HOW DID COVID-19 CRISIS AFFECT FINANCIAL RISK MODELING?

Gianluca Broll

Abstract

This brief article deals with the consequences carried out by the Covid-19 pandemic in the financial markets, focusing on the tail risk modelling. In details, I present a way of modelling the tail of returns distribution by means of a Generalized Pareto Distribution (*GPD*) along with an *ARMA-EGARCH*, analyzing how the pandemic affected the left tail of the distribution. In second place, I estimate a *VaR* forecasting model to show how this risk measure changed during the pandemic. Even in this case I implemented an *EVT-EGARCH* model. The choice of the *Exponential GARCH* accounts for the well-known evidence of asymmetric magnitude of shocks on volatility. The sample used is the S&P 500 over the period from 1990-01-01 to 2020-07-09.

Theoretical Framework

The *GPD* used to model the tail is estimated through the *peak over threshold* (POT) method, and the estimated x_α quantile of the distribution is defined as:

$$q_\alpha = u + \frac{\hat{\xi}}{\hat{\beta}}\left(\left(\frac{1-\alpha}{\hat{\overline{F}}(u)}\right)^{-\hat{\xi}} - 1\right) \tag{1}$$

Where $\hat{\xi}$ and $\hat{\beta}$ are the ML estimates for the shape and scale parameters, respectively, while $\hat{\overline{F}}(u)$ is a (non-parametric) estimate for the cumulative distribution of the exceedances over the threshold u.

The *i.i.d.* observations to which apply *POT* method are obtained by means of the implementation of a model for conditional mean and for conditional volatility. The former is an *ARMA* model having the following general form:

$$y_t = \mu + \sum_{j=1}^{p} \rho_j y_{t-j} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t$$

The latter is an *EGARCH* model (*Nelson,* 1991), that has the following general formulation:

$$\ln \sigma_t^2 = \omega + \sum_{j=1}^{m} \zeta_j \, v_{jt} + \sum_{j=1}^{q} \left(\alpha_j \frac{\varepsilon_{t-j}}{\sqrt{\sigma_{t-j}^2}} + \gamma_j \left(\frac{|\varepsilon_{t-j}|}{\sqrt{\sigma_{t-j}^2}} - E \frac{|\varepsilon_{t-j}|}{\sqrt{\sigma_{t-j}^2}} \right) \right) + \sum_{j=1}^{p} \beta_j \ln \sigma_{t-j}^2$$

This choice is convenient because allows to consider asymmetric impact of shocks. I implemented an *ARMA(1, 1)* for the mean and a *EGARCH(2, 2)* for the variance (without external regressors), i.e.:

$$y_t = \mu + \rho y_{t-1} \theta \varepsilon_{t-1} + \varepsilon_t \tag{2}$$

$$\varepsilon_t = \sigma_t z_t \qquad z_t \sim N(0,1)$$

$$\ln \sigma_t^2 = \omega + \sum_{j=1}^{2} \left(\alpha_j \frac{\varepsilon_{t-j}}{\sqrt{\sigma_{t-j}^2}} + \gamma_j \left(\frac{|\varepsilon_{t-j}|}{\sqrt{\sigma_{t-j}^2}} - E \frac{|\varepsilon_{t-j}|}{\sqrt{\sigma_{t-j}^2}} \right) \right) + \sum_{j=1}^{2} \beta_j \ln \sigma_{t-j}^2 \tag{3}$$

Where, if ε_t is assumed to be normally distributed, $E \frac{|\varepsilon_t|}{\sqrt{\sigma_t^2}} = E|z_t| = \sqrt{\frac{2}{\pi}}$, which is the expected value of a folded normal [1].

