

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO DE CIÊNCIAS JURÍDICAS E ECONÔMICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

PATRICK FERREIRA PATROCINIO

ROBUSTEZ EM PROCESSOS HETEROCEDÁSTICOS
CONTAMINADOS POR OUTLIERS ADITIVOS: UMA
ABORDAGEM M -QUANTILE

VITÓRIA
2021

PATRICK FERREIRA PATROCINIO

**ROBUSTEZ EM PROCESSOS HETEROCEDÁSTICOS
CONTAMINADOS POR OUTLIERS ADITIVOS: UMA
ABORDAGEM M -QUANTILE**

Dissertação apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal do Espírito Santo como requisito parcial para a obtenção do título de Mestre em Economia.

Área de Concentração: Teoria Econômica.

Orientador: Prof. Dr. Valdério Anselmo Reisen

Co-Orientador: Prof. Dr. Edson Zambom Monte

VITÓRIA
2021

Ficha catalográfica disponibilizada pelo Sistema Integrado de
Bibliotecas - SIBI/UFES e elaborada pelo autor

P314r Patrocinio, Patrick Ferreira, 1994-
Robustez em processos heterocedásticos contaminados por
outliers aditivos : uma abordagem M-Quantile / Patrick Ferreira
Patrocinio. - 2021.
38 f. : il.

Orientador: Valdério Anselmo Reisen.
Coorientador: Edson Zambon Monte.
Tese (Mestrado em Economia) - Universidade Federal do
Espírito Santo, Centro de Ciências Jurídicas e Econômicas.

1. Análise matemática. 2. Análise de séries temporais. 3.
Análise econômico-financeira. I. Reisen, Valdério Anselmo. II.
Monte, Edson Zambon. III. Universidade Federal do Espírito
Santo. Centro de Ciências Jurídicas e Econômicas. IV. Título.

CDU: 330

PATRICK FERREIRA PATROCINIO

**ROBUSTEZ EM PROCESSOS HETEROCEDÁSTICOS CONTAMINADOS POR
OUTLIERS ADITIVOS: UMA ABORDAGEM M-QUANTILE**

Dissertação apresentada ao Programa de Pós-graduação em Economia da Universidade Federal do Espírito Santo como requisito para a obtenção do título de Mestre em Economia.

Vitória, 26 de maio de 2021.

COMISSÃO EXAMINADORA

Prof. Dr. Valdério Anselmo Reisen
Universidade Federal do Espírito - Orientador

Prof. Dr. Edson Zambon Monte
Universidade Federal do Espírito Santo

Prof. Dr. Pascal Bondon
Centrale-Supélec – França

Prof^a. Dr^a. Verónica Andrea González-López
Universidade Estadual de Campinas



UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO

PROTOCOLO DE ASSINATURA



O documento acima foi assinado digitalmente com senha eletrônica através do Protocolo Web, conforme Portaria UFES nº 1.269 de 30/08/2018, por
EDSON ZAMBON MONTE - SIAPE 2650818
Departamento de Economia - DE/CCJE
Em 26/05/2021 às 13:00

Para verificar as assinaturas e visualizar o documento original acesse o link:
<https://api.lepisma.ufes.br/arquivos-assinados/197722?tipoArquivo=O>



UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO

PROTOCOLO DE ASSINATURA



O documento acima foi assinado digitalmente com senha eletrônica através do Protocolo Web, conforme Portaria UFES nº 1.269 de 30/08/2018, por
VALDERIO ANSELMO REISEN - SIAPE 297617
Departamento de Estatística - DE/CCE
Em 26/05/2021 às 15:13

Para verificar as assinaturas e visualizar o documento original acesse o link:
<https://api.lepisma.ufes.br/arquivos-assinados/197882?tipoArquivo=O>

AGRADECIMENTOS

Sem as intervenções divinas, meus projetos não teriam começo e muito menos fim, então à DEUS ofereço meu humilde obrigado.

À minha Mãe, que nunca deixou de me apoiar, em todas as áreas da minha vida.

Obrigado à minha avó Enadite, que nunca passou pela porta de uma universidade, mas me ensinou o que há de mais refinado em filosofia de vida.

Obrigado à minha companheira Merari Martins, que me apoiou incondicionalmente durante todo esse projeto, mesmo quando minha ausência física foi grande.

Ao Professor Valdério, pela humildade, pela forma poética de explicar, por aumentar exponencialmente minha relação com a pesquisa e por ser uma pessoa que sempre lembrarei (com probabilidade 1).

Ao professor Edson, pela (enorme) paciência, por nunca me deixar perder o foco e por toda a disponibilidade (mesmo com a agenda apertada).

Á Alexandra Elbakyan, por me proporcionar acesso a uma rica leitura, que fomentou a base teórica desse projeto.

RESUMO

Séries temporais financeiras possuem características que as distinguem de outras séries existentes, comumente conhecidas como fatos estilizados de séries financeiras. Uma dessas características é a presença de outliers na série o que dificulta sua modelagem e previsão. Adicionalmente, essas séries possuem substancial sensibilidade a decisões políticas, sociais e econômicas, gerando um aumento em sua volatilidade. Esta pesquisa tem o objetivo de apresentar o estimador de regressão quantílica M , discutido em Breckling e Chambers (1988), como abordagem alternativa para estimar os parâmetros do processo Autorregressivo Generalizado com Variância Heterocestática Condicional (GARCH), amplamente utilizado para modelar e prever séries temporais financeiras. É bem conhecido que o estimador de regressão quantílica M é consistente na presença de outliers aditivos, ou seja, possui a propriedade de robustez, e recentemente foi amplamente utilizado para estimar séries temporais lineares com diferentes estruturas de correlação, seja no domínio do tempo ou da frequência. Simulações serão realizadas para verificar o desempenho do método para pequenos tamanhos de amostra. Séries temporais com variância dependente do tempo (volatilidade) têm sido estudadas por vários autores em diferentes áreas da conhecimento. Aqui, em particular, a motivação do estudo proposto em um problema real é modelar e prever variáveis da área financeira, com atenção especial às variáveis de retorno de ativos que, em geral, não preservam a propriedade de variância constante ao longo do tempo.

Palavras Chave: Análise de séries temporais, outliers aditivos, GARCH, M -Quantile, regressão quantílica, parâmetros, robustez.

ABSTRACT

Financial time series have characteristics that distinguish them from other existing series, commonly known as stylized financial series facts. One of these characteristics is the presence of outliers in the series, which makes its modeling and forecasting difficult. Additionally, these series have substantial sensitivity to political, social and economic decisions, generating an increase in their volatility. This research has the goal to introduce the M quantile regression estimator, discussed in Breckling e Chambers (1988), as alternative approach to estimate the parameters of the Generalized Autoregressive with conditional heterocestatic variance (GARCH) process. It is well-know that the M -regression estimator is strong against additive outliers, that is, it has the robustness property, and it has recently been widely used to estimate linear time series with different correlation structures, either in the time or frequency-domain. Simulation will be carried out to verify the performance of the method for small sample sizes. Time series with time-dependent variance (volatility) have been studied by several authors in different areas of applications. Here, in particular, the motivation of the proposed study in a real-problem is to modelling and forecasting variables from financial data area, with special attention to assets returns variables which, in general, do not preserve the property of constant variance over time.

Keywords: Time series analysis, additive outliers, GARCH, M -Quantile, quantile regression, parameters, robustness.

SUMÁRIO

1	INTRODUÇÃO	10
1.1	EFEITOS DE OUTLIERS ADITIVOS NO MODELO ARCH(q) . . .	12
2	PROBLEMA	14
2.1	OBJETIVO GERAL	15
2.2	OBJETIVOS ESPECÍFICOS	15
3	CONCEITOS ESTATÍSTICOS	15
3.1	CONCEITOS BÁSICOS	15
3.2	O MODELO GARCH	16
4	MEDIDAS DE DESEMPENHO	17
5	MATERIAIS E MÉTODOS	18
5.1	DADOS	18
5.2	SOFTWARE ESTATÍSTICO	18
6	RESULTADOS	18
7	CONCLUSÃO GERAL	35

1 INTRODUÇÃO

A análise de séries temporais financeiras é amplamente utilizada para estudar dados coletados no mercado financeiro no tempo. Por exemplo, os retornos de ativos no tempo t , geralmente calculados como $r_t = \log(p_t) - \log(p_{t-1})$, onde p_t é o preço do ativo financeiro no tempo t . Devido à alta instabilidade que enfrenta o mercado financeiro por conta de decisões econômicas, políticas e sociais, retornos diários de ativos financeiros oscilam fortemente. Portanto, defini-se esta variabilidade a volatilidade dos retornos, que não é diretamente observável. Com o propósito de modelar a volatilidade defini-se a variância condicional como medida estatística da mesma, com o objetivo de obter informações precisas que possam ajudar na previsão do comportamento futuro dos retornos do mercado financeiro.

Desde a crise de 2008, o interesse pelo cálculo da volatilidade e seu respectivo uso na administração de riscos financeiros tem recebido crescente aumento entre analistas e acadêmicos. A literatura oferece uma rica fonte de métodos para o cálculo da volatilidade, com destaque para os métodos determinísticos, não determinísticos e de redes neurais.

Uma característica, ou fato estilizado, da volatilidade é que ela tende a apresentar-se em grupos numa série financeira, formando *clusters*, levando-se a notar um certo grau de dependência no tempo, de modo que é possível recorrer-se aos chamados modelos de heterocedasticidade condicionais, métodos determinísticos que distinguem a variância condicional, que é variável no tempo e condicionada às informações disponíveis até o dado instantâneo e a variância incondicional, que é constante em uma série observada. Esses modelos e suas extensões são amplamente empregados em econometria e finanças, pois possuem propriedades que conseguem explicar alguns fatos estilizados ou inerentes de séries temporais financeiras.

Um dos pioneiros e mais famosos estudos dessa classe de modelos foi Engle (1982), que propôs modelar e prever a variância condicional com os modelos Autorregressivos de Heterocedasticidade Condicional (ARCH), introduzidos nesse estudo. Uma generalização dos modelos ARCH foi proposta por Bollerslev (1986), nomeados modelos Generalizados Autorregressivos de Heterocedasticidade Condicional (GARCH), que são mais parcimoniosos no sentido de previsão do que os modelos ARCH, e conseqüentemente, são mais empregados na literatura eco-

nométrica e financeira.

