UNIVERSITY OF LAUSANNE
FACULTY OF BUSINESS AND ECONOMICS

Assignment 3

Group members: *Antoine Tissot, Florian Perusset, Florian Vogt, Liam Svoboda, Sebastien Gorgoni*

Course: *Empirical Methods in Finance* – Professor: *Dr. Florian Ielpo*
Due date: *June 6th 2021*

Contents

1. Introduction

This report consists of two different parts. In the first part, we discuss three different types of GARCH models (GARCH, EGARCH and GARCH-GJR) to estimate the time varying volatility of the log-returns of three different time series (S&P 500, BCOM Commodity index and Apple).
GARCH models allow us to perform a time-varying volatility estimation, as well as to take into account persistence in volatility. These are particularly interesting features of volatility that are present in empirical data. In addition to the standard GARCH model, we are using two extension, EGARCH and GARCH-GJR which allow us to take into account the asymmetry of volatility.

In the second part, we are using the famous Heston and Nandi model to price call options on the S&P 500. This model allows us to take into account time varying volatility in the pricing of the calls.
Using a time varying volatility model is important when it comes to option pricing because we know that empirically, volatility is indeed time varying. Thus, this kind of models offers a more precise estimation of the price of options than standards model like Black-Scholes or the Binomial Model who assume constant volatility.

2. Presentation of the time series

In the first part of the report, we focus on an equity index, a commodity index, and a single stock. For all of them, we collect daily data of the last 10 years. In the second part of the assignment, we focus on the S&P500 and expand the time span as we collect the price series (daily) from 1989 until 2021. Below, we provide some explanations of the different series.

For the stock index, we have decided to refer to the S&P500. It is a stock market index composed of the 500 largest US publicly-traded companies. It uses the capitalization weighting method[1]:

- Inclusion Criteria:
  1. Should be a US company
  2. Market Cap. $\geq$ \$ 8.2 Billion
  3. High liquidity of its shares
  4. Part of its outstanding shares available to public $\geq$ 50
  5. Last quarter's earnings must be positive
  6. Sum of last 4 quarter's earnings must be positive

[1] `https://corporatefinanceinstitute.com/resources/knowledge/trading-investing/sp-500-index/`, accessed 05.06.21

Hence, companies which stop conforming to these criteria become excluded from the index. There is however some leeway and selection by a committee in order to avoid turbulence caused by high turnover within the index

- Components are weighted based on free-float market capitalization

$$W_i = \frac{P_i \times Q_i}{\sum Q} \tag{1}$$

Where:

1. W_i is the weight of company i in the S&P500 index.
2. P_i is the stock price of company i.
3. Q_i is the number of free-floating shares of company i.

For the Commodity index, we decided to consider the Bloomberg Commodity Index (BCOM). As explained on Bloomberg's website[2], it "is calculated on an excess return basis and reflects commodity futures price movements. The index rebalances annually weighted 2/3 by trading volume and 1/3 by world production and weight-caps are applied at the commodity, sector and group level for diversification. Roll period typically occurs from 6th-10th business day based on the roll schedule."

Last but not least, We chose to look at Apple for the single equity as it is a main actor a the technological revolution taking place at the moment. They have definitely changed the world both in terms of technological breakthroughs and in terms of consumers habit creation. Therefore, we considered it would be legitimate to pay attention on what is going on behind the stock.

3. Part 1: GARCH models

(1) In this section, we will fit three different types of GARCH models (GARCH, EGARCH and GARCH-GJR) to each of our three series of log-returns. To be as parsimonious as possible, we decided to use only one lag and therefore fit GARCH/ EGARCH/ GARCH-GJR(1,1) models for each of the series.
From the lecture notes, recall that :

- A GARCH(1,1) model writes as :

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{2}$$

- An EGARCH(1,1) model writes as :

$$\log \sigma_t^2 = \omega + \alpha(|z_{t-1}| - \mathbb{E}|z_{t-1}|) + \gamma z_{t-1} + \beta \log \sigma_{t-1}^2 \tag{3}$$

where z_t follows a N(0,1), so that $\mathbb{E}|z_{t-1}| = \sqrt{2\pi}$.

[2] `https://www.bloomberg.com/quote/BCOM:IND`, accessed 05.06.21

– A GARCH-GJR(1,1) model writes as :

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \gamma\mathbb{1}_{t-1}^{-}\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{4}$$

where $\mathbb{1}_{t-1}^{-}$ is equal to 1 when $\epsilon_{t-1} < 0$ and 0 otherwise.

and in each case, we have:

$$r_t = \mu + \sigma_t\epsilon_t \tag{5}$$

where r_t is the return at time t of the series and μ is the mean of the series.
In each case, we use the maximum likelihood estimation to calibrate the parameters of the models to the series.
Firstly, let's focus on the log-returns of Apple. The results of the estimation of the three different models for Apple are the following :

Table 1: GARCH(1,1) - Apple

	Estimate	Std-deviation	t-stat
μ	0.0098628	0.0218109	0.452197
ω	0.000647216	0.000264841	2.44379
α	0.0298288	0.102318	0.291529
β	0.970171	0.0122228	79.3736

Table 2: EGARCH(1,1) - Apple

	Estimate	Std-deviation	t-stat
μ	0.00986281	2.24254e-05	439.806
ω	0.32697	3.00416	0.108839
α	0.0887094	0.84945	0.104432
γ	-0.0578101	1.04308	-0.0554224
β	0.904204	0.0164468	54.9774

Table 3: GARCH-GJR(1,1) - Apple

	Estimate	Std-deviation	t-stat
μ	0.00986004	0.0222716	0.442717
ω	0.000647228	0.000265571	2.43712
α	0.0188167	0.467507	0.040249
γ	0.0210117	0.636818	0.0329947
β	0.970678	0.0119056	81.5313

In each case, we can see that β is positive and close to one (even if it is a bit smaller in case of the EGARCH). In all the models, we can see that β is statistically different from 0 at 1% level (t-stats are bigger than 2.57) and positive, meaning that the lag 1 value of volatility has a positive impact on current volatility of Apple.

This result shows that volatility is a persistent phenomenon over time, meaning that if the volatility of Apple is high today, it will probably remain high tomorrow.
Concerning α, which measures in each model the effect of news on volatility, it is always positive but not statistically different from 0 in all models, meaning that a shock (positive or negative) on r_t doesn't have any impact on volatility.
In EGARCH and GARCH-GJR models, the parameter γ measures the effect on volatility of the sign of the innovations. We can see that for Apple, γ is not statistically significant, both for GARCH-GJR and EGARCH, meaning that the sign of past innovation has no influence on the volatility. In other words, a negative news has the same impact on volatility than a positive news.

Now, let's perform the same analysis for BCOM Commodity index :

Table 4: GARCH(1,1) - BCOM Commodity

	Estimate	Std-deviation	t-stat
μ	-0.00225864	5.90928e-05	-38.2219
ω	7.21075e-07	5.53438e-08	13.0289
α	0.03	0.0012729	23.5681
β	0.9	0.00401472	224.173

Table 5: EGARCH(1,1) - BCOM Commodity

	Estimate	Std-deviation	t-stat
μ	-0.00225864	4.02887e-07	-5606.13
ω	0.246173	0.00629021	39.136
α	0.0814352	0.0014762	55.1655
γ	-0.0251607	0.000790967	-31.8101
β	0.957254	0.000292187	3276.16

Table 6: GARCH-GJR(1,1) - BCOM Commodity

	Estimate	Std-deviation	t-stat
μ	-0.00225864	0.00448964	-0.503078
ω	0.000144215	7.45487e-05	1.93451
α	0.0269133	0.296465	0.0907808
γ	0.029701	0.497918	0.0596503
β	0.958232	0.021137	45.3342

The results that we obtain are a bit different compare to those that we get for Apple.
Once again, we can see that in every model, β is statistically different from 0 and positive, meaning that volatility is a persistent phenomenon. For all models, β is relatively close to one, meaning that the volatility is really persistent over time.

