

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
FACULDADE DE CIÊNCIAS ECONÔMICAS  
DEPARTAMENTO DE ECONOMIA

RELAÇÃO ENTRE AS VOLATILIDADES DA INFLAÇÃO E DO CÂMBIO NO  
BRASIL (1999-2010)

Paulo Ferreira Naibert

PORTO ALEGRE

2011

PAULO FERREIRA NAIBERT

RELAÇÃO ENTRE AS VOLATILIDADES DA INFLAÇÃO E DO CÂMBIO NO  
BRASIL (1999-2010)

Monografia submetida ao curso de graduação  
em Economia da Faculdade de Ciências  
Econômicas da UFRGS, como quesito  
parcial para obtenção do título de Bacharel  
em Ciências Econômicas

Orientador: Prof. Dr. Marcelo S. Portugal

PORTO ALEGRE

2011

PAULO FERREIRA NAIBERT

RELAÇÃO ENTRE AS VOLATILIDADES DA INFLAÇÃO E DO CÂMBIO NO  
BRASIL (1999-2010)

Monografia submetida ao curso de graduação  
em Economia da Faculdade de Ciências  
Econômicas da UFRGS, como quesito  
parcial para obtenção do título de Bacharel  
em Ciências Econômicas

Aprovado em: Porto Alegre, 28 de novembro de 2011.

---

Prof. Dr. Marcelo Savino Portugal - orientador  
UFRGS

---

Prof. Dr. Ronald Otto Hillbrecht  
UFRGS

---

Prof. Dr. Fernando Augusto Boeira Sabino da Silva  
UFRGS

## **AGRADECIMENTOS**

Agradeço aos meus pais por sempre terem me incentivado ao estudo.

Agradeço aos meus professores, especialmente ao professor Marcelo Portugal, o qual foi o meu orientador nessa monografia e na iniciação científica.

Agradeço aos colegas e amigos do Núcleo de Análise de Política Econômica (NAPE), Matheus Torma da Silveira, Affonso Reis, Luíza Andrade, Fernando Kuwer e Augusto Atz por ajudarem a criar um ambiente sério, calmo e propício ao estudo como criamos. Para alcançarmos essa proeza, agradeço em especial aos colegas de bolsa Victor de Fraga Sant'Ana, Bruno Martins, Fernando Schuck e Paulo Motula..

Também queria agradecer aos amigos do curso de Ciências Econômicas Alysson Portella e Guilherme Petersen, além do colega e amigo da UFRGS Cristian Alex Hoerlle, os quais foram muito importantes para mim nessa graduação.

## RESUMO

Há poucos estudos que tratam diretamente da relação entre a volatilidade da taxa de câmbio e da inflação, no Brasil; além disso há pouco consenso sobre esse assunto. Porém, com o início do regime de metas de inflação no Brasil, em 1999, é cada vez mais importante analisarmos as causas e os efeitos do nível de preços na economia brasileira. Tendo isso em vista, o presente trabalho se dispõe a analisar os efeitos da volatilidade da taxa de câmbio brasileira e suas relações com a volatilidade da inflação, no Brasil. Esse trabalho analisará a relação entre aquelas volatilidades por meio de um GARCH bivariado modelando as volatilidades condicionais de cada variável. Encontramos uma relação semi-côncava entre as séries, corroborando o trabalho já feito anteriormente por Albuquerque e Portugal (2006)

Palavras-chave: Taxa de Câmbio, Inflação, Volatilidade, Modelos GARCH.

**ABSTRACT**

There are few studies that directly address the relationship between exchange rate volatility and inflation volatility in Brazil. In addition, there is little consensus on this issue. But with the beginning of the inflation targeting regime in Brazil in 1999, it is increasingly important to analyze the causes and effects of the price level in the Brazilian economy. With this in mind, this paper sets out to analyze the effects of exchange rate volatility in Brazil and its relation with the volatility of inflation Brazilian inflation. This paper will analyze the relationship between these volatilities by a bivariate GARCH, which is able do the modeling of the conditional volatility of each variable. We found a semi-concave relation between the volatilities, confirming the work previously done by Albuquerque and Portugal (2006)

Key-words: Exchange rate, Inflation, Volatility, GARCH models.

## LISTA DE GRÁFICOS

Gráfico 4.1 – Taxa de inflação no Brasil pelo IPCA de Julho de 1999 a dezembro de 2010 (Índice base fixa Jul/1999=100).....	54
Gráfico 4.2 – Taxa de câmbio R\$/US\$, preços de venda, média mensal.....	54
Gráfico 4.3 – Primeira diferença dos logaritmos do IPCA.....	55
Gráfico 4.4 – Primeira diferença dos logaritmos da taxa de câmbio R\$/US\$.....	56
Gráfico 4.5 – Variâncias Condicionais.....	60
Gráfico 4.6 – Relação entre a variância condicional do IPCA e a variância condicional da taxa de câmbio.....	62
Gráfico 4.7 – Relação entre a variância condicional do IPCA e a variância condicional da taxa de câmbio com amostra reduzida.....	63
Gráfico 4.8 – Relação entre a variância condicional do IPCA e a variância condicional da taxa de câmbio com amostra reduzida.....	63

**LISTA DE TABELAS**

Tabela 3.1 – Resumo das FAC e FACP dos modelos ARMA.....	34
Tabela 4.1 – Teste ADF nas séries utilizadas.....	55
Tabela 4.2 – Equações da Média do IPCA.....	57
Tabela 4.3 – Equações da Média da taxa de câmbio.....	57
Tabela 4.4 – Coeficientes das matrizes do GARCH bivariado.....	59
Tabela 4.5 – Teste Ljung-Box dos resíduos do IPCA e do Câmbio.....	61
Tabela 4.6 – Teste ARCH-LM dos resíduos do IPCA.....	61
Tabela 4.7 – Teste ARCH-LM dos resíduos da taxa de câmbio.....	62

## SUMÁRIO

<b>1</b>	<b>Introdução.....</b>	<b>11</b>
<b>2</b>	<b>Inflação e taxa de câmbio.....</b>	<b>14</b>
2.1	Relação entre câmbio e inflação.....	14
2.2	Regime de câmbio, volatilidade cambial e volatilidade da inflação.....	16
2.3	Relação entre volatilidade do câmbio e volatilidade da inflação.....	20
<b>3</b>	<b>Modelos Econométricos de Séries de Tempo: ARIMA e GARCH.....</b>	<b>27</b>
3.1	Estacionariedade Fraca e Modelos ARMA.....	27
3.1.1	Processo de Médias Móveis.....	28
3.1.2	Processo Auto Regressivo.....	30
3.1.3	ARMA.....	32
3.1.4	Diferenciação de processos não estacionários.....	33
3.1.5	Critério de Informação.....	34
3.1.6	Processos Não-Estacionários.....	36
3.1.7	Raízes Unitárias.....	38
3.2	Modelos GARCH.....	41
3.2.1	ARCH.....	42
3.2.2	GARCH.....	43
3.2.3	Testes de detecção.....	45
3.3	GARCH Bivariado.....	47
<b>4</b>	<b>Aplicação na Economia Brasileira.....</b>	<b>53</b>
4.1	Dados.....	53

4.2	Formulação da equação da média.....	56
4.3	GARCH bivariado.....	58
<b>5</b>	<b>Conclusão.....</b>	<b>65</b>
<b>6</b>	<b>Referências.....</b>	<b>67</b>

## 1 Introdução

Após o início do regime de metas de inflação no Brasil, em 1999, é cada vez mais importante analisarmos as causas e os efeitos do nível de preços na economia brasileira. Com isso em mente, o presente trabalho se dispõe a analisar os efeitos da volatilidade da taxa de câmbio brasileira e suas relações com a volatilidade da inflação, no Brasil. Além disso, outras razões para tal investigação é que não há muito consenso, nem muitos estudos sobre este assunto. Entretanto, este tema já foi objeto de análise em Albuquerque e Portugal (2006), os quais fizeram uma análise para o Brasil no período que vai de 1999 a 2004. Aquele trabalho será a referência principal na qual este trabalho será baseado.

Portanto, como foi exposto, o objetivo desse trabalho é estimar a relação entre a volatilidade do câmbio e a da inflação, no Brasil. O período a ser utilizado vai de 1999 a 2010. Na literatura sobre o assunto, Albuquerque e Portugal (2006) dizem que acharam dois principais grupos com opiniões diversas sobre o efeito da volatilidade da taxa de câmbio sobre a volatilidade da inflação. O primeiro grupo diz que a volatilidade da taxa de câmbio não tem grandes efeitos para outras variáveis macroeconômicas. Já o segundo grupo encontra evidências da existência de uma relação, seja ela positiva ou negativa, nos estudos realizados sob diferentes enfoques e métodos.

Ademais, como o presente trabalho se baseia em Albuquerque e Portugal (2006), a hipótese inicial é que a relação entre as duas volatilidades é a mesma encontrada por Albuquerque e Portugal (2006), qual seja, a relação entre as volatilidades é semi-côncava. Isso, em termos econômicos, se traduziria em dizer que quando a volatilidade da taxa de câmbio é muito alta, o aumento da incerteza traz uma resposta à inflação reduzida, o que levaria a efeitos menores. Para verificarmos isso, pretendemos usar um GARCH bivariado para modelar a relação entre as volatilidades da inflação e do câmbio. Outro motivo para realizarmos tal análise é que, no trabalho de Albuquerque e Portugal (2006), é expressa a ideia de que os autores necessitariam de uma amostra maior para verificarem se o comportamento dessa relação se manteria ao longo do tempo, verificação que não poderia ser feita naquela época. No entanto, cinco anos passados desse trabalho, já podemos realizar esse tipo de teste.

No trabalho de Albuquerque e Portugal (2006), os autores estimaram a relação entre as volatilidades das variáveis citadas com dois métodos: volatilidade não-condicional (*rolling windows*) e volatilidade condicional (GARCH bivariado). O primeiro método consiste em pegar uma sub-amostra de tamanho fixo, estimar a volatilidade desta sub-amostra e fazer isso sucessivamente para toda amostra. Entretanto, isso traz o problema de se perder pontos da amostra e das volatilidades mudam bastante de acordo com os critérios empregados pelos pesquisadores, tais como tamanho de janelas e a particularidade da sub-amostra escolhida, o que traz estimações não muito confiáveis. Já com o segundo método, pelo menos teoricamente, a medida da volatilidade não se alterará de acordo com os critérios individuais de escolha do autor. Ademais, este último seria um método econométrico mais adequado para ser estimado. E é esse método que utilizaremos para estimarmos as volatilidades das variáveis de interesse.

A ideia básica de modelar a volatilidade é prever e analisar o tamanho dos erros dos modelos. Esses erros podem ser modelados com as ferramentas ARCH e GARCH. Esses modelos se tornaram importantes porque apesar de o modelo básico da econometria para a estimação de séries de tempo, o MQO (mínimos quadrados ordinários), modelar como uma variação de uma variável afeta a variação de outra, ele tem como hipótese básica que o valor esperado dos erros, quando elevado ao quadrado, é uma constante. Essa hipótese é chamada de homocedasticidade (Engle, 2001). E ela não se verifica em muitas séries de tempo, especialmente com as que iremos trabalhar (taxa de câmbio e inflação).

Além disso, outro problema de a homocedasticidade não ocorrer, e os erros forem heterocedásticos, como o valor esperado dos erros, quando elevado ao quadrado, não é o mesmo em qualquer ponto, quando formos verificar algum intervalo de confiança, ou tentar fazer alguma inferência estatística, erraremos o desvio quadrado da amostra e o nosso intervalo de confiança estará errado (o problema realmente aparece quando estimamos um desvio quadrado menor do que ele realmente é naquela hora e o intervalo de confiança fica muito estreito e temos uma falsa impressão de confiança). Então, em vez de considerarem isso como se fosse um erro a ser corrigido, os modelos GARCH e ARCH tratam a heterocedasticidade como uma variância a ser modelada. Disso, resulta que não apenas as deficiências do modelo de MQO são corrigidas, mas também a previsão da variância de cada termo do erro é computada (Engle, 2001).

Este trabalho é composto de três capítulos, além de uma introdução e uma conclusão. No primeiro capítulo apresentamos a revisão bibliográfica sobre as relações entre taxa de câmbio e inflação. No capítulo dois fazemos uma exposição da teoria econométrica que necessitaremos para a análise do capítulo três. No último capítulo antes da conclusão fazemos a análise sobre a relação entre a volatilidade da inflação e a volatilidade cambial a partir de um GARCH bivariado. E, finalmente, apresentamos as nossas conclusões.

## 2 Inflação e taxa de câmbio

Nesse capítulo, vamos apresentar a revisão bibliográfica para examinarmos a literatura sobre a relação entre taxa de câmbio e inflação e, principalmente entre as volatilidades dessas variáveis. Esse capítulo se divide em 3 seções: 1) Relação entre câmbio e inflação 2) Regimes de taxa de câmbio, volatilidade do câmbio e volatilidade da inflação 3) Relação entre volatilidade do câmbio e volatilidade da inflação.

### 2.1 Relação entre câmbio e inflação

Antes de entrarmos na discussão sobre a relação entre a volatilidade cambial e da inflação, faremos uma exposição sobre a relação entre câmbio e inflação. Quando discutimos a relação entre essas duas variáveis, temos de falar sobre a relação paridade poder de compra (PPC) que relaciona câmbio e nível de preços. Segundo Krugman e Obstfeld (2010), em sua forma absoluta, a teoria da paridade do poder de compra afirma que a taxa de câmbio entre as moedas dos países é igual a razão entre seus níveis de preços, medidos pelos preços nominais de uma cesta de mercadorias de referência. Apresentando a seguinte equação

$$P_A^i = (E_{A/B}) \times (P_B^i)$$

$$E_{A/B} = P_A^i / P_B^i$$

Onde  $P_A^i$  é o preço do bem  $i$  no país A,  $P_B^i$  é o preço do bem  $i$  no país B e  $E_{A/B}$  é a taxa de câmbio entre a moeda do país A e a moeda do país B.

Em outras palavras, poder de compra de qualquer moeda é o mesmo em qualquer país. Essa seria ideia por trás da lei do preço único, a qual afirma que, sob livre concorrência e na falta de impedimentos ao comércio, um bem será vendido a um preço único em qualquer parte do mundo.

Entretanto há alguns problemas com a PPC que Krugman e Obstfeld (2010) elencam. Problemas como barreiras comerciais e produtos não-comercializáveis, práticas monopolistas e oligopolistas, cestas de mercadorias diferentes nos índices de preços. O primeiro problema traz dificuldades para a relação da paridade do poder de compra visto que com barreiras comerciais, fica mais difícil para os agentes conseguirem comprar os bens nos outros países, praticar arbitragem e fazer a lei do preço único valer, o que leva a desvios da relação PPC. Além disso, o problema com os produtos não-comercializáveis é que em vez de ficar difícil de comprar os bens em outros países, fica simplesmente impossível de se fazer isso, e a lei do preço único passa a ser inválida e o poder de compra de qualquer moeda dada cairá nos países em que os preços dos bens não-comercializáveis se elevarem. O segundo problema traz dificuldades para a relação PPC visto que devido a essas estruturas de mercado, a lei do preço único não vale e há desvios da PPC. E finalmente, o terceiro problema é que uma vez que a PPC é feita com índices de preços ao consumidor, e cada país tem um índice de preços que usam cestas de bens diferentes, devido as diferenças entre os padrões de consumo de cada país – para uma discussão mais aprofundada sobre esse terceiro problema ver Frenkel (1976).

Além disso, a PPC absoluta implica uma segunda versão desta, a PPC relativa, segundo a qual as variações percentuais das taxas de câmbio eliminam as diferenças entre as taxas de inflação nacionais. Cujas equação é a seguinte:

$$\frac{\frac{E_A}{B^t} - \frac{E_A}{B^{t-1}}}{\frac{E_A}{B^{t-1}}} = \pi_{A,t} - \pi_{B,t}$$

Onde  $\pi_t$  é a taxa de inflação, ou seja, é a mudança no índice de preço. Ou seja:

$$\pi_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

A partir dessas equações, podemos observar que se a taxa de câmbio variar, ela pode levar os preços a variarem. Essa é a ideia por trás do repasse do câmbio para a inflação (Pass through). Isso é definido por Krugman e Obstfeld (2010) como sendo a porcentagem na qual os preços das importações aumentam quando a moeda local se deprecia em um por cento.

Esse repasse pode ser incompleto, um dos trabalhos que argumenta dessa forma é Devereux e Engel (2002), os quais dizem que “Evidências empíricas indicam que taxas de câmbio nominais não têm repasse completo do câmbio para os preços de bens. De fato, parece que os preços do consumidor não respondem a mudanças na taxa de câmbio”.