Muitos métodos usados para estimar a volatilidade são baseados na suposição de que os retornos possuem distribuição Gaussiana. Essa suposição, em geral, torna o modelo menos flexível, no sentido de que não serão capazes de capturar adequadamente a assimetria e as caudas pesadas que são características de séries temporais financeiras. A literatura oferece uma rica fonte de métodos paramétricos desse tipo. No entanto, conforme Koenker e Zhao (1996), não raramente, a hipótese de normalidade da distribuição é violada, gerando discrepâncias na volatilidade estimada. Recentemente, o método de regressão quantílica tem recebido crescente atenção entre pesquisadores de diversas áreas, em particular econometria e finanças, vide Koenker e Zhao (1996), Koenker e Xiao (2009) e Lee e Noh (2013). A aplicação do método de regressão quantílica possui vantagens substanciais, como: (i) fornece informações sutis sobre a distribuição de erros. A regressão quantílica ajuda na modelagem de distribuições condicionais, e o quantil condicional estimado pode ser aplicado ao *Value-at-Risk* (VaR) condicional; (ii) uma vez que os parâmetros de um modelo GARCH(p,q) é estimado, pode-se realizar estudos mais refinados, como, por exemplo, uma análise de assimetria na série; (iii) em casos de distribuições condicionais com caudas pesadas, a regressão quantílica é substancialmente útil na realização de uma estimativa robusta.

Outra característica de séries temporais financeiras é que os retornos exibem realizações atípicas (outliers), o que pode afetar adversamente o desempenho da estimativa e inferência projetada para as condições gaussianas. Isso é particularmente verdadeiro para os modelos ARCH e GARCH estimados por Mínimos Quadrados Ordinários (MQO) ou Máxima Verossimilhança (MV), que são muito sensíveis a outliers. Conforme Carnero, Peña e Ruiz (2005), quando modelos heterocedásticos são ajustados em séries temporais reais, os resíduos frequentemente exibem excesso de curtose, que pode ser explicado, entre outros motivos, pela presença de outliers.

Desde a década de 1970, tem sido dada atenção ao estudo de métodos no contexto de processos que contém outliers. Vários autores mostraram que a presença de outliers pode deteriorar dramaticamente a modelagem de séries temporais. Por exemplo, no padrão de estruturas de modelos Box-Jenkins, Ledolter (1989) mostrou que as previsões de intervalo em modelos Autorregressivos Integrados de Médias

Móveis (ARIMA) são consideravelmente sensíveis a outliers aditivos. Chang, Tiao e Chung Chen (1988) e C. Chen e Liu (1993) mostraram que os parâmetros estimados do modelo ARMA tornam-se mais tendenciosos quando os dados contém outliers. Os impactos de outliers aditivos em modelos heterocedásticos foi estudado em Carnero, Peña e Ruiz (2005), entre outros autores.

Como é bem discutido na literatura, os estimadores derivados do método M de regressão, de Huber (1964), são abordagens alternativas robustas para obter estimativas dos parâmetros do modelo, mesmo em séries contaminadas por outliers ou geradas por distribuições de probabilidade com caudas pesadas, veja por exemplo Bai, Rao e Y. H. Wu (1997), W. B. Wu (2007) e Li (2007). O uso do método de regressão M em processos de memórias curta e longa estão bem documentados nos artigos de Reisen e Fajardo (2012), Fajardo et al. (2016), entre outros.

O presente estudo tem o objetivo discutir os efeitos de outliers aditivos em modelos heterocedásticos e propor um estimador que seja robusto, com especial atenção ao modelo GARCH(p,q). Como é bem conhecido na literatura, outliers aditivos geram grandes impactos em modelos heterocedásticos, razão pela qual esse tipo de outlier é considerado nessa pesquisa.

Os efeitos dos outliers aditivos nos parâmetros do modelo GARCH(p,q) foi estudado pelos autores Franses e Ghijssels (1999), Mendes (2000), Carnero (2003), Carnero, Peña e Ruiz (2005), entre outros. Os autores supracitados mostraram que outliers aditivos podem distorcer substancialmente a estimativa dos parâmetros dos modelos ARCH(q) e GARCH(p,q). Com o objetivo de discutir e exemplificar esses efeitos, a seção seguinte apresenta o modelo ARCH(q) estimado pelo método de MQO na presença de outliers aditivos.

1.1 EFEITOS DE OUTLIERS ADITIVOS NO MODELO ARCH(q)

Seja uma sequência $(\epsilon_t)_{t \in \mathbb{Z}}$ estocástica, com $E(\epsilon_t) = 0$ e $E(\epsilon_t^4) < \infty$. O processo $(\epsilon_t)_{t \in \mathbb{Z}}$ é dito pertencer a classe de modelos Autorregressivos com Heterocedasticidade Condicional, de ordem $q \in \mathbb{N}^+$, isto é, $(\epsilon_t)_{t \in \mathbb{Z}}$ é um modelo ARCH(q), se satisfizer a seguinte equação

$$\begin{cases} \epsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2, \end{cases} \quad (1)$$

onde $\alpha_i, i = 1, \dots, q$ são constantes não negativas, α_0 é uma constante estritamente positiva, $\sigma_t > 0$ e η_t é uma sequência com média nula e variância unitária.

Seja $(\epsilon_t)_{t \in \mathbb{Z}}$ uma série estacionária, com $E(\epsilon_t^4) < \infty$, contaminada por k outliers, de magnitude ω no tempo $\tau_i, i = 1, \dots, k$. Então, a série observada é representada por

$$z_t = \begin{cases} \epsilon_t + \omega, \text{ se } t = \tau_i, i = 1, \dots, k \\ \epsilon_t, \text{ caso contrário} \end{cases} \quad (2)$$

A autocorrelação do quadrado das observações da série contaminada na equação (2) é estimada por

$$r(h) = \frac{\sum_{t=h+1}^T z_t^2 z_{t-h}^2 - \frac{T-h}{T^2} (\sum_{t=1}^T z_t^2)^2}{\sum_{t=1}^T z_t^4 - T^{-1} (\sum_{t=1}^T z_t^2)^2}, \quad (3)$$

com $h > 0$ a ordem de $r(h)$ e T o tamanho da amostra. Então

$$\lim_{\omega \rightarrow \infty} r(h) = \begin{cases} 1 - \frac{h}{k(1-\frac{k}{T})}, \text{ se } h < k \\ \frac{k}{k-T}, \text{ se } h \geq k \end{cases}. \quad (4)$$

Portanto, é possível inferir que para um único outlier ($k = 1$), a função $r(h)$ sempre tende para zero quando $\omega \rightarrow \infty$, enquanto que para um grande conjunto de k outliers a função $r(h)$ é sempre positiva quando $h < k$ e zero caso contrário. Adicionalmente, para $h < k$ e T suficientemente grande, as autocorrelações decaem linearmente na presença de outliers. Assim, se um processo heterocedástico é contaminado por um único outlier, a detecção de heterocedasticidade torna-se difícil. Por outro lado, quando um processo homocedástico é contaminado por vários outliers, as autocorrelações positivas geradas pelos outliers podem ser confundidas com heterocedasticidade condicional.

Seja $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_q)$ o vetor de parâmetros do modelo ARCH(q) expresso em (1). Seja $\hat{\boldsymbol{\alpha}}$ um estimador de $\boldsymbol{\alpha}$. Se $\boldsymbol{\alpha}$ é estimado pelo método de MQO, então

na presença de um único outlier aditivo, tem-se

$$\begin{aligned} \lim_{\omega \rightarrow \infty} \hat{\boldsymbol{\alpha}}^{MQO} &= \lim_{\omega \rightarrow \infty} \frac{1}{(T-2q)\omega^{4q}} \begin{pmatrix} \omega^{4q+2} + o(\omega^{4q+2}) \\ (-\omega^{4q} + o(\omega^{4q}))\mathbf{1} \end{pmatrix} \\ &= \begin{pmatrix} \infty \\ -\frac{1}{(T-2q)}\mathbf{1} \end{pmatrix}, \end{aligned} \quad (5)$$

onde $o(\cdot)$ denota a ordem de convergência de uma sequência e $\mathbf{1}$ um vetor de um 's. A equação (5) sugere que, se o tamanho da amostra é suficientemente grande e a magnitude do outlier tende para o infinito, todos os parâmetros estimados tendem para zero. A equação seguinte mostra o efeito de vários outliers em cada parâmetro do modelo.

$$\lim_{\omega \rightarrow \infty} \hat{\boldsymbol{\alpha}}_i^{MQO} = \begin{cases} \infty, & \text{para } i = 0 \\ \frac{-2k^2 + (2k-q)(T-q)}{-2k^2 + (2k-q+1)(T-q)}, & \text{para } i = 1 \\ 0 & \text{para } i = 2, \dots, q-1 \\ \frac{-(T-q)}{-2k^2 + (2k-q+1)(T-q)}, & \text{para } i = q \end{cases}. \quad (6)$$

A equação (6) sugere os parâmetros, $\hat{\alpha}_i$, tendem para zero, exceto $\hat{\alpha}_1$ e $\hat{\alpha}_q$. Se o número de outliers é grande em relação a ordem do modelo, então $\hat{\alpha}_1$ tende para um valor próximo de 1 e $\hat{\alpha}_q$ tende para zero.

Os efeitos de outliers aditivos no modelo ARCH(q) estimado por outros métodos, como por exemplo Mínimos Quadrados Generalizados, Máxima Verossimilhança e Quasi Máxima Verossimilhança são discutidos em Carnero (2003), Carnero, Peña e Ruiz (2005), entre outros. Como esses métodos não possuem solução fechada, os autores utilizaram resultados de simulação para a discussão.

2 PROBLEMA

O modelo GARCH(p, q) é extensamente utilizados no contexto financeiro. Franses e Ghijssels (1999), Mendes (2000) e Carnero, Peña e Ruiz (2005) mostraram que os parâmetros do modelo GARCH(p, q) estimado pelos métodos usuais são extremamente sensíveis a outliers aditivos e podem ser substancialmente dis-

torcidos nessa situação. De modo a mitigar esse problema faz-se necessário o uso de um estimador robusto no modelo GARCH(p,q). Entretanto, pouca atenção foi dada aos modelos GARCH(p,q) estimados por métodos robustos.

2.1 OBJETIVO GERAL

Propor um estimador robusto a outliers aditivos para os parâmetros do modelo GARCH(p,q) e aplica-lo em problemas da área de econometria e finanças.

