Mongolia's TOP-20 Index Risk Analysis, Pt. 2

Federico M. Massari

March 3, 2017

In the second part of our risk report on TOP-20 Index, Mongolia's main stock market indicator, we refine the models introduced in part one, accounting for several important features of the Mongolian stock market – most notably, positive skewness and negative leverage – and demonstrate why the index makes a very nice investment opportunity at the moment.

Overview

In part two of our risk report on TOP-20 Index we improve[1] the quality of the return forecast accounting for some non-normality features that remained after previous model calibration [4].

Non-normality (or non-Gaussianity) is a condition that prevents returns from being adequately described by the Gaussian density – the bell-shaped distribution –, making them difficult to forecast.

The most visible departures of TOP-20 Index from normality are the autocorrelation of both log and squared returns, the volatility clustering, the positive skewness/negative leverage effect, and the positive and large excess kurtosis.

We already discussed these features in part one, where we mostly relied on visual exploration of data, inspecting the empirical distribution of returns, as well as the plots of the autocorrelograms of both logs and squares. This time, we want to formalise the previous intuitions, performing tests of normality, autocorrelation, volatility clustering, and asymmetry.

Once we confirm that TOP-20 Index returns are non-Gaussian, we improve the GARCH-VT standardised t distribution model we relied on in past reports, starting from non-linear extensions of GARCH, then reconsidering the validity of the asymmetric t distribution. The fit we obtain in the end is visibly better, at least in the left tail.

In the last part of this work, we provide a one-month ahead forecast of the distribution of returns, together with measures of Value at Risk and Expected Shortfall.

Data

We use TOP-20 Index daily log returns* from August 13, 2007 to February 24, 2017. They are price returns (they only reflect capital appreciation and disregard dividends), gross of fees, expenses, and taxes.

We point out that, under a six-month renewal policy, the list of TOP-20 Index constituents slightly changed recently, with the removal of three companies (Genco Tour Bureau, Mik Holdings and Telecom Mongolia) and the inclusion of three others (Hermes Center, Khukh Gan and Aduun Chuluun). See [6] for additional information.

[1] Unless otherwise specified, all the techniques implemented come from [1], which we closely followed.

Stylised facts of TOP-20 Index returns

We present a list of empirical findings which are distinctive of TOP-20 Index, and which should be captured to the greatest extent to improve the quality of the return forecast.

Autocorrelation of log returns

This feature refers to the predictive power of past log returns (or residuals: demeaned returns) on today's one. Usually, at daily level, the autocorrelation of log returns should be close to zero, implying absence of predictability based on past data. However, TOP-20 returns do exhibit some linear dependence with their own history up to several weeks, an anomaly we connected to the thin trade, or scarce liquidity, of some of the index constituents. In addition, the sign and strength of this relationship vary in time (Figure 1).
We managed to capture this feature in full, using a GARCH filter.

Autocorrelation of squared returns and volatility clustering

The first refers to the ability of past squared returns to predict, to a certain extent, both the direction and the size of future squared returns. The second is the tendency of high (low) volatility days to be followed by days of similar intensity. Since squared returns are a valid proxy for variance, the presence of autocorrelation in squares also implies volatility clustering. These phenomena, very common in most equity markets even at daily level, became less significant in Mongolia after April 8, 2011 (Figure 2), the day in which an agreement between MSE and London Stock Exchange to bring the infrastructure, technology, and human resources of the Mongolian partner to international standards reduced market uncertainty, cutting volatility to one