[1] This because:

$$z_t \sim N(0,1) \quad \rightarrow \quad f(|z_t|; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(|z_t|)^2} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z_t)^2} + \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(-z_t)^2} = \frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}(z_t)^2}$$

Therefore:

$$E(|z_t|) = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} z_t e^{-\frac{1}{2}(z_t)^2} \, dz$$

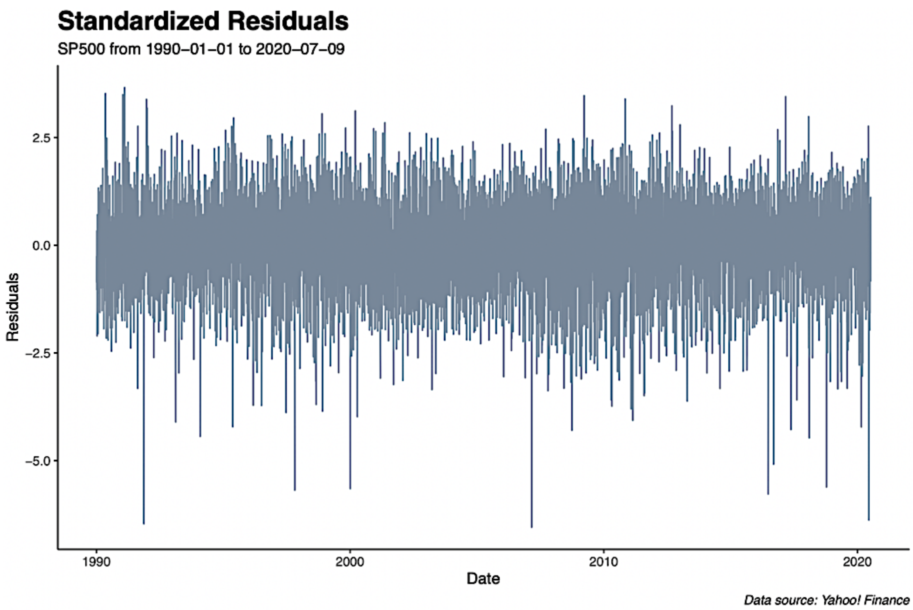
Given that:

$$\frac{\partial \left(e^{-\frac{1}{2} z_t^2} \right)}{\partial z_t} = -z_t e^{-\frac{1}{2} z_t^2} \quad \rightarrow \quad \int -z_t e^{-\frac{1}{2} z_t^2} \, dz = e^{-\frac{1}{2}(z_t)^2}$$

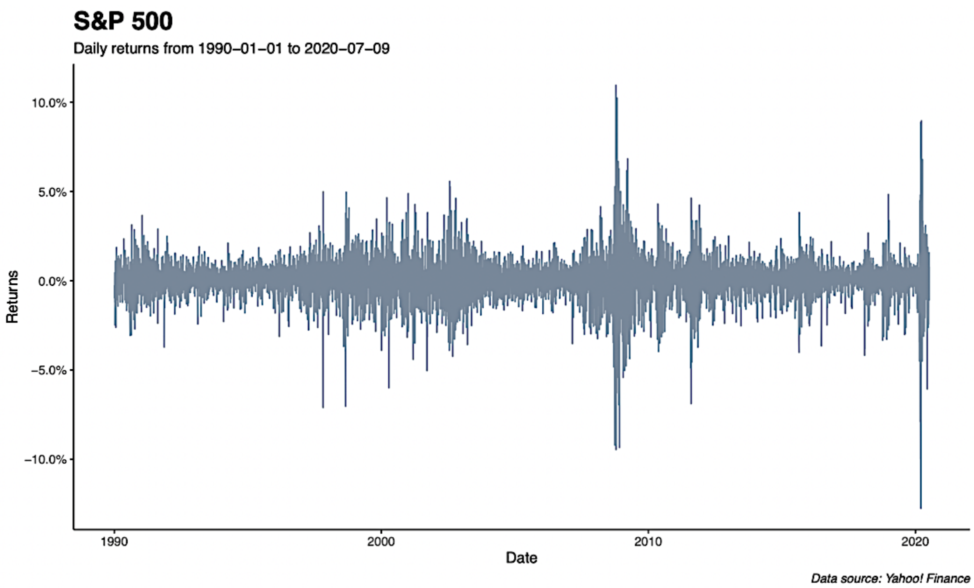
The expected value of the folded normal is then:

$$E(|z_t|) = -\frac{2}{\sqrt{2\pi}} \int_0^{\infty} -z_t e^{-\frac{1}{2}(z_t)^2} \, dz = -\frac{2}{\sqrt{2\pi}} \left[e^{-\frac{1}{2} z_t^2} \right]_0^{\infty} = -\frac{2}{\sqrt{2\pi}} (0 - 1) = \frac{2}{\sqrt{2\pi}} = \sqrt{\frac{2}{\pi}}$$

Once the model has been specified, the *i.i.d.* observations are obtained standardizing the inferred residuals by modeled conditional variance. The result of such procedure is reported in the following chart:



One can appreciate even more the benefits of standardization if compared with the return series (which clearly shows volatility clustering):



The procedure described up to now regards the estimate of the tails of the distribution for S&P 500 returns, i.e. the *GPD*. To determine the *VaR*, I developed two different approaches. The first one uses an *EGARCH* on a rolling window of 250 observations to forecast one-day-ahead volatility, focusing on the period from 2019-01-01 to 2020-07, and updating the *GPD* estimation each month (i.e., including the last month in the sample used for ML parameters estimation). The second approach uses an expanding window for the volatility forecasts. In both cases, daily *VaR* is determined by the following formula:

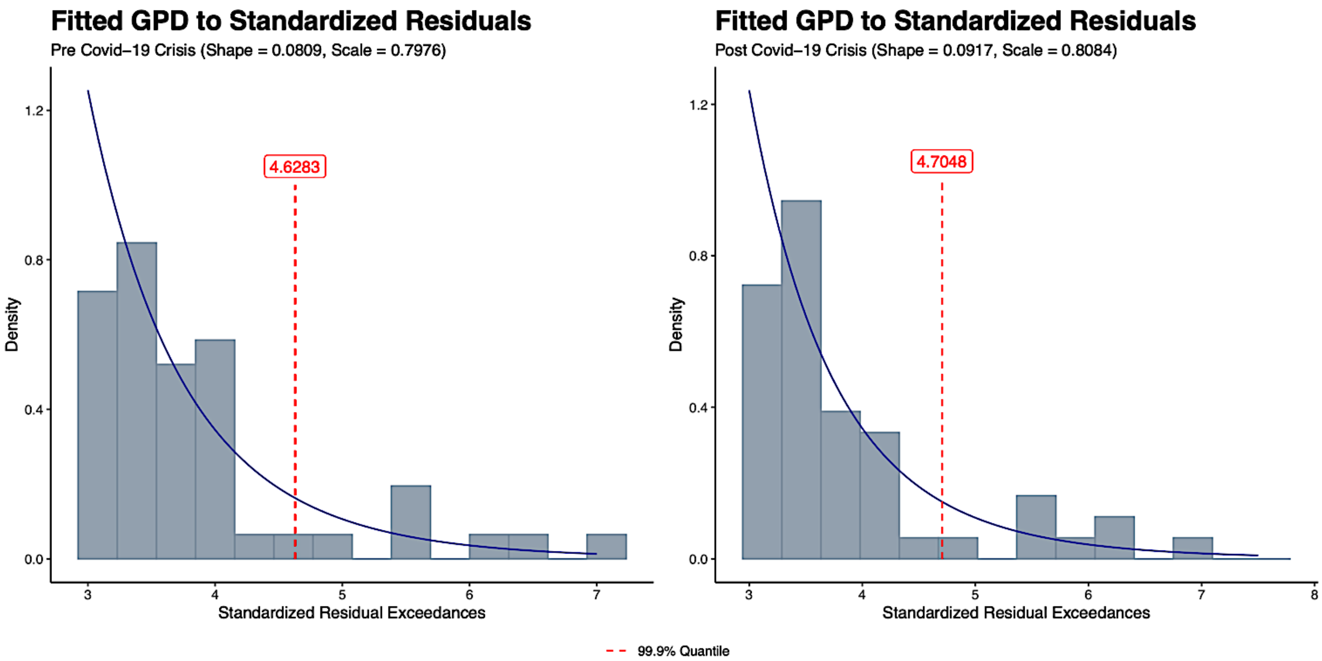
$$VaR_{T+1}^{\alpha} = y_{T+1} + \sigma_{T+1} \cdot q_{\alpha} \tag{4}$$