2.2 OBJETIVOS ESPECÍFICOS

1. Verificar o impacto de outliers aditivos nos parâmetros modelo GARCH(p,q), estimado pelo método de Quasi Máxima Verossimilhança (QMV);
2. Realizar estudos simulados dos parâmetros do modelo GARCH utilizando o estimador proposto baseado em M regressão quantílica;
3. Aplicar o modelo GARCH(p,q) robusto para analisar os retornos de dois índices financeiros.

3 CONCEITOS ESTATÍSTICOS

Nesta seção são apresentados de forma sumarizada os conceitos estatísticos e notações utilizadas no decorrer dessa dissertação. Para mais detalhes sobre os conceitos básicos e processos heterocedásticos, vide Francq e Zakoian (2019).

3.1 CONCEITOS BÁSICOS

Definição 1. Um vetor aleatório é um vetor cujos elementos são variáveis aleatórias. Se existirem X_1, \dots, X_p variáveis aleatórias em um vetor \mathbf{X} , então o vetor \mathbf{X} é um vetor aleatório denotado por

$$\begin{pmatrix} X_1 \\ \vdots \\ X_p \end{pmatrix}. \quad (7)$$

O vetor transposto do vetor aleatório \mathbf{X} é denotado por

$$\mathbf{X}' = (X_1, \dots, X_p). \quad (8)$$

Definição 2. Seja \mathbf{X} um vetor aleatório. O vetor $\boldsymbol{\mu} = E(\mathbf{X})$ é chamado de vetor de médias do vetor $\mathbf{X}' = (X_1, \dots, X_p)$ definido por

$$E(\mathbf{X}) = \begin{pmatrix} E(X_1) \\ \vdots \\ E(X_p) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix} = \boldsymbol{\mu}, \quad (9)$$

onde $E(X_i), i = 1, \dots, p$ é a esperança da variável $X_i, i = 1, \dots, p$.

Definição 3. A variância do i -ésimo componente do vetor \mathbf{X} é denotado por $VAR(X_i) = \sigma_i = \sigma_{ii}, i = 1, \dots, p$ e sua representação matricial é:

$$\begin{aligned} \Sigma &= E(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})' \\ &= E \begin{pmatrix} (X_1 - \mu_1) \\ \vdots \\ (X_p - \mu_p) \end{pmatrix} \begin{pmatrix} (X_1 - \mu_1, \dots, X_p - \mu_p) \end{pmatrix} \quad (10) \\ &= \begin{pmatrix} E(X_1 - \mu_1)^2 & E(X_1 - \mu_1)(X_2 - \mu_2) & \cdots & E(X_1 - \mu_1)(X_p - \mu_p) \\ \vdots & \vdots & \ddots & \vdots \\ E(X_p - \mu_p)(X_1 - \mu_1) & E(X_p - \mu_p)(X_2 - \mu_2) & \cdots & E(X_p - \mu_p)^2 \end{pmatrix} \end{aligned}$$

3.2 O MODELO GARCH

O modelo GARCH(p,q) é um importante modelo da teoria de séries temporais para o cálculo da volatilidade de dados. Nessa seção, o modelo GARCH é apresentado de forma sumarizada. Mais detalhes podem ser encontrados na parte I do livro de Francq e Zakoian (2019).

Seja uma sequência $(\epsilon_t)_{t \in \mathbb{Z}}$ estocástica, com $E(\epsilon_t) = 0$ e $E(\epsilon_t^4) < \infty$. O processo $(\epsilon_t)_{t \in \mathbb{Z}}$ é dito pertencer a classe de modelos Autorregressivos Generalizados com Heterocedasticidade Condicional, de ordem $p, q \in \mathbb{N}^+$, isto é, $(\epsilon_t)_{t \in \mathbb{Z}}$ é um modelo GARCH(p,q), se satisfizer a seguinte equação

$$\begin{cases} \epsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2, \end{cases} \quad (11)$$

onde α_i, β_j são constantes não negativas, α_0 é uma constante estritamente positiva e $\sigma_t > 0$. Adicionalmente $(\eta_t)_{t \in \mathbb{Z}}$ é uma sequência estocástica com média zero e variância unitária. Se os parâmetros do modelo GARCH(p,q) satisfazem $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1$, então o processo é estritamente estacionário e tem solução única. Veja Teorema 2.5 em Francq e Zakoian (2019). O modelo (11) foi introduzido na literatura por Bollerslev (1986) e desde então tem sido amplamente utilizado na literatura econométrica, financeira, entre outras. Por exemplo, no contexto da área de poluição do ar, pode-se citar Monte, Albuquerque e Reisen (2015) e Reisen, Sarnaglia et al. (2014).

4 MEDIDAS DE DESEMPENHO

As medidas de Erro Quadrático Médio (EQM), vício e desvio padrão são importantes para avaliar a qualidade e performance de um estimador.

Definição 4. O erro quadrático médio de um estimador $\hat{\theta}$ do parâmetro θ é dado por:

$$EQM(\hat{\theta}) = E(\hat{\theta} - \theta)^2. \quad (12)$$

Definição 5. Um estimador $\hat{\theta}$ é dito ser não viciado para θ se:

$$E(\hat{\theta}) = \theta \quad (13)$$

Então, tem-se que

$$EQM(\hat{\theta}) = VAR(\hat{\theta}) \quad (14)$$

Definição 6. O desvio padrão de um estimador $\hat{\theta}$ de θ é definido como

$$\sigma = \sqrt{VAR(\hat{\theta})} \quad (15)$$

5 MATERIAIS E MÉTODOS

5.1 DADOS

O presente trabalho foi realizado com os retornos de dois índices financeiros Coreanos, o Índice Composto de Preços de Ações da Coreia (KOSPI) e o índice de Cotação Corretores de Valores Mobiliários Coreanos Automatizados (KOSDAQ). As observações de cada série está entre o período de 2 de janeiro de 2004 e 29 de agosto de 2008. A escolha desse conjunto de dados foi motivada pelo fato de tornar possível uma comparação com entre o método porposto neste trabalho e o método porposto em Lee e Noh (2013).

5.2 SOFTWARE ESTATÍSTICO

A metodologia proposta e toda análise efetuada foi realizada por meio do Software R. O R possui um grande número de procedimentos estatísticos convencionais, entre eles estão os modelos lineares, modelos de regressão não linear, análise de séries temporais, testes estatísticos paramétricos e não paramétricos, análise multivariada, entre outros. Possui uma grande quantidade de funções para o desenvolvimento de ambiente gráfico e criação de diversos tipos de apresentação de dados, vide Reisen e Silva (2011).

6 RESULTADOS

Anexado ao presente trabalho econtra-se o artigo resultado desta dissertação.

Heteroscedastic Processes: An M-Quantile Approach

Patrick F Patrocínio¹, Valdério A Reisen^{1, 2, 3, 4}, Edson Z Monte ¹,
Ian M Danilevicz^{2, 3}e Pascal Bondon³

¹Federal University of Espírito Santo, Graduate Program in
Economics, Brazil.

²Federal University of Minas Gerais, Graduate Program in Statistics,
Brazil.

³University Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des
signaux et systèmes, 91190, Gif-sur-Yvette, France.

⁴Federal University of Espírito Santo, Department of Environmental
Engineering, Brazil.

June 20, 2021

Abstract

This research has the goal to introduce the M-regression estimator as alternative approach to estimate the parameters of the Generalized Autoregressive with conditional heteroscedastic variance (GARCH) process. It is well-known that the M-regression estimator is strong against additive outliers, that is, it has the robustness property, and it has recently been widely used to estimate linear time series with different correlation structures, either in the time or frequency-domain. Simulation will be carried out to verify the performance of the method for small sample sizes. Time series with time-dependent variance (volatility) have been studied by several authors in different areas of applications. Here, in particular, the motivation of the proposed study in a real-problem is to modelling and forecasting variables from financial data area, with special attention to assets returns variables which, in general, do not preserve the property of constant variance over time.

Keywords: Volatility, Robustness, GARCH, M-Quantile.

1 Introduction

Since the 1970s, attention has been paid to the study of methods in the context of processes that contain atypical observations (outliers) among other particularities. In the case of the effect of outliers in estimating time series models, several authors have showed that the presence of outliers can dramatically deteriorate the modeling of time series. For example, in the standard structure of Box-Jenkins models, Ledolter (1989) showed that the range predictions in Integrated Autoregressive

Moving Averages (ARIMA) models are considerably sensitive to additive outliers. Chang, Tiao, and Chung Chen (1988) and C. Chen and Liu (1993) showed that the estimated parameters of the ARMA model become biased when the data contains outliers. For time series with different correlation structures, see, for example, Reisen, Lévy-Leduc, and Taqqu (2017), Fajardo, Reisen, and Lévy-Leduc (2018) and Sarnaglia et al. (2021).

As is well discussed in the literature, the estimators derived from the M regression method, from Huber (1964), are robust alternative approaches to obtain estimates of the model parameters, even in series contaminated by outliers or generated by probability distributions with heavy tails, see for example Bai, Rao, and Y. H. Wu (1997), W. B. Wu (2007) and Li (2008). The use of the M regression method in short and long memory processes is well documented, see Reisen and Fajardo (2012), Fajardo, Reisen, Lévy-Leduc, et al. (2016), among others.

The effects of additive outliers on the estimates of the parameters of the GARCH (p,q) model were recently studied by Franses and Ghijssels (1999), Mendes (2000), Carnero (2003), Carnero, Peña, and Ruiz (2005) among others. In general, they show that additive outliers can substantially distort the estimation of the parameters of the ARCH (q) and GARCH (p, q) models in the same way as in the standard linear time series models. Since this research is concentrated on the model with conditional heteroscedasticity, in what follows the popular GARCH(p,q) model, which holds this property, is defined.

A thorough search of the relevant literature on robust estimation in heteroscedastic processes indicates that there are not much articles devoted to this topic. For the best knowledge of the authors, the most recent literature on the subject are Muler and Yohai (2008), Carnero, Peña, and Ruiz (2012) and Iqbal (2013). Therefore, to fill part of this gap in this research area, this study proposes the use of a quantile M-regression to estimate the parameters of the GARCH(p,q) model.

The contributions of this study be mainly follows these lines: (1) To propose an alternative estimation method based on M quantile regression to estimate the parameters of a GARCH(p,q) model, here denoted by MQGARCH(p,q); (2) To perform simulation study to verify the performance of the estimator in small samples of time series with and without additive outliers and to verify the usefulness of the estimator in a real problem, for examples, with financial index returns observed from Financial Time Series area.

