Causalidade de Granger em medidas de risco

Patricia Nagami Murakami

DISSERTAÇÃO APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO
DE
MESTRE EM CIÊNCIAS

Programa: Estatística

Orientador: Prof^a. Dr^a. Clelia Maria de Castro Tolo

São Paulo, Fevereiro de 2011

Causalidade de Granger em medidas de risco

Este exemplar corresponde à redação
final da dissertação devidamente corrigida
e defendida por Patricia Nagami Murakami
e aprovada pela Comissão Julgadora.

Banca Examinadora:

- Prof. Dra. Patricia Nagami Murakami - IME-USP.
- Prof. Dra. Chang Chiann - IME-USP.
- Prof. Dr. João Ricardo Sato - UFABC.

Agradecimentos

Agradeço a todos que de alguma forma participaram na jornada da elaboração desse trabalho. Em especial, agradeço à minha família, por todo o apoio ao longo não só desse período, mas sempre presente em todos os momentos. À minha orientadora Clélia Toloí, pelos ensinamentos, pela paciência e compreensão durante todo esse tempo. Ao Josué Rocha, por me incentivar e motivar nos estudos do mestrado. Às queridas colegas de (muitos) estudos e amigas Francielle Lima e Liliane Travassos, que sem dúvida me auxiliaram muito no desenvolvimento desse trabalho. Ao Rômulo, pelo apoio e compreensão nos momentos mais difíceis. Às amigas Verônica Amparo e Michele Figueiró que estiveram sempre presentes.

Resumo

Esse trabalho apresenta um estudo da causalidade de Granger em Risco bivariado aplicado a séries temporais financeiras. Os eventos de risco, no caso de séries financeiras, estão relacionados com a avaliação do Valor em Risco das posições em ativos. Para isso, os modelos CaViaR, que fazem parte do grupo de modelos de Regressão Quantílica, foram utilizado para identificação desses eventos. Foram expostos os conceitos principais envolvidos da modelagem, assim como as definições necessárias para entendê-las. Através da análise da causalidade de Granger em risco entre duas séries, podemos investigar se uma delas é capaz de prever a ocorrência de um valor extremo da outra. Foi realizada a análise de causalidade de Granger usual somente para como comparativo.

Palavras-chave: Regressão Quantílica, Valor em Risco, Modelos CAViaR, Causalidade de Granger, Causalidade de Granger em Risco.

Abstract

The present work presents a study of bivariate Granger causality in risk applied to financial time series. Risk events, considering financial series, are related to the VaR evaluation in the assets open interest. The CaViaR models, which are part of Quantilic Regression models, were used to identify these events. We exposed the main concepts involved at the modelling steps and some useful definitions in order to understand it. With the analysis of Granger causality in risk between two series, it is possible to investigate if one of the series is capable of predict the occurrence of a extreme value of the other. The study of the usual Granger causality was realized only to compare the results.

Keywords: Quantile Regression, Value at Risk, CAViaR Model, Granger Causality, Granger Causality in Risk.

Sumário

Lista de Figuras	xi
Lista de Tabelas	xiii
1 Introdução	1
2 Conceitos	3
2.1 Mensuração do Risco	3
2.1.1 Retornos	3
2.1.2 Volatilidade	4
2.1.3 Valor em Risco (VaR)	5
2.1.4 Conditional Autoregressive Value at Risk (CaViaR)	7
2.1.5 Estimação	9
2.1.6 Testes de Adequação do Modelo	11
2.1.7 Medidas auxiliares	15
2.2 Causalidade	16
2.2.1 Causalidade de Granger	17
2.2.2 Causalidade de Granger em Risco	20
3 Metodologia	23
3.1 Hipótese de Causalidade de Granger em Risco	23
3.2 Causalidade de Granger em Risco com Feedback	31

4	Aplicação	35
4.1	Descrição dos dados utilizados	35
4.2	Cálculo do VaR - CaViaR	39
4.3	Ajuste e Adequação para série de log-retornos da BBDC4	40
4.4	Ajuste e Adequação para série de log-retornos da PETR4	45
4.5	Análise de Causalidade de Granger em Risco	55
4.6	Análise de Causalidade de Granger	57
5	Considerações Finais	69
A	Suposições para Teoremas	71
B	Teoremas	73
	Referências Bibliográficas	77

Lista de Abreviaturas

VaR	Valor em Risco.
CaViaR	Conditional Autoregressive Value at Risk
BBDC4	Código de negociação da ação do Banco Bradesco
PETR4	Código de negociação da ação da Petrobrás
RV_{CI}	Razão de Verossimilhança do Teste de Cobertura Incondicional
RV_{ind}	Razão de Verossimilhança do Teste de Independência
RV_{CI}	Razão de Verossimilhança do Teste de Cobertura Condicional
DQ	Estatística do teste do Quantil Dinâmico
CQ	Critério Quantílico
FPE	Final Prediction Error
MQO	Mínimos Quadrados Ordinários

Lista de Figuras

2.1	VaR de posição comprada para retornos distribuídos com f.d.p Normal(0,1).	6
3.1	Núcleo Truncado (H=1).	27
3.2	Núcleo de Bartlett (H=1).	27
3.3	Núcleo de Daniell (H=1).	27
3.4	Núcleo de Parzen (H=1).	28
3.5	Núcleo Truncado (H=1; H=1,25 ; H=1,5 ; H=1,75 ; H=2).	29
3.6	Núcleo de Bartlett (H=1; H=1,25 ; H=1,5 ; H=1,75 ; H=2).	29
3.7	Núcleo de Daniell (H=1; H=1,25 ; H=1,5 ; H=1,75 ; H=2).	29
3.8	Núcleo de Parzen (H=1; H=1,25 ; H=1,5 ; H=1,75 ; H=2).	29
4.1	Série de preços das ações BBDC4 e PETR4.	36
4.2	Série de log-retornos das ações BBDC4 e PETR4	38
4.3	Histograma dos log-retornos para BBDC4 e PETR4.	39
4.4	Q-Q plot normal dos log-retornos para BBDC4 e PETR4.	39
4.5	VaR estimado por valor absoluto simétrico e série de log-retornos de BBDC4	41
4.6	VaR estimado por inclinação assimétrica e série de log-retornos de BBDC4	42
4.7	VaR estimado por GARCH(1,1) indireto e série de log-retornos de BBDC4	43
4.8	VaR estimado por modelo adaptativo e série de log-retornos de BBDC4	44

4.9	VaR estimado por valor absoluto simétrico e série de log-retornos de BBDC4	46
4.10	VaR estimado por inclinação assimétrica e série de log-retornos de BBDC4	47
4.11	VaR estimado por GARCH(1,1) indireto e série de log-retornos de BBDC4	48
4.12	VaR estimado por inclinação assimétrica e série de log-retornos de BBDC4	49
4.13	Valores estimados do FPE para $i = 1, 2, \dots, 24$ na equação (4.3)	59
4.14	Valores estimados do FPE para $M_{12} = 1, 2, \dots, 18$ na equação (4.5) . .	61
4.15	Valores estimados do FPE para $i \in I_1$ na equação (4.7)	63
4.16	Valores estimados do FPE para $i = 1, 2, \dots, 24$ na equação (4.8)	64
4.17	Valores estimados do critério FPE para $i = 1, 2, \dots, 18$	67

Lista de Tabelas

4.1	Datas e eventos para tratamento de BBDC4	37
4.2	Datas e eventos para tratamento de PETR4	37
4.3	Estatísticas descritivas amostrais das séries de log-retornos das ações BBDC4 e PETR4	38
4.4	Estimativas dos Parâmetros Estimados para os Modelos CAViaR para a série BBDC4.	50
4.5	continuação - Estimativas de Estatísticas de Teste dos Parâmetros Estimados para os Modelos CAViaR para a série BBDC4.	51
4.6	Estimativas dos Parâmetros Estimados para os Modelos CAViaR para a série PETR4.	52
4.7	continuação - Estimativas de Estatísticas de Teste dos Parâmetros Estimados para os Modelos CAViaR para a série PETR4.	54
4.8	Estimativas das Estatísticas de Teste de Causalidade de Granger em Risco (BBDC4 $\Rightarrow$ PETR4) para núcleos truncado, Bartlett, Daniell e Parzen com M variando em 5, 10, 20, 30 e 40	56
4.9	Estimativas das Estatísticas de Teste de Causalidade de Granger em Risco (PETR4 $\Rightarrow$ BBDC4) para núcleos truncado, Bartlett, Daniell e Parzen com M variando em 5, 10, 20, 30 e 40	56
4.10	Estimativas das Estatísticas de Teste de Causalidade de Granger em Risco (BBDC4 $\Leftrightarrow$ PETR4) para núcleos truncado, Bartlett, Daniell e Parzen com M variando em 5, 10, 20, 30 e 40	57
4.11	Valores estimados do FPE para $i = 1, 2, \dots, 24$ na equação (4.3)	58
4.12	Valores estimados do modelo completo (4.3) para s=18	60

4.13	Valores estimados do modelo reduzido (4.4)	60
4.14	Valores estimados do FPE para $i = 1, 2, \dots, 18$ na equação (4.5)	62
4.15	Valores estimados do modelo (4.5) para $n=5$	63
4.16	Valores estimados do FPE para $M_{22} = 1, 2, \dots, 24$ na equação (4.8)	65
4.17	Valores estimados do modelo completo (4.8) para $s=18$	66
4.18	Valores estimados do FPE para $s=18$ e $i = 1, 2, \dots, 18$ na equação (4.9)	66
4.19	Valores estimados do modelo completo (4.9) para $s = 18$ e $n = 1$	68

Capítulo 1

Introdução

Identificar, controlar e monitorar eventos extremos de investimentos financeiros são ações essenciais para o gerenciamento de riscos. Um exemplo comum é a avaliação da posição na carteira de ações de um investidor. Com flutuação de preços devido às movimentações do mercado, o ativo sofre valorizações e desvalorizações ao longo de um período. O histórico de variações de preços será útil para identificar certos padrões de comportamento dos ativos, auxiliando no estudo de ocorrência de eventos extremos.

Neste trabalho, o foco do estudo será entender o mecanismo de ocorrência de eventos extremos entre diferentes mercados. Será apresentada uma abordagem para Causalidade de Granger em detecção de riscos extremos entre duas séries financeiras. Como medida de evento extremo, utilizou-se o VaR (*Value at Risk*), que é uma medida amplamente conhecida no mercado como referência para identificação de um evento de risco.

A idéia de se utilizar preços de ações relacionados a duas empresas de ramos distintos é verificar se existe alguma relação de causalidade entre diferentes setores. Se isso for verificado, as informações de uma delas poderia ser utilizada para prever a ocorrência de algum evento extremo da outra.

O estudo de Causalidade de Granger em Risco deste trabalho está essencialmente dividido em três capítulos: o primeiro abrange os conceitos de risco e causalidade, o segundo expõe a metodologia de avaliação da Causalidade de Granger e o terceiro é

a aplicação da metodologia utilizando dados reais de séries de preços de ações.

O Capítulo 2 de Conceitos expõe: as definições de retorno de ativo, que é a forma de mensuração dos preços de ações; o VaR, que representa uma medida de evento de risco e os modelos CAViaR, que foram propostos por Engle and Manganelli (2004) [2] com o fim de modelar diretamente os quantis dos retornos, ao invés de considerar toda distribuição de retornos. Além disso, são expostos os testes de adequação do modelo e definidas medidas auxiliares para a escolha do melhor modelo.

A metodologia exposta no Capítulo 3 apresenta a formulação das hipóteses de Causalidade de Granger em Risco e Causalidade de Granger em Risco com *feedback*, propostas por Hong, Liu e Wang (2009) [7]. Nele, estão contidas as definições dos componentes da estatística de teste como as funções de autocorrelação cruzada e os núcleos estudados (truncado, de Bartlett, Daniell e Parzen).

No Capítulo 4, na aplicação da metodologia proposta, foram utilizados dados de preços das ações negociadas na bolsa de valores das empresas Bradesco e Petrobrás, para estudo da causalidade de Granger em risco. Primeiramente constam os tratamentos a serem realizados nos dados e análise descritiva. A seguir, constam os passos da modelagem do cálculo do VaR utilizando os modelos CaViaR. Após escolha do melhor modelo aplicado às duas séries, as séries transformadas são utilizadas para o teste de causalidade de Granger em Risco. Foi realizado o teste de causalidade de Granger usual somente para efeitos comparativos, seguindo metodologia proposta por Hsiao (1979) [6] e Pierce e Haugh (1977) [15].

Os códigos desenvolvidos nos softwares R e Matlab podem ser encontrados na página <http://patricianm.yolasite.com/>.

Finalmente, o Capítulo 5 apresenta as considerações finais do trabalho, com o resumo dos principais resultados obtidos.

Capítulo 2

Conceitos

2.1 Mensuração do Risco

O controle da efetividade do gerenciamento de risco é de grande importância para instituições financeiras. Como a natureza dos riscos muda no decorrer do tempo, os métodos para mensurá-los devem se adequar às experiências atuais.

Medidas quantitativas são ferramentas importantes para tomada de decisões de investimentos, alocação de capital e exigências regulatórias.

Em 1996, o Comitê de Supervisão Bancária, com o Acordo de Basiléia, passou a exigir que instituições financeiras tenham requisitos de capitais suficientes para cobrir os riscos que incorrem como resultado de suas negociações. Sendo assim, se os riscos não forem efetivamente mensurados, a alocação de capital para a cobertura dos mesmos pode ser subestimada ou superestimada.

O resultado de uma alocação não eficiente de recursos poderia ocasionar perdas, não só com a ocorrência de um evento de risco, mas como a perda de oportunidade de investimento.

2.1.1 Retornos

A avaliação do risco de uma carteira de ativos é comumente realizada sobre a variação de preços. Seja P_t o preço de um ativo no instante t . Essa variação pode ser vista de alguns modos:

- *retorno líquido simples* ou *taxa de retorno*: é a variação do preço relativamente

ao preço imediatamente anterior, geralmente expressa em percentagem

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

- *retorno bruto simples*: é a razão entre os preços atual e anterior

$$1 + R_t = \frac{P_t}{P_{t-1}}$$

- *retorno composto continuamente* ou *log-retorno*: é a forma mais comum e muitas vezes denominada somente de *retorno*. Reflete a diferença do log do preço atual com o log do preço imediatamente anterior

$$r_t = \log \left(\frac{P_t}{P_{t-1}} \right).$$

Por definição, os retornos não dependem da escala do ativo financeiro. Assim, é possível comparar retornos de ativos que têm patamares distintos, mas que sofrem oscilações de mercado semelhantes. Outra vantagem importante é que os retornos apresentam propriedades estatísticas importantes como a estacionariedade e ergodicidade, por exemplo.

2.1.2 Volatilidade

Volatilidade é o termo dado à variância condicional dos retornos. É comumente utilizada no mercado financeiro como uma medida que reflete o grau de aquecimento do mercado para uma janela de tempo. Séries de retornos podem apresentar agrupamentos de volatilidades ao longo do tempo, o que é explicado pelo comportamento do mercado persistente durante certos períodos. Como descrito em Morettin(2008) [13], os três enfoques para cálculo de volatilidades são:

- *volatilidade implícita*: volatilidade obtida em geral através da fórmula de Black-Scholes na precificação de opções européias;

- *volatilidade estatística*: volatilidade expressa por modelos estatísticos, como os modelos da família ARCH;
- *volatilidade histórica*: volatilidade obtida por uma média de uma função dos últimos k retornos:

$$v_t = \left[\sum_{j=0}^{k-1} |r_{t-j}|^p \right]^{1/p}.$$

2.1.3 Valor em Risco (VaR)

O VaR (*Value at Risk*) é uma medida de risco de mercado amplamente utilizada para controlar e monitorar o risco de mercado. Ela foi desenvolvida no começo da década de 90 pela indústria financeira com o objetivo de fornecer um valor que contivesse informação sobre o risco de uma carteira de ativos. Em Jorion (1999) [9], essa medida é amplamente discutida.

Essa medida expressa o grau de incerteza quanto aos possíveis valores de perda. Ela representa a variação potencial máxima, dado um horizonte de tempo, e uma probabilidade, de que essa variação possa ocorrer.

Embora não seja uma medida perfeita para acompanhar o risco de mercado com variações extremas, essa medida é calculada como procedimento padrão pela maioria das entidades financeiras.

A simplicidade conceitual fez com que o VaR se tornasse uma medida popular de risco, mesmo com suas fraquezas. Existem várias abordagens para cálculo do VaR. Na próxima seção, será descrita detalhadamente a escolhida para estudo neste trabalho.

