1.01	0.318
Cseasonal_8	U	0.0227030	0.01288	1.76	0.084

sigma = 0.0237405

IRPJ DOLR - Tabela 18

Mod. nº 1: LLR/LPIB

SYS(5) Estimating the system by OLS (using IRPJ.in7)
The estimation sample is: 2001 (2) to 2006 (12)

URF equation for: Llr

		Coefficient	Std.Error	t-value	t-prob
Llr_1		-0.0478216	0.1080	-0.443	0.660
Llr_2		-0.0488244	0.1004	-0.486	0.629
Llr_3		-0.116425	0.09918	-1.17	0.246
Llr_4		-0.153845	0.09719	-1.58	0.120
Lpib_I_1		-0.283081	0.7657	-0.370	0.713
Lpib_I_2		1.18378	0.8991	1.32	0.194
Lpib_I_3		-0.610909	0.7869	-0.776	0.441
Lpib_I_4		-0.462553	0.6662	-0.694	0.491
Trend		0.0124867	0.002342	5.33	0.000
Constant	U	7.98188	2.839	2.81	0.007
Cseasonal_1	U	0.141691	0.06457	2.19	0.033
Cseasonal_2	U	0.0595474	0.06194	0.961	0.341
Cseasonal_4	U	0.667295	0.07108	9.39	0.000
Cseasonal_5	U	0.193203	0.09997	1.93	0.059
Cseasonal_6	U	0.419907	0.1088	3.86	0.000
Cseasonal_7	U	0.450268	0.1059	4.25	0.000
Cseasonal_8	U	0.0835519	0.09757	0.856	0.396
Cseasonal_10	U	0.183708	0.06602	2.78	0.008
I200305_LR	U	0.584318	0.1316	4.44	0.000
L200408_LR	U	0.454159	0.08673	5.24	0.000
I200601_LR	U	-0.431338	0.1327	-3.25	0.002
I200106_PIB	U	0.0708151	0.1243	0.570	0.571

sigma = 0.113452 RSS = 0.6307009071

URF equation for: Lpib_I

		Coefficient	Std.Error	t-value	t-prob
Llr_1		-0.00620925	0.01428	-0.435	0.666

Llr_2		0.00280418	0.01327	0.211	0.834
Llr_3		0.0131209	0.01311	1.00	0.322
Llr_4		0.0276106	0.01285	2.15	0.037
Lpib_I_1		0.696448	0.1012	6.88	0.000
Lpib_I_2		0.204617	0.1189	1.72	0.091
Lpib_I_3		0.361539	0.1040	3.47	0.001
Lpib_I_4		-0.395265	0.08808	-4.49	0.000
Trend		-0.000136296	0.0003096	-0.440	0.662
Constant	U	0.420721	0.3753	1.12	0.268
Cseasonal_1	U	0.0571968	0.008537	6.70	0.000
Cseasonal_2	U	0.0316580	0.008189	3.87	0.000
Cseasonal_4	U	-0.0455979	0.009398	-4.85	0.000
Cseasonal_5	U	-0.0421076	0.01322	-3.19	0.003
Cseasonal_6	U	0.0527981	0.01438	3.67	0.001
Cseasonal_7	U	0.0353066	0.01400	2.52	0.015
Cseasonal_8	U	0.0494456	0.01290	3.83	0.000
Cseasonal_10	U	-0.0140519	0.008728	-1.61	0.114
I200305_LR	U	-0.0111174	0.01740	-0.639	0.526
L200408_LR	U	-0.00334903	0.01147	-0.292	0.771
I200601_LR	U	-0.0143750	0.01755	-0.819	0.417
I200106_PIB	U	-0.0628172	0.01643	-3.82	0.000

sigma = 0.0150002 RSS = 0.01102526519

Mod. n° 3: DLLR/DLPPIB

Identity for Cia

DLlr	1.0000
Dlpib_i	-0.097025
Cia_1	1.0000
Constant	-0.0088021

R² = 1 over 2001 (2) to 2006 (12)

MaxBFGS(): initial Hessian dimension error, using I

MOD(17) Estimating the model by FIML (using IRPJ.in7)
The estimation sample is: 2001 (2) to 2006 (12)

Equation for: DLlr

	Coefficient	Std.Error	t-value	t-prob
DLlr_1	0.307211	0.1366	2.25	0.029
DLlr_2	0.269282	0.1110	2.43	0.019
DLlr_3	0.173232	0.09057	1.91	0.061
Cia_1	-1.35480	0.1648	-8.22	0.000
Constant	6.52259	0.7937	8.22	0.000
I200305_LR	U 0.583679	0.1206	4.84	0.000
L200408_LR	U 0.443695	0.05584	7.95	0.000
I200601_LR	U -0.437459	0.1245	-3.51	0.001
I200106_PIB	U 0.0668626	0.1143	0.585	0.561
Cseasonal_1	U 0.142201	0.05551	2.56	0.013
Cseasonal_2	U 0.0193706	0.05274	0.367	0.715
Cseasonal_4	U 0.686599	0.06187	11.1	0.000
Cseasonal_5	U 0.185653	0.07922	2.34	0.023
Cseasonal_6	U 0.376525	0.06869	5.48	0.000
Cseasonal_7	U 0.371651	0.07291	5.10	0.000
Cseasonal_8	U 0.120549	0.08043	1.50	0.140
Cseasonal_10	U 0.217184	0.05798	3.75	0.000