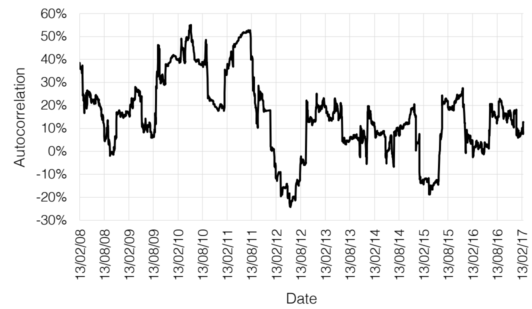


Figure 1: Time-varying autocorrelation of log returns (lag 1), rolling-window 6m

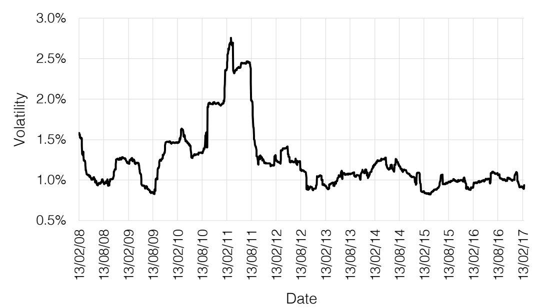


Figure 2: Time-varying volatility (standard deviation), rolling-window 6m

third of its initial value, and lowering both the frequency and the magnitude of extreme returns [5].
We also managed to capture these phenomena in full, using a GARCH filter.

Positive skewness and the negative leverage effect

Positive skewness refers to an asymmetry in the distribution of returns which makes small losses, as well as extreme gains, more frequent than, respectively, small gains and extreme losses. Negative leverage exists when positive returns boost variance (a proxy for risk) more than negative returns of the same size do, and it is the opposite of the so-called leverage effect, named after the fact that a decrease in the value of a company's equity provokes an increase in financial leverage (the debt-to-equity ratio), making the full repayment of debt uncertain and the firm more risky. Positive skewness and negative leverage are two sides of the same coin: a higher volatility of positive returns implies a greater dispersion of gains, with extreme gains both more likely and larger than extreme losses; a lower volatility of negative returns signifies a greater concentration of losses, with small losses both more frequent and closer to zero than small gains.

Although not rare in developing economies, these two features are considered anomalies, since most equity indices are characterised by the opposite phenomena: a higher frequency of small gains and, occasionally, big crashes. However, as they make large profits more likely, they are nice anomalies, and investors like them so much that they «may be willing to accept negative expected return in the presence of large, positive skewness [and negative leverage]»[a]. For instance, when we analysed monthly returns of TOP-20 Index mimicking portfolios [5], we concluded that, thanks to a combination of rising markets, positive skewness, and high kurtosis, the index makes a good investment opportunity at the moment, in spite of an average return, in the most recent five years, around –0.75%.

The simultaneous existence of negative average return and positive skewness/negative leverage is indeed an interesting phenomenon of TOP-20 Index. Considering rolling subsamples of six-month length, we see that skewness stayed positive about 75% of the times over the past ten years and, in the last five, only on very few occasions it dropped below zero (Figure 3). By contrast, mean return was negative about 55% of the times, often during periods of positive skewness (Figure 4), although its value was so small on average that it could have been forgone in view of the better chance to greatly profit from the position (Table 1).

Figure 3: Time-varying skewness, rolling-window 6m

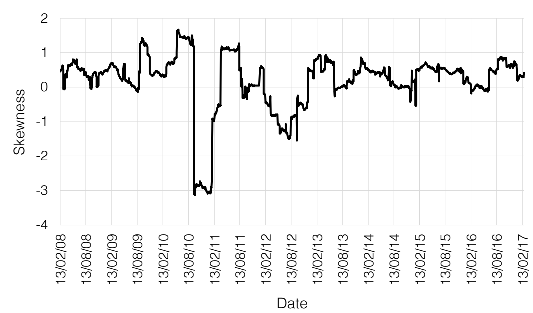


Figure 4: Time-varying mean, rolling-window 6m

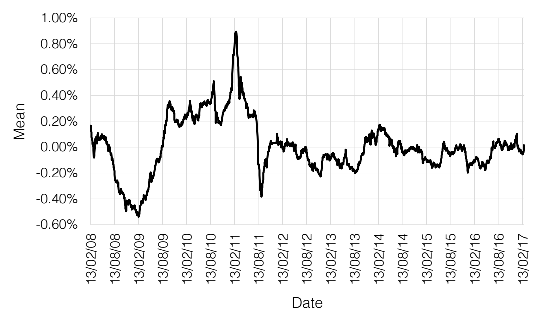


Table 1: Odd order moments, rolling-window 6m

	Negative	Positive
Mean	55.9%	44.1%
Skewness	24.4%	75.6%
Avg. mean (range)	0.00% (–0.54%; 0.89%)	
Avg. skewness (range)	0.20 (–3.14; 1.67)	

