XI Workshop em Estatística Aplicada e Biometria

Centro de Ciências Exatas e Tecnológicas - UFV

(08 de novembro de 2024)

Realização:

Utilizando o modelo APARCH para a modelagem da volatilidade dos preços da soja.

Lucas Pereira Belo[1]

Eduardo Campana Barbosa[1]

Paulo César Emiliano[1]

[1]Universidade Federal de Viçosa, Departamento de Estatística, lucas.p.belo@ufv.br.

INTRODUÇÃO

A soja é um dos principais produtos agrícolas do Brasil, sendo o segundo item mais exportado do país. Além da utilização do grão como alimento, a soja pode ser utilizada na produção de diversos outros produtos, dentre eles: Chocolate, temperos prontos, massas, alguns derivados de carne, mistura para bebidas, entre outros. O que reforça a importância deste produto para a economia brasileira.

OBJETIVOS

Este trabalho tem por objetivo modelar a média e a variância condicional da série logarítmica dos retornos de preços da soja, ajustando um modelo que capture as variações ao longo do tempo. Além de permitir identificar períodos de maior e menor risco, estudos voltados para este tema, permitem uma melhor orientação na tomada de decisões no mercado.

METODOLOGIA

Os dados avaliados referem-se aos preços semanais por 450 sacas de 50 kg de soja e contemplam o período de 30 de janeiro de 2011 até 07 de janeiro de 2024. A análise inicial ocorreu fazendo a série de log retornos, que é definida por Morettin e Toloi (2006) da seguinte forma:

Definição 1: Seja P_t o preço de um ativo em um instante t. Por conveniência supõe-se que não são pagos dividendos no intervalo $(t, t-1)$, definiu-se então o retorno simples como sendo a razão entre a variação dos preços do ativo no intervalo $(t, t-1)$ dividido pelo preço no instante $t-1$, este valor denotaremos de $\boldsymbol{R_t}$. O log retorno de um ativo é então definido como sendo: $\boldsymbol{r_t = log\left(\frac{P_t}{P_{t-1}}\right) = log(1 + R_t)}$.

Em seguida, verificou-se o ACF(Autocorrelation Function) e PACF(Partial Autocorrelation Function) das séries log retorno dos preços da soja. Para verificar a existência de tendência e periodicidade determinística nos dados, foram aplicados os testes de Cox-Stuart e G de Fisher. Posteriormente, ajustou-se um modelo ARMA(p,q) para modelar a média condicional da série.

Definição 2: Os autores Morettim e Toloi (2006) apresentam o processo autorregressivo e de médias móveis, ARMA(p,q) é da seguinte forma: $\tilde{Z}_t = \emptyset_1\tilde{Z}_{t-1} + \cdots + \emptyset_p\tilde{Z}_{t-p} + \varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q}$. Sendo que um modelo frequentemente utilizado é o ARMA(1,1), onde $p = q = 1$, daí a equação anterior se reduz a, $\tilde{Z}_t = \emptyset_1\tilde{Z}_{t-1} + \varepsilon_t - \theta_1\varepsilon_{t-1}$, em que, $\varepsilon_t \sim RB(0, \sigma^2)$. O termo $\emptyset$ é o parâmetro associado ao termo autorregressivo do modelo, θ é o termo associado a média móvel.

Na análise residual foram utilizados os testes de Shapiro Wilk para verificar a normalidade, o teste de Box-Pierce (LJUNG e BOX, 1978) para verificar a existencia de autocorrelação e heterocedasticidade nos resíduos. Com o intuito de modelar a volatilidade existente nos resíduos, ajustou-se um modelo autorregressivo de heteroscedasticidade condicional generalizada, conhecido como modelo GARCH(r,s), cuja definição pode ser encontrada em (MORETTIN E TOLOI, 2006). A alternativa proposta no estudo foi o modelo APARCH.

