SUMÁRIO

	LISTA DE TABELAS	2
	LISTA DE GRÁFICOS	3
1	INTRODUÇÃO	4
2	REVISÃO DE LITERATURA	6
2.1	Regras de Política Monetária	6
2.2	Regras de Política Monetária em Pequenas Economias Abertas	8
2.3	Avaliação de Regras Monetárias Aplicadas	11
3	METODOLOGIA	14
3.1	Modelos lineares univariados e multivariados	14
3.2	Modelos Multivariados com Markov Switching	16
3.3	Procedimentos estatísticos de estimação	19
3.4	Análise de Resposta a Impulso em modelos multivariados não-lineares	24
3.5	Metodologia *Markov Chain Monte Carlo* (MCMC)	28
4	SÉRIES ECONÔMICAS UTILIZADAS	37
4.1	Notas sobre a estacionariedade das séries utilizadas	37
5	APLICAÇÃO À POLÍTICA MONETÁRIA BRASILEIRA	46
5.1	Estimação para Economia Fechada	46
5.2	Estimação para Economia Aberta	59
5.3	Padrões de Resposta a Choques	68
	CONSIDERAÇÕES FINAIS	80
	REFERÊNCIAS	82
	ANEXO A	87

LISTA DE TABELAS

Tabela 1: Estatística descritiva das séries utilizadas 44
Tabela 2: Análise de raiz unitária 44
Tabela 3: Formas funcionais do modelo e critérios de seleção estimados 46
Tabela 4: Datas de cada regime e suas respectivas probabilidades de ocorrência 49
Tabela 5: Probabilidades estacionárias e duração dos regimes do modelo MSIAH (2) – VAR (1)..... 51
Tabela 6: Valores obtidos com a simulação de um modelo MSMA(2)-VAR(1) 54
Tabela 7: Parâmetros estimados a partir do modelo MSIAH (2) – VAR (1) para economia aberta 61
Tabela 8: Probabilidades estacionárias e duração dos regimes do modelo MSIAH (2) – VAR (1) 63
Tabela 9: Valores obtidos com a simulação de um modelo MSMA(2)-VAR(1) para a economia aberta 65
Tabela 10: Valores obtidos com a simulação de um modelo MSMA(2)-VAR(1) para a economia aberta 66

LISTA DE GRÁFICOS

Gráfico 1: Séries históricas das variáveis utilizadas 43
Gráfico 2: Probabilidades para os regimes do modelo MSIAH (2) – VAR (1) adotado à regra de política monetária no período do Plano Real 47
Gráfico 3: Probabilidades de transição preditas para cada regime $s_t = i$ 50
Gráfico 4: Probabilidades para os regimes do modelo MSMA (2) – VAR (1) adotado à regra de política monetária no período do Plano Real 53
Gráfico 5: Comportamento da amostragem após várias iterações do modelo para o regime de calmaria econômica 55
Gráfico 6: Comportamento da amostragem após várias iterações do modelo para o regime de crise econômica 56
Gráfico 7: Caminho da amostragem após várias iterações: Caso de calmaria econômica 57
Gráfico 8: Caminho da amostragem após várias iterações: Caso de crises econômicas 57
Gráfico 9: Densidades para a média das variáveis Selic e Selic(-1) 58
Gráfico 10: Comportamento da amostragem para correlações contemporâneas 59
Gráfico 11: Probabilidades para os regimes do modelo MSIAH (2) – VAR (1) adotado à regra de política monetária no período do Plano Real 60
Gráfico 12: Comportamento dos resíduos do modelo MSIAH (2) – VAR (1) estimado para a economia aberta 62
Gráfico 13: Probabilidades para os regimes do modelo MSMA (2) – VAR (1) adotado à regra de política monetária em uma economia aberta 63
Gráfico 14: Caminho da amostragem após várias iterações: Caso de calmaria econômica 66
Gráfico 15: Caminho da amostragem após várias iterações: Caso de crises econômicas 67
Gráfico 16: Densidades para a média das variáveis Selic e Selic(-1) 68
Gráfico 17: Funções de resposta a impulso, choque de 1% na taxa de juros nominal 70
Gráfico 18: Funções de resposta a impulso, choque de 1% na taxa de juros nominal 72
Gráfico 19: Funções de resposta a impulso, choque de 1% no hiato do produto 73
Gráfico 20: Funções de resposta a impulso, choque de 1% no hiato do produto 74
Gráfico 21: Funções de resposta a impulso, choque de 1% na taxa de inflação 76
Gráfico 22: Funções de resposta a impulso, choque de 1% na taxa de inflação 77
Gráfico 23: Funções de resposta a impulso, choque de 1% na taxa de câmbio 78
Gráfico 24: Funções de resposta a impulso, choque de 1% nas demais variáveis do modelo no período 1994-2005 79

1. INTRODUÇÃO

A partir da evolução da macroeconomia moderna nas últimas décadas, modelos estruturais cada vez mais complexos foram criados e métodos de solução tornaram-se muito mais potentes. A investigação de melhores políticas passa obrigatoriamente pela construção de modelos macroeconométricos baseados em microfundamentos e expectativas racionais.

A avaliação de políticas monetárias e seus instrumentos é geralmente feita por meio de modelos estruturais linearizados, que incluem uma regra de política monetária e onde o comportamento da Demanda Agregada é dado pela Equação de Euler e o da Oferta Agregada é dado pela Curva de Phillips. Entretanto, a formação das expectativas presentes nestas curvas e os coeficientes relativos à regra de política monetária são bastante sensíveis a mudanças no ambiente econômico, especialmente a mudanças na condução da política monetária. Devido à grande variabilidade da política monetária brasileira ao longo dos anos, a estimação baseada no modelo completo parece pouco atraente à medida que se torna mal especificada para um grande período de tempo, e seu poder de análise e sua confiabilidade ficam bastante comprometidos com tantas quebras de regimes monetários ao longo do tempo.

Deste modo, torna-se necessário modelar e estimar as equações que comporiam um modelo estrutural separadamente e de maneira independente da política monetária em qualquer período. No presente trabalho, se buscará modelar e estimar regras de política monetária que apresentem quebras estruturais através de metodologias não-lineares de estimação.

Serão modelados fenômenos econômicos que podem alterar o formato teórico original da curva de reação, tal como a inclusão da variável relativa ao câmbio nominal, mas que explicam características observáveis nos dados relativos à regra, como a flutuação cambial.

Nosso objetivo é estender o estudo sobre a dinâmica dos juros a partir de uma regra monetária. Partimos de uma motivação teórica onde justificamos a dinâmica não-linear a partir de dois modelos: no primeiro há um exercício *Markov Switching Vector AutoRegression,* MS (k) – VAR (p), em que a interpretação obtida é muito compatível com os dados brasileiros; já no segundo, explicitamos uma metodologia bayesiana de estimação, baseada na iteração de Gibbs (*Markov Chain Monte Carlo – MCMC*) e no algoritmo de Metropolis-Hasting para os mesmos modelos MS (k) – VAR (p) estimados na primeira seção. O objetivo é refletir da melhor maneira possível as relações existentes entre as variáveis do

sistema VAR(p) estimado, levando em consideração a mudança de regime associada à não-linearidade do comportamento destas relações.

A estimação de tal curva de reação do Banco Central, dentro do arcabouço novo-keynesiano através dos métodos estatísticos propostos, possibilitará no futuro uma melhor formulação e refino metodológico de modelos para a análise empírica de novas políticas monetárias e a investigação de resultados teóricos na realidade, que são ás vezes prejudicados pela ausência da aplicabilidade de métodos estatísticos mais robustos. As estimações realizadas neste trabalho demonstram que a economia brasileira realmente apresenta mudanças de regimes ao longo do Plano Real, tanto na estimação feita pelo algoritmo *Expectation-Maximization* (EM), quanto no método MCMC.

Esta dissertação inova, também, ao aplicar testes de raiz unitária baseados em testes PADF, segundo descrevem Lima e Xiao (2003). A presença de caudas grossas em algumas séries econômicas acaba por enfraquecer a hipótese nula de que há raiz unitária identificável para aquela série. Assim, vários trabalhos na literatura econômica têm discutido o problema da estacionariedade para algumas séries econômicas amostradas ao longo do Plano Real. Avaliamos com tal metodologia a presença ou não de raiz unitária, aperfeiçoando a forma de análise estatística. Ademais, aplicamos uma metodologia de análise das funções resposta a impulso que levasse em conta a dependência do choque atrelado a cada regime específico. As funções resposta a impulso generalizadas (GRI) provêem uma análise adequada para o MS (k) – VAR (P) estimado neste trabalho.

O trabalho está dividido da seguinte maneira: o capítulo 2 versa sobre a revisão de literatura acerca de regras de política monetária e sua avaliação; o capítulo 3 descreve a metodologia de Markov Switching, o desenvolvimento de funções resposta a impulso generalizadas (GRI), dependentes de cada regime e a metodologia de Markov Chain Monte Carlo (MCMC); o capítulo 4 apresenta uma discussão metodológica e os resultados obtidos na avaliação da estacionariedade das séries econômicas utilizadas neste trabalho; o capítulo 5 apresenta os resultados obtidos para a economia brasileira sendo utilizados os dois algoritmos propostos (EM e MCMC), além de toda avaliação dos mecanismos de transmissão da política monetária abordados pelas funções resposta a impulso.

2. REVISÃO DE LITERATURA

2.1. Regras de Política Monetária

O monetarismo de Milton Friedman (1968) propunha que a autoridade monetária deveria ter como meta uma taxa de crescimento de um agregado monetário sempre constante, igualando-a à taxa de crescimento do produto real. Para isto, o autor se baseava em uma demanda por moeda estável, com relação de causalidade no sentido moeda-preços, e independência dos fatores que afetam a demanda por moeda daqueles que determinam a oferta. A crítica deste autor se direcionava à condução discricionária da política monetária, sendo necessário o estabelecimento de regras que impunham disciplina ao governo, dando maior credibilidade a sua política econômica. Porém, estas propostas perderam força à medida que mudanças institucionais no mercado financeiro tornaram a velocidade de circulação dos vários agregados monetários mais variáveis e incertas do que anteriormente, em suma, a função demanda por moeda tornou-se mais instável.

Em vista das mudanças ocorridas, surgiram sugestões de utilização do produto nominal da economia como meta vis-à-vis a relação M*V.[1] Uma das propostas mais difundidas foi desenvolvida por Taylor (1993), sendo estimadas diversas versões com inclusão de novas variáveis em relação à proposta referencial (*benchmark*). Tendo em vista o amplo arcabouço de regras de política monetária já desenvolvidas na literatura, adotaremos neste trabalho a especificação de uma regra super-inercial conforme Rotemberg e Woodford (1999), que é uma extensão da regra clássica adotada por Taylor (1993).

Segundo Taylor (1999), o Banco Central deveria manter a inflação próxima da meta através de uma âncora aplicada à regra de política monetária, na qual seus instrumentos de política – geralmente a taxa de juros de curto prazo – são ajustados em resposta aos desenvolvimentos da economia. Em um modelo neo-wickselliano básico a regra de Taylor (1993) que corresponderia ao exposto seria dada por:

$$i_t = a\pi_t + bx_t \qquad (1)$$

[1] M = demanda por moeda
V = Velocidade de circulação da moeda.

Se assumirmos o modelo estrutural descrito em Woodford (2003), a determinação do equilíbrio implicaria que

$$a + \frac{1-\beta}{\kappa} b > 1 \qquad (2)$$

A relação (2) é conhecida na literatura como "princípio do maior que um" ou "princípio de Taylor", isto é, a taxa de juros nominal deve aumentar mais do que κ% em resposta a um aumento sustentado de κ% na inflação.[2] No modelo neo-wickselliano básico com a regra de Taylor clássica, cada ponto percentual de aumento permanente na inflação implica em um aumento de $\frac{1-\beta}{\kappa}$ pontos percentuais no hiato do produto e de $a + \frac{1-\beta}{\kappa} b$ pontos percentuais na taxa de juros nominal. Portanto, haverá determinação do equilíbrio no sistema formado pelas equações de demanda, oferta e regra monetária, se, e somente se, a condição (2) for atendida. Os coeficientes sugeridos por Taylor (1993) foram a = 1.5 e b = 0.5, os quais atendem tal condição de equilíbrio.

Já a regra super-inicial de Rotemberg e Woodford (1999) incorpora a tendência de os Bancos Centrais suavizarem os movimentos da meta para a taxa de juros básica. Assim, o modelo proposto é descrito a seguir:

$$i_t = a\pi_t + bx_t + ci_{t-1} \qquad (3)$$

Para esta regra, o sistema formado pelas equações de oferta, demanda e regra monetária possuirá uma solução local única se a seguinte relação for satisfeita:

$$a + \frac{1-\beta}{\kappa} b + c > 1; \text{ a,b} >> 0 \qquad (4)$$

De semelhante modo à regra de Taylor apresentada anteriormente, a condição (4) estende o conceito de "princípio do maior que um" incluindo o termo de suavização da taxa de juros nominal. Nos estudos desenvolvidos pelos autores a regra ótima exibe um coeficiente $c > 1$, sendo este fato o motivador de sua nomenclatura: super-inercial. Em nosso trabalho utilizaremos uma regra de política monetária bastante similar à utilizada por Rotemberg e

[2] Ver Woodford (2003).

Woodford (1999), com a diferença de estipularmos um coeficiente como intercepto para captar eventuais choques não-gaussianos. Portanto, ela terá a seguinte especificação:

$$i_t = a\pi_t + bx_t + ci_{t-1} + d_t \tag{5}$$

2.2. Regras de Política Monetária em Pequenas Economias Abertas

Como o Banco Central deve reagir a choques em uma economia aberta? A regra de política monetária de uma economia aberta tem o mesmo formato e os mesmos coeficientes da regra de uma economia fechada?

Para responder a estas perguntas precisamos ampliar a análise apresentada nas seções anteriores. A razão para o foco em uma pequena economia aberta é o interesse especial pela avaliação de um regime monetário e cambial de uma economia como a brasileira, sendo utilizados os modelos canônicos da literatura sobre regimes monetários em economias abertas. As economias emergentes são economias abertas, porque mantêm fluxos comerciais e financeiros com o resto do mundo, mas são pequenas, no sentido de que mudanças nos caminhos de equilíbrio de suas principais variáveis macroeconômicas (hiato do produto e inflação, taxa de juros nominal, etc) não têm efeitos significativos nos caminhos de equilíbrio das variáveis correspondentes no resto do mundo.

Poderíamos usar um modelo estrutural de pequena economia aberta como uma extensão do modelo de Galí e Monacelli (2002), reconhecido na literatura como um modelo canônico da "nova macroeconomia aberta", utilizando na estimação da regra de política monetária modelos com mudanças de regimes. A regra de política monetária derivada neste trabalho parte de uma função de perda social que é uma extensão para a economia aberta da função de perda proposta por Giannoni e Woodford (2003) para uma economia fechada. Seguindo o modelo proposto por Rigolon (2003), esta função perda social é interpretada como a função de perda social de um ditador benevolente internacional no caso especial de equilíbrio estacionário de longo prazo no resto do mundo. A sua minimização resulta em uma extensão da regra de Rotemberg e Woodford (1999) estudada anteriormente.[3]

O modelo estrutural completo necessita de uma regra de política monetária a qual o Banco Central se comprometa a implementar um equilíbrio ótimo sob a perspectiva

[3] A discussão dos modelos teóricos, bem como suas derivações podem ser encontradas em Galí e Monacelli (2002), Giannoni e Woodford (2003), Rigolon (2003) e Woodford (2003).

atemporal. Quando o fator de desconto β for próximo de 1, o objetivo do Banco Central será minimizar a esperança incondicional da função de perda intertemporal (E(W)), conforme:

$$\min E(W) = E\left[E_0 \{ \sum_{t=0}^{\infty} \beta^t \Psi_t \} \right] \qquad (6)$$

onde Ψ_t é a função de perda instantânea no contexto de uma pequena economia aberta, sendo descrita por:

$$\Psi_t = (\pi_t)^2 + \lambda_x x_t^2 + \lambda_i i_t^2 + \lambda_e (\Delta e_t)^2 \qquad (7)$$

onde π_t, x_t, i_t e Δe_t são os desvios percentuais das variáveis-objetivo em relação às suas metas de longo prazo e os pesos $\lambda_k \geq 0$ medem a importância relativa atribuída pelo Banco Central à estabilização de cada variável-objetivo.

Se o objetivo de estabilizar a inflação ao consumidor em torno da meta de longo prazo for alcançado, as perdas de bem-estar seriam minimizadas. Porém, como temos a inflação ao consumidor desdobrada em inflação doméstica e de importados nestes modelos de economia aberta, a estabilização da inflação ao consumidor requer que a política monetária responda aos desvios da inflação doméstica e da inflação das importações em relação aos seus equilíbrios de longo prazo, o que faz com que a hipótese de *pass-through* tenha relevante importância. Em uma economia aberta e com *pass-through* perfeito, não haveria a necessidade de o Banco Central se preocupar em estabilizar variáveis tais como o câmbio nominal ou os efeitos da inflação das importações. Porém, se a hipótese do *pass-through* perfeito não se verificar para a economia em questão, o Banco Central poderá atuar de forma ineficiente, já que haverá rigidez nominal de preços na determinação da inflação das importações, sendo que quanto maior o grau de abertura da economia, maior a participação da inflação das importações na determinação da inflação ao consumidor (π_t). Daí, para a economia brasileira ser necessária a inclusão de uma variável que meça as variações do câmbio na regra de política monetária.

Entretanto, alguns autores argumentam que não haveria necessidade de se ter o objetivo de estabilização da taxa de câmbio nominal em torno de uma meta de longo prazo ou

de metas de curto prazo. Para o caso da meta de longo prazo, o argumento da paridade de poder de compra (PPP) poderia justificar a busca de uma meta para a variação da taxa de câmbio nominal igual ao diferencial entre as metas para a inflação doméstica e externa. Mas o que justificaria a busca de metas para Δe_t, segundo esta corrente da literatura, seria a questão da fragilidade financeira da maioria das economias emergentes. Calvo e Reinhart (2000) mostraram que muitos países possuíam um medo generalizado de adotar um regime de câmbio flutuante devido à maior exposição de sua economia às volatilidades do mercado internacional, o que se traduzia em possíveis efeitos desfavoráveis para a balança comercial e ao nível de absorção interna destas economias. Daí, conforme aborda Rigolon (2003), a busca por metas que reduzam a volatilidade cambial pode ser justificada como uma saída para economias emergentes que lidam com fragilidade financeira, baixa credibilidade da política monetária, grau elevado de *pass-through* e efeitos adversos sobre suas balanças comerciais.

Enfim, qual deveria ser o formato de uma regra de política monetária em uma economia aberta que atendesse o exposto acima? Respostas ao câmbio devem ser consideradas ou elas perturbam a busca eficiente de outros objetivos, como a estabilização da inflação? Esta pergunta é interessante porque Taylor (2000) refuta a afirmativa de Ball (1999) sobre a inclusão da taxa de câmbio em uma regra *benchmark* como instrumento de aumento do desempenho macroeconômico em uma economia aberta e pequena. Segundo aquele autor, a interferência da taxa de câmbio nos demais objetivos da política pode ser prejudicial. Em seus estudos a inclusão de variáveis como a taxa real e nominal de câmbio corrente e/ou defasada na regra de política monetária proporcionou uma melhora pouco significativa em termos de variabilidade do produto e da inflação, podendo, em alguns casos, deteriorar a *performance* do modelo. O autor explica este resultado através da reação indireta da taxa de juros nominal às variações da taxa de câmbio real. Admitindo-se inércia no mecanismo de transmissão monetária, temos que a apreciação do câmbio corrente produz uma expectativa de redução do produto e da inflação no futuro, sinalizando uma possível redução das taxas de juros pelo Bacen nos períodos posteriores. Admitindo a hipótese de expectativas racionais, o possível cenário contingencial de redução das taxas de juros promove uma redução destas no período corrente. Em suma, a variável-instrumento pode ser afetada por mudanças no câmbio, mesmo que este não esteja presente na regra. Porém, quando se admite um desvio do câmbio muito discrepante em relação à Paridade do Poder de Compra (PPP), pode estar ocorrendo mudanças na produtividade ou aparição do fenômeno de expectativas irracionais (bolhas especulativas), o que resulta em taxas de sacrifício cada vez maiores para a estabilização das

taxas de juros. O custo relativo de suavizar as alterações da taxa de juros pode ser muito alto, sendo melhor extrair os efeitos do câmbio da regra monetária tendo em vista menores flutuações da variável-instrumento, o que minimiza os erros de fixação desta variável.

No entanto, a literatura internacional e, principalmente a nacional, ainda provêem estudos precários sobre a escolha de regras ótimas em economias abertas, indicando, geralmente, a ordenação de regras alternativas através da análise de critérios de eficiência obtidos em simulações numéricas, mas se limitando a comparar o desempenho de regras muitas vezes simples, com parâmetros arbitrários acoplados a modelos com baixo teor de explicação dos mecanismos de transmissão internos. Para o caso brasileiro a dificuldade se amplia bastante, visto que existem poucos trabalhos com este escopo. Uma variante seria o trabalho desenvolvido por Rigolon (2003) que avalia a eficiência de regras monetárias estendidas a partir da derivação das regras estabelecidas pela literatura internacional pelo método de resoluções numéricas. Em seu trabalho o autor afirma que regras que possuem "respostas ao câmbio" se aproximam de regras ótimas eficientes aplicadas a economias abertas. Com este espírito, este trabalho adotará uma versão da regra de Rotemberg e Woodford (1999) estendida para a economia aberta e que servirá como uma *proxy* para a regra ótima que minimiza as perdas de bem-estar descritas anteriormente. Interessante frisar que outras especificações da regra proposta foram avaliadas, mas apresentaram uma eficiência relativa medida pelos critérios estatísticos de modelos multivariados como de menor eficiência em relação à regra proposta neste trabalho.

Portanto, assumindo a função perda de bem-estar (6), a regra de Rotemberg e Woodford (1999) estendida para uma pequena economia aberta teria a seguinte especificação;

$$i_t = c + a\,\pi_t + \tau x_t + \gamma \Delta e_t + \psi i_{t-1} \qquad (8)$$

Lembrando que a_H, a_j, τ e ψ maiores que zero.

2.3. Avaliação de Regras Monetárias Aplicadas

O debate sobre questões de política monetária perpassa pela análise das relações existentes entre as variáveis que compõem uma regra de política monetária. Questões do tipo: "Como a economia responde aos choques exógenos da política monetária?" passam a ter

grande importância na avaliação dos mecanismos de transmissão da política monetária. Na literatura corrente, Christiano, Eichenbaum e Evans (1996, 1997 e 1998) apresentam várias formas de avaliação dos choques de política monetária, demonstrando que os resultados são sensíveis à especificação do modelo e a forma de se isolar os choques de política monetária.

Em relação à estratégia de se isolar os choques de política monetária, existem 3 estratégias gerais abordadas pela literatura corrente. A primeira se foca no conceito de análise. Ela envolve a realização de hipóteses de identificação suficientes para permitir ao analista a estimação dos parâmetros de uma curva de reação do Banco Central, isto é, a regra que relaciona as ações do *policymaker* com o estado da economia em geral. A premente identificação se pauta em hipóteses da forma funcional, das variáveis que o Banco Central observa enquanto utiliza seus instrumentos, e até, hipóteses relativas a quais instrumentos utilizar. Além disto, as hipóteses podem ser feitas sobre a natureza da interação do choque com as variáveis da curva de reação. Uma destas hipóteses é a de que o choque de política monetária é ortogonal às variáveis da regra monetária. A literatura se refere a esta hipótese como hipótese de "recursividade", o que implica dizer que, dada a linearidade da curva de reação do Banco Central, esta hipótese justifica a estimação de choques através de resíduos ajustados em uma estimação de Mínimos Quadrados Ordinários (MQO) dos instrumentos da curva de reação em um instante *t*. O foco principal da análise baseada na hipótese de "recursividade" é que as variáveis do conjunto de informação do Banco Central no tempo *t* não respondem no mesmo período *t* aos choques da política monetária. Por exemplo, Christiano, Eichenbaum e Evans (1996) assumem que o Banco Central americano (FED) acompanha os preços correntes e o produto, entre outras variáveis, quando fixa no tempo *t* os instrumentos de política monetária. Assim, a hipótese da "recursividade" implicava que o produto e os preços respondiam ao choque de política monetária com determinada defasagem.

A adoção da hipótese da "recursividade" também foi abortada por alguns pesquisadores, sendo adotada uma abordagem alternativa. Por exemplo, Sims (1986), Sims e Zha (1995) e Leeper, Sims e Zha (1996) abandonaram a hipótese de "recursividade", adotando outras hipóteses alternativas, tais como, a aplicação da idéia de que o FED não acompanha o produto e o nível de preços contemporâneos quando fixa seus instrumentos de política monetária e que movimentos contemporâneos na taxa de juros nominal não afetam diretamente o produto agregado (Sims e Zha, 1995).

A segunda estratégia envolve acompanhar os dados que aparentemente sinalizam ações de política monetária. Por exemplo, Rudebusch (1995) assume que, em certos períodos

amostrais, mudanças exógenas na política monetária são bem captadas pelas variações nas taxas dos *federal funds*.

A terceira estratégia se baseia no fato de que os choques de política monetária não afetam a atividade econômica no longo prazo. Para maiores detalhes desta variante de avaliação dos choques monetários, poderíamos citar Faust e Leeper (1997) e Pagan e Robertson (1995).

A literatura ainda não convergiu para um conjunto de hipóteses para identificar os efeitos de um choque exógeno da política monetária. Contudo, há uma considerável concordância em relação aos efeitos qualitativos de um choque de política monetária no sentido de que uma inferência robusta tem sido considerada na literatura através de um grande subconjunto de esquemas de identificação (Christiano, Eichenbaum e Evans, 1998). A base de concordância se exprime da seguinte forma: Após um choque de aperto monetário, a taxa de juros de curto prazo sobe, produto agregado, emprego, lucros e agregados monetários caem, o nível de preços responde lentamente, e várias medidas de salário caem, mesmo que em pequenas quantidades. Além disto, há um consenso de que os choques de política monetária são responsáveis por pequenas porcentagens da volatilidade do produto agregado e, menos ainda, dos movimentos no nível de preços agregado.

3. METODOLOGIA

3.1 Modelos lineares univariados e multivariados

O desenvolvimento da econometria moderna reduziu de forma substancial o poder de utilização de instrumentos econométricos tais como o método de equações simultâneas e os modelos de defasagem distribuída. Principalmente a partir das décadas de 1970 e 1980 com o choque do petróleo e a crítica de Lucas, viu-se que os parâmetros estimados eram não só instáveis no tempo, como também dependiam da política adotada *ipso facto* no período analisado. A partir do desenvolvimento da metodologia Box-Jenkins para a análise e previsão de séries temporais econômicas, explodiu a utilização de tais métodos na literatura econômica, bem como um maior aperfeiçoamento dos métodos estatísticos de estimação. Neste âmbito, os modelos univariados, tais como abordados pela metodologia Box-Jenkins (AR (p), MA (q), ARMA (p,q), ARIMA (p,d,q) e SARIMA (P,D,Q)(p,d,q)), já foram extensamente abordados pela literatura econômica, cabendo-nos poupar o leitor de longas discussões a respeito destes métodos.

Já os modelos de séries temporais multivariados, em especial a modelagem VAR (p), começaram a ser largamente utilizados após a crítica de Sims (1980) sobre modelagens macroeconômicas que não levavam em consideração a relação de causalidade entre as variáveis utilizadas. Segundo este autor, algumas variáveis eram tratadas como endógenas e outras como exógenas, sendo que esta predeterminação era freqüentemente alcançada quando se predeterminava algumas variáveis ausentes em certas equações de um sistema. Assim, esta predeterminação era feita de forma subjetiva pelo pesquisador, sem qualquer estabelecimento de igualdade entre as variáveis, ou seja, simultaneidade (Gujarati, 2000, p. 108).

A forma mais básica de um VAR (p) trata todas as variáveis de maneira simétrica sem estabelecer relações de dependência e independência. No entanto, é sabido que diversas ferramentas da análise do VAR (p) – causalidade de granger, exogeneidade, funções resposta a impulso e decomposição de variância – são utilizadas para explicar as relações entre as variáveis em um modelo econômico. O modelo VAR (p) básico possui um vetor de dimensão k, sendo que o vetor-linha de y_t é gerado por um processo de ordem p da forma:

$$y_t = A_0 + A_1 y_{t-1} + ... + A_p y_{t-p} + e_t \qquad (9)$$

em que t = 1,.....,T, sendo que A_0 é um vetor de k interceptos, A_1 são matrizes kxk de k^2 coeficientes de cada uma delas – sendo, portanto, estimados $k + k^2p$ termos – além do termo de perturbação estocástica $e_t \sim NID\ (o, \Sigma)$, dado que $\Sigma = E(e_te_t')$ a matriz de variância-covariância independente no tempo, positiva-definida e não-singular (Morais, 2003, p. 38).

Em relação aos procedimentos estatísticos para a estimação de um modelo VAR (p), deve-se em princípio acrescentar quantas variáveis e *lags* (defasagens) forem necessárias ao modelo, considerando a importância econômica de cada variável acrescentada. Para preservar a simetria do modelo aconselha-se usar o mesmo número de *lags* para todas as equações. Porém, a estimação de um modelo VAR (p) não-restrito pode gerar estimadores imprecisos devido ao substancial número de parâmetros presentes no modelo (lembremos da relação dada anteriormente, $k + k^2p$). Neste caso, devemos aplicar alguns critérios e estatísticas para definição do número correto de defasagens do modelo proposto. Isto pode ser feito de duas maneiras; a primeira é utilizar critérios de comparação de Akaike, Schwartz e Hannah-Quinn, onde é selecionado o modelo com o menor valor encontrado.[4]

Outra maneira é fazer de forma recursiva a seleção do menor número de defasagens. Estima-se um modelo VAR (p) com o maior número de *lags* possíveis, denominado de modelo não-restrito, e obtenha a matriz de variância-covariância Σ_{ij}. Em seguida são estimados outros modelos VAR (p) restritos, obtendo-se a matriz Σ_s. De posse dos resultados estimados, calcula-se o teste de razão de verossimilhança[5], verificando se a restrição imposta ao número de *lags* é válida. Este procedimento é repetido até que se encontre o número ideal de defasagens.