2 The GARCH Model

Let $(\epsilon_t)_{t \in \mathbb{Z}}$ be a process satisfying $E(\epsilon_t) = 0$ and $E(\epsilon_t^4) < \infty$. $(\epsilon_t)_{t \in \mathbb{Z}}$ is said to belong to the class of Generalized Auto-regressive models with Conditional Heteroscedasticity, of order $p, q \in \mathbb{N}^+$, that is, $(\epsilon_t)_{t \in \mathbb{Z}}$ is a GARCH (p, q) model, if it satisfies the following equation

$$\begin{cases} \epsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_t^2 + \sum_{j=1}^p \beta_j \sigma_t^2, \end{cases} \quad (1)$$

where α_i, β_j are non-negative constants, α_0 is a strictly positive constant and $\sigma_t > 0$. Additionally $(\eta_t)_{t \in \mathbb{Z}}$ is a stochastic sequence with zero mean and unit variance.

If the parameters of the GARCH(p,q) model satisfy $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1$, the process in (1) is strictly stationary and has a unique solution. See, for example, Theorem 2.5 in Francq and Zakoian (2019). The model (1) was introduced in the

literature by Bollerslev (1986) and since then it has been widely used in several areas of knowledge, such as in Econometrics, Finance and Air Quality. For Air Quality area, see, for example, Reisen, Sarnaglia, et al. (2014) and Monte, Albuquerque, and Reisen (2015).

The GARCH(p,q) model was introduced as an extension of the Auto-regressive model with Conditional Heteroscedasticity, ARCH (q), initially proposed by Engle (1982), which is defined as the model (1) with the parameters $\beta_j, j = 1, \dots, p$ nulls.

The ARMA (r, p) representation of the GARCH (p, q) model, where $r = \max(p, q)$, is given as follows

$$\epsilon_t^2 = \alpha_0 + v_t + \sum_{i=1}^r (a_i + \beta_i) \epsilon_t^2 - \sum_{j=1}^p \beta_j v_t, \quad (2)$$

with the $\alpha_i = 0$ ($\beta_j = 0$) convention if $i > q$ ($j > p$). The noise, $v_t = \epsilon_t^2 - \sigma_t^2$, is a non-correlated sequence with zero mean, not independent and not Gaussian. More details can be found on page 665 of Hamilton (1994). The representation in (2) is widely used to estimate the parameter of GARCH(p,q) model, because is more convenience mathematically. To estimate the GARCH models (p, q) there are several methods in the literature, with emphasis on Ordinary Least Squares (OLS), Generalized Least Squares (GLS), Maximun Likelihood (ML) and Quasi Maximun Likelihood (QML). Details can be found in part II of the book Francq and Zakoian (2019). In many cases, in addition to representation (2), the GARCH(p,q) model can be represented by a ARCH(∞) model. This because a ARCH representation have the follows asymptotic properties well-defined.

Let $\boldsymbol{\alpha}$ be the ARCH (q) model $q + 1$ dimension vector and $\hat{\boldsymbol{\alpha}}$ the $\boldsymbol{\alpha}$ estimator .
Asymptotic Properties of Estimators.

Under certain conditions, the properties of the OLS, GLS, ML, QML estimators have the following asymptotic properties.

$$\text{OLS} : \sqrt{T}(\hat{\boldsymbol{\alpha}}^{OLS} - \boldsymbol{\alpha}) \xrightarrow{d} N(0, \Sigma_{XX}^{-1} \Sigma_{X\Omega X} X X^{-1}),$$

where T is the sample size, X is an array of observations from the process $(\epsilon_t^2)_{t \in \mathbb{Z}}$, $\Sigma_{XX} = p \lim \frac{X X'}{T}$, and $\Sigma_{X\Omega X} = p \lim \frac{X' V V' X}{T}$, with V a process matrix $(v_t)_{t \in \mathbb{Z}}$ e $\Omega = \text{diag}(\sigma_{p+1}^4, \sigma_{p+2}^4, \dots, \sigma_T^4)$ (Engle (1982)).

$$\text{GLS} : \sqrt{T}(\hat{\boldsymbol{\alpha}}^{GLS} - \boldsymbol{\alpha}) \xrightarrow{d} N(0, \Sigma_{\Omega}^{-1}),$$

with $\Sigma_{\Omega} = p \lim \frac{X' \Omega X}{T}$. (Bose and Mukherjee (2003))

$$\text{ML} : \sqrt{T}(\hat{\boldsymbol{\alpha}}^{ML} - \boldsymbol{\alpha}) \xrightarrow{d} N(0, [I(\boldsymbol{\alpha})]^{-1}),$$

where $I(\alpha) = E[-\frac{\partial^2 L}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}'}]$ is an array of information and L a log-likelihood function. (Ling and McAleer (2003)).

QML :

$$\sqrt{T}(\hat{\boldsymbol{\alpha}}^{QMV} - \boldsymbol{\alpha}) \xrightarrow{d} N(0, (\kappa_{\eta} - 1)J^{-1}),$$

where $\kappa_{\eta} = E(\eta_t^4)$, $J := E(\frac{\partial l_t(\alpha_0)}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}'})$, with $l_t = \frac{\epsilon_t^2}{\sigma_t^2} + \log(\sigma_t^2)$ (Weiss (1986) and Berkes, Hovath, and Kokoszka (2003)).

3 The M-Quantile Estimator

This section is dedicated to the presentation, in a summarized form, of the main theoretical properties of the regression estimator M, initially proposed by Huber (1964). Additionally, the adaptation of this method is presented in the context of quantile regression, proposed by Breckling and Chambers (1988). This method has, so far, been used in linear time series models. As a scientific contribution, the M-Quantile estimation method is adapted for GARCH(p, q) models, in particular for the GARCH(1,1) model.

Some of the results presented here are widely discussed in the literature, see for example Weiss (1986), Berkes, Hovath, and Kokoszka (2003), Iqbal and Mukherjee (2007) and Mukherjee (2008).

Let $\theta = (\alpha_0 \alpha_1 \cdots \alpha_q \beta_1 \cdots \beta_p)'$ o vector of unknown parameters of the GARCH (p,q) model defined in the equation (1).

Let g be any function with finite derivatives. Then \dot{g} and \ddot{g} denote the first and second derivatives of g , respectively.

Let $\psi : \mathbb{R} \rightarrow \mathbb{R}$ be an odd function, that is, $\psi(-x) = -\psi(x), \forall x \in \mathbb{R} - \{0\}$, differentiable in almost the entire set of points. Let $\mathcal{D} \subset \mathbb{R}$ be the set of points where ψ is differentiable and $\overline{\mathcal{D}}$ its complement. Let $H(x) := x\psi(x)$ with $x \in \mathbb{R}$. Note that $H(-x) = H(x), \forall x$. The H function will be called the scoring function for regression M. As an example, the scoring function of Huber (1964) is described below. Let $\psi(x) = xI(|x| \leq k) + k\text{sign}(x)I(|x| > k)$, where $k > 0$ is known. So $\overline{\mathcal{D}} = \{-k, k\}$ and $H(x) = x^2I(|x| \leq k) + k|x|I(|x| > k)$. Next, the estimator M is defined. In this work the ψ function is the Huber function according to work Huber (1964), because the Huber function is robust in the presence of outliers.

Recalling that in a rental model, an M estimator is defined as the solution to a certain system of equations involving residual functions. To follow the same approach, we first discuss the concept of the variance function. Assume that for some $\kappa \geq 0$, you have $E(|\eta| < \infty)$. So from Lemma 2.3 and Theorem 2.1 presented in Berkes, Hovath, and Kokoszka (2003), σ_t^2 em (1) has the almost certain unique representation

$$\sigma_t^2 = c_0 + \sum_{j=1}^{\infty} c_j \epsilon_{t-j}^2; t \in \mathbb{Z} \quad (3)$$

where $c_j, j = 1, ..$ satisfies the non-negativity condition. For more details and conditions, see Berkes, Hovath, and Kokoszka (2003).

Note. The representation given in (3) is widely used to obtain estimators of the GARCH(p,q) model in univariate and multivariate cases. For example, according to Hu and Tsay (2014), the equation (3) in the multivariate context can be written as follows. Let $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t}, \dots, \epsilon_{kt})'$ a weakly stationary k-dimensional time series with null Mean and fourth finite moment. Let $\mathcal{F}_{t-1} = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \dots)$ a field of information available at time $t - 1$. The volatility matrix Σ_t of ϵ_t is defined as $Cov(\epsilon_t | \mathcal{F}_{t-1}) = E(\epsilon_t \epsilon_t' | \mathcal{F}_{t-1})$.

$$vec(\Sigma_t) = \mathbf{C}_0 + \sum_{i=1}^{\infty} \mathbf{C}_i vec(\epsilon_{t-i} \epsilon_{t-i}'), \quad (4)$$

where $vec(\mathbf{M})$ denotes a stacking array of columns in the array \mathbf{M} , \mathbf{C}_0 is a vector of constants k^2 -dimensional and \mathbf{C}_i are constant arrays of dimension $k^2 \times k^2$ for

$i > 0$. The arrays \mathbf{C}_i and \mathbf{C}_0 satisfy conditions that guarantee that Σ_t is positively defined for every t . Therefore, the equation (3) is a particular case of (4).

Let Θ be a compact subset of $(0, \infty)^{1+q} \times (0.1)^p$. Define $V_t(\boldsymbol{\theta})$ as a conditional variance function of $(\epsilon_t)_{t \in \mathbb{Z}}$ in Θ . Then, under the conditions presented in Mukherjee (2008), $V_t(\boldsymbol{\theta})$ can be represented as

$$V_t(\boldsymbol{\theta}) = c_0(\boldsymbol{\theta}) + \sum_{j=1}^{\infty} c_j(\boldsymbol{\theta}) \epsilon_{t-j}^2, \boldsymbol{\theta} \in \Theta, t \in \mathbb{Z}, \quad (5)$$

where $c_j(\boldsymbol{\theta}), j \geq 0$ satisfies the non-negativity condition. In particular, you can use the model (5) to estimate a GARCH(1,1) model, where $c_0(\boldsymbol{\theta}) = \frac{\alpha_0}{1-\beta_1}$ e $c_j(\boldsymbol{\theta}) = \alpha_1 \beta_1^{j-1}$. More details and examples are discussed in Berkes, Hovath, and Kokoszka (2003).