α is positive and statistically different from 0 in the GARCH and EGARCH case, meaning that according to these models, past innovations (positive or negative) have a positive influence on volatility.
Concerning γ, it's is not statistically different from 0 in GARCH-GJR model. However, it is negative and statistically different from 0 in the EGARCH case, meaning that according to the this model a negative news has more impact on the volatility than a positive one, which is usually what we can observe in the data.

Table 7: GARCH(1,1) - S&P 500

	Estimate	Std-deviation	t-stat
μ	0.00458084	0.000102468	44.7051
ω	1.20615e-06	5.82395e-08	20.7101
α	0.03	0.000845737	35.472
β	0.9	0.00252137	356.949

Table 8: EGARCH(1,1) - S&P 500

	Estimate	Std-deviation	t-stat
μ	0.00458084	2.46591e-06	1857.66
ω	0.484332	2.41176	0.200821
α	0.157438	0.680729	0.231278
γ	-0.118947	0.893466	-0.13313
β	0.897026	0.0200602	44.7168

Table 9: GARCH-GJR(1,1) - S&P 500

	Estimate	Std-deviation	t-stat
μ	0.00457283	0.00649774	0.703757
ω	0.000241135	0.000150335	1.60399
α	0.0373551	0.378287	0.0987479
γ	0.0699833	0.545621	0.128264
β	0.92757	0.0442012	20.9852

Once again by considering the log-returns of the S&P 500, we obtain similar results than before.
α is positive and statistically different from 0 in the GARCH model, meaning that the squared of innovations have an impact on volatility. In the two other models, α is positive but not statistically different from 0.
In each model, β is positive and statistically significant. This means that every model shows persistence in volatility.
Concerning the leverage effect, we can see that γ is not statistically different from 0, meaning that we can't conclude that there's an asymmetric effect coming from innovations.

Finally, we have seen quite consistent results through the analysis of the three different time series of log-returns using three different models to describe time varying volatility.
Indeed, for each time series and for each model, we can see that volatility is persistent over time since in every model and for every time series, β, the auto-regressive coefficient of past volatility is always statistically different from 0 and positive. This means that if the volatility is high today, it's probable that it will remain high in the next period and conversely, if volatility is low today, it's probable that it will remain low in the next period. However, the processes are covariance-stationary since $\beta < 1$, meaning that volatility is mean-reverting even though it is persistent.
One of the other interesting aspect of this exercise is to be able to assess whether the effect of innovations on volatility is symmetric or not. In a standard GARCH process, only the squared innovations have an impact on volatility (in addition to the past values of volatility), meaning that the sign of innovations doesn't matter but only their size.
However, we would expect that a negative news (negative innovation) in the return has a stronger impact on volatility due to **leverage effect**. This is why we use EGARCH and GARCH-GJR models, which take into account not only the size but also their signs.
Leverage effect means that large negative returns have a stronger impact on volatility than large positive returns. The explanation for this is that volatility is a proxy for risk and negative returns should cause a higher increase in risk than positive returns, since the majority of investors is long and therefore fears more negative than positive returns. In fact, the name leverage effect refers to the fact that when returns are negative, market leverage effect of firms tend to increase since the value of equity is decreasing, leading to a higher probability of bankruptcy, even though this causal link between negative returns and increase in the probability of bankruptcy isn't verified empirically.
Coming back to our results, we usually can't conclude that there is leverage effect in our time series since the coefficient that measures the effect of sign of innovations, γ, is not statistically different from 0 for each time series except in the case of BCOM Commoditiy when using the EGARCH model, we have that γ is negative and statically significant.
Therefore, we cannot reject the null hypothesis of symmetric effect in volatility at least in the case of Apple and S&P 500. This result is quite surprising since there is strong evidence of asymmetric effect in volatility but this empirical evidence isn't always verified in our data.

(2) In this part, we will compare the three models fitted on our data in question (1.1): GARCH, EGARCH and GARCH-GJR. To do so, we'll conduct some likelihood ratio (LR) test where possible (nested models) and some Vuong's test where LR are not possible (non-nested models) in order to conclude on the best fitting model our data. We'll be therefore able to conduct LR tests on GARCH and GARCH-GJR comparisons, while we'll have to derive Vuong's tests when EGARCH model is implied in the comparison.

Let's first see how the LR test works: it allows one to compare the goodness-of-fit between a simple model, the restricted model (GARCH), and a more complex one (EGARCH and GARCH-GJR), the unconstrained model. The aim is then to see whether the difference in likelihood when removing variables from more complex model is statistically significant or not. Therefore, if the likelihood difference between the two models is statistically significant, it would mean the more complex model better fits our data.

Here, we will therefore make one LR test comparing GARCH model (the restricted model) with the GARCH-GJR model (unconstrained model).

The test statistic of such a test is expressed as:

$$LR = -2(LL_R - LL_U) \tag{6}$$

Where LL_R is the log-likelihood of the restricted model (GARCH) and LL_U is the log-likelihood of the unconstrained model (EGARCH and GARCH-GJR). Moreover, under H_0 where the true model is the restricted one, we know that $LR \sim X_k^2$ where k is the number of constraints (in our case $k = 1$ as there is only one additional parameter in the GARCH-GJR (unconstrained model) compared the the restricted GARCH model.

When applying the LR test and computing the p-values for all our time series, we find the following results:

LR - p-values	S&P500	BCOM	Apple
GARCH VS GJR	1.0	1.0	1.0

Looking at the results above, we conclude that:

– Concerning S&P500, GARCH better fits our data compared to GARCH-GJR.
– Concerning BCOM, GARCH better fits our data compared to GARCH-GJR.
– Concerning Apple, GARCH better fits our data compared to GARCH-GJR.

In order to focus on non-nested models, we then run two Vuong tests:

– First one comparing GARCH model with the EGARCH model
– Second one comparing EGARCH model with the GARCH-GJR model

The Vuong's tests allows to compare the likelihood of two models (not necessarily nested) fitting our data and return which of those is better adapted to fit our data. Under H_0 (the two models are statistically equivalent in term of data fitting), the test statistic can be written as:

$$Z = \frac{\sum_{i=1}^{n} z_i}{\sqrt{n}\sigma(z_i)} \tag{7}$$

where $z_i = LL_1 - LL_2$ is the difference between the two model's log-likelihood for observation i.

Therefore, when computing the test statistics (Z) for all the time series, we get the following results:

Vuong test	S&P500	BCOM	Apple
GARCH;EGARCH	8.7213	1.7862	50.2787
EGARCH;GJR	-6.4037	-1.5988	-50.014

Knowing that under H_0, $Z \sim N(0,1)$, we can make the following conclusions:

- Concerning S&P500, we reject H_0 for both comparisons, meaning that GARCH and EGARCH aren't equivalent, and EGARCH and GJR aren't equivalent neither according to Vuong's test.
- Concerning BCOM, we fail to reject H_0 (5% confidence level) for both comparisons, meaning that GARCH, EGARCH and GJR are equivalent in terms of fitting according to Vuong's test.
- Concerning Apple, we reject H_0 for both comparisons, meaning that GARCH and EGARCH aren't equivalent, and EGARCH and GJR aren't equivalent neither according to Vuong's test.

We finally compute the AIC criterion for each model and each series to be able to select the best model when we reject H0 in the Vuong test. Recall that the Akaike criterion is defined as:

$$AIC = 2k - 2\ln \mathcal{L}_{\max} \tag{8}$$

where k is the number of parameters estimated and $\ln \mathcal{L}_{\max}$ is the maximized log-likelihood function.

