

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA**

**CONDIÇÃO DE MARSHALL-LERNER E QUEBRA ESTRUTURAL NA
ECONOMIA BRASILEIRA**

Autor: Guilherme Valle Moura

**Porto Alegre
2005**

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA**

**CONDIÇÃO DE MARSHALL-LERNER E QUEBRA ESTRUTURAL NA
ECONOMIA BRASILEIRA**

Autor: Guilherme Valle Moura

Orientador: Prof. Sergio da Silva

**Dissertação submetida ao
Programa de Pós-Graduação em
Economia da Faculdade de
Ciências Econômicas da UFRGS
como requisito para a obtenção
do Grau de Mestre em Economia.**

Porto Alegre

2005

Sumário

Introdução	7
Capítulo 1 Taxa de Câmbio e Balança Comercial.	11
1.1 Abordagem das Elasticidades.....	11
1.2 Abordagem da Absorção.....	17
1.3 Curva J.....	20
Capítulo 2 Metodologia	25
2.1 Raízes Unitárias.....	25
2.2 Cointegração.....	38
2.2.1 Testes de Cointegração	42
A Metodologia de Engle e Granger.....	42
A Metodologia de Johansen.....	43
2.3 Processos VAR Cointegrados com Mudanças de Regime Markoviano.	44
2.3.1 Mudança de Regime no <i>Drift</i> e na Média de Equilíbrio do VECM.....	46
2.3.2 Representação Estável de Espaço-Estado.....	48
2.3.3 Processos MS-VAR Mais Gerais.	49
2.3.4 Representação VARMA de Processos MS-VAR Cointegrados.	50
2.3.5 Estimação	52
Capítulo 3 Resultados	56
3.1 Dados.....	56
3.2 Modelo Linear	58
3.3 Modelo com Mudança de Regime Markoviano.	60
Conclusão	69
Referências Bibliográficas.	71
Apêndice	84

Lista de Ilustrações

Figura 1.1 Um modelo de dois bens e dois países.....	14
Figura 3.2.1 Resposta de $\ln x_m$ à inovação de um desvio padrão em $\ln \text{real}$	60
Tabela 3.3.2 Estimativas do MSMH(2)-VECM(4)	62
Tabela 3.3.3 Data dos regimes	63
Figura 3.3.1 Probabilidades suavizadas do regime 1	63
Figura 3.3.2 Probabilidades suavizadas do regime 2.....	63
Figura 3.4 Funções impulso-resposta do modelo MSMH(2)-VECM(4)	68
Tabela 3.1.1 Estatísticas descritivas	84
Tabela 3.1.2 Teste ADF	84
Tabela 3.1.3 Teste de raiz unitária baseado em Perron (1997)	84
Tabela 3.1.4 Datas das quebras estruturais dadas pelo teste Perron (1997)	84
Tabela 3.2.1 Critérios AIC e BIC para seleção do número de defasagens	85
Tabela 3.2.2 Testes de cointegração	85
Tabela 3.3.1 Critérios de informação para a seleção do modelo	85

RESUMO

O presente trabalho tem como objetivo analisar empiricamente o desempenho da balança comercial brasileira em resposta a depreciações cambiais no período entre janeiro de 1990 e dezembro de 2003. A condição de Bickerdike-Robinson-Metzler, bem como a de Marshall-Lerner afirmam que existe uma relação positiva entre estas variáveis. Porém, os trabalhos empíricos sobre o assunto têm obtido resultados divergentes, principalmente no que se refere à resposta de curto prazo. Vários autores estimam uma relação negativa entre a balança comercial e o câmbio no curto prazo, confirmando a hipótese da curva J. Utilizamos a metodologia MS-VECM (*Markov-switching vector error correction model*) para capturar os vários choques e mudanças ocorridos na economia brasileira. Concluímos que nos períodos de maior volatilidade, a resposta da balança comercial é menor. Porém, as condições de Bickerdike-Robinson-Metzler e de Marshall-Lerner são válidas para a economia brasileira no período analisado, independentemente do regime em vigor.

Palavras-Chave: Balança comercial, condição de Marshall-Lerner, curva J, cointegração, mudança markoviana

ABSTRACT

This work examines empirically the trade balance response to exchange rate depreciations. Both Bickerdike-Robinson-Metzler and Marshall-Lerner conditions state the existence of a positive relation between these variables. Yet empirical evidence has been mixed, especially that of the short run response. Several authors find a negative relationship, which confirms a J curve hypothesis. We employ a Markov switching vector error correction model to track shocks and structural changes hitting the Brazilian economy. We find that periods of greater volatility are accompanied by smaller response of the trade balance. However the results of both regimes are consistent with Bickerdike-Robinson-Metzler and Marshall-Lerner conditions.

Keywords: Trade balance, Marshall-Lerner condition, J curve, cointegration, Markov switching

Introdução

A taxa de câmbio é um preço importante em uma economia aberta, pois afeta diretamente vários negócios, investimentos e decisões de política. Dessa forma, não é surpreendente que o estudo das taxas de câmbio tenha se tornado uma área de pesquisa tão prolífica nas últimas décadas. Essa linha de pesquisa cresceu tremendamente depois de Bretton Woods, quando as taxas de câmbio se tornaram altamente voláteis devido à adoção do regime de taxa de câmbio flutuante.

Uma das áreas de pesquisa da taxa de câmbio que recebeu muita atenção de vários pesquisadores é a relação entre taxa de câmbio e balança comercial. O modelo da balança comercial baseado nas elasticidades, desenvolvido por Bickerdike (1920), Robinson (1947) e Metzler (1948), mostra a existência teórica dessa relação. Vários estudos empíricos foram feitos para descobrir os efeitos de variações cambiais sobre a balança comercial, com intuito de fornecer informações importantes para os formuladores de política. Principalmente sobre a eficácia de políticas cambiais, como desvalorizações nominais, para ajustar a balança comercial (Krugman e Baldwin 1987, Greenwood 1984, Hilmarios 1989, Rose e Yellen 1989, Bahmani-Oskooee 1991, 2001, Mahdavi e Sohrabian 1993, Arize 1994, Wei 1999, Baharumshah 2001, Singh 2002, Onafowora 2003).

Espera-se que a depreciação (apreciação) nominal altere a taxa de câmbio real (Hilmarios 1989, Bahmani-Oskooee 2001), gerando um efeito direto sobre a balança comercial. Especificamente, Bahmani-Oskooee (2001) diz que, com o objetivo de aumentar a competitividade internacional e melhorar a balança comercial, um país pode realizar uma desvalorização ou permitir que seu câmbio se deprecie.

Porém, existem evidências empíricas que mostram a persistência no ajuste tanto de preços quanto de quantidades, na balança comercial, após variações cambiais (Junz e Rhomberg 1973, Krugman e Baldwin 1987, Meade 1988). Além disso, muitos acreditam na existência de rigidez no ajuste de preços e quantidades devido à existência de contratos de câmbio, defasagens no processo de tomada de decisão dos agentes, preços viscosos, persistência nos hábitos de consumo e histerese (Magee 1973, Junz e Rhomberg 1973, Gerlach 1989, Mansoorian 1998, Dixit 1994). Este fenômeno é conhecido na literatura como curva J, devido ao efeito

perverso no curto prazo e a subsequente melhora da balança comercial no longo prazo, o que faz com que o gráfico de sua evolução ao longo do tempo se assemelhe à letra J.

A circunstância sob a qual uma desvalorização cambial melhora a balança comercial é conhecida como condição de Marshall-Lerner. O fenômeno da curva J pode ocorrer em um modelo de dois países mesmo se a condição de Marshall-Lerner for válida, pois o aumento da balança comercial postulado por ela só irá se manifestar no novo equilíbrio de longo prazo, uma vez que ela é derivada de uma análise de estática comparativa. Em um modelo de dois países, onde o país estrangeiro é o resto do mundo, a balança comercial é dada por

$$BC = PX(Y^*, R) - EP^*Q(Y, R) \quad (1.1)$$

onde P é o nível de preço do país local, $X(\cdot)$ é a demanda por exportações, $Q(\cdot)$ é a demanda por importações, Y^* é a renda do resto do mundo, R é a taxa de câmbio real $R \equiv EP^*/P$, E é a taxa de câmbio nominal, P^* é o nível de preço do resto do mundo e Y é a renda local.

Os efeitos de uma depreciação sobre a balança comercial são, portanto, incertos, uma vez que esta desencadeia três processos:

1. A depreciação torna os bens domésticos mais baratos para o estrangeiro, aumentando assim o volume das exportações. Com o nível de preços constante, a receita em moeda local das exportações também aumenta, elevando então a balança comercial.

2. Uma depreciação aumenta o preço doméstico das importações, diminuindo assim o volume demandado. Tudo o mais constante, isto também aumenta o saldo da balança comercial.

3. A depreciação faz com que o país local tenha que pagar mais por cada unidade importada remanescente. Isto piora a balança comercial.

Os processos 1 e 2 são conhecidos como efeito volume, já o processo três é o efeito preço. Combinando os efeitos, a depreciação cambial não melhora a balança comercial em todas as situações. A melhora só ocorre se 1 e 2 superarem 3. Para que uma depreciação cambial melhore a balança comercial, precisamos que a derivada de (1.1) em relação a E seja positiva:

$$P \frac{\partial X}{\partial E} - EP^* \frac{\partial Q}{\partial E} - P^* Q > 0. \quad (1.2)$$

Segundo a condição de Marshall-Lerner, uma depreciação cambial só melhora a balança comercial se, e somente se, a soma das elasticidades da taxa de câmbio das demandas por exportações e importações for, em termos absolutos, maior do que um, de forma a compensar o aumento no preço das importações causado pela depreciação. Porém, nada impede que ao longo do processo de ajustamento a balança comercial deteriore graças, por exemplo, ao efeito volume ou preço.

Krugman e Baldwin (1987), Foray e McMillan (1999) e Bahmani-Oskooee e Kantipong (2001) encontram evidências de curva J, com uma fase inicial de redução na balança comercial após a depreciação cambial e subsequente aumento. Porém, Moffet (1989), Rose e Yellin (1989) e Leonard e Stockman (2001) não encontram evidências robustas de curva J. Já Backus, Kehoe e Kydland (1994), adotam uma abordagem de ciclos reais internacionais e encontram grande dependência do tipo de choque nos resultados da balança comercial. De acordo com esse estudo, a balança comercial é contra-cíclica e, geralmente, negativamente correlacionada com movimentos presentes e futuros dos termos de troca, porém é positivamente correlacionada com movimentos passados nos termos de troca, o que gera o que eles chamaram de curva S.

Portanto, apesar de vários estudos já realizados sobre a dinâmica da balança comercial após uma depreciação cambial, não existe um consenso na literatura do tema. As diversas alterações cambiais ocorridas no Brasil, bem como a abertura comercial iniciada nos anos 90, oferecem uma grande oportunidade para a análise dos impactos de variações cambiais sobre a balança comercial brasileira.

Este trabalho tem como objetivo investigar os efeitos que variações na taxa de câmbio real exercem sobre a balança comercial brasileira. Pretende-se utilizar a metodologia de cointegração com o intuito de captar os efeitos de curto e longo prazo da taxa de câmbio sobre a balança comercial como tem sido feito na literatura (Antonucci 2003, Arize 1994, Baharumshah 2001, Bahmani-Oskooee 1991, 1994, 2001, Brada 1997, Mahdavi e Sohrabian 1993, Onafowora 2003, Singh 2002, Wei 1999).

Optamos pela adoção de modelos com mudança de regime markoviano para capturar as várias alterações no regime de câmbio, o grande número de planos econômicos adotados por diferentes governos e os choques externos que afetaram a economia brasileira.

O trabalho está dividido da seguinte maneira. No primeiro capítulo faremos uma descrição da evolução teórica do estudo da relação entre balança comercial e câmbio. Mais especificamente, apresentaremos as abordagens das elasticidades para a balança comercial, bem como a da absorção e o caso da curva J. No segundo capítulo discutiremos a metodologia econométrica usada. Começaremos pela análise de raiz unitária em séries macroeconômicas, apresentando os testes existentes para dados com e sem quebra estrutural. Em seguida exporemos a metodologia de cointegração, que possibilita a análise dos efeitos de curto e longo prazo como desejado. Finalmente, desenvolveremos a metodologia para análise e estimação de vetores auto-regressivos cointegrados com mudança de regime markoviano. No terceiro capítulo apresentaremos e interpretaremos os resultados obtidos para dados brasileiros usando tanto o modelo linear quanto o não linear.

Capítulo 1 Taxa de Câmbio e Balança Comercial

Os economistas sempre enfatizaram a interdependência entre a taxa de câmbio e a balança comercial e seus efeitos sobre as economias. Uma percepção rudimentar do ajuste das taxas de câmbio em resposta aos desequilíbrios nos pagamentos do comércio internacional pode ser encontrada pelo menos a partir do século XIV. Nesta época, as instituições bancárias européias começaram a emitir títulos de troca com o intuito de facilitar as transações nas feiras de comércio internacional. Estas feiras duravam muitas semanas, fazendo com que os mercadores pudessem fazer várias transações de compra e venda. Para facilitar as transações, as compras e vendas não eram efetuadas com meios de pagamento e os oficiais da feira supervisionavam um processo de compensação. Ao final do evento, validavam-se os pagamentos devidos e recebidos de cada mercador, reduzindo-se assim o número de transações. Os saldos restantes eram pagos através de títulos de troca, conversíveis em ouro ou prata, emitidos por bancos localizados nos grandes centros comerciais da Europa.

No início do século XVII, a influência das taxas de câmbio e da balança comercial nas condições domésticas foi reconhecida nos círculos políticos da Inglaterra e da Itália, onde dificuldades com a economia doméstica devido à saída de divisas geraram preocupações (para uma abordagem histórica das taxas de câmbio ver Enzig (1970)).

Desde meados do século XX, o desenvolvimento tanto na análise econômica quanto na evolução da economia mundial, alteraram a percepção das relações entre a taxa de câmbio e a balança comercial. Ao mesmo tempo, a análise do processo de ajustamento das taxas de câmbio tornou-se muito importante para as políticas públicas que visavam o equilíbrio interno e externo da economia doméstica.

1.1 Abordagem das Elasticidades

Considerando que até o final da década de 60 os fluxos internacionais de capitais eram pequenos em relação ao valor do comércio internacional, a maioria dos modelos de taxa de câmbio e balanço de pagamentos tratava a conta corrente – muitas vezes simplesmente a balança comercial – como o único componente

endógeno do balanço de pagamentos. As taxas de câmbio, por sua vez, eram tidas como exógenas, pois o câmbio era fixo, e as expectativas ainda não exerciam papel na modelagem econômica. Apesar destas limitações, algumas características dos primeiros modelos da conta corrente continuam presentes nos atuais modelos expectacionais de câmbio flexível.

O primeiro modelo relacionando a balança comercial e a taxa de câmbio seguia a abordagem das elasticidades, na tradição marshalliana de se considerar a taxa de câmbio como um preço que equilibra um mercado bem definido com curvas de demanda e oferta. Entre as contribuições mais citadas para o desenvolvimento da abordagem das elasticidades estão Bickerdike (1920), Marshall (1923), Lerner (1914), Robinson (1947) e Metzler (1948). Este modelo se baseia na idéia de um efeito substituição explícito no consumo e implícito na produção, introduzido por uma alteração no preço relativo dos bens domésticos em comparação com os bens estrangeiros após uma desvalorização cambial.

O modelo padrão analisa o efeito de uma alteração na taxa de câmbio sobre a balança comercial em termos de mercados separados para bens produzidos domesticamente e bens produzidos no estrangeiro, abstraindo a existência de bens não transacionáveis.

Seguindo a clássica exposição deste modelo feita por Dornbusch (1975), a oferta de exportações e a demanda por importações dependem somente dos preços nominais medidos em moeda local das exportações e das importações. As funções demanda são Marshallianas com elasticidade preço negativa. Apesar do modelo não ser construído com microfundamentação, pode-se assumir que estas funções de demanda são derivadas de um problema de maximização de utilidade, ou seja, que elas satisfazem as propriedades de homogeneidade de grau zero, de restrição orçamentária ativa e que a matriz de Slutsky é negativa semi-definida. O efeito preço cruzado entre os mercados é desprezado. A equação para a demanda doméstica por importações (exportações do país estrangeiro) é dada por:

$$M^d = M^d(P_m) \quad (1.1.1)$$

Observe que $P_m = EP_m^*$, onde E é a taxa de câmbio nominal em unidades de moeda local por uma unidade de moeda estrangeira e P_m^* é o nível de preço das importações domésticas no estrangeiro, ou seja, assume-se que a paridade do

poder de compra é válida. A equação para a demanda estrangeira por importações (exportações domésticas) é definida de forma similar:

$$M^{d*} = M^{d*}(P_x^*), \quad (1.1.2)$$

onde M^{d*} é a quantidade de importações estrangeiras e P_x^* é o nível de preço das exportações domésticas em moeda externa. Como na definição acima, temos que $P_x^* = \frac{P_x}{E}$, onde P_x é o nível de preço das exportações domésticas em moeda local.

Da mesma forma que se definiu as funções demanda, também as funções de oferta de exportação são definidas de forma a depender somente dos preços nominais:

$$X^s = X^s(P_x) \quad (1.1.3)$$

$$X^{s*} = X^{s*}(P_m^*), \quad (1.1.4)$$

onde X^s e X^{s*} são as quantidades de exportações ofertadas pelo país local e pelo estrangeiro, respectivamente. Portanto, as condições de equilíbrio para as exportações e importações são:

$$M^d = X^{s*} \quad (1.1.5)$$

$$M^{d*} = X^s \quad (1.1.6)$$

Dadas as equações (1.1.1), (1.1.2), (1.1.3) e (1.1.4), a balança comercial em moeda local é dada por:

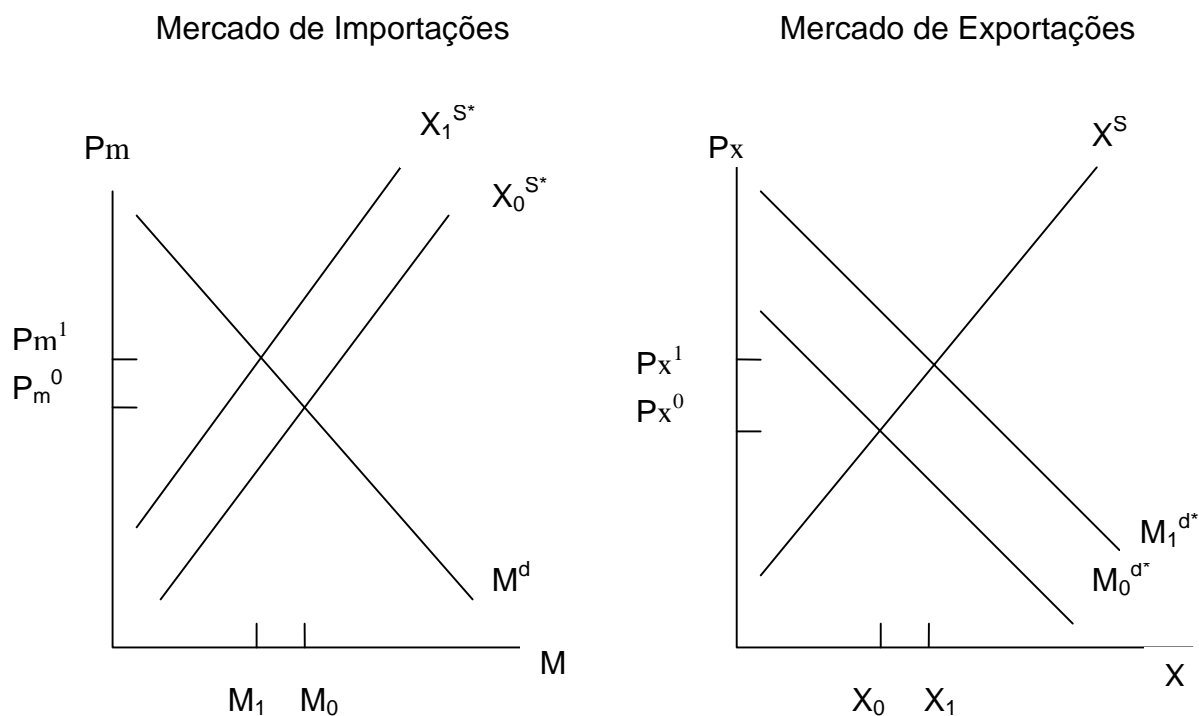
$$B = P_x X^s - P_m M^d \quad (1.1.7)$$

Devem-se notar dois pontos sobre as taxas de câmbio neste modelo. Primeiramente, como não existem bens não transacionáveis, a taxa de câmbio real é medida pelos termos de troca. Além disso, qualquer desvalorização nominal se torna uma desvalorização real. A explicação para isso é a hipótese implícita de que os níveis de preço locais e estrangeiros sejam determinados exogenamente. Kenen (1985) chama atenção para o fato de que isto torna este modelo keynesiano, uma vez que o mercado de bens é equilibrado por mudanças no produto e não por alterações nos preços.

Pode-se usar a estática comparativa para ilustrar as equações deste modelo. Existem dois mercados separados para demanda doméstica por importações e oferta de exportações. Supondo que exista equilíbrio, ou seja, $B=0$, o que acontece

com a balança comercial definida em (1.1.7) após uma desvalorização cambial doméstica? O efeito de uma desvalorização cambial neste modelo é mostrado na Figura 1.1.

Figura 1.1: Um modelo de dois bens e dois países.