Já a função resposta a impulso pode ser obtida a partir de um vetor de médias móveis (VMA (q)), em que as variáveis são expressas em termos dos valores passados e correntes do choque. Com o fito de ilustrarmos a análise, consideremos um modelo VAR (p) com apenas uma defasagem, ou seja, $y_t = A_0 + A_1 y_{t-1} + e_t$. De forma recursiva a partir de $t - 1$ temos:

$$y_t = A_0 + A_1(A_0 + A_1 y_{t-2} + e_{t-1}) + e_t \tag{10}$$

$$y_t = A_0(I + A_1) + A_1^2 y_{t-2} + A_1 e_{t-1} + e_t \tag{11}$$

[4] Segundo Enders (1995), a probabilidade do critério de informação de Schwartz é menor que o de Hanna-Quinn, e este é menor que a probabilidade de informação de Akaike, dado que o número de observações seja maior que 16.

[5] Este teste é calculado pela seguinte relação: $(T\ \ C)(\log|\Sigma_r| - \log|\Sigma_u| \sim \chi^2_{(r)}$, onde T é o número de observações e C o número de parâmetros do modelo não-restrito.

Substituindo $y_{t-2} = A_0 + A_1 y_{t-3} + e_t$ em (11):

$$y_t = A_0(I + A_1 + A_1^2) + A_1^3 y_{t-3} + A_1^2 e_{t-2} + e_t \tag{12}$$

Após *n* iterações tem-se:

$$y_t = A_0(I + A_1 + ... + A_n^2) + \sum_{t=0}^{n} A_1^i e_{t-1} + A_1^{n+1} y_{t-n+1} \tag{13}$$

Com $\lim_{n\to\infty} A_1^{n+1} y_{t-n+1} = 0$ e assumindo a condição de estabilidade, temos que $y_{t-2} = \mu + \sum_{t=0}^{n} A_1^i e_{t-1}$ de onde podemos derivar a função resposta a impulso a partir da manipulação do termo $\sum_{i=0}^{n} A_1^i e_{t-1}$, processo este denominado de ortogonalização. Morais (2003) afirma que "a ortogonalização se refere ao processo de seleção de muitas possíveis funções de resposta a impulso que podem ser encontradas". No entanto, conforme advertem Sims (1980) e Blanchard e Quah (1988), deve ser estipulada a hipótese de que os choques neste sistema sejam ortogonais, ou seja, não correlacionados.

Por fim, cabe ressaltar que um dos problemas da modelagem VAR (p) é justamente a identificação de sua estrutura. A literatura refere-se a diversas abordagens alternativas para o esquema de identificação do VAR (p). A mais tradicional é a decomposição de *Choleski*, que auxilia no fornecimento de um conjunto mínimo de hipóteses que podem ser úteis para o estabelecimento da identificação de um modelo VAR (p) em sua forma primitiva. Outra alternativa consiste em atribuir restrições de longo prazo, baseadas em interpretações teóricas *a priori*. Por fim, pode-se, também, atribuir restrições de não contemporaneidade a algumas variáveis. Considerando-se, ainda, que as citadas alternativas não são excludentes e, portanto, podem ser interpostas umas às outras, conclui-se que há uma variedade extensa de esquemas de identificação. A escolha irá depender, basicamente, dos objetivos específicos de cada investigação. No entanto, muitas vezes é premente estabelecer algumas restrições que tenham maior relação com os choques para que a identificação se torne precisa.

3.2. Modelo Multivariados com *Markov Switching*

Hamilton (1989), em seu artigo seminal, propôs a possibilidade de utilização da metodologia de cadeias de Markov na análise de dados econômicos, seja em modelos uni ou multivariados. Desde essa publicação tem crescido o interesse dos pesquisadores em utilizar a metodologia *Markov Switching*-VAR (MS – VAR) aplicada a séries temporais econômicas. Os modelos VAR estáveis propostos na seção anterior não são apropriados quando se consideram parâmetros variantes no tempo. Neste caso, o modelo com mudança de regime é mais apropriado para representar a trajetória de uma série econômica sujeita a choques exógenos. A idéia fundamental deste modelo é que os parâmetros do processo gerador de dados do vetor de séries de tempo observado y_t dependem de uma variável regime s_t, não observável, que representa a probabilidade dos diferentes estados contingenciais (Correa, 2003, p. 12). Assim, quando o processo está sujeito à mudança de regime, os parâmetros do modelo VAR tornam-se variantes no tempo, mas invariantes se condicionados à variável não-observada s_t, que aponta o regime prevalecente em t. Supondo que o número de regimes possíveis seja N, tal que $s_t \in \{1,2,...,N\}$, a densidade de probabilidade condicionada do vetor de séries de tempo observado y_t é dada por:

$$P(y_t|\psi_{t-1}, s_t) = \begin{cases} f(y_t/\psi_{t-1}, \theta_1) \text{ se } s_t = 1 \\ \text{.....} \\ f(y_t, \psi_{t-1}, \theta_n) \text{ se } s_t = N \end{cases} \tag{14}$$

onde θ_n é o vetor de parâmetros do VAR no regime n = 1,2,...., N e ψ_{t-1} são as observações das séries de tempo $\{y_{t-j}\}_{j=1}^{\infty}$. Tendo em vista a função densidade de probabilidade proposta acima, um processo MS – VAR ajustado à média teria a seguinte especificação: [6]

$$y_t = \mu(s_t) + \sum_{j=1}^{p} A_i(s_t)[y_{t-i} - \mu(s_{t-i})] + u_t \tag{15}$$

onde $u_t \sim$ IND (0, $\Sigma(s_t)$) e o vetor com médias é agora $\mu(s_t) = [I_k - \sum_{j=1}^{p} A_i(s_t)]^{-1} v(s_t)$. Esta é uma especificação geral do modelo, pois note que $\mu(s_t)$, $A_i(s_t)$ e $\Sigma(s_t)$ são parâmetros condicionados ao regime não-observável s_t. Por exemplo,

[6] Para maior detalhamento ver Correa (2003), Hamilton (1994) e Krolzig (1997b).

$$\mu(s_t) = \begin{cases} \mu_1 \text{ se } s_t = 1 \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ \mu_N \text{ se } s_t = \text{N} \end{cases} \tag{16}$$

O modelo MS – VAR descrito acima permite uma grande variedade de especificações que leva em conta a instabilidade dos parâmetros, mudanças no intercepto, na heteroscedasticidade e na média do processo. Pode-se indagar a possibilidade de diversas combinações de parâmetros que dependam da mudança de regime, conforme salientamos, desde a média do processo ao intercepto do modelo pode-se estipular dependência ou não do regime. Costuma-se em pesquisas empíricas fazer parte dos parâmetros variáveis e adotar formulações nas quais todos os parâmetros tenham o mesmo ponto de inflexão (*turning points*). As variadas formas de especificação são descritas no seguinte quadro[7]:

Quadro 1: Especificações especiais do modelo MS – VAR

		MSM		Especificações MSI	
		μ variante	μ invariante	*v* variante	*v* invariante
	Σ invariante	MSM-VAR	Linear MVAR	MSI-VAR	Linear VAR
A_i invariante	Σ variante	MSMH-VAR	MSH-MVAR	MSIH-VAR	MSH-VAR
	Σ invariante	MSMA-VAR	MSA-MVAR	MSIA-VAR	MSA-VAR
A_i variante	Σ variante	MSMAH-VAR	MSAH-MVAR	MSIAH-VAR	MSAH-VAR

M = mudança na média, A = mudança nos parâmetros, I = mudança no intercepto e H = mudança na heteroscedasticidade.

Completando o processo gerador de dados, faz-se mister adotarmos alguma hipótese para o comportamento estocástico de s_t, desde que os parâmetros da equação (9) sejam dependentes do regime, que é assumido ser estocástico e não observado. A partir destas hipóteses, torna-se possível derivar a densidade marginal de y_t e, concomitantemente, a função log-verossimilhança para a estimação dos parâmetros. Segundo Correa (2003, p. 31), "nos modelos MS – VAR é assumido que a variável estado não-observável $s_t \in \{1,2,\ldots,N\}$ segue uma cadeia de Markov ergódica, irredutível a tempo e espaços discretos, cujas probabilidades de transição são dadas por":

$$\Pr\{s_t = i \mid s_{t-1} = j, s_{t-2} = k, \ldots\} = \Pr\{s_t = i \mid s_t = j\} = p_{ij} \tag{17}$$

[7] Ver Krolzig (1997b).

A probabilidade p_{ij} representa a probabilidade de que no instante *t+1* a cadeia mude para o estado *i*, sendo que ela esteja no estado *j* no tempo *t*. É válido ressaltar também, conforme as abordagens clássica e freqüencialista de probabilidade, que o somatório das probabilidades da cadeia se encontrar em *i* ou *j* tem que ser igual a um, ou seja,

$$\sum_{j=1}^{N} p_{ij} = 1 \quad \forall i, j \in \{1,2,...,N\} \tag{18}$$

Esta característica gera uma matriz de transição (P) de dimensão NxN, denominada matriz coluna-estocástica; a soma dos elementos de cada coluna é igual a um. O elemento da *j*-ésima linha e *i*-ésima coluna desta matriz é a probabilidade de transição p_{ij}, a título ilustrativo, se considerarmos p_{12}, indica qual é a probabilidade do processo estar no regime 1 mas se transferir para o regime 2. Eis a matriz de transição;

$$P = \begin{bmatrix} p_{11} & p_{21} & .... & p_{N1} \\ p_{12} & p_{22} & .... & p_{N2} \\ .... & .... & .... & .... \\ p_{1N} & p_{2N} & .... & p_{NN} \end{bmatrix} \tag{19}$$

3.3. Procedimentos estatísticos de estimação

Considerando uma matriz com as séries de tempo $\psi_{t-1} = (y'_{t-1}, y'_{t-2}, ..., y'_0, ..., y'_{1-p})'$ sendo as variáveis endógenas utilizadas em uma análise, definimos ξ_t um determinado regime. Dado que o termo de perturbação estocástica u_t possui distribuição normal e que o processo esteja no regime $s_t = j$ na data *t*, então, a densidade condicional de y_t é dada por:

$$f(y_t / \xi_t = \iota_j, \psi_{t-1}; \lambda) = \ln(2\pi)^{-\frac{1}{2}} \ln|\Sigma|^{-\frac{1}{2}} \exp\{(y_t - y_{jt})' \sum_j^{-1} (y_t - y_{jt})\} \tag{20}$$

onde ι_j representa a *j*-ésima coluna da matriz identidade I_N, $y_{jt} = E[y_t / \xi_t, \psi_{t-1}]$ é a esperança condicionada de y_t dado o fato de que o processo se encontra em *j* e λ é um vetor que contém os parâmetros da população, que incluem os parâmetros da auto-regressão, θ, e as

probabilidades de transição que orientam a cadeia de Markov dos estados não observados (Krolzig, 1998).

A informação a respeito da realização dos estados da cadeia de Markov é coletada no vetor ξ_t, o qual consiste em variáveis binárias definidas a partir de uma função indicadora assumindo valores zero ou um. Assim,

$$\xi_t = \begin{bmatrix} I(s_t = 1) \\ \dots\dots\dots \\ I(s_t = M) \end{bmatrix}, \qquad I(s_t = m) = \begin{cases} 1 & \text{se } s_t = m \\ 0 & \text{em caso contrário} \end{cases} \tag{21}$$

De semelhante modo, as densidades condicionadas para os N possíveis regimes são definidas por:

$$\eta_t = \begin{bmatrix} f(y_t / \xi_t = \iota_t, \psi_{t-1}) \\ \dots\dots\dots\dots\dots\dots \\ f(y_t / \xi_t = \iota_m, \psi_{t-1}) \end{bmatrix} \tag{22}$$

Para se obter a função densidade marginal de y_t, deve-se utilizar a função densidade conjunta de y_t e ξ_t, integrando-a com relação a todos os regimes. Antes de avaliarmos tal procedimento, explicitaremos a mistura de distribuições i.i.d., que é um caso especial e mais simples dos processos de cadeias de Markov não observadas. Destarte, se considerarmos um processo dado em t onde existem N possíveis regimes (s_t = 1,2,...,N), quando o processo estiver no regime 1, a variável observável y_t será presumida ter vindo de uma distribuição normal $N(\mu_1,\sigma_1)$. Se o processo advém do regime 2, então y_t é derivado de uma distribuição normal $N(\mu_2,\sigma_2)$. Assim, a densidade de y_t condicionada à variável aleatória s_t no regime j é dada por:

$$f(y_t / s_t = i; \lambda) = \frac{1}{\sqrt{2\pi\sigma_i}} \exp\{\frac{-(y_t - \mu_i)^2}{2\sigma_i^2}\} \tag{23}$$

onde λ é o vetor dos parâmetros populacionais que incluem $\mu_1,\dots\dots,\mu_n$ e $\sigma_1,\dots\dots,\sigma_n$.

O regime não observado $\{s_t\}$ é presumido ter sido gerado por alguma distribuição de probabilidade, da qual a probabilidade incondicional de $s_t = i$ é denotada por π_i:

$$P\{s_t = i;\lambda\} = \pi_i \text{ para } i = 1, 2, ..., N. \quad (24)$$

Recorrendo à teoria básica das probabilidades, podemos encontrar a densidade marginal de y_t, integrando-se com relação a s_t e somando as densidades conjuntas para todos os estados possíveis:

$$f(y_t, s_t = i;\lambda) = f(y_t / s_t = i;\lambda) \bullet \Pr\{s_t = i;\lambda) \quad (25)$$

A expressão (25) é comumente chamada de "função densidade de probabilidade unificada (ajuntada)". A partir das equações (22) e (23) esta função é dada por:

$$f(y_t, s_t = i;\lambda) = \frac{\pi_i}{\sqrt{2\pi\sigma_i}} \exp\{\frac{-(y_t - \mu_i)^2}{2\sigma_i^2}\} \quad (26)$$

A densidade incondicional de y_t pode ser encontrada para todos os valores possíveis de i:

$$f(y_t;\lambda) = \sum_{j=1}^{n} p(y_t, s_t = i;\lambda) = \frac{\pi_1}{\sqrt{2\pi\sigma_1}} \exp\{\frac{-(y_1 - \mu_1)^2}{2\sigma_1^2}\} + \frac{\pi_2}{\sqrt{2\pi\sigma_2}} \exp\{\frac{-(y_2 - \mu_2)^2}{2\sigma_2^2}\} + ... + \frac{\pi_n}{\sqrt{2\pi\sigma_n}} \exp\{\frac{-(y_n - \mu_n)^2}{2\sigma_n^2}\}$$
(27)

Para derivar a densidade marginal de y_t e, conseqüentemente, a função verossimilhança, é necessário calcular os termos de peso $\Pr\{s_t = i;\lambda)$. Portanto, alguma inferência sobre o regime não observado deve ser feita. Uma vez que se tenha obtido uma estimativa de λ, é possível fazer uma inferência sobre qual regime mais provavelmente foi o responsável por gerar a observação y_t. Isso pode ser feito usando-se a definição de probabilidade condicional, tal como:[8]

$$\Pr\{s_t = i / y_t;\lambda\} = \frac{f(y_t / s_t = i;\lambda)}{f(y_t;\lambda)} = \frac{\Pr\{s_t = i;\lambda\} * f(y_t, s_t = i;\lambda)}{f(y_t;\lambda)} \quad (28)$$

O caso da mistura de distribuições i.i.d., conforme afirmamos alhures, é um caso especial de cadeia de Markov. Porém, o mesmo raciocínio pode ser aplicado para o caso mais

[8] Ver Correa (2003) e Hamilton (1994).

geral, sendo que a inferência sobre o estado seja dependente de todas as informações disponíveis. Desse modo, uma generalização da equação (28) é feita através do filtro e suavizador BHLK (*Baum-Lindgren-Hamilton-Kim*), que possibilita fazer inferências a respeito dos estados do processo via probabilidades filtradas e suavizadas.

Antes de explicar o processo, porém, deve-se explicitar as seguintes definições:

$\xi_{t/\tau}, \ \tau \leq t$	probabilidades preditas do regime (*predicted*)
$\xi_{t/\tau}, \ \tau = t$	probabilidades do filtradas regime (*filtered*)
$\xi_{t/\tau}, \mathrm{t} \leq \tau \leq T$	probabilidades suavizadas do regime (*smoothed*)

A partir da notação demonstrada, a inferência ótima e a previsão para cada data *t* na amostra pode ser derivada pela iteração do seguinte par de equações:

$$\xi_{t/t} = \frac{(\xi_{t/t-1} \otimes \eta_t)}{1'(\xi_{t/t-1} \otimes \eta_t)} \tag{29}$$

$$\xi_{t+1/t} = P * \xi_{t/t} \tag{30}$$

onde $\boldsymbol{\eta_t}$ representa o vetor com as densidades condicionais em (28), **P** é a matriz de transição dada em (19), **1** é um vetor de dimensão (Nx1) com números uns e o símbolo $\otimes$ denota a multiplicação de elemento por elemento. Das equações (29) e (30) são calculadas as probabilidades filtradas a partir do valor inicial $\xi_{1/0}$ e do valor do parâmetro populacional λ, iterage-se a equações para *t* = 1,2, ..., T, calculando-se os valores de $\xi_{t/t}$ e $\xi_{t+1/t}$ para cada data *t* da amostra.

A inferência sobre os regimes também pode ser feita através de probabilidades suavizadas. Neste caso, utiliza-se o algoritmo de *KIM* (que integra o filtro e suavizador BHLK). Não entraremos em detalhes a respeito da formulação deste algoritmo, contudo, na forma vetorial, ele pode ser escrito como:[9]

$$\xi_{t/T} = \xi_{t/t} \otimes \{P'[\xi_{t+1/T}(\div)\xi_{t+1/t}]\} \tag{31}$$

[9] Ver Kim(1993) e Hamilton (1994).

onde os símbolos $\otimes$ e $(\div)$ denotam, respectivamente, a multiplicação e a divisão de elemento por elemento. As probabilidades suavizadas $\xi_{t/T}$ são encontradas iteragindo (31) para trás, para t = T – 1, T – 2, ..., 1. Esta iteração é iniciada com $\xi_{T/T}$, que é obtida de (29), fazendo-se t = T.

A estimação de máxima verossimilhança do modelo é baseada na implementação do algoritmo de expectativa-maximização (*expectation-maximization – EM* algorithm). Cada iteração do algoritmo EM consiste em dois passos: um passo de *expectativa* e outro de *maximização*. No passo de expectativa os estados não observados s_t são estimados pelas probabilidades suavizadas $\Pr(s_t / \psi_T; \lambda^{j-1})$, onde todas as probabilidades condicionais $\Pr(s / \psi; \lambda^{j-1})$ são calculadas com as recursões filtradas e suavizadas usando o vetor de parâmetros λ^{j-1} estimado no último passo de maximização, já que o vetor de parâmetro λ é desconhecido. No passo de maximização, uma estimativa do vetor de parâmetro λ é derivada como uma solução $\hat{\lambda}$ das condições de primeira ordem associadas com a função de verossimilhança, onde as probabilidades condicionais desconhecidas dos regimes $\Pr(s / \psi; \lambda)$ são substituídas pelas probabilidades suavizadas $\Pr(s_t / \psi_T; \lambda^{j-1})$ derivadas no último passo de expectativa. Admitido o novo vetor de parâmetros $\hat{\lambda}$, as *probabilidades filtradas* $\Pr(s_t / \psi_T; \lambda^{i1})$ e as *probabilidades suavizadas* $\Pr(s_t / \psi_T; \lambda^{i1})$ são novamente incorporadas e assim sucessivamente.

Assim, começando de uma estimativa inicial arbitrária para o valor de λ, denotado por $\lambda^{(0)}$, é calculado o valor de $\Pr\{s_t = i / y_t; \lambda^{(0)}\}$, usando-se o filtro e suavizador BHLK. Das condições de maximização da função log-verossimilhança, usando $\lambda^{(0)}$ no lugar de $\hat{\lambda}$, uma nova estimativa do vetor de parâmetros $\lambda^{(1)}$ é gerada. Esta estimativa $\lambda^{(1)}$ é usada para reavaliar $\Pr\{s_t = i / y_t; \lambda^{(1)}\}$ e recalcular um novo vetor de parâmetros $\lambda^{(2)}$. Este processo iterativo é realizado até que a variação entre $\lambda^{(m+1)}$ e $\lambda^{(m)}$ seja menor que algum critério de convergência anteriormente estabelecido. Assim, cada iteração do algoritmo EM envolve uma passagem pela filtragem e suavização, seguida da resolução das condições de primeira ordem para a estimação do vetor de parâmetros, o que garante um acréscimo no valor da função de verossimilhança (Correa, 2003, p. 12).

O *software* utilizado na estimação dos modelos é o Ox versão 3.31 e seus componentes adicionais, o MSVAR 1.31 e as rotinas desenvolvidas por Krolzig (1997b)[10].

[10] Disponível em http://www.nuff.ox.ac.uk/Users/Doornik/.

3.4. Análise de Resposta a Impulso em modelos multivariados não-lineares

Muitos artigos têm buscado medir o efeito da persistência de choques sobre as variáveis macroeconômicas, utilizando modelos lineares univariados para provar esta persistência. No entanto, autores como Beaudry e Koop (1993), Potter (1995), Pesaran e Potter (1997) e Koop, Pesaran e Potter (1997, p. 25) têm colocado suas atenções em modelos multivariados não-lineares, argumentando que modelos lineares são muito restritivos. Em primeiro lugar, pelo fato de estes modelos apresentarem choques simétricos entre períodos, não distinguindo momentos de expansão e recessão, o que faz com que não capturem assimetrias intrínsecas às flutuações do ciclo de negócios.

Ademais, outros autores (Blanchard e Quah (1989) e Lee e Pesaran (1993)) vêm acrescentando ao debate o entendimento mais rico das relações econômicas e da persistência dos choques a partir da consideração de várias séries temporais macroeconômicas. Isto é possível graças à utilização de modelos multivariados não-lineares.

A presença de quebras estruturais e variáveis dependentes de regimes probabilísticos podem implicar vieses nas estimativas das funções resposta a impulso de um modelo multivariado. Conforme discutido em Potter (1995), modelos não-lineares produzem respostas a impulso que são dependentes do choque e da história das variáveis; isto implica em uma função resposta a impulso (FRI) tratada como uma variável aleatória. Esta classe de funções resposta a impulso é denominada de Função Resposta a Impulso Generalizada (*Generalized Impulse Response Function-GI*).

Uma Função resposta a impulso (FRI) mede o perfil temporal do efeito de um choque sobre o comportamento de uma série. Desta maneira, podemos caracterizar a função resposta a impulso como perfil temporal do efeito de um choque positivo unitário perturbando o sistema no tempo *t*, assumindo que nenhum outro choque afetará o sistema ao longo deste período. Em relação a uma função resposta a impulso tradicional, podemos assumir um modelo multivariado não-linear Markoviano de ordem *p* com a seguinte especificação:

$$y_t = F(y_{t-1},...,y_{t-p}) + H_t V_t \qquad (32)$$

onde F(.) é uma função conhecida, y_t é um vetor aleatório Kx1, V_t é um vetor Kx1 de distúrbios aleatórios i.i.d. e H_t é uma matriz KxK aleatória que é função de $\{y_{t-1},.....,y_{t-p}\}$.

Definimos a função resposta a impulso referente ao modelo dado em (15) como I_y. Ela é definida como sendo a diferença entre duas diferentes realizações de y_{t+n} em relação ao período anterior, $t - 1$. Assumimos que entre os períodos $t+n$ e t o sistema é perturbado apenas por um choque de tamanho δ e que as variáveis estão contidas em um conjunto de informações ω, então;

Se y_t é uma variável aleatória em (Ω,f,P) e f_t é uma seqüência de sigma-álgebras em f. Então existe uma seqüência de funções $\mu_n(t,H,\varpi)$ definidas para H em β(R), e ω em Ω, com as seguintes propriedades:

- Para cada ω em Ω $\mu_n(t,H,\varpi)$ é, como uma função de H, uma medida de probabilidade em R;
- Para cada H em β(R), $\mu_n(t,H,\varpi)$ é, como uma função de ω, uma versão de $P[y_{t+n} \in H | f_t]_\varpi$.

Usando estes resultados podemos definir a GI como uma seqüência de variáveis aleatórias sobre o espaço de probabilidade de uma série temporal da seguinte forma:

$$GI_n(t,\varpi) = \int y\mu_n(t,dy,\varpi) - \int y\mu_{n+1}(t-1,dy,\varpi) \qquad (33)$$

Por exemplo, no caso da equação (15), nós teríamos:

$$GI_y(n,\delta,\varpi_{t-1}) = E[y_{t+n} | V_t = \delta, V_{t+1} = 0,....,V_{t+n} = 0, \varpi_{t-1}] - E[V_{t+1} = 0,....,V_{t+n} = 0, \varpi_{t-1}] \quad (34)$$

O modelo acima busca responder a seguinte pergunta: Qual é o efeito de um choque de tamanho δ perturbando o sistema no tempo t, em relação ao período $t+n$, dado que nenhum outro choque aconteceu? Esta definição de uma função resposta a impulso captura a propriedade do mecanismo de propagação do modelo, comparando o valor de y_{t+n} após a ocorrência do choque com seu valor *benchmark* não sujeito a choques. A função resposta a impulso tradicional é mais comumente usada em modelos lineares, no entanto, é interessante notar que, no caso de modelos não-lineares, a função resposta a impulso depende do conjunto de informações ω_t, isto é, da história particular escolhida como base de comparação e do tamanho do choque δ, selecionado pelo pesquisador (Koop, Pesaran e Potter, 1997).

Já a Função Resposta a Impulso Generalizada (GI) é desenhada para resolver os problemas de composição de dependência citados acima (dependente do choque e da história da variável). O problema do tratamento do futuro é abordado pelo uso do operador de expectativas condicionado pela história e/ou choque, isto é, os choques futuros são *average out* (Koop, Pesaran e Potter, 1997). Assim, a resposta construída é uma média do que aconteceu dado o presente e o passado.

Em relação aos modelos com *Markov-Switching*, as funções resposta a impulso são dependentes dos regimes probabilísticos explicados anteriormente. Na verdade, esta função descreve a relação entre variáveis endógenas e distúrbios fundamentais dentro de um regime. As funções resposta a impulso dependentes de regimes são condicionais a um dado regime prevalente no tempo do distúrbio, avaliando a duração das respostas *dado cada regime*. Importante lembrar que a validade deste condicionamento da FRI a seu regime depende do horizonte temporal da resposta a impulso e da duração esperada do regime. Quanto menor o horizonte temporal e maior a persistência dos choques captada pela matriz de transição, maior a validade da função resposta a impulso como uma ferramenta analítica na avaliação de políticas econômicas (Ehrmann, Ellison e Valla, 2001, p. 9).

O modelo possui mK^2 funções resposta a impulso dependentes de regimes, correspondendo à reação de K variáveis a K distúrbios em m regimes. A equação (35) define matematicamente as funções resposta a impulso dependentes de regimes para cada regime i. Demonstram-se as mudanças esperadas nas variáveis endógenas no tempo $t+n$, fruto de um choque aleatório de x%, condicional ao regime i.

$$\frac{\partial E_t y_{t+n}}{\partial V_t}\Big|_{s_t,\ldots,s_{t+n}=i} = \theta_{ki.n} \quad \text{para } n > 0 \qquad (35)$$

onde $\theta_{ki,1}$, ..., $\theta_{ki,n}$ são vetores de resposta K-dimensionais, prevendo a resposta a impulso das variáveis endógenas.

Estimações dos vetores de resposta podem ser obtidas a partir da combinação de parâmetros estimados de um MS – VAR com a estimação da matriz A_i obtida através de restrições de identificação. As medidas dos vetores de resposta são derivadas a partir de um choque no desvio-padrão de um K-ésimo distúrbio fundamental, isto é, um vetor inicial de distúrbios com a seguinte forma; $u_0 = (0,\ldots,0,1,0,\ldots,0)$. Multiplicando este vetor pela matriz A_i dependente de regimes descrita na equação (32), temos as respostas de impacto. Assim,

$$\theta_{ki,0} = A_i V_0 \tag{36}$$

$$\theta_{ki,n} = \sum_{j=1}^{\min(n,p)} B_{ji}^{n-j+1} A_i V_0 \tag{37}$$

Podemos melhorar a precisão das estimativas dos vetores de resposta empregando técnicas de *bootstrap* padrão. As técnicas buscam criar histórias artificiais para as variáveis do modelo e então, submetemos estas histórias criadas aos mesmos procedimentos de estimação dos dados. As histórias artificiais são criadas através da recolocação de parâmetros no modelo com seus valores estimados, aproximando os resíduos através dos momentos estimados pela matriz de variância-covariância obtida anteriormente. Assim, recalculamos as variáveis endógenas. Claro que se a amostra for pequena, as séries históricas artificiais não coincidirão com os dados originais, porém, a técnica *bootstrap* buscará uma aproximação da distribuição dos parâmetros estimados (Ehrmann, Ellison e Valla, 2001).

Em modelos multivariados com *Markov-Switching* a aplicação de técnicas de *bootstrap* torna-se bastante complicada devido à presença de cadeias de Markov ocultas determinando os regimes. Para criar a história artificial da série é mister criar a priori o histórico dos regimes de transição. Ehrmann, Ellison e Valla (2001, p. 19) apontam 5 passos para a obtenção de estimações de funções resposta a impulso generalizadas (GI) com *Markov-Switching*. Os procedimentos são definidos a seguir:

1. Criando uma história para o regime oculto s_t: Usando a definição de um processo de Markov dada em (28), recursivamente obtemos a matriz de transição exógena. A cada período t sacamos um número aleatório de uma distribuição uniforme [0,1] e o comparamos com a probabilidade de transição condicional para determinar se há uma mudança de regime.
2. Criando uma história para as variáveis endógenas: Novamente isto é feito de forma recursiva. Todos os parâmetros são estimados e os resíduos fundamentais retirados de uma distribuição normal $V_t \sim N(0;I_K)$. A equação (40) pode ser aplicada recursivamente usando o histórico artificial dos regimes.
3. Estimando um MS – VAR: Usando os dados da história artificial, estime um MS (k) – VAR (p). Serão obtidos os vetores de parâmetros

$\{v_i, B_{1i}, ..., B_{pi}; \sum i\}$ para i = 1,...., n , a matriz de transição P e as probabilidades suavizadas $\xi_{i,t} = \Pr(s_t = i)$ para i = 1,..., n e t = 1,..., T.

