

Testando o poder preditivo do VIX: uma aplicação do modelo de erro multiplicativo*

Luis Fernando Pereira Azevedo[†]
EESP-FGV

Pedro L. Valls Pereira[‡]
CEQEF e EESP-FGV

9 de junho de 2011

Resumo

O VIX – Volatility Index – surgiu como uma alternativa no cálculo da volatilidade implícita, visando mitigar alguns problemas encontrados em modelos da família Black-Scholes. Este tipo de volatilidade é tida como a melhor previsora da volatilidade futura, dado que as expectativas dos operadores de opções se encontram embutidas em seus valores. O objetivo deste trabalho é testar se o VIX apresenta maior poder preditivo e informações relevantes não presentes em modelos de séries temporais para variáveis não-negativas, tratadas através do modelo de erro multiplicativo. Os resultados indicam que o VIX apresenta maior poder preditivo em períodos de estabilidade econômica, mas não contém informação relevante frente à real volatilidade. Em períodos de crise econômica o resultado se altera, com o VIX apresentando o mesmo poder explicativo, mas contém informações relevantes no curto prazo.

Palavras-chave: VIX, modelos GARCH, modelo de erro multiplicativo, GMM.

JEL Codes: C10, C12, C13, G10 e G15

Resumo

VIX - Volatility Index - emerged as an alternative calculation of implied volatility in order to mitigate some problems encountered in models of the Black-Scholes. This kind of volatility is seen as the best predictor of future volatility, given that option traders' expectations are embedded in their values. In this paper we test whether the VIX has more predictive power and relevant information not found in time series models time for non-negative variables, treated by multiplicative model error. The results indicate that the VIX has greater predictive power in periods of economic stability, but does not contain relevant information to the actual volatility ahead. In periods of economic crisis the result changes, with the VIX presenting the same explanatory power, but contains relevant information in the short term.

Key words: VIX, GARCH models, multiplicative error model, GMM.

JEL Codes: C10, C12, C13, G10 e G15

*Ambos os autores agradecem financiamento do CNPq.

[†]Rua Itapeva 474 sala 1214, 01332-000, São Paulo, S.P. Brasil. E-mail: fpavezedo@gmail.com

[‡]Rua Itapeva 474 sala 1307, 01332-000, São Paulo, S.P. Brasil. E-mail: pedro.valls@fgv.br

1 Introdução

A crise no setor imobiliário americano gerou sérias consequências para a economia global. A falência do mercado de crédito neste segmento rapidamente contaminou todo o sistema financeiro mundial, criando uma recessão comparada por muitos à crise de 1929. Segundo o FMI, Estados Unidos, Japão e Reino Unido apresentaram retrações no PIB de 2,6%, 5,2% e 4,9% em 2009, respectivamente. Somente em janeiro de 2009, as bolsas mundiais acumularam perdas de US\$ 5,2 trilhões, segundo a agência de classificação americana Standard & Poor's.

Para Gorton (2008), a crise se iniciou com o aumento da inadimplência em hipotecas de alto risco, os chamados *subprime mortgages*, ainda em 2006. Mesmo com esta indicação de problemas num setor de importância significativa, a economia mundial continuou a apresentar altas taxas de crescimento até meados de 2008. Os PIBs americano, japonês e inglês subiram 2,0%, 2,4% e 3,0% em 2007. Em outubro do mesmo ano, o índice S&P500 atingiu seu recorde histórico, acima dos 1.500 pontos. Em julho, o índice Nikkei 225 ultrapassava os 18 mil pontos. No dia 9 de março de 2009, os índices cederam para 676,53 e 7.083,06 pontos, respectivamente. O mês de outubro de 2008 concentrou grande parte das perdas, com o S&P500 caindo 27% e o Nikkei 37%¹.

Durante o período de alta nos preços dos ativos, o VIX, um índice de volatilidade implícita das opções do S&P500, apresentou dois momentos de alta, um em junho de 2006 e outro em março de 2007, como pode ser observado na figura 1.

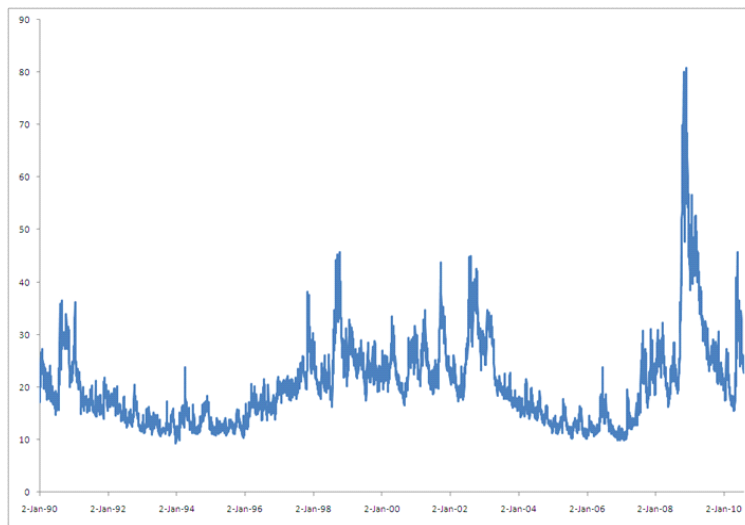


Figura 1: VIX - *Volatility Index*

Dado que a correlação entre volatilidade e preço dos ativos é negativa, o aumento em um índice de volatilidade significaria que o mercado estava esperando uma queda nos preços dos ativos, causada pelos distúrbios no mercado imobiliário americano.

¹Fonte: Yahoo Finance

Uma questão que se levanta é se a volatilidade obtida de forma implícita representa fielmente a volatilidade futura e, portanto, se é capaz de ajudar na previsão dos preços dos ativos. Engle e Gallo (2006) afirmam que como este tipo de volatilidade é construído a partir dos preços das opções negociadas pelos operadores, e estes embutem eventos futuros em suas análises, a volatilidade implícita supera, em termos de previsão, a volatilidade de séries temporais dado que o econometrista utiliza séries passadas, incapazes de absorver mudanças futuras. Por outro lado, modelos de volatilidade implícita estão disponíveis apenas para ativos que apresentam derivativos, como opções, por exemplo. Já modelos de séries temporais apenas necessitam da série de preços, estando então disponível para uma gama muito maior de ativos.

O objetivo deste artigo é comparar a aderência da volatilidade implícita e de modelos de volatilidade baseados em séries temporais à *proxies* da volatilidade real, como a volatilidade realizada, tanto dentro como fora da amostra. A volatilidade implícita utilizada foi o VIX, pois elimina uma série de problemas encontrados nos modelos Black-Scholes. Este índice será analisado como variável previsora, testando se ele contém informações relevantes para a volatilidade real que não são observáveis em modelos de erro multiplicativo para variáveis não-negativas. Poon e Granger (2003) citam que o período de análise pode influenciar a escolha do melhor modelo, como períodos de baixa e alta volatilidade, por exemplo. Por isso, os mesmos testes neste artigo são feitos para um período sem crise, onde o VIX permaneceu abaixo dos 20% e para período da crise imobiliária americana, onde o VIX ficou acima deste patamar.

O poder previsivo do índice cai ao se passar do período sem crise para o de crise, igualando-se aos modelos tradicionais de séries temporais. Porém, encontram-se evidências de que o VIX passa a conter informações relevantes para a volatilidade real no período de crise, pelo menos para um curto horizonte de tempo, que não são captadas pelos outros modelos, fato não presente no período sem crise.

Este artigo está dividido como se segue: na Seção 2 são apresentados os modelos de volatilidade mais conhecidos e a evolução cronológica do tema, assim como os principais trabalhos que realizam comparações entre eles. Modelos de volatilidade implícita são abordados na Seção 3, começando com o modelo Black-Scholes na Subseção 3.1, VIX na Subseção 3.2 e a comparação entre ambos na Subseção 3.3, demonstrando os problemas do Black-Scholes que o VIX tenta mitigar. A volatilidade realizada em tempo discreto é apresentada na Seção 4. A Seção 5 apresenta as séries utilizadas no artigo, assim como as justificativas na determinação do período sem crise e de crise. Na Seção 6 desenvolvem-se os testes do VIX como um predictor da volatilidade futura, com a demonstração do Modelo de Erro Multiplicativo, seguido da estimação na Subseção 6.1, a inferência na Subseção 6.2, o modelo utilizado nos testes empíricos, baseados no modelo de Engle e Gallo (2006), na Subseção 6.3 e finalmente os resultados obtidos na Subseção 6.4. A conclusão encontra-se na Seção 7, seguido das referências bibliográficas utilizadas na Seção 8. Por fim, é apresentado na Seção 9 o apêndice matemático, necessário às demonstrações do Modelo de Erro Multiplicativo da Seção 6.

2 Debate volatilidade implícita x volatilidade de séries temporais

Com a difusão dos modelos de séries temporais da família *Autoregressive Conditional Heteroskedasticity* (ARCH) durante a década de 80, a partir dos trabalhos seminais de Engle (1982) e Bollerslev (1986), surgiram concorrentes a altura na modelagem e previsão de volatilidade para os modelos de volatilidade implícita da década anterior, como o modelo de Black e Scholes (1973).

Antes desta classe de modelos, as previsões eram baseadas nos desvios padrões passados, fossem eles apenas uma média histórica ou uma média móvel dos desvios padrões ponderados exponencialmente, com maior peso para as observações mais recentes, conhecido como *Exponentially Weighted Moving Average* (EWMA). Ambos métodos podem carregar por muito tempo um choque passado já dissipado, e o EWMA apresenta o problema de que o parâmetro de persistência do modelo deve ser escolhido de maneira arbitrária, não sendo possível então fazer inferência sobre as estimativas do modelo. Taylor (1986) propôs utilizar modelos ARMA para o retorno absoluto e ao quadrado, podendo então realizar-se previsões para a volatilidade. Mais recentemente, como em Andersen e Bollerslev (1998), a volatilidade realizada passou a ser utilizada como *proxy* da volatilidade ao invés do retorno absoluto e ao quadrado.

Os modelos ARCH não utilizam os desvios padrões passados, mas atribuem uma lei de formação para a variância condicional não observada da série. Como se tem o valor da variância condicional estimada até a data t , é possível realizar diretamente a previsão para $t + 1$, apenas utilizando os parâmetros estimados da variância condicional e, realizando procedimentos iterativos, obtêm-se previsões para horizontes de tempo mais longos. Akgiray (1989), um dos primeiros a analisar o poder preditivo dos modelos GARCH, encontra evidências de que as previsões feitas por esses modelos superam de forma consistente os modelos EWMA e a simples volatilidade histórica. Ainda dentro da classe de modelos ARCH, existe o consenso de que modelos que permitem assimetria, ou seja, a relação negativa entre retorno e volatilidade, como os modelos EGARCH de Nelson (1991) e GJR de Glosten, Jagannathan e Runkle (1993), apresentam um melhor desempenho do que os modelos sem assimetria. Apesar de capturar a persistência da volatilidade, existem evidências empíricas de que a memória da volatilidade pode ser maior do que a apresentada nos modelos ARCH. Visando capturar este efeito, foram criados os modelos FIGARCH de Baillie, Bollerslev e Mikkelsen (1996) e o FIEGARCH de Bollerslev e Mikkelsen (1996) com ordem de integração entre 0 e 1.

No final da década de 80 surgiram outros modelos baseados em séries temporais: os modelos de volatilidade estocástica. Propostos por Hull e White (1987), estes modelos permitem que a volatilidade esteja sujeita a choques, que podem ou não estar ligados aos choques que afetam os próprios retornos, o que gera um maior grau de flexibilidade do modelo frente aos modelos ARCH. Segundo Poon e Granger (2003), esta flexibilidade gera dificuldades na estimação do modelo, pois não há uma forma fechada para a função de verossimilhança, ao contrário dos modelos ARCH, sendo então necessária a utilização de outros métodos de estimação, como o de quase-máxima verossimilhança, por exemplo. Devido a esta dificuldade no cálculo, a difusão destes modelos se deu apenas com o avanço computacional, durante a década de 90. Na comparação com modelos GARCH, não há um consenso de qual modelo é superior. Lehar, Scheicher e Schittenkopf (2002) afirmam que modelos GARCH são superiores aos modelos de volatilidade estocástica, enquanto que Nakajima (2010) dá suporte aos modelos de volatilidade estocástica, utilizando econometria *bayesiana*.

Em uma edição especial, a Revista de Econometria publicou em novembro de 1999 uma série de estudos comparando diversos modelos de volatilidade para os principais ativos brasileiros, incluindo ações, câmbio, títulos públicos e *commodities*. Issler (1999) encontra evidências de que o *EGARCH* (1,1) na sua versão gaussiana supera os demais modelos ARCH em se tratando de previsões. Utilizando diferentes funções perda para avaliar o melhor modelo, Valls Pereira et al. (1999) apresentam resultados em favor de modelos de volatilidade estocástica frente a modelos da família XARCH.