Definição 2.1.1 (Medida do Valor em Risco (VaR)) *Sejam r_{t+h} o retorno do ativo em horizonte h , τ a probabilidade de ocorrer um evento extremo, onde $0 < \tau < 1$ e $\mathcal{F}_t$ representa a informação disponível até tempo t . O valor em risco de um ativo detido em posição comprada, é o valor VaR tal que:*

$$VaR = \inf_x \{x \in \mathbb{R} : Prob(r_{t+h} \leq x | \mathcal{F}_t) \leq \tau\}$$

ou seja, o retorno do ativo seja inferior ao VaR em um determinado horizonte de

tempo h com probabilidade máxima igual a τ . Para uma posição vendida, a perda ocorre se o preço do ativo aumenta. Logo, deverá ser utilizada a cauda direita da distribuição:

$$VaR = \sup_x \{x \in \mathbb{R} : Prob(r_{t+h} \geq x | \mathcal{F}_t) \leq \tau\}.$$

A idéia é instituir uma medida de risco de perda, na qual define-se um valor limite de perda com uma probabilidade τ baixa de que haja perda que exceda esse limite. Em geral, utiliza-se uma probabilidade entre 1% e 5%. No gráfico 2.1, temos a representação do VaR de uma posição comprada de retornos distribuídos como Normal(0,1) para uma probabilidade $\tau = 5\%$: $VaR = -1,65$.

Figura 2.1: VaR de posição comprada para retornos distribuídos com f.d.p Normal(0,1).

O VaR representa um quantil particular dos valores futuros de uma carteira, dadas as informações disponíveis. A grande questão está em encontrar um modelo adequado para quantis que variam condicionalmente no decorrer do tempo.

Existem diversas metodologias para cálculo do VaR. Os passos básicos concentram-se em:

1. escolher a fórmula de cálculo para VaR, como função de variáveis conhecidas e observadas até o tempo $t-1$ e com parâmetros a serem estimados;
2. escolher um procedimento para estimar o conjunto de parâmetros, em geral através de um algoritmo de otimização e uma função de perda;

3. escolher um teste para avaliar a qualidade das estimativas.

2.1.4 Conditional Autoregressive Value at Risk (CaViaR)

Em geral, a volatilidade dos retornos de um ativo em um certo período está correlacionada com a volatilidade em período passado. Sendo assim, o VaR deve apresentar comportamento similar, já que é uma medida ligada ao desvio padrão da distribuição. Uma maneira de capturar essa autocorrelação é utilizar um modelo no qual a medida de VaR pode ser dependente de VaR passados.

O modelo CaViaR baseia-se na utilização de uma estrutura autorregressiva dos quantis, que representam o VaR. Ao invés de utilizar a distribuição de retornos, esse modelo utiliza diretamente os quantis.

Suponha que a série de retornos de um ativo em um período seja $\{r_t\}_{t=1}^n$. $\mathbf{x}_t$ é uma matriz de p variáveis observáveis que possam ter influência sobre os retornos. Para um nível τ de probabilidade de exceder o VaR, o τ -quantil da distribuição de retornos no tempo $t-1$ é $Q_{t-1}(\beta_\tau) = Q(x_{t-1}, \beta_\tau)$, no qual β_τ é o vetor de parâmetros do modelo. Por simplicidade notacional, será utilizado que $\beta_\tau = \beta$.

Como proposto por Engle e Manganelli (2002) [2], a especificação geral do modelo de quantis autorregressivos é:

$$Q_t(\beta) = \alpha + \sum_{i=1}^q \gamma_i Q_{t-i}(\beta) + \sum_{i=1}^p \phi_i L(\mathbf{x}_{t-i}) \quad (2.1)$$

em que $\beta = (\alpha, \gamma_1, \dots, \gamma_q, \phi_1, \dots, \phi_p)$ é o vetor de parâmetros a serem estimados e $L(\cdot)$ é uma função que gera defasagens.

Os termos dos quantis autorregressivos $\gamma_i Q_{t-i}(\beta)$, para $i = 1, \dots, q$, servem para adaptar a dependência serial na média e volatilidade, além de assegurar que a mudança no quantil seja suave ao longo do tempo. Os termos de valores observados defasados $\phi_i L(\mathbf{x}_{t-i})$, $i = 1, \dots, p$ têm o objetivo de ligar o quantil condicional às inovações no retorno do ativo. Em geral, utiliza-se a própria série de retornos do ativo.

Uma das vantagens é que não é necessário considerar alguma suposição sobre a distribuição de retornos. Além disso, há a flexibilidade em adaptar diversas especi-

ficações de modelo de acordo com o próprio comportamento da série de retornos. Engle e Manganelli (2002)[2] definem algumas especificações de referência para o modelo CaViaR.

Entre os diversos tipos de modelos que podem ser ajustados, temos:

1. valor absoluto simétrico:

$$Q_t(\boldsymbol{\beta}) = \alpha + \gamma Q_{t-1}(\boldsymbol{\beta}) + \phi |r_{t-1}|$$

2. inclinação assimétrica:

$$Q_t(\boldsymbol{\beta}) = \alpha + \gamma Q_{t-1}(\boldsymbol{\beta}) + \phi_1(r_{t-1})^+ + \phi_2(r_{t-1})^-$$

em que $(x)^+ = \max(x, 0)$ e $(x)^- = -\min(x, 0)$

3. GARCH(1,1) indireto:

$$Q_t(\boldsymbol{\beta}) = (\alpha + \gamma Q_{t-1}^2(\boldsymbol{\beta}) + \phi r_{t-1}^2)^{\frac{1}{2}}$$

4. adaptativo

$$Q_t(\beta_1) = Q_{t-1}(\beta_1) + \beta_1 \left\{ \frac{1}{[1 + e^{G(r_{t-1} - Q_{t-1}(\beta_1)) - \tau}]} \right\}$$

em que G é uma constante.

Eles podem ser escolhidos de acordo com o tratamento desejado a ser dado com relação aos retornos. A utilização do *inclinação assimétrica* permite que retornos positivos e negativos sejam tratados com pesos diferentes. Por outro lado, o impacto em módulo presente no *valor absoluto simétrico* e o quadrado do retorno passado na especificação *GARCH indireto* fazem com que o modelo responda simetricamente. No modelo *adaptativo*, o quantil no período t é imediatamente maior se o retorno

passado excede o VaR; se não excede, o quantil decresce lentamente. Para G grande, há pouca captura da influência dos retornos próximos ao VaR e dos extremamente positivos.

2.1.5 Estimação

A estimação dos parâmetros do modelo é realizada via método de *regressão quantílica*. Em Koenker (2005) [11] e Koenker e Bassett (1978) [10] constam os conceitos para modelagem dos quantis como um problema de minimização.

A regressão quantílica se traduz no problema de minimização da soma dos resíduos ponderados assimetricamente em valor absoluto.

Para entender a idéia, veja que para uma amostra aleatória $\{y_1, y_2, \dots, y_T\}$, a mediana é um valor ξ que minimiza a soma de valores absolutos dos desvios:

$$mediana = \operatorname{argmin}_{\xi \in \mathbb{R}} \sum_{i=1}^T |y_i - \xi|.$$

Atribui-se pesos para as observações a partir de uma função $\rho_\tau(w)$ tal que

$$\rho_\tau(w) = \begin{cases} w\tau & \text{se } w \geq 0 \\ w(\tau - 1) & \text{se } w < 0, \end{cases}$$

para $0 < \tau < 1$.

O τ -ésimo quantil amostral $\xi(\tau)$, é obtido através da minimização dos resíduos ponderados por $\rho_\tau(\cdot)$:

$$\xi(\tau) = \operatorname{argmin}_{\xi \in \mathbb{R}} \sum_{i=1}^T \rho_\tau(y_i - \xi).$$

A função atribui um peso τ para resíduos positivos $y_i - \xi$ e peso $\tau - 1$ para resíduos negativos.

Seja uma amostra de observações $\{y_1, y_2, \dots, y_T\}$ gerados pelo modelo

$$y_t = \mathbf{x}_t' \boldsymbol{\beta} + \epsilon_{\tau t}, \quad Q_\tau(\epsilon_{\tau t} | \mathbf{x}_t) = 0$$

em que $\mathbf{x}_t$ é a matriz de p regressores e $Q_\tau(\epsilon_{\tau t}|\mathbf{x}_t)$ é o τ -quantil de $\epsilon_{\tau t}$ condicional a $\mathbf{x}_t$.

Utilizando a função de perda, a função de quantil condicional estende o τ -ésimo quantil amostral $\xi(\tau)$ para a regressão da mesma maneira que a função de média amostral estende para a média condicional.

A regressão de MQO estima a média condicional da série $E(Y|X = x) = \mathbf{x}'\beta$ resolvendo o problema de minimização:

$$\hat{\beta} = \underset{\beta \in \mathbb{R}^p}{\operatorname{argmin}} \sum_{i=1}^T (y_i - \mathbf{x}_i'\beta)^2.$$

O parâmetro estimado $\hat{\beta}$ minimiza a soma de quadrados dos resíduos da regressão da mesma maneira que a média amostral $\hat{\mu}$ minimiza a soma de quadrados:

$$\hat{\mu} = \underset{\mu \in \mathbb{R}}{\operatorname{argmin}} \sum_{i=1}^T (y_i - \mu)^2.$$

A regressão quantílica estima a função condicional linear $f(\tau|X = x) = \mathbf{x}_t'\beta_\tau$ resolvendo o problema de minimização:

$$\hat{\beta}_\tau = \underset{\beta \in \mathbb{R}^p}{\operatorname{argmin}} E[\rho_\tau(y_i - \mathbf{x}_i'\beta)]$$

ou seja,

$$\hat{\beta}_\tau = \underset{\beta \in \mathbb{R}^p}{\operatorname{argmin}} \frac{1}{T} \sum_{i=1}^T \rho_\tau(y_i - \mathbf{x}_i'\beta)$$

para um quantil qualquer $\tau \in (0, 1)$. $\hat{\beta}_\tau$ é chamado de τ -ésimo quantil da regressão.

Outra maneira de expressar o problema é:

$$\hat{\beta}_\tau = \underset{\beta \in \mathbb{R}^p}{\operatorname{argmin}} \left[\sum_{i \in \{i: y_i \geq \mathbf{x}_i'\beta\}} \tau |y_i - \mathbf{x}_i'\beta| + \sum_{i \in \{i: y_i < \mathbf{x}_i'\beta\}} (1 - \tau) |y_i - \mathbf{x}_i'\beta| \right].$$

O τ -ésimo quantil da regressão é uma extensão do τ -ésimo quantil amostral, $\xi(\tau)$.

A regressão quantílica inclui o caso do modelo de regressão por menor desvio absoluto, LAD (*Least Absolute Deviation*). Isso ocorre quando $\tau = 0,5$. A regressão por LAD é conveniente por apresentar estimadores mais robustos que a por MQO. O problema de minimização é obtido através de algoritmos de programação linear.

Para relacionar a regressão quantílica com o modelo CaViar, temos que considerar o seguinte modelo:

$$y_t = Q(y_{t-1}, \mathbf{x}_{t-1}, \dots, y_1, \mathbf{x}_1; \boldsymbol{\beta}^0) + \epsilon_{t\tau} = Q_t(\boldsymbol{\beta}^0) + \epsilon_{t\tau} \quad \text{com} \quad Q_\tau(\epsilon_{\tau t} | \Omega_t) = 0 \quad (2.2)$$

em que $Q_1(\boldsymbol{\beta}^0)$ é condição inicial dada, $\Omega_t = \{y_{t-1}, \mathbf{x}_{t-1}, \dots, y_1, \mathbf{x}_1, Q_1(\boldsymbol{\beta}^0)\}$, $\mathbf{x}_t$ é a matriz de variáveis exógenas, $y_{t-1}, \dots, y_1$ são os fatores autorregressivos de y_t , incluídos agora no modelo, $\boldsymbol{\beta}^0 \in \mathbb{R}^p$ é o vetor de parâmetros desconhecidos a serem estimados.

O objetivo é encontrar $\hat{\boldsymbol{\beta}}$ tal que

$$\min_{\boldsymbol{\beta} \in \mathbb{R}^p} \frac{1}{T} \sum_{i=1}^T [\tau - \mathbb{I}(y_t < Q_t(\boldsymbol{\beta}))] [y_t - Q_t(\boldsymbol{\beta})]. \quad (2.3)$$

Existem dois teoremas, enunciados no Apêndice B, que têm como resultado que o estimador $\hat{\boldsymbol{\beta}}$ é consistente e assintoticamente normal.

2.1.6 Testes de Adequação do Modelo

Espera-se que a série de valores que ultrapassaram a estimativa do VaR, dado um modelo, tenha algumas características. Uma delas é que ela seja uma sequência serialmente independente. Um resumo dos testes que serão descritos a seguir consta em Kuester, Mittnik e Paolella (2005) [12].

Para avaliar a adequação do modelo, uma das abordagens é avaliar a acurácia da previsão fora do intervalo (*out-of-sample*).

A idéia básica é que a proporção de acertos do modelo deve ser equivalente ao percentual de corte τ . Seja $I_t = \mathbb{I}[y_t < Q_t(\boldsymbol{\beta}^0)]$, para $i = 1, 2, \dots$, a série de violações

ao VaR para uma posição comprada. Christoffersen(1998) [1] diz que as previsões são *eficientes* com respeito a $\mathcal{F}_{t-1}$ se

$$E[I_t | \mathcal{F}_{t-1}] = \tau \quad (2.4)$$

que, ao aplicar esperanças iterativamente, implica que $\{I_t\}$ é não correlacionada com qualquer função de variáveis no conjunto de informação disponível em t-1. Iremos destacar alguns testes para verificar a adequação do modelo.

TA1. Teste de Cobertura Incondicional

O teste de cobertura incondicional verifica se o modelo proporciona uma média (incondicional) de violações do VaR correspondente à probabilidade pré-especificada de ocorrer um evento extremo. Tomando as esperanças condicionais iterativamente em (2.4), Christoffersen(1998) [1] propõe o seguinte teste de hipóteses:

$$H_0 : E[I_t] \leq \tau$$

$$H_1 : E[I_t] > \tau$$

utilizando a seguinte de razão de verossimilhanças:

$$RV_{CI} = 2 [\mathcal{L}(\hat{\tau}; I_1, I_2, \dots, I_T) - \mathcal{L}(\tau; I_1, I_2, \dots, I_T)] \stackrel{a}{\sim} \chi_1^2$$

na qual $\mathcal{L}(\cdot)$ é o log da verossimilhança binomial, $\hat{\lambda}$ é a proporção do número de violações com relação ao total de observações T: $\hat{\lambda} = \frac{n_1}{T}$, para a qual $n_1 = \sum_{i=1}^T \mathbb{I}[I_i = 1]$.

TA2. Teste de Independência

Uma das características esperadas para as previsões obtidas para VaR, é que não haja dependência entre os valores que excederam a medida do VaR. Para detectar se há dependência entre as previsões de VaR, será utilizado o teste proposto por Christoffersen(1998) [1].

A hipótese nula é que não há dependência entre uma violação do VaR hoje e no dia seguinte. Para a alternativa, é definida uma cadeia de probabilidade de transições sobre os estados de $\{I_t\}$

$$\mathbf{\Pi} = \begin{bmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{bmatrix}, \quad \pi_{ij} = P(I_t = j | I_{t-1} = i)$$

em que π_{ij} é a probabilidade de se passar do estado i no tempo $t-1$ para o estado j em t .

A idéia é comparar a log-verossimilhança quando se supõe dependência ou quando se supõe independência da série de violações utilizando a seguinte estatística:

$$RV_{ind} = 2 \left[\mathcal{L}(\hat{\Pi}; I_2, \dots, I_T | I_1) - \mathcal{L}(\hat{\pi}_0; I_2, \dots, I_T | I_1) \right],$$

$$n_{ij} = \sum_{i=2}^T \mathbb{I}(I_t = i | I_{t-1} = j) \text{ para } i=0,1 \text{ e } j=0,1 \text{ e:}$$

1. a log-verossimilhança supondo a hipótese alternativa é

$$\mathcal{L}(\hat{\Pi}; I_2, \dots, I_T | I_1) = (1 - \hat{\pi}_{01})^{n_{00}} \hat{\pi}_{01}^{n_{01}} (1 - \hat{\pi}_{11})^{n_{10}} \hat{\pi}_{11}^{n_{11}}$$

$$\text{em que } \hat{\pi}_{01} = \frac{n_{01}}{n_{00} + n_{01}} \text{ e } \hat{\pi}_{11} = \frac{n_{11}}{n_{10} + n_{11}}$$

2. a log-verossimilhança supondo a hipótese nula de independência é

$$\mathcal{L}(\hat{\pi}_0; I_2, \dots, I_T | I_1) = (1 - \hat{\pi}_0)^{n_{00} + n_{10}} \hat{\pi}_0^{n_{01} + n_{11}}$$

$$\text{em que } \hat{\pi}_0 = \frac{n_{01} + n_{11}}{T}.$$

TA3. Teste de Cobertura Condicional

Uma versão combinada dos testes (TA1) e (TA2) é proposta por Christoffersen(1998)[1] afim de avaliar se a série de violações do VaR têm a propriedade de *cobertura condicional correta*, isto é, sob a hipótese nula:

$$I_t | \mathcal{F}_{t-1} \stackrel{iid}{\sim} Ber(\tau), \quad (2.5)$$

a estatística de teste proposta é:

$$RV_{cc} = 2 \left[\mathcal{L}(\hat{\Pi}; I_2, \dots, I_T | I_1) - \mathcal{L}(\tau; I_2, \dots, I_T | I_1) \right] \stackrel{a}{\sim} \chi_2^2$$

que pode ser expressa da seguinte forma:

$$RV_{cc} = RV_{CI} + RV_{ind}.$$

TA4. Teste do Quantil Dinâmico

Uma sequência de previsões é eficiente com relação a $\mathcal{F}_{t-1}$ se a equação (2.4) vale. Mas, isso implicaria que I_t não é correlacionado com qualquer função de variáveis no conjunto de informações até $t - 1$.