Let $\epsilon_1, \epsilon_2, \dots, \epsilon_n$ be a set of observations of the process in (1). Then a $(V_t(\tilde{\boldsymbol{\theta}}))_{t \in \mathbb{Z}}$ approximation of the $(V_t(\boldsymbol{\theta}))_{t \in \mathbb{Z}}$ process is given by

$$\tilde{V}_t(\boldsymbol{\theta}) = c_0(\boldsymbol{\theta}) + I(2 \leq t) \sum_{j=1}^{n-1} c_j(\boldsymbol{\theta}) \epsilon_{t-j}^2, \boldsymbol{\theta} \in \Theta,$$

So the $\boldsymbol{\theta}$ estimator, defined as $\hat{\boldsymbol{\theta}}_n$, is the solution of

$$\sum_{t=1}^n \frac{1}{2} \left[1 - H\left(\frac{\epsilon_t}{\tilde{V}_t^{\frac{1}{2}}(\boldsymbol{\theta})}\right) \right] \frac{\dot{\tilde{V}}_t(\boldsymbol{\theta})}{\tilde{V}_t(\boldsymbol{\theta})} = 0. \quad (6)$$

Under conditions presented in Mukherjee (2008),

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \rightarrow N(0, \sigma^2(H)G^{-1}),$$

where $G = E\left(\frac{\dot{\tilde{V}}_t(\boldsymbol{\theta})\dot{\tilde{V}}_t'(\boldsymbol{\theta})}{\tilde{V}_t^2(\boldsymbol{\theta})}\right)$.

According to Breckling and Chambers (1988), a ϕ function, in an M-Quantile context, can be written as

$$\phi_\tau(x) = \begin{cases} (1-\tau)\psi(x), & (x < 0) \\ \tau\psi(x), & \text{otherwise,} \end{cases} \quad (7)$$

where $0 < \tau < 1$ and $\psi(x)$ defined earlier. As an example the function $\psi(x)$ proposed by Huber (1964), in this context, is written as

$$\phi_\tau(x) = \begin{cases} -(1-\tau)k, & \text{if } x < -k, \\ (1-\tau)x, & \text{if } -k \leq x < 0, \\ \tau x, & \text{if } 0 \leq x < k, \\ \tau k, & \text{if } k \leq x. \end{cases} \quad (8)$$

Therefore, considering the estimator of $\hat{\boldsymbol{\theta}}_n$ in the expression (6), you must define a function $H(x) = x\phi_\tau(x)$ at point $x = \frac{\epsilon_t}{\tilde{V}_t(\boldsymbol{\theta})^{\frac{1}{2}}}$. So, there is

$$H\left(\frac{\epsilon_t}{\tilde{V}_t(\boldsymbol{\theta})^{\frac{1}{2}}}\right)\phi_\tau\left(\frac{\epsilon_t}{\tilde{V}_t(\boldsymbol{\theta})^{\frac{1}{2}}}\right) = \begin{cases} -(1-\tau)k\left(\frac{\epsilon_t}{\tilde{V}_t(\boldsymbol{\theta})^{\frac{1}{2}}}\right), & \text{if } \frac{\epsilon_t}{\tilde{V}_t(\boldsymbol{\theta})^{\frac{1}{2}}} < -k, \\ (1-\tau)\left(\frac{\epsilon_t}{\tilde{V}_t(\boldsymbol{\theta})^{\frac{1}{2}}}\right)^2, & \text{if } k \leq \frac{\epsilon_t}{\tilde{V}_t(\boldsymbol{\theta})^{\frac{1}{2}}} < 0, \\ \tau\left(\frac{\epsilon_t}{\tilde{V}_t(\boldsymbol{\theta})^{\frac{1}{2}}}\right)^2, & \text{if } 0 \leq \frac{\epsilon_t}{\tilde{V}_t(\boldsymbol{\theta})^{\frac{1}{2}}} < k, \\ \tau k\left(\frac{\epsilon_t}{\tilde{V}_t(\boldsymbol{\theta})^{\frac{1}{2}}}\right), & \text{if } k \leq \frac{\epsilon_t}{\tilde{V}_t(\boldsymbol{\theta})^{\frac{1}{2}}}. \end{cases} \quad (9)$$

4 Simulated Study

The purpose of this section is to evaluate, through various Monte Carlo experiments, the effects of additive outliers on the model's parameters GARCH(1,1) and MQGARCH(1,1). Additionally, the performance of the proposed robust model, MQGARCH(1,1), and the classic GARCH(1,1) model, estimated by ML, are also investigated. As measures of parameter performance, the mean, standard deviation (Std. Deviation), bias and mean squared error (MSE) were used.

Let $(\epsilon_t)_{t \in \mathbb{Z}}$ be the process generated by the equation (1), with $p = q = 1$, that is, $(\epsilon_t)_{t \in \mathbb{Z}}$ follows a GARCH(1,1) model. Let $(\eta_t)_{t \in \mathbb{Z}}$ be a sequence with zero mean, unit variance and Gaussian distribution. Consider that the series of interest, $(\epsilon_t)_{t \in \mathbb{Z}}$, stationary and with a finite fourth moment, is contaminated in time t^* by an additive outlier of magnitude ω . The observed series is given by

$$(z_t)_{t \in \mathbb{Z}} = \begin{cases} \epsilon_t + \omega, & \text{if } t = t^* \\ \epsilon_t, & \text{otherwise,} \end{cases} \quad (10)$$

The contaminated process $(z_t)_{t \in \mathbb{Z}}$ was simulated according to the equation (10). The probability of an outlier occurring in time t is $p = 0.01$ and, without loss of generality, the magnitude of the outlier is assumed to be $\omega = 4\sigma$. More details and other methods for including outliers in time series can be found at Carnero (2003), Carnero, Peña, and Ruiz (2005), among others.

For the experiment, 500 $(z_t)_{t \in \mathbb{Z}}$ processes of size $n = 100$ each were generated. Additionally, 4 sets of parameters were considered to generate the process $(z_t)_{t \in \mathbb{Z}}$ according to the equation (1), namely: C1 = $(\alpha_0 = 0.5 \alpha_1 = 0.4 \beta_1 = 0.4)$, C2 = $(\alpha_0 = 0.5 \alpha_1 = 0.2 \beta_1 = 0.6)$, C3 = $(\alpha_0 = 0.1 \alpha_1 = 0.1 \beta_1 = 0.8)$ and C4 = $(\alpha_0 = 0.1 \alpha_1 = 0.05 \beta_1 = 0.9)$.

In the context of the MQGARCH(1,1) model, $k = 1.5$ was considered, in the scoring function of the equation (9), conform simulations in Mukherjee (2008). Additionally, this model was estimated at the quantile $\tau = 0.5$. For this reason, the GARCH(1,1) and MQGARCH(1,1) models are expected to have similar parameter values when $(z_t)_{t \in \mathbb{Z}} \equiv (\epsilon_t)_{t \in \mathbb{Z}}$. It is also expected that the MQGARCH(1,1) model will have reasonable estimates considering the process $(z_t)_{t \in \mathbb{Z}}$ generated according to parameters C1-C3, as they satisfy stationarity conditions. For the C4 set, which is very close to not satisfying the stationarity conditions, that is $\alpha_1 + \beta_1 = 0.95$ (próximo de 1), it is expected that the proposed model does not provide good estimates.

The items (a) and (b) of the table (1) shows the performance measures of the parameters of the GARCH(1,1) model with and without additive outliers, respectively, where $(z_t)_{t \in \mathbb{Z}}$ are generated according to C1-C4. As can be seen, the empirical results are in line with the results of Section 5 of Carnero, Peña, and Ruiz (2005). In particular, the parameter most affected by the presence of outliers was the $\hat{\alpha}_1$ parameter what can be explained by the fact that the contamination is given directly in the $(\epsilon_t)_{t \in \mathbb{Z}}$ process. In this case, the the parameter of the ARCH effect ($\hat{\alpha}_1$) tends to decrease in the presence of additive outliers (see table (3)), as exposed in the mentioned study.

The results of item (a) of the table (2) show the performance measures of the parameters of the MQGARCH(1,1) model, where $(z_t)_{t \in \mathbb{Z}} \equiv (\epsilon_t)_{t \in \mathbb{Z}}$ are generated according to C1-C4. As mentioned earlier, in this case, the behavior of the proposed model is very close to that of the classic GARCH(1,1) model, which can be explained by the fact that both are estimated close to the center of symmetry of the conditional

data distribution.

The item (b) of the table (2), where the results are shown for the MQGARCH(1,1) model shows a high accuracy when compared with the item (b) of the table (1). The behavior of the MQGARCH(1,1) model, in this case, is substantially different from GARCH(1,1). Additionally, the M-quantile estimator provided to be more robust in terms of Bias and MSE quantities when the process is contaminated, what suggest that the M-quantile estimator is more efficient and accurate than the QML method widely used in econometrics and finance. In order to corroborate with the previous arguments, the percentage variability of the performance measures of the models estimated before and after contamination by additive outliers is reported in items (a) and (b) of the table (3).

Table 1: Mean, Std. Deviation, Bias and MSE of the estimated parameters of the GARCH(1,1) model

((a)) Without Outliers					((b)) Outliers: Probability 1 % and magnitude $\omega = 4\sigma$				
$N(0.1)$					$N(0.1)$				
100 observations and 500 repetitions					100 observations and 500 repetitions				
	Mean	Std. Deviation	Bias	MSE		Mean	Std. Deviation	Bias	MSE
	$\alpha_0 = 0.5 \alpha_1 = 0.4 \beta_1 = 0.4$					$\alpha_0 = 0.5 \alpha_1 = 0.4 \beta_1 = 0.4$			
$\hat{\alpha}_0$	0.6237	0.3611	0.1237	0.1304	$\hat{\alpha}_0$	0.6196	0.5085	0.1196	0.2586
$\hat{\alpha}_1$	0.4074	0.1890	0.0074	0.0357	$\hat{\alpha}_1$	0.3223	0.2373	0.0776	0.0563
$\hat{\beta}_1$	0.3460	0.2227	0.0539	0.0494	$\hat{\beta}_1$	0.4618	0.3080	0.0618	0.0948
	$\alpha_0 = 0.5 \alpha_1 = 0.2 \beta_1 = 0.6$					$\alpha_0 = 0.5 \alpha_1 = 0.2 \beta_1 = 0.6$			
$\hat{\alpha}_0$	0.6994	0.5312	0.1994	0.2822	$\hat{\alpha}_0$	0.6698	0.5521	0.1698	0.3103
$\hat{\alpha}_1$	0.2035	0.1289	0.0035	0.0192	$\hat{\alpha}_1$	0.1866	0.1506	0.0133	0.0226
$\hat{\beta}_1$	0.5116	0.2879	0.0883	0.0829	$\hat{\beta}_1$	0.5629	0.2887	0.0370	0.0833
	$\alpha_0 = 0.1 \alpha_1 = 0.1 \beta_1 = 0.8$					$\alpha_0 = 0.1 \alpha_1 = 0.1 \beta_1 = 0.8$			
$\hat{\alpha}_0$	0.2131	0.2207	0.1131	0.0496	$\hat{\alpha}_0$	0.2149	0.2534	0.1149	0.0642
$\hat{\alpha}_1$	0.1059	0.1007	0.0059	0.0101	$\hat{\alpha}_1$	0.1099	0.1516	0.0099	0.0230
$\hat{\beta}_1$	0.6807	0.2714	0.1192	0.0737	$\hat{\beta}_1$	0.7124	0.2740	0.0912	0.0751
	$\alpha_0 = 0.1 \alpha_1 = 0.05 \beta_1 = 0.9$					$\alpha_0 = 0.1 \alpha_1 = 0.05 \beta_1 = 0.9$			
$\hat{\alpha}_0$	0.3780	0.4820	0.2780	0.2323	$\hat{\alpha}_0$	0.4235	0.5160	0.3235	0.2662
$\hat{\alpha}_1$	0.0606	0.0848	0.0106	0.0072	$\hat{\alpha}_1$	0.0689	0.1124	0.0189	0.0126
$\hat{\beta}_1$	0.7537	0.2756	0.1462	0.0759	$\hat{\beta}_1$	0.7377	0.2864	0.1622	0.0820