Modelos de volatilidade implícita utilizam o preço de derivativos de um certo ativo para obter, implicitamente, a volatilidade que os operadores atribuem ao ativo. O modelo mais famoso para a

captura da volatilidade implícita é o modelo de Black-Scholes, onde o preço de uma ação segue um processo estocástico com uma distribuição *log-normal*. Sob uma série de hipóteses, dentre elas da volatilidade do ativo ser considerada constante ou ser uma função determinística do tempo, é possível obter os preços das opções. Como estes preços são observados, é possível retirar a volatilidade do ativo inferida pelos operadores². Outra forma mais recente de cálculo da volatilidade implícita é o VIX (sigla para *Volatility Index*), feito pela *Chicago Board Options Exchange*, visando eliminar a série de hipóteses assumidas no modelo Black-Scholes. Na comparação com os demais modelos citados, a maioria dos trabalhos encontra evidências a favor da volatilidade implícita na previsão de volatilidade. Hol e Koopman (2002) comparam as previsões de um modelo de volatilidade estocástica, com e sem a adição do VIX como variável explicativa, e a própria volatilidade implícita, com resultados a favor da última frente aos outros modelos.

Na comparação da volatilidade implícita com modelos GARCH, trabalhos como Canina e Figlewski (1993) afirmam que praticamente não existe correlação entre a volatilidade implícita e a futura volatilidade realizada. Ao incrementarem modelos de séries de tempo com dados intradiários em modelos de memória longa, Martens e Zein (2003) concluem que estes modelos competem de forma igual aos modelos de volatilidade implícita, chegando até mesmo a superá-los em alguns casos. Mas uma gama bem maior de trabalhos segue na direção oposta. Christensen e Prabhala (1998) utilizam amostras mais longas e mostram que a volatilidade implícita é superior aos demais modelos. Para eles, o não tratamento do *crash* de 1987 no trabalho de Canina e Figlewski (1993) levou ao resultado equivocado. Szakmary et al. (2002) encontra resultado semelhante analisando 35 mercados de opções e 8 mercados de câmbio diferentes. Além de testar o poder de cada modelo separadamente, alguns trabalhos agregam a volatilidade implícita a modelos de séries de tempo, visando aumentar o poder de previsão destes. Blair, Poon e Taylor (2001) adicionam o VIX e a volatilidade realizada ao modelo *GJR*(1, 1) de Glosten, Jagannathan e Runkle (1993) como variáveis explicativas na equação da variância condicional, encontrando evidências de que a primeira apresenta uma importância maior no modelo, tanto em análises dentro ou fora da amostra.

A explicação para esta superioridade da volatilidade implícita, segundo Engle e Gallo (2006), é que os operadores de opções são capazes de mensurar eventos futuros até a maturidade das opções e adicionar seus efeitos nos preços das mesmas, algo impossível para um economista que está analisando uma série passada de preços. Desta forma, a volatilidade implícita em si é um melhor predictor da volatilidade futura do que modelos de séries temporais. Neste trabalho, os autores buscam modelos de séries de tempo capazes de aderirem ao VIX, mas a regressão do VIX nas previsões obtidas por eles apresentam correlação serial, fato interpretado como uma superioridade do VIX na capacidade de previsão da volatilidade.

Utilizando outros modelos de previsão diferentes dos usados por Engle e Gallo (2006), Becker, Clements e White (2007) testam se este resíduo é ortogonal a volatilidade realizada real, ou seja, se a parte do VIX não explicada por modelos de séries temporais contém algum poder explicativo sobre os dados reais. A conclusão dos autores é de que o VIX não contém esta informação adicional.

²Mais detalhes sobre o modelo Black-Scholes serão abordados na Seção 3.1.

3 Volatilidade Implícita

3.1 Black-Scholes

Nesta seção será desenvolvida uma breve descrição do modelo Black-Scholes, evidenciando-se seus principais pressupostos para que seja possível apresentar as diferenças e as vantagens do VIX sobre esta fórmula de cálculo da volatilidade implícita. Conforme apresentado em Hull (2008), as hipóteses do modelo são:

1. O preço de um ativo segue um movimento *Browniano* geométrico com tendência μ e volatilidade σ constantes
2. É possível a utilização de venda a descoberto
3. Não há custos de transação ou taxas e os ativos são perfeitamente divisíveis
4. Não há o pagamento de dividendos durante a vida do derivativo
5. Não existe oportunidade para arbitragem
6. As transações ocorrem de forma contínua
7. Existe uma taxa de juros livre de risco r , constante e igual para todas as maturidades

Em um período de tempo pequeno, a variação do preço do ativo e do derivativo são perfeitamente correlacionados. É possível então formar um portfólio livre de risco, comprando uma certa quantidade do ativo e ficando vendido nas opções de compra do mesmo ativo.

Seja S o preço do ativo, $f(S, t)$ o preço de um derivativo em função do preço do ativo subjacente e do tempo, c e p os preços de opções de compra e venda de uma opção européia, respectivamente, K o preço de exercício da opção, r a taxa de juros livre de risco anualizada, com capitalização contínua, μ a tendência anualizada de S , σ a volatilidade do ativo, calculada como o desvio padrão do logaritmo de S , t o tempo em anos, sendo 0 o tempo inicial e T o tempo final e Π o valor de um portfólio.

O modelo parte do princípio que o preço S de um ativo segue a um movimento *Browniano* geométrico

$$dS = \mu S dt + \sigma S dW \quad (1)$$

onde W é o movimento *Browniano* geométrico. Aplicando o lema de Itô³ no preço do derivativo $f(S, T)$, no caso o preço da opção, tem-se

$$df = \left(\frac{\partial f}{\partial S} \mu S + \frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) dt + \frac{\partial f}{\partial S} \sigma S dW \quad (2)$$

O valor de uma carteira teórica na estratégia *delta-hedge* com uma posição vendida na opção e $\frac{\partial f}{\partial S}$ ações compradas é

$$\Pi = -f + \frac{\partial f}{\partial S} S \quad (3)$$

Logo, Π evolui como

$$d\Pi = -df + \frac{\partial f}{\partial S} dS \quad (4)$$

Substituindo (1) e (2) em (4) obtém-se

$$d\Pi = \left(-\frac{\partial f}{\partial t} - \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) dt \quad (5)$$

³Maiores detalhes sobre o lema de Itô podem ser obtidos em Itô (1951).

e como esta equação não contém o elemento estocástico dW , o portfólio não apresenta risco⁴. Logo, para não haver possibilidade de arbitragem, o retorno do portfólio deve ser o mesmo da taxa de juros livre de risco r .

$$d\Pi = r\Pi dt \quad (6)$$

Substituindo (3) e (5) em (6)

$$\left(\frac{\partial f}{\partial t} + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \sigma^2 S^2 \right) dt = r \left(f - \frac{\partial f}{\partial S} S \right) dt$$

ou

$$\frac{\partial f}{\partial t} + rS \frac{\partial f}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = rf \quad (7)$$

A equação (7) é conhecida como a equação diferencial de Black-Scholes-Merton e corresponde aos diferentes derivativos baseados no ativo subjacente. Resolvendo esta equação⁵ chega-se ao preço da opção de compra e de venda

$$\begin{aligned} c &= S_0 N(d_1) - K e^{-rT} N(d_2) \\ p &= K e^{-rT} N(-d_2) - S_0 N(-d_1) \end{aligned}$$

onde

$$\begin{aligned} d_1 &= \frac{\ln(S_0/k) + (r + \sigma^2/2)T}{\sigma\sqrt{T}} \\ d_2 &= \frac{\ln(S_0/k) + (r - \sigma^2/2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T} \end{aligned}$$

e $N(x)$ é a função de distribuição acumulada da normal padrão.

A volatilidade σ da série de preços utilizada na fórmula não é diretamente observada, sendo necessária sua estimação *a priori*. A estratégia utilizada pelos operadores é retirar esta volatilidade através dos preços observados das opções, ou seja, de forma implícita. Como não é possível isolar σ nas equações acima, utiliza-se de procedimentos iterativos para obter a volatilidade implícita do ativo, dado o seu preço, o preço de exercício da opção, a taxa de juros livre de risco e o tempo de maturidade.

3.2 VIX

O índice de volatilidade VIX, sigla para *Volatility Index*, surgiu em 1993 e é calculado pela *Chicago Board Options Exchange* através das opções do S&P500 com base na seguinte fórmula:

$$\sigma^2 = \frac{2}{T} \sum_i \frac{\Delta K_i}{K_i^2} e^{RT} Q(K_i) - \frac{1}{T} \left[\frac{F}{K_0} - 1 \right]^2$$

onde $VIX = \sigma * 100$, T é o tempo de maturidade, F é o valor do índice futuro do S&P500, K_0 é o primeiro preço de exercício (*strike*) abaixo de F , K_i é o preço de exercício da i -ésima opção fora

⁴O portfólio não apresenta risco apenas em um pequeno período de tempo. Conforme S e t variam, $\frac{\partial f}{\partial S}$ deve ser alterado para manter a portfólio neutro ao risco.

⁵A demonstração da solução da equação diferencial foge do escopo desta dissertação. Maiores detalhes podem ser obtidos em Black e Scholes (1973).

do dinheiro, sendo uma opção de compra se $K_i > K_0$, uma opção de venda se $K_0 > K_i$ ou ambas se $K_0 = K_i$, $\Delta K_i = \frac{K_{i+1} - K_i}{2}$, R é a taxa livre de risco até a maturidade, em torno de 30 dias corridos, ou seja, 1 mês e $Q(K_i)$ é o ponto médio da diferença entre o preço de compra e venda (*spread bid-ask*).

Foram criados derivativos baseados no índice, como contratos futuros em 2004 e opções em 2006, por exemplo. Como há uma relação negativa entre retorno e volatilidade, o índice e seus derivativos são usados como uma forma de reduzir a variância de carteiras de ativos.

3.3 Debate VIX x Black-Scholes

Com a posse das diferenças metodológicas do VIX e do Black-Scholes, é possível observar os pontos críticos no cálculo da volatilidade implícita do modelo Black-Scholes e as correções que a metodologia do VIX aplica neste cálculo. A primeira diferença aparece no fato do VIX usar uma soma ponderada de preços de opções, e não apenas uma série de preços como no Black-Scholes.

Exceto pela hipótese de volatilidade constante, a violação de qualquer outra hipótese transforma o preço teórico em uma banda de preços, sendo então impossível retirar a volatilidade implícita do modelo. Outro problema no cálculo da volatilidade implícita de um ativo é o fato de que opções com o mesmo tempo de vida para um mesmo ativo podem apresentar volatilidades diferentes conforme o preço de exercício é alterado, o chamado sorriso da volatilidade. Na teoria, a volatilidade deveria ser igual para todos os preços de exercício. As volatilidades das opções mais próximas do dinheiro são menores do que as opções dentro ou fora do dinheiro. A explicação para este efeito vão desde a hipótese de que os retornos são normais (fato nem sempre verdadeiro), da volatilidade da equação diferencial ser estocástica, problemas de microestrutura e erros de medida, como liquidez, *bid-ask spread*, espaço entre as observações e preferência pelo risco por parte dos investidores.

Estes problemas não estão presentes no cálculo do VIX. O fato de serem utilizadas tanto opções dentro ou fora, com ponderações que tornam o índice mais próximo das opções no dinheiro, acaba por eliminar o sorriso de volatilidade. O uso do ponto médio no *bid-ask spread* ao invés do próprio preço da negociação é usado no cálculo para mitigar o problema de mudanças espúrias de preços (*bid-ask bounce*). Segundo Fleming et al (1995), o modelo Black-Scholes também não leva em conta o exercício antecipado das opções e/ou dividendos pagos pelo ativo subjacente.

4 Volatilidade Realizada

Reproduzindo o modelo simples para tempo discreto encontrado em McAleer e Medeiros (2008), os retornos diários são dados por

$$r_t = h_t^{1/2} \eta_t$$

onde h_t é a variância condicional, $\{\eta_t\}_{t=1}^T$ é uma sequência de variáveis aleatórias independentes e normalmente distribuídas com média zero e variância unitária, $\eta_t \sim NID(0, 1)$.

Suponha que, em um dia de negociação t , o logaritmo dos preços é observado *tick-by-tick*. Considere a grade $\Lambda_t = \{\tau_0, \dots, \tau_{n_t}\}$ contendo todos os pontos de observação e o conjunto $p_{i,t}$, $i = 1, \dots, n_t$ sendo o i -ésimo preço observado durante o dia t , onde n_t é o número total de observações no dia t . Além disso, suponha que

$$r_{t,i} = h_{t,i}^{1/2} \eta_{t,i}$$

onde $\eta_{t,i} \sim NID(0, n_t^{-1})$, $r_{t,i} = p_{t,i} - p_{t,i-1}$ é o i -ésimo retorno intra-período do dia t tal que

$$r_t = \sum_{i=0}^{n_t} r_{t,i}$$

e

$$h_t = \frac{1}{n_t} \sum_{i=1}^{n_t} h_{t,i}$$

Definindo o conjunto de informação $\mathfrak{F}_{t,i} \equiv \mathfrak{F} \{p_{a,b}\}_{a=-\infty}^{a=t, b=i}$ como a σ -algebra gerado por toda a informação para o i -ésimo *tick* no dia t . Portanto, $\mathfrak{F}_{t,0}$ é o conjunto de informação disponível antes do início do dia t . Segue-se daí que $E(r_t^2 | \mathfrak{F}_{t,0}) = h_t$ e $V(r_t^2 | \mathfrak{F}_{t,0}) = 2h_t^2$.