Essa restrição é mais forte que a *cobertura condicional correta* (2.5). Em particular, em Engle e Manganelli (2002) [2], ressalta-se que ao analisar I_t é importante condicionar as violações passadas ao VaR no período correspondente.

A operacionalização do teste é realizada através da seguinte regressão:

$$I_t = \tau_0 + \sum_{i=1}^p \beta_i I_{t-i} + \beta_{p+1} \hat{Q}_t(\beta^0) + \sum_{j=1}^T \beta_{p+j+1} f(\omega_{t-1,j}) + u_t$$

em que I_{t-i} representa o fator autorregressivo da série de violações para $i=1,2,\dots,p$; $\hat{Q}_t(\beta^0)$ é o VaR estimado para o período t e $f(\omega_{t-1,j})$ é função dos valores $\omega_{t-1,j} \in \mathcal{F}_{t-1}$ que podem influenciar I_t e

$$u_t = \begin{cases} -\tau & \text{com probabilidade } 1 - \tau \\ 1 - \tau & \text{com probabilidade } \tau. \end{cases}$$

A hipótese nula é que os regressores não têm poder explicativo, já que deveria valer (2.4). Utilizando a forma matricial, a equação de regressão pode ser escrita da seguinte maneira:

$$\mathbf{H} - \tau \boldsymbol{\iota} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

em que $\boldsymbol{\iota}$ é um vetor de 1's. .

O teste a ser realizado é:

$$H_0 : \boldsymbol{\beta} = \mathbf{0}$$

$$H_1 : \boldsymbol{\beta} \neq \mathbf{0}$$

com a estatística de teste DQ:

$$DQ = \frac{\hat{\boldsymbol{\beta}}'_{MQO} \mathbf{X}' \mathbf{X} \hat{\boldsymbol{\beta}}_{MQO}}{\tau(1 - \tau)} \stackrel{a}{\sim} \chi^2_{p+n+2},$$

na qual o estimador de $\boldsymbol{\beta}_{MQO}$ por mínimos quadrados é:

$$\hat{\boldsymbol{\beta}}_{MQO} = (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' (\mathbf{I} - \tau \boldsymbol{\iota}) \stackrel{a}{\sim} \mathcal{N}(0, (\mathbf{X}' \mathbf{X})^{-1} \tau(1 - \tau)).$$

2.1.7 Medidas auxiliares

A escolha de um modelo satisfatório implica em selecionar aquele que, quando há violação ao VaR, não excede em um valor muito grande. Além disso, quando não há violação, as estimativas não poderiam estar longe das observadas. O *valor de violação* representa o quanto o valor observado excedeu a estimativa do VaR. No caso de uma posição comprada, a violação seria para retornos mais negativos que o VaR, para uma posição vendida, seria o oposto. Analogamente, o *valor de não-violação* será o quanto a estimativa do VaR excedeu o observado. Sejam V_t a série de VaR estimada pelo modelo e r_t a série de log-retornos observada para uma posição comprada. A série obtida pelos valores de violação ao VaR corresponde a:

$$G_t = \begin{cases} |r_t - V_t| & , V_t > r_t \\ 0 & , c.c. \end{cases}$$

e a série dos valores de não-violação:

$$H_t = \begin{cases} |V_t - r_t| & , V_t \leq r_t \\ 0 & , c.c. \end{cases}$$

Utilizaremos o cálculo do *valor acumulado de violações*

$$AcV = \sum_{t=1}^T G_t,$$

e o cálculo do *valor acumulado de não-violações*

$$AcNV = \sum_{t=1}^T H_t,$$

em que os modelos mais adequados serão aqueles que apresentem o menor valor acumulado de violações e não-violações.

2.2 Causalidade

O conceito de causalidade tem sido discutido há tempos. Esse tema se mostra relevante no entendimento de relações entre fatores sob estudo e está presente em extensa bibliografia.

Granger (1980) [4] discorre sobre a dificuldade em se chegar a um consenso entre as diferentes definições de causalidade. Por exemplo: existe o conceito de causalidade em situações determinísticas na qual se um evento A ocorre, então B deve ocorrer. Alternativamente, teríamos situações em que se A ocorre, então a probabilidade de ocorrer B aumenta (ou muda).

Neste estudo, será adotado o conceito de causalidade estabelecido por Granger (1969) [3]. Visto que será necessário estudar o relacionamento entre duas ou mais variáveis considerando-se o fator temporal, a causalidade de Granger está relacionada à previsibilidade.

2.2.1 Causalidade de Granger

A definição geral, estabelecida em Granger (1980) [4], diz que para Ω_t representando toda informação disponível até o período t , Y_t causa X_{t+1} se

$$Pr(X_{t+1} \in A | \Omega_t) \neq Pr(X_{t+1} \in A | \Omega_t - Y_t)$$

para algum A . Isto é, há causalidade de Y_t sobre X_{t+1} se a probabilidade é alterada quando se considera todo conjunto de informação disponível ou todo o conjunto retirando-se a informação de Y_t .

O conceito de causalidade Granger, atrelado à interpretação de previsibilidade, está relacionado às definições a seguir:

Seja A_t um processo estocástico estacionário.

- O conjunto de informação de valores passados de A_t será representado por $\bar{A}_t = \{A_{t-j}, j = 1, 2, \dots, \infty\}$.
- O conjunto de informação de valores passados e presente de A_t será representado por $\bar{\bar{A}}_t = \{A_{t-j}, j = 0, 1, \dots, \infty\}$.
- O conjunto de informação de valores passados a partir de uma defasagem k será representado por $\bar{A}(k) = \{A_{t-j}, j = k, k+1, \dots, \infty\}$.
- O preditor de EQM mínimo de A_t usando o conjunto de valores de B_t será representado por $P_t(A|B)$.
- O erro de previsão será representado por $\varepsilon_t(A|B) = A_t - P_t(A|B)$.
- A variância do erro de previsão será representada por $\sigma^2(A|B)$.

Definição 2.2.1 (Causalidade de Granger) *Seja $\bar{A}_t$ o conjunto de toda informação relevante até t (exclusive). Se $\sigma^2(X_t | \bar{A}_t) < \sigma^2(\bar{X}_t | \bar{A}_t - \bar{Y}_t)$ então Y causa X no sentido de Granger, ou Y Granger-causa X . Isto é, Y_t causa X_t no sentido de Granger se X_t é melhor prevista usando toda informação disponível, do que usando somente a informação sem considerar Y_t . Utiliza-se a notação: $Y_t \Rightarrow X_t$*

Definição 2.2.2 (Feedback) Dizemos que há feedback entre X_t e Y_t , se $Y_t \Rightarrow X_t$ e $X_t \Rightarrow Y_t$. A notação para feedback é: $Y_t \Leftrightarrow X_t$

Definição 2.2.3 (Causalidade Instantânea) Se $\sigma^2(X|\overline{A}_t, \overline{\overline{Y}}_t) < \sigma^2(X|\overline{A}_t, \overline{Y})$, então há causalidade instantânea. Ou seja, a previsão de X_t é melhor quando o valor presente de Y_t estiver incluído que a previsão quando não está incluso.

Existem várias técnicas para implementar o teste de causalidade de Granger. Iremos utilizar uma abordagem baseada na metodologia proposta por Hsiao (1979) [6].

Temos o modelo bivariado autorregressivo:

$$Y_t = \Psi_{11}^{M_{11}}(L)Y_t + \Psi_{12}^{M_{12}}(L)X_t + u_t \quad (2.6)$$

$$X_t = \Psi_{21}^{M_{21}}(L)Y_t + \Psi_{22}^{M_{22}}(L)X_t + v_t$$

no qual u_t é $RB(0, \sigma_u)$, v_t é $RB(0, \sigma_v)$, $\Psi_{ij}(L) = \sum_{t=1}^{M_{ij}} \Psi_{ij,t} L^t$ e L é o operador de defasagem, tal que $LY_t = Y_{t-1}$ e M_{ij} é a ordem dos termos autorregressivos.

Para descobrir as ordens dos termos das defasagens para $\Psi_{ij}^{M_{ij}}$, para $i=1,2$ e $j=1,2$, será utilizado o critério FPE (*Final Prediction Error*) de Akaike. Para a primeira equação em (2.6), por exemplo, podemos aplicar o método de MQO (Mínimos Quadrados Ordinários) para estimar os parâmetros para $M_{11} = m$ e $M_{12} = n$:

$$\hat{Y}_t = \hat{\alpha} + \hat{\Psi}_{11}^m(L)Y_t + \hat{\Psi}_{12}^n(L)X_t + \hat{u}_t.$$

A escolha de m e n será tal que o critério FPE seja minimizado:

$$FPE_Y(m, n) = \frac{T + m + n + 1}{T - m - n - 1} \frac{\sum_{t=1}^T (Y_t - \hat{Y}_t)}{T}. \quad (2.7)$$

A partir do modelo em (2.6), foi mostrado que dizer que Y não causa X no sentido de Granger, ou $Y \nRightarrow X$, é equivalente a $\Psi_{21}^{M_{21}}(L) = 0$. Analogamente, X não causa Y no sentido de Granger, ou $X \nRightarrow Y$, se somente se $\Psi_{12}^{M_{12}}(L) = 0$

Para a primeira equação em (2.6), os passos do ajustamento do modelo serão os seguintes:

1. *Escolha da ordem máxima M* : definir M .
2. *Escolha da ordem s* : considerar a primeira equação em (2.6) considerar Y_t como variável dependente. Encontrar s entre 1 e M_{11} que minimiza o critério FPE (ver equação 2.7), considerar $M_{11} = M$ e acrescentar os termos defasados de Y_t até essa ordem, representados por $\Psi_{11}^{M_{11}}(L)$.
3. *Reduzir o modelo*: re-estimar o modelo caso haja parâmetros estimados que não são significativos. Se o critério FPE for melhor, manter o modelo reduzido.
4. *Escolha da ordem n* : a partir da equação estimada no passo 1, adicionar os termos defasados de X_t , representados por $\Psi_{12}^{M_{12}}(L)$ até a ordem $M_{12} = M$, e encontrar a ordem n entre 1 e M_{12} que minimiza o critério FPE.
5. *Escolha da ordem m* : Para checar se algum valor defasado de Y_t possa ter sofrido a influência de alguma defasagem de X_t , com os termos $\Psi_{12}^n(L)$ fixos, e verificar, incluindo serialmente os termos defasados de Y_t , se há redução do FPE.

Para a segunda equação em (2.6), os passos do ajustamento do modelo serão os seguintes:

1. *Escolha da ordem s* : considerar a segunda equação em (2.6) considerar X_t como variável dependente. Encontrar s entre 1 e M_{22} que minimiza o critério FPE (ver equação 2.7), considerar $M_{22} = M$ e acrescentar os termos defasados de X_t até essa ordem, representados por $\Psi_{22}^{M_{22}}(L)$.
2. *Reduzir o modelo*: re-estimar o modelo caso haja parâmetros estimados que não são significativos. Se o critério FPE for melhor, manter o modelo reduzido.
3. *Escolha da ordem n* : a partir da equação estimada no passo 1, adicionar os termos defasados de Y_t , representados por $\Psi_{21}^{M_{21}}(L)$ até a ordem $M_{21} = M$, e encontrar a ordem n entre 1 e M_{21} que minimiza o critério FPE.

4. *Escolha da ordem m*: Para checar se algum valor defasado de X_t possa ter sofrido a influência de alguma defasagem de Y_t , com os termos $\Psi_{21}^n(L)$ fixos, e verificar, incluindo serialmente os termos defasados de X_t , se há redução do FPE.

As correlações cruzadas entre u_t e v_t são dadas por:

$$\rho_{uv}(k) = \frac{E(u_t v_{t+k})}{\sqrt{E(u_t^2)E(v_t^2)}}$$

que são estimadas por:

$$r_{\hat{u}\hat{v}}(k) = \frac{\sum_{t=1}^T \hat{u}_t \hat{v}_{t+k}}{\sqrt{\sum_{t=1}^T \hat{u}_t^2 \sum_{t=1}^T \hat{v}_t^2}}. \quad (2.8)$$

Pierce e Haugh (1977) [15] estabelecem que a restrição $\rho_{uv}(0) \neq 0$ indica *causalidade instantânea*. Assim, devemos realizar o teste utilizando (2.8). Foi provado que, sob a hipótese nula da não-causalidade, as correlações $r_{\hat{u}\hat{v}}(k)$ são assintoticamente independentes e tem distribuição $N(0, 1/T)$.

2.2.2 Causalidade de Granger em Risco

A causalidade de Granger em risco está focada nos movimentos entre as caudas das distribuições. Em particular, quando se trata de posições compradas, a importância está na cauda inferior, que está relacionada a possíveis valores de perda.

A vantagem de se analisar a causalidade de Granger em risco ao invés de se analisar a causalidade de Granger em variância é que a volatilidade é uma medida que trata ganhos e perdas de maneira simétrica. Uma vez que o risco de uma carteira está associada à perda do seu valor, movimentações em favor da cauda inferior devem ser priorizadas.

A causalidade de Granger em risco pode surgir na ausência da causalidade de Granger em média ou variância e também devido a movimentos em momentos de ordem maiores (ex.: curtose e assimetria).

Será checado se a ocorrência de um risco alto em um ativo pode prever a ocorrência

de outro risco alto em outro ativo. Sejam Y_{1t} e Y_{2t} os riscos obtidos no ativo 1 e 2, respectivamente. As informações passadas dessas séries serão representadas por $I_{1(t-1)} = \{Y_{1(t-1)}, Y_{1(t-2)}, \dots, Y_{11}\}$ para o ativo 1 e, equivalentemente, $I_{2(t-1)} = \{Y_{2(t-1)}, Y_{2(t-2)}, \dots, Y_{21}\}$ para o ativo 2. $\mathbf{I}_{t-1}$ é a matriz de informações passadas disponíveis em t-1, na qual $\mathbf{I}_{t-1} = (I_{1(t-1)}, I_{2(t-1)})$.

O VaR depende das informações disponíveis em t-1 e do nível de probabilidade em risco. Logo, será denotado sob a seguinte forma $V_{kt} := V(I_{k(t-1)}, \tau)$, onde $\tau \in (0, 1)$.

Então, as hipóteses sobre a causalidade serão representadas:

$$H_0 : Pr(Y_{1t} < V_{1t} | I_{1(t-1)}) = Pr(Y_{1t} < V_{1t} | \mathbf{I}_{t-1})$$

$$H_1 : Pr(Y_{1t} < V_{1t} | I_{1(t-1)}) \neq Pr(Y_{1t} < V_{1t} | \mathbf{I}_{t-1}),$$

ou seja, se H_0 valer, $\{Y_{2t}\}$ não causa $\{Y_{1t}\}$ no sentido de Granger em risco a um nível α com relação a $\mathbf{I}_{t-1}$. Se $\{Y_{2t}\}$ causa $\{Y_{1t}\}$ no sentido de Granger em risco. Caso contrário, a informação de ocorrência de um risco alto em $\{Y_{2t}\}$ pode ser utilizado para prever um risco futuro em $\{Y_{1t}\}$. Uma extensão do teste acima é representada pelas seguintes hipóteses:

$$H_0 : E(Y_{1t} | Y_{1t} < V_{1t}, I_{1(t-1)}) = E(Y_{1t} | Y_{1t} < V_{1t}, \mathbf{I}_{t-1}) \quad (2.9)$$

$$H_1 : E(Y_{1t} | Y_{1t} < -V_{1t}, I_{1(t-1)}) \neq E(Y_{1t} | Y_{1t} < -V_{1t}, \mathbf{I}_{t-1})$$

Capítulo 3

Metologia

3.1 Hipótese de Causalidade de Granger em Risco

Nesta seção, será exposta a metodologia para testar as hipóteses de causalidade de Granger em risco baseada no estudo de Hong, Liu e Wang (2009) [7].

Nessa metodologia, será verificado se há relação de causalidade a partir de um evento de risco.

Considere a seguinte transformação de variáveis sobre $\{Y_{1t}\}$ e $\{Y_{2t}\}$:

$$Z_{lt} = \mathbb{I}\{Y_{lt} < V_{lt}\}, l = 1, 2$$

onde $\mathbb{I}\{\cdot\}$ é a função indicadora. Assim, Z_{lt} é 1 quando a perda exceder ao VaR e 0 caso contrário. No caso de uma posição comprada, a perda excede ao VaR se o valor é menor que o VaR, para uma posição vendida, ocorre o oposto. Nesse trabalho, será considerada a posição comprada. O VaR, V_{lt} , é obtido a partir de algum modelo do tipo $V_{lt}(\beta_l)$. Com as novas variáveis, é possível reformular a hipótese anterior (2.9) da seguinte maneira:

$$H_0 : E(Z_{1t}|I_{1(t-1)}) = E(Z_{1t}|\mathbf{I}_{t-1}) \quad (3.1)$$

$$H_1 : E(Z_{1t}|I_{1(t-1)}) \neq E(Z_{1t}|\mathbf{I}_{t-1}).$$

Com esta formulação, a causalidade em risco entre $\{Y_{1t}\}$ e $\{Y_{2t}\}$ pode ser vista como causalidade de Granger em média entre $\{Z_{1t}\}$ e $\{Z_{2t}\}$

Em Granger(1969) [3], métodos utilizando o espectro cruzado e o espectro parcial cruzado podem ser utilizados para analisar efeitos de causalidade e *feedback*.