sigma = 0.108948

Equation for: Dlpib_I

	Coefficient	Std.Error	t-value	t-prob
Dlpib_I_1	-0.192633	0.09045	-2.13	0.038
Dlpib_I_3	0.448136	0.07462	6.01	0.000
I200305_LR	U -0.00496719	0.01707	-0.291	0.772
L200408_LR	U 0.00476507	0.003011	1.58	0.120
I200601_LR	U -0.0189278	0.01783	-1.06	0.293
I200106_PIB	U -0.0609920	0.01623	-3.76	0.000
Cseasonal_1	U 0.0564653	0.008023	7.04	0.000
Cseasonal_2	U 0.0280278	0.007748	3.62	0.001
Cseasonal_4	U -0.0464682	0.009449	-4.92	0.000
Cseasonal_5	U -0.0372427	0.008577	-4.34	0.000
Cseasonal_6	U 0.0634662	0.008772	7.24	0.000
Cseasonal_7	U 0.0483759	0.008867	5.46	0.000
Cseasonal_8	U 0.0713001	0.008678	8.22	0.000
Cseasonal_10	U -0.00244821	0.007367	-0.332	0.741

sigma = 0.0156074

Mod. n° 4: LLR/LIVV

SYS(13) Estimating the system by OLS (using IRPJ.in7)
The estimation sample is: 2001 (2) to 2006 (12)

URF equation for: Llr

		Coefficient	Std.Error	t-value	t-prob
Llr_1		0.186134	0.1077	1.73	0.090
Llr_2		0.163213	0.1046	1.56	0.125
Llr_3		0.0589420	0.1072	0.550	0.585
Llr_4		-0.135525	0.1132	-1.20	0.237
Llr_5		-0.0659376	0.08819	-0.748	0.458
Llr_6		0.312003	0.1038	3.01	0.004
LIVV_sp_1		-0.335269	0.5325	-0.630	0.532
LIVV_sp_2		0.102060	0.5541	0.184	0.855
LIVV_sp_3		0.477173	0.2110	2.26	0.028
LIVV_sp_4		0.418896	0.2677	1.56	0.124
LIVV_sp_5		0.0699693	0.2500	0.280	0.781
LIVV_sp_6		-0.206906	0.2390	-0.866	0.390
Trend		0.00723415	0.002426	2.98	0.004
I200305_LR	U	0.598319	0.1470	4.07	0.000
Cseasonal_1	U	0.0891475	0.07578	1.18	0.245
Cseasonal_3	U	0.0495473	0.08661	0.572	0.570
Cseasonal_4	U	0.642215	0.1901	3.38	0.001
Cseasonal_5	U	-0.0381322	0.2462	-0.155	0.878

sigma = 0.12974 RSS = 0.8921189817

URF equation for: LIVV_sp

		Coefficient	Std.Error	t-value	t-prob
Llr_1		-0.00169268	0.02017	-0.0839	0.933
Llr_2		-0.0242297	0.01958	-1.24	0.221
Llr_3		0.0321593	0.02007	1.60	0.115
Llr_4		0.0258626	0.02119	1.22	0.228
Llr_5		-0.0252786	0.01651	-1.53	0.132
Llr_6		0.00884804	0.01942	0.456	0.651
LIVV_sp_1		0.585070	0.09966	5.87	0.000
LIVV_sp_2		0.153977	0.1037	1.48	0.144
LIVV_sp_3		0.251736	0.03949	6.38	0.000
LIVV_sp_4		-0.0431763	0.05011	-0.862	0.393
LIVV_sp_5		0.127501	0.04679	2.73	0.009
LIVV_sp_6		-0.0940600	0.04473	-2.10	0.040
Trend		2.02468e-005	0.0004541	0.0446	0.965
I200305_LR	U	0.00272540	0.02751	0.0991	0.921
Cseasonal_1	U	0.0468222	0.01418	3.30	0.002
Cseasonal_3	U	0.318097	0.01621	19.6	0.000
Cseasonal_4	U	-0.260959	0.03558	-7.34	0.000
Cseasonal_5	U	-0.131310	0.04608	-2.85	0.006

sigma = 0.0242827 RSS = 0.03125140892

Mod. n° 5: DLLR/DLIVV

SYS(19) Estimating the system by OLS (using IRPJ.in7)
The estimation sample is: 2001 (2) to 2006 (12)

URF equation for: DLlr

		Coefficient	Std.Error	t-value	t-prob
DLlr_1		-0.668384	0.08954	-7.46	0.000
DLlr_2		-0.362914	0.09721	-3.73	0.000
DLlr_3		-0.213627	0.1157	-1.85	0.070
DLlr_4		-0.330707	0.1038	-3.19	0.002
DLlr_5		-0.351026	0.09044	-3.88	0.000
DLIVV_sp_1		0.204031	0.3353	0.609	0.545
DLIVV_sp_2		0.102752	0.3721	0.276	0.783
DLIVV_sp_3		0.504084	0.3469	1.45	0.152
DLIVV_sp_4		0.654205	0.2856	2.29	0.026
DLIVV_sp_5		0.519594	0.2273	2.29	0.026
Cseasonal_1	U	0.126335	0.07687	1.64	0.106
Cseasonal_3	U	0.0602958	0.07937	0.760	0.451
Cseasonal_4	U	0.485882	0.1482	3.28	0.002
I200305_LR	U	0.651803	0.1521	4.29	0.000

