Daniel de Almeida

Assimetrias na volatilidade e nas perturbações nos modelos de volatilidade

Campinas
2013

UNICAMP

UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E COMPUTAÇÃO CIENTÍFICA

Daniel de Almeida

Assimetrias na volatilidade e nas perturbações nos modelos de volatilidade

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Estatística

Orientador(a): Luiz Koodi Hotta

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO DEFENDIDA PELO ALUNO DANIEL DE ALMEIDA E ORIENTADA PELO PROF DR LUIZ KOODI HOTTA

Assinatura do Orientador

CAMPINAS
2013

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Ana Regina Machado - CRB 8/5467

AL64a

Almeida, Daniel de, 1989-
Assimetrias na volatilidade e nas perturbações nos modelos de volatilidade / Daniel de Almeida. – Campinas, SP : [s.n.], 2013.

Orientador: Luiz Koodi Hotta.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Alavancagem (Finanças). 2. Modelo GARCH. 3. Volatilidade (Finanças). I. Hotta, Luiz Koodi,1952-. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Leverage effect and asymmetry of the error distribution in volatility models
Palavras-chave em inglês:
Leverage (Finance)
GARCH model
Volatility (Finance)
Área de concentração: Estatística
Titulação: Mestre em Estatística
Banca examinadora:
Luiz Koodi Hotta [Orientador]
Mauricio Enrique Zevallos Herencia
Ricardo Sandes Ehlers
Data de defesa: 31-07-2013
Programa de Pós-Graduação: Estatística

Pela Banca Examinadora composta pelos Profs. Drs.

Prof(a). Dr(a). LUIZ KOODI HOTTA

Prof(a). Dr(a). MAURICIO ENRIQUE ZEVALLOS HERENCIA

Prof(a). Dr(a). RICARDO SANDES EHLERS

Abstract

The objective of this dissertation is to study volatility models that consider two types of asymmetry usually found in finance series, the skewness of the innovations and the leverage effect. Skewness means that the distribution of losses has a heavier tail than the distribution of gains. The leverage effect stems from the fact that losses have a greater influence on future volatilities than gains. It is considered univariate GARCH models that include both types of asymmetry, separately and jointly, and multivariate GARCH models that allow for leverage effects. The results are presented in two papers. The first one describes the main univariate models that consider these two stylized facts and analyzes, in detail, eight series: the Ibovespa, Nasdaq and S&P 500 indices, and the Itaú-Unibanco, Vale, Petrobras, Banco do Brasil and Bradesco stocks. The conclusion is that both stylized facts are present in some series, mostly simultaneously. The second paper reviews the main multivariate GARCH models, including models with asymmetric effects on conditional variances and covariance. Some of these models are analyzed in more detail through simulations. The most used models in the literature are applied to a three-dimensional time series, containing the Bovespa index and the Petrobras and Vale markets. The three models selected by AIC and BIC criteria allow for leverage effects.

Key words: leverage effect, asymmetries in Garch family models, volatility models

Resumo

O objetivo da dissertação é estudar modelos de volatilidade que consideram dois tipos de assimetria usualmente encontradas em séries de finanças, a assimetria das perturbações e o efeito de alavancagem. Perturbações assimétricas são utilizadas devido ao fato estilizado de que perdas têm distribuição com cauda mais pesada do que ganhos. Já o efeito de alavancagem leva em consideração que perdas têm maior influência na volatilidade do que os ganhos. São estudados os modelos GARCH univariados que contemplam os dois tipos de assimetria separadamente e conjuntamente e modelos GARCH multivariados que permitem o efeito de alavancagem. Os resultados são apresentados em dois artigos. O primeiro descreve os principais modelos univariados que possam explicar estes dois fatos estilizados e analisa, com detalhes, oito séries: os índices Ibovespa, Merval e S&P 500, e as ações Itaú-Unibanco, Vale, Petrobras, Banco do Brasil e do Bradesco. A conclusão é que os dois tipos de assimetria estão presentes nas séries, na maioria das vezes simultaneamente. O segundo artigo faz uma revisão dos principais modelos multivariados da família GARCH, incluindo modelos com efeitos assimétricos nas variâncias e nas covariâncias condicionais. Alguns destes modelos são analisados com mais detalhes através de simulações. Considerou-se as perdas de eficiência na estimativa da matriz de volatilidade ao se ter erros de especificação, isto é ajustar um determinado modelo a séries geradas por outros modelos. Os modelo mais utilizados na literatura são aplicados a uma série trivariada, contendo o índice Ibovespa e as ações Petrobras e Vale. Os três modelos selecionados pelos critérios AIC e BIC, possuem o efeito de alavancagem.

Palavras Chave: efeito de alavancagem, assimetrias em modelos da família GARCH, modelos de volatilidade.

Sumário

Agradecimentos

À FAPESP, processos 2011/02881-9 e 2012/09597-7, e à CAPES pelo apoio financeiro recebido durante o curso de mestrado.

Aos meus amigos do programa de Mestrado em Estatística, em especial, ao Carlos Mazza por ter me ajudado nos programas de simulação do modelo DCC multivariado.

À banca examinadora, composta pelo Prof. Dr. Mauricio Enrique Zevallos Herencia e pelo Prof. Dr. Ricardo Sandes Ehlers, e à Prof. Dra. Esther Ruiz pelas correções, sugestões e pela disponibilidade.

À Lívia pela ajuda nas atividades burocráticas da pós-graduação.

Finalmente, ao meu orientador Prof. Dr. Luiz Koodi Hotta, por ter me apoiado do começo ao fim do meu mestrado, não só na elaboração da dissertação, mas também em dúvidas de outras disciplinas e de finanças. Ele também me ajudou a conseguir um estágio de pesquisa de cinco meses em Madri, sempre me mandava mensagens me cobrando para fazer o projeto, me incentivou a assistir seminários de outras áreas e me apoiou a enviar projetos para congressos no país.

Lista de Tabelas

Lista de Figuras

Capítulo 1

Introdução

Em geral, séries de retornos financeiros são não correlacionadas serialmente ou possuem baixa correlação. No entanto, as autocorrelações da série dos quadrados (ou dos módulos) dos retornos são pequenas e têm queda lenta em função da defasagem. Assim, processos lineares como aqueles pertencentes à família ARMA, autorregressivos e de médias móveis, geralmente não são apropriados para descrever estas séries. Nos casos em que há alguma pequena autocorrelação serial, os modelos da classe ARMA podem ser inicialmente utilizados, a fim de remover esta dependência, antes de utilizar um modelo de volatilidade adequado.

Para a modelagem de dados financeiros, primeiramente é necessário definir o conceito de retorno. Este é frequentemente utilizado, ao invés dos preços, pelo fato de ser livre de escalas e de possuir propriedades usualmente necessárias para a modelagem da série, como estacionariedade e ergodicidade. Seja P_t o preço de um ativo no instante t. Supondo que não haja dividendos pagos no período, a variação de preços entre os instantes $t-1$ e t é dada por $\Delta P_t = P_t - P_{t-1}$, e a variação relativa de preços, ou retorno líquido simples, deste ativo, entre os mesmos instantes, é definido por:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{\Delta P_t}{P_{t-1}}. \tag{1.0.1}$$

Denotando $p_t = logP_t$, é possível escrever o retorno composto continuamente, ou log-retorno, como:

$$r_t = \log \frac{P_t}{P_{t-1}} = \log(1 + R_t) = p_t - p_{t-1}. \tag{1.0.2}$$

Esta definição geralmente é chamada simplesmente de retorno, pois é a mais utilizada na prática. Uma característica marcante em séries de retornos financeiros é a presença de conglomerados de volatilidade. Aqui, volatilidade é considerada como sendo a variância condicional da série, embora alguns textos definam volatilidade como o desvio padrão condicional. Uma das formas de se considerar a presença destes grupos de volatilidade neste tipo de série é através dos modelos heterocedásticos condicionais, cuja volatilidade de um retorno num dado instante, depende de retornos passados, além de outras informações disponíveis até o tempo considerado.

Atualmente os modelos mais utilizados para ajustar séries que apresentam heterocedasticidade condicional são os modelos ARMA-GARCH e suas várias extensões. Eles são importantes para descrever e fazer previsões do nível e volatilidade de séries de retornos financeiros. Entretanto, os modelos autorregressivos com heteroscedasticidade condicional generalizados (GARCH) tradicionais não conseguem explicar alguns fatos estilizados encontrados em séries financeiras. Dois destes fatos são a assimetria na distribuição das perturbações e o efeito de alavancagem. O primeiro consiste de que perdas têm uma distribuição com uma cauda mais pesada do que ganhos. Simkowitz e Beedles (1980), Kon (1984), dentre outros chamaram a atenção para estas assimetrias. French et al. (1987) encontrou um coeficiente condicional de assimetria significativamente diferente de zero nos resíduos padronizados quando os modelos da família ARCH foram ajustados para os retornos diários do índice *Standard & Poor 500* (S&P). O efeito de alavancagem, introduzido primeiramente por Black (1976), leva em consideração que perdas têm maior influência sobre volatilidades futuras do que ganhos de mesma magnitude. No entanto, nenhum estudo foi ainda testado para a presença simultânea destes dois efeitos, especialmente em séries relacionadas ao mercado brasileiro. A ênfase desta dissertação é estudar a presença simultânea destes dois fatos estilizados.

As volatilidades de diferentes mercados ou ativos dependem entre si ao longo do tempo. Assim, em várias situações, é preciso analisar várias séries conjuntamente, sendo natural utilizar uma abordagem multivariada. Tendo em vista este aspecto, nos últimos anos tem sido presenciado um grande avanço na modelagem multivariada de modelos de volatilidade, o que do ponto de vista financeiro, abriu um caminho para melhores decisões em várias áreas. O modelo GARCH multivariado (MGARCH) é atualmente um dos mais utilizados para explicar relações entre volatilidades e covolatilidades (covariâncias condicionais) de vários ativos. Por exemplo, quando é preciso calcular o valor em risco (VaR) considerando uma carteira de ações, a distribuição do retorno da carteira pode ser obtida diretamente utilizando a distribuição multivariada dos retornos, o que geralmente é mais eficiente do que ajustar modelos univariados para ação e encontrar a distribuição por uma média ponderada dos retornos. Outras aplicações onde é recomendado utilizar modelos de volatilidade multivariados são em precificação de ativos, gerenciamento de riscos e alocação de ativos; para referências, ver, por exemplo, Bollerslev et al. (1988), De Santis e Gerard (1997), Lien e Tse (2002) e Santos et at. (2013).

Os modelos GAHCH multivariados são atualmente um tema muito analisado na literatura econométrica financeira e têm recebido um crescente interesse nos últimos anos para representar a matriz de volatilidade variando no tempo; ver Bauwens et al. (2006) e Silvennoinen e Teräsvirta (2009b) para duas revisões sobre modelos estes modelos. A maioria dos modelos MGARCH propostos na literatura são baseados em respostas simétricas das variâncias e covariâncias condicionais em relação aos quadrados de retornos passados e aos produtos cruzados de retornos. Entretanto, há uma evidência empírica de assimetrias do efeito entre retornos positivos e negativos nas volatilidades e nas correlações entre as volatilidades. O fato estilizado mais conhecido é que retornos simultâneos negativos têm maior efeito nas volatilidades futuras do que retornos simultâneos positivos, existindo poucos estudos em relação ao efeito dos sinais dos retornos nas correlações.

Duas das principais dificuldades para construir um modelo GARCH multivariado alternativo são: garantir que a matriz de covariância condicional seja semidefinida positiva e a questão da dimensionalidade, inviabilizando adotar modelos complexos na análise de grandes sistemas. A maioria deste modelos tenta lidar com as duas questões utilizando restrições. Um ponto importante é analisar se estas restrições não incapacitam o modelo de explicar a dinâmica do sistema de interesse.

O objetivo da dissertação é estudar modelos de volatilidade que consideram dois tipos de assimetria usualmente encontrados em séries de finanças, a assimetria das perturbações e o efeito de alavancagem. O primeiro tipo de assimetria é utilizado para explicar o fato estilizado de que as perdas têm distribuição com cauda mais pesada do que os ganhos. O segundo tipo de assimetria, efeito de alavancagem, leva em consideração de que perdas têm uma influência maior na volatilidade do que os ganhos. Na dissertação é também investigado se estes fatos estilizados estão presentes em séries empíricas ligadas ao mercado brasileiro. Os principais resultados da dissertação são relatados em dois artigos, apresentados nos apêndices. O primeiro artigo, apresentado no apêndice A, analisa brevemente os modelos da família GARCH univariados, incluindo os dois tipos de assimetria descritos na dissertação, e aplica os modelos propostos a séries de retornos relacionadas ao mercado brasileiro. O segundo artigo, apresentado no Apêndice B, faz uma revisão dos modelos MGARCH simétricos, e os estendem de modo a permitir efeitos de alavancagem na matriz de covariâncias. Neste artigo, é estudada através de simulações a perda de eficiência quando há erro de especificação, isto é, quando é ajustado um modelo diferente do processo gerador de dados. Por exemplo, é possível estimar a perda de eficiência quando se ajusta um modelo de correlação constante, quando na verdade ela é dinâmica (restrição incorreta). Ao mesmo tempo, é verificada a perda de eficiência ao ajustar um modelo mais geral do que o necessário. Por exemplo, ao ajustar um modelo com correlação dinâmica quando, na verdade, ela é constante. Por fim, alguns modelos são aplicados a uma série de retornos trivariada. Como os artigos são escritos em inglês, nos Capítulos 2 e 3 são apresentados resumos estendidos dos artigos apresentados nos Apêndices A e B, respectivamente. Para facilitar o acompanhamento dos artigos, as notações utilizadas nos Capítulos 2 e 3 tem pequenas diferenças, pois elas seguem as notações dos artigos correspondente. No Capítulo 4 encontram-se as conclusões e os trabalhos futuros.

.

Capítulo 2

Modelos de volatilidade Univariados

2.1 Modelos ARMA-GARCH

Os processos GARCH (*generalized autoregressive conditional heteroskedastic*) de Bollerslev (1986) permitem capturar dois importantes fatos estilizados encontrados em série financeiras, os conglomerados de volatilidade e as altas kurtoses , por isso, se tornaram importante na análise deste tipo de séries, quando o objetivo é analizar e prever volatilidades. Nestes modelos, a variância condicional depende dos valores passados da série e das volatilidades nos tempos anteriores. Na prática é muito utilizado o processo autorregressivo e de médias móveis (ARMA) com perturbações autorregressivas condicionalmente heterodcedásticas (GARCH), pois permite filtrar, em conjunto, o nível e os conglomerados de volatilidades de uma série. O processo ARMA(m, n)-GARCH(p, q) é dado por

$$
\begin{aligned}
r_t &= \mu + \sum_{i=1}^{m} \phi_i r_{t-i} + \sum_{j=1}^{n} \theta_j \varepsilon_{t-j} + \varepsilon_t && (2.1.1) \\
\varepsilon_t &= \sigma_t z_t, \quad z_t \sim IID(0,1) && (2.1.2) \\
\sigma_t^2 &= \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, && (2.1.3)
\end{aligned}
$$

sendo μ, ϕ_i, θ_i, α_i e β_i os parâmetros do modelo, com as suposições usuais adotadas nos modelos ARMA e GARCH (Tsay, 2002, Seções 2.6 e 3.5). Note que a série é filtrada no nível pelo modelo ARMA (equação 2.1.1) e na volatilidade pelo modelo GARCH (equações 2.1.2 e 2.1.3).

Por exemplo, em um modelo AR(1)-GARCH(1,1), a média e a volatilidade dadas em (2.1.1) e (2.1.3), respectivamente, são

$$\mu_t = \mu + \phi r_{t-1} \tag{2.1.4}$$

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2, \tag{2.1.5}$$

onde $|\phi| < 1$ e $\omega > 0$. As condições $\alpha, \beta \geq 0$, $\alpha + \beta < 1$ são suficientes para o processo ser estacionário e ter variância finita, assim são geralmente adotadas.

2.2 Efeito de alavancagem

O efeito de alavancagem é motivada pelo conhecido fato da volatilidade ser maior quando há retornos negativos do que quando há positivos. Uma revisão dos principais modelos univariados que permitem este efeito pode ser encontrada em Rodríguez e Ruiz (2012). No modelo EGARCH (*exponencial* GARCH), proposto por Nelson (1991), a equação de volatilidade (2.1.3) é dada por

$$\ln(\sigma_t^2) = \omega + \gamma_E z_{t-1} + \alpha\{|z_{t-1}| - E(|z_{t-1}|)\} + \beta\ln(\sigma_{t-1}^2). \tag{2.2.1}$$

onde $|z_t| - E(|z_t|)$ é uma sequência de variáveis aleatórias i.i.d. com média zero. A positividade das variâncias condicionais é garantida pela especificação das volatilidades em termos de uma transformação logaritmica. Porém os parâmetros devem ser restritos para que o processo seja estacionário e possua curtose finita. Estas restrições variam de acordo com a distribuição das perturbações consideradas; ver Rodríguez e Ruiz (2012) para restrições usualmente adotadas na prática. O parâmetro γ_E controla o grau de assimetria do modelo. Se $\gamma_E < 0$, retornos negativos têm maior impacto em volatilidades futuras do que retornos positivos.

O modelo TGARCH de Zakoïan (1994), um caso particular do modelo ARCH não linear, ajusta o desvio padrão condicional ao invés da volatilidade. O modelo TGARCH(1,1) é dado por

$$\sigma_t = \omega + \alpha|\varepsilon_{t-1}| + \beta\sigma_{t-1} + \gamma_T\varepsilon_{t-1}, \tag{2.2.2}$$

A positividade de σ_t é garantida se $\omega > 0$, $\alpha \geq 0$ e $\gamma_T < \alpha$. Além disto, a restrição $\gamma_T^2 < 1 - \alpha^2 - \beta^2 - 2\alpha\beta E(|z_t|)$ é geralmente adotada para o processo ser estacionário.

O modelo GJR (de Glosten *et. al.*, 1993) especifica a variância condicional como

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma_G\mathbb{I}(z_t < 0)\varepsilon_{t-1}^2, \tag{2.2.3}$$

onde $\mathbb{I}(.)$ é a função indicadora que assume valor 1 quando a desigualdade é satisfeita e 0 caso contário. A volatilidade é positiva se $\omega > 0, \alpha, \beta, \gamma_G \geq 0$. Por outro lado, para o processo ser estacionário, é suficiente que $\gamma_G < 2(1 - \alpha - \beta)$.

2.3 Assimetria nas perturbações

A assimetria nas perturbações z_t começou a ser estudada, pois as perturbações geralmente possuem caudas mais pesadas para retornos negativos. Neste trabalho é considerada a distribuição t-Student assimétrica. Existem várias propostas para incluir a assimetria na distribuição t-Student, sendo que Fernández e Steel (1998) propuseram uma forma de introduzir assimetria em qualquer distribuição contínua simétrica e unimodal $g(.)$, mudando sua escala em cada lado da sua moda. Aplicando este procedimento para a distribuição t, encontrou-se outra densidade t-Student assimétrica, a qual pode ser assumida como sendo os erros z_t do modelo GARCH em (2.1.2). A fim de preservar as especificações do modelo GARCH, Lambert e Laurent (2001) modificaram esta densidade para padronizá-la, isto é, fazer com que tenha média zero e variância unitária.

Seguindo Lambert e Laurent (2001), a variável aleatória z_t é dita ser t-Student assimétrica padronizada, em notação SKST(0,1,ξ, v), com parâmetros $v > 2$ (graus de liberdade) e $\xi > 0$ (parâmetro associado à assimetria), se sua densidade é dada por

$$f(z_t|\xi, v) = \begin{cases} \dfrac{2}{\xi + \frac{1}{\xi}} \; s \; g[\xi(sz_t + m)|v] & \text{se} \quad z_t < -m/s \\ \dfrac{2}{\xi + \frac{1}{\xi}} \; s \; g[(sz_t + m)/\xi|v] & \text{se} \quad z_t \geq -m/s, \end{cases} \tag{2.3.1}$$

onde $g(.|v)$ é uma densidade t-Student simétrica com média zero, variância um e v (> 2) graus de liberdade, denotada por $X \sim ST(0, 1, v)$, e definida por

$$g(x|v) = \frac{\Gamma(\frac{v+1}{2})}{\sqrt{\pi(v-2)}\Gamma(\frac{v}{2})} \left[1 + \frac{x^2}{v-2}\right]^{-(v+1)/2}, \tag{2.3.2}$$

sendo $\Gamma(.)$ a função gama de Euler.

Em (2.3.1), as constantes $m = m(\xi, v)$ e $s = \sqrt{s^2(\xi, v)}$ são respectivamente a média e o desvio padrão da distribuição SKST(m, s^2, ξ, v) de Fernández e Steel (1998), e são definidas como segue

$$m(\xi, v) = \frac{\Gamma(\frac{v+1}{2})\sqrt{v-2}}{\sqrt{\pi}\ \Gamma(\frac{v}{2})} \left(\xi - \frac{1}{\xi}\right), \tag{2.3.3}$$

e

$$s^2(\xi, v) = \left(\xi^2 + \frac{1}{\xi^2} - 1\right) - m^2. \tag{2.3.4}$$

Ehlers (2012) ajusta modelos GARCH com perturbações seguindo esta distribuição e propõe uma abordagem Bayesiana para estimação dos parâmetros.

2.4 Aplicações

O objetivo das aplicações é verificar se o efeito de alavancagem e a assimetria nas perturbações estão presentes em séries reais, ligadas ao mercado brasileiro. Foram consideradas oito séries, correspondentes aos índices Ibovespa (IBV, Brasil), Merval (Argentina) e Standard & Poor 500 (S&P, EUA) e às ações Itau-Unibanco (Itau), Vale PNA, Petrobras PN (Petro), Banco do Brasil ON (BB) e Bradesco PN (Brad), no período entre 1 de fevereiro de 2000 e 1 de fevereiro de 2011. Modelos da forma AR(1)-GARCH(1,1), AR(1)-TARCH(1,1), AR(1)-EGARCH(1,1) e AR(1)-GJR-GARCH(1,1) com perturbações t-Student padronizada e t-Student assimétrica padronizada foram analisados. Estes modelos são comparados pelos critérios de informação, pelo teste de significância dos parâmetros de assimetria e, por fim, é utilizado o teste de razão de verossimilhanças (TRV) de Christoffersen, que utiliza a estimação do Valor em Risco (VaR). O objetivo deste teste é verificar se um determinado modelo é eficaz para fazer previsões.

Para cada conjunto de dados foi adotado o seguinte procedimento.

1. Considere as observações dos retornos dos oito primeiros anos e ajuste todos os oito modelos.

2. Verifique qual o modelo é selecionado pelos critérios AIC (Akaike, 1974), BIC (Schwarz, 1978), e HQ (Hannan e Quinn, 1979).

3. Para cada modelo, avalie o VaR de 95% e o de 99% um passo à frente para os

próximos cinco dias. Teste se os retornos estão abaixo dos valores dos VaR estimados. Note que apenas é feito o VaR um passo a frente, de modo que incluímos uma observação a cada a estimação, sem reestimar os parâmetros.

4. Inclua mais cinco observações e exclua as cinco primeiras observações.

5. Repita os passos (2) a (4), até ao fim do período.

Os resultados das aplicações são apresentdas no Apêndice A. As principais conclusões considerando os critérios de informação são:

- O modelo GARCH nunca é selecionado para o IBV, a S&P, o Itaú, a Petro e o Brad. Ele é selecionado algumas vezes para a série de outros. Portanto, há uma clara preferência do critério de informação para os modelos com efeito de alavancagem.

- O modelo GJR é o modelo mais vezes selecionado pelos três critérios de informação.

- Para as séries IBV, a S&P, Merval, Itaú e Brad, os critérios selecionam modelos com distribuições assimétricas na maioria das vezes. Já para os retornos da Vale, Petrobras e Banco do Brasil, os critérios indicam modelos com simetria nas distribuições.

Ao ajustar o modelo GJR-GARCH com distribuição t-Student assimétrica, temos como casos especiais o modelo sem alavancagem, quando $\gamma_G = 0$ e o modelo com inovações simétricas quando o parâmetro de assimetria (ξ) é igual a 1. São utilizados testes de hipótese para verificar a presença ou ausência desses dois fatos estilizados. O mesmo procedimento é feito com os modelos GARCH, TGARCH e EGARCH. Quando ajustamos modelos GJR, EGARCH e TGARCH, o efeito de alavancagem é selecionado na maioria dos casos, tanto com perturbações simétricas como assimétricas. A simetria dos erros foi rejeitada em todos os casos para o IBV e a S&P, e na maioria dos casos para o Merval, o Itaú, o Bradesco e o Banco do Brasil. Na Vale, hipótese nula de simetria nas perburbações nunca foi rejeitada e na Petro, foi rejeitada na maioria dos casos.

O TRV da avaliação da previsão condicional intervalar de Christoffersen (1998) utiliza o VaR de 95% e o de 99% para testar a precisão dos modelos em fazer previsões. É testada a hipótese nula de que o modelo é eficaz para fazer previsões. Um modelo adequado deve conter um percentual de valores maiores que o VaR de $k\%$ próximo ao valor $k\%$. As principais conclusões deste teste aplicados aos ajustes são:

- Não há diferenças significativas em termos de porcentagem, embora os modelos com distribuições assimétricas sejam, em geral, mais próximos dos valores nominais.