Na investigação de relação de causalidade entre duas variáveis, o espectro cruzado pode ser interpretado paralelamente como a análise do coeficiente de correlação, enquanto o espectro cruzado parcial como o coeficiente de correlação cruzado.

Antes de definirmos o espectro cruzado, precisamos definir os conceitos de *autocorrelação* e *espectro*. Esses conceitos estão expostos detalhadamente em Granger and Hatanaka (1964) [5] e Jenkins and Watts (1969) [8].

Definição 3.1.1 (Autocovariância) A função de autocovariância de um processo $\{Z_t\}$ nos tempos t_1 e t_2 é definida por:

$$\gamma(t_1, t_2) = E(Z_{t_1}Z_{t_2}) - E(Z_{t_1})E(Z_{t_2}).$$

Temos que $\gamma(t_1, t_2)$ fornece a covariância entre Z_{t_1} e Z_{t_2} . A partir dela, podemos estabelecer se a série está correlacionada com seus próprios valores passados ao utilizar $t_1 = t$ e $t_2 = t_1 - j$.

Um propriedade relevante de processos estacionários é que a autocovariância só depende do intervalo entre t_1 e t_2 que é $j = |t_1 - t_2|$. Logo, em geral, representa-se a autocovariância como $\gamma(j)$ para $j = -\infty, \dots, \infty$.

Definição 3.1.2 (Autocorrelação) A função de autocorrelação de um processo estacionário é definida por:

$$\rho(j) = \frac{\gamma(j)}{\gamma(0)}$$

para $j = -\infty, \dots, \infty$.

Veja que se $j = 0$ então $\gamma(0)$ é a variância, pois temos:

$$\gamma(0) = E(Z_t^2) - E(Z_t)^2.$$

Definição 3.1.3 (Autocovariância Cruzada) *Extendendo-se o conceito de autocovariância para avaliação de dependência linear entre dois processos $\{Z_{1t}\}$ e $\{Z_{2t}\}$, temos que a autocovariância cruzada nos tempos t_1 e t_2 é:*

$$\gamma_{12}(t_1, t_2) = E(Z_{1t_1}Z_{2t_2}) - E(Z_{1t_1})E(Z_{2t_2}).$$

Para processos estacionários, a notação será $\gamma_{12}(j)$ para $j = -\infty, \dots, \infty$:

$$\gamma_{12}(j) = E(Z_{1t}Z_{2(t-j)}) - E(Z_{1t})E(Z_{2(t-j)}).$$

Definição 3.1.4 (Autocorrelação Cruzada) *A função de autocorrelação cruzada entre dois processos estacionários $\{Z_{1t}\}$ e $\{Z_{2t}\}$ é definida por:*

$$\rho_{12}(j) = \frac{\gamma_{12}(j)}{\sqrt{\gamma_1(0)\gamma_2(0)}}.$$

para $j = -\infty, \dots, \infty$.

Definição 3.1.5 (Espectro ou Função de Densidade Espectral) *A função de densidade espectral ou espectro de um processo estacionário discreto, é:*

$$f(\lambda) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_j e^{-i\lambda j}, \quad -\pi \leq \lambda \leq \pi.$$

Definição 3.1.6 (Função de Densidade Espectral Cruzada) *A densidade espectral cruzada normalizada entre $\{Z_{1t}\}$ e $\{Z_{2t}\}$ é:*

$$f_{12}(\omega) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \rho_{12}(j) e^{-ij\omega},$$

em que $\omega \in [-\pi, \pi]$, $i = \sqrt{-1}$ e ρ_{12} é a autocorrelação cruzada entre as duas séries.

Temos que $\rho_{12}(j)$ e $f_{12}(\omega)$ são pares de Fourier, logo as duas contêm informação da correlação cruzada entre $\{Z_{1t}\}$ e $\{Z_{2t}\}$ e uma delas poderia ser utilizada para

testar a hipótese de causalidade. Neste estudo, será utilizada $f_{12}(\omega)$ na construção da estatística de teste.

Sob H_0 , temos que $\rho_{12}(j) = 0$ para todo $j > 0$, pois a hipótese é que não há causalidade de Granger entre as séries. Logo, a correlação cruzada é nula para todas as defasagens passadas de $\{Z_2\}$ que influenciariam $\{Z_{1t}\}$. Consequentemente, $f_{12}(\omega)$ se reduz a:

$$f_{12}^0(\omega) = \frac{1}{2\pi} \sum_{-\infty}^0 \rho_{12}(j) e^{-ij\omega}.$$

Assim, pode-se comparar $f_{12}(\omega)$ com $f_{12}^0(\omega)$ para testar H_0 . Se houver grande diferença entre as duas funções, há evidência de que H_0 deve ser rejeitada.

Tanto $f_{12}(\omega)$ quanto $f_{12}^0(\omega)$ precisam ser estimadas. Um método de estimação não paramétrico é o método por utilização de núcleos. A utilização de funções núcleo, para investigação dos efeitos, permite que sejam escolhidos pesos que sejam menores para eventos distantes. Essa possibilidade é consistente com o fato de que o mercado é mais afetado por informações recentes.

Alguns núcleos mais utilizados, cuja definição estão em Morettin e Tolo (2006) [14] que estão representados nas Figuras 3.1, 3.2, 3.3, 3.4, e 3.8 com a variação do parâmetro, são os seguintes:

- Núcleo truncado

$$k_H(z) = \begin{cases} 1 & , |z| \leq H \\ 0 & , |z| > H. \end{cases}$$

- Núcleo de Bartlett

$$k_H(z) = \begin{cases} 1 - \frac{|z|}{H} & , |z| \leq H \\ 0 & , |z| > H. \end{cases}$$

- Núcleo de Daniell / Núcleo retangular

$$k_H(z) = \frac{\sin\left(\frac{\pi z}{H}\right)}{\frac{\pi z}{H}}, \text{ para todo } z.$$

Figura 3.1: Núcleo Truncado (H=1).

Figura 3.2: Núcleo de Bartlett (H=1).

Figura 3.3: Núcleo de Daniell (H=1).

- Núcleo de Parzen

$$k_H(z) = \begin{cases} 1 - 6\left(\frac{z}{H}\right)^2 + 6\left(\frac{|z|}{H}\right)^3, & |z| \leq H/2 \\ 2\left(1 - \frac{|z|}{H}\right)^3, & H/2 \leq |z| \leq H \\ 0, & |z| > H. \end{cases}$$

Figura 3.4: Núcleo de Parzen (H=1).

Se o VaR for calculado a partir de um modelo paramétrico então V_{lt} é função de β , que corresponde ao vetor de parâmetros do modelo. Ou seja:

$$Z_{lt}(\beta) = \mathbb{I}\{Y_{lt} < V_{lt}(\beta)\}, l = 1, 2.$$

Então, temos que $\hat{Z}_{lt} = \mathbb{I}\{Y_{lt} < V_{lt}(\hat{\beta})\}, l = 1, 2$. Para efeito de simplicidade, iremos considerar $Z_{lt} = Z_{lt}(\beta)$.

A função de correlação cruzada amostral entre $\{Z_{1t}\}$ e $\{Z_{2t}\}$ na j -ésima defasagem é:

$$\hat{\rho}_{12}(j) = \frac{\hat{C}_{12}(j)}{\hat{S}_1 \hat{S}_2}$$

em que $\hat{C}_{12}(j)$ é a função de covariância cruzada amostral entre $\{Z_{1t}\}$ e $\{Z_{2t}\}$ e $\hat{S}_1, \hat{S}_2$ são os desvios amostrais de $\{Z_{1t}\}$ e $\{Z_{2t}\}$. São especificadas como segue:

Figura 3.5: Núcleo Truncado ($H=1$; $H=1,25$; $H=1,5$; $H=1,75$; $H=2$).

Figura 3.6: Núcleo de Bartlett ($H=1$; $H=1,25$; $H=1,5$; $H=1,75$; $H=2$).

Figura 3.7: Núcleo de Daniell ($H=1$; $H=1,25$; $H=1,5$; $H=1,75$; $H=2$).

Figura 3.8: Núcleo de Parzen ($H=1$; $H=1,25$; $H=1,5$; $H=1,75$; $H=2$).

$$\hat{C}_{12}(j) = \begin{cases} \frac{\sum_{i=1+j}^T (\hat{Z}_{1i} - \hat{\alpha}_1)(\hat{Z}_{2(t-j)} - \hat{\alpha}_2)}{T} & \text{se } 0 \leq j \leq T-1 \\ \frac{\sum_{i=1-j}^T (\hat{Z}_{1(t+j)} - \hat{\alpha}_1)(\hat{Z}_{2i} - \hat{\alpha}_2)}{T} & \text{se } 1-T \leq j < 0 \end{cases}$$

para $\hat{\alpha}_l = \frac{\sum_{i=1}^T \hat{Z}_{li}}{T}$ são as médias amostrais e $\hat{S}_l^2 = \hat{\alpha}_l(1 - \hat{\alpha}_l)$ são as variâncias amostrais, para $l = 1, 2$

Os estimadores para as densidades espectrais por kernel cruzadas $f_{12}(\omega)$ e $f_{12}^0(\omega)$ são os seguintes:

$$\hat{f}_{12}(\omega) = \frac{1}{2\pi} \sum_{j=1-T}^{T-1} k\left(\frac{j}{M}\right) \hat{\rho}_{12}(j) e^{-ij\omega}, \quad (3.2)$$

$$\hat{f}_{12}^0(\omega) = \frac{1}{2\pi} \sum_{j=1-T}^0 k\left(\frac{j}{M}\right) \hat{\rho}_{12}(j) e^{-ij\omega}.$$

A comparação entre $\hat{f}(\omega)$ e $\hat{f}_0(\omega)$ é feita utilizando a forma quadrática:

$$L^2(\hat{f}_{12}(\omega), \hat{f}_{12}^0(\omega)) = 2\pi \int_{-\pi}^{\pi} \left| \hat{f}_{12}(\omega) - \hat{f}_{12}^0(\omega) \right|^2 d\omega = \sum_{j=1}^{T-1} k^2\left(\frac{j}{M}\right) \hat{\rho}_{12}^2(j),$$

em que a última passagem é obtida a partir da identidade de Parseval. Assim, não é necessário calcular a integral sobre o domínio da frequência de ω .

A estatística para testar a hipótese (3.1) é uma versão padronizada da forma quadrática:

$$Q_1(M) = \frac{\left[T \sum_{j=1}^{T-1} k^2 \left(\frac{j}{M} \right) \hat{\rho}_{12}^2(j) - C_{1T}(M) \right]}{\sqrt{D_{1T}(M)}},$$

cujas constantes de centralização e padronização são:

$$C_{1T} = \sum_{j=1}^{T-1} \left(1 - \frac{j}{T} \right) k^2 \left(\frac{j}{M} \right)$$

$$D_{1T} = 2 \sum_{j=1}^{T-1} \left(1 - \frac{j}{T} \right) \left(1 - \frac{j+1}{T} \right) k^4 \left(\frac{j}{M} \right).$$

Teorema 3.1.1 (Teo) *Suponha que sejam válidas as suposições de 1 a 6, listadas no Apêndice A, e $M = cT^v$, onde $0 < c < \infty$, $0 < v < 1/2$, $v < \min \left(\frac{2}{b-2}, \frac{3}{b-1} \right)$, se $b = \max(b_1, b_2) > 2$, b_1 é a dimensão de β_1 e b_2 é a dimensão de β_2 . Então $Q_1(M) \xrightarrow{d} N(0, 1)$.*

3.2 Causalidade de Granger em Risco com Feedback

Hong, Liu e Wang(2009) [7] definem uma modalidade para verificação de causalidade de Granger em risco com *feedback*, isto é, a hipótese a ser considerada é a que nem Y_{1t} nem Y_{2t} causa no sentido de Granger uma com relação à outra.

$$H_0 : P(Y_{lt} < V_{lt} | I_{l(t-1)}) = P(Y_{lt} < V_{lt} | \mathbf{I}_{t-1}) \quad \text{para } l = 1, 2 \quad (3.3)$$

$$H_1 : P(Y_{lt} < V_{lt} | I_{l(t-1)}) \neq P(Y_{lt} < V_{lt} | \mathbf{I}_{t-1}) \text{ para pelo menos um } l.$$

Sendo assim, se o risco para pelo menos uma das séries for alterado quando con-

sideramos a informação das duas conjuntamente comparado quando só é utilizada a informação de uma das séries, então rejeitamos a hipótese que não há causalidade de Granger em risco bilateral.

Alternativamente, podemos utilizar as variáveis indicadoras formuladas no teste de hipóteses:

$$H_0 : E(Z_{lt}|I_{l(t-1)}) = E(Z_{lt}|\mathbf{I}_{t-1}) \quad \text{para } l = 1, 2$$

$$H_1 : E(Z_{lt}|I_{l(t-1)}) \neq E(Z_{lt}|\mathbf{I}_{t-1}) \quad \text{para pelo menos um } l.$$

Sob H_0 , a informação passada de uma das séries não é útil para prever o risco na outra. Assim, a densidade do espectro cruzado terá a seguinte forma:

$$f_{12}^0(\omega) = \frac{1}{2\pi} \rho_{12}(0), \omega \in [-\pi, \pi].$$

O estimador consistente correspondente é:

$$\hat{f}_{12}^0(\omega) = \frac{1}{2\pi} \hat{\rho}_{12}(0), \omega \in [-\pi, \pi]. \quad (3.4)$$

A estatística para testar a hipótese (3.3) é uma versão da forma quadrática padronizada entre (3.2) e (3.4):

$$Q_2(M) = \frac{\left[T \sum_{|j|=1}^{T-1} k^2 \left(\frac{j}{M} \right) \hat{\rho}_{12}^2(j) - C_{2T}(M) \right]}{\sqrt{D_{2T}(M)}},$$

cujas constantes de centralização e padronização são:

$$C_{2T} = \sum_{|j|=1}^{T-1} \left(1 - \frac{|j|}{T} \right) k^2 \left(\frac{j}{M} \right),$$

$$D_{2T} = 2[1 + \hat{\rho}^4(0)] \sum_{|j|=1}^{T-1} \left(1 - \frac{|j|}{T}\right) \left(1 - \frac{|j|+1}{T}\right) k^4 \left(\frac{j}{M}\right). \quad (3.5)$$

A constante de padronização D_{2T} em (3.5) apresenta o termo $\hat{\rho}(0)$, estimador de correlação cruzada, que mede possível correlação instantânea entre Z_{1t} e Z_{2t} sob H_0 em (3.3).

A estatística $Q_2(M)$ é assintoticamente Normal(0,1) sob H_0 , segundo teorema enunciado a seguir.

Teorema 3.2.1 (Teo) *Suponha que sejam válidas as suposições de 1 a 3(a), 4 a 6, listadas no Apêndice A, e $M = cT^v$, onde $0 < v < 1/2$, $v < \max\left(\frac{2}{b-2}, \frac{3}{b-1}\right)$, se $b = \max(b_1, b_2) > 2$, b_1 é a dimensão de β_1 e b_2 é a dimensão de β_2 . Então $Q_2(M) \xrightarrow{d} N(0, 1)$ sob H_0 .*

Capítulo 4

Aplicação

Nesta seção, será apresentada a aplicação da metodologia utilizando dados reais. Em primeiro lugar, serão descritos os dados utilizados para estudo; em segundo, será aplicada o modelo CaViaR de cálculo do VaR para realização da transformação das séries e por fim, a causalidade de Granger em risco será verificada com as devidas conclusões.

Para desenvolvimento da análise, foram utilizados os softwares Matlab e R. No Matlab, foi utilizada a rotina desenvolvida por Manganelli [2], com algumas adaptações. Para as demais análises, foi utilizado o software R. Os códigos estão disponíveis para utilização em: <http://patricianm.yolasite.com/>.

4.1 Descrição dos dados utilizados

Iremos analisar a relação de causalidade em risco entre duas séries utilizando dados reais para aplicação da metodologia exposta. Para isso, foram escolhidas as séries de preços das ações de empresas que são referência de mercado e que pertencem a segmentos distintos de mercado.

Como representante do setor financeiro, no segmento de bancos, foi selecionada a série de preços da ação do Banco Bradesco (BBND4). No setor petrolífero, foi escolhida a empresa de exploração e refino, Petrobrás (PETR4). As séries correspondem ao período de 01-jan-2000 a 16-jul-2010, representando 2732 observações. Os preços utilizados correspondem ao preço de fechamento, ajustados para eventos corporativos. A evolução dos preços das duas ações são ilustradas na Figura (4.1):

Figura 4.1: Série de preços das ações BBDC4 e PETR4.