4. Impondo restrições de identificação: Aplique as mesmas restrições de identificação anterior ao *bootstrap*.
5. Calculando as estimativas *bootstrap* dos vetores de resposta: Substitua os novos parâmetros B_{ii}, ..., B_{pi} e A_i nas equações (36) e (37) e estime os vetores de resposta $\theta_{ki,1}$, ..., $\theta_{ki,n}$ para cada regime i = 1,....,m.

Aplicando os 5 passos descritos anteriormente para um número suficientemente grande de replicações, podemos obter uma aproximação numérica da distribuição de estimativas originais de $\theta_{ki,1}$, ..., $\theta_{ki,n}$. Em uma análise de resposta a impulso esta distribuição forma a base para os intervalos de confiança de uma função resposta a impulso tradicional (Ehrmann, Ellison e Valla, 2001).

3.5. Metodologia *Markov Chain Monte Carlo* (MCMC)

Conforme já salientamos anteriormente, os parâmetros de um Vetor autoregressivo (VAR) são modelados como resultado de um processo discreto de Markov com probabilidades de transição desconhecidas. Os regimes não observados, juntamente com suas probabilidades de transição, determinam os parâmetros de um VAR dentro de cada regime, supondo linearidade intra-regime. Hamilton (1989) propôs um algoritmo de máxima verossimilhança denominado EM (*Expectation-Maximization*) para estimar um VAR com mudança de regime. Um procedimento bayesiano de estimação é desenvolvido nesta seção, tendo a utilização do algoritmo de *Markov Chain Monte Carlo* (MCMC) seu principal foco. A literatura demonstra que um procedimento de estimação baseado em MCMC bayesiano é mais informativo, flexível e eficiente que uma abordagem baseada na estimação por máxima verossimilhança.

A necessidade desta seção se justifica pelo fato de se aplicar inovações metodológicas para a estimação de regras de política monetária, principalmente para o caso do Plano Real brasileiro que apresentou grande variabilidade das séries econômicas devido às mudanças estruturais ocorridas no Brasil ao longo deste período.

Dados a distribuição *a posteriori* conjunta dos regimes e os parâmetros obtidos a partir dos dados iniciais, $p(\rho, \lambda \mid Y)$ pode ser simulada via *Gibbs Sampler* e algoritmo

Metropolis-Hasting. O algoritmo envolverá a geração repetida de várias densidades condicionais, consistindo nos seguintes passos:

$$\rho^{(c+1)} \leftarrow \mu^{(c)}, A^{(c)}, \Omega^{(c)}, P^{(c)} \ (Passo\ 1)$$
$$\mu^{(c+1)} \leftarrow \rho^{(c+1)}, A^{(c)}, \Omega^{(c)}, P^{(c)} \ (Passo\ 2)$$
$$A^{(c+1)} \leftarrow \rho^{(c+1)}, \mu^{(c+1)}, \Omega^{(c)}, P^{(c)} \ (Passo\ 3)$$
$$\Omega^{(c+1)} \leftarrow \rho^{(c+1)}, \mu^{(c+1)}, A^{(c+1)}, P^{(c)} \ (Passo\ 4)$$
$$P^{(c+1)} \leftarrow \rho^{(c+1)}, \mu^{(c+1)}, A^{(c+1)}, \Omega^{(c+1)} \ (Passo\ 5)$$

Em cada caso V será uma função de A e Ω. Sob condições normais de regularidade, a seqüência $\{\rho^{(c+1)}, \lambda^{(c+1)}\} \equiv \{\rho^{(c+1)}, \mu^{(c+1)}, A^{(c+1)}, \Omega^{(c+1)}, P^{(c+1)}\}$ formará uma cadeia de Markov cuja distribuição limite será $p(\rho, \lambda \mid Y)$.

A) Passo 1: Gerando os regimes ($\rho^{(c+1)} \leftarrow \mu^{(c)}, A^{(c)}, \Omega^{(c)}, P^{(c)}$)

Os regimes podem ser gerados conjuntamente de:

$$p(\rho \mid Y, \lambda) = p(\rho_N \mid Y, \lambda).\prod_{t=q}^{N-1} p(\rho_t \mid \rho_{t+1}, Y_t, \lambda). \tag{38}$$

O filtro de probabilidades, $(\rho_N \mid Y, \lambda)$, pode ser calculado de:

$$l(x_q, \rho_q \mid \lambda) = l(x_q \mid \rho_q, \lambda).p(\rho_q \mid \lambda)$$

$$l(x_q \mid \lambda) = \sum_{\rho_q=1}^{K} l(x_q, \rho_q \mid \lambda)$$

$$p(x_q \mid x_q, \lambda) = l(x_q, \rho_q \mid \lambda) / l(x_q \mid \lambda)$$

Para $t = q + 1, ..., N,$

$$l(x_t, \rho_t, \rho_{t-1} \mid Y_{t-1}, \lambda) = l(x_t \mid \rho_t, Y_{t-1}, \lambda).p(\rho_t \mid \rho_{t-1}).p(\rho_{t-1} \mid Y_{t-1}, \lambda)$$

$$l(x_t \mid Y_{t-1}, \lambda) = \sum_{\rho_t=1}^{K} \sum_{\rho_{t-1}=1}^{K} l(x_t, \rho_t, \rho_{t-1} \mid Y_{t-1}, \lambda)$$

$$p(\rho_t, \rho_{t-1} \mid Y_{t-1}, \lambda) = l(x_t, \rho_t, \rho_{t-1} \mid Y_{t-1}, \lambda) / l(x_t \mid Y_{t-1}, \lambda)$$

$$p(\rho_t \mid Y_t, \lambda) = \sum_{\rho_{t-1}=1}^{K} p(\rho_t, \rho_{t-1} \mid Y_t, \lambda)$$

Para $\forall \rho_t, \rho_{t-1} \in \{1,...,K\}$

Para o início das iterações, $p(\rho_q \mid \lambda)$ será necessário. Ele pode ser derivado como uma distribuição limite da cadeia de Markov. Defina o vetor coluna *Kx1* $\pi = \{p(\rho_t = i \mid \lambda),\ i = 1,...,K\}$, então, $\pi = P\pi$. π pode ser estimado através da iteração de $\pi^{(n+1)} = P\pi^{(n)}$ até se conseguir a convergência por determinado critério. Os elementos de $p(\rho_q \mid \lambda)$ são dados como elementos de π.

Assim, para gerar uma amostra da distribuição conjunta de ρ, deve-se primeiro gerar ρ_N a partir de $p(\rho_N \mid Y, \lambda)$. Então, para *t = q + 1, ...,N*, calcula-se $p(\rho_{t+1} \mid \rho_t, Y_t, \lambda)$ usando o valor mais recente de ρ_{t+1} e a probabilidade filtrada dada no momento anterior, conforme a seguinte relação:

$$p(\rho_{t+1}, \rho_t \mid Y_t, \lambda) = p(\rho_{t+1} \mid \rho_t, \lambda).p(\rho_t \mid Y_t, \lambda)$$

$$p(\rho_{t+1} \mid Y_t, \lambda) = \sum_{\rho\ =1}^{K} p(\rho_{t+1}, \rho_t \mid Y_t, \lambda)$$

$$p(\rho_t \mid \rho_{t+1}, Y_t, \lambda) = \frac{p(\rho_{t+1}\rho_t \mid \lambda)}{p(\rho_{t+1} \mid Y_t, \lambda)}$$

Após as probabilidades $p(\rho_t \mid \rho_{t+1}, Y_t, \lambda)$ terem sido calculadas, ρ_t pode ser facilmente gerada de $p(\rho_t \mid \rho_{t+1}, Y_t, \lambda)$, desde que a densidade seja discreta. Para que o processo de mudança de regime seja definido, cada um dos *K* regimes precisa ser visitado.

B) Passo 2: Gerando os parâmetros ($\mu^{(c+1)} \leftarrow \rho^{(c+1)}, A^{(c)}, \Omega^{(c)}, P^{(c)}$)

Suponha que μ_r representa um dos valores possíveis de $\mu_{(\rho_t)}$. As distribuições independentes a priori podem ser utilizadas. O vetor de parâmetros para as médias condicionadas a cada regime seria dado por:

$$p(\mu_1,...,\mu_K \mid \rho_t, \lambda_{=\mu}, Y) \propto L(Y \mid \rho, \lambda) xp(\mu_1,....,\mu_k) \qquad (39)$$

sendo o produto de K densidades normais multivariadas independentes. A expressão acima pode ser rearranjada conforme desenvolvimento abaixo:

$$-\frac{1}{2}\xi_q^T V_{\rho_q}^{-1}\xi_q - \frac{1}{2}\sum_{t=q+1}^{N}\xi_t^T \Omega_{\rho_t}^{-1}\xi_t$$
$$= -\frac{1}{2}(x_q - \mu_{\rho_q})^T V_{\rho_t}^{-1}(x_q - \mu_{\rho_q}) -$$
$$-\frac{1}{2}\sum_{t=q+1}^{N}(x_t - \sum_{h=1}^{q} A_{\rho_t}^h x_{t-h} - (I_m - \sum_{h=1}^{q} A_{\rho_t}^h)\mu_{\rho_t})^T \Omega_{\rho_t}^{-1}(x_t - \sum_{h=1}^{q} A_{\rho_t}^h x_{t-h} - (I_m - \sum_{h=1}^{q} A_{\rho_t}^h)\mu_{\rho_t})$$
$$= -\frac{1}{2}\xi_q^T V_k^{-1}\xi_q$$
$$-\frac{1}{2}\left(\mu_k - \frac{1}{n_k - 1}(I_m - \sum_{h=1}^{q} A_{k_t}^h)^{-1}\sum_{\rho_t = k}(x_t - \sum_{h=1}^{q} A_{k\rho}^h x_{t-h})\right)^T (n_k - 1)W_k^{-1}\left(\mu_k - \frac{1}{n_k - 1}(I_m - \sum_{h=1}^{q} A_{k_t}^h)^{-1}\sum_{\rho_t = k}(x_t - \sum_{h=1}^{q} A_{k\rho}^h x_t\right.$$
$$-\frac{1}{2}\sum_{r \neq k}\left(\mu_r - \frac{1}{n_r}(I_m - \sum_{h=1}^{q} A_r^h)^{-1}\sum_{\rho_t = r}(x_t - \sum_{h=1}^{q} A_r^h x_{t-h})\right)^T (n_r)W_r^{-1}\left(\mu_r - \frac{1}{n_r}(I_m - \sum_{h=1}^{q} A_r^h)^{-1}\sum_{\rho_t = r}(x_t - \sum_{h=1}^{q} A_r^h x_{t-h})\right)$$

onde $W_{\rho_t}^{-1} = \left(I_m - \sum_{h=1}^{q} A_{\rho_t}^h\right)^T \Omega_{\rho_t}^{-1}\left(I_m - \sum_{h=1}^{q} A_{\rho_t}^h\right)$

Ignorando o primeiro termo em V^{-1} , a expressão acima denota um vetor de k densidades normais independentes em μ_r. Assim,

$$\therefore p\left(\mu_1,...,\mu_k \mid \rho, \lambda_{\neq\mu}, Y\right) \propto \left|V_k^{-1}\right|^{\frac{1}{2}} \exp\{-\frac{1}{2}\xi_q^T V_k^{-1}\xi_q\} \times N\left(\frac{1}{n_k - 1}(I_m - \sum_{h=1}^{q} A_{k_t}^h)^{-1}\sum_{\rho_t = k}(x_t - \sum_{h=1}^{q} A_{k\rho}^h x_{t-h}), \frac{1}{n_k - 1}W_k\right)$$
$$\times \prod_{r \neq k} N\left(\frac{1}{n_r}(I_m - \sum_{h=1}^{q} A_{r_t}^h)^{-1}\sum_{\rho_t = r}(x_t - \sum_{h=1}^{q} A_r^h x_{t-h}), \frac{1}{n_r}W_r\right) \times p(\mu_1,...,\mu_k).$$

onde N(.) é um vetor de densidades distribuídas como uma normal. Os valores para μ_r podem ser independentemente gerados da densidade normal multivariada descrita acima. Assintoticamente as médias das densidades acima, para cada regime, são a média do vetor de dados em cada regime.

Interessante frisar que um passo do algoritmo Metropolis-Hasting pode ser usado para gerar μ_r. Em primeiro lugar, geramos um candidato para μ_r conforme a aproximação $\mu_k^* \sim N\left(\frac{1}{n_k-1}(I_m - \sum_{h=1}^{q} A_{k_t}^h)^{-1} \sum_{\rho_t=k}(x_t - \sum_{h=1}^{q} A_{k\rho}^h x_{t-h}), \frac{1}{n_k-1}W_k\right)$ e aceito com probabilidade $\min\{\sqrt{\frac{|V_k^{-1*}|}{|V_k^{-1(c)}|}} \exp\{-\frac{1}{2}\xi_q^T(V_k^{-1*} - V_k^{-1(c)})\xi_q\},1\}$, para $\mu_k^* = \mu_k^{c+1}$. Caso contrário, o valor anterior da iteração de Gibbs é mantido, isto é, $\mu_k^c = \mu_k^{c+1}$. Assim, a geração de k μ_r continua até que as condições a priori representadas por $p(\mu_1,...,\mu_k)$ são satisfeitas.

Em relação à geração dos parâmetros a partir dos procedimentos de geração dos regimes, as densidades condicionais dos parâmetros são dadas por:

$$p(\lambda_j \mid \rho, \lambda_{\neq j}, Y) \propto L(Y \mid \rho, \lambda).p(\rho \mid \lambda).p(\lambda_j) \tag{40}$$

A) *Passo 3: Gerando os parâmetros regressivos de correlação* ($A^{(c+1)} \leftarrow \rho^{(c+1)}, \mu^{(c+1)}, \Omega^{(c)}, P^{(c)}$)

Assuma que A_r representa um dos valores possíveis de A_ρ. As matrizes de parâmetros de correlação podem ser gerados conjuntamente de:

$$p(A_1,....,A_k \mid \rho, \lambda_{\neq A}, Y) \propto L(Y \mid \rho, \lambda).p(A_1,....,A_k) \tag{41}$$

Relembrando que

$$A_{\rho_t} = \begin{pmatrix} A_{\rho_t}^1 & A_{\rho_t}^2 & \cdots & A_{\rho_t}^q \\ I_m & 0_m & \cdots & 0_m \\ 0_m & I_m & \cdots & 0_m \\ \vdots & \vdots & \ddots & \vdots \\ 0_m & 0_m & \cdots & 0_m \end{pmatrix}$$

As primeiras m linhas precisam ser geradas. Defina o operador matricial m x mq $\vartheta = (I_m, 0_m, ..., 0_m)$, tal que $\vartheta A = (A^1, A^2, ..., A^q)$. Este vetor também é preciso ser gerado.

Suponha $\rho_t = k$, considere o regime $\rho_t = r\ (\neq k)$ e a matriz de resíduos do regime *r*, $\varsigma_r = (\xi_t : \rho_t = r)$, além das matrizes de desvios das variáveis, $\tilde{\chi}_r = \left((x_{t-1} - \mu_r) : \rho_t = r\right)$ e $\chi_r = \left((x_t - \mu_r) : \rho_t = r\right)$. Portanto,

$$
\begin{aligned}
-\varsigma &= \vartheta A \tilde{\chi} - \vartheta \chi \\
-\varsigma \tilde{\chi}^T &= \vartheta A \tilde{\chi} \tilde{\chi}^T - \vartheta \chi \tilde{\chi}^T \\
-\varsigma \tilde{\chi}^T &= \left(\vartheta A - \vartheta (\chi \tilde{\chi}^T)(\tilde{\chi} \tilde{\chi}^T)^{-1} \right) (\tilde{\chi} \tilde{\chi}^T) \\
-\varsigma &= \left(\vartheta A - \vartheta (\chi \tilde{\chi}^T)(\tilde{\chi} \tilde{\chi}^T)^{-1} \right) \tilde{\chi} \\
-vec\varsigma &= \left(\tilde{\chi}^T \otimes I_m \right) \left[\left(vec\,\vartheta A - vec\,\vartheta (\chi \tilde{\chi}^T)(\tilde{\chi} \tilde{\chi}^T)^{-1} \right) \right] \\
&= \left(\tilde{\chi}^T \otimes I_m \right) \left[\left(vec\,\vartheta A - vecC \right) \right]
\end{aligned}
$$

onde $C = \vartheta (\chi \tilde{\chi}^T)(\tilde{\chi} \tilde{\chi}^T)^{-1}$ é uma matriz de correlação no *lag* 1. Perceba que $vec(AB) = (B^T \otimes I) vec(A)$. Lütkepohl (1991) apresenta uma visão sobre as propriedades do produto de Kronecker e seus operadores.

A partir dos desenvolvimentos acima mencionados, o termo de verossimilhança pode ser expresso como:

$$
\begin{aligned}
& -\frac{1}{2} \xi_q^T V_{\rho_t}^{-1} \xi_q - \frac{1}{2} \sum_{t=q+1}^{N} \xi_t^T \Omega_{\rho_t}^{-1} \xi_t \\
&= -\frac{1}{2} \xi_q^T V_k^{-1} \xi_q - \frac{1}{2} vec\varsigma_k^{\,T} \left(I_{n_{k-1}} \otimes \Omega_k^{-1} \right) vec\varsigma_k - \frac{1}{2} \sum_{r \neq k} vec\varsigma_r^{\,T} \left(I_{n_r} \otimes \Omega_r^{-1} \right) vec\varsigma_r
\end{aligned}
$$

Incluindo os dados a priori, a expressão acima se torna:

$$-\frac{1}{2}\xi_q^T V_{k\rho_t}^{-1}\xi_q - \sum_r \left(vec\,\vartheta A_r - vecC_r\right)^T \left((\tilde{\chi}_r \tilde{\chi}_r^T) \otimes \Omega_r^{-1} \right) \left[vec\,\vartheta A_r - vecC_r\right]$$

$$-\frac{1}{2}\sum_r (vec\,\vartheta A_r - vecB_r)^T \nu_r I_{m^2 q} \left[vec\,\vartheta A_r - vecB_r\right]$$

$$= -\frac{1}{2}\xi_q^T V_{k\rho_t}^{-1}\xi_q - \sum_r (vec\,\vartheta A_r - \alpha_r)^T \Psi_r^{-1} \left[vec\,\vartheta A_r - \alpha_r\right]$$

onde $\Psi_r^{-1} = \left((\tilde{\chi}_r \tilde{\chi}_r^T) \otimes \Omega_r^{-1} \right) + \nu_r I_{m^2 q}$ e $\alpha_r = \Psi_r \left[((\tilde{\chi}_r \tilde{\chi}_r^T) \otimes \Omega_r^{-1}) vecC_r + \nu_r vecB \right]_r$.

Perceba que o termo envolvendo V^{-1} é também função de A_k. Excluindo o termo em V^{-1}, a expressão anterior está na forma de densidades normais independentes em A_r. As densidades podem ser reescritas como:

$$\therefore\ p(A_1,...,A_k \mid \rho, \lambda_{\neq A}, Y) \propto \left|V_k^{-1}\right|^{\frac{1}{2}} \exp\{-\frac{1}{2}\xi_q^T V_{k\rho_t}^{-1}\xi_q\} \times \prod_r N(\alpha_r, \Psi_r) \times g(A_r) \qquad (42)$$

As médias das densidades acima para A_r são médias ponderadas pela matriz de correlação dentro de cada regime e pelas estimações a priori de A_r. Um método de rejeição/não rejeição pode ser usado para gerar matrizes de regressão independentemente. Os candidatos A_r são gerados a partir de uma densidade normal até que eles caiam dentro de uma região estável.

D) Passo 4: Gerando os parâmetros de variância e covariância ($\Omega^{(c+1)} \leftarrow \rho^{(c+1)}, \mu^{(c+1)}, A^{(c+1)}, P^{(c)}$)

Nesta seção, em conformidade com as demais seções, Ω_r representa um dos valores possíveis de Ω_{ρ_t}. É mais conveniente gerar o inverso das matrizes de variância-covariância do que tentar gerar Ω_r diretamente.

Desta forma, os dados a priori deveriam ter densidades de Wishart com parâmetros η_r e F^{-1}, onde F é uma matriz diagonal com elementos na diagonal iguais a $\eta_r s_i^2$, onde s_i^2 é o erro da variância dado a priori para cada uma das i-ésimas séries. η_r pode ser interpretado como o número equivalente de observações a priori em cada regime. A geração das matrizes

não apresenta problemas de estabilidade, tal que a distribuição a priori precisa ser difusa, assim η_r será provavelmente menor que ν_r. As probabilidades a priori seriam assim definidas como $p(\Omega_1^{-1},...,\Omega_k^{-1}) = \prod_{r=1}^{k} W_m\left(\eta_r, F_r^{-1}\right) \times h\left(\Omega_1,...,\Omega_k\right)$, onde $h\left(\Omega_1,...,\Omega_k\right)$ captura as restrições a priori.

As matrizes podem ser geradas conjuntamente de:

$$
\begin{aligned}
& p(\Omega_1^{-1},...,\Omega_k^{-1} \mid \rho, \lambda_{\neq\Omega}, Y) \propto L(Y \mid \rho, \lambda).p(\Omega_1^{-1},...,\Omega_k^{-1}) \\
& \propto \left|V_k^{-1}\right|^{\frac{1}{2}} \exp\{-\frac{1}{2}\xi_q^T V_k^{-1} \xi_q\} \times \left|\Omega_k^{-1}\right|^{\frac{n_k-1}{2}} \exp\{-\frac{1}{2} vec\varsigma_k^T \left(I_{n_k-1} \otimes \Omega_k^{-1}\right) vec\varsigma_k\} \times \\
& \times \prod_{r \neq k} \left|\Omega_r^{-1}\right|^{\frac{n_r}{2}} \exp\{-\frac{1}{2} vec\varsigma_r^T \left(I_{n_r} \otimes \Omega_r^{-1}\right) vec\varsigma_r\} \times p(\Omega_1^{-1},...,\Omega_k^{-1})
\end{aligned}
$$

$$
\begin{aligned}
& \propto \left|V_k^{-1}\right|^{\frac{1}{2}} \exp\{-\frac{1}{2}\xi_q^T V_k^{-1} \xi_q\} \times \left|\Omega_k^{-1}\right|^{\frac{n_k+\eta_k-m-2}{2}} \exp\{-\frac{1}{2} tr\left((\varsigma_k \varsigma_k^T + F_k)\Omega_k^{-1}\right)\} \times \\
& \times \prod_{r \neq k} \left|\Omega_r^{-1}\right|^{\frac{n_r+\eta_r-m-1}{2}} \exp\{-\frac{1}{2} tr\left((\varsigma_r \varsigma_r^T + F_r)\Omega_r^{-1}\right)\} \times h(\Omega_1,...,\Omega_k)
\end{aligned}
$$

$$
\begin{aligned}
& \therefore\ p(\Omega_1^{-1},..., \Omega_k^{-1} \mid \rho, \lambda_{\neq\Omega}, Y) \\
& \propto \left|V_k^{-1}\right|^{\frac{1}{2}} \exp\{-\frac{1}{2}\xi_q^T V_k^{-1} \xi_q\} \times W_m\left(n_k + \eta_k - 1, (\varsigma_k \varsigma_k^T + F_k)^{-1}\right) \times \\
& \times \prod_{r \neq k} W_m\left(n_k + \eta_k, (\varsigma_r \varsigma_r^T + F_r)^{-1}\right) \times h(\Omega_1,..., \Omega_k)
\end{aligned} \tag{43}
$$

Assim, as matrizes podem ser geradas independentemente de uma densidade de Wishart. Ω_k^{-1} pode ser gerada via Metropolis-Hasting com uma densidade Wishart como candidata, isto é, geramos um candidato Ω_k^{-1}, $\Omega_k^{-1*} \sim W_m\left(n_k + \eta_k - 1, (\varsigma_k \varsigma_k^T + F_k)^{-1}\right)$ e o aceitamos com probabilidade $\min\{\sqrt{\frac{\left|V_k^{-1*}\right|}{\left|V_k^{-1(c)}\right|}} \exp\{-\frac{1}{2}\xi_q^T (V_k^{-1*} - V_k^{-1(c)})\xi_q\}, 1\}$, para $\Omega_k^{-1*} = \Omega_k^{-1(c+1)}$, conforme já demonstramos anteriormente.

E)Passo 5: Gerando as probabilidades de transição ($P^{(c+1)} \leftarrow \rho^{(c+1)}, \mu^{(c+1)}, A^{(c+1)}, \Omega^{(c+1)}$)

A matriz de probabilidade de transição pode ser gerada de:

$$p(P \mid \rho, \Theta, Y) \propto p(\rho_q \mid P).\prod_{t=q+1}^{N} p(\rho_t \mid \rho_{t-1}, P).p(P) \qquad (44)$$

Suponha que ρ representa n_{ij} transições do regime i para o regime j. Defina os dados a priori para p_{ij} com uma distribuição Beta (m_{ij} + *1*, m_{ii} +*1*), onde m_{ij} é o número de transições a priori. Então;

$$p(P \mid \rho, \Theta, Y) \propto p(\rho_q \mid P).\prod_{t=q+1}^{N} p_{ij}^{n_{ij}+m_{ij}} \times \prod_{i=1}^{k} (1 - \sum_{i \neq j} p_{ij})^{n_{ij}+m_{ij}}$$

$$\propto p(\rho_q \mid P) \times \prod_{i=1}^{k} \left((\prod_{i \neq j} p_{ij}^{n_{ij}+m_{ij}})(1 - \sum_{i \neq j} p_{ij})^{n_{ij}+m_{ij}} \right)$$

Na expressão acima, $p(\rho_q \mid P)$ é uma função da cada p_{ij}. Ao simular a densidade conjunta acima através dos passos do algoritmo de Metropolis-Hasting, as densidades Beta são geradas como as densidades candidatas, até que se convirja para um determinado critério. Seguindo o mesmo argumento apresentado nas seções anteriores, o candidato gerado P, P^*, a partir de $p_{ij}^* \sim Beta(n_{ij} + m_{ij} + 1,, n_{ii} + m_{ii} + 1)$ e $p_{ii}^* = 1 - \sum_{i \neq j} p_{ij}$, será aceito ou rejeitado com probabilidade $\min\{\frac{\pi^* / q^*}{\pi^c / q^c}, 1\}$, onde $\pi / q = p(\rho_q \mid P) \left(\frac{p_{ii}}{\prod (1 - p_{ij})} \right)^{n_{ii}+m_{ii}}$.

4. SÉRIES ECONÔMICAS UTILIZADAS

4.1. Notas sobre a estacionariedade das séries utilizadas

É usual a utilização de testes de Dickey-Fuller para a avaliação de presença de raiz unitária em modelos de séries temporais. Estes procedimentos são baseados em métodos de mínimos quadrados com a presença de choques gaussianos. Porém, muitas séries econômicas são afetadas por infreqüentes e importantes eventos, tais como guerras, desastres naturais, choques do petróleo, mudanças de política, entre outros, que acabam por determinar um componente não-gaussiano no comportamento das séries econômicas. Algumas séries, como as taxas de juros e de câmbio, possuem distribuições com caudas grossas, o que indica desvios da série em relação a um comportamento de uma distribuição normal padrão.

Por isto, é mister que se considere tais inovações nos procedimentos de estimação e inferência de modelos com a utilização destas séries, e uma das primeiras propostas para tal problema foi a utilização da estimação adaptativa baseada em técnicas não-paramétricas.[11] Sob hipóteses apropriadas os testes baseados em estimação adaptativa usando métodos de Kernel não-paramétricas puderam ser construídos, apesar da grande dificuldade de sua aplicação (Lima e Xiao, 2003).

Uma abordagem alternativa foi a utilização da estimação M.[12] As funções critério são assumidas ser conhecidas e as inferências associadas são geralmente eficientes quando as verdadeiras funções de probabilidade são usadas. Na prática, a distribuição do erro é desconhecida, o que enseja a utilização de uma função critério que possui característica similar à distribuição dos dados. Portanto, através do uso de uma classe intermediária de procedimentos de teste para raiz unitária, estes são mais eficientes do que os métodos tradicionais baseados em estimações OLS, no caso de séries com distribuição com caudas grossas, além de serem de mais fácil implementação que os testes baseados em estimação inteiramente adaptativa via métodos não-paramétricos (Lima e Xiao, 2003).

O teste proposto nesta seção segue Lima e Xiao (2003) e é baseado na estimação parcialmente adaptativa de um modelo *Augmented Dickey-Fuller* (ADF). Os autores demonstram que a distribuição limite da correspondente estatística t é uma mistura entre uma distribuição Dickey-Fuller e uma distribuição normal padrão independente da anterior. Os autores utilizam a família de distribuições *t-student* porque ela representa uma dimensão

[11] Para maiores detalhes ver Beelders (1998).
[12] Ver Phillips (1995), Xiao (2001) e Koencker e Xiao (2003).

importante do espaço de distribuições, incluindo a distribuição normal como um caso limite e a distribuição de *Cauchy* como caso especial.

Seguindo Dickey e Fuller (1979), o modelo de regressão ADF conhecido é especificado por:

$$\Delta y_t = \rho y_{t-1} + \sum_{j=1}^{k} \psi_j \Delta y_{t-j} + \varepsilon_t \qquad (45)$$

Na presença de raiz unitária, ρ = 0 na regressão ADF acima.

Geralmente, inclui-se um componente de tendência determinística na regressão ADF, sendo definida como:

$$\Delta y_t = \gamma' x_t + \rho y_{t-1} + \sum_{j=1}^{k} \psi_j \Delta y_{t-j} + \varepsilon_t \qquad (46)$$

onde x_t é um componente determinístico conhecido e γ é um vetor de parâmetros desconhecidos. A variável x_t pode assumir ter um termo constante $x_t = 1$ ou uma tendência linear no tempo $x_t = (1,t)'$.

