

EA-S6



UNIVERSIDADE EDUARDO MONDLANE

FACULDADE DE CIÊNCIAS

**DEPARTAMENTO DE MATEMÁTICA E
INFORMÁTICA**

TRABALHO DE LICENCIATURA EM ESTATÍSTICA

TEMA

**CRESCIMENTO MONETÁRIO E SEU IMPACTO NA
INFLAÇÃO E TAXAS DE JURO**

**ESTUDO APLICADO AO MERCADO MONETÁRIO MOÇAMBICANO
PARA O PERÍODO DE 01/96 À 06/08**

ELÍSIO JOAQUIM JOSÉ MABASSO



UNIVERSIDADE EDUARDO MONDLANE

FACULDADE DE CIÊNCIAS

**DEPARTAMENTO DE MATEMÁTICA E
INFORMÁTICA**

TRABALHO DE LICENCIATURA EM ESTATÍSTICA

TEMA

**CRESCIMENTO MONETÁRIO E SEU IMPACTO NA
INFLAÇÃO E TAXAS DE JURO**

**ESTUDO APLICADO AO MERCADO MONETÁRIO MOÇAMBICANO
PARA O PERÍODO DE 01/96 À 06/08**

SUPERVISORA: Dra. RAFICA ABDUL RAZAC

AUTOR: ELÍSIO JOAQUIM JOSÉ MABASSO

MAPUTO, FEVEREIRO DE 2009

Agradecimentos

Inicialmente agradeço a Deus e ao divino Espírito Santo, que me iluminaram e guiaram ao longo dessa empreitada e me mantiveram firme e determinado para executar um bom trabalho.

À minha família, que me apoiou ao longo de toda a caminhada. Em especial, aos meus pais, minha avó e meus irmãos que me acompanharam, ajudaram, aconselharam e incentivaram em todos os momentos de minha vida para que hoje eu chegasse a essa ocasião tão especial. Esta é uma conquista conjunta e que devo tudo ao exemplo e integridade que aprendi e aprendo com vocês.

À senhora Ruth Sacate, minha tia, pela atenção, apoio material e compreensão diante dos sacrifícios impostos pela jornada.

À professora Dra. Rafica Abdul Razac, minha orientadora, pelas observações, pela atenção, empenho, seriedade e paciência demonstradas, pelas sugestões e comentários ao presente trabalho e pela disponibilidade e humildade para ajudar e buscar o melhor caminho possível.

Meus agradecimentos se estendem a todo o corpo docente do DMI/UEM que me permitiu melhorar e crescer como teórico e pessoa. Aos funcionários minha gratidão pela disposição, atenção e profissionalismo no desempenho de suas atribuições, especialmente ao Sr. Augusto, à dona Zulmira e ao Sr. Fanheiro da Biblioteca do departamento.

Aos meus colegas e amigos das turmas de 2004 a 2008 do curso de estatística, com os quais dividi momentos de alegria e tristeza ao longo da caminhada de desenvolvimento intelectual. Aos meus amigos: Venâncio Bata, Sérgio Fumo, Maurício Chemane, Gabriel Cossa, Paulo Ziaveia, Armando Macamo, Egídio Chaimite, Etelvino Langa e outros.

Ao Dr. Sérgio e o Dr. Inácio Manjama, economistas do Banco de Moçambique, pela disponibilidade e paciência nas entrevistas prestadas, o auxílio no acesso aos dados das séries da base monetária e taxas de juro. À Mónica, minha colega do turno pós-laboral, pelo apoio na obtenção da base de dados da variação mensal do IPC no Instituto Nacional de Estatística.

Por fim, ressalto que os erros que persistirem ao longo do presente trabalho são de minha completa responsabilidade.

Declaração de honra

Eu, Elísio Joaquim José Mabasso, filho de Joaquim José Mabasso e de Salva António Simbine, residente na Cidade de Maputo, Av. de Moçambique, Bairro de Nsalene, portador do B.I. nº. 110005537x emitido pelo Arquivo de Identificação de Maputo, declaro por minha honra que o presente trabalho apresentado ao curso de licenciatura do Departamento de Matemática e Informática da Faculdade de Ciências da Universidade Eduardo Mondlane, como requisito parcial à obtenção do título de licenciado em Estatística é fruto de uma grande e empenhada jornada de investigação científica, que culminou na sua concepção e não constitui, por conseguinte, nenhum plágio ou algo semelhante.

“Economists frequently formulate an economically well-specified model as the empirical model and apply statistical methods to estimate its parameters. In contrast, statisticians might formulate a statistically well-specified model for the data and analyze the statistical model to answer the economic questions of interest.”

Katarina Juselius

Resumo

O objectivo deste estudo é analisar as características dinâmicas da tendência do mercado monetário, especificamente o efeito do crescimento monetário sobre a taxa de inflação e taxas de juro de mercado na economia de Moçambique. Para contemplar estes objectivos, realiza-se a análise das funções impulso-resposta e decomposição da variância oriundas da estimação de regressões pelo método dos vectores autoregressivos (VAR). Os resultados empíricos do presente trabalho sugerem, segundo o teste de causalidade de Granger, que a base monetária é a variável mais exógena do sistema (ou seja, a que não pode ser explicada por qualquer outra variável no sistema), seguida da taxa de juro das operações sobre o activo dos bancos comerciais e taxa de inflação. A taxa de juro das operações do passivo revela-se a variável mais endógena e, por conseguinte, afectada por todas outras variáveis do sistema. A análise das relações moeda-juro e moeda-inflação pelas funções resposta-impulso e decomposição de variância levou às seguintes conclusões: (1) apenas a base monetária e a taxa de juro das operações do activo é que influem significativamente na determinação da taxa de juro das operações do passivo, enquanto a taxa de inflação explica uma margem bastante reduzida e não significativa da variação dessa taxa de juro; (2) na relação moeda-inflação, nota-se uma participação ligeira da base monetária na explicação da variância do erro de previsão da taxa de inflação, e uma participação quase nula da taxa de juro activa. Todavia, a expectativa que se estabelece, previamente, deduzida a partir da teoria económica inerente a relação de longo prazo do crescimento monetário e inflação contrasta com os resultados observados.

Palavras chave: Base monetária, Inflação, Taxas de juro e VAR

Sumário

CAPÍTULO I	9
1. Introdução.....	9
1.1 Definição do Problema.....	11
1.2 Questão de pesquisa.....	12
1.3 Hipótese básica.....	12
1.3.1 Hipótese secundária.....	13
1.4 Relevância do estudo.....	13
1.5 Objectivos.....	14
1.5.1 Objectivos específicos.....	14
CAPÍTULO II	15
2. Revisão da literatura.....	15
2.1 Mercado monetário.....	15
2.1.1 Banco central e oferta de moeda.....	15
2.2 Mercado monetário moçambicano.....	17
2.3 Procedimentos econométricos.....	19
2.3.1 Modelos macroeconómicos – enquadramento histórico.....	20
2.3.2 Estacionaridade.....	26
2.3.2.1 Testes de Raiz unitária.....	32
2.4 Aspectos teóricos sobre os modelos VAR.....	36
2.4.1 Derivação do modelo.....	38
2.4.2 Propriedades estocásticas do modelo VAR.....	41
2.4.3 Causalidade de Granger.....	43
2.4.4 Funções de resposta a impulso e decomposição de variância.....	44
2.5 Cointegração e mecanismo de correcção de erro (ECM).....	46
2.5.1 Tendências comuns.....	48
2.5.2 Determinação de cointegração no modelo VAR.....	49
2.5.3 Mecanismo de correcção de erro (ECM).....	49
2.5.4 Componentes determinísticas no modelo de cointegração.....	52
Cinco casos para inclusão de interceptos e tendências.....	53
CAPÍTULO III	57
3. Metodologia.....	57
3.1 Material e modelo de pesquisa.....	57
3.1.1 Fonte de dados.....	58
3.2 Métodos.....	58
3.2.1 Análise de dados.....	58
CAPÍTULO IV	64
4. Análise de resultados.....	64

4.1 Estacionaridade.....	64
4.2 Mercado monetário: modelo VAR	67
CAPÍTULO V	75
5. Conclusões.....	75
5.1 Recomendações	76
5.2 Limitações do estudo	76
Bibliografia.....	77
Glossário	80
Anexos.....	84

Lista de tabelas

Tabela 4.1: Teste sobre inclusão de intercepto e tendência.....	66
Tabela 4.2: Teste de raiz unitária baseado em ADF.....	67
Tabela 4.3: Teste de raiz unitária baseado em PP.....	67
Tabela 4.4: Critério de selecção do <i>lag</i>	68
Tabela 4.5: Teste de autocorrelação serial dos resíduos LM.....	69
Tabela 4.6: Teste de causalidade de Granger.....	70
Tabela 4.7: Decomposição de variância da taxa de juro passiva.....	72
Tabela 4.8: Decomposição de variância da taxa de inflação.....	74

Lista de figuras

Figura 4.1: Análise descritiva das series em níveis.....	65
Figura 4.2: Análise descritiva das séries em diferenças.....	66
Figura 4.3: Estabilidade do modelo.....	68
Figura 4.4: Resposta-impulso na taxa de juro passiva.....	71
Figura 4.5: Resposta-impulso na taxa de inflação.....	73

CAPÍTULO I

1. Introdução

O surgimento de um ambiente altamente competitivo, decorrente da globalização dos mercados e da necessidade de desenvolvimento, faz com que governos e agentes económicos tomem, a um determinado período, decisões visando o reajustamento das suas economias. Esse facto pode levar a que a curto e médio prazos se gere uma certa instabilidade nessas economias, na medida em que os analistas das políticas governamentais tomam suas decisões baseando-se em expectativas sobre o futuro e estas caracterizam-se fundamentalmente pela incerteza, isto é, pelo facto de ser impossível conhecer o futuro.

Assim, por exemplo, as decisões de como estabelecer um crescimento monetário de modo a estimular uma maior produção sem que no longo prazo se reflita em uma alta inflação baseiam-se em factos e informações, mas não escapam de avaliações e expectativas subjectivas. São estas avaliações que governam a opinião prevalecente e assim o funcionamento dos mercados e que são baseadas em elementos práticos e simplificadores tais como: experiência passada, instintos e desejos, factos e interesses, opiniões de especialistas etc.

Uma questão de interesse no estudo dos mercados é a seguinte: se o aumento sustentado no crescimento monetário, em última instância, se traduz em um aumento da inflação, o que determinará a taxa de crescimento da moeda? Um argumento frequente, segundo Sachs e Larrain (2000), é de que o crescimento monetário é resultado dos deficits de orçamento do governo.

Por conseguinte, variáveis económicas como inflação, procura de moeda, taxas de juro etc., vistas como séries temporais provavelmente alterarão, no longo prazo, o seu percurso *normal* em virtude desses choques monetários à economia. Se por um lado a teoria económica sugere a existência de uma relação de interdependência entre essas variáveis, no longo prazo, poder-se-ia esperar que exista nelas uma relação de equilíbrio¹. No entanto, os choques monetários, podem gerar situações de

¹ O conceito de equilíbrio abordado ao longo deste texto refere-se a séries temporais com uma evolução semelhante ao longo do tempo. Embora nalgumas situações, a teoria económica se refira ao equilíbrio do mercado como uma situação na qual determinada quantidade transaccionada num certo mercado, constitui, nesse mercado, a quantidade *ótima* a qual tanto as firmas/empresas e os agentes económicos desejam transaccionar. Ou seja, o ponto de equilíbrio refere-se a situação de um mercado em que a quantidade ofertada é igual a quantidade procurada.

desequilíbrio na relação entre as variáveis. Em tais situações supõe-se que a “economic theory will propose forces which tend to keep such series together.” (Engle & Granger, 1987, p. 251).

Segundo Engle and Granger (1987), “... equilibrium is a stationary point characterized by forces which tend to push the economy back toward equilibrium whenever it moves away.” (p. 251). Portanto, situações de desequilíbrio não são fenómenos permanentes, mas sim passageiras, pois, sempre que ocorrem aparecem forças no sentido de eliminá-las. É por isso que o ponto de equilíbrio pode ser considerado estável, embora isso não signifique necessariamente que os pontos de equilíbrio não sejam mutáveis, ou seja, sempre que ocorrer alguma situação de desequilíbrio em um determinado mercado a economia proporá forças no sentido de restabelecer o equilíbrio a um maior ou menor preço e/ou quantidade. (Silva, 1999).

Assim, sempre que um sistema em equilíbrio é alterado por um factor externo ocorre um deslocamento para o sentido que tende a anular a acção deste factor de modo a atingir um novo estado de equilíbrio.

O presente trabalho pretende apresentar uma síntese de modelos que possibilitem responder os mecanismos, ainda que apenas parcialmente, da dinâmica do mercado monetário, através da análise dos efeitos do crescimento monetário sobre a inflação e taxas de juro do mercado, utilizando uma abordagem econométrica baseada em modelos autoregressivos.

No capítulo I é feita a apresentação e fundamentação do tema bem como a definição do problema de estudo. São também apresentados os objectivos gerais e específicos a alcançar e as principais motivações que conduziram a pesquisa

O enquadramento teórico é apresentado no capítulo II. Neste capítulo são introduzidos os mecanismos de funcionamento de um banco central, em particular o Banco de Moçambique, assim como a base teórica que fundamenta a relação de longo prazo entre a moeda, inflação e taxas de juro nominais. Essa relação conduz à especificação de uma equação de longo prazo obtida através de um modelo VAR.

Dado que a grande maioria das séries económicas e financeiras apresenta evidências claras de não estacionaridade, introduzir-se-á este conceito e deduzir-se-ão testes que permitam a análise das

características não estacionárias dessas séries. Por causa da natureza do fenómeno estudado, são também analisadas com algum detalhe aspectos relevantes que conduzem ao conceito de cointegração.

O capítulo III apresenta uma descrição dos dados utilizados. Neste trabalho foram usadas as séries da taxa de inflação, os logaritmos naturais da base monetária e taxas de juro activas e passivas, incluindo um termo de tendência linear que mostra o comportamento da base monetária e duas *dummies* exógenas que captam os *outliers* nas séries da inflação e taxa passiva de juro.

Após a apresentação dos dados e da metodologia utilizada no capítulo III, o capítulo IV mostra os resultados dos testes empíricos realizados para verificar os conceitos deduzidos no enquadramento teórico. E finalmente apresentam-se no capítulo V as conclusões gerais.

1.1 Definição do Problema

Estando a economia em uma situação de equilíbrio, poderá haver um *timing* não previsível entre a instabilidade gerada pelos choques no mercado e o restabelecimento desse equilíbrio. Assim, vários factores podem concorrer para proporcionar melhor qualidade e quantidade de estímulos capazes de dinamizar o processo de desenvolvimento do equilíbrio. No mercado monetário, o banco central actua com o objectivo primordial de preservação do valor da moeda nacional e, portanto, tem a obrigação de disciplinar a actividade bancária.

Segundo Silva (1999), no cumprimento do seu papel o banco central deve ter um comportamento activo de modo a manter maior controle possível da evolução da base monetária, tentando evitar que a moeda se torne abundante ou rara ². Assim, muitos podem ser os focos da política das autoridades monetárias. Nos EUA, por exemplo, houve uma longa discussão sobre qual variável monetária o FED devia tentar controlar e, durante muito tempo, o FED procurou influenciar as taxas de juros de mercado sendo este o foco da política monetária ³. (Sachs & Larrain, 2000).

Um dos instrumentos mais usados para garantir a estabilidade da moeda nas economias modernas são as operações de *open market*, pois, segundo Silva (1999), todos os países têm dívidas públicas

² Situação essa que pode acarretar a perda do valor da moeda (eventualmente causada pela elevação contínua e persistente do nível geral de preços) ou uma valorização em demasia decorrente da sua escassez.

³ A designação FED ou Federal Reserve Board/Bank refere-se ao banco central norte americano.

bastante significativas representadas por títulos públicos e, por outro lado, pela exactidão mostrada no quanto o banco central expande ou contrai base monetária. Assim, se o banco central necessitar contrair ou expandir base monetária só precisaria vender ou comprar títulos públicos a um valor correspondente à magnitude dos meios de pagamento que pretendam *enxugar* ou emitir à economia.

Note-se que uma aplicação *indevida* das operações de *open market* pode gerar uma contracção ou expansão demasiada de base monetária. Segundo Cleto (2001), caso exista grande quantidade de moeda em circulação ocorre, para além da elevação do nível de inflação, uma queda na taxa de juros, sendo o inverso também verdadeiro. Mas o problema em elevar ou manter elevada a taxa de juros, é que inevitavelmente ocorre uma retracção no crescimento económico, aumentando a pobreza do país e de sua população.

Em muitas nações em desenvolvimento as negociações de títulos governamentais não são suficientes para tornar as operações de *open market* possíveis, pois, o mercado da dívida governamental costuma ser fraco. Assim como nas economias de inflação alta e imprevisível, e naquelas em que as pessoas não confiam na capacidade do governo em pagar as dívidas (Sachs & Larrain, 2000). Neste último caso a população não estará disposta a manter a dívida no sector público, ou então exigirá um prémio de juros muito alto para compensar o risco.

1.2 Questão de pesquisa

- ➔ Qual é o impacto do comportamento de longo prazo do crescimento monetário sobre as taxas de juro de mercado e inflação na economia moçambicana?

1.3 Hipótese básica

De modo a garantir a efectividade e idoneidade das decisões das autoridades monetárias, considere-se o seguinte:

H1: Se a economia tem sempre em vista alguma situação de equilíbrio, e sendo o ponto de equilíbrio mutável, isso sugere haver um *timing* entre dois pontos de equilíbrio sucessivos. Por essas e outras razões advindas das próprias características da teoria de cointegração, as variáveis dessa economia

podem ser vistas como sendo cointegradas. O *timing* é representativo dos desequilíbrios de curto prazo na relação de cointegração de longo prazo, que, por conseguinte, são negligenciáveis.

1.3.1 Hipótese secundária

A hipótese colocada, segundo Marconi e Lakatos (1991), deriva duma dedução lógica da teoria (a teoria da cointegração) ou de uma analogia, observações causais da natureza do fenómeno de outra ciência (da economia a econometria). Com o objectivo de englobar os aspectos não especificados na hipótese básica, de clarificar e de decompor pormenorizadamente a afirmação feita (Marconi & Lakatos, 1991), a hipótese secundária é explicitada nos seguintes termos:

h1: O conceito de cointegração ou de equilíbrio de longo prazo aponta apenas para as trajectórias comuns alcançadas por distintas variáveis ao longo do tempo. Neste caso, para que o mercado seja cointegrado será necessário que a base monetária, as taxas de juro e a inflação caminhem juntas ao longo do tempo, e a uma mesma *taxa* aproximadamente constante. É importante destacar que é possível que exista também uma componente aleatória na determinação das trajectórias das variáveis, e que essa componente pode amenizar ou aprofundar a influência das séries umas nas outras.

A razão principal para essa suposição é que, segundo Sachs e Larrain (2000), um aumento sustentado na taxa de crescimento de moeda irá, no longo prazo, quando todos os ajustamentos tiverem ocorrido, levar a um igual aumento na taxa de inflação. E também pelo facto de a equação de Fisher defender a existência de uma ligação positiva entre as taxas de juro nominais e a inflação ⁴.

1.4 Relevância do estudo

A taxa de juros definida por Leite (2000), como o preço pago para se tomar dinheiro emprestado. É uma das variáveis mais acompanhadas na economia, já que seu comportamento afecta as decisões de consumo, investimento, taxa de câmbio, a procura de moeda etc. Um aumento no *stock* de moeda

⁴ A equação de Fisher mostra que a taxa real (esperada) de juros é a taxa nominal de juros menos a taxa esperada de inflação. Com uma taxa de juros real aproximadamente constante no longo prazo, e com as expectativas da inflação se ajustando à inflação real, a taxa de juros nominal ajusta-se à taxa de inflação prevalecente. (Sachs & Larrain, 2000).

provoca, primeiro, a diminuição das taxas de juro e como resultado dessa diminuição, ocorre o aumento da procura agregada (Dornbush & Fisher, 1982).

A necessidade de um estudo quantitativo e sucinto do comportamento dinâmico da relação de interdependência entre as variáveis do mercado monetário é de extrema importância sobretudo para auxiliar as autoridades monetárias no processo de tomada de decisões sobre a conjuntura do mercado financeiro dispondo de informação mais acurada possível.

Enquanto alguns macroeconomistas acham que a relação entre a política e a economia sejam tão imprevisíveis e instáveis, Sachs e Larrain (2000), reiteram que outros defendem que o governo deve administrar a política orçamentária e a monetária de modo que esta influa na tendência geral da economia.

Esse estudo poderá servir como material de consulta a estudantes de graduação do curso de Estatística, e não só, no tratamento e análise de séries temporais económicas integradas e cointegradas, ou seja, com comportamento de longo prazo de equilíbrio. Bem como sobre a dinâmica dos ajustamentos de curto prazo na relação de estabilidade de longo prazo, no contexto das autoregressões vectoriais.

1.5 Objectivos

- ➔ O estudo presente tem como principal objectivo analisar as características dinâmicas da tendência do mercado monetário no período de 01/96 a 06/08 com suporte aos modelos econométricos de vectores autoregressivos (VAR) e mecanismo de correcção de erro (ECM).

1.5.1 Objectivos específicos

- ➔ Avaliar o quanto a série temporal da base monetária afecta as variâncias da inflação e taxas de juro;
- ➔ Sugerir, segundo a teoria estatística, qual variável pode constituir-se numa forte candidata como foco da política monetária do banco central moçambicano.

CAPÍTULO II

2. Revisão da literatura

Neste capítulo realiza-se uma abordagem básica do mecanismo de funcionamento do mercado monetário, seus instrumentos e em particular o mercado monetário moçambicano. Subsequentemente, trata-se da metodologia de séries temporais, especialmente a modelagem dos vectores autoregressivos (VAR) e autoregressivos de correcção de erro (ECM).

2.1 Mercado monetário

Existe em função de que todas as operações comerciais da economia são realizadas através da moeda, “pelo menos como determinante de valores, como no caso de troca directa de mercadorias, a moeda sempre está presente” (Silva, 1999, p. 37). Por sua vez, o conceito de moeda pode ser entendido, segundo Leite (2000), como o valor dos instrumentos que podem ser usados como meios de pagamento. Os principais meios de pagamento são o papel moeda em poder do público e os depósitos à vista (que formam o conceito M_1 do banco central). Outros tipos de haveres financeiros como: depósitos de poupança, depósitos a prazo e títulos do governo são incluídos em conceitos mais gerais (M_2 , M_3 e M_4).