sigma = 0.136535 RSS = 1.062580192

URF equation for: DLIVV_sp

		Coefficient	Std.Error	t-value	t-prob
DLlr_1		-0.0287619	0.01716	-1.68	0.099
DLlr_2		-0.0502649	0.01863	-2.70	0.009
DLlr_3		-0.00738177	0.02217	-0.333	0.740
DLlr_4		0.0106687	0.01990	0.536	0.594
DLlr_5		-0.0249119	0.01734	-1.44	0.156
DLIVV_sp_1		-0.158430	0.06426	-2.47	0.017
DLIVV_sp_2		-0.272909	0.07132	-3.83	0.000
DLIVV_sp_3		0.0359651	0.06649	0.541	0.591
DLIVV_sp_4		-0.0771189	0.05474	-1.41	0.164
DLIVV_sp_5		0.116653	0.04357	2.68	0.010
Cseasonal_1	U	0.0530175	0.01473	3.60	0.001
Cseasonal_3	U	0.330341	0.01521	21.7	0.000
Cseasonal_4	U	-0.342526	0.02840	-12.1	0.000
I200305_LR	U	0.000441306	0.02915	0.0151	0.988

sigma = 0.0261701 RSS = 0.03903776996

Mod. n° 6: DLLR/DLIVV

MOD(22) Estimating the model by FIML (using IRPJ.in7)
The estimation sample is: 2001 (2) to 2006 (12)

Equation for: DLlr

		Coefficient	Std.Error	t-value	t-prob
DLlr_1		-0.702006	0.07333	-9.57	0.000
DLlr_2		-0.395245	0.08956	-4.41	0.000
DLlr_3		-0.242378	0.09454	-2.56	0.013
DLlr_4		-0.344311	0.09692	-3.55	0.001
DLlr_5		-0.352602	0.08366	-4.21	0.000
DLIVV_sp_4		0.604512	0.1893	3.19	0.002
DLIVV_sp_5		0.482024	0.1504	3.20	0.002
DLIVV_sp_3		0.385359	0.1617	2.38	0.020
Cseasonal_1	U	0.133115	0.07405	1.80	0.078
Cseasonal_3	U	0.0700382	0.06178	1.13	0.262
Cseasonal_4	U	0.561842	0.07252	7.75	0.000
I200305_LR	U	0.656437	0.1459	4.50	0.000

sigma = 0.133795

Equation for: DLIVV_sp

		Coefficient	Std.Error	t-value	t-prob
DLlr_2		-0.0348187	0.01416	-2.46	0.017
DLIVV_sp_1		-0.174090	0.03600	-4.84	0.000
DLIVV_sp_2		-0.319099	0.03025	-10.5	0.000
DLIVV_sp_5		0.0932753	0.02601	3.59	0.001
DLlr_5		-0.0397972	0.01290	-3.08	0.003
DLIVV_sp_4		-0.131215	0.03201	-4.10	0.000
Cseasonal_1	U	0.0460239	0.01219	3.77	0.000
Cseasonal_3	U	0.337343	0.01252	27.0	0.000
Cseasonal_4	U	-0.339034	0.01623	-20.9	0.000
I200305_LR	U	-0.00600897	0.02800	-0.215	0.831

sigma = 0.0262113

Mod. n° 7: LLR/LIPI

SYS(6) Estimating the system by OLS (using IRPJ.in7)
The estimation sample is: 2001 (2) to 2006 (12)

URF equation for: Llr

		Coefficient	Std.Error	t-value	t-prob
Llr_1		0.0120995	0.1088	0.111	0.912
Llr_2		0.0707877	0.1129	0.627	0.533
Llr_3		0.144452	0.1121	1.29	0.203
Llr_4		-0.0732126	0.09047	-0.809	0.422
LIPI_g_sp_1		0.621116	0.5147	1.21	0.233
LIPI_g_sp_2		1.34125	0.5134	2.61	0.012
LIPI_g_sp_3		0.0454344	0.4512	0.101	0.920
LIPI_g_sp_4		-0.314858	0.4538	-0.694	0.491
Trend		0.00777491	0.002297	3.39	0.001
Constant	U	-3.74087	1.948	-1.92	0.060
Cseasonal_2	U	-0.0965078	0.06841	-1.41	0.164
Cseasonal_3	U	-0.109681	0.07505	-1.46	0.150

Cseasonal_4	U	0.567747	0.1115	5.09	0.000
Cseasonal_5	U	0.318162	0.1507	2.11	0.039
Cseasonal_6	U	0.514029	0.1392	3.69	0.001
Cseasonal_7	U	0.332482	0.1311	2.54	0.014
Cseasonal_9	U	-0.137806	0.1022	-1.35	0.183
I200305_LR	U	0.672921	0.1508	4.46	0.000

sigma = 0.133311 RSS = 0.9419051834

URF equation for: LIPI_g_sp

		Coefficient	Std.Error	t-value	t-prob
Llr_1		0.0182613	0.02551	0.716	0.477
Llr_2		-0.0232743	0.02648	-0.879	0.383
Llr_3		0.0203523	0.02630	0.774	0.443
Llr_4		0.0456609	0.02122	2.15	0.036
LIPI_g_sp_1		0.447507	0.1207	3.71	0.001
LIPI_g_sp_2		0.230258	0.1204	1.91	0.061
LIPI_g_sp_3		0.277992	0.1058	2.63	0.011
LIPI_g_sp_4		-0.335052	0.1064	-3.15	0.003
Trend		0.000368637	0.0005387	0.684	0.497
Constant	U	1.48889	0.4570	3.26	0.002
Cseasonal_2	U	-0.0501514	0.01605	-3.13	0.003
Cseasonal_3	U	-0.133709	0.01760	-7.60	0.000
Cseasonal_4	U	-0.118709	0.02616	-4.54	0.000
Cseasonal_5	U	-0.0881836	0.03534	-2.50	0.016
Cseasonal_6	U	0.0979333	0.03264	3.00	0.004
Cseasonal_7	U	-0.0613944	0.03076	-2.00	0.051
Cseasonal_9	U	-0.0776830	0.02397	-3.24	0.002
I200305_LR	U	-0.0428194	0.03537	-1.21	0.231

sigma = 0.03127 RSS = 0.05182395683

Mod. n° 9: DLLR/DLIPI

Identity for Cia

DLlr	1.0000
DLIPI_g_sp	-2.7424
Cia_1	1.0000
Constant	-0.0065143

R² = 1 over 2001 (2) to 2006 (12)

