

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/343162874>

A GARCH Tutorial with R

Article in *Revista de Administração Contemporânea* · January 2021

DOI: 10.1590/1982-7849rac2021200088

CITATIONS

0

READS

5,048

4 authors, including:



Marcelo Scherer Perlin

Universidade Federal do Rio Grande do Sul

43 PUBLICATIONS 351 CITATIONS

[SEE PROFILE](#)



Mauro Mastella

Universidade Federal de Ciências da Saúde de Porto Alegre

10 PUBLICATIONS 9 CITATIONS

[SEE PROFILE](#)



Henrique Ramos

Universidade Federal do Rio Grande do Sul

11 PUBLICATIONS 24 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Research Publication Analysis [View project](#)



Public policy [View project](#)

Tutorial Article

A GARCH Tutorial with R

Um Tutorial sobre Modelos Garch no R



Marcelo Scherer Perlin¹
Mauro Mastella²
Daniel Francisco Vancin³
Henrique Pinto Ramos¹

ABSTRACT

Context: modeling volatility is an advanced technique in financial econometrics, with several applications for academic research. **Objective:** in this tutorial paper, we will address the topic of volatility modeling in R. We will discuss the underlying logic of GARCH models, their representation and estimation process, along with a descriptive example of a real-world application of volatility modeling. **Methods:** we use a GARCH model to predict how much time it will take, after the latest crisis, for the Ibovespa index to reach its historical peak once again. The empirical data covers the period between years 2000 and 2020, including the 2009 financial crisis and the current 2020's episode of the COVID-19 pandemic. **Conclusion:** we find that, according to our GARCH model, Ibovespa is more likely than not to reach its peak once again in one year and four months from June 2020. All data and R code used to produce this tutorial are freely available on the internet and all results can be easily replicated.

Keywords: volatility; GARCH; Ibovespa; tutorial.

RESUMO

Contexto: a modelagem de volatilidade é uma técnica avançada em econometria financeira, com diversas aplicações em pesquisa acadêmica. **Objetivo:** neste artigo tutorial abordaremos o tópico da modelagem de volatilidade na plataforma R. Discutiremos a lógica subjacente dos modelos GARCH, seus processos de representação e estimação, juntamente com um exemplo descritivo de uma aplicação no mundo real. **Métodos:** usamos um modelo GARCH para investigar quanto tempo levará, após a última crise, para que o índice Ibovespa volte a atingir seu pico histórico mais uma vez. Os dados empíricos cobrem o período entre os anos 2000 e 2020, incluindo a crise financeira de 2009 e o episódio atual de 2020 da pandemia do COVID-19. **Conclusão:** de acordo com nosso modelo GARCH, as chances de o Ibovespa atingir o seu pico passam de 50% um ano e seis meses após junho de 2020. Todos os dados e códigos R usados para produzir este tutorial estão disponíveis gratuitamente na internet e todos os resultados podem ser facilmente replicados.

Palavras-chave: volatilidade; GARCH; Ibovespa; tutorial.







1. Universidade Federal do Rio Grande do Sul, Escola de Administração, Porto Alegre, RS, Brazil.

2. Universidade Federal de Ciências da Saúde de Porto Alegre, Porto Alegre, RS, Brazil.

3. Universidade do Vale do Rio dos Sinos, Programa de Pós Graduação em Ciências Contábeis, Porto Alegre, RS, Brazil.

Cite as: Perlin, M. S., Mastella, M., Vancin, D. F., & Ramos, H. P. (2021). A GARCH tutorial with R. *Revista de Administração Contemporânea*, 25(1), e200088. <https://doi.org/10.1590/1982-7849rac2021200088>

of invited reviewers until the decision:

	1	2	3	4	5	6	7	8	9
1 st round									
2 nd round									
3 rd round									

JEL Code: A2, H12, G1.

Editor-in-chief: Wesley Mendes-Da-Silva (Fundação Getúlio Vargas, EAESP, Brazil)

Associate Editor: Henrique Castro Martins (PUC Rio, IAG, Brazil)

Reviewers: Pedro Raffy Vartanian (Universidade Presbiteriana Mackenzie, Brazil)

Júlio César Araújo da Silva Junior (Universidade Federal de Viçosa, Departamento de Economia, Brazil)

One of the reviewers chose not to disclose his/her identity.

Received: March 29, 2020

Last version received: July 06, 2020

Accepted: July 06, 2020

INTRODUCTION

Modeling uncertainty is a certain element of the financial practice, with important applications in portfolio allocation, risk management, and pricing of financial contracts. The simple question of how much uncertainty we can expect for future prices of financial contracts resulted in a large body of literature interested in understanding the statistical properties of price changes and how we can use them to make better predictions (Brockwell & Davis, 2016; Francq & Zakoian, 2019). When studying financial time series, researchers can regularly observe common characteristics (Cont, 2001). The main stylized facts for financial time series data are the absence of linear autocorrelation, heavy tails, asymmetry for gains or losses, agglomeration of volatilities, and leverage effect.

The ARCH (autoregressive conditional heteroscedasticity) and GARCH (generalized autoregressive conditional heteroscedasticity) family of models constitute a seminal innovation in the field of financial modeling by taking into account some of the stylized effects of financial data. Among other contributions, it resulted in the 2003 Nobel Prize for its creator, Robert Engle¹. Its main contribution is to allow uncertainty to be a dynamic process. That is, instead of assuming that the level of future volatility is constant, the ARCH and GARCH family of models acknowledge its time-varying process.

In practice, the implications of dynamic volatility are as follows: financial returns will more often than expected result in large losses — the ‘fat tail’ effect; uncertain periods tend to cluster, with large price variations within days. As a recent and practical example, on 03/09/2020 the Ibovespa index lost 13.98% of its value in a single day due mostly to the COVID-19 episode. The index recovered 13% of its value just a couple of days later, in 03/13/2020. When considering its historical distribution of price changes, both events are highly unexpected and happened just a couple of days apart.

Ignoring the ARCH effect and underplaying short-term uncertainty can be very costly (Bollerslev, Engle, & Nelson, 1994). Risk control measures in investment portfolios need to assess how likely a large loss can happen. By assuming constant volatility, an analyst underplays the inner risk of financial contracts and, possibly, is surprised by extreme unexpected losses in the investment portfolio. Likewise, banking regulations, such as Basel III, require that banks report their level of portfolio risk systematically and periodically. Thus, given that banks act as liquidity hubs, an incorrect calculation of volatility and risk can threaten the stability of a whole financial system as an unexpected financial shock can force banks to quickly liquidate financial contracts and increase their cash position.

In the scientific side, there is an extensive body of literature in Brazil relying on GARCH models. Volatility models can be used in any empirical study, such as scientific articles, thesis, and dissertations, where the prediction of uncertainty plays an important role. A practical example is the pricing of derivative contracts where, in order to calculate the fair price of future cash flows from a derivative product, we need to understand the future distribution of the underlying asset using a volatility model. Another example is the estimation of starting parameters in a more complex specification, such as a copula model.

A non-exhaustive overview of topics recently covered would include the covariance of Brazilian stock exchange (Mastella & Coster, 2014), optimal minimum variance portfolios (Caldeira, Moura, Perlin, & Santos, 2017), risk exposure in Brazilian sectorial stock indices (Bernardino, Brito, Ospina, & Melo, 2018; Lobão & Fernandes, 2018), correlation between stock returns (Costa, Porto Junior & Menezes, 2018), among others. As for the international literature, we suggest the following articles: covariance in the stock exchange (Byström, 2004; De Goeij & Marquering, 2004), optimal portfolios (Deng, Ma, & Yang, 2011; Varga-Haszonits & Kondor, 2007), risk exposure (Carson, Elyasiani, & Mansur, 2008), correlation between stock returns (Omran & McKenzie, 2000). And for a better understanding of the applicability of multivariate GARCH models, see Bauwens, Laurent and Rombouts (2006) — this paper surveys the most important developments in multivariate ARCH-type modeling.

In this paper, we present a practical tutorial for the ARCH/GARCH family of models, introducing its underlying motivations and a reproducible study case in R, with real data and reproducible code. Given the recent 2020's turmoil in financial markets and huge drop in stock prices on the stock exchange, we offer a step-by-step guide to model financial returns. Our goal is to use a simulated GARCH model to assess when will Ibovespa reach its historical peak once again and, consequently, the expected time it will take for the market to recover from the current crisis. Thus, our contribution to researchers is to provide a step-by-step tutorial for beginners to investigate a classical issue in finance and, for experienced researchers, the tutorial allows to easily reformat the code for other applications and models.

This tutorial differs from other econometric manuals based in R since it brings a step-by-step guide that covers not only the description of actions necessary to obtain a secure database from public data and the estimation of models, but also all tests and steps necessary for its correct estimation and interpretation of results. All steps of the tutorial are organized in six R scripts, including helpful code comments that will guide the novice reader.