2.1.1 Banco central e oferta de moeda

Um factor que influi na estabilidade dos preços é a inflação. É ela a responsável pelo aumento contínuo e generalizado no nível de preços. Contudo, aceita-se que um pouco de inflação seja integrante dos ajustes de uma sociedade em crescimento, porque esse avanço económico dificilmente se realiza sem que ocorram elevações dos preços (Leite, 2000).

Tal facto se sustenta na ideia de que variações positivas na taxa de crescimento monetário aumentam primeiro a produção e por último se reflectem totalmente em inflação mais alta ⁵. (Sachs & Larrain, 2000). Portanto, essa relação apenas se estabelece no longo prazo. O impacto de curto prazo de uma variação no crescimento monetário pode diferir, dependendo do estado das expectativas de inflação.

⁵ Os monetaristas argumentam que a inflação é um fenómeno monetário, ou seja, altas taxas sustentadas de crescimento monetário produzem altas taxas de inflação. Além disso, essa afirmação significa que altas taxas de inflação não podem continuar por muito tempo sem altas taxas de crescimento monetário. (Sachs & Larrain, 2000).

Nas modernas economias capitalistas o banco central é uma instituição, normalmente pública, cuja principal tarefa é a manutenção do valor da moeda nacional, ou seja, da estabilidade dos preços dos bens e serviços transaccionados nos limites do país (Silva, 1999).

O *stock* líquido de papel-moeda (cédulas e moedas metálicas) emitido pelo banco central desde os primórdios da história económica de um país, a base monetária, representa um montante que se acumulou ao longo do tempo. O banco central somente emite base monetária aos agentes económicos em troca de algo que tenha ou represente algum valor, que passa a ser sua propriedade, tornando-se parte do seu activo, enquanto o papel-moeda entregue ao agente (a base monetária emitida) representa um acréscimo do passivo (compromisso). (Silva, 1999).

Como o objectivo do banco central é o de controlar a moeda na sua utilização final como meio de pagamento, já que é nesse momento que ela pode se tornar abundante ou rara, devido a indisposição de meios directos de controlar a moeda em poder do público e o volume de depósitos à vista, na execução de tal função ele procura actuar na sua origem (emissão de base monetária) e na sua trajectória (no processo de multiplicação através dos bancos comerciais). (Silva, 1999).

Para realizar esse controlo, o banco central usa, segundo Silva (1999), conjunta ou separadamente, mecanismos tradicionais chamados instrumentos de política monetária:

(1) *Taxa de reservas obrigatórias* — refere-se a parcela dos depósitos à vista dos bancos comerciais que, por exigência legal, fica retida no banco central. Alterando essa taxa de depósito, o banco reduz ou aumenta a disponibilidade dos bancos comerciais para emprestar dinheiro aos agentes económicos. O que por consequência, limitaria ou expandiria o processo de multiplicação monetária.

(2) *Taxa de juro de redesconto* — o banco central pode influir na oferta monetária através dos empréstimos ao sector privado. Os bancos comerciais, por exemplo, podem usar essa opção de crédito com o propósito de ajustar suas reservas de dinheiro para o caso de ficarem abaixo do nível desejado ou exigido pelas regulamentações do banco central, ou para obter fundos que os bancos possam emprestar aos clientes, se as condições de mercado forem favoráveis.

(3) *Operações de mercado aberto* — através do processo de venda ou compra de títulos, o banco central emite ou contrai base monetária e, por conseguinte, os meios de pagamento.

(4) *Operações de câmbio* — os importadores e exportadores necessitam quase sempre adquirir moeda nacional ou estrangeira para poder viabilizar seus negócios, vendendo ou comprando moeda. Assim, se em um determinado ano o banco central comprar mais do que vender moeda estrangeira, essas operações vão gerar um aumento líquido da base monetária sobre a economia. Mas, se a autoridade monetária vender mais do que comprar moeda estrangeira, a base monetária vai diminuir em consequência dessas operações ⁶.

Um dos grandes desafios com o qual os bancos centrais se deparam e que, por sinal, dificulta o seu funcionamento, é que a origem do crescimento monetário se deve fundamentalmente ao financiamento dos deficits do orçamento do governo. Ora, a possível ligação entre os deficits orçamentais e o crescimento monetário pode ser a intenção do governo em aumentar deliberadamente o *stock* monetário como um meio de obtenção de receitas para o governo no longo prazo. (Sachs & Larrain, 2000).

2.2 Mercado monetário moçambicano

O BM – Banco de Moçambique é o banco central do país com sede em Maputo. O seu capital está integralmente subscrito e realizado pelo Estado da República de Moçambique (BM, 2003).

O programa monetário anual moçambicano é desenhado em consonância com os objectivos finais da política económica do governo de: (1) crescimento real do PIB a uma determinada ordem, (2) manutenção da inflação média anual não superior a um certo valor e (3) constituição de reservas internacionais líquidas estipuladas a um montante (normalmente em milhões, USD), de modo a garantir algum período de importações de bens e serviços não factoriais pelas reservas internacionais brutas no final do ano ⁷.

Tendo em conta a relação funcional entre a base monetária (BaM) e a inflação, objectivo final da política monetária do BM, esse agregado monetário passou a ser, a partir de 2006, a variável

⁶ É necessário realçar que essas relações ocorrem sobretudo em países onde o regime de câmbio adoptado é fixo, isto é, onde o banco central seria obrigado a vender ou comprar moeda estrangeira da população a uma taxa fixa. Pois quando o regime é flutuante, o banco central não intervém no mercado de câmbio e deste modo as transacções de câmbio não tem efeito sobre a base monetária (Sachs & Larrain, 2000).

⁷ Em economia, os serviços são denominados “serviços dos factores” quando se referem aos factores de produção (trabalho, capital, natureza e capacidade empresarial) e suas respectivas remunerações (salários, juros, alugueis e lucros), ou de “serviços não factores”, no caso de todos os demais, tais como fretes, seguros, transportes etc. (Silva, 1999).

operacional da política monetária do BM (BM, 2006). Segundo o BM, o princípio da BaM como variável operacional da política monetária está consagrado na estratégia de política monetária de médio e longo prazos com o objectivo de “preservação do valor da moeda nacional”, traduzindo-o em inflação baixa e estável, e define as operações dos mercados interbancários como o principal instrumento da política monetária.

Para garantir o cumprimento destes objectivos e, sobretudo o de inflação, a política monetária continuou, de 2006 a 2007, a privilegiar as operações de mercado aberto, feitas através do Mercado Monetário Interbancário (MMI), para a regulação da liquidez no sistema, por via da emissão e colocação de Bilhetes de Tesouro (BT's), e a venda de divisas no Mercado Cambial Interbancário (MCI) (BM, 2007).

Os BT's são valores mobiliários escriturais representativos de empréstimos de curto prazo da República de Moçambique, denominados em moeda nacional e destinam-se ao financiamento do défice da tesouraria corrente do Estado e às intervenções do BM no âmbito da política monetária. Os BT's são vendidos em sistema de leilão aos operadores do mercado monetário (BM, 2003).

O BM actua no MMI com as seguintes taxas de referência: (1) taxa de juro de facilidade permanente de depósito (FPD) – a taxa a qual o BM aplica aos depósitos de moeda nacional efectuados pelos operadores do mercado que se encontrem com excessos temporários de liquidez e por iniciativa destes, com o prazo de um dia útil; (2) a taxa de juro de facilidade permanente de cedência (FPC) – a taxa que o BM aplica nas cedências de liquidez aos operadores do mercado que se encontrem com défices temporários de liquidez, por iniciativa destes, essa taxa é usada pelo BM para sinalizar a direcção da política monetária em função do desempenho dos indicadores fundamentais da economia; (3) taxa de juro da facilidade de última hora (FUH) – a taxa que o BM aplica nas cedências de liquidez aos operadores que necessitem de fundos para cumprir com o requisito de reservas obrigatórias no último dia do período de constituição durante a meia hora a seguir ao fecho normal do mercado; (4) taxa de juro de cedência por leilão – é a taxa que o BM aplica nas cedências de liquidez, por sua própria iniciativa, aos operadores do mercado através de leilão.

O MMI é um segmento do mercado monetário do Metical, regulamento, no qual as instituições autorizadas permutam fundos representados por saldos das suas contas de

depósito à ordem no BM ou valores mobiliários desmaterializados inscritos em contas-título neste mesmo banco, visando equilibrar os excedentes e necessidades de moeda primária entre as instituições monetárias. Neste mercado o BM pode também intervir, absorvendo ou cedendo liquidez, sendo estas operações sempre realizadas através da compra, venda ou emissão de títulos (BM, 2003, p. 72).

Dado que o banco central não concede crédito a sectores não ligados ao sistema financeiro, então, é razoável supor que uma variação nas taxas de juro de intervenção do BM reflectir-se-ia nas taxas de juro das operações (activas e passivas) dos bancos comerciais⁸. Ou seja, as taxas das operações bancárias em que os bancos comerciais actuam como *captadores* de fundos, assumindo a qualidade de devedores perante os depositantes a quem pagam um certo juro (operação passiva). Esses fundos são, por sua vez, geradores de crédito (operação activa).

Por essa razão, a taxa de juro a cobrar pelos bancos pelo empréstimo de dinheiro designa-se por activa porque se inscreve nas contas do activo dos bancos. Enquanto que a taxa de juro que remunera os depósitos bancários designa-se por passiva porque se inscreve nas contas do passivo dos bancos, visto representar uma responsabilidade. O interesse por essas taxas, é que são elas que determinam a expansão da base monetária na economia e, segundo Maleiane (2001), a subida das taxas activas afecta imediatamente o crédito e cresce a propensão pelo crédito mal-parado.

2.3 Procedimentos econométricos

O exame empírico das inter-relações macroeconómicas entre as variáveis utilizadas pode ser realizado através do ramo da economia chamado econometria⁹. No presente caso, a econometria de séries temporais.

⁸ Além das instituições do mercado financeiro, o BM concede crédito ao governo moçambicano para financiar os défices do seu orçamento. Maleiane (2001), refere que na questão do crédito líquido ao governo, o estado deveria procurar fontes alternativas de financiamento e não utilizar aquilo que não aplicou no passado sem resultar de esforço próprio – de outra forma tal actuação será emissão porque, se não resultou dos impostos – a quantidade de dinheiro injectado na economia há-de aumentar.

⁹ Apesar de a econometria não representar um fim em si mesma, não se constituindo um mecanismo de aferição da validade das teorias, confirma-se seu carácter de auxílio à prática da ciência como ferramenta para observação da aplicabilidade da teoria económica à vida prática (Resende, 2006).

O facto de as variáveis serem medidas segundo uma sequência cronológica configura-se como um elemento característico das mesmas e, por isso, requer uma estrutura de análise totalmente particular como é o caso da análise das séries temporais. Série temporal é constituída de observações sequenciais, ou seja, de dados a respeito de uma determinada variável oriundos de um processo aleatório de geração de dados. O objecto de estudo das séries temporais nada mais é que o estudo do processo estocástico gerador dessas realizações particulares, sendo que segundo Mills (1993, citado por Resende, 2006) a relação entre a realização e um processo estocástico é análoga à existente entre a amostra e a população na estatística clássica.

Tal processo, por sua vez, representa o mecanismo de geração da série estudada e, assim, através da realização, o estudo de série temporal visa levantar as características determinantes do processo gerador de dados. Cabe ressaltar, contudo, que uma crítica recorrente ao uso de séries temporais é que este tem como hipótese central a validade do princípio da *ergodicidade*¹⁰ (Camuri, 2007).

2.3.1 Modelos macroeconómicos – enquadramento histórico

De acordo com Stock and Watson (2001), os macroeconometricistas encarregam-se de quatro tarefas: descrever e resumir dados macroeconómicos, fazer previsões, quantificar o que se sabe ou o que não se sabe em relação a estrutura da macroeconomia, e aconselhar os planificadores das políticas macroeconómicas.

No cumprimento dessas tarefas, os macroeconometricistas dispõem essencialmente de cinco abordagens com base em séries temporais: métodos de suavização exponencial, modelos de regressão com única equação, modelos de regressão com equações simultâneas, modelos autoregressivos integrados e de média móvel, e autoregressões vectoriais (VAR) (Gujarati, 2006).

¹⁰ Um processo estocástico é dito *ergódico* se uma única realização do processo é suficiente para caracterizá-lo. Na análise de séries temporais existe uma única realização do processo disponível e, portanto, precisamos supor que o processo subjacente é *ergódico*, pois, usa-se apenas uma das suas realizações para caracterizá-lo. O facto de se utilizar realizações particulares com objectivo de se estudar o processo estocástico só é válido caso o processo seja *ergódico*, ou seja, segundo Mills (1993, citado por Resende, 2006), que “os momentos da amostra para intervalos finitos da realização aproximam-se dos valores da população à medida que o tamanho da realização torna-se infinito” (p. 60). Diante do exposto assume-se que as séries de tempo possuem esta propriedade.

Por volta dos anos 60, os modelos usados na macroeconomia descreviam a economia agregada consistindo de um sistema de equações estruturais simultâneas: uma equação para descrever a procura de moeda, uma para o consumo, uma para o investimento etc. (Chari, 1999).

O modelo é estrutural porque cada equação do sistema representa o comportamento de um conjunto de agentes económicos e, simultaneamente, porque os valores correntes das variáveis aparecem nas equações. O conceito “estrutural”, segundo Tesfatsion (2007), refere-se a uma situação em que tendo estimado formas funcionais para um determinado modelo de economia, tais formas são ditas estruturais se forem invariantes à mudanças na política governamental.

Segundo Chari (1999), a formulação do modelo era considerada bastante atractiva porque o modelo era matematicamente explícito e os parâmetros das equações poderiam ser estimados pelos procedimentos econométricos até aí desenvolvidos sob patrocínio da Comissão Cowles (baseada nas teorias *revolucionárias* keynesianas do período pós-depressão). Segundo o autor, estes modelos eram amplamente usados para responder questões como: “como é que a administração da política monetária afecta a inflação, o desemprego e o produto?” (p. 6).

Um dos aspectos interessantes nestes modelos é que mostravam um reconhecimento claro do desejo de tornar precisa a relação entre a macroeconometria e a teoria macroeconómica, o que segundo Simkins (1999), constitui o resultado de um grande debate que desde os meados do século 20 ocupou os economistas sobre o papel da medição e da teoria nas investigações macroeconómicas, amplamente caracterizado como uma controvérsia sobre *medição sem teoria e teoria sem medição*.

A principal dificuldade em desenvolver os fundamentos dessa inter-relação foi o facto de as várias questões macroeconómicas lidarem com dinâmica e incerteza, pois, uma escolha individual de quanto consumir hoje, é necessariamente uma escolha de quanto consumir no futuro, uma decisão de investimento é baseada na expectativa do ganho futuro, portanto, como modelar esta decisão e qual o espaço no qual as antecipações são feitas e revisadas (Chari, 1999).

Todavia, na década de 70 esses modelos foram fortemente criticados, pois, infelizmente nem sempre as teorias económicas são ricas o suficiente para fornecer especificações dinâmicas capazes de identificar todas as relações entre os agentes económicos. Lucas (1976, as cited in Chari, 1999)

sugere que é ingénuo tentar prever os efeitos de uma mudança na política económica completamente na base de relações de dados históricos, especialmente dados históricos altamente agregados.

Este contributo invalidava assim todos os conselhos políticos advindos de conclusões baseadas em modelos de sistemas de equações simultâneas. Como consequência, isso mostrava que os parâmetros desses modelos não eram estruturais (invariantes com a política), ou seja, mudariam sempre que a política mudasse e, por isso, conclusões políticas baseadas nesses modelos seriam potencialmente incorrectas (Tsfatsion, 2007).

Outro grande contributo foi o de Sargent and Sims (1977), que criticaram os modelos estruturais do ponto de vista da confiança nas “incríveis restrições” feitas para identificar estatisticamente os modelos, o que até certo ponto colocava em causa as teorias estatísticas aplicadas, ou seja, *medição sem teoria*.

Sims (1980, as cited in Simkins, 1999) criticou as restrições de identificação embutidas nesses modelos em três grandes frentes: (1) enquanto as restrições usadas para identificar equações individuais em modelos estruturais extensos pode parecer razoável quando analisada uma equação de cada vez num equilíbrio parcial fixo, quando tomadas juntas como sistema de equações, tais restrições parecem pouco razoáveis; (2) afirmou que os seguidores dessas abordagens faziam suposições enganosas sobre a exogeneidade das variáveis que aparecem nos modelos, mas que não eram justificadas pela teoria económica nem estatística; (3) mostrou que as restrições de exclusão frequentemente empregues em modelos estruturais extensos não eram consistentes com o tratamento racional das expectativas ¹¹.

Segundo Simkins (1999), muitas variáveis rotineiramente excluídas das equações estruturais, para propósitos de identificação, poderiam ser razoavelmente incluídas se os indivíduos exibissem “expectativas racionais”. Por isso, Sargent and Sims (1977), sugeriram que fosse necessário considerar a estimação de uma representação geral das variáveis nos modelos como um vector de processos estocásticos. Pois, segundo eles muito pouca teoria subjacente aos modelos até então usados era explicitamente estocástica.

¹¹ Um modelo que não exhibe expectativas racionais mostra o comportamento da acção política da economia como se fosse não estocástico do ponto de vista do público, ou seja, ignorando as expectativas do sector privado antes e/ou depois da mudança na política económica.

O resultado desses desafios colocados aos modelos estruturais Keynesianos, anteriormente propostos, era de mudar o foco das investigações macroeconómicas voltadas aos ciclos de negócio e desenvolver modelos macroeconómicos capazes de produzir comportamentos cíclicos.

O mais *brilhante* defensor das novas evidências empíricas emergentes foi Christopher Sims. Que na base da crítica sobre os modelos estruturais focou sua atenção em modelos estatísticos do comportamento económico que pudessem descrever os ciclos de negócios e analisar os resultados da política económica. Tais modelos ficaram conhecidos como Vectores Autoregressivos (do inglês *Vector Autoregressive*) – VAR – focados nas propriedades dinâmicas de séries temporais de um sistema irrestrito (pela teoria) de equações. Assim, Sims desviou-se da ênfase nos trabalhos macroeconómicos em testar hipóteses para *arquitectar* as propriedades dinâmicas subjacentes aos modelos estatísticos, “deixando os dados falarem por si”.

O nome autoregressivo advém do facto de as variáveis serem explicadas pelos seus valores passados, e passados das restantes variáveis incluídas no modelo. Nestes modelos não existe nenhuma distinção *a priori* sobre quais variáveis são endógenas ou exógenas, tratando-as de forma simultânea.

Segundo Greene (1997), “...the VAR is simply an overfit reduced form of some simultaneous equation models. The overfitting results from the possible inclusion of more lags than would be appropriate in the original model ¹²” (p. 815). Os defensores dos modelos VAR ressaltam como vantagens de sua utilização: (1) a simplicidade do método, uma vez que não parte da determinação *a priori* de quais variáveis são endógenas ou exógenas; (2) a simplicidade de estimativa, dado que o método OLS (*ordinary least squares*) pode ser aplicado a cada equação de forma separada; (3) geralmente, as previsões obtidas com esses modelos são melhores que as alcançadas com modelos mais complexos. Além disso, os modelos VAR utilizam variáveis desfasadas como explicativas, o que por definição garante que não estejam correlacionadas com os resíduos, ou seja, não há autocorrelação dos resíduos.

Durante o ano de 1980, o crescente apoio no uso de modelos estatísticos na forma reduzida como o de Sims, para descrever o comportamento da macroeconomia, tornaram do modelo VAR um

23

¹² Dados dois vectores, Y_t de variáveis endógenas e X_t exógenas. Chama-se forma reduzida do modelo estrutural $BY_t + CX_t = \mu_t$; a equação $Y_t = \Pi X_t + v_t$, onde $\Pi = -B^{-1}C$ e $v_t = B^{-1}\mu_t$.

instrumento comum na descrição das propriedades dinâmicas de séries temporais económicas, e deste modo era visto por alguns economistas como substituto competitivo para os modelos estruturais até então usados (Simkins, 1999).

Segundo Greene (1997), alguns autores defendem que os modelos VAR sem qualquer restrição têm um melhor desempenho na previsão do que os modelos estruturais de múltiplas equações, embora essa visão, segundo ele, seria puramente empírica. Mesmo assim, o modelo desenvolvido por Sims sofreu vários ataques, dentre os quais, os argumentos dos críticos a modelagem VAR são: (1) o modelo não é teórico uma vez que parte de menos informação prévia, isto é, *measurement without theory*; (2) o problema prático da determinação do tamanho da defasagem a incluir; (3) caso as variáveis do modelo não se mostrem conjuntamente estacionárias, torna-se necessário proceder a uma transformação adequada de modo a alcançar tal condição, o que pode se mostrar muito trabalhoso ¹³. (Gujarati, 2006).

Cooley and LeRoy (1985, as cited in Simkins, 1999), criticaram severamente o modelo proposto por Sims argumentando que enquanto este modelo não é bom nem mau, a interpretação do comportamento dinâmico imposto por ele é frequentemente defeituosa. Eles argumentaram ainda que, por exemplo, a interpretação dada as funções de resposta a impulso, um instrumento *standard* usado para resumir o comportamento dinâmico das variáveis do modelo, é incorrecta porque depende da escolha da ordenação das variáveis por parte do investigador via resultados observados sob as correlações condicionadas nos testes de causalidade de Granger. De acordo com Cooley and LeRoy (1985, as cited in Simkins, 1999), “identifying these conditional correlations with causal orderings ... is justified only under predeterminedness assumptions which is untestable in the absence of prior restrictions derived from theory” (p. 16).

Estas críticas focavam a necessidade do uso da teoria económica para restringir a interacção entre as variáveis no modelo VAR de modo a obter interpretações válidas acerca do seu comportamento dinâmico, dado que a interpretação dada pelas funções de resposta a impulso e decomposição de variância é estrutural enquanto que as restrições necessárias para sua identificação são *ad hoc*.

¹³ Uma série temporal é estacionária, no sentido estrito, quando todos os momentos de sua distribuição de probabilidade não variam ao longo do tempo. O conceito de estacionaridade é tratado em pormenor na próxima secção.