A variância realizada é definida como a soma de todos os quadrados dos retornos intradiários de alta frequência:

$$RV_t^{(all)} = \sum_{i=0}^{n_t} r_{t,i}^2 \quad (8)$$

O retorno diário ao quadrado pode ser escrito como

$$r_t^2 = \left(\sum_{i=0}^{n_t} r_{t,i}^2 + 2 \sum_{i=0}^{n_t-1} \sum_{j=i+1}^{n_t} r_{t,i} r_{t,j} \right)$$

tal que

$$\begin{aligned} E(r_t^2 | \mathfrak{F}_{t,0}) &= E\left(\sum_{i=0}^{n_t} r_{t,i}^2 | \mathfrak{F}_{t,0}\right) + 2E\left(\sum_{i=0}^{n_t-1} \sum_{j=i+1}^{n_t} r_{t,i} r_{t,j} | \mathfrak{F}_{t,0}\right) \\ &= E(RV_t^{(all)} | \mathfrak{F}_{t,0}) + 2E\left(\sum_{i=0}^{n_t-1} \sum_{j=i+1}^{n_t} r_{t,i} r_{t,j} | \mathfrak{F}_{t,0}\right) \end{aligned}$$

e se os retornos intradiários são não-correlacionados

$$E(r_t^2 | \mathfrak{F}_{t,0}) = E(RV_t^{(all)} | \mathfrak{F}_{t,0}) = h_t$$

Tem-se então dois estimadores não viesados para a variância média do retorno no dia t : o retorno ao quadrado do dia t e a variância realizada definida em (8). Entretanto, pode-se mostrar que

$$V(RV_t^{(all)} | \mathfrak{F}_{t,0}) = \frac{2}{n_t} \sum_{i=0}^{n_t} \frac{h_{t,i}^2}{n_t} < \frac{2}{n_t} \left(\sum_{i=0}^{n_t} \frac{h_{t,i}}{\sqrt{n_t}} \right)^2 = V(r_t^2 | \mathfrak{F}_{t,0})$$

dado que

$$E\left[\left(\sum_{i=1}^{n_t} h_{t,i} \eta_{t,i}^2\right)^2 | \mathfrak{F}_{t,0}\right] = \frac{3}{n_t^2} \sum_{i=0}^{n_t} h_{t,i}^2 + \frac{2}{n_t^2} \sum_{i=1}^{n_t-1} \sum_{j=i+1}^{n_t} h_{t,i} h_{t,j}$$

ou seja, a variância média dos retornos diários pode ser estimada com maior acurácia somando-se os retornos intradiários ao quadrado do que apenas calcular o retorno diário ao quadrado. Além

disso, quando os retornos são observados em qualquer frequência arbitrária, é possível estimar a variância média diária livre de erros de medida como

$$\lim_{n_t \rightarrow \infty} V \left(RV_t^{(all)} \mid \mathfrak{S}_{t,0} \right) = 0$$

O único requisito necessário sobre a dinâmica da variância dos retornos intradiários para que o resultado acima seja válido é que

$$\sum_{i=1}^{n_t} h_{t,i}^2 \propto n_t^{1+c}$$

onde $0 \leq c < 1$. A metodologia em tempo contínuo, baseada na variação quadrática, pode ser encontrada no mesmo trabalho dos autores.

Pela própria definição da metodologia, quanto maior a frequência dos dados, maior a acurácia na estimação da volatilidade. O ideal, portanto, é utilizar todos os *ticks* no cálculo. Mas podem existir ruídos de microestrutura nestes dados (como descasamento de compra e venda (*bid-ask bounce*), negociação assíncrona, negociações infrequentes, preços discretos, dentre outros fatores). Estes problemas de microestrutura são minimizados ao aumentar o tempo de coleta de dados. Temos então um *trade-off* entre acurácia e ruídos de microestrutura. Andersen e Bollerslev (1998), Madhavan (2000) e Biais, Glosten e Spatt (2005) mostram que o intervalo ótimo de dados para o S&P500, índice utilizado neste artigo, é de 5 em 5 minutos.

O critério adotado para a definição do período de crise foi o patamar de 20% para o VIX. Desconsiderando alguns picos de curto prazo, de 9 de maio de 2003 a 25 de julho de 2007 o índice permaneceu abaixo do valor de 20%, ultrapassando a marca a partir do dia 26 de julho de 2007, ficando acima deste valor até o final da amostra em 30 de julho de 2010, novamente desconsiderando alguns poucos dias onde o valor foi inferior a 20%. Portanto temos 1060 dias no período sem crise e 760 no período de crise.

5 Modelo de Erro Multiplicativo

Modelos GARCH podem ser utilizados em outras variáveis que não simplesmente o retorno das séries. Um dos pioneiros e o mais conhecido é o *Autoregressive Conditional Duration* (ACD) de Engle e Russell (1998), onde a duração irregular entre eventos é tratada através de modelos ARCH. A duração é um caso específico de variáveis não-negativas. Uma generalização do ACD é o modelo de erro multiplicativo (*Multiplicative Error Model – MEM*) desenvolvido por Engle (2002). A especificação deste modelo é dada por

$$x_t = \mu_t \varepsilon_t$$

onde x_t é um processo univariado não-negativo. Condicional ao conjunto de informação \mathfrak{S}_{t-1} , com informação disponível até $t-1$, μ_t é um processo não-negativo previsível, dependente do vetor de parâmetros desconhecidos θ

$$\mu_t = \mu_t(\theta)$$

e o erro ε_t é um processo i.i.d. estocástico condicional, com suporte não-negativo para a densidade, com média 1 e variância desconhecida

$$\sigma_t^2 \varepsilon_t \mid \mathfrak{S}_{t-1} \sim D(1, \sigma_t^2)$$

$$E(x_t | \mathfrak{S}_{t-1}) = \mu_t$$

$$V(x_t | \mathfrak{S}_{t-1}) = \sigma_t^2 \mu_t^2$$

Torna-se necessário então adotar uma função de densidade paramétrica para ε_t e especificar uma equação para μ_t . Assumindo uma distribuição Gama para o erro com parâmetros ϕ tem-se

$$\varepsilon_t | \mathfrak{S}_{t-1} \sim Gama(\phi, \phi)$$

com $E(\varepsilon_t | \mathfrak{S}_{t-1}) = 1$ e $V(\varepsilon_t | \mathfrak{S}_{t-1}) = 1/\phi$. Logo, utilizando-se das propriedades da função de distribuição Gama, listadas no Apêndice 10.1

$$x_t | \mathfrak{S}_{t-1} \sim Gama(\phi, \phi/\mu_t) \quad (9)$$

Para facilitar a estimação, a relação dada entre a função Gama e a distribuição generalizada do erro (GED), apresentada nos Apêndices 10.2. e 10.3., nos dá a seguinte relação

$$x_t | \mathfrak{S}_{t-1} \sim Gama(\phi, \phi/\mu_t) \Leftrightarrow x_t^\phi | \mathfrak{S}_{t-1} \sim 0.5 - GED\left(0, \mu_t^\phi, \phi\right)$$

Logo, a correspondência da densidade condicional de x_t e x_t^ϕ e dado por

$$x_t^\phi = \mu_t^\phi v_t$$

onde $v_t | \mathfrak{S}_{t-1} \sim 0.5 - GED(0, 1, \phi)$

Fazendo $\phi = 1/2$ temos

$$\sqrt{x_t} = \sqrt{\mu_t} v_t$$

onde

$$v_t | \mathfrak{S}_{t-1} \sim 0.5 - Normal(0, 1)$$

ou seja, os modelos devem ser estimados utilizando os segundos momentos condicionais da raiz quadrada das variáveis, impondo-se média zero e assumindo normalidade nos erros, usando rotinas habituais de estimações de modelos GARCH.

A média pode ser especificada na forma (p, q) como em Engle (2002)

$$\mu_t = \omega + \sum_{j=1}^p \alpha_j x_{t-j} + \sum_{j=1}^q \beta_j \mu_{t-j} + \gamma' z_t$$

com z_t sendo um vetor de variáveis exógenas fracas ou, como no caso deste artigo, variáveis pre-determinadas, dado que as variáveis a serem incluídas são defasadas em um período.

Uma condição suficiente para garantir que x_t seja não-negativo, caso todas variáveis em z_t sejam positivas, é que os parâmetros do modelo sejam positivos. Para assegurar que x_t seja estacionário em covariância, é condição suficiente que também z_t o seja e

$$\sum_{j=1}^p \alpha_j + \sum_{j=1}^q \beta_j < 1$$

Neste artigo será usado o GARCH na forma $(1, 1)$.

5.1 Estimação

Dado (9), a contribuição de x_t para o log da função de verossimilhança é dado por

$$l_t = \ln L_t = \phi \ln \phi - \ln \Gamma(\phi) + (\phi - 1) \ln x_t - \phi (\ln \mu_t + x_t/\mu_t)$$

enquanto que a contribuição para o *score* é

$$S_t = \begin{pmatrix} S_{t,\theta} \\ S_{t,\phi} \end{pmatrix}$$

sendo

$$\begin{aligned} S_{t,\theta} &= \nabla_{\theta} l_t = \phi \nabla_{\theta} \mu_t \left(\frac{x_t - \mu_t}{\mu_t^2} \right) \\ S_{t,\phi} &= \nabla_{\phi} l_t = \ln \phi + 1 - \psi(\phi) + \ln \left(\frac{x_t}{\mu_t} \right) - \frac{x_t}{\mu_t} \end{aligned}$$

onde $\psi(\phi) = \frac{\Gamma'(\phi)}{\Gamma(\phi)}$ é a função *Digama* e o operador ∇_{λ} são as derivadas em relação aos componentes de λ , com $\lambda = (\theta; \phi)$.

A contribuição de x_t para a matriz *Hessiana* é

$$H_t = \begin{pmatrix} H_{t,\theta\theta'} & H_{t,\theta\phi'} \\ H'_{t,\theta\phi} & H_{t,\phi\phi} \end{pmatrix}$$

sendo

$$\begin{aligned} H_{t,\theta\theta'} &= \nabla_{\theta\theta'} l_t = \phi \left(\frac{-2x_t + \mu_t}{\mu_t^3} \nabla_{\theta} \mu_t \nabla_{\theta'} \mu_t + \frac{x_t - \mu_t}{\mu_t^2} \nabla_{\theta\theta'} \mu_t \right) \\ H_{t,\theta\phi} &= \nabla_{\theta\phi} l_t = \frac{x_t - \mu_t}{\mu_t^2} \nabla_{\theta} \mu_t \\ H_{t,\phi\phi} &= \nabla_{\phi\phi} l_t = \frac{1}{\phi} - \psi'(\phi) \end{aligned}$$

onde $\psi'(\phi)$ é a função *Trigama*.

As condições de primeira ordem para θ e ϕ são dadas por

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \nabla_{\theta} \mu_t \frac{x_t - \mu_t}{\mu_t^2} &= 0 \\ \ln \phi + 1 - \psi(\phi) + \frac{1}{T} \sum_{t=1}^T \left[\ln \left(\frac{x_t}{\mu_t} \right) - \frac{x_t}{\mu_t} \right] &= 0 \end{aligned}$$

Observa-se que as condições de primeira ordem para θ não dependem de ϕ , o que implica que, independentemente do valor assumido de ϕ , as estimativas para θ serão as mesmas. A estimação por máxima verossimilhança de ϕ pode ser feita após a estimação de θ .

Além disso, se $\mu_t = E(x_t | \mathfrak{F}_{t-1})$, o valor esperado do *score* de θ avaliado no parâmetro verdadeiro é zero, independentemente da hipótese de $\varepsilon_t | \mathfrak{F}_{t-1}$ ter distribuição Gama. Desta forma, estas funções de máxima verossimilhança podem ser interpretadas como quase-máxima verossimilhança e, portanto, o estimador $\hat{\theta}$ é um estimador de quase-máxima verossimilhança.

5.2 Inferência

A matriz de variância-covariância assintótica do estimador de máxima verossimilhança é dada por⁶

$$V_\infty = \begin{pmatrix} \phi \lim_{t \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \frac{1}{\mu_t^2} \nabla_\theta \mu_t \nabla_{\theta'} \mu_t & 0 \\ 0 & \psi'(\phi) - \frac{1}{\phi} \end{pmatrix}^{-1}$$

Nota-se que mesmo ϕ não estando envolvido na estimativa de θ , a variância de $\hat{\theta}$ é proporcional a $1/\phi$. Também pode-se observar que os estimadores de máxima verossimilhança de $\hat{\theta}$ e $\hat{\phi}$ não são correlacionados assintoticamente.

Explorando a ortogonalidade de θ e ϕ , é possível o uso de diversos estimadores de ϕ . Um deles, proposto por Cipollini et al. (2006), após definir $u_t = x_t/\mu_t - 1$, tem-se $V(u_t | \mathfrak{S}_{t-1}) = \phi^{-1}$, o que leva ao estimador de método de momentos

$$\hat{\phi}^{-1} = \frac{1}{T} \sum_{t=1}^T \hat{u}_t^2$$

A inferência sobre θ e ϕ é baseada nas estimativas de V_∞ , que podem ser obtidas através das médias das Hessinas ou do produto dos gradientes avaliado nas estimativas $(\hat{\theta}, \hat{\phi})$.