MaxBFGS(): initial Hessian dimension error, using I

MOD(20) Estimating the model by FIML (using IRPJ.in7)
The estimation sample is: 2001 (2) to 2006 (12)

Equation for: DLlr

		Coefficient	Std.Error	t-value	t-prob
DLIPI_g_sp_1		-1.87092	0.4026	-4.65	0.000
Cia_1		-0.894092	0.08774	-10.2	0.000
Constant		-7.07447	0.6951	-10.2	0.000
I200305_LR	U	0.702555	0.1433	4.90	0.000
Cseasonal_2	U	-0.142585	0.06047	-2.36	0.022
Cseasonal_3	U	-0.139105	0.06264	-2.22	0.031
Cseasonal_4	U	0.561000	0.07949	7.06	0.000
Cseasonal_5	U	0.276099	0.09871	2.80	0.007
Cseasonal_6	U	0.601616	0.07152	8.41	0.000
Cseasonal_7	U	0.489217	0.09045	5.41	0.000
Cseasonal_9	U	-0.00949343	0.06730	-0.141	0.888

sigma = 0.132695

Equation for: DLIPI_g_sp

		Coefficient	Std.Error	t-value	t-prob
DLIPI_g_sp_1		-0.388495	0.1005	-3.87	0.000
DLlr_2		-0.0501636	0.02018	-2.49	0.016
DLlr_3		-0.0547567	0.01988	-2.75	0.008
DLIPI_g_sp_3		0.249659	0.08045	3.10	0.003
Cia_1		0.0536848	0.02210	2.43	0.018
Constant		0.429781	0.1750	2.46	0.017
I200305_LR	U	-0.0257082	0.03517	-0.731	0.468
Cseasonal_2	U	-0.0636688	0.01510	-4.22	0.000
Cseasonal_3	U	-0.142538	0.01603	-8.89	0.000
Cseasonal_4	U	-0.120010	0.01966	-6.10	0.000
Cseasonal_5	U	-0.0978789	0.02476	-3.95	0.000
Cseasonal_6	U	0.128541	0.02431	5.29	0.000
Cseasonal_7	U	-0.0284563	0.02822	-1.01	0.318

Cseasonal_9 U -0.0445373 0.01987 -2.24 0.029

sigma = 0.0320042

COFINS DE - Tabela 23

Mod. nº 1: LCF/LPIB

SYS(8) Estimating the system by OLS (using Cofins.in7)
The estimation sample is: 2001 (2) to 2006 (12)

URF equation for: Lcf

		Coefficient	Std.Error	t-value	t-prob
Lcf_1		0.367658	0.1322	2.78	0.007
Lcf_2		-0.0597019	0.1371	-0.435	0.665
Lcf_3		-0.0535394	0.1406	-0.381	0.705
Lcf_4		0.0208540	0.1085	0.192	0.848
Lpib_i_1		0.726241	0.2889	2.51	0.015
Lpib_i_2		-0.344619	0.3964	-0.869	0.389
Lpib_i_3		0.286576	0.3405	0.842	0.404
Lpib_i_4		0.417456	0.3093	1.35	0.183
CSeasonal_1	U	-0.0382514	0.02930	-1.31	0.197
CSeasonal_2	U	-0.00492771	0.02737	-0.180	0.858
CSeasonal_4	U	0.0825157	0.02733	3.02	0.004
CSeasonal_5	U	-0.136914	0.03291	-4.16	0.000
CSeasonal_6	U	-0.0354507	0.03997	-0.887	0.379
CSeasonal_7	U	0.000865202	0.03374	0.0256	0.980
CSeasonal_8	U	-0.0294371	0.03604	-0.817	0.418
CSeasonal_10	U	-0.0631539	0.02513	-2.51	0.015
L200407_CF	U	-0.148826	0.02963	-5.02	0.000
I200106_PIB	U	0.0699028	0.05267	1.33	0.190

sigma = 0.0482198 RSS = 0.1232327476

URF equation for: Lpib_i

		Coefficient	Std.Error	t-value	t-prob
Lcf_1		-0.0278660	0.04391	-0.635	0.528
Lcf_2		0.0399031	0.04555	0.876	0.385
Lcf_3		-0.0139493	0.04672	-0.299	0.766
Lcf_4		-0.000165631	0.03603	-0.00460	0.996
Lpib_i_1		0.795374	0.09597	8.29	0.000
Lpib_i_2		0.214551	0.1317	1.63	0.109
Lpib_i_3		0.404488	0.1131	3.58	0.001
Lpib_i_4		-0.410704	0.1028	-4.00	0.000
CSeasonal_1	U	0.0554620	0.009733	5.70	0.000
CSeasonal_2	U	0.0262619	0.009093	2.89	0.006
CSeasonal_4	U	-0.0503644	0.009078	-5.55	0.000
CSeasonal_5	U	-0.0369959	0.01093	-3.38	0.001
CSeasonal_6	U	0.0549647	0.01328	4.14	0.000
CSeasonal_7	U	0.0493760	0.01121	4.41	0.000
CSeasonal_8	U	0.0693655	0.01197	5.79	0.000
CSeasonal_10	U	-0.00292313	0.008348	-0.350	0.728
L200407_CF	U	-0.000100642	0.009842	-0.0102	0.992
I200106_PIB	U	-0.0658324	0.01749	-3.76	0.000

sigma = 0.0160167 RSS = 0.01359629003

Mod. nº 3: DLCF/DLPIB

Identity for CIa

DLcf 1.0000

DLpib_i -1.4976

CIa_1 1.0000

R² = 1 over 2001 (2) to 2006 (12)