O conjunto de todas essas críticas impulsionou um certo progresso no desenvolvimento de modelos de séries temporais fracamente identificados, também conhecidos por VAR estrutural (do inglês *structural vector autoregressive*). Sims (1996, as cited in Simkins, 1999), refere que:

Esta maneira de trabalhar, em oposição aos modelos de equações simultâneas, começa com uma modelagem, cuidadosa, multivariada de dados de séries temporais, desenvolvendo evidências sobre regularidades proeminentes. As restrições baseadas em um raciocínio económico real, somente são impostas se necessário para interpretar os dados e sempre com vista a evitar qualquer distorção no ajustamento do modelo. (p. 21)

O foco de Sims permanece em deixar os “dados falarem por si”, mas impondo restrições da teoria económica suficientes para tornar as interpretações esboçadas, nos modelos, teoricamente honestas. Esses modelos estruturais constituíam então, modelos promissores nas investigações económicas que segundo Sims (1996 as cited in Simkins, 1999), superavam as críticas dos modelos estruturais:

... Esta literatura visa minimizar a confiança em modelagens *ad hoc* convencionais de ambos modelos de equações simultâneas tradicionais e os novos, estocásticos, modelos de equilíbrio geral, por forma a focar claramente o ponto central da identificação da distinção entre variações geradas por acções políticas premeditadas e variações advindas de distúrbios além do processo político. (p. 21)

A metodologia de Sims fundamentava-se em inferência probabilística, impondo somente restrições da teoria económica suficientes para identificar estatisticamente os modelos e levar a cabo análises sobre a efectividade da política. Em muitos aspectos isto mostra um movimento em direcção a concordância no debate sobre *medição versus teoria*, onde a medição e a teoria combinam para alcançar o entendimento da economia.

Uma outra situação que se tornou alvo de crítica no VAR tem haver com a estabilidade do modelo, ou seja, a verificação da estacionaridade conjunta das variáveis. Pois em sua ausência tornar-se-ia necessário transformar os dados de modo a alcançá-la, dado que a maioria das séries económicas e financeiras caracterizam-se pela não estacionaridade em média, isso dificulta a tarefa do analista em prever o seu comportamento futuro.

“Regressões de series temporais que incluem variáveis integradas podem comportar-se muito distintamente que os modelos de regressão *standards*.” (Watson, 1994, p. 2848). Segundo Watson, coeficientes estimados em modelos VAR com componentes integradas (não estacionárias) podem comportar-se diferentemente das estimativas de um VAR estacionário em covariâncias, em particular, alguns coeficientes estimados podem desviar-se das suas distribuições assintóticas, enquanto os outros permanecem *standards*.

Segundo Greene (1997), com o progresso da literatura sobre os modelos econométricos, desenvolveu-se uma técnica bastante poderosa para lidar com tais situações, chamada cointegração, introduzida por Engle and Granger (1987) e posteriormente generalizada para o caso multivariado por Johansen (1988) e Stock and Watson (1988), para analisar as relações de equilíbrio entre séries temporais económicas com comportamento semelhante ao longo do tempo (cointegração).

A ideia por detrás deste método, é o facto de a combinação linear de todas variáveis em um modelo de regressão não poder ser estacionária, mas podendo haver algumas variáveis no modelo cuja combinação o seja. Portanto, além da diferenciação dos dados ¹⁴, “cointegration analysis is designed to find linear combinations of variables that also remove unit roots.” (Hendry & Juselius, 2000, p. 2).

Se duas ou mais variáveis económicas são integradas da mesma ordem, e *flutuam* partilhando uma tendência comum, com a “mesma intensidade”, então em tal caso, podemos distinguir entre relações de longo prazo entre elas e a dinâmica de curto prazo, isto é, a relação entre desvios de ambas variáveis das suas tendências de longo prazo (Greene, 1997). Se este for o caso, diferenciação dos dados seria contraproducente, dado que obscureceria a relação de longo prazo entre as variáveis. Assim, estudos de cointegração e uma técnica relacionada, correcção do erro (que mostra a dinâmica de curto prazo do modelo), interessam-se com métodos de estimação que preservam a informação sobre ambas formas de covariação.

2.3.2 Estacionaridade

Para que as conclusões advindas de um processo estocástico tenham confiabilidade, as séries temporais utilizadas devem apresentar a característica da estacionaridade. Segundo Mills (1993,

¹⁴ Os conceitos de ordem de integração e diferenciação serão tratados detalhadamente na secção 2.3.2.

citado por Resende, 2006), a estacionaridade representa um estado particular de “equilíbrio estatístico”, pois qualquer mudança de origem de tempo, mantido constante o tamanho do intervalo, não altera as propriedades da série.

Um processo estocástico é dito estritamente estacionário se suas propriedades não são afectadas por uma mudança de origem do tempo. Em outras palavras, a distribuição de probabilidade conjunta em qualquer conjunto de tempos t_1, t_2, \dots, t_m deve ser a mesma dos tempos $t_{1+k}, t_{2+k}, \dots, t_{m+k}$, em que k é uma mudança arbitrária no tempo. (p. 10).

Na maior parte das situações práticas, não é necessário que o processo estocástico seja estritamente estacionário. Segundo Gujarati (2006), torna-se quase sempre suficiente garantir apenas que o processo seja *fracamente estacionário*, ou *estacionário em covariâncias*, ou *estacionário de segunda ordem*. A *estacionaridade fraca* pode ser definida como:

um processo estocástico cuja média e variância são constantes ao longo do tempo, e quando o valor da covariância entre dois períodos de tempo depende apenas da distância, do intervalo ou da defasagem entre dois períodos de tempo, e não do próprio tempo em que a covariância é calculada. (Gujarati, 2006, p. 639).

Portanto, se uma determinada série temporal Y_t for estacionária, então terá as seguintes propriedades:

$$\text{Média:} \quad E(Y_t) = \mu \quad (2.1)$$

$$\text{Variância:} \quad V(Y_t) = E(Y_t - \mu)^2 = \sigma^2 \quad (2.2)$$

$$\text{Covariância:} \quad \gamma_k = E[(Y_t - \mu)(Y_{t+k} - \mu)] \quad (2.3)$$

Onde γ_k , a covariância (ou autocovariância) na defasagem k , entre os valores de Y_t e Y_{t+k} , isto é, entre dois valores de Y separados por k períodos. Se $k = 0$, obtemos γ_0 que é simplesmente a variância de Y (σ^2); se $k = 1$, γ_1 é a covariância entre dois valores adjacentes de Y .

“ Em suma, se uma série temporal é estacionária, sua média, variância e autocovariância (em diferentes defasagens) permanecem as mesmas, não importa qual seja o ponto em que as medimos: isto é, elas não variam com o tempo.” (Gujarati, 2006, p. 639). Em contrapartida, se a série não for estacionária, terá uma média ou uma variância que varia com o tempo, ou ambas.

Segundo Gujarati (2006), importa referir um tipo especial de processo estocástico, um processo puramente aleatório, dito ruído branco, ou seja, um processo cuja média é zero e a variância é constante, e ao mesmo tempo não autocorrelacionado. Onde a autocorrelação é dada por:

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad (2.4)$$

O motivo pelo qual se exige que uma série temporal seja estacionária, é que: (1) se ela for não estacionária só poderemos estudar seu comportamento para o período considerado, cada conjunto de dados da série constituirá um acontecimento específico e, portanto, não seria possível generalizá-lo para outros períodos de tempo; (2) a utilização de séries não estacionárias pode gerar o problema de regressão espúria caso a ordem de integração das mesmas seja diferente.

Assim, ao regredir duas séries temporais, é relativamente comum obter-se um coeficiente de determinação (R^2) bastante alto, sem que necessariamente haja uma relação significativa entre as duas variáveis em questão¹⁵. Tal exemplifica o problema da regressão espúria, situação na qual as séries temporais envolvidas exibem tendências similares. Nesses casos, o R^2 elevado se deve à presença de tendências, e não à verdadeira relação entre as variáveis¹⁶. Pelo facto de os resíduos da regressão espúria serem não estacionários os testes usuais (t , F , R^2) não são válidos. Por isso, é importante verificar previamente se a relação entre as variáveis económicas é verdadeira ou espúria.

Segundo Camuri (2005), a presença de tendência em séries temporais reporta ao facto da média variar com o tempo. Tal média, nesses casos, pode ser descrita por um polinómio de ordem d ($d = 1, 2, \dots$) qualquer. Uma tentativa de modelar esse tipo de série pode se dar via decomposição da tendência de modo a separar os componentes estocásticos estacionários do polinómio representativo da tendência, isso pode ser alcançado através da transformação da série em primeiras diferenças.

¹⁵ O coeficiente de determinação (R^2) mostra o percentual da variável dependente que é explicado pelas variáveis independentes. É calculado pelo quociente entre a soma do quadrado dos resíduos explicado e a soma do quadrado dos resíduos total (SQE/SQT). Para mais esclarecimentos ver Gujarati (2006).

¹⁶ A regra prática para detectar a possível existência de regressões espúrias se dá mediante a comparação da estatística D de Durbin Watson e o valor do R^2 . Caso o R^2 seja maior que a estatística D há a possibilidade de regressões espúrias. A consequência de regressões espúrias é o facto de tomar inválidas as estatísticas de teste t , F e R^2 (Gujarati, 2006).

Similarmente, caso a variável não constante no tempo seja a variância, Box and Cox (1964, citados por Camuri, 2005) sugerem aplicar uma transformação logarítmica de modo a linearizar a mesma. Esse procedimento de estabilização da variância da série ficou conhecido na literatura de séries temporais, segundo o autor, como *Transformação de Box and Cox*.

Por conseguinte, para entender a exposição dos procedimentos necessários para detectar a presença da estacionaridade numa série temporal é preciso conhecer, inicialmente, a definição de um processo gerador de dados (DGP) autoregressivo. Um processo autoregressivo de primeira ordem, AR(1), pode ser representado por:

$$Y_t = \alpha + \beta Y_{t-1} + \varepsilon_t \quad (2.5)$$

Onde: $t = \dots -1, 0, 1 \dots$

α — é um termo constante;

β — parâmetro que relaciona o valores presentes da série, aos seus valores passados;

ε_t — é um termo de erro estocástico, considerado ruído branco, identicamente e independentemente distribuído, isto é, $\varepsilon_t \sim IID(0, \sigma^2)$.

Para entender a forma pela qual (2.5) será estacionária, Johnston e Dinardo (2001), sugerem efectuar substituições sucessivas na equação (2.5) do tipo:

$$\begin{aligned} Y_1 &= \alpha + \beta Y_0 + \varepsilon_1 \\ Y_2 &= \alpha + \beta Y_1 + \varepsilon_2 = \alpha(1 + \beta) + \beta^2 Y_0 + (\varepsilon_2 + \beta \varepsilon_1) \end{aligned}$$

Procedendo da mesma maneira, obtém-se uma equação geral, considerando que o processo tenha começado a muito tempo atrás e não em Y_0 teremos:

$$Y_t = \alpha(1 + \beta + \beta^2 + \dots) + (\varepsilon_t + \beta \varepsilon_{t-1} + \beta^2 \varepsilon_{t-2} + \dots) \quad (2.6)$$

Note-se que as propriedades estocásticas de Y são determinadas pelas propriedades da sucessão ε . Aplicando valores esperados a ambos os membros da equação (2.6), obtemos

$$E(Y_t) = \alpha(1 + \beta + \beta^2 + \dots)$$

O valor esperado apenas será absoluto se a série geométrica no segundo membro tiver limite. Assim a condição necessária e suficiente para a existência desse limite será

$$|\beta| < 1 \quad (2.7)$$

E o valor esperado é, então

$$E(Y_t) = \mu = \frac{\alpha}{1 - \beta} \quad (2.8)$$

Que é uma média não condicionada para todos pontos do tempo. A variância não condicionada é:

$$\text{Var}(Y_t) = \frac{\sigma^2}{1 - \beta^2} \quad (2.9)$$

A autocovariância de defasagem k , seria dada por

$$\gamma_k = \beta^k \frac{\sigma^2}{1 - \beta^2} \quad (2.10)$$

Assim, estabelecida a condição em (2.7), a média, variância e covariância da sucessão Y são constantes e independentes do tempo. Então diz-se que a sucessão Y é *fracamente estacionária* ou tem *covariância estacionária* (Johnston e Dinardo, 2001).

Segundo Resende (2006), a série temporal é dita estacionária uma vez que os choques externos são dissipados ao longo do tempo, por choque externo ou exógeno, entende-se qualquer perturbação ou inovação que incida sobre a variável de modo a, mesmo que por um curto espaço de tempo, alterar sua média e variância. O modo como um choque exógeno é repassado aos momentos futuros constitui-se no determinante acerca da estacionaridade da série e é expresso pelo termo $|\beta|$.

De acordo com Camuri (2005) e Resende (2006), caso $|\beta| < 1$, um choque exógeno no momento t altera a média e variância da série temporal, mas o efeito desse choque será cada vez menor a medida que o tempo passa. Além disso, quando $|\beta| < 1$ a covariância entre as observações de momentos distintos será função somente da defasagem de tempo, independentemente do momento específico considerado. Tal comportamento conduzirá a média e a variância de volta a seus valores originais e, portanto, a série é dita estacionária.

Nessas circunstâncias a inferência estatística se enquadra no facto da série analisada preservar, no longo prazo, suas características originais. Em outras palavras, caso $|\beta| < 1$ os parâmetros estimados no processo AR(1) correspondem aos parâmetros do processo estocástico da variável estudada e no caso em que $|\beta|$ assume valores menores ou iguais a um, a média e variância da série temporal deixarão de ser constantes, e a covariância entre as observações será sensível às mudanças no tempo específico considerado (Camuri, 2005). Isso caracteriza séries temporais não estacionárias.

Quando $|\beta| > 1$, um choque exógeno tem impacto crescente sobre a variável à medida que o tempo passa, ou seja, a variável é dita não estacionária e, portanto, as inferências estatísticas deixam de ser confiáveis. Por fim, caso $|\beta|$ seja igual a um num processo AR(1) a série também é não estocástica uma vez que o efeito de um choque exógeno é *perpetuado* ao longo do tempo. Apesar disso, é possível torná-la estacionária através do cálculo de sua primeira diferença, ou seja:

$$\Delta Y_t = Y_t - Y_{t-1} \quad (2.11)$$

Substituindo (2.11) em (2.5):

$$\Delta Y_t = \alpha + \delta Y_{t-1} + \varepsilon_t \quad (2.12)$$

Onde, $\delta = \beta - 1$ e como $\beta = 1$, $\delta = 0$ (caso em que [2.5] torna-se um *passeio aleatório com deslocamento*, este último, designado pelo parâmetro α), então teremos

$$\Delta Y_t = \alpha + \varepsilon_t \quad (2.13)$$

Como ε_t é um processo estocástico de ruído branco, segundo Gujarati (2006), isso significa que Y_t exibirá uma tendência positiva ($\alpha > 0$) ou negativa ($\alpha < 0$). Essa tendência é também denominada *tendência estocástica*. Sendo assim, (2.13) é um processo estacionário em diferenças porque a não estacionaridade de Y_t pode ser eliminada tomando-se as primeiras diferenças da série temporal.

Portanto, os elementos aqui expostos caracterizam o processo AR(1) como integrado de ordem um, $I(1)$, ou alternativamente, como um passeio aleatório¹⁷. A série temporal descrita por um processo AR(1) com $\beta=1$ é chamada de série temporal com raiz unitária¹⁸. Dado que a eficiência das estimações dos modelos que utilizam séries temporais depende da estacionaridade da série, passamos subsequentemente para o tratamento dos testes capazes de detectar se o processo estocástico é ou não estacionário, testes estes, conhecidos como testes de raiz unitária.

2.3.2.1 Testes de Raiz unitária

Os testes, aqui usados, para detectar a presença de raiz unitária e, por conseguinte, a ordem de integração das séries, são o Dickey-Fuller (DF), o Dickey-Fuller aumentado (ADF) e o Phillips-Perron (PP). O teste DF consiste na estimação, através da regressão por OLS, do coeficiente δ da equação (2.12) e no cálculo da estatística do teste:

$$t = \frac{\hat{\delta}}{S.E.(\hat{\delta})} \quad (2.14)$$

Onde, $\hat{\delta}$ é a estimativa de δ , e $S.E.(\hat{\delta})$ é o erro padrão da estimativa desse coeficiente. E testa-se:

$H_0: \delta = 0$, (presença de raiz unitária, série não estacionária)

$H_1: \delta < 0$, (ausência de raiz unitária, série estacionária)¹⁹

Observe-se que o teste acima representa a estimação do coeficiente β da equação (2.5), correspondendo às seguintes hipóteses:

$H_0: \beta = 0$

$H_1: \beta < 0$

Segundo Gujarati (2006), o procedimento concreto para aplicar um teste de Dickey-Fuller envolve várias decisões. Assim, ao discutir a natureza do processo de raiz unitária, pode-se notar que um

¹⁷ A ordem de integração de determinado processo está relacionada ao número de diferenças que se faz necessário submeter determinada variável até que esta se apresente estacionária. Assim, uma série será estacionária de ordem d ($d = 0, 1, 2, 3, \dots$) quando diferenciada d vezes.

¹⁸ Toda série com raiz unitária precisa ser diferenciada d vezes até que se alcance sua estacionaridade.

¹⁹ Exclui-se a possibilidade de que $\delta > 0$, pois, nesse caso, $\beta > 0$ a série seria explosiva.

processo de passeio aleatório pode ou não ter deslocamento, pode ter deslocamento ou pode ter tendências tanto determinísticas quanto estocásticas²⁰. Para isso, várias possibilidades podem ser tomadas em conta no teste de Dickey-Fuller:

$$Y_t \text{ é um passeio aleatório: } \Delta Y_t = \delta Y_{t-1} + \varepsilon_t \quad (2.15)$$

$$Y_t \text{ é um passeio aleatório com deslocamento: } \Delta Y_t = \alpha_1 + \delta Y_{t-1} + \varepsilon_t \quad (2.16)$$

Y_t é um passeio aleatório com deslocamento em torno de uma tendência estocástica:

$$\Delta Y_t = \alpha_1 + \alpha_2 t + \delta Y_{t-1} + \varepsilon_t \quad (2.17)$$

Onde t é o tempo ou variável de tendência. A questão que suscita alguma dúvida é se no teste dessas hipóteses, pode ainda ser usada a estatística t - *student* habitual. Infelizmente, sob a hipótese nula de que $\delta = 0$, o valor t do coeficiente estimado de Y_{t-1} não segue a distribuição t , nem mesmo em amostras grandes, isto é, não tem uma distribuição normal assintótica, porque a estacionaridade é necessária para a dedução das distribuições *standard*. (Johnston & Dinardo, 2001; Gujarati, 2006).

Dickey e Fuller mostraram que, sob a hipótese nula de que $\delta = 0$, o valor t estimado do coeficiente de Y_{t-1} , segue a estatística τ (*tau*). Esses autores calcularam os valores críticos da estatística *tau* com base em simulações de Monte Carlo. Devido a limitação das tabelas por eles preparadas, James Mackinnon elaborou tabelas mais extensas. É extremamente importante notar que os valores críticos do teste τ usado para verificar a hipótese nula, diferem para cada uma das especificações anteriores do teste de Dickey-Fuller.

O defeito do teste DF está no facto de supor que os resíduos são estritamente *ruído branco*, satisfazendo as três propriedades da média zero, variância constante e não autocorrelação, além de serem independentes. Quando o termo ε_t é autocorrelacionado inclui-se na especificação anterior termos de diferença desfasados de modo que o termo de erro torne-se serialmente independente (Gujarati, 2006). Charemza and Deadman (1992, citados por Resende, 2006), afirmam que:

²⁰ Um processo com tendência determinística, ou *estacionário em tendência* é um processo no qual somente a variância é constante, mas a média varia com o tempo. Por conseguinte, ao subtrair-se de Y_t a média de Y_t , a série resultante será estacionária, daí o nome estacionária *pós-remoção da tendência*. Para mais detalhes, ver Gujarati (2006).

A weakness of the original Dickey-Fuller test is that it does not take account of possible autocorrelation in the error process ε_t . If ε_t is autocorrelated (that is, it is not white noise) then the ordinary least squares estimates of equation $[\Delta Y_t = \delta Y_{t-1} + \varepsilon_t]$, and of its variants described above, are not efficient. A simple solution... is to use lagged left-hand side variables as additional explanatory variables to approximate the autocorrelation. (p. 135).

A inclusão dos termos desfasados dá origem aos seguintes modelos que constituem o teste aumentado de Dickey-Fuller (ADF):

$$\Delta Y_t = \delta Y_{t-1} + \sum_{i=2}^m \alpha_i \Delta Y_{t-i+1} + \varepsilon_t \quad (2.18)$$

$$\Delta Y_t = \alpha_1 + \delta Y_{t-1} + \sum_{i=2}^m \alpha_i \Delta Y_{t-i+1} + \varepsilon_t \quad (2.19)$$

$$\Delta Y_t = \alpha_1 + \alpha_2 t + \delta Y_{t-1} + \sum_{i=2}^m \alpha_i \Delta Y_{t-i+1} + \varepsilon_t \quad (2.20)$$

Onde, m = número de desfasagens necessárias para tornar o termo de erro serialmente independente. A execução do teste ADF segue a mesma lógica do teste DF e os valores críticos para o teste ADF são os mesmos do teste DF, dado que ambas estatísticas de teste possuem a mesma distribuição assintótica (Gujarati, 2006).

O número de desfasagens utilizadas, segundo Enders (1995, citado por Resende, 2006), parte de um número grande de desfasagens n^* e avalia-se o modelo pelo teste t e/ou teste F comuns. Caso a estatística t na desfasagem n^* seja insignificante para um dado valor crítico, se estima novamente a regressão usando n^*-1 desfasagens. Esse processo continua até que a última desfasagem se torne significativamente diferente de zero.

Outra forma de testar a presença de estacionaridade, é pelo teste de Phillips-Perron (PP). Enquanto o teste de ADF ajusta o teste DF para levar em conta uma possível correlação serial nos termos de diferenças desfasadas do regressando, o teste PP usa métodos estatísticos não paramétricos para levar em conta a correlação serial nos termos de erro sem somar termos de diferenças desfasados (Gujarati, 2006). O método do teste PP é o DF não aumentado, que modifica o valor da estatística t

calculada de δ para que a autocorrelação serial não afecte a distribuição assintótica da estatística do teste ²¹:

$$Z = \frac{\hat{\delta}}{S.E(\hat{\delta})} \times \left(\frac{\varphi_0}{\eta_0} \right)^{1/2} - \frac{T \times (\varphi_0 - \eta_0) \times S.E.(\hat{\delta})}{2\eta_0^{1/2} \times s} \quad (2.21)$$

Onde, T corresponde ao tamanho da amostra em análise, s , o erro padrão da regressão. φ_0 é uma estimativa consistente da variância do erro ε_t , calculado como $(T-k) \times s^2 / T$, onde K mostra o número de regressores. Por outro lado, η_0 é um estimador espectral dos resíduos em frequência zero ²². Embora a estatística do teste PP seja Z , sua distribuição assintótica coincide com a de ADF.