Queremos testar a hipótese de raiz unitária baseada na estimação de ρ. No caso simples de ε_t ser normalmente distribuído, dadas as observações de y_t, os estimadores de máxima verossimilhança de ρ e $\langle \psi_j \rangle_{j=1}^{k}$ são simplesmente os estimadores de mínimos quadrados (OLS) obtidos pela minimização da soma dos resíduos ao quadrado. Na ausência de gaussianidade em ε_t, os estimadores de mínimos quadrados são inconsistentes. Neste caso, Huber (1973) introduziu uma classe de estimadores M, os quais geralmente têm boas propriedades em várias distribuições estatísticas. Os M-estimadores são obtidos a partir da resolução do problema de recolocação da função critério quadrática com alguma generalização da função critério φ na estimação por OLS. No caso de φ ser a verdadeira função log densidade dos resíduos, o M-estimador é o estimador de máxima verossimilhança.[13]

Para introduzir o teste de raiz unitária proposto, consideremos o M-estimador para $\left(\gamma, \rho, \langle \psi_j \rangle_{j=1}^{k}\right)$ definido como a solução do seguinte problema de otimização:

[13] Para maiores detalhes ver Lima e Xiao (2003), Huber (1973) e Postcher e Prucha (1986).

$$\left(\gamma, \rho, \left\langle \psi_j \right\rangle_{j=1}^{k}\right) = \arg\max Q\left(\gamma, \rho, \left\langle \psi_j \right\rangle_{j=1}^{k}\right) \tag{47}$$

onde

$$Q\left(\gamma, \rho, \left\langle \psi_j \right\rangle_{j=1}^{k}\right) = \sum_{t=k+1}^{n} \varphi\left(\Delta y_t - \gamma' x_t - \rho y_{t-1} - \sum_{j=1}^{k} \psi_j \Delta y_{t-j}\right) \tag{48}$$

para qualquer função critério φ. Quando φ é a verdadeira função log densidade de ε_t, $Q\left(\gamma, \rho, \left\langle \psi_j \right\rangle_{j=1}^{k}\right)$ é a função log-verossimilhança e o estimador $\left(\gamma, \rho, \left\langle \psi_j \right\rangle_{j=1}^{k}\right)$ dado por (47) é o estimador de máxima verossimilhança.

O M-estimador é assintoticamente eficiente quando é um estimador de máxima verossimilhança. Na prática, mesmo que a distribuição exata das inovações seja desconhecida, se os dados possuírem similar comportamento nas caudas, tal como a função de densidade usada na estimação, então a inferência baseada nos métodos aqui propostos ainda terá ótimo desempenho de amostragem. Desta forma, é importante selecionar uma função critério que tenha características similares aos dados da distribuição (Lima e Xiao, 2003).

A M-estimação parcialmente adaptativa considera uma família paramétrica de distribuições. Cada membro desta família é indexado a algum parâmetro de adaptação. Dada a amostra observada, é possível estimar os parâmetros de adaptação de tal forma que a função de densidade que melhor aproxima a distribuição dos dados (dentro da família paramétrica) é selecionada. Na literatura, diferentes classes de distribuição foram estudadas com o intuito de se verificar sua aderência à estimação parcialmente adaptativa.[14] Considerando a bem conhecida característica das "caudas grossas" presentes em séries econômicas e financeiras, a literatura advoga em prol da utilização de distribuições *t-student* na estimação parcialmente adaptativa. A distribuição *t-student* é uma importante classe de distribuição porque possui a distribuição Cauchy como um caso especial e a distribuição normal como um caso limite, além de possuir uma variedade de aplicações em dados econômicos. Os parâmetros de adaptação desta distribuição são dependentes dos parâmetros de escala e "fineza" das caudas, os quais podem ser facilmente estimados a partir dos dados. Os estimadores parcialmente adaptativos desta classe de distribuição são razoavelmente robustos (Postcher e Prucha, 1986).

Dado o modelo ADF proposto em (46), na presença de inovações *t*-distribuídas, a função log-verossimilhança será dada por:

[14] Ver Postcher e Prucha (1986) e Phillips (1995).

$$L = cons\tan te + \frac{n}{2}\ln\Theta - \frac{\nu+1}{2}\sum_{t=j+2}^{n}\ln\{1+\frac{\Theta}{\nu}\left(\Delta y_t - \gamma' x_t - \rho y_{t-1} - \sum_{j=1}^{k}\psi_j \Delta y_{t-j}\right)^2\} \qquad (49)$$

onde o parâmetro Θ mede o espalhamento da distribuição dos distúrbios e ν é o grau de liberdade que mede a "fineza" da cauda. Grandes valores de ν correspondem a caudas mais finas na distribuição. Para valores dados de ν e Θ, denotando Θ/ ν = θ, o estimador de máxima verossimilhança de $\left(\gamma, \rho, \langle\psi_j\rangle_{j=1}^{k}\right)$ é a solução do seguinte problema de otimização:

$$\min \sum_{t=j+2}^{n}\ln\{1+\vartheta\left(\Delta y_t - \gamma' x_t - \rho y_{t-1} - \sum_{j=1}^{k}\psi_j \Delta y_{t-j}\right)^2\} \qquad (50)$$

Seguindo Postcher e Prucha (1986), deixe Ц ser o estimador de Π e Θ/ ν = θ ser o parâmetro de adaptação da distribuição *t*, então, o M-estimador adaptativo para o modelo ADF será:

$$\Pi = Ц + \left[\frac{1}{n}\sum Z_t'(w_t - 2\vartheta w_t \varepsilon_t^2)Z_t\right]^{-1}\frac{1}{n}\sum Z_t' w_t \varepsilon_t \qquad (51)$$

onde

$$w_t = (1+\vartheta\varepsilon_t^2)^{-1} \qquad Z_t = (x_t', y_{t-1}, \Delta y_{t-1}, \ldots, \Delta y_{t-k})'$$

$$\varepsilon_t = \Delta y_t - Z_t Ц \qquad \Pi = \left(\gamma, \rho, \langle\psi_j\rangle_{j=1}^{k}\right)$$

Na análise prática , os parâmetros ν e Θ são desconhecidos e necessitam ser estimados. Nós aplicaremos uma estimação em dois passos para o estimador parcialmente adaptativo de $\left(\gamma, \rho, \langle\psi_j\rangle_{j=1}^{k}\right)$. O primeiro passo consistirá em uma estimação preliminar dos parâmetros ν e Θ. Em seguida, recolocamos θ em (51) e fazemos uma segunda estimação para $\left(\gamma, \rho, \langle\psi_j\rangle_{j=1}^{k}\right)$. Na presença de distúrbios gaussianos, os valores de ν e Θ perderão seu significado. Caso contrário, poderão ser interpretados, respectivamente, como estimadores da medida da "fineza" da cauda e espalhamento dos distúrbios da distribuição (Lima e Xiao, 2003).

Postcher e Prucha (1986) discutiram a estimação dos parâmetros de adaptação ν e Θ. Em particular, se denotarmos $E(|u_t|^k)$ como σ_k, então, para ν > 2, teremos:

$$\frac{\sigma_2}{\sigma_1^2} = \frac{\pi}{\nu - 2}\frac{\Gamma(\nu/2)^2}{\Gamma[(\nu-1)/2]^2} = \rho(\nu) \tag{52}$$

e

$$\Theta = \frac{1}{\pi}\left[\frac{\sigma_1^2\Gamma(\nu/2)^2}{\nu\Gamma[(\nu-1)/2]^2}\right]^{-1} = q(\nu,\sigma_1) \tag{53}$$

Os autores mostram que $\rho(\nu)$ é analítico e monotonicamente decrescente em (2, ∞) com $\rho(2)$ = ∞ e $\rho(\infty)$ = π/2. Assim, dado o estimador σ_1 e σ_2, ν pode ser estimado invertendo-se $\rho(\nu)$ em (53) e desta forma, o estimador de θ pode ser obtido por:

$$\vartheta = \frac{1}{\pi}\left[\frac{\sigma_1^2\Gamma(\nu/2)^2}{\Gamma[(\nu-1)/2]^2}\right]^{-1} = \frac{q(\nu,\sigma_1)}{\nu} \tag{54}$$

Para a estimação de σ_1 e σ_2 podemos usar a seguinte função geradora de momentos:

$$\sigma_k = \frac{1}{n}\sum_t |u_t|^k \tag{55}$$

Seguindo Lima e Xiao (2003), indicaremos os passos que seguimos para avaliar a presença de raiz unitária a partir da utilização da estimação parcialmente adaptativa:

Passo 1: Estimamos os resíduos ε_t a partir de uma regressão ADF preliminar:

$$\Delta y_t = \gamma' x_t + \rho y_{t-1} + \sum_{j=1}^{k}\psi_j \Delta y_{t-j} + \varepsilon_t$$

Passo 2: Obtemos os parâmetros adaptativos levando em conta a classe de distribuições *t-student*. Usamos os passos dados para a estimação dos parâmetros ν e Θ utilizando os resíduos estimados no passo 1.

Passo 3: Selecionamos uma função critério seguindo o modelo adotado por Lima e Xiao (2003), sabendo que este modelo se adequou bem aos dados testados neste trabalho. Assim, a função critério escolhida tem a forma:

$$\varphi(\varepsilon) = \ln\{1 + \vartheta[\varepsilon]^2\}$$

O M-estimador para $\left(\gamma, \rho, \langle \psi_j \rangle_{j=1}^{k}\right)$ é obtido pela relação:

$$\max \sum_{t=2}^{n} \varphi\left(\Delta y_t - \gamma' x_t - \rho y_{t-1} - \sum_{j=1}^{k} \psi_j \Delta y_{t-j}\right)$$

O valor obtido neste problema de otimização numérica corresponde a estatística *t*.

Passo 4: Obtemos o λ^2 a partir do estimador de variância de ε_t e $\varphi'(\varepsilon)$. Assim,

$$\sigma^2 = \frac{1}{n-k-1}\sum_{t=2}^{n} \varepsilon_t^2 \text{ e } \omega_\varphi^2 = \frac{1}{n-k-1}\sum_{t=k+1}^{n} \varphi'(\varepsilon_t^2).$$

Como $\operatorname{cov}(\varepsilon_t, \varphi'(\varepsilon_t)) = \frac{1}{n-k-1}\sum_{t=2}^{n} \varepsilon_t \varphi'(\varepsilon_t)$, devemos estimar a seguinte relação:

$$\lambda^2 = \frac{\sigma_{u\varphi}^2}{\omega_\varphi^2 \omega_u^2}$$

sendo $\omega_u^2 = \frac{\sigma^2}{A(1)^2}$ e $\sigma_{u\varphi}^2 = \frac{\operatorname{cov}(\varepsilon_t, \varphi'(\varepsilon_t))}{A(1)}$, onde $A(1) = 1 - \sum_j \psi_j$.

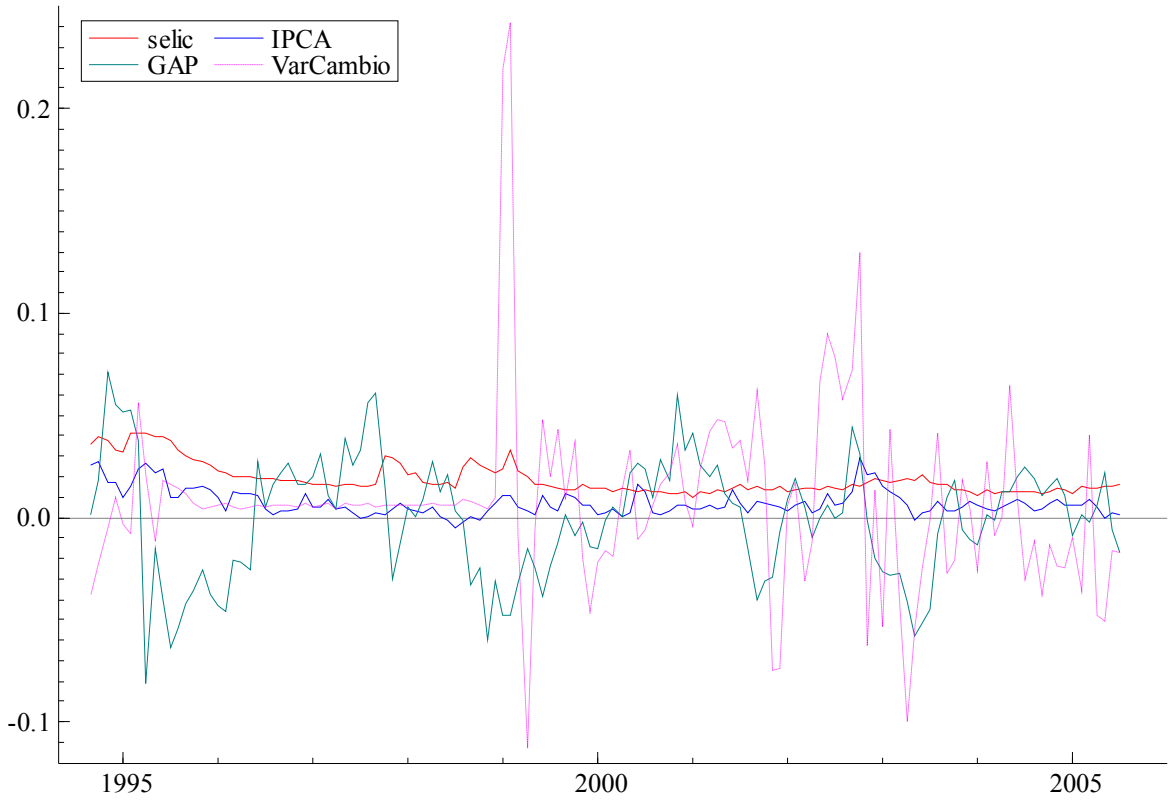
Usando os valores estimados para λ^2, avaliamos os valores críticos da tabela 1 constante do anexo 1.

Utilizamos as séries do hiato do produto (definido como o logaritmo da razão do produto industrial dessazonalizado e o produto industrial potencial obtido pelo filtro HP), da taxa de juros Selic (obtida a partir do logaritmo das variações mensais), da inflação como

sendo o IPCA o deflator escolhido (definido como o logaritmo da razão da inflação no período *t* e no período *t + 1*), e da variação do câmbio (definido como o logaritmo da variação da taxa de câmbio comercial – Dólar americano(U$)/Real (R$) – fim de período). [15]

No caso das séries econômicas utilizadas, aplicamos os testes de estacionariedade para as séries do *gap* do produto[16], inflação, taxa de juros Selic e desvalorização nominal da taxa de câmbio, cujas séries são mostradas no gráfico 1 abaixo.[17] A priori, devemos avaliar as estatísticas descritivas das séries utilizadas com o intuito de averiguar a presença de caudas grossas em suas respectivas distribuições. Os resultados estão reportados na tabela 1.

Gráfico 1: Séries históricas das variáveis utilizadas



Podemos ver que, com exceção da série do *gap* do produto, todas as demais séries possuem curtose maior que 3. Isto significa que as séries possuem excesso de curtose, ou seja, sua distribuição é leptocúrtica, possuindo "caudas grossas". Ademais, o teste de Jarque-Bera

[15] Dados mensais para o período 08/1994 até 08/2005.
[16] Apesar da discussão acerca do uso de filtros para a obtenção do hiato do produto, como não há consenso na literatura sobre a aplicação de uma única metodologia, adotamos a medida padrão do hiato do produto calculado a partir do uso do filtro de Hodrick-Prescott (HP).
[17] Dados obtidos nos relatórios do Banco Central do Brasil (BCB) e IPEA-DATA 4.0.

para a hipótese de normalidade rejeitou a hipótese nula de que a série possui uma distribuição normal, com exceção para a série do hiato do produto.

O parâmetro de *thickness* termina por confirmar a suspeição de presença de "caudas grossas" nos dados utilizados neste trabalho. Todas as séries possuem parâmetros próximos do valor zero, o que indica uma medida para a presença de "caudas grossas". No caso limite do parâmetro ser próximo de um valor infinitesimal tem-se a distribuição normal.

TABELA 1: Estatística descritiva das séries utilizadas

Séries	Parâmetro de thickness	Curtose	Jarque-Bera
Gap do Produto (x_t)	-	2.32	4.65
Inflação (π_t)	4.31	4.529	45.447*
Taxa de juros (i_t)	2.91	8.39	246.1062*
Desvalorização nominal do câmbio (Δe_t)	2.94	13.89	735.07*

* Rejeição da hipótese nula ao nível de significância de 1%.

A tabela abaixo apresenta os valores para os testes ADF e P-ADF para as séries utilizadas.

TABELA 2: Análise de raiz unitária

Séries	θ	λ^2	ADF	P-ADF
Gap do Produto (x_t)	-	-	-6.6813*	-
Inflação (π_t)	2.3195	0.7550	-3.9156*	-4.8120*
Taxa de juros (i_t)	15.14	0.3936	-2.817**	-3.4402*
Desvalorização nominal do câmbio (Δe_t)	0.0684	0.4707	-7.450*	-4.7296*

* Rejeição da hipótese nula ao nível de significância de 1%.
**Não rejeição da hipótese nula-Presença de raiz unitária.
Obs.: testes ADF feitos com intercepto e tendência.

A tabela acima demonstra situações interessantes. Em primeiro lugar, a série da inflação (π_t) rejeitou a hipótese nula da presença de raiz unitária tanto no teste ADF convencional, quanto no teste P-ADF. A presença de "caudas grossas" nesta distribuição não distorceu os valores para definição da não-estacionariedade da série. Para a série do *gap* (x_t)

foi aplicado apenas o teste ADF convencional, corroborando o resultado da literatura acerca da estacionariedade destes dados para a economia brasileira.

Porém, os dados relativos á taxa de juros selic (i_t) demonstraram resultados divergentes. Para a taxa de juros, o teste ADF indicou a presença de raiz unitária, por outro lado, o teste P-ADF rejeitou esta hipótese. O fato da série da taxa de juros ter sido sujeita a fortes choques exógenos ao longo do Plano Real pode ter gerado problemas na normalidade da distribuição dos dados, ocasionando erros por parte do cômputo da estatística de teste baseada nos métodos Dickey-Fuller. Portanto, podemos afirmar que tal série é estacionária.

5. APLICAÇÃO À POLÍTICA MONETÁRIA BRASILEIRA

5.1. Estimação para economia fechada

O primeiro passo foi examinar a melhor escolha da ordem do componente auto-regressivo do modelo VAR (p). Esta escolha foi feita com base nos critérios de informação de Akaike (AIC), de Schwartz (SC) e no de Hannah-Quinn (HQ), sendo selecionado o modelo que possuísse os menores valores para tais estatísticas. Começamos com um modelo MSM (2) – VAR (1) e fizemos várias tentativas de especificação, prefixando a presença de dois regimes. Em seguida, estimamos o melhor modelo com o auxílio do software *PcGets 3.0*. Os resultados são reportados à tabela 3, sendo MSIAH (2) – VAR (1) o modelo selecionado.

TABELA 3: Formas funcionais do modelo e critérios de seleção estimados

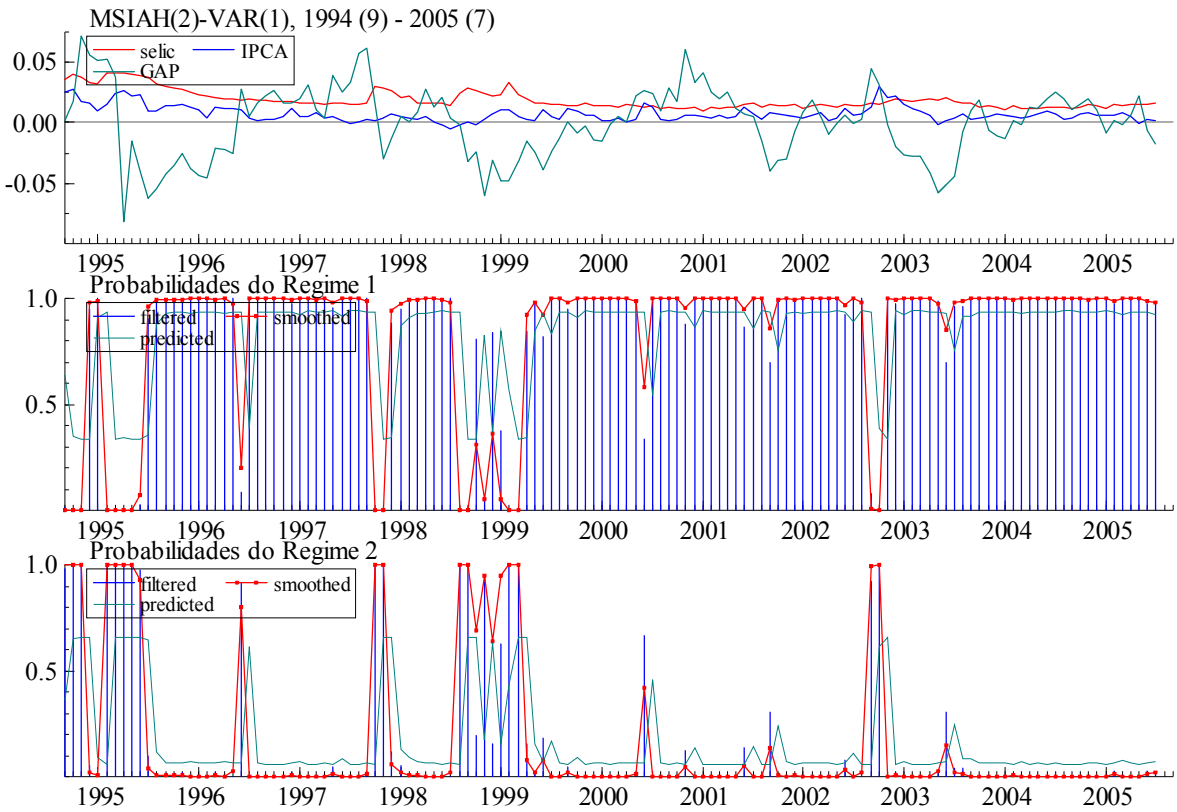
MS (2) - VAR (1)			
	AIC	*H-Q*	*SC*
MSIA	7.325	7.4711	7.4944
MSIH	7.0213	7.2852	7.3611
MSI	8.0195	8.2288	8.5649
MSA	7.0895	7.3547	7.7423
MSAH	7.3047	7.6231	7.731
MSH	6.9982	7.2327	7.5624
MSM	8.0371	8.2463	8.3285
MSMAH	11.7448	12.0906	12.6784
MSMA	11.6503	11.9415	12.3224
MSMH	7.0178	7.2817	7.6673
MSIAH	6.747	7.098	7.603

O modelo estimado possui regimes diferenciados tanto para o intercepto e variância quanto para os parâmetros estimados. A dependência da variância quanto aos estados da cadeia de Markov foi testada através de um teste de razão de verossimilhança (LR), cuja estatística de teste obtida foi 91.75, muito superior ao valor tabelado (3.84).

Vejamos as probabilidades suavizadas, filtradas e previstas para os regimes, comparando os resultados estimados com os fatos estilizados à luz da história econômica brasileira. O gráfico 1 nos mostra a evolução das séries do hiato, da inflação e da taxa de juros nominal, bem como as probabilidades para os regimes do modelo MSIAH (2) – VAR (1). Tendo como base os trabalhos de Hamilton (1989) e Krolzig (1997b), inicialmente vamos

considerar um modelo contendo dois regimes; o regime $s_t = 1$ indicará estabilidade econômica, enquanto $s_t = 2$ implicará em períodos de crises internas e externas.

Gráfico 2: Probabilidades para os regimes do modelo MSIAH (2) – VAR (1) adotado à regra de política monetária no período do Plano Real



O gráfico das probabilidades nos conta o período do Plano Real e suas crises internas e externas. Em 1994, os objetivos iniciais de estabilização via transferência da economia para uma nova moeda "limpa" do componente inflacionário inercial requereu a implementação de um regime de câmbio fixo que expôs o país ao processo especulativo internacional, que combinado com o desajuste fiscal, gerou fortes oscilações das taxas de juros nominais, conforme vemos a trajetória temporal da taxa de juros Selic. A vulnerabilidade a crises de confiança e a comportamentos irracionais (bolhas financeiras) resultou a grandes fugas de capitais externos e forte pressão no mercado cambial, o que levava o governo a subir as taxas de juros internas com a pretensão de atrair um novo influxo de capitais e manter o regime de bandas cambiais, tido como imprescindível para o mecanismo de estabilização econômica. O crescimento econômico herdado do final de 1996

prosseguiu em 1997 sustentado, principalmente, pelo consumo. O volume de importações tornara-se muito alto, gerando um déficit comercial de aproximadamente US$ 1 bilhão no mês de março alimentando projeções de que, no ano, o saldo negativo ultrapassaria os US$ 10 bilhões. Ao final do ano de 1997 (novembro), com a deterioração das condições internacionais em razão da crise asiática, as taxas de juros voltaram a subir por causa da necessidade de garantir um cupom cambial mais alto e evitar assim a perda de capitais externos ocorridas a partir do aumento do risco-país e do atraso cambial na rolagem dos títulos brasileiros. O governo tentou debelar a crise com o aumento de juros referenciais e com a implantação de um pacote fiscal (o pacote "51"), utilizando medidas para aumentar receitas, cortar gastos, restringir importações e incentivar as exportações. O período de ajustamento pós-crise asiática ocorreu quando o Banco Central do Brasil (BCB) reduziu a taxa de juros e a pressão sobre o câmbio. Porém, ocorreu o retorno do nervosismo nos mercados internacionais com a declaração da moratória russa por volta de agosto de 1998, fato acrescido do processo eleitoral brasileiro. As eleições transcorreram em um ambiente economicamente instável, com juros bastante elevados e acompanhados por diversas medidas para controlar o câmbio. As autoridades econômicas argumentavam que a política cambial poderia se manter inalterada com uma taxa de juros alta e o ajuste fiscal acelerado seria suficiente, como nas outras crises, para atravessar o período de maior turbulência internacional.

Este processo culminou, em 1998, em uma situação externa completamente desfavorável a partir da moratória russa e perda de confiança dos investidores nos países emergentes, fazendo com que o governo elevasse substancialmente as taxas de juros de curto prazo e anunciasse a permanência de um aperto fiscal tendo em vista a redução da absorção interna e a melhora no saldo do balanço de pagamentos. Mesmo com certos resultados obtidos, a baixa confiança dos mercados permaneceu até que o Brasil substituiu, em 15 de janeiro de 1999, o regime de bandas cambiais por um regime cambial de flutuação suja. Além disso, foi introduzido o sistema de metas de inflação, tendo o Índice de Preços ao Consumidor Amplo (IPCA) o status de âncora nominal deste sistema.

Em abril de 2000 ocorreram os problemas advindos da economia Argentina com conseqüente desconfiança dos mercados internacionais de que o Brasil poderia ser atingido. O Banco Central interveio no mercado cambial vendendo dólares e aumentando a taxa de juros referencial na tentativa de evitar a fuga de capitais.

Em 2001 as dificuldades da crise energética e do atentado terrorista aos EUA implicaram em mais uma crise interna combinada com uma crise mundial. As taxas de juros

novamente foram aumentadas com o fito de suavizar a fuga de capitais e compensar a subida da taxa de câmbio, evitando desvios da meta inflacionária estabelecida.

Em meados de 2002 o Banco Central subiu a taxa de juros de curto prazo (selic) para evitar as incertezas presentes no processo eleitoral brasileiro. Esta incerteza quanto ao novo governo disparou a fuga de capitais, ocasionando um *overshooting* da taxa de câmbio que atingiu seu valor mais alto em todo o Plano Real: R$ 3.80/dólar em outubro de 2002. Esta pressão cambial resultou em um aumento generalizado dos índices de preços da economia brasileira. No âmbito externo, o baixo crescimento mundial, o surgimento de problemas contábeis em grandes empresas americanas, as crises observadas em mercados emergentes e a perspectiva de mais uma guerra no Golfo provocaram aumento na aversão ao risco, com conseqüente diminuição na liquidez internacional.