The following table summarizes the results we have:

Model	S&P500	BCOM	Apple
GARCH(1,1)	-15480.42	-15887.98	-5128.25
EGARCH(1,1)	5415.65	1073.04	5356.67
GARCH-GJR(1,1)	-9857.00	-9865.91	-5072.22

Remembering the results of the LR and Vuong tests, and minimizing the AIC criterion, we finally choose:

- S&P500 : GARCH Model
- BCOM : GARCH Model
- Apple : GARCH Model

(3) In this part, we will firstly filter returns for each series given the model we have chosen, that is, a GARCH(1,1), and then test for Gaussianity of the residuals. To filter returns, let's recall that for all considered models:

$$r_t = \mu_t + \sigma_t \epsilon_t \tag{9}$$

Then, filtering returns from their conditional moments means making the following transformation to get filtered returns:

$$\hat{\epsilon}_t = \frac{r_t - \mu_t}{\sigma_t} \tag{10}$$

When transformation is done, we are able to test for gaussianity of filtered returns. To do so, we run a Kolmogorov and Smirnov Test for filtered returns of all series. Recall that the test statistic of a Kolmogorov & Smirnov test is calculated as follows:

$$KS = \max_{t=1,..,T} |G_T(x) - F^*(x;\theta)| \tag{11}$$

This test statistic is based on the discrepancy between the theoretical cumulative distribution function (CDF) $F^*(x;\theta)$ that prevails under the null hypothesis (in this particular case, Gaussianity) and the empirical cumulative distribution function $G_T(x)$, that is the sample CDF of the filtered returns.

Computing the KS test statistics for all our time series, we find the following results:

	S&P500	BCOM	Apple
KS Test	0.259209	0.201265	0.406922
p-value	0	0	0

We can therefore reject H_0 (that the filtered returns are Gaussian) for all considered time series since all p-values are negligible.

(4) Having demonstrated that the residuals most likely don't follow a Gaussian distribution, we attempt to fit them to 5 different distribution functions using the log-likelihood to estimate the appropriate parameters for each distribution.
The distributions and their respective parameters are the following:

– **Generalised Normal Distribution** (GenNorm):
Departs slightly from the basic Normal distribution in that a 3^{rd} parameter is used to calibrate the kurtosis of the distribution meaning that we can alter the size of the tails.

	S&P500	Apple	BCOM
Shape (β)	1.272768	1.023232	1.442427
Location (μ)	-0.676810	-0.060690	0.440389
Scale (α)	1.342447	0.087378	1.738450

– **Normal Inverse Distribution** (NormInvGauss):
Strays from the standard Normal Distribution by adding a 3^{rd} parameter to once again tailor the size of tails and a 4^{th} parameter to be able to change the symmetry of the distribution.

	S&P500	Apple	BCOM
Tail Heaviness (α)	1.792111	0.632563	2.360610
Asymmetry (β)	-0.345311	-0.028891	-0.379291
Location (μ)	-0.358689	-0.056131	0.776246
Scale (δ)	1.765726	0.095741	2.330106

– **Laplace Distribution**:
Although only being characterized with 2 parameters like the Normal Distribution, it differs on the basis that instead of emanating from the squared distance from the mean it takes its absolute distance. This leads to a kurtosis

twice as high as the normal distribution meaning that the Laplace Distribution is leptokurtic.

	S&P500	Apple	BCOM
Location (μ)	-0.678818	-0.060800	0.492469
Scale (b)	1.013287	0.084622	1.182948

– **Johnson S_U Distribution** (JohnSU):
Its probability density function makes use of the pdf of the Standard Normal Distribution applied to a heavily transformed input variable. Among these transformations, some are non-parametric but there exists 2 new parameters which are used to tweak the final distribution. The result can present skewness as well as excess kurtosis.

	S\&P500	Apple	BCOM
Shape 1 (a)	0.366629	0.048929	0.338697
Shape 2 (b)	1.807596	1.230138	1.988006
Location (μ)	-0.215384	-0.054480	0.912821
Scale (σ)	2.041595	0.104201	2.653818

– **Double Gamma Distribution** (DGamma):
Like some of the previous distribution functions, its shape is defined by 3 parameters determining the location, scale and shape of the curve.

	S&P500	Apple	BCOM
Shape (a)	1.214185	1.059147	1.328192
Location (μ)	-0.503392	-0.065386	0.324451
Scale (α)	0.844011	0.079966	0.895806

Given these parameters, we perform a KS test like in the previous section but this time $F^*(x, \theta)$ in (11) is one of the 5 aforementioned distributions instead of just the Gaussian distribution.

Table 10: KS tests for S&P500

	GenNorm	NormInvGauss	Laplace	JohnSU	DGamma
KS statistic	0.016671	0.020519	0.033106	0.021738	6.793599e-02
p-val	0.460785	0.221025	0.006574	0.169097	7.188877e-11

Table 11: KS tests for Apple

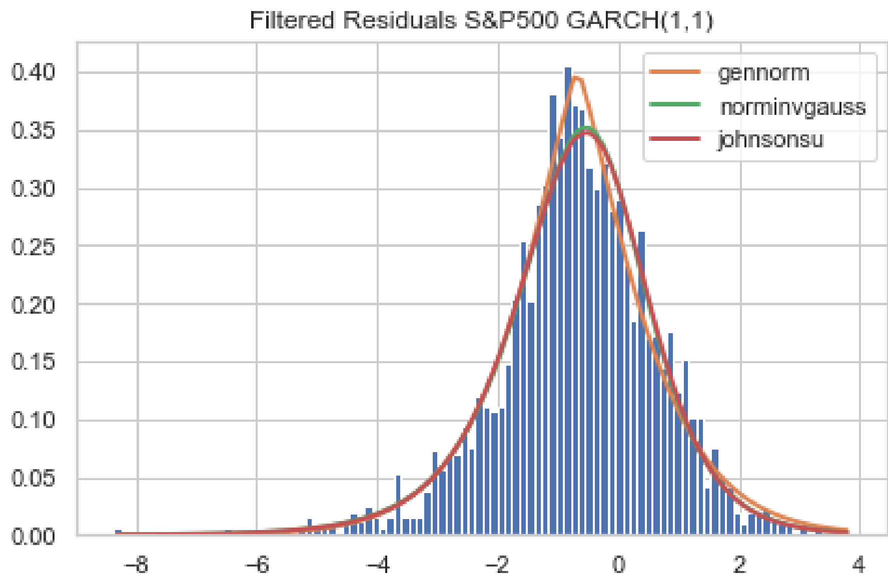
	GenNorm	NormInvGauss	Laplace	JohnSU	DGamma
KS statistic	0.025188	0.013883	0.027351	0.015158	0.039302
p-val	0.072777	0.693118	0.040265	0.583916	0.000636