The prevalence of an asymmetry of TOP-20 Index towards large gains indicates that the public react to positive news more than they do to negative news. In other words, *investors are not as scared after negative events as they are excited after positive ones*. Under these circumstances, it is possible to make good use of traditionally disliked even moments, variance and kurtosis, as well as the strongly positive autocorrelation in log returns: a higher volatility of gains, combined with fatter tails and return clustering may be beneficial when the market is strong, even in presence of negative average return.

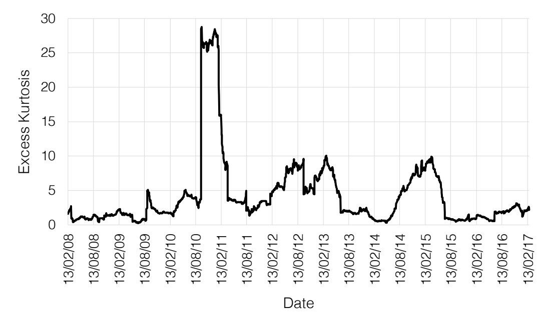
> *“ Investors are not as scared after negative events as they are excited after positive ones. „*

Incorporating positive skewness and negative leverage in forecasting models is of paramount importance to improve the accuracy of predictions, and it is the primary focus of this work. Obviously, we could not account for these features in full using GARCH and the standardised t distribution, because they are symmetric models. We will come back to this point in a while, after introducing non-linear GARCH models and the concept of news impact curve.

Positive excess kurtosis

In part one of this risk report we showed that the distribution of TOP-20 Index returns has positive excess kurtosis. Kurtosis is a shape parameter that measures the tail thickness of a distribution: the larger its value, the thicker the tails, and the higher the frequency of extreme events. Any kurtosis above 3, the value for the Gaussian density, is called positive excess kurtosis, and it is a synonym for fat tails. Investors usually dislike this parameter, which is associated to the higher likelihood of market crashes. However, if the return distribution is positively skewed, excess kurtosis may be beneficial because it amplifies gains more than it does losses. Over the whole ten-year period, excess kurtosis has always been positive (although, on average, not too large) and frequently clustered.
We managed to capture much of the excess

Figure 5: Time-varying excess kurtosis, rolling-window 6m



kurtosis with the GARCH standardised t model, and we now consider refining our proxy.

We surveyed the most important features of TOP-20 Index: the autocorrelation of log and squared returns, volatility clustering, positive skewness and the negative leverage effect, positive excess kurtosis. While it is useful to model all features to the greatest extent, we point out that their relative importance might vary in time, together with the development of the Mongolian stock exchange.

Tests

We now test the following assumptions: normality, absence of autocorrelation, absence of volatility clustering (no ARCH effect), absence of asymmetry in the distribution of returns.

Normality tests

When we previously questioned the normality assumption, we did so mainly relying on visual tools, such as the histogram of returns and the autocorrelograms (the plots of autocorrelation functions) of logs and squares.
This time, we formally test the null hypothesis that TOP-20 Index log returns come from a Gaussian density against the alternative that they do not:

H_0 : Log returns are Gaussian distributed
H_1 : Log returns are non-Gaussian

We perform two normality tests. The first one, the Jarque-Bera test, checks whether the higher moments of the empirical distribution of returns, skewness and excess kurtosis, are jointly zero, as they are in the Gaussian density. If they are not, we reject the normality assumption. The JB statistic is distributed according to a chi-squared density with two degrees of freedom (at least two observations are needed in order to find two parameters – only in this case a solution exists for two equations in two unknowns).