Em equilíbrio, as exportações domésticas são X_0 e as importações são M_0 , os preços são P_x^0 e P_m^0 , respectivamente. A desvalorização não desloca a oferta doméstica de exportações e nem a demanda por importações, pois estas são afetadas somente pelos preços locais, que estão no eixo vertical dos gráficos da Figura 1.1. O que ocorre é apenas um movimento ao longo das curvas M^d e X^S , de forma que a oferta de exportações domésticas aumenta e a demanda por importações diminui. Porém a desvalorização cambial altera o nível de preço das exportações locais em moeda externa, bem como o nível de preço das importações domésticas em moeda externa, que não estão explícitos na Figura 1.1. Por isso, as curvas de demanda externa por importações domésticas e de oferta externa de exportações são deslocadas para cima de M_0^{d*} para M_1^{d*} e de X_0^{S*} para X_1^{S*} , respectivamente. De forma a manter o preço em moeda externa dos bens, como definido acima, os preços em moeda local têm que aumentar na mesma proporção da desvalorização. O novo equilíbrio se estabelece quando ambos os mercados se

equilibram com as novas quantidades X_1 e M_1 e os novos preços P_m^1 e P_x^1 . Então, a desvalorização aumenta os preços de equilíbrio em moeda local nos dois mercados, aumentando o volume de exportações domésticas e reduzindo o volume de importações do país local. O que acontece é uma substituição no consumo entre bens domésticos e estrangeiros induzida pela alteração na taxa de câmbio. Dessa forma, o valor das exportações domésticas, $P_x X^s$, aumenta, enquanto o valor das importações, $P_m M^d$, pode aumentar ou diminuir dependendo da elasticidade preço doméstico da demanda. Isto implica que o efeito de uma desvalorização na balança comercial, neste modelo, é ambíguo.

Uma condição suficiente para que a balança comercial melhore após uma desvalorização cambial, na abordagem das elasticidades, foi desenvolvida por Bickerdike (1920) e exposta por Robinson (1947) e Metzler (1949). Por isso ficou conhecida como a condição de Bickerdike-Robinson-Metzler, ou simplesmente condição BRM. Derivando (1.1.7) e colocando os resultados em forma de elasticidade, pode-se conseguir esta condição algébrica. Ela relaciona a resposta da balança comercial às alterações na taxa de câmbio e as elasticidades preço doméstico e externo das importações e exportações:

$$\frac{dB}{dE} = P_x X^s \left[\frac{(1+\varepsilon)\eta^*}{(\varepsilon+\eta^*)} \right] - P_m M^d \left[\frac{(1-\eta)\varepsilon^*}{(\varepsilon^*+\eta)} \right], \quad (1.1.8)$$

onde η e ε denotam as elasticidades (em valor absoluto) da demanda doméstica por importações e oferta doméstica de exportações. Existem duas hipóteses implícitas na derivação das elasticidades demanda. A primeira é que as rendas nominais do país local e do estrangeiro são mantidas constantes. A segunda hipótese é que o nível geral de preços do país local é constante. A interpretação de Dornbusch (1975) da primeira hipótese é de que estas elasticidades são, então, elasticidades compensadas. Kemp (1970) entre outros, enfatizou que, além dessas duas hipóteses, o modelo também assume implicitamente que todas as elasticidades preço cruzadas (entre exportações e importações) são zero. Logo, a matriz de Slutsky é uma matriz diagonal.

Pode-se mostrar que, se $B=0$, então $\frac{dB}{dE} > 0$ se e somente se:

$$\frac{\eta\eta^*(1+\varepsilon+\varepsilon^*)-\varepsilon\varepsilon^*(1-\eta-\eta^*)}{(\varepsilon+\eta^*)(\varepsilon^*+\eta)} > 0. \quad (1.1.9)$$

Note que quando $\eta^* = \varepsilon^* = \infty$, o país em questão não afeta os preços internacionais, pois a demanda por importações e a oferta externa de exportações são perfeitamente elásticas. Este é o caso do “país pequeno”, no qual o Brasil parece se enquadrar. Nesta situação, o país é tomador de preço tanto no mercado de exportação quanto no mercado de importação. Dessa forma, uma desvalorização local não afeta os preços internacionais em moeda externa dos bens exportados e importados. Isto implica que apenas mudanças de volume afetam a balança comercial em moeda externa. Logo, o efeito de uma desvalorização cambial sobre a balança comercial seria o seguinte. Com a desvalorização, os exportadores irão receber mais unidades de moeda local por suas exportações. Com isso, é de se esperar que a resposta destes agentes seja o aumento da quantidade ofertada a um dado preço em moeda externa. Por outro lado, os importadores irão se defrontar com preços mais altos em moeda local para suas importações. Conseqüentemente, eles irão reduzir suas importações. Logo, com volume de exportação aumentando e volume de importações diminuindo a preços em moeda externa fixos, a desvalorização irá melhorar a balança comercial (Lindert e Kindleberger, 1982). Portanto, no caso do país pequeno, uma desvalorização cambial deve melhorar a balança comercial em moeda externa.

Todavia, se a balança comercial é medida em moeda local, a estória pode ser bem diferente. A razão para isto é que o aumento no valor das exportações domésticas pode ser menor que a redução no valor das importações domésticas, ou seja, o efeito final sobre a balança comercial depende das elasticidades preço domésticas. A desvalorização doméstica deve melhorar a balança comercial, em moeda local, se $\varepsilon > |\eta|$ (por hipótese não existem barreiras comerciais qualitativas nem quantitativas).

Outro resultado que pode ser derivado da condição (1.1.9) é a chamada condição de Marshall-Lerner (Marshall, 1923; Lerner, 1944). Esta condição é obtida fazendo $\varepsilon, \varepsilon^* \rightarrow \infty$. Tal hipótese de elasticidade infinita para a oferta de exportações domésticas e externas faz com que o lado esquerdo de (1.1.9) fique $\eta^* + \eta - 1$. Logo, para uma melhora da balança comercial após uma desvalorização cambial, $\eta^* + \eta > 1$ deve valer ou, como na apresentação tradicional, $|\eta^* + \eta| > 1$. A condição de Marshall-Lerner estabelece que quando as elasticidades de oferta de exportação

doméstica e externa são estritamente elásticas e a renda dos países permanece constante, uma desvalorização cambial gera uma melhora do saldo da balança comercial apenas se a soma das elasticidades demanda por importação local e estrangeira, em valor absoluto, é maior que um. Esta condição tem sido considerada na literatura como suficiente para assegurar a estabilidade do mercado de câmbio. Logo, se a condição de Marshall-Lerner é válida, existe um excesso de demanda por moeda estrangeira se esta está abaixo do valor de equilíbrio e um excesso de oferta quando a taxa de câmbio está acima do preço de equilíbrio. Nessas circunstâncias, a taxa de câmbio irá se mover para o valor de equilíbrio e o mercado irá “zerar”. Muitos trabalhos usaram argumentos de equilíbrio parcial para argumentar que a condição de Marshall-Lerner pode não ser satisfeita, entre eles se destacam Dornbusch, 1987; Krugman, 1987 e Krugman e Baldwin, 1987. Os argumentos se baseiam em falhas de mercados tais como histerese, comportamento de *pricing to market* ou incerteza. Desta maneira, uma questão relevante para o presente trabalho é saber se a condição de Marshall-Lerner é válida para o Brasil.

Como descrito pelo conjunto de equações (1.1.1) – (1.1.8), a abordagem das elasticidades tem várias limitações. Primeiro, as funções demanda por importações e oferta de exportações dependem somente dos preços nominais dos bens em questão, ao invés dos preços relativos e de variáveis tais como renda ou capacidade produtiva. Segundo, mudanças na balança comercial correspondem a alterações idênticas nas contas nacionais na diferença entre produção e absorção interna, porém nenhuma destas variáveis foi incluída na análise explicitamente. Terceiro, o conceito de desequilíbrio comercial implica que os bens são pagos com algum ativo que não foi explicitamente incluído na análise.

1.2 Abordagem da Absorção.

Os teóricos da abordagem das elasticidades reconheceram as limitações de sua modelagem. No começo dos anos 50, os efeitos de desvalorizações cambiais sobre a renda nacional e o emprego foram modelados por Robinson (1947), Harberger (1950), Meade (1951), Alexander (1952, 1959), entre outros. Este novo corpo de análise, conhecido como abordagem da absorção, não é uma rejeição da abordagem das elasticidades, mas uma tentativa de integrar esta última com as idéias keynesianas e seu foco nas contas nacionais.

O núcleo desta abordagem é a proposição de que qualquer melhora na balança comercial requer um aumento da renda nacional sobre as despesas domésticas totais. Tal teoria da balança comercial pode ser definida através de identidades macroeconômicas básicas, que expressam as diferentes ligações entre a balança comercial e os agregados macroeconômicos. Duas semelhanças entre as abordagens das elasticidades e da absorção são as hipóteses de países “grandes” e de que a conta corrente se reduz à balança comercial. Duas diferenças são a introdução da renda e da moeda, apesar da última ser muito pouco discutida.

Assumindo ausência de transferências e serviços, de forma que a renda nacional é o PIB e a conta corrente é a balança comercial, pode-se escrever:

$$Y - A = TBDC = XDC - MDC \quad (1.2.1)$$

onde Y é o PIB, $TBDC$ é a balança comercial em moeda local e XDC e MDC são os valores em moeda local das exportações e importações respectivamente. Esta identidade mostra que a balança comercial é apenas um lado da moeda e que a abordagem das elasticidades esqueceu de analisar o outro lado. Ou seja, o que a abordagem da absorção faz é analisar a economia do ponto de vista das despesas agregadas e, especialmente, analisar os efeitos diretos da taxa de câmbio nos preços relativos, renda, absorção e balança comercial. Esta abordagem assume implicitamente a hipótese keynesiana de que o volume de exportações é independente da renda nacional e que as importações dependem positivamente desta. Esta dependência positiva, segundo Alexander (1952), acontece porque a produção de um país depende de insumos que são importados e também porque a importação responde à absorção total.

Esta abordagem integrada enfatiza que a desvalorização da moeda local leva à redução do preço relativo do bem doméstico. Tal redução produz dois efeitos diretos. Primeiro, gera um efeito substituição que altera a composição da demanda, transferindo a demanda por bens estrangeiros para os bens domésticos. Até agora esta análise produz os mesmos resultados da abordagem das elasticidades. Porém, nesta abordagem, existe um efeito renda que aumenta a absorção interna, reduzindo assim a balança comercial. Assumindo a existência de desemprego, como é tradicional na análise keynesiana, este efeito substituição leva a um aumento do produto local e uma redução do produto estrangeiro. O efeito renda está relacionado tanto com o aumento na renda doméstica, que age através da propensão marginal a consumir e a investir, quanto com a redução nos termos de troca. A abordagem da

absorção argumenta que, em geral, uma desvalorização cambial gera uma piora nos termos de troca. A hipótese é de que uma desvalorização irá levar a uma redução no preço das exportações em moeda estrangeira. Como os países são “grandes”, com ofertas elásticas, então, sob a hipótese de preços domésticos constantes (oferta de exportação estritamente elástica), uma desvalorização irá reduzir o preço doméstico das exportações em moeda estrangeira. O preço doméstico das importações em moeda externa fica constante, ou pode diminuir se a oferta externa de exportações não é estritamente elástica. A condição para uma piora nos termos de troca é que o decréscimo no preço das exportações seja maior que o decréscimo no preço das importações.

Porém, o fato de que os termos de troca irão decrescer não implica que a balança comercial irá piorar. Pode haver a piora na balança comercial se a redução nos termos de troca for grande o suficiente para compensar a melhora na balança comercial gerada pelo aumento no volume de exportações e a redução do volume de importações (Lindert e Kindleberger, 1982). Apesar de que a desvalorização aumenta o produto real medido em termos do bem doméstico, ela pode não melhorar o bem estar. Usando os termos de Bhagwati (1958), uma desvalorização pode ser empobrecedora se causa uma queda muito grande nos termos de troca. Pode-se construir modelos onde a desvalorização reduz a renda doméstica, ver Salop (1974). No final, o efeito líquido da desvalorização cambial sobre a balança comercial irá depender da combinação dos efeitos substituição e renda.

A abordagem da absorção leva a quatro conclusões:

1. Uma desvalorização cambial aumenta o produto local quando não há pleno emprego e reduz o produto externo.
2. Por causa dos efeitos sobre as rendas local e estrangeira, e nos fluxos de comércio, a desvalorização melhora a balança comercial em uma quantidade menor do que a prevista pela abordagem das elasticidades.
3. Mudança na taxa de câmbio é uma política ótima em resposta às alterações espontâneas na despesa local ou às alterações na despesa externa.
4. Se os países estão em pleno emprego, entretanto, mudanças na taxa de câmbio por si só não são suficientes para melhorar a balança comercial.

A primeira conclusão é resultado do efeito substituição na despesa doméstica. O aumento na demanda por bens locais faz com que o produto doméstico cresça para eliminar o excesso de demanda, enquanto a redução na

demanda por bens externos faz com que a renda estrangeira decresça para equilibrar este mercado. Isto implica que a desvalorização configura-se como uma política de empobrecer o vizinho (*beggar my neighbour*) para resolver o problema do desemprego local.

A segunda conclusão decorre da primeira, pois o aumento da renda local gera aumento das importações domésticas, enquanto a redução na renda externa causa redução nas importações do exterior e, portanto, nas exportações locais. Logo, o efeito na balança comercial é menor do que a alteração inicial nas despesas com bens domésticos e estrangeiros. Neste caso, o aumento é permanente, porque o processo de equilíbrio é incompleto neste modelo, pois as reservas foram esterilizadas. Em modelos monetários do balanço de pagamentos, por sua vez, a desvalorização leva a melhora apenas temporária na balança comercial, pois as alterações no fluxo de reservas levam à alterações na oferta de moeda, o que desloca as despesas (Dornbusch, 1973).

Os efeitos de uma alteração na taxa de câmbio sempre levam a mudanças na despesa. Portanto, mudança no câmbio é uma resposta ótima à alterações espontâneas na despesa, pois ela pode manter tanto equilíbrio interno nos países quanto o equilíbrio externo. Além disso, é a política ótima para um país que enfrenta uma alteração na despesa estrangeira.

Se a economia doméstica está em pleno emprego, uma desvalorização cambial não melhora a balança comercial, a não ser que esta seja acompanhada por uma redução na demanda pelo bem doméstico. Isto se dá porque a acomodação do excesso de demanda gerado pelo efeito substituição na despesa não pode ser acomodado por aumento na renda. Ao invés, este excesso de demanda é acomodado por aumento nos preços de forma a manter constante a taxa de câmbio real, impedindo qualquer alteração na estrutura das despesas.

1.3 Curva J.

As mudanças nas taxa de câmbio que se seguiram após o colapso do sistema de Bretton Woods, em particular a depreciação do dólar americano, renovaram o interesse no perfil temporal das respostas dos preços e quantidades de bens transacionados às variações cambiais. Passou a ser freqüente o uso do conceito de curva J entre os economistas. A idéia da curva J surge da constatação empírica de

que a balança comercial medida em moeda local pode, inicialmente, deteriorar-se após uma desvalorização cambial, sendo seguida por uma melhora. A idéia é que no curto prazo, o preço em moeda local das importações aumenta mais rápido que o preço das exportações, sendo que os volumes só se ajustam após alguma defasagem. Porém, a idéia de um padrão para a resposta da balança comercial à desvalorização do câmbio começou a ficar mais obscura depois que os economistas começaram a analisar com mais cuidado o processo. A moeda na qual os contratos são fechados, as defasagens no processo de *pass-through* dos preços e quantidades, as incertezas associadas à desvalorização cambial, entre outros fatores, são importantes na determinação do perfil temporal do ajustamento da balança comercial após uma desvalorização.

As explicações teóricas para tal fenômeno têm variado bastante. A primeira explicação teórica para este perfil temporal do ajustamento da balança comercial em resposta a variações cambiais foi feita por Magee (1973). Em uma estrutura que usa a abordagem das elasticidades, abstraindo os efeitos renda, e adotando a hipótese de que tanto a demanda quanto a oferta de importações e exportações dependem somente dos preços relativos, ele identificou três períodos distintos após uma desvalorização cambial. Estes períodos são definidos de acordo com os diferentes fatores afetando a balança comercial: o período de contrato de moedas, o período de *pass through* e o período de ajuste nas quantidades.

O período de contrato de moedas é definido como o período logo após a depreciação, quando contratos fechados antes da alteração no câmbio ainda estão sendo realizados. Como neste horizonte de tempo tanto os preços quanto as quantidades estão fixos, uma depreciação cambial com o intuito de reduzir o déficit da balança comercial irá aumentar ou diminuir o saldo da última, dependendo da proporção dos contratos estipulados em moeda local e moeda estrangeira. Como na maioria das vezes, os contratos de exportação bem como os de importação são fixados em moeda estrangeira, de forma a evitar uma perda cambial, a depreciação irá aumentar ainda mais o déficit neste período.

O período do “pass-through” é definido como o período após a depreciação no qual os preços começam a variar devido à alteração cambial, mas as quantidades permanecem fixas devido a várias restrições atingindo o lado da oferta e o da demanda de importações e exportações. Dessa forma, o valor das importações irá aumentar devido a depreciação cambial, entretanto, a quantidade

importada ficará constante, aumentando o valor das importações totais. Por outro lado, o preço das exportações em moeda estrangeira diminui na mesma proporção da desvalorização - assumindo que os exportadores ajustam o preço em moeda estrangeira de acordo com a desvalorização - mas a demanda não irá se alterar, de forma que a receita em moeda estrangeira irá diminuir e a receita em moeda local permanecerá constante. Dessa forma, a balança comercial medida em moeda local irá deteriorar seguindo o padrão de uma curva J.

O período de ajuste nas quantidades é definido como o período no qual tanto preços quanto quantidades ajustam-se livremente e, dado que a CML é válida, a balança comercial irá certamente melhorar. Entretanto, isto é verdade somente para uma análise de estática comparativa. A mesma análise feita em contexto dinâmico mostra que na transição do equilíbrio antigo para o novo, uma menor velocidade de ajustamento para volume em relação aos preços é suficiente para causar o efeito curva J na balança comercial.

Entre os possíveis fatores causadores de tal rigidez, Junz e Rhomberg (1973) identificam uma defasagem de reconhecimento, onde os agentes demoram a perceber a mudança no ambiente de competição. Após a percepção da nova situação, existe uma defasagem de decisão, que dura do momento no qual a nova situação foi reconhecida até o momento da ação, onde os produtores precisam ser convencidos de que a nova situação será duradoura e rentável o suficiente para compensar os esforços e custos de aumentar a capacidade produtiva ou transferência de recursos. Finalmente, existe uma defasagem na produção e entrega de mercadorias cujos preços relativos foram alterados, de forma que mesmo ordens de compra emitidas logo antes da desvalorização só irão afetar o volume de comércio após seu pagamento. Mas o pagamento, normalmente, só ocorre após a produção e o embarque da mercadoria no porto, o que demanda tempo, devido a dificuldades técnicas e logísticas.

Estas explicações tradicionais para a curva J têm sido criticadas. O argumento usado é de que maior velocidade no ajustamento de preços em relação às quantidades não é o único motivo para que, após uma depreciação cambial, tenha-se a deterioração da balança comercial no curto prazo. Tais efeitos também ocorrem na presença de preços viscosos e quantidades livres para se ajustar. Dessa maneira, o fenômeno da curva J não necessariamente implica um rápido *pass through*. Se os preços dos bens importados são viscosos, os consumidores irão

antecipar o aumento futuro destes, revisando suas compras futuras, o que pode levar a uma dinâmica em J para a balança comercial. É nesse contexto que Gerlach (1989) distingue entre o efeito preço relativo e o efeito de realocação intertemporal. Ele enfatiza que preços mais altos em moeda local de bens importados tendem a melhorar a balança comercial, porém, se os preços são viscosos, os consumidores terão incentivos para antecipar suas compras futuras de bens importados, o que tende a piorar a balança comercial.

Mansoorian (1998) mostra que a deterioração inicial da balança comercial em consequência de uma depreciação do câmbio pode ser atribuída à persistência nos hábitos de consumo. Também neste caso, os efeitos de substituição intertemporal, além de outros fenômenos de sobreposição, podem afetar os preços e os fluxos de comércio.

Outra explicação para os efeitos de curto prazo na balança comercial é a teoria da histerese, como mostrado por Dixit (1994) e Teles (2003). Como operações de importação e exportação geram custos irrecuperáveis, os agentes econômicos valorizam a oportunidade de esperar e ver se a variação cambial é apenas passageira. Se este for o caso, os agentes evitam os custos irrecuperáveis. De fato, o incentivo para que se iniciem ou encerrem negócios internacionais ocorre somente se a variação cambial excede um certo limite. Neste caso, as defasagens de ajustamento não são definidas com relação ao tempo, mas sim em termos de um estado econômico representado pela distância entre a taxa de câmbio corrente e a que faz os agentes reagirem.

Sem recorrer a argumentos de rigidez de preços, Backus et. al. (1994) usa um modelo de ciclos reais de negócios com dois países para explicar os movimentos no saldo da balança comercial em resposta a variações cambiais. Neste modelo, um choque positivo de produtividade, inicialmente, gera um aumento da importação de bens de investimento, causando um déficit na balança comercial. Porém, quando o estoque de capital torna-se condizente com o novo nível de produtividade, o “boom” das importações é dissipado e a balança comercial passa a ser superavitária. O padrão da função impulso resposta para a economia teórica mostra um efeito contemporâneo negativo para a correlação entre os termos de troca e a balança comercial. Esta correlação tende a aumentar com o tempo, gerando uma função de correlação cruzada similar a um S deitado, sendo chamada pelos autores de curva

S. Este padrão é justificado com base na influência da formação de capital na balança comercial.

A teoria da curva J tem sido alvo de vários estudos empíricos, alguns dos quais confirmaram a presença de defasagens de ajustamento na balança comercial após uma depreciação cambial, como por exemplo, Bahmani-Oskooee (1985), Moffet (1989) e Onafowora (2003), enquanto outros, como por exemplo Rose e Yellen (1989) e Hsing e Savvides (1996), rejeitaram a hipótese de curva J. Estes estudos podem ser classificados em dois grupos. No primeiro grupo o fenômeno da curva J é analisado em um modelo de dois países, sendo um o país local e o outro o resto do mundo, como em Felmingham (1988). No segundo grupo o fluxo de comércio bilateral entre grandes parceiros comerciais é levado em conta para a análise, como em Bahmani-Oskooee e Brooks (1999).