- Para o VaR 99%, em geral, os modelos com perturbações assimétricas passam pelo teste TRV, enquanto os modelos com pertubações simétricas falham para os retornos do IBV e do S&P.

- Para o VaR de 95%, a maioria dos modelos com distribuição de erro assimétrico passam no teste TRV. Quando consideramos inovações simétricas, todos os modelos falham para as séries IBV, Merval, S&P, Vale (exceto para GJR), e Petro.

Considerando os três métodos de comparações, pode-se dizer que a assimetria nas perturbações e o efeito de alavancagem estão presentes na maioria das séries analisadas e que os modelos incorporando estes dois fatos estilizados têm maior precisão para estimar o VaR.

.

Capítulo 3

Modelos de volatilidade multivariados

3.1 Introdução

Considere um processo estocástico multivariado $N \times 1$ $\{\mathbf{r}_t\}$, onde $\mathbb{E}(\mathbf{r}_t) = 0$. Por simplicidade, será apresentado o caso mais simples de especificação, que inclui apenas uma defasagem de retornos, variâncias condicionais e covariâncias condicionais passadas. O modelo original GARCH multivariado (MGARCH), uma generalização direta do GARCH univariado, proposto por Bollerslev et al. (1988) é dado por

$$\mathbf{r}_t = \mathbf{H}_t^{1/2}\eta_t, \tag{3.1.1}$$

$$\text{vech}(\mathbf{H}_t) = \mathbf{c} + \mathbf{A}\ \text{vech}(\mathbf{r}_{t-1}\mathbf{r}'_{t-1}) + \mathbf{B}\ \text{vech}(\mathbf{H}_{t-1}), \tag{3.1.2}$$

onde $\mathbf{H}_t$ é uma matriz positiva definida $N \times N$, representando a covariância condicional de $\mathbf{r}_t$, dada a informação até o tempo $t-1$, $\{\eta_t\}$ é uma sequência independente de vetores de perturbações com média zero e matriz de covariância $\mathbf{I}_N$, onde $\mathbf{I}_N$ é a matriz identidade de ordem k, vech(.) é o operador que empilha as colunas da parte triangular inferior de uma matriz quadrada, $\mathbf{c}$ é um vetor de ordem $N(N+1)/2$ e $\mathbf{A}$ e $\mathbf{B}$ são matrizes de parâmetros de ordens $N(N+1)/2 \times N(N+1)/2$. O modelo VEC é suficientemente flexível para representar respostas simétricas das volatilidades em relação aos quadrados de retornos passados e aos produtos cruzados de retornos. As volatilidades dependem das volatilidades e covolatilidades passadas de todas as séries.

O modelo VEC tem covariância estacionária se o módulo dos autovalores de $\mathbf{A}+\mathbf{B}$ são menores do que 1. Há condições suficientes para a positividade de $\mathbf{H}_t$, mas não há condições necessárias fechadas. As desvantagens deste modelo são o elevado número de parâmetros (por exemplo, para $N=3$, é igual a 78), o que torna sua aplicação impraticável para grandes sistemas, e a dificuldade de garantir a positividade das matrizes de variância condicional $\mathbf{H}_t$. Além disso, a estimativa dos parâmetros envolve cálculos pesados.

Por estes inconvenientes, é preciso encontrar modelos parcimoniosos. Diferentes especificações de $\mathbf{H}_t$ dão origem a diferentes modelos. Incluindo estas especificações, os modelos MGARCH podem ser divididos em quatro categorias não excludentes: modelos de matriz de covariância condicional; modelos de fatores; modelos de correlação condicional; abordagens semiparamétricas e não paramétricas. A seguir, estas quatro especificações são apresentadas com maiores detalhes.

3.2 Especificações com restrições nas matrizes de covariâncias condicionais

Vários modelos têm sido sugeridos na literatura para solucionar os problemas do modelo VEC. Por exemplo, o modelo VEC diagonal (DVEC) de Bollerslev et al. (1988) impõe que as matrizes $\mathbf{A}$ e $\mathbf{B}$ sejam diagonais, de modo que cada elemento $\text{vech}(\mathbf{H}_t)_{ij}$ depende apenas de sua própria defasagem e do valor anterior de $r_{it}r_{jt}$. A estimativa dos parâmetros é mais simples do que no modelo VEC completo, porque o número de parâmetros é menor e cada equação pode ser estimada de forma independente. Por outro lado, o modelo de DVEC, em geral, é restritivo, podendo ser inadequado para representar as transmissões de volatilidade entre ativos.

Engle e Kroner (1995) propuseram uma modificação do modelo VEC, a fim de assegurar que as matrizes de covariância são definidas positivas. O modelo BEKK de Baba-Engle-Kraft-Kroner, em notação BEKK(1,1,K), é definido como

$$\mathbf{H}_t = \mathbf{C} + \sum_{k=1}^{K} \mathbf{A}_k^{*'} r_{t-1} r_{t-1}' \mathbf{A}_k^* + \sum_{k=1}^{K} \mathbf{B}_k^{*'} \; \mathbf{H}_{t-1} \mathbf{B}_k^*, \tag{3.2.1}$$

em que $\mathbf{C}$, $\mathbf{A}_k^*$ e $\mathbf{B}_k^*$ são matrizes de ordem $N \times N$, de modo que $\mathbf{C}$ é positiva definida e simétrica. A positividade de $\mathbf{H}_t$ é garantida pela parametrização simétrica do modelo e a ordem K é a generalidade do modelo.

O modelo BEKK tem covariância estacionária se, e somente se, todos os autovalores de $\sum_{k=1}^{K} \mathbf{A}_k^* \otimes \mathbf{A}_k^* + \sum_{k=1}^{K} \mathbf{B}_k^* \otimes \mathbf{B}_k^*$ são menores do que um, em módulo. Embora este modelo permita uma dependência dinâmica entre as volatilidades da série, os parâmetros de (3.2.1) não têm uma interpretação fácil e, como o modelo VEC, o número de parâmetros cresce rapidamente, a medida em que é aumentada a dimensionalidade (N) e a estimativa do modelo é computacionalmente exigente, devido a inversões de várias matrizes. Ademais, é difícil de obter-se convergência na estimação dos parâmetros por máxima verossimilhança, pois o modelo não é linear nos parâmetros. Se for imposto que $\mathbf{A}_k^*$ e $\mathbf{B}_k^*$ são matrizes diagonais, obtem-se o modelo BEKK diagonal (DBEKK), um caso particular do modelo DVEC, porém mais restrito. O DBEKK garante que as matrizes de volatilidades sejam definidas positivas, o que não ocorre no DVEC.

Outras opções para garantir a positividade de $\mathbf{H}_t$ podem ser encontradas em Kawakatsu (2003) e em Tsay (2002).

3.3 Modelos de fatores

Estes modelos são motivados pela parcimônia, ou seja, o processo $\{\mathbf{r}_t\}$ é suposto ser gerado por um (pequeno) número de fatores heteroscedásticos não observados. Nesta classe estão os modelos ortogonais, de componentes principais, de fatores latentes, dentre outros. Todos são combinações lineares de modelos do tipo GARCH univariados.

Há várias modificações destes modelos na literatura; por exemplo, Lin (1992), Bollerslev e Engle (1993) e Vrontos et al. (2003). No modelo GARCH ortogonal (O-GARCH), a série é suposta ser gerada por uma transformação ortogonal de m ($m \leq N$) processos do tipo GARCH univariados, os fatores não observados. O modelo O-GARCH (1,1,m) de Alexander e Chibumba (1997) é definido como

$$\mathbf{V}^{-1/2}\mathbf{r}_t = \mathbf{u}_t = \mathbf{\Lambda}_m \mathbf{f}_t, \tag{3.3.1}$$

onde $\mathbf{V} = \text{diag}(v_1, v_2, \cdots, v_N)$, sendo v_i a variância amostral de r_{it}. A matriz $\mathbf{\Lambda}_m$ de ordem $N \times m$ é dada por

$$\mathbf{\Lambda}_m = \mathbf{P}_m \text{diag}(l_1^{1/2} \cdots l_m^{1/2}), \tag{3.3.2}$$

onde $l_1 \geq \cdots \geq l_m > 0$ são os m maiores autovalores da matriz de correlação populacional de $\mathbf{u}_t$, e a matriz da transformação linear, $\mathbf{P}_m$, é uma matriz ortogonal de autovetores associados. O vetor $\mathbf{f}_t = (f_{1t}, \cdots, f_{mt})'$ é um processo aleatório satisfazendo a $\mathbb{E}(\mathbf{f}_t|\Omega_{t-1}) = \mathbf{0}$, $\text{Var}(\mathbf{f}_t|\Omega_{t-1}) = \mathbf{\Sigma}_t = \text{diag}(\sigma^2_{f1t}, \cdots, \sigma^2_{fmt})$, onde Ω_t é a informação até o tempo t, e

$$\sigma^2_{fit} = (1 - \alpha_i - \beta_i) + \alpha_i f^2_{i,t-1} + \beta_i \sigma^2_{fi,t-1} \quad i = 1, \cdots, m. \tag{3.3.3}$$

Segue que $\mathbf{H}_t = \mathbf{V}^{1/2}\mathbf{V}_t\mathbf{V}^{1/2}$, com $\mathbf{V}_t = \text{Var}(\mathbf{u}_t|\Omega_{t-1}) = \mathbf{\Lambda}_m\mathbf{\Sigma}_t\mathbf{\Lambda}'_m$. Os parâmetros do modelo são $\mathbf{V}$, $\mathbf{\Lambda}_m$ e os parâmetros dos fatores GARCH, α_i e β_i, $i = 1, \cdots, m$. Na prática, $\mathbf{V}$ e $\mathbf{\Lambda}_m$ são substituídos por seus estimadores de momentos, e m é escolhido por análise de componentes principais aplicada aos resíduos padronizados.

O modelo GO-GARCH(1,1) é definido como em (3.3.1), mas, neste caso, com $m = N$ e $\mathbf{\Lambda}$ uma matriz não singular de parâmetros. A matriz de correlação condicional de $\mathbf{r}_t$ é expressa por

$$\mathbf{R}_t = \mathbf{J}_t^{-1}\mathbf{V}_t\mathbf{J}_r^{-1}, \text{ onde } \mathbf{J}_t = (\mathbf{V}_t \odot \mathbf{I}_m)^{1/2} \text{ e } \mathbf{V}_t = \mathbf{\Lambda}\mathbf{\Sigma}_t\mathbf{\Lambda}'. \tag{3.3.4}$$

Em van der Weide (2002), a matriz $\mathbf{\Lambda}$ é decomposta em valores singulares, $\mathbf{\Lambda} = \mathbf{P}\mathbf{L}^{1/2}\mathbf{U}$, onde a matriz $\mathbf{U}$ é ortogonal, e $\mathbf{P}$ e $\mathbf{L}$ são definidos de modelo análogo ao modelo O-GARCH, isto é, como autovetores e autovalores, respecticamente. O modelo O-GARCH (quando $m = N$) corresponde à escolha particular $\mathbf{U} = \mathbf{I}_N$. De forma mais geral, van der Weide (2012) expressa $\mathbf{U}$ como o produto de $N(N-1)/2$ matrizes de rotação

$$\mathbf{U} = \prod_{i<j} G_{ij}(\delta_{ij}), \quad -\pi \le \delta_{ij} \le \pi, \quad i, j = 1, 2, \cdots, n, \tag{3.3.5}$$

onde $G_{ij}(\delta_{ij})$ faz uma rotação no plano gerado pelos i-ésimo e j-ésimo vetores da base canônica do $\mathbb{R}^N$ sob um ângulo δ_{ij}.

3.4 Modelos de correlação condicional

Nesta categoria, a matriz $\mathbf{H}_t$ é especificada de forma hierárquica: primeiramente, é especificado um modelo univariado para cada variância condicional, a seguir, são estimatadas as covariâncias condicionais, através das séries normalizadas por seus respectivos desvios padrões condicionais. A especificação hierárquica torna o modelo atraente, pois possibilita adotar especificações diferentes para a variância condicional de cada série. Além disso, o processo de estimação pode ser realizado em mais de uma fase, assim, não é difícil estimar os parâmetros do modelo, podendo utilizá-lo em sistemas de alta dimensão.

A matriz $\mathbf{H}_t$ é decomposta da forma

$$\mathbf{H}_t = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t = (\rho_{ijt} \sqrt{h_{iit} h_{jjt}}), \tag{3.4.1}$$

onde $\mathbf{D}_t = diag(h_{11t}^{1/2}, \cdots, h_{NNt}^{1/2})$ é uma matriz diagonal $N \times N$ cujos elementos são os desvios padrões condicionais de cada série, de modo que h_{iit} é ajustada por qualquer modelo do tipo GARCH, e $\mathbf{R}_t = (\rho_{ijt})$ é uma matriz simétrica positiva definida com $\rho_{iit} = 1$, $\forall i$. A matriz $\mathbf{H}_t$ é positiva definida se, e somente se, todas as variâncias condicionais h_{iit}, $i = 1, \cdots, N$, são positivas e $\mathbf{R}_t$ é uma matriz positiva definida.

Diferentes especificações para $\mathbf{R}_t$ resultam em diferentes modelos de correlação condicional. Bollerslev (1990) introduziu o modelo de correlação condicional constante (CCC), assumindo que a matriz de correlação condicional é independente do tempo, isto é, $\mathbf{R}_t = \mathbf{R}_0 = (\rho_{ij})$, sendo $\mathbf{R}_0$ geralmente estimada pela matriz de correlação amostral dos retornos padronizados.

Como a suposição de correlações condicionais constantes não é razoável em muitas situações práticas, Engle (2002) e Tse e Tsui (2002) propuseram, de forma independente, os modelos de correlação condicional dinâmica (DCC), em que a matriz de correlação condicional é variável no tempo. No modelo DCC de Tse e Tsui (2002), DCC$_T$(M), a matriz $\mathbf{R}_t$ é dada por

$$\mathbf{R}_t = (1 - \theta_1 - \theta_2)\mathbf{R} + \theta_1 \mathbf{\Psi}_{t-1} + \theta_2 \mathbf{R}_{t-1}, \tag{3.4.2}$$

onde θ_1 e θ_2 são parâmetros não negativos satisfazendo a desigualdade $\theta_1 + \theta_2 < 1$, $\mathbf{R}$ é uma matriz de parâmetros simétrica, positiva definida e de dimensão $N \times N$, de forma que $\rho_{ii} = 1$ e $\mathbf{\Psi}_{t-1}$ é a matriz de correlação amostral de $\{r_{\tau-M}, \cdots, r_{\tau-1}\}$.

O modelo de DCC de Engle (2002), $\text{DCC}_E(1,1)$, é definido como

$$\mathbf{R}_t = diag(q_{11,t}^{-1/2}, \cdots, q_{NN,t}^{-1/2})\mathbf{Q}_t \; diag(q_{11,t}^{-1/2}, \cdots, q_{NN,t}^{-1/2}), \tag{3.4.3}$$

onde a matriz $N \times N$ $\mathbf{Q}_t = (q_{ij,t})$ é positiva definida e segue a forma

$$\mathbf{Q}_t = (1 - a - b)\bar{\mathbf{Q}} + a\mathbf{u}_{t-1}\mathbf{u}'_{t-1} + b\mathbf{Q}_{t-1}, \tag{3.4.4}$$

sendo $u_{it} = r_{it}/\sqrt{h_{iit}}$ os retornos padronizados, $\bar{\mathbf{Q}} = E(\mathbf{u}_t\mathbf{u}'_t)$ a matriz de variância incondicional de $\mathbf{u}_t$, e a e b parâmetros não negativos tal que $a + b < 1$.

Outros modelos de correlação condicional podem ser encontrados em Silvennoinen e Teräsvirta (2005), Billio et al. (2006), Silvennoinen e Teräsvirta (2009a) dentre outros.

3.5 Abordagens semiparamétricas e não paramétricas

A maioria dos modelos paramétricos possui uma interpretação da estrutura dinâmica da volatilidade e o estimador de quase máxima verossimilhança é consistente. No entanto, deve-se assumir uma estrutura particular no modelo, que pode não ser correta e, assim, ocorrer considerável perda de eficiência. Modelos semiparamétricos e não paramétricos têm a vantagem de não supor uma especificação para modelo. Nos modelos não paramétricos, há a desvantagem de seus parâmetros não terem interpretação simples. Por outro lado, os modelos semiparamétricos combinam as vantagens dos modelos paramétricos (consistência e interpretabilidade) e dos modelos não paramétricos, robustez contra má especificação da distribuição. Para uma revisão completa de modelos semiparamétricos e não paramétricos, ver Franke et al. (2008) e Linton (2008).

3.6 Modelos MGARCH com efeito de alavancagem

Em séries financeiras multivariadas, as variâncias e covariâncias podem reagir de forma diferente a um choque positivo ou a um negativo de mesma magnitute. A seguir, são apresentados alguns modelos assimétricos.

De Goeji e Maquering (2004) generalizam o modelo GJR, através do modelo VECH diagonal assimétrico, ao escrever os elementos da matriz de covariância condicional como

$$h_{ij,t+1} = \gamma_{ij} + \beta_{ij} h_{ij,t} + \alpha_{1ij} r_{i,t} r_{j,t} + \alpha_{2ij} r^-_{i,t} r^-_{j,t} + \alpha_{3ij} r^-_{i,t} r^+_{j,t} + \alpha_{4ij} r^+_{i,t} r^-_{j,t}, \qquad (3.6.1)$$

onde $i, j = 1, \cdots, N,$, $r^+_{i,t} = \max(0, r_{i,t})$ e $r^-_{i,t} = \max(0, -r_{i,t})$. De Goeji e Maquering (2009) generalizaram esse modelo adicionando um efeito de nível na equação (B.3.3).

A generalização do modelo EGARCH multivariado foi utilizada por Braun et al. (1995) para séries bivariada, por Rossi et al. (2009), dentre outros. Aqui é considerado o modelo proposto por Kawakatsu (2006) e definido como

$$\text{vech}(\ln \mathbf{H}_t - \mathbf{C}) = \mathbf{A}\ \eta_{t-1} + \mathbf{F}\ (|\eta_{t-1}| - \mathbb{E}|\eta_{t-1}|) + \mathbf{B}\text{vech}(\ln \mathbf{H}_{t-1} - \mathbf{C}), \qquad (3.6.2)$$

onde o logaritmo de uma matriz é definido como a função inversa da exponencial, dada na equação (B.3.6), $\mathbf{C}$ é uma matriz $N \times N$ simétrica, $\mathbf{A}, \mathbf{F}$ e $\mathbf{B}$ são matrizes de parâmetros, tal que $\mathbf{A}$ e $\mathbf{F}$ são da ordem $N(N+1)/2 \times N$ e $\mathbf{B}$ da ordem $N(N+1)/2 \times N(N+1)/2$. Restrições nos parâmetros para garantir a positividade definida das matrizes de covariância condicional não são necessárias, uma vez que a matriz $\ln\mathbf{H}_t$ não precisa ser positiva definida. A positividade da matriz de covariância $\mathbf{H}_t$ decorre do fato que, para qualquer matriz simétrica $\mathbf{S}$, a matriz exponencial de $\mathbf{S}$ também é positiva definida. A matriz exponencial é dada por

$$\exp(\mathbf{S}) = \sum_{i=0}^{\infty} \frac{\mathbf{S}^i}{i!} \tag{3.6.3}$$

Para a inclusão do efeito de alavancagem nos modelos de fatores, basta impor que os fatores são combinações lineares de modelos do tipo GARCH univariados que permitem efeito de alavancagem, como o modelo EGARCH e o modelo TGARCH. Por exemplo, no modelo O-GARCH (1,1, m) de Alexander e Chibumba (1997), a equação (3.3.3) pode ser substituída por

$$\sigma^2_{fit} = (1 - \alpha_i - \beta_i) + \alpha_i f^2_{i,t-1} + \beta_i \sigma^2_{fi,t-1} + \gamma_i \mathbb{I}(f_{i,t-1} < 0) \quad i = 1, \cdots, m, \tag{3.6.4}$$

que permite efeitos assimétricos na variância condicional. $\mathbb{I}(r_{i,t} < 0)$ é igual a 1 se $r_{i,t} < 0$, e zero caso contrário.

Nos modelos de correlação condicional, devido a especificação hierárquica da matriz $\mathbf{H}_t$, pode-se escolher modelos assimétricos para variância condicional. Na verdade, inicialmente, as variâncias condicionais são especificadas por modelos do tipo GARCH univariado, que são estimados individualmente e de forma independente para cada série. Por exemplo, na primeira série pode-se ajustar um modelo GARCH convencional, na segunda série ajustar um modelo EGARCH, e na última série um modelo GJR univariado, os dois últimos modelos incluem o efeito de alavancagem.

O modelo assimétrico generalizado DCC (AG-DCC) GARCH de Cappiello et al. (2006) é uma generalização de modelos DCC, que impõe uma estrutura BEKK sobre as correlações condicionais e inclui efeitos assimétricos. A matriz Q_t em (3.4.4) é reescrita como

$$\mathbf{Q}_t = (\bar{\mathbf{Q}} - \mathbf{A}'\bar{\mathbf{Q}}\mathbf{A} - \mathbf{B}'\bar{\mathbf{Q}}\mathbf{B} - \mathbf{G}'\bar{\mathbf{N}}\mathbf{G}) + \mathbf{A}'\mathbf{u}_{t-1}\mathbf{u}'_{t-1}\mathbf{A} + \mathbf{B}'\mathbf{Q}_{t-1}\mathbf{B} \tag{3.6.5}$$

$$+\mathbf{G}'\mathbf{u}^{-}_{t-1}\mathbf{u}'^{-}_{t-1}\mathbf{G},$$

onde A, B, G são matrizes $N \times N$ de parâmetros, $\mathbf{u}_t^- = \mathbb{I}_{\{\mathbf{u}_t<0\}} \odot \mathbf{u}_t$, $\bar{\mathbf{Q}}$ e $\bar{\mathbf{N}}$ são as matrizes de covariância incondicional de $\mathbf{u}_t$ e $\mathbf{u}_t^-$, respectivamente. O número de parâmetros que regem as correlações neste modelo aumenta rapidamente com a dimensão da série. O modelo DCC assimétrico (ADCC) é um caso especial do modelo AG-DCC, obtido quando as matrizes $\mathbf{A}$, $\mathbf{B}$, e $\mathbf{G}$ são substituídas por escalares. Da forma análoga, o modelo DCC generalizado (G-DCC) é um caso especial do AG-DCC, quando $\mathbf{G} = \mathbf{0}$.

3.7 Simulações

O estudo de simulação compara os modelos BEKK, BEKK diagonal (DBEKK), GO-GARCH, CCC-GARCH, DCC-GARCH e ADCC-GJR, em um caso bivariado e considerando inovações gaussianas. Somente o modelo ADCC-GJR permite o efeito de assimetria. O objetivo da simulação é investigar a capacidade dos critérios AIC e BIC de identificar o modelo correto, e estimar a perda de eficiência quando o modelo utilizado para ajustar os dados é mal especificado ou tem parâmetros desnecessários. Uma má especificação ocorre quando os dados são gerados por um processo diferente do modelo utilizado na estimação. Também é verificada a perda de eficiência ao se ajustar um modelo com mais parâmetros que o necessário, por exemplo, estimar os dados por um modelo ADCC-GARCH quando na verdade não há assimetrias na matriz de volatilidades, sendo o modelo DCC-GARCH o correto. Os resultados são baseados em séries de tamanho 1000, de modo que, para cada modelo, foram geradas 1000 séries. Os modelos BEKK, DBEKK e GO-GARCH foram estimados por máxima verossimilhança, já os demais foram estimados em dois passos, por quase máxima verossimilhança. Os resultados quando o modelo BEKK é ajustado não são apresentados devido a problemas de convergência na estimação.

A bondade do ajuste é medida através do vício, erro quadrático médio (MSE), erro absoluto médio (MAE), média dos quadrados dos erros relativos (MSRE) e média dos valores absolutos dos relativos, todas estas estatísticas calculadas em relação à matriz de volatilidades, definidos como

$$\begin{aligned}
\text{bias}(i,j) &: \sum_{t=101}^{1000} (\hat{h}_{ijt} - h_{ijt})/900 && (3.7.1)\\
\text{MSE}(i,j) &: \sum_{t=101}^{1000} (\hat{h}_{ijt} - h_{ijt})^2/900 && (3.7.2)\\
\text{MAE}(i,j) &: \sum_{t=101}^{1000} |\hat{h}_{ijt} - h_{ijt}|/900 && (3.7.3)\\
\text{MSRE}(i,j) &: \sum_{t=101}^{1000} [(\hat{h}_{ijt} - h_{ijt})/h_{ijt}]^2/900 && (3.7.4)\\
\text{MARE}(i,j) &: \sum_{t=101}^{1000} |(\hat{h}_{ijt} - h_{ijt})/h_{ijt}|/900, && (3.7.5)\\
& && (3.7.6)
\end{aligned}$$

onde $\hat{h}_{ijt}$, é a estimativa da volatilidade verdadeira h_{ijt}, $i = 1, 2,\ j = 1, 2$. Os erros são calculados após o centésimo valor para evitar influência dos valores iniciais. Seja (A) o processo gerador de dados correto, e suponha que o objetivo é avaliar a eficiência do modelo (B) na estimativa h_{ij}. Neste projeto, considera-se a eficiência do modelo incorreto (B), em relação ao modelo correto (A), em termos de uma medida de bondade de ajuste, digamos MSE(i,j) , como sendo a razão entre os MSE(i,j) das estimativas dos modelos

(A) e (B). Ao se ajustar os modelos CCC-GARCH e DCC-GARCH, as estimativas das volatilidades h_{11} e h_{22} são iguais porque as especificações das marginais são as mesmas e nos dois casos utiliza-se o método de estimação por dois estágios. As principais conclusões foram:

- Em geral, o menor vício não ocorre quando se ajusta o modelo correto. Em muitos casos, o menor vício ocorre quando é ajustado o modelo GO-GARCH. Como em todos os casos o vício, em termos práticos, é muito pequeno, todos os comentários a seguir são relativos às outras medidas de bondade de ajuste definidas anteriormente.