O estimador sanduíche de *Huber-Eicker-White* é dado por

$$\hat{V}_\infty = \hat{H}_T^{-1} \widehat{OPG}_T \hat{H}_T^{-1}$$

onde

$$\begin{aligned} \hat{H}_T &= \frac{1}{T} \sum_{t=1}^T \hat{H}_T \\ \widehat{OPG}_T &= \frac{1}{T} \sum_{t=1}^T \hat{s}_t \hat{s}_t' \end{aligned}$$

elimina a dependência da submatriz relativa a θ e ϕ .

Portanto, para a estimação, basta impor média zero à raiz quadrada da variável de interesse e estimar um GARCH em qualquer software econométrico. Os erros padrões utilizados devem ser os erros padrões robustos de Bollerslev-Wooldridge, pois se a equação da média e da volatilidade estiverem corretamente especificadas, o estimador de quase-máxima verossimilhança é consistente e assintoticamente normalmente distribuído⁷.

5.3 Modelo de Engle e Gallo

Nesta seção, a metodologia utilizada para testar o poder preditivo do VIX será semelhante a apresentada por Engle e Gallo (2006). A proposta dos autores é criar um modelo GARCH, cuja sua previsão seja capaz de aderir a volatilidade implícita dada pelo VIX. As variáveis utilizadas em conjunto no trabalho foram retorno absoluto diário, o intervalo entre o máximo e o mínimo diário

⁶Cipollini, Engle e Gallo (2006)

⁷Bollerslev e Wooldridge (1992).

e a volatilidade realizada diária, todas não-negativas. A justificativa para o uso de tais variáveis é que a dinâmica entre as três pode variar significativamente ao longo dos dias. Como exemplo, pode-se ter um retorno diário zero, mas com alta ou baixa volatilidade durante o dia. Além disso, a amplitude pode ser alta, mesmo num dia de baixa volatilidade. A figura 2 sintetiza a discussão:

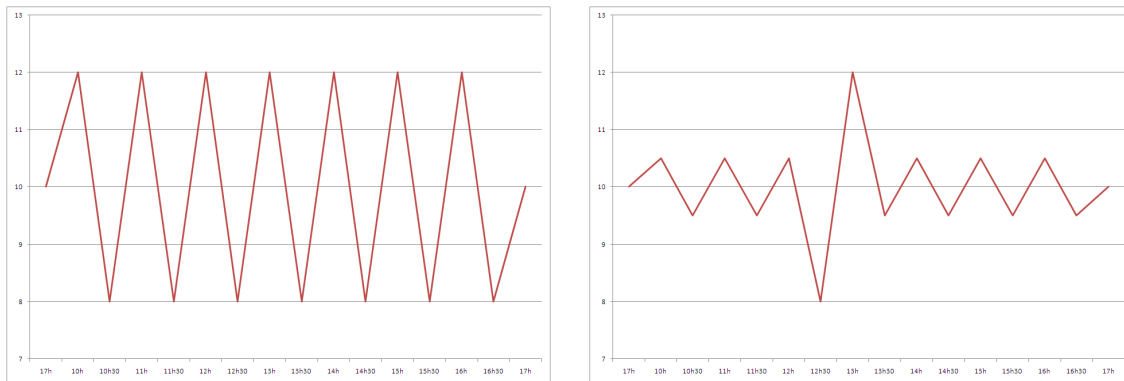


Figura 2: Diferentes dinâmicas diárias

Cada gráfico apresenta a dinâmica de um particular dia, com as cotações dadas a cada meia hora. Ambos apresentam retornos absolutos diários iguais a zero, com o preço de fechamento sendo igual ao fechamento do dia anterior. O máximo e o mínimo dos dias também é o mesmo, 12 e 8, respectivamente. Mas a volatilidade durante o dia é muito maior no primeiro gráfico. Outras combinações podem acontecer, mostrando a importância de cada variável na caracterização das negociações durante o dia.

Para cada uma das três variáveis tem-se

$$r_t^2 = h_t^r \varepsilon_t, \varepsilon_t \mid \mathfrak{S} \sim i.i.d.D(1, \xi^r)$$

$$hl_t^2 = h_t^h \eta_t, \eta_t \mid \mathfrak{S} \sim i.i.d.D(1, \xi^h)$$

$$v_t^2 = h_t^v \zeta_t, \zeta_t \mid \mathfrak{S} \sim i.i.d.D(1, \xi^v)$$

onde $r_t = \ln(C_t/C_{t-1})$, com C_t sendo o preço de fechamento do dia, $hl_t = \ln(H_t/L_t)$, onde H_t e L_t são o máximo e o mínimo dos preços durante o dia, respectivamente, e v_t é a raiz quadrada da variância realizada dada por $RV_t = \sum_{j=1}^M r_{t,j}^2$.

O sistema delas dado por

$$\begin{pmatrix} r_t^2 \\ hl_t^2 \\ v_t^2 \end{pmatrix} = \begin{pmatrix} h_t^r \\ h_t^h \\ h_t^v \end{pmatrix} \odot \begin{pmatrix} \varepsilon_t \\ \eta_t \\ \zeta_t \end{pmatrix}, \begin{pmatrix} \varepsilon_t \\ \eta_t \\ \zeta_t \end{pmatrix} \mid \mathfrak{S}_{t-1} \sim i.i.d.D(\nu, \Xi)$$

onde \odot indica o produto de Hadamard e Ξ é uma matriz de variância-covariância diagonal 3×3 , não apresentando perda de consistência, mas de eficiência no modelo. Segundo Davidian e

Carroll (1987), a utilização do retorno absoluto apresenta maior robustez frente a assimetria e não-normalidade do que os retornos ao quadrado na estimação da variância. Um dos primeiros trabalhos a utilizar os dados de preço máximo e mínimo diário foi Garman e Klass (1980).

Cada variável é modelada com a inclusão das demais variáveis defasadas, além de permitir assimetria nos retornos, seja com o uso da variável dummy $d_t = I(r_t < 0)$ e/ou o retorno defasado r_{t-1} , como no modelo APARCH de Ding et al. (1993).

$$h_t^r = (\omega_r + \alpha_r r_{t-1}^2 + \beta_r h_{t-1}^r) + \gamma_r r_{t-1}^2 d_{t-1} + \delta_r r_{t-1} + \varphi_r h_{t-1}^2 + \vartheta_r h_{t-1}^2 d_{t-1} + \psi_r v_{t-1}^2 + \lambda_r v_{t-1}^2 d_{t-1}$$

$$h_t^h = (\omega_h + \alpha_h h_{t-1}^2 + \beta_h h_{t-1}^h) + \delta_h r_{t-1} + \gamma_h h_{t-1}^2 d_{t-1} + \varphi_h r_{t-1}^2 + \vartheta_h r_{t-1}^2 d_{t-1} + \psi_h v_{t-1}^2 + \lambda_h v_{t-1}^2 d_{t-1}$$

$$h_t^v = (\omega_v + \alpha_v v_{t-1}^2 + \beta_v h_{t-1}^v) + \delta_v r_{t-1} + \gamma_v v_{t-1}^2 d_{t-1} + \varphi_v r_{t-1}^2 + \vartheta_v r_{t-1}^2 d_{t-1} + \psi_v h_{t-1}^2 + \lambda_v h_{t-1}^2 d_{t-1}$$

Pelas equações acima, podemos observar que até seis variáveis predeterminadas são introduzidas na especificação $GARCH(1,1)$. Este último modelo é deixado como base para a posterior comparação com os modelos aumentados pelas variáveis predeterminadas. O total de modelos a serem estimados é de $2^6 = 64$. O melhor modelo foi escolhido conforme o menor critério de Schwartz (BIC).

Com posse dos melhores modelos, são feitas previsões para 1 dia, 2 dias, ..., 20 dias a frente, considerando a lei de formação em que cada variável predeterminada é mantida constante ao longo do período de previsão (*náive*). As previsões são somadas a fim de se obter a previsão de um mês a frente, para fins de comparação com o horizonte temporal do VIX.

Os resultados encontrados Engle e Gallo (2006) indicam que o modelo $GARCH(1,1)$ aumentado é superior ao $GARCH(1,1)$ não aumentado, apresentando uma maior aderência ao VIX. Mesmo assim, o modelo apresenta correlação nos resíduos, um sinal para os autores de que o VIX realmente contém informações adicionais não captados pelos modelos GARCH. Tem-se então

$$VIX_t = VIX_t^{MST} + VIX_t^*$$

ou seja, o VIX contém dois componentes: um que pode ser captado por modelos de séries temporais (MST) e outro componente próprio, atribuído a precificação dos operadores de opções. Ambos os componentes são ortogonais entre si.

Uma forma equivalente desta equação é

$$VIX_t = \gamma_0 + \gamma_1 \omega_t + \varepsilon_t \tag{10}$$

com $VIX_t^{MST} = \hat{\gamma}_0 + \hat{\gamma}_1 \omega_t$ e $VIX_t^* = \hat{\varepsilon}_t$

Portanto, o VIX apresentará mais informações do que modelos de séries temporais se VIX_t^* apresentar algum poder explicativo da volatilidade real.

Esta hipótese de informação adicional do VIX foi testada por Becker, Clements e White (2007) utilizando estimações por GMM. Os autores estimam 8 modelos de volatilidade e obtêm a previsão dos mesmos. Os modelos são: i) GARCH (GAR_t), ii) GARCH adicionado de volatilidade realizada como variável explicativa na equação da variância condicional (GAR_t^+), iii) volatilidade estocástica (SV_t), iv) volatilidade estocástica com a adição da volatilidade realizada como variável explicativa (SV_t^+), v) modelo ARMA (AR_t), vi) modelo ARFIMA (ARF_t), vii) MIDAS (MAR_t), viii) volatilidade realizada (RV_t).

$$\omega_t = (GAR_t, GAR_t^+, SV_t, SV_t^+, AR_t, ARF_t, MAR_t, RV_t)$$

Com posse das 8 previsões, é possível testar diretamente a ortogonalidade de VIX_t^* com a *proxy* da volatilidade real, no caso a volatilidade realizada, pelo método de momentos generalizados (GMM).

O estimador de GMM de $\gamma = (\gamma_0, \gamma_1)'$ em (10) minimiza $V = M'HM$, onde $M = T^{-1} (\varepsilon_t (\gamma)') Z_t$ é um vetor $k \times 1$ de condições de momentos, H é uma matriz de pesos $k \times k$ e Z_t é um vetor de instrumentos. Para minimizar a variância dos coeficientes, H é escolhida como sendo a matriz de variância-covariância da condição de momento k em M , permitindo-se correlação nos resíduos. Sempre que k for maior que a dimensão de γ , o teste de restrições sobre-identificadas $J = TM'HM$ é distribuído como uma χ^2 com k menos a dimensão de γ graus de liberdade sob a hipótese nula de que os resíduos em (10) não apresentam correlação com os elementos de Z_t . Neste vetor de instrumentos estão contidas as variáveis de ω_t e RV_{t+20} , a real volatilidade futura para diversos períodos, dada pela variância realizada calculada com dados intradiários

$$Z_t = (\omega_t', RV_{t+20}')'$$

Chernov (2002) afirma que a volatilidade implícita reflete a volatilidade do prêmio de risco. Isto significa que o VIX pode ter um viés positivo em relação a futura volatilidade. Por isso, faz-se necessário a inclusão de uma constante em ω_t . Além disso, o autor sugere que o prêmio de risco pode variar e ser correlacionado com o nível corrente da volatilidade. Para evitar este tipo de problema, uma *proxy* da volatilidade corrente, a volatilidade realizada, é adicionada ao vetor de instrumentos.

Já o vetor RV_{t+20} contém os valores realizados da volatilidade no período de até 20 dias após a data t . A inclusão de mais dias, e não somente o valor do vigésimo dia, foi utilizada para testar se o VIX possui algum poder explicativo também para horizontes mais curtos. Assim, RV_{t+20} é definido como

$$RV_{t+20} = \{ \overline{RV}_{t+1}, \overline{RV}_{t+5}, \overline{RV}_{t+10}, \overline{RV}_{t+15}, \overline{RV}_{t+20}, \}$$

onde \overline{RV}_{t+j} é a média dos dias $t+1$ a $t+j$.

Basta então regressir o VIX sobre as Z_t . Caso o teste de sobre-identificação apresente um p-valor alto para a estatística J , conclui-se que o componente não explicado é ortogonal ao valor real da variância realizada, ou seja, o VIX não contém nenhuma informação adicional que os demais modelos na previsão da volatilidade.

Utilizando-se dos 8 modelos citados, os autores chegam a conclusão de que o VIX não contém maiores informações do que os modelos tradicionais já proporcionam. O mesmo teste será realizado neste artigo, mas levando em conta as previsões dos 3 modelos propostos por Engle e Gallo (2006) e a soma das previsões, não mais a média.