Como uma moeda pode, simultaneamente, apreciar em relação à outra moeda e depreciar com respeito à uma terceira, a abordagem de comércio bilateral é usada para evitar erros de interpretação devido a este fato. Porém, a curva J é um fenômeno agregado e não apenas bilateral, de forma que o uso de dados agregados em um modelo de dois países também parece justificado.

Bahmani-Oskooee e Alse (1994) fazem uma análise de cointegração para a taxa de câmbio, renda local, renda externa e balança comercial bilateral para 41 países, entre eles o Brasil. Eles encontram evidência de cointegração para estas variáveis usando a metodologia de Engle e Granger, estimam um modelo de correção de erros linear e, além disso, não utilizam funções impulso resposta.

Teles (2003) faz uma análise do equilíbrio externo da economia brasileira usando a abordagem de histerese introduzida por Dixit (1994). Neste contexto, ele conclui que uma política de juros altos afeta negativamente a decisão da firma doméstica de entrar no mercado externo, pois aumenta o nível da taxa de câmbio a partir do qual a firma passa a agir. Outra conclusão é que a mudança do regime de câmbio fixo para câmbio flutuante também alterou o limite da taxa de câmbio a partir do qual a firma entra no mercado externo. A mudança de regime alterou os parâmetros do processo estocástico ao qual a taxa de câmbio está sujeita, como a tendência e a variância. Dessa forma, não só a taxa de câmbio esperada mudou, mas também sua variância esperada. Portanto, a política monetária, bem como a política cambial, afetam a defasagem descrita pela curva J.

Capítulo 2 Metodologia

2.1 Raízes Unitárias

Os resultados assintóticos usuais não são aplicáveis se algumas variáveis do modelo de regressão são não estacionárias. No caso do modelo de regressão linear $Y = X\beta + U$, os resultados usuais dependem da hipótese de que $n^{-1}X'X$ tenda para uma matriz positiva definida finita, quando n tende para infinito. Quando esta hipótese é violada pode-se ter regressões espúrias, ou seja, obter estimativas significativas de relações entre variáveis totalmente não relacionadas. Este é um problema prático sério, pois grande parte das séries de tempo macroeconômicas aumentam ao longo do tempo e, portanto, violam esta hipótese (ver Nelson e Plosser (1982)).

Duas formas de evitar que esta hipótese seja violada são retirar a tendência da série ou diferenciá-la. Mas estas são duas operações distintas. Se a primeira é apropriada, a segunda não será e vice-versa. Retirar a tendência da série de tempo y_t será apropriado quando o processo gerador dos dados puder ser descrito como $y_t = \gamma_0 + \gamma_1 t + z_t$. Onde t é a tendência de tempo e z_t segue um processo ARMA estacionário. Por outro lado, diferenciar a série é apropriado quando o processo gerador dos dados puder ser descrito por $y_t = \gamma_1 + y_{t-1} + z_t$. Onde, novamente, z_t segue um processo ARMA estacionário.

A escolha entre diferenciar e retirar a tendência de uma série é a escolha entre estes dois modelos. As principais formas para esta escolha são os testes de raiz unitária.

- Representação de uma série de tempo com e sem raiz unitária

Pode ser útil pensar uma série de tempo macroeconômica y_t como a soma de vários componentes com propriedades diferentes.

$$y_t = TD_t + z_t \quad (2.1.1)$$

Aqui TD_t é uma função que descreve a tendência determinística de y_t e z_t é uma função ruído ou componente estocástico de y_t . A hipótese de raiz unitária diz respeito ao comportamento da função ruído, mas a especificação da tendência determinística é crucial para os testes de raiz unitária. Em princípio, uma grande variedade de especificações são possíveis, mas normalmente postula-se que TD_t é linear no tempo, isto é:

$$TD_t = k + \delta t \quad (2.1.2)$$

De forma a simplificar a exposição, assume-se que a função ruído pode ser descrita por um processo autoregressivo com médias móveis.

$$A(L)z_t = B(L)e_t \quad (2.1.3)$$

Onde $A(L)$ e $B(L)$ são polinômios no operador de defasagem L , de ordem p e q respectivamente, e e_t é uma seqüência de inovações i.i.d.. A função ruído tem média zero, uma vez que a tendência determinística inclui a média da série no termo k . Outra hipótese é de que as raízes do polinômio de médias móveis estão estritamente fora do círculo unitário. A equação (2.1.3) resume a dinâmica univariada do processo z_t . Aqui o sistema (2.1.1)-(2.1.3) será chamado de processo gerador dos dados (PGD), mesmo que ele só resume as implicações univariadas de um sistema multivariado mais completo.

Pode-se agora distinguir dois modelos alternativos para y_t . No modelo tendência-estacionário, as raízes do polinômio autoregressivo estão estritamente fora do círculo unitário, de forma que z_t é um processo estacionário e y_t é estacionário em torno de uma tendência. No modelo diferença-estacionário, z_t tem uma raiz unitária autoregressiva e todas as outras raízes estão estritamente fora do círculo unitário. Neste caso, $\Delta z_t \equiv (1-L)z_t$ é um processo estacionário e Δy_t é estacionária em torno de uma média fixa. A hipótese de raiz unitária é que y_t seja um processo diferença-estacionário. Os modelos tendência-estacionário e diferença-estacionário são comumente chamados de modelos integrados de ordem zero e de ordem um, respectivamente.

Para entender o significado da hipótese de raiz unitária, é útil decompor a função ruído z_t em um componente cíclico c_t e uma tendência estocástica TS_t . O

componente cíclico é, por hipótese, um processo estacionário de média zero, de forma que choques em c_t não têm efeito de longo prazo no nível de y_t . A tendência estocástica incorpora todos os choques que têm efeitos permanentes sobre o nível de y_t . A soma da tendência determinística TD_t com a tendência estocástica TS_t é a tendência total da série. É comum em macroeconomia empírica tentar isolar o componente cíclico c_t , subtraindo de y_t as tendências TD_t e TS_t .

Nos modelos tendência-estacionário, a decomposição de z_t em tendência estocástica e ciclo é trivial, uma vez que z_t é estacionário por hipótese, de forma que ele já satisfaz as condições para o ciclo c_t . Neste caso, a tendência estocástica TS_t é zero e o ciclo é igual ao ruído z_t . No modelo diferença-estacionário, a decomposição é um pouco mais complicada. Quando o polinômio $A(L)$ em (3) possui uma raiz unitária, pode-se escrever $A(L) = (1-L)A^*(L)$, onde $A^*(L)$ tem raízes fora do círculo unitário. A primeira diferença Δz_t segue um processo ARMA estacionário $A^*(L)\Delta z_t = B(L)e_t$. Seguindo Beveridge e Nelson (1981), pode-se construir a seguinte decomposição. Seja $\psi(L) = A^*(L)^{-1}B(L)$ a representação de médias móveis da primeira diferença de z_t . A notação $\psi(1)$ denota a soma dos coeficientes de média móvel. Define-se $\psi^*(L) = (1-L)^{-1}[\psi(L) - \psi(1)]$, de forma que Δz_t satisfaz $\Delta z_t = [\psi(1) + (1-L)\psi^*(L)]e_t$. Então, multiplicando os dois lados por $(1-L)^{-1}$, pode-se escrever:

$$z_t = \psi(1) \frac{e_t}{(1-L)} + \psi^*(L)e_t$$

Mas $\frac{e_t}{(1-L)}$ é uma progressão geométrica. Supondo que o processo se iniciou em e_1 , obtém-se:

$$\frac{e_t}{(1-L)} = \sum_{j=1}^t e_j = s_t$$

Então:

$$z_t = \psi(1)s_t + \psi^*(L)e_t \quad (2.1.4)$$

Onde s_t é um passeio aleatório de média zero. Aqui a função de tendência para a variável y_t contém não somente a tendência determinística TD_t , mas também o componente estocástico de z_t , $TS_t = \psi(1)s_t$ que afeta o intercepto da tendência em cada período. Esta tendência estocástica é obtida através da soma dos coeficientes de médias móveis de Δz_t , que é equivalente ao efeito de longo prazo de um choque unitário em e_t sobre o nível de z_t . O ruído ou componente cíclico é $c_t = \psi^*(L)e_t$, construído de forma a não ter efeito de longo prazo em z_t .

A decomposição (2.1.4) pode ser usada para desenvolver uma medida da importância da tendência estocástica TS_t para o comportamento da série y_t . Campbell e Mankiw (1987) propuseram que o coeficiente $\psi(1)$ é uma medida natural da persistência da série, porque ele é a razão entre o efeito de longo prazo e o efeito imediato de uma inovação e_t . Quando $\psi(1) > 1$, o efeito de longo prazo de um choque univariado em y_t é maior que o efeito imediato; quando $\psi(1) < 1$, por outro lado, o choque tende a diminuir. O modelo de passeio aleatório tem $\psi(1) = 1$, enquanto o modelo tendência-estacionário é o caso limite onde $\psi(1) = 0$.

Os processos tendência-estacionário e diferença-estacionário descritos acima podem ser considerados modelos na forma reduzida. É possível derivar estes processos como formas reduzidas de modelos estruturais com componentes não observáveis, como em Harvey (1985). Considere, por exemplo, um modelo de componentes não observáveis que representa y_t como a soma de um passeio aleatório com *drift* e um processo estocástico independente. Quando a inovação do passeio aleatório é zero, y_t é tendência-estacionário. Generalizando, a forma reduzida deste modelo é um processo diferença-estacionário com restrições, como em Clark (1987) e Watson (1986). De especial relevância é a restrição de $\psi(1)$ em (2.1.4) ser menor que um, isto é, que o efeito de longo prazo das inovações não seja maior que o efeito imediato.

Ultimamente, vários modelos estruturais não lineares foram propostos. Estes levam a formas reduzidas não lineares ao invés de formas reduzidas lineares tendência-estacionária e diferença-estacionária. Estes modelos tentam capturar a idéia de que dois tipos fundamentais de choques estão presentes. Os grandes

choques ocorrem com menor freqüência e afetam a função tendência da série de forma permanente. Já os choques regulares ocorrem todos os períodos e podem ou não afetar o nível da série permanentemente. Neste caso, o problema da raiz unitária é saber se os choques regulares têm efeitos permanentes sobre a série.

Uma classe destes modelos foi proposta por Hamilton (1989). O seu modelo estrutural faz com que y_t seja a soma de uma função tendência não linear e um processo ARIMA linear com raiz no círculo unitário. A função tendência é um passeio aleatório com *drift*, que muda de valores altos para baixos de acordo com um processo de Markov de primeira ordem.

Perron (1989) sugeriu que uma estrutura de série de tempo com mudanças infrequentes na inclinação pode ser uma boa aproximação em aplicações empíricas. Na verdade, uma única alteração na inclinação pode ser suficiente para caracterizar muitas séries de interesse. Restringir o número de mudanças na inclinação a priori, é uma solução para as dificuldades técnicas dos testes de raiz unitária na estrutura proposta por Hamilton. Desta forma, pode-se obter testes válidos assintoticamente da hipótese nula de que a parte linear do processo contém uma raiz unitária. Nesta estrutura restritiva, mas útil empiricamente, a forma reduzida das séries é descrita por (2.1.1) com componente determinístico dado por:

$$TD_t = k + \delta_0 t + \delta_1 (t - T_B) \cdot 1(t > T_B),$$

onde $1(\cdot)$ é a função indicador e T_B é o momento da alteração na inclinação. Se z_t contém uma raiz unitária, a função tendência também contém um componente estocástico, de forma análoga aos processos diferença-estacionários usuais.

Um modelo similar pode ser derivado para uma série com mudanças infrequentes no intercepto. Neste caso, é possível testar a hipótese nula de raiz unitária através da especificação do componente determinístico da função tendência como:

$$TD_t = k_0 + k_1 \cdot 1(t > T_B) \quad (2.1.5)$$

Logo, os modelos reduzidos descritos por Perron (1989) podem ser vistos como aproximações dos modelos estruturais onde mudanças infrequentes no intercepto ou na inclinação são modelados estocasticamente como em Hamilton (1989). A hipótese implícita é de que, dado o conjunto de dados de interesse, existe

apenas um grande choque. Porém, se os dados forem relativos a um período muito longo, pode ser necessário mais de uma mudança.

Porém, a teoria da distribuição assintótica por trás dos valores críticos obtidos por Perron (1989) depende da hipótese de que a data da mudança no intercepto ou na inclinação da série é conhecida *a priori*. Christiano (1992) argumenta que a escolha das datas das quebras estruturais é quase sempre correlacionada com os dados. Este é um problema importante, pois tanto a distribuição de amostra finita das estatísticas de teste, quanto a assintótica, dependem do grau de correlação entre a data da quebra e os dados.

Para solucionar este problema, Perron (1997) propõe um teste que escolhe o ponto de quebra estrutural de forma que a estatística t para testar a hipótese nula de raiz unitária seja a menor possível. Dessa forma, a escolha do ponto de quebra é perfeitamente correlacionada com os dados. Perron (1997) argumenta que se for possível rejeitar a hipótese nula neste contexto, será possível rejeita-la também quando a correlação for menor que um. Através de experimentos de Monte Carlo e comparações com os resultados obtidos em Perron (1989), ele conclui que os resultados de seu trabalho anterior são robustos. Perron (1997) mostra que os valores críticos baseados na hipótese de nenhuma correlação entre a data da quebra e os dados, pode ser uma boa aproximação para os valores críticos reais. Porém, cada pesquisador pode diferir em relação à quantidade de informação *a priori* que deseja usar. Parece existir um *trade-off* entre poder do teste e quantidade de informação *a priori* a respeito da escolha do ponto de quebra que se deseja utilizar. Um procedimento que não impõe informações *a priori* tem poder relativamente baixo.

- Testando a presença de uma raiz unitária.

Considerando primeiramente o caso mais simples, onde o componente de ruído z_t (que é a série y_t menos a tendência determinística) é um processo AR(1), isto é, $z_t = \phi z_{t-1} + e_t$. Este processo pode ser descrito como:

$$\Delta z_t = \pi z_{t-1} + e_t \quad (2.1.6)$$

Onde $\pi = \phi - 1$. Aqui a hipótese de raiz unitária é dada por $\pi = 0$, enquanto tendência-estacionariedade implica $\pi < 0$.

Ao testar a hipótese de raiz unitária, é importante distinguir entre o processo gerador dos dados e a equação de regressão usada no teste.

Seja DV_t o conjunto de variáveis que aparecem na tendência determinística do PGD. Na maioria das aplicações $DV_t = \{1\}$, uma constante, ou $DV_t = \{1, t\}$, permitindo um polinômio de 1ª ordem em t . Entretanto, DV_t pode ser mais complicado; por exemplo, o modelo estrutural não linear com mudança determinística no intercepto na data T_B tem $DV_t = \{1, t, 1(t - T_B)\}$. Como o interesse é a função ruído, uma estratégia natural é primeiro retirar a tendência da série e analisar o componente dos resíduos estimados. Usa-se a notação \tilde{y} para os resíduos da projeção de y_t em um conjunto de regressores determinísticos DR_t . A hipótese de raiz unitária pode ser testada através da estimação do par de regressões:

$$y_t = \tau' DR_t + \tilde{y} \qquad \Delta \tilde{y}_t = \pi \tilde{y}_{t-1} + u_t \qquad (2.1.7)$$

Usa-se então a estatística t para testar $\pi = 0$, denotada t_{π} . A escolha natural para o conjunto de regressores DR_t é o conjunto de variáveis DV_t que aparecem na tendência determinística do PGD.

Quando a tendência determinística é linear no tempo ($DV_t = \{1, t\}$ ou $DV_t = \{1\}$) este procedimento de dois estágios é assintoticamente equivalente ao procedimento de um estágio, onde os regressores determinísticos DR_t^* são incluídos na seguinte autoregressão:

$$\Delta y_t = \tau' DR_t^* + \pi y_{t-1} + u_t \qquad (2.1.8)$$

Onde $DR_t^* = DR_t$. O conjunto de regressores DR_t^* deve incluir todos os elementos DR_t para que a equivalência assintótica valha. Em particular, se $DV_t = DR_t = \{1, t\}$, o procedimento dois estágios será equivalente ao procedimento um estágio somente se DR_t^* incluir a tendência t . O coeficiente desta variável é $-\delta\pi$, que será igual a zero sob H_0 , mas diferente de zero sob H_1 de que y_t é tendência-estacionária. Logo, a tendência t deve ser incluída de forma a permitir que a equação (2.1.8) englobe tanto a hipótese nula quanto a alternativa.

Quando a função tendência determinística TD_t é não linear, a relação entre o procedimento dois estágios de (2.1.7) e o procedimento um estágio de (2.1.8) é mais complicada. No caso de tendência com uma única mudança no intercepto, como descrito em (2.1.5), onde $DV_t = \{1, t, 1(t - T_B)\}$, o procedimento em dois estágios com $DR_t = DV_t$ será equivalente ao procedimento um estágio somente se $DR_t^* = \{1, t, 1(t - T_B), D(T_B)_t\}$, onde $D(T_B)_t$ é um para $t = T_B + 1$ e zero nos outros casos. O regressor extra $D(T_B)_t$ deve ser incluído para que o procedimento em um estágio englobe adequadamente as hipóteses nula e alternativa do teste, porém não é necessário no procedimento dois estágios. No caso de tendência com uma alteração na inclinação, o procedimento dois estágios e um estágio podem não ser equivalentes nem assintoticamente.

Uma vez determinados os regressores determinísticos, pode-se testar as implicações da hipótese de raiz unitária para as regressões (2.1.7) e (2.1.8).

Um ponto importante a se notar é que a distribuição assintótica de t_{π} , sob a hipótese nula de raiz unitária, é não normal e depende dos elementos incluídos no conjunto de regressores determinísticos.

A razão básica para a dependência da distribuição assintótica sob a hipótese nula nos regressores determinísticos incluídos é o fato de que a função tendência especificada precisa ser estimada. Se os verdadeiros coeficientes da tendência do PGD fossem conhecidos, somente um conjunto de valores críticos seria necessários, o conjunto para $DR_t = \phi$. Os valores críticos tabelados também revelam importantes implicações para o poder dos testes de raiz unitária, ou seja, para a probabilidade do teste rejeitar a hipótese de raiz unitária quando a hipótese alternativa de tendência-estacionariedade é verdadeira. Pois, sob H_0 de raiz unitária, os valores críticos da cauda esquerda da distribuição assintótica de t_{π} aumenta, em valor absoluto, com o número de regressores determinísticos incluídos.

A situação é outra quando o conjunto de regressores determinísticos incluídos não contém todos os componentes da tendência determinística do PGD. Suponha que DR_t omita uma variável de DV_t que esteja crescendo a uma taxa pelo menos tão rápida quanto qualquer outro elemento de DR_t . Então, sob H_0 de raiz unitária, a

estatística t_{π} em (2.1.7) pode ser normalizada de tal forma que sua distribuição assintótica seja normal.

Este é o caso onde uma tendência linear diferente de zero está presente no PGD, mas não está contida no conjunto de regressores determinísticos DR_t . É também o caso onde o PGD contém um polinômio de tendência de ordem maior que os contidos em DR_t .

Porém, é preciso cuidado ao interpretar esta afirmação. Ela parece sugerir que pode-se aumentar o poder do teste de raiz unitária omitindo-se certos regressores determinísticos que estão presentes no PGD. Se a estatística t for usada para testar $\pi = 0$ em uma regressão sem tendência no caso onde o PGD é um passeio aleatório com “drift”, a distribuição assintótica da estatística de teste é normal e os valores críticos são menores, em valor absoluto, que os valores críticos obtidos quando a tendência é incluída como regressor. Entretanto, a distribuição para amostras finitas de t_{π} não é invariante aos valores do parâmetro de tendência e, para valores pequenos deste parâmetro, a aproximação pela normal pode não ser válida. Além disso, este procedimento leva a testes onde o poder vai para zero quando a amostra cresce para infinito. Esta é uma forma extrema de inconsistência, sendo teste inconsistente definido como aquele cujo poder em relação a uma hipótese alternativa fixa, não vai para um quando a amostra vai para infinito.

No caso onde o PGD é tendência estacionário ao longo da função tendência determinística $TD_t = k_t + \delta t$; e somente uma constante é incluída como regressor. Ao aplicar a regressão $\Delta y_t = c + \pi y_{t-1} + u_t$, a única maneira de considerar a tendência é fazendo $\pi = 0$. De forma similar, se o PGD especifica uma função tendência com mudança no coeficiente de inclinação, o teste de raiz unitário usando somente uma constante e uma tendência determinística como regressores irá gerar um teste inconsistente. Já se o PGD especifica uma alteração no intercepto da função tendência em alguma data e não for incluído um regressor para captar esta quebra, o poder do teste irá decrescer à medida que a magnitude da mudança aumenta.

Portanto, percebe-se a importância da inclusão de tantos regressores determinísticos quanto forem os componentes da função tendência do PGD. Porém, é desejável que não se inclua regressores determinísticos em excesso.

Suponha que t_{π} é construída usando um conjunto de regressores determinísticos DR_t , que inclua pelo menos todos os componentes determinísticos do PGD. O poder do teste de raiz unitária contra a hipótese alternativa de estacionariedade decresce à medida que novos regressores determinísticos são incluídos.

A afirmação acima é parcialmente justificada pela afirmação já feita de que os valores críticos aumentam, em valor absoluto, com o número de regressores determinísticos incluídos. Entretanto, em amostras finitas, existe um viés para baixo na estimativa de π e o viés cresce à medida que o número de regressores determinísticos não contidos no PGD aumenta.