- A bondade do ajuste do GO-GARCH é sempre muito ruim, qualquer que seja o modelo correto, indicando que ele não é adequado quando os outros modelos são adequados. Desta forma, ele não será considerado nos próximos comentários.

- Nas simulações dos modelos DBEKK, CCC, DCC e ADCC, os melhores resultados ocorrem quando é ajustado o modelo correto, exceto para alguns casos, nas estimativas das volatilidades h_{11} e h_{22} pelo modelo DBEKK. Isto ocorre em todos os critérios quando o modelo correto é o DCC. Mesmo neste caso o ganho de eficiência não é muito grande. Este ganho de eficiência pode ser explicado pelo fato de que as estimativas nos modelos CCC e DCC são realizadas em dois passos, enquanto a estimativa do BEKK é por máxima verossimilhança, ou seja, ela aproveita as informações dadas pela outra série.

- Nas simulações de modelos BEKK, os modelos CCC, DCC e DBEKK, têm eficiências muito próximas nas estimativas de h_{11} e h_{22}, mas em todos os casos as eficiências dos modelos CCC e DCC foram melhores. O modelo ADCC tem os valores das medidas de (perda de) bondade de ajuste aumentadas em torno do intervalo (10%, 30%), provavelmente devido à generalização desnecessária para permitir o efeito da assimetria. Na estimativa da covolatilidade h_{12}, o modelo DCC é o melhor em todos os critérios, exceto, pelo critério MSRE, quando o DBEKK é melhor. A eficiência do modelo ADCC é muito próxima ao do modelo DCC. O desempenho do modelo CCC é muito ruim.

- Quando o modelo verdadeiro é o CCC, as estimativas de todos os modelos, exceto o DCC-GARCH, perdem em eficiência. As eficiências dos modelos DBEKK e ADCC-GARCH, em relação às estimativas utilizando o verdadeiro modelo, giram em torno de 80%. Isto indica que, exceto pelo GO-GARCH, os outros modelos podem aproximar os modelos CCC, mas perdem em eficiência ao introduzir parâmetros supérfluos. A exceção é a estimativa de h_{11} e h_{22} pelo modelo DCC, que é igual, já que este modelo foi ajustado pelo método de dois estágios.

- Quando o modelo verdadeiro é o DCC, como comentado anteriormente, o modelo DBEKK é o mais eficiente na estimação de h_{11} e h_{22}. A eficiência do modelo ADCC, que tem os parâmetros de assimetria desnecessários, tem uma eficiência em torno de 80%. A eficiência do modelo DBEKK para estimar h_{12} também é em torno de 80. Já, o modelo CCC tem uma eficiência muito baixa, como esperado, porque considera a correlação constante.

- Quando o modelo correto é o ADCC-GJR, todos os modelos perdem mais de 50% para todas as medidas de bondade de ajuste. Isto indica que, para os valores dos parâmetros utilizados, nenhum dos outros modelos consegue aproximar o modelo

ADCC-GJR.

3.8 Aplicações

Na aplicação são analisados os retornos da série trivariada, contendo o índice Ibovespa e as ações Petrobras e Vale, de 15 de abril de 2002 a 12 de abril de 2013, num total de 2725 retornos. As datas com falta de informação foram ignoradas. Foram ajustados 13 modelos: DBEKK, CCC - GARCH, DCC - GARCH, O-GARCH, GO-GARCH, CCC-GJR, CCC-TGARCH, DCC-GJR, DCC-TGARCH, DBEKK assimétrico, ADCC-GARCH, ADCC-GJR e ADCC-TGARCH, todos com as ordens mais simples. As principais conclusões foram:

- O melhor modelo segundo os critérios AIC e BIC é o DBEKK assimétrico, que possui efeito de alavancagem em todos os elementos da matriz de volatilidades. Além disso, todos os parâmetros relativos a assimetrias são estatisticamente significativos.

- De acordo com o AIC, o segundo, o terceiro e o quarto modelos selecionados são ADCC-GJR (4), DCC-GJR (2) e DCC-TGARCH (3) e, de acordo com o BIC foram DCC-GJR (3) (em parênteses a ordem do modelo pelo BIC), DCC-TGARCH (4) e ADCC-GJR (2). Nota-se que os quatro modelos selecionados pelo AIC e pelo BIC possuem efeito de alavancagem nas volatilidades, sendo que dois deles possuem assimetria nas covolatilidades também. Todos os parâmetros relacionados à assimetria são significativos a um nível de 5%.

- Os piores modelos, de acordo os dois critérios, são CCC-GARCH, O-GARCH, CCC-GJR e CCC-TGARCH.

- Analisando as estimativas dos principais modelos selecionados pelo AIC e BIC, DBEKK assimétrico, DCC-GJR e ADCC-GJR, notamos que as maiores diferenças estão nos picos e vales e na velocidade de decaimento da volatilidade após um pico. As estimativas das volatilidade e covolatilidades pelo modelo DCC-GJR tem picos e vales mais acentuados e decaimento mais rápido do que as estimativas do modelo BEKK assimétrico.

.

Capítulo 4

Conclusões e trabalhos futuros

No caso univariado foram analisadas oito séries, a fim de testar se dois fatos estilizados estão presentes: a assimetria nas distribuições das perturbações e o efeito alavancagem. Primeiramente, os modelos foram comparados utilizando os critérios de informação AIC, BIC e HQ, e considerando testes de hipóteses. Em ambos os métodos, encontrou-se evidências de que estes dois fatos estilizados estão presentes na maioria das séries analisadas. A fim de avaliar a eficácia dos modelos em fazer previsão, foi utilizado o TRV de Christoffesen, com base nas estimativas do VaR. Verificou-se que os modelos com perturbações assimétricas apresentam resultados melhores em relação à modelos com distribuições simétricas, em termos do teste de razão de verossimilhanças. Dentre os modelos de alavancagem, o que mais foi selecionado pelos critérios de informação e que passou um número razoável de vezes pelo TRV para o cobertura condicional foi o GJR, que também possui a vantagem de ser um caso particular do modelo GARCH, quando $\gamma_G = 0$.

A segunda parte da dissertação faz uma revisão dos modelos GARCH multivariados, com foco nos modelos com efeitos de alavancagem nas variâncias e covariâncias condicionais. Alguns modelos foram simulados e estimados por modelos diferentes, a fim de estimar a perda de eficiência quando há erro de especificação, isto é, quando o processo gerador de dados não é o mesmo do modelo ajustado. Também são ajustados vários modelos a uma série de retornos trivariada, correspondente ao índice Ibovespa e as ações Petrobras PN e Vale PN. Os três modelos selecionados pelos critérios de informação permitem o efeito de alavancagem.

Os próximos estudos incluem a implementação de outros modelos com alavancagem, análise das diferenças nas assimetrias observadas ao lidar com os mercados internacionais, ou ao lidar com grandes carteiras de ativos, e propostas de novos modelos capazes de evitar os efeitos da má especificação, e que sejam flexíveis o suficiente para representar as assimetrias observadas em dados reais. O número de variáveis consideradas nos dois casos (mercado internacional e carteira de ativos) demandará modelos diferentes. Alguns dos pontos a serem considerados:

a. Propriedades e comparação com outros modelos propostos na literatura;

b. Estimação e testes;

c. Ajuste aos dados reais: previsão de volatilidades e correlações; e

d. Preparação de programas computacionais para implementar os itens anteriores.

Referências bibliográficas

[1] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**, 716-722.

[2] Alexander, C. O. e A. M. Chibumba (1997). Multivariate Orthogonal Factor GARCH, Mimeo, University of Sussex, Grã Bretanha.

[3] Ang, A. e J. Chen (2002). Asymmetric correlations of equity portfolios, *Journal of Financial Economics* **76**, 471–508.

[4] Bauwens, L., S. Laurent e J. V. K. Rombouts (2006). Multivariate GARCH models: A survey, *Journal of Applied Econometrics* **21**, 79-109.

[5] Bekaert, G. e G. Wu (2000). Asymmetric volatility and risk in equity markets, *Review of Financial Studies* **13**, 1–42.

[6] Billio, M., M. Caporin e M. Gobbo (2006). Flexible dynamic conditional correlation multivariate GARCH models for asset allocation, *Applied Financial Economics Letters* **2**, 123–130.

[7] Black, R. (1976). Studies in stock price volatility changes, *Proceedings of the 1976 Meeting of the American Statistical Association, Business and Economics Statistics Section*, 177–181.

[8] Bollerslev T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **31**: 307-327.

[9] Bollerslev, T. (1990). Modeling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model, *Review of Economics and Statistics* **72**, 498–505.

[10] Bollerslev, T. e R.F. Engle (1993). Common persistence in conditional variances, *Econometrica* **61**, 167–186.

[11] Bollerslev, T., R. F. Engle e J. M. Wooldridge (1988). A capital asset pricing model with time varying covariances, *Journal of Political Economy* **96**, 116–131.

[12] Braun, P., D. Nelson e A. Sunier (1995). Good news, bad news, volatility, and betas, *Journal of Finance* **50**, 1575-1603.

[13] Cappiello, L., R. F. Engle e K. Sheppard (2006). Asymmetric dynamics in the correlations of global equity and bond returns, *Journal of Financial Econometrics* **4**, 537–572.

[14] Christoffersen, P. (1998). Evaluating interval forecasts, *International Economic Review*, **39**: 841-862.

[15] de Goeij, P.D. e W. Marquering (2004). Modeling the conditional covariance between stock and bond returns: A multivariate GARCH approach, *Journal of Financial Econometrics* **16**, 531–564.

[16] de Goeij, P.D. e W. Marquering (2009). Stock and bond market interactions with

level and asymmetry dynamics: An out-of-sample application, *Journal of Empirical Finance* **2**, 318–329.

[17] De Santis, G. e B. Gerard (1997). International Asset Pricing and Portfolio Diversification with Time-Varying Risk, *The Journal of Finance* **52**, 1881-1912.

[18] Ehlers, R. S. (2012). Computational tools for comparing asymmetric GARCH models via Bayes factors, *Mathematics and Computers in Simulation* **82**, 858–867.

[19] Engle, R. F. (2002). Dynamic conditional correlation - a simple class of multivariate GARCH models, *Journal of Business and Economic Statistics* **20**,339-350.

[20] Engle, R. F. e F. K. Kroner (1995). Multivariate simultaneous generalized ARCH, *Econometric Theory* **11**, 122–150.

[21] Fernández C. e M. Steel (1998). On Bayesian modelling of fat tails and skewness. *Journal of the American Statistical Association* **93**, 359-371.

[22] Franke J., H. Wolfgang e H. Christian (2004). *Introduction to Statistics of Financial Markets*, e-book. http://sfb649.wiwi.hu-berlin.de/fedchomepage/xplore/ebooks/html/sfe/.

[23] French K.G., W. Schwert e R.F. Stambaugh (1987). Expected Stock Returns and Volatility, *Journal of Financial Economics* **19**, 3-29.

[24] Glosten, L.W., R. Jagannathan e D. E. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance* **48**, 1779–1801.

[25] Hannan, E.J. e B.G. Quinn (1979). The determination of the order of an autoregression, *Journal of the Royal Statistical Society B* **41**, 190-195.

[26] Kawakatsu, H. (2003). Cholesky factor GARCH. Mimeo, Quantitative Micro Software, Irvine, CA, EUA.

[27] Kawakatsu, H. (2006). Matrix exponential GARCH, *Journal of Econometrics* **134**, 95–128.

[28] Kon, S. (1982). Models of stock returns, a comparison, *Journal of Finance* **39**, 147–165.

[29] Lambert, P. e S. Laurent (2001). Modelling financial time series using GARCH-type models and a skewed Student density, Mimeo, Université de Liège.

[30] Lien, D. e Y.K. Tse (2002). Some recent developments in futures hedging, *Journal of Economic Surveys* **16**, 357–396.

[31] Lin, W.L. (1992). Alternative estimators for factor GARCH models - a Monte Carlo comparison, *Journal of Applied Econometrics* **7**, 259–279.

[32] Linton, O. B. (2008). Semiparametric and nonparametric ARCH modelling, *Handbook of Financial Times Series*, T.G. Andersen, R.A. Davis, J-P. Kreiss e T. Mikosch (eds.), 157–167. Nova Iorque: Springer.

[33] Nelson, D.B. (1991). Conditional heteroskedasticity in asset returns: A new approach, *Econometrica* **59**, 347–370.

[34] Rodríguez, M. J. e E. Ruiz (2012). Revisiting several popular GARCH models with leverage effect: differences and similarities, *Journal of Financial Econometrics* **10(4)**, 637–668.

[35] Rossi, E. e F. Spazzini (2009). Model and distribution uncertainty in multivariate GARCH estimation: a Monte Carlo analysis, http://mpra.ub.uni-muenchen.de/12260/ MPRA Paper No. 12260.

[36] Santos, A. A. P., E. Ruiz e F. J. Nogales (2013). Comparing univariate and multivariate models to forecast portfolio value-at-risk, *Journal of Financial Econometrics* **11**, 400–441.

[37] Schwarz G.E. (1978). Estimating the dimension of a model. *Annals of Statistics* **6**(2), 461–464.

[38] Silvennoinen, A. e T. Teräsvirta (2009a). Modeling multivariate autoregressive conditional heteroskedasticity with the double smooth transition conditional correlation GARCH model, *Journal of Financial Econometrics* **7**, 373–411.

[39] Silvennoinen, A. e T. Teräsvirta (2009b). Multivariate GARCH models. Em: *Handbook of Financial Times Series*, T.G. Andersen, R.A. Davis, J-P. Kreiss e T. Mikosch (eds.), 201–232. New York: Springer.

[40] Simkowitz M e W. Beedles (1980). Asymmetric stable distributed security returns, *Journal of the American Statistical Association* **75**, 306−312.

[41] Tsay, R. S. (2002). *Analysis of Financial Time Series*. Nova Iorque: John Wiley.

[42] Tse, Y.K. e A. K. C. Tsui (2002). A multivariate GARCH model with time-varying correlations, *Journal of Business and Economic Statistics* **20**, 351–362.

[43] van der Weide, R. (2002). GO-GARCH: a multivariate generalized orthogonal GARCH model, *Journal of Applied Econometrics* **17**: 549–564.

[44] Vrontos, I.D., P. Dellaportas, D.N. Politis (2003). A full-factor multivariate GARCH model, *Econometrics Journal* **6**, 311–333.

[45] Zakoïan, J. M. (1994). Threshold heteroskedastic models, *Journal of Economic Dynamics and control* **18**, 931–955.

Apêndice A

The leverage effect and the asymmetry of the error distribution in Garch-based models: the case of Brazilian market related series

Leverage effect and the asymmetry of the error distribution in Garch-based models: the case of Brazilian market related series

Daniel de Almeida and Luiz Koodi Hotta [1]

University of Campinas

ABSTRACT. Traditional GARCH models fail to explain at least two of the stylized facts found in financial series: the asymmetry of the distribution of errors and the leverage effect. The leverage effect stems from the fact that losses have a greater influence on future volatilities than do gains. Asymmetry means that the distribution of losses has a heavier tail than the distribution of gains. We test whether these features are present in some series related to the Brazilian market. To test for the presence of these features, the series were fitted by GARCH(1,1), TGARCH(1,1), EGARCH(1,1), and GJR-GARCH(1,1) models with standardized Student t distribution errors with and without asymmetry. Information criteria and statistical tests of the significance of the symmetry and leverage parameters are used to compare the models. The estimates of the VaR (value-at-risk) are also used in the comparison. The conclusion is that both stylized facts are present in some series, mostly simultaneously.

Keywords: asymmetry in volatility models, asymmetric Garch family models, VaR (Value-at-Risk).

[1] Corresponding author: Rua Sérgio Buarque de Holanda, 651; Distr. Barão Geraldo; Campinas, São Paulo; Brazil; CEP 13083-859 ; T: 55-19-35216081 and 55-19-93101120
E-mails: dani.d.almeida89@gmail.com; hotta@ime.unicamp.br

A.1 Introduction

Two important features usually found in time series of asset returns are the presence of volatility clustering and the high kurtosis. Here volatility is considered as the conditional variance, although some authors define it as the conditional standard deviation. The most popular model used in the literature to explain these two stylized fats is the generalized autoregressive conditional heteroskedastic (GARCH) model of Bollerslev (1986) with symmetric errors (normal or Student's t distributions). However, these traditional GARCH models can not explain some stylized facts found in financial time series. Two important unexplained facts are the skewness, or asymmetry, in the distribution of the errors and the leverage effect. The former consists of losses having a distribution with a heavier tail than gains. Simkowitz and Beedles (1980) , Kon (1984), among others have drawn attention to the skewness in such distributions. French et al. (1987) found that the conditional asymmetry coefficient significantly differs from zero in the standardized residuals when ARCH family models were fitted to the daily returns of the S&P500 series. The leverage effect, originally introduced by Black (1976), takes into account that losses have a greater influence on future volatility than do the gains. However, no study has tested yet for the simultaneous presence of these two effects, especially for Brazilian related series.

The aim of the present paper is to verify if these stylized facts are present in some market indices related to the Brazilian market and five of the most important stocks traded in the São Paulo Stock Exchange (BOVESPA). The indices considered are the Ibovespa (IBV, Brazil), MERVAL (Argentina), and Standard & Poor 500 (S&P, USA), and the five stocks are Itaú-Unibanco (Itaú), Vale PNA (Vale), Petrobras PN (Petro), Banco do Brasil ON (BB), and Bradesco PN (Brad), in the period from February 1st, 2000 to February 1st, 2011. After filtering the return series with AR(1) models, we fitted the GARCH(1,1), TGARCH(1,1), EGARCH(1,1), and GJR-GARCH(1,1) models with standardized Student t and standardized asymmetric Student t innovations, for a total of eight models.

Three methods are used to compare the models. The first one uses the Akaike information criterion (AIC) (Akaike, 1974), the Bayesian information criterion (BIC) (Schwarz, 1978), and the Hannan and Quin information criterion (HQ) (Hannan & Quinn, 1979), to select the best model. The second method tests the significance of the symmetry and leverage parameters. The third method compares the value at risk (VaR) estimated by the eight models treated. A model is considered adequate if the VaR estimates have the desired properties. Section A.2 presents three GARCH family models which have leverage effect and the asymmetric distribution used to model the error term. Section A.3 presents the methods used to compare these models and Section A.4 presents some applications. Our concluding remarks are in Section A.5.

A.2 ARMA-GARCH Models

Denoting the returns by r_t, this series is first filtered by an ARMA model (A.2.1), yielding residuals ε_t, serially uncorrelated, but not necessarily independent. In (A.2.2)–(A.2.4), the series ε_t is fitted by a conditional volatility model. We can write this class of models as

$$
\begin{aligned}
r_t &= \mu_t + \varepsilon_t && \text{(A.2.1)}\\
\varepsilon_t &= \sigma_t z_t && \text{(A.2.2)}\\
\mu_t &= c(\mu|\Omega_{t-1}) && \text{(A.2.3)}\\
\sigma_t &= h(\mu, \eta|\Omega_{t-1}), && \text{(A.2.4)}
\end{aligned}
$$

where $c(.|\Omega_{t-1})$ and $h(.|\Omega_{t-1})$ are functions of $\Omega_{t-1} = \{r_j, j \leq t-1\}$, and z_t is an independent and identically distributed (i.i.d.) process, independent of Ω_{t-1}, with $\mathrm{E}(z_t) = 0$ and $\mathrm{Var}(z_t) = 1$. In the ARMA-GARCH model, the residuals are modeled by the generalized autoregressive conditional heteroskedasticity (GARCH) models and the shapes of the functions $c(.|\Omega_{t-1})$ and $h(.|\Omega_{t-1})$ are defined by the orders of the ARMA and GARCH models, respectively. Assuming their existence, μ_t and σ_t^2 are the conditional mean and variance of r_t, respectively.

For example, in the AR(1)-GARCH(1,1) model, the mean and the volatility given by (A.2.3) and (A.2.4), respectively, are

$$
\begin{aligned}
\mu_t &= \mu + \phi r_{t-1} && \text{(A.2.5)}\\
\sigma_t^2 &= \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2, && \text{(A.2.6)}
\end{aligned}
$$

with $|\phi| < 1$ and $\omega > 0$. The conditions $\alpha, \beta \geq 0$, $\alpha + \beta < 1$, which are sufficient conditions for the process to be stationary and have finite variance, are usually adopted.

A.2.1 The leverage effect

The leverage effect is caused by the fact that negative returns have a greater influence on future volatility than do positive returns. For a good comparison among several

GARCH models with leverage effect, see Rodríguez and Ruiz (2012). In this paper, we consider three of the most popular models to represent it: the EGARCH, TGARCH, and GJR models.

In the EGARCH model (Nelson, 1991), the conditional volatility is given by

$$\ln(\sigma_t^2) = \omega + \gamma_E z_{t-1} + \alpha\{|z_{t-1}| - E(|z_{t-1}|)\} + \beta \ln(\sigma_{t-1}^2). \tag{A.2.7}$$

Since z_t is an i.i.d. sequence, $|\varepsilon_t| - E(|\varepsilon_t|)$ is also a sequence of i.i.d. random variables with zero mean. γ_E is a real parameter, such that $\gamma_E < 0$ when negative returns have a greater impact on future volatility than positive returns. Due to the volatility specification in terms of the logarithmic transformation, there are no restrictions on the parameters to ensure positive variance. A sufficient condition for stationarity and finite kurtosis is $|\beta| < 1$.

The Threshold GARCH (TGARCH)(see Zakoïan, 1994) is a particular case of a non-linear ARCH model and models the conditional standard deviation instead of the conditional variance. The TGARCH(1,1) is written as

$$\sigma_t = \omega + \alpha|\varepsilon_{t-1}| + \beta\sigma_{t-1} + \gamma_T \varepsilon_{t-1}. \tag{A.2.8}$$

Ding et al. (1993) proved that, in order to guarantee the positivity of σ_t, it is sufficient that $\omega > 0$, $\alpha \geq 0$ and $\gamma_T < \alpha$. Furthermore, the model is stationarity if $\gamma_T^2 < 1 - \alpha^2 - \beta^2 - 2\alpha\beta E(|z_t|)$. For example, if z_t is Gaussian, then $E(|z_t|) = \sqrt{\frac{2}{\pi}}$.

The GJR model of Glosten et al. (1993) specifies the conditional variance by

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma_G \mathbb{I}(z_{t-1} < 0)\varepsilon_{t-1}^2, \tag{A.2.9}$$

where $\mathbb{I}(.)$ is equal to 1 when the inequality is satisfied and 0 otherwise. Hentschel (1995) showed that σ^2 is positive if

$$\omega > 0, \alpha, \beta, \gamma_G \geq 0. \tag{A.2.10}$$

A sufficient condition for stationarity and finite variance is

$$\gamma_G < 2(1 - \alpha - \beta). \tag{A.2.11}$$

A.2.2 Asymmetry in the errors

In practice, it is generally assumed that $z_t \sim N(0,1)$ or $z_t \sim t_v$ standardized, or any distribution that describes the heavy tails of financial time series. For normal errors and GARCH(1,1) (A.2.6), the kurtosis is equal to

$$K = \frac{\mathbb{E}(r_t^4)}{[\mathbb{E}(r_t^2)]^2} = \frac{3[1 - (\alpha + \beta)]}{1 - (\alpha + \beta)^2 - 2\alpha_1^2} > 3, \tag{A.2.12}$$

when the fourth moment is defined, i.e., when the denominator is positive. This shows that even when the error z_t has a standard normal distribution and ε_t follows a GARCH process, the tails of ε_t are heavier than normal. However, in empirical series it is often found that the distribution of the error term z_t has heavier tails than the normal distribution, and is often replaced by the standardized Student's t distribution (see, for example, Bollerslev, 1986).

The standardized Student t distribution with ν ($\nu > 2$) degrees of freedom is given by

$$g(x|\nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{(\nu - 2)\pi}\Gamma(\nu/2)} \left(1 + \frac{x^2}{(\nu - 2)}\right)^{-(\frac{\nu+1}{2})}, \tag{A.2.13}$$

where Γ is the gamma function.