∞

Table 2: Mean, Std. Deviation, Bias and MSE of the parameters of the MQGARCH(1,1) model, with $\tau = 0.5$

((a)) Without Outliers					((b)) Outliers: Probability 1 % and magnitude $\omega = 4\sigma$				
$N(0.1)$					$N(0.1)$				
100 observations and 500 repetitions					100 observations and 500 repetitions				
	Mean	Std. Deviation	Bias	MSE		Mean	Std. Deviation	Bias	MSE
	$\alpha_0 = 0.5 \alpha_1 = 0.4 \beta_1 = 0.4$					$\alpha_0 = 0.5 \alpha_1 = 0.4 \beta_1 = 0.4$			
$\hat{\alpha}_0$	0.5981	0.2872	0.1840	0.1113	$\hat{\alpha}_0$	0.6003	0.2468	0.1089	0.1452
$\hat{\alpha}_1$	0.3971	0.1276	0.0080	0.0167	$\hat{\alpha}_1$	0.3922	0.1263	0.0189	0.0172
$\hat{\beta}_1$	0.4101	0.2056	0.0123	0.0575	$\hat{\beta}_1$	0.4219	0.2451	0.0137	0.0789
	$\alpha_0 = 0.5 \alpha_1 = 0.2 \beta_1 = 0.6$					$\alpha_0 = 0.5 \alpha_1 = 0.2 \beta_1 = 0.6$			
$\hat{\alpha}_0$	0.6704	0.4863	0.1947	0.2034	$\hat{\alpha}_0$	0.6713	0.4063	0.1678	0.2274
$\hat{\alpha}_1$	0.2021	0.1009	0.0091	0.0102	$\hat{\alpha}_1$	0.1999	0.1177	0.0112	0.0102
$\hat{\beta}_1$	0.5478	0.2531	0.0169	0.0643	$\hat{\beta}_1$	0.5507	0.2109	0.0321	0.0656
	$\alpha_0 = 0.1 \alpha_1 = 0.1 \beta_1 = 0.8$					$\alpha_0 = 0.1 \alpha_1 = 0.1 \beta_1 = 0.8$			
$\hat{\alpha}_0$	0.2092	0.2146	0.1014	0.0875	$\hat{\alpha}_0$	0.2108	0.2252	0.1043	0.0691
$\hat{\alpha}_1$	0.1045	0.1099	0.0085	0.0101	$\hat{\alpha}_1$	0.1001	0.1324	0.0107	0.0109
$\hat{\beta}_1$	0.7441	0.2152	0.0395	0.0478	$\hat{\beta}_1$	0.7807	0.2073	0.0473	0.0562
	$\alpha_0 = 0.1 \alpha_1 = 0.05 \beta_1 = 0.9$					$\alpha_0 = 0.1 \alpha_1 = 0.05 \beta_1 = 0.9$			
$\hat{\alpha}_0$	0.2909	0.5002	0.2669	0.2501	$\hat{\alpha}_0$	0.3112	0.4524	0.2410	0.2794
$\hat{\alpha}_1$	0.0605	0.0747	0.0494	0.0080	$\hat{\alpha}_1$	0.0654	0.1095	0.0548	0.0081
$\hat{\beta}_1$	0.8212	0.2364	0.1000	0.0654	$\hat{\beta}_1$	0.8476	0.2351	0.1007	0.0659

Table 3: Percentage change variation of the Mean, Std. Deviation, Bias and MSE of the parameters of the GARCH(1,1) and MQGARCH(1,1) models before and after contamination

((a)) GARCH(1,1)					((b)) MQGARCH(1,1), $\tau = 0.5$				
$N(0, 1)$					$N(0, 1)$				
Percentage change					Percentage change				
	Mean	Std. Deviation	Bias	MSE		Mean	Std. Deviation	Bias	MSE
	$\alpha_0 = 0.5 \alpha_1 = 0.4 \beta_1 = 0.4$					$\alpha_0 = 0.5 \alpha_1 = 0.4 \beta_1 = 0.4$			
$\hat{\alpha}_0$	-0.6573	40.8197	-3.3144	98.3128	$\hat{\alpha}_0$	0.3678	-14.0668	-1.0188	30.4582
$\hat{\alpha}_1$	-0.2088	25.5555	948.6486	57.7030	$\hat{\alpha}_1$	-1.2339	14.6551	136.25	2.9940
$\hat{\beta}_1$	33.4682	38.3026	14.6567	91.9028	$\hat{\beta}_1$	2.8773	19.2120	11.3821	37.2172
	$\alpha_0 = 0.5 \alpha_1 = 0.2 \beta_1 = 0.6$					$\alpha_0 = 0.5 \alpha_1 = 0.2 \beta_1 = 0.6$			
$\hat{\alpha}_0$	-4.2321	3.9344	-14.8445	9.8866	$\hat{\alpha}_0$	0.1342	-16.4517	-13.8161	11.7994
$\hat{\alpha}_1$	-8.3046	16.8347	280.0000	17.7083	$\hat{\alpha}_1$	-0.2200	16.6501	23.0769	0.0000
$\hat{\beta}_1$	10.0273	0.2778	-58.0973	6.5138	$\hat{\beta}_1$	0.5293	-16.6732	89.9408	2.0217
	$\alpha_0 = 0.1 \alpha_1 = 0.1 \beta_1 = 0.8$					$\alpha_0 = 0.1 \alpha_1 = 0.1 \beta_1 = 0.8$			
$\hat{\alpha}_0$	0.8446	14.8164	1.5915	29.4353	$\hat{\alpha}_0$	0.7648	4.9394	2.8599	-21.0285
$\hat{\alpha}_1$	-3.7771	50.5461	67.7966	127.7277	$\hat{\alpha}_1$	-4.2105	20.4731	25.8823	7.9207
$\hat{\beta}_1$	4.6569	0.9579	23.4899	1.8995	$\hat{\beta}_1$	4.9186	-3.6710	19.7468	17.5732
	$\alpha_0 = 0.1 \alpha_1 = 0.05 \beta_1 = 0.9$					$\alpha_0 = 0.1 \alpha_1 = 0.05 \beta_1 = 0.9$			
$\hat{\alpha}_0$	12.0370	7.0539	16.3669	14.5931	$\hat{\alpha}_0$	6.9783	-9.5561	-9.7040	11.4408
$\hat{\alpha}_1$	13.6963	32.7431	78.3018	75.0000	$\hat{\alpha}_1$	8.0991	46.5863	10.9311	1.2500
$\hat{\beta}_1$	-2.1228	3.9187	10.9439	8.0368	$\hat{\beta}_1$	2.1679	-0.5499	0.7000	0.7665

5 Real Data Application

In this section, we analyzed the daily returns of two Korean financial indices, the Korea Composite Price Index (KOSPI) and the Automated Korean Securities Price Index (KOSDAQ). 1155 observations were considered in each series between the period from January 2, 2004 to August 29, 2008. Let p_t be the price of the financial index in time t and r_t the log return defined as $r_t = \log \frac{p_t}{p_{t-1}} 100$. It is assumed that r_t follows a GARCH process according to the definition given in (1), where $p = q = 1$, that is, a GARCH(1,1) model.

Table 4: Descriptive statistics of the logarithm of the daily returns of the financial indexes (KOSPI and KOSDAQ)

	KOSPI	KOSDAQ
Mean	0.0219	0.0015
Std. Deviation	0.5743	0.6149
Maximum	2.4032	2.9965
Minimum	-3.172	-4.6586
Asymmetry	-0.5004	-1.4592
Kurtosis	1.9523	8.2567

A statistical summary of the data can be seen in the table (4). The series from the period from January 2, 2004 to August 31, 2007 was used to estimate the parameters of the MQGARCH(1,1) model and the series for the next twelve months was considered as a post-sample. The temporal behaviour of both series is shown in the figure (1). As previously discussed, there are possible outliers in both

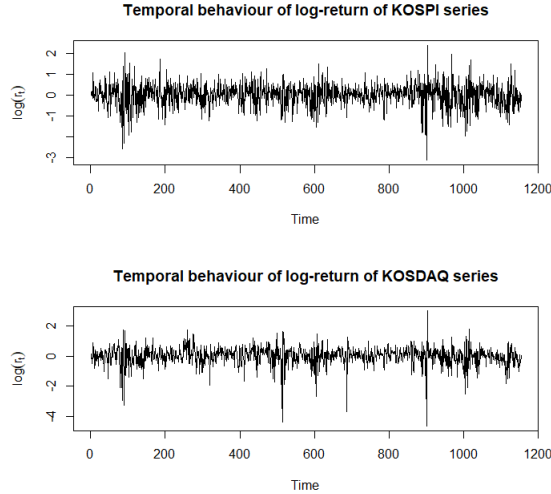


Figure 1: Temporal behavior of log-returns of KOSPI and KOSDAQ series

series, conform the figure (2), which is usual phenomenon in financial time series. The KOSDAQ series have outliers of large magnitude compared with KOSPI series. Than, as shown in the Section 4 this study, is expected that the QML method be less efficient and accurate than MQ method for estimated the parameters of GARCH(1,1) model.