A discussão acima sugere que deve-se ficar atento para a escolha dos regressores determinísticos a incluir, de forma a obter-se propriedades de poder razoáveis. Quando não é claro qual é o conjunto de regressores determinísticos adequado, uma seqüência de testes pode ser útil. Em Perron (1988), é discutida uma estratégia de teste seqüencial para os casos onde a função tendência do PGD contém nenhum componente, uma constante ou uma constante e tendência. Argumenta-se que a estratégia de teste deve iniciar na especificação mais geral da tendência e continuar na direção de especificações mais restritivas.

A má especificação dos componentes determinísticos incluídos como regressores no teste de raiz unitária, pode levar a não rejeição da hipótese nula quando esta é falsa.

- Tamanho e período da amostra.

Muitas vezes, se obtém vários conjuntos de dados para a mesma série de tempo. Isto ocorre, principalmente, quando os dados são disponíveis em freqüências amostrais diferentes. Normalmente, uma série anual terá menos observações que uma série trimestral, que por sua vez, terá menos observações que uma série mensal. Porém, as séries anuais costumam cobrir um período maior que as séries trimestrais e estas últimas cobrem um período maior que as séries mensais. É

natural perguntar qual conjunto de dados permite maior poder. Um número maior de observações é melhor em termos de poder?

Os testes de raiz unitária contra a alternativa estacionária dependem muito pouco do número de observações em si, mas sim do período coberto pelos dados. Para um dado número de observações, o poder é maior quanto maior for o período. Para um dado período, observações adicionais usando dados mais freqüentes leva a um acréscimo apenas marginal no poder. Na maioria das aplicações de interesse, um conjunto de dados contendo dados anuais de um longo período levará a testes com maior poder do que usando mais observações por um período mais curto.

Porém, dados históricos muito longos podem ter qualidade questionada devido às antigas metodologias de construção das estatísticas. Além disso, uma amostra grande aumenta a chance de que a série de interesse tenha quebras estruturais de grande porte. A presença de uma quebra desse tipo irá viesar o teste em favor da hipótese de raiz unitária.

- Extensão para processos com correlação serial.

No caso onde a função ruído z_t obedece um processo ARMA(p,q) como em (2.1.3), ao invés de um AR(1) como em (2.1.6), os pontos levantados anteriormente continuam válidos, porém novos tópicos aparecem. A distribuição assintótica da estatística t_{π} nas autoregressões de primeira ordem como em (2.1.7) e (2.1.8) depende da estrutura de correlação dos dados. Portanto, modificações são necessárias para livrar-se desta dependência incômoda.

Dickey e Fuller (1979) e Said e Dickey (1984) fazem uma correção paramétrica motivada no caso de um processo AR(p). Neste caso, pode-se escrever

$\Delta z_t = \pi z_{t-1} + \sum_{j=1}^{p-1} \gamma_j \Delta z_{t-j}$, onde $\pi = \sum_{i=1}^p a_i - 1$ é a diferença entre a soma dos coeficientes

autoregressivos e um, enquanto $\gamma_j = -\sum_{i=j+1}^p a_i$. Como antes, o componente de ruído z_t

tem raiz unitária se $\pi = 0$. A regressão (estimada por mínimos quadrados) é então:

$$y_t = \tau' DR_t + \tilde{y}_t \quad ; \quad \Delta \tilde{y}_t = \pi \tilde{y}_{t-1} + \sum_{j=1}^k \gamma_j \Delta \tilde{y}_{t-j} \quad (2.1.9)$$

ou

$$\Delta y_t = \tau' DR_t^* + \pi y_{t-1} + \sum_{j=1}^k \gamma_j \Delta y_{t-j} + u_t \quad (2.1.10)$$

Onde $k=p-1$. Aqui DR_t e DR_t^* também são vetores de regressores determinísticos. No caso de um AR(p) puro, a distribuição assintótica de t_{π} obtida de (2.1.9) e (2.1.10) é a mesma obtida para o caso do AR(1). No caso mais geral, onde o componente de ruído é um ARMA(p,q), Said e Dickey (1984) sugerem que o processo pode ser aproximado por um processo autoregressivo de maior ordem. Neste caso, as regressões (2.1.9) e (2.1.10) continuam apropriadas. A condição técnica para que tal procedimento continue assintoticamente válido é que a ordem da autoregressão estimada, k , cresça a uma taxa razoável quando a amostra vai para infinito.

Na prática, a escolha do parâmetro de defasagem para a truncagem é importante. Primeiro, mesmo no caso do AR(p) puro, a ordem p é normalmente desconhecida. No caso mais geral do ARMA(p,q), as condições assintóticas teóricas para a validade do procedimento não são informativas o suficiente para guiar a escolha em amostras finitas. Este problema é importante porque o resultado do teste depende da escolha do parâmetro de truncagem das defasagens. Vários fatores podem explicar essa sensibilidade. Primeiro, poucas defasagens podem afetar negativamente o tamanho do teste. Segundo, a introdução de muitas defasagens reduz o poder, pois mais parâmetros devem ser estimados com o mesmo número de observações. Finalmente, quando k se altera, as condições iniciais também se alteram. Este último fator é de grande importância, dada a não invariância da função de poder da estatística do teste às condições iniciais, como discutido em DeJong, Nankervis, Savin, e Whiteman (1990). Estes fatores mostram a importância de se escolher o parâmetro de truncagem das defasagens com bom senso.

A escolha de um k fixo, escolhido independentemente dos dados, provavelmente, será inadequada. Existem evidências de que um método para a seleção do parâmetro de defasagens do teste que seja dependente dos dados leva a estatísticas com melhores propriedades de tamanho e poder (ver Ng e Perron (1995), Perron e Vogelsang (1992) e Hall (1994))

Existem alguns procedimentos para a escolha de k que são dependentes dos dados. Qualquer procedimento que escolha a ordem autoregressiva assintoticamente, por exemplo, usando o critério de informação de Akaike ou um

teste F sobre a significância conjunta das defasagens adicionais será adequado. É importante, porém, notar que o método de seleção deve partir do modelo geral para o específico.

Uma outra alternativa para lidar com o problema da correlação serial no processo de ruído z_t foi proposta por Phillips (1987) e Phillips e Perron (1988). Esta abordagem consiste em adicionar à estatística do teste de raiz unitária original, um fator de correção que elimina a dependência de sua distribuição assintótica na correlação serial de z_t . A correção utiliza uma estimativa não paramétrica da densidade espectral de Δz_t à frequência zero, medida em relação à variância amostral de Δz_t . Esta estimativa não paramétrica é uma soma ponderada da autocovariância de Δz_t , onde os pesos são escolhidos de tal forma que a densidade espectral é positiva por construção.

Esta estatística de teste é facilmente implementada e assintoticamente válida sob condições muito gerais. Porém, estudos e simulações como Schwert (1989) e Phillips e Perron (1988) mostraram que ela tem sérios problemas de distorção e tamanho em amostras finitas quando o PGD tem uma predominância de autocorrelação negativa na primeira diferença. Isto sugere que o teste de Phillips e Perron pode ser menos confiável que a metodologia de Dickey e Fuller, onde uma correção paramétrica é usada. Um ponto importante, que dá alguma esperança para esta classe de estatísticas é que resultados de simulação também sugerem que o tamanho, ajustado ao poder, delas é bem maior que o da estatística de Dickey e Fuller.

- Abordagens alternativas

Grande parte da literatura existente se preocupa com as propriedades dos coeficientes e da estatística t em autoregressões para a variável y_t . Alguns autores, entretanto, exploram as implicações de modelos de raiz unitária em comparação com modelos tendência-estacionários através da observação do comportamento assintótico das séries de tempo em questão.

Suponha que o PGD não contenha componente determinístico, de forma que $y_t = z_t$ seja um modelo ARMA(p,q) com média zero. Se y_t contém raiz unitária, sob

condições gerais, tem-se que $T^{-1/2}y_T$ converge para um movimento browniano. Sob a hipótese de que y_t não contém raiz unitária, $T^{-1/2}y_t$ converge para zero. Stock (1990) usou esta idéia para desenvolver uma classe de testes estatísticos da hipótese nula de raiz unitária. Simulações sugerem que o problema de tamanho é aliviado, enquanto o poder deste teste é maior que o da maioria das estatísticas disponíveis.

2.2 Cointegração

Um vetor de séries de tempo y_t é dito cointegrado se todas as série individuais são $I(1)$, ou seja, não estacionárias com uma raiz unitária, enquanto uma combinação linear das séries $a'Y_t$ é estacionária, ou $I(0)$, para algum vetor não nulo a .

Cointegração significa que apesar de muitos acontecimentos causarem alterações permanentes em y_t , existe uma relação de equilíbrio de longo prazo unindo seus componentes individuais, representada pela combinação $a'Y_t$. Um exemplo de tal sistema é o modelo de gastos com consumo proposto por Davidson, Hendry, Srba e Yeo (1978). O resultado deste estudo sugere que apesar do consumo e da renda serem processos não estacionários, no longo prazo, o consumo tende a ser uma porção constante da renda, de forma que a diferença entre o logaritmo do consumo e da renda é constante ao longo do tempo.

A análise de cointegração formal introduzida por Engle e Granger (1987) começa por considerar um conjunto de variáveis econômicas em equilíbrio de longo prazo, quando:

$$\beta_1 y_{1t} + \beta_2 y_{2t} + \dots + \beta_n y_{nt} = 0 \quad (2.2.1)$$

Em notação vetorial, o sistema está em equilíbrio de longo prazo quando $\beta y_t = 0$. O desvio do equilíbrio de longo prazo, chamado de erro de equilíbrio e denotado por e_t é tal que:

$$e_t = \beta y_t \quad (2.2.2)$$

Se a relação de equilíbrio faz sentido, os desvios do equilíbrio devem ser estacionários. Engle e Granger (1987) dão a seguinte definição de cointegração.

Os componentes do vetor $y_t = (y_{1t}, y_{2t}, \dots, y_{nt})'$ são ditos cointegrados de ordem d, b , denotado por $y_t \sim CI(d, b)$ se:

1. Todos os componentes de y_t são integrados de ordem d .
2. Existe um vetor $\beta = (\beta_1, \beta_2, \dots, \beta_n)$ tal que a combinação linear $\beta y_t = \beta_1 y_{1t} + \beta_2 y_{2t} + \dots + \beta_n y_{nt}$ é integrada de ordem $(d-b)$, onde $b > 0$.

O vetor β é dito vetor de cointegração.

Se o vetor y_t é cointegrado, então não é correto ajustar um VAR com os dados diferenciados. Suponha o seguinte exemplo de um vetor cointegrado:

$$y_{1t} = \gamma y_{2t} + u_{1t} \quad (2.2.3)$$

$$y_{2t} = y_{2t-1} + u_{2t} \quad (2.2.4)$$

onde u_{1t} e u_{2t} são ruídos brancos não correlacionados. A representação univariada de y_{2t} é um passeio aleatório:

$$\Delta y_{2t} = u_{2t} \quad (2.2.5)$$

enquanto diferenciando (2.2.3), temos:

$$\Delta y_{1t} = \gamma \Delta y_{2t} + \Delta u_{1t} = \gamma u_{2t} + u_{1t} - u_{1t-1} \quad (2.2.6)$$

Reescrevendo (2.2.6), temos:

$$(\gamma u_{2t} + u_{1t}) - u_{1t-1} = \varepsilon_{1t} - (\varepsilon_{1t-1} - \gamma \varepsilon_{2t-1}) = (1-L)\varepsilon_{1t} + \gamma L \varepsilon_{2t}$$

onde $\varepsilon_{2t} = u_{2t}$ é o erro de previsão de y_{2t} com base em valores passados de y_{1t} e y_{2t} , e $\varepsilon_{1t} = \gamma u_{2t} + u_{1t}$ é o erro de previsão de y_{1t} .

Substituindo este resultado em (2.2.6) e formando um vetor com (2.2.5), obtém-se a representação VARMA de $(\Delta y_{1t}, \Delta y_{2t})'$:

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \Psi(L) \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \quad (2.2.7)$$

$$\text{onde } \Psi(L) = \begin{bmatrix} 1-L & \gamma L \\ 0 & 1 \end{bmatrix} \quad (2.2.8)$$

O VAR para os dados diferenciados, se existisse, seria:

$$\Phi(L) \Delta Y_t = \varepsilon_t \quad (2.2.9)$$

onde $\Phi(L) = [\Psi(L)]^{-1}$. Mas a matriz polinomial associada com a representação VMA para este processo, $\Psi(L)$, tem uma raiz unitária:

$$|\Psi(L)| = \begin{vmatrix} 1-L & \gamma L \\ 0 & 1 \end{vmatrix} = 0 \quad (2.2.10)$$

Isto faz com que a matriz do operador de média móvel seja não invertível, de forma que nenhum VAR de ordem finita pode descrever ΔY_t .

A razão pela a qual o VAR em diferenças não é uma boa aproximação para o sistema cointegrado contido em (2.2.3) e (2.2.4) é que y_{2t} em nível contém informações relevantes para a previsão de y_{1t} , e que são descartadas ao se diferenciar y_{2t} .

Ao se modificar o VAR de forma a incluir defasagens das variáveis em nível, é fácil de encontrar uma representação estacionária similar a um VAR para ΔY_t . Lembrando que $u_{1t-1} = y_{1,t-1} - \gamma y_{2,t-2}$, note que podemos reescrever (2.2.5) e (2.2.6) de forma a obtermos:

$$\begin{bmatrix} \Delta y_{1t} \\ \Delta y_{2t} \end{bmatrix} = \begin{bmatrix} -1 & \gamma \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \gamma u_{2,t-1} + u_{1t} \\ u_{2t} \end{bmatrix} \quad (2.2.11)$$

O sistema em (2.2.11) ilustra o fato de que em um sistema cointegrado, devemos incluir valores defasados das variáveis em nível no VAR explicando ΔY_t . As defasagens em nível devem aparecer na forma das combinações lineares de Y_t que são estacionárias.

Uma característica de variáveis cointegradas é que suas trajetórias ao longo do tempo são influenciadas pelos desvios do equilíbrio de longo prazo. Logo, a dinâmica de curto prazo deve ser influenciada pelos desvios da relação de equilíbrio de longo prazo.

Formalmente, o vetor $Y_t = (y_{1t}, y_{2t}, \dots, y_{mt})'$ tem uma representação de correção de erros se pode ser expresso na forma:

$$\Delta y_t = \pi_0 - \pi y_{t-1} + \pi_1 \Delta y_{t-1} + \pi_2 \Delta y_{t-2} + \dots + \pi_p y_{t-p} + \varepsilon_t \quad (2.2.12)$$

onde π_0 é um vetor de interceptos, π_i são matrizes de coeficientes, π é uma matriz com um ou mais elementos diferentes de zero e ε_t é um vetor de distúrbios com elementos ε_{it} , onde ε_{it} e ε_{jt} podem ser correlacionados.

Sejam todas as variáveis contidas em Y_t I(1). Se existe uma representação de correção de erros como em (2.2.12), existe necessariamente uma combinação linear das variáveis em Y_t que é estacionária. Resolvendo para πy_{t-1} , temos:

$$\pi y_{t-1} = \Delta y_t - \pi_0 - \sum_{i=1}^p \pi_i \Delta y_{t-i} - \varepsilon_t \quad (2.2.13)$$

Como todas as expressões do lado direito são estacionárias, πy_{t-1} é necessariamente estacionária. Como π contém apenas constantes, cada linha de π é um vetor de cointegração de Y_t , pois cada série $y_{i,t-1}$ é $I(1)$. Logo, um modelo de correção de erros para variáveis $I(1)$ implica, necessariamente, cointegração. Pode-se mostrar que o contrário também é verdadeiro. Este resultado é conhecido como Teorema da Representação de Granger, estabelecendo que para um conjunto de variáveis $I(1)$, correção de erros e cointegração são representações equivalentes.

Existem três pontos a serem destacados sobre a matriz π :

1. Se todos os elementos de π são zero, (2.2.12) é um VAR tradicional em primeira diferença, de forma que não há representação de correção de erros bem como cointegração.
2. Se um ou mais elementos de π for diferente de zero, ΔY_t responde a desvios do período passado em relação ao equilíbrio de longo prazo. Logo, estimar um VAR em primeira diferença é inapropriado. A omissão do termo πy_{t-1} leva a um erro de especificação, como foi visto anteriormente.
3. Este resultado ilustra as idéias de Johansen (1988) e Stock e Watson (1988) de que pode-se usar o posto da matriz π para determinar se as variáveis são ou não cointegradas.

Considere o caso de um VAR de primeira ordem:

$$y_t = A_1 y_{t-1} + \varepsilon_t \quad (2.2.14)$$

onde A_1 é uma matriz de parâmetros.

Subtraindo y_{t-1} dos dois lados de (2.2.14), obtém-se:

$$\begin{aligned} \Delta y_t &= -(I - A_1) y_{t-1} + \varepsilon_t \\ \Delta y_t &= \pi y_{t-1} + \varepsilon_t \end{aligned} \quad (2.2.15)$$

É fácil perceber que (2.2.15) é um caso especial de (2.2.12). Novamente o ponto crucial em relação a cointegração está ligado ao posto de π . Se o posto de π é zero, cada elemento de π deve ser zero. Neste caso, (2.2.15) é equivalente a um VAR em primeira diferença.

No outro extremo está o caso onde π tem posto completo. A solução de longo prazo para (2.2.15) é dada por n equações independentes. Neste caso, todas as variáveis de y_t são estacionárias.

Em casos intermediários, onde o posto de π é igual a r , para $r < n$, existem r vetores de cointegração.

2.2.1 Testes de Cointegração

- **A metodologia de Engle e Granger.**

Engle e Granger (1987) propõe uma maneira simples de testar se duas variáveis $I(1)$ são $CI(1,1)$. Sejam y_t e z_t duas variáveis $I(1)$, deve-se estimar a relação de longo prazo entre elas na forma:

$$y_t = \beta_0 + \beta_1 z_t + e_t \quad (2.2.16)$$

Se as variáveis são cointegradas, uma regressão por mínimos quadrados ordinários leva a estimadores super consistentes, como demonstrado por Stock (1988). Para determinar se as variáveis são mesmo cointegradas, deve-se testar se a série de resíduos estimados \hat{e}_t contém uma raiz unitária. Se os resíduos estimados forem estacionários, conclue-se que y_t e z_t são cointegradas de ordem $(1,1)$. Para isto, utiliza-se teste de raiz unitária sem intercepto, pois os resíduos devem ter médias iguais a zero:

$$\Delta \hat{e}_t = a_1 \hat{e}_{t-1} + \varepsilon_t \quad (2.2.17)$$

Não se pode utilizar a tabela de Dickey-Fuller, pois a seqüência $\{\hat{e}_t\}$ vem de uma regressão, ou seja, são estimativas e o pesquisador não sabe o verdadeiro valor de e_t , apenas sua estimativa. A tabela apropriada foi desenvolvida por Engle e Yoo (1987). Se os resíduos não parecerem ser ruídos brancos, pode-se realizar o teste de Dickey-Fuller aumentado ao invés de (2.2.17):

$$\Delta \hat{e}_t = a_1 \hat{e}_{t-1} + \sum_{i=0}^p a_{i+1} \Delta \hat{e}_{t-i} + \varepsilon_t \quad (2.2.18)$$

Se as variáveis são cointegradas, os resíduos da equação de equilíbrio de longo prazo podem ser usados para estimar o modelo de correção de erros:

$$\Delta y_t = \alpha_1 + \alpha_y \hat{e}_{t-1} + \sum_{i=1}^p \alpha_{11}(i) \Delta y_{t-i} + \sum_{i=1}^p \alpha_{12}(i) \Delta z_{t-i} + \varepsilon_{yt} \quad (2.2.19)$$

$$\Delta z_t = \alpha_2 + \alpha_z \hat{e}_{t-1} + \sum_{i=1}^p \alpha_{21}(i) \Delta y_{t-i} + \sum_{i=1}^p \alpha_{22}(i) \Delta z_{t-i} + \varepsilon_{zt} \quad (2.2.20)$$

As equações (2.2.19) e (2.2.20) podem ser estimadas usando a metodologia VAR, uma vez que o sistema é um VAR com um termo de correção de erro.

- **A metodologia de Johansen.**

Apesar de ser facilmente implementado, o procedimento de Engle e Granger tem algumas limitações. A estimação do equilíbrio de longo prazo requer que o pesquisador selecione uma variável como dependente e outra como explicativa. Em alguns casos, acontece de uma regressão indicar que as variáveis são cointegradas e quando se inverte a ordem, as variáveis não são mais cointegradas. Além disso, se tivermos mais de duas variáveis, podemos ter mais de um vetor de cointegração e o método não apresenta forma de separar estes vetores. Outra limitação é o fato de que o procedimento de Engle e Granger é feito através de estimação em dois estágios. Isto faz com que o erro de estimação do primeiro passo seja transferido para o segundo passo.

O método de estimação por máxima verossimilhança, introduzido por Johansen (1988), não necessita de dois estágios e, além do mais, possibilita testes para múltiplos vetores de cointegração. Outra vantagem deste procedimento é a possibilidade de restringir parâmetros no vetor de cointegração.

Considere a equação (2.2.15). O posto de π é igual ao número de vetores de cointegração.

O número de vetores de cointegração independentes pode ser obtido checando o nível de significância dos auto valores de π . É um resultado de álgebra matricial que o posto de uma matriz é igual ao número de seus auto-valores que são diferentes de zero. Suponha que obtenha-se a matriz π e ordene-se seus n auto-valores de forma que $\lambda_1 > \lambda_2 > \dots > \lambda_n$. Se as variáveis em y_t não são cointegradas, o posto de π será zero e todos os auto-valores serão iguais a um. Se as variáveis não são cointegradas, como $\ln(1)=0$, cada expressão $\ln(1-\lambda_i)$ será igual a zero. Se o posto de π for um, a primeira expressão $\ln(1-\lambda_1)$ será diferente de zero e as outras expressões $\ln(1-\lambda_2)=\ln(1-\lambda_3)=\dots=\ln(1-\lambda_n)=0$.