MaxBFGS(): initial Hessian dimension error, using I

MOD(20) Estimating the model by FIML (using Cofins.in7)
The estimation sample is: 2001 (2) to 2006 (12)

Equation for: DLcf

		Coefficient	Std.Error	t-value	t-prob
DLpib_i_2		-0.438413	0.2536	-1.73	0.090
CIa_1		-0.599512	0.09996	-6.00	0.000
L200407_CF	U	-0.124634	0.02293	-5.44	0.000
I200106_PIB	U	0.0469256	0.04799	0.978	0.332

CSeasonal_1	U	-0.0221898	0.02449	-0.906	0.369
CSeasonal_2	U	-0.0100401	0.02315	-0.434	0.666
CSeasonal_4	U	0.0762209	0.02425	3.14	0.003
CSeasonal_5	U	-0.128410	0.03039	-4.23	0.000
CSeasonal_6	U	-0.0142376	0.02574	-0.553	0.582
CSeasonal_7	U	0.0166093	0.02462	0.675	0.503
CSeasonal_8	U	-0.0128046	0.02455	-0.521	0.604
CSeasonal_10	U	-0.0669229	0.02327	-2.88	0.006

sigma = 0.047275

Equation for: DLpib_i

		Coefficient	Std.Error	t-value	t-prob
DLpib_i_1		-0.197842	0.08835	-2.24	0.029
DLpib_i_3		0.473345	0.07285	6.50	0.000
L200407_Cf	U	0.00303423	0.002865	1.06	0.294
I200106_PIB	U	-0.0623159	0.01609	-3.87	0.000
CSeasonal_1	U	0.0569222	0.007935	7.17	0.000
CSeasonal_2	U	0.0283398	0.007665	3.70	0.001
CSeasonal_4	U	-0.0513672	0.008439	-6.09	0.000
CSeasonal_5	U	-0.0376724	0.008460	-4.45	0.000
CSeasonal_6	U	0.0637847	0.008656	7.37	0.000
CSeasonal_7	U	0.0496917	0.008750	5.68	0.000
CSeasonal_8	U	0.0714944	0.008007	8.93	0.000
CSeasonal_10	U	-0.00315915	0.007301	-0.433	0.667

sigma = 0.0154776

Mod. n° 4: LCF/LIVV

SYS(6) Estimating the system by OLS (using Cofins.in7)
The estimation sample is: 2001 (2) to 2006 (12)

URF equation for: Lcf

		Coefficient	Std.Error	t-value	t-prob
Lcf_1		0.250995	0.1492	1.68	0.098
Lcf_2		-0.0363537	0.1436	-0.253	0.801
Lcf_3		-0.0318785	0.1450	-0.220	0.827
Lcf_4		-0.0290019	0.1275	-0.227	0.821
Lcf_5		-0.0730217	0.1069	-0.683	0.497
LIVV_sp_1		0.925799	0.2477	3.74	0.000
LIVV_sp_2		-0.252653	0.1067	-2.37	0.022
LIVV_sp_3		0.563798	0.2226	2.53	0.014
LIVV_sp_4		0.0251263	0.09512	0.264	0.793
LIVV_sp_5		0.116632	0.09431	1.24	0.222
Trend		0.000974970	0.0006196	1.57	0.122
CSeasonal	U	0.0631030	0.02675	2.36	0.022
CSeasonal_1	U	0.00923552	0.02828	0.327	0.745
CSeasonal_3	U	0.0297471	0.02839	1.05	0.300
CSeasonal_4	U	-0.215130	0.07852	-2.74	0.008
CSeasonal_6	U	-0.162950	0.09431	-1.73	0.090
CSeasonal_9	U	-0.0167447	0.03426	-0.489	0.627
CSeasonal_10	U	0.000621646	0.02685	0.0232	0.982
L200407_Cf	U	-0.223405	0.04476	-4.99	0.000

sigma = 0.0495167 RSS = 0.1274989199

URF equation for: LIVV_sp

		Coefficient	Std.Error	t-value	t-prob
Lcf_1		0.0399709	0.06526	0.612	0.543
Lcf_2		-0.0331567	0.06281	-0.528	0.600
Lcf_3		0.135486	0.06342	2.14	0.037
Lcf_4		-0.0869944	0.05578	-1.56	0.125
Lcf_5		0.00153477	0.04674	0.0328	0.974
LIVV_sp_1		0.566525	0.1083	5.23	0.000
LIVV_sp_2		-0.252594	0.04667	-5.41	0.000
LIVV_sp_3		0.460355	0.09737	4.73	0.000
LIVV_sp_4		-0.134547	0.04161	-3.23	0.002
LIVV_sp_5		0.270310	0.04126	6.55	0.000
Trend		0.000622403	0.0002710	2.30	0.026
CSeasonal	U	-0.0132458	0.01170	-1.13	0.263
CSeasonal_1	U	0.0591023	0.01237	4.78	0.000
CSeasonal_3	U	0.331381	0.01242	26.7	0.000
CSeasonal_4	U	-0.243640	0.03435	-7.09	0.000
CSeasonal_6	U	-0.103803	0.04126	-2.52	0.015
CSeasonal_9	U	-0.0478071	0.01499	-3.19	0.002

```
CSeasonal_10  U    0.0467523    0.01175    3.98    0.000
L200407_CF    U   -0.00364473    0.01958   -0.186    0.853
```

sigma = 0.0216599 RSS = 0.02439583701

Mod. n° 6: DLCF/DLIVV

Identity for CIa

```
DLcf          1.0000
DLIVV_sp     -1.5072
CIa_1        1.0000
```