The distribution given in (A.2.13) has skewness coefficient equal to zero and the excess of kurtosis equal to $6/(\nu - 4)$, for $\nu > 4$.

While the high kurtosis of returns is a well established fact, the situation is much more obscure for the symmetry of the distribution of z_t. In this paper, we consider the asymmetric Student's t distribution. There have been several proposals to include asymmetry in the Student's t distribution. Hansen (1994) was the first to use an asymmetric Student's t distribution in modeling financial data. Fernández and Steel (1998) proposed a way of introducing asymmetry into any symmetric and unimodal continuous distribution $g(.)$, changing its scale on each side of the mode. Applying this procedure to the Student t distribution, one obtains an asymmetric Student t density. In order to preserve the specifications of the GARCH model, Lambert and Laurent (2001) modified this density to standardize it, that is, to have zero mean and unit variance.

Following Lambert and Laurent (2001), the random variable z_t is said to follow the standardized asymmetric Student t, denoted by SKST(0,1,ξ, v), with parameters $v > 2$ (the number of degrees of freedom) and $\xi > 0$ (the parameter associated with the skewness), if its density is of the form

$$f(z_t|\xi, v) = \begin{cases} \dfrac{2}{\xi + \frac{1}{\xi}} \; s \; g[\xi(sz_t + m)|v] & \text{if} \quad z_t < -m/s \\ \dfrac{2}{\xi + \frac{1}{\xi}} \; s \; g[(sz_t + m)/\xi|v] & \text{if} \quad z_t \geq -m/s, \end{cases} \tag{A.2.14}$$

where $g(.|v)$ is the density of the standardized symmetric Student's t given by (A.2.13), and the constants $m = m(\xi, v)$ and $s = \sqrt{s^2(\xi, v)}$ are, respectively, the mean and standard deviation of the SKST(m, s^2, ξ, v) distribution and can be expressed by

$$m(\xi, v) = \frac{\Gamma(\frac{v+1}{2})\sqrt{v-2}}{\sqrt{\pi}\;\Gamma(\frac{v}{2})}\left(\xi - \frac{1}{\xi}\right) \tag{A.2.15}$$

and

$$s^2(\xi, v) = \left(\xi^2 + \frac{1}{\xi^2} - 1\right) - m^2, \tag{A.2.16}$$

respectively (Fernández and Steel, 1998). The main advantages of this density are its easy implementation and the clear interpretation of its parameters. Ehlers (2012) modeled GARCH model with the error term errors with this distribution and proposed a fully Bayesian approach to estimate the model.

A.3 Criteria for Comparison of Models

Consider n observations of a volatility process and suppose that we want to verify the presence of the leverage effect and of asymmetry in the perturbations. In order to do this, we use the following eight models: GARCH, TGARCH, EGARCH, and GJR-GARCH with standardized symmetric and asymmetric Student t distributions. In this section, we present the three criteria used to select the most appropriate model.

Information criteria: There are several information criteria suggested in the literature to select a model. In this paper, we consider the AIC, BIC, and HQ criteria. These criteria are the likelihood penalized by different functions of the number of parameters of the model.

Testing hypotheses: By fitting the GJR-GARCH model with asymmetric Student t distribution, for example, we have as special cases a model without leverage when $\gamma_G = 0$ and a model with symmetric innovations when the skewness parameter (ξ) is equal to 1. Thus we can use hypothesis testing to verify the presence or absence of these two stylized facts. We can follow the same procedure with GARCH, TGARCH, and EGARCH models.

The third criterion uses the VaR at the 95% and 99% levels to test the accuracy of the models in making predictions. We use the conditional prediction interval evaluation procedure of Christoffersen (1998). He proposed a likelihood ratio (LR) test to test the null hypothesis that a statistical method (the model) is good for prediction purpose. This test is defined as follows.

A.3.1 The likelihood ratio test for the conditional coverage

The VaR can be viewed as a prediction interval. One of the methods to evaluate prediction interval is the LR test of Christoffersen (1998). In the VaR case, it tests whether the sequence of losses smaller than the VaR comes from a random sample of the Bernoulli distribution with probability equal to the nominal value.

Let $(r_t)_{1\leq t\leq T}$ be the realization of a series of returns of any financial asset and let $[L(p)_{t|t-1}, U(p))_{t|t-1}]$ be the corresponding sequence of interval forecast outside the sample, where $L(p)_{t|t-1}$ and $U(p))_{t|t-1}$ are the lower and upper limits of the forecast intervals at time t, given the information until time $t-1$, at the confidence level p. Set the indicator variable I_t at time t, given information until time $t-1$, as

$$I(t) = \begin{cases} 1, & \text{if} \quad r_t \in [L(p)_{t|t-1}, U(p))_{t|t-1}] \\ 0, & \text{if} \quad r_t \notin [L(p)_{t|t-1}, U(p)_{t|t-1}]. \end{cases} \tag{A.3.1}$$

We say that the sequence of prediction interval, $[L(p)_{t|t-1}, U(p))_{t|t-1}]$, is efficient with respect to the information set at time $t-1$ (Ψ_{t-1}), if $E(I_t|\Psi_{t-1}) = p$, $\forall t$ if it passes the LR test. Christoffersen (1998) showed that testing $E(I_t|\Psi_{t-1}) = p$, for all t, is equivalent to testing if the sequence $(I_t)_{1\leq t\leq T}$ is i.i.d. with a Bernoulli distribution with parameter p , i.e., $I_t \sim$ i.i.d. Ber(p). Therefore, a sequence of prediction intervals, $[L(p)_{t|t-1}, U(p))_{t|t-1}]$, has a correct conditional coverage if $I_t \sim$ Ber(p) i.i.d., $\forall t$.

In the conditional coverage test, the null hypothesis is that $(I_t)_{1\leq t\leq T}$ is independent and $E(I_t|\Phi_{t-1}) = p$. The test statistics is

$$LR_{cc} = -l(p; I_1, ..., I_T) - l(\hat{\pi}_1; I_1, ..., I_T)], \tag{A.3.2}$$

where $l(\theta;; I_1, ..., I_T)$ is the log likelihood function, i.e., $l(p; I_1, ..., I_T) = (n_T)\,\theta + (T - n_T)\,(1-\theta)$ with $n_T = \sum_{i=1}^{T} I_i$, and $\hat{\pi}_1 = n_T/T$. The statistics LR_{cc} has a χ^2_2 distribution under the null hypothesis.

Equation (A.3.2) can be written as the sum of the LR test statistics for the correct unconditional coverage and the LR test statistics for independence (Christoffersen, 1998). Rejecting the null hypothesis implies that the model is not good for prediction purpose.

A.4 Applications

In this section, we analyze the series of returns of IBV, Merval, S&P, Itaú, Vale, Petro, BB, and Brad, from February 1st, 2000 to February 1st, 2011, with a total of 12 years.

For each dataset we adopted the following procedure.

1. Consider the observations of the returns of the first eight years.

2. Fit all eight models.

3. Verify which model is selected by the AIC, BIC and HQ criteria.

4. For each estimated model, evaluate the one-step-ahead 95%-VaR and 99%-VaR for the next five days. Test whether the returns are below the estimated VaR values. Note that we are always doing one-step-ahead estimation of the VaR, but the model is not re-estimated every time we include one observation.

5. Include five more observations and exclude the first five observations.

6. Repeat steps (2) to (5) until the end of the period.

For each series and each model, we fitted around 200 models and estimated around $1,000$ VaR values. The number of models and VaR estimates depend upon each series, because we ignored non-trading days.

Tables 1 and 2 indicate how many times each of the eight models were selected by the AIC, BIC, and HQ criteria. The main conclusions are:

- The GARCH model was never selected by any criterion for the IBV, S&P, Itaú, Petro, or Brad series. For the Merval and Vale series, the GARCH model was only selected by the BIC (60% of the time for the Merval series and 21% for the Vale series); for the BB series the GARCH model was only selected 31% of the time. This means that there is a clear preference of the information criteria for models with the leverage effect.

- For all of the stocks, the GJR was the most selected model by all the criteria. For the Petro and Brad series, it was always the model selected. For the Merval series, the GJR model was always selected by the AIC, in 91% of the cases by the HQ criterion, and 40% by the BIC. The TGARCH was selected most of the time for the IBV series by all criteria, and the EGARCH model was selected most of the time for the S&P series by all criteria.

- For the IBV and S&P series, the criteria selected models with leverage and asymmetric distributions almost all the time. For the Merval, Itaú, and Brad series, the criteria selected models with leverage and asymmetric distributions most of the time. For the Vale, Petro, and BB series, the criteria selected models with leverage and symmetric distributions most of the time.

Tabela A.1: Number of times the model was selected by the AIC, BIC, and HQ criteria. Panels 1–8 correspond to the IBV, Merval, S&P, Itaú, Vale, Petro, BB, and Brad series, respectively. sym., asym. = standardized symmetric and asymmetric Student's t innovations, respectively

Crite-	GARCH		GJR		EGARCH		TGARCH	
rion	sym.	asym.	sym.	asym.	sym.	asym.	sym.	asym.
AIC	0	0	0	30	0	26	0	141
BIC	0	0	0	30	8	18	3	138
HQ	0	0	0	30	0	26	0	141
AIC	0	0	20	176	0	0	0	0
BIC	66	52	37	41	0	0	0	0
HQ	0	17	46	133	0	0	0	0
AIC	0	0	0	13	0	142	0	46
BIC	0	0	0	13	22	120	0	46
HQ	0	0	0	13	6	136	0	46
AIC	0	0	0	92	0	40	0	65
BIC	0	0	25	53	49	0	38	32
HQ	0	0	9	81	10	30	6	61
AIC	0	0	183	0	0	0	14	0
BIC	42	0	141	0	0	0	14	0
HQ	0	0	183	0	0	0	14	0
AIC	0	0	38	159	0	0	0	0
BIC	0	0	197	0	0	0	0	0
HQ	0	0	152	45	0	0	0	0
AIC	0	42	69	86	0	0	0	0
BIC	81	0	116	0	0	0	0	0
HQ	1	58	102	36	0	0	0	0
AIC	0	0	12	185	0	0	0	0
BIC	0	0	111	86	0	0	0	0
HQ	0	0	62	135	0	0	0	0

Tables 3 and 4 present, respectively, the percentage of cases where the asymmetry and leverage parameters were significant at the 5% level. Figure A.1 presents the estimated asymmetry and leverage parameters in the GJR-GARCH asymmetric model for the IBV,

Tabela A.2: Percentage of selection of a model with leverage (GJR, EGARCH, TGARCH) and without leverage (GARCH), and with and without asymmetric innovations. The left side panels 1–4 correspond to the IBV, S&P, Vale, and BB series, respectively. The right side panels 1–4 correspond to the Merval, Itaú, Petro, and Brad series, respectively. sym., asym. = standardized symmetric and asymmetric Student t innovations, respectively

Crite-	Left panel				Right panel			
	Leverage		Innovation		Leverage		Innovation	
rion	without	with	sym.	asym.	without	with	sym.	asym.
AIC	0.00	100.0	0.00	100.0	0.00	100.00	10.20	89.80
BIC	0.00	100.0	5.58	94.42	59.90	40.10	52.55	47.45
HQ	0.00	100.0	0.00	100.0	23.35	76.65	23.47	76.53
AIC	0.00	100.0	0.00	100.0	0.00	100.0	0.00	100.0
BIC	0.00	100.0	10.95	89.05	0.00	100.0	56.85	43.15
HQ	0.00	100.0	2.99	97.01	0.00	100.0	12.69	87.31
AIC	0.00	100.0	100.0	0.00	0.00	100.0	19.29	80.71
BIC	21.32	78.7	100.0	0.0	0.00	100.0	100.0	0.00
HQ	0.00	100.00	100.00	0.00	0.00	100.0	77.16	22.84
AIC	21.32	78.7	35.03	64.97	0.00	100.0	6.09	93.91
BIC	41.12	58.9	100.0	0.00	0.00	100.0	56.35	43.65
HQ	29.95	70.05	52.28	47.72	0.00	100.0	31.47	68.53

Merval, Vale, and BB series, while Figure A.2 presents the results for the Itaú, S&P, Petro, and Brad series. We do not present the equivalent graphs for the other models, since their behavior is very similar to that of the GJR model. Under the null hypothesis of no asymmetry, one has $\xi = 1$; and under the null hypothesis of no leverage effect, $\gamma_G = 0$. We use full symbols to indicate rejection of the null hypotheses at the 5% level. The main conclusions are:

- For the GJR model, the leverage effect was detected in the models with symmetric and asymmetric errors in all cases for the IBV, S&P, Itaú, Petro, and Brad series, and in practically all cases for the Vale series. For the Merval and BB series, the leverage effect was detected in approximately 80% and 67% of the cases, respectively.

The results were similar in the EGARCH and TGARCH models.

- The asymmetry in the errors was detected in all the cases for all models for the IBV and S&P series, and in approximately 75%, 70%, and 50% of the cases for the Merval, Brad, and BB series, respectively. For the Vale series, the null hypothesis was never rejected. For the Itaú and Petro series, the percentage depended on the model. For the Itaú series, the detection of asymmetry varied from 99.5% for the GARCH model to 74.6% for the TGARCH model, while for the Petro series the percentages varied from 34.5% for the GJR model to 9.1% for the EGARCH model.

- From the figures we can observe that there is a certain stability in time and that in most of the cases the leverage effect and the asymmetry are simultaneously significant most of the time.

Tabela A.3: Percentage of times the skewness parameter of the asymmetric Student t distribution were significant at the 5% level.

	IBV	Merval	S&P	Itaú	Vale	Petro	BB	Brad.
GARCH	99.49	77.04	92.54	99.49	0.00	11.17	57.87	72.08
GJR	100.0	79.08	99.50	89.34	0.00	34.52	44.16	68.53
EGARCH	100.0	72.45	99.00	77.66	0.00	9.14	51.78	68.53
TGARCH	100.0	72.96	99.00	74.62	0.00	14.21	49.24	68.53

Tabela A.4: Percentage of times the leverage parameter of the GJR model was significant at the 5% level.

distr.	IBV	Merval	S&P	Itaú	Vale	Petro	BB	Brad.
sym.	100.0	79.70	100.0	100.0	98.98	100.0	68.02	100.0
asym.	100.0	82.74	100.0	100.0	95.94	100.0	66.50	100.0

Table 5 presents the percentage of cases with returns loss smaller than the one-step-ahead 95%-VaR and 99%-VaR. A good model should give a percentage close to the nominal value. It is prefered to have percentage smaller than larger that the nominal values. A good model should also have a large p-value for the LR test. The main conclusions are:

- There is no meaningful difference in terms of percentage, although the models with asymmetric distributions are generally slightly better.

- For the 99%-VaR, the models with asymmetric error distribution, except for S&P series (for EGARCH and TGARCH), pass the LR test, with the smallest p-value equal to 0.15. When we consider the symmetric error distribution, all the models fail for the IBV and S&P series.

- For the 95%-VaR, the models with asymmetric error distribution, except for the Vale, Petro (for GARCH model), and S&P series, pass the LR test at the 5% level. When we consider the symmetric error distribution, all the models fail for the IBV, Merval, S&P, Vale (except for GJR), and Petro series.

Considering the three methods of comparison we can say that the two stylized facts are present in most of the series analyzed, and that models taking into account these two stylized facts improve the estimation of the VaR.

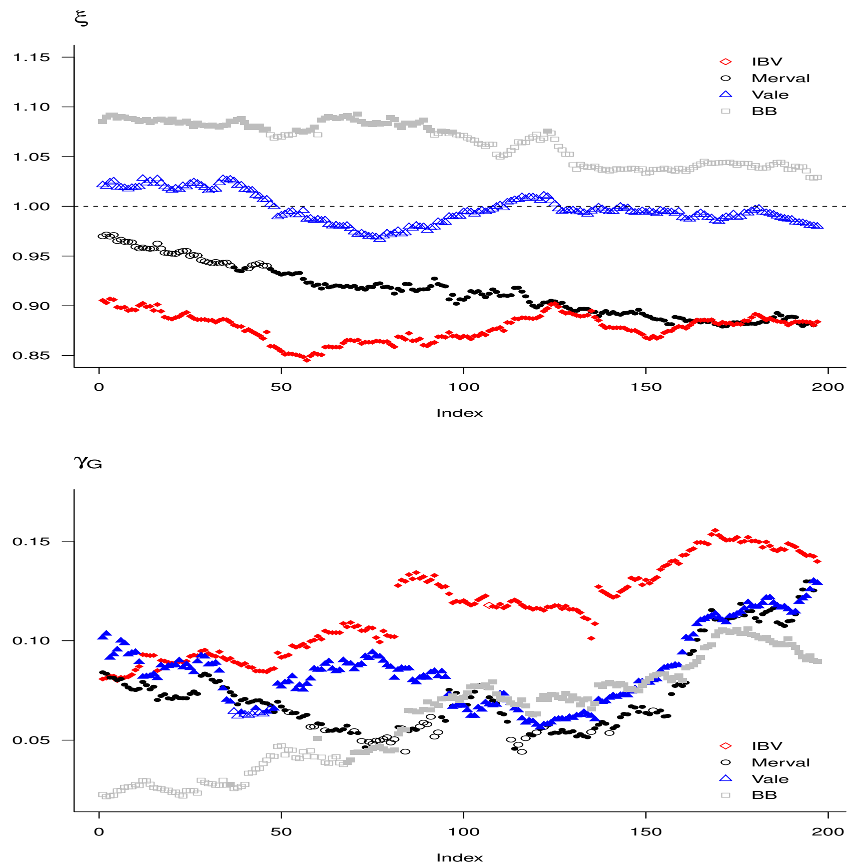

Figura A.1: Estimates of the asymmetry parameter of the error distributions (ξ) and of the leverage parameter (γ_G) of the GJR-GARCH model for the IBV, Merval, Vale, and BB series. Full symbols mean rejection of the null hypothesis at 5%. Under the null hypothesis of no asymmetry, one has ($\xi = 1$); and under the null hypothesis of no leverage effect, ($\gamma_G = 0$).

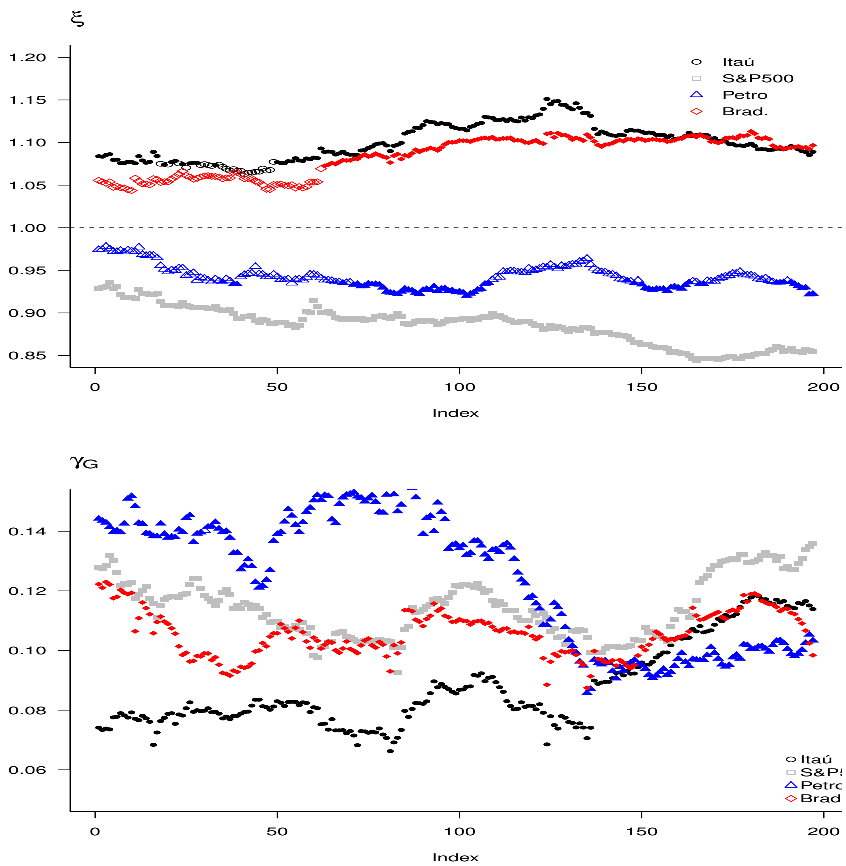

Figura A.2: Estimates of the asymmetry parameter of the error distributions (ξ) and of the leverage parameter (γ_G) of the GJR-GARCH model for the Itaú, S&P, Petro and Brad series. Full symbols mean rejection of the null hypothesis at 5%. Under the null hypothesis of no asymmetry, one has ($\xi = 1$); and under the null hypothesis of no leverage effect, ($\gamma_G = 0$).

Tabela A.5: Percentage of returns loss smaller than the VaR and the p-value of the LR test for the conditional 95%-VaR and 99%-VaR. Panels 1–8 correspond to the IBV, Merval, S&P, Itaú, Vale, Petro, BB, and Brad series, respectively.

	VaR 95%				VaR 99%			
	Percentage		p-value		Percentage		p-value	
Model	sym.	asym.	sym.	asym.	sym.	asym.	sym.	asym.
GARCH	93.10	94.11	0.032	0.447	98.27	98.98	0.086	0.901
GJR	93.10	93.91	0.023	0.293	98.27	98.88	0.086	0.827
EGARCH	92.79	93.81	0.006	0.144	98.07	98.68	0.024	0.529
TGARCH	92.99	94.11	0.015	0.447	98.07	98.98	0.0235	0.901
GARCH	92.96	93.47	0.022	0.118	98.47	98.61	0.245	0.545
GJR	93.06	93.88	0.023	0.292	98.37	98.89	0.191	0.629
EGARCH	92.76	93.67	0.010	0.200	98.16	98.77	0.1463	0.474
TGARCH	92.96	93.67	0.022	0.088	98.27	98.77	0.186	0.474
GARCH	92.84	93.23	0.002	0.025	97.61	98.61	0.001	0.287
GJR	92.94	93.53	0.001	0.032	97.81	98.47	0.002	0.247
EGARCH	91.74	92.74	<0.001	0.012	96.82	97.91	<0.001	0.002
TGARCH	91.84	92.94	<0.001	0.020	96.92	98.31	<0.001	0.022
GARCH	94.42	93.50	0.561	0.097	99.39	98.68	0.399	0.529
GJR	94.92	93.81	0.928	0.229	99.19	98.98	0.776	0.901
EGARCH	94.31	93.50	0.473	0.097	99.29	99.09	0.599	0.886
TGARCH	94.62	93.60	0.737	0.132	99.29	99.19	0.599	0.776
GARCH	93.20	93.20	0.046	0.046	98.48	98.48	0.246	0.246
GJR	93.50	93.50	0.118	0.118	98.38	98.38	0.150	0.150
EGARCH	93.10	93.20	0.023	0.034	98.38	98.38	0.150	0.150
TGARCH	92.89	92.99	0.010	0.015	98.48	98.58	0.246	0.374
GARCH	92.99	93.10	0.022	0.032	98.78	98.78	0.691	0.691
GJR	93.30	93.91	0.026	0.071	98.48	98.68	0.246	0.529
EGARCH	93.10	93.50	0.023	0.097	98.58	98.68	0.374	0.529
TGARCH	93.30	93.50	0.026	0.055	98.68	98.68	0.529	0.529
GARCH	95.63	95.23	0.235	0.080	98.88	98.58	0.827	0.374
GJR	95.74	95.43	0.062	0.149	98.88	98.68	0.827	0.529
EGARCH	95.74	95.33	0.381	0.436	98.88	98.48	0.827	0.246
TGARCH	95.74	95.33	0.381	0.436	98.98	98.58	0.901	0.374
GARCH	95.23	94.42	0.591	0.561	99.29	99.19	0.599	0.776
GJR	95.33	94.72	0.588	0.415	99.09	98.98	0.886	0.901
EGARCH	95.13	94.62	0.084	0.360	99.09	98.88	0.886	0.827
TGARCH	95.13	94.62	0.084	0.360	99.19	98.88	0.776	0.827

A.5 Concluding Remarks

In this paper we analyzed eight series in order to test whether two stylized facts are present: asymmetry in the error distributions and the leverage effect. We first compared the models using the AIC, BIC, and HQ information criteria, and by using hypothesis testing. In both methods, we found evidence that the two stylized facts are present in most of the series analyzed. In the third method, we compared the VaR estimates and found that in VaR estimation, the models with asymmetric errors performed much better than those with symmetric distributions, in terms of the LR test.

Acknowledgements

This work was partially supported by grants from São Paulo Research Foundation (FAPESP) grants 2011/02881-9 and 2008/51097-6, CAPES and CNPq.

References

[1] Akaike-74 Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19, 716-722.

[2] Black Black, R. (1976). Studies in stock price volatility changes. Proceedings of the 1976 Meeting of the American Statistical Association, Business and Economics Statistics Section, 177–181.

[3] Bollerslev-86 Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31, 307-327.

[4] Christoffersen-98 Christoffersen, P. (1998). Evaluating interval forecasts. International Economic Review, 39, 841-862.