Ambos os testes (PP e ADF) são baseados em teoria assintótica e, portanto, a escolha do teste mais adequado dependerá de como as estatísticas aproximam a distribuição para amostra finita da estatística relevante. Quando o termo de resíduo tiver um componente de média móvel positivo o poder do teste PP é maior que o poder do teste ADF, porém se o termo de média móvel for negativo o teste ADF revela-se preferível (Resende, 2006). A sugestão de Enders (1995, citado por Resende, 2006) quanto a qual teste utilizar é o uso de ambos para verificar se os mesmos se reforçam e, eventualmente, utilizar a teoria económica para a sugestão de qual o teste mais adequado.

É preciso prestar atenção na análise de séries temporais, quando há a suspeita de quebras estruturais em um ou mais pontos da série. Pois, segundo Perron (1989, citado por Johnston e Dinardo, 2001), as mudanças estruturais invalidam os testes convencionais de raiz unitária. A quebra estrutural tem por característica uma mudança permanente no comportamento da série e pode corresponder a uma mudança na inclinação da tendência, no intercepto ou em ambos (Camuri, 2005).

Enders (1995, citado por Camuri, 2005), sugere que na presença de uma quebra estrutural, dois procedimentos podem ser adoptados. O primeiro consiste em testar a presença de raiz unitária em cada sub-período da série. Embora isso acabe por gerar um sério problema de perda de graus de

²¹ A peculiaridade do teste PP está no facto de a estatística calculada estar baseada em Z , que é uma correcção da estatística t incluindo as possibilidades de resíduos autocorrelacionados e/ou heteroscedásticos. Mais detalhes sobre o teste Z , ver Mills (1993, citado por Camuri, 2005).

²² Consideram-se basicamente duas classes de estimadores espectrais em frequência zero, um baseado na estimação do núcleo da soma de covariâncias, e outro, um estimador baseado na densidade espectral autorregressiva. Mais detalhes ver Quantitative Micro Software [QMS] (2002).

liberdade à medida que se subdivide a amostra. Alternativamente, sugere o uso do teste de Perron para a quebra. Tal método reporta ao uso do teste ADF modificado pela inclusão de uma *dummy* representativa da quebra estrutural. Caso a *dummy* mostre-se significativa e dê mais poder ao teste no sentido de rejeitar a hipótese nula, o teste ADF tradicional deve ser deixado de lado e passa-se a considerar a série estacionária.

A presença de mudanças estruturais reporta o facto de que mudanças nas políticas macroeconómicas podem ter impactos permanentes sobre as variáveis. Em função disso muito se tem discutido acerca da presença ou não de quebras estruturais nas séries uma vez que a estabilidade do modelo é essencial para a avaliação do impacto dessas mudanças. Além do teste de Perron, existe um teste de raiz unitária recentemente desenvolvido denominado teste da presença de quebra estrutural endógena de Lee e Strazicich. Todavia, no presente estudo não se adopta a metodologia desses testes.

2.4 Aspectos teóricos sobre os modelos VAR

Os vectores autoregressivos são normalmente usados para fazer previsões de sistemas de séries temporais interrelacionadas e para analisar a dinâmica dos distúrbios aleatórios no sistema de variáveis (QMS, 2002).

Um VAR é um sistema de k – equações, modelo linear de k – variáveis no qual cada uma é explicada pelos seus valores desfasados, e pelos valores correntes e desfasados das restantes variáveis (Stock & Watson, 2001).

Esse simples modelo traz consigo um modo sistemático de *capturar* efeitos dinâmicos em séries temporais múltiplas, as ferramentas estatísticas inerentes ao VAR são fáceis de usar e interpretar. Como Sims (1980) e outros argumentaram em uma série de documentos influentes, o VAR concretizou a promessa de prover uma aproximação coerente e credível para a descrição dos dados, previsão e análises políticas. (Stock & Watson, 2001, p. 2).

Alguns autores como Greene (1997), Johnston e Dinardo (2001) e Gujarati (2006) não fazem qualquer referência a possível inclusão de valores correntes das variáveis no sistema de equações.

Stock and Watson (2001) referem que os modelos VAR podem ser analisados de três formas: *forma reduzida*, recursiva e estrutural.

- i. *Forma reduzida* — o modelo VAR na forma reduzida expressa cada variável como uma função linear dos seus valores passados e dos valores passados das restantes variáveis no modelo, com um termo de erro não correlacionado. Cada equação é estimada por uma regressão OLS. Segundo Stock and Watson (2001), “the error term are the *surprise* movements in the variables, after taking its past values into account.” (p. 3). Se as diferentes variáveis estiverem correlacionadas entre si, como habitualmente acontece, então os termos de erro na forma reduzida do modelo podem estar contemporaneamente correlacionados.
- ii. *Forma recursiva* — esse modelo produz em cada equação de regressão, um termo de erro não correlacionado com os termos de erro das outras equações. Isto é feito prudentemente incluindo alguns valores contemporâneos como regressores. Portanto, é preciso prestar bastante atenção no sentido de determinar sucintamente quais variáveis correntes afectam as outras. A estimação de cada equação por OLS produz, evidentemente, resíduos não correlacionados contemporaneamente, mas os resultados dependem da ordenação desses valores correntes no modelo, pois mudando a ordem, mudam também as estimações do modelo. O número total de possíveis modelos VAR recursivos é $k!$
- iii. *Forma estrutural* — os modelos estruturais utilizam a teoria económica para classificar as relações contemporâneas entre as variáveis. Esses modelos exigem *suposições de identificação* que permitem que as correlações sejam interpretadas como causalidade. As suposições de identificação podem envolver todo VAR, para que todas relações de causalidade no modelo sejam mostradas. Isto faz com que sejam produzidas variáveis instrumentais que permitam a estimação das relações contemporâneas através dos métodos de regressão com variáveis instrumentais.

A justificação para a aplicação dos três modelos, segundo Stock and Watson, é que: (1) a forma reduzida e a recursiva resumem os movimentos comuns entre as variáveis; (2) a forma reduzida é usada para prever os valores das variáveis, e sua performance é avaliada comparando-o com alguns modelos alternativos referenciais; (3) o VAR estrutural é usado, alternativamente, para avaliar as

políticas macroeconómicas. Para os propósitos desse estudo importa apenas o desenvolvimento específico dos modelos na forma reduzida e recursiva.

2.4.1 Derivação do modelo

Considere-se inicialmente um vector de k variáveis, Y_t , observadas no tempo. Visto que em sistemas macroeconómicos muitas variáveis são, provavelmente, interdependentes (Sims, 1980, citado por Fedderke, 2003), então torna-se necessário analisar as covariâncias entre as variáveis no tempo t , assim como as covariâncias nos tempos t e $t - h$ ($h = 1, 2, \dots$). Segundo Juselius (2003), as covariâncias contêm informações das relações estática e dinâmica entre as variáveis que são descobertas através dos métodos econométricos.

Seja dado o vector

$$Y_t = \begin{bmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{k,t} \end{bmatrix}, t = 1 \dots T$$

Sem que se estabeleça qualquer suposição, podemos definir o vector de médias e a matriz de covariâncias como segue:

$$E(Y_t) = \begin{bmatrix} \mu_{1,t} \\ \mu_{2,t} \\ \vdots \\ \mu_{k,t} \end{bmatrix} = \mu_t, \quad Cov(Y_t, Y_{t-h}) = \begin{bmatrix} \sigma_{11,h} & \sigma_{12,h} & \dots & \sigma_{1k,h} \\ \sigma_{21,h} & \sigma_{22,h} & \dots & \sigma_{2k,h} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{k1,h} & \sigma_{k2,h} & \dots & \sigma_{kk} \end{bmatrix} = \Sigma_{t,h}$$

Posto isto, “assume that the same distribution applies for all Y_t and that it is approximately normal, i.e. $Y_t \sim N(\mu_t; \Sigma_t)$. Under the normality assumption the first two moments around the mean (central moments) are sufficient to describe the variation in the data.” (Juselius, 2003, p.41). Por outro lado, “since there are far more parameters than observations available for estimation, it has no meaning from a practical point of view. Therefore, we have to make simplifying assumptions to reduce the number of parameters.” (Juselius, 2003, p. 42).

As suposições tipicamente feitas aos modelos empíricos, baseiam-se em assumir que:

$$\rightarrow \Sigma_{t,h} = \Sigma_h, \forall t \in T, h = \dots, -1, 0, 1, \dots$$

$$\rightarrow \mu_t = \mu, \forall t \in T$$

Essas duas suposições são necessárias para assegurar a constância dos parâmetros do modelo VAR.

Em Johansen (1995) e Juselius (2003), demonstra-se que a função de densidade de probabilidade conjunta de máxima verosimilhança $P(\mathbf{Y}; \theta)$ pode ser decomposta em T probabilidades condicionais $P(Y_t | Y_{t-1}, \dots, Y_1; \mathbf{Y}_0, \theta)$, e que o processo condicional $(Y_t | Y_{t-1}, \dots, Y_1; \mathbf{Y}_0)$ contém uma reparametrização que corresponde ao modelo vectorial autoregressivo. Assim, em uma análise empírica na qual é dada a matriz $\mathbf{Y} = [Y_1, \dots, Y_T]'$, onde cada Y_t é um vector $k \times 1$ de variáveis, sob as suposições de que os dados observados \mathbf{Y} constituem uma realização de um processo estocástico, é possível expressar a probabilidade conjunta de \mathbf{Y} dado valor inicial \mathbf{Y}_0 e o valor do parâmetro θ descrevendo o processo estocástico:

$$P(\mathbf{Y} | \mathbf{Y}_0; \theta) = P(Y_1, \dots, Y_T | \mathbf{Y}_0; \theta)$$

Para uma dada função de probabilidade, estimativas máximo verosímeis podem ser encontradas maximizando a função de verosimilhança. Impondo a restrição de que o DGP é normal multivariado, Juselius (2003), faz uma decomposição do processo conjunto em um processo condicional e um processo marginal. De modo simplificado, o modelo VAR pode ser expresso como ²³:

$$Y_t = \pi + \Pi_1 Y_{t-1} + \Pi_2 Y_{t-2} + \dots + \Pi_p Y_{t-p} + \varepsilon_t \quad (2.22)$$

Onde, π é um vector $k \times 1$ de constantes. Π_i , $i = 1, \dots, p$ é uma matriz de dimensão $k \times k$ de parâmetros de todas variáveis incluídas no modelo, desfasadas até a p -ésima ordem. ε_t é um vector $k \times 1$ de resíduos, o qual assume-se normal, identicamente e independentemente distribuído com média zero e uma matriz de covariância definida positiva, ou seja, $\varepsilon_t \sim IIDN(0; \Omega)$.

Se as suposições de normalidade multivariada com parâmetros (μ, Σ) forem correctas, então (2.22) terá as seguintes características:

- linear nos parâmetros
- parâmetros constantes
- erros ε_t distribuídos normalmente

²³ Onde os parâmetros do modelo, Π , são matrizes condicionadas à decomposição sobre a função de probabilidade conjunta decomposta em um processo condicionado e outro marginal. Mais detalhes, ver Juselius (2003).

↘ independência entre ε_t e ε_{t-h} .

O facto de os parâmetros do modelo serem derivados de processos condicionais e marginais, faz com que a constância desses mesmos parâmetros esteja sujeita a constância das matrizes de covariância condicionais, isto é, se por algum motivo essas matrizes variarem os seus valores, os parâmetros do modelo provavelmente variarão em resposta a essas mudanças.

Uma advertência feita por Hendry and Juselius (2000) acerca da suposição da normalidade multivariada, é que em aplicações económicas ela é raramente satisfeita, por conseguinte, “ Este é um sério problema, dado que a derivação do modelo de um DGP geral sustenta-se na normalidade multivariada, e inferências são somente válidas tendo em conta que as suposições subjacentes ao modelo são correctas.” (p. 7).

A questão que então se coloca é como modificaríamos o modelo VAR *standard* na prática, preservando sua atractividade e minuciosidade na descrição das características básicas dos dados, e ao mesmo tempo alcançar inferências válidas.

Estudos simulados demonstraram que a inferência estatística é sensível à validade de algumas suposições, tais como, a não constância dos parâmetros, resíduos serialmente correlacionados e assimetria, enquanto moderadamente robusta a outras, tais como curtose e heteroscedasticidade dos resíduos. Assim, parece aconselhável assegurar que as primeiras três sejam válidas. (Hendry and Juselius, 2000, p.7).

Uma forma alternativa exposta por Hendry and Juselius (2000) e Juselius (2003) para *superar* esses problemas é: (1) adicionando *dummies* para considerar eventos significativos relacionados a mudanças políticas ou institucionais; (2) condicionando o modelo a variáveis fracamente exógenas; (3) inspeccionando as medições das variáveis escolhidas; (4) mudando o período amostral para evitar mudanças de regime ou dividindo a amostra em períodos mais homogéneos.

2.4.2 Propriedades estocásticas do modelo VAR

O modelo em (2.22) foi até aqui descrito como estacionário, isto é, sem considerar raízes unitárias na combinação entre as variáveis²⁴. As propriedades dinâmicas de estabilidade do processo podem ser investigadas calculando as raízes do processo VAR (Hendry & Juselius 2000; Juselius, 2003). Segundo os autores, é conveniente formular o modelo como um polinómio no operador *Lag* (L), onde $L'Y_t = Y_{t-1}$:

$$\begin{aligned} (I_p - \Pi_1 L - \dots - \Pi_p L^p) Y_t &= \Phi A_t + \varepsilon_t, \\ \Pi(L) Y_t &= \Phi A_t + \varepsilon_t, \end{aligned} \quad (2.23)$$

Onde, A_t é um vector representativo de todas componentes determinísticas incluídas no modelo, tais como, constantes, *dummies*, ou tendências temporais. Segundo Juselius (2003), o modelo VAR pode ser representado sob a forma autoregressiva, que é útil na especificação de hipóteses sobre o comportamento económico das variáveis, ou sob a forma de componentes de médias móveis, que é útil no exame das propriedades do processo. Segundo Juselius, quando o processo é estacionário, a última representação pode ser determinada directamente invertendo o modelo VAR de tal modo que Y_t seja expresso como uma função dos choques passados e presentes, ε_{t-j} , $j = 0, 1, \dots$, valores iniciais Y^0 , e possíveis componentes determinísticas A_t .

$$\begin{aligned} Y_t &= \Pi^{-1}(L)(\Phi A_t + \varepsilon_t) \\ &= |\Pi(L)|^{-1} \Pi^a(L)(\Phi A_t + \varepsilon_t) \\ &= (I_p + C_1 L + C_2 L^2 \dots)(\Phi A_t + \varepsilon_t) \end{aligned} \quad (2.24)$$

Onde, $|\Pi(L)| = \det[\Pi(L)]$ e $\Pi^a(L)$ é a matriz adjunta de $\Pi(L)$. Na presença de estacionaridade do processo VAR, $C_j = f(\Pi_1, \dots, \Pi_p)$ pode ser calculado de forma recursiva²⁵. Se o processo VAR for não estacionário, Y_t não é invertível e, por conseguinte, as matrizes C_j só poderão ser derivadas sob a suposição de existência de cointegração nas combinações lineares entre as variáveis.

²⁴ Uma tentativa de estimar o VAR com OLS sem considerar a existência de raízes unitárias, enquanto realmente existem, produziria inferências enviesadas. (Hendry & Juselius, 2000).

²⁵ Mais detalhes sobre a fórmula recursiva para o cálculo das matrizes C_j , ver Johansen (1995).

Para a determinação das raízes características do processo considere-se, inicialmente, o polinómio característico em z : $\Pi(z) = I_p - \Pi_1 z - \dots - \Pi_p z^p$, $[\Pi(z)]^{-1} = |\Pi(z)|^{-1} \Pi^\alpha(z)$

A informação obtida através do cálculo das raízes da equação característica, $|\Pi(z)| = 0$, resume as propriedades dinâmicas e a estabilidade do processo. Sem querer limitar a essência, com o intuito de mostrar claramente o mecanismo de obtenção das raízes características, desenvolver-se-á o processo para o caso em que $p = 2$, isto é, um VAR(2):

$$(I_2 - \Pi_1 L - \Pi_2 L^2)Y_t = \Phi A_t + \varepsilon_t$$

A função característica correspondente é:

$$\Pi(z) = I_2 - \begin{bmatrix} \pi_{1.11} & \pi_{1.12} \\ \pi_{1.21} & \pi_{1.22} \end{bmatrix} z - \begin{bmatrix} \pi_{2.11} & \pi_{2.12} \\ \pi_{2.21} & \pi_{2.22} \end{bmatrix} z^2 = \begin{bmatrix} 1 - \pi_{1.11}z - \pi_{2.11}z^2 & -\pi_{1.12}z - \pi_{2.12}z^2 \\ -\pi_{1.21}z - \pi_{2.21}z^2 & 1 - \pi_{1.22}z - \pi_{2.22}z^2 \end{bmatrix}$$

e

$$\begin{aligned} |\Pi(z)| &= (1 - \pi_{1.11}z - \pi_{2.11}z^2)(1 - \pi_{1.22}z - \pi_{2.22}z^2) - (-\pi_{1.12}z - \pi_{2.12}z^2)(-\pi_{1.21}z - \pi_{2.21}z^2) \\ &= 1 - a_1z - a_2z^2 - a_3z^3 - a_4z^4 \\ &= (1 - \rho_1z)(1 - \rho_2z)(1 - \rho_3z)(1 - \rho_4z) \end{aligned}$$

Obtém-se um polinómio característico de quarta ordem, que produz quatro raízes (inversas), $z_1 = 1/\rho_1, \dots, z_4 = 1/\rho_4$. Para entender a utilidade das raízes características na descrição do comportamento dinâmico do processo, consideremos o modelo VAR(2):

$$Y_t = \frac{\Pi^\alpha(L)(\Phi A + \varepsilon_t)}{(1 - \rho_1L)(1 - \rho_2L)(1 - \rho_3L)(1 - \rho_4L)} = \left(\frac{\Pi^\alpha(L)}{(1 - \rho_2L)(1 - \rho_3L)(1 - \rho_4L)} \right) \left(\frac{\varepsilon_t + \Phi A}{1 - \rho_1L} \right)$$

Expandindo o componente da primeira raiz, $(1 - \rho_1L)^{-1}(\varepsilon_t + \Phi A) = (1 + \rho_1L + \dots)(\varepsilon_t + \Phi A)$. Assim, cada choque ε_t afectará de *forma dinâmica* ambos valores presentes e futuros das variáveis em Y_t . Note-se que isso também se sustenta para qualquer variável *dummy* incluída em A , (Juselius, 2003). A persistência desse efeito depende da magnitude de $|\rho_1|$.

Segundo Juselius (2003), uma raiz real ρ_j gera um comportamento de declínio exponencial, um par de raízes complexas $\rho_j = \rho_{real} \pm i\rho_{complexo}$ gera um comportamento cíclico de declínio exponencial. Se uma raiz se encontrar no círculo unitário ($\rho = 1$, $\rho = -1$), gerará um comportamento de não estacionaridade, isto é, uma tendência estocástica em Y_t . Se o módulo de uma raiz complexa for igual a unidade, estaremos perante um comportamento sazonal não estacionário.

“In econometrics, it is more usual to discuss stability in terms of the companion matrix of the system, obtained by stacking the variables such that a first-order system results.”²⁶ (Hendry & Juselius, 2000, p. 9). Por fim, tendo extraído as raízes características, usando qualquer um dos métodos, as conclusões tiradas são: (1) se as raízes da equação característica se encontrarem fora do círculo unitário (ou alternativamente se os valores próprios da matriz companheira se encontrarem dentro do círculo unitário) o processo $\{Y_t\}$ é estacionário; (2) se as raízes estiverem fora ou no círculo unitário (ou se os valores próprios estiverem dentro ou na linha do círculo unitário) $\{Y_t\}$ é não estacionário; (3) se qualquer uma das raízes estiver dentro do círculo unitário (ou se qualquer valor próprio estiver fora do círculo unitário) $\{Y_t\}$ é explosivo.

Depois de estimar e verificar a estabilidade do processo VAR, o passo seguinte seria o de previsão, mas para além disso, como observa Greene (1997), “in addition to forecasting, VAR’s have been used for two primary functions, testing Granger causality and studying impulse response functions.” (p. 816).

2.4.3 Causalidade de Granger

“Causality in the sense defined by Granger (1969) and Sims (1980) is inferred when lagged values of a variable, say $y_{1,t}$, have explanatory power in a regression of a variable $y_{2,t}$ on lagged values of $y_{2,t}$ and $y_{1,t}$. The VAR can be used to test the hypothesis.” (Greene, 1997, p. 816). Segundo Greene, essa ideia pode ser estendida num sistema de equações para verificar se uma determinada variável é fracamente exógena para o sistema. Se os valores desfasados duma dessa variável, $y_{i,t}$,

²⁶ Mais detalhes sobre esse assunto, ver Hendry and Juselius (2000) e Juselius (2003).

não têm nenhuma explicação para qualquer das variáveis do sistema, poder-se-ia ver $y_{i,t}$ como fracamente exógena para o sistema.

Todavia, Greene (1997), aponta uma complicação nos testes de causalidade pois:

O VAR é um artigo de fé. Não existe nenhuma teoria por detrás da formulação. Como tal, os testes de causalidade são atribuídos em modelos que podem, de facto, estar omitindo variáveis ou defasagens adicionais que deveriam estar presentes mas não estão. Primeiramente, o problema é que um resultado de efeitos causais pode igualmente resultar da omissão de variáveis que estejam correlacionadas com algumas ou todas variáveis no modelo. A segunda falha é mais sociável se se estiver disposto a assumir a normalidade dos distúrbios. (p. 817).

2.4.4 Funções de resposta a impulso e decomposição de variância

As características dinâmicas do processo VAR são melhor estudadas recorrendo à representação do processo em componentes de médias móveis. As ferramentas que permitem melhor interpretação das propriedades dinâmicas obtidas no VAR são, as funções de resposta-impulso e decomposição da variância dos erros de previsão (Resende, 2006).