6 Resultados

A figura 3 e a tabela 2 de estatísticas descritivas apresentam as séries utilizadas no modelo, tanto para o período sem crise quanto para o de crise. As séries foram transformadas para apresentar a mesma média que a série do S&P500.

Os modelos escolhidos para o período sem crise conforme o critério de informação de Schwarz são:

$$\begin{aligned} h_t^r &= (\omega_r + \alpha_r r_{t-1}^2 + \beta_r h_{t-1}^r) + \psi_r v_{t-1}^2 \\ h_t^h &= (\omega_h + \alpha_h h_{t-1}^2 + \beta_h h_{t-1}^h) + \lambda_h v_{t-1}^2 d_{t-1} \end{aligned}$$

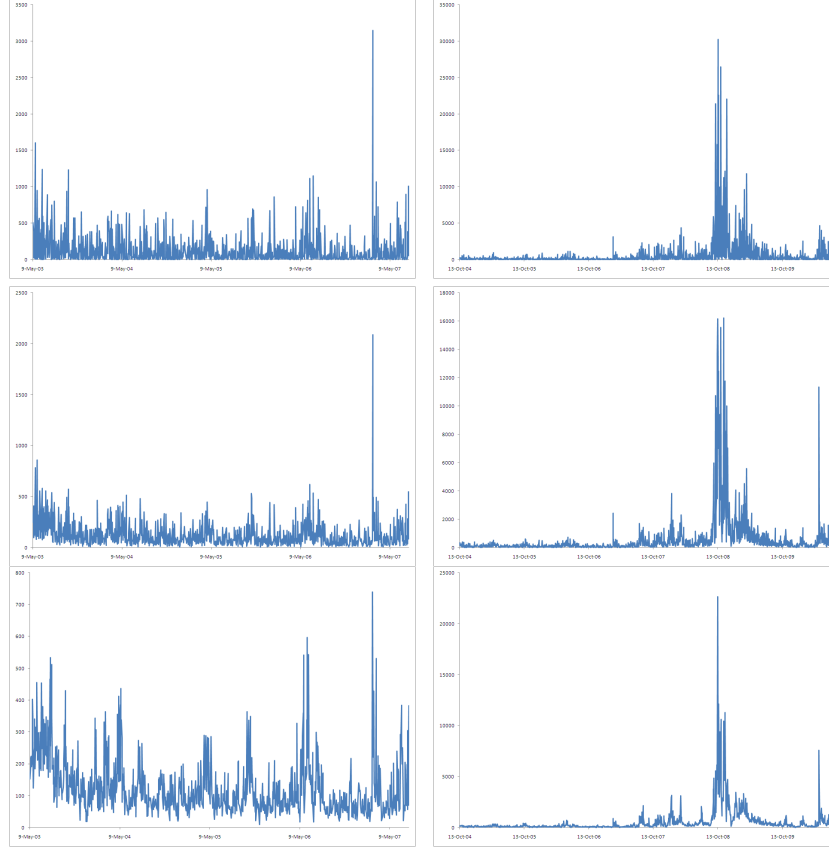


Figura 3: Retorno Absoluto, Intervalo entre Máximo e Mínimo e Volatilidade Realizada - Período sem crise (coluna da esquerda) e de crise (coluna da direita)

$$h_t^v = (\omega_v + \alpha_v v_{t-1}^2 + \beta_v h_{t-1}^v) + \vartheta_v r_{t-1}^2 d_{t-1}$$

enquanto que para o período de crise são:

$$h_t^r = (\omega_r + \alpha_r r_{t-1}^2 + \beta_r h_{t-1}^r) + \gamma_r r_{t-1}^2 d_{t-1} + \delta_r r_{t-1} + \psi_r v_{t-1}^2$$

$$h_t^h = (\omega_h + \alpha_h h_{t-1}^2 + \beta_h h_{t-1}^h) + \delta_h r_{t-1} + \varphi_h r_{t-1}^2 + \vartheta_h r_{t-1}^2 d_{t-1}$$

$$h_t^v = (\omega_v + \alpha_v v_{t-1}^2 + \beta_v h_{t-1}^v) + \vartheta_v r_{t-1}^2 d_{t-1}$$

Para fins de comparação com o VIX, que representa opções com um mês de vida, foi feita a previsão até 20 dias à frente para cada equação, período relativo a um mês de negociações, tanto para o modelo aumentado quanto para o modelo GARCH não aumentado, chamado de modelo base. Feita a previsão para os dias selecionados, a amostra é deslocada um período a frente, mantendo-se

o tamanho de 700 observações, sendo obtidas novas previsões. O procedimento é repetido então até o final da amostra, 360 vezes para o período sem crise e 760 para o período de crise. As previsões e o VIX são mostrados na figura 4.

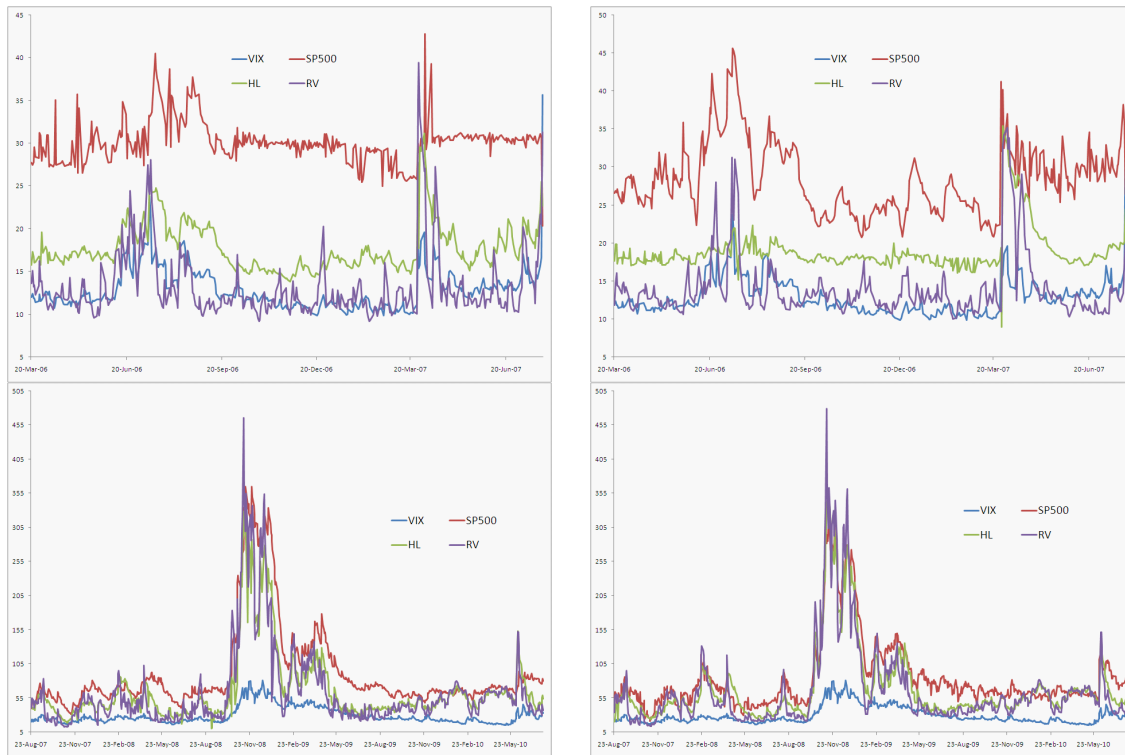


Figura 4: VIX e Previsões - Primeira linha: modelos aumentados a esquerda e modelos base a direita para o período sem crise. Segunda linha: modelos aumentados a esquerda e modelos base a direita para o período de crise.

Para testar a qualidade das previsões tendo como referência o VIX, realiza-se uma regressão do VIX em um $AR(1)$, na constante e nas previsões de volatilidade. A tabela 3 mostra os resultados para o período sem crise

A afirmação de Engle e Gallo (2006) de que as previsões dos modelos GARCH aumentados com as variáveis predeterminadas apresentam maior aderência ao VIX do que os modelos base é confirmada no período sem crise. A constante e o termo autorregressivo são significantes em todas as regressões. O BIC da primeira regressão é de 2,868, o R^2 é de 0.808 e o termo $AR(1)$ ultrapassa os 0.9. Ao serem adicionadas as previsões dos modelos aumentados, há um ganho tanto no R^2 , que chega a 0.903, quanto no BIC, que é de 2.235. O coeficiente autorregressivo cai para 0.817 e das três previsões, apenas $SP500_a$ não é significativa a 5%. Se forem adicionadas previsões obtidas de modelos base, há uma piora no critério de informação, que fica em 2.906, sinalizando não só que a

escolha do melhor modelo entre os 64 modelos possíveis com a adição das variáveis predeterminadas contribuiu para uma melhor aderência ao VIX, como também que o $AR(1)$ é superior aos modelos base, devido ao fato das três previsões não serem significativas a 5%. A última linha apresenta a inclusão de todas as variáveis no modelo, e as previsões do SP500, seja no modelo aumentado ou no base, não são significativas.

Observa-se que os resíduos apresentam correlação serial em todas as regressões, como verificado pelo teste LM de Breusch-Godfrey com quatro defasagens. Para os autores, esta auto-correlação é um indício de que o VIX apresenta um componente extra que não pode ser capturado por modelos tradicionais de volatilidade, fato que será discutido mais a frente.

Nos dados para o período de crise, mostrados na tabela 4, as constantes e os coeficientes autorregressivos continuam significativos, e os últimos apresentam um aumento frente ao período sem crise, chegando até 0.974 na primeira e na última regressão. As três previsões dos modelos aumentados são significativas a 5%, mas o ganho no R^2 e no BIC são mínimos frente a primeira regressão. Já em relação a terceira regressão, mesmo as previsões não sendo significativas, o poder explicativo é o mesmo da primeira regressão e apenas um pouco inferior a segunda regressão. Na última regressão, apenas a previsão obtida do modelo do $SP500_b$ não aumentado não é significativa. Pode-se afirmar então que a escolha do melhor modelo aumentado não gera ganhos tão grandes no período de crise como no período sem crise.

Também foi analisada a aderência de todas as previsões, incluindo o VIX, frente a variância realizada acumulada durante 20 dias, uma *proxy* da verdadeira volatilidade ocorrida em cada período, utilizando o método de mínimos quadrados ordinários, regredindo a variância realizada acumulada em 20 dias sobre cada variável de forma individual e com a adição do VIX como segunda variável explicativa, além da constante.

Como apontado na tabela 5, no período sem crise, todas as quatro variáveis e as constantes são positivas e significativas a 5% nas regressões com apenas uma variável explicativa, e o VIX apresenta apenas o terceiro maior poder explicativo, com um R^2 de 0.161. A partir da quinta regressão, o VIX é adicionado como segunda variável explicativa. Sua adição aumenta o poder explicativo de $SP500_a$ em 2,5% (0,144 para 0,169 de R^2), de HL_a em 1,2% e RV_a em 2,8%. O VIX deixa de ser significativo quando adicionado a HL_a e RV_a . Na adição ao $SP500_a$, o VIX continua significativo, a contrário do próprio $SP500_a$. A parte final da tabela apresenta a matriz de correlação das quatro variáveis. Como há uma grande correlação entre elas, $SP500_a$, HL_a e VIX acabam perdendo significância na regressão completa, não sendo possível a análise da mesma. A última linha da primeira parte da tabela mostra que apenas RV_a é significativa, e apesar do R^2 ser o maior de todos os modelos, o BIC é apenas o terceiro menor.

Nos modelos base, apresentados na tabela 6, apenas o VIX e RV_b são significativos nas regressões com apenas uma variável explicativa, e o R^2 e o BIC do modelo com o VIX apresentam grande superioridade aos demais. Todas as constantes são significativas. A adição do VIX como segunda explicativa aumenta o poder explicativo das três previsões em mais de 10 pontos percentuais. Além disso, o VIX concentra todo o poder explicativo, dado que as previsões não são significantes. Conclui-se então que aumentar os modelos no período sem crise gera um ganho significativo no poder explicativo frente a volatilidade realizada. É possível observar também que a correlação entre as variáveis apresenta grande redução, não sendo superior a 66% entre HL_b e RV_b . Apenas o VIX e $SP500_b$ são significativos, e o segundo apresenta sinal negativo. O R^2 é o maior de todos e o BIC é semelhante aos modelos com a presença do VIX.

A tabela 7 apresenta os mesmos dados da tabela 5, mas agora para o período de crise. De forma individual, é possível observar que o VIX não apresenta grandes diferenças frente as outras

três previsões, apresentando um R^2 maior e um BIC menor apenas em relação a HL_a . Já a adição como segunda variável explicativa apresenta ganhos significativos de R^2 em HL_a e RV_a e ganhos mais modestos em $SP500_a$. As constantes nas regressões com apenas o VIX como explicativa e do VIX com $SP500_a$ e HL_a não são significativas, indicando que não há a presença de viés nestes modelos. A correlação agora é muito alta, chegando a atingir 95% entre $SP500_a$ e HL_a , sinalizando problemas de multicolinearidade, possível causador do sinal negativo em HL_a . Apenas $SP500_a$ não é significativa, e o R^2 chega a 0.618, o maior dentre todos. O BIC também sinaliza que o modelo é o melhor na tabela.