R^2 = 1 over 2001 (2) to 2006 (12)

MaxBFGS(): initial Hessian dimension error, using I

MOD(18) Estimating the model by FIML (using Cofins.in7)
The estimation sample is: 2001 (2) to 2006 (12)

Equation for: DLcf

		Coefficient	Std.Error	t-value	t-prob
DLIVV_sp_2		-0.235545	0.09602	-2.45	0.017
CIa_1		-0.736425	0.1136	-6.48	0.000
L200407_CF	U	-0.154795	0.02574	-6.01	0.000
CSeasonal	U	0.0459258	0.02200	2.09	0.042
CSeasonal_1	U	0.0122816	0.02556	0.480	0.633
CSeasonal_3	U	0.0180076	0.02201	0.818	0.417
CSeasonal_4	U	-0.265013	0.05184	-5.11	0.000
CSeasonal_6	U	-0.0360298	0.05138	-0.701	0.486
CSeasonal_9	U	-0.0268396	0.02247	-1.19	0.238
CSeasonal_10	U	-0.00943288	0.02408	-0.392	0.697

sigma = 0.0489958

Equation for: DLIVV_sp

		Coefficient	Std.Error	t-value	t-prob
DLIVV_sp_1		-0.313106	0.07858	-3.98	0.000
DLIVV_sp_2		-0.531690	0.09439	-5.63	0.000
DLcf_3		0.110954	0.04270	2.60	0.012
DLIVV_sp_3		-0.149659	0.03944	-3.79	0.000
DLIVV_sp_4		-0.266178	0.03546	-7.51	0.000
L200407_CF	U	0.00751546	0.004050	1.86	0.069
CSeasonal	U	-0.0135350	0.01038	-1.30	0.198
CSeasonal_1	U	0.0649980	0.01047	6.21	0.000
CSeasonal_3	U	0.330636	0.01091	30.3	0.000
CSeasonal_4	U	-0.276146	0.02637	-10.5	0.000
CSeasonal_6	U	-0.0682377	0.03665	-1.86	0.068
CSeasonal_9	U	-0.0492402	0.01141	-4.32	0.000
CSeasonal_10	U	0.0491053	0.01024	4.80	0.000

sigma = 0.02158

Mod. n° 7: LCF/LIPI

SYS(12) Estimating the system by OLS (using Cofins.in7)
The estimation sample is: 2001 (2) to 2006 (12)

URF equation for: Lcf

		Coefficient	Std.Error	t-value	t-prob
Lcf_1		0.486669	0.1318	3.69	0.001
Lcf_2		0.0274924	0.1222	0.225	0.823
Lcf_3		-0.224477	0.1295	-1.73	0.089
Lcf_4		0.165570	0.09087	1.82	0.074
LIPI_g_sp_1		0.762810	0.1686	4.52	0.000
LIPI_g_sp_2		-0.459709	0.2077	-2.21	0.031
LIPI_g_sp_3		0.117125	0.1797	0.652	0.517
LIPI_g_sp_4		0.305592	0.1569	1.95	0.057
Trend		0.000369814	0.0006045	0.612	0.543
Constant	U	0.296884	0.9883	0.300	0.765
CSeasonal_1	U	-0.0598246	0.02513	-2.38	0.021
CSeasonal_3	U	0.0111527	0.02783	0.401	0.690
CSeasonal_4	U	0.129053	0.03902	3.31	0.002
CSeasonal_5	U	-0.127073	0.04618	-2.75	0.008
CSeasonal_7	U	0.0305214	0.03896	0.783	0.437
CSeasonal_9	U	-0.00717102	0.03145	-0.228	0.821
L200407_CF	U	-0.161333	0.03894	-4.14	0.000

sigma = 0.0496347 RSS = 0.1330344314

```

URF equation for: LIPI_g_sp
      Coefficient  Std.Error  t-value  t-prob
Lcf_1          0.157041   0.08512   1.84    0.071
Lcf_2          0.0613852  0.07894   0.778   0.440
Lcf_3          0.0627834  0.08360   0.751   0.456
Lcf_4        -0.00545555  0.05869  -0.0930  0.926
LIPI_g_sp_1    0.209860   0.1089    1.93    0.059
LIPI_g_sp_2   -0.0722680   0.1341   -0.539   0.592
LIPI_g_sp_3    0.195143   0.1160    1.68    0.098
LIPI_g_sp_4   -0.338662   0.1014   -3.34    0.002
Trend          0.00102853  0.0003904  2.63    0.011
Constant      U    2.90099    0.6383    4.55    0.000
CSeasonal_1   U    0.0212063  0.01623    1.31    0.197
CSeasonal_3   U   -0.120379  0.01797   -6.70    0.000
CSeasonal_4   U   -0.133827  0.02520   -5.31    0.000
CSeasonal_5   U   -0.151234  0.02982   -5.07    0.000
CSeasonal_7   U   -0.0900380  0.02516   -3.58    0.001
CSeasonal_9   U   -0.0791025  0.02031   -3.89    0.000
L200407_CF    U    0.126352  0.02515    5.02    0.000