Definição 3: O modelo (*Asymmetric Power* ARCH) que é uma extensão dos modelos GARCH (*Generalized Autoregressive Conditional Heteroskedasticity*) apresentado por Ding, Granger e Engle (1993). Os autores apresentam o modelo que tem a seguinte forma:

$$\sigma_t^\delta = \omega + \sum_{i=1}^{p} \alpha_i \cdot (|\epsilon_{t-1}| - \gamma_i \cdot \epsilon_{t-1})^\delta + \sum_{j=1}^{p} \beta_j \cdot \sigma_{t-1}^\delta,$$

em que ω, α, ε, γ, β, δ, são parâmetros a serem estimados pelo modelo. O modelo pressupões as seguintes condições: $\omega > 0$, $0 < \sum\alpha_i < 1$, $0 < \sum\beta_j < 1$ e $\sum\alpha_i + \sum\beta_j < 1$. A escolha dos modelos durante todo o processo foi baseada no maior valor da função log verossimilhança ajustada, com estimativas realizadas pelo

RESULTADOS E DISCUSSÕES

Figura 1: Série histórica de preços

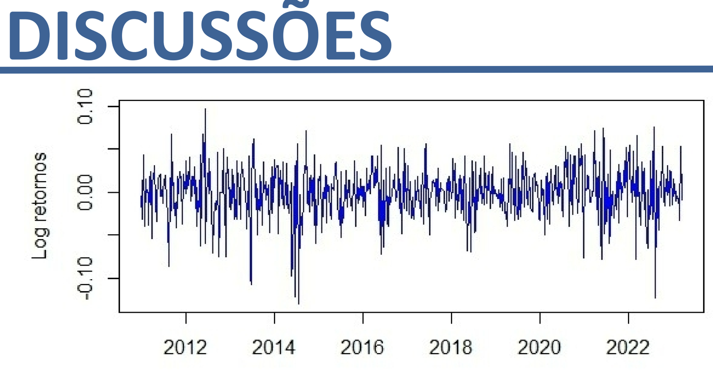

Figura 2: Série log retorno dos preços

Os testes de Cox-Stuart e G de Fisher apresentaram valores p de 28,66% e 91,88%, respectivamente, indicando que, ao nível de significância de 5%, não há evidências suficientes para rejeitar as hipóteses nulas de ausência de tendência e de periodicidade determinística nos dados respectivamente. Assim, foi ajustado um modelo ARMA(1,1), com um parâmetro autorregressivo e um de médias móveis, cujas estimativas dos parâmetros foram: $\hat{\emptyset} = -0{,}8679$, $\hat{\theta} = 0{,}8081$, para modelar a média condicional da série de log-retornos (todos estes significativos a 5% de probabilidade). Para os resíduos, foi ajustado um modelo GARCH(1,1), que captura a variância condicional, assumindo uma distribuição de erro generalizada enviesada (Skewed Generalized Error Distribution-SGED) para os resíduos. Os parâmetros estimados (todos estes significativos a 5% de probabilidade) foram: $\hat{\alpha}_1 = 0{,}1764$, $\hat{\beta}^1 = 0{,}6754$, $\hat{\vartheta} = 0{,}8733$ (parâmetro de assimetria dos resíduos), $\hat{\nu} = 1{,}7054$ (está relacionado à curtose da distribuição).

Embora o modelo tenha capturado parte da heterocedasticidade, o teste de Ljung-Box aplicado aos resíduos quadráticos mostrou um valor-p de 6,63% até a 10ª defasagem, sugerindo potencial refinamento. Como alternativa, ajustou-se um modelo APARCH(1,1) ajustado para variância da série, juntamente com um ARMA(1,1) para a média e uma distribuição de erro skew generalizada para os resíduos. Os parâmetros estimados, foram: $\hat{\emptyset}^1 = -0{,}5023$, $\hat{\theta}_1 = 0{,}4090$, $\hat{\alpha}_1 = 0{,}1436$, $\hat{\beta}^1 = 0{,}6258$, $\hat{\gamma} = -0{,}1113$, $\hat{\vartheta} = 0{,}8826$, $\hat{\nu} = 1{,}6916$, $\hat{\delta} = 3{,}4552$. A especificação final do modelo é então dada por:

$$Z_t = -0{,}5023Z_{t-1} + \varepsilon_t - 0{,}4090\varepsilon_{t-1},$$
$$\sigma_t^{3,3552} = 0{,}1436 \cdot (|\epsilon_{t-1}| - 0{,}1113 \cdot \epsilon_{t-1})^{3,3552} + 0{,}6258 \cdot \sigma_{t-1}^{3,3552},$$

com $\varepsilon_t \sim \mathrm{SGED} - (0{,}8826,\ 1{,}6916)$.