A tabela 8 para o período de crise com as previsões dos modelos base confirmam os dados apresentados nas tabelas 3 e 4, pois não há novamente uma melhora significativa dos modelos base para os aumentados. Os R^2 de $SP500_a$, HL_a e RV_a são de 0.550, 0.481 e 0.544 na tabela 7, enquanto os R^2 de $SP500_b$, HL_b e RV_b são de 0.559, 0.524 e 0.515 na tabela 8, respectivamente, mostrando que apenas RV apresentou melhora ao aumentar o modelo. A correlação entre HL_b e RV_b ultrapassa os 95%, mais uma vez apontando multicolinearidade, fazendo com que a análise individual das variáveis seja mais relevante. Todas as variáveis são significativas e os critérios de informação desta regressão são os melhores da tabela. Novamente HL_a apresenta sinal negativo, causado pelo problema de multicolinearidade.

Pela metodologia de Becker, Clements e White (2007), a afirmação de Engle e Gallo (2006) de que o resíduo apresentado nas tabelas 3 e 4 contém informações relevantes que os operadores são capazes de embutir nos preços dos ativos é testada na tabela 9 para os modelos aumentados no período sem crise

Na primeira equação, como não há variáveis instrumentais, os estimadores são os mesmos que uma regressão por mínimos quadrados ordinários pois o sistema é exatamente identificado, não apresentando, portanto, a estatística J . Este modelo foi colocado na tabela para servir como referência na comparação com os demais. As variáveis explicativas no modelo são as previsões obtidas nos modelos GARCH aumentados e o valor real da variância realizada, além da constante.

Para a estimação da matriz de pesos H , foi utilizado o algoritmo de Andrews e Monahan. Como teste de robustez, a matriz H também foi calculada utilizando-se do estimador de Newey-West com a largura da banda de 20 e 40, além da escolha automática definida conforme a amostra. Assim como o encontrado em Becker, Clements e White (2007) para os 8 modelos, os 3 modelos de Engle e Gallo (2006) conseguem captar as mesmas informações sobre a volatilidade que o VIX é capaz. O p-valor da estatística J , apresentado na última coluna, indica que não é possível rejeitar a hipótese de que o VIX não contém informação relevante frente as previsões dos 3 modelos, a um nível de significância de 5%. Na última regressão, a estatística J é de 0.062, indicando que o VIX pode ter alguma informação a mais do que os modelos de série de tempo no curto prazo, ainda que apenas marginalmente. Exceto pelas constantes e por RV, todas as variáveis são significantes e positivas.

Na tabela 10 são apresentadas as estatísticas J nas quatro formas estimadas da matriz H . O resultado é mantido, independentemente do critério de estimação.

Na tabela 11 foram utilizadas as previsões obtidas através dos modelos base, ou seja, um $GARCH(1,1)$ sem variáveis explicativas predeterminadas. Observa-se que a hipótese nula é rejeitada a 5% para todos os períodos utilizados como instrumentos, ou seja, deixando-se de melhorar o modelo com a inclusão das variáveis explicativas, perde-se informação relevante sobre a volatilidade futura, comprovando a importância citada por Engle e Gallo (2006) de aumentar o modelo GARCH. A previsão HL_b é a única variável não significativa, além da constante.

Mas, ao realizar a estimação da matriz H com outro método, o resultado não se mantém, como observado na tabela 12. A estatística J aumenta da esquerda para a direita, ou seja, conforme o

método de estimação, as previsões dos modelos base podem ou não conter todas as informações relevantes que o VIX contém. Mesmo assim, é possível observar que o p-valor é reduzido conforme a redução do horizonte de previsão, constatado pela queda do valor ao longo das colunas, sinalizando que o VIX pode conter informações relevantes no curto prazo.

As estimativas para os modelos aumentados para o período de crise são apresentadas na tabela 13. Observa-se que, durante a crise, o VIX apresenta informações frente a real volatilidade que os modelos utilizados não são capazes de capturar, dado o baixo p-valor apresentado em todos os períodos. HL_a não é significativo em nenhum dos horizontes, enquanto que $SP500_a$ não o é apenas na regressão com RV_{t+20} como instrumento. Os coeficientes de RV_a são negativos e estão mais próximas de zero do que os coeficientes das demais variáveis.

Assim como na análise anterior dos modelos base para o período sem crise, o método de estimação de H altera os resultados, como mostrado na tabela 14. Para a estimação com todos os instrumentos e para o horizonte de 20 dias, a estatística J ultrapassa os 5% ao utilizar o método de Newey-West. Já para horizontes inferiores, o VIX apresenta informações a mais, independente do método de estimação.

Os dados para o período de crise com previsões dos modelos base são apresentados na tabela 15. Nas regressões com RV_{t+20} , RV_{t+15} e RV_{t+10} como instrumentos individuais não é possível rejeitar a hipótese de que o VIX não contém mais informações do que os modelos GARCH, dado os p-valores de 0.836, 0.739 e 0.318 da estatística J . Já para RV_{t+5} e RV_{t+1} , o VIX contém estas informações, fazendo com que o p-valor da regressão com todos os instrumentos também seja baixo. RV_b não é significativo em nenhum momento, enquanto HL_b não apresenta significância com todos os instrumentos e com apenas RV_{t+20} como instrumento. A variável também é a única negativa na tabela, para qualquer que seja o horizonte.

A tabela 16 registra os testes de robustez para os modelos base durante a crise. Observa-se que a informação extra que o VIX tem a mais que as previsões não é mantida ao se mudar o método de estimação. O padrão de redução da estatística J com a diminuição do horizonte permanece.

Conclui-se então que aumentar os modelos GARCH com a adição das variáveis predeterminadas apresenta resultados positivos apenas no período sem crise, pois os modelos aumentados apresentaram uma melhor aderência ao VIX, assim como o encontrado por Engle e Gallo (2006). Já no período de crise, nem os modelos aumentados quanto os base conseguiram captar toda a informação relevante que o VIX contém, pelo menos para curtos horizontes de tempo.

7 Conclusão

Este artigo buscou testar se a volatilidade implícita é a melhor previsora da volatilidade futura em comparação a modelos de séries temporais durante o período da crise imobiliária americana. Para realizar a comparação foi necessária a definição de quais modelos seriam utilizados dentro de cada grupo. Para a volatilidade implícita, o índice de volatilidade VIX foi escolhido por mitigar diversas hipóteses necessárias na formulação do tradicional modelo de volatilidade implícita de Black-Scholes.

O VIX foi testado como uma possível variável preditiva da volatilidade futura. Seguindo Engle e Gallo (2006), foram criados modelos GARCH aplicados a variáveis não-negativas, com a permissão de assimetria e adição de variáveis explicativas na variância condicional. Assim como encontrado pelos autores, estes modelos apresentam maior poder preditivo frente a simples modelos GARCH sem assimetria ou não adição de variáveis explicativas, mas apenas para o período sem crise. No período de crise, não há ganhos no aumento dos modelos GARCH. Os mesmos autores afirmam

que, por conter as visões dos operadores de opções sobre eventos futuros, a volatilidade implícita é o melhor previsor da volatilidade. Esta afirmação foi analisada segundo a metodologia de Becker, Clements e White (2007), onde foi testado se o componente do VIX não explicado pelos modelos GARCH, sejam aumentados ou não, apresenta algum poder preditivo. Encontram-se evidências de que isto é verdade apenas para curtos horizontes de tempo durante o período de crise.

Tem-se então que a conclusão sobre a qualidade das previsões pode variar significativamente conforme o período de análise. No período onde a volatilidade é baixa, os resultados encontrados corroboram as opiniões dos autores citados, fato não obtido de meados de 2007 a meados de 2010.

8 Referência Bibliográfica

AKGIRAY, Vedat. *Conditional Heteroskedasticity in Time Series of Stock Returns: Evidence and Forecasts*. **Journal of Business**. Vol. 62, p. 55-80, 1989.

ANDERSEN, Torben G.; BOLLERSLEV, Tim. *Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts*. **International Economic Review**. Vol. 39, p. 885-905, 1998.

BAILLIE, Richard T.; BOLLERSLEV, Tim; MIKKELSEN, Hans O. *Fractionally integrated generalized autoregressive conditional heteroskedasticity*. **Journal of Econometrics**. Vol. 74, n.1, p. 3-30, 1996.

BECKER, Ralf; CLEMENTS, Adam E.; WHITE, Scott I. *Does implied volatility provide any information beyond that captured in model-based volatility forecasts?*. **Journal of Banking and Finance**. Vol. 31, p. 2535-2549, 2007.

BIAIS, Bruno R.; GLOSTEN, Larry; SPATT, Chester. *Market microstructure: A survey of microfoundations, empirical results, and policy implications*. **Journal of Financial Markets**. Vol. 8, p. 217-264, 2005.

BLACK, Fischer; SCHOLES, Myron. *The Pricing of Options and Corporate Liabilities*. **Journal of Political Economy**. Vol. 81, n.3, p. 637-654, 1973.

BLAIR, Bevan J.; POON, Ser-Huang; TAYLOR, Stephen. *J. Forecasting S&P 100 volatility: The incremental information content of implied volatilities and high frequency index returns*. **Journal of Econometrics**. Vol. 105, p. 5-26, 2001.

BOLLERSLEV, Tim. *Generalized Autoregressive Conditional Heteroskedasticity*. **Journal of Econometrics**. Vol. 31, p. 307-327, 1986.

BOLLERSLEV, Tim; MIKKELSEN, Hans O. *Modeling and pricing long memory in stock market volatility*. **Journal of Econometrics**. Vol. 73, n.1, p. 151-184, 1996.

BOLLERSLEV, Tim; WOOLDRIDGE, Jeffrey M. *Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances*. **Econometric Reviews**. Vol. 11, p. 143-172, 1992.

CANINA, Linda; FIGLEWSKI, Stephen. *The informational content of implied volatility*. **Review of Financial Studies**. Vol. 6, n.3, p. 659-681, 1993.

CHERNOV, Mikhail. *On the role of volatility risk premia in implied volatilities based forecasting regressions*. **Columbia Business School**. 2002.

CHRISTENSEN, Bent J.; PRABHALA, N. R. *The relation between implied and realized volatility*. **Journal of Financial Economics**. Vol. 50, n.2, p. 125-150, 1998.

CIPOLLINI, Fabrizio; ENGLE, Robert F.; GALLO, Giampiero M. *Vector Multiplicative Error Models: Representation and Inference*. NBER Working Papers 12690, **National Bureau of Economic Research, Inc.** 2006.

- DAVIDIAN, M.; CARROLL, R.J. *Variance function estimation*. **Journal of American Statistical Association**. Vol. 82, p. 1079-1091, 1987.
- DING, Zhuangxin; GRANGER, Clive W. J.; ENGLE, Robert. F. *A long memory property of stock market returns and a new model*. **Journal of Empirical Finance**. Vol. 1, n.1, p. 83-106, 1993.
- ENGLE, Robert. F. *Autoregressive Conditional Heteroscedasticity with Estimates of Variance of United Kingdom Inflation*. **Econometrica**. Vol. 50, p. 987-1008, 1982.
- ENGLE, Robert. F. *New frontiers for ARCH models*. **Journal of Applied Econometrics**. Vol. 17, p. 425-446, 2002.
- ENGLE, Robert. F.; GALLO, Giampiero M. *A multiple indicators model for volatility using intra-daily data*. **Journal of Econometrics**. Elsevier, Vol. 131, n.1-2, p. 3-27, 2006.
- ENGLE, Robert. F.; RUSSELL, Jeffrey R. *Autoregressive conditional duration: A new model for irregularly spaced transaction data*. **Econometrica**, Vol. 66, n.5, p. 1127-1162, 1998.
- FLEMING, Jeff; OSTDIEK, Barbara; WHALEY, Robert.E. *Predicting stock market volatility: A new measure*. **Journal of Futures Market**. Vol. 15, n.3, p. 265-302, 1995.
- GARMAN, Mark B.; KLASS, Michael J. *On the estimation of security price volatilities from historical data*. **Journal of Business**. Vol. 53, p. 67-78, 1980.
- GLOSTEN, Lawrence R.; JAGANNATHAN, Ravi; RUNKLE, David E. *On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks*. **Journal of Finance**. American Finance Association, Vol. 48, n.5, p. 1779-1801, 1993.
- GORTON, Gary B. *The Panic of 2007*. **Yale ICF Working Paper No. 08-24**, 2008. Disponível em: <<http://ssrn.com/abstract=1255362>>. Acesso em: Fevereiro de 2011.
- HOL, Eugenie; KOOPMAN, Siem J. *Forecasting the variability of stock index returns with stochastic volatility models and implied volatility*. **Working Paper, Free University Amsterdam**, 2001.
- HULL, John C. **Options, Futures and Other Derivatives**. Prentice Hall, 7^a ed., 744p, 2008.
- HULL, John C.; WHITE, Alan. *The pricing of options on assets with stochastic volatilities*. **Journal of Finance**. Vol. 42, n.2, p. 281-300, 1987.
- ISSLER, João V. *Estimating and Forecasting the Volatility of Brazilian Finance Series Using ARCH Models*. **Revista de Econometria**. Número Especial, 1999.
- ITÔ, Kiyoshi. *On stochastic differential equations*. Memoirs, **American Mathematical Society**. Vol. 4, p. 1-51, 1951.
- LEHAR, Alfred; SCHEICHER, Martin; SCHITTENKOPF, Christian. *GARCH vs. stochastic volatility: Option pricing and risk management*. **Journal of Banking & Finance**, Vol. 26, n.2-3, p. 323-345, 2002.
- MADHAVAN, Ananth. *Market microstructure: a survey*. **Journal of Financial Markets**. Vol. 3, p. 205-258, 2000.
- MARTENS, Martin; ZEIN, Jason. *Predicting Financial Volatility: Highfrequency Time-Series Forecast vis-à-vis Implied Volatility*. **Journal of Futures Markets**. Vol. 24, n.11, p. 1005-1029, 2004.
- McALEER, Michael; MEDEIROS, Marcelo C. (2008). *Realized Volatility: A Review*. **Econometric Reviews**, Vol. 27, n.1, p. 10-45, 2008.
- NAKAJIMA, Jouchi. *EGARCH and Stochastic Volatility: Modeling Jumps and Heavy-tails for Stock Returns*. **IMES Discussion Paper Series 08-E-23**. Institute for Monetary and Economic Studies, Bank of Japan, 2008.