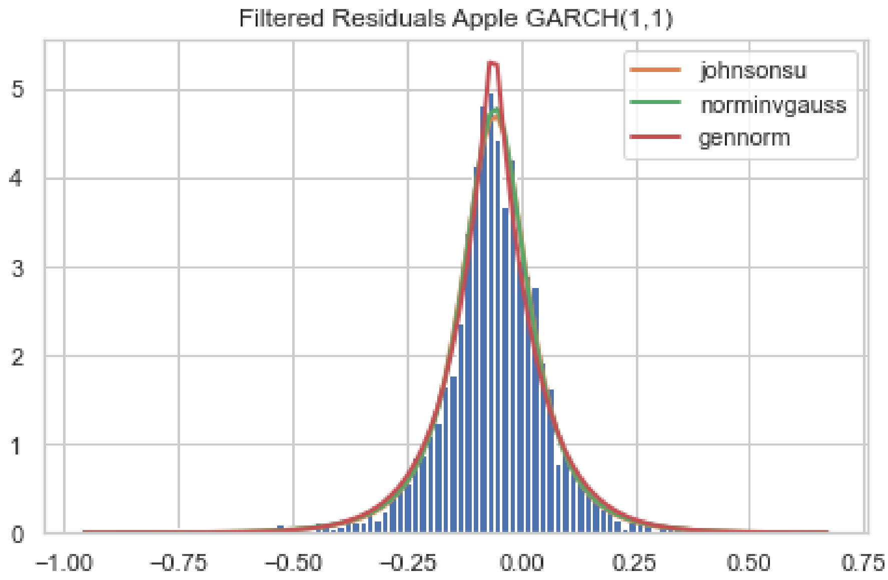
Table 12: KS tests for BCOM

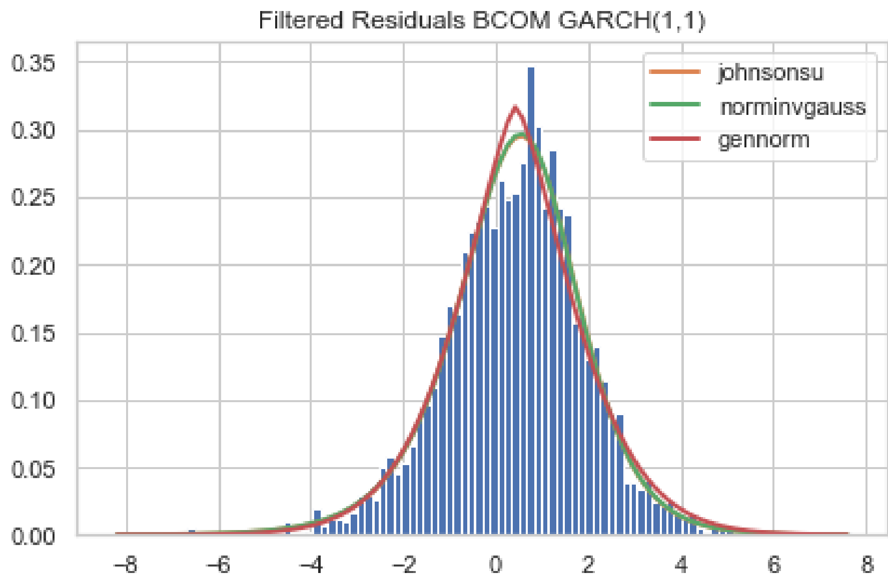
	GenNorm	NormInvGauss	Laplace	JohnSU	DGamma
KS statistic	0.023599	0.014718	0.052638	0.014715	0.049007
p-val	0.109113	0.621615	0.000001	0.621924	0.000007

Sifting through the results of tables 10 to 12, more specifically the p-values, we can conclude that the residuals of all 3 of our assets appear to be adequately described by GenNorm, NormInvGauss and JohnSU given that we do not reject the null hypothesis for the KS test at a 5% confidence level as opposed to the ones for Laplace and DGamma which yield a very low p-value making us reject the null hypothesis.

Hence, the KS tests show that residuals of our 3 time series could be described by multiple distributions from the ones we proposed leaving our options open. The following figures illustrate that the distributions at hand do seem to fit the residuals quite well.







We might however not be satisfied by this notion and want to find the best one. To answer this question, we computed 2 different information criteria, the Akaike (AIC) and the Bayesian Information Criteria (BIC) as well as the sum of squared errors (SSE) between the true distribution and the hypothetical distribution we observe. For ease of computation, we use the fitter to calculate these values since the estimated parameters with this method are very close to the ones we find with our MLE.

This leads to the following results:

Table 13: Information Criteria for S&P500 Residuals

	sumsquare_error	aic	bic
gennorm	0.038701	852.940816	-28851.609970
norminvgauss	0.046370	815.502832	-28374.062862
johnsonsu	0.047191	816.305003	-28328.441358

Table 14: Information Criteria for Apple Residuals

	sumsquare_error	aic	bic
johnsonsu	2.186391	565.654516	-18363.025981
norminvgauss	2.216782	578.792601	-18327.162477
gennorm	3.454188	643.463995	-17182.732501

Table 15: Information Criteria for BCOM Residuals

	sumsquare_error	aic	bic
johnsonsu	0.016576	937.646561	-31033.659009
norminvgauss	0.016925	943.442887	-30979.522222
gennorm	0.026634	982.176067	-29809.917869

The objective being to minimize the 3 measures, we use the SSE to adjudicate when the two information criteria are in contradiction. Hence, we arrive at the following conclusions:

S&P500 residuals favour the NormInvGauss when using the AIC and the GenNorm when looking at the BIC. We therefore decide to use the SSE to make the final judgment and settle for the GenNorm distribution.

Apple residuals favour the Johnson S_U with all 3 metrics.

BCOM residuals favour the Johnson S_U with all 3 metrics yet again.

Although we can in no way assert that our residuals definitely follow these distributions, we can at least claim that they seem to fit better than the Gaussian distribution and the 4 other distributions they were pitted against.

4. Part 2: Option Pricing Model

This section aims to implement a Heston and Nandi GARCH option pricing model, using the daily log-returns of the S&P500 from the 31st of December 2020 to the 31st of December 2020.

(1) Let S_t be the underlying asset price at time t, r the risk-free rate, $\epsilon_t \sim \mathcal{N}(0, 1)$ the standard normal random variable, h_t the conditional variance of the log-return

between $t-\Delta$ and t (known at $t-\Delta$) and λh_t the risk premium. Then the Heston-Nandi process is defined as:

$$\begin{gathered} logS_t = logS_{t-\Delta} + r + \lambda h_t + \sqrt{h_t}\epsilon_t \\ \rightarrow R_t = log\left(\frac{S_t}{S_{t-\Delta}}\right) = r + \lambda h_t + \sqrt{h_t}\epsilon_t \end{gathered} \tag{12}$$

The conditional variance h_t is defined as a GARCH(1, 1):

$$h_t = \omega + \beta h_{t-\Delta} + \alpha\left(\epsilon_{t-\Delta} - \gamma\sqrt{h_{t-\Delta}}\right)^2 \tag{13}$$

To estimate the parameters $\theta = (\alpha, \beta, \gamma, \omega, \lambda)$, we will apply the maximum likelihood estimation (MLE) method by maximizing the likelihood function:

$$\max_\theta L = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi h_t}} exp\left[-\frac{(log\left(\frac{S_t}{S_{t-\Delta}}\right) - r - \lambda h_t)^2}{2h_t}\right] \tag{14}$$

This is equivalent to maximize the log-likelihoof function:

$$\max_\theta logL = -\sum_{t=1}^{T} 0.5\left[log(2\pi h_t) + \frac{(log\left(\frac{S_t}{S_{t-\Delta}}\right) - r - \lambda h_t)^2}{h_t} +\right] \tag{15}$$

Where we included the constraint for stationarity as being $\beta + \alpha * \gamma^2 < 1$, as described on the paper of Heston and Nandi. For simplification, we assumed a low interest rate of 0.25%, representative of the current economic situtation. Consequently, after using the Nelder-Mead algorithm in our optimization problem, and using the initial parameters as being $(\omega_0, \beta_0, \alpha_0, \gamma_0, \lambda_0) = (9.765e-10, 0.90, 2.194e-06, 100.15, 10)$[3], we obtained the following estimates:

Table 16: Heston-Nandi Parameters Estimation

	Parameter	Std. Err.	T-stat
Omega	-1.314799e-08	3.000727e-07	-0.043816
Beta	8.757779e-01	6.006347e-03	145.808749
Alpha	7.259519e-06	4.184156e-07	17.350020
Gamma	8.551729e+01	6.354834e+00	13.457046
Lambda	-2.159823e+01	1.032712e+00	-20.914094