Em 2004, o governo do presidente Lula verificou um bom crescimento da economia que desencadeou um aumento do nível de inflação devido, grande parte, ao aquecimento da economia via aumento do comércio exterior e à busca de recomposição das margens de lucro por parte do setor empresarial. Nosso modelo identificou este período como de crise (com baixa probabilidade de ocorrência) devido à subida da taxa de inflação acompanhada da taxa de juros, apesar do hiato do produto apresentar relevante crescimento. A tabela 4 sintetiza as datas de cada regime e suas respectivas probabilidades de ocorrência:

TABELA 4: Datas de cada regime e suas respectivas probabilidades de ocorrência

Regime 1	**Regime 2**
Ausência de crises	*Presença de crises*
1994:9 - 1994:9 [0.9996]	1994:10 - 1994:11 [0.9991] (México)
1994:12 - 1995:2 [0.9728]	1995:3 - 1995:5 [0.9567] (México)
1995:6 - 1997:10 [0.9792]	1997:11 - 1997:11 [1.0000] (Ásia)
1997:12 - 1998:8 [0.9574]	1998:9 - 1998:10 [1.0000] (Rússia)
1998:11 - 1999:2 [0.9838]	1999:3 - 1999:4 [1.0000] (Desvalorização)
1999:5 - 1999:10 [0.9290]	2000:7 - 2000:7 [0.9938] (Argentina)
1999:11 - 2000:6 [0.9890]	2001:7 - 2001:7 [0.5367] (Crise energética)
2000:8 - 2001:6 [0.9925]	2002:11 - 2002:11 [1.0000] (Eleições)
2001:8 - 2002:10 [0.9593]	
2002:12 - 2005:2 [0.9705]	

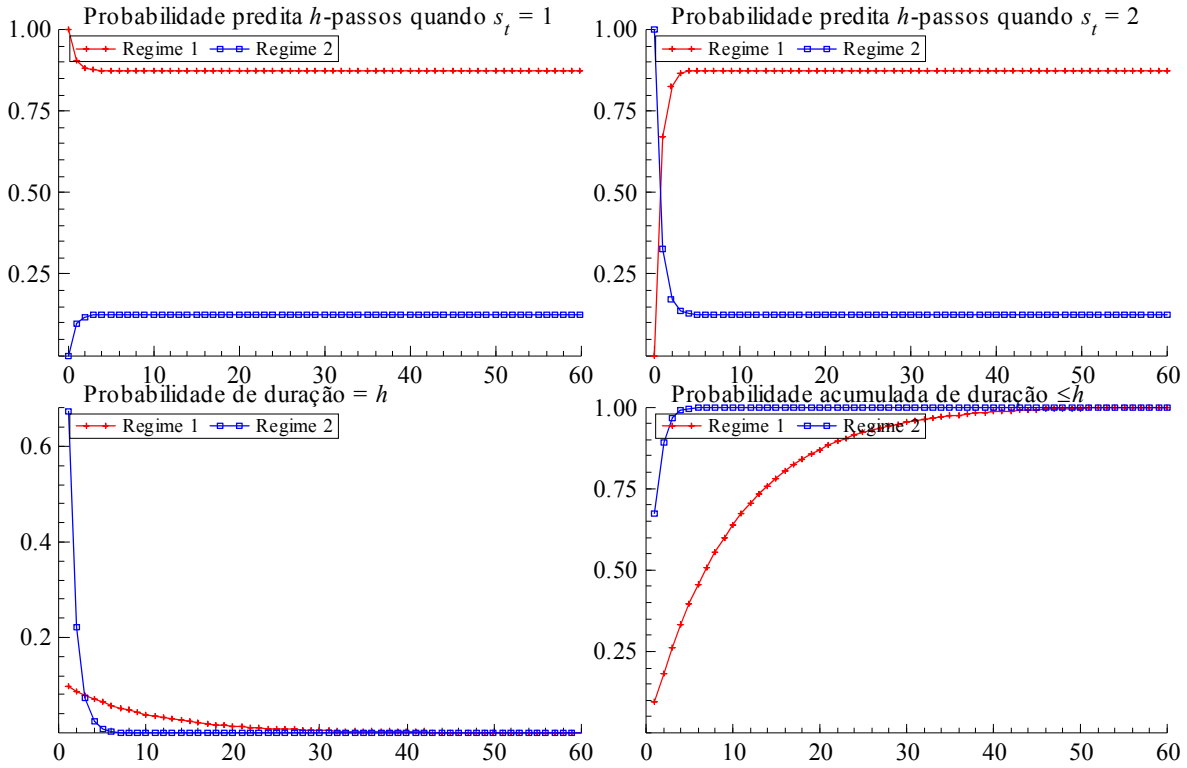
Nota: Probabilidades entre colchetes.

Quanto às probabilidades de transição, a matriz de transição (linha estocástica) com distribuição estacionária obtida no modelo de dois regimes é a seguinte:

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} 0.9395 & 0.0605 \\ 0.3390 & 0.6610 \end{bmatrix}$$

A matriz nos indica que a probabilidade de permanência em períodos de estabilidade econômica (p_{11} = 0.9395) é maior do que a de crises agudas (p_{22} = 0.6610). A permanência em períodos de estabilidade econômica é bem maior do que em crises agudas, mesmo sabendo que a economia brasileira enfrentou neste período várias crises de curta duração. Em outras palavras, a probabilidade de manutenção das taxas de juros em patamares médios dado que a economia esteja em períodos de ausência ou presença de crises é muito significativo, pois a probabilidade de ocorrência de crises não é desprezível. No entanto, a probabilidade de se estar em um período de fortes crises e mudar para o de ausência de crises é mais alta do que a probabilidade de permanência neste período de fortes crises (0.3390 contra 0.3280). Este resultado é muito interessante, pois demonstra o processo de defasagem interna e externa da política monetária frente a choques adversos na economia. A probabilidade de transição da prática de juros altos pela regra de política monetária especificada apresenta-se como temporária em períodos de fortes crises, já que a maior probabilidade é de mudança para períodos com taxas de juros mais suavizadas. O gráfico 3 demonstra as probabilidades de mudança dos regimes discutidos aqui.

Gráfico 3: Probabilidades de transição preditas para cada regime $s_t = i$



No caso dos gráficos superiores vemos o comportamento das probabilidades de transição se alterando conforme o regime fixado (s_t = 1,....,N). Por exemplo, no gráfico superior à esquerda é fixado o regime 1 e feita a previsão de probabilidades h-passos à frente, vista pelo eixo das abscissas. Neste caso a probabilidade de permanência no estado 1 (regime 1) praticamente fica constante à medida que a previsão é feita vários passos à frente, ou seja, a previsão mostra probabilidade alta de permanência da economia no regime de estabilidade econômica dado que se esteja neste regime. Por outro lado, dado que a economia esteja no regime de crises econômicas, à medida que a previsão é feita, a probabilidade de que a economia mude para um regime de estabilidade econômica vai aumentando de forma acelerada nos primeiros 5 meses, estabilizando com o aumento do *forecasting*. O gráfico inferior à esquerda mostra que a probabilidade de duração do regime de estabilidade econômica é maior do que a probabilidade de duração das crises, dada a previsão h-passos à frente. Esta informação é corroborada pela tabela 5 que identifica as probabilidades estacionárias e a duração dos regimes para o modelo estimado.

TABELA 5: Probabilidades estacionárias e duração dos regimes do modelo MSIAH (2) – VAR (1)

	Nº de observações	Probabilidade	Duração

Regime 1: Estabilidade econômica	110	0.875	16.54
Regime 2: Presença de crises	21.5	0.125	2.95

Vemos que a duração do regime de estabilidade econômica é maior que a de crises. O regime de estabilidade econômica tem probabilidade de ocorrência de 87.5% com duração média de 16.54 meses. Já a presença de crises tem probabilidade de ocorrência de 12.5% com duração média de 2.95 meses.

Aplicando a mesma metodologia adotada anteriormente, estimamos um modelo MCMC com especificação MSMA(2)-VAR(1) através de 2500 amostras geradas dos parâmetros de densidade conjunta. A especificação acima foi definida porque os modelos MSIA, MSIAH e MSMAH não apresentaram solução estável para a decomposição de Choleski, o que prejudicou um pouco a comparação da primeira com a segunda parte desta seção, visto que o interessante seria conseguirmos gerar as médias via *Gibbs Sampler* de um modelo nos mesmos moldes da primeira parte.

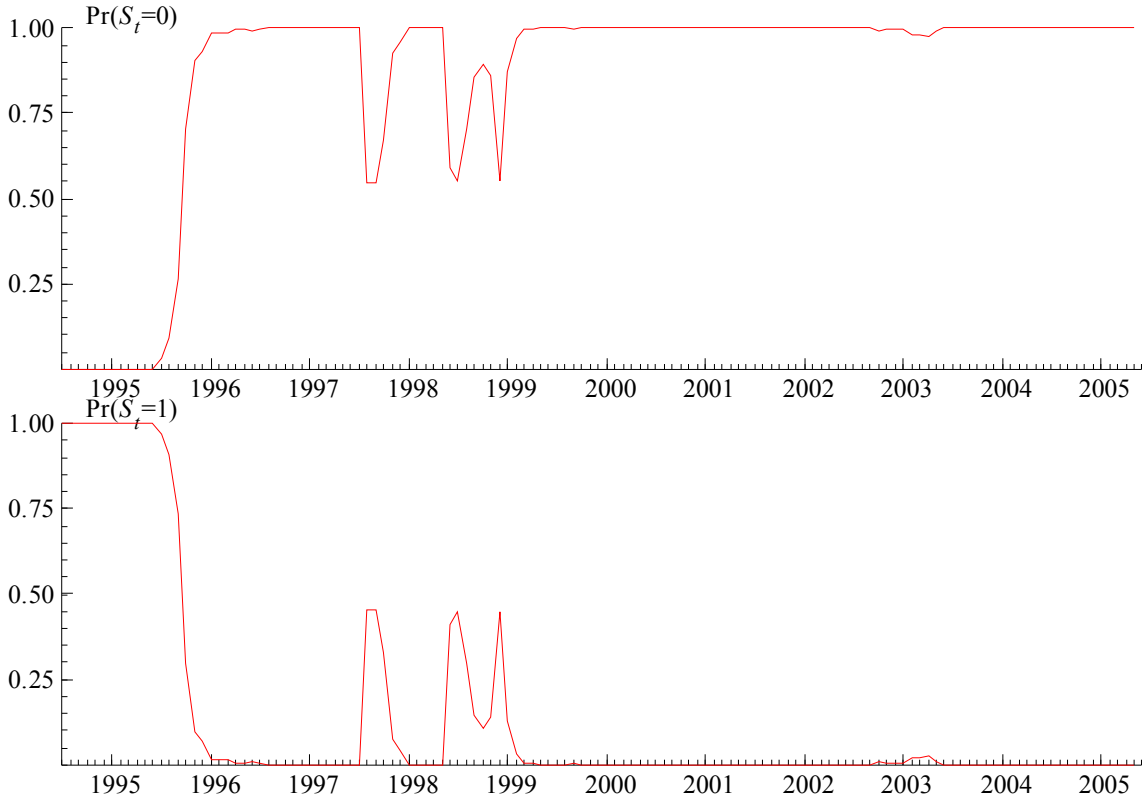
Os parâmetros das médias convergiram extremamente rápido, mesmo quando os parâmetros iniciais dados a priori eram muito distantes dos valores reais estimados, ou quando a ordem do modelo ajustado era incorreta. Os efeitos negativos dos parâmetros dados a priori eram dissipados após várias iterações. Portanto, o procedimento de MCMC pode ser esperado como uma boa estimativa para os parâmetros de média, independente dos valores iniciais, em contraste com a abordagem EM utilizada anteriormente. Além disto, as primeiras 100 amostras foram descartadas e as 2400 amostras restantes foram utilizadas para descrever os parâmetros de densidade conjunta.

Assim, estimamos um MSMA(2)-VAR(1) da seguinte regra de política monetária:

$$i_t = a\,\pi_t + \tau x_t + \psi i_{t-1} \qquad (76)$$

O gráfico 4 nos mostra as probabilidades para os regimes do modelo MSMA (2) – VAR (1) estimado pelo método MCMC proposto. Os resultados obtidos se assemelham bastante aos anteriores, demonstrando os períodos de forte crise como a crise mexicana em 1994, a crise da Ásia em 1997, a crise da Rússia em 1998, a desvalorização do câmbio em 1999 e a crise das eleições presidenciais em 2002/2003. O método MCMC identificou os mesmos regimes de crise obtidos pela estimação *Markov Switching* através do método EM.

Gráfico 4: Probabilidades para os regimes do modelo MSMA (2) – VAR (1) adotado à regra de política monetária no período do Plano Real



Os coeficientes para a média estimados, bem como os valores iniciais dados a priori para o início da iteração são mostrados na tabela 6 abaixo. É importante frisar que os valores dados a priori foram definidos a partir dos valores médios das séries utilizadas, assim como suas variâncias. Os valores em parênteses indicam o desvio-padrão e postamos, também, os intervalos de confiança entre 0.05% e 99.5%, tendo sido adotado um critério de convergência baixo. Os resultados abaixo demonstram que as médias da curva de reação apresentam clara distinção entre os regimes, sendo que o regime de crise demonstra médias mais altas das variáveis, visto que a necessidade do Banco Central reagir a choques de curto prazo tende a ser muito mais enérgica.

TABELA 6: Valores obtidos com a simulação de um modelo MSMA(2)-VAR(1)

	Priors	**Média**	**0.05%**	**50%**	**99.5%**
		Regime 1			
Selic	0.013	0.016821	0.0152	0.0168	0.0181
	(0.001)	(0.00048)			
Selic(-1)	0.05	0.016960	0.0155	0.0169	0.0183
	(0.001)	(0.00051)			
IPCA	0.015	0.006126	0.0047	0.00608	0.0073
	(0.001)	(0.00049)			
IPCA(-1)	0.0	0.006181	0.0045	0.00619	0.0073
	(0.009)	(0.00047)			
Gap	0.0045	0.00072	-0.0094	0.0007	0.0061
	(0.001)	(0.00276)			
Gap(-1)	0.0	-0.00009	-0.010	-0.0009	0.0063
	(0.001)	(0.00272)			
		Regime 2			
Selic	0.013	0.02051	0.01611	0.02046	0.0244
	(0.001)	(0.00153)			
Selic(-1)	0.05	0.02089	0.01620	0.02084	0.0250
	(0.001)	(0.00151)			
IPCA	0.1	0.01273	0.00725	0.01269	0.0173
	(0.001)	(0.00170)			
IPCA(-1)	0.0	0.01284	0.00741	0.01276	0.0177
	(0.009)	(0.00169)			
Gap	0.005	0.00371	-0.0293	0.0030	0.0329
	(0.001)	(0.01078)			
Gap(-1)	0.002	0.00974	-0.0246	0.0094	0.0391
	(0.001)	(0.0115)			

* Desvio-padrão em parênteses.

O comportamento de cada média no processo de iteração é mostrado nos gráficos 5 e 6. O primeiro demonstra o caminho da amostra para o regime de calmaria econômica,

resultando nos valores apresentados na tabela anterior. O segundo possui a mesma estrutura de construção, apresentando os valores para o regime de crise econômica.

Gráfico 5: Comportamento da amostragem após várias iterações do modelo para o regime de calmaria econômica

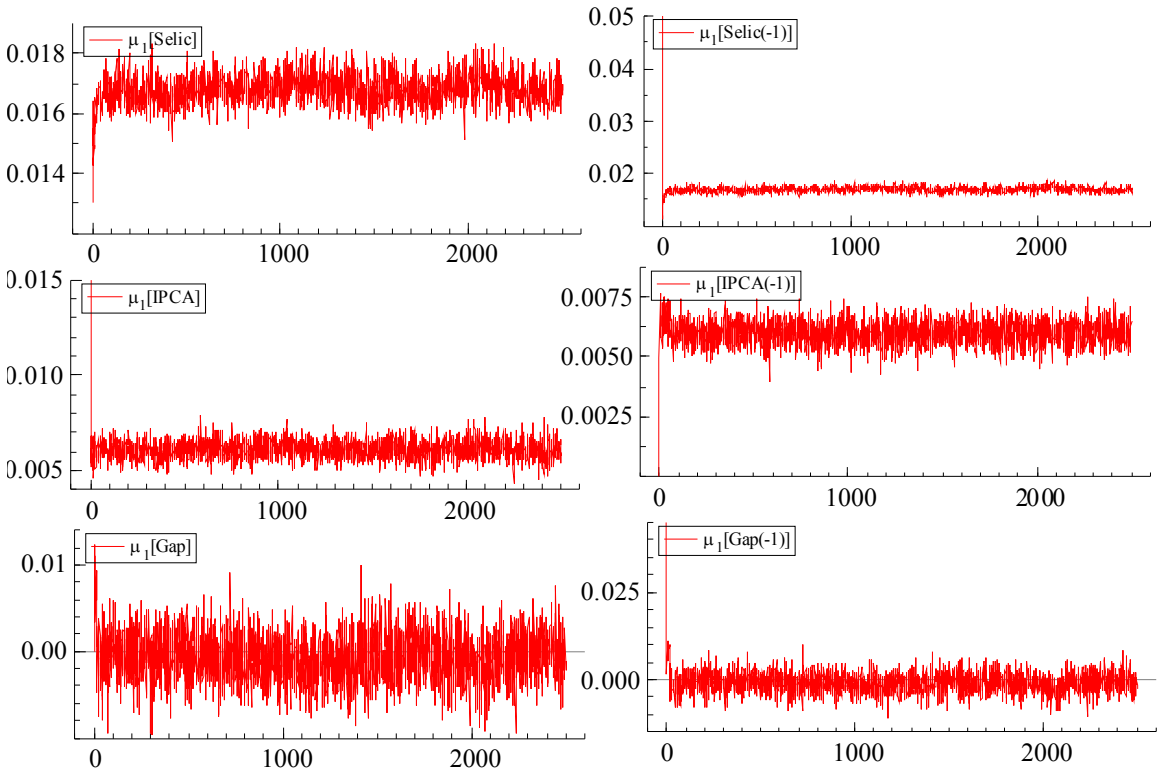
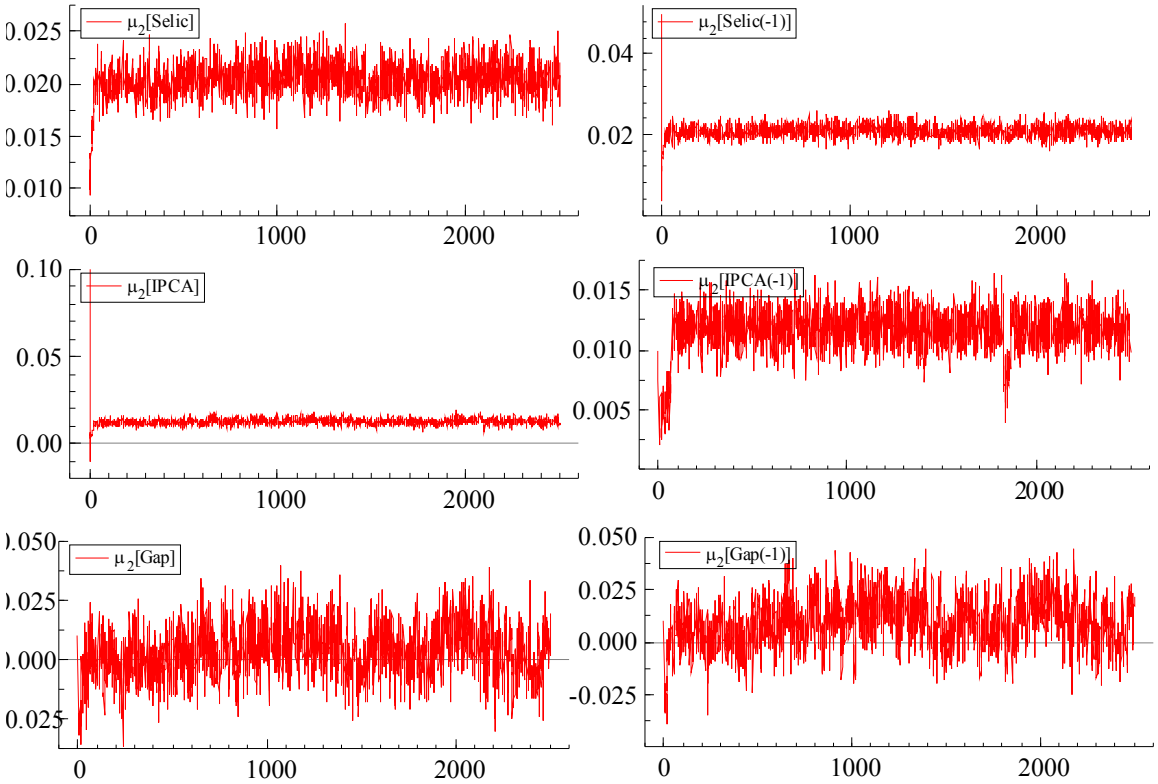


Gráfico 6: Comportamento da amostragem após várias iterações do modelo para o regime de crise econômica



Seguindo o mesmo raciocínio acima, é mister apresentarmos também o "caminho" da média na medida em que ocorrem as iterações de Gibbs. O intuito é avaliar a possibilidade de convergência dos valores através de um método visual, bem como analisar discrepâncias no processo de estabilização dos valores obtidos. Os gráficos 7 e 8 abaixo mostram tais caminhos, sendo o primeiro para o regime 1 e o segundo para o regime 2 (de crises).

Gráfico 7: Caminho da amostragem após várias iterações: Caso de calmaria econômica

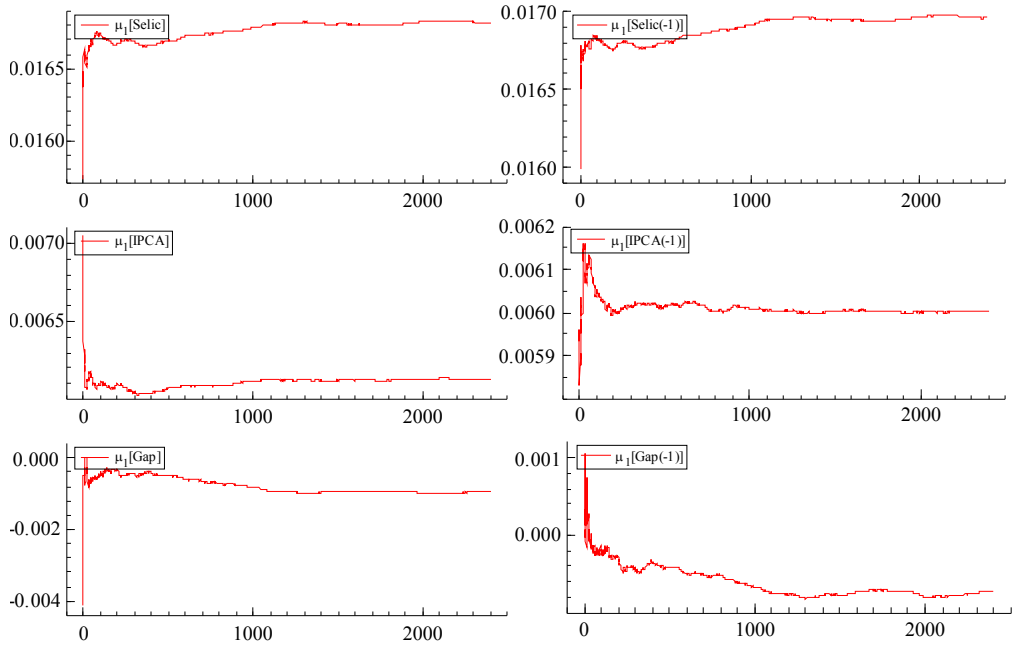
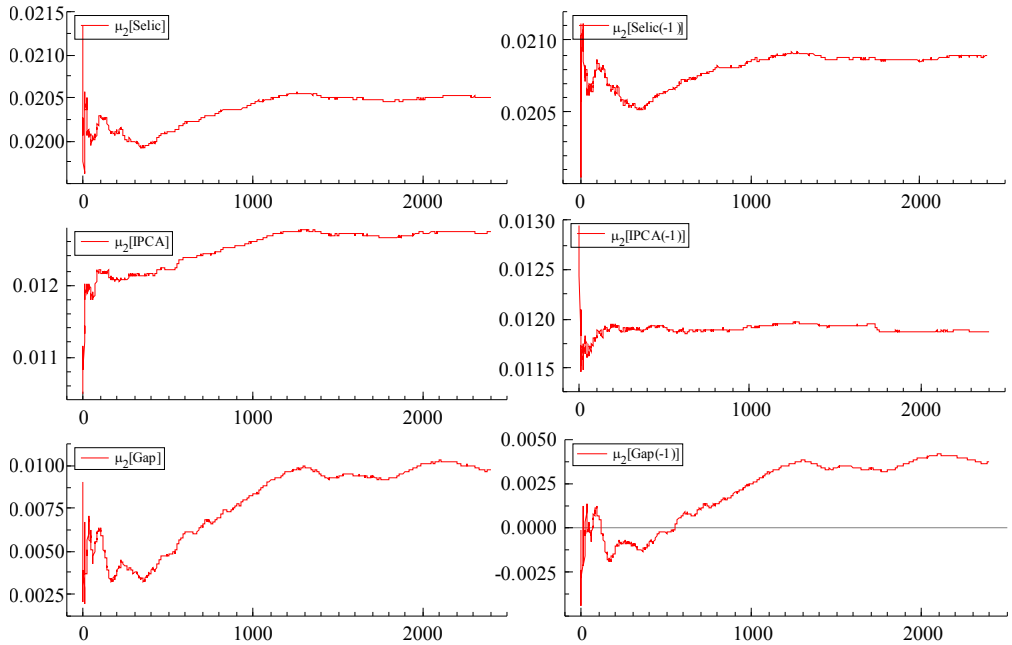


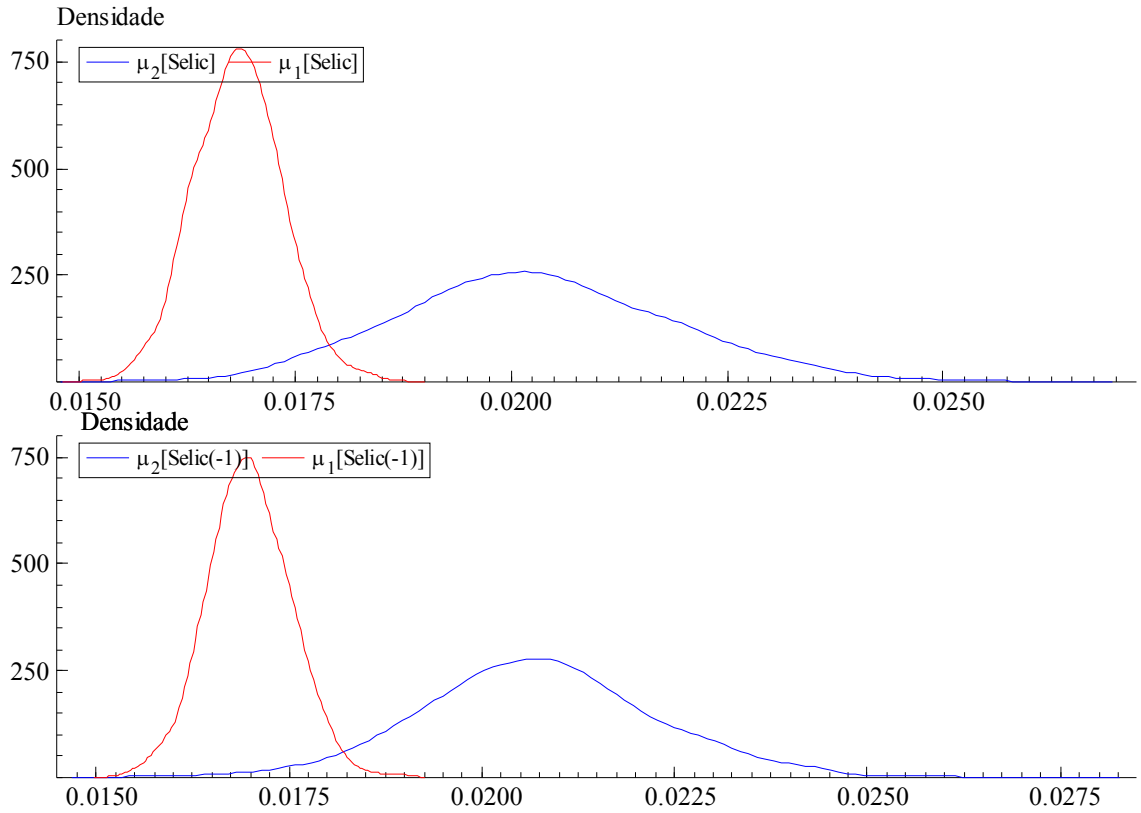
Gráfico 8: Caminho da amostragem após várias iterações: Caso de crises econômicas



Apenas a título ilustrativo, apresentamos as densidades da média para a variável relativa à taxa de juros (Selic) corrente e defasada, visto que as demais séries apresentaram comportamento semelhante em suas densidades. O gráfico 9 apresenta a densidade da média

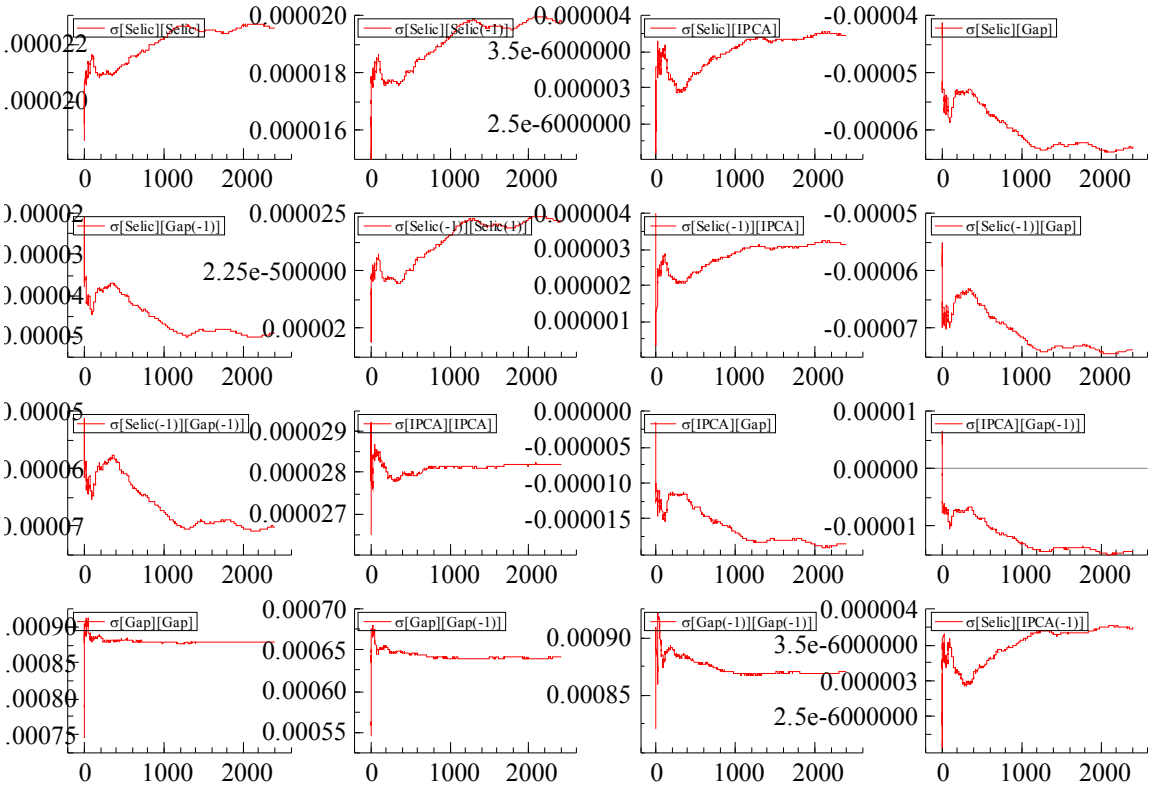
para a Selic corrente e defasada um período para os regimes 1 e 2. Perceba que há uma diferença entre as caudas da densidade para o regime 1 em relação ao 2, demonstrando a necessidade de uma modelagem que reflita as diferenças existentes entre os regimes. O uso de técnicas estatísticas de *Markov Switching* combinadas com o uso de misturas de distribuições ou distribuições multivariadas do tipo Wishart podem corrigir tais problemas de estimação, levando em consideração os problemas econométricos arrolados ao longo deste trabalho.