```

sigma = 0.0320544 RSS = 0.05548409676

Mod. n° 9: DLCF/DLIPI

```

Identity for CIa
DLcf          1.0000
DLIPI_g_sp   -2.4800
CIa_1         1.0000
Constant      0.0015061
R^2 = 1 over 2001 (2) to 2006 (12)
MaxBFGS(): initial Hessian dimension error, using I

MOD(19) Estimating the model by FIML (using Cofins.in7)
      The estimation sample is: 2001 (2) to 2006 (12)

```

```

Equation for: DLcf
      Coefficient  Std.Error  t-value  t-prob
DLIPI_g_sp_2  -0.571957   0.1176   -4.86    0.000
DLcf_3        -0.219491   0.07391  -2.97    0.004
DLIPI_g_sp_3  -0.368520   0.1277   -2.88    0.006
CIa_1         -0.310774   0.05332  -5.83    0.000
Constant      U   -1.53130   0.2631   -5.82    0.000
L200407_CF    U   -0.120421  0.02409   -5.00    0.000
CSeasonal_1   U   -0.0721685  0.02304   -3.13    0.003
CSeasonal_3   U  -0.00366270  0.02243  -0.163   0.871
CSeasonal_4   U    0.108634  0.02973    3.65    0.001
CSeasonal_5   U   -0.173591  0.03430   -5.06    0.000
CSeasonal_7   U    0.0341507  0.02309    1.48    0.145
CSeasonal_9   U    0.000778469  0.02711  0.0287   0.977

```

sigma = 0.0499918

```

Equation for: DLIPI_g_sp
      Coefficient  Std.Error  t-value  t-prob
DLcf_1        -0.104728   0.04261   -2.46    0.017
DLIPI_g_sp_3  0.232399   0.07691    3.02    0.004
CIa_1         0.341588   0.03291   10.4    0.000
Constant      U    1.69014   0.1624   10.4    0.000
L200407_CF    U    0.127842  0.01511    8.46    0.000
CSeasonal_1   U    0.00733607  0.01491    0.492   0.625
CSeasonal_3   U   -0.136641  0.01461   -9.35    0.000
CSeasonal_4   U   -0.159747  0.01870   -8.54    0.000
CSeasonal_5   U   -0.174964  0.01924   -9.09    0.000
CSeasonal_7   U   -0.0554589  0.01453   -3.82    0.000
CSeasonal_9   U   -0.0664086  0.01723   -3.85    0.000

```

sigma = 0.0322838

Especificação modelos estruturais - EM/EMX – II/IRPJ DOLR/COFINS DE - Tabela 28

Mod. nº 2: LII

Equation 6.

Lii = Level + Fixed seasonal + Expl vars + Irregular

Estimation report

Model with 2 parameters (1 restrictions).
 Parameter estimation sample is 2000. 7 - 2006.12. (T = 78).
 Log-likelihood kernel is 1.817765.
 Very strong convergence in 3 iterations.
 (likelihood cvg 6.107625e-016
 gradient cvg 0
 parameter cvg 9.474936e-010)

Eq 6 : Diagnostic summary report.

Estimation sample is 2000. 7 - 2006.12. (T = 78, n = 77).
 Log-Likelihood is 141.786 (-2 LogL = -283.571).
 Prediction error variance is 0.00707017

Summary statistics

	Lii
Std.Error	0.084084
Normality	0.35495
H(25)	0.51770
r(1)	-0.015132
r(7)	-0.019536
DW	2.0061
Q(7, 6)	3.6110
Rs^2	0.25042

Eq 6 : Estimated variances of disturbances.

Component	Lii (q-ratio)
Irr	0.0037657 (1.0000)
Lvl	0.0026242 (0.6969)

Eq 6 : Estimated standard deviations of disturbances.

Component	Lii (q-ratio)
Irr	0.061365 (1.0000)
Lvl	0.051227 (0.8348)

Eq 6 : Estimated coefficients of final state vector.

Variable	Coefficient	R.m.s.e.	t-value	
Lvl	3.9450	0.64964	6.0726	[0.0000]
Sea_ 1	0.062246	0.031329	1.9868	[0.0505]
Sea_ 2	0.085580	0.029922	2.8601	[0.0054]
Sea_ 3	0.032659	0.029820	1.0952	[0.2768]
Sea_ 4	0.085265	0.030220	2.8214	[0.0061]
Sea_ 5	0.043309	0.031562	1.3722	[0.1740]
Sea_ 6	-0.022062	0.030369	-0.72648	[0.4697]
Sea_ 7	-0.072724	0.033818	-2.1505	[0.0347]
Sea_ 8	-0.20928	0.032781	-6.384	[0.0000]
Sea_ 9	0.021205	0.031841	0.66596	[0.5074]
Sea_10	-0.048410	0.032204	-1.5032	[0.1369]
Sea_11	0.040791	0.038908	1.0484	[0.2977]

Eq 6 : Estimated coefficients of explanatory variables.

Variable	Coefficient	R.m.s.e.	t-value	
Lii_3	0.28880	0.11770	2.4536	[0.0164]

Mod. nº 4: LLR

Equation 1.

Llr = Trend + Trigo seasonal + Expl vars + Interv + Irregular

Estimation report

Model with 3 parameters (1 restrictions).
 Parameter estimation sample is 2001. 2 - 2006.12. (T = 71).
 Log-likelihood kernel is 1.289261.
 Very strong convergence in 7 iterations.
 (likelihood cvg 2.871012e-013
 gradient cvg 8.21565e-010
 parameter cvg 4.720598e-007)

Eq 1 : Diagnostic summary report.