Krugman e Obstfeld (2010) oferecem a explicação de mercados monopolisticamente competitivos e medo da empresa de perder fatias de mercado no país importador (empresas determinam o preço onde os produtos são produzidos e pode haver demora entre a produção e a comercialização do produto).

Outro aspecto que influencia o repasse cambial para a inflação é o grau de desenvolvimento do país, por exemplo, Calvo e Reinhart (2000) acham evidências de que “O repasse cambial à inflação é bem maior em mercados emergentes do que em economias desenvolvidas”.

Entretanto, a discussão acima diz respeito apenas como o nível da taxa de câmbio afeta o nível de preços. A nossa discussão será sobre a relação entre as volatilidades dessas variáveis, ou seja, sobre a variância delas. No fundo, o que vamos querer ver é o quanto um aumento da incerteza de uma variável afeta a incerteza sobre a outra variável. Para isso, usaremos uma técnica mais moderna para verificar empiricamente como essas variâncias se relacionam entre si. Essa técnica é o GARCH bivariado, que terá uma apresentação mais extensa no capítulo 2.

Na literatura sobre esse assunto, Albuquerque e Portugal (2006) dizem que há dois principais grupos com opiniões diversas sobre o efeito da volatilidade da taxa de câmbio sobre a volatilidade da inflação, um diz que não há tal relação, enquanto o outro diz que há. Antes de esmiuçarmos mais essa discussão nós vamos expor outra discussão, que diz respeito a como o regime de câmbio afeta a volatilidade do câmbio e da inflação.

## **2.2 Regime de câmbio, volatilidade cambial e volatilidade da inflação**

A discussão acerca dos regimes de taxa de câmbio é bem antiga. Porém, um trabalho mais recente sobre as taxas de câmbio é o de Seabra (1998), o qual faz um bom

resumo de como cada regime de câmbio funciona, seus principais aspectos, vantagens e desvantagens. Outro trabalho interessante é o de Friedman (1953), o qual faz uma defesa das taxas de câmbio livre.

Em seu trabalho, Friedman defende o regime de taxas de câmbio livre, visto que com esse regime as taxas vão se movimentando conforme o mercado, ficando sempre em equilíbrio, caso que não aconteceria com as taxas de câmbio fixo, uma vez que estas só se alterariam quando um ajuste não pudesse ser mais postergado. Friedman também fala que as taxa de câmbio livre beneficiam o comércio, visto que os ajustes são equilibrantes. Em suas palavras

“Sob regimes de taxas de câmbio flexíveis determinadas livremente em mercados abertos, o primeiro impacto de qualquer tendência a um superávit ou déficit na balança de pagamentos será sobre as taxas de câmbio. Se um país tem um incipiente superávit de recebimento de pagamentos – um excesso de demanda por sua moeda – a taxa de câmbio tenderá a subir. Se o país tem um incipiente déficit, a taxa de câmbio tenderá a cair. Se as condições responsáveis pela subida ou queda das taxas são geralmente tratadas como temporárias, portadores, de fato ou potenciais, de moeda do país tenderam a mudar as suas colocações de capital de uma forma que modere os movimentos da taxa de câmbio. Se um aumento temporário nas taxas de câmbio é esperado, por exemplo, há um incentivo para portadores da moeda do país venderem as suas posições, comprarem moedas estrangeiras e comprarem de volta as moedas antigas por um preço menor. Ao fazerem isso, ele suprem o excesso de demanda por moeda responsável pelo aumento inicial da taxa de câmbio, ou seja, eles absorvem parte do que seria um superávit de recebimentos de moeda estrangeira na antiga taxa de câmbio. Da mesma forma, se um declínio temporário é esperado, há um incentivo para comprar a moeda doméstica para vendê-la a uma preço maior. Tais compras de moeda domestica fornecem a moeda estrangeira necessária para suprir o que seria um déficit de moeda estrangeira à taxa de câmbio antiga. Dessa forma, tais transações “especulativas” fornecem reservas ao país para ele absorver superávits ou suprir déficits temporários de moeda estrangeira. Por outro lado, se uma mudança na taxa de câmbio é geralmente tratada como produzida por fundamentos que provavelmente serão permanentes, os incentivos serão o contrário dos que listamos e transações especulativas vão trazer maior presteza ao aumento ou queda da taxa de câmbio e, portanto, acelerar a chegada a posição final. (FRIEDMAN, 1953)”

Outro ponto que Friedman faz é que quando a taxa de câmbio é fixa, há uma maior regulação do governo sobre o comércio internacional e os fluxos de capitais, de modo que a taxa de câmbio se mantenha no mesmo patamar. Outra manipulação do câmbio citada por Friedman é que impostos podem agir como depreciações ou apreciação da taxa de câmbio, em suas palavras:

Note que uma tarifa com porcentagem uniforme em todas as importações que paga um subsídio em todas as exportações é equivalente a uma depreciação na taxa de câmbio pela porcentagem correspondente a tarifa. Da mesma forma, um subsídio na forma de uma porcentagem uniforme em todas as importações financiada por um imposto em todas as exportações é equivalente a uma apreciação na taxa de câmbio correspondente a mesma porcentagem. Portanto, dispositivos como esses deveriam ser classificados como alterações na taxa de câmbio, em vez de controles diretos (FRIEDMAN, 1953)

Sobre o assunto de regimes de taxa de câmbio *de jure* funcionarem diferentemente do *de facto*, há dois trabalhos muito interessantes. Obstfeld e Rogoff (1995), Reinhart (2000).

O primeiro trabalho fala sobre como seria muito difícil um regime de câmbio fixo conseguir manter a sua taxa de câmbio fixa durante um período muito longo. Em seu texto eles encontram evidências de que “Tirando algumas pequenas economias de turismo, monarquias de xeques com petróleo e principados altamente dependentes, há, literalmente, apenas um punhado de países que tem mantido a sua taxa de câmbio realmente fixa contra qualquer outra moeda por cinco anos ou mais.” Além disso, esse artigo é uma ótima apresentação dos câmbios fixos, descrevendo como eles funcionam, os custos, os benefícios e os motivos pelos quais um governo pode os adotar.

Apesar de o autor apresentar a ideia dessa difícil implementação, ele fala que é errado o pensamento de que com mercados de capitais internacionalizados e conectados seja impossível para fixarem a taxa de câmbio. Ele escreve que para um país evitar um choque especulativo, basta ele enxugar a sua base monetária e oferecer juros altos pelos depósitos em seu país, ações que a maioria dos países que sofreram ataques na década de 1990 poderiam fazer. Entretanto, tais países tinham outros objetivos em sua política macroeconômica e simplesmente era mais barato eles pararem de defender o seu câmbio e o deixarem flutuar, do que continuar a defesa das taxas de câmbio.

Obstfeld e Rogoff (1995) também falam sobre a diferença dos regimes de câmbio *de jure* e *de facto*:

Algumas das evidências empíricas mais importantes são: países que dizem que permitem a taxa de câmbio flutuar, na maioria das vezes, não deixam – parece haver uma epidemia de “medo de flutuar”. Relativamente aos flutuadores mais comprometidos – tais como Estados Unidos, Austrália e Japão – a variabilidade da taxa de câmbio é bem baixa. A baixa variabilidade da taxa de câmbio nominal não se deve a falta de choques nominais nessas economias – de fato, relativamente aos Estados Unidos e Japão, a maioria desses países está sujeito a choques maiores e mais freqüentes nos seus termos de troca. Dificilmente, isso é surpreendente, dada à alta composição de *commodities* primárias das exportações desses países, em muitos casos. (OBSTFELD e ROGOFF, 1995)

O segundo trabalho se contrapõe a Obstfeld e Rogoff (1995), visto que a autora apresenta evidências empíricas que os países que “flutuam” não flutuam tanto assim e que a maioria deles (principalmente os subdesenvolvidos) usa um “colete salva-vidas” quando o faz, ideia que Hausmann et al (2001) resgata em seu trabalho. Essa “secreta” fixação da taxa de câmbio também é explorada em Ghosh et al. (1997), que diferencia regimes cambiais *de jure* e *de facto*.

Este último trabalho destaca como o regime de taxas de câmbio afeta as variáveis macroeconômicas de um país, assunto que Bleaney e Fielding (2002) também exploram. No primeiro trabalho, os autores investigam a importância do regime de câmbio que os países seguem, ou declaram seguir, sobre a inflação e o produto. Os dados usados cobrem nove tipos de regime de taxas de câmbio, 140 países e mais de trinta anos de abrangência. A sua conclusão é que tanto o nível quanto a volatilidade da inflação é mais baixa quando a taxa de câmbio é fixa.

Já o segundo trabalho colhe dados de 80 países em desenvolvimento com abrangência temporal de 1980 até 1989 com o objetivo de analisar as diferenças entre a volatilidade e a média de inflação e produto entre os países menos desenvolvidos e os países desenvolvidos. Um dos resultados apontado pelos dados é que há um *trade-off* na escolha do regime de taxas de câmbio entre redução da inflação e a estabilidade do produto (e da inflação). Com os regimes de câmbio flutuante tendo menor volatilidade de inflação e produto, mas maior nível inflacionário; e os regimes de bandas cambiais tendo menor nível

de inflação e maior variância da inflação e do produto. Já comparados com os regimes de taxa de inflação, os fatores estruturais parecem ter um papel menor na explicação, entre os países, das diferenças entre os níveis de inflação, porém são mais importantes na explicação tanto na volatilidade do produto quanto da inflação.

### 2.3 Relação entre volatilidade do câmbio e volatilidade da inflação

Agora voltamos à exposição de Albuquerque e Portugal (2006) e começaremos a discussão sobre a relação entre as volatilidades do câmbio e da inflação. Em seu trabalho distinguem dois grupos diferentes sobre o assunto, um diz que não há tal relação, enquanto o outro diz que há seja ela positiva ou negativa. Então, a seguir veremos mais a fundo os pontos defendidos por cada lado. Primeiro apresentaremos o grupo que diz que não relação entre as volatilidades do câmbio e da inflação e depois apresentaremos o grupo que diz que há relação entre essas volatilidades.

O grupo que defende que não há relação entre as volatilidades da taxa de câmbio e da inflação diz que a volatilidade da taxa de câmbio não tem grandes efeitos para outras variáveis macroeconômicas, uma vez que a evidência empírica mostra um acréscimo substancial na volatilidade daquela taxa desde a predominância do regime de taxas de câmbio flutuantes, enquanto as volatilidades das outras variáveis macroeconômicas não apresentaram um aumento de mesma magnitude.

Um dos trabalhos mais interessantes desse grupo é de Flood e Rose (1999), que mostra que se considerarmos um equilíbrio no mercado de ativos da seguinte forma:

$$m_t - p_t = \beta y_t - \alpha i_t + \varepsilon_t \quad (1)$$

$$p_t = e_t + p_t^* + v_t \quad (2)$$

Onde  $m$  é o estoque monetário doméstico,  $p$  é o nível de preço,  $y$  é o produto real,  $i$  é o nível da taxa de juros,  $e$  é a taxa de câmbio doméstica e  $\alpha$  e  $\beta$  são os valores dos parâmetros, o asterisco denota a variável estrangeiras, todas as variáveis (exceto a taxa de juros) são logaritmos naturais,  $\varepsilon$  é um choque no mercado monetário, e  $v$  é um desvio estacionário da PPP. Lembrando que o modelo é estrutural, o que leva os choques e

parâmetros a serem não dependentes da política econômica. Se assumirmos uma equação estrangeira idêntica e análoga a (1), subtraindo isso de (1) e substituindo em (2) chegamos a seguinte equação:

$$e_t = (m - m^*)_t - \beta(y - y^*)_t + \alpha(i - i^*)_t - (\varepsilon - \varepsilon^*)_t - (v - v^*)_t \quad (3)$$

Essa equação expressa as mudanças da taxa de câmbio como função dos “fundamentos macroeconômicos”. Então nos concentramos em um aspecto fundamental dos dados, as volatilidades do lado esquerdo e do direito. Se esse modelo funcionasse bem, ele capturaria as mudanças nas variáveis e as relações entre as volatilidades do lado esquerdo e do direito. Entretanto, de acordo com o regime de taxas de câmbio que cada país adota, a volatilidade do lado esquerdo varia, o que não ocorre com o lado direito. Ou seja, as variáveis macroeconômicas não seriam importantes para a explicação da volatilidade da taxa de câmbio.

Esses autores também tem um trabalho mais antigo, o Flood e Rose (1995). Nele, os autores fazem modelos para tentar explicar a volatilidade cambial com base em fundamentos monetários, entretanto, mesmo com o regime cambial sendo alterado e a volatilidade cambial se alterando também, a volatilidade as variáveis macroeconômicas não mudam, o que dá evidências para a chamada desconexão da taxa de câmbio. Em suas palavras “(...) volatilidade de variáveis como dinheiro e produção não parece ser significativamente diferentes durante regimes de câmbio fixo ou flutuantes, e raramente é considerado ser diferente por pesquisadores empíricos de macroeconomia”. Ou seja, para esses autores não há um *trade-off* claro entre volatilidade reduzida das taxas de câmbio e estabilidade macroeconômica. Embora eles tenham achado que há um *trade-off* entre volatilidade do PIB e da taxa de câmbio. O que os leva a pensar que como os “mercados de taxa de câmbio são mercados onde se pode comprar e vender riscos facilmente, pode ser mais prudente concentrar os distúrbios nesse mercado, em vez de transferi-los para outros mercados, como o de trabalho, onde não se pode lidar com risco de uma maneira tão eficiente”

Outros autores que defendem que há desconexão da taxa de câmbio são Obstfeld e Rogoff (2000). Nesse trabalho, esses autores expõem que “Taxas de câmbio são notoriamente voláteis em relação a qualquer modelo que usamos com fundamentos

implícitos como taxas de juros, produção e oferta de moeda; e nenhum modelo parece explicar bem as taxas de câmbio, nem ex-post”. Nesse trabalho os autores também ensaiam uma resposta para tal desconexão falando que “os mercados para a maioria dos bens comercializáveis não são completamente integrados, e segmentação devida a vários custos de comércio podem ser bem difundidos”. O que explicaria que “ajustes na taxa de câmbio têm pouco efeito econômico no curto prazo (...). Apenas gradualmente que a resposta dos importadores e exportadores é repassada para o varejo. Portanto, para esses autores a volatilidade alta e desconexão da taxa de câmbio provêm de uma combinação de custos de comércio, monopólio e precificação para o mercado em moeda local.

Devereux e Engel (2002) também argumentam que a volatilidade do câmbio não afeta a inflação; nesse trabalho, os autores montam modelos de forma a testar algumas hipóteses como precificação para o mercado e precificação em moeda doméstica para examinar a desconexão da taxa de câmbio. Eles descobrem que “taxas de câmbio podem ser altamente voláteis, visto que, de certa forma, elas tem pouco efeito nas variáveis macroeconômicas”. Eles também falam que “Evidências empíricas indicam que taxas de câmbio nominais não têm repasse completo do câmbio para os preços de bens. De fato, parece que os preços do consumidor não respondem a mudanças na taxa de câmbio”. Ou seja, “volatilidade da taxa de câmbio pode ser alta e ainda assim não importar para as variáveis reais”.

Eichengreen (2002) também destaca que um motivo que pode afetar a volatilidade cambial é que “porque a informação é custosa para se adquirir e processar, investidores não tem informação perfeita. Eles, portanto, exibem comportamento de manada ao entrar e sair de mercados em resposta a movimentos de outros investidores, o que amplifica a volatilidade”. Obstfeld e Rogoff (1995) também defendem a desconexão entre a volatilidade cambial e a volatilidade dos fundamentos macroeconômicos visto que ele é um preço relativo entre duas moedas que muito do seu valor é formado pelas expectativas dos agentes.

Agora, nos concentraremos no segundo grupo que defende que há relação entre as volatilidades da taxa de câmbio e das variáveis macroeconômicas, seja ela positiva ou negativa. Albuquerque e Portugal (2006) falam que as descobertas do seu trabalho se associam a essa corrente de descobertas. Nesse trabalho, aqueles autores estimaram a

relação entre as volatilidades das variáveis citadas com dois métodos: volatilidade não-condicional (*rolling windows*) e volatilidade condicional (GARCH bivariado) e acham uma relação entre as volatilidades semi-côncava (diferentemente dos casos de séries econômicas, onde as relações entre as volatilidades são convexas, levando ao “sorriso da volatilidade”). Em termos econômicos, essa relação semi-côncava se traduziria em dizer que quando a volatilidade da taxa de câmbio é muito alta, o aumento da incerteza traz uma resposta à inflação reduzida, o que levaria a efeitos menores. Então, em períodos de alta volatilidade, agentes não responderiam com a mesma intensidade que eles fazem em períodos de estabilidade devido a falta de conhecimento relativo a duração dos movimentos na taxa de câmbio. No entanto, quando a volatilidade é menos, a inflação responderia mais prontamente.