De maneira geral, é possível notar um período de valorização dos preços das ações até meados do segundo bimestre de 2008, ano em que o mercado financeiro sofreu os impactos da crise internacional econômica, também conhecida como "crise dos subprimes".

Os dados obtidos de preços de fechamento das ações BBDC4 e PETR4 precisaram sofrer ajustes devido a alguns eventos corporativos de desdobramento e grupamento de ações que ocorreram no período de análise. Esses eventos alteram o patamar de preços das ações.

As ações são negociadas em lotes. O desdobramento, ou *split*, causa o aumento de ações negociadas no mercado. Ele ocorre geralmente quando o preço da ação costuma subir muito. Assim, a liquidez tende a diminuir já que preços altos atraem menos investidores. Ao realizar um desdobramento de 1 para 2 (1:2), a quantidade de ações irá dobrar e o preço cairá pela metade. Com essa operação, o valor da empresa não é alterado.

O grupamento ocorre de maneira inversa, isto é, uma determinada quantidade de ações é agrupada, com respectivo aumento proporcional do preço. Em geral, isso ocorre quando os preços das ações estão sendo cotados a valores baixos.

Outro tratamento foi necessário para ajustar as séries devido à mudança no fator

de cotação. Em algumas ações, o valor obtido em cotações são, na verdade, o resultado da multiplicação do preço real da ação pelo fator de cotação da mesma.

As datas em que foram necessários os ajustes para os eventos citados constam nas Tabelas 4.1 e 4.2.

Data	Evento	Proporção
22/03/2004	fator de cotação	1:1000
13/12/2004	desdobramento	1:2
07/04/2008	desdobramento	2:3
09/06/2009	grupamento	1:50

Tabela 4.1: Datas e eventos para tratamento de BBDC4

Data	Evento	Proporção
22/06/2000	desdobramento	1:10
01/09/2005	desdobramento	1:4
28/04/2008	desdobramento	1:2

Tabela 4.2: Datas e eventos para tratamento de PETR4

Para análise, serão utilizados os log-retornos dos preços das ações. Como definidos na seção de Conceitos, utilizaremos também a seguinte notação, quando conveniente: r_{1t} equivale ao log-retorno da ação BBDC4 e r_{2t} equivale ao log-retorno da ação PETR4, ambas com base no tempo t para $t = 1, 2, \dots, 2731$. Em Tsay (2005) [16] e Zivot (2006) [17] são encontrados alguns exemplos de análises descritivas e tratamentos a serem realizados com séries financeiras.

Através das estatísticas descritivas amostrais da Tabela 4.3, temos que as séries de log-retornos de BBDC4 e PETR4 apresentam mediana estimada igual à zero. Além disso, as médias amostrais das séries de log-retornos das ações BBDC4 e PETR4 são: $\bar{r}_{1t} = 0.001$ e $\bar{r}_{2t} = 0.001$, sendo que o desvio padrão estimado de PETR4 é menor que o desvio padrão estimado de BBDC4.

As estimativas de excesso de curtose 9,590 para BBDC4 e 3,886 para PETR4 indicam que as duas séries tem caudas pesadas. Sob hipótese de normalidade, o excesso

Figura 4.2: Série de log-retornos das ações BBDC4 e PETR4

Tabela 4.3: Estatísticas descritivas amostrais das séries de log-retornos das ações BBDC4 e PETR4

Série	BBDC4	PETR4
Mínimo	-0,187	-0,148
1º Quartil	-0,013	-0,012
Média	0,001	0,001
Mediana	0,000	0,000
Desvio Padrão	0,025	0,023
3º Quartil	0,014	0,013
Máximo	0,277	0,132
Assimetria	0,542	-0,097
Curtose	9,590	3,886

de curtose de dados com distribuição normal seria igual a 0. Assim, se fizemos um teste para normalidade, veremos que as curtoses apresentam medidas muito superiores à este valor. Esse fato também pode ser notado nos histogramas da Figura 4.3. A linha contínua representa a densidade teórica para uma distribuição Normal com a mesma média e variância dos dados. Vemos que uma boa parte dos valores que se afastam da média tem densidade superior à normal.

Pelo QQ-plot na Figura 4.4, vemos grande quantidade de observações que estão

Figura 4.3: Histograma dos log-retornos para BBDC4 e PETR4.

Figura 4.4: Q-Q plot normal dos log-retornos para BBDC4 e PETR4.

longe da reta diagonal, o que nos indica que não há indícios de normalidade.

O coeficiente de assimetria estimado para a série BBDC4 foi de 0,542 o que indica que há uma possível assimetria à direita. Para a série PETR4 o coeficiente de assimetria estimado foi de -0,097, logo há indício de assimetria à esquerda.

4.2 Cálculo do VaR - CaViaR

A partir das séries de log-retornos das duas ações, foram selecionadas as primeiras 2460 observações para realizar a estimação do modelo e 271 observações para testar a adequação. Os modelos testados serão: valor absoluto simétrico, inclinação

assimétrica, GARCH(1,1) indireto e adaptativo, segundo especificações expostas na seção 1.

O valor inicial de $Q_1(\beta)$ necessário nos modelos foi obtido a partir dos quantis empíricos das primeiras 300 observações de cada série.

A obtenção das estimativas dos parâmetros é realizada a partir de uma rotina de otimização, como descrita por Engle e Manganelli (2004) [2]. Para cada modelo, essa rotina que foi desenvolvida em Matlab consiste basicamente nos seguintes passos:

1. Geração de n vetores com valores entre 0 e 1 com distribuição uniforme como valores iniciais para β ;
2. Calcular o valor do critério quantílico para os valores iniciais, segundo a Equação 2.3;
3. Escolher os m vetores que produziram o menor valor do critério;
4. Utilizar os valores iniciais obtidos em 3 para executar o algoritmo simplex e o algoritmo quasi-Newton para obter os parâmetros otimizados;
5. Repetir a otimização até satisfazer o critério de convergência.

As estimativas para os modelos do tipo CaViaR foram obtidas para as séries de BBDC4 e PETR4.

4.3 Ajuste e Adequação para série de log-retornos da BBDC4

O primeiro modelo ajustado foi o **valor absoluto simétrico**:

$$\hat{Q}_{1,t}(\hat{\beta}) = 0,0692 + 0,8899\hat{Q}_{1,t-1}(\hat{\beta}) + 0,1397|r_{1,t-1}|$$

na qual temos que $\hat{\beta} = (0,0692; 0,8899; 0,1397)$ com p-valores = $(0,03; 0; 0)$. As estimativas dos parâmetros do modelo são significantes ao nível de 5%.

Na Figura 4.5 constam a série do VaR estimado pelo modelo e a série de log-retornos.

4.3. AJUSTE E ADEQUAÇÃO PARA SÉRIE DE LOG-RETORNOS DA BBDC441

A proporção de violações na amostra foi de 5,04% que indica um bom ajuste pois está em torno de 5%. Já a proporção de violações pós-amostrais foi menor, com um valor de 3,32%.

A estatística de teste de cobertura incondicional para as violações pós-amostrais $RC_{CI} = 1,81$ apresenta p-valor maior que 5% de significância. Logo, não rejeitamos a hipótese de que a média incondicional de violações corresponde à θ .

O teste de independência para as violações pós-amostrais $RC_{Ind} = 0,62$ é não significativo ao nível de 5%. Assim, não rejeitamos a hipótese de independência serial entre as violações.

Figura 4.5: VaR estimado por valor absoluto simétrico e série de log-retornos de BBDC4

O teste de cobertura condicional, que é uma mistura dos dois testes anteriores, isto é, checa a média e independência, é favorável à adequação do modelo, já que não rejeitamos a hipótese de que a proporção de violações é igual à proporção esperada. A estatística de teste obtida foi $RC_{CC} = 3,63$, com p-valor maior que 5%.

Para o teste de quantil dinâmico obtivemos as estatísticas $QD_a = 0,011$ e $QD_{pa} = 0,747$, com p-valores respectivamente iguais a 0,99 e 0,83. Assim, não rejeitamos a hipótese de que não haja autocorrelação entre as violações e que a proporção de violações seja a esperada, tanto para a amostra de ajuste, quanto para a de validação.

A soma dos valores de violação obtida foi de 85,41%, sendo que 89,51% é devido às violações amostrais. Esses valores serão utilizado para comparação entre os demais modelos.

O segundo modelo foi o de **inclinação assimétrica**:

$$\hat{Q}_{1,t}(\hat{\beta}) = 0,0565 + 0,9048\hat{Q}_{1,t-1}(\hat{\beta}) + 0,0658(r_{1,t-1})^+ - 0,1915(r_{1,t-1})^-$$

que apresentou todos coeficientes significativos. As estimativas obtidas para os parâmetros foram $\hat{\beta} = (0,0565; 0,9048; 0,0658; -0,1915)$ com p-valores, para testar se são nulos, iguais a $(0; 0; 0,01; 0)$, respectivamente.

Na Figura 4.6 constam a série do VaR estimado pelo modelo e a série de log-retornos.

Figura 4.6: VaR estimado por inclinação assimétrica e série de log-retornos de BBDC4

A proporção de violações amostrais e pós-amostrais foi equivalente ao modelo anterior, 5,04% e 3,32%, respectivamente.

Os testes de adequação de cobertura incondicional, independência e cobertura condicional apresentam os mesmos valores de estatísticas de teste com relação ao modelo anterior: $RC_{CI} = 1,81$, $RC_{Ind} = 0,62$ e $RC_{CC} = 3,63$, respectivamente. Isso porque a proporção de violações é a mesma e as violações que são divergentes quanto ao primeiro modelo não impactam no teste de independência. Todas elas apresentam p-valor maior que 5%.

Os p-valores para as estatísticas $QD_a = 0,361$ e $QD_{pa} = 0,876$ mostram que o ajuste é adequado, pois são bem maiores que 5%.

A soma dos valores de violação foi de 80,90, na qual 88,66% foi na fase amostral.

4.3. AJUSTE E ADEQUAÇÃO PARA SÉRIE DE LOG-RETORNOS DA BBDC443

Isso mostra que o ajuste do modelo foi melhor que o anterior, já que a ordem de grandeza dos dias em que houve violação foi menor.

O próximo modelo estimado foi o **GARCH(1,1) indireto**:

$$\hat{Q}_{1,t}(\hat{\beta}) = \left(0,1129 + 0,8702\hat{Q}_{1,t-1}^2(\hat{\beta}) + 0,2157r_{1,t-1}^2 \right)^{\frac{1}{2}}.$$

Os três parâmetros estimados do modelo podem ser considerados não nulos $\hat{\beta} = (0,1129; 0,8702; 0,2157)$ com p-valores iguais a $(0,003; 0; 0,01)$.

A proporção de violações ao VaR amostrais e pós-amostrais foi de 5% e 4%, respectivamente. Entre os dois modelos anteriores, é o que apresentou maior proporção pós-amostral.

Figura 4.7: VaR estimado por GARCH(1,1) indireto e série de log-retornos de BBDC4

Na Figura 4.7 constam a série do VaR estimado pelo modelo e a série de log-retornos.

Os testes de adequação de cobertura incondicional, independência e cobertura condicional mostram que o modelo é adequado. As estatísticas $RC_{CI} = 0,54$, $RC_{Ind} = 0,93$ e $RC_{CC} = 1,08$. Os respectivos p-valores $(0,463; 0,334; 0,584)$ não rejeitam as hipóteses de cobertura incondicional, independência e cobertura condicional.

Para o teste de quantil dinâmico, foram obtidos $QD_a = 0,0424$ e $QD_{pa} = 0,768$, ambos significativamente diferente de zero. Logo, enfatiza a adequação do modelo.

A soma dos valores de violação foi de 81,78, com 88,36% devido à parte amostral. Relativamente aos outros modelos, é melhor apenas que o primeiro.

Para o modelo **adaptativo**:

$$\hat{Q}_{1,t}(\hat{\beta}) = \hat{Q}_{1,t-1}(\hat{\beta}) + 0,3919 \left\{ \frac{1}{1 + \exp(10[r_{1,t-1} - \hat{Q}_{1,t-1}(\hat{\beta})])} \right\}$$

na qual o parâmetro estimado $\hat{\beta} = 0,3919$ foi significativo (o p-valor obtido foi aproximadamente igual a zero).

Figura 4.8: VaR estimado por modelo adaptativo e série de log-retornos de BBDC4

Na Figura 4.8 constam a série do VaR estimado pelo modelo e a série de log-retornos.

A proporção de violações amostrais foi de 4,43% e pós-amostrais foi de 5,53%, ambas próximas de 5%. Além disso, a soma dos valores de violação foi de 78,83 no total das observações, sendo que 86,4% é devida à parte amostral. Com relação aos modelos anteriores, é o menor valor obtido.

Não há indício de dependência serial entre as violações, pois $RC_{Ind} = 1,37$ tem p-valor 0,241. O teste de cobertura incondicional apresenta estatística de teste $RC_{CI} = 0,16$ com p-valor 0,691 que o indica que a proporção de violações não pode ser considerada diferente de 5%. A estatística de teste para cobertura condicional $RC_{CC} = 0,32$, com p-valor de 0,854, indica que o modelo está adequado.

4.4. AJUSTE E ADEQUAÇÃO PARA SÉRIE DE LOG-RETORNOS DA PETR445

No teste de quantil dinâmico, foram obtidos $QD_a = 0,2211$ e $QD_{pa} = 0,2167$, ambos significantemente diferente de zero.

A soma dos valores de violação foi de 78,83, com 86,44% devido à parte amostral. Relativamente aos outros modelos, é melhor que todos os outros 3 modelos, pois tem o menor valor de violações acumuladas.

Analisando os resultados dos quatro modelos vistos, iremos considerar como o modelo de **inclinação assimétrica** como o mais adequado para obtenção do VaR a partir a série de log-retornos de BBDC4. A Tabela 4.4 apresenta as estimativas dos modelos considerados. Na Tabela 4.5, as estimativas dos testes de adequação dos modelos.

Assim, o modelo escolhido é:

$$\hat{Q}_{1,t}(\hat{\beta}) = 0,0565 + 0,9048\hat{Q}_{1,t-1}(\hat{\beta}) + 0,0658(r_{1,t-1})^+ - 0,1915(r_{1,t-1})^-$$

A medida de critério quantílico foi a menor, entre todos os modelos. Além disso, a proporção de violações amostrais ao VaR esteve próxima de 5% na fração amostral, embora a proporção pós-amostral esteja próxima de 3,32%. O valor acumulado das violações foi o segundo menor (o menor obtido foi o modelo adaptativo, mas que apresenta alto valor do critério quantílico). Além disso, o valor acumulado dos desvios, quando não ha violação do VaR, foi o segundo menor obtido. Os testes de cobertura incondicional, independência e cobertura condicional mostram que o modelo é adequado.

4.4 Ajuste e Adequação para série de log-retornos da PETR4

Analogamente à análise para a série BBDC4, os mesmos modelos foram gerados para a série de log-retornos de PETR4. Para o primeiro modelo, **valor absoluto simétrico**:

$$\hat{Q}_{2,t}(\hat{\beta}) = 0,0228 + 0,9043\hat{Q}_{2,t-1}(\hat{\beta}) + 0,1745|r_{2,t-1}|$$

, foram obtidas as seguintes estimativas dos parâmetros $\hat{\beta} = (0,0228; 0,9043; 0,1745)$, com p-valores de $(0,033; 0; 0,064)$. O último parâmetro não é significativo a um nível de 5% de significância pois seu p-valor foi de 0,064. Isso indica que o fator de retorno passado não tem influência para o quantil no tempo atual.

Figura 4.9: VaR estimado por valor absoluto simétrico e série de log-retornos de BBDC4

Na Figura 4.9 constam a série do VaR estimado pelo modelo e a série de log-retornos.

A proporção de violações amostrais foi de 5% e de pós-amostrais de 5,9%. A proporção amostral é mais próxima de 5%, que mostra melhor adequação, que o desempenho pós-amostal. O valor do critério quantílico foi de $CQ = 271,86$. A soma dos valores de violação totalizou 80,71, sendo que 93% foi devido à parte amostral.

O teste de quantil dinâmico indica adequação do modelo, pois ambas estatísticas amostral $QD_a = 0,011$ e e pós-amostal $QD_{pa} = 0,747$ apresentam p-valores altos.

Os testes de cobertura incondicional, independência e cobertura condicional apresentam estatísticas de teste e respectivos p-valores iguais a: $RC_{CI} = 0,44(p - valor : 0,506)$, $RC_{Ind} = 0(p - valor : 0,955)$ e $RC_{CC} = 0,88(p - valor : 0,643)$. Eles indicam a adequação de ajuste da série. Todas as três estatísticas de teste apresentam p-valores maior que um nível de significância de 5%.