Those researchers and students hungry for more relevant complementary books and manuals may also find shelter in the work of [Tsay \(2005\)](#), [Ruppert and Matteson \(2015\)](#), [Ferreira \(2018\)](#), and [Perlin \(2018\)](#).

The article is organized as follows: we start with a description of the method and brief presentation of the underlying theory, followed by an empirical application of the model in real world data. We end the paper with the usual conclusion section.

A GARCH MODEL

As noted in the previously cited literature, financial time series such as stock prices, inflation rates, and exchange rates present the phenomenon of volatility agglomeration. In other words, they present periods in which the prices/values of these series show significant fluctuation, followed by periods in which there is low variation.

In the modeling side, the most basic quantitative tool for researchers in finance is the ordinary least squares (OLS) model. This is a natural choice, because applied econometricians are typically called upon to determine how much one variable will change in response to a change in some other variable ([Engle, 2001](#)). Stationarity becomes an issue when modeling financial returns, as they tend to have non-constant volatility (heteroscedasticity). In these situations, the use of OLS models could draw misleading conclusions in an empirical study where volatility plays an important role.

The typical warning is that, in the existence of heteroscedasticity, the regression coefficients for an ordinary least squares regression are still unbiased, but the standard errors and confidence intervals estimated by traditional procedures will be too narrow, giving a false sense of accuracy ([Engle, 2001](#)). Instead of considering it a problem to be fixed (through the use of heteroscedasticity-consistent standard errors, for example), ARCH and GARCH models care for heteroscedasticity as a variance to be modeled ([Engle, Focardi, & Fabozzi, 2012](#)). Therefore, not only are the shortcomings of least squares corrected, but also a forecast is computed for the variance of each error term.

Formally, the assumption of constant variance in a model is called homoscedasticity. The opposite, non-constant variance, is the focus of ARCH/GARCH models, especially in the univariate case. Based on this fact, the estimation and prediction of volatility should be different from those observed in classic time series models, such as the ARMA (autoregressive moving average) by [Box and Jenkins \(1976\)](#). The later does not accept some of the mentioned stylized facts about the volatility, such as conditional/unconditional

non-normality and non-constant conditional variance over time.

To address the issues of heteroscedasticity, [Engle \(1982\)](#) presented the ARCH models in a study of inflation rates. These models seek to estimate time-dependent volatility as a function of the previously observed volatility. The original ARCH model proposed by [Engle \(1982\)](#) modeled the variance of errors in a regression model as a linear function of lagged values of squared regression errors. Using the mathematical notation of [Tsay \(2005\)](#) on forward, we can write an ARCH (m) model as:

$$R_t = \sigma_t \epsilon_t \text{ (conditional average)} \quad (1)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 R_{t-1}^2 + \dots + \alpha_q R_{t-q}^2 \text{ (conditional variance)} \quad (2)$$

The dependent variable R_t represents the returns of a financial asset in a given frequency, that is, the percentage (or log difference) of prices from one period to the next. The term σ_t^2 is the conditional volatility at time t , while α_q are the different parameters of the ARCH models, usually estimated from real data. As noted, the ARCH model has a specification for both the conditional average and the conditional variance. Specifically, an ARCH method models the variance at a time step as a function of the residual errors from a mean process.

Although the ARCH model may be simple, it often demands many parameters/lags to properly explain the volatility process of an asset return ([Tsay, 2005](#)). To solve this problem, [Bollerslev \(1986\)](#) extended Engle's original work by developing a technique that allows the conditional variance to be an ARMA process. In other words, a GARCH model is equivalent to an ARCH model with many, many lags. Formally, we define a GARCH (p,q) model as follows:

$$R_t = \sigma_t \epsilon_t \text{ (conditional average)} \quad (3)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i R_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \text{ (conditional variance)} \quad (4)$$

where ϵ_t is a sequence of independent and identically distributed random variables with mean zero, variance equal to one, and $\alpha_i > 0$ for $i > 0$. With such properties, the coefficients α_i satisfy conditions to ensure that the unconditional variance is finite and positive. In practice, ϵ_t is often assumed to follow the standard normal, a standardized Student's t or a generalized error distribution. Hence, if all values of β_j are equal to zero, the GARCH (q,p) model is equivalent to an ARCH (q) model. The benefits of the GARCH model should be clear: a high-order ARCH model may have a more parsimonious GARCH representation that is much easier to identify and estimate. This is particularly true because all coefficients in equation (4) must be positive. Likewise, to ensure that the variance is

finite, all characteristic roots of the GARCH equation have to lie inside the unit circle.

In practice, the parameters of a GARCH model are estimated from the data. The standard method for the estimation of parameters of a GARCH model is called maximum likelihood (ML). The main idea of the method is to find parameters that match, as close as possible, the distribution of predictions from the model against the distribution of the real data. Nowadays, several econometric software, including R, have a GARCH toolbox, and ML estimation of standard GARCH models takes just a few seconds on a modern computer. From a theoretical viewpoint, ML estimators benefit from being asymptotically optimal under certain conditions. Moreover, alternative estimation methods such as Bayesian methods (Fioruci, Ehlers, & Andrade Filho, 2014) also exist.

Alternative volatility models

As the ARCH and GARCH models set the cornerstone framework for volatility modeling, the reader must be aware

that further research improved and extended these models in order to incorporate other stylized facts noted in financial markets. One possible critique to ARCH-GARCH family is that it models the conditional variance as a linear function of the squared past innovations and its implicit symmetry of impact, where a positive shock affects volatility the same way as a negative shock (Francq & Zakoian, 2019).

The exponential GARCH model (EGARCH) proposed by Nelson (1991) allows asymmetric effects depending on the sign of the random innovation (error term). This follows the idea that volatility may rise in response to ‘bad news’ and can be reduced after ‘good news’ (Nelson, 1991). The GJR-GARCH model (Glosten, Jagannathan, & Runkle, 1993) also deals with the different asymmetric effects of past innovations, but in a simpler way. Table 1 presents the main formulation of ARCH, GARCH, EGARCH, and GJR-GARCH models. Although the scope of this tutorial is not to explore in detail the variety of conditional heteroscedastic models, we believe a novice researcher should be familiar with other types of models.

Table 1. A selection of conditional heteroscedastic models.

Model	Variance equation
ARCH (q)	$R_t = \sigma_t \epsilon_t, \sigma_t^2 = \alpha_0 + \alpha_1 R_{t-1}^2 + \dots + \alpha_q R_{t-q}^2$
GARCH (p,q)	$R_t = \sigma_t \epsilon_t, \sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i R_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2$
EGARCH (p,q)	$R_t = \sigma_t \epsilon_t, \ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^q \frac{ R_{t-i} + \delta_i R_{t-i}}{\sigma_{t-i}} + \sum_{j=1}^p \beta_j \ln(\sigma_{t-j}^2)$
GJR-GARCH (p,q)	$R_t = \sigma_t \epsilon_t, \sigma_t^2 = \alpha_0 + \sum_{i=1}^q (\alpha_i + \gamma_i N_{t-i}) R_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2$

Note. Table 1 presents variance formulas of four different GARCH models. Be aware this is a non-exhaustive list and many models were left out.

Regarding the EGARCH model, the δ_i parameter implies the leverage effect of R_{t-i} , in which volatility tends to increase more after negative returns than in positive returns (Black, 1976). In the GJR-GARCH model, N_{t-i} is an indicator variable that assumes value of one if $R_{t-i} < 0$ and zero otherwise. Thus, a positive R_{t-i} contributes $\alpha_i R_{t-i}^2$ to volatility (σ_t^2), whereas a negative R_{t-i} contributes $(\alpha_i + \gamma_i) R_{t-i}^2$ to volatility. Considering the parameter $\gamma_i > 0$, the model uses zero as the threshold to separate the impact of past shocks (Tsay, 2005).

We do not intend to exhaust the possible explanations and applications of the models described in this tutorial. Even so, some essential textbooks could be highlighted for students in the need for a deeper understanding of GARCH models: Nelson (1991),

Hamilton (1994), Tsay (2005), Morettin and Tolo (2006), Bueno (2011), and Härdle, Chen and Overbeck (2017).

APPLICATION OF A GARCH MODEL

In this section, we will present a practical application of a GARCH model, with a real-world dataset. We will use a GARCH model to answer a simple question: **given the most recent financial crisis of 2020, when will Ibovespa reach its historical peak once again?** We can try to answer this question by estimating a GARCH model for the index, and use it to simulate many different future price paths.