[5] Ding Ding, Z., Engle, R. F. & Granger, C.W. J. (1993). A long memory property of stock market return and a new model. Journal of Empirical Finance, 1, 83–106.

[6] Ehlers-12 Ehlers, R.S. (2012). Computational tools for comparing asymmetric GARCH models via Bayes factors. Mathematics and Computers in Simulation, 82, 858–867.

[7] Fernandez-Steel-98 Fernández C. & Steel, M. (1998). On Bayesian modelling of fat tails and skewness. Journal of the American Statistical Association, 93, 359-371.

[8] French-al-87 French, K.G., Schwert, W. & Stambaugh, R.F. (1987). Expected Stock Returns and Volatility. Journal of Financial Economics, 19, 3-29.

[9] GJR-93 Glosten, L.R., Jagannathan, R & Runkle, D.E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. Journal of Finance, 48, 1779-1801.

[10] Kon Kon, S. (1982). Models of stock returns, a comparison. Journal of Finance, 39, 147–165.

[11] Hannan-Quinn-79 Hannan, E.J. & Quinn, B.G. (1979). The determination of the order of an autoregression. Journal of the Royal Statistical Society B, 41, 190-195.

[12] Hansen-94 Hansen, B. (1994). Autoregressive conditional density estimation. International Economic Review, 35, 705–730.

[13] Hentschel-95 Hentschel, L. (1995). All in the family nesting symmetric and asymmetric GARCH models. Journal of Financial Economics, 39, 71-104.

[14] Lambert-Laurent-01 Lambert, P & Laurent, S. (2001). Modelling financial time series using GARCH-type models and a skewed Student density, Mimeo, Université de Liège.

[15] Nelson-91 Nelson, .DB. (1991). Conditional heteroskedasticity in asset returns: A new approach. Econometrica, 59, 347-370.

[16] Rodriguez-Ruiz-12 Rodríguez, M.J. & Ruiz, E. (2012). Revisiting Several Popular GARCH Models with Leverage Effect: Differences and Similarities. Journal of Financial Econometrics, 10, 637-668.

[17] Schwarz-78 Schwarz, G.E. (1978). Estimating the dimension of a model. Annals of Statistics, 6, 461-464

[18] Simkowitz-Beedles-80 Simkowitz, M. & Beedles, W. (1980). Asymmetric stable distributed security returns. Journal of the American Statistical Association, 75, 306−312.

[19] Zakoian Zakoïan, J. M. (1994). Threshold heteroskedastic models. Journal of Economic Dynamics and control, 18, 931–955.

Apêndice B

Leverage effects in multivariate GARCH models

Leverage effects in multivariate GARCH Models

Daniel de Almeida[a], Luiz K. Hotta[a1] and Esther Ruiz[b2]

[a]University of Campinas, [b]University Carlos III of Madrid

ABSTRACT: Multivariate GARCH (MGARCH) models are popular to represent relationships between levels and volatilities of systems of asset returns. The original MGARCH models are symmetric in terms of volatilities and covolatilities. Furthermore, they have to be strongly restricted to reduce the number of parameters so that their estimation is feasible and positive definiteness of the conditional covariance matrices is guaranteed. However, these restrictions limit the dynamics that the models can represent. For example, it is often observed that volatilities of different assets are correlated and often asymmetric in relation to the sign of past returns. Also, many authors observe that correlation between returns may be asymmetric. This paper surveys the main MGARCH models updating previous surveys. The focus is to analyze the limitation of the main MGARCH models often implemented in the literature to represent conditionally heterocedastic systems of returns with asymmetric volatilities, covariances and covolatilities. Simulations are carried out to compare some of the main models and an empirical analysis of a three-dimensional Brazilian time series is presented.

Keywords: asymmetry in multivariate volatility models, asymmetry in covariances, MGARCH.

[1]Corresponding author: Rua Sérgio Buarque de Holanda, 651; Distr. Barão Geraldo; Campinas, São Paulo; Brazil; CEP 13083-859 ; T: 55-19-35216081 and 55-19-93101120
E-mails: dani.d.almeida89@gmail.com; hotta@ime.unicamp.br

[2]Calle de Madrid, 126; Getafe, Madrid; Spain; CEP 28903; T: 34-91-6249851
E-mail: ortega@est-econ.uc3m.es

B.1 Introduction

Univariate generalized autoregressive conditional heteroskedastic (GARCH) models were a big advance in the statistical analysis of financial returns and have often been fitted to describe and predict changes in their volatility. Furthermore, there are many financial applications that require multivariate analysis. Asset pricing, risk management and asset allocation models depend on the covariances between the assets in a portfolio; see, for example, Bollerslev et al. (1988), De Santis and Gerard (1997), Lien and Tse (2002), and Santos et at. (2013), among many others.

Multivariate conditionally heteroskedastic models are now a hot topic in the financial econometrics literature; see Bauwens et al. (2006) and Silvennoinen and Teräsvirta (2009b) for two recent surveys on MGARCH models. Many multivariate GARCH models already available to represent systems of financial returns are based on symmetric responses of conditional variances and covariances to past square returns and cross-products of returns. However, there is a strong empirical evidence of asymmetries. In the context of conditional variances, it is often observed that their responses are larger when returns are negative; see, for example, Bollerslev et al. (2006) for a comprehensive list of references with empirical evidence about the asymmetric response of volatility to past returns. In addition, when looking at conditional covariances, it is common for simultaneous negative returns in different assets to have greater effect on covariances than simultaneous positive returns; see, for example, Bekaert and Wu (2000), Ang and Chen (2002), Capiello et al. (2006) and Yoldas (2012). De Goeij and Marquering (2004) also find cross-asymmetries generated by stocks of different signs. Covariances between stock and bond returns tend to be relatively low after bad news in the stock market and after good news in the bond market.

MGARCH models have also often been restricted so that the number of estimated parameters is reduced to render feasible estimations in large systems. In addition, restrictions are often imposed to ensure positive definiteness of the conditional covariance matrices. Most of these restriction are based on the assumption that volatilities and covariances do not depend on each other so that two step procedures can be implemented. Nevertheless, the empirical evidence on volatility transmissions between assets is plentiful; see, for example, Jeantheau (1998), He and Teräsvirta (2004), Diebold and Yilmaz (2009), Nakatani and Teräsvirta (2009), Diebold and Nerlove (1989) and Engle et al. (1990). In this paper, we survey the main developments of multivariate GARCH models with leverage effect. We have two main objectives. First, we review how the different types of asymmetries described above have been incorporated in the MGARCH models in the literature. Our second objective is to analyze the empirical implications of the restrictions imposed in the multivariate models to reduce dimensionality and guarantee positiveness. The paper is organized as follows. Section 2 describes popular alternative specifications of MGARCH models. The main focus is on the restrictions often imposed to enable estimation, while also guaranteeing positivity. Section 3 discusses how the leverage effect has been incorporated in these models. Section B.4 uses simulation to compare some of the proposed models, by modeling the effect of misspecification in the estimation of the volatility matrix. In Section B.5, we present an application to a three-dimensional time series, corresponding to the following Brazilian series: Ibovespa (São Paulo Stock Exchange Index), Petrobras PN and Vale PNA. Finally, Section B.6 concludes the paper.

B.2 Symmetric Multivariate GARCH models

B.2.1 Introduction

Consider a stochastic $N \times 1$ vector process $\{\mathbf{r}_t\}$, where $\mathbb{E}(\mathbf{r}_t) = 0$. For the sake of simplicity, in this paper we focus on the simplest specification, which only includes one lag of past returns, conditional variances and covariances. The original multivariate GARCH model, a direct generalization of the univariate GARCH, proposed by Bollerslev et al. (1988) is given by

$$\mathbf{r}_t = \mathbf{H}_t^{1/2} \eta_t, \tag{B.2.1}$$

$$\text{vech}(\mathbf{H}_t) = \mathbf{c} + \mathbf{A}\ \text{vech}(\mathbf{r}_{t-1}\mathbf{r}'_{t-1}) + \mathbf{B}\ \text{vech}(\mathbf{H}_{t-1}), \tag{B.2.2}$$

where $\mathbf{H}_t$ is an $N \times N$ positive definite matrix, which represents the conditional covariance of $\mathbf{r}_t$, given the information until time $t-1$, and $\{\eta_t\}$ is an independent sequence of vectors of errors with zero mean and covariance matrix $\mathbf{I}_N$ (identity matrix of order N), vech(.) is the operator that stacks the columns of the lower triangular part of a square matrix, and $\mathbf{c}$ is an $N(N+1)/2$ vector and $\mathbf{A}$ and $\mathbf{B}$ are $N(N+1)/2 \times N(N+1)/2$ parameter matrices. For example, in the bivariate model, $N = 2$, the expression (B.2.2) is given by

$$\begin{aligned}
\begin{pmatrix} h_{11t} \\ h_{12t} \\ h_{22t} \end{pmatrix} &= \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} r^2_{1,t-1} \\ r_{1,t-1} r_{2,t-1} \\ r^2_{2,t-1} \end{pmatrix} + \\
&+ \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \begin{pmatrix} h_{11,t-1} \\ h_{12,t-1} \\ h_{22,t-1} \end{pmatrix}.
\end{aligned}$$

Therefore, the VEC model is very flexible to represent the symmetric response of conditional variances and covariances to past square returns and cross-products of returns. Note that the conditional variances depend on each other and on past covariances. Similarly, the conditional covariance depends not only on past cross-products of returns but also on past conditional variances.

The VEC model is covariance stationary if the moduli of the eigenvalues of $\mathbf{A}+\mathbf{B}$ are less than one. There are sufficient conditions for the positivity of $\mathbf{H}_t$, but there are not closed necessary conditions; see Gourieroux (1997) for the sufficient conditions. Analytical expressions of the fourth-order moments can be found in Hafner (2003).

The disadvantages of this general model are the high number of parameters, $N(N+1)[N(N+1)+1]/2$, making its implementation impracticable for large systems, and the difficulty to ensure the positive definiteness of the conditional variance matrices $\mathbf{H}_t$, since one needs to impose strong restrictions on the parameters in the VEC representation. This is often infeasible in practice. Besides this, the estimation of the parameters involves heavy computations.

The log-likelihood of the model in (B.2.1)-(B.2.2), supposing η_t have a multivariate Gaussian distribution, is given by

$$l(\theta|\mathbf{r}) = c - \frac{1}{2}\sum_{t=1}^{T}\ln|\mathbf{H}_t| - \frac{1}{2}\sum_{t=1}^{T}\mathbf{r}_t'\mathbf{H}_t^{-1}\mathbf{r}_t. \tag{B.2.3}$$

The vector θ contains the model parameters and must be estimated by iterative methods. Note that the matrix $\mathbf{H}_t$ needs to be inverted T times in each iteration, which is why the estimation process is computationally demanding. This takes a very long time when N is large.

As the VEC model is highly parameterized, we have to find parsimonious models. Different specifications of equation $\mathbf{H}_t$ give rise to different models. Including these models, MGARCH models can be divided into four non-exclusive categories:

- Models of the conditional covariance matrix;

- Factor models;
- Conditional correlation models;
- Semiparametric and nonparametric approaches.

Briefly, the first category is related to the early parametric MGARCH models based on directly restricting the parameter in model (B.2.2). The most popular is the BEKK which ensures positive definiteness. The second category models the covariance in a parsimonious way, supposing that the comovements of stock returns are generated by a small number of unobserved factors. In the third one the conditional variances and correlations are modeled rather than direct modeling of the conditional covariance matrix. Examples of these models are constant conditional correlation (CCC) models and their extensions. Finally, the last group contains semiparametric and nonparametric approaches.

The two main difficulties to construct a multivariate GARCH model are basically the curse of dimensionality and the need to guarantee positive definiteness of the conditional variance matrix. Since the number of parameters in MGARCH models increase fast according to the dimensionality of the series, the model specification should be parsimonious enough not to hamper estimating and interpreting the model parameters. On the other hand, parsimonious models might possess few parameters, but could fail to explain the relevant dynamics of the data. The other point that deserves attention is ensuring that the conditional covariance matrices are positive. This problem can be solved in two ways: finding conditions under which the conditional covariance matrices implied by the model are positive definite, or defining a model whose conditional covariance matrices are positive definite by the model structure, in addition to some constraints. In practice the second approach is preferred, since it is more feasible.

B.2.2 Restricted specifications based on conditional covariance matrices

The diagonal VEC (DVEC) model of Bollerslev et al. (1988) imposes diagonality of the A and B matrices, so each element $\text{vech}(\mathbf{H}_t)_{ij}$ depends only on its own lag and on the previous value of $r_{it}r_{jt}$. Therefore, the conditional variances and covariances have a GARCH specification. In the bivariate model, for example,

$$\begin{pmatrix} h_{11t} \\ h_{12t} \\ h_{22t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} + \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix} \begin{pmatrix} r^2_{1,t-1} \\ r_{1,t-1}r_{2,t-1} \\ r^2_{2,t-1} \end{pmatrix} + \\ + \begin{pmatrix} b_{11} & 0 & 0 \\ 0 & b_{22} & 0 \\ 0 & 0 & b_{33} \end{pmatrix} \begin{pmatrix} h_{11,t-1} \\ h_{12,t-1} \\ h_{22,t-1} \end{pmatrix}.$$

Estimation is less difficult than in the complete VEC model, because the number of parameters is smaller and each equation can be estimated independently. In contrast, the DVEC model in general is too restrictive, since it contains $3(N+1)N/2$ parameters, fewer compared with the complete model and it is not suitable to represent volatility transmissions between assets.

Engle and Kroner (1995) proposed a modification of the VEC model in order to ensure that the covariance matrices are positive definite, the BEKK model of Baba-Engle-Kraft-Kroner. The BEKK(1,1,K) is given by

$$\mathbf{H}_t = \mathbf{C} + \sum_{k=1}^{K} \mathbf{A}_k^{*'} \mathbf{r}_{t-1}\mathbf{r}'_{t-1}\mathbf{A}_k^* + \sum_{k=1}^{K} \mathbf{B}_k^{*'}\ \mathbf{H}_{t-1}\mathbf{B}_k^*, \tag{B.2.4}$$

where $\mathbf{C}$, $\mathbf{A}_k^*$ and $\mathbf{B}_k^*$ are square parameter matrices of order N, so that $\mathbf{C}$ is a positive definite symmetric matrix. The positivity of $\mathbf{H}_t$ is guaranteed by symmetric parameterization of the model and the value of k represents the generality of the process. For $N = 2$ and $k = 1$,

$$\begin{bmatrix} h_{11t} & \\ h_{12t} & h_{22t} \end{bmatrix} = \begin{bmatrix} c_{11} & \\ c_{21} & c_{22} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} r_{1,t-1}^2 & \\ r_{1,t-1}r_{2,t-1} & r_{2,t-1}^2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} +$$

$$\begin{bmatrix} b_{11} & b_{21} \\ b_{12} & b_{22} \end{bmatrix} \begin{bmatrix} h_{11,t-1} & \\ h_{12,t-1} & h_{22,t-1} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}.$$

Then,

$$h_{11t} = c_{11} + a_{11}^2 r_{1,t-1}^2 + a_{12}^2 r_{2,t-1}^2 + 2a_{11}a_{12}r_{1,t-1}r_{2,t-1} + b_{11}^2 h_{11,t-1} + b_{12}^2 h_{22,t-1} + 2b_{11}b_{12}h_{12,t-1}$$

$$h_{12t} = c_{12} + a_{11}a_{21}r_{1,t-1}^2 + a_{12}a_{22}r_{2,t-1} + (a_{11}a_{22} + a_{12}a_{21})r_{1,t-1}r_{2,t-1} + b_{11}b_{21}h_{11,t-1}$$

$$+ b_{12}b_{22}h_{22,t-1} + (b_{11}b_{22} + b_{12}b_{21})h_{12,t-1}$$

$$h_{22t} = c_{22} + a_{21}^2 r_{1,t-1}^2 + a_{22}^2 r_{2,t-1}^2 + 2a_{21}a_{22}r_{1,t-1}r_{2,t-1} + b_{21}^2 h_{11,t-1} + b_{22}^2 h_{22,t-1} + 2b_{21}b_{22}h_{12,t-1}.$$

The BEKK model is covariance stationary if and only if the eigenvalues of $\sum_{k=1}^{K} \mathbf{A}_k^* \otimes \mathbf{A}_k^* + \sum_{k=1}^{K} \mathbf{B}_k^* \otimes \mathbf{B}_k^*$ are less than one in modulus. The number of parameters in the BEKK(1,1,1) model is $N(5N+1)/2$. While this model allows for a dynamic dependence between the volatility of the series, the parameters of (B.2.4) do not have an easy interpretation and, like the VEC model, the number of parameters still increases when one increases the number of series (N) and the model estimation by maximum likelihood is computationally demanding due to the need for several matrix inversions. Furthermore, it is difficult to obtain convergence, because the models is not linear in the parameters.

By requiring that $\mathbf{A}_k^*$ and $\mathbf{B}_k^*$ to be diagonal matrices, one obtains the diagonal BEKK (DBEKK) model. This model is also a DVEC model but it is less general, although it is guaranteed to be positive definite while the DVEC is not.

The model of Lin (1992) is a special case of the BEKK model. The BEKK(1,1,K) defined in (B.2.4) is a k factor model if for each $k = 1, 2, \cdots, K$, $\mathbf{A}_k^*$ and $\mathbf{B}_k^*$ have rank one and have the same left and right eigenvectors, i.e., $\mathbf{A}_k^* = \alpha_k \mathbf{w}_k \lambda_k'$ and $\mathbf{B}_k^* = \beta_k \mathbf{w}_k \lambda_k'$, where α_k and β_k are scalar, and λ_k and $\mathbf{w}_k$, $k = 1, 2, \cdots, K$ are $N \times 1$ vectors, such that

$$\mathbf{w}_k' \lambda_i = \begin{cases} 0 & \text{if } k \neq i \\ 1 & \text{if } k = i \end{cases} \tag{B.2.5}$$

$$\sum_{n=1}^{N} w_{kn} = 1. \tag{B.2.6}$$

Besides the BEKK model, other options to ensure the positivity of $\mathbf{H}_t$ are given by Kawakatsu (2003), who proposes the model MGARCH with Cholesky factorization, and by Tsay (2002).

B.2.3 Factor models

Factor models are designed to represent the commonalities in volatility, being simultaneously parsimonious and guaranteeing the positiveness of the conditional covariances. The process $\{\mathbf{r}_t\}$ is supposed to be generated by a (small) number of unobserved heteroskedastic factors, which are linear combinations of univariate GARCH-type models. There are various alternative specifications of these models proposed in the literature; see for example, Diebold and Nerlove (1989), Engle et al. (1990), Bollerslev and Engle (1993) and Vrontos et al. (2003). Sentana (1998) presents the differences between factor GARCH models and BEKK models.

Engle et al. (1990) proposed the first factor GARCH model (F-GARCH), assuming that these factors have a first-order GARCH structure. The F-GARCH(1,1,K) model is given by

$$\mathbf{H}_t = \mathbf{\Omega} + \sum_{k=1}^{K} \mathbf{c}_k \mathbf{c}_k' [\alpha_k (\mathbf{d}_k' \mathbf{r}_{t-1})^2 + \beta_k (\mathbf{d}_k' \mathbf{H}_{t-1} \mathbf{d}_k)] \tag{B.2.7}$$

where $\mathbf{\Omega}$ is an $N \times N$ symmetric positive semi-definite matrix of rank $N_1 \leq N$, $\mathbf{C} = (\mathbf{c}_1, \cdots, \mathbf{c}_k)$ and $\mathbf{D} = (\mathbf{d}_1, \cdots, \mathbf{d}_k)$ are $N \times K$ full column rank coefficient matrices satisfying $\mathbf{D}'\mathbf{C} = \mathbf{I}_K$ and $\sum_{n=1}^{N} d_{nk} = 1 \; \forall k$. If $k = 1$ the number of parameters in this model is $N(N+5)/2+2$. We can see that the time-variation in the conditional covariance matrix can be summarized by a few (K) linear combinations of returns. The F-GARCH model is a special case of the BEKK model, therefore the covariance matrices are positive semi-definite. Bollerslev and Engle (1993) show that $\mathbf{r}_t$ is covariance stationary if and only if $\alpha_k + \beta_k < 1 \; \forall k$. Note that the F-GARCH models differ from the factor models with heteroskedastic disturbances; see Sentana (1998).

The factors in the model of Engle et al. (1990) are generally correlated, which may be undesirable since several of the factors can represent very similar characteristics of the data. Still, if the factors were uncorrelated they would represent entirely different common components driving the returns. For this reason, several factor models with uncorrelated factors have been proposed in the literature.

In the orthogonal GARCH (O-GARCH) model, the series are supposed to be generated by an orthogonal transformation of m ($m \leq N$) univariate GARCH-type processes, the unobserved factors. The O-GARCH(1,1,m) model of Alexander and Chibumba (1997) is defined as

$$\mathbf{V}^{-1/2} \mathbf{r}_t = \mathbf{u}_t = \mathbf{\Lambda}_m \mathbf{f}_t, \tag{B.2.8}$$

where $\mathbf{V} = \text{diag}(v_1, v_2, \cdots, v_N)$, with v_i being the sample variance of r_{it}. The $N \times m$

matrix $\mathbf{\Lambda}_m$ is given by

$$\mathbf{\Lambda}_m = \mathbf{P}_m \text{diag}(l_1^{1/2}, \cdots, l_m^{1/2}), \tag{B.2.9}$$

where $l_1 \geq \cdots \geq l_m > 0$ are the m largest eigenvalues of the population correlation matrix of $\mathbf{u}_t$, and the matrix of the linear transformation, $\mathbf{P}_m$, is an orthogonal matrix of the associated eigenvectors. The vector $\mathbf{f}_t = (f_{1t}, \cdots, f_{mt})'$ is a random process satisfying $\mathbb{E}_{t-1}(\mathbf{f}_t) = \mathbf{0}$, $\text{Var}_{t-1}(\mathbf{f}_t) = \mathbf{\Sigma}_t = \text{diag}(\sigma^2_{f1t}, \cdots, \sigma^2_{fmt})$ and

$$\sigma^2_{fit} = (1 - \alpha_i - \beta_i) + \alpha_i f^2_{i,t-1} + \beta_i \sigma^2_{fi,t-1} \quad i = 1, \cdots, m. \tag{B.2.10}$$

It follows that $\mathbf{H}_t = \mathbf{V}^{1/2}\mathbf{V}_t\mathbf{V}^{1/2}$, with $\mathbf{V}_t = \text{Var}_{t-1}(\mathbf{u}_t) = \mathbf{\Lambda}_m\mathbf{\Sigma}_t\mathbf{\Lambda}'_m$. The parameters of the model are $\mathbf{V}$, $\mathbf{\Lambda}_m$ and the parameters of the GARCH factors, α_i and β_i, $i = 1, \cdots, m$. If $m = N$, the number of parameters is $N \times (N + 5)/2$. In practice, $\mathbf{V}$ and $\mathbf{\Lambda}_m$ are replaced by their sample counterparts, and m is chosen by principal component analysis applied to the standardized residuals.

The GO-GARCH(1,1) model is defined as in (B.2.8), but in this case with $m = N$ and $\mathbf{\Lambda}$ a nonsingular matrix of parameters. The conditional correlation matrix of $\mathbf{r}_t$ is given by

$$\mathbf{R}_t = \mathbf{J}_t^{-1}\mathbf{V}_t\mathbf{J}_r^{-1}, \text{ where } \mathbf{J}_t = (\mathbf{V}_t \odot \mathbf{I}_m)^{1/2} \text{ and } \mathbf{V}_t = \mathbf{\Lambda}\Sigma_t\mathbf{\Lambda}'. \tag{B.2.11}$$

In van der Weide (2002), the singular value decomposition of the matrix $\mathbf{\Lambda}$ is used as a parameterization, i.e. $\mathbf{\Lambda} = \mathbf{P}\mathbf{L}^{1/2}\mathbf{U}$, where the matrix $\mathbf{U}$ is orthogonal, and $\mathbf{P}$ and $\mathbf{L}$ are defined as above (from the eigenvectors and eigenvalues). The O-GARCH model (when $m = N$) corresponds then to the particular choice $\mathbf{U} = \mathbf{I}_N$. More generally, van der Weide (2002) expresses $\mathbf{U}$ as the product of $N(N-1)/2$ rotation matrices

$$\mathbf{U} = \prod_{i<j} G_{ij}(\delta_{ij}), \quad -\pi \leq \delta_{ij} \leq \pi, \quad i, j = 1, 2, \cdots, n, \tag{B.2.12}$$

where $G_{ij}(\delta_{ij})$ performs a rotation in the plane spanned by the ith and jth vectors of the canonical basis of $\mathbb{R}^N$ over an angle δ_{ij}.

The orthogonal models are also nested in the BEKK model and, consequently, they guarantee that the (G)O-GARCH model is covariance-stationary if all the m univariate GARCH processes are stationary.