Para o modelo de **inclinação assimétrica**:

$$\hat{Q}_{2,t}(\hat{\beta}) = 0,0481 + 0,8803\hat{Q}_{2,t-1}(\hat{\beta}) + 0,1161(r_{2,t-1}) - 0,2571(r_{2,t-1})^-$$

4.4. AJUSTE E ADEQUAÇÃO PARA SÉRIE DE LOG-RETORNOS DA PETR447

que apresentou todos coeficientes estimados de $\hat{\beta} = (0,0481; 0,8803; 0,1161; -0,2571)$ significativos, os p-valores obtidos foram respectivamente $(0,001; 0; 0,005; 0)$.

Figura 4.10: VaR estimado por inclinação assimétrica e série de log-retornos de BBDC4

Na Figura 4.10 constam a série do VaR estimado pelo modelo e a série de log-retornos.

A proporção de violações amostrais está em torno de 5%, mas a pós amostral foi bem menor, com um valor de 3,3%, o que pode indicar uma superestimação para as previsões. O valor do critério quantílico $CQ = 270,6$ foi o menor até o momento. O valor acumulado das violações de 84,58 tem 94,45% devido à parte amostral.

As estatísticas de teste do quantil dinâmico, $QD_a = 0,457$ e $QD_{pa} = 0,061$ com p-valores altos, indicam que o modelo é adequado.

Os demais testes de cobertura condicional, incondicional e independência enfatizam a adequação. Os valores obtidos para as estatísticas de teste, com respectivos p-valores foram: $RC_{CI} = 0,86(p - valor : 0,354)$, $RC_{Ind} = 0,01(p - valor : 0,942)$ e $RC_{CC} = 1,72(p - valor : 0,424)$. Assim, não há indícios de dependência serial e a proporção condicional e incondicional não estão distantes da desejada de 5%.

O modelo **GARCH(1,1) indireto**

$$\hat{Q}_{2,t}(\hat{\beta}) = \left(0,0602 + 0,8860\hat{Q}_{2,t-1}^2(\hat{\beta}) + 0,2345r_{2,t-1}^2 \right)^{\frac{1}{2}}$$

apresentou coeficientes estimados $\hat{\beta} = (0,0602; 0,8860; 0,2345)$, nos quais o último

coeficiente não é significativo. Os p-valores obtidos para as estimativas foram de $(0,03; 0; 0,23)$

Figura 4.11: VaR estimado por GARCH(1,1) indireto e série de log-retornos de BBDC4

Na Figura 4.11 constam a série do VaR estimado pelo modelo e a série de log-retornos.

A proporção de violações ao VaR foi de 5% amostrais e 5,9% pós amostrais, ambas podem ser consideradas próximas da proporção esperada. O critério quantílico $CQ = 273,8$ para este modelo foi maior que os dois anteriores. O valor acumulado de violações ao VaR foi de 83,46, para o qual 93,60% é devido à parte amostral.

Todas as estatísticas de teste de adequação para cobertura condicional, incondicional e dependência mostram que o ajuste foi adequado. As estatísticas obtidas foram $RC_{CI} = 0,44(p - valor : 0,506)$, $RC_{Ind} = 0(p - valor : 0,955)$ e $RC_{CC} = 0,88(p - valor : 0,643)$.

O modelo **adaptativo**

$$\hat{Q}_{2,t}(\hat{\beta}) = \hat{Q}_{2,t-1}(\hat{\beta}) + 0,3836 \left\{ \frac{1}{1 + \exp(10[r_{2,t-1} - \hat{Q}_{2,t-1}(\hat{\beta})])} \right\}$$

apresentou parâmetro estimado significativo $\hat{\beta} = 0,3919$, com p-valor aproximadamente zero.

Na Figura 4.12 constam a série do VaR estimado pelo modelo e a série de log-

4.4. AJUSTE E ADEQUAÇÃO PARA SÉRIE DE LOG-RETORNOS DA PETR449

Figura 4.12: VaR estimado por inclinação assimétrica e série de log-retornos de BBDC4

retornos.

A proporção de violações amostrais e pós-amostrais foram as menores obtidas, 4,8% e 3,69% respectivamente, comparando-se com os modelos anteriores. A proporção de violações amostrais foi de 4,76% e pós-amostal foi de 3,69. Por outro lado o critério quantílico $CQ = 281,73$ foi o maior entre os quatro.

Os testes de adequação indicam que o ajuste do modelo foi adequado. As estatísticas de teste foram as seguintes: para o teste de cobertura condicional $RC_{CI} = 1,07(p - valor : 0,3)$, $RC_{Ind} = 0,77(p - valor : 0,38)$ e $RC_{Ind} = 2,15(p - valor : 0,342)$.

Tabela 4.4: Estimativas dos Parâmetros Estimados para os Modelos CAViaR para a série BBDC4.

Estimativas e Testes					
Modelo	Valor Absoluto Simétrico	Inclinação Assimétrica	GARCH(1,1)Indireto	Adaptativo	
$\hat{\beta}_1$	0,0692	0,0565	0,1130	0,3919	
Erro Padrão	0,0386	0,0104	0,0421	0,0427	
<i>P-valor</i>	0,0367	0	0,0036	0	
$\hat{\beta}_2$	0,8899	0,9048	0,8702	-	
Erro Padrão	0,0328	0,0120	0,0178	-	
<i>P-valor</i>	0	0	0	-	
$\hat{\beta}_3$	0,1397	0,0658	0,2157	-	
Erro Padrão	0,0266	0,0285	0,0983	-	
<i>P-valor</i>	0,0001	0,0103	0,0141	-	
$\hat{\beta}_4$	-	-0,1915	-	-	
Erro Padrão	-	0,0225	-	-	
<i>P-valor</i>	-	0	-	-	

4.4. AJUSTE E ADEQUAÇÃO PARA SÉRIE DE LOG-RETORNOS DA PETR451

Tabela 4.5: continuação - Estimativas de Estatísticas de Teste dos Parâmetros Estimados para os Modelos CAViaR para a série BBDC4.

Estimativas e Testes					
Modelo	Valor Absoluto Simétrico	Inclinação Assimétrica	GARCH(1,1)Indireto	Adaptativo	
Hit_{am}	5,04%	5,04%	5,00%	4,43%	
Hit_{pam}	3,32%	3,32%	4,06%	5,54%	
QD_{am}	0,0107	0,3608	0,0424	0,2211	
QD_{am} : p-valor	0,9997	0,9856	0,9977	0,6382	
QD_{pam}	0,7469	0,8759	0,7685	0,2167	
QD_{pam} : p-valor	0,8621	0,9280	0,8570	0,6416	
CQ	279,12	275,90	278,00	282,20	
Teste CI: LR	1,8150	1,8150	0,5384	0,1581	
Teste CI: p-valor	0,1779	0,1779	0,4631	0,6909	
Teste IND: LR	0,6208	0,6208	0,9346	1,3718	
Teste IND: p-valor	0,4307	0,4307	0,3337	0,2415	
Teste CC: LR	3,6301	3,6301	1,0767	0,3162	
Teste CC: p-valor	0,1628	0,1628	0,5837	0,8538	
Valor Acum. Viol.	85,41	80,90	81,77	78,83	
Valor Acum. Não Viol.	2.587,20	2.597,66	2.638,05	2.762,77	

Tabela 4.6: Estimativas dos Parâmetros Estimados para os Modelos CAViaR para a série PETR4.

Estimativas e Testes					
Modelo	Valor Absoluto Simétrico	Inclinação Assimétrica	GARCH(1,1)Indireto	Adaptativo	
$\hat{\beta}_1$	0,0228	0,0480	0,0602	0,3836	
Erro Padrão	0,0124	0,0157	0,0324	0,0496	
<i>P-valor</i>	0,0331	0,0011	0,0315	0	
$\hat{\beta}_2$	0,9044	0,8803	0,8860	-	
Erro Padrão	0,0231	0,0242	0,0121	-	
<i>P-valor</i>	0	0	0	-	
$\hat{\beta}_3$	0,1745	0,1161	0,2345	-	
Erro Padrão	0,0400	0,0461	0,3249	-	
<i>P-valor</i>	0,0641	0,0059	0,2352	-	
$\hat{\beta}_4$	-	-0,2571	-	-	
Erro Padrão	-	0,0585	-	-	
<i>P-valor</i>	-	0	-	-	

4.4. AJUSTE E ADEQUAÇÃO PARA SÉRIE DE LOG-RETORNOS DA PETR453

Por fim, o modelo mais adequado entre os quatro vistos é o **valor absoluto simétrico**. As Tabelas 4.6 e 4.7 apresentam as estimativas dos modelos e dos testes de adequação. O modelo final, portanto, é:

$$\hat{Q}_{2,t}(\hat{\beta}) = 0,0228 + 0,9043\hat{Q}_{2,t-1}(\hat{\beta}) + 0,1745|r_{2,t-1}|.$$

Temos que a proporção de violações amostrais é bem próxima de 5%, e o valores acumulado de violações foi o menor deles. O valor acumulados dos desvios dos valores que não violaram o VaR foi o segundo menor, o que indica foi um dos que menos superestimou os valores ajustados. Todos os testes de adequação indicaram que a série de violações não apresenta dependência condicional e a proporção condicional e incondicional de violações foram próximas da esperada.

Tabela 4.7: continuação - Estimativas de Estatísticas de Teste dos Parâmetros Estimados para os Modelos CAViaR para a série PETR4.

Estimativas e Testes				
Modelo	Valor Absoluto Simétrico	Inclinação Assimétrica	GARCH(1,1)Indireto	Adaptativo
Hit_{am}	5,00%	5,04%	5,00%	4,80%
Hit_{pam}	5,99%	6,27%	5,90%	3,69%
QD_{am}	0,0104	0,4572	0,0132	0,1973
QD_{am} : p-valor	0,9997	0,9993	0,9775	0,6569
QD_{pam}	0,0186	0,0615	0,1074	0,1030
QD_{pam} : p-valor	0,9996	0,9996	0,9909	0,7483
CQ	271,86	270,60	273,83	281,73
Teste CI: LR	0,4419	0,8586	0,4419	1,0727
Teste CI: p-valor	0,5062	0,3541	0,5062	0,3003
Teste IND: LR	0,0032	0,0054	0,0032	0,7694
Teste IND: p-valor	0,9552	0,9416	0,9552	0,3804
Teste CC: LR	0,8837	1,7171	0,8837	2,1454
Teste CC: p-valor	0,6428	0,4238	0,6428	0,3421
Valor Acum. Viol.	80,71	84,59	83,46	76,55
Valor Acum. Não Viol.	2.495,28	2.389,57	2.470,85	2.747,98

4.5 Análise de Causalidade de Granger em Risco

A hipótese de causalidade de Granger em Risco será verificada a partir da série indicadora de violações ao VaR. As séries de VaR estimadas pelos modelos de inclinação assimétrica para BBDC4 e valor absoluto simétrico para PETR4 são obtidas respectivamente por $V_{1,t} = \hat{Q}_{1,t}(\hat{\beta})$ e $V_{2,t} = \hat{Q}_{2,t}(\hat{\beta})$.

A seguir, serão consideradas as seguintes transformações de variáveis sobre $r_{1,t}$, $r_{2,t}$, $V_{1,t}$ e $V_{2,t}$:

$$Z_{lt} = \mathbb{I}\{r_{lt} < V_{lt}\}, l = 1, 2.$$

Assim, Z_{lt} é 1 quando a perda exceder ao VaR e 0 caso contrário.

Para identificar a causalidade de Granger em risco unidirecional, foram calculadas as estatísticas de teste para os núcleos truncado, Bartlett, Daniell, Parzen com $M = 5, 10, 20, 30$ e 40 . Esses valores para M foram selecionados para capturar as múltiplos de semanas de dias úteis (1 semana, 2 semanas, 1 mês, 1 mês e meio, 2 meses).

Foram testados três tipos de causalidade de Granger em risco. Por exemplo, a notação $\text{BBDC4} \Rightarrow \text{PETR4}$ indica o teste que verifica se BBDC4 causa PETR4 em risco. A notação $\text{BBDC4} \Leftrightarrow \text{PETR4}$, representa a relação de causalidade de Granger em risco com *feedback*. As Tabelas 4.8, 4.9 e 4.10 a seguir apresentam os resultados obtidos.

Na Tabela 4.8, temos os resultados das estimativas das estatísticas de teste de $\text{BBDC4} \Rightarrow \text{PETR4}$. Para todos as variações de núcleos e M , não rejeitamos a hipótese nula que BBDC4 não Granger-cause PETR4 em risco. Assim, utilizando diferentes formas de núcleos, podemos concluir que não há causalidade de Granger em risco, isto é, $\text{BBDC4} \nRightarrow \text{PETR4}$.

Na Tabela 4.9, temos os resultados para das estimativas das estatísticas de teste para $\text{PETR4} \Rightarrow \text{BBDC4}$.

Utilizando o núcleo de Bartlett com $M=5$, rejeitamos a hipótese que não PETR4 Granger-cause BBDC4 em risco. Logo, $\text{PETR4} \Rightarrow \text{BBDC4}$. O p-valor obtido foi de 4,98%, sendo menor que o nível de significância de 5%. Para o núcleo de Parzen, o p-valor aproximado foi de 5% (com uma precisão maior, o p-valor foi 4,9976%).

Tabela 4.8: Estimativas das Estatísticas de Teste de Causalidade de Granger em Risco (BBDC4 $\Rightarrow$ PETR4) para núcleos truncado, Bartlett, Daniell e Parzen com M variando em 5, 10, 20, 30 e 40

BBDC4 $\Rightarrow$ PETR4					
M	5	10	20	30	40
Q_{trunc} ($p - valor$)	-0,6812 (0,7521)	0,1823 (0,4277)	0,6104 (0,2708)	-1,3702 (0,9147)	0,7746 (0,2193)
Q_{bart} ($p - valor$)	0,2609 (0,3971)	-0,7333 (0,7683)	0,1222 (0,4514)	-0,0435 (0,5173)	-0,5315 (0,7025)
Q_{dan} ($p - valor$)	0,0934 (0,4628)	-0,8558 (0,8039)	0,2101 (0,4168)	0,1139 (0,4547)	-0,6082 (0,7285)
Q_{parz} ($p - valor$)	0,7090 (0,2392)	-0,8833 (0,8115)	-0,0311 (0,5124)	0,3384 (0,3675)	-0,1894 (0,5751)

Logo, também rejeitaríamos a hipótese de não causalidade de Granger em risco. Para os demais efeitos de M e núcleos, não rejeitamos a hipótese nula que BBDC4 não Granger-cause PETR4 em risco.

Tabela 4.9: Estimativas das Estatísticas de Teste de Causalidade de Granger em Risco (PETR4 $\Rightarrow$ BBDC4) para núcleos truncado, Bartlett, Daniell e Parzen com M variando em 5, 10, 20, 30 e 40

PETR4 $\Rightarrow$ BBDC4					
M	5	10	20	30	40
Q_{trunc} ($p - valor$)	0,7284 (0,2332)	0,6030 (0,2733)	-0,2273 (0,5899)	-2,4793 (0,9934)	-4,1607 (1,0000)
Q_{bart} ($p - valor$)	1,6473 (0,0498)	0,9978 (0,1592)	0,3865 (0,3496)	-0,3291 (0,6290)	-1,5762 (0,9425)
Q_{dan} ($p - valor$)	1,5894 (0,0560)	0,7437 (0,2285)	0,1597 (0,4365)	-0,6174 (0,7315)	-2,0137 (0,9780)
Q_{parz} ($p - valor$)	1,6451 (0,0500)	1,1841 (0,1182)	0,2535 (0,4000)	0,1413 (0,4438)	-0,7139 (0,7623)

Na Tabela 4.10, temos os resultados para das estimativas das estatísticas de teste de BBDC4 $\Leftrightarrow$ PETR4. Para todos as variações de núcleos e M, não rejeitamos a

Tabela 4.10: Estimativas das Estatísticas de Teste de Causalidade de Granger em Risco (BBDC4 $\Leftrightarrow$ PETR4) para núcleos truncado, Bartlett, Daniell e Parzen com M variando em 5, 10, 20, 30 e 40

BBDC4 $\Leftrightarrow$ PETR4					
M	5	10	20	30	40
Q_{trunc} ($p - valor$)	0,0333 (0,4867)	0,5543 (0,2897)	0,2704 (0,3934)	-2,7174 (0,9967)	-2,3903 (0,9916)
Q_{bart} ($p - valor$)	1,3470 (0,0890)	0,1867 (0,4260)	0,3591 (0,3598)	-0,2630 (0,6037)	-1,4878 (0,9316)
Q_{dan} ($p - valor$)	1,1879 (0,1174)	-0,0791 (0,5315)	0,2611 (0,3970)	-0,3555 (0,6389)	-1,8508 (0,9679)
Q_{parz} ($p - valor$)	1,5088 (0,0657)	0,1225 (0,4513)	0,5244 (0,3000)	-0,2736 (0,6078)	-1,0256 (0,8475)

hipótese nula que BBDC4 não Granger-causea PETR4 em risco com *feedback*. Assim, utilizando diferentes formas de núcleos, podemos concluir que não há causalidade de Granger em risco com feedback, ou seja, BBDC4 $\nRightarrow$ PETR4 .

4.6 Análise de Causalidade de Granger

Nessa seção serão expostos os resultados da análise de Causalidade de Granger usual. A partir dela, verificaremos se a causalidade de Granger é confirmada sem considerar os efeitos gerados pelos eventos de risco.