Where q_{α} is the *alpha*-quantile of the *GPD* as expressed by (1), recalculated each month, and y_{T+1} and σ_{T+1} are the one-day-ahead forecasts for the conditional mean and conditional volatility, respectively.

Therefore, the analysis splits in two parts, the first one evaluates how the Covid-19 crisis has changed the tail of the distribution, comparing the *GPD* estimates obtained with the whole sample including the first six months of 2020 with the ones obtained with a sample excluding them (up to 2019-12-31). The second part focuses on the evaluation of the risk of a long position in the whole US market.

Empirical Results

The following charts report the result of the fitted *GPD* to the standardized return for the two subsamples analyzed.



As it can be seen, the *GPD* fits quite well the tail distribution of the standardized returns. As the quantile highlighted in red suggests, the Covid-19 crisis modified slightly the estimated distribution, making it more heavy-tailed. This is remarkable if we think that the dotted line represents the 99.9 cumulative percentile, as saying that the very extreme (negative) events are slightly more probable after the pandemic. Even the shape parameter suggests us that the distribution is now more leptokurtic (even though its variation is well within the confidence interval), as it has increased after the crisis.

As previously said, the second part of the analysis ascertains the impact of the crisis on *VaR* forecast, for this purpose I selected a lower threshold for the *POT* procedure to estimate the *GDP* [2] for determining a slightly less conservative VaR measure (99% instead of 99.9%). As anticipated above, two ways are possible in this case: expanding window and rolling window. The former uses a sample that becomes larger and larger at each iteration, including the new information to the sample used, while the latter uses the same number of observations, including the newest and discarding the oldest one at each iteration. The difference between the two approaches is that the expanding windows models is less reactive to new information as it is conditioned also by the very far in the past one, while the rolling method quickly adapts to new market conditions, producing more reactive forecasts. In addition, *EGARCH* order for the expanding window is (2, 2), while for the rolling analysis the simpler specification (1, 1) is sufficient to obtain a proper fit of the data (with a particular focus on sign bias test statistic, whose null hypothesis is not rejected). The latter is also less time-consuming in the estimate procedure. For both strategies, I determined the risk measure for the period that goes from 2020-01-01 to 2020-07-09. The *GPD* is recalculated each month (in this case with an expanding window) to account for the changes in the tails that happened during the pandemic.