Segundo Camuri (2005), uma particularidade do modelo (2.24) é que apresenta erros contemporaneamente correlacionados, e para que este seja utilizado na análise de políticas, é preciso transformá-lo de modo a deixar de apresentar erros relacionados. Uma transformação desse tipo permite que as equações que compõem o modelo possam ser utilizadas separadamente, pois a presença de erros relacionados indica que um choque numa determinada variável resultará em variações nas demais variáveis impossibilitando, por sua vez, isolar o efeito sobre o modelo, de uma determinada variável.

Camuri (2005) ressalta que a modelagem VAR parte da hipótese de ortogonalidade entre as inovações estruturais para equacionar o problema de identificação. O que significa que a matriz de covariâncias das inovações estruturais é diagonal, ou seja, restringem-se a zero as covariâncias entre

os diferentes tipos de inovação indicando, por sua vez, que as inovações estruturais são choques não directamente observáveis e que apresentam causas distintas.

Partindo de uma matriz simétrica definida positiva Ω , que representa a matriz de covariâncias, é possível encontrar uma matriz não singular P , de modo que $\Omega = P^{-1}(P^{-1})'$, de onde os novos resíduos não contemporaneamente correlacionados, η_t , obter-se-iam fazendo $\eta_t = P\varepsilon_t$. Esse processo de ortogonalização ilustra a **factorização de Choleski** da matriz definida positiva Ω . Que apresenta o produto duma matriz triangular inferior, P^{-1} , pela sua transposta que é triangular superior (Johnston & Dinardo, 2001).

Assim, cabe ressaltar a importância que cada variável assume perante as demais no modelo e analisar a sensibilidade das variáveis diante de inovações nelas ocorridas de forma isolada. As funções de resposta-impulso são utilizadas para analisar a sensibilidade das variáveis económicas a determinados choques, ou seja, ao se considerar um experimento conceptual de *perturbar* um sistema em equilíbrio. Supondo que $\varepsilon_{i,t}$ tenha igualado a zero por muito tempo que $y_{i,t}$ alcançou o equilíbrio. Ao injectar um choque para o sistema, mudando um dos $\varepsilon_{i,t}$, por um período, retornando subsequentemente a zero depois disso, $y_{i,t}$ mover-se-á, mas posteriormente, retornará ao seu equilíbrio. Portanto, o caminho pelo qual as variáveis retornam ao seu equilíbrio é chamado resposta a impulso do VAR.

Outro instrumento de análise do VAR é a decomposição da variância do erro de previsão cujo objectivo é explicitar a participação de cada variável do modelo na variância dos resíduos $\varepsilon_{i,t}$ das demais variáveis²⁷. De acordo com Enders (1995, citado por Camuri, 2005) essa técnica contribui ao identificar as relações de causalidade existentes dentro do modelo. Caso um choque exógeno no

²⁷ Recorde-se que para além do uso do processo de ortogonalização de Cholesky, o modelo contendo resíduos contemporaneamente não correlacionados obter-se-ia pela estimação de um modelo VAR recursivo, ou seja, incluindo, adequadamente, alguns valores correntes como regressores e estimando cada equação do modelo por OLS. Se as funções resposta-impulso e decomposição de variância assumem como hipótese a ortogonalização dos resíduos, o sentido da sua interpretação será sensível à ordem de entrada das variáveis no modelo, pois, uma estimação adequada do VAR recursivo ou a aplicação da fórmula de factorização de Cholesky, passam por uma sucinta escolha da ordem de entrada das variáveis no modelo.

resíduo de uma variável não explique nenhuma parcela da variância dos demais resíduos pode-se afirmar que essa variável é exógena no modelo.

De acordo com Runkle (1986), as funções resposta-impulso e a decomposição de variância são simplesmente funções não lineares dos parâmetros autoregressivos e sua matriz de covariância. Se os coeficientes individuais no VAR irrestrito forem insignificantes, seus erros padrão implicam largos e crescentes erros padrão para as estimativas da decomposição de variância e funções de resposta-impulso.

2.5 Cointegração e mecanismo de correcção de erro (ECM)

Estudos empíricos em macroeconomia quase sempre envolvem variáveis não estacionárias e/ou com tendência, tais como rendimento, consumo, procura de moeda, nível de preços, fluxos comerciais, e taxas de câmbio (Greene, 1997).

Como vimos na secção 2.3.2, se as variáveis forem não estacionárias devido a raízes unitárias, elas poderiam ser trazidas a estacionaridade através de transformações lineares de diferenciação dos dados. Por essa razão, segundo Hendry and Juselius (2000), "it is natural to inquire if other linear transformations than differencing will also induce stationarity." (p. 2). Segundo os autores, a resposta é "possivelmente", mas ao contrário da diferenciação dos dados, não existe nenhuma garantia que o resultado seja $I(0)$: "cointegration analysis is designed to find linear combination of the variables that also remove unit roots." (p. 2).

Matematicamente, a cointegração é definida por: seja β um vector de k parâmetros $(\beta_1, \beta_2, \dots, \beta_k)$ e Y_t um vector de k variáveis económicas $(y_{1,t}, y_{2,t}, \dots, y_{k,t})$ com características de equilíbrio de longo prazo, tal que a situação de equilíbrio ocorre quando se verifica a restrição

$$\beta Y_t = 0, \quad (2.25)$$

na maior parte do tempo, Y_t encontra-se fora da situação de equilíbrio, sendo a distância a que o sistema se encontra da situação de equilíbrio, que se designa por erro de equilíbrio v_t , dada por

$$\beta Y_t = v_t, \quad (2.26)$$

Os desvios devem ser limitados de modo que o processo de erro de equilíbrio seja estacionário. Assim, diz-se que as componentes do vector $(y_{1,t}, y_{2,t}, \dots, y_{k,t})$ são cointegradas de ordem (d, b) , $(d, b = 1, 2, \dots)$, e designa-se $Y_t \sim CI(d, b)$, se:

- todas as componentes do vector Y_t são integradas de ordem d ou $I(d)$, isto é, as variáveis precisam ser diferenciadas d vezes para que sejam estacionárias;
- existir um vector β tal que a combinação linear βY_t seja integrada de ordem $(d - b)$.

A essa combinação linear representada pelo vector β , chama-se vector cointegrante. Na definição de cointegração é necessário tomar em conta os seguintes aspectos:

- A cointegração refere-se à combinação linear de variáveis económicas não estacionárias; todas as variáveis devem ser integradas da mesma ordem, o que não implica necessariamente que estas sejam cointegradas (por outro lado, se as variáveis forem integradas de diferentes ordens, estas não podem ser cointegradas);
- Se o vector Y_t tem k componentes, pode existir até $k - 1$ (inclusive) vectores cointegrantes linearmente independentes (no caso em que se tem apenas duas variáveis, se existir um vector cointegrante este tem de ser único).

Supondo que as variáveis em Y_t sejam $I(d)$, e que existe um vector β tal que a combinação linear βY_t é $I(0)$, então, esse vector é o único capaz de produzir uma combinação das variáveis que seja estacionária. Dito de outra forma, não existirá nenhuma raiz unitária na combinação entre essas variáveis. Consequentemente, cointegração é uma restrição em um modelo dinâmico e, por conseguinte, pode ser testada (Hendry & Juselius, 2000).

Intuitivamente, se duas ou mais séries são integradas da mesma ordem, sua combinação poderia ser estável em redor de uma média fixada. A implicação seria que as séries caminham juntas, aproximadamente, à *mesma taxa*. Portanto, séries que satisfazem essa condição são ditas cointegradas. Em tal caso, segundo Greene (1997), podemos distinguir entre relação de longo prazo entre as variáveis em Y_t , que é o modo no qual elas caminham juntas, e a dinâmica de curto prazo, isto é, a relação entre os desvios das variáveis em torno da sua tendência de longo prazo. Por essa

razão, a tentativa de diferenciar os dados antes da estimação do modelo por forma a garantir a estabilidade do mesmo, seria contraproducente dado que obscureceria a relação de longo prazo.

A cointegração e uma técnica relacionada, o mecanismo de correcção de erro (ECM), interessam-se em métodos de estimação de modelos que preservem a informação sobre as duas formas mencionadas de covariação dos dados. Num sistema de k variáveis, pelo menos em princípio, poderiam existir no máximo $k-1$ vectores cointegrantes linearmente independentes²⁸. O número de vectores cointegrantes linearmente independentes que existem em um sistema de equilíbrio é chamado característica cointegrante. Greene (1997), aponta que a consequência desse resultado é que infelizmente a característica cointegrante pode superar a unidade, então sem o conhecimento exacto da relação fora da amostra, não é possível estimar relações como vectores cointegrantes.

2.5.1 Tendências comuns

Se um determinado conjunto de variáveis $I(1)$ for cointegrado, intuitivamente, isso sugere que a combinação linear não cria misteriosamente uma nova variável bem comportada, mas algo presente nas variáveis originais deve estar faltando no modelo. (Greene, 1997).

Considerando que o conjunto de variáveis seja modelado de tal modo que cada uma delas tenha uma tendência linear nos dados, e uma série de resíduos estacionários. Uma combinação linear dessas variáveis produzirá uma nova variável que, em geral, é ainda $I(1)$. Assim, o único caminho pelo qual essa variável será estacionária é impondo uma restrição sobre a combinação da componente de tendência linear por forma que ela se anule. Portanto, o efeito de combinar as variáveis linearmente é remover a tendência linear comum.

Segundo Stock and Watson (1988, as cited in Greene, 1997), a única maneira para que um conjunto de variáveis seja cointegrado, para começar, é que elas tenham uma tendência comum, pois, caso elas partilhem mais de uma tendência não será possível encontrar uma combinação linear das variáveis que seja cointegrada mesmo que elas partilhem tendências comuns. Por isso, se duas ou mais variáveis são cointegradas, então elas devem partilhar exactamente uma tendência comum.

²⁸ A prova para essa proposição pode ser encontrada em Greene (1997).

Um conjunto de k variáveis que são cointegradas pode ser escrito como uma componente estacionária mais combinações lineares de um pequeno conjunto de tendências comuns. (Stock & Watson, 1988, as cited in Greene, 1997). Se a característica cointegrante do sistema for r , então podem existir até $k - r$ tendências lineares e $k - r$ passeios aleatórios comuns. O efeito da cointegração é evacuar essas tendências comuns nas variáveis resultantes.

2.5.2 Determinação de cointegração no modelo VAR

Na secção 2.4.2 discutiu-se em que condições um processo VAR seria estacionário. Viu-se que quando algumas das raízes características do sistema (2.24) estiverem no limite da linha do círculo unitário, o processo $\{Y_t\}$ é não estacionário. Todavia, algumas combinações lineares podem ser estacionárias, embora as variáveis por si sejam não estacionárias.

Segundo Juselius (2003), "the presence of unit roots in the unrestricted VAR model corresponds to nonstationary stochastic behavior which can be accounted for by a reduced rank ($r < k$) restriction of the long-run levels matrix $\Pi = \alpha\beta'$." (p. 85).

A cointegração significa que certas combinações lineares das variáveis no processo são integradas de ordem mais baixa que a do próprio processo.

When two or several variables have common stochastic (and deterministic) trends, they will show a tendency to move together in the long-run. Such cointegrated relations, $\beta'Y_t$, can often be interpreted as long-run economic steady-state relations and are, therefore, of considerable economic interest. (Juselius, 2003, p. 86).

No modelo VAR a hipótese de cointegração pode ser formulada como uma restrição da característica reduzida na matriz Π estimada por OLS.

2.5.3 Mecanismo de correcção de erro (ECM)

O modelo VAR irrestrito (2.22) pode ser reformulado em diferentes parametrizações, sem impor qualquer restrição nos parâmetros do modelo, isto é, sem que haja uma alteração no valor da função de verosimilhança, chamadas modelos vectoriais de correcção de erro, que dão uma formulação do

modelo em termos de diferenças, diferenças desfasadas e valores em níveis das variáveis do processo. Segundo Juselius (2003), existem inúmeras vantagens nessas formulações:

- i. o efeito da multicolinearidade, que tipicamente está presente nos dados de séries temporais, é significativamente reduzida na forma de ECM;
- ii. toda informação do efeito de longo prazo está contida na matriz Π que pode, por conseguinte, receber especial atenção no caso da resolução do problema de cointegração;
- iii. a interpretação dos parâmetros é mais intuitiva, a medida que os coeficientes podem ser naturalmente classificados em efeitos de curto e longo prazos;
- iv. geralmente o interesse é saber porque uma determinada variável mudou de um período precedente para o subsequente. O ECM responde a esta questão directamente.

Podem ser consideradas três versões diferentes da formulação ECM

$$\Delta Y_t = \Gamma_1^{(m)} \Delta Y_{t-1} + \Gamma_2^{(m)} \Delta Y_{t-2} + \dots + \Gamma_{p-1}^{(m)} \Delta Y_{t-p+1} + \Pi Y_{t-m} + \Phi A_t + \varepsilon_t \quad (2.27)$$

Onde m é um valor inteiro entre 1 e p . Assim, mesmo que o valor de m altere, o valor da função de verosimilhança não mudará. Supondo $p = 2$, ilustra-se a reformulação do modelo ECM:

(1) Formulação do ECM com $m = 1$

$$\Delta Y_t = \Gamma_1^{(1)} \Delta Y_{t-1} + \Pi Y_{t-1} + \Phi A_t + \varepsilon_t \quad (2.28)$$

Onde, $\Pi = I_2 - \Pi_1 - \Pi_2$, e $\Gamma_1^{(1)} = -\Pi_2$. Os valores das variáveis em níveis na matriz Π estão determinados no tempo $t - 1$. A inclusão dessa última componente no modelo levanta a questão de como resolver o problema da não estacionaridade. Dado que um processo estacionário, no primeiro membro, não pode ser igualado a outro não estacionário, segundo membro (que é uma combinação linear de variáveis estacionárias e não estacionárias), os resultados da estimação a penas fazem sentido se ΠY_{t-1} define uma combinação linear estacionária das variáveis. Segundo Juselius (2003):

This is, in a nutshell, what cointegration analysis does: it identifies stationary linear combinations between nonstationary variables so that an $I(1)$ model can be reformulated exclusively in stationary variables. Our task is to give the stationary linear combinations an economically meaningful interpretation by imposing relevant identifying or over-identifying restrictions on the coefficients. (p. 67).

(2) Formulação do ECM com $m = 2$

$$\Delta Y_t = \Gamma_1^{(2)} \Delta Y_{t-1} + \Pi Y_{t-2} + \Phi A_t + \varepsilon_t \quad (2.29)$$

Com $\Pi = I_2 - \Pi_1 - \Pi_2$, e $\Gamma_1^{(2)} = (I_2 - \Pi_1)$. Assim, a matriz Π permanece como na situação anterior, mas $\Gamma_1^{(2)}$ não. Essa última matriz, mede os efeitos de longo prazo acumulados, enquanto $\Gamma_1^{(1)}$ descreve os efeitos transitórios medidos pelas mudanças nas defasagens das variáveis. Mesmo que o poder de explanação dos dois modelos seja idêntico, os coeficientes estimados e sua significância podem variar consideravelmente.

Frequentemente boa parte de coeficientes estatisticamente significantes são obtidos com a formulação (2.29) do que com (2.28). Todavia, a determinação de coeficientes significativos em demasia não implica necessariamente alto poder de explanação, pode também advir da forma de reparametrização do modelo. Isso mostra que a interpretação dos coeficientes estimados nos modelos dinâmicos é menos clara que nos modelos de regressão estáticos (Juselius, 2003).

(3) Formulação do ECM em taxas de aceleração, mudanças e níveis

Uma outra forma conveniente de representar o VAR é em diferenças de segunda ordem (taxas de aceleração), mudanças e níveis.

$$\Delta^2 Y_t = \Gamma \Delta Y_{t-1} + \Pi Y_{t-2} + \Phi A_t + \varepsilon_t \quad (2.30)$$

Esta formulação é muito útil no caso em que Y_t contém variáveis $I(2)$. Mas é, em geral, útil quando o período amostral em análise contiver épocas de rápidas mudanças, assim, taxas de aceleração tornam-se determinantes relevantes no comportamento dos agentes. $\Gamma = I_2 - \Gamma_1$, e Π permanece

inalterado. Note-se que diferentemente dos casos anteriores, em Γ_1 adicionou-se um factor constante, -1, nos elementos da diagonal. Por essa razão, a significância dos elementos da diagonal é somente uma consequência da aplicação do operador de primeiras diferenças uma vez mais a ΔY_t .

No modelo de ECM em sua formulação geral (2.27), se $Y_t \sim I(1)$, então $\Delta Y_t \sim I(0)$ o que implica que Π não pode ter característica completa, já que isso conduziria a uma inconsistência no modelo (2.27). Portanto, Π terá de ter uma característica reduzida, isto é:

$$\Pi = \alpha\beta' \quad (2.31)$$

onde, α e β são matrizes de dimensão $k \times r$, $r < k$. Se $Y_t \sim I(1)$, então a hipótese de que o VAR seja cointegrado é dada por:

$$\Delta Y_t = \Gamma_1 \Delta Y_{t-1} + \dots + \Gamma_{p-1} \Delta Y_{t-p+1} + \alpha\beta' Y_{t-1} + \Phi A_t + \varepsilon_t \quad (2.32)$$

$\beta' Y_t$ é $I(0)$, e cada coluna de β é um vector cointegrante. Os elementos da matriz α são obtidos pela estimação do modelo ECM. Deste modo, a metodologia de Johansen (1995) para o teste de cointegração consiste em estimar a matriz Π do VAR irrestrito e testar se se pode rejeitar a restrição imposta (de característica reduzida).

2.5.4 Componentes determinísticas no modelo de cointegração

Uma das principais dificuldades encontradas na especificação de modelos ECM refere-se à interpretação dos coeficientes em termos dos efeitos dinâmicos. Segundo Hendry and Juselius (2000), "this is also true for trend and the constant term, as well as other deterministic terms like dummy variables." (p. 15).

Quando um conjunto de variáveis partilha tendências estocásticas e determinísticas, é possível encontrar uma combinação linear que cancele ambas tendências. Tal relação de cointegração resultante não contém nenhuma tendência. A combinação linear das variáveis, no VAR, pode remover apenas a tendência estocástica, mas não a determinística, assim, torna-se necessário incluir uma tendência linear no espaço cointegrante. Argumentos similares são aplicáveis a *dummies*.

Segundo Johansen (1995), Hendry and Juselius (2000) e Juselius (2003), o papel das componentes determinísticas e estocásticas no modelo VAR, cointegrado, é bastante complicado. Todavia, é importante entender sua importância no modelo, pois, pode-se obter erros nos parâmetros estimados se as componentes determinísticas forem incorrectamente formuladas. Em parte porque as distribuições assintóticas dos testes de cointegração não são invariantes à essas especificações. Além disso, as propriedades das formulações resultantes podem ser indesejáveis para previsões, por inadvertidamente reter componentes desnecessárias, como é o exemplo de tendências quadráticas.

Cinco casos para inclusão de interceptos e tendências

“Many economic variables typically exhibit linear deterministic growth (at least locally over the sample period) in addition to stochastic growth.” (Juselius, 2003, p. 102). Todavia, nem sempre é possível fazer uma distinção estatística entre as duas tendências.

Por vezes é preferível aproximar o comportamento da tendência com uma tendência estocástica, as vezes é melhor modelar com uma tendência determinística, e na maioria dos casos necessitamos uma combinação de ambas (Juselius, 2003). Para melhor entender a necessidade de inclusão de componentes determinísticas, segue-se o procedimento adoptado em Hendry and Juselius (2000) que consiste em decompor, inicialmente, o intercepto, π , em componentes que induzem a um crescimento no sistema, e em componentes que tomam as médias da relação de cointegração. Considerando a representação VAR (2.27), no caso em que $m = 1$ e $p = 2$:

$$\Delta Y_t = \Gamma_1 \Delta Y_{t-1} + \alpha(\beta' Y_{t-1}) + \pi + \varepsilon_t \quad (2.33)$$

Dado que as r relações de cointegração $\beta' Y_{t-1}$ são estacionárias, cada uma delas tem uma média constante. Similarmente, ΔY_t é estacionário com uma média constante, que se denota por $E(\Delta Y_t) = \zeta$, que descreve um vector ($k \times 1$) de taxas de crescimento. Além disso, considere-se $E(\beta' Y_{t-1}) = \xi$ descrevendo um vector ($r \times 1$) de interceptos na relação de cointegração. Tomando (2.33) em valor esperado:

$$(I_k - \Gamma_1)\zeta = \alpha E(\beta' Y_{t-1}) + \pi = \alpha \xi + \pi$$

Por conseguinte, $\pi = (I_k - \Gamma_1)\zeta - \alpha\xi$. Note-se que o termo constante π no modelo consiste em duas componentes: uma relacionada às taxas de crescimento linear nos dados, e a outra relacionada aos valores médios da relação de cointegração (isto é, os interceptos na relação de longo prazo). De (2.33) pode-se constatar que:

$$\beta'E(\Delta Y_t) = E(\Delta\beta'Y_t) = \beta'\zeta = 0$$

Assim, (2.33) pode ser expresso sob forma de desvios médios:

$$(\Delta Y_t - \zeta) = \Gamma_1(\Delta Y_{t-1} - \zeta) + \alpha(\beta'Y_{t-1} - \xi) + \xi_t \quad (2.34)$$

Existem duas formas de correcção de equilíbrio em (2.34): uma de crescimento de ΔY_t no sistema em relação à sua média ζ ; e dos vectores cointegrantes $\beta'Y_{t-1}$ para sua média ξ . Os dois valores médios, ζ e ξ , têm um papel importante no modelo VAR, cointegrado, e é importante averiguar se eles são significativamente diferentes de zero ou não no início da análise empírica.