Gráfico 9: Densidades para a média das variáveis Selic e Selic(-1)



Por fim, apresentamos também o comportamento da amostragem dos coeficientes de correlação contemporânea dos erros (Ω) ao longo das n iterações propostas (médias). A correlação contemporânea dos erros apresentou muitos parâmetros não significativos, mas nenhuma correlação cruzada pareceu ser importante na dinâmica das mudanças de regimes. O gráfico 10 mostra os resultados obtidos.

Gráfico 10: Comportamento da amostragem para correlações contemporâneas

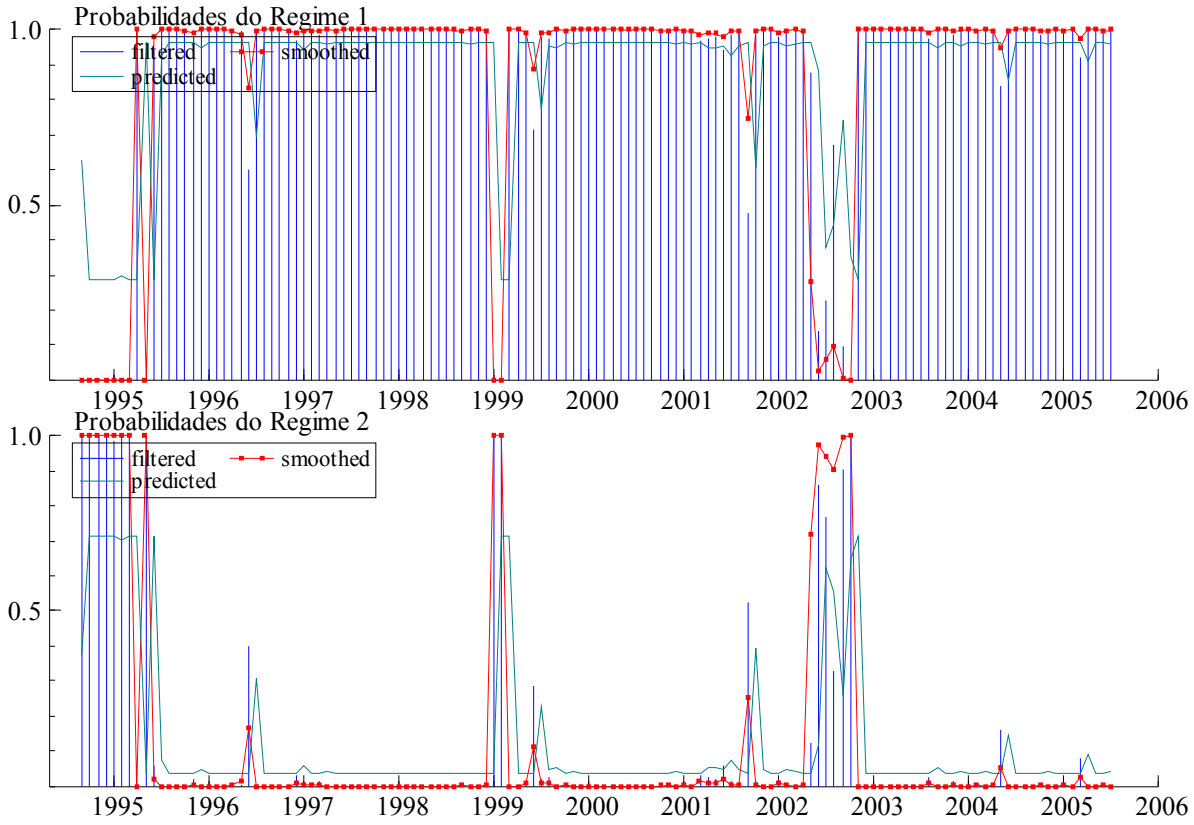


5.2. Estimação para Economia Aberta

O modelo estimado foi selecionado da mesma forma que o modelo para economia fechada o foi. Os critérios de seleção indicaram também um modelo MSIAH (2) – VAR (1) para a regra de política monetária para a economia aberta.

Vejamos as probabilidades suavizadas, filtradas e previstas para os regimes em estudo, nos moldes da primeira parte do trabalho. O gráfico 11 apresenta as probabilidades para cada um regimes ao longo do período amostral. Os resultados obtidos se assemelham bastante aos da seção que versa sobre a economia fechada. É possível avaliar os mesmos períodos de crise para o modelo com economia aberta; as crises do México, Ásia, Rússia, a desvalorização cambial brasileira, a crise Argentina e a crise das eleições de 2002.

Gráfico 11: Probabilidades para os regimes do modelo MSIAH (2) – VAR (1) adotado à regra de política monetária no período do Plano Real



Quanto às probabilidades de transição, a matriz de transição (linha estocástica) com distribuição estacionária obtida no modelo de dois regimes é a seguinte:

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} 0.9266 & 0.0734 \\ 0.4022 & 0.5978 \end{bmatrix}$$

A matriz nos indica que a probabilidade de permanência em períodos de estabilidade econômica (p_{11} = 0.9266) é maior do que a de crises agudas (p_{22} = 0.5978). A permanência em períodos de estabilidade econômica é bem maior do que em crises agudas, mesmo sabendo que a economia brasileira enfrentou neste período várias crises de curta duração. Em outras palavras, a probabilidade de manutenção das taxas de juros em patamares médios dado que a economia esteja em períodos de ausência ou presença de crises é muito significativo, pois a probabilidade de ocorrência de crises não é desprezível. Porém, as probabilidades de transição entre regimes indicam valores maiores do que os encontrados para a economia fechada. Em termos percentuais, ocorreu um aumento de 21%, em relação ao modelo para economia fechada, na probabilidade da economia sair de uma crise e passar para um período de maior estabilidade (0.4022 contra 0.33), mas a probabilidade da transição de um regime de

estabilidade econômica para um regime de crise também aumentou, o que reflete as possíveis alterações que a inclusão da taxa de câmbio produziu para o modelo como um todo (p_{12} = 0.0.0734).

A título ilustrativo, optamos por apresentar os coeficientes gerados com o MSIAH(2)- VAR(1) para a regra de política monetária estudada. Importante salientar que os coeficientes de uma estimação VAR(p) são sujeitos à crítica da literatura econômica, por isso, a sua importância se restringe à avaliação das relações entre as variáveis de um dado sistema. É possível identificarmos várias relações econômicas estabelecidas na literatura, tais como o aumento da taxa de juros referencial (Selic) no período *t - 1* implicando em redução do hiato do produto no momento *t*, ou o aumento da taxa de inflação (IPCA) em *t - 1* resultando no aumento do hiato do produto em *t*. É premente afirmarmos acerca da maior magnitude dos valores dos parâmetros para o regime de crises econômicas (regime 2), visto que já seria esperado relações de maior impacto entre as variáveis em períodos de choques e instabilidades macroeconômicas. Os resultados são bastante próximos aos obtidos por Bueno (2005) e Moreira (2004), apesar destes autores estimarem uma regra de política monetária através de um *Markov* Switching univariado. Os valores dos parâmetros são reportados na tabela 7 abaixo e o comportamento dos resíduos é visualizado no gráfico 12.

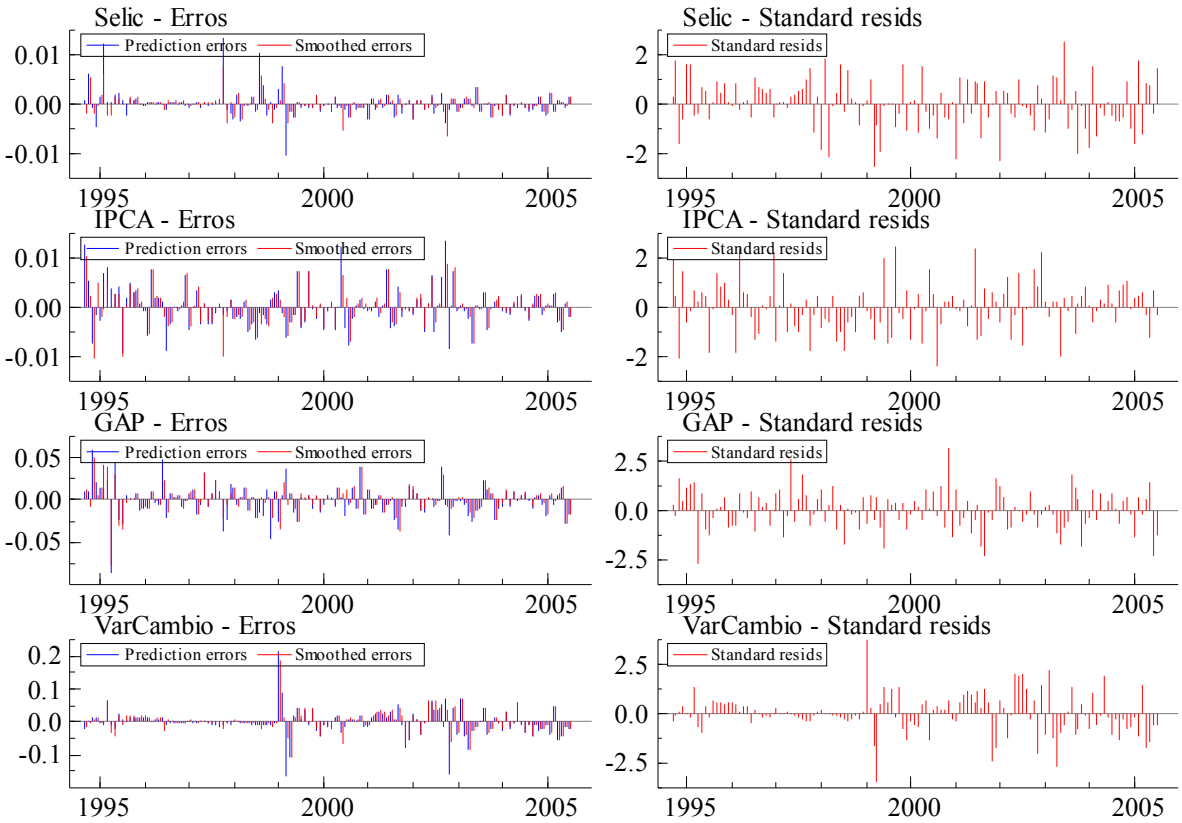
TABELA 7: Parâmetros estimados a partir do modelo MSIAH (2) – VAR (1) para economia aberta

		Regime 1		
	Selic	IPCA	Gap	Câmbio
Constante	0.001282	0.000677	0.029290	-0.006157
Selic(-1)	0.906942	0.075537	-1.477090	0.690355
IPCA(-1)	0.019416	0.620921	-0.785680	-0.919126
Gap(-1)	0.025816	-0.002402	0.533569	0.174903
Câmbio(-1)	-0.008640	0.006901	0.072082	0.124505
Desvio-Padrão	0.001356	0.003200	0.012468	0.030953
		Regime 2		
Constante	-0.000817	0.006774	0.003954	0.074444
Selic(-1)	1.158184	0.095410	-0.899920	0.428098
IPCA(-1)	-0.242340	0.480874	2.429105	-3.698131

Gap(-1)	-0.008170	0.022247	0.688238	-0.259779
Câmbio(-1)	0.027097	-0.006384	-0.083466	0.739227
Desvio-Padrão	0.003925	0.005468	0.030849	0.049340

Em relação aos resíduos do modelo, o gráfico 12 apresenta o comportamento do resíduo de cada série, tanto o predito quanto o suavizado, demonstrando a volatilidade para cada variável.

Gráfico 12: Comportamento dos resíduos do modelo MSIAH (2) – VAR (1) estimado para a economia aberta



Por fim, a tabela 8 condensa as informações sobre a duração e a probabilidade de ocorrência dos regimes. Vemos que a duração do regime de estabilidade econômica é de 13.63 meses. A presença de fortes crises tem menor probabilidade de ocorrência (15,43%), com duração, em média, de 2.49 meses.

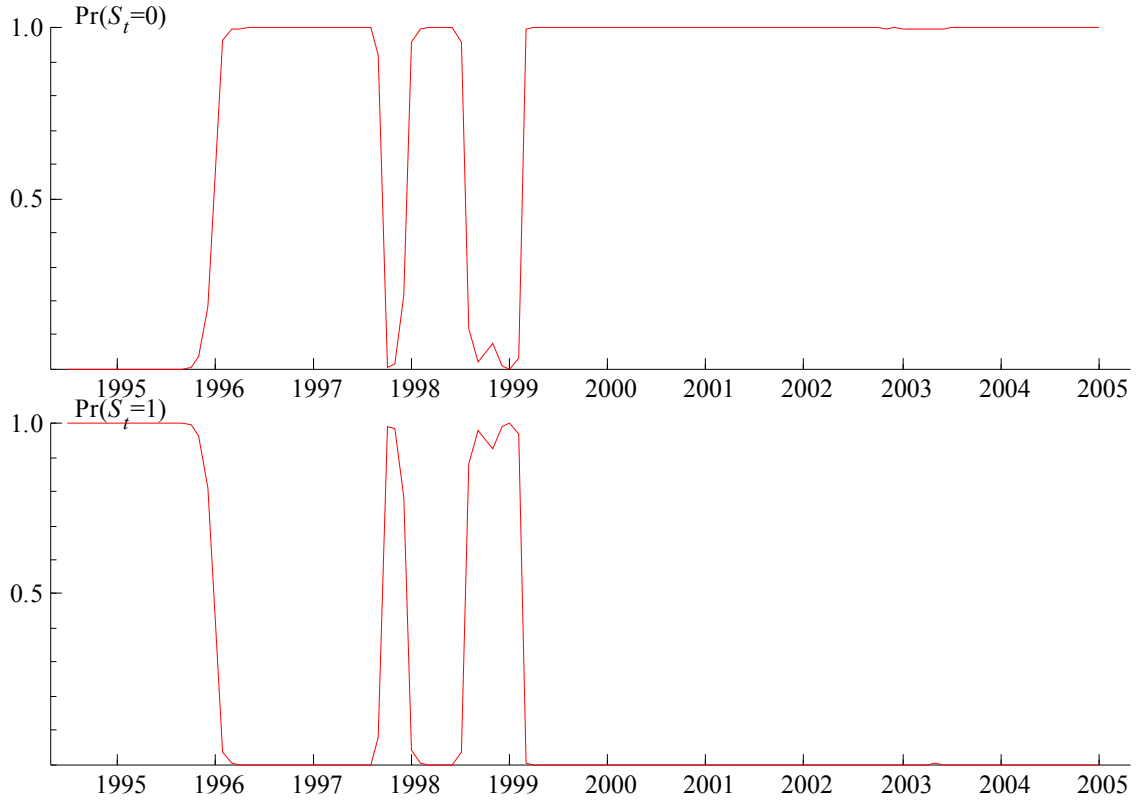
TABELA 8: Probabilidades estacionárias e duração dos regimes do modelo MSIAH (2) – VAR (1)

	Nº de observações	Probabilidade	Duração
Regime 1: Estabilidade econômica	109.4	0.8457	13.63
Regime 2: Presença de crises	21.6	0.1543	2.49

Seguindo a mesma abordagem da seção anterior, estimaremos um modelo para a regra de política monetária dada para uma economia aberta usando a metodologia de MCMC. Foram realizadas 1500 iterações, sendo descartadas as 100 primeiras. Os dados considerados nesta seção são os mesmos das anteriores, sendo 1994/8-2005/8 o período coberto pela amostra. O algoritmo convergiu com 1 hora e 15 minutos de iterações e novamente várias especificações não puderam ser geradas devido a problemas de inversão de matrizes e instabilidade nos autovalores no processo de decomposição de Choleski. Assim, adotamos o mesmo modelo da seção anterior, um MSMA(2) - VAR(1), sendo que os valores para as médias convergiram rapidamente, independentemente dos valores dados inicialmente. Os resultados são reportados a seguir.

O gráfico 13 nos mostra as probabilidades para os regimes do modelo MSMA (2) – VAR (1) estimado pelo método MCMC proposto para uma regra de política monetária em uma pequena economia aberta.

Gráfico 13: Probabilidades para os regimes do modelo MSMA (2) – VAR (1) adotado à regra de política monetária em uma economia aberta



Os resultados obtidos novamente são coincidentes com as demais partes deste trabalho. A diferença é que a inclusão da variável câmbio aumentou a probabilidade de ocorrência de crises econômicas, visto que parcela importante das flutuações das variáveis macroeconômicas ao longo do plano real está intrinsecamente relacionada aos choques externos ocorridos ao longo deste período. Ora, como o câmbio nominal é uma variável que capta estes choques externos, a probabilidade de ocorrência de crises aumentou, mas os períodos de crise permaneceram os mesmo (Crise do México – 1994, Crise Asiática – 1997, Crise Russa – 1998, Desvalorização cambial – 1999 e Crise das eleições – 2002).

A tabela 9 apresenta os valores obtidos para a média, bem como o intervalo de confiança para os valores médios obtidos pelas iterações do algoritmo de Gibbs. Novamente as médias das variáveis para o regime de crise econômica apresentam valores maiores em comparação ao regime com ausência de crises, o que demonstra a maior efetividade da política monetária em uma economia aberta com câmbio flexível. Interessante apontar a grande variabilidade do câmbio, principalmente no regime de crise. No período de crise a média para o câmbio apresenta valores altos (mesmo as variáveis sendo log-linearizadas) e negativos, já que as fortes desvalorizações cambiais dos períodos de crise afetam diretamente esta variável. Outro problema foi o valor negativo para a média do hiato do produto para o regime de crise econômica. Seria esperado um valor positivo, já que aumentos da taxa de

juros conduzem à desaceleração da atividade econômica. Porém, a média para esta variável é fracamente significativa, o que pode sugerir um resultado espúrio para ela.

TABELA 9: Valores obtidos com a simulação de um modelo MSMA(2)-VAR(1) para a economia aberta

	Priors	**Média**	**0.05%**	**50%**	**99.5%**
		Regime 1			
Selic	0.013 (0.001)	0.01553 (0.00045)	0.0140	0.0155	0.0167
Selic(-1)	0.05 (0.001)	0.01531 (0.00045)	0.0141	0.0155	0.0168
IPCA	0.015 (0.001)	0.00604 (0.00061)	0.0041	0.00604	0.00761
IPCA(-1)	0.0 (0.009)	0.006181 (0.00047)	0.0045	0.00619	0.0073
Gap	0.0045 (0.001)	0.00398 (0.00299)	-0.0045	0.00387	0.0112
Gap(-1)	0.0 (0.001)	-0.002 (0.00254)	-0.0032	-0.0018	0.009
Câmbio	0.1 (0.01)	-0.0602 (0.4739)	-1.673	-0.051	1.217
Câmbio(-1)	-0.01 (0.01)	0.0321 (0.443)	-1.017	0.029	1.055
		Regime 2			
Selic	0.013 (0.001)	0.0159 (0.00102)	0.01285	0.01594	0.01897
Selic(-1)	0.05 (0.001)	0.0168 (0.00091)	0.01361	0.01688	0.01946
IPCA	0.1 (0.001)	0.00695 (0.00170)	0.00259	0.00694	0.0109
IPCA(-1)	0.0 (0.009)	0.01284 (0.00169)	0.00741	0.01276	0.0177
Gap	0.005 (0.001)	-0.0206 (0.00649)	-0.0338	-0.0207	-0.0027
Gap(-1)	0.002 (0.001)	0.00974 (0.0115)	-0.0246	0.0094	0.0391
Câmbio	0.1 (0.01)	-0.542 (1.699)	-3.316	-0.532	1.659
Câmbio(-1)	-0.01 (0.01)	-0.521 (1.737)	-3.738	-0.503	1.504

Aproveitando os valores mostrados acima, os gráficos 14 e 15 apresentam o caminho das médias das variáveis após as 1500 iterações. É possível visualizar a convergência de todas as médias para um determinado valor após a iteração de número 1000. A estabilidade obtida foi testada para iterações maiores do que 1500, tal como 3000, e a

trajetória das médias apresentou a mesma aderência, sendo os valores divergentes apenas na 4º casa decimal.

Gráfico 14: Caminho da amostragem após várias iterações: Caso de calmaria econômica

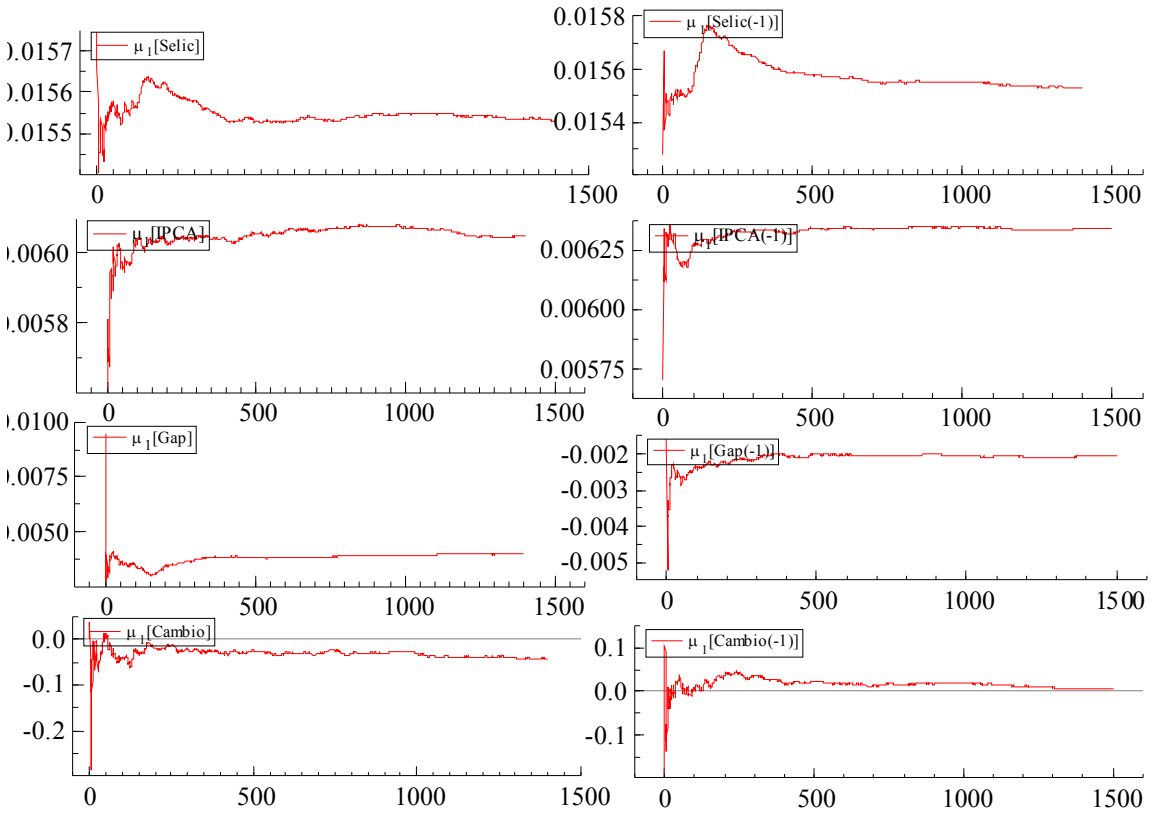
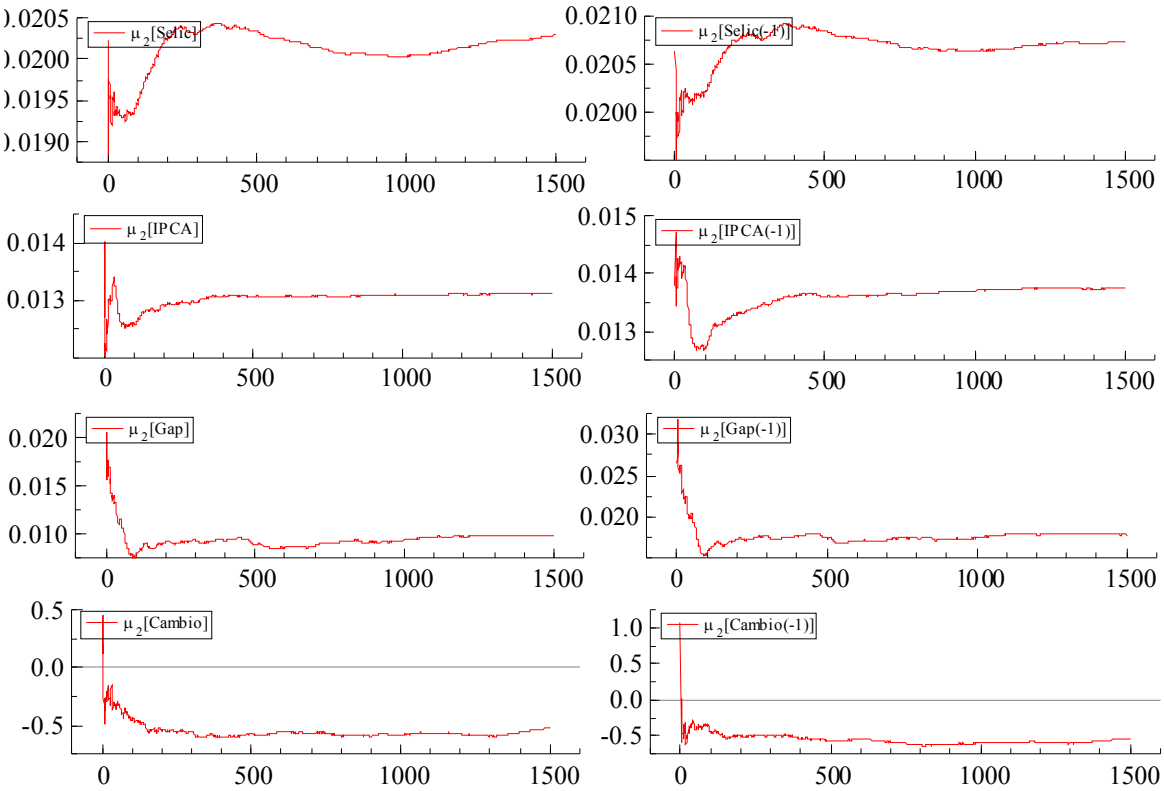
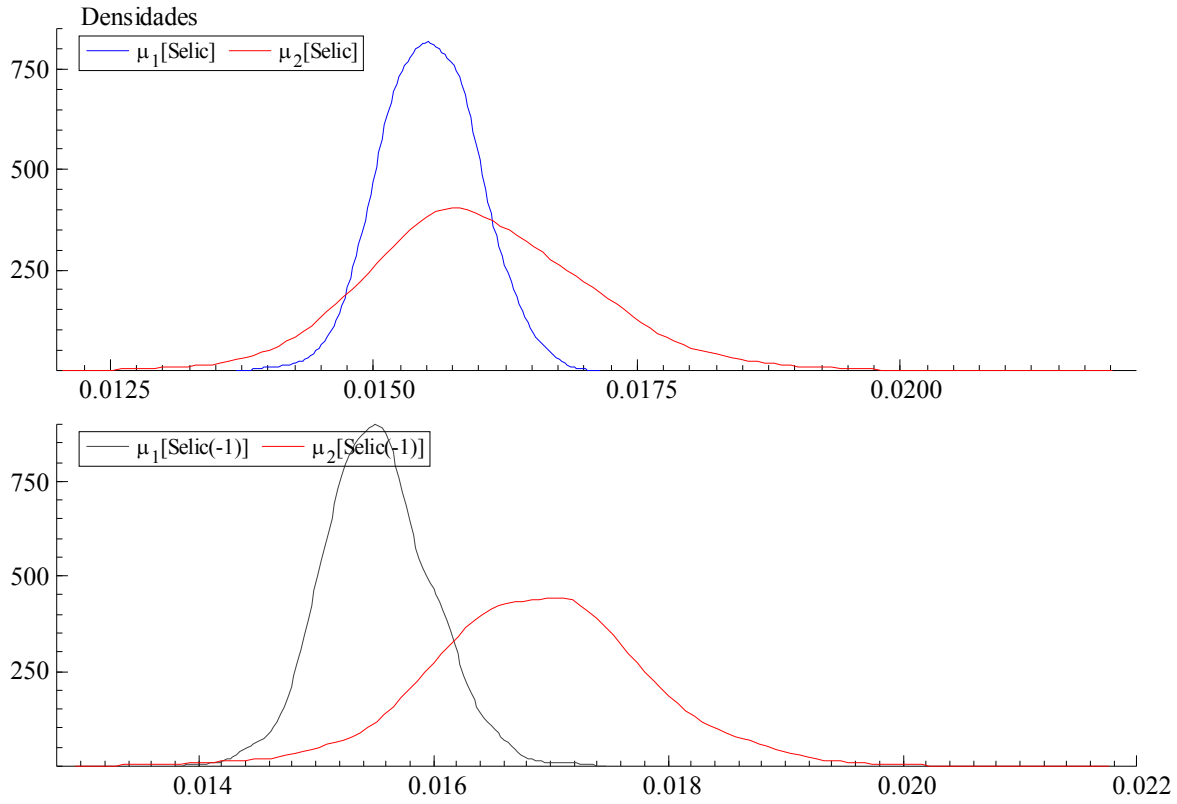


Gráfico 15: Caminho da amostragem após várias iterações: Caso de crises econômicas



Por fim, cabe mostrar as densidades da variável Selic e Selic(-1) para os dois regimes estudados. A escolha da variável se justifica para mostrar as diferentes densidades geradas para cada regime, o que sugere a utilização de métodos estatísticos com mudança de regime em modelos que utilizem tais variáveis. Como a mudança de regime através da metodologia de *Markov Switching* é fruto da mistura de distribuições de probabilidade das variáveis em questão, é mister apresentarmos algumas densidades para ilustrarmos o problema de densidades dependentes de regimes a que nos referimos no início deste trabalho. O gráfico 16 apresenta as diferentes densidades dependentes de cada regime estudado.