Na prática, obtém-se apenas estimativas dos auto-valores de π . O teste para o número de auto-valores significativamente diferente de um pode ser conduzido usando as seguintes estatísticas testes:

$$\lambda_{\text{traço}}(r) = -T \sum_{i=r+1}^n \ln(1 - \hat{\lambda}_i)$$

$$\lambda_{\text{max}}(r, r+1) = -T \ln(1 - \hat{\lambda}_{r+1})$$

A primeira estatística testa a hipótese nula de que o número de vetores de cointegração é menor ou igual a r , contra uma alternativa geral. Dessa forma, quanto mais longe os auto-valores estão de zero, maior é $\lambda_{traço}$.

A segunda estatística testa a hipótese nula de r vetores de cointegração, contra a hipótese alternativa de $r+1$ vetores de cointegração. Se as estimativas dos auto-valores forem próximas de zero, λ_{\max} será pequena. Os valores críticos para estas estatísticas são apresentados por Osterwald-Lenum (1992) e dependem do número de componentes não estacionários sob a hipótese nula e da inclusão ou não de constantes no vetor de cointegração.

2.3 Processos VAR Cointegrados com Mudanças de Regime Markoviano.

A idéia por trás da análise de sistemas lineares multivariados de séries de tempo é que os parâmetros do processo estocástico são variantes no tempo, porém condicionais à variável de regime, s_t . Essa variável indica a possibilidade de existência de diferentes estados de mundo. Seja M o número de regimes possíveis de forma que $s_t \in \{1, \dots, M\}$. Então a densidade de probabilidade condicional do vetor de série de tempo observado, y_t , é dada por:

$$p(y_t | Y_{t-1}, s_t) = \begin{cases} f(y_t | Y_{t-1}, \theta_1) & \text{se } s_t = 1 \\ \vdots & \vdots \\ f(y_t | Y_{t-1}, \theta_M) & \text{se } s_t = M \end{cases} \quad (2.3.1)$$

onde Y_{t-1} são as observações $\{y_{t-j}\}_{j=1}^{\infty}$. O vetor de parâmetros do VAR, θ , depende do regime que prevalece em t , de forma que θ_m é o vetor de parâmetros do VAR no regime $m=1, \dots, M$.

Logo, para um dado regime s_t , o vetor de série de tempo K dimensional $y_t = (y_{1,t}, \dots, y_{K,t})'$, $t=1, \dots, T$, é gerado por um vetor autoregressivo de ordem p (VAR(p)), tal que:

$$E[y_t | Y_{t-1}, s_t] = v(s_t) + \sum_{j=1}^p A_j(s_t) y_{t-j}$$

As inovações em y_t são dadas por:

$$u_t = y_t - E[y_t | Y_{t-1}, s_t]$$

O processo das inovações u_t é um ruído branco gaussiano com média zero e matriz de variância covariância $\Sigma(s_t)$.

$$u_t \square NID(0, \Sigma(s_t))$$

Os modelos de vetores autoregressivos com mudanças de regime markoviano (MS-VAR) são uma generalização dos modelos VAR de ordem finita:

$$y_t = v(s_t) + A_1(s_t)y_{t-1} + \dots + A_p(s_t)y_{t-p} + u_t \quad (2.3.2)$$

onde os valores amostrais y_0, \dots, y_{1-p} são fixos. As funções de mudança nos parâmetros $v(s_t), A_1(s_t), \dots, A_p(s_t)$ e $\Sigma(s_t)$ descrevem a dependência dos parâmetros do processo em relação ao regime em questão.

$$v(s_t) = \begin{cases} v_1 & \text{se } s_t = 1 \\ \vdots \\ v_M & \text{se } s_t = M \end{cases} \quad (2.3.3)$$

Como os parâmetros do processo condicional (2.3.2) dependem do regime vigente, que é, por hipótese, estocástico e não observável, uma descrição completa do processo gerador dos dados requer uma formulação do processo gerador dos regimes. Uma vez formulada uma lei para a geração dos estados, a evolução dos regimes pode ser inferida dos dados. Em modelos de mudança de regime markovianos, o processo gerador dos regimes é uma cadeia de Markov ergódica com número finito de estados $s_t = 1, \dots, M$ que são definidos pelas probabilidades de transição.

$$p_{ij} = \Pr(s_{t+1} = j | s_t = i), \sum_{j=1}^M p_{ij} = 1 \quad \forall i, j \in \{1, \dots, M\} \quad (2.3.4)$$

Mais precisamente, assume-se que s_t segue uma cadeia de Markov ergódica com M estados e matriz de transição P, tal que:

$$P = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1M} \\ p_{21} & p_{22} & \dots & p_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ p_{M1} & p_{M2} & \dots & p_{MM} \end{bmatrix} \quad (2.3.5)$$

onde $p_{iM} = 1 - p_{i1} - \dots - p_{iM-1}$ para $i = 1, \dots, M$.

Um VAR(p) com M regimes com mudança markoviana é chamado MS(M)-VAR(p). Uma grande vantagem dos modelos MS-VAR é sua grande flexibilidade (ver Krolzig(1997)). Na sua especificação mais geral, todos os parâmetros do MS-VAR são condicionais ao estado s_t da cadeia de Markov. No caso onde se têm apenas mudanças no intercepto do processo VAR(p), obtêm-se:

$$y_t = \sum_{i=1}^p A_i y_{t-1} + u_t + v(s_t) \quad (2.3.6)$$

onde $y_t = [y_{1t}, \dots, y_{kt}]'$, $v(s_t) = [v_1(s_t), \dots, v_k(s_t)]'$, os A_i 's são matrizes (kxk) de coeficientes e $u_t = [u_{1t}, \dots, u_{kt}]'$.

Subtraindo y_{t-1} dos dois lados de (2.3.2), se obtém o modelo de correção de erros vetorial com mudança markoviana de regime (MS-VECM).

$$\Delta y_t = v(s_t) + \sum_{i=1}^{p-1} \Gamma_i(s_t) \Delta y_{t-1} + \alpha(s_t) \beta' y_{t-p} + u_t \quad (2.3.7)$$

Onde $\Gamma_i = -\left(I_k - \sum_{j=1}^i A_j(s_t) \right)$, $i = 1, \dots, p-1$ e $\alpha(s_t) \beta' = \Pi = I_k - \sum_{j=1}^p A_j$. A matriz β'

é a matriz de cointegração.

2.3.1 Mudança de regime no *Drift* e na média de equilíbrio do VECM

Em modelos VAR(p) cointegrados, o intercepto v reflete, em geral, duas quantidades diferentes. Aplicando o operador de expectativa ao VECM em (2.3.7), obtém-se:

$$\Gamma(1) E[\Delta y_t] = v + \alpha E[\beta' y_t]$$

onde $\Gamma(1) = I_k - \Gamma_1 - \dots - \Gamma_{p-1}$. Logo, o intercepto é igual a $v = -\alpha \delta + \Gamma(1) \mu$, onde $\mu = E[\Delta y_t]$ e $\delta = E[cy_t]$ é uma constante determinando o equilíbrio de longo prazo e incluída na relação de cointegração. A cointegração implica na seguinte restrição à esperança do sistema em primeira diferença

$$E[\beta' \Delta y_t] = \beta' \mu = 0 \quad (2.3.8)$$

mostrando que m contém apenas k-r parâmetros livres, refletindo as tendências lineares comuns no sistema. Logo, μ pode ser parametrizado como $\mu = \beta'_{\perp} \mu^*$,

onde β'_{\perp} é uma matriz com posto completo ($k \times [k-r]$), ortogonal a β' , $\beta'\beta'_{\perp} = 0$, e μ^* é um vetor ($[k-r] \times 1$).

Se o intercepto pode ser absorvido pela relação de cointegração, as variáveis não têm tendências lineares determinísticas. Por outro lado, na ausência de qualquer restrição em v , existem $k-r$ tendências temporais produzindo um *drift* em y_t , tal que:

$$v = -\alpha\delta + \Gamma(1)\beta'_{\perp}\mu^* \quad (2.3.9)$$

Em sistemas cointegrados é, portanto, útil discriminar modelos com restrição absorvente, $v = -\alpha\delta$, e modelos irretritos com $\alpha \perp v \neq 0$.

Analogamente, uma mudança de regime no intercepto pode mudar a taxa de crescimento média e média de equilíbrio. Em modelos MS-VAR cada regime $m=1, \dots, M$ é associado com um atrator (μ_m^*, δ_m) .

$$v_m = -\alpha\delta_m + \Gamma(1)\beta'_{\perp}\mu_m^*$$

Portanto, diferentes especificações de modelos MS-VECM podem ser de interesse:

- Mudança irrestrita no intercepto $v(s_t)$

$$\Gamma(1)\Delta y_t = v(s_t) + \alpha\beta' y_{t-p} + u_t \quad (2.3.10)$$

- Mudança no drift $\mu(s_t)$

$$\Gamma(1)(\Delta y_t - \mu(s_t)) = \alpha(\beta' y_{t-p} - \delta) + u_t \quad (2.3.11)$$

onde δ é um vetor ($rx1$) de interceptos de cointegração e $\mu = \beta'_{\perp}\mu^*$ é um vetor ($kx1$) com $k-r$ parâmetros livres μ_m^* para cada regime m .

- Mudança no equilíbrio de longo prazo $\delta(s_t)$

$$\Gamma(1)(\Delta y_t - \mu) = \alpha(\beta' y_{t-p} - \delta(s_t)) + u_t \quad (2.3.12)$$

onde $\delta(s_t)$ é um vetor ($rx1$) de interceptos de cointegração e $\mu = \beta'_{\perp}\mu^*$ $\mu = \beta'_{\perp}\mu^*$ é um vetor ($kx1$) com $k-r$ parâmetros livres μ_m^* .

- Mudança contemporânea no *drift* e no equilíbrio de longo prazo.

$$\Gamma(1)(\Delta y_t - \mu(s_t)) = \alpha(\beta' y_{t-p} - \delta(s_t)) + u_t \quad (2.3.13)$$

onde $\delta(s_t)$ e $\mu(s_t)$ são definidos como em (2.3.11) e (2.3.12). A diferença para o modelo (2.3.10) é que em (2.3.13), após uma mudança de regime, tem-se uma alteração tanto no *drift* quanto na média do processo.

2.3.2 Representação Estável de Espaço-Estado.

Para introduzir a representação de estado-espço dos MS-VECM, faz-se uso da representação VAR(1) de uma cadeia de Markov (Hamilton, 1994). Introduce-se então as variáveis indicadores:

$$I(s_t = m) = \begin{cases} 1 & \text{se } s_t = m \\ 0 & \text{se } s_t \neq m \end{cases}$$

para $m=1, \dots, M$. A informação sobre o acontecimento de s_t pode ser coletada no vetor de regimes ξ_t :

$$\xi_t = \begin{bmatrix} I(s_t = 1) \\ \vdots \\ I(s_t = M) \end{bmatrix}$$

O vetor de regime ξ_t denota o estado não observável do sistema. Em Hamilton (1994), a representação VAR(1) é de uma cadeia de Markov ergódica, de forma que $\bar{1}_M \xi_t = 1$. Esta restrição de soma assegura a estabilidade. Para propósitos analíticos, elimina-se esta restrição de forma a obter-se uma formulação um pouco diferente da equação de transição. Este procedimento altera a representação de estado-espço considerada por Hamilton (1994), pois o novo vetor ζ_t tem dimensão $M-1$.

$$\zeta_t = \begin{bmatrix} \xi_{1,t} \\ \vdots \\ \xi_{M-1,t} \end{bmatrix} - \begin{bmatrix} \bar{\xi}_1 \\ \vdots \\ \bar{\xi}_{M-1} \end{bmatrix} \quad (2.3.14)$$

onde $\bar{\xi}$ é o vetor de probabilidades ergódicas da cadeia de Markov. O processo estocástico do vetor de regimes ζ_t é uma cadeia de Markov escondida, que pode ser representada por um processo VAR(1).

$$\zeta_t = F\zeta_{t-1} + e_t \quad (2.3.15)$$

onde a matriz $([M-1] \times [M-1])$ F , corresponde à matriz transposta de probabilidades de transição.

$$F = \begin{bmatrix} p_{1,1} - p_{M,1} & \cdots & p_{M-1,1} - p_{M,1} \\ \vdots & \ddots & \vdots \\ p_{1,M-1} - p_{M,M-1} & \cdots & p_{M-1,M-1} - p_{M,M-1} \end{bmatrix}$$

e v_t é um processo ruído branco não normal de média zero.

Conseqüentemente o termo de intercepto $v(s_t)$ em (2.3.7) é gerado por um processo estocástico do tipo:

$$v(s_t) = \bar{v} + M \zeta_t \quad (2.3.16)$$

onde \bar{v} é a média não condicionada de $v(s_t)$ e $M = [v_1 - v_M \quad \cdots \quad v_{M-1} - v_M]$ é uma matriz $(K \times [M-1])$.

Reparametrizando o mecanismo do VECM em (2.3.7) e combinando com (2.3.14) e (2.3.16), obtém-se a seguinte representação de estado-espço para MS(M)-VECM(p):

$$\begin{bmatrix} \Delta y_t \\ \Delta y_{t-1} \\ \vdots \\ \Delta y_{t-p+1} \\ \beta' y_{t-p} \\ \zeta_t \end{bmatrix} = \begin{bmatrix} \Gamma_1 & \cdots & \Gamma_{p-1} & \alpha\beta' & \alpha & M \\ I_K & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & I_K & 0 & 0 & \vdots \\ 0 & \cdots & 0 & \beta' & I_r & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & F \end{bmatrix} \begin{bmatrix} \Delta y_{t-1} \\ \Delta y_{t-2} \\ \vdots \\ \Delta y_{t-p} \\ \beta' y_{t-p} \\ \zeta_{t-1} \end{bmatrix} + \begin{bmatrix} \bar{v} \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} u_t \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ e_t \end{bmatrix} \quad (2.3.17)$$

2.3.3 Processos MS-VAR mais gerais.

Os modelos MS-VAR cointegrados podem ser generalizados de forma a representar diferentes tipos de mudanças de regimes no sistema. Além de mudança no intercepto, as matrizes de coeficientes, bem como a matriz de covariância também podem ser dependentes do regime.

$\Gamma_i(s_t)$ mudança nos coeficientes autoregressivos.

$\Sigma_i(s_t)$ heterocedasticidade dependente do regime.

$\beta'(s_t)$ mudanças na relação de cointegração, ou seja, no equilíbrio de longo prazo.

Enquanto mudanças nos coeficientes autoregressivos Γ_i são apenas uma extensão dos modelos MS-VECM tratados até aqui, mudanças na relação de cointegração ainda necessitam de mais investigação. Se a matriz de covariância é dependente dos regimes, então o termo de erro w_t do modelo VAR resultante se torna bilinear nas inovações e_t e u_t , de forma que:

$$w_t = \bar{\Sigma}^{\frac{1}{2}} e_t + M \left(\sum_{i=0}^{\infty} F^i e_{t-i} \right) + S \left(\sum_{i=0}^{\infty} F^i e_{t-i} \otimes e_t \right) \quad (2.3.18)$$

onde $e_t \square NID(0, I_K)$, $\Sigma_m^{\frac{1}{2}}$ é uma decomposição de Choleski triangular inferior de Σ_m , $\bar{\Sigma} = [\Sigma_1 \ \dots \ \Sigma_M] (\bar{\xi} \otimes I_K)$ é a média incondicional de $\Sigma_t = \Sigma(s_t)$ e S é definida como:

$$S = \begin{bmatrix} \Sigma_1^{\frac{1}{2}} - \bar{\Sigma} & \dots & \Sigma_{M-1}^{\frac{1}{2}} - \bar{\Sigma} \end{bmatrix}$$

A bilinearidade de w_t pode afetar o uso da metodologia de Johansen na estimação do modelo, apesar de que Johansen (1991) mostra que a hipótese de distribuição normal não compromete seriamente os resultados da análise assintótica.

2.3.4 Representação VARMA de processos MS-VAR cointegrados.

Em contraste com modelos gaussianos considerados por Johansen (1995), o intercepto aqui não é um simples parâmetro que adiciona inovações não-normais ao modelo gaussiano. O intercepto v é determinado pela função de mudança

$$v(s_t) = (v_1, \dots, v_M) \xi_t,$$

$$y_t = \sum_{i=1}^p A_i y_{t-p} + u_t + v(s_t). \quad (2.3.19)$$

Por isso o termo de intercepto não é um simples parâmetro, mas é gerado por um processo estocástico dado por (2.3.16). A representação VAR(1) da cadeia de Markov geradora dos regimes é dada em (2.3.15), daí pode-se chegar à representação MA(∞),

$$\xi_t = \sum_{j=0}^{\infty} F^j e_{t-j}. \quad (2.3.20)$$

Substituindo (2.3.20) e (2.3.16) em (2.3.19), tem-se

$$y_t = \sum_{i=1}^p A_i y_{t-p} + u_t + \bar{v} + M \sum_{j=0}^{\infty} F^j e_{t-j}. \quad (2.3.21)$$

Portanto a mudança de regime no intercepto implica um processo VAR cointegrado onde o termo de equilíbrio w_t é a soma de dois processos independentes, o ruído branco gaussiano u_t e o processo auto-correlacionado não normal $M \sum_{j=0}^{\infty} F^j e_{t-j}$,

$$y_t = \bar{v} + \sum_{i=1}^p A_i y_{t-p} + w_t \quad (2.3.22)$$

$$w_t = u_t + M \sum_{j=0}^{\infty} F^j e_{t-j} \quad (2.3.23)$$

Denotando $F(L) = I_{M-1} - FL$, o processo w_t pode ser escrito como:

$$w_t = u_t + MF(L)^{-1} e_t.$$

Usando a definição de matriz adjunta, $F(L)^* = |F(L)|F(L)^{-1}$, obtém-se:

$$|F(L)|w_t = |F(L)|u_t + MF(L)^* e_t.$$

Krolzig (1995) mostra que um processo estacionário w_t possui uma representação VARMA (M-1, M-1) com um ruído branco de média zero não normal ε_t ,

$$\gamma(L)w_t = \alpha(L)\varepsilon_t \quad (2.3.24)$$

onde $\gamma(L) = |F(L)| = |I_{M-1} - FL|$ e $\alpha(L) = I_K - \alpha_1 L - \dots - \alpha_{M-1} L^{M-1}$. Logo a equação (2.3.24) é a equação na forma final, que é identificável. Considere agora as implicações de (2.3.24) para as propriedades do processo gerador da variável observável y_t :

$$\begin{aligned} A(L)y_t &= \bar{v} + w_t \\ \gamma(L)w_t &= \alpha(L)\varepsilon_t \end{aligned}$$

A estrutura das inovações w_t resulta em um processo VARMA cointegrado de y_t da forma:

$$\gamma(L)A(L)y_t = \gamma(1)\bar{v} + \alpha(L)\varepsilon_t \quad (2.3.25)$$

ou escrito como o polinômio reduzido (KxK) $\Lambda(L) = \gamma(L)A(L)$ e com o vetor (Kx1) de constantes $a_0 = \gamma(1)\bar{v}$,

$$y_t = a_0 + \sum_{j=1}^{p+M-1} \Lambda_j y_{t-j} + \varepsilon_t + \sum_{i=1}^{M-1} \alpha_i \varepsilon_{t-i}. \quad (2.3.26)$$

Como na representação VARMA dos teoremas de representação de processos MS-VAR estacionários em Krolzig (1995), a ordem dada dos processos são apenas limites superiores, ou seja, não é assegurado que a matriz de transição irrestrita F é não singular, tendo posto $M-1$.

Pode-se perceber de (2.3.25) que um processo MS(M) VAR(p) cointegrado pode ser escrito na forma de um vetor autoregressivo de ordem infinita. Para perceber isto, multiplica-se os dois lados de (2.3.25) pelo polinômio $\alpha(L)^{-1}$, de forma a obter

$$\gamma(L)\alpha(L)^{-1}A(L)y_t = \gamma(1)\alpha(1)^{-1}\bar{v} + \varepsilon_t. \quad (2.3.27)$$

Definindo $\Psi(L) = I_K - \sum_{i=1}^{\infty} \Psi_i L^i = \gamma(L)\alpha(L)^{-1}A(L)$ como um polinômio AR infinito, obtém-se um sistema cointegrado de ordem infinita com inovações não normais,

$$y_t = \psi + \sum_{i=0}^{\infty} \Psi_i y_{t-i} + \varepsilon_t, \quad (2.3.28)$$

onde o termo de intercepto $\psi = \gamma(1)\alpha(1)^{-1}\bar{v}$ reflete a média incondicional de $v(s_t)$ e $\Psi(L)$ exhibe somente as raízes unitárias introduzidas por $A(L)$.

A idéia principal da equação (2.3.28) é que ela representa o sistema cointegrado com mudança markoviana de regime apresentado em (2.3.7) como um vetor autoregressivo não normal de ordem infinita. Tal processo gerador de dados permite que a análise de cointegração seja baseada nos procedimentos existentes para modelos VAR de ordem infinita.

2.3.5 Estimação

A técnica proposta para a estimação de máxima verossimilhança de processos MS-VECM é um procedimento em dois estágios. Primeiramente usa-se a metodologia máxima verossimilhança proposta por Johansen para determinar o número de vetores de cointegração, bem como para estimar a matriz de cointegração β' . Então se estima o restante dos parâmetros do MS VECM com o algoritmo EM para uma estimação de máxima verossimilhança.

A análise de sistemas cointegrados usando máxima verossimilhança com informação total desenvolvida por Johansen (1988, 1991) para vetores autoregressivos de ordem finita e gaussianos foi estendida para classes mais gerais de modelos. Saikkonen (1992) e Saikkonen e Luukkonen (1995) mostram que o uso de versões análogas do teste de razão de verossimilhança desenvolvido para vetores autoregressivos de ordem finita e gaussianos é justificado mesmo quando os dados são gerados por um processo de ordem infinita e não gaussiano. Saikkonen (1992) apresenta alguns resultados assintóticos gerais para processos VAR de ordem infinita que mostram que a maioria dos resultados assintóticos de Johansen (1988, 1991) para as relações de cointegração estimadas e para a matriz de pesos estimada continuam válidos. As propriedades assintóticas dos parâmetros de curto prazo estimados, bem como das funções impulso resposta para essa classe de mais geral de modelos é derivada em Saikkonen e Lütkepohl (1994) e Lütkepohl e Saikkonen (1996). O trabalho de Lütkepohl e Saikkonen (1996) é de especial importância para os resultados que serão apresentados aqui, uma vez que demonstra que a interpretação usual de VECM através de funções impulso resposta pode ser justificada mesmo quando os dados são gerados por um processo VAR de ordem infinita, mas o modelo ajustado é um VAR de ordem finita.