To estimate the standard error and T-stat of our estimates, we applied the following procedure. Let's consider the information matrix, defined as:

$$I(\theta) = -\frac{1}{T}E\left[\frac{\partial^2(L_T(\theta))}{\partial\theta\partial\theta'}\right] \tag{16}$$

[3] `https://quant.stackexchange.com/questions/63830/heston-nandi-garch-implementation-problem-for-python`

Now in practice, the information matrix can be computed as:

$$I(\theta) = \frac{1}{T} \nabla_\theta log L_T(\theta) \nabla_\theta log L_T(\theta)^T \tag{17}$$

Where $\nabla_\theta log L_T(\theta)$ can be considered as the first partial derivative of the log-likelihood function. To do some, we implemented a numerical approximation of the derivation for a small scalar h:

$$f'(x) = \frac{f(x + h/2) - f(x - h/2)}{h} \tag{18}$$

Thereafter, we have been able to compute the standard error of our parameters given the fact that the asymptotic distribution of the maximum likelihood estimators is the following:

$$\sqrt{T}(\hat{\theta} - \theta) \overset{d}{\sim} N(0, I(\theta)^{-1}) \tag{19}$$

Additionally, if we consider a random variable Z and s the standard deviation, then we define the test statistic as follow:

$$t = \frac{Z}{s} \tag{20}$$

Consequently, we divided our estimates by their pairwise variance, which is the element in the diagonal of $I(\theta)^{-1}$ to obtain their t-test.
From these results, we notice that α, β, γ and λ are significant to a 1% level of significant since we can reject the null-hypothesis of being non-significant. Nevertheless, ω is the only parameter which is not significant. We know that α and γ define the kurtosis and skewness respectively. As we have in our results that $\alpha \approx 0$ and $\gamma \approx 85.52$, we can notice that the distribution of our time series does not have a kurtosis but does have a skewness. From the paper of Heston-Nandi (2020), as α and β parameters approach to zero, it is equivalent to a Black Scholes model observed at discrete intervals[4]. However, as $\beta \approx 0.88$ and is significant to a $1\% level of significance$, we can conclude that our model is indeed not equivalent to the Black Scholes Model. The implication would mean that our time series is not Gaussian. From the same paper, we know that a time series is stationary if:

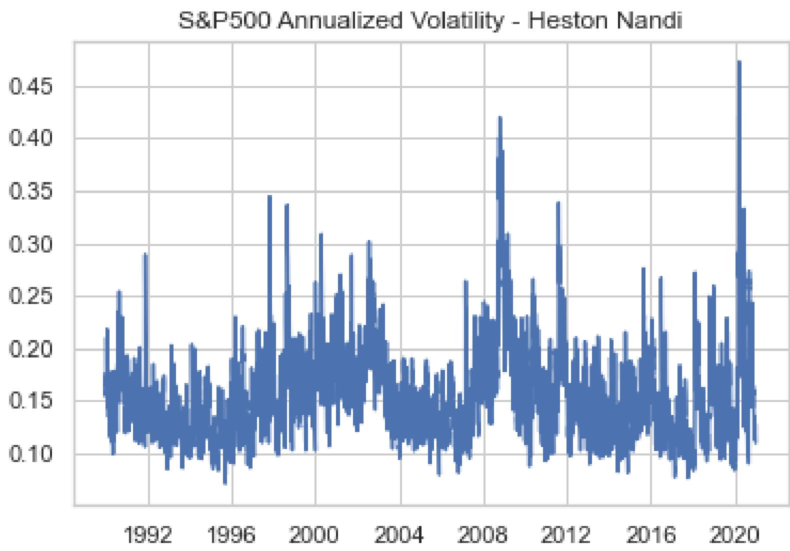
$$\beta + \alpha \times \gamma^2 < 1 \tag{21}$$

With our results, we have that $\beta + \alpha \times \gamma^2 = 0.9288 < 1$, which we can conclude that, based on the Heston-Nandi model, our time series is stationary.

Additionally, we were able to estimate the S&P500 annualized volatility using the Heston-Nandi Model:

[4] `https://www.econstor.eu/bitstream/10419/100805/1/wp1997-09.pdf`

Figure 1: SP500 Annualized Volatility using Heston-Nandi Model



We notice that the annualized volatility of the S&P500 is clearly higher in periods of crisis (e.g. DotCom Bubble, GFC, COVID crisis), highlighting the economic uncertainties arising during these unprecedented events.

(2) Now that we estimated the parameters of the Heston-Nandi Model, we will first compute the option prices on December 31st by ranging the time to maturity from 3 months, 6 months and 1 year as well as the moneyness (i.e. strike divided by the spot price) from 0.9 to 1.1 using 10,000 Monte Carlo Simulations. Using the previously defined equation **12** and **13**, we were able to simulate the returns of our underlying asset (i.e. SP500 Index) with the aforementioned maturities. Thereafter, we determined the expected price of our asset at the time of our different maturities maturities $T = [3M, 6M, 1Y]$ as follow:

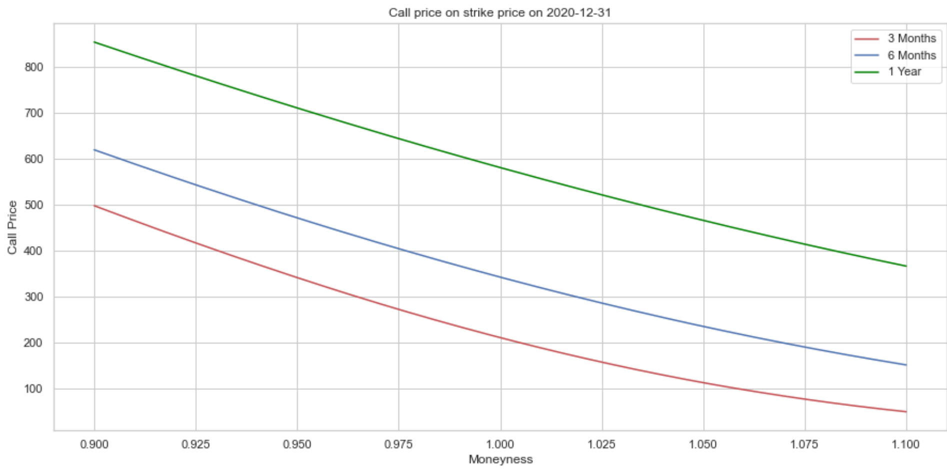
$$S_T = S_0 \times exp(R_T) \tag{22}$$

Consequently, after ranging the moneyness from 0.9 to 1.1, the price of the call as a function of the strike price and time to maturity is defined as:

$$C_0 = e^{-rT} E_0[max(S_T - K, 0)] \tag{23}$$

After an implementation using our time series, we obtained the following results.

Figure 2: Estimation of Call Price



From these outputs, we can conclude two crucial elements in call options:

- **As the strike price increases, the call price decreases:** As the strike price increases, the probability of being in the money (i.e. $S_T > K$) decreases, thus decreasing the value of the call option.
- **As time to maturity increases, the call price increases:** As the time to maturity increases, the probability of being in the money (i.e. $S_T > K$) increases, thus increasing the value of the call option.