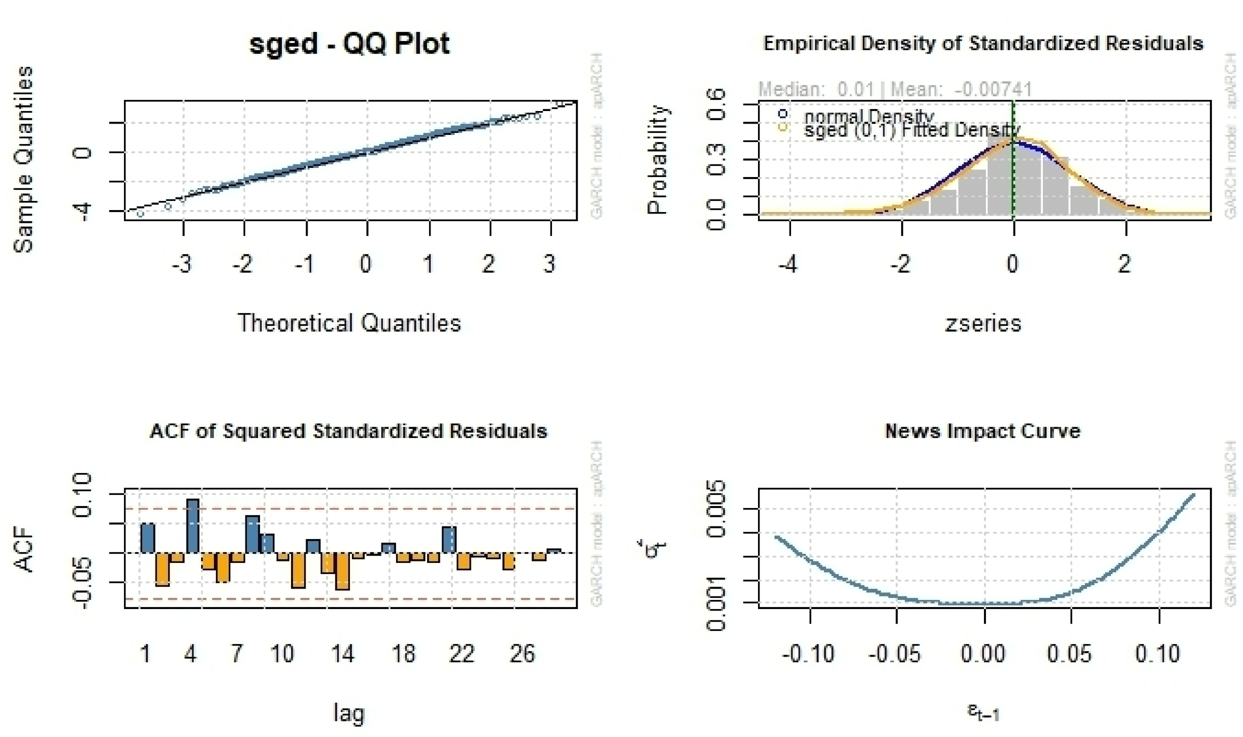

Figura 3: Gráfico da análise de resíduos do modelo APARCH(1,1)

CONCLUSÃO

O modelo APARCH apresentou um ajuste superior ao GARCH, com funções log-verossimilhança ajustadas, $\widehat{l_{GARCH}} = 1371.768$, $\widehat{L_{APARCH}} = 1373.79$ além de melhores resultados na mitigação da heterocedasticidade, evidenciados pelo teste de Ljung-Box, que indicou um p-valor > 0,0920 para a 9ª defasagem. Além disso, não foram evidenciados efeitos de alavancagem significativos nos dados.

REFERÊNCIAS BIBLIOGRÁFICAS

DING, Z.; GRANGER, C.W; ENGLE, R.F Uma propriedade de memória longa de retornos do mercado de ações e um novo modelo. **Revista de Finanças Empíricas**, v.1,n.1, págs. 83-106, 1993.

LJUNG, G. M.; BOX, G. E. On a measure of lack of fit in time series models. **Biometrika**, 174 v. 65, n. 2, p. 297-303, 1978.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de séries temporais.** 2. ed. São Paulo: E.Blücher, 538 p. 2006.

NELSON, D. B.; Conditional heteroskedasticity in asset returns: A new approach. **Econometrica: Journal of the econometric society**, p. 347-370, 1991.

AGRADECIMENTOS

Os autores agradecem a Capes, ao programa de pós graduação em estatística aplicada e biometria e a Universidade Federal de Viçosa que deram suporte financeiro para realização deste trabalho.

CAPES

PPESTBIO-UFV

UFV Universidade Federal de Viçosa