The second one is a variation of the Kolmogorov-Smirnov test due to Lilliefors. The KS test verifies whether the greatest distance between the empirical CDF and the theoretical, Gaussian CDF is within the limits (so that we can reasonably say the data come from a normal density). If it is not, we reject the normality assumption. The Gaussian CDF to which the empirical distribution is compared requires knowledge of the population mean and standard deviation, two unobserved parameters. The Lilliefors test simply replaces the unknown values with the estimated, sample moments. Basically, what changes are the tabulated critical values.

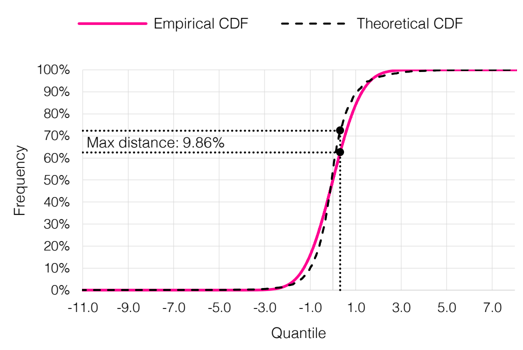
The Jarque-Bera statistic we obtain is well above the critical value at a significance level of 1% (the p-value, the minimum threshold at which it is possible to reject the null hypothesis, is practically zero), and it confirms our previous intuition (Table 2).

The maximum distance between the empirical and the Gaussian distributions is 9.86%, also well above the one allowed: the tabulated critical value for the Lilliefors test, $\alpha = 1\%$, is $1.031/\sqrt{\text{sample size}} \approx 2.11\% < 9.86\%$ (Table 3, Figure 6).

Table 2: Jarque-Bera test of normality, α=0.01	
Sample skewness	0.2666
Sample excess kurtosis	11.7589
Jarque-Bera statistic	13,746.01
Critical value	9.21
Reject normality hypothesis?	YES

Table 3: Lilliefors test of normality, α=0.01	
Maximum distance	9.86%
Maximum distance allowed	2.11%
Reject normality hypothesis?	YES

Figure 6: Lilliefors test of normality



In this case, too, we reject the normality assumption.

Hence, we formally reject H_0, the null hypothesis of Gaussian distributed log returns, in favour of H_1, the alternative hypothesis of non-normal log returns.

Autocorrelation tests

After inspection of the autocorrelation functions, we also concluded that TOP-20 Index returns, as well as their squares, were visibly correlated with past values at several lags of data (even weeks). We formally check for the presence of autocorrelation using the Ljung-Box test. LB tests the null hypothesis that the autocorrelation coefficients for both log and squared returns are zero at all lags, against the alternative that at least one of them is different from zero. The analysis is performed on residuals (demeaned log returns) and on squared residuals:

H_0 : Residuals/squared residuals are non-correlated to their past values at all lags

H_1 : Residuals/squared residuals are correlated to at least one of their past values

We use 20 lags of past data, the value suggested by Ljung and Box. Testing at different lags (10, 50, 100) leads to the same conclusion.

The autocorrelation of log returns is strong. At 1% significance, the LB statistic is 445.92, well above 37.57, the critical value from a chi-squared distribution with 20 degrees of freedom (Table 4). As we previously pointed out, this is an anomaly which may be related to the thin trade of some of the index constituents.

The autocorrelation of squared returns is even stronger. The LB statistic is larger at all lags, and equal to 507.34 at 20.

Hence, we formally reject H_0, the null hypothesis of absence of correlation of log returns and squared returns at all lags, in favour of H_1, the alternative hypothesis of autocorrelation in at least one lag of past data.

ARCH effect tests

We already hinted at the presence of volatility clustering when we performed the Ljung-Box test on squared residuals (a proxy for variance) and concluded that, in agreement with the analysis of the autocorrelation function made in part one, squared returns are strongly autocorrelated up to several weeks. We now test the volatility clustering assumption using Engle's ARCH effect test (also valid for GARCH). We regress squared residuals against their past four values. If there is no ARCH effect, the coefficients of the past

Table 4: Ljung-Box AC test, α=0.01, lags=20	
LB statistic (log returns)	445.92
LB statistic (squares)	507.34
Critical value	37.57
Reject absence of AC?	YES