Now that we estimated the call prices, we will recover the implied volatility associated. Before doing so, let's first define a fundamental equation in the option market. Let S_t be the price of an underlying asset at time t, K the strike price at time T, r the continuously compounded rate of interest, σ the volatility of the underlying asset and Φ the cumulative standard Normal distribution function. Then, the Black-Scholes formula define a call price at time t:

$$C_t(S, K, T, r, \sigma) = S_t\Phi(d_1) - e^{-r(T-t)}K\Phi(d_2) \tag{24}$$

$$d_1 = \frac{log\left(\frac{S_t}{K}\right) + (r + 0.5 \times \sigma^2)}{\sigma\sqrt{T-t}} \tag{25}$$

$$d_2 = d_1 - \sigma\sqrt{T-t} \tag{26}$$

In the classical Black-Scholes model, all aspects of risk are captured by the single volatility parameter σ. Implied volatility (IV) is the volatility input that makes the model value equal the option's market price:

$$C(S, K, T, r, IV) = C_{Market} \tag{27}$$

In other words, the IV determines a value of σ that explains the option price we observe in the market. Implied volatility can be considered an estimation of the expected volatility over a future period of time. In practice, given option market prices, it is possible to invert numerically the equation **24** to obtain the volatility implied by both the market price of the option and of the other elements involved into the following equation:

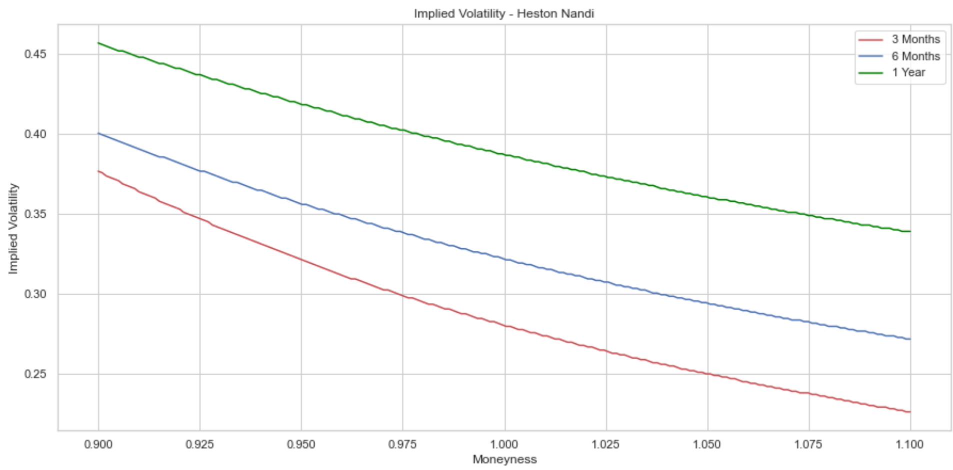
$$\sigma_{BS} = C_t(S, K, T, r, \sigma)^{-1} \tag{28}$$

Then, the IV is determined by minimizing a quadratic criterion such that:

$$\min_{\sigma}(C_t(S, K, T, r, \sigma) - C_{Market})^2 \tag{29}$$

Therefore, the implied volatility associated to the previously computed call prices, as a function to the time to maturity and the moneyness are as follow:

Figure 3: Estimation of Call Price

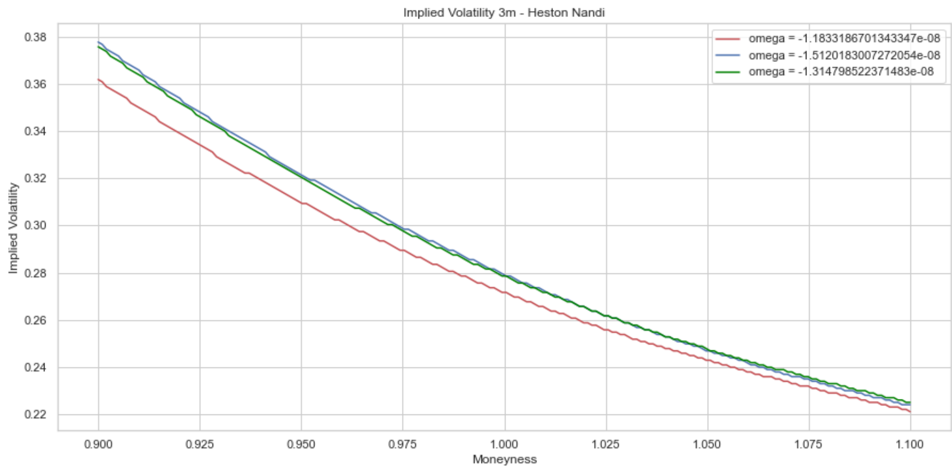


From our results, we notice a skew (or smirk) in the relation between implied volatility and the moneyness. Additionally, as the time to maturity increases, the plot is less curved which may suggest that we may find a volatility smile as time to maturity decreases. Neverthess, as we notice that implied volatilities are not constant over time and across strike prices, this may suggest that we are departing from the Gaussian hypothesis of the Black-Scholes (their model assumed that the volatility is a linear function of moneyness).

(3) To understand in more depth the impact of the Heston-Nandi model parameters, we determined the impact of each parameters on the implied volatility curve first, and then on the implied risk neutral distribution.

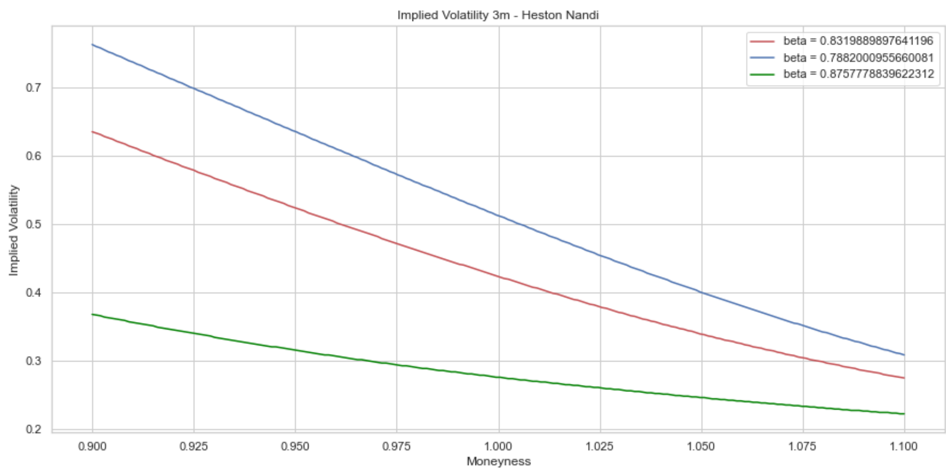
Varying Omega: Before looking at the results, we would expect an increase of implied volatility as omega increases. Indeed, in the GARCH(1, 1) of the Heston-Nandi model, ω is a constant defining the variance. Therefore, a lower omega would imply a lower variance, consequently a lower implied volatility as well since we would obtain a lower value for a call option. However, from our optimisation, we notice that the implied volatility decreases as omega increases, which is against our assumption.

Figure 4: IV using Heston Nandi - Variation of Omega



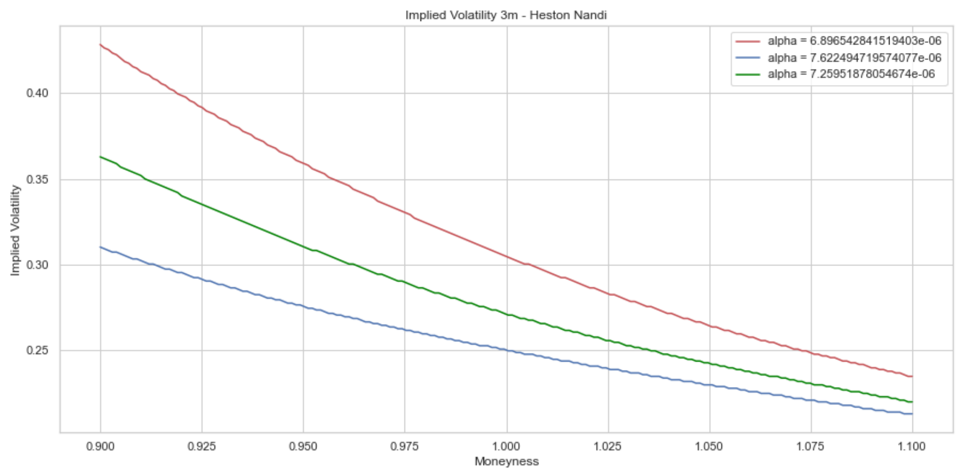
Varying Beta: We would expect the implied volatility to increases as β increases, since we can consider this parameter to be a measure past volatility's persistence. However, in our simulation we notice that the implied volatility decreases as beta increases.