In this application, we will explore alternative GARCH models and will provide a hands-on experience in dealing with econometric modeling and simulation in R, a widely used

programming platform in academia and financial industry. Like R, other software can also estimate and simulate GARCH models: EViews, Matlab, OXmetrics, Python, SAS, Stata, among others. Our choice for using R is justified by its compatibility with different operating systems, large user base, the absence of license fees, and easy distribution of external modules through CRAN, a user contributed repository of packages (R Core Team, 2017).

Choosing an R module that fits the overall needs of this tutorial is an important step. We search for established packages, with resources — R functions — for the estimation, simulation, and forecasting of GARCH models, which are still actively maintained by their authors. As such, we find two main candidates, fGarch (Wuertz, Setz, Chalabi, Boudt, Chausse, & Miklovac, 2020) and rugarch (Ghalanos, 2020). We choose the

rugarch due to its support of a larger family of GARCH models. The estimation of advanced GARCH specifications, such as regime switching volatility models, is available in R, but not used in this tutorial. As a reference, see the work of Ardia, Blueau, Boudt, Catania and Trottier (2019).

All the data, code, and results presented in this section are available in Github (<https://github.com/msperlin/GARCH-RAC>; retrieved in June 15, 2020) and in the journal's Dataverse (Perlin, Mastella, Vancin, & Ramos, 2020). Simply download the zip file in your computer and extract it to a personal folder. All scripts are distinctively named and ordered, representing the different and sequential stages of the data-based research. All R script files include section with a description, making it clear which research content, such as a figure or table, is being produced by the script. We organized the code as follows.

Table 2. Filename and description of R scripts used in the study.

Filename	Description
00-Prepare_Computer.R	Mandatory step. Setup the computer by installing all required R packages. Also provides further steps for alternative operating systems such as Linux.
01-Get_Index_Data.R	Using the internet and R package BatchGetSymbols, imports a dataset of prices of the Ibovespa index.
02-Do_Descriptive_Figures.R	Creates and saves the descriptive figure presented in the paper, Figure 2 and Figure 3.
03-Do_ARCH_Test.R	Performs the ARCH test in the data and exports Table 3.
04-Estimate_Garch_Model.R	Estimate an introductory GARCH model in three different volatility formulas and exports the estimation results into a html file (Table 4).
05-Find_Best_Garch_Model.R	Finds the best ARMA(ar,ma)-GARCH(p,q) model for the dataset, including changes in variance equation and distribution parameter. Also produces Figure 4.
06-Simulate_Garch_Model.R	Simulates the previous GARCH model and plot simulated paths and probabilities. Produces Figures 5 and 6.

Note. Table 2 reports the R scripts related to this tutorial paper, including description and produced material (all figures and tables). You can download a compressed file with all scripts at Github (<https://github.com/msperlin/GARCH-RAC>; retrieved in June 15, 2020). More instructions on how to execute the code in your own computer can be found at the code page.

Being able to run the R code in your own computer is an important step for making the most out of this tutorial paper. For that, in the first script *00-Prepare_Computer.R* and in Github's *Readme.md* file we offer clear steps to execute all scripts without any error, including support for alternative operating systems such as Linux and Mac. If you followed the steps and still have not been able to execute the script, please let the authors know by using the Github issue system or directly contacting the first author of the paper.

A simple way of executing this project is with RStudio Cloud. At link <https://rstudio.cloud/project/1371589> (retrieved in June 15, 2020) you will find the full working project, executable in a browser, without the need of any local preparation such as package installation. However, while this alternative is practical, we encourage readers to independently

execute the code in their own machine, which is likely how the work will be completed in the future.

The first script, *00-Prepare_Computer.R*, will make sure all R dependencies are available in the computer. The second script, *01-Get_Index_Data.R*, will import the index data from the internet and store it as a local file, which will then be used in other scripts. The alphabetical order is not accidental; one should execute each script in the same order. We also took special attention in writing code comments that will guide the user throughout the scripts and help the learning process. It is important to notice that other files and folders available in Github should be downloaded as well as the codes described in Table 2.

Next, Figure 1, we present a flowchart that can guide upcoming users to visualize and understand the steps required to reproduce the results in this section.

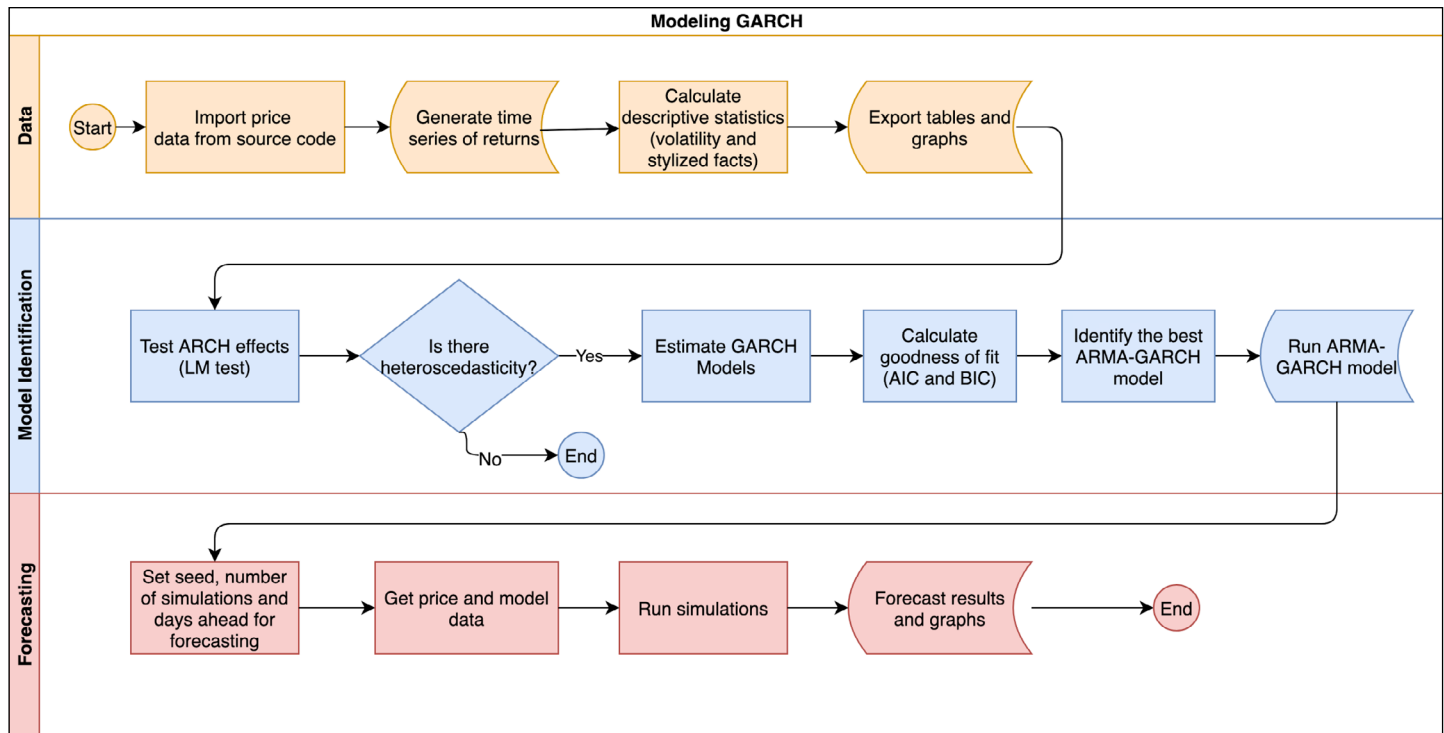


Figure 1. Flowchart of R code.

Figure 1 shows the stages of the R code, from data importation, testing for ARCH effects, estimation, and selection of models and results presentation.

All reported tests and estimations were executed in a Linux Ubuntu 20.03 machine, with R version 4.0.1. Be aware that some parameters, especially in the estimation of GARCH models, might be slightly different due to changes in computer's specifications.

The data

The chosen data for this example is Ibovespa, a broad market index for the Brazilian equity market. Currently, 06/15/2020, it is composed of approximately 70 stocks and regarded as the main thermometer of the local market, serving as a benchmark for investments and derivatives contracts. Worth pointing out that, without loss, the analysis in this study could be conducted using an individual stock or other international stock index. In fact, the R code was designed so that the user only needs to change the ticker symbol and the name of the series in script *01-Get_Index_Data.R*.

The price data is composed of daily closing values of the index from 01/01/2000 to 06/15/2020, including the 2009 financial crisis and the current 2020's episode of the COVID-19 pandemic. We opted for daily data because of its frequent use in volatility studies and to avoid possible microstructure noise from higher frequency data (Aït-Sahalia, Mykland, & Zhang, 2011). Regarding the choice of the period, we seek a window of time large enough to

encompass different market conditions and volatility regimes. The origin of the data is Yahoo Finance, a vast public repository of financial data (<https://br.financas.yahoo.com/> retrieved in June 15, 2020). The choice is justified by its open access nature - anyone can use R or other platform to download daily stock prices from Yahoo Finance.