Portanto, como a não normalidade de ε_t não é essencial para a análise. Como enfatizado por Johansen (1991) e demonstrado em Saikkonen e Lütkepohl (1994), a hipótese de distribuição gaussiana não é muito importante no que concerne aos resultados da análise assintótica. Como os modelos MS VECM implicam que um processo VARMA gerou os dados, o procedimento de máxima verossimilhança de Johansen parece ser adequado. Em outras palavras, sob a hipótese de que o P.G.D. é um processo MS(M) VAR(p), não existe obstáculo em estudar as propriedades de longo prazo do sistema através da metodologia de Johansen para sistemas lineares.

A análise de cointegração é baseada na aproximação a um sistema linear. A partir dos resultados de longo prazo obtidos nesta estimação, constrói-se o modelo de correção de erros do processo MS(M) VAR(p) gerador dos dados:

$$\Delta y_t = \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \alpha(s_t) z_{t-p} + M \xi_t + u_t \quad (2.3.29)$$

onde $M = [v_1 - v_M \quad \dots \quad v_{M-1} - v_M]$ e $z_t = \beta' y_t$. Pode-se perceber que (2.3.29) corresponde a um modelo MS(M) VAR(p) com um regressor exógeno, z_t . A estimação de máxima verossimilhança de (2.3.29), portanto, pode ser feita através de um algoritmo EM como uma extensão do modelo sugerido por Hamilton (1990). A estimação do modelo é condicionada à matriz de cointegração estimada no primeiro período $\hat{\beta}'$.

Seja η_t o vetor das densidades de y_t , condicional a ξ_t e Y_{t-1}

$$\eta_t = \begin{bmatrix} p(y_t | Y_{t-1}; v_1, \Gamma_1, \dots, \Gamma_{p-1}, \alpha, \Sigma) \\ \vdots \\ p(y_t | Y_{t-1}; v_M, \Gamma_1, \dots, \Gamma_{p-1}, \alpha, \Sigma) \end{bmatrix} = \begin{bmatrix} p(y_t | \xi_t = \iota_1; Y_{t-1}; \theta) \\ \vdots \\ p(y_t | \xi_t = \iota_M; Y_{t-1}; \theta) \end{bmatrix},$$

onde $\theta = \text{vec}(M, \Gamma_1, \dots, \Gamma_{p-1}, \alpha, \Sigma)$ é o vetor de parâmetros do VECM, de forma que a densidade de y_t condicional a Y_{t-1} é dada por

$$p(y_t | Y_{t-1}) = \eta_t' \hat{\xi}_{t-1} = 1_M' (\eta_t \square \hat{\xi}_{t-1})$$

onde $\hat{\xi}_{t-1}$ é a inferência estatística $E[\xi_t | Y_t]$ sobre o vetor de estado não observável ξ dado o conjunto de informações Y_t . Então, condicional à matriz de cointegração β' , a função de verossimilhança do vetor de parâmetros λ , que contém o vetor de parâmetros do VECM, θ , os parâmetros da cadeia de Markov, $\rho = \text{vec}(P)$, e o estado inicial ξ_0 , é dada por:

$$L(\lambda | Y_T) = p(Y_T | \lambda) = \int p(Y_T, \xi | \lambda) d\xi = \int p(Y_T | \xi, \theta) \Pr(\xi | \rho, \xi_0) d\xi. \quad (2.3.30)$$

Devido à não linearidade resultante das condições de primeira ordem de λ , a maximização da função de verossimilhança é um problema de otimização não linear. Sua solução requer o uso de técnicas de solução numérica que maximizam a função log-verossimilhança de forma iterativa. Um algoritmo que resolve este problema é o algoritmo *Expectation-Maximization* (EM) introduzido por Dempster, Laird e Rubin (1977). Este algoritmo foi desenvolvido para uma classe geral de modelos onde a série de tempo observada depende de alguma variável estocástica não observável. Cada iteração deste algoritmo consiste de dois passos, o passo da expectativa e o passo da maximização. O primeiro passo é o passo da expectativa, onde o estado não observável ξ_t é estimado por suas probabilidades suavizadas $\hat{\xi}_{t|T} = E[\xi_t | Y_T]$;

onde todas as probabilidades condicionais $\Pr(\xi|Y, \lambda^{(j-1)})$ são calculadas usando o vetor de parâmetros $\lambda^{(j-1)}$ calculado no último passo de maximização. No passo de maximização, uma estimativa do vetor de parâmetros λ é derivada como solução das condições de primeira ordem associadas a (2.3.30), onde as probabilidades condicionais dos regimes $\Pr(\xi_t|Y, \lambda)$ são substituídas pelas probabilidades suavizadas $\Pr(\xi_t|Y, \lambda^{(j-i)})$ derivadas no último passo de expectativa. A cada passo obtém-se um novo vetor de parâmetros λ com o qual se atualizam as probabilidades filtradas e suavizadas. Então, cada iteração do algoritmo EM envolve uma passagem pelas iterações de filtragem e suavização, seguidas por uma atualização nas condições de primeira ordem e nos parâmetros estimados, garantindo um aumento na função de verossimilhança.

Capítulo 3 Resultados

Aqui serão apresentados os resultados dos modelos lineares e não lineares bem como a análise destes, com o intuito de entender o processo dinâmico de resposta da balança comercial brasileira a variações cambiais.

3.1 Dados

Os dados utilizados são mensais com início em janeiro de 1990 e término em dezembro de 2003, perfazendo um total de 168 observações. As variáveis utilizadas são a razão exportações/importações, a taxa de câmbio real, o Produto Interno Bruto brasileiro, o Índice Nacional de Preços ao Consumidor (INPC), as importações mundiais e o índice de preço das exportações mundiais.

Os dados para exportações totais e importações totais, usados para calcular a razão exportações/importações, usada como índice para o saldo da balança comercial, foram cedidos pela Secretaria de Comércio Exterior do Ministério do Desenvolvimento, Indústria e Comércio Exterior (MDIC/Secex). A taxa de câmbio real utilizada é calculada pelo Instituto de Pesquisa Econômica Aplicada (IPEA) através do expurgo do INPC e dos Índices de Preços por Atacado (IPAs) dos 16 mais importantes parceiros comerciais do Brasil da série nominal de taxa de câmbio (R\$ / US\$), ponderada pela participação de cada parceiro na pauta do total das exportações brasileiras em 2001. O PIB mensal utilizado foi estimado pelo Banco Central do Brasil (BC), deflacionado pelo INPC calculado pelo Instituto Brasileiro de Geografia e estatística (IBGE). As importações mundiais foram obtidas no *International Financial Statistics* do Fundo Monetário Internacional (IFS/FMI) e deflacionadas pelo índice de preços das importações mundiais, também do IFS/FMI. Algumas estatísticas descritivas das séries utilizadas estão descritas na Tabela 3.1.1 do Anexo.

A definição de balança comercial usada segue Haynes e Stone (1982), Bahmani-Oskoei (1991), Bahmani-Oskoei e Alse (1994), Brada, Kutan e Zhou (1997), Rincón(1998), Bahmani-Oskoei e Kantipong (2001), Antonuci (2003), entre outros. É utilizada a razão exportações/importações por vários motivos. Primeiramente, este índice permite expressar a balança comercial em logaritmo (Brada, Kutan e Zhou, 1997), de forma que a primeira diferença das variáveis reflète

a taxa de variação destas. Outra vantagem deste índice é sua invariância a alterações nas unidades de medida. Além disso, quando a balança comercial é definida como a diferença entre exportações e importações, utiliza-se um deflator para obter a balança comercial real. Logo, essa medida fica sensível ao deflator utilizado. O índice usado aqui resolve este problema, pois independe do deflator utilizado. Por isso, seguindo Banmani-Oskoe (1991) pode-se interpretar este índice como a balança comercial nominal ou real. As importações mundiais são usadas como *proxi* para a renda do resto do mundo.

Figura 3.1.1 - Gráfico da série $Inxm$.

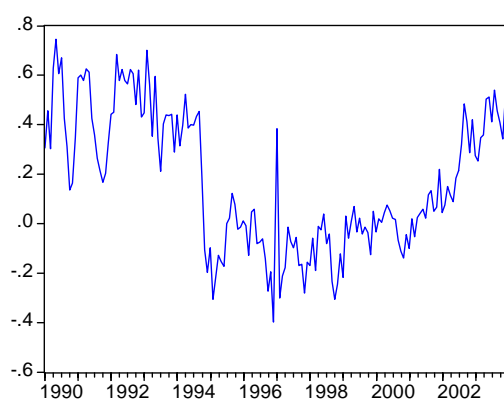


Figura 3.1.2 - Gráfico da série $Inreal$.

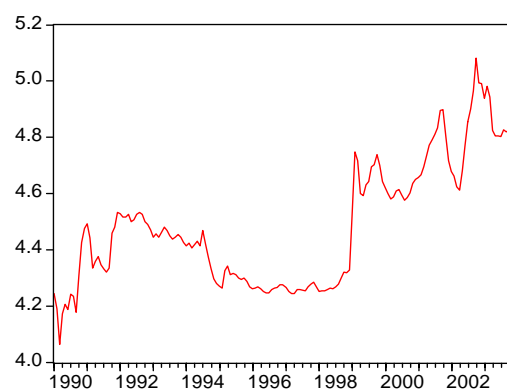


Figura 3.1.3 - Gráfico da série $Inpibr$.

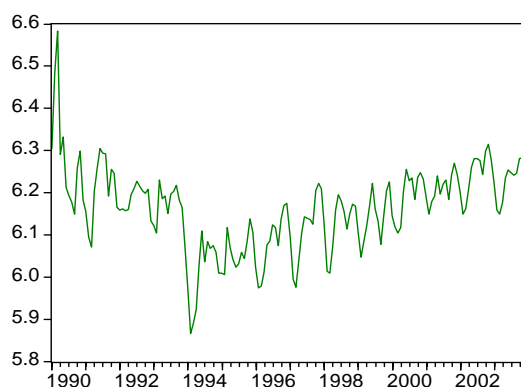
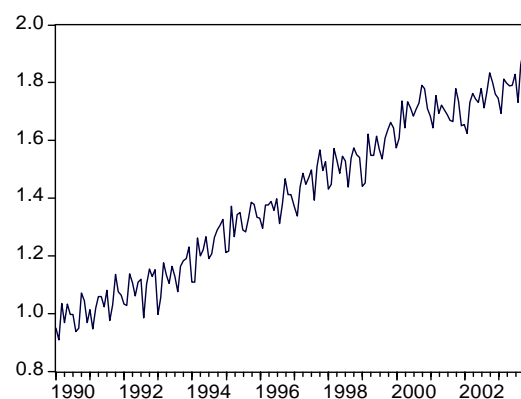


Figura 3.1.4 - Gráfico da série $Inimpr$.



A análise gráfica das séries sugere que elas são não estacionárias. Para confirmar esta suspeita, foram realizados testes de raiz unitária do tipo ADF nas séries em nível e em primeira diferença, os resultados se encontram na Tabela 3.1.2 do Anexo. Os resultados dos testes confirmam a análise gráfica, sendo que todas as séries contêm uma raiz unitária em nível e são estacionárias em primeira diferença.

A análise gráfica, bem como as mudanças na economia brasileira ocorridas no período de análise, sugerem que algumas séries possuem quebras estruturais, principalmente as séries do câmbio e da balança comercial. Como visto anteriormente, estas quebras estruturais podem distorcer os testes de raiz unitária. Neste caso, o teste mais indicado é o proposto por Perron(1997), que escolhe uma data para a quebra estrutural de forma endógena. Os resultados deste teste, bem como as datas de quebra estrutural indicadas por este se encontram nas Tabelas 3.1.3 e 3.1.4 do anexo. Os resultados reforçam a hipótese de não estacionariedade das séries.

3.2 Modelo Linear

Dado que as séries são não estacionárias, é necessário testar se estas cointegram, o que possibilitaria a estimação de um modelo de correção de erros. Antes disso é necessário saber quantas defasagens deve-se usar no teste. Para isso, estima-se um VAR com as variáveis em nível e escolhe-se o melhor modelo segundo os critérios de informação Akaike (AIC) e de Schwarz (BIC) (Ver Tabela 3.2.1 no Anexo). Seleciona-se quatro defasagens, pois este modelo apresenta valores satisfatórios para os dois critérios de seleção, além de possibilitar uma dinâmica para o modelo sem exagerar no número de parâmetros a ser estimados. Na seqüência, utiliza-se o teste de cointegração de Johansen (1998) com quatro defasagens, onde se constata a existência de um vetor de cointegração, os resultados estão na tabela 3.2.2 do Anexo. O vetor de cointegração encontrado, normalizado para $\ln x_m$, é dado por:

$$\beta = (1, -1.811856, 1.158394, 1.383791, -2.176329).$$

Ou seja, a equação que representa a relação de longo prazo entre as variáveis em questão é:

$$\ln x_m = 2.176329 + 1.811856 \ln \text{real} - 1.158394 \ln \text{impr} - 1.383791 \ln \text{pibr}. \quad ()$$

Esta equação revela que a elasticidade de longo prazo taxa de câmbio da balança comercial estimada é positiva. Dessa maneira, uma desvalorização (real) melhora a balança comercial (real). O coeficiente estimado mostra que para uma desvalorização cambial de 1%, *ceteris paribus*, a balança comercial aumenta 1.811856%, na média. Portanto, essa evidência empírica mostra que a condição de

Marshall-Lerner é válida para o Brasil. O sinal negativo do coeficiente da renda local é consistente com a abordagem da absorção, pois mostra que a renda tem uma relação negativa com a balança comercial. Nota-se, porém, que o coeficiente para a variável de renda externa é inconsistente com a abordagem da absorção, que prevê uma relação positiva entre balança comercial e renda externa.

Partindo da evidência de que existe apenas uma relação de longo prazo entre as variáveis estudadas, estima-se o modelo vetorial de correção de erros com quatro defasagens. Os resultados da estimação do modelo da balança comercial são:

Tabela 3.2.3 – Estimativas do VECM.

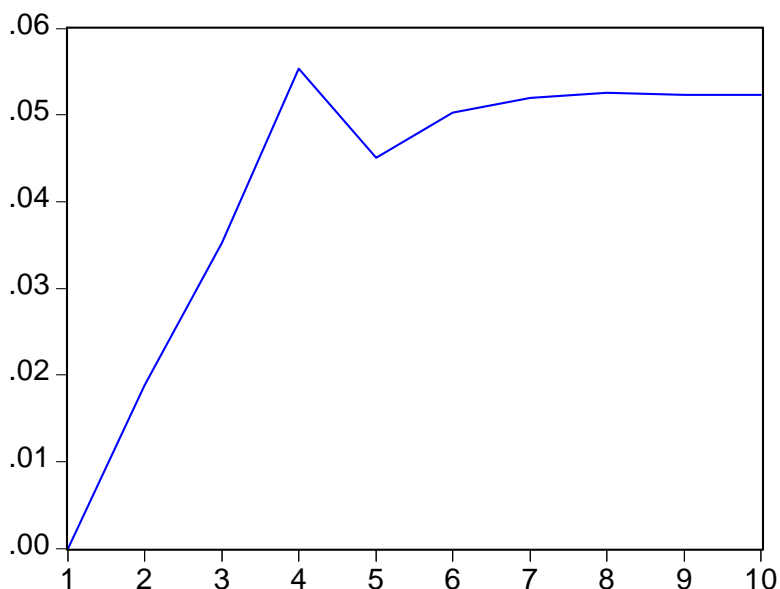
Variável	Coeficiente	Desvio Padrão	Estatística t
Constante	-0.0075	-0.01039	[-0.72165]
D(LNXM(-1))	-0.345346	-0.08286	[-4.16791]
D(LNXM(-2))	-0.005552	-0.08486	[-0.06543]
D(LNXM(-3))	0.078976	-0.08396	[0.94066]
D(LNXM(-4))	0.059713	-0.07797	[0.76580]
D(LNREAL(-1))	0.105307	-0.24902	[0.42288]
D(LNREAL(-2))	0.158494	-0.27091	[0.58505]
D(LNREAL(-3))	0.296454	-0.25768	[1.15046]
D(LNREAL(-4))	-0.377392	-0.24223	[-1.55799]
D(LNIMPR(-1))	-0.071862	-0.21965	[-0.32717]
D(LNIMPR(-2))	0.244983	-0.25481	[0.96142]
D(LNIMPR(-3))	0.34979	-0.23943	[1.46093]
D(LNIMPR(-4))	0.258443	-0.19742	[1.30909]
D(LNPIBR(-1))	0.303673	-0.25922	[1.17150]
D(LNPIBR(-2))	0.022641	-0.20727	[0.10923]
D(LNPIBR(-3))	-0.210954	-0.20083	[-1.05040]
D(LNPIBR(-4))	-0.029536	-0.20425	[-0.14461]
CointEq1	-0.207923	-0.04995	[-4.16258]

O coeficiente do termo de correção de erros é significativo a 1%. Isto significa que a velocidade na qual a variação da balança comercial, $\Delta \ln xm$, ajusta em direção ao único equilíbrio de longo prazo existente difere de zero. Ou seja, a equação para a $\Delta \ln xm$ contém informação sobre a relação de longo prazo entre as variáveis, uma vez que o termo de correção de erro entra nesta equação. De acordo com as estimativas, os desequilíbrios de curto prazo na balança comercial são corrigidos a uma velocidade de 20% por mês.

Uma forma de interpretar os coeficientes estimados em um vetor autoregressivo é usar funções de impulso resposta, que resumem o significado

destes coeficientes em gráficos. O maior interesse deste trabalho é analisar a dinâmica da resposta da balança comercial a variações no câmbio, o que pode ser analisado através de funções impulso resposta.

Figura 3.2.1 - Resposta de $lnxm$ a inovação de um desvio padrão em $lnreal$.



A função impulso resposta da balança comercial em resposta a um choque no câmbio mostra um aumento da balança comercial nos períodos subseqüentes ao choque no câmbio, sendo que este aumento atinge o máximo depois de aproximadamente quatro meses. Parece haver uma espécie de *overshooting* da balança comercial, pois após atingir o máximo quatro meses depois do choque no câmbio, há uma pequena retração na balança comercial, sendo que esta se estabiliza em um novo nível após aproximadamente dez meses. A função impulso resposta para este modelo linear de correção de erros rejeita a hipótese da curva J, uma vez que a balança comercial não apresenta deterioração no curto prazo.

3.3 Modelo com Mudança de Regime Markoviano.

O fato de que a economia brasileira passou por várias mudanças no período abordado, que afetaram as variáveis utilizadas neste trabalho, além da constatação visual (Figura 3.1) de que as séries parecem possuir quebras estruturais, sugerem a

utilização de modelos não lineares que contabilizem estas mudanças. Desta maneira, propõe-se a utilização de modelos MS-VEC.

Um problema que surge nos modelos MS-VEC é a determinação do número de regimes. Os testes de razão de verossimilhança para a determinação do número de regimes não têm distribuição assintótica padrão devido à existência de parâmetros incômodos. Hansen (1992) e Garcia (1993) desenvolveram um procedimento para a derivação da distribuição assintótica, porém esta não pode ser padronizada para todos os modelos. A distribuição que pode ser derivada é dependente dos dados e dos parâmetros do modelo, de forma que tem-se que calcular uma distribuição assintótica para cada teste. Além, o cálculo destas distribuições exige uma demorada simulação que se repete para cada especificação desejada, o que limita muito seu uso em trabalhos empíricos com modelos muito parametrizados (Krolzig,1997). Para a determinação do número de regimes foram utilizados critérios de informação, já que é demonstrado que os critérios de informação de Akaike e Schwartz nunca subestimam o número mínimo de regimes (Rydén,1995). Os resultados se encontram na Tabela 3.3.1 do Anexo.

O modelo mais apropriado parece ser o MSMH(2)-VECM(4), onde a média e a variância são dependentes do regime vigente. Neste caso, após uma mudança de regime há um salto imediato na média do processo:

$$\Delta y_t - \mu(s_t) = v + \sum_{i=1}^4 \Gamma_i [\Delta y_{t-i} - \mu(s_{t-i})] + \alpha z_{t-1} + u_t,$$

onde $s_t=1,2$, $z_t = \beta \hat{y}_t$ e $u_t \sim NID(0, \Sigma(s_t))$. Após a normalização do modelo não linear de forma a colocar $\ln x_m$ como variável dependente, obtém-se as seguintes estimativas para as elasticidades:

Tabela 3.3.2 – Estimativas do MSMH(2)-VECM(4).

	dxm	Desvio Padrão	t-valor
Média (Reg.1)	0.795951	0.2065	3.8544
Média (Reg.2)	0.788046	0.2044	3.8554
dxm_1	-0.3183	0.094	-3.3852
dxm_2	0.009631	0.0768	0.1253
dxm_3	0.119602	0.077	1.5541
dxm_4	0.121797	0.0738	1.6494
dreal_1	0.012341	0.2768	-0.0446
dreal_2	0.29702	0.2955	1.0053
dreal_3	0.350168	0.2881	1.2154
dreal_4	-0.32299	0.2394	-1.3492
dimpr_1	-0.12236	0.1965	-0.6227
dimpr_2	0.136474	0.2329	0.5860
dimpr_3	0.147492	0.237	0.6223
dimpr_4	0.180132	0.18	1.0010
dpibr_1	0.224533	0.2437	0.9214
dpibr_2	0.002029	0.2105	0.0096
dpibr_3	-0.19808	0.1994	-0.9932
dpibr_4	-0.16155	0.1981	-0.8154
z_1	-0.21396	0.0502	-4.2631
Desvio (Reg.1)	0.100029		
Desvio (Reg.2)	0.137412		

O coeficiente para a série exógena contendo as informações sobre a relação de longo prazo entre as variáveis é significativo ao nível de significância de 1%. Percebe-se que desvio padrão do segundo regime é 37% maior do que a do primeiro regime. Percebe-se também que a média do primeiro regime é maior que a do segundo regime. A maioria dos coeficientes apresenta sinal compatível com a abordagem da absorção para a balança comercial, porém apenas o coeficiente da primeira defasagem da balança comercial é significativo a 5%.