Para efeitos comparativos com os resultados apresentados para causalidade de Granger em risco, serão utilizados os mesmos fatores de ordem autorregressiva.

Para testar se BBDC4 $\Rightarrow$ PETR4, teremos os modelos baseados na seguinte estrutura:

$$r_{1,t} = \sum_{i=1}^{M_{11}} \alpha_{1,i} r_{1,t-i} + \sum_{i=1}^{M_{12}} \beta_{1,i} r_{2,t-i} + u_t \quad (4.1)$$

e, analogamente, para testar se PETR4 $\Rightarrow$ BBDC4, temos:

$$r_{2,t} = \sum_{i=1}^{M_{21}} \alpha_{2,i} r_{1,t-i} + \sum_{i=1}^{M_{22}} \beta_{2,i} r_{2,t-i} + v_t \quad (4.2)$$

nas quais u_t é $RB(0, \sigma_u)$ e $RB(0, \sigma_v)$.

Iremos aplicar os passos descritos na seção de Conceitos. Para verificar a relação BBDC4 $\Rightarrow$ PETR4, primeiramente, devemos encontrar a ordem máxima de defasagem autorregressiva s , variando $M_{11} = 1, 2, \dots, 24$ na equação:

$$r_{1,t} = \sum_{i=1}^{M_{11}} \alpha_{1,i} r_{1,t-i} + u_t. \quad (4.3)$$

A seguir, na Tabela 4.11 e Figura 4.13, temos os valores do FPE estimados para inclusão das defasagens até 24.

Tabela 4.11: Valores estimados do FPE para $i = 1, 2, \dots, 24$ na equação (4.3)

Defasagem	FPE	Defasagem	FPE
1	6,342.10 ⁻⁴	13	6,231.10 ⁻⁴
2	6,326.10 ⁻⁴	14	6,225.10 ⁻⁴
3	6,317.10 ⁻⁴	15	6,227.10 ⁻⁴
4	6,323.10 ⁻⁴	16	6,232.10 ⁻⁴
5	6,314.10 ⁻⁴	17	6,237.10 ⁻⁴
6	6,305.10 ⁻⁴	18*	6,210.10 ⁻⁴
7	6,304.10 ⁻⁴	19	6,215.10 ⁻⁴
8	6,291.10 ⁻⁴	20	6,222.10 ⁻⁴
9	6,256.10 ⁻⁴	21	6,224.10 ⁻⁴
10	6,253.10 ⁻⁴	22	6,231.10 ⁻⁴
11	6,246.10 ⁻⁴	23	6,230.10 ⁻⁴
12	6,229.10 ⁻⁴	24	6,236.10 ⁻⁴

O menor valor estimado do FPE ocorre para a defasagem $s=18$. As estimativas

Figura 4.13: Valores estimados do FPE para $i = 1, 2, \dots, 24$ na equação (4.3)

dos parâmetros da equação estão na Tabela 4.12.

Retirando os parâmetros estimados não significativos, foi obtida a seguinte equação reduzida:

$$\begin{aligned} \hat{r}_{1,t} = & -0,0438r_{1,t-2} - 0,0535r_{1,t-3} - 0,0493r_{1,t-5} - 0,0370r_{1,t-6} - \\ & -0,0366r_{1,t-7} + 0,0406r_{1,t-12} + 0,0458r_{1,t-14} - 0,0708r_{1,t-18} + \hat{u}_t \end{aligned} \quad (4.4)$$

cujas estimativas estão apresentadas na Tabela 4.13. O conjunto de defasagens do modelo reduzido é: $I_1 = \{2, 3, 5, 6, 7, 12, 14, 18\}$.

O FPE estimado da equação (4.4) reduziu para 0,0006177 (o modelo completo apresentava FPE= 0,0006210).

Tabela 4.12: Valores estimados do modelo completo (4.3) para s=18

Parâmetro	Estimativa	Erro Pad.	Est. Teste	p-valor	Signif. ¹
α_1	0,0252	0,0192	1,313	0,1893	
α_2	-0,044	0,0192	-2,339	0,0194	*
α_3	-0,053	0,0192	-2,788	0,0053	**
α_4	-0,000	0,0193	-0,01	0,9921	
α_5	-0,050	0,0192	-2,636	0,0084	**
α_6	-0,035	0,0193	-1,837	0,0662	.
α_7	-0,035	0,0192	-1,828	0,0677	.
α_8	-0,001	0,0192	-0,091	0,9275	
α_9	0,0108	0,0192	0,563	0,5732	
α_{10}	-0,011	0,0192	-0,602	0,5472	
α_{11}	-0,016	0,0191	-0,86	0,3897	
α_{12}	0,0426	0,0191	2,221	0,0264	*
α_{13}	0,0149	0,0191	0,776	0,4376	
α_{14}	0,0448	0,0191	2,34	0,0193	*
α_{15}	0,0168	0,0191	0,877	0,3807	
α_{16}	0,0110	0,0191	0,575	0,5653	
α_{17}	0,0107	0,0191	0,561	0,5751	
α_{18}	-0,069	0,0191	-3,654	0,0002	***

¹Códigos de Significância: 0(***) 0.001(**) 0.01(*) 0.05(.) 0.1()

Tabela 4.13: Valores estimados do modelo reduzido (4.4)

Parâmetro	Estimativa	Erro Pad.	Est. Teste	p-valor	Signif. ¹
α_1	-0,043	0,0191	-2,293	0,0219	*
α_2	-0,053	0,0191	-2,796	0,0052	**
α_3	-0,049	0,0191	-2,572	0,0101	*
α_4	-0,036	0,0191	-1,928	0,0539	.
α_5	-0,036	0,0191	-1,909	0,0564	.
α_6	0,0405	0,0190	2,129	0,0333	*
α_7	0,0458	0,0190	2,408	0,0160	*
α_8	-0,070	0,0189	-3,729	0,0001	***

¹Códigos de Significância: 0(***) 0.001(**) 0.01(*) 0.05(.) 0.1()

O próximo passo é acrescentar os termos defasados $r_{2,t-i}$, para $i = 1, \dots, M_{12}$, e encontrar n que apresente o menor valor estimado para o critério FPE. O modelo é:

$$r_{1,t} = \sum_{i \in I_1} \alpha_{1,i} r_{1,t-i} + \sum_{i=1}^{M_{12}} \beta_{1,i} r_{2,t-i} + u_t. \quad (4.5)$$

Figura 4.14: Valores estimados do FPE para $M_{12} = 1, 2, \dots, 18$ na equação (4.5)

Na Tabela 4.14 e Figura 4.14, temos os valores do FPE estimados para inclusão das defasagens $M_{12} = 1, 2, \dots, 18$. Temos que na defasagem $n = 5$ o critério FPE é o menor entre eles (0,0006181).

A equação para $n=5$ é a seguinte:

$$\begin{aligned} \hat{r}_{1,t} = & -0,0438r_{1,t-2} - 0,0535r_{1,t-3} - 0,0493r_{1,t-5} - 0,0370r_{1,t-6} - 0,0366r_{1,t-7} + \\ & + 0,0406r_{1,t-12} + 0,0458r_{1,t-14} - 0,0708r_{1,t-18} + 0,0092r_{2,t-1} \end{aligned}$$

Tabela 4.14: Valores estimados do FPE para $i = 1, 2, \dots, 18$ na equação (4.5)

Defasagem	FPE	Defasagem	FPE
1	$6,181.10^{-4}$	10	$6,231.10^{-4}$
2	$6,181.10^{-4}$	11	$6,225.10^{-4}$
3	$6,182.10^{-4}$	12	$6,227.10^{-4}$
4	$6,182.10^{-4}$	13	$6,232.10^{-4}$
5*	$6,181.10^{-4}$	14	$6,237.10^{-4}$
6	$6,185.10^{-4}$	15	$6,210.10^{-4}$
7	$6,186.10^{-4}$	16	$6,215.10^{-4}$
8	$6,184.10^{-4}$	17	$6,222.10^{-4}$
9	$6,188.10^{-4}$	18	$6,224.10^{-4}$

$$+0,0352r_{2,t-2} - 0,031r_{2,t-3} + 0,0283r_{2,t-4} + 0,0358r_{2,t-5} + \hat{u}_t. \quad (4.6)$$

As estimativas dos parâmetros na Tabela 4.15 mostra que todos os coeficientes dos termos defasados de $r_{2,t-i}$ para $i = 1, 2, 3, 4, 5$ não são significativos. O critério FPE para esse modelo foi de 0,0006181, que é pouco maior que o modelo reduzido estimado anteriormente. Considerando esses termos, será verificado se ao incluir serialmente os termos autoregressivos $r_{1,t-i}$, há alguma diminuição no critério FPE.

A partir do modelo (4.1), para $n = 5$, são incluídos um a um os termos autoregressivos de $r_{1,t}$ conforme a especificação:

$$r_{1,t} = \sum_{i \in I_1} \alpha_i r_{1,t-i} + \sum_{i=1}^5 \beta_i r_{2,t-i} + u_t. \quad (4.7)$$

A Figura 4.15 nos indica que quando incluímos serialmente os 8 termos autorregressivos de $r_{1,t-i}$, para $i \in I_1$, o FPE é o menor.

Assim, como o modelo não apresenta os parâmetros dos termos $r_{2,t-i}$ significativos para $i = 1, 2, 3, 4, 5$ e o critério FPE do modelo reduzido (0,0006177) foi menor que o do modelo completo, não há causalidade de Granger da série PETR4 para BBDC4.

Para testar se $PETR4 \Rightarrow BBDC4$, realizamos os mesmos passos. Primeiro, devemos encontrar a ordem máxima de defasagem autorregressiva s em (4.6), variando

Tabela 4.15: Valores estimados do modelo (4.5) para $n=5$

Parâmetro	Estimativa	Erro Pad.	Est. Teste	p-valor	Signif. ¹
α_3	-0,059	0,0219	-2,713	0,0067	**
α_4	-0,039	0,0220	-1,812	0,0701	.
α_6	-0,066	0,0220	-3,012	0,0026	**
α_7	-0,037	0,0191	-1,954	0,0507	.
α_8	-0,035	0,0191	-1,869	0,0617	.
α_{13}	0,0391	0,0190	2,052	0,0402	*
α_{15}	0,0465	0,0190	2,445	0,0145	*
α_{19}	-0,071	0,0189	-3,772	0,0001	***
β_1	0,0092	0,0205	0,452	0,6511	
β_2	0,0352	0,0236	1,493	0,1355	
β_3	-0,031	0,0236	-1,334	0,1823	
β_4	0,0283	0,0205	1,378	0,1683	
β_5	0,0358	0,0236	1,518	0,1290	

¹Códigos de Significância: 0(***) 0.001(**) 0.01(*) 0.05(.) 0.1()

Figura 4.15: Valores estimados do FPE para $i \in I_1$ na equação (4.7)

$M_{22} = 1, 2, \dots, 24$ na equação:

$$r_{2,t} = \sum_{i=1}^{M_{22}} \alpha_{2,i} r_{2,t-i} + v_t. \quad (4.8)$$

A seguir, na Tabela 4.16 e Figura 4.16, temos os valores do FPE estimados para inclusão das defasagens até 24.

Figura 4.16: Valores estimados do FPE para $i = 1, 2, \dots, 24$ na equação (4.8)

O menor valor estimado do FPE ocorre para a defasagem $s=18$. Na Tabela 4.17, constam as estimativas dos parâmetros da equação.

O critério FPE para o modelo completo foi de 0,0005364. Ao retirar os parâmetros não significativos, não há redução do FPE. Sendo assim $s=18$ e foram considerados todos os fatores autorregressivos no modelo.

No próximo passo, os fatores defasados da outra série $r_{1,t}$ foram adicionados gradualmente no modelo, para $M_{21} = 18$:

Tabela 4.16: Valores estimados do FPE para $M_{22} = 1, 2, \dots, 24$ na equação (4.8)

Defasagem	FPE	Defasagem	FPE
1	$5,381.10^{-4}$	13	$5,393.10^{-4}$
2	$5,381.10^{-4}$	14	$5,397.10^{-4}$
3	$5,372.10^{-4}$	15	$5,402.10^{-4}$
4	$5,376.10^{-4}$	16	$5,396.10^{-4}$
5	$5,381.10^{-4}$	17	$5,394.10^{-4}$
6	$5,382.10^{-4}$	18*	$5,363.10^{-4}$
7	$5,388.10^{-4}$	19	$5,369.10^{-4}$
8	$5,384.10^{-4}$	20	$5,373.10^{-4}$
9	$5,389.10^{-4}$	21	$5,378.10^{-4}$
10	$5,394.10^{-4}$	22	$5,379.10^{-4}$
11	$5,392.10^{-4}$	23	$5,382.10^{-4}$
12	$5,387.10^{-4}$	24	$5,385.10^{-4}$

$$r_{2,t} = \sum_{i=1}^{M_{21}} \alpha_{2,i} r_{1,t-i} + \sum_{i=1}^{18} \beta_{1,i} r_{1,t-i} + v_t. \quad (4.9)$$

Nesse caso, para $i = 1, \dots, 18$ nos termos de $r_{1,t-i}$, somente o primeiro termo apresenta o melhor valor estimado do critério FPE, como pode ser visto na Figura 4.17 ou na Tabela 4.18.

As estimativas dos parâmetros podem ser vistos na Tabela 4.19. A equação completa para a equação (4.9) é a seguinte:

$$\begin{aligned} \hat{r}_{2,t} = & 0,0050r_{1,t-1} + 0,0297r_{2,t-1} - 0,027r_{2,t-2} - 0,054r_{2,t-3} + 0,0168r_{2,t-4} + 0,0022r_{2,t-5} - \\ & - 0,018r_{1,t-6} + 0,0020r_{2,t-7} - 0,035r_{2,t-8} - 0,009r_{2,t-9} + 0,0123r_{2,t-10} - 0,037r_{2,t-11} + \\ & + 0,0419r_{1,t-12} + 0,0033r_{2,t-13} - 0,004r_{2,t-14} - 0,003r_{2,t-15} + 0,0446r_{2,t-16} + 0,0382r_{2,t-17} - \end{aligned}$$

Tabela 4.17: Valores estimados do modelo completo (4.8) para $s=18$

Parâmetro	Estimativa	Erro Pad.	Est. Teste	p-valor	Signif. ¹
α_1	0,0324	0,0191	1,689	0,0913	.
α_2	-0,027	0,0191	-1,419	0,1560	
α_3	-0,054	0,0191	-2,82	0,0048	**
α_4	0,0166	0,0192	0,869	0,3847	
α_5	0,0023	0,0192	0,121	0,9036	
α_6	-0,018	0,0192	-0,944	0,3451	
α_7	0,0020	0,0191	0,106	0,9152	
α_8	-0,035	0,0191	-1,859	0,0631	.
α_9	-0,009	0,0191	-0,479	0,6322	
α_{10}	0,0123	0,0191	0,643	0,5200	
α_{11}	-0,037	0,0191	-1,941	0,0523	.
α_{12}	0,0419	0,0191	2,183	0,0291	*
α_{13}	0,0033	0,0192	0,175	0,8613	
α_{14}	-0,004	0,0192	-0,251	0,8020	
α_{15}	-0,003	0,0192	-0,193	0,8468	
α_{16}	0,0446	0,0191	2,326	0,0200	*
α_{17}	0,0382	0,0192	1,993	0,0463	*
α_{18}	-0,082	0,0191	-4,316	0,0000	***

¹Códigos de Significância: 0(***) 0.001(**) 0.01(*) 0.05(.) 0.1()

Tabela 4.18: Valores estimados do FPE para $s=18$ e $i = 1, 2, \dots, 18$ na equação (4.9)

Defasagem	FPE	Defasagem	FPE
1*	$0,0005371.10^{-4}$	10	$5,399.10^{-4}$
2	$0,0005375.10^{-4}$	11	$5,399.10^{-4}$
3	$0,0005379.10^{-4}$	12	$5,403.10^{-4}$
4	$0,0005380.10^{-4}$	13	$5,406.10^{-4}$
5	$0,0005383.10^{-4}$	14	$5,406.10^{-4}$
6	$0,0005385.10^{-4}$	15	$5,405.10^{-4}$
7	$0,0005389.10^{-4}$	16	$5,409.10^{-4}$
8	$0,0005393.10^{-4}$	17	$5,412.10^{-4}$
9	$0,0005396.10^{-4}$	18	$5,415.10^{-4}$

Figura 4.17: Valores estimados do critério FPE para $i = 1, 2, \dots, 18$

$$-0,082r_{2,t-18} + \hat{v}_t. \quad (4.10)$$

O critério FPE para o modelo completo foi 0,0005371, que é maior que o modelo ao se considerar somente os termos autorregressivos de $r_{2,t}$. Como o parâmetro de $r_{1,t-1}$ não é significativo, isso nos leva à conclusão de que não há causalidade de Granger da série BBDC4 em PETR4.