The first step of the study is to download the dataset of daily values of Ibovespa and manipulate the data. In this example, we are interested in the vector of daily log returns, calling it R_t . Such vector represents the log returns of Ibovespa at time t , where t goes from one to the number of observations in the sample. We calculate it based on P_t , a vector of prices, using the following formula:

$$R_t = \log\left(\frac{P_t}{P_{t-1}}\right) \quad (5)$$

The log return is simply the log difference of the value/price from one day to the next. We use log return due to its statistical properties, such as stationarity and ergodicity. Log returns have some more favorable properties for statistical analysis than the simple net returns as shown by Quigley and Ramsey (2008). The continuously compounded multi-period return is simply the sum of the continuously compounded one-period returns.

Aas (2004) points out that (a) arithmetic returns will follow a lognormal distribution if the geometric returns follow a normal distribution and (b) the differences between arithmetic and geometric return distributions will grow larger as the volatility of the returns increases. For temporal aggregation, it is more convenient to work with log returns,

while for cross-section aggregation simple returns are more convenient as advocated by Morettin (2017). Next, in Figure 2, we show the evolution of Ibovespa, panel A, along with its daily log returns, panel B, produced by the execution of script *02-Do_Descriptive_Figures.R*.

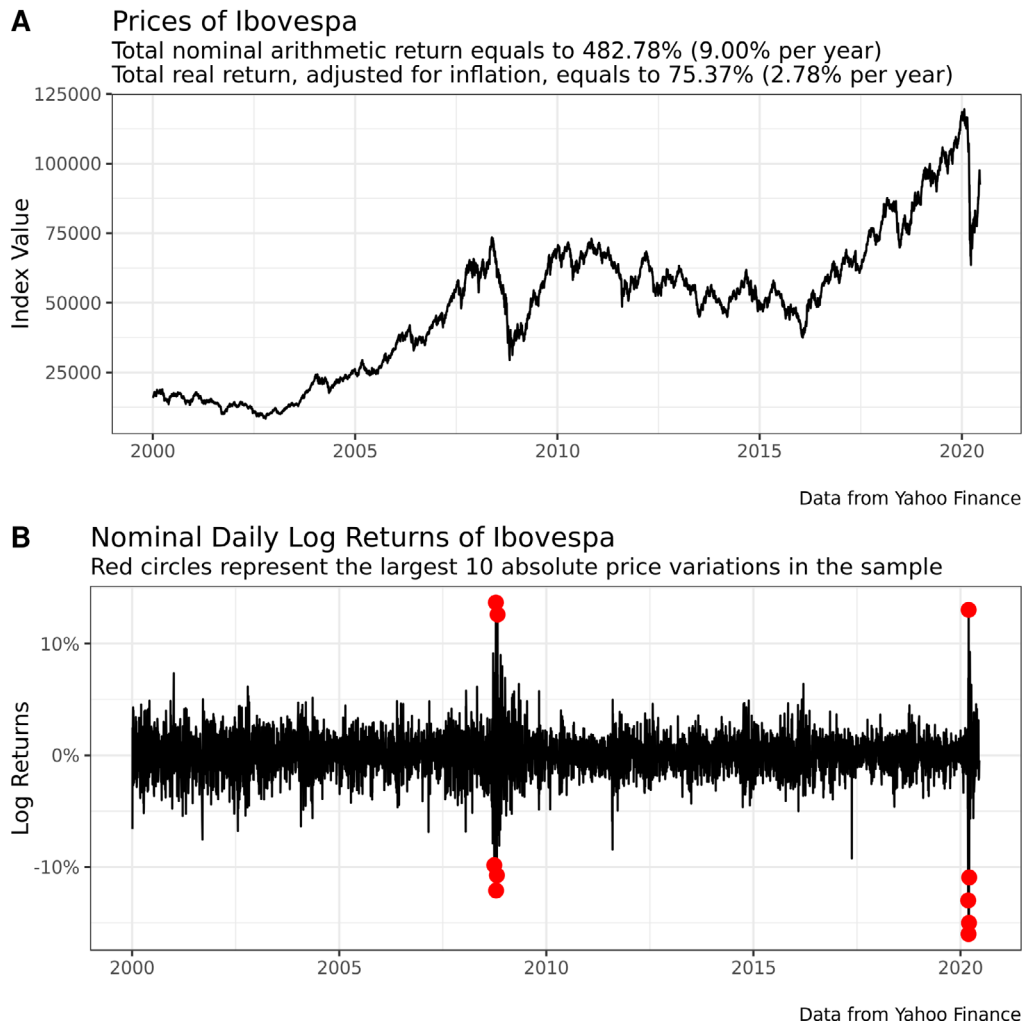


Figure 2. Prices and returns of the Ibovespa index (2000-2020).

Source: elaborated by the authors. Figure 2 presents the value of the Ibovespa index in Panel A and its corresponding log return in Panel B. All price data was imported in the daily frequency from Yahoo Finance, covering the period between January 2000 and June 2020. In panel A, we use the IPCA index and the inflation benchmark in the calculation of real returns in the period. The red circles in Panel B indicate the largest 10 absolute (positive or negative) price changes found in the sample of log returns. Figure 2 is produced by script *02-Do_Descriptive_Figures.R*.

Historically, an investor of the Brazilian equity market will likely see a positive nominal (without considering inflation) return in his investments (see panel A of Figure 2). Overall, if someone mirrored the Ibovespa composition at the beginning of 2000, he/she would have seen a total return of 482.78%, equivalent to 9.00% per year. When adjusting for inflation using Fisher equation (Fisher, 1896), however, the annual

return drops to a meager 2.78% per year. Not surprisingly, the 2020's COVID-19 event of global pandemic has hurt investor's historical returns significantly. While writing this paper, prices have not yet recovered from the big drop in March 2020. Such an episode resulted in the largest price drop in the sample, with a -14.77% price variation in 03/12/2020.

Looking at panel B of Figure 2, we see that most of the returns are centered around the value of zero. Big price changes — up or down — tend to happen within a close period. This is what is commonly called ‘volatility clustering’. As an example, notice that, out of the ten greatest absolute price changes (red points in the chart), five occurred in the 2009 financial crisis, two of them with positive return and three with negative. Not surprisingly, such property is not specific to the Brazilian market. We are confident that the clustering of volatility will also happen to broad equity indices from other countries (Francq & Zakoian, 2019).

The lesson here is that returns of financial assets present large changes that tend to agglomerate. In the sampled data, the ten most extreme price changes have happened mostly in two episodes, 2009’s financial crisis and 2020’s COVID-19 pandemic. When modeling returns, we should take such an effect into consideration, and not make the mistake of considering a constant volatility of returns. This is exactly the solution that an ARCH/GARCH model offers.

Additionally, it is always helpful to check the autocorrelation pattern of the analyzed financial series, which will show how an observation at time t is correlated to an observation at time $t-k$. The autocorrelogram can indicate potential problems in the dataset and will help the future stage of the modeling process. For financial returns, we expect to find low values of positive or negative correlations. In other words, past financial returns have very low explanatory power over future returns. In Figure 3, we can see that, as expected, the log return series of Ibovespa presents small absolute values of autocorrelation for lags up to ten.

Testing for ARCH effects

Before we estimate a GARCH model, we need to make sure the effect exists in the dataset. For that, we use the Lagrange multiplier (LM) test for ARCH effects (Engle, 1982). It works by regressing the squared errors on its lags and testing the hypothesis that all coefficients of the lagged regression are equal to zero.

The test takes as input a time series of returns and a given lag. With both information, it tests the null hypothesis that there are no ARCH effects in the data. As a user, we need to pay attention to the p-value resulting from the test, which will indicate the likelihood of no ARCH effects in the data, that is, the coefficients in the test are close to zero. The lower the p-value, the higher the chances of finding the ARCH effect. In order to execute this procedure, the reader should run the code in file *03-Do_ARCH_Test.R*. In Table 3, we provide the result of the ARCH LM test for lags up to five, which, in our experience, is more than enough to capture the volatility memory in daily returns. The use of higher lags is possible, but not recommended as the ARCH effect can usually be detected in the first or second lag, motivating the use of a conditional volatility model.