Estimation sample is 2001. 2 - 2006.12. (T = 71, n = 58).
 Log-Likelihood is 91.5375 (-2 LogL = -183.075).
 Prediction error variance is 0.0173061

Summary statistics

	Llr
Std.Error	0.13155
Normality	1.3319
H(19)	1.8376
r(1)	0.058036
r(8)	-0.11684
DW	1.7970
Q(8, 6)	3.9194
Rs^2	0.44213

Eq 1 : Estimated variances of disturbances.

Component	Llr (q-ratio)
Irr	0.0071571 (1.0000)
Lvl	0.0018616 (0.2601)
Sea	5.3022e-005 (0.0074)

Eq 1 : Estimated standard deviations of disturbances.

Component	Llr (q-ratio)
Irr	0.084600 (1.0000)
Lvl	0.043146 (0.5100)
Sea	0.0072817 (0.0861)

Eq 1 : Estimated coefficients of final state vector.

Variable	Coefficient	R.m.s.e.	t-value	
Lvl	0.12860	3.4280	0.037514	[0.9702]
Slp	0.010286	0.0054410	1.8904	[0.0637]
Sea_ 1	0.026774	0.049373	0.54228	[0.5897]
Sea_ 2	0.051232	0.043996	1.1644	[0.2490]
Sea_ 3	-0.092420	0.031848	-2.9019	[0.0052]
Sea_ 4	-0.018936	0.058165	-0.32556	[0.7459]
Sea_ 5	0.10333	0.061222	1.6878	[0.0968]
Sea_ 6	0.021273	0.031817	0.6686	[0.5064]
Sea_ 7	-0.0023995	0.033361	-0.071925	[0.9429]
Sea_ 8	0.16192	0.048078	3.368	[0.0014]
Sea_ 9	-0.10239	0.046567	-2.1987	[0.0319]
Sea_10	0.077184	0.041822	1.8455	[0.0701]
Sea_11	-0.076390	0.029450	-2.5938	[0.0120]

Eq 1 : Estimated coefficients of explanatory variables.

Variable	Coefficient	R.m.s.e.	t-value	
LIVV_sp_2	1.2869	0.71583	1.7978	[0.0774]
Irr 2003. 5	0.71602	0.12504	5.7263	[0.0000]

Eq 1 : Seasonal analysis (at end of period).

Seasonal Chi^2(11) test is 53.2695 [0.0000].

Value	Seas 1	Seas 2	Seas 3	Seas 4	Seas 5
	0.35255	-0.34922	0.27355	0.24631	-0.11883
Value	Seas 6	Seas 7	Seas 8	Seas 9	Seas 10
	-0.19893	-0.042128	-0.021996	0.059268	0.0089867
Value	Seas 11	Seas 12			
	-0.066074	-0.14349			

Mod. n° 7: LCF

Equation 14.

Lcf = Trend + Fixed seasonal + Expl vars + Interv + Irregular

Estimation report

Model with 3 parameters (1 restrictions).

Parameter estimation sample is 2001. 2 - 2006.12. (T = 71).

Log-likelihood kernel is 2.130569.

Very strong convergence in 6 iterations.

```
( likelihood cvg 8.700158e-013
  gradient cvg 2.748912e-007
  parameter cvg 5.112257e-012 )
```

Eq 14 : Diagnostic summary report.

Estimation sample is 2001. 2 - 2006.12. (T = 71, n = 69).

Log-Likelihood is 151.27 (-2 LogL = -302.541).

Prediction error variance is 0.00213188

Summary statistics

	Lcf
Std.Error	0.046172
Normality	0.24776
H(23)	1.1607
r(1)	0.037747
r(8)	0.18230
DW	1.8832
Q(8, 6)	5.1960
Rs^2	0.28850

Eq 14 : Estimated variances of disturbances.

Component	Lcf (q-ratio)
Irr	0.0017193 (1.0000)
Lvl	0.00024131 (0.1404)
Slp	1.8800e-006 (0.0011)

Eq 14 : Estimated standard deviations of disturbances.

Component	Lcf (q-ratio)
Irr	0.041465 (1.0000)
Lvl	0.015534 (0.3746)
Slp	0.0013711 (0.0331)

Eq 14 : Estimated coefficients of final state vector.

Variable	Coefficient	R.m.s.e.	t-value	
Lvl	6.0915	0.84395	7.2178	[0.0000]
Slp	0.0044259	0.0051648	0.85694	[0.3944]
Sea_ 1	-0.10814	0.021970	-4.9222	[0.0000]
Sea_ 2	-0.073759	0.021017	-3.5095	[0.0008]
Sea_ 3	0.0011203	0.020478	0.054708	[0.9565]
Sea_ 4	-0.040150	0.017727	-2.2649	[0.0267]
Sea_ 5	-0.010742	0.018425	-0.58304	[0.5618]
Sea_ 6	-0.045512	0.017977	-2.5317	[0.0136]
Sea_ 7	0.022363	0.018570	1.2042	[0.2326]
Sea_ 8	0.059170	0.017659	3.3506	[0.0013]
Sea_ 9	-0.0053564	0.018843	-0.28427	[0.7771]
Sea_10	0.043501	0.017657	2.4637	[0.0162]
Sea_11	0.044988	0.018361	2.4502	[0.0168]

Eq 14 : Estimated coefficients of explanatory variables.

Variable	Coefficient	R.m.s.e.	t-value	
Lcf_1	0.13072	0.12047	1.085	[0.2817]
Lvl 2004. 7	-0.24325	0.051229	-4.7484	[0.0000]