Figure 5: IV using Heston Nandi - Variation of Beta



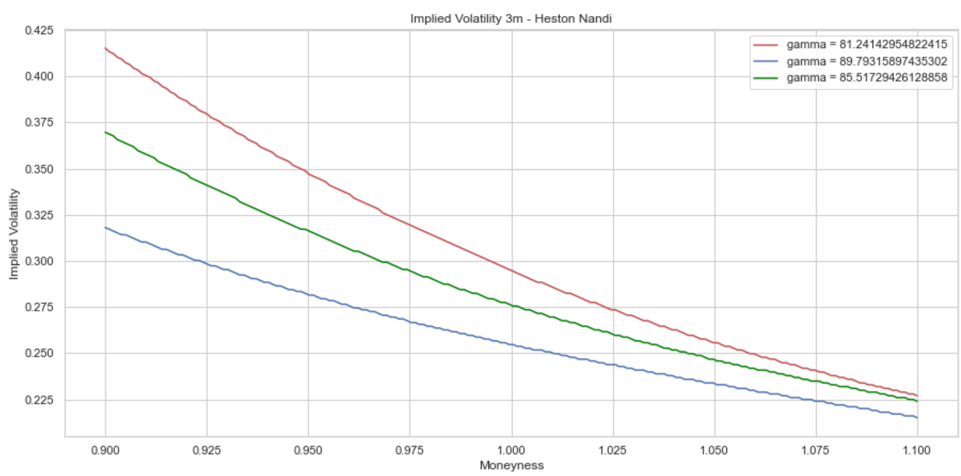
Varying Alpha: We notice that the implied volatility decreases as alpha increases. We know that α determines the kurtosis of the distribution. We have a positive excess kurtosis when the tails of the distribution are fatter and the distribution itself is thinner. But we have a negative excess kurtosis when the the tails of the distribution are flat and the distribution is wider. Therefore, as the distribution is thinner with a positive excess kurtosis, the volatility is lower, thus lowering the call option values. Consequently the implied volatility will also decrease.

Figure 6: IV using Heston Nandi - Variation of Alpha



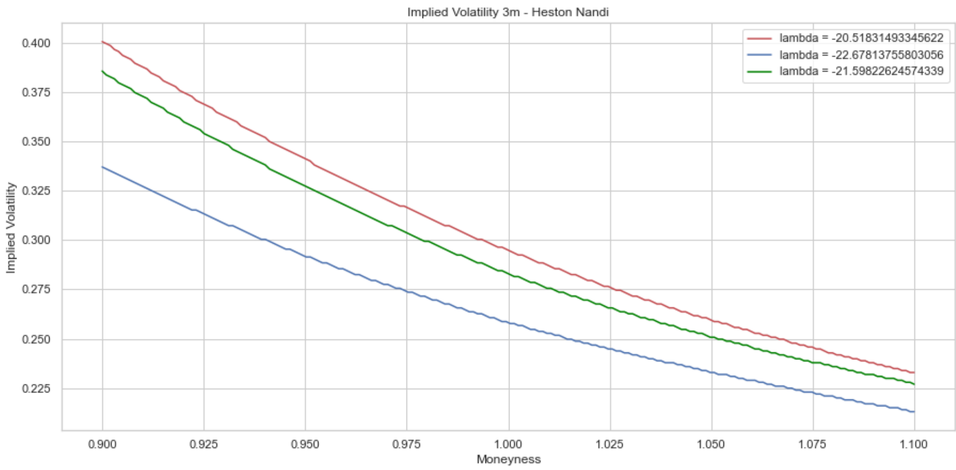
Varying Gamma: We notice that the implied volatility decreases as gamma increases. As γ defines the skewness of a distribution, when we have a positive skewness, one can expect moderate negative returns frequently and occasionally some substantial positive outcomes, and when we have a negative skewness, one can expect moderate positive returns frequently and occasionally some substantial negative outcomes. Consequently, risk averse investors would prefer a positive skew to avoid having unexpected substantial losses, thus limiting the volatility on a position. Therefore, a higher γ would imply a higher skewness as well, thus a lower volatility (less big crashes), which also entail a lower implied volatility since we would obtain a lower value for a call option.

Figure 7: IV using Heston Nandi - Variation of Gamma



Varying Lambda: We notice that the implied volatility decreases as lambda decreases. As λ is a determinant of the risk premium, any riskier investments would require an additional premium. Therefore a lower risk premium (i.e. lower lambda) would suggest that the overall volatility is indeed lower, thus entailing to a lower implied volatility since we would obtain a lower value for a call option.

Figure 8: IV using Heston Nandi - Variation of Lambda



Before analyzing the impact of a variation of each parameters on the implied risk neutral distribution, let's first define the latter. The risk neutral process of the Heston-Nandi model can be depicted as:

$$R_t = log\left(\frac{S_t}{S_{t-\Delta}}\right) = r - 0.5h_t + \sqrt{h_t}\epsilon_t^* \tag{30}$$

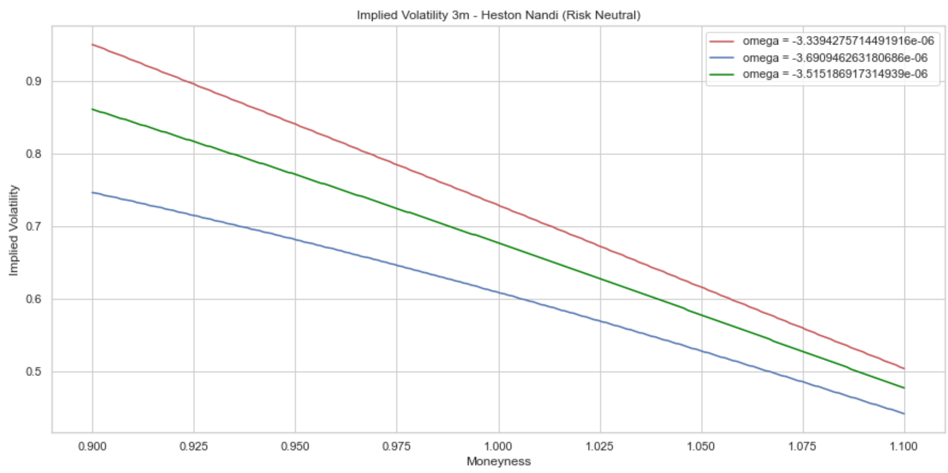
$$h_t = \omega + \beta h_{t-\Delta} + \alpha\left(\epsilon_{t-\Delta}^* - \gamma^*\sqrt{h_{t-\Delta}}\right)^2 \tag{31}$$

Where ϵ_t^* is the standard normal random variable in the risk neutral world and $\gamma^* = \gamma + \lambda + 0.5$. Before going into details for the sensitivity of the IV when varying the parameters, we will notice that overall, the IV is clearly concave in the risk neutral setting. This is because in December 2020, investors were fearing a sudden crash as inflation was rising and concerns on whether the Fed would raise the interest rate were raised, thus explaining the concavity of our IV. A recent paper by Alexiou, Goyal, Kostakis and Rompolis (2021) have found that "implied volatility (IV) curves extracted from short-term equity options frequently become concave prior to the earnings announcement day (EAD) reflecting a bimodal risk-neutral distribution for the underlying stock price"[5].