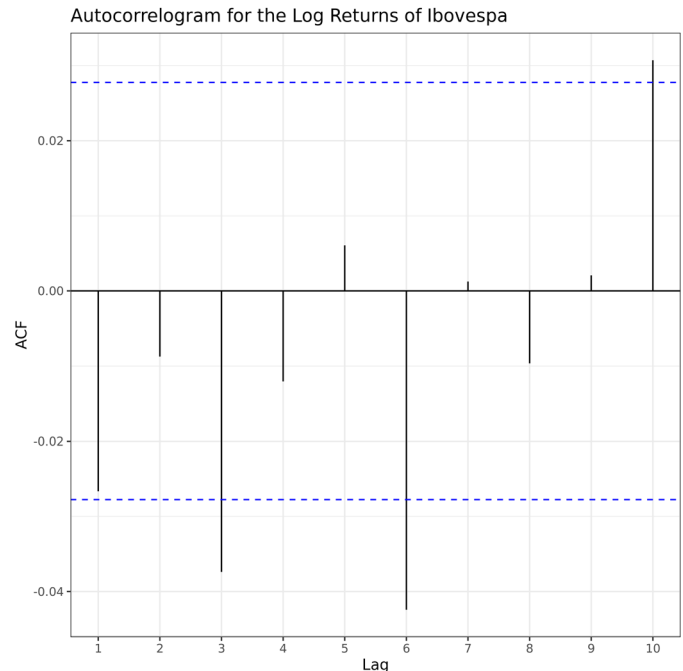


Figure 3. Autocorrelation of log returns of Ibovespa.

The autocorrelation at lag k (x-axis) shows the linear correlation between log returns at time t and $t-k$ for the Ibovespa index. The blue dashed lines represent numerical limits for the statistical significance of correlation. The autocorrelogram can be replicated using *Rscript02-Do_Descriptive_Figures.R*.

Table 3. Results of ARCH LM test for Ibovespa.

Lag	LMStatistic	pvalue
1	597,2498568	6,6371E-132
2	1234,384594	9,053E-269
3	1280,125377	3,0214E-277
4	1293,000984	1,0956E-278
5	1326,368275	1,2376E-284

Note. Table 3 reports the results of the ARCH LM test applied to the log returns of Ibovespa. Each lag has a corresponding LM statistic and associated p-value. The results can be replicated with R script *03-Do_ARCH_Test.R*.

Estimating a GARCH model

The first step in estimating a GARCH model is identifying the model, that is, to define the number of used lags in each part, the variance equation, and distribution parameters. For simplicity, we will estimate three different versions of a GARCH model, each one with a different volatility formula, but same number of lags and distribution assumption. We only leave a constant parameter in the mean equation (no ARMA coefficients), and one lag for each term of the volatility model. In the next section, we will go deeper into this topic and let the data select the best model by using goodness-of-fit indicators.

Next, Table 4, we present the estimated results and statistics generated by *04-Estimate_Simple_Garch_Model.R*.

Table 4. Estimation results of GARCH models.

	Model 1	Model 2	Model 3
mu	0.001** (0.000)	0.000 (0.000)	0.000 (0.000)
omega	0.000*** (0.000)	-0.234*** (0.009)	0.000*** (0.000)
alpha1	0.081*** (0.004)	-0.081*** (0.008)	0.021*** (0.003)
beta1	0.894*** (0.006)	0.971*** (0.001)	0.899*** (0.005)
gamma1		0.145*** (0.012)	0.102*** (0.011)
Variance Model	sGARCH	eGARCH	gjrGARCH
Distribution	norm	norm	norm
Log likelihood	13494.206	13528.679	13532.310
AIC	-5.412	-5.426	-5.427
BIC	-5.407	-5.419	-5.421

Note. ***p < 0.001; **p < 0.01; *p < 0.05.

Table 4 presents the results of the estimation of three different GARCH models to the log returns of Ibovespa. Each model has different variance equation but the same assumed distribution of residuals. Likewise, all lags in variance equation are the same. The results were found executing script named *04-Estimate_Garch_Model.R*.

The first step after estimating a GARCH model is looking at the significance of its parameters. In Table 4, model 1, we see that all coefficients are statistically significant at the 5% level (see asterisks next to the parameter's values). In the mean equation, the value of the intercept, *mu*, is positive. This implies that, as expected, Ibovespa is likely to have a positive return and increase its value in the long run.

In the variance equation, we must pay attention to the value of the *alpha1* and *beta1* in the simplest case, model 1. We expect them to be positive and their sum should not be more than one. In our case, their sum equals to 0.975. If, by any chance, this is not the case and the sum of ARCH and GARCH parameters is higher than one, it is very likely that the estimation of the model has failed. If the sum of ARCH and GARCH parameters is higher than one, the conditional probability can increase indefinitely and the unconditional volatility becomes negative, which breaks its positive condition.

As for the asymmetric models EGARCH and GJR-GARCH, models 2 and 3 in Table 4, we see that the *gamma1* coefficient is positive and statistically significant, clearly showing how the volatility reacts differently to bad news

with respect to good news. Thus, when bad news hit the market and returns are negative, volatility increases strongly.

The last three statistics, log-likelihood, Akaike information criteria (AIC), and Bayesian information criterion (BIC), are related to the estimation of the model. AIC and BIC are measures of goodness of fit (Kuha, 2004), and can be used to select an optimal lag, as we will explain in the next section of the tutorial. The log-likelihood indicates the final value of the log-likelihood after being maximized in search of all model's parameters (Meng & Rubin, 1992).

The high and significant values of the volatility parameters *alpha1* and *beta1* in all models reflect one of the characteristics of the Brazilian economic system, its high degree of uncertainty and shock propagation. When an unexpected external shock, such as a worldwide infectious disease, hits the Brazilian economy, its effects tend to last longer than in more stable countries.

Finding the best GARCH specification

In practice, for every set of data we can find the most appropriate GARCH model by comparing measures of goodness of fit. Thus, instead of manually choosing a GARCH type of model, its lags and distribution assumption, we let the data speak for itself. In practice, performing an automatic search for parameters is good research policy as it removes potential bias from the researcher. As a non-exhaustive list of studies comparing different GARCH models, we suggest Hansen and Lunde (2005) and Katsiampa (2017).

The most commonly used indicators for selecting models are the AIC and BIC, calculated using the following formulas (Tsay, 2005):

$$AIC = -2(\loglik) + 2K \quad (6)$$

$$BIC = -2(\loglik) + K \cdot \ln(N) \quad (7)$$

where *loglik* is the resulting log-likelihood of the estimated model, *K* is the number of estimated parameters, and *N* is the number of observations. The rule is: the better the model, the lower the value of AIC or BIC. One must, however, choose which criteria to use. The difference from one to the other is how they penalize the number of coefficients in the model. From Equations (6) and (7), we see that AIC is more flexible than BIC and tends to pick models with a large number of parameters. On the other hand, the BIC penalizes extra coefficients and additional observations, resulting in the choice of parsimonious models with a relatively small number of coefficients (Burnham & Anderson, 2004).

Next, Figure 4, we show the result of AIC and BIC values for different GARCH models estimated from the

data. The specification of the model is presented in the vertical axis. Do notice that the results in Figure 4 report a small number of combinations of lags and distribution parameters. We do so for simplicity and to keep the number of results manageable. In the R code, script *05-Find_Best_Garch_Model.R*, the user can choose the maximum lag for each parameter and the range of tested distributions.

From Figure 4, we see the best model as an ARMA(0,0)-eGARCH(2,1) specification with the

Student distribution. Comparing panels AIC and BIC, do notice a staircase pattern for the BIC panel, which is explained by the penalty on the extra parameters (see Equation (6)). In the AIC panel, we see the opposite effect: models with higher number of parameters tend to present better fit and lower value of AIC. This is explained by the fact that the AIC penalizes extra coefficients more softly than the BIC criteria, especially in the case of using a large number of observations in the estimation of the model.