Table 5: Engle's ARCH effect test, α=0.01, lags=4			
Adjusted R-squared	5.61%		
F-test (significance)	36.33 (1.65E-29)		
	Coefficient	p-value	Semi-partial R^2
ε^2_{t-1}	0.1636	2.51E-15	2.52%
ε^2_{t-2}	0.1080	2.11E-07	1.08%
ε^2_{t-3}	0.0698	7.85E-04	0.45%
ε^2_{t-4}	0.0090	0.6605	7.66E-05

squared residuals should all be zero. Formally, we test:

H_0 : past squared residuals do not have any predictive power over the most recent one (absence of ARCH effect)

H_1 : past squared residuals have predictive power over the most recent one (presence of ARCH effect)

The coefficient of determination (adjusted R^2) tells that nearly 5.6% of the variability of today's squared residual around its mean value is explained by the linear combination of past variables (Table 5). The results are statistically significant and did not occur by chance (the F-test p-value is close to zero) so we can safely reject the null hypothesis for α = 1%.

The variables are not only significant but also relevant, according to the semi-partial R^2. A semi-partial R^2 shows how much the total R^2 of the regression increases thanks to the addition of a particular explanatory variable. In other words, it tells the specific relevance of a given regressor, keeping else constant. Including the first three lags of squared residuals, for instance, boosts total R^2 by 2.52%, 1.08%, and 0.45%, respectively, and shows that these variables have some predictive power on the variability of today's squared residual around its average. For example, an increase in the value of ε^2_{t-1} by 1 causes ε^2_t to go up by $0.1636 \times 2.52\% \approx 0.41\%$. It may be small, but it is quite satisfactory, given the complexity of the process. Interestingly, while all lags greater than three are progressively less important in the regression, so that including them does not add much to the overall R^2, the ninth term has half, and the tenth term the same, predictive power as the first one: when up to ten lags of data are included, both the first and the tenth semi-partial R^2 are about 1.78%, and the ninth about 0.87% (the spike in the autocorrelogram of squared returns, Figure 6 in part one of the risk report).

Hence, we reject H_0, the null hypothesis of absence of ARCH effects (volatility clustering) at all lags of the squared residuals, in favour of H_1, the alternative hypothesis of presence of ARCH effect in at least one previous lag of the same variable.

Asymmetry tests

Finally, we test for the presence of asymmetry in the distribution of returns. Specifically, we want to prove that negative leverage, the fact that positive news boost variance more than negative news do, persists even after a GARCH fit. To do so, we rely on a slight variation of Engle and Ng's sign and size bias test. We regress squared GARCH shocks (residuals standardised using GARCH conditional volatility, then squared) on three variables. The first one is a dummy that takes value 1 if previous day's residual is positive and zero otherwise (positive sign), and it is useful to test whether a positive return, albeit small, is enough to raise volatility. The second and third ones are dummies that take value equal to that of previous day's residual if it is, respectively, negative or positive, and zero otherwise, and they are helpful to test whether the size of the residual also matters. If the distribution of GARCH-VT shocks is symmetric, none of these three variables should have any predictive power, and their coefficients should be jointly null.

Table 6: Engle-Ng sign and size bias test, α=0.01

Adjusted R-squared			0.53%
F-test (significance)			5.25 (1.31E-03)
	Coefficient	p-value	Semi-partial R^2
Positive Sign Bias	−0.2718	0.0376	0.18%
Negative Size Bias	12.1324	0.1054	0.11%
Positive Size Bias	23.7186	3.45E-04	0.54%

Formally, we test:

H_0 : absence of asymmetry (skewness), absence of leverage or negative leverage

H_1 : presence of at least one of the following: negative or positive skewness, leverage or negative leverage