Para verificar se existe causalidade instantânea entre as duas séries, foi calculada a estimativa da função de autocorrelação cruzada entre os resíduos estimados u_t e v_t das equações completas (4.6) e (4.10). O valor obtido foi 0,4854. Logo, devemos rejeitar a hipótese que $r_{\hat{u}\hat{v}}(0) = 0$. Isso significa que há causalidade de Granger instantânea entre BBDC4 e PETR4.

Tabela 4.19: Valores estimados do modelo completo (4.9) para $s = 18$ e $n = 1$

Parâmetro	Estimativa	Erro Pad.	Est. Teste	p-valor	Signif. ¹
α_1	0,0297	0,0220	1,353	0,1763	
α_2	-0,027	0,0191	-1,416	0,1568	
α_3	-0,054	0,0191	-2,824	0,0047	**
α_4	0,0168	0,0192	0,876	0,3811	
α_5	0,0022	0,0192	0,117	0,9069	
α_6	-0,018	0,0192	-0,943	0,3459	
α_7	0,0020	0,0191	0,107	0,9150	
α_8	-0,035	0,0192	-1,845	0,0651	,
α_9	-0,009	0,0192	-0,474	0,6354	
α_{10}	0,0123	0,0191	0,642	0,5206	
α_{11}	-0,037	0,0191	-1,942	0,0522	,
α_{12}	0,0419	0,0191	2,183	0,0291	*
α_{13}	0,0033	0,0192	0,174	0,8621	
α_{14}	-0,004	0,0192	-0,256	0,7981	
α_{15}	-0,003	0,0192	-0,195	0,8455	
α_{16}	0,0446	0,0191	2,325	0,0201	*
α_{17}	0,0382	0,0192	1,989	0,0467	*
α_{18}	-0,082	0,0191	-4,318	0,0000	***
β_1	0,0050	0,0204	0,246	0,8057	

¹Códigos de Significância: 0(***) 0.001(**) 0.01(*) 0.05(.) 0.1()

Capítulo 5

Considerações Finais

Neste trabalho foi apresentada uma metodologia de análise de causalidade de Granger em Risco. O objetivo foi estabelecer passos para identificação de eventos considerados de risco a partir da avaliação da medida do VaR e assim verificar através das hipóteses formuladas, se uma série causa a outra no sentido de Granger por eventos extremos. Foi possível observar que o resultado é sensível às variações de núcleos e janelas no estimador da função de densidade espectral cruzada. Essa escolha deve ser dada de acordo com a importância que se queira dar para as observações mais recentes ou distantes no tempo.

O primeiro passo foi realizar o tratamento dos dados, ajustando a série para eventos corporativos e mudança no fator de cotação de negociação. Através da análise descritiva das duas séries de log-retornos de preços, foi visto que ambas possuem caudas pesadas e assimetria.

Os modelos CaViaR foram utilizados para ajustar o VaR ao longo do período para as duas séries. Foram escolhidos alguns tipos de especificação de modelos CaViaR para se analisar qual era o mais adequado, dado o comportamento das séries. O ajuste do modelo foi avaliado segundo alguns testes de adequação e medidas auxiliares para verificar a ordem de grandeza das violações ao VaR. Foi priorizado a escolha do modelo que apresentasse as condições satisfatórias para os testes de adequação e que melhor fosse avaliado no quesito do valor acumulado de violações. Para a série de log-retornos de preços da ação BBDC4 o modelo escolhido foi o inclinação assimétrica.

Já o modelo ajustado para PETR4 foi o valor absoluto simétrico.

Ajustado o modelo, as duas séries foram transformadas em séries de indicadores de eventos de risco, cujas ocorrências se dão através das violações ao VaR. A partir delas, foram testadas as hipóteses de causalidade de Granger em risco. Não foi verificada a causalidade de Granger em risco significativa entre as duas séries, exceto na utilização do núcleo de Bartlett com $M=5$ na qual $PETR4 \Rightarrow BBDC4$.

Somente para efeito comparativo, foi avaliado se há causalidade de Granger usual entre as séries. Foi escolhida a metodologia proposta por Hsiao (1979) [6] para realizar esse teste. Não houve evidência de causalidade de Granger entre as séries. Através do teste de causalidade instantânea, proposto por Pierce e Haugh (1977) [15], vimos que há causalidade instantânea entre as séries.

Apêndice A

Suposições para Teoremas

Suposições utilizadas para os Teoremas B.0.1 e B.0.2

1. Para $l = 1, 2$, Y_{lt} é um processo estocástico com função de distribuição desconhecida, continuamente diferenciável duas vezes, $F_{lt}(y) = P(Y_{lt} \leq y | \mathbf{I}_{t-1})$, onde $y \in \mathbb{R}$ e $\mathbf{I}_{t-1}$ é toda informação disponível até o tempo $t - 1$.
2. Para $\beta_l \in \Theta \subset R^{d_l}$, onde d_l é um inteiro positivo, $l = 1, 2$, $V_{lt}(\beta_l) = V_l(\mathbf{I}_{l(t-1)}, \beta_l)$ é um modelo de VaR ao nível $\tau \in (0, 1)$ tal que:
 - (a) para cada $\beta_l \in \Theta$, $V_{lt}(\beta_l)$ é função mensurável de $\mathbb{I}_{l(t-1)}$;
 - (b) com probabilidade 1, $V_{lt}(\cdot)$ é duas vezes continuamente diferenciável com relação a $\beta_l \in \Theta$, com

$$\lim_{T \rightarrow \infty} T^{-1} \sum_{t=1}^T E \left[\sup_{\beta_l \in \Theta} \left\| \frac{\partial}{\partial \beta} F_{lt}[-V_{lt}(\beta_l)] \right\|^4 \right] < \infty$$

e

$$\lim_{T \rightarrow \infty} T^{-1} \sum_{t=1}^T E \left[\sup_{\beta_l \in \Theta} \left\| \frac{\partial^2}{\partial \beta \partial \beta'} F_{lt}[-V_{lt}(\beta_l)] \right\|^2 \right] < \infty.$$

3. Para $l = 1, 2$, existe algum $\beta_l^0 \in \Theta$ tal que:

- (a) $P(Y_{lt} < -V_{lt}(\beta_l^0) | \mathbb{I}_{l(t-1)}) = \tau$;

- (b) o indicador de risco $Z_{2t}(\beta_2^0) = \mathbb{I}[Y_{2s} < -V_{2s}(\beta_2^0)]$ depende do passado de comprimento finito e longo arbitrariamente de $\{Z_{1s}(\beta_1^0) = \mathbb{I}[Y_{1s} < -V_{1s}(\beta_1^0)], s < t\}$.
4. $T^{\frac{1}{2}}(\hat{\beta}_l - \beta_l^*) = O_p(1)$ para $l = 1, 2$, onde $\beta_l^* = \text{plim} \hat{\beta}_l$ e $\beta_l^* = \beta_l^0$ sob a hipótese nula de interesse.
5. Para $S_t(\beta) = [S_{1t}(\beta_1)', S_{2t}(\beta_2)']'$, onde $S_{lt}(\beta_l) = \frac{\partial F_{lt}[-V_{lt}(\beta_l)]}{\partial \beta_l}$.
- (a) $\sum_{j=0}^{\infty} \|\Gamma(j)\| \leq \Delta$, onde $\Gamma(j) = \text{cov}[S_t(\beta)', Z_{t-j}(\beta^*)']$;
- (b) $\sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \|\kappa_0(j, k, l)\| \leq \Delta$, onde $\kappa_0(j, k, l)$ é a função de distribuição acumulada de quarta ordem de $S_t(\beta^*)' - ES_t(\beta^*)'$.
6. Para $k: \mathbb{R} \rightarrow [-1, 1]$ é função simétrica que é contínua em zero e em todos os pontos exceto em um número finito de pontos em $\mathbb{R}$, com $k(0) = 1$ e $\int_{-\infty}^{\infty} k^2(z) dz$.

Apêndice B

Teoremas

Teorema B.0.1 (Consistência) *No modelo (2.2), sobre as suposições que seguem, temos que $\hat{\beta} \xrightarrow{d} \beta$, onde $\hat{\beta}$ é a solução de (2.3). As suposições para que o teorema seja válido são as seguintes:*

1. $\{\epsilon_{t\tau}, \mathbf{x}_t\}$, $t = 1, 2, \dots$ são vetores aleatórios do espaço completo de probabilidade (Ω_t, F, P) ;
2. A função $Q_t(\beta) : \mathbb{R}^{k_t} \times B \rightarrow \mathbb{R}$ é tal que para cada $\beta \in B$, um subconjunto compacto de $\mathbb{R}^p$, $Q_t(\beta)$ é mensurável com relação ao conjunto Ω_t e $Q_t(\cdot)$ é contínua em B , $t=1,2,\dots$ para uma escolha dada de variáveis explicativas $\{y_{t-1}, \mathbf{x}_{t-1}, \dots, y_1, \mathbf{x}_1\}$;
3. O erro $\epsilon_{t\tau}$ condicionado a toda informação passada Ω_t é processo estacionário, com distribuição condicional de densidade $h_t(\epsilon_{t\tau}|\Omega_t)$;
4. Existe $h > 0$ tal que para todo t , $h_t(0|\Omega_t) \geq h$;
5. $|Q_t(\beta)| < K(\Omega_t)$ para cada $\beta \in B$ e para todo t , onde $K(\Omega_t)$ é uma função (possivelmente) estocástica de variáveis que pertencem ao conjunto de informação tal que $E(|K(\Omega_t)|) \leq K_0 < \infty$ para alguma constante K_0 ;
6. $E(|\epsilon_{t\tau}|) < \infty$ para todo t ;
7. $[\tau - \mathbb{I}(y_t < Q_t(\beta))] [y_t - Q_t(\beta)]$ obedece a lei uniforme dos grandes números;

8. Para todo $\xi > 0$, existe $\delta > 0$ tal que se $\|\beta - \beta^0\| \geq \xi$, então $\liminf_{n \rightarrow \infty} n^{-1} \sum P[|Q_t(\beta) - Q_t(\beta^0)| > \delta] > 0$.

Teorema B.0.2 (Normalidade assintótica) No modelo (2.2), com as seguintes suposições e condições do Teorema (B.0.1), temos que

$$\sqrt{n} \mathbf{A}_n^{-\frac{1}{2}} \mathbf{D}_t(\hat{\beta} - \beta^0) \xrightarrow{d} N(\mathbf{0}, \mathbb{I})$$

onde β é obtido por (2.3) e

$$A_n = E \left[n^{-1} \tau(1 - \tau) \sum_{i=1}^n \nabla' Q_t(\beta^0) \nabla Q_t(\beta^0) \right],$$

$$D_n = E \left[n^{-1} \sum_{i=1}^n h_t(0|\Omega_t) \nabla' Q_t(\beta^0) \nabla Q_t(\beta^0) \right].$$

Além das condições do Teorema (B.0.1), devem também ser satisfeitas as seguintes condições:

1. $Q_t(\beta)$ é diferenciável em B para todo β e γ em uma vizinhança v_0 de β^0 , tal que $\|\beta - \gamma\| \leq d$ para d suficientemente pequeno e para todo t :
 - (a) $\|\nabla Q_t(\beta)\| \leq F(\Omega_t)$, onde $F(\Omega_t)$ é alguma função (possivelmente) estocástica que pertence ao conjunto de informações e $E[F(\Omega_t)^3] \leq F_0 < \infty$ para alguma constante F_0 ;
 - (b) $\|\nabla Q_t(\beta) - \nabla Q_t(\gamma)\| \leq M(\Omega_t, \beta, \gamma) = O(\|\beta - \gamma\|)$, onde $M(\Omega_t, \beta, \gamma)$ é alguma função tal que $E[M(\Omega_t, \beta, \gamma)]^2 \leq M_0 \|\beta - \gamma\| < \infty$ e $E[M(\Omega_t, \beta, \gamma)] \leq M_1 \|\beta - \gamma\| < \infty$ para algumas constantes M_0 e M_1 .
2. (a) $h_t(\epsilon|\Omega_t) \leq N < \infty \forall t$, para alguma constante N ;
- (b) $h_t(\epsilon|\Omega_t)$ satisfaz a condição Lipschitziana $|h_t(\lambda_1|\Omega_t) - h_t(\lambda_2|\Omega_t)| \leq L|\lambda_1 - \lambda_2|$ para alguma constante $L < \infty$;

(c) As matrizes

$$A_n = E \left[n^{-1} \tau (1 - \tau) \sum_{i=1}^n \nabla' Q_t(\beta^0) \nabla Q_t(\beta^0) \right],$$

$$D_n = E \left[n^{-1} \sum_{i=1}^n h_t(0|\Omega_t) \nabla' Q_t(\beta^0) \nabla Q_t(\beta^0) \right]$$

tem os menores auto-valores limitados inferiormente por uma constante T suficientemente grande.

(d)

$$\{n^{-\frac{1}{2}} \sum_{i=1}^n [\tau - \mathbb{I}(y_t < Q_t(\beta^0))] \nabla' Q_t(\beta^0)\}.$$

Teorema B.0.3 (Estimação da matriz de variância-covariância) No modelo (2.2), com as suposições a seguir e as condições do Teoremas (B.0.1) e (B.0.2), temos que:

$\hat{A}_n \xrightarrow{p} A_n$ e $\hat{D}_n \xrightarrow{p} D_n$, onde

$$\hat{A}_n = \frac{\nabla' Q_t(\hat{\beta}) \nabla Q_t(\hat{\beta})}{n}$$

e

$$\hat{D}_n = \frac{\sum_{i=1}^n \mathbb{I}(|y_t - Q_t(\hat{\beta})| < \hat{c}_n) \nabla' Q_t(\hat{\beta}) \nabla Q_t(\hat{\beta})}{2n\hat{c}_n}.$$

1. $\frac{\hat{c}_n}{c_n} \xrightarrow{p} 1$, na qual a sequência não estocástica c_n satisfaz $c_n = O(1)$ e $c_t^{-1} = O\left(n^{\frac{1}{2}}\right)$.

2. $E(|F(\Omega)_t|^4) \leq F_1 \leq \infty$, para todo t e alguma constante F_1 , onde $F(\Omega)$ foi definida na condição 1.a do Teorema (B.0.1).

3.

$$\frac{\tau(1 - \tau) \sum_{i=1}^n \nabla' Q_t(\beta^0) \nabla Q_t(\beta^0)}{n} - A_n \xrightarrow{p} \mathbf{0}$$

e

$$\frac{h_t(0|\Omega_t)\nabla'Q_t(\boldsymbol{\beta}^0)\nabla Q_t(\boldsymbol{\beta}^0)}{n} - A_n \xrightarrow{p} \mathbf{0}.$$

Referências Bibliográficas

- [1] Christoffersen, P. F. (1998), Evaluating Interval Forecasts, *International Economic Review*, **39**, 841-862.
- [2] Engle, R. F., and Manganelli, S. (2004), CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles, *Journal of Business and Economic Statistics*, **22**, 367-381.
- [3] Granger, C.W.J (1969), Investigating causal relations by econometric models and cross-spectral methods, *Econometrica*, **37**, 424-438.
- [4] Granger, C.W.J (1980), Testing for causality: A personal view, *Journal of Economic Dynamics and Control*, **2**, 329-352.
- [5] Granger, C. W. J. and Hatanaka, M., *Spectral Analysis of Economic Time Series*, Princeton University Press, third Ed., New Jersey, 1964.
- [6] Hsiao, C. (1979), Autoregressive Modeling of Canadian Money and Income Data, *Journal of the American Statistical Association*, **367**, 553-560.
- [7] Hong, Y., Liu, I., Wang, S. (2010), Granger causality in risk and detection of extreme risk spillover between financial markets, *Journal of Econometrics*, **150**, 271-287.
- [8] Jenkins, G. M. and Watts, D. G., *Spectral Analysis and its applications*, Holden-Day Series, first Ed., London, Holden-Day, 1969.

- [9] Jorion, P., *Value at Risk: A Nova Fonte de Referência para o Controle do Risco de Mercado*, BM&F - Bolsa de Mercadorias & Futuros, São Paulo, 1999.
- [10] Koenker, R. and Basset, G. (1978), Regression Quantiles, *Econometrica*, **46**, 33-50.
- [11] Koenker, R. W. (2005), *Quantile Regression*, Cambridge, UK: Cambridge University Press.
- [12] Kuester, K., Mittnik, S., and Paolella, M., S. (2005), Value-at-Risk Prediction: A Comparison of Alternative Strategies, *Journal of Financial Econometrics*, **4**, 53-89.
- [13] Morettin, P. A., *Econometria Financeira*, Segunda Ed., São Paulo, 2008.
- [14] Morettin, P. A. e Toloi, C. M. C. (2006), *Análise de Séries Temporais*, Segunda Ed., Editora Edgard Blucher, São Paulo, 2006.
- [15] Pierce, D.A. and Haugh, L. (1977), Causality in temporal systems: Characterizations and a survey, *Journal of Econometrics*, **5**, 265-293.
- [16] Tsay, R. S. , *Analysis of Financial Time Series*, Wiley-Interscience, Second Ed., New Jersey, 2005.
- [17] Zivot, W., and Wang, J. , *Modeling Financial Time Series With S-Plus*, Springer, second Ed., New York, 2006.