Além disso, Albuquerque e Portugal (2006) falam que acharam resultados semelhantes ao de Dixit (1989) e ao de Seabra (1996). O primeiro trabalho desenvolve um modelo de otimização da firma estrangeira com custos de entrada e saída no mercado doméstico, esse modelo nos mostra que os fluxos de troca e preços dependem dos investimentos feitos olhando para o futuro, e por consequência, das expectativas e dos momentos mais altos das distribuições envolvidas. Por isso, o ambiente macroeconômico afeta o padrão de alterações nos preços. Portanto, não só o nível de desvalorização, mas também a volatilidade da taxa de câmbio afetariam o seu repasse para os preços. Além disso, dependendo do ambiente macroeconômico (o que o autor simula por alteração de certos parâmetros do modelo), a entrada de novas firmas estrangeiras pode ser duradoura, tendo que a taxa de câmbio voltar a um nível menor do que era quando a firma entrou no mercado doméstico (fenômeno descrito como histerese, ou seja, um efeito que persiste, mesmo depois que a causa de seu início tenha sido removida). Já o segundo autor mostra dois modelos de otimização com bases teóricas diferentes: uma neoclássica e outra de investimento irreversível. Com a primeira abordagem, o autor chega à conclusão de que o efeito total de uma variação na incerteza cambial sobre o investimento é positiva. Já a abordagem de investimento irreversível com assimetria de custos de adaptação e mostra que o efeito total de uma variação na incerteza cambial sobre o investimento é negativa. Ou seja, se a incerteza é elevada, a melhor decisão será de esperar antes de fazer alguma ação (estratégia *wait-and-see*), mesmo que a taxa de câmbio esteja em um patamar que

torna o investimento rentável. Esta atitude gera impactos na oferta agregada e, destarte, sobre a inflação.

Outros trabalhos interessantes desse grupo são o de Smith (1999), o de Calvo e Reinhart (2000), Bleaney e Fielding (2002) e Ghosh et al (1997). A exposição desses últimos dois trabalhos já foi feita acima. Nesse momento, nos concentraremos nos dois primeiros.

Smith (1999) começa o seu artigo dizendo que “Uma das desvantagens mais citadas de um regime de taxas de câmbio flutuante é que as flutuações nessa taxa levam a maior volatilidade no preço para os exportadores e importadores e, portanto, em uma maior incerteza de preços”. Nesse trabalho aquele autor tenta mostrar como variações na taxa de câmbio afetam o preço de *commodities*, em vez de índices de preço do consumidor. Isso porque quando medimos a variação dos preços das *commodities*, em vez de índices de preços ao consumidor tiramos os preços administrados e podemos ver como se comportam os preços de bens negociados em mercados internacionais, o que pode trazer uma evidência de como realmente se comportam esses preços. Uma de suas conclusões é que “A variação de preços depende da variação da taxa de câmbio, mas também da variação do preço do bem em moeda estrangeira, variação do preço do bem em moeda doméstica e na covariância entre essas moedas”. Outra conclusão bem importante que ele destaca é que as “variações da taxa de câmbio é um indicador inadequado do risco cambial (*price risk*) enfrentado pelas firmas (...) Variações maiores da taxa de câmbio não necessariamente implica em variações maiores dos preço real da moeda doméstica”. Também é importante destacar que aquele autor também fala que as taxas de câmbio fixas não eliminam completamente o risco cambial devido a reavaliações periódicas da taxa de câmbio.

Calvo e Reinhart (2000), em seu trabalho não acham que a volatilidade do câmbio está desconectada da volatilidade das outras variáveis macroeconômicas. Nesse trabalho os autores analisam o comportamento de taxas de câmbio, taxas de câmbio estrangeiras, agregados monetários, e preços de *commodities* através do espectro de arranjos de taxa de câmbio para avaliar se os rótulos oficiais fornecem uma das práticas reais dos países. Uma de suas conclusões é que, nos países analisados, a autoridade monetária quando se depara com choques, prefere suavizar a taxa de câmbio, em vez da inflação por uma série de motivos, o que leva ao “medo de flutuar”, nas palavras dos autores “a relativamente baixa

volatilidade das taxas de câmbio provém de políticas deliberadas para estabilizar a taxa de câmbio. A volatilidade de reservas (...) é bem alta. A volatilidade de juros (...) é significativamente – e em um outro patamar ainda – dos flutuadores (mais) verdadeiros”.

Entretanto, essa ideia é debatida em Eichengreen (2002), o qual mostra que uma vez que a economia é aberta, isso dificulta a avaliação da função objetivo do banco central com as metas de inflação. Ele cita um exemplo em que apesar de ter havido um choque externo que depreciou o câmbio, o banco central pode elevar a taxa de juros não para suavizar o efeito na taxa de câmbio, mas para impedir o repasse cambial para inflação que essa mudança pode trazer. Então apesar de parecer que o banco central nacional tem “medo de flutuar” e que o câmbio tem um peso grande na sua função objetivo, na verdade a inflação tem um peso maior e foi ela que levou o banco central a elevar a taxa de juros.

Concluindo, o que podemos afirmar sobre a PPC é que quando um país importa um produto, obviamente a taxa de câmbio é relevante e mudanças nela podem levar a alterações no nível de preços de um país. Porém, fatores como barreiras comerciais, estruturas de mercado e diferenças no padrão de consumo podem levar o repasse da taxa de câmbio para a inflação a ser incompleto. Outra discussão sobre a taxa de câmbio e seus efeitos é como o regime de câmbio afeta o próprio câmbio e a inflação. Nessa discussão Friedman (1953) defende que as taxas de câmbio flutuantes seriam menos voláteis e seriam auto-ajustáveis, diferentemente das taxas de câmbio fixas, as quais só se ajustariam quando a mudança não pudesse ser evitada, o que traria mais danos ao sistema do que as pequenas variações do regime de câmbio flutuante. Também sobre esse assunto há a discussão entre Obstfeld e Rogoff (1995) fala sobre como seria muito difícil um regime de câmbio fixo conseguir manter a sua taxa de câmbio fixa durante um período muito longo. Em seu texto eles encontram evidências de que “Tirando algumas pequenas economias de turismo, monarquias de xeqes com petróleo e principados altamente dependentes, há, literalmente, apenas um punhado de países que tem mantido a sua taxa de câmbio realmente fixa contra qualquer outra moeda por cinco anos ou mais”. Ideia essa que é contraposta por Reinhart (2000), a qual apresenta evidências empíricas que os países que “flutuam” não flutuam tanto assim e que a maioria deles (principalmente os subdesenvolvidos) usa um colete salva-vidas quando o faz. Essa “secreta” fixação da taxa de câmbio também é explorada em Ghosh et al. (1997), que diferencia regimes cambiais *de jure* e as *de facto*. Por fim,

concluimos que não podemos fazer asserções muito fortes sobre a relação entre as volatilidades do câmbio e da inflação, visto que os trabalhos analisados chegaram a resultados diferentes usando métodos diferentes.

### 3 Modelos Econométricos de Séries de Tempo: ARIMA e GARCH

Nesse capítulo, vamos apresentar a revisão metodológica com o intento de apresentar as técnicas econométricas que serão usadas nesse trabalho da forma mais geral possível; contudo, sem a pretensão de exaurir os assuntos por completo. Esse capítulo se baseia em Enders (2011) e Bueno (2008). Ele se divide em 3 seções:; 1) Estacionariedade e Modelos ARMA 2) Heterocedasticidade e Modelos ARCH e GARCH univariados e 3) GARCH bivariado

#### 3.1 Estacionariedade Fraca e Modelos ARMA

Antes de falarmos dos modelos ARMA, falaremos brevemente sobre estacionariedade. Segundo Enders, um processo estocástico que tem média e variância finita é fracamente estacionário se para todo  $t$  e  $t - s$ :

$$E(y_t) = E(y_{t-s}) = \mu < \infty;$$

Ou seja,  $\mu$  é constante, finito e não depende do tempo;

$$E[(y_t - \mu)^2] = E[(y_{t-s} - \mu)^2] = \sigma_y^2$$

$$[var(y_t) = var(y_{t-s}) = \sigma_y^2];$$

Isso é,  $\sigma_y^2$  é constante, finito e não depende do tempo;

$$E[(y_t - \mu)(y_{t-s} - \mu)] = E[(y_{t-j} - \mu)(y_{t-j-s} - \mu)] = \gamma_s$$

$$[cov(y_t, y_{t-s}) = cov(y_{t-j}, y_{t-j-s}) = \gamma_s]$$

O que quer dizer que a covariância é finita e não depende do tempo, o que também implica que as realizações do processo estocástico só depende da distância entre as observações.

Portanto, como definem Cryer e Chan (2008), a idéia básica por trás da estacionariedade é que as leis de probabilidade que governam o processo estocástico a ser estudado não mudam durante o tempo decorrido. Ou seja, se houver estacionariedade, não

importa qual realização observada, as leis serão as mesmas para todo o processo estocástico.

Para a série fracamente estacionária, podemos definir a autocorrelação entre  $y_t$  e  $y_{t-s}$  como:

$$\rho_s \equiv \gamma_s / \gamma_0$$

Onde  $\gamma_0$  equivale à variância de  $y_t$ .

Uma vez que  $\gamma_s$  e  $\gamma_0$  são independentes do tempo, os coeficientes de autocorrelação  $\rho_s$  também não dependem do tempo. Então, embora a autocorrelação entre  $y_t$  e  $y_{t-1}$  possa ser diferente entre a autocorrelação entre  $y_t$  e  $y_{t-2}$ , a autocorrelação entre  $y_t$  e  $y_{t-1}$  deve ser idêntica a entre  $y_{t-s}$  e  $y_{t-s-1}$ . E  $\rho_0 = 1$ .

### 3.1.1 Processo de Médias Móveis (MA)

Vamos começar o nosso estudo dos processos de médias móveis analisando um processo de médias móveis de ordem 1 – MA (1). Então, considerando o seguinte processo estocástico:

$$y_t = \mu + \varepsilon_t + \theta \varepsilon_{t-1}$$

Em que  $\varepsilon_t$  é um processo de ruído branco. Uma vez que  $y_t$  depende do erro contemporâneo,  $\varepsilon_t$ , e do erro imediatamente passado, então o processo é chamada de médias móveis de ordem 1 e denotados por MA(1). Se o processo dependesse de  $\varepsilon_{t-2}$ , então seria chamado de MA (2) e assim por diante. Nosso próximo passo será verificar se esse processo é estacionário.

$$E(y_t) = \mu + E(\varepsilon_t) + \theta E(\varepsilon_{t-1}) = \mu$$

$$\begin{aligned} \text{Var}(y_t) &= E(y_t - \mu)^2 = E(\varepsilon_t + \theta \varepsilon_{t-1})^2 = E(\varepsilon_t^2 + 2\theta \varepsilon_t \varepsilon_{t-1} + \theta^2 \varepsilon_{t-1}^2) \\ &= \sigma^2 + 0 + \theta^2 \sigma^2 = (1 + \theta^2) \sigma^2 \end{aligned}$$

$$\begin{aligned} E[(y_t - \mu)(y_{t-1} - \mu)] &= E[(\varepsilon_t + \theta \varepsilon_{t-1})(\varepsilon_{t-1} + \theta \varepsilon_{t-2})] \\ &= E(\varepsilon_t \varepsilon_{t-1} + \theta \varepsilon_t \varepsilon_{t-2} + \theta \varepsilon_{t-1}^2 + \theta^2 \varepsilon_{t-1} \varepsilon_{t-2}) = \sigma^2 \theta \end{aligned}$$

$$\begin{aligned}
E[(y_t - \mu)(y_{t-j} - \mu)] &= E[(\varepsilon_t + \theta\varepsilon_{t-1})(\varepsilon_{t-j} + \theta\varepsilon_{t-j-1})] \\
&= E(\varepsilon_t\varepsilon_{t-j} + \theta\varepsilon_t\varepsilon_{t-j-1} + \theta\varepsilon_{t-1}\varepsilon_{t-j} + \theta^2\varepsilon_{t-1}\varepsilon_{t-j-1}) = 0, \text{ para } j > 1
\end{aligned}$$

A esperança é constante e finita para  $t$ . A variância é finita. A autocovariância não depende de  $t$ . Como a esperança e autocovariâncias não são funções do tempo, o processo é fracamente estacionário, independentemente do valor de  $\theta$ .

Agora, generalizamos o processo de médias móveis para  $q$  defasagens.

$$y_t = \mu + \sum_{j=0}^q \theta_j \varepsilon_{t-j}, \theta_0 = 1$$

Agora, verificamos se esse processo atende às condições para estacionariedade.

$$E(y_t) = \mu + \sum_{j=0}^q \theta_j E(\varepsilon_{t-j}) = \mu$$

$$Var(y_t) = E(y_t - \mu)^2 = E\left(\sum_{j=0}^q \theta_j \varepsilon_{t-j}\right)^2 = \sum_{j=0}^q \theta_j^2 E(\varepsilon_{t-j}^2) = \sigma^2 \sum_{j=0}^q \theta_j^2$$

Para  $j = 1, 2, 3, \dots, q$  temos:

$$\begin{aligned}
\gamma_j &= E\left(\sum_{i=0}^q \theta_i \varepsilon_{t-i} \sum_{i=0}^q \theta_i \varepsilon_{t-i-j}\right) \\
&= E[\theta_j \varepsilon_{t-j}^2 + \theta_{j+1} \theta_1 \varepsilon_{t-j-1}^2 + \theta_{j+2} \theta_2 \varepsilon_{t-j-2}^2 + \dots + \theta_q \theta_{q-j} \varepsilon_{t-q}^2]
\end{aligned}$$

Para  $j > q$ , não haverá  $\varepsilon$ 's em datas comuns. Logo a esperança será nula. Portanto:

$$\gamma_j = \begin{cases} [\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \dots + \theta_q\theta_{q-j}]\sigma^2, & \text{para } j = 1, 2, 3, \dots, q; \\ 0, & \text{para } j > q \end{cases}$$

Assim como os processos auto-regressivos tem convergir para serem estacionários, os processos devem ser invertíveis. E da mesma forma que um AR(p) pode se transformar em um MA( $\infty$ ), um MA(q) pode ser transformado em um AR( $\infty$ ), se certas condições forem satisfeitas. Seja um processo MA(1):

$$y_t - \mu = \varepsilon_t + \theta \varepsilon_{t-1} = (1 + \theta L) \varepsilon_t$$

Para esse polinômio não “explodir” e a série  $y_t$  convergir para ser estacionária precisamos ter  $|\theta| < 1$ . Então, podemos escrever esse processo da seguinte forma:

$$\begin{aligned} (y_t - \mu)(1 + \theta L)^{-1} = \varepsilon_t &\Rightarrow (y_t - \mu)(1 - \theta L + \theta^2 L^2 - \dots) = \varepsilon_t \Rightarrow y_t - \mu \\ &= - \sum_{j=1}^{\infty} (-\theta)^j (y_{t-j} - \mu) + \varepsilon_t \end{aligned}$$

No caso do MA(q), haverá invertibilidade se as raízes características do polinômio  $(1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q)$  estiverem fora do círculo unitário. Isso é necessário para que a série possa ser estimada recursivamente e haja unicidade dos resultados.

### 3.1.2 Processo Auto Regressivo – AR

Começamos o nosso estudo dos processos auto-regressivos pelo mais simples deles, o de ordem 1. Então, consideramos o seguinte processo estocástico:

$$y_t = c + \phi y_{t-1} + \varepsilon_t$$

Onde  $\varepsilon_t$  é um ruído branco.

O processo assim definido é chamado de auto-regressivo de ordem 1 e denotado como AR (1). Então, verificamos se esse processo é estacionário. Para isso, primeiro supomos que  $|\phi| < 1$ . Isso é importante para que a série não “exploda” e a sequência  $\{y_t\}$  convirja e seja estacionaria. Depois usamos o operador de defasagem para transformarmos esse AR (1) em um MA ( $\infty$ ).

$$y_t = c + \phi y_{t-1} + \varepsilon_t \Rightarrow (1 - \phi L)y_t = c + \varepsilon_t \Rightarrow$$

$$y_t = \frac{c}{1 - \phi} + \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j} = \mu + \psi(L)\varepsilon_t$$

Em que

$$\mu = \frac{c}{1 - \phi} \quad e \quad \psi(L) = (1 - \phi L)^{-1} = 1 + \phi L + \phi^2 L^2 + \dots$$

Então calculamos:

$$E(y_t) = \mu + \sum_{j=0}^{\infty} \phi^j E(\varepsilon_{t-j}) = \mu$$

$$Var(y_t) = E(y_t - \mu)^2 = E\left(\sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}\right)^2 = \sum_{j=0}^{\infty} \phi^j E(\varepsilon_{t-j}^2) = \frac{\sigma^2}{1 - \phi^2}$$

A autocovariância de defasagem  $j$  é:

$$\begin{aligned} E[(y_t - \mu)(y_{t-j} - \mu)] &= E\left[\left(\sum_{s=0}^{\infty} \phi^s \varepsilon_{t-s}\right)\left(\sum_{s=0}^{\infty} \phi^s \varepsilon_{t-s-j}\right)\right] \\ &= \sigma^2(\phi^j + \phi^{j+2} + \phi^{j+4} + \dots) = \left(\frac{\phi^j}{1 - \phi^2}\right)\sigma^2 \end{aligned}$$

Como a média e a autocovariância não são funções do tempo, o processo é fracamente estacionário.