Gráfico 16: Densidades para a média das variáveis Selic e Selic(-1)



Conforme se pode ver, existem distribuições leptocúrticas e platicúrticas, dependendo do regime. No caso da variável Selic, o regime de calmaria econômica apresenta menor variabilidade, sendo os valores mais próximos da média da distribuição. Por outro lado, para o regime de crises econômicas é nítida a diferença não só no valor da média que se desloca um pouco, mas na variância que tende a ser maior. Isto é um indicativo da necessidade de maiores mudanças na variável-instrumento da política monetária, o que gera o fenômeno de distribuições platicúrticas após as *n* iterações geradas.

Resultado semelhante pode ser visto para a variável Selic(-1).

5.3. Padrões de Resposta a Choques

Conforme observamos na seção anterior, o comportamento da economia tem mudado de forma substancial nas últimas décadas. Assim tem crescido o consentimento de que os efeitos da política monetária também têm mudado. Autores como Boivin e Giannoni (2003) têm destacado as possíveis mudanças destes efeitos, vaticinando que uma possível

interpretação seria a perda da influência da política monetária sobre a economia, quiçá, induzida pelo progresso tecnológico e inovações financeiras, permitindo aos consumidores absorverem melhor os impactos das flutuações da taxa de juros.

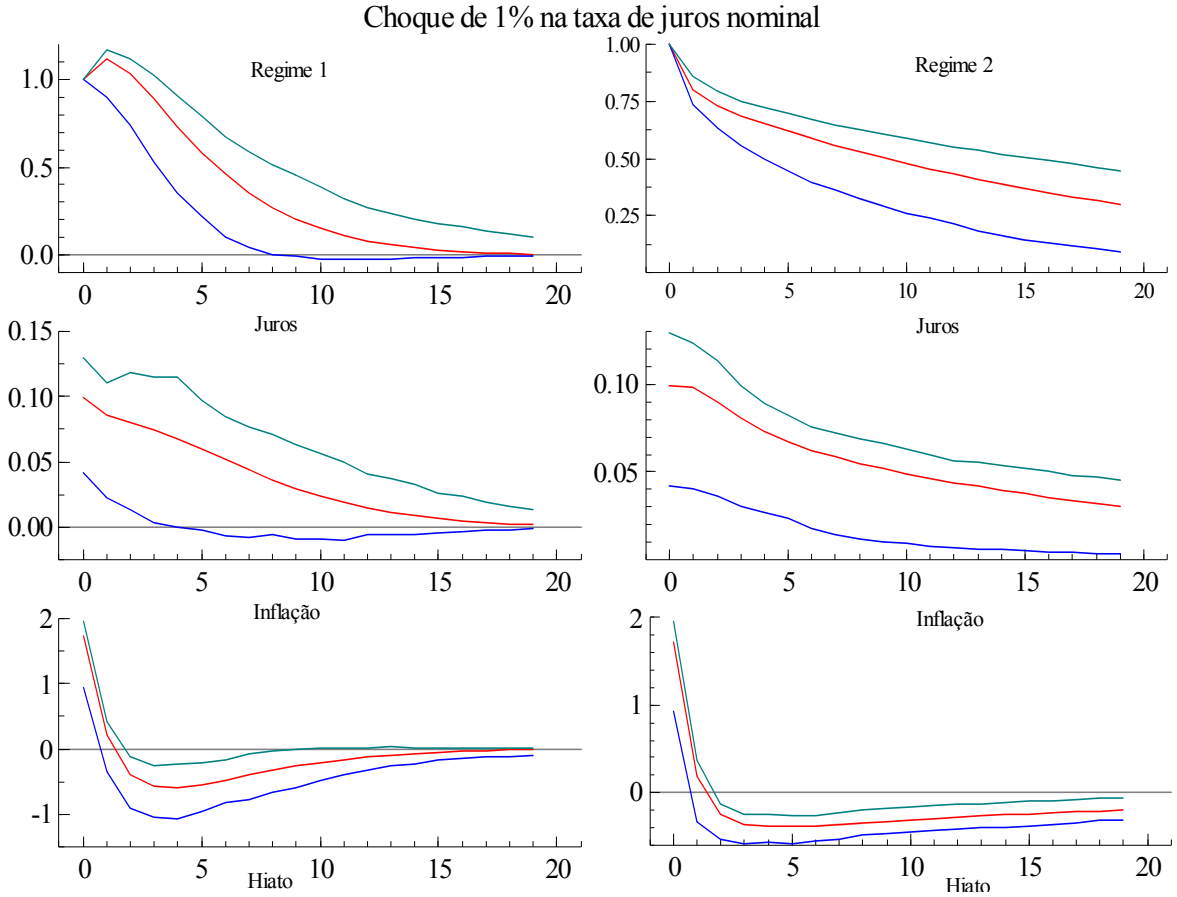
Motivados por argumentos de que a política monetária tem mudado, para o caso brasileiro, vimos que há existência de dois regimes específicos de política monetária adotados ao longo do plano real. Assim, a utilização de funções resposta a impulso generalizadas dependentes de cada regime obtido na seção anterior atendem perfeitamente o objetivo de avaliarmos as mudanças nos mecanismos de transmissão da política monetária e as conseqüentes implicações para a efetividade desta política.

No caso do presente trabalho, onde o interesse primário recai sobre a resposta dos níveis do produto e de preços ao impulso monetário, é razoável supor que um choque na taxa de juros não tenha efeitos imediatos sobre as variáveis-alvo da política monetária, devido à rigidez dos preços nominais no horizonte de curto prazo. Esta alternativa pode ser adequadamente representada pela identificação, via decomposição de Cholesky, ordenando as equações de forma recursiva, a partir da variável "mais exógena" (afetada contemporaneamente apenas pelo seu próprio choque estrutural) até a variável "mais endógena" (afetada contemporaneamente por todos os choques). As outras técnicas foram discutidas na seção 2.3.

No que tange as funções resposta a impulso analisamos nesta seção os padrões de resposta da política monetária a diferentes choques, tanto para a economia fechada quanto para a aberta. O objetivo desta seção é ilustrar, em um contexto dinâmico, como a regra de política monetária, tanto em um modelo de economia fechada quanto aberta, suaviza os caminhos das variáveis-objetivo e como se dão as relações entre estas variáveis da regra de política monetária. Os gráficos a seguir mostram as funções de resposta a impulso das variáveis-objetivo para a regra de política monetária brasileira, e no caso da economia aberta, sob regimes de câmbio diferentes. No último caso, a idéia é avaliar como se dariam as relações entre a taxa de juros nominal e as demais variáveis do sistema considerando flutuações do câmbio nominal, principalmente porque sabemos a priori que a economia brasileira ao longo do Plano Real apresentou dois regimes cambiais distintos; câmbio fixo até dez/1998 e flutuante após 1999.

O gráfico 17 mostra as funções de resposta do hiato do produto, da inflação e da taxa de juros nominal a choques positivos nesta variável.

Gráfico 17: Funções de resposta a impulso, choque de 1 p.p. na taxa de juros nominal



Vamos considerar inicialmente o choque não persistente de 1 ponto percentual (p.p.) na taxa de juros nominal. Na data 0, há um choque positivo de 1 p.p. na taxa de juros nominal, demonstrado pelos primeiros gráficos superiores, que representam o choque no regime 1 e no regime 2. Nas datas posteriores, no caso da taxa de juros nominal, vemos que nos períodos de estabilidade econômica há uma redução gradual da variação causada pelo choque, ocorrendo a zeragem dos efeitos, *ceteris paribus*, em torno da data 15. Por outro lado, nos regimes de crise, não ocorre redução gradativa nem zeragem dos efeitos do choque positivo sobre a taxa de juros. Como é mister o controle das variáveis-meta pelo Banco Central, a permanência de uma taxa de juros mais alta busca estabilizar as demais variáveis em torno de suas metas de equilíbrio.

Em relação à inflação e ao hiato do produto, o choque positivo na taxa de juros de curto prazo demonstra aumenta dos valores destas variáveis nos períodos estudados. Aqui cabe uma explicação complementar. Se o setor privado acredita no compromisso do Banco Central com regras super-inerciais, ele tem motivos para esperar taxas de juros positivas no

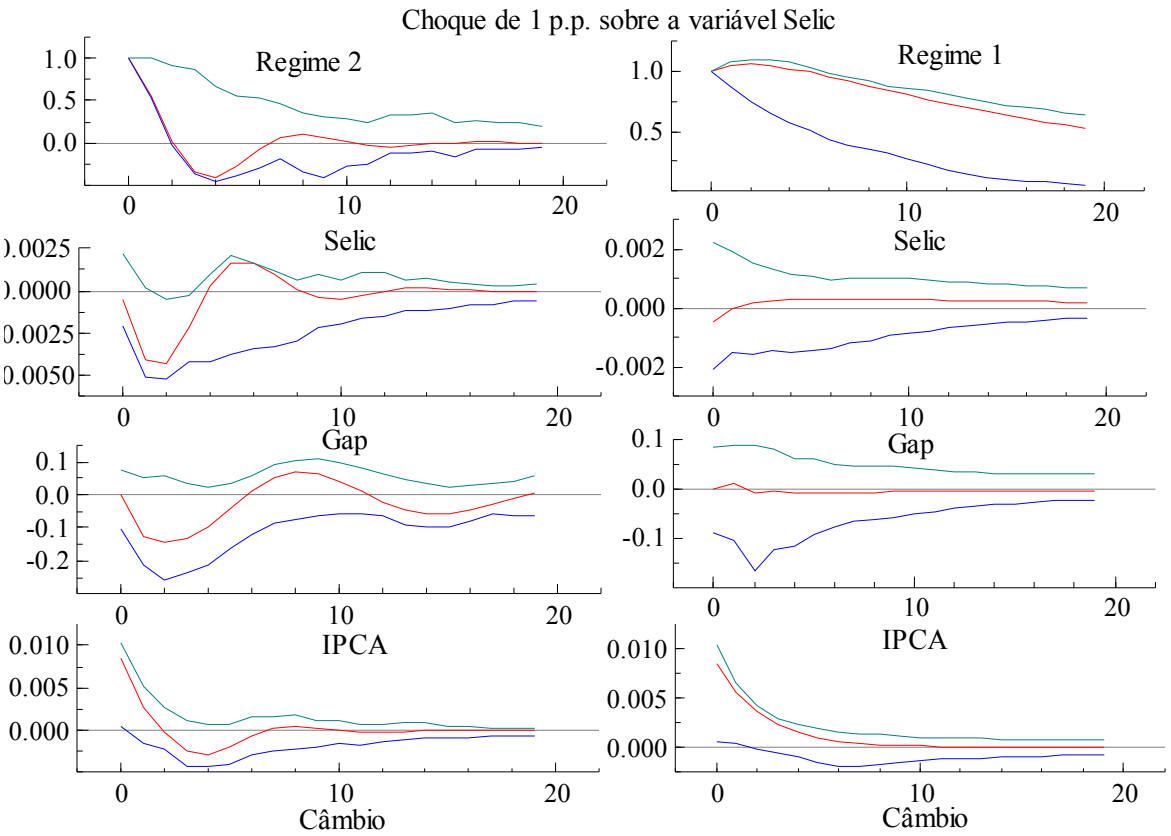
futuro, mesmo após o desaparecimento dos efeitos diretos do choque nas variáveis endógenas. Mais ainda: como as regras têm elementos inerciais, o setor privado tem motivos para esperar trajetórias explosivas das taxa de juros nominal no futuro, a não ser que o *overshooting* inicial da inflação sobre o seu valor de *steady-state* seja seguido por um *undershooting* nos próximos períodos. Para evitar o cenário de trajetórias explosivas da taxa de juros nominal, o setor privado coordena suas expectativas em torno dos equilíbrios com *undershooting* da inflação, *undershooting* do hiato do produto e taxas de juros nominais positivas, porém decrescentes. A inércia da taxa de juros nominal e o *undershooting* da inflação esperada levam a taxa de juros reais esperadas positivas e a hiatos do produto esperados negativos a partir da data 2. Expectativas de hiatos do produto negativos no futuro mais do que compensam o impacto positivo do crescimento do hiato do produto, mas o resultado alcançado pelo trabalho implica em um aumento da inflação causado por um choque da taxa de juros nominal, resultado espúrio do ponto de vista teórico. Aqui já vemos um indício de que funções resposta a impulso dependentes de cada regime e da história das variáveis podem não ser aderentes para um modelo de economia fechada, talvez pela identificação usando decomposição de Cholesky não ser aderente ao modelo brasileiro para economia fechada, conforme ressalta Christiano, Eichenbaum e Evans (1998).

Estes resultados podem ser vistos nos gráficos colocados acima, *mutatis mutandis*, em cada regime adotado. Nos regimes de crise, devido ao comportamento *forward looking* da economia brasileira, é visto que as expectativas do setor privado de uma regra de política monetária crível fazem com que a taxa de juros nominal no final do período de choque fique com zeragem dos efeitos sobre as demais variáveis permanece positiva. Porém, conforme ressaltamos anteriormente, os resultados para a economia fechada não foram satisfatórios, o que indica a necessidade de se avaliar os choques para uma economia aberta.

No gráfico 18 estão as respostas a choques positivos na taxa de juros nominal (selic), tendo a magnitude de 1 p.p., aplicados para o modelo de economia aberta. A primeira coluna mostra o caminho das variáveis para o regime de crise econômica (regime 2), enquanto a segunda coluna apresenta a trajetória das variáveis para o regime de estabilidade econômica. O choque positivo na taxa de juros apresentou efeitos diferentes para os regimes considerados. No caso do regime de estabilidade econômica (regime 1), a redução dos efeitos da variável referente à taxa de juros nominal vão gradualmente diminuindo, o hiato do produto cai discretamente no início, se estabilizando próximo de zero ao longo do processo de zeragem dos efeitos gerados pela taxa de juros nominal. A taxa de inflação manteve-se estável, mas o câmbio apresentou desvalorização. Resultado diverso do esperado já se

esperava uma valorização cambial, visto que os aumentos da taxa de juros nominais combinados com a manutenção da taxa de inflação poderiam gerar um grande fluxo de capitais externos, valorizando a moeda local.

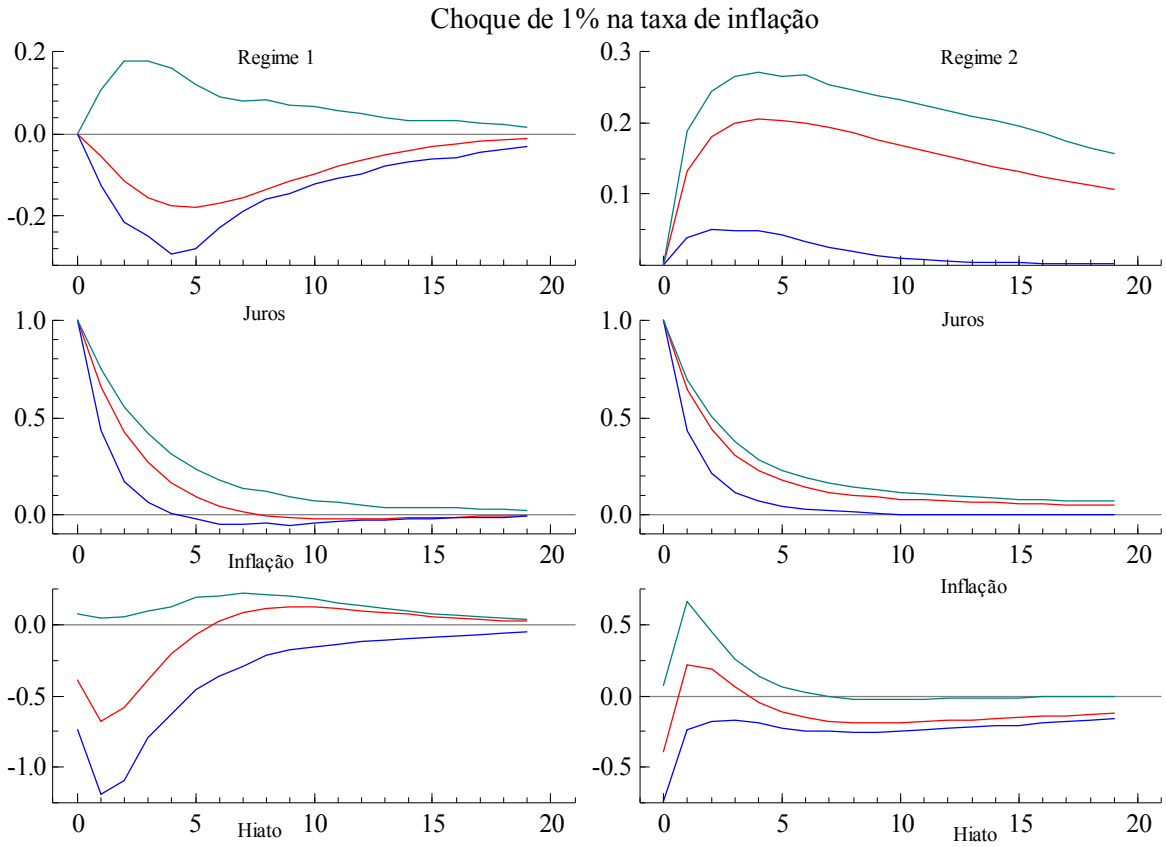
Gráfico 18: Funções de resposta a impulso, choque de 1 p.p. na taxa de juros nominal



Para o regime de crise econômica, o aumento brusco da taxa de juros nominal levou a uma forte queda do hiato do produto, acompanhado por uma redução da taxa de inflação. A redução da inflação e do hiato do produto faz com que o Banco Central reduza as taxas de juros, estimulando a demanda doméstica e gerando um resultado oposto ao mencionado acima. Após a data 6 a autoridade monetária volta a subir as taxas de juros nominais, zerando os efeitos iniciais discutidos. O câmbio se desvaloriza em um primeiro momento, o que não esperávamos acontecer, mas a partir da data 2 ocorre uma valorização cambial, visto que o aumento da taxa de juros nominal juntamente com a redução da inflação afetam positivamente a entrada de divisas.

No gráfico 19 podem ser vistos os efeitos das funções resposta a impulso a um choque positivo de 1 p.p. sobre a taxa de inflação para o modelo de economia fechada. No regime 1, o hiato do produto cresce a partir de um valor inicial negativo, devido ao impacto da inflação corrente sobre o crescimento do hiato. Interessante notar que a taxa de juros nominal apresenta valores negativos nos primeiros períodos, fruto, quiçá, da inércia da taxa de juros na regra de política monetária, resultando em mais um impacto espúrio. Como se está em períodos de estabilidade econômica, o Banco Central suaviza os efeitos do aumento da inflação ao longo do tempo, buscando aumentar o crescimento da economia. Já no regime de crises, um choque de 1 p.p. na inflação leva o Bacen a aumentar a taxa de juros nominal com o fito de controlar os aumentos da inflação e do hiato do produto, mesmo após a zeragem dos efeitos sobre as variáveis-meta, a autoridade monetária mantém taxas de juros nominais positivas. A resposta ao choque via aumento de juros nominais surte efeitos imediatos sobre o hiato produto, que em 4 períodos já apresenta sinais negativos. Este resultado é creditado à redução das expectativas do setor privado quanto ao crescimento da inflação e do produto.

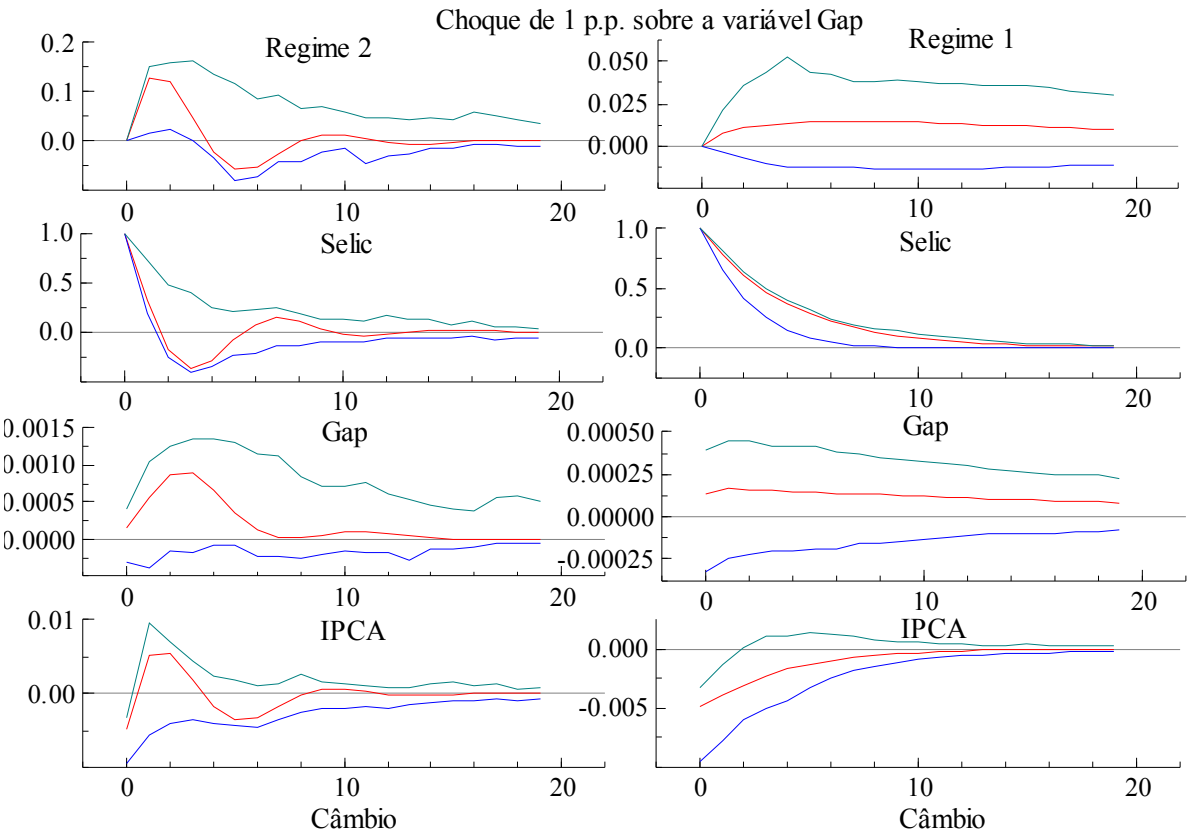
Gráfico 19: Funções de resposta a impulso, choque de 1 p.p. na taxa de inflação



Para o modelo de economia aberta, o gráfico 20 apresenta os resultados de um choque de 1 p.p. sobre o hiato do produto. É mister avaliarmos que para o regime de crises (regime 2) a elevação provocada pelo aumento do hiato do produto (por exemplo, um choque negativo na produtividade doméstica que diminui o produto natural e aumenta o hiato do produto) leva a uma desvalorização cambial (nominal) que desloca inicialmente a demanda na direção do produto doméstico, aumentando a inflação doméstica. No entanto, o Banco Central aumenta a taxa de juros nominal até o período 3, o que faz com que os fluxos de capitais internacionais retornem ao país, reduzindo os efeitos da desvalorização cambial. No âmbito interno, o aumento da taxa de juros reduz o nível de atividade interna, reduzindo a inflação e o hiato do produto. Este processo em regimes de crises faz com que as variações da taxa de juros nominal sejam reduzidas pelo Banco Central na medida em que o hiato do produto e a inflação se reduzem, mas por volta do período 5, com a inflação em patamares baixos e o hiato do produto apresentando resultados negativos, o Banco Central reduz mais a taxa de juros nominal, rompendo o valor aplicado antes do choque. O mesmo raciocínio anterior é empregado, fazendo com que os efeitos sobre as demais variáveis sejam zerados após o 11º período.

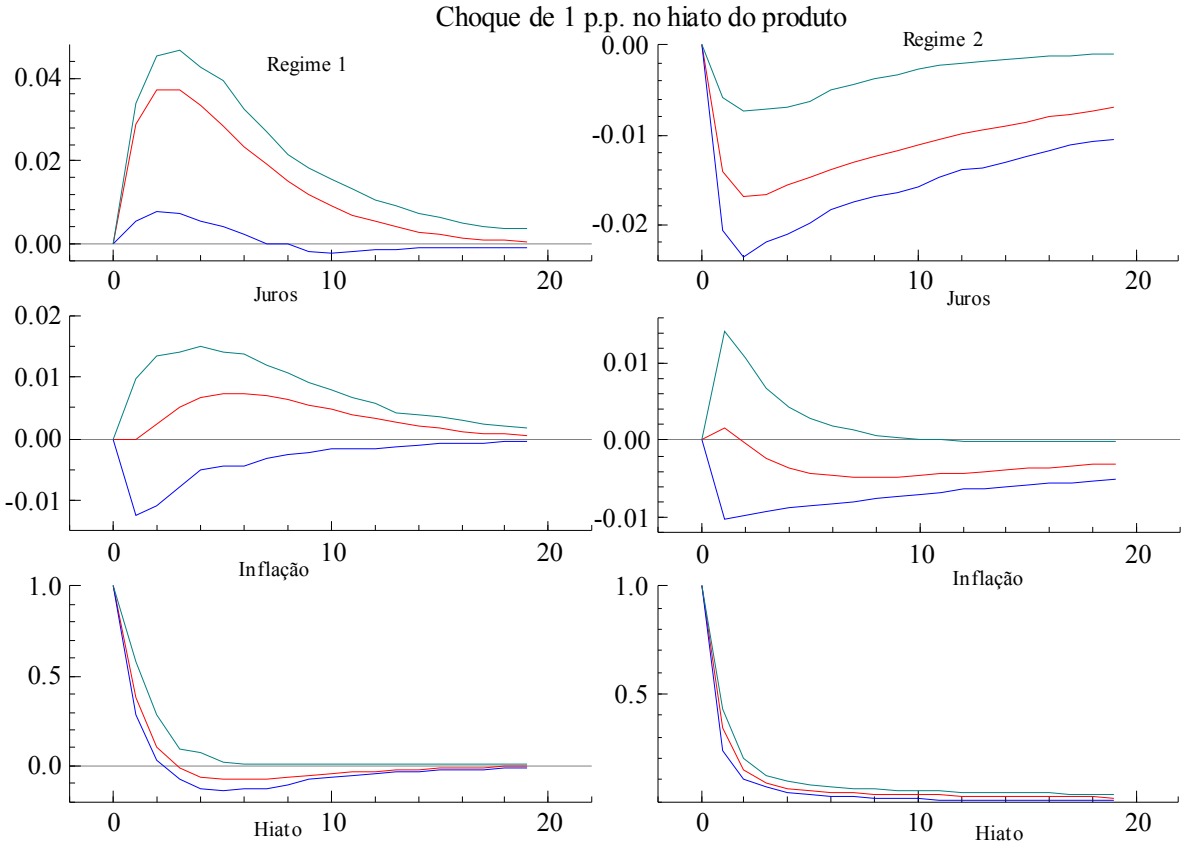
Em relação ao regime de estabilidade econômica (regime 1), é fácil notar o comportamento mais estável e *flat* das trajetórias das variáveis em questão. O aumento do hiato do produto leva a um aumento *flat* da taxa de inflação, o que implica em um aumento, também estável, da taxa de juros nominal. O aumento da taxa de juros em um ambiente com ausência de crises levou a um processo de valorização cambial, visto que ocorre um processo de atração de capitais, afetando o mercado de câmbio. Mais uma vez os resultados para o modelo de economia aberta se apresentaram mais coerentes e robustos em relação à teoria.

Gráfico 20: Funções de resposta a impulso, choque de 1 p.p. no hiato do produto



No gráfico 21 tem-se os efeitos de um choque de 1 p.p. no hiato do produto. No caso do regime 1, o Banco Central responde com aumento da taxa de juros nominais, dado que parte do aumento do hiato do produto se traduz em elevação da inflação. À medida que os efeitos dos choques vão se aproximando de zero, a variação da taxa de juros nominal vai diminuindo até se aproximar do valor zero. Já no período de crises, vemos a ocorrência de um fenômeno descrito anteriormente. O choque de 1 ponto percentual no hiato do produto é parcialmente repassado aos preços nos primeiros períodos, mas ocorre uma súbita zeragem dos efeitos do hiato do produto a da inflação a partir da data 4. Porém, não era de se esperar uma taxa de juros nominal negativa, e sim, positiva com o intuito de estabilizar os efeitos gerados a partir do choque. Porém, a rápida zeragem dos efeitos do hiato resulta em uma inflação negativa para o período, sendo premente uma intervenção do Banco Central no sentido de evitar deflação e estagnação econômica. Com isso, as variações na taxa de juros nominal passam a ser negativas, refletindo a necessidade de ajuste do Bacen para estabilizar as variáveis-meta em torno de seus valores de longo prazo. Interessante dizer que estes efeitos não seriam esperados em períodos de crise, até porque nestes períodos não se costuma verificar choques positivos no hiato do produto, o que novamente corrobora a baixa explicação das relações entre as variáveis econômicas para o modelo de economia fechada.