Varying Omega in Risk Neutral Setting: We notice that the implied volatility decreases as omega decreases.

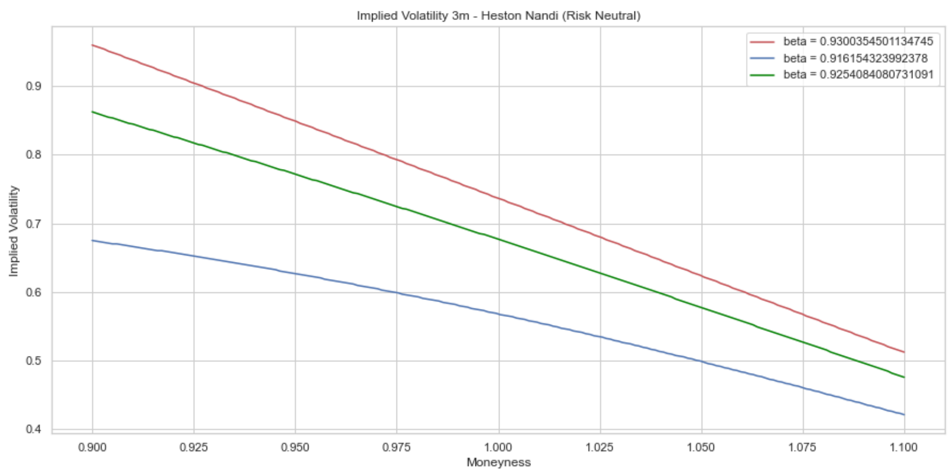
[5] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3840081

Figure 9: IV using Heston Nandi - Variation of Omega (Risk Neutral)



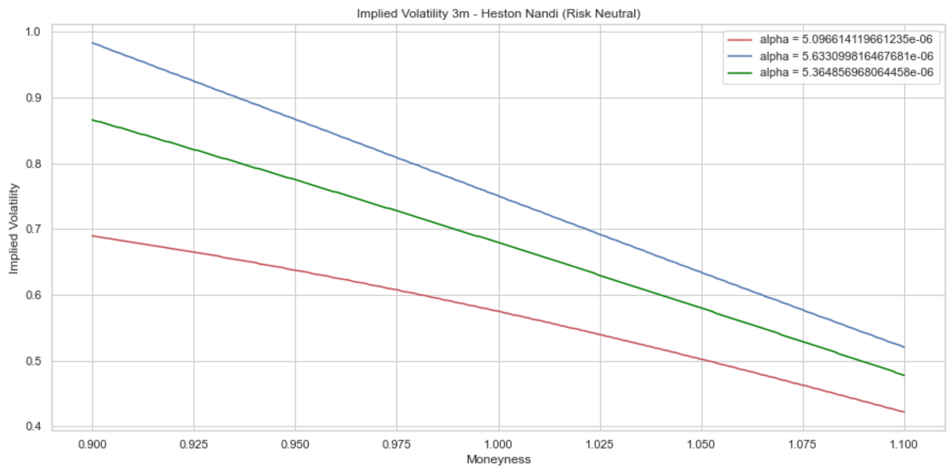
Varying Beta in Risk Neutral Setting: We notice that the implied volatility decreases as beta (i.e. the persistence of past volatility) increases, which is again against our aforementioned assumption.

Figure 10: IV using Heston Nandi - Variation of Beta (Risk Neutral)



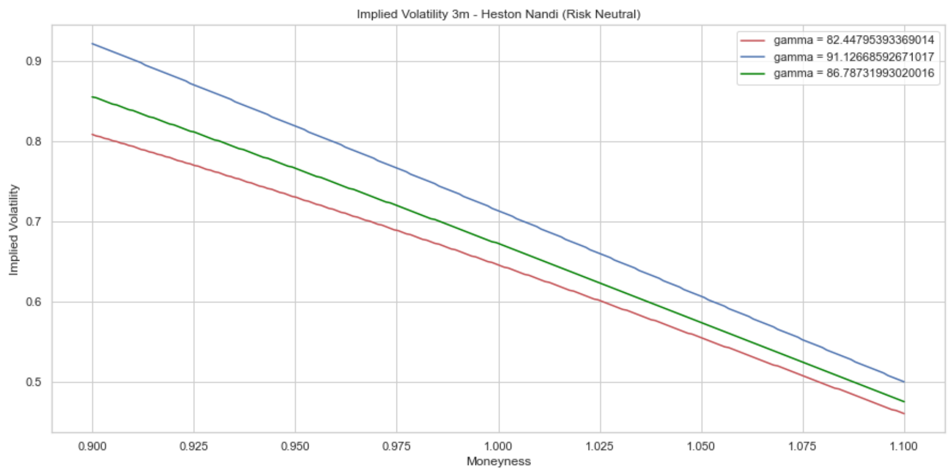
Varying Alpha in Risk Neutral Setting: We notice that the implied volatility decreases as alpha increases. That is, it decreases when the kurtosis increases.

Figure 11: IV using Heston Nandi - Variation of Alpha (Risk Neutral)



Varying Gamma in Risk Neutral Setting: We notice that the implied volatility decreases as gamma increases. That is, it decreases when the skewness increases.

Figure 12: IV using Heston Nandi - Variation of Gamma (Risk Neutral)



5. Conclusion

In Part 1 of the assignment, we have investigated the properties of the log-returns of three particular time series (Apple, BCOM Commodity index and S&P 500) over a time span of 10 years.

For each of the time series, we fitted three different GARCH models (GARCH, EGARCH and GARCH-GJR) and we compared the models with each other, using a Vuong test and a LR test. For each of the considered time series, we retain the standard GARCH model as being the one that fits the best the data.

Having estimated the time-varying volatility of each of the series with GARCH (the best model for each series), we recover the filtered residuals of each series. This allows us to test the Gaussianity of the filtered residuals by applying a Kolmogorov and Smirnov test. For each series, we can clearly reject the Gaussianity of the filtered residuals, meaning that returns are conditionally not Gaussian.

Finally, we tried to fit 5 different distribution to our filtered residuals. Then, we test whether these distributions are accurately fitting our residuals by using a Kolmogorov and Smirnov test. We conclude that for three of the five distributions (GenNorm, NormInvGauss and JohnSU) and for every series, the residuals were accurately fitted. Then for each of the time series, we used information criteria and the sum of squared residuals to decide which of the three distributions that fits well our residuals is the best. For the S&P 500, we retain the NormInvGauss distribution while for Apple and BCOM Commodity, we retain the Johnson Su to be the best distribution.

In part 2 of this report, we have been able to implement the Heston & Nandi model option pricing model for the S&P500 Index. After conducting a maximum likelihood estimation, we have been able to estimate all parameters (α, β, δ, γ, λ) of our model. We observed that all parameters are significant to a high level of significance, except omega.

Afterwards, we simulated the option prices on the 31st of December 2020 using $10'000$ Monte Carlo simulations and a risk free-rate of 0.25% for different maturities and different moneyness. We observed that the call price decreases as the moneyness increases, since a rise in the latter implies a higher strike price, which decreases the probability of being in-the-money. A higher time to maturity implies a higher option price since the probability of being in-the-money increases as well. We then computed the implied volatility of our options, which depends negatively to the moneyness and positively to the time to maturity. The important result is that the relation between moneyness and IV is not constant through time, which is a departure of Black-Scholes' assumption that returns are Gaussian.

Finally, we computed the sensitivity of the 3-month IV with the parameters of the Heston-Nandi model, under historical probability and risk neutral probability. The interesting result in this part is that, under risk neutral setting, the IV is concave, which is an unusual result. This may be caused by the general fear of a potential crash in the next 3 months (point of view in December 2020), given the latest rise in inflation and the fear of a potential rise of interest rates by the Fed as counter response.