Agora que vimos como funciona o AR (1), vamos para o caso mais geral, onde veremos um processo auto-regressivo de ordem  $p$  – AR ( $p$ ). Esse processo é definido da seguinte forma:

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t = c + \sum_{j=0}^p \phi_j y_{t-j} + \varepsilon_t$$

Se as raízes do polinômio  $(1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p)$  estiverem fora do círculo unitário ou, equivalentemente, se as raízes do polinômio  $(\lambda^p - \phi_1 \lambda^{p-1} - \phi_2 \lambda^{p-2} - \dots - \phi_p)$  estiverem dentro do círculo unitário, o processo será fracamente estacionário e poderá ser representado como um MA ( $\infty$ ).

$$y_t = \mu + \psi(L)\varepsilon_t$$

Em que

$$\mu = \frac{c}{1 - (\phi_1 + \phi_2 + \dots + \phi_p)}$$

$$\psi(L) = (1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)^{-1}$$

Sob estacionariedade, é verdade que  $E(y_t) = E(y_{t-1}) = \dots$ . Logo, podemos calcular a esperança não condicional de  $y_t$ ;

$$\begin{aligned} E(y_t) &= c + \phi_1 E(y_{t-1}) + \phi_2 E(y_{t-2}) + \dots + \phi_p E(y_{t-p}) + E(\varepsilon_t) \Rightarrow E(y_t) \equiv \mu \\ &= \frac{c}{1 - (\phi_1 + \phi_2 + \dots + \phi_p)} \end{aligned}$$

Dada a esperança não condicional do processo, é conveniente reescrevê-lo de outra forma, para tornar alguns cálculos mais fáceis:

$$y_t - \mu = \phi_1(y_{t-1} - \mu) + \phi_2(y_{t-2} - \mu) + \dots + \phi_p(y_{t-p} - \mu) + \varepsilon_t$$

Multiplicando os dois lados por  $(y_{t-j} - \mu)$  e tomando as esperanças:

$$\begin{aligned} E(y_t - \mu)(y_{t-j} - \mu) &= \phi_1 E(y_{t-1} - \mu)(y_{t-j} - \mu) + \phi_2 E(y_{t-2} - \mu)(y_{t-j} - \mu) + \dots \\ &+ \phi_p E(y_{t-p} - \mu)(y_{t-j} - \mu) + E[\varepsilon_t(y_{t-j} - \mu)] \end{aligned}$$

Teremos:

$$\gamma_j = \begin{cases} \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} + \dots + \phi_p \gamma_{j-p}, & j = 1, 2, \dots \\ \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p + \sigma^2, & j = 0 \end{cases}$$

### 3.1.3 ARMA

Sabendo como funcionam esses processos, podemos combiná-los de modo a formarmos um processo auto-regressivo de médias móveis –  $ARMA(p, q)$ . Esse processo pode ser escrito da seguinte forma:

$$y_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=0}^q \theta_j \varepsilon_{t-j} \quad (1)$$

Para estimarmos os modelos  $ARMA$  é preciso que sejam “estáveis”, significando que as raízes das equações características dos processos  $AR$  e  $MA$  estejam simultaneamente fora do círculo unitário. Para verificarmos esse fato, aplicamos o

operador de defasagem e separarmos os processos AR e MA e encontrarmos as raízes características de cada processo.

$$y_t(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_q L^q) = c + \varepsilon_t(1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q)$$

E achamos as raízes características de cada equação. Aqui, assumimos que não existem raízes comuns entre os processos AR e MA, visto que se houvessem nós poderíamos “cortar” as raízes e reduzir os polinômios para ordens menores. Particularmente, para um ARMA (1,1), isso significa que  $\phi \neq \theta$ .

### 3.1.4 Diferenciação de processos não estacionários

Agora que já vimos como esses modelos se comportam, vem a pergunta: “e se a série não for estacionária”? Segundo a metodologia de Box, Jenkins e Reinsel (1994), se uma série for considerada não estacionária, ela deve ser diferenciada. Caso contrário, pode-se usar a série original, estacionária, e aplicar a seguinte metodologia para modelá-las:

1. Identificar as ordens p e q do modelo
2. Estimar o modelo

O próximo passo seria verificar se os resíduos estimados não rejeitam a hipótese nula de que são um ruído branco. Entretanto, como o escopo desse trabalho é estimar modelos GARCH, nós pularemos essa parte e o nosso próximo passo seria verificar a presença de resíduos ARCH e GARCH pelos testes de Ljung-Box e Multiplicador de Lagrange.

Então, voltando à identificação das ordens p e q do modelo, nós as identificaremos por meio da função de autocorrelação (FAC) e da função de autocorrelação parcial (FACP). A FAC é o gráfico da autocorrelação contra a defasagem. Conseguimos obter essa função dividindo a autocovariância com distância j ( $\gamma_j$ , para  $j > 0$ ) pela autocovariância de distância 0 –  $\gamma_0$ . Já a FACP é o gráfico de  $\hat{\phi}_{j,j}$  contra j estimado a partir das seguintes regressões em que a série original tem a sua média subtraída.

$$y_t = \phi_{j,1} y_{t-1} + \phi_{j,2} y_{t-2} + \dots + \phi_{j,j} y_{t-j} + e_t, \quad j = 1, 2, \dots$$

Em que  $e_t$  é um erro.

Em outras palavras, o procedimento consiste em regredir  $y_t$  contra  $y_{t-1}$  e obter  $\hat{\phi}_{1,1}$ . Em seguida, deve-se regredir  $y_t$  contra  $y_{t-1}$  e  $y_{t-2}$ . Serão obtidos os coeficientes  $\hat{\phi}_{2,1}$  e  $\hat{\phi}_{2,2}$ , dos quais interessa apenas este último; e assim por diante.

A partir dessas idéias, podemos identificar mais facilmente os modelos AR (p) e MA (q) e, com dificuldades, o modelo ARMA (p,q). A FACP define a defasagem do AR. A FAC define a defasagem do MA. No primeiro caso, sabe-se que a FAC decai exponencialmente com o aumento de defasagens, e a FACP é truncada a partir da defasagem p. No segundo caso, ocorre o inverso. A FAC é truncada na defasagem q, e a FACP decai exponencialmente. No caso de um ARMA (p,q), ambas as funções decaem exponencialmente a partir da defasagem de truncagem. Essa defasagem é frequentemente difícil de reconhecer visualmente. A FAC começará a decair a partir da defasagem q, e a FACP, a partir da defasagem p. O quadro a seguir resume as condições apresentadas:

**TABELA 3.1**  
**Resumo das FAC e FACP dos modelos ARMA**

Modelo	FAC	FACP
AR (p)	Decai exponencialmente	Truncada na defasagem p
MA (q)	Truncada na defasagem q	Decai exponencialmente
ARMA (p,q)	Decai exponencialmente se $j > q$	Decai exponencialmente se $j > p$

### 3.1.5 Critério de Informação

O critério de informação para seleção do modelo é uma forma de encontrar o número ideal de parâmetros de um modelo. Nele, a cada regressor adicional a soma dos resíduos não vai aumentar. A redução dessa soma se dá a custo de mais regressores. Para balancear a redução dos erros e o aumento do número de regressores, o critério de informação associa uma penalidade a esse aumento. Se a penalidade for menor que a diminuição da soma de resíduos, o regressor adicional deve ser adicionado ao modelo (ainda assim, mesmo que o benefício marginal líquido desse regressor for positivo, fica a cargo do pesquisador se decide adicioná-lo ou não). Se a penalidade for maior que a diminuição da soma, o regressor adicional traz mais custos que benefícios. Então a idéia do critério de informação é minimizar uma função baseada nos resíduos penalizada pelo número de regressores.

É difícil identificar o modelo ARMA (p,q) visualmente. Frequentemente, dois ou mais modelos possíveis geram resíduos cujos testes indicam ser um ruído branco. O melhor modelo será o mais parcimonioso, satisfeito que os resíduos sejam os menores possíveis. Isto é, o modelo mais parcimonioso, portanto com menor número de parâmetros, deverá gerar menos imprecisão de estimativas justamente porque tem menos parâmetros que um modelo com mais parâmetros.

As estatísticas descritas a seguir tratam desse problema e são usadas para escolher o melhor modelo. O critério de especificação tem, em geral, a seguinte forma:

$$C = \ln \hat{\sigma}^2(T) + c_T \varphi(T)$$

Em que

$\hat{\sigma}^2(T) = T^{-1} \sum_{t=1}^T \hat{\varepsilon}_t^2$  é a variância estimada dos resíduos;

$c_T$  representa o número de parâmetros estimados;

$\varphi(T)$  é a ordem do processo, que penaliza a falta de parcimônia.

O primeiro termo da equação mede a adequação do processo. É claro que, quanto menor a variância dos resíduos, melhor. Entretanto, tal redução de variância foi obtida a custo da imposição de mais parâmetros ao processo, cuja penalização ocorre na segunda parcela. Há três principais critérios de informação. A estatística de Schwarz, a de Akaike, e a de Hannan-Quinn. A primeira é denotada por BIC (Bayesian Information Criterion) ou SBC (Schwarz Bayesian Criterion) e é dada pela seguinte expressão:

$$BIC(p, q) = \ln \hat{\sigma}^2 + n \frac{\ln T}{T}$$

Em que

$n = p + q$ , se o modelo não tem constante, e

$n = p + q + 1$ , se há constante no modelo,

$$\hat{\sigma}^2 = \frac{\sum_{t=1}^T \hat{\varepsilon}_t^2}{T}$$

A segunda é denotada por AIC (Akaike Information Criterion), é dada por:

$$AIC(p, q) = \ln \hat{\sigma}^2 + n \frac{2}{T}$$

Finalmente, a última é denotado por HQ e é dada por:

$$HQ(p, q) = \ln \hat{\sigma}^2 + n \frac{2}{T} \ln \ln T$$

Quanto mais parâmetros são estimados no mesmo período de amostra, menor será o erro estimado, mas isso será penalizado na segunda parcela da estatística. Por isso, deseje-se o menor AIC, HQ, ou BIC possível.

### 3.1.6 Processos Não-Estacionários

Até agora, analisamos apenas processos estacionários, porém quando a série temporal não é estacionária, não se pode estimá-la trivialmente. O problema essencial é que impossível estimar todos os momentos da série. Em particular, é impossível fazer inferências estatísticas usando o procedimento de Box e Jenkins. Para visualizarmos melhor essa idéia, lembremos que a variância não condicional de um AR (1) é:

$$\text{var}(y_t) = \frac{1}{1 - \phi^2}$$

Nessa equação podemos observar que se  $\phi = 1$ , o que caracteriza uma série não estacionária de raiz unitária, a variância tende ao infinito. No entanto, podemos resolver esse problema. Em séries univariadas, o procedimento básico é diferenciar a série tantas vezes quantas sejam necessárias para a estacionarizar. O caso mais comum é encontrar uma série com raiz unitária, de forma que basta a primeira diferença dessa série para estacionarizá-la, porém mais diferenças poderão ser necessárias.

De modo mais geral, uma série econômica tem três partes distintas:

$$y_t = \textit{tendência} + \textit{componente estacionario} + \textit{ruído}$$

Considerando o seguinte modelo:

$$y_t = y_0 + \delta t + \psi(L)\varepsilon_t$$

Esse modelo é chamado de *tendência estacionária*, uma vez que ele flutua em torno de uma tendência determinística, sem se distanciar de tal tendência. Também podemos enxergar esse fato ao observarmos que a subtração da tendência determinística gera um componente estacionário. Além disso, essa série poderia ser estacionarizada pela primeira diferença da seguinte forma:

$$\Delta y_t \equiv (1 - L)y_t = y_t - y_{t-1} = \delta + (1 - L)\psi(L)\varepsilon_t$$

Essa diferenciação estacionariza a série; todavia, introduz um problema por tornar o erro não invertível. Logo, se uma série é tendência estacionária, é melhor estimá-la usando a variável explicativa  $t$ . Considerando agora o seguinte modelo, em que se supõe que as variações esperadas de  $y$  são dadas por:

$$\Delta y_t = \delta + \varepsilon_t \Rightarrow y_t = y_{t-1}\delta + \varepsilon_t$$

Compondo  $y$  recursivamente, obtemos:

$$y_t = y_0 + \delta t + \sum_{i=1}^t \varepsilon_i$$

Veja que a variável aleatória  $y$  é dada pela composição de todos os choques havidos,  $\sum_{i=1}^t \varepsilon_i$ . Define-se tal série como *tendência estocástica* ou *diferença estacionária*, pois a diferença da série é estacionária. Nesse caso, os choques produzem mudanças permanentes na série  $y$ , ainda que aleatórias. Séries cuja tendência é estocástica são chamadas de *séries integradas*. A ordem  $d$ , de integração, depende do número de somatórios dos erros. Séries integradas são denotadas por  $I(d)$ . Séries integradas com erros estacionários são chamadas séries  $ARIMA(p, d, q)$ . Diferenciando  $d$  vezes a série, obtém uma série estacionária. No caso anterior, diz-se que a série é integrada de ordem 1,  $I(1)$ , porque os erros são todos somados uma única vez, como indica a presença do somatório.

### 3.1.7 Raízes Unitárias

A inspeção visual de uma série raramente permite distingui-la como de tendência estocástica ou tendência determinística. Então, foram desenvolvidos testes para a verificação de existência de raízes unitárias. O primeiro desses testes foi feito por Dickey e Fuller (1979, 1981), apud Bueno (2011). Esse teste segue da seguinte forma, primeiro consideramos o modelo:

$$y_t = \phi y_{t-1} + \varepsilon_t$$

A tendência inicial é estimar esse modelo e usar um teste convencional de estatística  $t$  sobre  $\phi$ , tendo hipótese nula  $H_0: \phi = 1$ . Isso equivale a testar se  $\alpha = 0$  na seguinte equação:

$$\Delta y_t = \alpha y_{t-1} + \varepsilon_t \quad (3.1)$$

Onde ocorreu a seguinte transformação:

$$\Delta y_t = (1 - \phi)y_{t-1} + \varepsilon_t = \alpha y_{t-1} + \varepsilon_t$$

E  $\alpha = (1 - \phi)$ . Assim,  $H_0: \phi = 1$  é equivalente a  $H_0: \alpha = 0$ .

Porém, surge o seguinte problema: sob a hipótese nula, a distribuição do teste não é convencional, isso é, ela não é igual a distribuição  $t$  usual. Sendo assim, Dickey e Fuller (1979), apud Bueno (2011) recalcularam o valor da estatística  $t$  usando experimentos de Monte Carlo. O valor dessa estatística se altera, conforme se define a equação de regressão e segundo o tamanho da amostra. A primeira estatística desconsidera a existência de *drift* e tendência determinística; a segunda desconsidera a existência de tendência determinística, e a terceira há um intercepto e tendência determinística,  $\mu + \delta t$ . Esse é o teste Dickey-Fuller

Nesse teste assume-se 3 formas para a equação (3.1)

$$\Delta y_t = \alpha y_{t-1} + \varepsilon_t \rightarrow \tau \text{ Passeio aleatório puro}$$

$$\Delta y_t = \mu + \alpha y_{t-1} + \varepsilon_t \rightarrow \tau_\mu \text{ Passeio aleatório com } drift$$

$$\Delta y_t = \mu + \delta t + \alpha y_{t-1} + \varepsilon_t \rightarrow \tau_\tau \text{ Passeio aleatório com } drift \text{ e tendência linear}$$

Entretanto, e se quisermos estimar um modelo com ordens superiores? Nesse caso a melhor metodologia é a do teste que se convencionou chamar de teste de Dickey-Fuller aumentado (ADF). Nesse teste, estimamos o seguinte modelo com  $p$  variáveis auto-regressivas:

$$y_t = \mu + \phi_1 y_{t-1} + \dots + \phi_{p-1} y_{t-p+1} + \phi_p y_{t-p} + \varepsilon_t$$

Para entendermos melhor esse teste, adicionamos e subtraímos  $\phi_p y_{t-p+1}$  e obtemos:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_{p-2} y_{t-p+2} + (\phi_{p-1} + \phi_p) y_{t-p+1} + \phi_p \Delta y_{t-p+1} + \varepsilon_t$$

E, continuando dessa maneira, obtemos:

$$\Delta y_t = \mu + \alpha y_{t-1} + \sum_{i=1}^p \lambda_i \Delta y_{t-i} + \varepsilon_t$$

Em que

$$\alpha = -\left(1 - \sum_{i=1}^p \phi_i\right); \quad \lambda_i = -\sum_{j=1}^p \phi_{j+1}$$

O teste então pode ser feito, usando os mesmos valores críticos encontrados por Dickey e Fuller. Ou seja, desde que se faça a correção do modelo, de forma a considerar as demais variáveis defasadas, em nada se alteram os valores críticos do teste ou a interpretação do modelo. Além disso, outra questão se apresenta: e se o modelo for ainda mais complexo, com termos de médias móveis? Bem, como sempre podemos transformar um modelo  $MA(q)$  em um  $AR(\infty)$ , então podemos usar o mesmo teste para um  $ARIMA(p, 1, q)$ . Porém, surge o problema de se estimar um modelo com infinitas defasagens. Sucede daí que é provado que um modelo  $ARIMA(p, 1, q)$  pode ser bem aproximado por um  $ARIMA(n, 1, 0)$  em que  $n \leq T^{1/3}$ .