B.2.4 Restricted specification based on conditional correlation matrices

In this category, the matrix $\mathbf{H}_t$ is specified in a hierarchical form: first a univariate GARCH-based model is specified for each conditional variance; and second the standardized series by their conditional standard deviations are used in order to determine the conditional covariances. This kind of hierarchical specification for the matrix $\mathbf{H}_t$ makes the conditional correlation models very attractive, since it is possible to adopt different specifications for each conditional variance, which is estimated in a univariate fashion. Moreover, the estimation process can be conducted in more than one stage, therefore it is not difficult to estimate the model and one can work with a high dimension system. On the other hand, theoretical results on moments, ergodicity and stationarity may not be obtained in a straightforward way.

The matrix $\mathbf{H}_t$ is decomposed as

$$\mathbf{H}_t = \mathbf{D}_t\mathbf{R}_t\mathbf{D}_t = (\rho_{ijt}\sqrt{h_{iit}h_{jjt}}), \tag{B.2.13}$$

where $\mathbf{D}_t = diag(h_{11t}^{1/2}, \cdots, h_{NNt}^{1/2})$ is a $N \times N$ diagonal matrix, whose elements are the conditional standard deviations of each series, such that h_{iit} can be adjusted by any

GARCH-type model and the conditional correlation matrix $\mathbf{R}_t$ is positive definite at each point in time.

Different specifications for $\mathbf{R}_t$ result in different conditional correlation models. Bollerslev (1990) introduced the constant conditional correlation (CCC) GARCH model, assuming that the conditional correlation matrix is time-independent, that is, $\mathbf{R}_t = \mathbf{R}_0 = (\rho_{ij})$ is a symmetric positive definite matrix with $\rho_{ii} = 1$, $\forall i$. Therefore,

$$\mathbf{H}_t = \mathbf{D}_t \mathbf{R}_0 \mathbf{D}_t. \tag{B.2.14}$$

This model contains $N(N+5)/2$ parameters. The parameter estimation can be done in two steps: first, the conditional variances for each series are estimated in a univariate fashion; then, the matrix $\mathbf{R}_0$ can be estimated through the unconditional correlation matrix of standardized returns. He and Terasvirta (2004) propose an extension of the original CCC model to allow interactions between the conditional variances. This extension is known as the extended CCC model.

The assumption that the conditional correlation between the returns are constant makes the estimation process simple even for very large systems. Besides that, $\mathbf{H}_t$ is definite positive if and only if all the conditional variance h_{iit}, $i = 1, \cdots, N$, are positive and $\mathbf{R}_0$ is a positive definite matrix.

Nevertheless, assuming constant conditional correlations is not reasonable in many practical situations; see, for example, Longin and Solnik (1995). This fact has given rise to a class of models introduced by Engle (2002) and Tse and Tsui (2002), independently, and known in the literature as dynamic conditional correlation (DCC) models, in which the conditional correlation matrix is a time-dependent variable.

In the DCC model of Tse and Tsui (2002), DCC$_T$(M), the matrix $\mathbf{R}_t$ is given by

$$\mathbf{R}_t = (1-\theta_1-\theta_2)\mathbf{R} + \theta_1\mathbf{\Psi}_{t-1} + \theta_2\mathbf{R}_{t-1}, \tag{B.2.15}$$

where θ_1 and θ_2 are nonnegative scalar parameters satisfying the inequality $\theta_1+\theta_2<1$, $\mathbf{R}$ is a symmetric positive definite parameter matrix of dimension $N \times N$ such that $\rho_{ii}=1$ and $\mathbf{\Psi}_{t-1}$ is the sample correlation matrix of $\{r_{t-M}, \cdots, r_{t-1}\}$, i.e,

$$\mathbf{\Psi}_{ij,t-1} = \frac{\sum_{m=1}^{M} u_{i,t-m}u_{j,t-m}}{\sqrt{\sum_{m=1}^{M} u_{i,t-m}^2 \sum_{m=1}^{M} u_{j,t-m}^2}}, \tag{B.2.16}$$

with $u_{it} = r_{it}/\sqrt{h_{iit}}$. For $\mathbf{\Psi}_{t-1}$ to be positive definite, it is necessary that $M \geq N$.

The matrix $\mathbf{R}$ can be replaced by the sample correlation matrix of u_t. It follows from (B.2.15) that

$$\rho_{it,j} = (1-\theta_1-\theta_2)\rho_{ij} + \theta_2\rho_{ij,t-1} + \theta_1\frac{\sum_{m=1}^{M} u_{i,t-m}u_{j,t-m}}{\sqrt{\sum_{m=1}^{M} u_{i,t-m}^2 \sum_{m=1}^{M} u_{j,t-m}^2}}. \tag{B.2.17}$$

The DCC model of Engle (2002), DCC$_E(1,1)$, is defined as

$$\mathbf{R}_t = diag(q_{11,t}^{-1/2} \cdots q_{NN,t}^{-1/2})\mathbf{Q}_t \; diag(q_{11,t}^{-1/2} \cdots q_{NN,t}^{-1/2}), \tag{B.2.18}$$

where the $N \times N$ positive definite matrix $\mathbf{Q}_t = (q_{ij,t})$ is of the form

$$\mathbf{Q}_t = (1-a-b)\overline{Q} + a\mathbf{u}_{t-1}u'_{t-1} + b\mathbf{Q}_{t-1}, \tag{B.2.19}$$

where $u_{it} = r_{it}/\sqrt{h_{iit}}$ is the standardized returns, $\bar{\mathbf{Q}} = E(\mathbf{u}_t\mathbf{u}'_t)$ the unconditional variance matrix of $\mathbf{u}_t$, and a and b nonnegative scalar parameters satisfying $a+b<1$.

The standardization made in equation (B.2.18) ensures that the estimated conditional correlations are in the $[-1, 1]$ interval.

According to this model, the conditional correlation coefficient $\rho_{ij,t}$ is given by $q_{ij,t}/\sqrt{q_{ii,t}q_{jj,t}}$, with $q_{ij,t} = (1-a-b) + au_{i,t-1}u_{j,t-1} + bq_{ij,t-1}$. Note that $q_{ij,t}$ is the ARMA(1,1) model. In order for $\mathbf{H}_t$ to be a positive definite matrix, it is sufficient that all conditional variances be positive and that $\mathbf{R}_t$ be a positive definite matrix. $\mathbf{R}_t$ is positive definite, since $a, b > 0$ and $a + b < 1$. Engle and Sheppard (2001) prove consistency and asymptotic normality of the two-step estimator of the DCC_E model.

The basic difference between the calculated conditional correlations according to the DCC_E and DCC_T models is that in the model of Tse and Tsui they do not react so sharply to a shock in standardized returns as in the model of Engle. Both of the models have $(N + 1)(N + 4)/2$ parameters. Furthermore, they impose a common dynamic structure, governed by the parameters a and b in the model DCC_E and θ_1 and θ_2 in the DCC_T model, for all conditional correlations. This might not be true when the correlations between returns of a given asset have different behavior than the correlations between returns of an other asset.

To avoid the constraint that every series must have the same dynamic in the correlations, several generalizations of the DCC-GARCH model have been proposed. Billio et al. (2006) introduce a block-diagonal structure, where the dynamic is restricted to be equal just among certain groups of variables, and a BEKK structure on the conditional correlations is proposed. The Quadratic Flexible DCC (GFDCC) GARCH model is given by

$$\mathbf{Q}_t = \mathbf{C}'\mathbf{S}C + \mathbf{A}'\mathbf{u}_{t-1}\mathbf{u}'_{t-1}\mathbf{A} + \mathbf{B}'\mathbf{Q}_{t-1}\mathbf{B}, \tag{B.2.20}$$

where $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ are $N \times N$ parameter matrices, and $\mathbf{S}$ is the unconditional covariance matrix of the standardized errors $\mathbf{u}_t$. The model is stationary if $\mathbf{C}'\mathbf{S}C$ is positive definite and the eigenvalues of $\mathbf{A} + \mathbf{B}$ are less than one in modulus. The number of parameters, excluding the univariate GARCH equations, is $3N(N+1)/2$.

In the smooth transition conditional correlation (STCC) GARCH model of Silvennoinen and Teräsvirta (2005) the correlations vary smoothly between two different states. The time-varying correlation structure is defined as

$$\mathbf{R}_t = (1 - G_t)\mathbf{R}_{(1)} + G_t\mathbf{R}_{(2)}, \tag{B.2.21}$$

where $\mathbf{R}_{(1)}$ and $\mathbf{R}_{(2)}$ represent the two extreme states of correlations, say 1 and 2, respectively, and the transition function G_t is a logistic function given by

$$G_t = G(s_t, \gamma, c) = (1 + e^{-\gamma(s_t - c)})^{-1}, \quad \gamma > 0, \tag{B.2.22}$$

which is restricted on the interval $(0, 1)$. The conditional correlations can vary between the states 1 and 2 over time according to the observable transition variable s_t, which is chosen by the modeler. The parameter c define the location of the transition. When $s_t > c$, the correlations are closer to the state 1 than state 2, and when $s_t > c$, the opposite occurs. The parameter γ controls the smoothness of the transition between the two states, so it defines the speed of transition. The transition will be slower as γ is closer to zero and the transition function is faster as $\gamma \to \infty$. The STCC-GARCH model contains $N(N+2)+2$ parameters. The matrix $\mathbf{R}_t$ is positive definite $\forall t$ if the two correlation matrices $\mathbf{R}_{(1)}$ and $\mathbf{R}_{(2)}$ are positive definite. One particular case of the STCC-GARCH model is the time-varying conditional correlation (TVCC) GARCH model, obtained by defining the transition variable to be the calendar time, i.e., $s_t = t/T$.

Silvennoinen and Teräsvirta (2009a) extended the original STCC-GARCH model by allowing for another transition around the first one. The time-varying correlation structure in the Double Smooth Transition Conditional Correlation (DSTCC) GARCH model is defined as

$$\mathbf{R}_t = (1 - G_{1t})\mathbf{R}_{(1)t} + G_{1t}\mathbf{R}_{(2)t} \tag{B.2.23}$$

$$\mathbf{R}_{(i)t} = (1 - G_{2t})\mathbf{R}_{(i1)} + G_{2t}\mathbf{R}_{(i2)}, \quad i = 1, 2, \tag{B.2.24}$$

where the transition functions G_{1t} and G_{1t} are the logistic functions as in (B.2.22), and s_{1t} and s_{2t} are transition variables that can be either stochastic or deterministic. The parameters γ_i and c_i determine, respectively, the speed and the location of the transition $i,\ i = 1, 2$.

The correlation matrix $\mathbf{R}_t$ is thus a convex combination of four positive definite matrices, $\mathbf{R}_{(11)}$, $\mathbf{R}_{(12)}$, $\mathbf{R}_{(21)}$, and $\mathbf{R}_{(22)}$, each of which defines an extreme state of constant correlations. The positive definiteness of $\mathbf{R}_t$ at each point in time follows from the positive definiteness of these four matrices. As in the STCC-GARCH model, the values of these variables are assumed to be known at time t.

B.2.5 Semiparametric and nonparametric models

The majority of parametric models possess an interpretation of dynamic structure of the conditional covariance or correlation matrix and, furthermore, the quasi-maximum likelihood estimator is consistent. However, we must assume a particular structure of the data, which might not be correct, so we loss considerable efficiency. Semiparametric and nonparametric models have the advantage of not supposing a model specification, but nonparametric models do not have an easy interpretation. Semiparametric models combine the advantages of parametric models (consistency and interpretation) and those of nonparametric model, robust against distribution misspecification. For a complete review of semiparametric and nonparametric models see Linton (2008) and Franke et al. (2008).

The semiparametric conditional correlation (SPCC) MGARCH model of Hafner et al. (2005) adjusts the conditional variance parametrically by a univariate GARCH model and the conditional correlations $\mathbf{R}_t$ are then estimated using a transformed Nadaraya-Watson estimator

$$\mathbf{R}_t = (\mathbf{I} \odot \mathbf{Q}_t)^{-1/2}\mathbf{Q}_t(\mathbf{I} \odot \mathbf{Q}_t)^{-1/2}, \tag{B.2.25}$$

where

$$\mathbf{Q}_t = \frac{\sum_{\tau=1}^{T} \hat{\mathbf{u}}_\tau \hat{\mathbf{u}}_\tau' K_h(s_\tau - s_t)}{\sum_{\tau=1}^{T} K_h(s_\tau - s_t)}, \tag{B.2.26}$$

and $\hat{\mathbf{u}}_t = \hat{\mathbf{D}}_t^{-1}\mathbf{r}_t$ is the vector consisting of the standardized residuals, $s_t \in \mathbf{F}_{t-1}$ is a conditional variable, $K_h(.) = K(./h)/h$ is the kernel function and h is the bandwidth parameter.

B.3 MGARCH models with leverage effect

The leverage effect is motivated by the fact that negative returns have generally greater influence on future volatilities than positive returns. Some of the univariate models that consider the leverage effect are the EGARCH, TGARCH, GJR and QGARCH models. Rodríguez and Ruiz (2012) presented a review of univariate models that allow for this effect. For multivariate series the same rule applies: the variances and covariances may react differently to a positive and to a negative shock. Next we extend each of the approaches for constructing MGARCH models, including the leverage effect.

B.3.1 Models of conditional covariance matrices

The extension of the DVEC model (see, for example, Kroner and Ng, 1998) including asymmetric effects is given by the following two equations

$$h_{ii,t} = c_{ii} + a_i^2 r_{i,t-1}^2 + b_i^2 h_{ii,t-1} + \gamma_i^2 r_{i,t-1}^2 \mathbb{I}(r_{i,t-1} < 0) \tag{B.3.1}$$

$$h_{ij,t} = \phi_{ij} c_{ij} + \phi_{ij} a_i a_j r_{i,t-1} r_{j,t-1} + \phi_{ij} b_i b_j h_{ij,t-1} + \phi_{ij} \gamma_i \gamma_j r_{i,t-1}^- r_{j,t-1}^-, \ \forall i \neq j, \tag{B.3.2}$$

where $r_{i,t}^- = \max(0, -r_{i,t})$. Note that equation (B.3.1) corresponds to a GJR asymmetric function.

De Goeij and Marquering (2004) extend the GJR model by writing the elements of the conditional covariance matrix as

$$h_{ij,t+1} = \gamma_{ij} + \beta_{ij} h_{ij,t} + \alpha_{1ij} r_{i,t} r_{j,t} + \alpha_{2ij} r_{i,t}^- r_{j,t}^- + \alpha_{3ij} r_{i,t}^- r_{j,t}^+ + \alpha_{4ij} r_{i,t}^+ r_{j,t}^-, \tag{B.3.3}$$

where $i, j = 1, \cdots, N$, and $r_{i,t}^+$ is defined as $r_{i,t}^+ = \max(0, r_{i,t})$. De Goeij and Marquering (2004) call this model asymmetric diagonal VECH. They include several types of asymmetries in the model in equation (B.3.3). The term $\alpha_{2ij} r_{i,t}^- r_{j,t}^-$ assigns an asymmetric covariance effect on shocks in the same direction (simultaneous positive shocks versus simultaneous negative shocks). The terms $\alpha_{3ij} r_{i,t}^- r_{j,t}^+$ and $\alpha_{4ij} r_{i,t}^+ r_{j,t}^-$ considers different effects for opposite shocks in asset returns in addition to the existing negative return shock effects.

De Goeij and Marquering (2009) generalize this model first with the multivariate level effect as in, e.g., Christiansen (2005) and, thus, add level effects and cross-asymmetries in conditional variances and covariances

$$h_{ij,t+1} = |r_{i,t}r_{j,t}|^{\gamma_{ij}}[\omega_{ij} + \beta_{ij}h_{ij,t} + \alpha_{1ij}r_{i,t}r_{j,t} + \alpha_{2ij}r^{-}_{i,t}r^{-}_{j,t} + \alpha_{3ij}r^{-}_{i,t}r^{+}_{j,t} + \alpha_{4ij}r^{+}_{i,t}r^{-}_{j,t}], \quad \text{(B.3.4)}$$

where the conditional volatility of the assets depends on the level of the bond and stock returns, more specifically on the γ_{ij}th power of their level. This is known as the level effect. Consequently, the dynamics of the volatility depend on both return levels and information shocks. The larger γ_{ij} is, the more sensitive the (co)variance is to the levels of returns.

The generalization of the multivariate EGARCH model was used by Braun et al. (1995) for bivariate series and by Rossi et al. (2009), among others. Here we consider the model proposed by Kawakatsu (2006), which is a generalization of the univariate exponential GARCH model of Nelson (1991). The conditional covariance matrix is obtained from the equation

$$\text{vech}(\ln\mathbf{H}_t - \mathbf{C}) = \mathbf{A}\ \eta_{t-1} + \mathbf{F}\ (|\eta_{t-1}| - \mathbb{E}|\eta_{t-1}|) + \mathbf{B}\text{vech}(\ln\mathbf{H}_{t-1} - \mathbf{C}), \quad \text{(B.3.5)}$$

where the logarithm of a matrix is defined as the inverse of the exponential function, which is defined in Equation (B.3.6), $\mathbf{C}$ is a symmetric $N \times N$ matrix, $\mathbf{A}, \mathbf{F}$ are $N(N+1)/2 \times N$ parameter matrices and $\mathbf{B}$ is an $N(N+1)/2 \times N(N+1)/2$ parameter matrix. No restrictions are necessary in the parameter to ensure the positive definiteness because the matrix ln $\mathbf{H}_t$ does not need to be positive definite. The positive definiteness of the covariance matrix $\mathbf{H}_t$ follows from the fact that for any symmetric matrix $\mathbf{S}$, the matrix exponential defined as

$$\exp(\mathbf{S}) = \sum_{i=0}^{\infty} \frac{\mathbf{S}^i}{i!} \tag{B.3.6}$$

is positive definite.

Hafner and Herwartz (1998) proposed an asymmetric version of the bivariate BEKK(1,1,1). Here we present the approach of Kroner and Ng (1998), which can be used for any dimension. The asymmetric BEKK(1,1,1) (ABEKK) is given by

$$\mathbf{H}_t = \mathbf{C} + \mathbf{A}^{*'}\mathbf{r}_{t-1}\mathbf{r}'_{t-1}\mathbf{A}^* + \mathbf{B}^{*'}\mathbf{H}_{t-1}\mathbf{B}^* + \mathbf{G}^{*'}\mathbf{r}^-_{t-1}\mathbf{r}'^-_{t-1}\mathbf{G}^*, \tag{B.3.7}$$

where $\mathbf{r}^-_t = (r^-_{1,t}, \cdots, r^-_{N,t})'$. This model extends the original BEKK model by adding a quadratic form, depending on the vector $\mathbf{r}^-_t$, which contains just negative returns, so allowing for leverage effects.

B.3.2 Factor models

The models in this category can be chosen as being linear combinations of univariate GARCH-type models which allow leverage effect, such as the EGARCH model of Nelson (1991) and the TGARCH model (see Zakoïan, 1994). For example, in the O-GARCH(1,1,m) model of Alexander and Chibumba (1997), equation (B.2.10) can be replaced by

$$\sigma^2_{fit} = (1 - \alpha_i - \beta_i) + \alpha_i f^2_{i,t-1} + \beta_i \sigma^2_{fi,t-1} + \gamma_i \mathbb{I}(f_{i,t-1} < 0) \quad i = 1, \cdots, m, \tag{B.3.8}$$

where $\mathbb{I}(r_{i,t-1} < 0)$ equals 1 if $r_{i,t} < 0$, and zero otherwise. This model allows for asymmetric effects in the conditional variances.

The asymmetric F-GARCH($k = 1$) considers a bad news portfolio (x_p), which is constructed by taking a weighted average of the individual asset of bad news, such that the weights are the same as the factor weights. This model is written as

$$h_{ij,t} = \sigma_{ij} + c_i c_j h_{p,t}, \quad \forall i,j \tag{B.3.9}$$

$$h_{p,t} = \omega_p + \beta h_{p,t-1} + \alpha r_{p,t-1}^2 + \gamma x_{p,t-1}^2, \tag{B.3.10}$$

where $h_{p,t} = \mathbf{d}'\mathbf{H}_t\mathbf{d}$, $\omega_p = \mathbf{d}'\mathbf{\Omega}\mathbf{d}$, $r_{p,t} = \mathbf{d}'\mathbf{r}_t$, $x_{p,t} = \mathbf{d}'\mathbf{r}_t^-$, $\sigma_{ij} = \omega_{ij} - c_i c_j \mathbf{d}'\mathbf{\Omega}\mathbf{d}$.

B.3.3 Restricted specification based on conditional correlation matrices

Due to the hierarchical specification of the matrix $\mathbf{H}_t$, we can choose asymmetric models on conditional variance easily. Indeed, initially a univariate GARCH-based model is specified for conditional variances, which is estimated individually and independently. For example, McAleer, et al. (2009) proposed the CCC asymmetric GARCH (CCC-AGARCH) where they allow for asymmetric impact of past returns on the volatility. This model is a generalization of the extended GARCH of He and Terasvirta (2004), but, unlike the extended GARCH, it allows for leverage effect in volatilities. There are similar extensions using different specifications. For example, Koutmos and Booth (1995) consider the CCC model for investigating the dynamic interaction of the three major stock markets, Tokyo, London and New York. They adjusted each conditional variance by an extended EGARCH model process that allows its own lagged standardized returns as well as cross-market standardized returns to exert an asymmetric impact on the volatility of market i, $i = 1, 2, 3$.

The m-dimensional vector process $\{\varepsilon_t\}$ is a CCC-AGARCH(1,1) model if it satisfies (B.2.14) and

$$\mathbf{D}_t = \omega + \mathbf{A}_+\mathbf{r}_{t-1}^{2+} + \mathbf{A}_-\mathbf{r}_{t-1}^{2-} + \mathbf{B}\mathbf{D}_{t-1}, \tag{B.3.11}$$

where $\mathbf{D}_t = diag(h_{11t}^{1/2} \cdots h_{NNt}^{1/2})$, $r_{i,t}^{2+} = (r_{it}^{+})^2$, $r_{i,t}^{2-} = (r_{it}^{-})^2$, $\mathbf{r}_t^{2+} = (r_{1,t}^{2+}, \cdots, r_{N,t}^{2+})'$ and $\mathbf{r}_t^{2-} = (r_{1,t}^{2-}, \cdots, r_{N,t}^{2-})'$. The $m \times m$ matrices $\mathbf{A}_+$, $\mathbf{A}_-$ and $\mathbf{B}$ have positive coefficients, the $m \times 1$ vector ω has strictly positive coefficients and $\mathbf{R}_0$ is the correlation matrix.

The AGARCH model allows for leverage effect by choosing $\mathbf{A}_+ \neq \mathbf{A}_-$. If we impose $\mathbf{A}_+ < \mathbf{A}_-$ element by element, the AGARCH model has higher increase of volatility after negative returns rather than positive returns of the same magnitude. If $\mathbf{A}_+ = \mathbf{A}_-$, the model reduces to the popular CCC-GARCH(p, q), which does not allow for leverage effect. In the original CCC-GARCH model of Bollerslev (1990), it is assumed that the matrices $\mathbf{A}$ and $\mathbf{B}$ are diagonal. In (B.3.11) the conditional variance $h_{kk,t}$ depends not only on its past values but also on the past values of the other components.

Francq and Zakoïan (2012) provide a necessary and sufficient strict stationarity condition for CCC-AGARCH and CCC-GARCH models and put some restrictions on the matrices $\mathbf{A}_+$, $\mathbf{A}_-$ and $\mathbf{B}$ to ensure the uniqueness of the parameterization. The strict stationarity assumption and the identifiability conditions are explicit. Franq and Zaköian (2012) establish the strong consistency and asymptotic normality (CAN) of the quasi-maximum likelihood estimator (QMLE) of the parameters of CCC-AGARCH models. The conditions required are mild and coincide with the minimal ones in the univariate case. For example, no assumption is made related to the existence of moments of the observed process; the only moment assumption is based on the i.i.d. process. Instead, strict stationarity is required, the parametric space must be compact and $\mathbf{R}_0$ has to be a positive-definite correlation matrix for all possible parameters in the parametric space.

The asymmetric generalized DCC (AGDCC) GARCH model of Cappiello et al. (2006) is a generalization of DCC models, which imposes a BEKK structure on the conditional correlations and includes asymmetric effects. It is an extension of the original model along two dimensions: asset-specific correlation evolution parameters and conditional asymmetries in correlation. The matrix $\mathbf{Q}_t$ in (B.2.19) is rewritten as

$$\mathbf{Q}_t = (\bar{\mathbf{Q}} - \mathbf{A}'\bar{\mathbf{Q}}\mathbf{A} - \mathbf{B}'\bar{\mathbf{Q}}\mathbf{B} - \mathbf{G}'\bar{\mathbf{N}}\mathbf{G}) + \mathbf{A}'\mathbf{u}_{t-1}\mathbf{u}'_{t-1}\mathbf{A} + \mathbf{B}'\mathbf{Q}_{t-1}\mathbf{B} \qquad \text{(B.3.12)}$$
$$+\mathbf{G}'\mathbf{u}^{-}_{t-1}\mathbf{u}'^{-}_{t-1}\mathbf{G},$$

where $\mathbf{A}$, $\mathbf{B}$, $\mathbf{G}$ are $N \times N$ parameter matrices, $\mathbf{u}^{-}_t$ is defined as before, $\bar{\mathbf{Q}}$ and $\bar{\mathbf{N}}$ are the unconditional covariance matrices of $\mathbf{u}_t$ and $\mathbf{u}^{-}_t$, respectively. The number of parameters governing the correlations in this model increases rapidly with the dimension of the series. The asymmetric DCC (ADCC) is a special case of the AGDCC obtained when the matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{G}$ are replaced by scalars. Analogously, the generalized DCC (GDCC) is a special case of the AG-DCC when $\mathbf{G} = \mathbf{0}$.