The tables (5) and (6) list the parameter estimates of the GARCH(1,1) model obtained based on the method QML estimator and the M-Quantile estimator, respectively. In the Table (6), it is noted that for the KOSPI series, the quantile

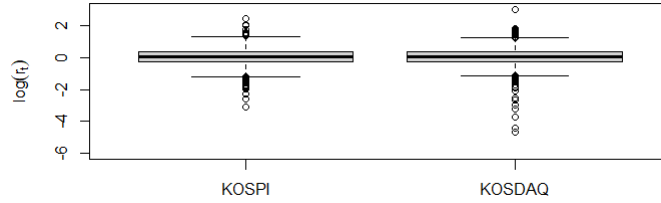


Figure 2: Boxplot of log-returns of KOSPI and KOSDAQ series

regression estimates from α_1 to $\tau < 0.55$ tend to be higher than those for $\tau > 0.55$. This fact suggests the presence of asymmetry in the KOSPI series, which is in line with the results discussed in Lee and Noh (2013).

Table 5: GARCH model parameters (1,1) estimated by the QML method

	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$
KOSPI	0.0088	0.0859	0.8876
KOSDAQ	0.0541	0.3207	0.5676

Table 6: Parameters of the MQGARCH(1,1) model estimated at different quantiles

τ	0.01	0.05	0.25	0.45	0.55	0.75	0.95	0.99
KOSPI								
$\hat{\alpha}_0$	0.4194	0.1333	0.0791	0.0568	0.0314	0.0022	0.0108	0.0445
$\hat{\alpha}_1$	0.0233	0.0084	0.0089	0.0192	0.0126	0.0059	0.0066	0.0071
$\hat{\beta}_1$	0.4343	0.9864	0.9902	0.9135	0.8969	0.6650	0.4936	0.5644
KOSDAQ								
$\hat{\alpha}_0$	0.7381	0.5011	0.1773	0.0800	0.0475	0.0242	0.0317	0.1240
$\hat{\alpha}_1$	0.0211	0.0309	0.0334	0.0292	0.0375	0.0356	0.0207	0.0591
$\hat{\beta}_1$	0.7407	0.9495	0.8800	0.7062	0.6581	0.5180	0.4810	0.5223

Heteroscedastic models based on the quantile regression method make it possible to directly estimate the Value-at-risk (VaR), widely used in econometric literature for the management of financial risks. This is because estimating VaR is explicitly a problem in estimating conditional quantiles. The following definition of VaR can be found at Gouriéroux and Jasiak (2001) and Franke, Härdle, and Hafner (2004).

Definition. The $(1 - \tau)$ -quantile of the conditional loss distribution is called VaR at the τ level:

$$VaR_{t,h}(\tau) := \inf\{x \in \mathbb{R} | P_t[L_{t,t+h} \leq x] \geq 1 - \tau\}, \quad (11)$$

where $\tau \in (0,1)$, P_t is the conditional probability given the information available at time t , and $L_{t,t+h} = -\sum_{i=1}^d a_i p_{i,t} (e^{r_{i,t,t+h}} - 1)$ denotes the loss of a portfolio of d shares, with a_i the number of shares i in the portfolio, $p_{i,t}$ the share price i in time t and $r_{i,t,t+h}$ the log return in a horizon h given by $\log(p_{i,t+h}) - \log(p_{i,t})$. In particular, $VaR_{t,h}(\tau)$ is a decreasing function of τ . Given σ_t as defined in (1), according to Francq and Zakoian (2019) the $VaR_{t,h}(\tau)$ can be calculated as

$$VaR_{t,h}(\tau) = m_{t,t+h} + \sigma_{t,t+h} F_{\eta}^{-1}(1 - \tau), \quad (12)$$

where $m_{t,t+h} = E(L_{t,t+h})$, $\sigma_{t,t+h}^2$ is the conditional volatility of $L_{t,t+h}$ and $F_{\eta}^{-1}(\cdot)$ denotes the conditional quantile of the variable η . Therefore, the MQGARCH(1,1) model, the main subject of this article, has a substantial advantage in estimating VaR.

VaR based on the MQGARCH(1,1) model for the KOSPI and KOSDAQ series are presented in the tables (7) and (8). In the Table (7) the VaR's of the returns in the training sample are reported. A significantly difference between the truth quantile and the estimated is not observed.

Table 7: Proportion of returns in the training sample below estimated conditional quantiles

τ	0.01	0.05	0.95	0.99
KOSPI	0.0000	0.0876	0.9270	0.9804
KOSDAQ	0.0000	0.1344	0.9573	0.9793

In the case of the post-sample, VaR was predicted to be one step ahead, where the parameters are obtained based on the previously defined training sample. The table (8) lists the results of the predictions made using the quantiles $\tau \in \{0.01, 0.05, 0.95, 0.99\}$. It is noted that the proportions of returns in the post-sample below the expected quantiles do not exhibit a significant deviation of τ in the KOSPI and KOSDAQ series, which is in line with the results of Section 4.2 of Lee and Noh (2013), suggesting that the GARCH(1,1) model based on the M-Quantile method can also be used to calculate the VaR of returns on financial assets.

Table 8: Proportion of returns in the post-sample below the predicted conditional quantiles

τ	0.01	0.05	0.95	0.99
KOSPI	0.0000	0.0227	0.9530	0.9857
KOSDAQ	0.0000	0.0353	0.9521	0.9842

Note. There are other methods to calculate VaR, in addition to the one illustrated in the equation (12). However, Lee and Noh (2013) mention that methods based on the GARCH model perform better in the sense of forecasting.

6 Conclusions

This article proposed a robust estimator for additive outliers for the GARCH(1,1) model based on the M quantile regression method, that is, the MQGARCH(1,1) model. In addition, the MQGARCH(1,1) was used to measure the risk of returns on financial asset price indices, which helps in investment decisions. The article considers the effects of additive outliers on the parameters of a GARCH(1,1) model and their implications for the analysis and interpretations of these models.

A stylized fact about Financial Time Series is that the series have aberrant or atypical observations, that is, outliers. It was shown that the parameters of the GARCH(1,1) model, widely used in the literature, suffer dramatic effects due to the contamination of the series, which in this case can lead to erroneous interpretations. On the other hand, it was also shown that the parameters of the MQGARCH(1,1)

model, the main theme of this study, remained robust when the series was contaminated, showing that the M-Quantile estimator, in the context of heteroscedastic models, was parsimonious in that case.

It was also mentioned that the presence of additive outliers affects the estimated autocorrelation function $r(h)$ of a time series. As a result, heteroscedasticity may not be detected, taking into account that the values of $r(h)$ may be close to zero. In this case, the use of these usual estimators can cause spurious results. As a direct consequence, the interpretations of the ARCH and GARCH effect parameters in a time series, as well as the model's stationarity analysis and its autocorrelation function, may be wrong. Therefore, the use of the MQGARCH(1,1) model is strongly encouraged

References

- Bai, Z. D., C. R. Rao, and Y. H. Wu (1997). *M-estimation of multivariate linear regression by minimizing the difference of two convex functions*. Elsevier Science.
- Berkes, I., L. Hovath, and P. Kokoszka (2003). "GARCH process: Structure and estimation". In: *Bernoulli* 9(2), pp. 201–227. DOI: 10.3150/bj/1068128975.
- Bollerslev, Tim. (1986). "Generalized autoregressive conditional heteroscedasticity". In: *Journal of Econometrics*(31), pp. 307–327.
- Bose, A. and K. Mukherjee (2003). "Estimating the ARCH parameters by solving linear equations". In: *Journal of Time Series Analysis*(24), pp. 127–136.
- Breckling, Jens and Ray Chambers (1988). "M-quantiles". In: *Biometrika* 75(4), pp. 761–771. DOI: 10.1093/biomet/75.4.761.
- Carnero, M. A. (2003). "Heterocedasticidad condicional, Atípicos y Cambios de nivel en series temporales financieras". Universidad Carlos III de Madrid.
- Carnero, M. A., Daniel. Peña, and Esther Ruiz (2005). "Effects of Outliers on the Identification and Estimation of GARCH Models". In: *Journal of Time Series Analysis* 28, pp. 471–497.
- Carnero, M. A., Daniel. Peña, and Esther Ruiz (2012). "Estimating GARCH volatility in the presence of outliers". In: *Econometrics Letters* 114, pp. 86–90.
- Chang, Ih, George C. Tiao, and Chung Chen (1988). "Estimation of time series parameters in presence of outliers". In: *Technometrics* 30, pp. 1936–204.
- Chen, C. and L. Liu (1993). "Joint estimation of model parameters and outlier effects in time series". In: *Journal of the American Statistical Association* 89, pp. 284–297.
- Engle, R. F. (1982). "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation". In: *Econometrica*(50), pp. 987–1007.
- Fajardo, F., V. A. Reisen, C. Lévy-Leduc, et al. (2016). "Robust periodogram methods for time series with long-range dependence: An application to pollution levels". In: *Journal of statistical planning and inference*.
- Fajardo, F., V. A. Reisen, and Céline Lévy-Leduc (2018). "M-periodogram for the analysis of long-range-dependent time series". In: *Statistics A Journal of Theoretical and Applied Statistics* 52, pp. 665–683. DOI: <https://doi.org/10.1080/02331888.2018.1427751>.
- Francq, C. and J.-M. Zakoian (2019). *GARCH Models: Structure, Statistical Inference and Financial Applications*. Wiley.
- Franke, J., W. Härdle, and C. Hafner (2004). *Statistics of Financial Markets*. Berlin: Springer.