A Tabela 3.3.3 apresenta as datas estimadas de início e término de cada regime, enquanto os gráficos na Figura 3.3.1 e 3.3.2 abaixo apresentam as probabilidades suavizadas dos dois regimes:

Tabela 3.3.3 – Data dos Regimes.

Regime 1	Regime 2
Baixa Volatilidade	Alta Volatilidade
1991:01 - 1991:2	1990:06 - 1990:12
1991:07 - 1991:09	1991:03 - 1991:06
1992:01 - 1994:06	1991:10 - 1991:12
1994:08 - 1995:02	1994:07 - 1994:07
1995:07 - 1998:12	1995:03 - 1995:06
1999:11 - 2001:02	1999:01 - 1999:10
2003:05 - 2003:12	2001:03 - 2003:04

Figura 3.3.1 – Probabilidades suavizadas do regime 1.

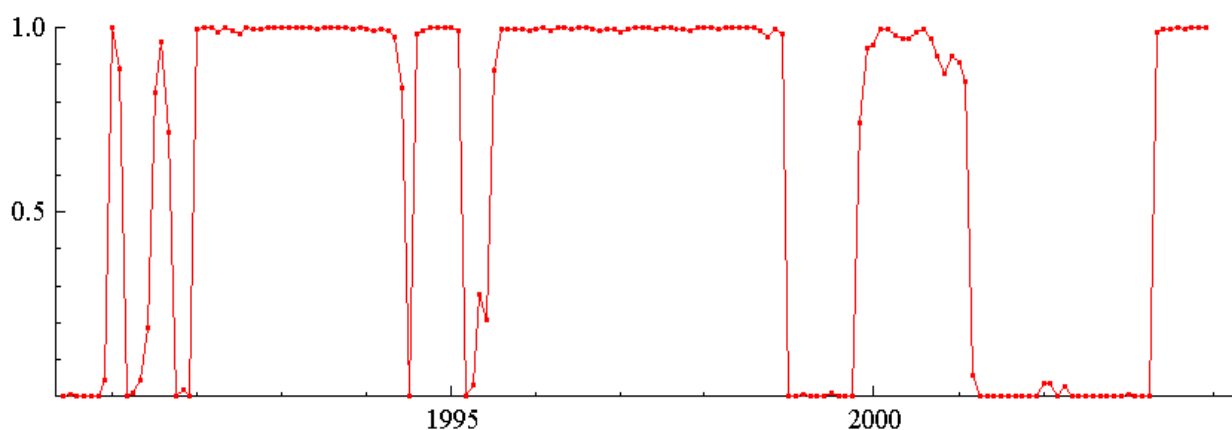
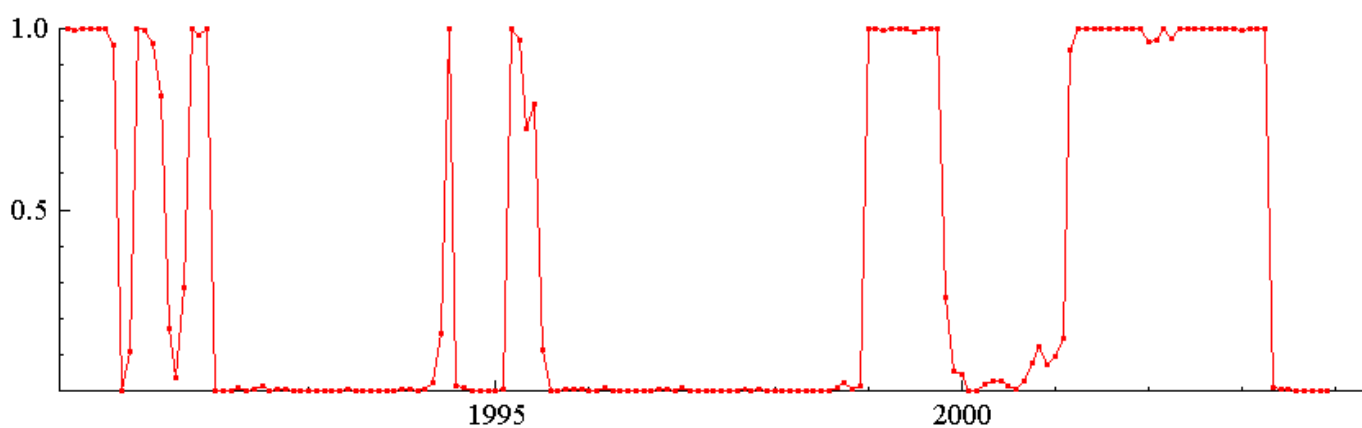


Figura 3.3.2 – Probabilidades suavizadas do regime 2.



O período analisado começa em janeiro de 1990 e termina em dezembro de 2003, porém, devido às quatro defasagens do modelo e à diferenciação dos dados,

se perde as cinco observações iniciais. Para uma recuperação da política cambial brasileira dessa época, utiliza-se Bonomo e Terra (1999), que fazem uma síntese dos acontecimentos que influenciaram a taxa de câmbio no período de 1964 a 1997, Garcia (2004), bem como várias edições dos relatórios trimestrais de análise do mercado de câmbio do Banco Central do Brasil.

Durante o período que vai de junho de 1990 a dezembro de 1991 esteve em vigor o Plano Collor 1, houve de confisco grande parte dos ativos líquidos da economia brasileira e suspensão dos pagamentos da dívida do governo. Neste período a taxa de câmbio real se manteve apreciada por grande parte do tempo, devido à alta taxa de inflação, o que diminuiu os superávits comerciais.

Em janeiro de 1991 a inflação ultrapassava 20% ao mês e o governo lançou o Plano Collor 2, baseado em congelamento de preços.

Em maio de 1991 a ministra da fazenda foi substituída e o controle de preços foi suspenso. Após a entrada da nova equipe econômica, a taxa de câmbio foi levemente desvalorizada e mantida constante até o final de setembro, quando houve uma desvalorização de 14% devido à queda nas reservas durante os meses anteriores.

A partir de setembro de 1991 o governo continuou desvalorizando lentamente a moeda até fevereiro de 1992, quando se interromperam as desvalorizações e o câmbio foi mantido nos mesmos níveis até meados de 1993.

De fevereiro de 1992 a meados de 1993 não houve desvalorizações. Em maio de 1993 Fernando Henrique Cardoso assumiu o ministério da fazenda e sua equipe econômica começou a formular o Plano Real. Em março de 1994 foi criada a unidade de medida URV, que rapidamente foi aceita como indexador de preços.

Em 1 de julho o Cruzeiro Real, foi extinto e substituído pelo Real, que era baseado na URV.

Na introdução da nova moeda, a paridade com o dólar era unitária. O governo decidiu não intervir no câmbio, o que gerou uma apreciação do Real que chegou a ser cotado US\$1,20 no final de outubro, devido ao influxo de capitais. Este nível da taxa de câmbio gerou uma queda nas exportações brasileiras, bem como um aumento das importações. O superávit comercial de US\$1 bilhão se transformou em um déficit de US\$700 milhões.

A partir de outubro de 1994 o governo começou a intervir no mercado de câmbio de forma a manter a taxa de câmbio dentro de uma mini-banda informal. Esta política se estendeu até fevereiro de 1995.

Em março de 1995 o câmbio foi desvalorizado em 6% e um regime de bandas cambiais foi anunciado. A taxa de câmbio poderia variar 5% dentro desta banda e o governo interviria no mercado se a taxa ultrapassasse este limite.

Após o estabelecimento de leilões de *spread* periódicos em julho de 1995, o câmbio se tornou muito estável e o Banco Central foi capaz de conduzir a taxa de câmbio dentro de limites muito menores que os estabelecidos pelo regime de bandas. Porém, devido ao grande déficit em conta corrente gerado pela política cambial, iniciou-se um debate sobre a sustentabilidade do regime de câmbio adotado.

A taxa de câmbio fixa e o grande déficit comercial gerado por esta, tornaram a economia brasileira extremamente dependente da entrada de capitais. As crises do México, da Ásia e da Rússia afetaram o influxo de capital para o Brasil, reduzindo as possibilidades de financiamento do déficit externo e gerando uma rápida queda nas reservas nacionais. O mercado de câmbio foi pressionado pela queda no ingresso de recursos financeiros via conta de capital e pelo fato de que serviços e vencimentos de obrigações externas aumentavam a demanda por moeda estrangeira. Nos últimos meses de 1998 o volume de emissão de bônus sofreu violenta contração, com estreitamento dos prazos e elevação dos *spreads*. É nesse contexto que se dá a desvalorização cambial de 15 de janeiro de 1999.

A desvalorização cambial foi seguida por especulações sobre a revisão do acordo com o FMI e sobre mudanças no ministério da fazenda e no Banco Central, declaração de moratória pelo governador de Minas Gerais, Itamar Franco e quedas históricas no índice BOVESPA. Devido a este quadro negativo passou a haver forte demanda por dólares, o que obrigou o Banco Central a vender grandes quantidades de moeda estrangeira. Mesmo com a intervenção do BACEN, no final de fevereiro a taxa de câmbio Real/Dólar já ultrapassava R\$2,00. O mercado de câmbio continuou volátil durante a maior parte do ano. Nos meses de março e abril a taxa recuou para R\$1,70, mas em agosto já estava em R\$1,95 devido a especulações sobre mudanças no regime cambial argentino, sendo que no final de setembro o dólar era negociado a R\$1,9223.

A partir de outubro de 1999 o câmbio deu sinais de recuperação, sendo que em 31 de dezembro o dólar era cotado a R\$1,7890. Até o início do primeiro trimestre de 2001 o mercado cambial operou relativamente estável, sendo que a cotação mensal média deste período foi R\$1,8279, com variância de apenas R\$0,0041.

No mês de março, as duas trocas de ministro da economia na Argentina em menos de um mês, como também as incertezas quanto aos efeitos da desaceleração da economia americana trouxeram uma maior instabilidade para o mercado de câmbio. No final de março a moeda americana era negociada novamente acima de R\$2,00. O dólar continuou a trajetória ascendente nos meses seguintes, sendo que em 19 de julho de 2001, o dólar chegou à cotação de R\$2,4675. No mês de setembro, os atentados terroristas nos E.U.A. trouxeram ainda mais instabilidade para o mercado cambial brasileiro, sendo que de janeiro a setembro de 2001 o Banco Central já tinha gasto US\$5,3 bilhões com intervenções cambiais, enquanto em todo o ano de 2002 as intervenções somaram US\$4,6 bilhões. Em fevereiro de 2002 o fim do regime de câmbio fixo na Argentina gerou muita incerteza quanto à reação do peso ao regime de câmbio flutuante. Neste contexto, os bancos no Brasil mantinham posições compradas em dólar como forma de proteção, o que continuou a pressionar o câmbio. No mês de julho o dólar fechou cotado a R\$2,84, refletindo as incertezas do mercado em relação ao crescimento da candidatura da oposição para as eleições de outubro. Os analistas financeiros temiam que, com a eventual vitória do candidato petista, poderia haver uma descontinuidade da ortodoxia fiscal e monetária utilizada pelo governo vigente. Com o aumento das incertezas no cenário político e a subsequente recomendação por parte de alguns bancos de investimento estrangeiros para que seus clientes diminuíssem a exposição a títulos da dívida externa brasileira, aumentou a demanda por dólar tanto pelo lado das empresas quanto pelo das tesourarias dos bancos. A vitória do partido oposicionista nas eleições de outubro de 2002 aumentou ainda mais a volatilidade do mercado de câmbio, sendo que o dólar chegou a ser cotado a R\$3,75. Esta incerteza só se dissipou a partir do segundo trimestre de 2003.

No final de março o comportamento da taxa de câmbio se tornou mais estável, devido à confiança dos agentes externos na determinação do novo governo em manter a austeridade econômica, inclusive com a realização de reformas estruturais importantes como a da previdência. O processo de reversão das expectativas culminou com uma acentuada desvalorização do dólar, impulsionada

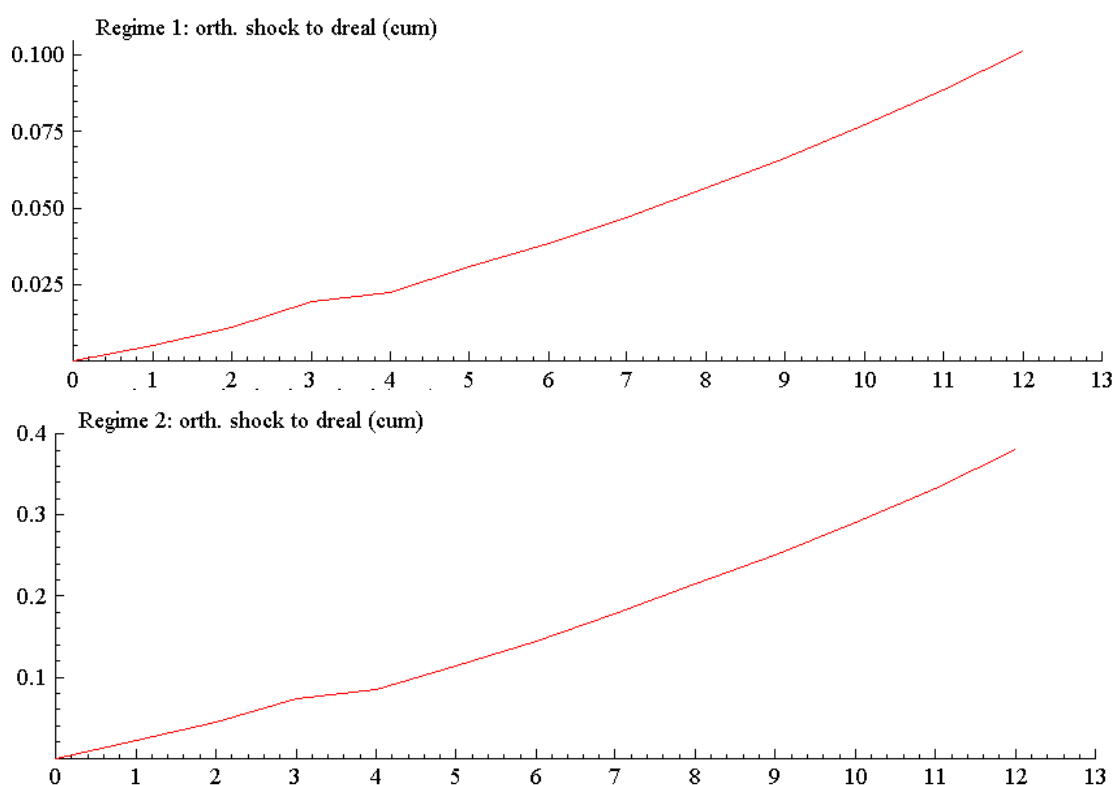
pela entrada de capitais estrangeiros motivados pelo alto diferencial entre os juros brasileiros e internacionais, sendo que em abril o dólar era negociado a R\$2,889. A confiança no novo governo e a estabilidade do cenário externo contribuíram para um período de tranquilidade no mercado cambial até o final de 2003.

O regime 2 parece coincidir com momentos de maior volatilidade no mercado de câmbio e instabilidade econômica. As datas estimadas para este regime são relativas a períodos de alterações no regime cambial, que acabam por gerar instabilidade no mercado de câmbio. Fazem parte deste regime os períodos de mini-desvalorizações do Plano Collor 1 e 2, o início do Plano Real e a primeira desvalorização da nova moeda que levou ao estabelecimento do regime de bandas cambiais, a liberalização do câmbio de 1999, bem como o período da crise Argentina que foi seguido pelo atentado de 11 de setembro e pela instabilidade do período eleitoral. Já o regime 1 se encaixa nos momentos de maior estabilidade, como o primeiro mandato do presidente Fernando Henrique Cardoso e o segundo semestre de 2003.

O fato do regime de alta volatilidade apresentar média mais baixa mostra que períodos de grande volatilidade cambial são caracterizados por um saldo menor na balança comercial. Teles (2003), usando um modelo teórico baseado na hipótese de histerese e abordando apenas o período entre julho de 1994 e maio de 2003, concluiu que a mudança de um regime de câmbio fixo para o regime de flutuação cambial alterou a taxa de câmbio limite sob a qual as firmas domésticas decidem aumentar sua produção para exportação. Esta parece ser uma explicação para a menor média durante o período de alta volatilidade.

Para analisar os resultados das estimativas do modelo MSMH(2) VECM(4), novamente se utiliza as funções impulso resposta. Como não há mudança nos parâmetros auto-regressivos do modelo, as funções impulso resposta apresentam a mesma dinâmica. Entretanto, a variância do primeiro regime é menor, o que faz com que o choque de um desvio padrão do segundo regime seja maior.

Figura 3.4 – Funções Impulso-Resposta do modelo MSMH(2)-VECM(4).



Como no modelo linear, não há evidência de curva J na dinâmica de ajustamento da balança comercial após uma desvalorização cambial. De acordo com a função impulso resposta, a condição de Marshall-Lerner é válida mesmo no curto prazo, apesar do impacto total do choque na taxa de câmbio sobre a balança comercial não se realizar no mesmo período da desvalorização, mas somente após mais de um ano.

Portanto, percebe-se que a modelagem não linear da relação entre a balança comercial e o câmbio é a mais adequada para o caso do Brasil. Isto ficou evidente através dos critérios de informação, bem como pelo ajustamento dos regimes aos períodos de alterações ou turbulências no mercado cambial brasileiro.

Conclusão

O objetivo deste trabalho é a análise empírica da relação de curto e longo prazo entre balança comercial e taxa de câmbio para a economia brasileira, no período compreendido entre janeiro de 1990 e dezembro de 2003, com o intuito de verificar a validade da hipótese da curva J.

Para atingir o primeiro objetivo foram usados modelos VECM, que possibilitam a análise de curto e longo prazo da relação entre as variáveis em questão. Como o coeficiente para o termo de correção de erro foi significativo ao nível de 1% de significância, existe forte evidência de uma relação de longo prazo entre balança comercial e taxa de câmbio. A relação de longo prazo estimada é positiva, o que confirma a condição de Marshall-Lerner. Para a análise da hipótese da curva J, se utiliza funções de impulso resposta, de forma a verificar se estas possuem um padrão temporal semelhante à letra J. Porém, não foi possível verificar uma piora na balança comercial após a desvalorização cambial, o que contradiz a hipótese da curva J.

Analisando-se os resultados dos coeficientes estimados para renda local, foi observado que somente no longo prazo eles estão em conformidade com a abordagem da absorção, que prevê coeficientes negativos para a renda local. No curto prazo o coeficiente estimado é positivo, o que contraria a teoria. O coeficiente de longo prazo estimado, bem como o de curto prazo, para a renda externa não estão de acordo com abordagem da absorção, uma vez que apresenta sinal negativo, contrariando a teoria.

Para contabilizar pelas várias alterações no regime cambial brasileiro no período, bem como pelos planos econômicos e choques externos ocorridos, utilizou-se a metodologia de modelos MS VECM. A vantagem deste método é poder estimar variações dos parâmetros ao longo do tempo, o que pode captar mudanças na relação entre as variáveis no período analisado. Além disso, esta é uma metodologia mais geral, sendo que o modelo linear VECM é apenas um caso dentre todas as especificações possíveis.

Os critérios de informação mostraram que o melhor modelo para os dados analisados é o MSMH(2) VECM(4). Este é um modelo vetorial de correção de erros com dois regimes na média e na matriz de variância. O primeiro regime é caracterizado por uma média maior e uma menor variância, captando os períodos de

maior estabilidade na economia brasileira e mundial. O segundo regime é caracterizado por menor média e maior variância, e se encaixa nos períodos de maior turbulência da economia brasileira.

Portanto, parece que quanto maior a volatilidade nas variáveis, menor é o saldo da balança comercial. Logo, uma menor volatilidade do câmbio parece ser mais favorável ao comércio internacional. Este resultado pode ser ligado a Teles (2003), que usando uma abordagem de histerese, chegou à conclusão de que o aumento na volatilidade do câmbio faz com que os exportadores reajam mais cautelosamente a uma desvalorização cambial.

Os resultados do modelo não linear confirmaram os resultados do modelo linear e mostra que a condição de Marshall-Lerner é válida mesmo no curto prazo. Além disso, as funções de impulso resposta não confirmam a hipótese da curva J.

Referências Bibliográficas.

Alexander, S. S. (1952) "Effects of a Devaluation on the Trade Balance", *IMF Staff Papers*, 2, 263-78.

Alexander, S. S. (1959) "Effects of a Devaluation: a simplified synthesis of the elasticities and absorption approaches", *American Economic Review*, 49, 21-42.

Antonucci, D. (2003). Effects of exchange rate changes on the Italian trade balance: the J-Curve. *CSC working paper*, n. 39.

Arize, A.C. (1994) Cointegration test of a long-run relation between the real effective exchange rate and the trade balance. *International Economic Journal*, 8, 1-9.

Backus D. K., Kehoe P. J., Kydland F. E. (1994), Dynamics of the Trade Balance and the Terms of Trade: the J-Curve? *American Economic Review* 84, 84-103.

Bacchetta, P., Gerlach, S., 1994. Sticky import prices and J-curves. *Economics Letters* 44 (3), 281–285.