Os casos básicos que descrevem como uma tendência linear e um intercepto podem entrar na especificação VAR, podem ser ilustrados, segundo Juselius (2003), por um simples VAR(1) contendo uma constante e uma tendência linear, e os efeitos dinâmicos de curto prazo igualados a zero, isto é, $\Gamma_t = 0$:

$$\Delta Y_t = \alpha\beta'Y_{t-1} + \pi + \delta t + \varepsilon_t \quad (2.35)$$

Os dois vectores π e δ de dimensão $(k \times 1)$ podem ser decompostos em dois novos vectores, dos quais um está relacionado com o valor médio das relações cointegrantes, $\beta'Y_{t-1}$, e outras duas taxas de crescimento em ΔY_t :

$$\begin{cases} \pi = \alpha\xi + \zeta_0 \\ \delta = \alpha\rho + \tau \end{cases} \quad (2.36)$$

Substituindo (2.36) em (2.35) tem-se:

$$\Delta Y_t = \alpha \beta' Y_{t-1} + \alpha \xi + \zeta + \alpha \rho t + \tau t + \varepsilon_t \quad (2.37)$$

Segundo Hendry and Juselius (2000), pode-se, sempre, escolher ξ e ρ de modo que a combinação linear em (2.37) seja expressa sob a forma de valor esperado por:

$$E(\Delta Y_t) = \zeta + \tau t \quad (2.38)$$

Assim, um valor de $\zeta \neq 0$ corresponde a um crescimento constante nas variáveis Y_t , enquanto que $\tau \neq 0$ corresponde a tendências lineares no crescimento, e assim tendências quadráticas nas variáveis. Por isso, o termo constante e a tendência linear determinística representam um papel duplo no modelo cointegrado: “in the α directions, they describe a linear trend and an intercept in the steady-state relations; in the remaining directions, they describe quadratic and linear trends in the data.” (Hendry & Juselius, 2000, p. 17).

Os cinco casos frequentemente usados, encontrados em Johansen (1995) são:

Caso 1: não existem restrições em π e δ , assim o termo de tendência e o intercepto são irrestritos no modelo VAR. Com parâmetros irrestritos, π , δ , o modelo é consistente com tendências lineares nas séries diferenciadas ΔY_t , e, assim, tendências quadráticas em Y_t . Embora tendências quadráticas possam, por vezes, melhorar o ajustamento dentro da amostra, é provável que as previsões fora da amostra produzam resultados incertos.

Caso 2: $\tau = 0$, mas ζ , ξ , ρ permanecem irrestritos, assim a tendência está restrita apenas ao espaço de cointegração, mas a constante não contém nenhuma restrição no modelo. Assim, sendo τ igual a zero, o modelo permite ainda uma tendência linear, mas evita tendências quadráticas nos dados. $E(\Delta Y_t) = \zeta \neq 0$, implica uma tendência linear determinística nos níveis das variáveis, Y_t . Se adicionalmente, $\rho \neq 0$, essas tendências lineares nas variáveis não se cancelam na relação de cointegração, deste modo, o modelo contém relações de tendência estacionária que podem descrever uma simples variável com tendência estacionária, $(y_{1,t} - b_1 t) \sim I(0)$, ou uma relação de equilíbrio

$(\beta_1'Y_t - b_2t) \sim I(0)$. Por conseguinte, a hipótese de que uma variável contém tendência estacionária pode ser testada neste modelo ²⁹.

Caso 3: $\delta = 0$, assim não existem tendências lineares em (2.35). Dado que o termo constante π é irrestrito, existem ainda tendências lineares nos dados, mas não há nenhuma tendência determinística em qualquer relação cointegrante. Também, $E(\Delta Y_t) = \zeta \neq 0$, é consistente com as tendências lineares determinísticas nas variáveis em níveis Y_t , mas se $\rho = 0$, essas tendências são canceladas na relação de cointegração. No caso em que $\pi \neq 0$, leva em consideração ambas tendências lineares no DGP e um intercepto não nulo nas relações de cointegração.

Caso 4: $\delta = 0$, $\zeta = 0$, mas $\xi \neq 0$, assim o termo constante está restrito ao espaço de cointegração em (2.35). Neste caso, não existem tendências lineares determinísticas nos dados, consistentes com $E(\Delta Y_t) = 0$. As únicas componentes determinísticas no modelo são os interceptos em qualquer relação de cointegração, implicando que algumas médias de equilíbrio são diferentes de zero.

Caso 5: $\delta = 0$ e $\pi = 0$, assim o modelo exclui todas componentes determinísticas nos dados, com ambas $E(\Delta Y_t) = 0$ e $E(\beta_1'Y_{t-1}) = 0$, implicando a inexistência de crescimentos e interceptos em todas relações cointegrantes. Dado que um intercepto é geralmente necessário para tomar em conta os níveis iniciais de medição, Y_0 , somente em um caso excepcional quando as medições iniciam em zero, ou quando se cancelam nas relações de cointegração, pode a restrição $\pi = 0$ ser justificada.

²⁹ Uma variável é dita estacionária em tendência, se ela for estacionária em torno duma tendência determinística, isto é, estacionária após removida a tendência.

CAPÍTULO III

3. Metodologia

3.1 Material e modelo de pesquisa

O estudo pode ser caracterizado como uma pesquisa conclusiva. Segundo Forte (2004), as pesquisas conclusivas podem ser descritivas ou causais, e as pesquisas descritivas podem ser transversais ou longitudinais. Este estudo caracteriza-se como uma pesquisa conclusiva descritiva longitudinal. Como pesquisa descritiva, pretendeu -se descrever algo, no caso as relações de interdependência entre as taxas de juro das operações activas e passivas (R_a e R_p), base monetária (BaM) e taxa de inflação (π). Uma importante diferença entre as pesquisas exploratória e descritiva é que a segunda se caracteriza pela formulação prévia de hipóteses específicas, assim as informações necessárias são claramente definidas (Marconi & Lakatos, 1991).

A pesquisa descritiva utilizada para este estudo também pode ser classificada como longitudinal, pois neste caso, quatro amostras fixas (R_a e R_p , BaM e π) são medidas repetidamente. Um estudo longitudinal difere de um transversal porque as amostras permanecem as mesmas ao longo do tempo, somente modificando-se os seus valores. Trata-se também, de acordo com Marconi e Lakatos (1991), de uma pesquisa indutiva, pois a cointegração é estabelecida como uma relação entre as variáveis. As considerações extraídas da amostra são generalizadas para a população e o futuro é interpretado como sendo, de algum modo, uma consequência do passado.

Os dados para este estudo são, de acordo com Forte (2004), classificados como secundários, pois já foram colectados para objectivos que não os do problema em pauta. Séries históricas mensais no período entre Janeiro de 1996, um ano antes da introdução do mercado monetário interbancário (MMI), até Junho de 2008, perfazendo um intervalo de aproximadamente dez anos.

3.1.1 Fonte de dados

Os dados referentes às taxas de inflação, variações mensais do índice de preços ao consumidor com base em Dezembro de 2004, foram obtidos no Instituto Nacional de Estatística (INE), no departamento de preços e conjuntura. As séries da BaM e das taxas de juro activas e passivas (taxas médias de juros aplicadas no mercado e compiladas pelo BM com prazos de 180 dias) foram obtidas no departamento de estudos, Divisão de Estatística do BM. Também disponíveis nos endereços: www.ine.gov.mz e www.bancomoc.mz.

3.2 Métodos

3.2.1 Análise de dados

Os resultados empíricos do estudo foram obtidos através do *software* estatístico Eviews versão 4.0. Segundo Fedderke (2003), a obtenção de um determinado nível de sucesso na aplicação da metodologia de Johansen, passa por diferentes passos, desde a estimação do modelo na forma reduzida, de modo a analisar as suas propriedades, incluindo: (1) um teste prévio da ordem de integração de cada uma das variáveis; (2) uma selecção adequada do tamanho da desfasagem, para assegurar que o modelo de ECM tenha erros Gaussianos; (3) determinar se o sistema precisa estar condicionado a quaisquer variáveis estacionárias, incluindo variáveis *dummy*; (4) teste da característica reduzida do modelo, o que inclui a necessidade de saber se o sistema é $I(1)$ ou $I(2)$; (5) certificar-se da existência de tendências nos dados, dado que isso determina se o sistema será estimado com variáveis determinísticas (constantes e tendências), ou não; (6) teste de presença de exogeneidade fraca.

I. Estacionaridade

Procede-se aos testes, para identificar a possível existência de raízes unitárias nas séries em análise. Todavia, testa-se previamente a presença de interceptos e tendências³⁰. A partir dos dados da base monetária, taxas de juro e taxa de inflação, testam-se as hipóteses:

³⁰ A opção por esse procedimento baseou-se no teste de presença de tendência aplicado por Mills (1993, citado por Camuri, 2005). No entanto, é preciso ressaltar que o teste adoptado não tem por objectivo inferir sobre a existência ou não de intercepto e tendência nas variáveis estudadas. O objectivo do uso do teste foi exclusivamente dar um suporte estatístico à inclusão ou não de termos de intercepto e/ou tendência nos testes ADF e PP.

$$\begin{cases} H_0 : \mu = 0 \\ H_a : \mu \neq 0 \end{cases} \quad (3.1)$$

Através dos valores da média e desvio padrão fornecidos pela amostra, calcula-se a estatística *t* - *student*, no caso do intercepto, e compara-se aos valores críticos a 5% e 1% de significância

$$t = \frac{(\bar{y}_{i,t} - 0)}{(s / \sqrt{T})} \quad (3.2)$$

Para o caso de inclusão de tendências criam-se, a partir das variáveis iniciais, novas variáveis ΔR_a , ΔR_p , ΔBaM e $\Delta \pi$, e pelo mesmo procedimento calculam-se as médias e desvios padrão, testando-se subsequentemente (3.1) para verificar a significância da estatística *t*.

Reconsiderando as hipóteses a serem testadas a partir dos testes ADF e PP

$$\begin{cases} H_0 : \delta = 0 \\ H_a : \delta < 0 \end{cases}$$

Se os dois testes apresentarem conclusões distintas quanto à presença de raízes unitárias, através de uma inspeção visual nos gráficos das variáveis, investigar-se-á a presença de quebras estruturais nos dados, e procurar-se-ão as prováveis datas de ocorrência dessas quebras. Tendo identificado as datas, segue-se a uma das sugestões de Enders que consiste em testar a presença de raízes unitárias em cada sub-período da amostra.

II. Ordem da defasagem a incluir no modelo

A determinação da ordem de defasagem consiste em estimar o modelo na *p* - ésima ordem (ordem máxima) e testar recursivamente a significância geral do modelo da ordem máxima à mínima, que é um no caso de inexistência de qualquer variável exógena e zero caso contrário.

O método é baseado em um teste sequencial de razão de verosimilhança (*LR*) com distribuição χ^2 :

$$LR = (T - m) \{ \log |\Omega_{p-1}| - \log |\Omega_p| \} \sim \chi^2(k^2) \quad (3.3)$$

Onde m é o número de parâmetros por equação. A hipótese nula em ambos casos é a de nulidade de todos os coeficientes correspondentes a um determinado *lag*.

III. Estabilidade do modelo

A estabilidade do modelo será verificada a partir das raízes inversas do polinómio característico $\Pi(L)$. A condição de estabilidade é sempre garantida no caso em que todas raízes forem, em módulo, inferiores a unidade.

IV. Teste da característica cointegrante

A imposição de restrições nos parâmetros de longo prazo do modelo, β , e finalmente restrições nos parâmetros de ajustamento de curto prazo α e Γ , passa por uma escolha apropriada da restrição da característica cointegrante, r , na matriz Π , dado que:

- Se se subestimar r , os mecanismos empíricos de correcção do erro de equilíbrio ficarão omitidos;
- Se se super-estimar r , as distribuições de algumas estatísticas tornar-se-ão não *standards*, assim inferências incorrectas podem resultar do uso dos valores críticos convencionais (baseados em t , F , χ^2);
- As previsões serão menos precisas devido a retenção de componentes $I(1)$ que irão incrementar a variância de previsão.

O teste de presença de r vectores cointegrantes é baseado em uma aproximação proposta por Johansen (1988, as cited in Hendry and Juselius, 2000). Onde r é determinada por um procedimento de razão de verosimilhanças que testa sequencialmente a significância de um determinado número de vectores cointegrantes, de $r = 0$ até $r = k - 1$:

$$LR_r(r | k) = -T \sum_{i=r+1}^k \log(1 - \lambda_i) \quad (3.4)$$

Onde λ_i , $i = 1, \dots, r$, é o i -ésimo valor próprio obtido na matriz Π . Com as hipóteses:

$$\begin{cases} H_k : r = k \\ H_r : r < k \end{cases}$$

A segunda hipótese corresponde a um processo $\{Y_t\}$ estacionário. Os valores críticos da distribuição dessa estatística estão disponíveis até $k = 10$ variáveis. O facto desses críticos serem sensíveis a algumas suposições inerentes à tendência, faz com que não seja apropriado aplicar esses métodos a modelos que contenham outros regressores determinísticos. Por exemplo, a inclusão de *dummies* no modelo pode fazer com que haja uma quebra da tendência linear nos dados (QMS, 2002).

V. Presença de tendências e interceptos no modelo

Segundo a QMS (2002), caso o investigador não esteja certo de qual das suposições sobre a inclusão de termos determinísticos incluir no modelo, secção 2.5.4, aconselha-se a optar por uma solução prévia que consiste na estimação de um sumário de todas possíveis soluções do problema, e daí analisar, com o auxílio de critérios de informação, para escolher a mais parcimoniosa e altamente significativa sob ponto de vista estatístico. Os critérios usados na escolha da forma funcional do teste de relações cointegrantes são:

(1) Critério de informação de Akaike (AIC):

$$AIC = -2l/T + 2n/T \quad (3.5)$$

Onde

$$l = -\frac{T}{2} \{k(1 + \log 2\pi) + \log |\hat{\Omega}|\}, \text{ e } |\hat{\Omega}| = \det\left(\frac{1}{T-p} \sum_i e_i e_i'\right), \quad (3.6)$$

e n corresponde ao número total de parâmetros estimados no modelo.

(2) Critério de Schwarz (SC):

$$SC = -2l/T + n \log T/T \quad (3.7)$$

De modo geral, no processo de escolha de modelos, considera-se quase sempre o modelo no qual o critério de informação se apresenta com o menor valor possível.

VI. Causalidade de Granger/Block Exogeneity Wald Test

Na estimação de um modelo VAR, um passo importante para a interpretação das funções de resposta a impulso e decomposição de variância, é o conhecimento prévio de como as variáveis afectam umas as outras. Pois a determinação da ordenação das variáveis em termos de *importância* é fundamental para a verificação da hipótese de ortogonalidade dos resíduos obtidos pela decomposição de Cholesky.

Um procedimento utilizado nessas situações, consiste na determinação da ordenação das variáveis a partir do teste de Causalidade de Granger, com a denominação de **Block Exogeneity Wald Test**. O teste calcula para cada equação do VAR a significância conjunta de cada variável endógena desfasada nessa equação, bem como uma estatística conjunta para a significância de todas as variáveis endógenas desfasadas na equação.

Após a ordenação das variáveis pelo critério do teste dos valores críticos obtidos a partir do teste de Causalidade de Granger, a análise estrutural feita pelas funções de resposta-impulso e decomposição de variância, passa pela estimação prévia de um modelo VAR recursivo, ou pelo processo de ortogonalização dos resíduos do modelo da forma reduzida pelo procedimento de Cholesky.

Na estimação do modelo (2.22), considerou-se apenas os logaritmos naturais da base monetária e taxas de juro. A inflação mensal não pode ser logaritmizada por conter valores negativos em sua série ³¹. Por outro lado, embora os modelos VAR sejam derivados sob a suposição de normalidade multivariada e, por conseguinte, os estimadores obtidos por métodos de máxima verosimilhança, o modelo será estimado por regressões OLS aplicadas a cada equação.

“Se os modelos econométricos são ajustados por meio de métodos *standards*, invariantes, ou seja, sem a utilização de uma distribuição teórica prévia, não haverá, por conseguinte, uma tendência de ‘mau’ comportamento no momento de realização de previsões cada vez mais distantes dos valores amostrais?”. Espera-se que essa tendência se torne pior se tais métodos não forem usados na

³¹ O processo de logaritmizar as séries tem o principal objectivo de estabilizar a variância dos dados. Segundo Hendry and Juselius (2000), uma das razões de assumir que as variáveis sejam analisadas em logaritmos, é que evita com que as taxas de crescimento no modelo dependam dos níveis das variáveis.

estimação e se contínuas re-especificações *ad hoc*'s em equações individuais forem intensivas por forma a melhorar o ajustamento ³². (Runkle, 1987)

Diagnóstico do modelo VAR

(1) **Autocorrelação serial dos resíduos** - A presença de autocorrelação serial dos resíduos é testada através de um teste *LM* generalizado para o caso multivariado. O procedimento do teste consiste em testar a presença de autocorrelação dos resíduos do primeiro *lag* até uma determinada ordem específica *h*, através de uma regressão auxiliar de ε_t contra ε_{t-h} , onde os primeiros *h* componentes de ε_{t-h} foram igualados a zero. Sob a hipótese nula de ausência de autocorrelação serial de ordem *h*, *LM* tem uma distribuição assintoticamente χ^2 com k^2 graus de liberdade ³³.

(2) **Normalidade multivariada** - Dado que o modelo é derivado sob a suposição de existência de uma distribuição normal multivariada das variáveis, essa conclusão é obtida através de uma extensão multivariada do teste de Jacque-Berra em que, como no caso das funções de resposta a impulso, os resíduos precisam ser previamente ortogonalizados. Optou-se pela factorização de Cholesky. A hipótese nula para o teste é de presença de normalidade, com distribuição χ^2 e $k(k+1)(k+2)(k+7)/24$ graus de liberdade.

(3) **Heteroscedasticidade** - Aplicou-se o teste de heteroscedasticidade de White, generalizado, que consiste na estimação de uma regressão auxiliar dos resíduos OLS, regredindo cada produto cruzado dos resíduos com os produtos cruzados dos regressores e testando a significância conjunta da regressão. Os termos cruzados dos regressores correspondem apenas aos níveis e quadrados dos regressores. Em cada equação um termo constante é incluído como regressor. Sob a hipótese nula de homoscedasticidade, a estatística do teste, *LM*, é distribuída assintoticamente χ^2 com *mn* graus de liberdade. Onde $m = k(k+1)/2$ e *n* corresponde ao total dos regressores incluídos na regressão do teste ³⁴.

³² Mas, segundo Juselius (2003), "OLS regression equation by equation are Maximum Likelihood estimates as long as no restrictions have been imposed on the VAR model" (p. 27).

³³ A fórmula do teste pode ser consultada em Johansen (1995).

³⁴ Mais detalhes sobre o teste de heteroscedasticidade de White ver Johnston e Dinardo (2001).

CAPÍTULO IV

4. Análise de resultados

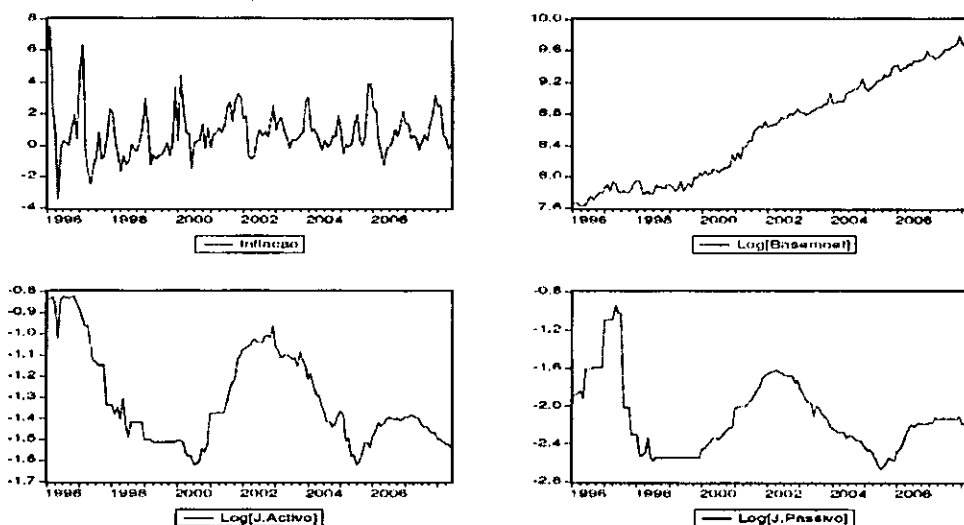
Este capítulo apresenta os resultados obtidos no presente trabalho. Primeiramente serão apresentados os testes necessários à determinação da ordem de integração das séries estudadas. Em seguida faz-se a análise estrutural do modelo, pelas funções de resposta a impulso e decomposição de variância.

4.1 Estacionaridade

O primeiro passo no processo na averiguação da existência de raízes unitárias, é a verificação da presença de intercepto e tendência nas séries estudadas. Para conferir a necessidade de inclusão de termo de intercepto no teste de raiz unitária utiliza-se o seguinte procedimento:

1) A partir da observação da variabilidade dos dados pela representação gráfica das séries.

Figura 4.1: Análise descritiva (séries em níveis)

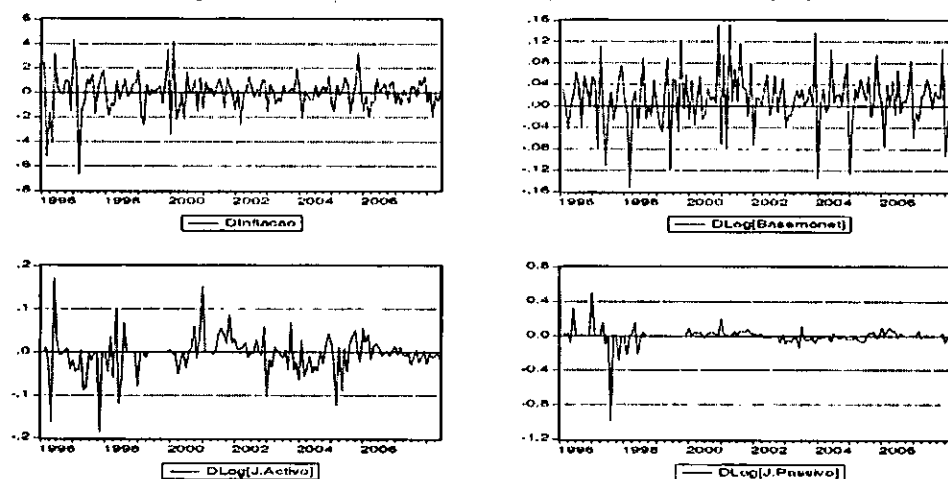


Fonte: Output – Eviews

2) Com base nos dados das estatísticas descritivas, da média e do desvio padrão, calcula-se a estatística t e testa-se a nulidade dessa média a 5% de significância.

3) Pelo mesmo procedimento, conclui-se sobre a presença de termos de tendência nos testes de raiz unitária através das séries diferenciadas.