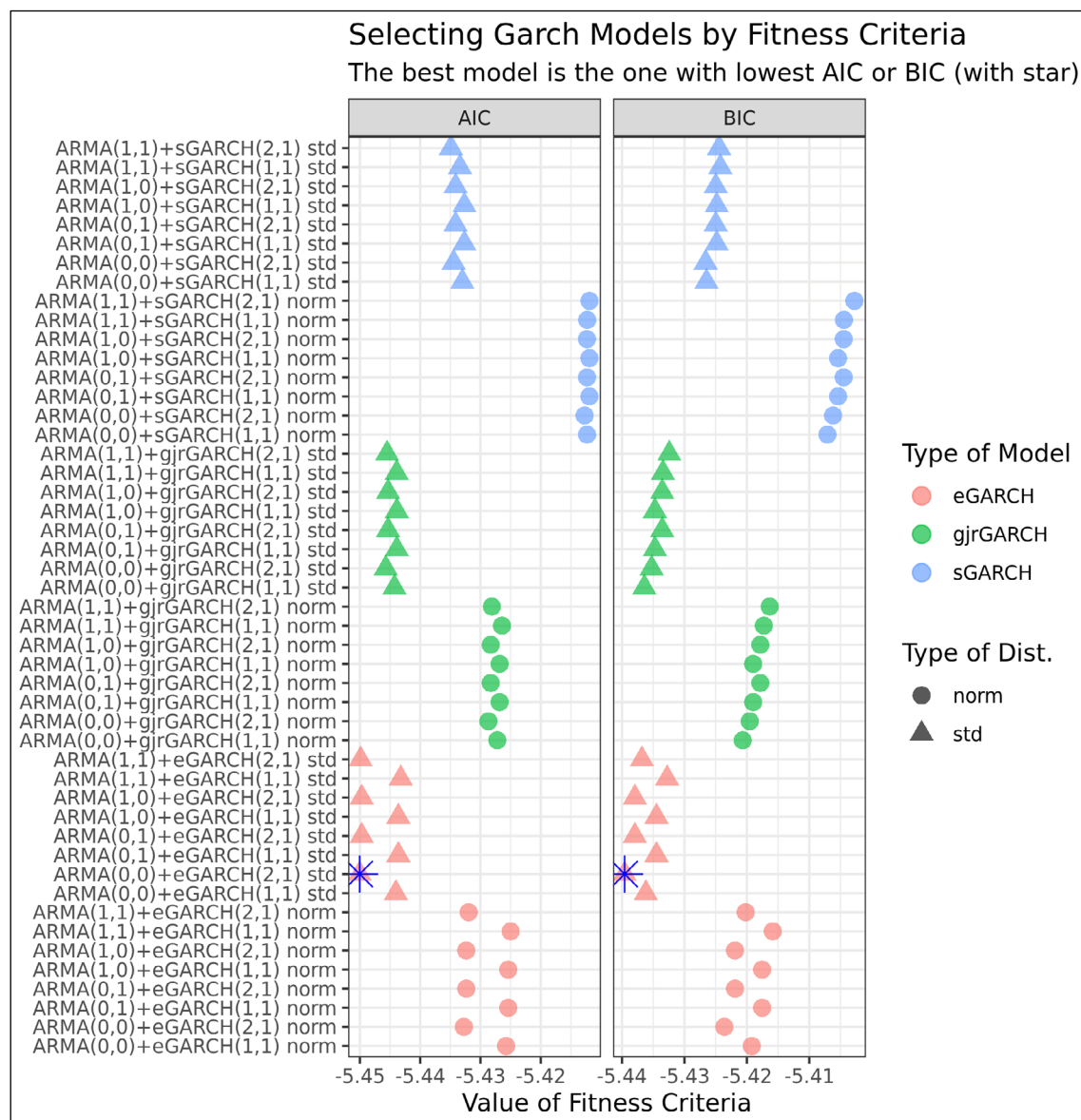


Figure 4. Selecting GARCH models using BIC and AIC.

Figure 4 reports AIC (left panel) and BIC (right panel) values for different ARMA-GARCH models, changing lags, variance equation, and distribution parameters. The best model is the one with the lowest value of AIC and BIC, highlighted by the blue star. All results of the figure can be replicated using R script *05-Find_Best_Garch_Model.R*.

Another interesting result is the greater degree of fitness for the Student distribution — triangles in Figure 4. For all different lags and variance formulas, the case with the Student distribution always presented lower value of AIC or BIC. Even more impressive is that, for all variance models, the AIC/BIC difference is much higher for changing distribution assumption than changing lags of the mean and variance equation. Thus, we find evidence that the choice of distribution assumption is very important for selecting a GARCH model using AIC and BIC.

In general, it is best advised to keep GARCH models simple and parsimonious. The benefits come from fast estimations and better volatility forecasts (Hansen & Lunde, 2005). With that in mind, we chose to use the BIC criteria and select the ARMA(0,0)-eGARCH(2,1) model for the next section of the article.

Simulating a GARCH model

Now that we have the best GARCH specification with parameters estimated from the data, we can use it to simulate the future time series of returns and possible paths for the Ibovespa index in the upcoming years. This is accomplished by script *06-Simulate_Garch_Model.R*.

Simulation with GARCH models works by sequentially inputting the first value of returns in a pre-existing model specification and drawing samples from the distribution of residuals. By doing so, we can build a time series of returns of any length, and the simulated series will have the same properties as the underlying model. This is very convenient as we can project the future for any time frame.

In this application, we are interested in simulating many time series and future paths for Ibovespa. Once that is complete, we will use the simulation paths to better understand the likelihood of the index crossing its historical peak once again. Next, in Figure 5, we show the results for 5,000 simulations.

In Figure 5, we restrict the data just after 2015 in order to reduce the scale in the y-axis. The historical peak of Ibovespa was reached in the beginning of 2020, with the closing value of 119,528. The last day of the sample of real prices is 06/15/2020, which defines the starting point of the simulation. All gray lines are different and independent price simulations using returns of the same ARMA(0,0)-eGARCH(2,1) model estimated in the previous section.

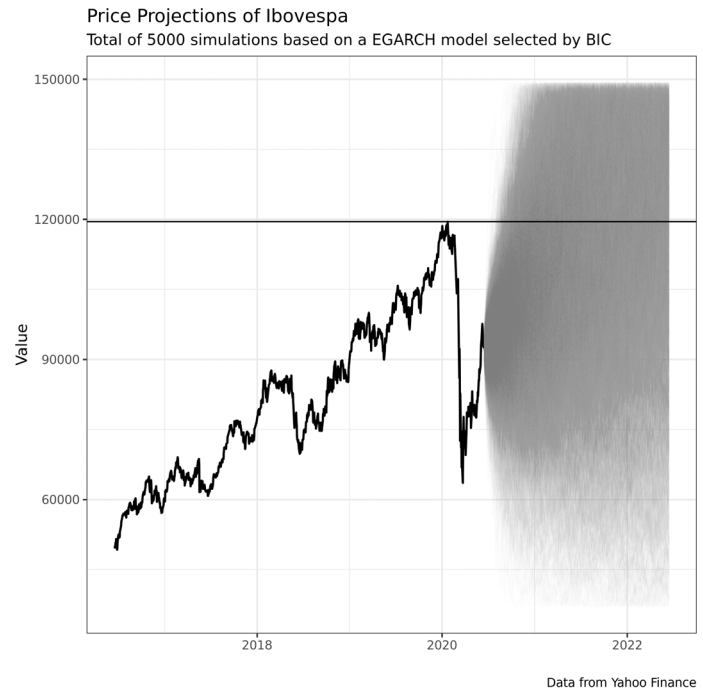


Figure 5. Simulating the Ibovespa index (2020-2025).

Figure 5 shows the simulation results using a GARCH model. After estimating the model selected by BIC, we simulate its forward log returns and use them to construct projections of the future value of the index. The solid black line is the index itself, while the gray lines are the different paths of the simulation. Figure 5 can be replicated with script *06-Simulate_Garch_Model.R*.

The first striking result of the simulation is the upward pattern. In the long run, prices of financial indexes tend to increase in value, and the estimated model was able to capture such an effect. In some unusual cases, the index reached the value of more than 400,000 points. While the simulated index value may drop in the short run, its chances of passing the historical peak increase with time. From the plot, we see the first simulated price that crosses the peak is just a couple months from June 2020. While very unlikely, our simulation shows that it can happen under some rare conditions.

For a better presentation of the results and accuracy, next, Figure 6, we show the actual probabilities of the future values of Ibovespa reaching the peak. The values in Figure 5 are calculated in the following way: for each point in time across different simulations, we find the proportion of cases where the peak value of 119,528 points was crossed. We repeat the calculations for all future dates, resulting in a vector of probabilities presented in Figure 5.

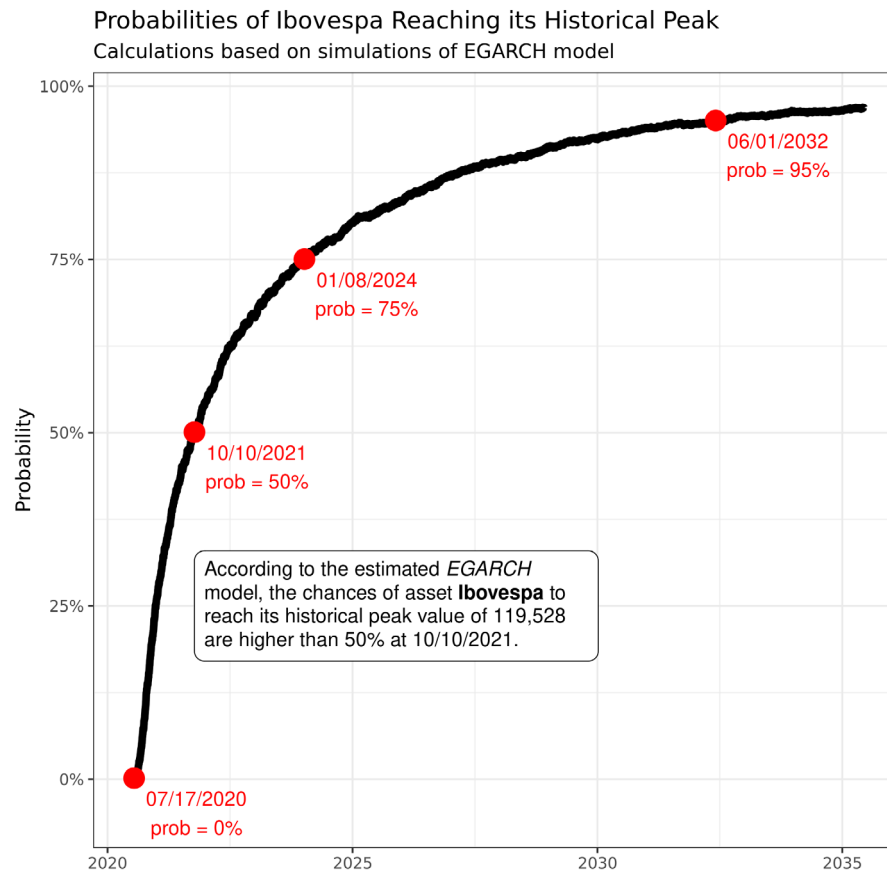


Figure 6. Probabilities of Ibovespa reaching its peak (2020-2035).