NELSON, Daniel B. *Conditional heteroskedasticity in asset returns: A new approach*. **Econometrica**. Vol. 59, p. 347-370, 1991.

POON, Ser-Huang; GRANGER, Clive W. J. *Forecasting volatility in financial markets: a review*. **Journal of Economic Literature**. Vol. 41, n.2, p. 478-539, 2003.

SZAKMARY, Andrew; ORS, Evren; KIM, Jin K.; DAVIDSON III, Wallace N. *The predictive power of implied volatility: Evidence from 35 futures markets*. **Journal of Banking and Finance**. Vol. 27, p. 2151-2175, 2003.

TAYLOR, Stephen. **Modelling financial time series**. Wiley, 1986.

VALLS PEREIRA, Pedro L.; HOTTA, Luiz K.; SOUZA, Luiz A. R.; ALMEIDA, Nuno M. C. G. *Modelos Alternativos para Extração da Volatilidade de Ativos: um Estudo Comparativo*. **Revista de Econometria**. Número Especial, 1999.

(Chapter head:)Apêndice

9 Distribuição Gama

A função de densidade de probabilidade de uma variável aleatória que segue uma distribuição Gama é dada por

$$f(x) = \frac{\beta^\phi}{\Gamma(\phi)} x^{\phi-1} e^{-\beta x}$$

para $x > 0$ ou 0 caso contrário e $\phi, \beta > 0$. Os dois primeiros momentos desta distribuição são dados por

$$\begin{aligned} E(X) &= \frac{\phi}{\beta} \\ V(X) &= \frac{\phi}{\beta^2} \end{aligned}$$

Esta distribuição apresenta a seguinte propriedade e os seguintes casos particulares

$$X \sim Gama(\phi, \beta) \text{ e } c > 0 \Rightarrow Y = cX \sim Gama\left(\phi, \frac{\beta}{c}\right)$$

$$Exponencial(\beta) = Gama\left(\frac{1}{\beta}\right)$$

$$\chi_{(n)}^2 = Gama\left(\frac{n}{2}, \frac{1}{2}\right)$$

10 Distribuição GED

A distribuição de erro generalizado (GED) com os parâmetros (μ, σ, ϕ) apresenta a seguinte função de distribuição de probabilidade

$$f(x) = \frac{\phi^{\phi-1}}{2\Gamma(\phi)\sigma} \exp\left(-\phi \left|\frac{x-\mu}{\sigma}\right|^{1/\phi}\right)$$

para $x \in \mathbb{R}$, onde $\mu \in \mathbb{R}, \sigma, \phi > 0$. Seus dois primeiros momentos são dados por

$$\begin{aligned} E(X) &= \mu \\ V(X) &= \frac{\Gamma(3\phi)}{\phi^{2\phi}\Gamma(\phi)}\sigma^2 \end{aligned}$$

Uma propriedade útil da distribuição é

$$X \sim GED(\mu, \sigma, \phi) \text{ e } c_1, c_2 \in \mathbb{R} \Rightarrow Y = c_1 + c_2X \sim GED(c_1 + c_2\mu, |c_2|\sigma, \phi)$$

Ela também apresenta os seguintes casos particulares

$$\begin{aligned} Laplace(\mu, \sigma) &= GED(\mu, \sigma, 1) \\ N(\mu, \sigma) &= GED\left(\mu, \sigma, \frac{1}{2}\right) \end{aligned}$$

11 Relação entre Gama e GED

Ambas distribuições são ligadas pela seguinte relação

$$X \sim GED(0, \sigma, \phi) \Rightarrow Y = \left| \frac{X - \mu}{\sigma} \right|^{1/\phi} \sim Gama(\phi, \phi)$$

e alguns casos particulares são

$$\begin{aligned} X \sim GED(0, 1, \phi) &\Rightarrow Y = |X|^{1/\phi} \sim Gama(\phi, \phi) \\ X \sim Laplace(0, 1) &\Rightarrow Y = |X| \sim Exponencial(1) \\ X \sim N(0, 1) &\Rightarrow Y = X^2 \sim \chi^2(1) \end{aligned}$$

Tabela 1: Estatísticas Descritivas

Estatística	Período sem Crise			Período de Crise			Amostra Completa		
	S&P500	VIX	RVol	S&P500	VIX	RVol	S&P500	VIX	RVol
Média	1218.71	14.46	5.49E-03	1149.70	29.12	1.21E-02	1189.89	20.58	8.24E-03
Mediana	1203.30	13.97	5.13E-03	1114.45	24.97	9.87E-03	1187.73	17.55	6.40E-03
Máximo	1553.08	23.81	1.41E-02	1565.15	80.86	7.13E-02	1565.15	80.86	7.13E-02
Mínimo	919.73	9.89	1.66E-03	676.53	15.58	2.53E-03	676.53	9.89	1.66E-03
Desvio Padrão	142.60	2.86	1.78E-03	221.22	11.97	7.89E-03	182.80	10.81	6.20E-03
Assimetria	0.36	0.60	1.00	0.06	1.76	2.50	-0.10	2.25	3.41
Curtose	2.58	2.54	4.26	1.94	5.93	11.95	2.55	9.09	20.11
Jarque Bera	30.05	73.89	247.84	35.99	665.94	3327.12	18.39	4344.90	25722.62
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
No. Obs.	1060	1060	1060	760	760	760	1820	1820	1820

Obs: Valores entre colchetes representam o p-valor da estatística de Jarque-Bera. O período sem crise vai de 9 de maio de 2003 a 25 de julho de 2007, enquanto que o período de crise inicia-se em 26 de julho de 2007 e segue até o final da amostra, em 30 de julho de 2010.

Tabela 2: Estatísticas Descritivas - Modelo de Engle e Gallo

Estatística	Período sem Crise			Período de Crise		
	S&P500	HL	RV	S&P500	HL	RV
Média	124.51	124.51	124.51	544.91	544.91	544.91
Mediana	42.27	88.03	98.39	79.77	171.10	192.37
Máximo	3147.75	2090.30	739.31	30255.16	16215.23	22666.11
Mínimo	0.00	6.73	10.32	0.00	7.80	12.33
Desvio Padrão	209.91	126.16	86.92	1890.86	1410.55	1262.48
Assimetria	4.82	4.88	2.04	9.08	6.83	7.61
Curtose	49.24	60.54	9.09	109.08	59.30	91.10
Jarque Bera	98542.02	150439.90	2372.71	704639.90	204209.40	486231.30
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
No. Obs.	1060	1060	1060	1460	1460	1460

Obs: Valores entre colchetes representam o p-valor das estatísticas Jarque-Bera. O período sem crise vai de 9 de maio de 2003 a 25 de julho de 2007, enquanto que o período de crise vai 13 de outubro de 2004 a 30 de julho de 2010.

Tabela 3: Período sem crise - Aderência ao VIX

Variável Dependente: VIX										
Constante	$AR(1)$	$SP500_a$	HL_a	RV_a	$SP500_b$	HL_b	RV_b	R^2	BIC	Teste LM
13.139 (22.31)	0.908 (37.78)							0.808	2.868	3.40E-03
3.965 (6.47)	0.817 (24.59)	0.008 (0.36)	0.390 (11.18)	0.136 (8.11)				0.903	2.235	3.04E-06
11.067 (9.64)	0.900 (34.71)				0.032 (1.29)	0.059 (1.87)	-0.003 (-0.11)	0.811	2.906	1.74E-04
6.094 (6.98)	0.863 (29.24)	-0.012 (-0.63)	0.456 (10.82)	0.167 (7.59)	0.015 (0.88)	-0.123 (-4.43)	-0.090 (-3.72)	0.915	2.160	6.42E-03

Obs.: Regressões por MQO. O subscrito 'a' refere-se às previsões dos modelos aumentados, enquanto que 'b' refere-se aos modelos base.

O Teste LM para correlação serial foi realizado com 4 defasagens e o valor apresentado é o p-valor. A estatística t-robusta encontra-se entre parênteses.

Tabela 4: Período de crise - Aderência ao VIX

Variável Dependente: VIX										
Constante	$AR(1)$	$SP500_a$	HL_a	RV_a	$SP500_b$	HL_b	RV_b	R^2	BIC	Teste LM
29.998 (7.84)	0.974 (118.00)							0.950	4.855	1.58E-08
20.859 (11.72)	0.936 (68.17)	0.063 (5.08)	0.030 (3.32)	0.022 (3.96)				0.955	4.769	4.23E-14
29.245 (7.53)	0.972 (108.69)				0.012 (0.57)	-0.017 (-0.91)	0.012 (1.35)	0.950	4.881	2.44E-11
27.241 (7.44)	0.974 (114.34)	0.067 (5.85)	0.112 (9.78)	0.062 (7.85)	-0.031 (-1.72)	-0.122 (-6.12)	-0.045 (-4.30)	0.962	4.627	4.62E-14

Obs.: Regressões por MQO. O subscrito 'a' refere-se às previsões dos modelos aumentados, enquanto que 'b' refere-se aos modelos base.

O Teste LM para correlação serial foi realizado com 4 defasagens e o valor apresentado são os p-valores. A estatística t-robusta encontra-se entre parênteses.

Tabela 5: Período sem crise - Poder Preditivo dos Modelos Aumentados

Variável dependente: RV_{20d}						
Constante	$SP500_a$	HL_a	RV_a	VIX	R^2	BIC
22.791 (3.23)	0.803 (3.36)				0.144	7.321
20.173 (3.66)		1.431 (4.75)			0.165	7.296
32.138 (9.82)			1.019 (4.38)		0.171	7.288
22.844 (3.75)				1.764 (3.88)	0.161	7.300
20.793 (2.85)	0.318 (1.07)			1.222 (2.42)	0.169	7.308
19.391 (3.40)		0.820 (2.09)		0.901 (1.21)	0.177	7.298
24.311 (3.83)			0.646 (2.63)	0.987 (1.67)	0.199	7.271
20.813 (6.44)	0.243 (1.25)	0.439 (1.26)	0.608 (3.69)	0.158 (0.34)	0.211	7.290
<hr/> <hr/>						
$SP500_a$	1	0.805	0.547	0.823		
HL_a	0.805	1	0.644	0.845		
RV_a	0.547	0.644	1	0.674		
VIX	0.823	0.845	0.674	1		
<hr/> <hr/>						

Obs.: A estatística t-robusta encontra-se entre parênteses.

RV_{20d} é a variância realizada acumulada em 20 dias.

A correlação entre as variáveis é apresentada na parte inferior da tabela.

Tabela 6: Período sem crise - Poder Preditivo dos Modelos Base

Variável dependente: RV_{20d}						
Constante	$SP500_b$	HL_b	RV_b	VIX	R^2	BIC
38.039 (3.05)	0.253 (0.63)				0.004	7.472
38.317 (7.54)		0.390 (1.58)			0.011	7.464
36.889 (13.61)			0.625 (3.80)		0.065	7.408
22.844 (3.75)				1.764 (3.88)	0.161	7.300
40.137 (3.06)	-0.750 (-1.72)			2.175 (4.98)	0.187	7.285
26.682 (4.00)		-0.321 (-1.15)		1.935 (3.89)	0.168	7.309
22.637 (3.74)			0.134 (0.69)	1.634 (3.00)	0.163	7.314
43.616 (6.95)	-0.669 (-2.89)	-0.479 (-1.99)	0.281 (1.64)	2.111 (7.29)	0.198	7.308

$SP500_b$	1	0.294	0.208	0.503
HL_b	0.294	1	0.651	0.443
RV_b	0.208	0.651	1	0.539
VIX	0.503	0.443	0.539	1

Obs.: A estatística t-robusta encontra-se entre parênteses.

RV_{20d} é a variância realizada acumulada em 20 dias.

A correlação entre as variáveis é apresentada na parte inferior da tabela.