Figura 4.2: Análise descritiva (séries em diferenças)



Fonte: Output – Eviews

As principais conclusões do teste são apresentadas resumidamente na tabela 4.1:

Tabela 4.1: Teste sobre inclusão de intercepto e tendência

Teste <i>t</i> a 5% de significância				
Variável	<i>t</i> calculado	Hipótese nula	Intercepto	Tendência
Log[base monetária]	154,6306	Rejeitar	Sim	
1ª diferença	3,2982	Rejeitar		Sim
Inflação	6,4095	Rejeitar	Sim	
1ª diferença	-0,2730	Não rejeitar		Não
Log[J. Activo]	-68,3539	Rejeitar	Sim	
1ª diferença	-1,2997	Não rejeitar		Não
Log[J. Passivo]	25,6923	Rejeitar	Sim	
1ª diferença	-0,2295	Não rejeitar		Não

Fonte: Elaboração própria

De acordo com os resultados reportados na tabela 4.1, nota-se que, ao nível de 5% de significância, não há evidências suficientes para rejeitar a hipótese da nulidade da média do logaritmo da base monetária, em níveis, bem como em primeiras diferenças e, por conseguinte, a possibilidade de inclusão de termos de tendência e intercepto nos testes ADF e PP. Nas restantes variáveis, o teste não rejeita apenas a possibilidade de inclusão de termos de intercepto nos testes ADF e PP.

A seguir, são apresentados nas tabelas 4.2 e 4.3 os testes ADF e PP para presença de raiz unitária. Como na presença de mudança estrutural esses testes tendem a não rejeitar a hipótese de não

estacionaridade, adopta-se a seguinte regra de decisão para determinação da ordem de integração das séries: i) caso haja consenso entre os resultados dos testes opta-se por este consenso; ii) havendo divergência entre os testes inspeccionam-se os dados para identificar as datas de ocorrência dessas quebras, e realizam-se os testes para os sub-períodos encontrados.

Tabela 4.2: Tese de raiz unitária baseado em ADF

Teste ADF*				
Variável	Lag	t calculado	Valores Críticos	
			1%	5%
Log[Basemot]	1	-2,4269	-4,0213	-3,4405
Inflação	0	-6,9243	-3,4746	-2,8809
Log[J.Activo]	0	-1,8927	-3,4746	-2,8809
Log[J.Passivo]	0	-1,8101	-3,4746	-2,8809

* O número de Lag's foi determinado segundo o critério SC
Fonte: Elaboração própria

Tabela 4.3: Tese de raiz unitária baseado no procedimento PP

Teste PP*				
Variável	Bandwidth	t calculado	Valores Críticos	
			1%	5%
Log[Basemot]	3	-2,6453	-4,0208	-3,4403
Inflação	11	-6,2206	-3,4746	-2,8809
Log[J.Activo]	6	-1,9707	-3,4746	-2,8809
Log[J.Passivo]	6	-2,1796	-3,4746	-2,8809

*Os Bandwidth obtiveram-se pelo critério de Newey-West
Fonte: Elaboração própria

Para os dois testes realizados, a série da taxa de inflação rejeitou a hipótese de existência de uma raiz unitária para os dois níveis de significância, ou seja, não há nenhuma evidência que justifique a ignorância da estacionaridade dessa série, em níveis. Por conseguinte, a série da inflação é $I(0)$. O mesmo não acontece com as séries dos logaritmos da base monetária e taxas de juro, em níveis, onde embora não haja discordância nos dois níveis significância, a hipótese de existência de raízes unitárias não pode ser rejeitada. Realizando os mesmos testes para as variáveis em primeiras diferenças nota-se que elas são estacionárias, conforme sugerem seus gráficos na Figura 4.2. Logo elas são $I(1)$.

4.2 Mercado monetário: modelo VAR

O presente trabalho adoptou os critérios AIC e SC para a escolha da ordem do *lag* no modelo. De acordo com o critério AIC o modelo mais adequado é o de 8 *lags* e, o critério SC aponta para o modelo com apenas uma desfasagem. Pelo facto de o modelo com 8 desfasagem gerar o problema da proliferação de parâmetros a serem estimados e, por conseguinte, a perda de graus de liberdade optou-se pela estimação do modelo VAR com um *lag*.

Tabela 4.4: Critério de selecção do *Lag*

<i>Lag</i>	AIC	SC
0	-0,0031	0,3300
1	-5,4569	-4,7980*
2	-5,3802	-4,3810
3	-5,3789	-4,0467
4	-5,5317	-3,8665
5	-5,5386	-3,5403
6	-5,5393	-3,2080
7	-5,4728	-2,8084
8	-5,6242*	-2,6267

* Ordem seleccionada pelo critério

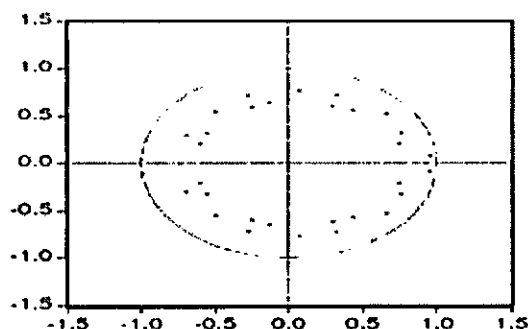
Fonte: Elaboração própria

Mas estimação do modelo com apenas uma desfasagem mostrou, para além do problema da autocorrelação, uma instabilidade, isto é, uma das raízes características ultrapassava ligeiramente a unidade, encontrando-se deste modo fora do círculo unitário. Assim, optou-se pela estimação do modelo com uma desfasagem adicional, a oitava, visto que o modelo com dois *lag*'s é o que maximiza o logaritmo de verosimilhança e minimiza os critérios AIC e SC ³⁵.

³⁵ A possibilidade de inclusão de um *lag* superior a oito foi descartada pelo facto de este ser o máximo *lag* escolhido dentre os dois critérios usados. E dentre as várias tentativas de inclusão de uma desfasagem diferente da oitava ou de uma outra adicional, o modelo mostrou-se bastante sensível à verificação de algumas suposições como a de ausência de autocorrelação serial.

Outro aspecto que merece realce é o melhoramento da estabilidade do modelo com a inclusão de um termo de tendência linear no modelo, bem como duas *dummies* representativas de cinco *outliers* detectados nas séries da inflação e da taxa de juro passiva³⁶.

Figura 4.3: Raízes inversas do polinómio característico



O teste de autocorrelação baseado no procedimento *LM* é apresentado na tabela 4.5:

Tabela 4.5: Teste de autocorrelação serial dos resíduos LM

<i>Lag's</i>	LM-Stat	Prob
1	34,39301	0,00480
2	13,09477	0,66580
3	23,60478	0,09850
4	26,85130	0,04320
5	20,95924	0,18010
6	35,57350	0,00330
7	41,43119	0,00050
8	10,56169	0,83570
9	21,52348	0,15930
10	11,50230	0,77750

Fonte: Elaboração própria

Embora o modelo estimado com duas defasagens tenha amenizado o problema da autocorrelação, ela prevalece para o caso do primeiro, o sexto e sétimo *lag's* que tanto a 1% como a 5% de

³⁶ Essas observações atípicas, as que se afastam substancialmente dos intervalos estabelecidos pelo primeiro e terceiro quartis $Q_1 / Q_3 \pm 1.5 \times (Q_3 - Q_1)$, e que correspondem a taxas de inflação registadas no mês de Fevereiro de 1996 e 1997, de 7,52 e 6,33 respectivamente. E taxas de juro passivas nos meses de Fevereiro, Maio e Junho de 1997, cujas taxas foram respectivamente de 39,7%, 35,1% e 32,3%. Representadas por unidades nas variáveis criadas.

significância não mostram evidências para a não rejeição da hipótese de ausência de autocorrelação serial dos resíduos. A presença de autocorrelação dos resíduos, segundo Gujarati (2006), faz com que os estimadores deixem de ser eficientes (variância mínima). Apesar disso, continuam a ser lineares e não enviesados. O facto de se estimar via OLS desconsiderando autocorrelação faz com que se subestime a variância dos resíduos. E além disso, a autocorrelação fará com que os testes de significância t e F não sejam mais válidos.

A estatística do teste de Jacque-Berra resultou em um valor $\chi^2 = 4312,807$ para 8 graus de liberdade e um valor de $p = 0,00$ que sugere a rejeição da hipótese nula, ou seja, não mostra evidências para a sustentabilidade da suposição de normalidade multivariada. Por outro lado, o teste de heteroscedasticidade de White, com $\chi^2 = 305,1252$ para 200 graus de liberdade leva também à conclusão de rejeição da hipótese nula. Mas como se viu na secção 2.4.1, inferências a partir de um modelo VAR são moderadamente robustas à ambas suposições.

Entretanto, as inferências obtidas a partir das funções de resposta-impulso e decomposição de variância mostram-se, em parte, muito sensíveis à ordenação das variáveis no modelo VAR, diante disso, o presente trabalho valeu-se do Teste de Causalidade de Granger/Block Exogeneity Wald Test para definir uma ordenação *estatisticamente consistente*, como segue:

Tabela 4.6: Teste de Causalidade de Granger/Block Exogeneity Wald Test

	Variável dependente							
	Log[Basemonet]		Log[JuroActivo]		Log[JuroPassivo]		Inflação	
	χ^2	Prob.	χ^2	Prob.	χ^2	Prob.	χ^2	Prob.
Log[Basemonet]			7,068	0,029	1,212	0,546	14,829	0,001
Log[JuroActivo]	1,540	0,463			14,410	0,001	0,809	0,667
Log[JuroPassivo]	2,300	0,317	1,904	0,386			1,432	0,489
Inflação	2,678	0,262	5,953	0,051	0,413	0,814		
Total	13,905	0,031	16,307	0,012	22,497	0,001	19,644	0,003

Fonte: Elaboração própria

Partindo dos valores da estatística χ^2 ordena-se as variáveis da *mais exógena* (menor valor da estatística) à *mais endógena* (maior valor). Com esse procedimento, obtém-se a seguinte ordenação: base monetária (13,905), taxa de juro activa (16,307), inflação (19,644) e taxa de juro passiva (

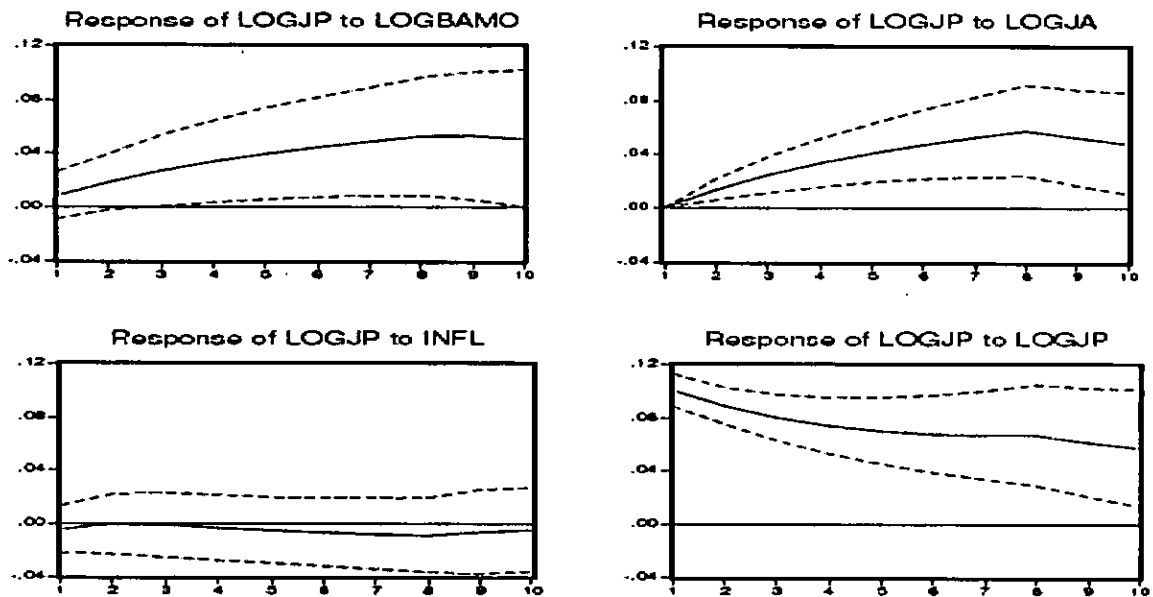
22,947). A base monetária não responde contemporaneamente às inovações nas demais variáveis, dado que ela é segundo a ordenação a variável *mais exógena* do modelo. O que realmente faz sentido já que a emissão de moeda é total e exclusiva responsabilidade do banco central, sendo assim, a base monetária apenas sofre variações exógenas. A variável de interesse, taxa de juro passiva é afectada contemporaneamente por todas as demais variáveis. Ou por outra, as inovações em qualquer das variáveis afectam a taxa de juro passiva no mesmo período em que ocorrem os choques.

É importante destacar que de facto a taxa de juro passiva depende, em parte, do nível da taxa de inflação vigente na economia, ou mais concretamente da sua expectativa futura; do montante da carteira de empréstimos de que o banco comercial dispõe, que traduz a quantidade de retornos a partir da taxa de juro activa, já que isso mostra a expectativa do nível quitação das suas dívidas a partir desses retornos; e também do montante de moeda em circulação no mercado, pois o montante dos activos monetários detidos pelos agentes económicos ou população em geral é que determinará as diversas formas de aplicação dessa riqueza de acordo com as suas expectativas.

A seguir são apresentados os resultados das funções impulso-resposta. A linha contínua representa a função impulso-resposta, enquanto as linhas pontilhadas representam o intervalo de ± 2 desvios-padrão. Os desvios-padrão foram gerados por meio de simulação de Monte Carlo com 1000 repetições³⁷. A intuição dessa ferramenta é que qualquer resposta que leve o intervalo de confiança de ± 2 desvios-padrão a se afastar do valor zero representa uma resposta significativa da variável a alterações na outra variável independente contemplada (Resende, 2006).

³⁷ O objecto de interesse, nas experiências de Monte Carlo, é usualmente um estimador ou uma estatística de teste com propriedades de amostra finita desconhecidas (embora se possa saber sobre as suas propriedades assintóticas). Embora possam existir processos analíticos para estudar a distribuição da amostra finita dos estimadores, é muitas vezes mais fácil executar experiências de Monte Carlo (Johnston & Dinardo, 2001).

Figura 4.4:Resposta-impulso nas taxa de juro passiva



Fonte: Output Eviews

Os choques nas variáveis, taxas de inflação, base monetária e taxa de juro activa, não só afectam directamente a elas mesmas, como também são transmitidos para a taxa de juro das operações passivas e, em geral, a todas variáveis através da estrutura dinâmica do *lag*. As respostas da taxa de juro passiva a impulsos nessas variáveis, mostram o comportamento dos choques em cada uma das inovações nos valores correntes e futuros da taxa de juro passiva.

A taxa de juro passiva responde significativamente a variações em si própria e aos choques na taxa de juro activa durante aproximadamente dez períodos. Os choques na base monetária mostram uma reacção positiva e significativa sobre a taxa de juro passiva durante dois períodos, já a taxa de inflação não mostra reacções significativas sobre a taxa de juro passiva, pelo menos, no intervalo de dez períodos de previsão.

Note-se que o limite inferior das respostas da taxa de juro aos choques na base monetária encontra-se muito próximo do eixo, o que significa que a hipótese de que as inovações na base monetária não têm, praticamente, nenhum efeito na taxa de juro passiva não pode ser rejeitada.

71

Assim, as funções de resposta-impulso mostram, inicialmente, uma influência positiva da base monetária e da taxa de juro activa na determinação das taxas de juro passiva, e um resultado

inconclusivo quanto ao efeito da taxa de inflação. Para complementar a análise é ilustrada, na tabela 4.7, a decomposição de variância da taxa de juro passiva:

Tabela 4.7: Decomposição de variância da taxa de juro passiva

Decomposição de variância de Log[JuroPassivo] em (%):					
Período	S.E.	Log[Basemonet]	Inflação	Log[J.Passivo]	Log[J.Activo]
1	0.037337	0,706	0,097	88,137	11,060
2	0.052662	2,220	0,058	81,935	15,788
3	0.065061	4,315	0,041	74,965	20,680
4	0.076225	6,690	0,039	67,901	25,371
5	0.086735	9,100	0,056	61,198	29,647
6	0.096875	11,391	0,088	55,098	33,423
7	0.106811	13,486	0,129	49,688	36,697
8	0.116659	15,357	0,174	44,959	39,510
9	0.125451	17,053	0,178	41,750	41,019
10	0.133599	18,391	0,171	39,515	41,923

Fonte: Elaboração própria

A análise de decomposição da variância dos erros de previsão permite quantificar o peso de cada variável independente na determinação da variância do erro de previsão da variável dependente ao longo do tempo. A decomposição de variância indica que grande parte da variação da taxa de juro passiva se dá, inicialmente, em função dos choques em si mesma, seguido dos choques na taxa de juro activa (11,06%), base monetária (0,706%) e inflação (0,097%). No primeiro período 88,137% da variação na taxa de juro passiva é resultado de choques nela própria, o que sugere uma alta *inércia* na taxa de juro passiva.

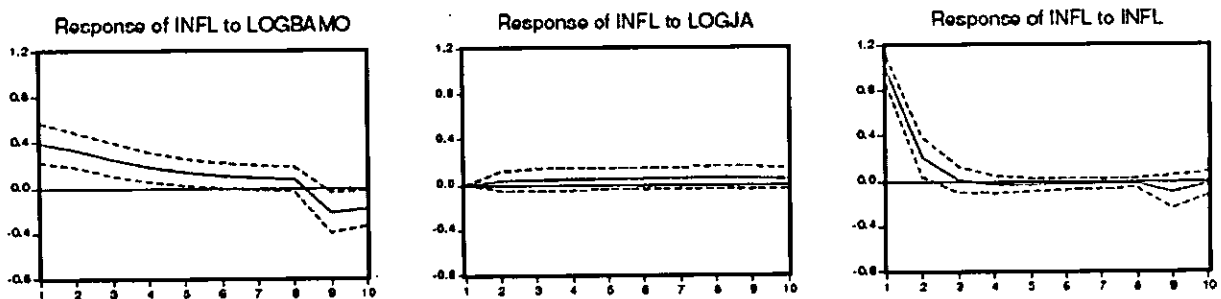
Porém, esta participação se reduz gradativamente até chegar a 44,959% no oitavo período, onde ainda nesse período a base monetária e a taxa de juro activa passam a ter uma influência mais significativa na determinação da taxa de juro passiva, especialmente esta última variável que a partir do décimo período passa a explicar grande parte da variação da taxa passiva, cerca de 41,9%, que supera a variação explicada pela própria taxa passiva (39,515%).

Os resultados alcançados pela decomposição de variância corroboram, em parte, com as constatações obtidas pelas funções de resposta-impulso, dado que apenas a base monetária e a taxa

de juro activa têm uma influência positiva e significativa na determinação da taxa de juro passiva, enquanto que a taxa de inflação não mostra qualquer influência na sua determinação³⁸.

A segunda variável de interesse, segundo o critério de ordenamento das variáveis, das *mais exógenas* às *mais endógenas*, é a taxa de inflação. E, por conseguinte, uma análise de interesse decorre do facto de que a inflação é sem dúvida primordialmente um fenómeno monetário no sentido de que ela não poderia continuar sem que houvesse crescimento monetário (Sachs & Larrain, 2000).

Figura 4.5: Resposta-impulso na taxa de inflação



Fonte: Output Eviews

A taxa de inflação responde significativamente aos choques na base monetária durante aproximadamente seis períodos, e também de forma significativa à variações nela mesma durante dois períodos. Enquanto, a taxa de juro activa não mostra nenhuma influência significativa na determinação da taxa de inflação.

³⁸ Corroboram em parte já que as funções de resposta a impulso não evidenciam claramente a primazia das taxas de juro activas sobre a base monetária na determinação da taxa passiva.

Tabela 4.8: Decomposição de variância da taxa de Inflação

Decomposição de Variância da taxa de Inflação (%)				
Período	S.E.	Log[JA]	Log[B.monet]	Inflação
1	0,0496	0,070	13,787	86,144
2	0,0613	0,072	20,788	78,932
3	0,0672	0,149	24,411	75,144
4	0,0709	0,299	26,178	73,224
5	0,0738	0,515	27,084	72,105
6	0,0763	0,791	27,580	71,312
7	0,0788	1,122	27,868	70,646
8	0,0813	1,504	28,044	70,021
9	0,0843	1,912	29,910	67,508
10	0,0875	2,165	31,284	65,701

Fonte: Elaboração própria

A decomposição da variância do erro de previsão da taxa de inflação mostra que até o último período de previsão, grande parte da variação da inflação é devida aos choques nela mesma. Nesse intervalo, nota-se que a taxa de inflação explica pelo menos cerca de 65,7% da sua variância. Contudo, os choques na base monetária têm um grande contributo para determinação da variância do erro de previsão da taxa de inflação, chegando a explicar cerca de 31,3% da sua variação a partir do décimo período de previsão.

Na relação entre a base monetária e a taxa de inflação, um aspecto precisa ser clarificado. Pois, embora a base monetária tenha uma certa influência na explicação da variância da inflação, essa relação pode ser apenas uma “coincidência” dado que os monetaristas, segundo Sachs e Larrain (2000), argumentam que a moeda explica a maior parte do que está acontecendo à inflação no longo prazo. No curto prazo, outros distúrbios além das variações no *stock* monetário afectam a inflação e, por outro lado, variações no *stock* monetário têm realmente efeitos reais. E mesmo se os distúrbios forem puramente monetários, ainda levará um certo tempo para que eles se reflectam totalmente na inflação e somente na inflação.

CAPÍTULO V

5. Conclusões

Neste trabalho fez-se uma análise das características dinâmicas da tendência do mercado monetário moçambicano com recurso aos modelos econométricos de vectores autoregressivos. A base teórica que fundamenta a relação de longo prazo entre a base monetária, taxas de juro e inflação, conduziu à especificação de um modelo VAR, na forma reduzida, com dois desfasamentos, duas dummies e uma tendência linear.