We already know (Table 10 in part one) that GARCH-VT is quite effective at reducing the asymmetry (skewness ≈0.01), so we do not expect the fit of the regression to be high.
However, results show that there may still be room for improvement on the symmetric model. About 0.5% of the variability of squared GARCH-VT shocks is explained by the regression against the variables (Table 6). The estimated R^2 is statistically significant and it did not occur by chance (the F-test p-value is quite small) so we can safely reject the null hypothesis at $\alpha = 1\%$. Among the variables included in the regression, only positive size is both relevant and statistically significant, as expected: its p-value is small, so that we

are confident that its coefficient will not oscillate between negative and positive values (99% confidence interval: 6.66; 40.78), and its semi-partial R^2 is quite high, with a sizeable impact on the regressed variable. For instance, an increase in the value of positive size by one should cause future squared shocks to go up by $23.72 \times 0.54\% \approx 12.74\%$, all else constant. The presence of positive size bias also implies the existence of negative leverage. By contrast, positive sign and negative size are neither statistically significant nor relevant. Both variables are subject to great estimation error, so that it is impossible to reject the null hypothesis based on them alone. Moreover, the positive sign coefficient has puzzling sign (we would expect positive residuals to be followed by ones of equal sign), although the coefficient is so unreliable that is may oscillate between negative and positive values (99% confidence interval: –0.61; 0.07).

Hence, given both the statistical significance and the relevance of the positive size variable, we reject H_0, the null hypothesis of absence of residual asymmetry in the distribution of GARCH-VT shocks, in favour of H_1, the alternative hypothesis of presence of the following: positive skewness, negative leverage.

GARCH-VT standardised t model

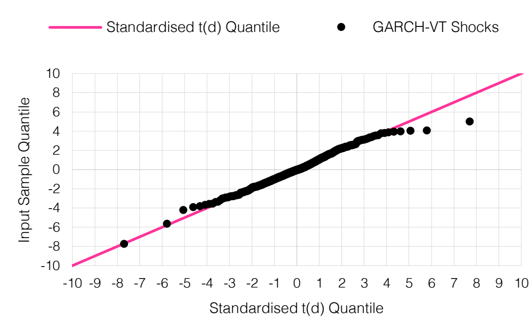
Before considering asymmetric extensions, we briefly review the GARCH-VT standardised t distribution model presented in part one of this risk report. GARCH-VT is a conditional variance model fully specified by three parameters: α and β, estimated via maximum likelihood (a procedure which looks for the values of the parameters most likely to have generated the sample at hand), and ω,

Table 7: GARCH-VT parameter estimation	
ω (omega)	1.16E-05
α (alpha)	0.1593
β (beta)	0.7734
Persistence	0.9328
Maximised log-likelihood	9444.54
R-squared	6.65%
N. estimated parameters	2
AICc	–18885.08
BIC	–18873.54

Table 8: GARCH-VT shocks statistics	
Mean	–0.0223
Variance	0.9986
Standard deviation	0.9993
Skewness	0.0294
Excess kurtosis	3.6721

Table 9: GARCH-VT standardised t parameters	
Degrees of freedom	4.08
Maximised log-likelihood	–3239.81
N. estimated parameters	1
AICc	6481.62
BIC	6487.39

Figure 7: QQ-plot of GARCH-VT shocks against the standardised t distribution



computed using the previous two as well as the sample variance (variance targeting).
The output of the model is a set of conditional variances useful to create shocks – standardised residuals (demeaned returns) or standardised returns (if the mean is approximately zero) – which are then fit to a symmetric density centered at the origin, the standardised t distribution.
We produce updated statistics to the model (Tables 7-9). ω = 1.16E-05, α = 0.1593, and

β = 0.7734. The goodness of fit of conditional variances to squared returns, the R^2, is equal to 6.65%, in line with empirical evidence (5-10%). Shocks have smaller higher moments than returns (skewness is down to 0.0294 from 0.2666, excess kurtosis down to 3.6721 from 11.7589) and their distribution is well-described by a standardised t density with 4.08 degrees of freedom (Figure 7).

Overall, the model fully captures some non-normality features of TOP-20 Index (autocorrelation of log and squared returns, volatility clustering) and partially some others (positive skewness, negative leverage, positive excess kurtosis). We previously tried to minimise the latter features, fitting an asymmetric t distribution to GARCH-VT shocks, though we obtained a worse result, in terms of likelihood ratio: GARCH is designed to respond equally to negative and positive news, so it is best paired with a symmetric distribution.