Outra extensão do teste de Dickey Fuller é que ele pode ser feito conjuntamente para dois ou três coeficientes. Considerando as seguintes especificações:

$$\Delta y_t = \alpha y_{t-1} + \sum_{i=1}^p \lambda_i \Delta y_{t-i} + \varepsilon_t$$

$$\Delta y_t = \mu + \alpha y_{t-1} + \sum_{i=1}^p \lambda_i \Delta y_{t-i} + \varepsilon_t$$

$$\Delta y_t = \mu + \delta t + \alpha y_{t-1} + \sum_{i=1}^p \lambda_i \Delta y_{t-i} + \varepsilon_t$$

Dickey e Fuller (1981), apud Bueno (2011) calcularam as estatísticas F para testes conjuntos, chamando-as de  $\Phi_i$ ,  $i = 1, 2, 3$ . Essas estatísticas não têm distribuição convencional, daí a necessidade de as renomear. Hipóteses nulas e respectivas estatísticas são fornecidas a seguir:

$$H_0: \alpha = \mu = 0 \rightarrow \Phi_1;$$

$$H_0: \alpha = \delta = \mu = 0 \rightarrow \Phi_2;$$

$$H_0: \alpha = \delta = 0 \rightarrow \Phi_3$$

Essas estatísticas são construídas da mesma forma que os testes convencionais:

$$\Phi_i = \frac{(\hat{\varepsilon}' \hat{\varepsilon}_{restrita} - \hat{\varepsilon}' \hat{\varepsilon}_{n\tilde{a}o\ restrita})/r}{\hat{\varepsilon}' \hat{\varepsilon}_{n\tilde{a}o\ restrita}/(T - k)}$$

Em que

$r$  é o numero de restrições, igual a 2 ou 3;

$T$  é o numero de observações;

$k$  é o numero de parâmetros estimados no modelo não restrito.

### 3.2 Modelos ARCH e GARCH univariados

Agora que já vimos os modelos ARMA e as condições de estacionariedade de uma série temporal, vamos começar a exposição dos modelos ARCH e GARCH. Segundo Engle (2001), homocedasticidade é quando os termos do erro elevado ao quadrado são iguais em todos os pontos de uma série de dados. Já heterocedasticidade seria quando em uma série de dados, os termos do erro elevado ao quadrado podem ser diferentes em pontos ou intervalos distintos. Nas palavras do autor:

A versão básica dos modelos de mínimos quadrados (MQO), tem como pressuposto que o valor esperado de todos os termos do erro, quando elevados ao quadrado, é o mesmo em todos os pontos esse pressuposto é chamado de homocedasticidade, e é nesse pressuposto que os modelos ARCH/GARCH se focam. Quando uma série de dados, em que a variância dos termos do erro não são iguais e cujos termos do erro podem razoavelmente ser esperados a serem maiores em alguns pontos ou intervalos da série, do que em outros pontos ou intervalos dessa série, dizemos que esses dados tem heterocedasticidade. (ENGLE, 2001)

Além disso, para Engle (2001) “O problema padrão é que, embora na presença de heterocedasticidade, os coeficientes de uma regressão de MQO não possuem viés, os erros-padrão e os intervalos de confiança, por procedimentos convencionais serão muito estreitos, o que dará uma falsa impressão de precisão”. Ademais, Engle (2001) também fala que “em vez de considerar isso como um problema a ser corrigido, os modelos ARCH e GARCH tratam a heterocedasticidade como uma variância a ser modelada. Disso, resulta que não apenas as deficiências do MQO são corrigidas, mas uma previsão para a variância é calculada para cada termo do erro.”

No seu artigo de 2001, Engle fala um pouco da história dos modelos ARCH e GARCH e qual seria a ideia por trás desses modelos. Nesse artigo, ele fala que, antes desses modelos, praticamente não havia métodos para prever o comportamento da variância e que a técnica disponível era o “*rolling standard deviation*”, a qual consistia em usar o desvio padrão calculado a partir de um número fixo de observações mais recentes (aqui, não desenvolveremos mais essa ideia visto que não é o objetivo desse trabalho utilizar esse tipo de modelo). Entretanto, esse tipo de modelo tem como suposição que a variância do período a ser previsto é uma média ponderada dos resíduos quadrados do número fixo de períodos passados; porém essa suposição de pesos iguais não é muito boa,

visto que acontecimentos mais recentes têm um efeito maior na volatilidade a ser estimada, do que acontecimentos mais antigos, destarte eles deveriam ter um peso maior; além de a suposição de que acontecimentos mais antigos tenham peso zero na estimação da variância. Então, a vantagem do ARCH é possibilitar a estimação desses pesos como parâmetros a serem estimados, de forma que o modelo permita que os dados determinem os melhores pesos a serem usados na previsão da variância. Já os modelos GARCH, que seriam a generalização dos modelos ARCH, considera a variância como uma média ponderada dos resíduos quadrados passados, mas com pesos declinantes que nunca chegam completamente a zero. Nas palavras de Engle:

(...) virtualmente, nenhum método estava disponível para prever o comportamento da variância antes da introdução dos modelos ARCH. A principal técnica para isso era o “*rolling standard deviation*”. Isso é, o desvio padrão era calculado usando um número fixo das observações mais recentes. Por exemplo, isso poderia ser calculado todos os dias usando o mês mais recente (22 dias úteis) de dados. É conveniente pensar nesse tipo formulação como o primeiro modelo ARCH; é suposto que a variância do retorno de amanhã seja uma média ponderada dos resíduos quadrados dos últimos 22 dias. Essa suposição de pesos iguais não é atraente, uma vez que podemos pensar que os eventos mais recentes podem ser de maior relevância e portanto deveriam ter pesos maiores. Além disso, a suposição de um peso de zero para as observações mais antigas do que um mês de idade também não é atraente. O modelo ARCH proposto por Engle (1982) deixam esses pesos serem parâmetros a serem estimados. Então, o modelo permite que os dados determinem os melhores pesos a serem usados na previsão da variância.

Uma generalização útil desses modelos é a parametrização GARCH introduzida por Bollerslev (1986). Esse modelo também tem uma média ponderada dos resíduos quadrados passados, mas com pesos declinantes que nunca chegam completamente a zero. Isso dá uma parcimônia aos modelos que são fáceis de estimar e que, mesmo em suas formas mais simples, têm se provado surpreendentemente bem sucedidos na previsão de variâncias condicionais. (ENGLE, 2001)

### 3.2.1 ARCH(q)

Depois dessa introdução com um pouco da história e da intuição por trás dos modelos ARCH e GARCH vamos apresentar a técnica para estimarmos esses modelos. Iniciamos a apresentação pelo modelo mais simples que é o modelo ARCH de Engle

(1982). Nesse modelo consideramos que  $\varepsilon_t$  é um processo estocástico real em tempo discreto, condicional à informação em  $t - 1$ . Supondo que queremos estimar o modelo  $ARIMA(p, 0, q)$ . Se o erro desse processo segue um processo  $ARCH$ , então ele é definido da seguinte forma:

$$\varepsilon_t = \sigma_t u_t, u_t \sim i. i. d. (0, 1);$$

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2.$$

A variância incondicional desse processo é dada por:

$$E(\varepsilon_t^2) = E(\sigma_t^2 u_t^2) = E(\sigma_t^2) E(u_t^2) = \sigma_\varepsilon^2 = \frac{\omega}{1 - \sum_{i=1}^q \alpha_i}.$$

Algumas restrições a esse modelo são necessárias para garantir estacionariedade e positividade da variância condicional (que não pode ser negativa). Em primeiro lugar, uma condição suficiente para que a variância condicional seja sempre positiva é que  $\omega > 0$  e  $\alpha_i \geq 0, \forall i, i = 1, 2, \dots, q$ . Além disso,  $\sum_{i=1}^q \alpha_i < 1$  é condição necessária para que a série seja estacionária. Isso significa que as raízes do polinômio  $(1 - \sum_{i=1}^q \alpha_i L^i)$  devem estar fora do círculo unitário. Podemos representar o modelo condicional no seu segundo momento da seguinte maneira:

$$E(\sigma_t^2 | I_{t-1}) = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2$$

### 3.2.2 GARCH(p,q)

Uma vez que os modelos  $ARCH(q)$  precisam de muitos parâmetros para serem ajustados corretamente, Bollerslev (1986) tentou reduzir esse problema e sugeriu a formulação de um modelo mais parcimonioso, o modelo de heterocedasticidade condicional auto-regressivo generalizado (GARCH na sigla inglesa), no qual a equação da variância muda para:

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \quad (3.2)$$

Embora párea que nessa formulação haja mais parâmetros que no ARCH, a verdade é que, em geral,  $q_{ARCH} > q_{GARCH} + p_{GARCH}$ . De novo, a condição suficiente para estacionariedade é que as raízes do polinômio da variância estejam fora do círculo unitário, como no caso dos modelos *ARIMA*.

O modelo GARCH (p,q) pode ser interpretado como um processo auto-regressivo em  $\varepsilon_t^2$ . Para vermos isso, definimos:

$$v_t = \varepsilon_t^2 - \sigma_t^2 = (u_t^2 - 1)\sigma_t^2$$

$$v_t = \varepsilon_t^2 - \sigma_t^2 \Rightarrow \sigma_t^2 = \varepsilon_t^2 - v_t$$

Então, substituímos esse processo na equação (3.2)

$$\varepsilon_t^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j (\varepsilon_{t-j}^2 - v_{t-j}) + v_t$$

Reescrevendo:

$$\varepsilon_t^2 = \omega + \sum_{i=1}^{\max[p,q]} (\alpha_i + \beta_i) \varepsilon_{t-i}^2 - \sum_{j=1}^p \beta_j v_{t-j} + v_t$$

Essa ultima equação lembra um modelo ARMA (p,q), o que sugere que a identificação do modelo GARCH pode seguir a mesma metodologia de Box-Jenkins. Porém, essa metodologia só vale para a identificação de modelos GARCH se se trata de fato de um modelo GARCH. Outro aspecto importante desse modelo é que as restrições impostas sobre os coeficientes dos modelos GARCH, de que eles devem ser todos maiores que zero, pode ser relaxada, uma vez que as condições gerais de estacionariedade que devem ser impostas sobre os coeficientes não foram derivadas ainda.

### 3.2.3 Testes para detecção de modelos GARCH

Agora que já vimos a formulação desses modelos, veremos como podemos identificá-los. Para verificarmos se há a presença de erros GARCH há três testes principais: o de FAC e FACP, o teste Ljung-Box e teste de multiplicador de Lagrange.

O modelo GARCH (p,q) se assemelha ao modelo ARMA (p,q). Assim, as funções de autocorrelação, FAC, e de autocorrelação parcial FACP, devem sugerir se a série é heterocedástica, assim como devem dar uma ideia sobre as ordens p e q de um modelo ARMA. O modo de procedermos para realizarmos o teste é o seguinte: primeiro obtemos os quadrados dos erros estimados pela regressão do modelo estabelecido, como se a variância condicional fosse constante:

$$\bar{\sigma}^2 = \frac{1}{T} \sum_{t=1}^T \hat{\sigma}_t^2,$$

Onde T é o número de resíduos.

Em seguida, calculamos a FAC amostral para o quadrado dos resíduos, sem esquecer de a representarmos graficamente. A FAC é dada por:

$$\hat{\rho}_j = \frac{\frac{\sum_{t=j+1}^T (\hat{\epsilon}_t^2 - \bar{\sigma}^2)(\hat{\epsilon}_{t-j}^2 - \bar{\sigma}^2)}{T}}{\frac{\sum_{t=j+1}^T (\hat{\epsilon}_t^2 - \bar{\sigma}^2)}{T}}, \quad j = 1, 2, \dots$$

Podemos aproximar o desvio-padrão de  $\hat{\rho}_j$  em grandes amostras por  $T^{-0,5}$ . Diz-se que algum modelo ARCH está presente se os valores de  $\hat{\rho}_j$  forem estatisticamente diferentes de zero. A FACP é obtida pondo em um gráfico o coeficiente estimado,  $\phi_{s,s}$  contra a defasagem s, a partir da seguinte equação:

$$\hat{\epsilon}_t^2 = \phi_s + \phi_{s,1} \hat{\epsilon}_{t-1}^2 + \phi_{s,2} \hat{\epsilon}_{t-2}^2 + \dots + \phi_{s,s} \hat{\epsilon}_{t-s}^2 + u_t, \quad s = 1, 2, \dots, q.$$

De modo semelhante, o desvio padrão pode ser aproximado por  $T^{-0,5}$ .

Diferentemente dos modelos ARMA, a FAC dá a ordem máxima da parte autorregressiva do GARCH, representada pelo termo  $\hat{\epsilon}_{t-i}^2$ , e a FACP dá a ordem p das médias móveis do “GARCH” do termo  $\sigma_{t-j}^2$ . Se  $p < \max [p, q]$ , sabemos exatamente as ordens p

e  $q$  do modelo. DO contrário, encontra-se a ordem  $q$  por tentativa e erro. Esse método de estimação apresenta dificuldades da mesma natureza da identificação nos modelos ARMA. Por essa razão usamos testes mais “rigorosos”, como o de Ljung-Box e o de Multiplicadores de Lagrange em busca da detecção de erros GARCH.

A estatística de Ljung-Box,  $Q$ , pode ser usada para testarmos a presença de heterocedasticidade condicional. Sua hipótese nula é que a soma das autocorrelações é estatisticamente diferente de zero. Ou seja:

$$H_0: \sum_{j=1}^n \rho_j = 0$$

$$H_1: \hat{\rho}_j \neq 0, \text{ para algum } j,$$

Em que

$$Q = T(T + 2) \sum_{j=1}^n \frac{\rho_j^2}{T - j} \xrightarrow{d} \chi_n^2$$

Onde  $\xrightarrow{d} \chi_n^2$  indica a convergência em distribuição para uma distribuição qui-quadrada com  $n$  graus de liberdade.

O teste multiplicador de Lagrange (LM), é bastante tradicional, inicialmente foi proposto por Engle (1982) para detectar processo ARCH. Definimos o teste da seguinte maneira:

$$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_q = 0$$

$$H_1: \alpha_i \neq 0, \text{ para algum } i = 1, 2, \dots, q$$

Em que

$$\hat{\varepsilon}_t^2 = \alpha_1 \hat{\varepsilon}_{t-1}^2 + \alpha_2 \hat{\varepsilon}_{t-2}^2 + \dots + \alpha_q \hat{\varepsilon}_{t-q}^2 + u_t,$$

Cuja estatística é dada por  $TR^2 \xrightarrow{d} \chi_q^2$ , em que  $R^2$  representa o coeficiente de determinação da equação estimada. A regressão terá pouco poder explicativo, se os coeficientes são conjuntamente próximos de zero, de forma que o  $R^2$  será baixo. Portanto, se a hipótese nula não é rejeitada, não há presença de erros ARCH. Já se a hipótese nula for

rejeitada, o teste não determina qual tipo de modelo ARCH deverá ser usado. Esses testes indicam apenas se a covariância condicional é, ou não, constante.

Uma vez que é uma tarefa muito trabalhosa definir que tipo de modelo GARCH melhor se adapta aos dados e qual é o número de defasagens que devemos utilizar, a alternativa usual da literatura para contornar esse problema é utilizar os critérios de Akaike, Schwarz ou Hannan-Quinn para selecionar o modelo a ser utilizado.