A sufficient condition to guarantee the positive definiteness of $\mathbf{Q}_t$ is that the intercept, $\bar{\mathbf{Q}} - \mathbf{A}'\bar{\mathbf{Q}}\mathbf{A} - \mathbf{B}'\bar{\mathbf{Q}}\mathbf{B} - \mathbf{G}'\bar{\mathbf{N}}\mathbf{G}$, is positive semi-definite and the initial covariance matrix $\mathbf{Q}_0$ is positive definite; see Ding and Engle (2001) for further details.

Table B.3.3 presents a brief comparison of some of the models presented in this article, with respect with the number of parameters, whether the model allows for dependence of between volatilities and of covariances and whether the model allows for leverage effect of the volatilities and of the covolatilities (conditional covariances).

Tabela B.1: Comparison between the models

Model	N^0. of parameters			Dependence		Asymmetry	
	$N = 2$	$N = 5$	$N = 20$	bt. vol.	of cov.	of vol.	of cov.
VEC	21	465	88410	yes	yes	no	no
DVEC	9	45	630	no	no	no	no
BEKK($k = 1$)	11	65	1010	yes	yes	no	no
DBEKK($k = 1$)	7	25	250	yes	yes	no	no
ABEKK($k = 1$)	15	90	1410	yes	yes	no	no
F-GARCH ($k = 1$)	9	27	252	maybe	maybe	no	no
F-GARCH ($k = 3$)	-	51	336	maybe	maybe	no	no
O-GARCH($m = n$)	7	25	250	maybe	maybe	no	no
CCC-GARCH	7	25	250	no	yes	no	no
DCC-$_E$GARCH	9	27	252	no	yes	no	no
DCC-$_T$GARCH	9	27	252	no	yes	no	no
STCC-GARCH	10	37	442	no	yes	no	no
DSTCC-GARCH	14	59	824	no	yes	no	no
MEGARCH	25	400	52900	yes	yes	yes	yes
CCC-AGARCH	15	90	1410	yes	yes	yes	no
ADCC-GARCH	13	43	463	no	yes	no	yes
AGDCC-GARCH	18	90	1260	yes	yes	no	yes

* bt. vol. = between volatility.

** of cov. = of covariances.

B.4 Simulation

The simulation compares the bivariate case of BEKK, DBEKK, GO-GARCH, CCC-GARCH, DCC-GARCH and ADCC-GJR models. In all the models, we include just one lag in relation to the past returns, conditional variances and covariances, considering, Gaussian innovations in all of them. Only the ADCC-GJR allows for leverage effects. The BEKK model was only used to simulate series. The results, when the BEKK model is fitted, are not shown because it was difficult to control the convergence problem during the estimation process. The simulation investigates whether the AIC and BIC criteria are able to select the right model and the loss of efficiency when one used a misspecified model to estimate the conditional volatility. The goodness of fit was measured using the bias, the mean square error (MSE), the mean absolute error (MAE), the mean square relative error (MSRE) and the mean absolute relative error (MARE), all these statistics related to the estimation of the volatilities, defined, respectively, as

$$\text{bias}(i,j) \quad : \quad \sum_{t=101}^{1000} (\hat{h}_{ijt} - h_{ijt})/900 \tag{B.4.1}$$

$$\text{MSE}(i,j) \quad : \quad \sum_{t=101}^{1000} (\hat{h}_{ijt} - h_{ijt})^2/900 \tag{B.4.2}$$

$$\text{MAE}(i,j) \quad : \quad \sum_{t=101}^{1000} |\hat{h}_{ijt} - h_{ijt}|/900 \tag{B.4.3}$$

$$\text{MSRE}(i,j) \quad : \quad \sum_{t=101}^{1000} [(\hat{h}_{ijt} - h_{ijt})/h_{ijt}]^2/900 \tag{B.4.4}$$

$$\text{MARE}(i,j) \quad : \quad \sum_{t=101}^{1000} |(\hat{h}_{ijt} - h_{ijt})/h_{ijt}|/900, \tag{B.4.5}$$

where $\hat{h}_{ijt}$, is the estimated volatility and h_{ijt} is the true volatility, $i = 1,2$, $j = 1,2$. All these errors were computed from the 101-st observation to avoid problems of initial values.

For each model, we generated 1000 replications of bivariate series of size 1000. For each of the bivariate generated series, we followed the four-step procedure:

- Fit all the models.
- Select the best model, among all the six fitted models, according to the AIC and BIC.
- For the six adjusted models, estimate the volatilities and evaluate the five goodness of fit statistics defined in (B.4.1)-(43).

The parameters of the BEKK model were chosen like those found in an application of Franke et al. (2004, Section 13.3.4). In order to compare the other models with similar volatilities and covolatility, we simulated a series of size $100,000$ using the BEKK model with the same parameters of Franke et al. (2004, Section 13.3.4). The estimated parameter for DBEKK, GO-GARCH, CCC-GARCH and DCC-GARCH models are used to simulate the series for each case. The ADCC-GJR model was an exception, since it has leverage effects. In this case, the parameters were chosen similarly to those of Cappiello et al. (2006) for the bivariate times series (Belgian stocks, French stocks). The simulation and estimation were done using the Sheppard's MFE Toolbox package, written in MATLAB, and downloaded on June 11, 2013, for all models, except for the simulations of the DCC-GARCH and ADCC-GJR models, which was done by the rmgarch package of the R-project software. All models were estimated by the maximum likelihood method considering Gaussian errors. For the CCC-GARCH, DCC-GARCH and ADCC-GJR models, we used the two-stage estimation method, i.e., a quasi maximum likelihood method.

B.4.1 Performance of the AIC and BIC for model selection

The number of times that each model is selected by the AIC and BIC are shown in Table B.2. The analysis of the results clearly shows the parsimony feature of the BIC. The brackets after each model's name contain its respective number of parameters.

When the right model is the BEKK (11) or the DBEKK (7), the BIC and AIC select the DBEKK model in approximately 99% and 93% of the cases, respectively. In the CCC (7) simulation, the BIC and AIC select the right model in approximately 94% and 79%

of the cases, respectively; the DCC (9) model in 0% and 14.5% of the cases, respectively; the GO-GARCH (7) in 5.2% and 4.6%, respectively; the ADCC-GJR (15) model is only selected by the AIC model and only in 2.4% of the replications. When the right model is the DCC-GARCH (9), the BIC selects the DBEKK (7) model, which has a small number of parameters, in 99% of the cases. The results presented in the next section show that the DBEKK model is able to approximate the DCC-GARCH (9) process very well, at least for the parameters used in the simulation. The DBEKK(7) is also selected by AIC in 52% of the cases. The right model is chosen in 45% of replications by the AIC and never by the BIC. For the ADCC-GJR (15) simulation, a model with leverage effect and large number of parameters, the AIC selects the right model in 89% of the cases, while the BIC in only 52%. The DCC-GARCH model, which is an ADCC-GJR model without the leverage effect in the covolatiliy and in the volatilities, is selected 43% of the times by the BIC, and 11% by the AIC. In summary, the BIC performs well when the true model is more simple, and the AIC performs better for more complex models.

B.4.2 Performance in volatility estimation

Suppose that a model A is generated and the aim is to evaluate the efficiency of the model B in estimating h_{ij}, according to some measures of efficiency. In this application, we define efficiency as the ratio of the goodness of fit defined at the beginning of the section when the correct model (A) is fitted compared to the goodness of fit by model B. The mean of the goodness of fit measured is presented in Tables B.3-B.7, when the data generating model is BEKK, DBEKK, CCC-GARCH, DCC-GARCH and ADCC-GJR, respectively. The values in brackets represent the efficiency of misspecified model estimators, according

Tabela B.2: Frequency of the cases when the AIC and BIC selected each model. Results for AIC are in the first line and for BIC are in the second. For each model, we generated $1,000$ series of size $1,000$.

Selected model	True data generating model				
	BEKK	DBEKK	CCC-GARCH	DCC-GARCH	ADCC-GJR
DBEKK	898	951	1	506	6
	977	994	2	964	36
GO-GARCH	17	5	39	7	0
	18	5	45	11	9
CCC-GARCH	0	0	807	0	0
	0	0	950	0	3
DCC-GARCH	77	41	123	425	108
	5	1	3	25	432
ADCC-GJR	8	3	30	62	886
	0	0	0	0	520

to the measures of the goodness of fit defined previously, compared to the goodness of fit of the right model. The BEKK model was only used in the generating process. By adjusting the CCC-GARCH and DCC-GARCH models, the estimates of volatilities h_{11} and h_{22} are the same as the univariate specification, since we used two step estimators for both models. The main conclusions are

- In general, the smallest bias does not occur when the correct model is fitted. In many cases the smallest bias occurs when fitting the GO-GARCH model. Because in all cases the bias is negligible, all the comments below are related to the measures of efficiency of other goodness of fit statistics defined previously.

- The goodness of fit by the GO-GARCH model is always very bad, whatever the correct model is, indicating that this model is not suitable when the other models reflect the true data generating process. Thus, the GO-GARCH is not taken into

Tabela B.3: Goodness of fit statistics when the series are generated by the BEKK model. -G means GARCH.

Statistics	Volatility	DBEKK	GO-G	CCC-G	DCC-G	ADCC-GJR
bias	h_{11}	-2.566	**-1.258**	*-2.461*	*-2.461*	-2.513
($\times 10^{-7}$)	h_{22}	-4.887	**-1.112**	*-3.438*	*-3.438*	-3.566
	h_{12}	-4.208	**-1.331**	9.724	*-3.487*	-3.333
MSE	h_{11}	*0.465*	1.330	**0.413**	**0.413**	0.536
($\times 10^{-10}$)	h_{22}	*0.608*	2.110	**0.537**	**0.537**	0.704
	h_{12}	0.461	0.756	0.790	**0.298**	*0.362*
MAE	h_{11}	*4.551*	7.152	**4.253**	**4.253**	4.877
($\times 10^{-6}$)	h_{22}	*4.823*	9.777	**4.384**	**4.384**	4.946
	h_{12}	4.433	5.017	6.113	**3.524**	*3.873*
MSRE	h_{11}	*0.988*	2.590	**0.882**	**0.882**	1.150
($\times 10^{-2}$)	h_{22}	*1.078*	6.241	**0.870**	**0.870**	1.083
	h_{12}	**214.4**	494.0	16852	*435.9*	437.5
MARE	h_{11}	*0.769*	1.193	**0.719**	**0.719**	0.821
($\times 10^{-1}$)	h_{22}	*0.803*	1.761	**0.721**	**0.721**	0.807
	h_{12}	1.083	1.332	2.364	**0.882**	*0.959*

In bold the model with best performance.
In italic the model with the second best performance.

account in the following.

- As expected, the efficiencies of the estimates of h_{11} and h_{22} in relation to DBEKK, CCC-GARCH and ADCC-GARCH models are always very close to each other.

- When the BEKK is the true model, the DCC-GARCH model is the best to estimate the volatilities and the covolatility, according to any measure of goodness of fit, except for the MSRE in the covolatility estimation. In this case the performance of the DBEKK is the best, and is also the best to estimate the volatilities h_{11} and h_{22}. The performance of the CCC is very bad, as expected, since it considers constant correlation.

Tabela B.4: Goodness of fit statistics when the series are generated by the DBEKK model. The brackets contain the relative efficiency in relation to the estimation by the true model. -G means GARCH.

Statistics	Volati lity	Fitted model DBEKK	GO-G	CCC-G	DCC-G	ADCC-GJR
bias	h_{11}	2.717	1.521	*0.707*	*0.707*	**0.612**
($\times 10^{-7}$)	h_{22}	*1.641*	**0.897**	-2.125	-2.125	-2.211
	h_{12}	2.745	1.042	12.125	*-0.337*	**-0.132**
MSE	h_{11}	**0.192**	1.812(0.11)	*0.361*(0.53)	*0.361*(0.53)	0.511(0.38)
($\times 10^{-10}$)	h_{22}	**0.341**	3.586(0.10)	*0.677*(0.50)	*0.677*(0.50)	0.963(0.35)
	h_{12}	**0.170**	1.307(0.13)	0.875(0.19)	*0.299*(0.57)	0.398(0.43)
MAE	h_{11}	**3.019**	8.411(0.36)	*3.910*(0.77)	*3.910*(0.77)	4.652(0.65)
($\times 10^{-6}$)	h_{22}	**3.523**	12.833(0.27)	*4.431*(0.80)	*4.431*(0.80)	5.342(0.66)
	h_{12}	**2.742**	6.684(0.41)	6.464(0.42)	*3.410*(0.80)	3.917(0.70)
MSRE	h_{11}	**0.392**	3.466(0.11)	*0.667*(0.59)	0.667(0.59)	0.923(0.42)
($\times 10^{-2}$)	h_{22}	**0.408**	7.147(0.06)	*0.638*(0.64)	*0.638*(0.64)	0.885(0.46)
	h_{12}	**3.18**	110.8(0.03)	1710(0.002)	12.08(0.26)	*6.27*(0.51)
MARE	h_{11}	**0.487**	1.325(0.37)	*0.618*(0.79)	*0.618*(0.79)	0.730(0.67)
($\times 10^{-1}$)	h_{22}	**0.497**	1.928(0.26)	*0.603*(0.82)	*0.603*(0.82)	0.719(0.69)
	h_{12}	**0.572**	1.517(0.38)	1.792(0.32)	*0.715*(0.80)	0.796(0.72)

In bold the model with best performance.
In italic the model with the second best performance.

- In the case of DBEKK, CCC-GARCH, DCC-GARCH and ADCC-GJR models, the best results occur when fitting the correct model, except for some cases in the estimation of volatilities h_{11} and h_{22} by model DBEKK. This occurs in all the measures of efficiency when DCC is the correct model. Even in this case, the efficiency gain is not very big. This gain in efficiency can be explained by the fact that the estimates in the CCC and DCC models are performed in two steps, while the ones from DBEKK are done by maximum likelihood, i.e., it takes the information given by the other series.

- When the CCC-GARCH is the true model, the efficiencies of CCC-GARCH, DCC-

Tabela B.5: Goodness of fit statistics when the series are generated by the CCC-GARCH model. The brackets contain the relative efficiency in relation to the estimation by the true model. -G means GARCH.

Statistics	Volatility	DBEKK	GO-G	CCC-G	DCC-G	ADCC-GJR
		Fitted model				
bias	h_{11}	7.86	**0.019**	*-0.674*	*-0.674*	-0.742
($\times 10^{-7}$)	h_{22}	10.87	0.796	*-0.733*	*-0.733*	**-0.689**
	h_{12}	12.14	0.266	*0.094*	-0.387	**-0.076**
MSE	h_{11}	*0.470*(0.75)	1.688(0.21)	**0.352**	**0.352**(1.0)	0.492(0.72)
($\times 10^{-10}$)	h_{22}	*0.452*(0.90)	1.892(0.22)	**0.409**	**0.409**(1.0)	0.546(0.75)
	h_{12}	0.286(0.55)	0.536(0.29)	**0.156**	*0.169*(0.92)	0.223(0.70)
MAE	h_{11}	4.689(0.83)	8.183(0.48)	**3.912**	**3.912**(1.0)	*4.621*(0.85)
($\times 10^{-6}$)	h_{22}	*4.335*(0.86)	9.074(0.41)	**3.733**	**3.733**(1.0)	4.393(0.85)
	h_{12}	3.661(0.73)	4.514(0.59)	**2.685**	*2.744*(0.98)	3.138(0.86)
MSRE	h_{11}	1.027(0.65)	3.259(0.20)	**0.666**	**0.666**(1.0)	*0.915*(0.73)
($\times 10^{-2}$)	h_{22}	0.942(0.68)	4.802(0.13)	**0.641**	**0.641**(1.0)	*0.866*(0.74)
	h_{12}	1.028(0.53)	1.525(0.36)	**0.543**	*0.564*(0.96)	0.734(0.74)
MARE	h_{11}	0.758(0.82)	1.294(0.48)	**0.622**	**0.622**(1.0)	*0.730*(0.85)
($\times 10^{-1}$)	h_{22}	0.730(0.83)	1.580(0.38)	**0.608**	**0.608**(1.0)	*0.712*(0.85)
	h_{12}	0.790(0.73)	0.948(0.61)	**0.580**	*0.591*(0.98)	0.672(0.86)

In bold the model with best performance.
In italic the model with the second best performance.

GARCH and DBEKK models in estimating h_{11} and h_{22} are approximately equal, but in all the cases the efficiencies of CCC-GARCH and DCC-GARCH models are better. The loss of efficiency of the ADCC-GJR model is around the interval (10%, 30%), probably due to unnecessary generalization to allow for the effect of asymmetry. Except for the GO-GARCH, the other models can approximate the CCC-GARCH models, but lose considerable efficiency in estimating h_{12} due to the introduction of superfluous parameters.

- In the DCC-GARCH simulation, the DBEKK model is the most efficient in estimating h_{11} and h_{22}, as mentioned previously. The efficiency of the ADCC-GJR

Tabela B.6: Goodness of fit statistics when the series are generated by the DCC-GARCH model. The brackets contain the relative efficiency in relation to the estimation by the true model.-G means GARCH.

Statistics	Volatility	Fitted model DBEKK	GO-G	CCC-G	DCC-G	ADCC-GJR
bias	h_{11}	-2.694	**0.708**	*-0.737*	*-0.737*	-0.848
($\times 10^{-7}$)	h_{22}	-2.753	0.944	**-0.585**	**-0.585**	*-0.708*
	h_{12}	*0.182*	0.766	0.153	**0.149**	-0.370
MSE	h_{11}	**0.398**(1.05)	1.750(0.24)	*0.416*(1.0)	*0.416*	0.561(0.74)
($\times 10^{-10}$)	h_{22}	**0.347**(1.00)	1.846(0.19)	*0.348*(1.0)	*0.348*	0.490(0.71)
	h_{12}	0.138 (0.74)	1.870(0.05)	4.020(0.03)	**0.102**	*0.124*(0.82)
MAE	h_{11}	**3.815**(1.05)	8.898(0.45)	*4.019*(1.0)	*4.019*	4.747(0.85)
($\times 10^{-6}$)	h_{22}	**3.359**(1.06)	8.934(0.40)	*3.555*(1.0)	*3.555*	4.209(0.84)
	h_{12}	2.619(0.86)	9.122(0.25)	14.526(0.15)	**2.240**	*2.468*(0.91)
MSRE	h_{11}	**0.0572**(1.19)	0.4476(0.15)	*0.0678*(1.0)	*0.0678*	0.0922(0.74)
($\times 10^{-1}$)	h_{22}	**0.0537**(1.22)	0.6105(0.11)	*0.0656*(1.0)	*0.0656*	0.0893(0.73)
	h_{12}	*27740*(0.53)	222613(0.07)	201328(0.07)	**14719**	117859(0.12)
MARE	h_{11}	**0.580**(1.07)	1.448(0.43)	*0.622*(1.0)	*0.622*	0.733(0.85)
($\times 10^{-1}$)	h_{22}	**0.563**(1.09)	1.662(0.37)	*0.612*(1.0)	*0.612*	0.719(0.85)
	h_{12}	*11.590*(0.83)	29.127(0.33)	28.905(0.33)	**9.636**	11.828(0.81)

In bold the model with best performance.
In italic the model with the second best performance.

model, which has unnecessary asymmetry parameters, is around 80%. The efficiency of the DBEKK model to estimate h_{12} is also around 80%. The efficiency of the CCC-GARCH model, as expected, is low, since it considers constant correlation.

- When ADCC-GJR is the true model, all models lose more than 50%, according to all goodness of fit measures. This is an indication that none of the other models can emulate the features of the ADCC-GJR model.

Tabela B.7: Goodness of fit statistics when the series are generated by the ADCC-GJR model. The brackets contain the relative efficiency in relation to the estimation by the true model. -G means GARCH.

Statistics	Volatility	Fitted model DBEKK	GO-G	CCC-G	DCC-G	ADCC-GJR
bias	h_{11}	-63.54	1.311	0.591	0.591	-1.016
($\times 10^{-6}$)	h_{22}	-64.36	1.754	1.149	1.149	-0.311
	h_{12}	-64.68	2.496	8.478	4.248	-0.659
MSE	h_{11}	11.73(0.09)	8.324(0.13)	*5.234*(0.13)	*5.234*(0.13)	**1.054**
($\times 10^{-8}$)	h_{22}	9.718(0.07)	7.135(0.10)	*3.749*(0.19)	*3.749*(0.19)	**0.711**
	h_{12}	9.215(0.08)	5.766(0.12)	4.952(0.14)	*3.766*(0.19)	**0.701**
MAE	h_{11}	18.13(0.28)	14.46(0.35)	*12.61*(0.40)	*12.61*(0.40)	**5.128**
($\times 10^{-5}$)	h_{22}	18.94(0.26)	14.43(0.35)	*12.65*(0.39)	*12.65*(0.39)	**4.996**
	h_{12}	17.32(0.26)	12.54(0.36)	14.13(0.32)	*11.78*(0.39)	**4.567**
MSRE	h_{11}	1.397(0.06)	0.705(0.12)	*0.514*(0.16)	*0.514*(0.16)	**0.085**
($\times 10^{-1}$)	h_{22}	1.521(0.05)	0.825(0.09)	*0.526*(0.14)	*0.526*(0.14)	**0.074**
	h_{12}	931.2(0.25)	1514(0.16)	78039(0.003)	*3215*(0.07)	**236.6**
MARE	h_{11}	2.660(0.26)	1.986(0.35)	*1.693*(0.41)	*1.693*(0.41)	**0.694**
($\times 10^{-1}$)	h_{22}	2.669(0.24)	1.938(0.34)	*1.685*(0.39)	*1.685*(0.39)	**0.650**
	h_{12}	3.738(0.28)	3.066(0.34)	7.845(0.13)	*2.620*(0.40)	**1.051**

In bold the model with best performance.
In italic the model with the second best performance.

B.5 Application

Some of the models presented in the proposal are fitted to a three dimensional returns of the following Brazilian series: Ibovespa (IBV), Petrobras PN (Petro) and Vale PNA in the period from April 15, 2002 to April 12, 2013, with a total of 2725 observations. Holidays and days without trading were ignored. The fitted models are the diagonal BEKK (symmetric and asymmetric), O-GARCH and GO-GARCH, CCC , DCC and ADCC. In the last three cases, the marginal series are adjusted by GARCH, GJR and TGARCH models. The first two allow for leverage effect, while GARCH model does not.

Tabela B.8: Results of the estimate of the VAR(1) model to the data example. IBV, Petro and Vale are series 1, 2 and 3, respectively

Parameter	Estimate	S.D.	p-value
μ_1	0.00053	0.00036	0.141
ϕ_{11}	0.03744	0.04206	0.373
ϕ_{12}	-0.01733	0.02034	0.394
ϕ_{13}	-0.02294	0.03274	0.483
μ_2	0.00124	0.00051	0.014
ϕ_{21}	0.10719	0.05740	0.062
ϕ_{22}	-0.15158	0.02981	$<$0.001
ϕ_{23}	-0.04987	0.04140	0.228
μ_3	0.00050	0.00044	0.256
ϕ_{31}	0.04107	0.05313	0.439
ϕ_{32}	-0.02856	0.02408	0.236
ϕ_{33}	0.04170	0.04008	0.298

In all the cases, the innovations are from the multivariate normal distribution.

To remove the serial correlation in the level of the series they were filtered by a VAR(1) model. This model is able to remove the autocorrelation present in the series. Table B.8 shows the results of the estimated model. The autoregressive parameter ϕ_{22} and the constant μ_2 are significant at the 5% level, while the autoregressive parameter ϕ_{21} is significant at the 10% level. We opted to leave all the parameters in the model.