Tabela 7: Período de crise - Poder Preditivo dos Modelos Aumentados

Variável dependente: RV_{20d}						
Constante	$SP500_a$	HL_a	RV_a	VIX	R^2	BIC
39.432 (5.32)	0.961 (9.72)				0.550	10.466
59.592 (9.61)		0.918 (9.60)			0.481	10.609
67.211 (14.10)			0.847 (10.64)		0.544	10.481
1.018 (0.09)				4.055 (10.11)	0.536	10.497
19.112 (1.90)	0.574 (3.62)			1.798 (3.27)	0.567	10.438
14.283 (1.46)		0.352 (2.88)		2.815 (6.44)	0.557	10.461
24.973 (2.29)			0.486 (4.62)	2.208 (4.89)	0.604	10.348
23.092 (4.96)	0.258 (1.87)	-0.572 (-5.00)	0.726 (9.93)	2.306 (6.88)	0.618	10.330
<hr/>						
$SP500_a$	1	0.950	0.905	0.920		
HL_a	0.950	1	0.930	0.842		
RV_a	0.905	0.930	1	0.788		
VIX	0.920	0.842	0.788	1		

Obs.: A estatística t-robusta encontra-se entre parênteses.

RV_{20d} é a variância realizada acumulada em 20 dias.

A correlação entre as variáveis é apresentada na parte inferior da tabela.

Tabela 8: Período de crise - Poder Preditivo dos Modelos Base

Variável dependente: RV_{20d}						
Constante	$SP500_b$	HL_b	RV_b	VIX	R^2	BIC
47.823 (6.63)	0.783 (8.85)				0.559	10.446
57.635 (9.67)		0.907 (9.55)			0.524	10.522
67.954 (14.14)			0.808 (10.72)		0.515	10.542
1.018 (0.09)				4.055 (10.11)	0.536	10.497
23.505 (2.42)	0.485 (3.29)			1.765 (3.22)	0.580	10.407
18.757 (1.84)		0.475 (3.90)		2.340 (5.53)	0.584	10.398
21.099 (1.97)			0.424 (4.46)	2.443 (5.66)	0.593	10.376
29.728 (6.01)	0.414 (4.25)	-0.375 (-2.48)	0.502 (5.45)	1.553 (5.28)	0.604	10.368
$SP500_b$	1	0.939	0.863	0.892		
HL_b	0.939	1	0.952	0.817		
RV_b	0.863	0.952	1	0.773		
VIX	0.892	0.817	0.773	1		

Obs.: A estatística t-robusta encontra-se entre parênteses.

RV_{20d} é a variância realizada acumulada em 20 dias.

A correlação entre as variáveis é apresentada na parte inferior da tabela.

Tabela 9: Período sem crise - Condição de Ortogonalidade - Modelos Aumentados

Variável dependente: VIX					
Constante	$SP500_a$	HL_a	RV_a	RV	J
$z = \{c, \omega\}$					
0.460	0.181	0.320	0.109	0.018	NA
(0.40)	(3.11)	(3.25)	(2.21)	(0.29)	
$z = \{c, \omega, RV_{t+1}, \dots, RV_{t+20}\}$					
-0.029	0.165	0.370	0.111	0.012	0.251
(-0.02)	(2.85)	(3.95)	(2.24)	(0.21)	(5)
$z = \{c, \omega, RV_{t+20}\}$					
0.415	0.180	0.324	0.109	0.019	0.877
(0.37)	(3.06)	(3.39)	(2.18)	(0.29)	(1)
$z = \{c, \omega, RV_{t+15}\}$					
0.227	0.173	0.340	0.112	0.023	0.415
(0.20)	(3.01)	(3.51)	(2.16)	(0.36)	(1)
$z = \{c, \omega, RV_{t+10}\}$					
0.192	0.172	0.343	0.111	0.024	0.288
(0.17)	(3.02)	(3.52)	(2.16)	(0.40)	(1)
$z = \{c, \omega, RV_{t+5}\}$					
0.002	0.177	0.350	0.111	0.012	0.201
(0.00)	(3.06)	(3.66)	(2.25)	(0.20)	(1)
$z = \{c, \omega, RV_{t+1}\}$					
-0.221	0.163	0.390	0.115	-0.002	0.062
(-0.19)	(2.69)	(3.98)	(2.31)	(-0.02)	(1)

Obs.: Estatística t-robusta entre parênteses

Abaixo da estatística J encontra-se o número de graus de liberdade.

Tabela 10: Período sem crise - Estatística J com diferentes métodos de estimação de H - Modelos Aumentados

	Andrew e Monahan	NW - automático	NW - banda de 20	NW - banda de 40
$z = \{c, \omega\}$				
	NA	NA	NA	NA
$z = \{c, \omega, RV_{t+1}, \dots, RV_{t+20}\}$				
	0.251	0.345	0.381	0.594
$z = \{c, \omega, RV_{t+20}\}$				
	0.877	0.874	0.871	0.872
$z = \{c, \omega, RV_{t+15}\}$				
	0.415	0.388	0.386	0.412
$z = \{c, \omega, RV_{t+10}\}$				
	0.288	0.280	0.284	0.338
$z = \{c, \omega, RV_{t+5}\}$				
	0.201	0.180	0.202	0.270
$z = \{c, \omega, RV_{t+1}\}$				
	0.062	0.075	0.089	0.119

Tabela 11: Período sem crise - Condição de Ortogonalidade - Modelos Base

Variável dependente: VIX					
Constante	$SP500_b$	HL_b	RV_b	RV	J
$z = \{c, \omega\}$					
-1.034	0.282	0.021	0.136	0.325	NA
(-0.45)	(4.29)	(0.22)	(2.24)	(6.44)	
$z = \{c, \omega, RV_{t+1}, \dots, RV_{t+20}\}$					
-2.172	0.294	0.074	0.125	0.310	0.053
(-0.95)	(4.29)	(0.91)	(2.22)	(6.21)	(5)
$z = \{c, \omega, RV_{t+20}\}$					
-0.637	0.268	-0.001	0.150	0.349	0.011
(-0.27)	(4.07)	(-0.00)	(2.26)	(7.20)	(1)
$z = \{c, \omega, RV_{t+15}\}$					
-1.141	0.270	0.032	0.137	0.345	0.006
(-0.50)	(4.10)	(0.32)	(2.22)	(7.15)	(1)
$z = \{c, \omega, RV_{t+10}\}$					
-1.684	0.286	0.047	0.129	0.334	0.010
(-0.74)	(4.30)	(0.52)	(2.21)	(6.76)	(1)
$z = \{c, \omega, RV_{t+5}\}$					
-2.058	0.296	0.058	0.129	0.316	0.013
(-0.89)	(4.30)	(0.68)	(2.26)	(6.13)	(1)
$z = \{c, \omega, RV_{t+1}\}$					
-2.836	0.318	0.082	0.124	0.288	0.004
(-1.16)	(4.34)	(1.03)	(2.23)	(5.27)	(1)

Obs.: Estatística t-robusta entre parênteses

Abaixo da estatística J encontra-se o número de graus de liberdade.

Tabela 12: Período sem crise - Estatística J com diferentes métodos de estimação de H - Modelos Base

	Andrew e Monahan	NW - automático	NW - banda de 20	NW - banda de 40
$z = \{c, \omega\}$				
	NA	NA	NA	NA
$z = \{c, \omega, RV_{t+1}, \dots, RV_{t+20}\}$				
	0.053	0.247	0.417	0.577
$z = \{c, \omega, RV_{t+20}\}$				
	0.011	0.069	0.097	0.132
$z = \{c, \omega, RV_{t+15}\}$				
	0.006	0.058	0.086	0.128
$z = \{c, \omega, RV_{t+10}\}$				
	0.010	0.066	0.090	0.135
$z = \{c, \omega, RV_{t+5}\}$				
	0.013	0.058	0.095	0.151
$z = \{c, \omega, RV_{t+1}\}$				
	0.004	0.020	0.037	0.070

Tabela 13: Período de crise - Condição de Ortogonalidade - Modelos Aumentados

Variável dependente: VIX					
Constante	$SP500_a$	HL_a	RV_a	RV	J
$z = \{c, \omega\}$					
9.335 (15.66)	0.232 (10.21)	-0.024 (-0.94)	-0.050 (-3.24)	0.201 (4.99)	NA
$z = \{c, \omega, RV_{t+1}, \dots, RV_{t+20}\}$					
8.252 (9.25)	0.208 (7.06)	0.001 (0.03)	-0.029 (-1.14)	0.219 (4.18)	0.037 (5)
$z = \{c, \omega, RV_{t+20}\}$					
9.196 (9.19)	0.213 (0.21)	-0.009 (-0.00)	-0.051 (-3.51)	0.228 (5.78)	0.019 (1)
$z = \{c, \omega, RV_{t+15}\}$					
9.179 (15.51)	0.210 (10.19)	-0.009 (-0.39)	-0.049 (-3.37)	0.233 (5.81)	0.004 (1)
$z = \{c, \omega, RV_{t+10}\}$					
9.222 (15.97)	0.208 (10.40)	-0.006 (-0.27)	-0.052 (-3.69)	0.238 (6.01)	0.001 (1)
$z = \{c, \omega, RV_{t+5}\}$					
8.968 (16.29)	0.199 (10.74)	0.008 (0.39)	-0.064 (-4.87)	0.271 (6.84)	0.000 (1)
$z = \{c, \omega, RV_{t+1}\}$					
8.601 (9.10)	0.208 (6.16)	-0.006 (-0.16)	-0.018 (-0.58)	0.198 (3.32)	0.003 (1)

Obs.: Estatística t-robusta entre parênteses

Abaixo da estatística J encontra-se o número de graus de liberdade.

Tabela 14: Período de crise - Estatística J com diferentes métodos de estimação de H - Modelos Aumentados

	Andrew e Monahan	NW - automático	NW - banda de 20	NW - banda de 40
$z = \{c, \omega\}$				
	NA	NA	NA	NA
$z = \{c, \omega, RV_{t+1}, \dots, RV_{t+20}\}$				
	0.037	0.233	0.233	0.512
$z = \{c, \omega, RV_{t+20}\}$				
	0.019	0.061	0.063	0.098
$z = \{c, \omega, RV_{t+15}\}$				
	0.004	0.024	0.025	0.055
$z = \{c, \omega, RV_{t+10}\}$				
	0.001	0.012	0.013	0.034
$z = \{c, \omega, RV_{t+5}\}$				
	0.000	0.004	0.005	0.032
$z = \{c, \omega, RV_{t+1}\}$				
	0.003	0.015	0.017	0.038

Tabela 15: Período de crise - Condição de Ortogonalidade - Modelos Base

Variável dependente: VIX					
Constante	$SP500_b$	HL_b	RV_b	RV	J
$z = \{c, \omega\}$					
11.334	0.169	-0.097	0.013	0.320	NA
(18.21)	(9.79)	(-3.14)	(0.73)	(8.88)	
$z = \{c, \omega, RV_{t+1}, \dots, RV_{t+20}\}$					
10.165	0.166	-0.070	0.022	0.296	0.041
(10.54)	(5.99)	(-1.40)	(0.85)	(6.56)	(5)
$z = \{c, \omega, RV_{t+20}\}$					
11.338	0.170	-0.099	0.013	0.320	0.836
(11.33)	(0.17)	(-0.09)	(0.73)	(8.47)	(1)
$z = \{c, \omega, RV_{t+15}\}$					
11.332	0.167	-0.095	0.012	0.321	0.739
(17.97)	(9.85)	(-3.01)	(0.67)	(8.66)	(1)
$z = \{c, \omega, RV_{t+10}\}$					
11.333	0.165	-0.092	0.011	0.320	0.318
(18.20)	(9.86)	(-2.95)	(0.64)	(8.80)	(1)
$z = \{c, \omega, RV_{t+5}\}$					
11.293	0.160	-0.082	0.007	0.318	0.014
(18.38)	(9.91)	(-2.69)	(0.42)	(8.93)	(1)
$z = \{c, \omega, RV_{t+1}\}$					
10.801	0.156	-0.072	0.003	0.345	0.000
(17.38)	(9.08)	(-2.30)	(0.16)	(9.01)	(1)

Obs.: Estatística t-robusta entre parênteses

Abaixo da estatística J encontra-se o número de graus de liberdade.

Tabela 16: Período de crise - Estatística J com diferentes métodos de estimação de H - Modelos Base

	Andrew e Monahan	NW - automático	NW - banda de 20	NW - banda de 40
$z = \{c, \omega\}$				
	NA	NA	NA	NA
$z = \{c, \omega, RV_{t+1}, \dots, RV_{t+20}\}$				
	0.041	0.370	0.270	0.545
$z = \{c, \omega, RV_{t+20}\}$				
	0.836	0.895	0.893	0.911
$z = \{c, \omega, RV_{t+15}\}$				
	0.739	0.821	0.821	0.847
$z = \{c, \omega, RV_{t+10}\}$				
	0.318	0.445	0.445	0.506
$z = \{c, \omega, RV_{t+5}\}$				
	0.014	0.036	0.038	0.067
$z = \{c, \omega, RV_{t+1}\}$				
	0.000	0.025	0.022	0.074