Baharumshah, A.Z. (2001) The effect of exchange rate on bilateral trade balance: new evidence from Malaysia and Thailand. *Asian Economic Journal*, 15, 291-312.

Bahmani-Oskooee M. (1985), Devaluation and the J-Curve: some Evidence from LDCs, *Review of Economics and Statistics* 67, 500-504.

Bahmani-Oskooee, M. (1991) Is there a long-run relation between the trade balance and the real effective exchange rate of LDCs? *Economics Letters*, 36, 403-407.

Bahmani-Oskooee M., Alse J. (1994), Short-Run Versus Long-Run Effects of Devaluation: Error-Correction Modeling and Cointegration, *Eastern Economic Journal* 20, 453-464.

Bahmani-Oskooee, M., Kantipong, T. (2001) Bilateral J-Curve between Thailand and her trading partners. *Journal of Economic Development*, 6, 107-117.

Baldwin, A. (1988) "Hysteresis in Import Prices: The Beachhead Effect" *American Economic Review*, 78, p.773-785.

Banerjee, A., Dolado, J.J., Galbraith, W., Hendry, D.F., 1993. Cointegration, error-correction, and the econometric analysis of non-stationary data. Oxford University Press, Oxford.

Beveridge, S., Nelson, C. R. (1981). A new approach to decomposition of economic time series into permanent and transitory component with particular attention to measurement of the `business cycle`. *Journal of Monetary Economics*. 7, 151-174.

Bhagwati, J. (1958). Immiserizing Growth: A Geometrical Note. *Review of Economic Studies*, vol. 25.

Bickerdike, C. F. (1920).The Instability of Foreign Exchanges. *The Economic Journal*. Março.

Bonomo, M. , Terra, C. (1999). The Political Economy of Exchange Rate Policy in Brazil: 1964-1997. *FGV-Rio, texto para discussão*.

Box, G.E.P. and G.M. Jenkins (1976) *Time Series Analysis: Forecasting and Control. Revised Edition*. San Francisco: Holden-Day.

Brada J. C., Kutan A. M., Zhou S. (1993), China's Exchange Rate and the Balance of Trade, *Economics of Planning* 26, 229-242.

Brada, J.C., Kutan A.M., S. Zhou (1997), "The Exchange Rate and the Balance of Trade: The Turkish Experience," *Journal of Development Studies*, 33, 675-692.

Buluswar, M.D., Thompson, H. and Upadhyaya, K.P. (1996) Devaluation and the trade balance in India: stationarity and cointegration. *Applied Economics*, 28, 429-432.

Campbell, J.Y., P. Perron (1991) "Pitfalls and Opportunities: What Macroeconomists Should Know about Unit Roots", *NBER Macroeconomics Annual*, 141-200.

Campbell, J.Y. , Mankiw, N. G. (1987). Are output fluctuations transitory? *Quarterly Journal of Economics*. 102, 857-880.

Clark, P. K. (1987). The cyclical component of U.S. economic activity. *Quarterly Journal of Economics*. 102, 798-814.

Christiano, L. J. (1992). Searching for breaks in the GNP. *Journal of business and Economic Statistics*. 10, 237-250.

Davidson, J.E.H., D.F. Hendry, F. Srba and S. Yeo (1978) Econometric Modelling of the Aggregate Time-Series Relationship Between Consumer's Expenditure and Income in the United Kingdom, *Economic Journal*, 86, 661&692.

DeJong, D. N., Nankervis, J. C., Savin, N. E., Whiteman, C. H. (1990). Integration versus trend-stationarity in time series. *Mimeo, University of Iowa*.

Diebold, F.X. (1990) On 'Unit Root Econometrics': Discussion of Geweke, Sims, and DeJong and Whiteman, in: *Proceedings of the American Statistical Association Business and Economics Statistics Session*. Washington, D.C.: American Statistical Association.

Dickey, D.A., W.A. Fuller, 1979. Distributioir of the estimators for autoregressive time series with unit root. *Journal of the American Statistical Association*. 74,427-431.

Diebold, F.X. and M. Nerlove (1990) Unit Roots in Economic Time Series: A Selective Survey, in: T.B. Fomby and G.F. Rhodes, eds., *Advances in Econometrics: Co-Integration, Spurious Regressions, and Unit Roots*. Greenwich, CT: JAI Press.

Dixit A (1994), Hysteresis and the Duration of the J-curve, *Japan and the World Economy* 6, 105-115.

Demirden, T., I. Pastine (1995), "Flexible Exchange Rates and the J-Curve: An Alternative Approach," *Economics Letters*, 48, 373-377.

Dornbusch, R. (1973). Devaluation, Money and Nontraded Goods. *American Economic Review*. 5, 871-880.

Dornbusch, R. (1975). Exchange rates and fiscal policy in a popular model of international trade. *American Economic Review* 65, 859-871.

Dornbusch, R. (1987). Purchasing power Parity. *The New Palgrave Dictionary*, Stockton Press.

Dombusch. R., P. Krugman (1976). Flexible exchange rates in the short run. *Brookings Papers on Economic Activity*, 537-575.

Engle, R.F. (1984) Wald, Likelihood Ratio, and Lagrange Multiplier Tests in Econometrics, in: J. Griliches and M. Intriligator, eds., *Handbook of Econometrics*. North-Holland: New York, Vol. 2, 775-826.

Engle, R.F. and C.W.J. Granger, 1987, Co-integration and error correction: Representation, estimation and testing, *Econometrica* 55, 1251-1276.

Engle, R.F. e Yoo, B.S.(1987). Forecasting and testing in cointegrated systems. *Journal of Econometrics*, 35.

Evans, G.B.A. and N.E. Savin (1981), Testing for Unit Roots: I, *Econometrica*, 49, 753-779.

Evans, G.B.A. and N.E. Savin (1984), Testing for Unit Roots: II, *Econometrica*, 52, 1241-1260.

Felmingham B. S. (1988), Where is the Australian J-Curve?, *Bulletin of Economic Research* 40, 43-56.

Felmingham, B.S., and S. Divisekera (1986), "The Response of Australia's Trade Balance Under Different Exchange Rate Regimes," *Australian Economic Papers*, 25, June, 33-46.

Fialho, M. (2004). Interação entre as Políticas Monetária e Fiscal no Brasil sob a Ótica da Teoria Fiscal do Nível de Preços. *UFRGS*, Dissertação de mestrado.

Fuller, W.A. (1976) *Introduction to Statistical Time Series*. New York:Wiley.

Garcia, M. (2004) Vulnerabilidade externa e as crises do Plano Real sob câmbio flutuante: como escapar da armadilha dos juros reais muito elevados? Disponível em <http://www.econ.puc-rio.br/Mgarcia/>.

Garcia, R. (1993). Asynptotic Null Distribution of Likelihood Ratio Test in Markov Switching Models. *Université de Montréal, working paper*.

Granger, C.W.J. (1983) Co-Integrated Variables and Error-Correcting Models, UCSD Discussion Paper83-13.

Granger, C.W.J. and T.-H. Lee (1990) Multicointegration, *Advances in Econometrics*, 8, 71-84.Reprinted in: R.F. Engle and C.W.J. Granger, eds., *Long-Run Economic Relations: Readings in Cointegration*. Oxford University Press: New York, 1991.

Granger, C.W.J. and P. Newbold (1974) Spurious Regressions in Econometrics, *Journal of Econometrics*, 2, 11 I-20.

Granger, C.W.J. and P. Newbold (1976) *Forecasting Economic Time Series*. Academic Press: New York.

Granger, C.W.J. and A.A. Weiss (1983) Time Series Analysis of Error-Correcting Models, in: *Studies in Econometrics, Time Series and Multivariate Statistics*. Academic Press: New York, 255-78.

Gerlach S. (1989), Intertemporal Speculation, Devaluation, and the J-curve, *Journal of International Economics* 27, 335-345.

Greenwood, J. (1984) Non-traded goods, the trade balance, and the balance of payments. *Canadian Journal of Economics*, 17, 806-823.

Gupta-Kapoor, A. e Ramakrishnan, U. (1999) Is there a J-Curve? A new estimation for Japan. *International Economic Journal*, 4.

Hamilton, J.D. (1989).A new approach to the economic analysis of non stationary time series and business cycle. *Econometrica*. 57, 357-384.

Hamilton, J.D. (1994) *Time Series Analysis*. Princeton University Press: Princeton, NJ.

Hansen, B. E. (1992). The likelihood ratio test under non-standard conditions: testing the markov switching model of GNP. *Journal of Applied Econometrics*, 7, S61-S82.

Harberger, A. C. (1950). Currency Depreciation, Income, and the Balance of Trade. *Journal of Political Economy*. 58, 47-60.

Harvey, A. C. (1985). Trends and cycles in macroeconomic time series. *Journal of Business and Economics Statistics*. 3(3), 216-227.

Hendry, D.F. (1986), Econometric modelling with cointegrated variables: An overview. *Oxford Bulletin of Economics and Statistics* 48, 201–212.

Haynes, S. E., Stone, J. A. (1982). Impacts of the terms of trade on the U.S. trade balance: a reexamination. *The Review of Economics and Statistics*. 702-706.

Himarios, D. (1989), Do Devaluations Improve the Trade Balance? The Evidence Revisited, *Economic Inquiry*, 27, January, 143-168.

Hsing H., Savvides A. (1996), Does J-Curve Exist for Korea and Taiwan? *Open Economies Review* 7, 127-145.

Johansen, S. (1988). *Statistical Analysis of Cointegration Vectors*. *Journal of Economic Dynamics and Control*, 12.

Johansen, S., Juselius, K., 1990. Maximum likelihood estimation and inference on cointegration — with application to the demand for money. *Oxford Bulletin of Economics and Statistics* 52, 169–210.

Johansen, S. (1991). Estimation and hypothesis testing of cointegrating vectors in gaussian vector autoregressive models. *Econometrica*. 59, 1551-1580.

Johansen, S. (1995). Likelihood-Based Inference in Cointegrating Vector Autoregressive Models. *Oxford University Press*.

Juselius, K., 1995. Do purchasing power parity and uncovered interest rate parity hold in the long run? An example of likelihood inference in a multivariate time-series model. *Journal of Econometrics* 69, 211–240.

Junz H. B., Rhomberg R. R. (1973), Price Competitiveness in Export Trade Among Industrial Countries, *American Economic Review* 63, 412-418.

Kemp, M. C. (1970). The Balance of Payments and the Terms of Trade in Relation to Financial Control. *Review of Economic Studies*. 37, 25-31.

Koray, F. and McMillin, W. Douglas (1999). Monetary Shocks, the Exchange Rate, and the Trade Balance, *Journal of International Money and Finance* 18, no. 6, pp. 925-40.

Kremers, J.J., N.R. Ericson, and J.J. Dolado (1992). The Power of Cointegration Tests, *Oxford Bulletin of Economics and Statistics*, 54, 325-347.

Krolzig, H. M. (1995). Specification of Autoregressive Process with Markov Switching Regime Based ARMA Representation. *Humboldt Universität zu Berlin*, texto para discussão.

Krolzig, H. M. (1997). Markov-Switching Vector Autoregressions. *Springer-Verlag*, Berlin.

Krolzig, H. M. (1998). Econometric Modelling of Markov-Switching Vector Autoregression using MSVAR for Ox. *Institute of Economics and Statistics, and Nuffield College*, texto didático.

Krolzig, H. M., Toro, J. (2000). A New Approach to the Analysis of Business Cycle Transitions in a Model of Output and Employment. *Institute of Economics and Statistics, Oxford*, texto para discussão.

Krueger, A.D. (1983). *Exchange Rate Determination*, Cambridge University Press.

Krugman, P. (1987). Pricing to Market When the Exchange Rate Changes. *Em Real Financial Linkages Among Open Economies*, MIT Press.

Krugman P. R., Baldwin R. E. (1987). The Persistence of U.S. Trade Deficit, *Brookings Papers on Economic Activity* 1, 1-43.

Krugman, P.; Obstfeld, M. (1991). *International Economics*, 2nd ed., Harper Collins, NY.

Lal, A.K. and Lowinger, T.C. (2002) Nominal effective exchange rate and trade balance adjustment in South Asia countries. *Journal of Asian Economics*, 13, 371-383.

Lerner, A. P. (1944). *The Economics of Control: Principles of Welfare Economics*. *The Macmillan Company*, N.Y..

Lindert, P. H., Kindleberger, C.P. (1982). *International Economics*. *Irwin Series in Economics*, Homewood, IL..

Lütkepohl, H., Saikkonen, P.(1996). Infinite order cointegrated vector autoregressive processes: Estimation and inference. *Econometric Theory* 12, 814-844.

Lütkepohl, H., P. Saikkonen (1997). Impulse response analysis in infinite

order cointegrated vector autoregressive processes. *Journal of Econometrics* 81, 127-157.

Luukkonen, R. , P. Saikkonen (1995). Power of the Lagrange multiplier test for testing an autoregressive unit root. *Economics Letters* 51, 27-35.

Magee S. P. (1973), Currency Contracts, Pass-through, and Devaluation, *Brookings Papers on Economic Activity* 1, 303-323.

Mahdavi, S. and Sohrabian, A. (1993) The exchange value of the dollar and the US trade balance: an empirical investigation based on cointegration and Granger causality tests. *Quarterly Review of Economics and Finance*, 33, 343-358.

Mansoorian A. (1998), Habits and Durability in Consumption, and the Dynamics of the Current Account, *Journal of International Economics* 44, 69-81.

Marshall, A. (1923). Money, Credit and Commerce. *Macmillan*, London.

Marwah, K., and L.R. Klein (1996), Estimation of J-Curve: United States and Canada, *Canadian Journal of Economics*, 29, August, 523-539.

Meade, J. (1951). The Balance of Payments. *Oxford University Press*.

Metzler, L. (1948). A Survey of Contemporary Economics, vol. I. *Richard Irwin*, Homewood, IL..

Miles, M.A. (1979), "The Effects of Devaluation on the Trade Balance and the Balance of Payments: Some New Results," *Journal of Political Economy*, 87, June, 600-620.

Moffett M. H. (1989), The J-Curve Revisited: an Empirical Examination for the United States, *Journal of International Money and Finance* 8, 425-444.

Moraes, I., Portugal, M. (2003). Structural Change in the Brazilian Demand for imports: A Regime Switching Approach. *UFRGS*, texto para discussão, 2003/13.

Nelson, C. R., Plosser, C. I. (1982), Trends and random walks in macroeconomic time series. *Journal of Monetary Economics*. 10, 139-162.

Ng, S., P. Perron, (1995). Unit root tests in ARMA models with data dependent methods for the selection of the truncation lag. *Journal of the American Statistical Association* . 90, 268-281.

Onafowora, O. (2003), Exchange Rate and Trade Balance in East Asia: Is There a J-Curve?, *Economics Bulletin*.

Osterwald-Lenum, M. (1992). "A note with fractiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics: Four cases". *Oxford Bulletin of Economics and Statistics*, 54.

Peron, P. (1988). Trends and random walks in macroeconomic time series: Further evidence from a new approach. *Journal of Economic Dynamics and Control*. 12, 297-332.

Peron, P. (1989). The great crash, the oil price shock and the unit root hypothesis. *Econometrica*. 57, 1361-1401.

Peron, P. (1997). Further evidence on breaking trends functions in macroeconomic variables. *Journal of Econometrics*. 80, 355-385.

Perron, P., T-J. Vogelsang (1992). Nonstationarity and level shifts with an application to purchasing power parity. *Journal of Business and Economic Statistics*. 10, 301-320.

Pesaran, M.H., Shin, Y. (1996), Cointegration and speed of convergence to equilibrium. *Journal of Econometrics* 71, 117–143.

Phillips, P.C.B., Perron, P. (1988), Testing for a unit root in time series regression. *Biometrika* 75, 335–346.

Rahman, M. and Mustafa, M. (1996) The dancing of the real exchange rate of US dollar and the US real trade balance. *Applied Economics Letters*, 3, 807-808.

Rahman, M., Mustafa, M. and Burckel, D.V. (1997) Dynamics of the yen-dollar real exchange rate and the US-Japan real trade balance. *Applied Economics*, 29, 661-664.

Relatório Trimestral de Análise do Mercado de Câmbio. *Banco Central do Brasil*. Várias edições.

Rincón, H. (1998) Testing the short and long run exchange rate effects on the trade balance: The case of Colombia. *working paper University of Illinois*.

Robinson, J. (1947). *Essays in the Theory of Employment*. Basil Blackwell, London.

Rose A. K., Yellen J. L. (1989), Is There a J-Curve?, *Journal of Monetary Economics* 24, 53-68.

Rosensweig, J.A., and P.D. Koch (1988), The U.S. Dollar and the Delayed J-Curve, *Economic Review*, 73, Federal Reserve Bank of Atlanta, July/August, 2-15.

Rydén, T. (1995). Estimating the order of hidden markov models. *Statistics*, 26, 345-354.

Said, S.E., D.A. Dickey (1984). Testing for unit roots in autoregressive - moving average models of unknown order. *Biometrika* 71, 599-608.

Saikkonen, P. (1992). Estimation and testing of cointegrated systems by na autoregressive approximation. *Econometric Theory*. 8, 1-27.

Saikkonen, P., Luukkonen, R. (1997). Testing cointegration in infinite order cointegrated vector autoregressive processes. *Journal of Econometrics* 81, 93-126.

Salop, J.K. (1974). Devaluation and Balance of Trade Under Flexible Wages. *F.E.D. International Finance Discussion Paper*, n. 27.

Senhadji, A.S. (1998), Dynamics of the Trade Balance and the Terms of Trade in LDCs: The S-Curve, *Journal of International Economics*, 46, 105-131.

Stock, J. e Watson, M. (1988). Testing for comon trends. *Journal of the American Statistical Association*, 83.

Teles, Vladimir Kühl (2003), Choques Cambiais: Política Monetária e Equilíbrio Externo da Economia Brasileira em um Ambiente de *Hysteresis*, *Anais do XXV Encontro Brasileiro de Econometria*.

Watson, M. W. (1986). Univariate detrending methods with stochastic trends. *Journal of Monetary Economics*. 18, 1-27.

Wei, W.X. (1999) An empirical study of the foreign trade balance in China. *Applied Economics Letters*, 6, 485-490.

Apêndice

Tabela 3.1.1 – Estatísticas Descritivas.

Série	Obs	Mean	Std Error	Minimum	Maximum
Inxm	168	0.169773	0.276411	-0.396838	0.745195
Inreal	168	4.486688	0.222783	4.062519	5.081669
Inpibr	168	6.162157	0.098654	5.86521	6.583862
Inimpr	168	1.405084	0.275638	0.90829	1.916716

Tabela 3.1.2 – Teste ADF.

Série	Nível			Primeira diferença		
	τ_{τ}	τ_{μ}	τ	τ_{τ}	τ_{μ}	τ
Inxm	-2.118219	-2.261872	-1.888342	-18.7365	-18.74535	-18.80222
Inreal	-2.12053	-1.665742	1.171722	-9.561508	-9.597206	-9.502779
Inpibr	-3.413373	-3.212559	-0.305698	-7.547178	-10.62408	-10.65189
Inimpr	-2.533028	-0.359855	1.322549	-6.868682	-3.921123	-16.24827

Nota: Os valores em negrito denotam rejeição da hipótese nula de raiz unitária ao nível de significância de 5%.

Tabela 3.1.3 Teste de raiz unitária baseado em Perron (1997).

Modelo	Inxm	Inreal	Inpibr	Inimpr
1	-4.26519	-3.89801	-6.86492	-3.75716
2	-4.52955	-5.33981	-8.78159	-3.39941
3	-2.96843	-4.32458	-4.23698	-2.69416

Nota: Os valores em negrito denotam rejeição da hipótese nula de raiz unitária ao nível de significância de 5%.

Tabela 3.1.4 – Datas das quebras estruturais dadas pelo teste Perron (1997).

Modelo	Inxm	Inreal	Inpibr	Inimpr
1	1994-08	1994-06	1993-10	2001-09
2	1996-12	1998-11	1993-10	1999-11
3	1998-03	1997-05	1995-01	1999-10

Tabela 3.2.1 – Critérios AIC e BIC para seleção do número de defasagens.

Nº de Lags	AIC	BIC
8	-10.8387	-8.301666
7	-10.9106	-8.690424
6	-11.1417	-9.235736
5	-11.0883	-9.493935
4	-10.8902	-9.604841
3	-10.7539	-9.775046
2	-10.5183	-9.843404
1	-10.1921	-9.818636

Tabela 3.2.2 – Testes de Cointegração.

H0	$r = 0$	$r \leq 1$	$r \leq 2$
Autovalor	0.190723	0.134642	0.02508
Traço	62.23588	27.7427	4.170876
Valor crítico a 5%	47.21	29.68	15.41

Tabela 3.3.1 – Critérios de informação para a seleção do modelo.

Modelo	Critério de Informação		
	AIC	HQ	BIC
MSIAH(3)	-12,558	-10,616	-7,775
MSMAH(3)	-7,293	-5,351	-2,510
MSIA(3)	-12,020	-10,233	-7,617
MSMA(3)	-7,539	-5,751	-3,135
MSIH(3)
MSMH(3)	-11,863	-10,969	-9,661
MSI(3)	-11,761	-11,021	-9,939
MSM(3)
MSIAH(2)	-11,754	-10,475	-8,604
MSMAH(2)	-8,023	-6,744	-4,872
MSIA(2)	-11,710	-10,508	-8,749
MSMA(2)	-8,146	-6,944	-5,185
MSIH(2)	-12,090	-11,335	-10,230
MSMH(2)	-12,115	-11,360	-10,255
MSI(2)	-11,571	-10,892	-9,900
MSM(2)	-11,479	-10,801	-9,809
Linear	-11,494	-10,862	-9,937

Tabela 3.3.4 - Matriz de probabilidades de transição e duração dos regimes.

	Regime 1	Regime 2	n. obs	Prob.	Duração
Regime 1	0.9381	0.0619	106.8	0.6797	16.14
Regime 2	0.1315	0.8685	56.2	0.3203	7.61