Gráfico 21: Funções de resposta a impulso, choque de 1 p.p. no hiato do produto

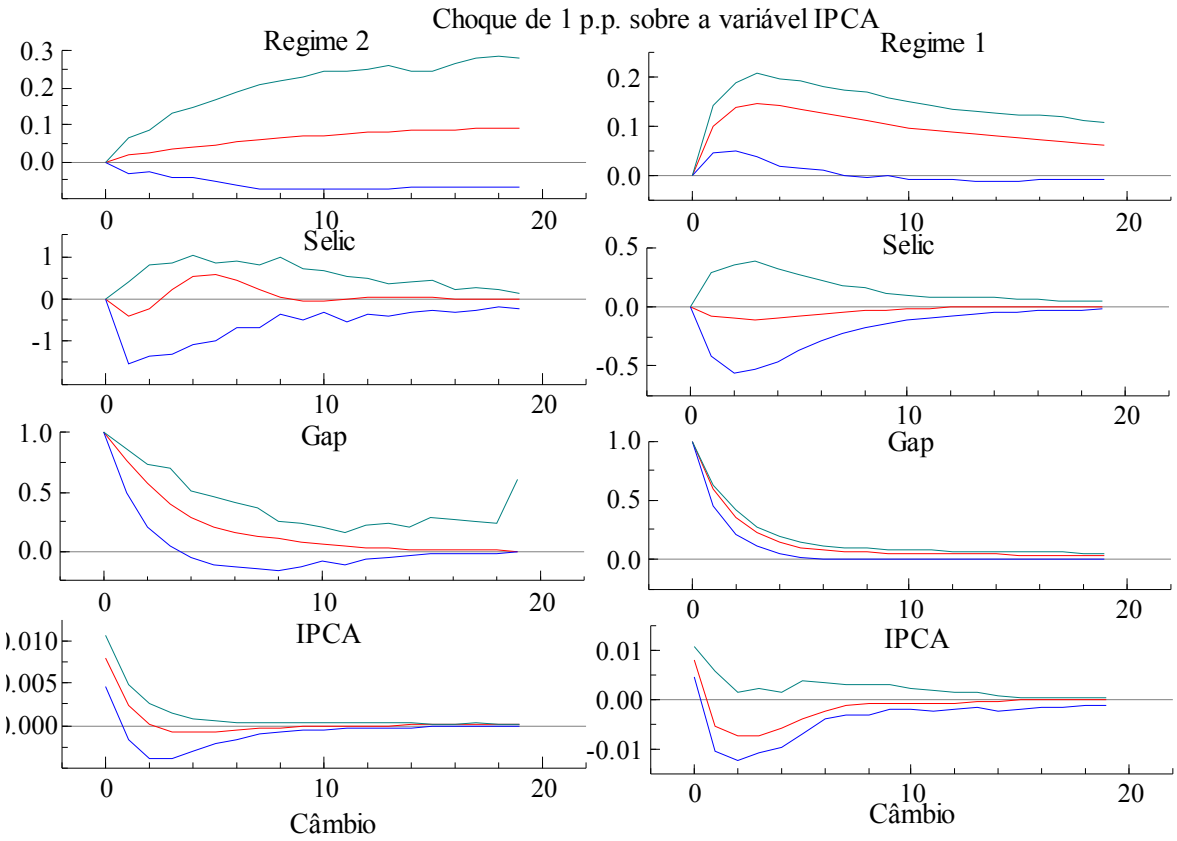


Por outro lado, no gráfico 22 é apresentado resultados para o choque de 1 p.p. sobre a taxa de inflação (IPCA). Para o regime de estabilidade econômica é perceptível que o aumento da taxa de inflação implica em um aumento de maior magnitude na taxa de juros nominal por parte da autoridade monetária, tendo em vista o controle monetário. Os efeitos sobre o hiato do produto são muito pequenos, visto que o aumento da taxa de juros já absorve os impactos do aumento da inflação, além do que o hiato do produto depende da seqüência esperada das taxas de juros. Com aumentos sucessivos da taxa de juros o hiato do produto continua caindo até retornar ao seu equilíbrio de longo prazo. O aumento da inflação combinado com o aumento da taxa de juros nominal gera uma desvalorização do câmbio nominal, que impacta a inflação doméstica. Este efeito é minimizado pelo aumento da taxa de juros que produz efeitos sobre o mercado cambial, o que demonstra a rápida zeragem dos efeitos cambiais.

Para o regime de crises os resultados são próximos, mas os fatos da presença de crises combinados com o aumento inflacionário fazem com que a taxa de câmbio se

desvalorize em um primeiro momento, porém, entende-se que a partir da data 2 os riscos associados ao mercado brasileiro são minimizados pelo comportamento da política monetária ortodoxa que aumenta bruscamente a taxa de juros, cessando a saída de capitais e revertendo o quadro de desvalorização cambial. A partir da data 7 vê-se que as variáveis voltam para suas trajetórias de longo prazo.

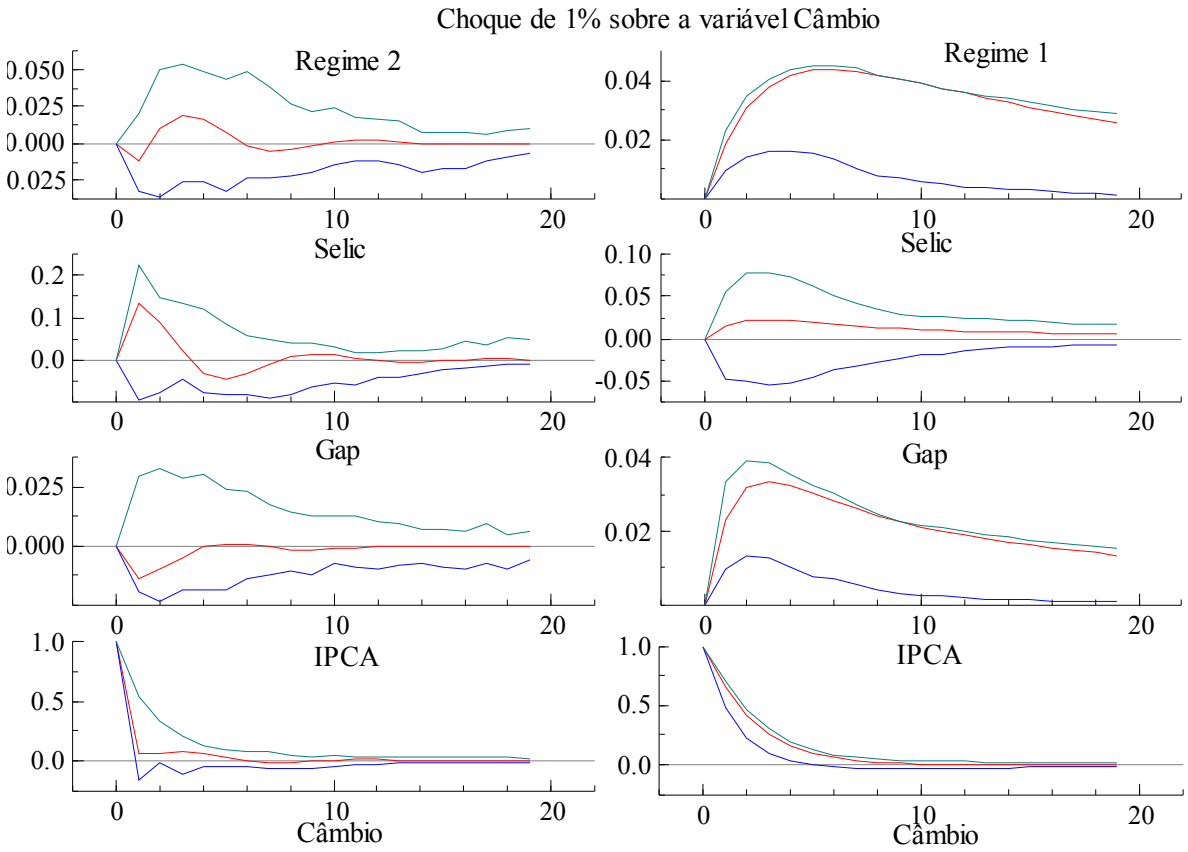
Gráfico 22: Funções de resposta a impulso, choque de 1 p.p. na taxa de inflação



Por fim, o choque de 1 p.p. sobre a variável câmbio é mostrado na figura 23. Para o regime de crise econômica pode-se ver que a desvalorização cambial desloca o hiato do produto, visto que os custos dos insumos importados caem e há um aumento das importações. Conseqüentemente, esperava-se um aumento da taxa de inflação doméstica visto os mecanismos de *pass-through* atrelados ao processo de desvalorização cambial, porém, os efeitos sobre a taxa de inflação são praticamente nulos. Tendo em vista este ambiente, a autoridade monetária fixa uma taxa de juros nominal mais alta, o que novamente surte efeitos sobre o mercado cambial, reduzindo os impactos do choque cambial após o período 7.

No que tange o regime de estabilidade econômica, vemos que o choque cambial é inteiramente absorvido pela taxa de inflação doméstica, que *per se* ativa a fixação de uma taxa de juros nominal mais alta, provocando a redução gradativa dos efeitos. Interessante frisar que o gráfico acima nos indica um comportamento de repasse mais imediato das variações da taxa de câmbio nominal para os preços domésticos.

Gráfico 23: Funções de resposta a impulso, choque de 1 p.p. na taxa de câmbio



A título ilustrativo, expomos as funções resposta a impulso sem a aplicação da metodologia de GI e adotando as hipóteses de um modelo de economia fechada e que os valores dos parâmetros são estáveis ao longo do período estudado[18]; o gráfico 24 descreve tais relações.

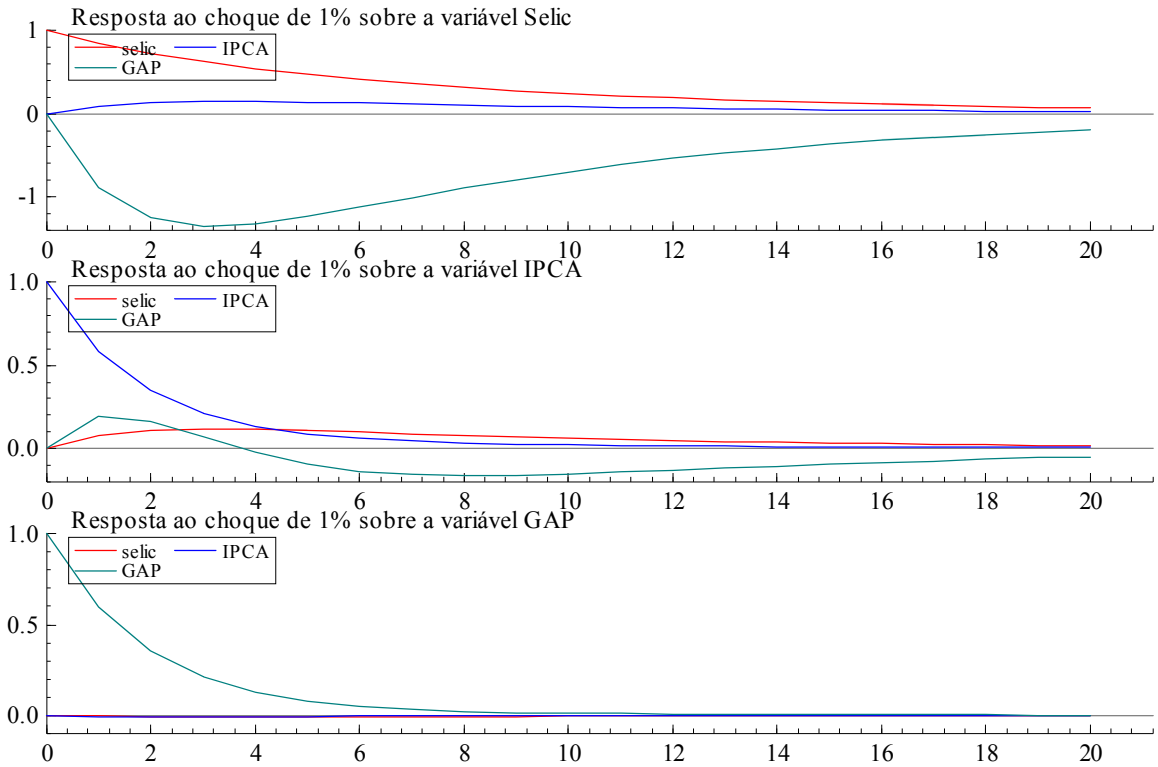
Pode-se avaliar que um choque de 1 p.p. na taxa de juros nominal (Selic) produz uma forte redução do *gap* do produto e a estabilidade no valor da inflação no período. Comparando com os demais padrões de resposta a choques em uma abordagem GI, averigua-

[18] Estimamos um modelo MSMH(2) – VAR(1) para o modelo macroeconômico proposto e obtemos seus padrões de resposta a choques.

se a menor aderência daquele modelo, generalizando para todos os períodos de recessão e expansão da economia brasileira os padrões de tais choques. Daí, supondo a ausência de diversos regimes, o modelo estimado MSMH(2)-VAR(1) apresenta problemas na interpretação dos choques calculados para o período 1994-2005, pois a instabilidade dos parâmetros não é considerada. Porém, é possível ver as inter-relações entre os mecanismos de transmissão da política monetária quando um choque positivo de 1 p.p. na inflação resulta, em um primeiro momento, no aumento da taxa de juros nominal e do *gap* do produto; e em seguida, o aumento da taxa de juros nominal reduz o *gap* do produto em períodos posteriores. Em outras simulações com períodos temporais de choques mais longos, obtém-se o tempo 26 para a zeragem de tal choque sobre as demais variáveis do sistema.

Já o choque de 1 p.p. sobre o hiato do produto não produz os efeitos esperados. A suposição da ausência de regimes que afetam os valores dos parâmetros pode explicar o comportamento ininteligível das demais variáveis do modelo ao longo do tempo. Portanto, o método GI empregado anteriormente produz resultados mais condizentes com a economia brasileira, além de ser um aperfeiçoamento metodológico para a estimação de padrões de resposta a choques aplicados à ela.

Gráfico 24: Funções de resposta a impulso, choque de 1 p.p. nas demais variáveis do modelo no período 1994-2005



CONSIDERAÇÕES FINAIS

Os resultados alcançados contam a história da economia brasileira ao longo do Plano Real. Com a Crise do México no fim de 1994 e problemas de natureza política no início de 1995, há interrupção dos influxos financeiros. A conseqüência imediata é o início de um processo de depreciação cambial, revertendo o movimento de apreciação do Real, presente desde julho de 1994. Como a maior parte destes movimentos derivava de entradas e saídas de capitais de curto prazo, o BC começa a intervir na liquidez do mercado de câmbio, ora comprando ora vendendo divisas, instituindo então um regime de *dirty floating*.

A extensão natural do processo foi a adoção do regime de *Crawling Peg*. Neste sistema, o BC definia "Faixas de Flutuação" dentro das quais o preço do dólar poderia oscilar e, caso a cotação fosse pressionada para cima ou para baixo, o BC intervinha vendendo ou comprando reservas. Porém, vale ressaltar, o objetivo da autoridade monetária é a estabilidade do nível de preços, sendo a defesa do regime cambial uma forma de alcançar este objetivo.

Lembrando que nos anos de *Crawling Peg* a política fiscal era mais frouxa do que deveria ser para manter o controle da inflação, a alternativa viável foi praticar uma política monetária mais restritiva, direcionada para reprimir a absorção interna. As variações na taxa de juros nominal passaram a ser bruscas, o que indicou não-linearidade no comportamento desta série, conforme discutido anteriormente. Em 1997 ocorreu a crise asiática, em 1998 a crise Russa, em 1999 a desvalorização cambial brasileira e em 2002 a crise das eleições presidenciais. Ocorreu grande coincidência na determinação dos períodos de crise do modelo TAR desenvolvido por Salgado, Garcia e Medeiros (2005) com os períodos associados ao regime 2 na metodologia de *Markov Switching* utilizada neste trabalho. Tais períodos estão identificados como os meses das crises de 1994, 1997, 1998, 1999 e 2002, conforme descrito acima, o que motiva a interpretação de chaveamento da função de reação, independente do algoritmo usado para a estimação da mudança de regime, seja ele o EM ou o MCMC. Devemos ter em mente também que a mistura de dados antes e depois de 1999 (período de desvalorização) implicou em conclusões mais suaves, visto que há o problema da adoção de políticas monetárias distintas para o período completo de análise.

Todas estas evidências nos fazem concluir que, de fato, o BCB fazia uso da política monetária com fins distintos durante os anos avaliados. Isto é, em alguns momentos os juros eram fixados de forma passiva, restritos pela condição de paridade, objetivando dar

credibilidade ao regime cambial e, quando os fluxos eram positivos, o controle inflacionário exigia juros mais altos que, para serem implementados, demandavam a imposição conjunta de medidas mais ortodoxas. Isto é corroborado pela maior magnitude dos coeficientes nos períodos de crise, demonstrando que a curva de reação era mais instável. As funções resposta a impulso geradas indicaram comportamentos diferentes para uma mesma série para cada regime proposto.

O trabalho ficou um pouco prejudicado pelo problema de instabilidade no processo de decomposição de Cholesky para os modelos MSIAH (k) – VAR (p) e MSIA (k) – VAR (p), reduzindo a possibilidade de comparação entre os parâmetros estimados pelo método MCMC em relação ao EM. Portanto, uma das agendas para o presente trabalho seria a melhor estruturação do algoritmo de simulação, podendo ser adotado um aperfeiçoamento a partir de "*Particle Filters*" com o intuito de se resolver o problema mencionado. Outra agenda de pesquisa é estimar outros formatos de regras monetárias e comparar a eficiência delas, anelando-se encontrar regras ótimas robustas.

REFERÊNCIAS

ALBERT, J. H. e CHIB, S. Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts. *Journal of Business and Economics Statistics*, 11, p. 1 – 15, 1993.

BALL, L. *Policy Rules for Open Economies*. In John Taylor, Monetary Policy Rules, University of Chicago Press, pg. 127-144, 1999.

BATINI, N. e HALDANE, A. G. *Forward-Looking Rules for Monetary Policy*. NBER Working Papers 6543, National Bureau of Economic Research, Inc. 1998

BEAUDRY, P e KOOP M. Do recessions permanently change output? *Journal of Monetary Economics*, Elsevier, vol. 31(2), pg. 149-163, 1993.

BEELDERS, O. Adaptative estimation and unit root tests. *Working papers*, Emory University, 1998.

BLANCHARD, O. e QUAH, D. *The Dynamic Effects of Aggregate Demand and Supply Disturbance*," Working papers 497, Massachusetts Institute of Technology (MIT), Department of Economics, 1988.

BLINDER, A. *Central Banking in Theory and Practice*. MIT Press: Cambridge, 1998.

BONOMO, M e BRITO, R. *Regras Monetárias e Dinâmica Macroeconômica no Brasil: Uma Abordagem de Expectativas Racionais.* Ensaios Econômicos da EPGE, nº 410, Escola de Pós-Graduação em Economia, Fundação Getúlio Vargas, 2001.

BUENO, R. *Taylor Rule under inquiry*. Seminários Acadêmicos PUC-RJ (mimeo). 2005.

CALVO, G. Staggered prices in a utility-maximizing framework, *Journal of Monetary Economics*, v. 12, pg. 383–98, 1983.

CALVO, G. e REINHART, C. *Fear of Floating*. NBER Working Papers 7993, National Bureau of Economic Research, 2000.

CHIB, S, Marginal Likelihood From the Gibbs Output, *Journal of the American Statistical Association*, 90, p. 1313-1321, 1995.

CHIB, S. e E. GREENBERG. Understanding the Metropolis-Hastings algorithm. *The American Statistican* 49, p. 327-335, 1995.

CLARIDA, R., GALI J. e GERTLER M. The Science of Monetary Policy: a New Keynesian Perspective. *Journal of Economic Literature*, v. 37 (4), p. 1661-1707, 1999.

CLARIDA, R. e GALI, J. Optimal monetary policy in open versus closed economies: An integrated approach. *American Economic Review*, v. 91(2), maio, p. 248-252, 2001.

CORRÊA, A. S. *Diferenças e Semelhanças entre Países da América Latina: Uma Análise de Markov Switching para os Ciclos Econômicos de Brasil e Argentina*. Banco Central do Brasil, Working Paper nº 80, 2003.

CHRISTIANO, J. L., EICHENBAUM M., EVANS, C. The effects of monetary policy shocks: Evidence from the Flow of Funds. *Review of economics and statistics*, v. 78, nº 1, p. 16-34, 1996.

_____. Sticky price and limited participation models: A comparison. *European Economic Review*, v. 41, nº 6, p. 1201-1249, 1997.

_____. *Monetary policy shocks: What have we learned and to what end?* NBER Working Papers 6400, National Bureau of Economic Research, 1998.

DEVEREUX, M. e LANE, P. *Exchange Rate Regime and Monetary Policy Rules for Emerging Markets*. mimeo, Unversity of British Columbia, 2000.

DICKEY, D. A. FULLER, W.A. Distribution of the Estimators for Autoregressive Time Series with a Unit Root, *J. American Statistical Assn.*, 74, 427-431, 1979.

EHRMANN, M., ELLISON, M., VALLA, N. *Regime-dependent Impulse Response Functions in a Markov-switching Vector Autoregression Model*, Bank of Finland, Discussion Papers 11/2001, Bank of Finland - Research Department, 2001.

ENDERS, W. *Applied Econometric Time Series*. Wiley Series in Probability and Mathematical Statistics, 1995.

EVANS, G. W. e HONKAPOHJA, S. Expectational stability of stationary sunspot equilibria in a forward-looking linear model. *Journal of Economic Dynamics and Control*, Elsevier, vol. 28(1), p. 171-181, 2003.

FAIA, E. e MONACELLI T. *Financial Exposure, Exchange Rate Regime and Fear of Floating*. mimeo, New York university e Boston College. 2002.

FAUST, J. e LEEPER, E. When do long-run identifying restrictions give reliable results? *Journal of Business and Economics Statistics*, v. 15, nº 3, p. 345-353, 1997.

FRIEDMAN, M. The role of monetary policy. *American Economic Review*, 1968.

GALI, J. e MONACELLI T. *Optimal Monetary Policy and Exchange Rate Variability in a Small Open Economy. Mimeo*, Boston College, 1999.

GIANNONI, Marc. P., and WOODFORD M. *Optimal Interest-Rate Rules: I. General Theory*. NBER working paper no. 9419, dez., 2002.

GUJARATI, D. *Econometria Básica*. McGraw-Hill, 2000.

HAMILTON, J. *Time Series Analysis.* Princeton: Princeton University Press, 1994.

HANSEN, B. E. *Rethinking the univariate approach to unit root tests: How to use covariates to increase power*, Econometric Theory, 1995.

HARRIS, G. *Regime Switching Vector Autoregressions: A Bayesian Markov Chain Monte Carlo approach*. Australian Mutual Provident Society, 2000. Disponível em www.actuaries.org/AFIR/colloquia/Cairns/Harris.pdf . Acesso em: <15/12/2005>.

_____. *Markov Chain Monte Carlo estimation of Regime Switching Vector Autoregressions*. ASIIN Bulletin, Vol 29, Nº I, 2000. Disponível em: < http://www.casact.org/library/astin/vol29no1/47.pdf > Acesso em: <15/12/2005>.

HUBER, P. J. Robust regression: Asymptotics, conjectures, and Monte Carlo: *Ann. Statist*., 1, 799-821, 1973

KING, M., The Inflation Target Five Years On, *Bank of England Quarterly Bulletin*, v. 37(4), 1997.

KLEIN, P. Using the generalized Schur form to solve a multivariate linear rational expectations model, *Journal of Economic Dynamics and Control*, 24(10), set., pag. 1405-1423, 2000.

KROLZIG, H.M. *Markov-Switching Vector Autoregressions. Modelling, Statistical Inference and Application to Business Cycle Analysis*, Lecture Notes in Economics and Mathematical Systems, Berlin: Springer, Volume 454, 1997a.

_____. *Econometric Modelling of Markov-Switching Vector Autoregressions using MSVAR for Ox*, Department of Economics, University of Oxford, 1997b.

_____. *Business Cycle Asymmetries: Characterisationand Testing Based on Markov-Switching Autoregression*,. The Warwick Economics Research Paper Series (TWERPS), nº 522, Departamento de Economia, University of Warwick, 1998.

LEE, K. e PESARAN M. Persistence Profiles and Business Cycle Fluctuations in a Disaggregated Model of UK Output Growth", *Ricerche Economiche*, , v. 47, pg. 293-322, 1993.

LEEPER, E., SIMS, C. e ZHA, T. What does monetary policy do? *Brooking papers on economic activity*, v. 2, p. 1-63, 1996.

LIMA, L. R, XIAO, Z. *Testing Unit Root Based on Partially Adaptive Estimation*, Econometric Society 2004 Latin American Meetings 63, Econometric Society, 2004.

LÜTKEPOHL, H. *Introduction to Multiple Time Series Analysis*. Springer-Verlag, 1991.

MIRA, A. *Introduction to Monte Carlo and MCMC methods*. University of Insubria, Italy. (mimeo). Disponível em <venda.uku.fi/research/FIPS/BMIP/pdffiles/**mcmc**-uk.pdf>. Acesso em <10/08/2005>.

MORAIS, I. A. *Mudança de Regime Markoviano: Uma aplicação a séries econômicas brasileiras*. 2003. Tese (Doutorado em economia) – Universidade Federal do Rio Grande do Sul.

MOREIRA, R. H. R. *Uma análise de Markov-Switching aplicada à regra de Taylor para a política monetária brasileira no período 1994-2003*. Associação Brasileira de Estatística (ABE)/SINAPI, 2004.

PAGAN, A. e ROBERTSON, J. Resolving the liquidity effect. *Federal Reserve Bank of St. Louis Review*, v. 77, nº 3, p. 33-54, 1995.

PESARAN, M, POTTER, S e KOOP, G. Impulse Response Analysis in Nonlinear Multivariate Models, *Journal of Econometrics,*Vol. 74, Nº 1, pg. 119-147, 1997.

PESARAN, M. e POTTER, S. A Floor and Ceiling Model of US Output, *Journal of Economic Dynamics and Control*, V. 21, Nº.4, pg. 661-695, 1997.

PHILLIPS, P. C. B. *Robust Nonstationary Regression*, Econometric Theory, Cambridge University Press, vol. 11(5), p. 912-51, 1995.

POTSCHER, B., I. PRUCHA. A Class of Partially Adaptive One-Step M-Estimators for the Nonlinear Regression Model with Dependent Observations, *Journal of Econometrics*, v. 32, 1986.

POTTER, S. M. A Nonlinear Approach to US GNP, *Journal of Applied Econometrics*, John Wiley & Sons, v. 10(2), pg. 109-125, 1995.

RIGOLON, F. *Regras de política monetária ótimas com expectativas racionais*. 2003. Tese (Doutorado em Economia) - Universidade de São Paulo.

ROBERTS, J. New Keynesian Economics and the Phillips Curve, *Journal of Money Credit and Banking* v. 27, pg. 975-984, 1995.

ROTEMBERG, Julio J., and WOODFORD M. *An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy in* B.S. BERNANKE and ROTEMBERG (eds.), NBER Macroeconomic Annual 1997, Cambridge, MA: MIT Press. 297-346, 1997.

RUDEBUSCH, G. Federal Reserve interest rate targeting, rational expectations end the term structure. *Journal of Monetary Economics*, v. 35, nº 2, p. 245-274, 1995.

SALGADO, M. J.; GARCIA, M. G. P.; MEDEIROS, M. C. *Monetary Policy During Brazil's Real Plan: Estimating the Central Bank's Reaction Function*. Revista Brasileira de Economia, 2005.

SIMS, C. A. *Macroeconomics and Reality*, Econometrica, Econometric Society, vol. 48(1), 1980.

_____. Are forecasting models usable for policy analysis? *Federal Reserve Bank of Minneapolis Quartely Review*, v. 10, nº 1, p. 2-16, 1986.

SIMS, C. e ZHA, T. *Errors bands for impulse responses.* Working papers 95, Federal Reserve Bank of Atlanta, 1995.

SÖDERLIND, P. Solution and Estimation of RE Macromodels with Optimal Policy. *European Economic Review,* 43, 813-823, 1999.

SVENSSON, Lars E.O. What Is Wrong with Taylor Rules? Using Judgment in Monetary Policy through Targeting Rules. *Journal of Economic Literature*, 2003.

TAYLOR, J. B. *Discretion versus Policy Rules in Practice*. Carnegie-Rochester Conference Series on Public Policy 39, p. 195—214, 1993.

_____. The Robustness and Efficiency of Monetary Policy Rules as Guidelines for Interest Rate Setting by the European Central Bank, *Journal of Monetary Economics*, v. 43, pg. 655-679, 1999.

_____. Low Inflation, Pass-Through, and the Pricing Power of Firms, *European Economic Review*, v. 44, pg. 1389-1408, 2000.

XIAO, Z. *Likelihood-Based Inference in Trending Time Series with a Root Near Unity*, Econometric Theory, Cambridge University Press, vol. 17(6), p 1082-1112, 2001.

WALSH, C. *Monetary Theory and Policy*. Cambridge, 2º ed., MIT Press. 2003.

WOODFORD, M. *Optimal Monetary Policy Inertia*. NBER Working Paper No. 7261, jul., 1999a.

_____. *Commentary: How Should Monetary Policy Be Conducted in an Era of Price Stability? In* Federal Reserve Bank of Kansas City, New Challenges for Monetary Policy, 1999b.

_____. Pitfalls of Forward-Looking Monetary Policy. *American Economic Review Papers* and Proceedings, v. 90, p. 100-104, 2000.

_____. Inflation Stabilization and Welfare. *Contributions to Macroeconomics*, v. 2(1), 2000. Disponível em < www.bepress.com >. Acesso em < 12/08/2005>.

_____. Optimal Interest-Rate Smoothing. *Review of Economic Studies*, v. 70, p. 861-886, 2003a.

_____. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press. 2003b.

WICKSELL, K. A influência da taxa de juros sobre os preços. In: Carneiro, R. (organizador). *Os Clássicos da Economia*. São Paulo: Editora Ática, 1997.

ANEXO A

Valores Críticos Assintóticos para P-ADF t-statistic

Tabela 1.Valores Críticos Assintóticos para P-ADF t-statistic

	Standard			Demeaned			Detrended		
λ^2	1%	5%	10%	1%	5%	10%	1%	5%	10%
1.0	-2.57	-1.94	-1.62	-3.43	-2.86	-2.57	-3.96	-3.41	-3.13
0.9	-2.57	-1.94	-1.61	-3.39	-2.81	-2.50	-3.88	-3.33	-3.04
0.8	-2.57	-1.94	-1.60	-3.36	-2.75	-2.46	-3.83	-3.27	-2.97
0.7	-2.55	-1.93	-1.59	-3.30	-2.72	-2.41	-3.76	-3.18	-2.87
0.6	-2.55	-1.90	-1.56	-3.24	-2.64	-2.32	-3.68	-3.10	-2.78
0.5	-2.55	-1.89	-1.54	-3.19	-2.58	-2.25	-3.60	-2.99	-2.67
0.4	-2.55	-1.89	-1.53	-3.14	-2.51	-2.17	-3.49	-2.87	-2.53
0.3	-2.52	-1.85	-1.51	-3.06	-2.40	-2.06	-3.37	-2.73	-2.38
0.2	-2.49	-1.82	-1.46	-2.91	-2.28	-1.92	-3.19	-2.55	-2.20
0.1	-2.46	-1.78	-1.42	-2.78	-2.12	-1.75	-2.97	-2.31	-1.95

Fonte: Hansen (1995).