Depois de estimados os modelos GARCH, devemos refazer os teste de erros ARCH para verificarmos se os erros ARCH desapareceram, assim como fazer um teste de autocorrelação dos resíduos. Um teste popular entre esse é o de Breusch-Godfrey. Para implementarmos esse teste, fazemos a seguinte regressão:

$$\hat{\varepsilon}_t = \alpha_1 \hat{\varepsilon}_{t-1} + \alpha_2 \hat{\varepsilon}_{t-2} + \dots + \alpha_q \hat{\varepsilon}_{t-q} + u_t$$

E então testamos as seguintes hipóteses:

$$H_0: \alpha_1 = \alpha_2 = \dots = \alpha_q = 0$$

$$H_1: \alpha_i \neq 0, \text{ para algum } i = 1, 2, \dots, q$$

Cuja estatística é dada por  $LM_q = TR^2 \xrightarrow{d} \chi_q^2$ , em que  $R^2$  representa o coeficiente de determinação da equação estimada. A regressão terá pouco poder explicativo, se os coeficientes são conjuntamente próximos de zero, de forma que o  $R^2$  será baixo. Portanto, se a hipótese nula não é rejeitada, não há indícios para se afirmar que exista autocorrelação dos resíduos.

### 3.3 Modelos GARCH bivariados

Agora que vimos como funcionam o modelos GARCH para séries univariadas vamos passar para os modelos GARCH par série multivariadas. Bem, se a série de dados tem várias variáveis, podemos estimar as volatilidades condicionais de cada variável simultaneamente. Modelos GARCH multivariados se aproveitam do fato de que choques contemporâneos às variáveis podem ser relacionados entre eles. Além disso, essa classe de modelos permite “transbordamento” de volatilidade, de forma a choques de volatilidade para uma variável podem afetar a volatilidade de outras variáveis relacionadas.

Nesse trabalho, por exemplo, tratamos de duas variáveis importantes de política macroeconômica, inflação e taxa de câmbio. Nós podíamos modelar apenas a inflação, ou então modelar apenas a taxa de câmbio e então, modelar as variâncias dessas séries separadamente. Contudo, a hipótese desse trabalho é que essas séries são inter-relacionadas. Hipótese que testaremos com evidência empírica.

Supondo que elas sejam inter-relacionadas, é de se esperar que choques que aumentem a incerteza de uma série, deve aumentar a incerteza de outra série. Para deixarmos a análise o mais simples possível, suponhamos que existem duas variáveis  $y_{1t}$  e  $y_{2t}$ . Por enquanto, não vamos nos preocupar com a média dessas séries para nos concentrarmos apenas nos seguintes dois processos de erro:

$$\varepsilon_{1t} = v_{1t}(h_{11t})^{0,5}$$

$$\varepsilon_{2t} = v_{2t}(h_{22t})^{0,5}$$

De forma similar ao do caso univariado, se assumimos que  $var(v_{1t}) = var(v_{2t}) = 1$ . Então, podemos pensar nas variáveis  $h_{11t}$  e  $h_{22t}$  como sendo as variâncias condicionais de  $\varepsilon_{1t}$  e de  $\varepsilon_{2t}$ , respectivamente. Uma vez que queremos que haja a possibilidade de os choques serem correlacionados. Vamos deixar que  $h_{12t}$  seja a notação para a covariância condicional entre os dois choques. Especificamente,  $h_{12t} = E_{t-1}\varepsilon_{1t}\varepsilon_{2t}$ .

Uma maneira natural de se construir um GARCH (1,1) multivariado é permitir que todos os termos de volatilidade interajam entre si. Considerando o modelo *vech*:

$$h_{11t} = c_{10} + \alpha_{11}\varepsilon_{1t-1}^2 + \alpha_{12}\varepsilon_{1t-1}\varepsilon_{2t-1} + \alpha_{13}\varepsilon_{2t-1}^2 + \beta_{11}h_{11t-1} + \beta_{12}h_{12t-1} + \beta_{13}h_{22t-1} \quad (3.3)$$

$$h_{12t} = c_{20} + \alpha_{21}\varepsilon_{1t-1}^2 + \alpha_{22}\varepsilon_{1t-1}\varepsilon_{2t-1} + \alpha_{23}\varepsilon_{2t-1}^2 + \beta_{21}h_{11t-1} + \beta_{22}h_{12t-1} + \beta_{23}h_{22t-1} \quad (3.4)$$

$$h_{22t} = c_{30} + \alpha_{31}\varepsilon_{1t-1}^2 + \alpha_{32}\varepsilon_{1t-1}\varepsilon_{2t-1} + \alpha_{33}\varepsilon_{2t-1}^2 + \beta_{31}h_{11t-1} + \beta_{32}h_{12t-1} + \beta_{33}h_{22t-1} \quad (3.5)$$

usando notação matricial:

$$vech(H_t) = C + A vech(\varepsilon_{t-1}\varepsilon'_{t-1}) + B vech(H_{t-1})$$

Onde:

$$\text{vech}(H_t) = \begin{bmatrix} h_{11t} \\ h_{12t} \\ h_{22t} \end{bmatrix}, \quad \varepsilon_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}, \quad \varepsilon_{t-1}\varepsilon'_{t-1} = \begin{bmatrix} \varepsilon_{1t}^2 & \varepsilon_{1t}\varepsilon_{2t} \\ \varepsilon_{1t}\varepsilon_{2t} & \varepsilon_{2t}^2 \end{bmatrix}, \quad \text{vech}(\varepsilon_{t-1}\varepsilon'_{t-1}) = \begin{bmatrix} \varepsilon_{1t}^2 \\ \varepsilon_{1t}\varepsilon_{2t} \\ \varepsilon_{2t}^2 \end{bmatrix}, \quad C = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}, \quad A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \text{ e } B = \begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \\ \beta_{31} & \beta_{32} & \beta_{33} \end{bmatrix}$$

Nesse modelo, a variância condicional de cada variável ( $h_{11t}$  e  $h_{22t}$ ) depende de suas realizações passadas, a covariância condicional entre as duas variáveis  $h_{12t}$ , os quadrados dos seus erros defasados ( $\varepsilon_{1t-1}^2, \varepsilon_{2t-1}^2$ ), e o produto de seus erros defasados  $\varepsilon_{1t-1}\varepsilon_{2t-1}$ . Reparamos que a covariância condicional depende dessas mesmas variáveis. Entretanto, ao observarmos essas equações reparamos que surgem alguns problemas como equações demais e parâmetros demais e falta de solução analítica para o problema de maximização.

Para contornar esses problemas, muitos dos recentes trabalhos sugerem restrições para a equação (3.3)-(3.5). Um modelo que ficou popular na literatura foi o chamado vech diagonal. A ideia é diagonalizar o sistema de tal forma que  $h_{ijt}$  contenha apenas as suas defasagens e o produto cruzado  $\varepsilon_{it}\varepsilon_{jt}$ . Um exemplo da versão diagonalizada de (3.3)-(3.5) é:

$$h_{11t} = c_{10} + \alpha_{11}\varepsilon_{1t-1}^2 + \beta_{11}h_{11t-1}$$

$$h_{12t} = c_{20} + \alpha_{22}\varepsilon_{1t-1}\varepsilon_{2t-1} + \beta_{22}h_{12t-1}$$

$$h_{22t} = c_{30} + \alpha_{33}\varepsilon_{2t-1}^2 + \beta_{33}h_{22t-1}$$

Como nesse modelo só se usa os valores da diagonal principal de A e B, em notação matricial, o modelo fica da mesma maneira que do modelo vech mostrado anteriormente:

$$\text{vech}(H_t) = C + A \text{vech}(\varepsilon_{t-1}\varepsilon'_{t-1}) + B \text{vech}(H_{t-1})$$

Mas as matrizes A e B ficam da seguinte maneira:

$$A = \begin{bmatrix} \alpha_{11} & 0 & 0 \\ 0 & \alpha_{22} & 0 \\ 0 & 0 & \alpha_{33} \end{bmatrix} \text{ e } B = \begin{bmatrix} \beta_{11} & 0 & 0 \\ 0 & \beta_{22} & 0 \\ 0 & 0 & \beta_{33} \end{bmatrix}$$

Dado o numero de restrições, esse modelo é relativamente fácil de ser estimado. Cada variância condicional é equivalente ao processo GARCH univariado e a covariância condicional, também é bastante parcimoniosa. O problema desse modelo é se igualarmos  $\alpha_{ij} = \beta_{ij} = 0$  (quando  $i \neq j$ ), significa que não há interação entre as variâncias. Por exemplo, um choque  $\varepsilon_{1t-1}$  afeta  $h_{11t}$  e  $h_{12t}$  porém não afeta a variância condicional  $h_{22t}$ .

Para que possamos maximizar a verossimilhança, é necessário que imponhamos restrições para garantir que a matriz de covariância seja positiva definida. Por isso, estima-se a função de verossimilhança sujeita à restrição de que os autovalores da matriz de covariância condicional sejam simultaneamente maiores que zero. Para estacionariedade, é preciso que as raízes do seguinte polinômio estejam fora do círculo unitário:

$$\left| I - \sum_{i=1}^{\max [p,q]} (A_i + B_i)\lambda^i \right| = 0$$

Onde  $\lambda$  é o autovalor associado a cada um dos autovetores.

Outro modelo inventado foi o BEKK (Acrônimo de Baba Engle Kraft e Kroner). Esse modelo assegura que as variâncias condicionais serão positivas. A ideia central desse modelo é forçar que todos os parâmetros entrem no modelo via formas quadráticas, o que asseguraria que todas as variâncias serão positivas. Apesar de esse modelo ter várias formas que variam entre si, consideremos a seguinte especificação:

$$H_t = C'C + A'\varepsilon_{t-1}\varepsilon_{t-1}'A + B'H_{t-1}B$$

Onde para o caso de 2 variáveis considerado nas equações (3.3)-(3.5)

$$H_t = \begin{bmatrix} h_{11t} & h_{12t} \\ h_{12t} & h_{22t} \end{bmatrix}; C = \begin{bmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{bmatrix}; A = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix}; B = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix}$$

Se realizarmos as multiplicações matriciais indicadas, vamos encontrar a seguinte equação:

$$h_{11t} = (c_{11}^2 + c_{12}^2) + (\alpha_{11}^2 \varepsilon_{1t-1}^2 + 2\alpha_{11}\alpha_{21}\varepsilon_{1t-1}\varepsilon_{2t-1} + \alpha_{21}^2 \varepsilon_{2t-1}^2) \\ + (\beta_{11}^2 h_{11t-1} + 2\beta_{11}\beta_{21}h_{12t-1} + \beta_{21}^2 h_{22t-1})$$

$$h_{12t} = c_{12}(c_{11} + c_{22}) + \alpha_{12}\alpha_{11}\varepsilon_{1t-1}^2 + (\alpha_{11}\alpha_{22} + \alpha_{12}\alpha_{21})\varepsilon_{1t-1}\varepsilon_{2t-1} \\ + \alpha_{22}\alpha_{22}\varepsilon_{2t-1}^2 + \beta_{11}\beta_{12}h_{11t-1} + (\beta_{11}\beta_{22} + \beta_{12}\beta_{22})h_{12t-1} \\ + \beta_{21}\beta_{22}h_{11t-1}$$

$$h_{22t} = (c_{22}^2 + c_{12}^2) + (\alpha_{12}^2 \varepsilon_{1t-1}^2 + 2\alpha_{12}\alpha_{22}\varepsilon_{1t-1}\varepsilon_{2t-1} + \alpha_{22}^2 \varepsilon_{2t-1}^2) \\ + (\beta_{12}^2 h_{11t-1} + 2\beta_{12}\beta_{22}h_{12t-1} + \beta_{22}^2 h_{22t-1})$$

A condição de estacionariedade para essa formulação dos GARCH multivariados, segundo Engle e Kroner (1995) é que os autovalores da seguinte matriz sejam menores um.

$$\sum_{i=1}^q \sum_{k=1}^K (B_{ik} \otimes B_{ik}) + \sum_{i=1}^p \sum_{k=1}^K (A_{ik} \otimes A_{ik})$$

Onde  $\otimes$  é o produto de Kronecker.

Em geral,  $h_{ij}$  dependerá do quadrado de seus próprios resíduos, produtos cruzados dos resíduos e a covariância de todas de todas as variáveis do sistema. Dessa forma, o modelo permite que choques à variância de uma das variáveis “transbordem” sobre os outros. O problema é que a formulação BEKK pode ser difícil de ser estimada. O modelo tem um numero grande de parâmetros que não são globalmente identificados. Se mudarmos os sinais de todos os elementos de A, B ou C não afetará o valor da função de verossimilhança, de tal forma que pode ser difícil de fazê-la convergir. Uma maneira de simplificar esse modelo é impondo que as matrizes A e B sejam diagonais de modo que devemos estimar sete parâmetros apenas.

Outro modelo popular da classe dos GARCH multivariados é a especificação Correlação Condicional Constante. Como o nome sugere, o modelo de coeficiente constante de correlação (CCC) restringe os coeficientes de correlação a serem constantes. Assim, para cada  $i \neq j$ , o modelo CCC assume que  $h_{ijt} = \rho_{ij}(h_{iit}h_{jtt})^{0,5}$ . De modo que, o modelo é um meio termo, onde os termos da variância não precisam ser diagonalizados, mas os termos da covariância são sempre proporcionais a  $(h_{iit}h_{jtt})^{0,5}$ . Por exemplo, o modelo CCC pode consistir em (3.3)-(3.5) e  $h_{12t} = \rho_{12}(h_{11t}h_{22t})^{0,5}$ . Portanto, a equação

da covariância implica em apenas um parâmetro em vez de sete parâmetros aparecendo em (2.4). A formulação CCC é claramente um caso especial do modelo GARCH multivariado mais geral, no caso de duas variáveis, nós podemos escrever  $H_t$  como:

$$H_t = \begin{bmatrix} h_{11t} & \rho_{12}(h_{11t}h_{22t})^{0,5} \\ \rho_{12}(h_{11t}h_{22t})^{0,5} & h_{22t} \end{bmatrix}$$

Se  $h_{11t}$  e  $h_{22t}$  são ambos os processos GARCH(1,1) devemos estimar sete parâmetros (os seis valores de  $c_i$ ,  $\alpha_{ii}$ ,  $\beta_{ii}$  e  $\rho_{12}$ ).

Concluindo, o que vamos precisar desse capítulo é o conceito de estacionariedade, do qual não podemos prescindir para modelarmos as nossas séries de interesse. Esse conceito será verificado por meio do teste de raiz unitária de Dickey e Fuller. Após isso, iremos modelar a média das séries de interesse utilizando modelos ARIMA, e modelaremos a volatilidade condicional das séries com modelos GARCH bivariados, mais precisamente o BEKK diagonal. Finalmente analisaremos a qualidade do ajuste dos modelos GARCH ao realizarmos os testes Q de Ljung-Box para autocorrelação dos resíduos, assim como o teste de Breusch-Godfrey para autocorrelação dos resíduos e finalmente o teste ARCH-LM para ver se os erros ARCH foram eliminados após a aplicação dos modelos GARCH.

## Capítulo 4

### Aplicação para a Economia Brasileira

Nesse capítulo vamos fazer a estimação do GARCH bivariado, para isso usaremos a formulação BEKK diagonal (DBEKK). Apesar de que com essa formulação perderemos a interação entre a variância de uma variável com outra, ganhamos em economia de parâmetros. Como o presente trabalho se baseia em Albuquerque e Portugal (2006), a hipótese inicial é que a relação entre as duas volatilidades é a mesma encontrada por Albuquerque e Portugal (2006), qual seja, a relação entre as volatilidades é semi-côncava. Isso, em termos econômicos, se traduziria em dizer que quando a volatilidade da taxa de câmbio é muito alta, o aumento da incerteza traz uma resposta à inflação reduzida, o que levaria a efeitos menores. Outro ponto importante para a verificação dessa relação é que poderemos verificar se houve alguma mudança nela, ou se ela manteve estável ao longo do tempo. Essa seção se dividirá em I) Apresentação dos dados, II) Formulação das equações da média, III) Formulação do GARCH bivariado.