Embora a natureza do fenómeno estudado tenha sugerido como ponto de partida, segundo alguns fundamentos da teoria económica, uma análise detalhada de alguns aspectos inerentes à teoria de cointegração, não se procedeu a nenhum teste de existência de características cointegrantes, dado que os testes de raízes unitárias empregues para verificar a ordem de integração das variáveis mostraram a presença de estacionaridade na série de inflação e raízes unitárias nas séries da base monetária e taxas de juro. Por conseguinte, o facto de se estar perante uma mistura de variáveis $I(0)$ e $I(1)$, inviabiliza o desenvolvimento dos testes de cointegração. E, assim sendo, asseguram-se as bases para a rejeição da hipótese previamente estabelecida no capítulo I.

A análise das inter-relações constituídas no modelo da forma reduzida, evidencia que a relação que se estabelece entre as variáveis é bastante complexa. A taxa de juro passiva desempenha um papel importante na determinação da trajectória da política monetária, dado que ela é segundo o critério do Teste de Causalidade de Granger a variável mais endógena no modelo VAR estimado. Essa taxa é contemporaneamente afectada pelas inovações que ocorrem nas demais variáveis no mesmo período que ocorrem os choques. As funções resposta-impulso e decomposição de variância da taxa de juro passiva apontam para a conclusão de que apenas a taxa de juro activa e a base monetária contribuem significativamente na determinação da trajectória da taxa de juro passiva. Todavia, a expectativa que se estabelece pela teoria económica inerente a relação de longo prazo do crescimento monetário e inflação contrasta com os resultados observados.

5.1 Recomendações

Os resultados obtidos pela análise da relação proposta, mostram uma certa coerência na necessidade de tentar influenciar as taxas de juro de mercado como foco da política monetária do banco central. É recomendável que em pesquisas futuras seja analisada essa relação substituindo as taxas de juro de mercado pelas taxas de juro de intervenção do Banco de Moçambique (FPC, FPD, FUH), assim como a análise da relação entre essas taxas de juro e as taxas de mercado.

5.2 Limitações do estudo

As limitações encontradas durante a realização do estudo foram:

- Dificuldades no acesso as fontes bibliográficas primárias o que levou a que se recorresse constantemente à fontes bibliográficas secundárias;
- Falta de estudos similares em Moçambique que possibilitassem a comparação dos resultados.

Bibliografia

- Banco de Moçambique[BM] (2003). *Relatório Anual: CDI – Banco de Moçambique*. Maputo
- BM (2006). *Relatório Anual: CDI – Banco de Moçambique*. Maputo
- BM (2007). *Relatório Anual: CDI – Banco de Moçambique*. Maputo
- Caiado, J. (2003). Premio Nobel de Economia: *o desenvolvimento de métodos estatísticos para a análise de séries económicas e financeiras*. pp 6-14.
- Camuri, Paulo André (2005). *Dívida pública, Política Fiscal e Restrição Externa no Brasil: 1994 – 2004*. Tese de mestrado em Economia. Faculdade de Ciências Económicas – Universidade Federal de Minas Gerais. Belo Horizonte. 97 pp.
- Chari, V. V. (1999). Nobel Laureate Robert E. Lucas, Jr.: Architect of Modern Macroeconomics. *Federal Reserve Bank of Minneapolis Quarterly review*. 2: 2 – 12.
- Cleto, Carlos Ilton (2001). Taxa de Juro: de onde vem e para onde nos leva. *Fae Business*. 1: 1 – 3.
- Dornbush, Rudiger & Fisher, Stanley (1982). *Macroeconomia* (Tradução). 2ª edição, McGraw-Hill. São Paulo.
- Engle, Robert F. & Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*. 2: 251-276.
- Fedderke, Johannes (2003). *Applied Time Series Analysis*. University of Cape Town.
- Forte, Sérgio Henrique A. C. (2004). *Manual De Elaboração de Tese, Dissertação E Monografia*. 4ª edição, Universidade de Fortaleza. Fortaleza.
- Greene, William H. (1997). *Econometrics analysis*. 3rd edition, McGraw-Hill. New York
- Griffiths, William E., Hill, R. Carter & Judge, George G. (1993). *Learning and practicing econometrics*. McGraw-Hill. New York..
- Gujarati, D.N (2006). *Econometria básica* (Tradução). 4ª edição, Campus. São Paulo.

- Hendry, D. F. & Juselius, K. (2000). Explaining cointegration analysis: Part II. *Energy Journal*. pp.33.
- Johansen, Soren (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.
- Johnston, Jack & Dinardo, John (2001). *Métodos econométricos* (Tradução). 4ª edição, McGraw-Hill. Lisboa.
- Juselius, Katarina (2003). *The cointegrated VAR Model: Econometric Methodology and Macroeconomic Application*. pp. 257.
- Keating, John W. (1992). Structural Approaches to Vector Autoregressions. *American Statistical Association*. pp. 57.
- Lakatos, Eva Maria & Marconi Marina de Andrade (1991). *Fundamentos de Metodologia Científica*. 3ª edição, Atlas. São Paulo.
- Leite, José Alfredo A. (2000). *Macroeconomia*. Atlas. São Paulo.
- Maleiane, Adriano Afonso (2001). Mecanismos de transmissão da política monetária. *Documento de estudo* 8: 17-35.
- Quantitative Micro Software (2002). *Eviews 4 User's Guide*. LCC. USA
- Resende, João Luís (2006). *Metas de inflação, câmbio flexível e autonomia da política monetária*. Tese de mestrado em Economia. Faculdade de Ciências Económicas – Universidade Federal de Minas Gerais. Belo Horizonte. pp. 135
- Runkle, David E. (1987). Vector Autoregression and Reality. *Research Department Staff Report*. Federal Reserve Bank of Minneapolis 107: 1-20
- Sachs, Jeffrey D. & B., Larrain Felipe (2000). *Macroeconomia* (Tradução). McGraw-Hill. São Paulo.

Silva, José Cláudio Ferreira Da (1999). *Modelos de Análise Macroeconómica: um curso completo de Macroeconomia*. Campus. Rio de Janeiro.

Simkins, Scott (1999). *Measurement and Theory in Macroeconomics*. Acedido a 21 de Agosto de 2008, no *Web site* da: North Carolina A&T State University: <http://www.ncat.edu>.

Stock, James H. & Watson, Mark W. (2001). Vector Autoregressions. *Journal of Economic Perspectives*. 15: 101 – 106.

Watson, Mark W. 1994. Vector Autoregressions and Cointegration. *Handbook of Econometrics, volume IV*. Robert Engle and Daniel McFadden, editors. Amsterdam: Elsevier. 2844-2915.

Sargent, Thomas J. & Sims, Christopher (1977). Business Cycle Modeling Without Pretending to Have Too much a Priori Economic Theory. *American Statistical Association*. 55: 17 – 89.

Tesfatsion, Leigh (2007). *Notes on the Lucas Critique, Time Inconsistency, and Related Issues*. Consultado em 25 de Agosto de 2008, no *Web site*: <http://www.feweb.vu.nl/econometriclinks/>.

Glossário

1. **ACTIVO** – Algo que pertence a um indivíduo ou sociedade; instrumento financeiro; bem económico; equipamento; rubrica do balanço.
2. **ACTIVO FINANCEIRO** – Valor mobiliário que, geralmente, confere ao seu titular um crédito ou um direito de propriedade sobre a entidade emissora, tais como as obrigações e as acções.
3. **AMOSTRA** - Qualquer subconjunto de elementos de uma população.
4. **AUTOCORRELAÇÃO** - a correlação interna entre membros de uma série de observações ordenadas no tempo ou espaço.
5. **AUTOREGRESSÃO** - a geração de uma série de observações pela qual o valor de cada uma é parcialmente dependente nos valores daquelas que imediatamente a precederam.
6. **BILHETES DE TESOIRO** – são títulos de dívida pública emitidos a desconto por prazos até um ano.
7. **CARACTERÍSTICA DE UMA MATRIZ** – é o número máximo de colunas (ou linhas) linearmente independentes.
8. **CURTOSE** - um termo usado para descrever a extensão pela qual a curva de frequência unimodal é “pontaguda”; ou seja, a extensão relativa do declive na vizinhança da moda. O termo foi introduzido por Karl Pearson em 1906 e, ele propôs a razão de momentos μ^4/μ_2^2 como uma medida de curtose.
9. **DEPÓSITO**- Operação bancária em que os bancos actuam como captadores de fundos, assumindo a qualidade de devedores perante os depositantes a quem pagam um certo juro (operações passivas). Esses fundos são, por sua vez, geradores de crédito (operações activas).Consoante o grau de exigibilidade e o tempo de utilização, podemos distinguir as seguintes modalidades: depósitos à ordem, depósitos a prazo e depósitos com aviso prévio
10. **DEPÓSITO À ORDEM** — Entrega de dinheiro a uma instituição para incremento de uma conta de disponibilidade imediata. Neste tipo de depósito, o depositante tem o direito, salvo legislação especial, de exigir o reembolso das importâncias entregues ao Banco, em qualquer momento.
11. **DISTRIBUIÇÃO ASSIMPTÓTICA** - a forma limite de uma distribuição de frequência ou probabilidade dependente de um parâmetro, tal como tamanho de amostra ou tempo, quando o parâmetro tende ao limite, geralmente infinito.

12. EQUAÇÃO CARACTERÍSTICA - equação característica da matriz A é um polinómio de grau k em λ , desconhecido, que pode ser resolvido para determinar as k raízes. Estes λ são os valores próprios de A , também conhecidos como as raízes latentes ou raízes características.

13. ESTATÍSTICA DE DURBIN – WATSON - uma estatística de teste da independência dos erros em uma regressão de mínimos quadrados contra uma alternativa de correlação serial.

14. ESTIMADOR - um estimador é uma regra ou método de estimar uma constante de uma população. É geralmente expressado como uma função de valores de uma amostra e, conseqüentemente, uma variável cuja distribuição é de grande importância no cálculo da confiabilidade da estimativa a ser calculada a partir dele.

15. ESTIMADOR CONSISTENTE - um estimador é dito consistente se ele converge em probabilidade, quando o tamanho de amostra cresce, ao parâmetro do qual é estimador.

16. ESTIMADOR EFICIENTE - um estimador é dito eficiente se ele é de variância mínima.

17. ESTIMATIVA - de forma restrita, uma estimativa é um valor particular obtido para um estimador em determinada amostra.

18. ESTIMAÇÃO - estimação está relacionada com a inferência sobre o valor numérico de valores desconhecidos da população baseados em dados incompletos como uma amostra. Se um único valor for calculado para cada parâmetro desconhecido, o processo é chamado estimação pontual. O processo de estimação por intervalos consiste em calcular um intervalo de forma que, com certo grau de certeza, ele contenha o parâmetro de interesse.

19. FUNÇÃO DE AUTOCORRELAÇÃO - a função de autocorrelação de um processo estocástico estacionário é a autocovariância dividida pela variância.

20. FUNÇÃO DE AUTOCOVARIÂNCIA - para qualquer processo estocástico estacionário a função $\gamma(k) = C(x_t, x_{t+k})$, onde C representa a covariância dos termos entre parêntese é conhecida como função de autocovariância.

21. FUNÇÃO DE DENSIDADE CONDICIONAL - se um conjunto de variáveis $x_1, \dots, x_p, x_{p+1}, \dots, x_q$ têm densidade conjunta $f(x_1, \dots, x_p, x_{p+1}, \dots, x_q)$, a sub-densidade obtida mantendo constantes algumas das variáveis é chamada de condicional; assim, a densidade de x_1, \dots, x_p para x_{p+1}, \dots, x_q fixos é comumente representada por $f(x_1, \dots, x_p | x_{p+1}, \dots, x_q)$.

22. FUNÇÃO DE VEROSSIMILHANÇA - se a função de distribuição de variáveis aleatórias x_1, \dots, x_n , que dependem dos parâmetros $\theta_1, \dots, \theta_k$, é representada como $dF = f(x_1, \dots, x_n; \theta_1, \dots, \theta_k) dx_1 \dots dx_n$, a função $f(x_1, \dots, x_n; \theta_1, \dots, \theta_k)$ considerada como função dos θ 's para valores fixos dos x 's é chamada função de verossimilhança.

23. HOMOCEDESTICIDADE - Para muitas variáveis aleatórias com distribuição conjunta dada, se a variância de uma variável é a mesma para todas as outras, a distribuição é chamada homocedástica

24. **INFLAÇÃO** - Ascensão continuada no nível geral de preços. A consequência é uma diminuição no poder aquisitivo (do inglês *purchasing power*). Há diversos tipos de inflação, porém existe indicação de que a mais nociva é a de preços.

25. **INTERVALO DE CONFIANÇA** - se for possível definir duas estatísticas t_1 e t_2 (funções de valores da amostra somente) de forma que, sendo θ o parâmetro a estimar, $Pr(t_1 \leq \theta \leq t_2) = \alpha$, onde α é o coeficiente de confiança, o intervalo entre t_1 e t_2 é chamado intervalo de confiança.

26. **MATRIZ ADJUNTA** - é uma matriz transposta formada à custa dos elementos de uma matriz A através de um processo de eliminação das linhas e colunas de cada elemento de A , e escrevendo-se o elemento restante afectado do sinal positivo ou negativo.

27. **MATRIZ COMPANHEIRA** - colocando c_1, \dots, c_n na n -ésima linha, e colocando $n - 1$ 1's na diagonal I. a seguir $\det(A - \lambda I) = \pm(c_1 + c_2\lambda + c_3\lambda^2 + \dots)$

28. **MATRIZ DEFINIDA POSITIVA** - seja $q = b'Ab$, onde q é um escalar e b um vector $k \times 1$ não nulo, portanto, a matriz A diz-se definida positiva se q for estritamente positivo para qualquer vector b não nulo.

29. **MATRIZ DIAGONAL** - é como a matriz identidade, na medida em que fora da diagonal todos os elementos são zero, mas agora os elementos da diagonal principal são escalares, sendo pelo menos um diferente de zero.

30. **MATRIZ IDENTIDADE (I)** - é uma matriz de dimensão $n \times n$ com uns na diagonal principal e zeros fora da diagonal.

31. **MATRIZ INVERSA** - se A for uma matriz de dimensão $n \times n$, e existir uma matriz B , com a mesma dimensão, tal que $AB = I$, então B será a matriz inversa de A e designa-se por A^{-1} .

32. **MATRIZ ORTOGONAL** - matriz cuja inversa é igual a sua transposta.

33. **MATRIZ SIMÉTRICA** - é uma matriz que satisfaz a condição: $A = A'$.

34. **MATRIZ TRANSPOSTA** - a transposta de uma matriz A designa-se por A' . A primeira linha de A torna-se a primeira coluna da transposta, a segunda linha de A torna-se a segunda coluna da transposta, e por aí fora.

35. **MÉTODO DE MÍNIMO QUADRADOS ORDINÁRIOS [MQO/OLS]** - uma técnica de estimação pela qual as quantidades a estimar são determinadas minimizando-se uma certa forma quadrática das observações e aquelas quantidades.

36. **NORMALIDADE ASSIMPTÓTICA** - uma distribuição dependente de um parâmetro n , geralmente um tamanho de amostra, é definida como assintoticamente normal se, quando n tende ao infinito, então, a distribuição tende para a forma Normal.

37. **PREVISÃO** - em geral, previsão é o processo de determinar a magnitude de variáveis estatísticas em alguns pontos de tempo futuros.

38. **RESÍDUO** - um termo geral que denota a quantidade restante depois de alguma outra quantidade ter sido subtraída. Ocorre em uma variedade de contextos específicos.

39. **RETORNO** - dinheiro ganho em um investimento, muitas vezes expresso como uma porcentagem.

40. **TAXA DE JURO** - Quem coloca o seu dinheiro no banco, espera receber uma remuneração, pois está a disponibilizar recursos que são seus para serem utilizados por outras pessoas ou empresas.

41. **TAXA DE JURO ACTIVA** - Taxa de juro a cobrar pelos bancos pelo empréstimo de dinheiro. Designa-se por activa porque se inscreve nas contas do Activo dos bancos.

42. **TAXA DE JURO PASSIVA** - Taxa de juro que remunera os depósitos bancários. Designa-se por passiva porque se inscreve nas contas do Passivo dos bancos, visto representar uma responsabilidade.

ANEXOS

Tabela A1: Estatísticas descritivas

	LOGJP	LOGJA	LOGBAMO	INFL
Mean	-2.125740	-1.304233	8.614001	0.806636
Median	-2.194092	-1.387108	8.684493	0.551182
Maximum	-0.933946	-0.820981	9.788186	7.515977
Minimum	-2.664075	-1.620193	7.624655	-3.421098
Std. Dev.	0.380884	0.224560	0.682268	1.541350
Skewness	1.019247	0.684450	0.099316	1.099863
Kurtosis	3.831874	2.287889	1.549712	5.877965
Jarque-Bera Probability	30.29669 0.000000	14.88119 0.000587	13.39243 0.001236	82.00922 0.000000
Sum	-318.8610	-195.6350	1292.100	120.9954
Sum Sq. Dev.	21.61581	7.513653	69.35796	353.9882
Observations	150	150	150	150

Tabela A2: Estimação do modelo VAR

	LOGJP	LOGJA	LOGBAMO	INFL
LOGJP(-1)	0.834233 (0.04582) [18.2079]	0.021674 (0.01692) [1.28133]	0.033038 (0.02249) [1.46915]	-0.545125 (0.48122) [-1.13281]
LOGJP(-8)	-0.026502 (0.04069) [-0.65124]	0.002599 (0.01502) [0.17298]	-0.000107 (0.01997) [-0.00537]	0.281702 (0.42742) [0.65908]
LOGJA(-1)	0.394076 (0.10849) [3.63244]	0.933824 (0.04005) [23.3149]	0.065843 (0.05325) [1.23654]	0.978810 (1.13945) [0.85902]
LOGJA(-8)	-0.273096 (0.09921) [-2.75275]	-0.046956 (0.03663) [-1.28202]	-0.034133 (0.04869) [-0.70098]	-0.200884 (1.04198) [-0.19279]
LOGBAMO(-1)	0.171628 (0.15649) [1.09876]	0.151316 (0.05777) [2.61914]	0.745269 (0.07681) [9.70330]	5.075123 (1.64357) [3.08787]
LOGBAMO(-8)	-0.064871 (0.14190) [-0.45715]	-0.088359 (0.05239) [-1.68658]	0.081122 (0.06965) [1.16473]	-5.250559 (1.49042) [-3.52288]
INFL(-1)	0.003600 (0.00815) [0.44195]	-0.000766 (0.00301) [-0.25474]	-0.006126 (0.00400) [-1.53215]	0.213274 (0.08556) [2.49256]
INFL(-8)	0.003198 (0.00621) [0.51531]	0.005451 (0.00229) [2.37954]	-0.002290 (0.00305) [-0.75197]	-0.081753 (0.06517) [-1.25442]
C	-1.037198 (1.18578) [-0.87469]	-0.564950 (0.43778) [-1.29049]	1.410891 (0.58200) [2.42421]	1.674889 (12.4543) [0.13448]
D1	-0.118690 (0.13652) [-0.86938]	0.058256 (0.05040) [1.15581]	0.139198 (0.06701) [2.07736]	6.509637 (1.43389) [4.53985]
D2	0.207495 (0.09021) [2.30016]	-0.099430 (0.03330) [-2.98551]	-0.072907 (0.04428) [-1.64664]	-2.103106 (0.94746) [-2.21972]
TEMPO	-0.001869 (0.00251) [-0.74388]	-0.001158 (0.00093) [-1.24885]	0.002946 (0.00123) [2.38905]	0.005616 (0.02638) [0.21285]

Análise residual

Figura A1: Representação gráfica

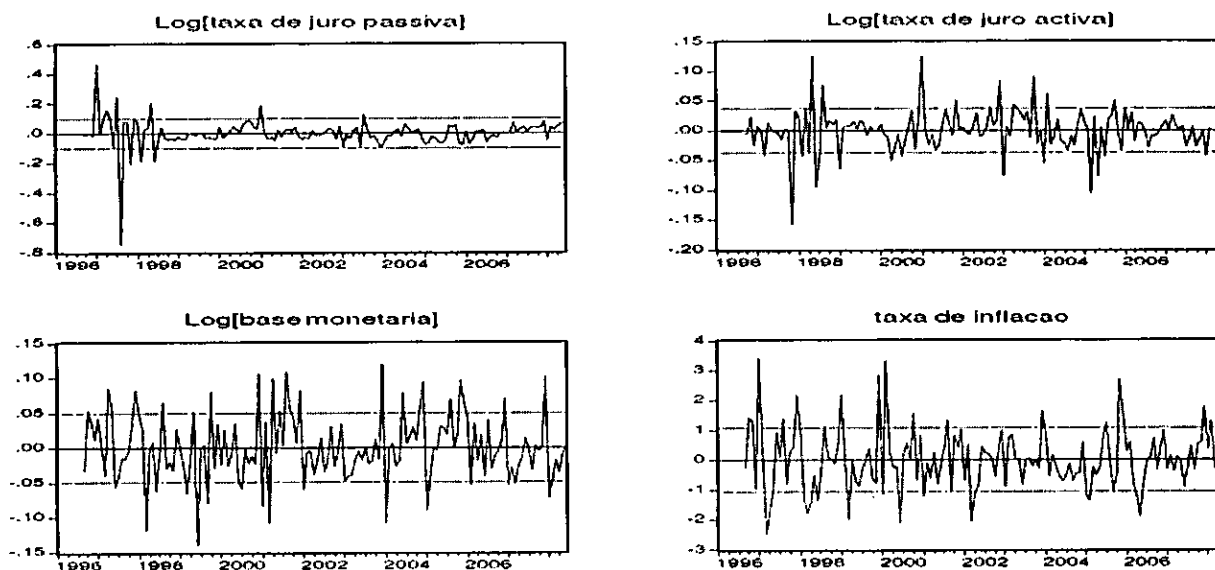


Tabela A3: Estatísticas descritivas

Mean	-2.33E-16	6.87E-16	2.31E-15	2.25E-14
Median	-9.72E-05	0.001562	-0.004210	-0.037043
Maximum	0.473210	0.126359	0.118678	3.400494
Minimum	-0.746913	-0.158042	-0.140151	-2.463549
Std. Dev.	0.097106	0.035851	0.047661	1.019904
Skewness	-2.311312	-0.206210	0.105809	0.622897
Kurtosis	30.08511	7.258439	3.333875	4.228626
Jarque-Bera	4466.917	108.3010	0.924508	18.11402
Probability	0.000000	0.000000	0.629862	0.000117
Sum	-3.31E-14	9.77E-14	3.27E-13	3.20E-12
Sum Sq. Dev.	1.329577	0.181224	0.320294	146.6688
Observations	142	142	142	142