A total of 13 MGARCH models were fitted to the filtered data. Table B.9 presents the AIC and BIC of the fitted models. The best model according to both criteria is the asymmetric DBEKK, which has leverage effect in the volatilities and in the covolatilities. The second, third and fourth best models selected by AIC are ADCC-GJR (4) (order according to BIC in brackets), DCC-GJR (2) and DCC-TGARCH(3),respectively, while for the BIC they are DCC-GJR (3) (order according to AIC in brackets), DCC-TGARCH (4) and ADCC-GJR (2). The DCC-GJR and the DCC-TGARCH allow for leverage effect

Tabela B.9: AIC and BIC of the models adjusted to the data example. The orders according to AIC and BIC are given in brackets. The top, middle and bottom panels contain models without leverage effect, leverage only in the diagonal of the volatility matrix and leverage in all elements of the matrix, respectively

Model	AIC	BIC
DBEKK	-44024.7 (9)	-43953.8 (7)
CCC-GARCH	-43959.0 (12)	-43888.1 (11)
DCC-GARCH	-44049.9 (6)	-43967.1 (5)
O-GARCH	-43673.0 (13)	-43602.1 (13)
GO-GARCH	-44039.6 (7)	-43951.0 (8)
CCC-GJR	-43993.2 (10)	-43902.6 (10)
CCC-TGARCH	-43969.3 (11)	-43880.7 (12)
DCC-GJR	-44125.4 (3)	-44025.0 (2)
DCC-TGARCH	-44099.3 (4)	-43998.9 (3)
Asymmetric DBEKK	-44134.8 (1)	-44046.2 (1)
ADCC-GARCH	-44036.4 (8)	-43912.3 (9)
ADCC-GJR	-44127.9 (2)	-43986.1 (4)
ADCC-TGARCH	-44098.2 (5)	-43956.4 (6)

only in the diagonal of the volatility matrix and the ADCC-GJR allows in all elements of the volatility matrix. Note that the best four models following the AIC and BIC have asymmetric effects at least in volatilities of the univariate series. In general, there is good agreement between the criteria, with a slight tendency of the AIC to select a model with more parameters. The worst models are CCC-GARCH, OGARCH, CCC-GJR and CCC-TGARCH, which are ranked as the four worst models according to both criteria. We selected two models without leverage effect (DBEKK and DCC-GARCH), one with leverage effect only in the diagonal (DCC-GJR) and two with leverage effect in all elements of the volatility matrix (ADBEKK and ADCC-GJR) to present their estimates of the volatility. The estimates of these models, and of the ADCC-GARCH model, are presented in Tables B.10, B.11 and B.12. All the models passed the usual autocorrelation

Tabela B.10: Results of the estimation of the DBEKK model to the filtered data

	Symmetric DBEKK			Asymmetric DBEKK		
Parameter	Estimate	S.D.	p-value	Estimate	S.D.	p-value
c_{11}	0.0023	0.0003	<0.001	0.0026	0.0005	<0.001
c_{21}	0.0032	0.0012	0.009	0.0031	0.0011	0.004
c_{31}	0.0024	0.0004	<0.001	0.0027	0.0005	<0.001
c_{22}	0.0041	0.0007	<0.001	0.0039	0.0008	<0.001
c_{32}	-0.0002	0.0002	0.523	-0.0006	0.0003	0.028
c_{33}	0.0020	0.0003	<0.001	0.0023	0.0006	<0.001
a_{11}	0.2154	0.0157	<0.001	0.1177	0.0238	<0.001
a_{22}	0.2580	0.0425	<0.001	0.2315	0.0522	<0.001
a_{33}	0.2262	0.0220	<0.001	0.1689	0.0233	<0.001
g_{11}	-	-	-	0.2676	0.0320	<0.001
g_{22}	-	-	-	0.1843	0.0348	<0.001
g_{33}	-	-	-	0.2547	0.0514	<0.001
b_{11}	0.9669	0.0050	<0.001	0.9618	0.0104	<0.001
b_{22}	0.9446	0.0215	<0.001	0.9447	0.0217	<0.001
b_{33}	0.9634	0.0067	<0.001	0.9544	0.0125	<0.001

See Equation B.2.4.

diagnostics tests. All the leverage parameters of the DCC-GJR, ADCC-GJR, ADCC-GARCH and DBEKK models are statistically significant at the 5% level. To compare the volatility matrix estimates among the models, Figures B.1 and B.2 present the estimation of the last 1000 conditional standard deviation and the last 1000 conditional correlation of the Asymmetric DBEKK, DCC-GJR and ADCC-GJR models. In the case of the conditional standard deviations, the residuals of each series, which were evaluated as the returns after having been filtered by a VAR(1) model, are plotted in the same graphic. The DCC-GJR allows for the leverage only in the volatilities and the other two models allow for asymmetry in all elements of the volatility matrix. In general, the estimates are close, but the estimates of the DCC-GJR tend to have larger (smaller) peaks (valleys) than the values given by the other models. In general, the estimates given by the asymmetric

Tabela B.11: Results of the estimation of the DCC model to the filtered data

	DCC - GARCH			DCC - GJR		
Parameter	Estimate	S.D.	p-value	Estimate	S.D.	p-value
Ω(1)-Ibov	0.0003	0.0001	0.005	0.00001	<0.0001	0.002
ARCH(1)-Ibov	0.0803	0.0150	<0.001	0.0105	0.0066	0.115
GJR(1)-Ibov	-	-	-	0.1089	0.0257	<0.001
GARCH(1)-Ibov	0.9184	0.0157	<0.001	0.9063	0.0192	0.000
Ω(1)-Petro	0.0013	0.0005	0.006	0.00003	0.0000	0.002
ARCH(1)-Petro	0.1131	0.0235	<0.001	0.0726	0.0261	0.005
GJR(1)-Petro	-	-	-	0.0716	0.0241	0.003
GARCH(1)-Petro	0.8612	0.0347	<0.001	0.8509	0.0333	<0.001
Ω(1)-Vale	0.0006	0.0002	0.002	0.00002	0.0000	0.001
ARCH(1)-Vale	0.0901	0.0150	<0.001	0.0298	0.0103	0.004
GJR(1)-Vale	-	-	-	0.0997	0.0257	<0.001
GARCH(1)-Vale	0.9005	0.0189	<0.001	0.8796	0.0221	<0.001
$\rho_{2,1}$	0.5767	0.0287	<0.001	0.5690	0.0260	<0.001
$\rho_{3,1}$	0.7685	0.0322	<0.001	0.7638	0.0289	<0.001
$\rho_{3,2}$	0.4199	0.0268	<0.001	0.4121	0.0248	<0.001
a	0.0262	0.0051	<0.001	0.0323	0.0057	<0.001
b	0.9466	0.0145	<0.001	0.9349	0.0157	<0.001

DBEKK model, decay slower than the estimates given by the DCC-GJR model.

Tabela B.12: Results of the estimation of the ADCC model to the filtered data

	ADCC-GARCH			ADCC-GJR		
Parameter	Estimate	S.D.	p-value	Estimate	S.D.	p-value
Omega(1)-IBV	0.0003	0.0001	0.005	0.00001	<0.0001	0.002
ARCH(1)-IBV	0.0803	0.0150	<0.001	0.0105	0.0066	0.115
GJR(1)-IBV	-	-	-	0.1089	0.0257	<0.001
GARCH(1)-IBV	0.9184	0.0157	<0.001	0.9063	0.0192	<0.001
Omega(1)-Petro	0.0013	0.0005	0.006	0.00003	<0.0001	0.002
ARCH(1)-Petro	0.1131	0.0235	<0.001	0.0726	0.0261	0.005
GJR(1)-Petro	-	-	-	0.0716	0.0241	0.003
GARCH(1)-Petro	0.8612	0.0347	<0.001	0.8509	0.0333	<0.001
Omega(1)-Vale	0.0006	0.0002	0.002	0.00002	<0.0001	0.001
ARCH(1)-Vale	0.0901	0.0150	<0.001	0.0298	0.0103	0.004
GJR(1)-Vale	-	-	-	0.0997	0.0257	<0.001
GARCH(1)-Vale	0.9005	0.0189	<0.001	0.8796	0.0221	<0.001
$\rho_{2,1}$	0.5767	0.0286	<0.001	0.5690	0.0263	<0.001
$\rho_{3,1}$	0.7685	0.0322	<0.001	0.7638	0.0290	<0.001
$\rho_{3,2}$	0.4199	0.0266	<0.001	0.4121	0.0250	<0.001
$\bar{N}_{1,1}$	0.5360	0.0178	<0.001	0.5311	0.0182	<0.001
$\bar{N}_{2,1}$	0.3543	0.0160	<0.001	0.3518	0.0163	<0.001
$\bar{N}_{3,1}$	0.4359	0.0175	<0.001	0.4321	0.0177	<0.001
$\bar{N}_{2,2}$	0.4797	0.0202	<0.001	0.4820	0.0200	<0.001
$\bar{N}_{2,3}$	0.2960	0.0152	<0.001	0.2946	0.0153	<0.001
$\bar{N}_{3,3}$	0.5215	0.0190	<0.001	0.5226	0.0192	<0.001
a	0.0226	0.0086	0.009	0.0167	0.0082	0.043
g	0.0057	0.0111	0.605	0.0342	0.0114	0.003
b	0.9496	0.0174	<0.001	0.9386	0.0210	<0.001

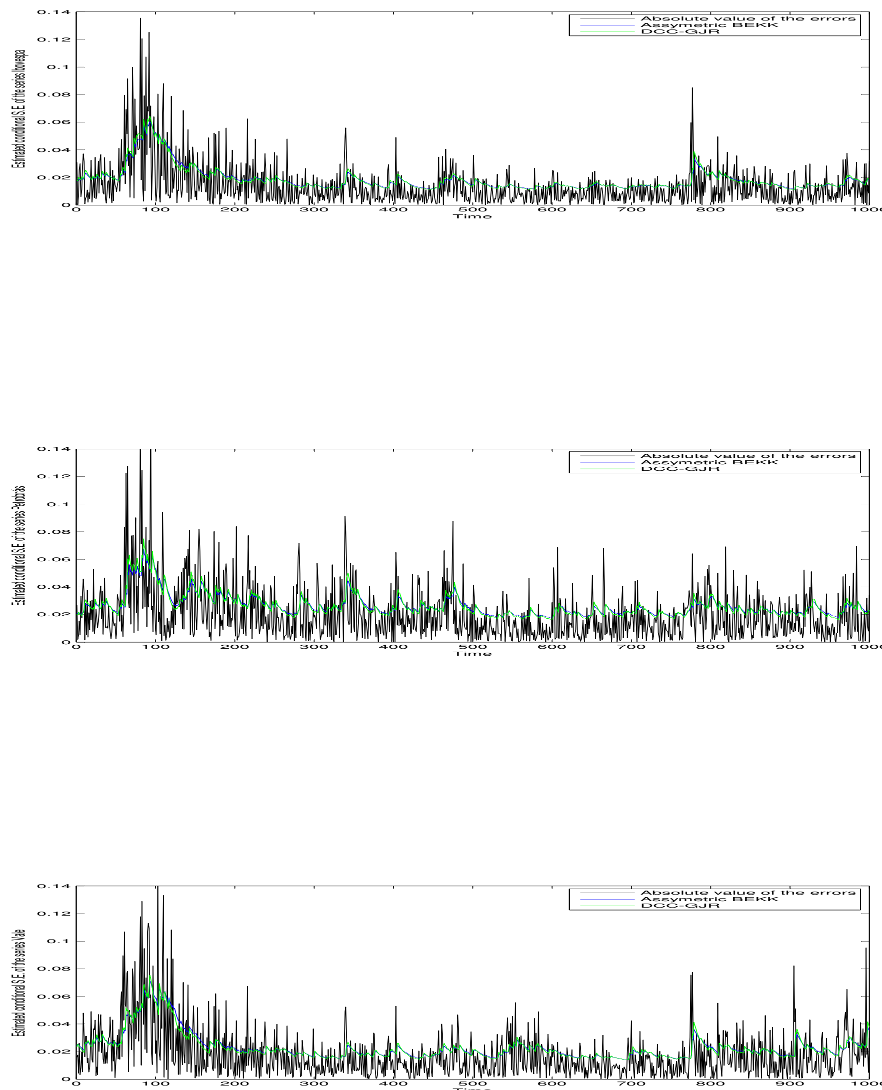

Figura B.1: Estimation of the conditional standard deviation of the example. Absolute value of the errors in black, Asymmetric DBEKK in blue, DCC-GJR in green.

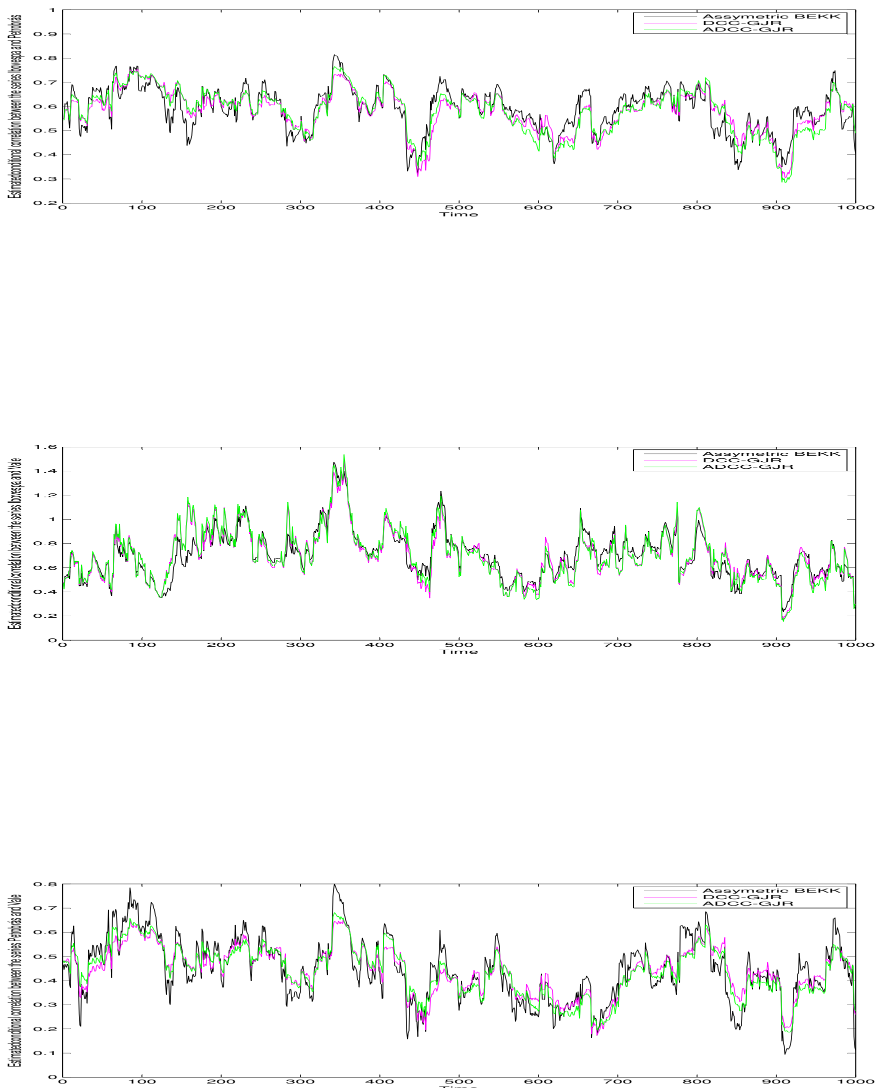

Figura B.2: Estimation of the conditional of the example. Asymmetric DBEKK in black, DCC-GJR in magenta, ADCC-GJR in green.

B.6 Concluding Remarks

In this paper we presented the main multivariate GARCH models available in the literature, with and without leverage effect, with a brief description of their features and their advantages and disadvantages of these models. A simulation study was carried out to compare some models. We used as data generating process BEKK, DBEKK, CCC, DCC and ADCC-GJR models. All these models, except the BEKK model, and including the GO-GARCH model are used to estimate the volatility. The main conclusions of this study are that the BIC performs well when the right model is simple; the AIC performs better than BIC when the true model is more complex; the estimation of models with unnecessary parameters, for instance ADCC-GARCH, when DCC-GARCH is the right model, decreased the efficiency by around 20%; none of the tested models is able to approximate the ADCC-GJR model; and the performance of the GO-GARCH model is poor when the data are generated by other models. An application to the returns of three Brazilian series (Ibovespa, Petrobras and Vale) shows that asymmetric effects are present in both volatilities and covolatilities and that the main difference among the estimated volatilities by different models are in the peaks and valleys and in the decay speed of the volatility from peak. Much progress has been made in terms of MGARCH models, but there is no statistical theory covering all of them. Although we have shed some light in this direction, there are still many issues concerning MGARCH models.

Acknowledgements

This work was partially supported by grants from São Paulo Research Foundation (FAPESP) grants 2011/02881-9, 2012/09597-7 and 2013/00506-1, and CNPq.

References

[1] Alexander, C. O. and A. M. Chibumba (1997). Multivariate Orthogonal Factor GARCH, Mimeo, University of Sussex.

[2] Ang, A. and J. Chen (2002). Asymmetric correlations of equity portfolios, *Journal of Financial Economics* **76**, 471–508.

[3] Bauwens, L., S. Laurent and J. V. K. Rombouts (2006). Multivariate GARCH models: A survey, *Journal of Applied Econometrics* **21**, 79-109.

[4] Bekaert, G. and G. Wu (2000). Asymmetric volatility and risk in equity markets, *Review of Financial Studies* **13**, 1–42.

[5] Billio, M., M. Caporin and M. Gobbo (2006). Flexible dynamic conditional correlation multivariate GARCH models for asset allocation, *Applied Financial Economics Letters* **2**, 123–130.

[6] Bollerslev, T. (1990). Modeling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model, *Review of Economics and Statistics* **72**, 498–505.

[7] Bollerslev, T. and R.F. Engle (1993). Common persistence in conditional variances, *Econometrica* **61**, 167–186.

[8] Bollerslev, T., R. F. Engle and J. M. Wooldridge (1988). A capital asset pricing model with time varying covariances, *Journal of Political Economy* **96**, 116–131.

[9] Bollerslev, T., J. Litvinova and G. Tauchen (2006). Leverage and volatility feedback effects in high-frequency data, *Journal of Financial Econometrics* **4**, 353–384.

[10] Braun, P., D. Nelson and A. Sunier (1995). Good news, bad news, volatility, and betas, *Journal of Finance* **50**, 1575-1603.

[11] Cappiello, L., R. F. Engle and K. Sheppard (2006). Asymmetric dynamics in the correlations of global equity and bond returns, *Journal of Financial Econometrics* **4**, 537–572.

[12] Christiansen, C. (2005). Multivariate term structure models with level and heteroskedasticity effect, *Journal of Banking and Finance* **29**, 1037–1057.

[13] de Goeij, P.D. and W. Marquering (2004). Modeling the conditional covariance between stock and bond returns: A multivariate GARCH approach, *Journal of Financial Econometrics* **16**, 531–564.

[14] de Goeij, P.D. and W. Marquering (2009). Stock and bond market interactions with level and asymmetry dynamics: An out-of-sample application, *Journal of Empirical Finance* **2**, 318–329.

[15] De Santis, G. and B. Gerard (1997). International Asset Pricing and Portfolio Diversification with Time-Varying Risk, *The Journal of Finance* **52**, 1881-1912.

[16] Diebold, F. X. and M. Nerlove (1989). The Dynamics of exchange rate volatility: a multivariate latent factor ARCH model, *Journal of Applied Econometrics* **4**, 1–21.

[17] Diebold, F. X. and K. Yilmaz (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets, *The Economic Journal* **119**, 158–171.

[18] Ding, Z. and Engle, R. F. (2001). Large scale conditional covariance matrix modeling, estimation and testing, *NYU Working Paper No. FIN-01-029*

[19] Engle, R. F. (2002). Dynamic conditional correlation - a simple class of multivariate GARCH models, *Journal of Business and Economic Statistics* **20**,339-350.

[20] Engle, R. F. and F. K. Kroner (1995). Multivariate simultaneous generalized ARCH, *Econometric Theory* **11**, 122–150.

[21] Engle, R.F. and K. Sheppard (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH, *UCSD Discussion Paper* **2001-15**.

[22] Engle R. F., V. K. Ng and M. Rothschild (1990). Asset pricing with a factor ARCH covariance structure: empirical estimates for treasury bills, *Journal of Econometrics* **45**, 213–238.

[23] Francq, C., and J. M. Zakoïan, (2012). Strict stationarity testing and estimation of explosive and stationary generalized autoregressive conditional heteroscedasticity models, *Econometrica* **80**, 821–861.

[24] Franke J., Wolfgang, H. and Christian, H. (2004). *Introduction to Statistics of Financial Markets*, e-book. http://sfb649.wiwi.hu-berlin.de/fedchomepage/xplore/ebooks/html/sfe/

[25] Franke J., Wolfgang, H. and Christian, H. (2008). *Introduction to Statistics of Financial Markets*, Springer.

[26] Gourieroux, C. (1997) *ARCH Models and Financial Applications*, Springer-Verlag, New York.

[27] Hafner C. M. (2003). Fourth Moment Structure of Multivariate GARCH Models, *Journal of Financial Econometrics, Society for Financial Econometrics* **1**, 26–54.

[28] Hafner, C.M. and Herwartz, H. (1998). Structural Analysis of Portfolio Risk using Beta Impulse Response Functions, *Statistica Neerlandica* **52**, 336–355.

[29] Hafner, C., D. van Dijk and P. H. Franses (2006). Semi-Parametric Modelling of Correlation Dynamics. In: *Econometric Analysis of Financial and Economic Time Series*, D. Terrell and T. B. Fomby (eds.), (Advances in Econometrics, Volume 20), 59-103, Emerald Group Publishing Ltda.

[30] He, C. and T. Teräsvirta (2004). An extended constant conditional correlation GARCH model and its fourth-moment structure, *Econometric Theory* **20(5)**, 904–926.

[31] Jeantheau, T. (1998). Strong consistency of estimators for multivariate ARCH models, *Econometric Theory* **14**, 70–86.

[32] Kawakatsu, H. (2003). Cholesky factor GARCH. Mimeo, Quantitative Micro Software, Irvine, CA.

[33] Kawakatsu, H. (2006). Matrix exponential GARCH, *Journal of Econometrics* **134**, 95–128.

[34] Koutmos, G. and G. G. Booth (1995). Asymmetric volatility transmission in international stock markets, *Journal of international Money and Finance* **14**, 747–762.

[35] Kroner and Ng (1998). Modeling asymmetric comovements of asset returns, *Review of Financial Studies* **11**, 817–844.

[36] Lien, D. and Y.K. Tse (2002). Some recent developments in futures hedging, *Journal of Economic Surveys* **16**, 357–396.

[37] Lin, W.L. (1992). Alternative estimators for factor GARCH models - a Monte Carlo comparison, *Journal of Applied Econometrics* **7**, 259–279.

[38] Linton, O. B. (2008). Semiparametric and nonparametric ARCH modelling, *Handbook of Financial Times Series*, T.G. Andersen, R.A. Davis, J-P. Kreiss and T. Mikosch (eds.), 157–167. New York: Springer.

[39] Longin, F. and B. Solnik (1995). Is the correlation in international equity returns constant, *Journal of international money and finance* **14**, 3–26.

[40] McAleer, M., S. Hoti and F. Chan (2009). Structure and asymptotic theory for multivariate asymmetric conditional volatility, *Econometric Reviews* 28, 422–440.

[41] Nakatani, T. and T. Teräsvirta (2009). Testing for volatility interactions in the constant conditional correlation GARCH model, *The Econometrics Journal* **12**, 147–163.

[42] Nelson, D.B. (1991). Conditional heteroskedasticity in asset returns: A new approach, *Econometrica* **59**, 347–370.

[43] Rodríguez, M. J. and E. Ruiz (2012). Revisiting several popular GARCH models with leverage effect: differences and similarities, *Journal of Financial Econometrics* **10(4)**, 637–668.

[44] Rossi, E. and F. Spazzini (2009). Model and distribution uncertainty in multivariate GARCH estimation: A Monte Carlo analysis, http://mpra.ub.uni-muenchen.de/12260/ MPRA Paper No. 12260.

[45] Santos, A. A. P., E. Ruiz and F. J. Nogales (2013). Comparing univariate and multivariate models to forecast portfolio value-at-risk, *Journal of Financial Econometrics*, **11**, 400–441.

[46] Sentana, E. (1998). The relation between conditionally heteroskedastic factor models and factor GARCH models, *The Econometrics Journal* **1**, 1–9.

[47] Silvennoinen, A. and T. Teräsvirta (2005). Multivariate Autoregressive Conditional Heteroskedasticity with Smooth Transitions in Conditional Correlations. SSE/EFI Working Paper Series in Economics and Finance No. 577.

[48] Silvennoinen, A. and T. Teräsvirta (2009a). Modeling multivariate autoregressive conditional heteroskedasticity with the double smooth transition conditional correlation GARCH model, *Journal of Financial Econometrics* **7**, 373–411.

[49] Silvennoinen, A. and T. Teräsvirta (2009b). Multivariate GARCH models. In: *Handbook of Financial Times Series*, T.G. Andersen, R.A. Davis, J-P. Kreiss and T. Mikosch (eds.), 201–232. New York: Springer.

[50] Tsay, R. S. (2002). *Analysis of Financial Time Series*. New York: John Wiley.

[51] Tse, Y.K. and A. K. C. Tsui (2002). A multivariate GARCH model with time-varying correlations, *Journal of Business and Economic Statistics* **20**, 351–362.

[52] van der Weide, R. (2002). GO-GARCH: a multivariate generalized orthogonal GARCH model, *Journal of Applied Econometrics* **17**, 549–564.

[53] Vrontos, I.D., P. Dellaportas, D.N. Politis (2003). A full-factor multivariate GARCH model, *Econometrics Journal* **6**, 311–333.

[54] Yoldas, E. (2012). Threshold asymmetries in equity return distributions: statistical tests and investment implications, *Studies in Nonlinear Dynamics and Econometrics* **16**, doi:10.1515/1558-3708.1799.

[55] Zakoïan, J. M. (1994). Threshold heteroskedastic models, *Journal of Economic Dynamics and control* **18**, 931–955.