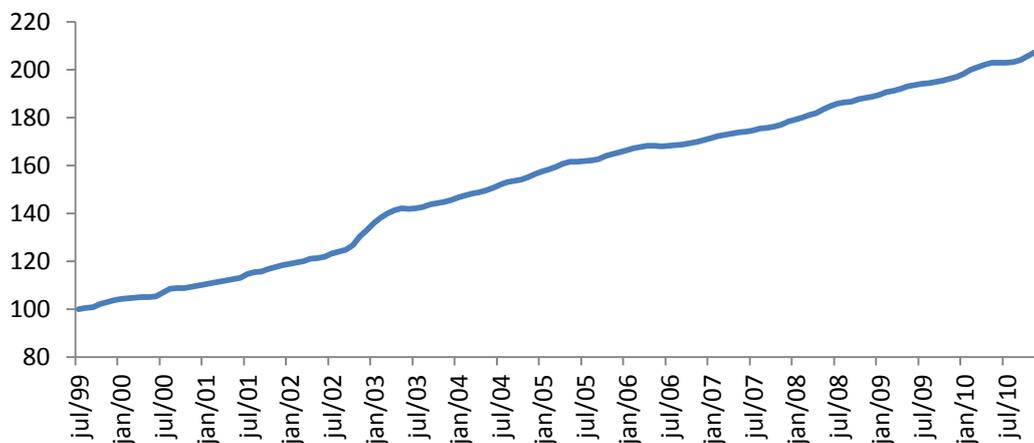
#### 4.1 Dados

As amostras das séries que usaremos serão mensais de julho de 1999 a dezembro de 2010. As séries que usaremos serão:

- (a) Índice de preços ao consumidor amplo (IPCA), esse índice de preços ao consumidor que é publicado pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Julho/1999=100. Podemos ver essa série no Gráfico 1

**GRÁFICO 4.1**  
**Taxa de inflação no Brasil pelo IPCA de Julho de 1999 a dezembro de 2010 (Índice base fixa Jul/1999=100)**

IPCA jul/99=100



(b) Taxa de câmbio R\$/US\$, preços de venda, média mensal. Podemos ver essa série no Gráfico 2.

**GRÁFICO 4.2**  
**Taxa de câmbio R\$/US\$, preços de venda, média mensal**

R\$/US\$



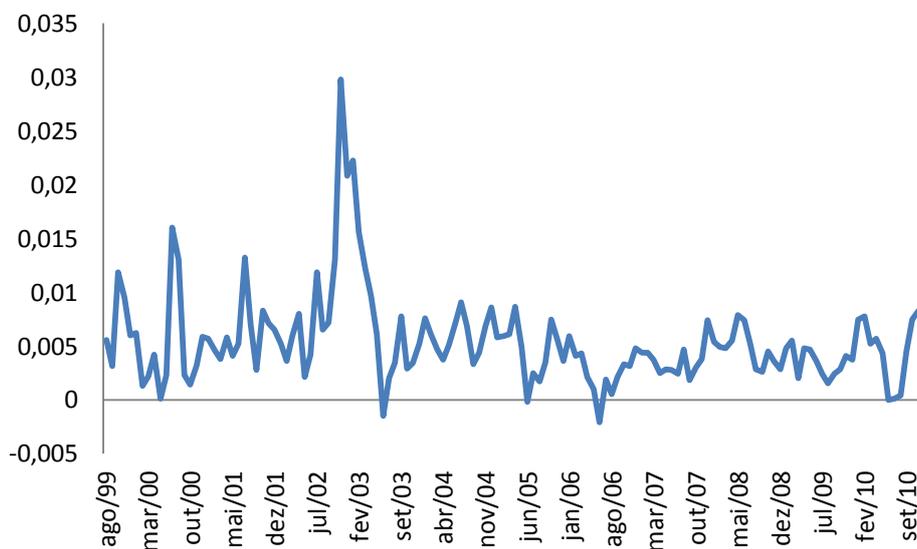
Essas duas séries foram transformadas em logaritmos naturais. Depois, foi feito testes de raiz unitária, levando em conta um intercepto e uma tendência. Nesses testes, as duas séries apresentaram raízes unitárias (dando a entender que elas não eram estacionárias). Apresentamos os testes de raízes unitária a seguir:

**TABELA 4.1**  
Teste ADF nas séries utilizadas

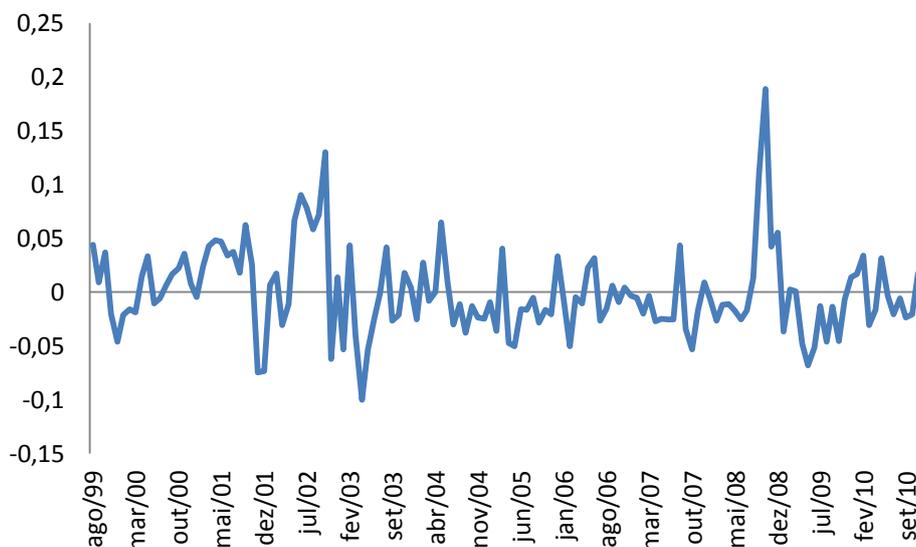
Vaiável	Estatística do Teste ADF	Valor Crítico em 5%	Estatística do Teste ADF - Primeira diferença da variável	Valor Crítico em 5%
Índice de Preços	-1.3332	-3.4432	-5.5393	-3.4432
Taxa de câmbio	-2.0167	-3.4432	-7.5634	-3.4432

Portanto, tiramos a primeira diferença de cada uma dessas séries, e agora a primeira diferença dos logaritmos do Índice de Preço e Taxa de câmbio serão chamadas de *IPCA* e *E*, respectivamente. A seguir apresentamos os gráficos das séries *IPCA* e *E*:

**GRÁFICO 4.3**  
Primeira diferença dos logaritmos do *IPCA*



**GRÁFICO 4.4**  
**Primeira diferença dos logaritmos da taxa de câmbio R\$/US\$**



#### 4.2 Formulação das equações da média

Antes de partirmos para a estimação dos modelos GARCH temos de fazer algumas hipóteses sobre as equações da média e ver se realmente há erros do tipo ARCH em nossas séries. Aqui testamos 4 formulações para a equação da média para cada variável (estamos testando essas formulações baseados na análise dos correlogramas de cada série, os quais indicam que cada uma é um AR puro). A primeira é um passeio aleatório com um drift e nas outras formulações foram adicionados termos defasados das variáveis. A seguir apresentamos um quadro de resumo das estimações com os coeficientes estimados e Critério de Akaike.

**TABELA 4.2**  
**Equações da Média do IPCA**

IPCA				
Variáveis	Caso 1	Caso 2	Caso 3	Caso 4
AIC	-8,0829	-8,6083	-8,5974	-8,6368
C	0,0054	0,0019	0,0021	0,0017
DP	0,0004	0,0004	0,0005	0,0005
IPCA(-1)		0,6494	0,7051	0,7289
DP		0,0657	0,0866	0,0851
IPCA(-2)			-0,0855	-0,1757
DP			0,0867	0,1038
IPCA(-3)				0,1141
DP				0,0850

**TABELA 4.3**  
**Equações da Média da taxa de câmbio**

E				
Variáveis	Caso 1	Caso 2	Caso 3	Caso 4
AIC	-3.5821	-3.7581	-3.7372	-3.7204
C	-0.0004	-0.0006	-0.0006	-0.0008
DP	0.0034	0.0031	0.0032	0.0032
IPCA(-1)		0.4135	0.4017	0.4031
DP		0.0783	0.0871	0.0875
IPCA(-2)			0.0321	0.0224
DP			0.0867	0.0945
IPCA(-3)				0.0061
DP				0.0874

Então, nos baseando pelos quadros de resumo acima escolhemos, para as equações de média, o modelo com um termo defasado da própria variável para cada uma das equações. Dessa forma, cada equação ficará da seguinte forma:

$$IPCA_t = \delta_0 + \delta_1 IPCA_{t-1} + \varepsilon_{1,t}$$

$$E_t = \gamma_0 + \gamma_1 E_{t-1} + \varepsilon_{2,t}$$

Escolhemos essas formulações para essas equações com base no Critério de Akaike e na parcimônia da formulação. Por isso que, apesar de a formulação com três defasagens para a equação do IPCA, apesar de ter um valor do AIC menor, escolhemos a formulação com apenas uma variável defasado, visto que o acréscimo no AIC é bem pequeno para adicionarmos mais duas variáveis. Já para a equação de E, a formulação com apenas uma defasagem é a mesma que minimiza o valor do AIC.

### 4.3 GARCH Bivariado

Agora, que já fizemos as hipóteses sobre as equações da média, faremos a estimação de cada variável por modelos GARCH bivariados. A formulação que consideramos foi o BEKK diagonal. Que tem a seguinte formulação:

$$H_t = M + A'(\varepsilon_{t-1}\varepsilon'_{t-1})A + B'H_{t-1}B$$

Onde  $A$  e  $B$  são matrizes diagonais, ou seja:

$$h_{11t} = m_{11} + a_{11}^2\varepsilon_{1t-1}^2 + b_{11}^2h_{11t-1}$$

$$h_{12t} = m_{12} + a_{11}a_{22}\varepsilon_{1t-1}\varepsilon_{2t-1} + b_{11}b_{22}h_{12t-1}$$

$$h_{22t} = m_{22} + a_{22}^2\varepsilon_{2t-1}^2 + b_{22}^2h_{22t-1}$$

Escolhemos esse modelo devido a praticidade que ele oferece, uma vez que temos de estimar apenas sete parâmetros. Além disso, essa formulação é bem parcimoniosa, o que deixa o nosso modelo bem simples. Então, usando a equação da média já destacada anteriormente, partimos para a estimação dos coeficientes do modelo GARCH, os quais são os que seguem na tabela 4.4, e os coeficientes das equações da média são apresentados a seguir:

$$IPCA_t = 0,0019 + 0,6494IPCA_{t-1}$$

$$E_t = -0,0006 + 0,4135E_{t-1}$$

**TABELA 4.4**  
**Coefficientes das matrizes do GARCH bivariado**

Valor da Função	857.0795
M(11)	0.0000
DP	0.0000
M(12)	0.0000
DP	0.0000
M(22)	0.0005
DP	0.0002
A(11)	0.3540
DP	0.1033
A(22)	0.5769
DP	0.1286
B(11)	0.9288
DP	0.0282
B(22)	0.5541
DP	0.1735

Depois disso, temos de calcular os autovalores para ver se a condição de estacionariedade foi obedecida. Para isso, os autovalores da seguinte matriz devem, em módulo, serem menores do que um:

$$\sum_{i=1}^q \sum_{k=1}^K (B_{ik} \otimes B_{ik}) + \sum_{i=1}^p \sum_{k=1}^K (A_{ik} \otimes A_{ik})$$

Em outras palavras, nós temos de calcular os autovalores da matriz  $X$ .

$$X = \begin{bmatrix} 0.9879 & 0 & 0 & 0 \\ 0 & 0,7189 & 0 & 0 \\ 0 & 0 & 0,7189 & 0 \\ 0 & 0 & 0 & 0,6399 \end{bmatrix}$$

E constatamos que os autovalores são:

$$\lambda_1 = 0.9879$$

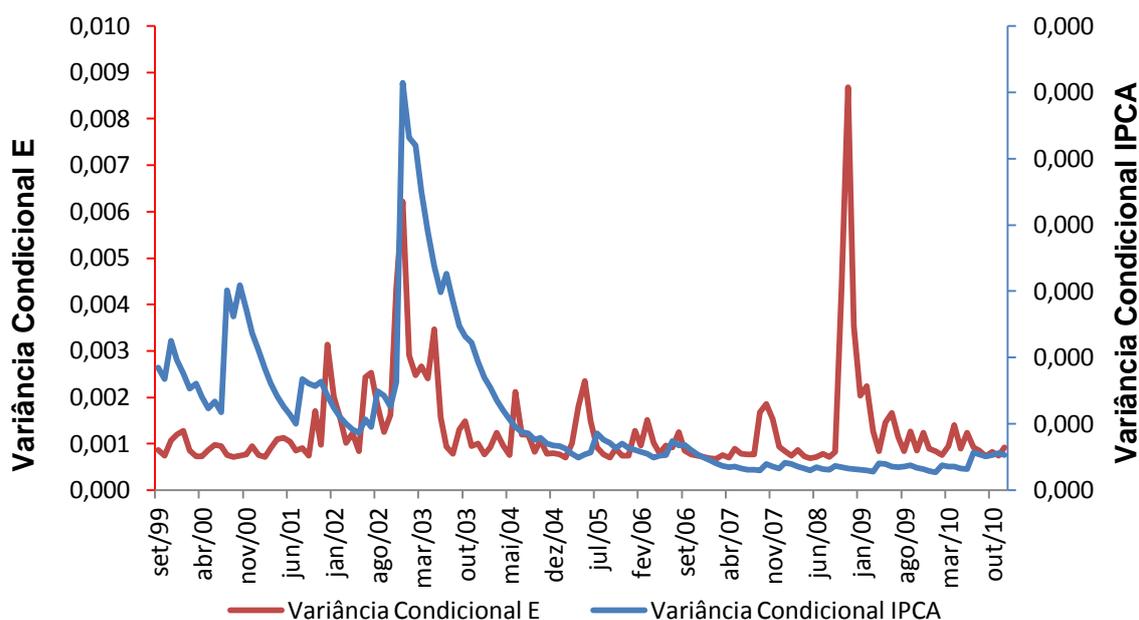
$$\lambda_2 = 0,7189$$

$$\lambda_3 = 0,7189$$

$$\lambda_4 = 0,6399$$

Ou seja, os quatro autovalores, em módulo, são menores que um, o que nos permite afirmar que a equação obedece a condição de estacionariedade. Os gráficos das variâncias condicionais são os seguintes:

**GRÁFICO 4.5**  
**Variâncias Condicionais**



Depois disso, temos de verificar se o modelo GARCH ficou bem ajustado. Para isso vamos fazer os testes de Ljung-Box e o teste ARCH-LM para verificarmos se não há autocorrelação dos resíduos e se conseguimos “limpar” os erros ARCH com a nossa modelagem. Primeiro fazemos o teste de Ljung Box que é apresentado na tabelas 4.5 abaixo.