---

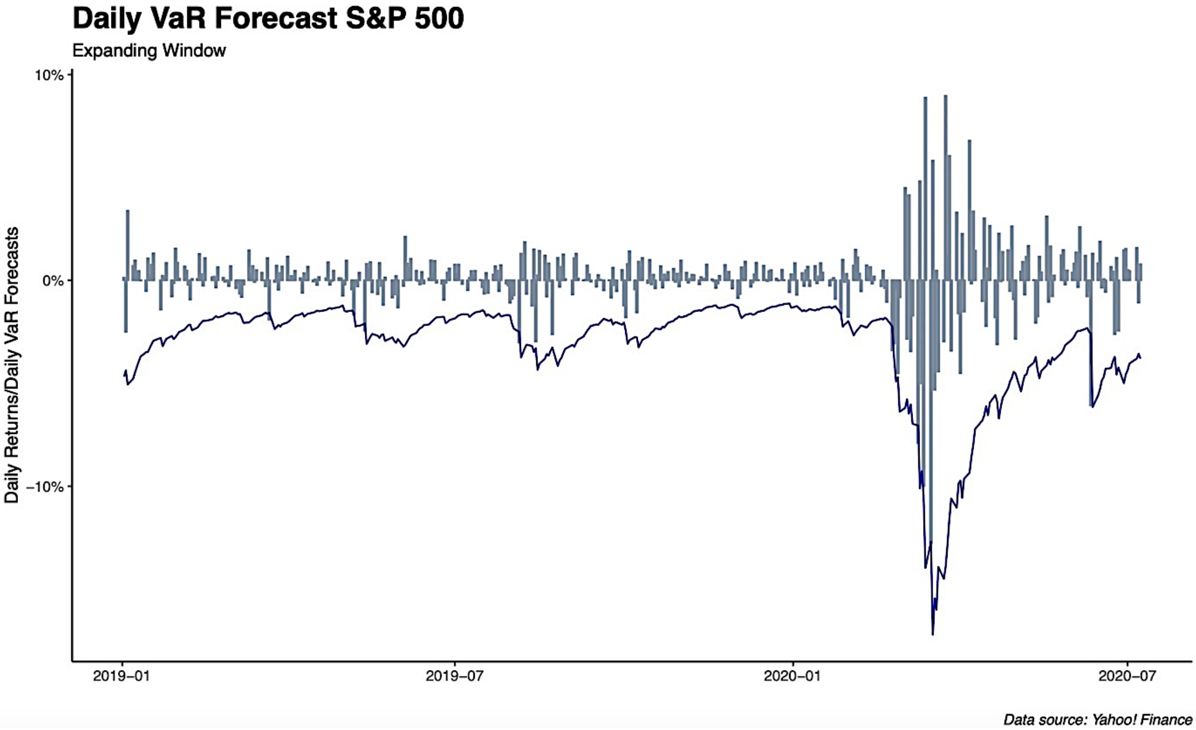
[2] A trade-off comes along with the choice of the threshold. As stated by *Pickand-Balkema-de Haan* theorem, the goodness of the *GPD* approximation depends on the choice of the threshold: the higher it is, the better the approximation is. In fact, if and only if $F \in MDA(H)$, where H is the *Generalized Extreme Value (GEV) Distribution*, it holds that:

$$\lim_{u \to x_F} \sup_{0<x<x_F-u} |F_u(x) - GPD(x)| = 0$$

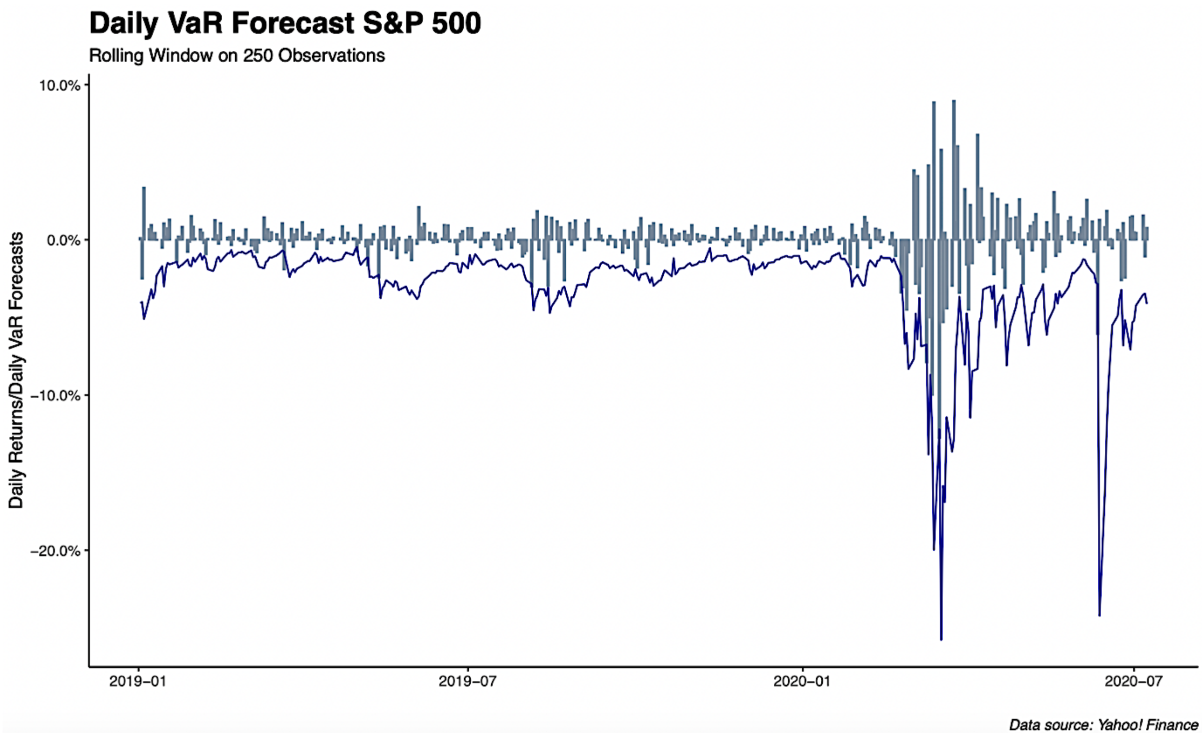
$$x_F = \sup\{x \in \mathbb{R} : F(x) < 1\}$$

On the other hand, as a result of a very high threshold, the exceedances with which estimating the *GPD* parameters might be too little to ensure a sufficient precision of the ML estimator. While in the first part of the analysis I was more interested in looking what happened in the extreme tail of the distribution (and thus wanting the best approximation possible), in the second part I am willing to scarify a little bit of approximation precision to have more stable parameter estimates.

The following chart reports the result for the expanding window model. The dark blue line represents the forecasted *VaR*, while the bars indicates the S&P 500 returns.



The next graph displays the result from the rolling window model, with the very same data represented.



The situation depicted is very similar, despite some differences in the adaptive capacity already mentioned. The main takeaway is the uncanny increase in volatility during the worst months of the pandemic (Mach-April), with the *VaR* forecast increased up to 20% (and beyond). This is astonishing if we recognize these values represents the maximum daily losses (with a 99% confidence) and that usually *VaR* is used to place limit to trading operations: with such increase in volatility the risk-taking capacity of a trading desk dramatically decreases.

Conclusions

This brief article attempts to give a glance at the consequences of the worst pandemic ever faced in the globalized world. Despite being coarse in some assumptions, the model presented gives an insight of the financial consequences of the Covid-19 crisis in tail modelling and risk forecasting. Financial markets are weird (and wild!) creatures, but surely reflects investors' behavior: the increase in volatility and the plunge in prices are clear evidence of widespread uncertainty caused by the pandemic. In addition, it is worth mentioning that such thrill on financial market has not an endogenous origin, but comes from the real sector, i.e. the real economy. The production halt and the consequent losses (look at the airlines sector, for instance) had an immediate feedback on the financial markets but will probably have more severe real consequences in the long term.

References

[1] Alexios Ghalanos (2020). *rugarch*: Univariate GARCH models. R package version 1.4-2.

[2] Alec G. Stephenson. (2018). *ismev*: An Introduction to Statistical Modeling of Extreme Values. R package version 1.42.

[3] McNeil, A. J., Embrechts, P., & Frey, R. (2015). *Quantitative risk management: Concepts, techniques and tools*. Princeton, NJ: Princeton Univ. Press.

[4] Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica, 59*(2), 347.