- Franses, Philip Hans and Hendrik Ghijsels (1999). “Additive Outliers, GARCH and Forecasting Volatility”. In: *International Journal Forecasting* 15, pp. 1–9.
- Gouriéroux, C. and J. Jasiak (2001). *Financial Econometrics*. Princeton University Press.
- Hamilton, James D. (1994). *Time Series Analysis*. Princeton University Press.
- Hu, Yu-Pin and Ruey S. Tsay (2014). “Principal Volatility Component Analysis”. In: *Journal of Business and Economic Statistics* 32. DOI: 10.1080/07350015.2013.818006.
- Huber, Peter J. (1964). “Robust Estimation of a Location Parameter”. In: *Annals of Statistics* 1(53), pp. 73–101. DOI: 10.1214/aoms/1177703732.
- Iqbal, F. (2013). “ROBUST ESTIMATION FOR THE ORTHOGONAL GARCH MODEL”. In: *The Manchester school* 81, pp. 904–924.
- Iqbal, F. and K. Mukherjee (2007). “Computation of M-Estimates in GARCH Models”. In: *Department of Mathematical Sciences, Liverpool University*.
- Ledolter, Johannes (1989). “The effect of additive outliers on the forecast from ARIMA models”. In: *International Journal of Forecasting* 5, pp. 231–240.
- Lee, Sangyeol and Jungsik Noh (2013). “Quantile Regression Estimator for GARCH Models”. In: *Scandinavian Journal of Statistics* 40(1), pp. 2–20. ISSN: 03036898, 14679469. DOI: 10.1111/j.1467-9469.2011.00759.x.
- Li, T. H. (2008). “Laplace periodogram for time series analysis”. In: *Journal of the American Statistical Association* 103, pp. 757–768.
- Ling, S. and M. McAleer (2003). “Asymptotic theory for a vector ARMA-GARCH model”. In: *Econometric Theory*(19), pp. 280–310.
- Mendes, S. V. M. (2000). “Assessing the bias of maximum likelihood estimates of contaminated GARCH models”. In: *Journal of Statistical Computation and Simulation* 67, pp. 359–376.
- Monte, Edson Zambon, Taciana Toledo de Almeida Albuquerque, and V. A. Reisen (2015). “Previsão da Concentração de Ozônio na Região da Grande Vitória, Espírito Santo, Brasil, Utilizando o Model ARMAX-GARCH”. In: *Revista Brasileira de Meteorologia* 30, pp. 285–294. DOI: 10.1590/0102-/10.1590/0102-778620140060.
- Mukherjee, K. (2008). “M-Estimation in GARCH Models”. In: *Econometric Theory* 24, pp. 1530–1553.
- Muler, N. and Victor J. Yohai (2008). “Robust Estimates for GARCH models”. In: *Journal of Statistical Planning and Inference* 138, pp. 2918–2940. DOI: <https://doi.org/10.1016/j.jspi.2007.11.003>.
- Reisen, V. A. and F. Fajardo (2012). “Robust estimation in time series with long and short memory properties”. In: *Annales Mathematicae et Informaticae* 39, pp. 207–224.
- Reisen, V. A., Céline Lévy-Leduc, and M.S. Taqqu (2017). “An M-estimator for the long-memory parameter”. In: *Journal of Statistical Planning and Inference* 187, pp. 44–55. DOI: <https://doi.org/10.1016/j.jspi.2017.02.008>.
- Reisen, V. A., Alessandro José Queiroz Sarnaglia, et al. (2014). “Modeling and forecasting daily average PM10 concentrations by a seasonal long-memory model with volatility”. In: *Environmental Modelling Software* 51, pp. 286–295. DOI: <https://doi.org/10.1016/j.envsoft.2013.09.027>.
- Sarnaglia, Alessandro José Queiroz et al. (2021). “M-regression spectral estimator for periodic ARMA models. An empirical investigation”. In: *Stochastic Environmental Research and Risk Assessment volume* 35, pp. 653–664.
- Weiss, Andrew A. (1986). “Asymptotic Theory for ARCH Models: Estimation and Testing”. In: *Econometric Theory* 2, pp. 107–131.

Wu, W. B. (2007). “M-estimation of linear models with dependent errors”. In: *The annals of statistics* 35, pp. 495–521.

7 CONCLUSÃO GERAL

Os estudos e resultados aqui apresentados foram motivados pela investigação de um estimador robusto para modelos heterocedásticos no contexto de séries temporais financeiras, em especial retornos de ativos financeiros. Os dados obtidos no mercados financeiros possuem a característica de conter outliers. A não consideração dessa característica pode levar a resultados equivocados que não retratam de modo real a verdadeira natureza das informações estudadas. Adicionalmente, modelos heterocedásticos baseiam-se no pressuposto de variáveis com distribuição condicional Gaussiana, propriedades também não observadas nos dados provenientes do mercado financeiro.

Os resultados deste trabalho estendem e corroboram com os resultados obtidos por Muler e Yohai (2008), Carnero, Peña e Ruiz (2012) e Iqbal (2013) e mostraram que os parâmetros de modelos heterocedásticos estimados por métodos usuais na literatura podem levar a interpretações espúrias. Por exemplo, a presença de outliers pode dificultar a detecção de heterocasticidade.

A contribuição científica dessa dissertação foi apresentada no artigo resultante do presente estudo. O artigo apresenta a fundamentação empírica do impacto de outliers aditivos no modelo GARCH(1,1) estimado pelo método M-Quantile, isto é, o modelo MQGARCH(1,1). Essa fundamentação foi feita por meio de experimentos de Monte Carlo. Foi mostrado que os parâmetros do modelo GARCH(1,1), amplamente utilizado na literatura, sofre efeitos notáveis pela contaminação da série, o que nesse caso pode levar a intepretações errôneas. Por outro lado, também foi mostrado que os parâmetros do modelo MQGARCH(1,1), principal tema dessa dissertação, manteve-se robusto quando a série foi contaminada, mostrando que o estimador M-Quantile, no contexto de modelos heterocedásticos, foi parcimonioso nesse caso.

Referências

- Bai, Z. D., C. R. Rao e Y. H. Wu (1997). *M-estimation of multivariate linear regression by minimizing the difference of two convex functions*. Elsevier Science.
- Bollerslev, Tim. (1986). «Generalized autoregressive conditional heteroscedasticity». Em: *Journal of Econometrics*(31), pp. 307–327.
- Breckling, Jens e Ray Chambers (1988). «M-quantiles». Em: *Biometrika* 75(4), pp. 761–771. DOI: [10.1093/biomet/75.4.761](https://doi.org/10.1093/biomet/75.4.761).
- Carnero, M. A. (2003). «Heterocedasticidad condicional, Atípicos y Cambios de nivel en series temporales financieras». Universidad Carlos III de Madrid.
- Carnero, M. A., Daniel. Peña e Esther Ruiz (2005). «Effects of Outliers on the Identification and Estimation of GARCH Models». Em: *Journal of Time Series Analysis* 28, pp. 471–497.
- Carnero, M. A., Daniel. Peña e Esther Ruiz (2012). «Estimating GARCH volatility in the presence of outliers». Em: *Econometrics Letters* 114, pp. 86–90.
- Chang, Ih, George C. Tiao e Chung Chen (1988). «Estimation of time series parameters in presence of outliers». Em: *Technometrics* 30, pp. 1936–204.
- Chen, C. e L. Liu (1993). «Joint estimation of model parameters and outlier effects in time series». Em: *Journal of the American Statistical Association* 89, pp. 284–297.
- Engle, R. F. (1982). «Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation». Em: *Econometrica*(50), pp. 987–1007.
- Fajardo, F. et al. (2016). «Robust periodogram methods for time series with long-range dependence: An application to pollution levels». Em: *Journal of statistical planning and inference*.
- Francq, C. e J.-M. Zakoian (2019). *GARCH Models: Structure, Statistical Inference and Financial Applications*. Wiley.
- Franses, Philip Hans e Hendrik Ghijssels (1999). «Additive Outliers, GARCH and Forecasting Volatility». Em: *Internatonal Journal Forecasting* 15, pp. 1–9.
- Huber, Peter J. (1964). «Robust Estimation of a Location Parameter». Em: *Annals of Statistics* 1(53), pp. 73–101. DOI: [10.1214/aoms/1177703732](https://doi.org/10.1214/aoms/1177703732).

- Iqbal, F. (2013). «ROBUST ESTIMATION FOR THE ORTHOGONAL GARCH MODEL». Em: *The Manchester school* 81, pp. 904–924.
- Koenker, Roger e Zhijie Xiao (2009). «Conditional Quantile Estimation for Generalized Autoregressive Conditional Heteroscedasticity Models». Em: *Taylor Francis, Ltd* 104(488), pp. 1696–1712. DOI: [10.1198/jasa.2009.tm09170](https://doi.org/10.1198/jasa.2009.tm09170).
- Koenker, Roger e Quanshui Zhao (1996). «Conditional Quantile Estimation and Inference for Arch Models». Em: *Econometric Theory* 12(5), pp. 793–813. DOI: [10.1017/S0266466600007167](https://doi.org/10.1017/S0266466600007167).
- Ledolter, Johannes (1989). «The effect of additive outliers on the forecast from ARIMA models». Em: *International Journal of Forecasting* 5, pp. 231–240.
- Lee, Sangyeol e Jungsik Noh (2013). «Quantile Regression Estimator for GARCH Models». Em: *Scandinavian Journal of Statistics* 40(1), pp. 2–20. ISSN: 03036898, 14679469. DOI: [10.1111/j.1467-9469.2011.00759.x](https://doi.org/10.1111/j.1467-9469.2011.00759.x).
- Li, T. H. (2007). «Laplace periodogram for time series analysis». Em: *Journal of the American Statistical Association* 103, pp. 757–768.
- Mendes, S. V. M. (2000). «Assessing the bias of maximum likelihood estimates of contaminated GARCH models». Em: *Journal of Statistical Computation and Simulation* 67, pp. 359–376.
- Monte, Edson Zambon, Taciana Toledo de Almeida Albuquerque e V. A. Reisen (2015). «Previsão da Concentração de Ozônio na Região da Grande Vitória, Espírito Santo, Brasil, Utilizando o Model ARMAX-GARCH». Em: *Revista Brasileira de Meteorologia* 30, pp. 285–294. DOI: [10.1590/0102-778620140060](https://doi.org/10.1590/0102-778620140060).
- Muler, N. e Victor J. Yohai (2008). «Robust Estimates for GARCH models». Em: *Journal of Statistical Planning and Inference* 138, pp. 2918–2940. DOI: <https://doi.org/10.1016/j.jspi.2007.11.003>.
- Reisen, V. A. e F. Fajardo (2012). «Robust estimation in time series with long and short memory properties». Em: *Annales Mathematicae et Informaticae* 39, pp. 207–224.
- Reisen, V. A., Alessandro José Queiroz Sarnaglia et al. (2014). «Modeling and forecasting daily average PM10 concentrations by a seasonal long-memory model with volatility». Em: *Environmental Modelling Software* 51, pp. 286–295. DOI: <https://doi.org/10.1016/j.envsoft.2013.09.027>.

Reisen, V. A. e A. N. Silva (2011). *O uso da linguagem R para cálculos de estatística básica*. EDUFES.

Wu, W. B. (2007). «M-estimation of linear models with dependent errors». Em: *The annals of statistics* 35, pp. 495–521.