**TABELA 4.5**  
**Teste Ljung-Box dos resíduos do IPCA e do Câmbio**

Defasagem	Resíduos da Equação da Inflação		Resíduos da Equação da Taxa de Câmbio	
	Q-Stat	Prob	Q-Stat	Prob
1	0,2305	0,6310	0,4798	0,4890
2	3,5469	0,1700	1,1204	0,5710
3	3,6047	0,3070	2,1945	0,5330
4	4,0069	0,4050	2,2771	0,6850
5	4,8779	0,4310	4,1366	0,5300
6	4,9798	0,5460	5,3389	0,5010
7	7,8468	0,3460	9,8570	0,1970
8	7,8914	0,4440	14,0210	0,0810
9	11,5900	0,2370	14,6330	0,1020
10	11,6290	0,3110	14,7050	0,1430
11	11,6350	0,3920	16,3890	0,1270
12	17,2150	0,1420	16,4070	0,1730
13	17,6470	0,1710	16,4160	0,2270
14	19,1850	0,1580	19,8680	0,1340
15	21,4230	0,1240	20,8110	0,1430
16	21,5400	0,1590	20,8860	0,1830
17	21,7400	0,1950	21,7310	0,1950
18	22,0850	0,2280	22,4410	0,2130
19	22,1080	0,2790	24,1850	0,1890
20	22,4540	0,3160	25,2160	0,1930

Depois disso, realizaremos o teste ARCH para vermos se conseguimos “limpar” os erros ARCH, os resultados desse teste é apresentado nas tabelas 4.6 e 4.7 abaixo:

**TABELA 4.6**  
**Teste ARCH-LM dos resíduos do IPCA**

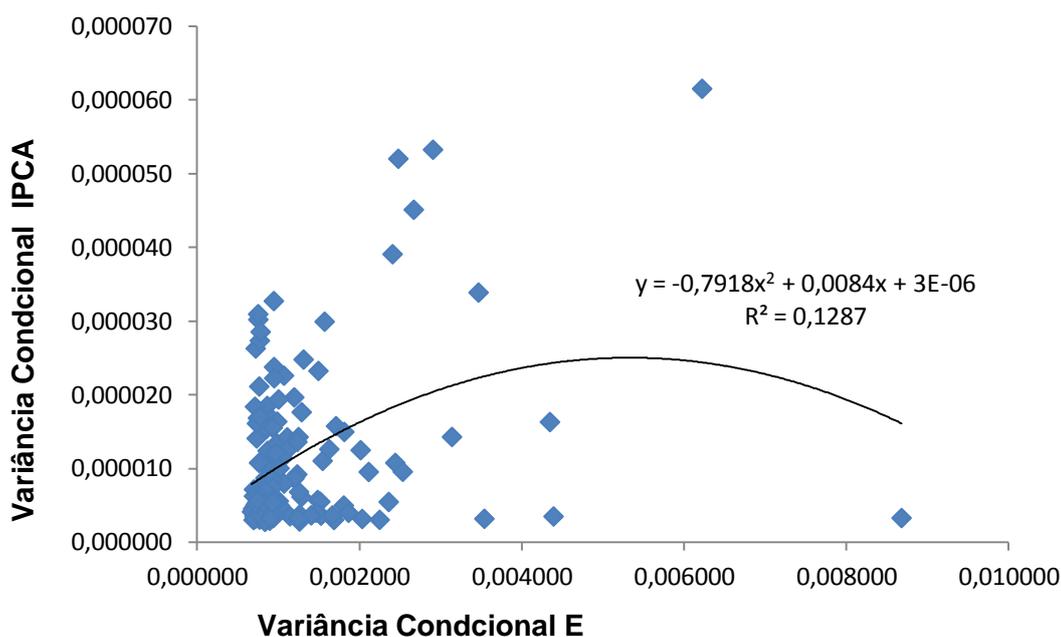
Teste ARCH-LM do IPCA			
Estatística F	0.0534	Prob. F	0.9837
TxR <sup>2</sup>	0.1649	Prob. Qui-quadrado	0.9830
Variáveis	Coefficiente	Desvio Padrão	p-valor
C	1.0571	0.2768	0.0002
WGT_RESID <sup>2</sup> (-1)	0.0012	0.0880	0.9895
WGT_RESID <sup>2</sup> (-2)	-0.0005	0.0876	0.9958
WGT_RESID <sup>2</sup> (-3)	-0.0350	0.0876	0.6899

**TABELA 4.7**  
**Teste ARCH-LM dos resíduos da taxa de câmbio**

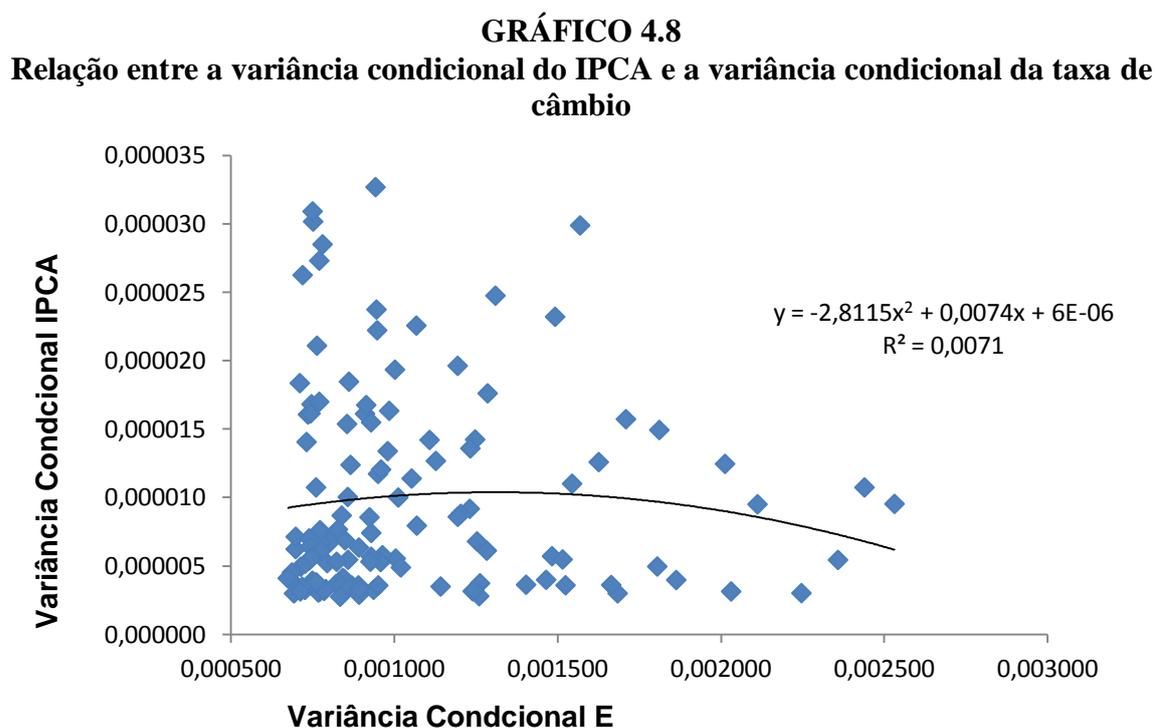
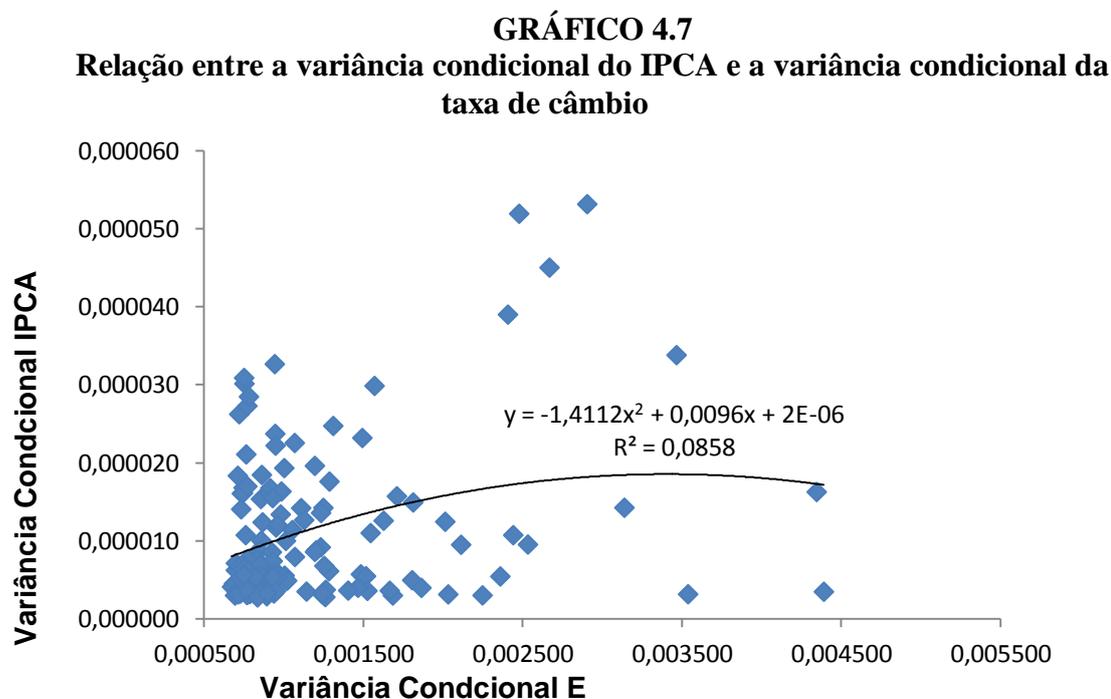
Teste ARCH-LM de E			
Estatística F	0.8882	Prob. F	0.4492
TxR <sup>2</sup>	2.6916	Prob. Qui-quadrado	0.4417
Variáveis	Coefficiente	Desvio Padrão	p-valor
C	0.9994	0.2084	0.0000
WGT_RESID <sup>2</sup> (-1)	0.1108	0.0878	0.2091
WGT_RESID <sup>2</sup> (-2)	-0.0246	0.0882	0.7812
WGT_RESID <sup>2</sup> (-3)	-0.0842	0.0877	0.3388

Como podemos observar, a partir das tabelas 4.5, 4.6 e 4.7, não há evidências de que os resíduos das equações possuem autocorrelação, assim como não há evidências que ainda há os erros ARCH, o que indica que o modelo GARCH está bem ajustado. Agora, examinamos os gráficos de dispersão das variâncias condicionais de *E* e do *IPCA*, onde este é representado no eixo vertical e aquele é representado no eixo horizontal. Nesse gráfico, fazemos uma regressão por MQO para acharmos a relação entre essas variáveis.

**GRÁFICO 4.6**  
**Relação entre a variância condicional do IPCA e a variância condicional da taxa de câmbio**



Esse gráfico, claramente, apresenta 2 outliers, que representam os períodos de Dez/02 e Nov/08. Retiramos esses pontos e obtemos o seguinte gráfico 4.7, que apresenta 9 outliers - Dez/01, Nov/02, Jan/03, Fev/03, Mar/03, Abr/03, Mai/03 Out/08 Dez/08, os quais retiramos e obtemos o gráfico 4.8.



A partir da análise desses gráficos, podemos chegar a conclusão que a relação entre as variâncias condicionais é semi-côncava, assim como em Albuquerque e Portugal (2006), ou seja, as evidências empíricas apresentadas por esse trabalho nos levam a acreditar que a volatilidade da taxa de câmbio e da inflação tem uma relação semi-côncava e que, apesar de a resposta da volatilidade da inflação à volatilidade da taxa de câmbio ser positiva, a proporção na sua variação diminui quando a volatilidade da taxa de câmbio cresce. Além disso, associado a esse trabalho, podemos afirmar que esse padrão –que já foi verificado em Albuquerque e Portugal (2006) – está se repetindo no tempo.

## 5 Conclusão

A partir da análise do capítulo anterior, podemos chegar a conclusão que a relação entre as variâncias condicionais é semi-côncava, assim como em Albuquerque e Portugal (2006). Essa conclusão se associa com o grupo que reconhece que há relação entre a volatilidade cambial e a volatilidade de outras variáveis macroeconômicas. Mais especificamente, ela se associa com os trabalhos de Dixit (1989) e Seabra (1996).

Então, as evidências que encontramos nesse trabalho nos levam a crer que quando a volatilidade cambial está muito alta, a incerteza aumenta e a resposta da inflação pode ser reduzida levando a efeitos menores, o que também pode ser confundido com a hipótese da desconexão entre a volatilidade cambial e a volatilidade de outras variáveis macroeconômicas, a qual é apresentada por Flood e Rose (1999). Entretanto, rejeitamos a hipótese da desconexão da volatilidade da taxa de câmbio com a volatilidade da inflação visto que o que se verifica é que, em períodos de alta volatilidade, os agentes não respondem com a mesma intensidade que eles reagiriam quando a volatilidade é baixa. Isso pode ter várias razões. Uma hipótese pode ser que devido a sua falta de informação relativa à duração dos movimentos da taxa de câmbio, os agentes adotam a estratégia “wait-and-see”. Outra hipótese é que, quando os agentes percebem o período de grande volatilidade cambial, eles podem fazer hedges cambiais para se protegerem dessa volatilidade, o que estabiliza a inflação do país.

Porém, a relação entre as volatilidades que encontramos também implica que quando a volatilidade cambial estiver menor, a inflação irá responder mais prontamente aos movimentos do câmbio. Dessa forma, a hipótese dos agentes que fazem hedge ganha força, pois quando os movimentos não são muito grandes os agentes não têm incentivos para fazer hedge e, quando há um movimento na taxa de câmbio, ele tenta repassar esse custo na venda dos produtos, o que eleva a volatilidade da inflação.

Destarte, as evidências empíricas apresentadas por esse trabalho nos levam a acreditar que a volatilidade da taxa de câmbio e da inflação tem uma relação semi-côncava e que, apesar de a resposta da volatilidade da inflação à volatilidade da taxa de câmbio ser positiva, a proporção na sua variação diminui quando a volatilidade da taxa de câmbio

crece. Além disso, associado a esse trabalho, podemos afirmar que esse padrão, o qual já foi verificado em Albuquerque e Portugal (2006), está se repetindo no tempo.

Outro fato que destacamos é que esse trabalho, assim como o de Albuquerque e Portugal (2006), utiliza um método relativamente novo para modelar a variância das variáveis ao usar um GARCH bivariado, em vez de sub-amostras da série completa, o que pode afetar os resultados dependendo dos critérios usados pelo pesquisador e do tamanho das sub-amostras. Portanto, ao aplicarmos aquele método, conseguimos dados mais confiáveis para a nossa pesquisa.

## 6 Referências:

ALBUQUERQUE, C. R.; PORTUGAL, M. S. Testing Nonlinearities Between Brazilian Exchange-Rate and Inflation Volatilities. **Revista Brasileira de Economia**, Rio de Janeiro, v. 60, n. 4, 2006.

BLEANEY, M. & FIELDING, D. (2002). Exchange rate regimes, inflation and output volatility in developing countries. **Journal of Development Economics**, 68(1):233–245.

BOLLERSLEV, T. Generalized Autoregressive Conditional Heteroskedasticity, **Journal of Econometrics**. April, 31:3, pp. 307-27, 1986.

BUENO, Rodrigo De Losso da Silveira, **Econometria de Séries Temporais**, São Paulo,:Cenage Learning, 2011

CALVO, G.; REINHART, C.; Fear of floating; **NBER Working Papers 7993**, National Bureau of Economic Research, Inc., 2000.

CRYER, J. D.; CHAN, K. **Time series analysis: with applications in R**. New York, Springer Science+Business Media, 2008.

DEVEREUX, M. B.; ENGEL, C. Exchange rate pass-through, exchange rate volatility and Exchange rate disconnect, **Journal of Monetary Economics**, v.40, p.913-940, 2002

DIXIT, A. K. Hysteresis, import penetration, and exchange rate pass-through. **The Quarterly Journal of Economics**, 104(2):205–28, 1989.

EICHENGREEN, B. Can emerging markets float the way they float? Should they inflation target?, **Banco Central do Brasil**, fevereiro, 2002, trabalhos para discussão n.36

ENDERS, W.. **Applied econometric time series**. 2011, John Wiley & Sons, Nova York.

ENGLE, R. F. Autoregressive Conditional Heteroskedasticity with estimates of the variance of the United Kingdom, **Econometrica**, vol. 50, pp 987-1007, 1982.

ENGLE, R. F. GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics, **Journal of Economic Perspectives**, American Economic Association, vol. 15(4), pages 157-168, 2001.

ENGLE, R. F.; KRONER, K. F. Multivariate Simultaneous Generalized ARCH, **Econometric Theory**, Cambridge University Press, vol. 11(01), pages 122-150, February, 1995.

FLOOD, R. P.; ROSE A. K., Fixing exchange rates – a virtual quest for fundamentals, **Journal of Monetary Economics**, v. 36, p.3-37, 1995

FLOOD, R. P.; ROSE, A. K., Understanding Exchange Rate Volatility Without The Contrivance of Macroeconomics, **The Economic Journal**, Vol. 109, po. F660-F672, Novembro, 1999.

FRIEDMAN, M; “The case for flexible exchange rates”; in **Essays in Positive Economics**; Chicago: University of Chicago Press, 1953

GHOSH, A. R., GULDE, A.-M., OSTRY, J. D., & WOLF, H. C. **Does the nominal exchange rate regime matter?** NBER Working Papers 5874, National Bureau of Economic Research, Inc., 1997.

HAUSMANN, R.; PANIZZA U.; STEIN, E.; Why do countries float the way they float, **Journal of Development Economics**, v. 66, p.387-414, 2001

KRUGMAN, P.; OBSTFELD, M.; **Economia Internacional**, São Paulo, Pearson Prentice Hall, 2010

LÜTKEPOHL, H., **New Introduction to Multiple Time Series Analysis**, Berlin: Springer-Verlag, 2007

MORETTIN, P. A., **Econometria Financeira**, São Paulo: Blucher, 2008

OBSTFELD, M.; ROGOFF, K.; The mirage of fixed exchange rates; NBER Working Papers 5191, National Bureau of Economic Research, Inc., 1995.

OBSTFELD, M.; ROGOFF, K.; The six major puzzles in international macroeconomics: is there a common cause; **NBER Working Papers 7777**, National Bureau of Economic Research, Inc., 2000.

REINHART, C. M., The mirage of floating exchange rates. **American Economic Review** 90, 65-70, 2000.

SEABRA, F. A relação teórica entre incerteza cambial e investimento: os modelos neoclássico e de investimento irreversível. **Política e Planejamento Econômico**, 26(2):183–202, 1996.

SEABRA, F. O modelo de Bandas Cambiais e a Variabilidade da Taxa de Câmbio. **Estudos Econômicos**. São Paulo: v.28, n.2, p.199 - 224, 1998.

SMITH, C. E. Exchange rate variation, commodity price variation and the implications for international trade. **Journal of International Money and Finance**, 18(3):471–491, 1999.