Figure 6 shows the probabilities of Ibovespa reaching its peak value once again, calculated using the simulated prices from the GARCH model. The x-axis represents the forward time period, while the y-axis is the probability itself. We calculate the probability as the proportion of times where the forward simulated index value is higher than the historical peak value of 119,528.

First, as expected, the probabilities increase with time. The first date, 07/17/2020, represents the first probability larger than 0.1%. This means that, according to the model, the chances of Ibovespa passing its peak before the upcoming month after 06/15/2020 is almost nil. The odds start to shift after 10/10/2021, where the probabilities of crossing the peak are higher than 50%. The chances keep increasing as time evolves, reaching a 95% event probability at 06/01/2032, approximately twelve years from 06/15/2020, the last day in the sample of real prices.

One of the messages from this empirical exercise is the destructive nature of a financial crisis. Prices of equity contracts dropped very fast, activating B3's circuit break² many times within the same week of March 2020. According to our GARCH model, the lost valuation due to the crisis will only be reached with a reasonable level of certainty after approximately three and a half years of trading, in January 2024.

Back to our first question: 'given the most recent financial crisis of 2020, when will Ibovespa reach its historical peak once again?', we believe the estimated GARCH model and subsequent simulation exercise provide an answer to the question. However, as with any simulation exercise, we have no guarantee that our answer is the best answer, and whether the calculated odds are realistic, especially for such a long time horizon of twelve years. Our objective here is to provide a tutorial of GARCH models in R. We feel that the answer to the question itself could be immensely improved with new models and a different research setup.

Additionally, it is important to emphasize that different GARCH models can yield different simulation results and different probabilities in our study. As an exercise, we motivate readers to use the provided R code to increase the variety of GARCH models and check the new results.

CONCLUSIONS

Volatility models and ARCH/GARCH specifications are one of the main innovations in financial modeling in the last decades, being used extensively in the industry and academic research. In this tutorial paper, we present a brief introduction to the motivation and theory behind the ARCH/GARCH family of models, with an example of empirical application for the Brazilian equity market. Based on a GARCH model and taking into consideration the recent COVID-19 crisis, we investigated how much time it would take for the index to reach its peak value once again. Our GARCH model forecasts that it will more likely than not reach its peak in about one year and four months.

But as a word of caution, we must be honest about the restrictions of our econometric study. The GARCH model is a limited representation of financial returns and no model can perfectly grasp the market participants' state of mind. Reproducing a coined phrase in statistics: 'all models are wrong, but some are useful.' Despite its shortcomings, the GARCH model has its merits by being able to provide a ballpark answer to our question, meaning that we can expect

that the Ibovespa market will recover to the COVID-19 crisis in the near future.

All the code and data used in this study is available on the internet. We motivate the readers to run the scripts and reproduce all results in their own computer. Going further, we also suggest the reader to change the financial data in script *01-Get_Index_Data.R* and reproduce the results for other market index such as SP500 (USA), FTSE (UK), or any other asset, including individual stocks, available in Yahoo Finance.

However, we must point out that only a portion of the topic of volatility modeling was covered here. The idea in this tutorial paper was to introduce the reader to the topic of volatility modeling and provide material for the reproduction of an empirical example in R. We left out many other topics such as: Bayesian estimation, multivariate GARCH, among others.

We hope this material will serve as a starting point for many students that are learning volatility modeling and financial econometrics. By combining text with actual R code, readers will be able to understand how each content was produced and, more importantly, use the code as reference for other volatility studies.

ENDNOTES

1. It is important to note that Robert Engle received the Nobel Prize for the creation of the ARCH models. Subsequently, [Bollerslev \(1986\)](#) expanded his work by developing a methodology that makes the conditional variance in these models an ARMA process. Thus, he created the GARCH models, the subject of this research.
2. Circuit break is an internal mechanism that halts trading for 30 minutes once the Ibovespa index reaches a 10% negative return within a day.

REFERENCES

- Aas, K. (2004). *To log or not to log: The distribution of asset returns* (Technical Report SAMBA n° 03/04), Oslo, Norway, Norwegian Computing Center, Applied Research and Development. Retrieved from <https://www.nr.no/files/samba/bff/SAMBA0304.pdf>
- Aït-Sahalia, Y., Mykland, P. A., & Zhang, L. (2011). Ultra high frequency volatility estimation with dependent microstructure noise. *Journal of Econometrics*, 160(1), 160-175. <https://doi.org/10.1016/j.jeconom.2010.03.028>
- Ardia, D., Bluteau, K., Boudt, K., Catania, L., & Trottier, D.-A. (2019). Markov-switching GARCH models in R: The MSGARCH package. *Journal of Statistical Software*, 91(4), 1-38. <http://dx.doi.org/10.18637/jss.v091.i04>
- Bauwens, L., Laurent, S., & Rombouts, J. V. K. (2006). Multivariate GARCH models: A survey. *Journal of Applied Econometrics*, 21(1), 79-109. <https://doi.org/10.1002/jae.842>
- Bernardino, W., Brito, L., Ospina, R., & Melo, S. (2018). A GARCH-VaR investigation on the Brazilian sectoral stock indices. *Brazilian Review of Finance*, 16(4), 573-610. <http://dx.doi.org/10.12660/rbfin.v16n4.2018.74676>
- Black, F. (1976, August). Studies of stock market volatility changes. *Proceedings of the 1976 Meeting of the Business and Economic Statistics Section, American Statistical Association*, Washington, DC, USA.
- Bollerslev, T., Engle, R. F., & Nelson, D. B. (1994). Arch models. *Handbook of Econometrics*, 4, 2959-3038.

- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307-327. [https://doi.org/10.1016/0304-4076\(86\)90063-1](https://doi.org/10.1016/0304-4076(86)90063-1)
- Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control*. San Francisco, CA: Holden Day.
- Brockwell, P. J., & Davis, R. A. (2016). *Introduction to time series and forecasting*. New York, NY: Springer International Publishing
- Bueno, R. D. S. (2011). *Econometria de séries temporais*. São Paulo, SP: Cengage Learning
- Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33(2), 261-304. <https://doi.org/10.1177%2F0049124104268644>
- Byström, H. N. E. (2004). Orthogonal GARCH and covariance matrix forecasting: The Nordic stock markets during the Asian financial crisis 1997-1998. *The European Journal of Finance*, 10(1), 44-67. <https://doi.org/10.1080/1351847032000061379>
- Caldeira, J. F., Moura, G. V., Perlin, M. S., & Santos, A. A. P. (2017). Portfolio management using realized covariances: Evidence from Brazil. *EconomiA*, 18(3), 328-343. <https://doi.org/10.1016/j.econ.2017.04.002>
- Carson, J. M., Elyasiani, E., & Mansur, I. (2008). Market risk, interest rate risk, and interdependencies in insurer stock returns: A system-GARCH model. *Journal of Risk and Insurance*, 75(4), 873-891. <https://doi.org/10.1111/j.1539-6975.2008.00289.x>
- Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1(2), 223-236. <https://doi.org/10.1080/713665670>
- Costa, C. H., Porto Junior, S. S., & Menezes, G. R. (2018). Um estudo empírico da dinâmica da correlação do retorno das ações do Brasil. *Brazilian Review of Finance*, 16(4), 635-667. <http://dx.doi.org/10.12660/rbfin.v16n4.2018.72142>
- De Goeij, P., & Marquering, W. (2004). Modeling the conditional covariance between stock and bond returns: A multivariate GARCH approach. *Journal of Financial Econometrics*, 2(4), 531-564. <https://doi.org/10.1093/jjfinec/nbh021>
- Deng, L., Ma, C., & Yang, W. (2011). Portfolio optimization via pair copula-GARCH-EVT-CVaR model. *Systems Engineering Procedia*, 2, 171-181. <https://doi.org/10.1016/j.sepro.2011.10.020>
- Engle, R. F., Focardi, S. M., & Fabozzi, F. J. (2012). ARCH/GARCH models in applied financial econometrics. In F. J. Fabozzi (Ed.), *Encyclopedia of financial models*. Hoboken, NJ: John Wiley & Sons
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987-1007. <https://doi.org/10.2307/1912773>
- Engle, R. (2001). GARCH 101: The use of ARCH/GARCH models in applied econometrics. *Journal of Economic Perspectives*, 15(4), 157-168. <https://doi.org/10.1257/jep.15.4.157>
- Ferreira, P. G. C. (2018). *Análise de séries temporais em R: Curso introdutório*. São Paulo, SP: GEN Atlas
- Fioruci, J. A., Ehlers, R. S., & Andrade Filho, M. G. (2014). Bayesian multivariate GARCH models with dynamic correlations and asymmetric error distributions. *Journal of Applied Statistics*, 41(2), 320-331. <https://doi.org/10.1080/02664763.2013.839635>
- Fisher, I. (1896). *Appreciation and interest: A study of the influence of monetary appreciation and depreciation on the rate of interest with applications to the bimetallic controversy and the theory of interest*. New York: Macmillan Company.
- Francq, C., & Zakoian, J. M. (2019). *GARCH models: Structure, statistical inference and financial applications*. Hoboken, NJ: John Wiley & Sons.
- Ghalanos, A. (2020). *Introduction to the rugarch package* (version 1.3-8). Retrieved from <http://cran.r-project.org/web/packages/rugarch>
- Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, 48(5), 1779-1801. <https://doi.org/10.1111/j.1540-6261.1993.tb05128.x>
- Hamilton, J. D. (1994). *Time series analysis*. Princeton, NJ: Princeton University Press.
- Hansen, P. R., & Lunde, A. (2005). A forecast comparison of volatility models: Does anything beat a GARCH (1, 1)? *Journal of Applied Econometrics*, 20(7), 873-889. <https://doi.org/10.1002/jae.800>
- Härdle, W. K., Chen, C. Y. H., & Overbeck, L. (2017). *Applied quantitative finance*. Heidelberg, Germany: Springer-Verlag
- Katsiampa, P. (2017). Volatility estimation for bitcoin: A comparison of GARCH models. *Economics Letters*, 158(C), 3-6. <https://doi.org/10.1016/j.econlet.2017.06.023>
- Kuha, J. (2004). AIC and BIC: Comparisons of assumptions and performance. *Sociological Methods & Research*, 33(2), 188-229. <https://doi.org/10.1177%2F0049124103262065>
- Lobão, J., & Fernandes, J. (2018). Psychological barriers in single stock prices: Evidence from three emerging markets. *Revista Brasileira de Gestão de Negócios*, 20(2), 248-272. <https://doi.org/10.7819/rbgn.v20i2.3049>
- Mastella, M., & Coster, R. (2014). O impacto da crise de 2008 na estrutura temporal de correlação condicional da BM&FBovespa. *Revista Brasileira de Gestão de Negócios*, 16(50), 110-123. <https://doi.org/10.7819/rbgn.v16i50.1534>
- Meng, X. L., & Rubin, D. B. (1992). Performing likelihood ratio tests with multiply-imputed data sets. *Biometrika*, 79(1), 103-111. <https://doi.org/10.2307/2337151>
- Morettin, P. A., & Toloi, C. (2006). *Análise de séries temporais*. São Paulo, SP: Blucher
- Morettin, P. A. (2017). *Econometria financeira: Um curso de séries temporais financeiras*. São Paulo, SP: Blucher


- Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2), 347-370. <https://doi.org/10.2307/2938260>
- Omran, M. F., & McKenzie, E. (2000). Heteroscedasticity in stock returns data revisited: Volume versus GARCH effects. *Applied Financial Economics*, 10(5), 553-560. <https://doi.org/10.1080/096031000416433>
- Perlin, M., Mastella, M., Vancin, D., & Ramos, H. (2020). Replication data for: A garch tutorial with R. *Harvard Dataverse*, v1. <https://doi.org/10.7910/DVN/C4WHUJ>
- Perlin, M. S. (2018) Processamento e análise de dados financeiros e econômicos com o R (2nd ed.). Porto Alegre, RS: Author
- Quigley, L., & Ramsey, D. (2008). *Statistical analysis of the log returns of financial assets* (Bachelor dissertation). University of Limerick, Ireland. Retrieved from <https://www.uni-muenster.de/Stochastik/paulsen/Abschlussarbeiten/Diplomarbeiten/Quigley.pdf>
- R Core Team (2017). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.r-project.org/>
- Ruppert, D., & Matteson, D. S. (2015). GARCH models. In D. Ruppert, D. S. Matteson (Eds.), *Statistics and data analysis for financial engineering* (pp. 405-452). New York, NY: Springer.
- Tsay, R. S. (2005). *Analysis of financial time series* (3rd ed.). Hoboken, NJ: John Wiley & Sons.
- Varga-Haszonits, I., & Kondor, I. (2007). Noise sensitivity of portfolio selection in constant conditional correlation GARCH models. *Physica A: Statistical Mechanics and its Applications*, 385(1), 307-318. <https://doi.org/10.1016/j.physa.2007.06.017>
- Wuertz, D., Setz, T., Chalabi, Y., Boudt, C., Chausse, P., & Miklovac, M. (2020). *fGarch: Rmetrics - Autoregressive conditional heteroskedastic modelling* [R package version 3042.83.2]. Retrieved from <https://CRAN.R-project.org/package=fGarch>

Authorship

Marcelo Scherer Perlin*

Universidade Federal do Rio Grande do Sul, Escola de Administração.
Rua Washington Luiz, nº 855, Centro, 90040-060, Porto Alegre, RS, Brazil.


E-mail address: marceloperlin@gmail.com

 <https://orcid.org/0000-0002-9839-4268>

Mauro Mastella

Universidade Federal de Ciências da Saúde de Porto Alegre.
Rua Sarmiento Leite, nº 245, Centro Histórico, 90050-170, Porto Alegre, RS, Brazil.


E-mail address: mmastella@gmail.com

 <https://orcid.org/0000-0002-7163-9448>

Daniel Francisco Vancin

Universidade do Vale do Rio dos Sinos, Programa de Pós-graduação em Ciências Contábeis.
Av. Dr. Nilo Peçanha, nº 1600, Boa Vista, 91331-002, Porto Alegre, RS, Brazil.

E-mail address: daniel_vancin@hotmail.com

 <https://orcid.org/0000-0001-6303-0555>

Henrique Pinto Ramos

Universidade Federal do Rio Grande do Sul, Escola de Administração.
Rua Washington Luiz, nº 855, Centro, 90040-060, Porto Alegre, RS, Brazil.

E-mail address: hpramos4@gmail.com

 <https://orcid.org/0000-0002-7998-7033>

* Corresponding Author

Authors' Contributions

1st author: project administration (lead); writing – original draft (lead); writing – review & editing (lead); software (lead); data curation (lead); formal analysis (lead); investigation (equal); visualization (equal); validation (equal); conceptualization (equal); methodology (equal); resources (equal).

2nd author: visualization (equal); validation (equal); methodology (equal); investigation (equal), resources (equal), writing – original draft (supporting); writing – review & editing (supporting); data curation (supporting); formal analysis (supporting).

3rd author: visualization (equal); validation (equal); methodology (equal); investigation (equal), resources (equal), writing – original draft (supporting); writing – review & editing (supporting); data curation (supporting); formal analysis (supporting).

4th author: visualization (equal); validation (equal); methodology (equal); investigation (equal), resources (equal), writing – original draft (supporting); writing – review & editing (supporting); data curation (supporting); formal analysis (supporting).

Funding

There are no funders to report for this article.

Conflict of Interests

The authors have stated that there is no conflict of interest.

Copyrights

RAC owns the copyright to this content.

Plagiarism Check

The RAC maintains the practice of submitting all documents approved for publication to the plagiarism check, using specific tools, e.g.: iThenticate.

Peer Review Method

This content was evaluated using the double-blind peer review process. The disclosure of the reviewers' information

on the first page is made only after concluding the evaluation process, and with the voluntary consent of the respective reviewers.

Data Availability

All data and materials were made publicly available through the Harvard Dataverse platform and can be accessed at:



Perlin, M., Mastella, M., Vancin, D., & Ramos, H. (2020). Replication data for: A garch tutorial with R. *Harvard Dataverse*, v1. <https://doi.org/10.7910/DVN/C4WHUJ>