As the asymmetric t seems to be a powerful model in light of the stylised characteristics of TOP-20 Index, we now look for non-linear extensions of GARCH which can provide a better fit to sample data.

Asymmetric GARCH models

We present two popular GARCH extensions best suited to capture the different response of conditional volatility to good and bad market news: EGARCH and GJR-GARCH.

For ease of comparison with the previous model, we use variance targeting for both.

EGARCH-VT

EGARCH[c] is a model for the logarithmic variance, and it is completely specified by the following parameters: γ and β, comparable to α and β in GARCH, though not subject to the same restrictions (the only requirement is $\beta < 1$), ω, whose value, when exponentiated, is similar to that of ω in GARCH, and ψ, an additional variable that captures the leverage (or negative leverage) effect. ψ is obviously the most interesting parameter and, in EGARCH:

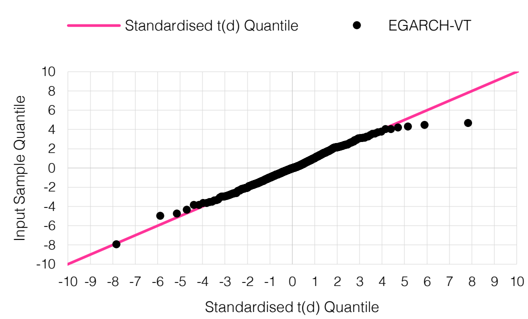
$\psi < 0$: presence of leverage effect

Table 10: EGARCH-VT parameter estimation	
ω (omega)	–0.5392
γ (gamma)	0.2250
β (beta)	0.9378
ψ (psi)	0.0667
Persistence	0.9378
Maximised log-likelihood	9451.63
R-squared	9.68%
N. estimated parameters	3
Likelihood ratio test, 99%	14.16 (6.63)
Improves on GARCH-VT?	YES
AICc	–18897.25
BIC	–18879.92

Table 11: EGARCH-VT shocks statistics	
Mean	–0.0335
Variance	1.0280
Standard deviation	1.0139
Skewness	–0.0556
Excess kurtosis	3.6477

Table 12: EGARCH-VT standardised t parameters	
Degrees of freedom	4.24
Maximised log-likelihood	–3277.34
N. estimated parameters	1
AICc	6556.69
BIC	6562.46

Figure 8: QQ-plot of EGARCH-VT shocks against the standardised t distribution



$\psi = 0$: symmetric response to news, equivalent to GARCH

$\psi > 0$: presence of negative leverage effect

We find the optimal values of γ, β, and ψ via maximum likelihood (MLE), then use them to retrieve that of ω (Table 10). $\gamma = 0.2250$, $\beta = 0.9378$, $\psi = 0.0667$, $\omega = -0.5392$. A positive value for ψ implies the presence of negative leverage, as expected.

Considering the regression R^2, the likelihood ratio, and the information criteria (AICc and BIC), EGARCH appears to be a superior model. The R^2 is about three percentage points higher than that of GARCH-VT, the likelihood ratio is positive and significant ($\alpha = 1\%$), and both AICc and BIC are smaller than those of the symmetric model. AICc and BIC are, respectively, the Akaike Information Criterion corrected and the Bayesian Information Criterion, two measures of the quality of a statistical model relative to others, based on the log-likelihood function and on the number of estimated parameters. The best model is the one minimising the information criteria: basically the one with the largest value of the log-likelihood function and the smallest number of parameters (lack of parsimony is penalised).

Despite the improved fit, EGARCH is not an entirely satisfactory model. For example, the skewness of the distribution of shocks is now negative and larger than before (–0.0550), while excess kurtosis is only mildly reduced (3.6477, down 0.67%) (Table 11). Moreover, the t density fit is worse than that for GARCH both in the symmetric (Table 12, Figure 8) and in the asymmetric (Table 13, Figure 9) cases (for asymmetric GARCH-VT, AICc = 6465.34; BIC = 6476.89). As expected, however, EGARCH-VT asymmetric t has a better

Table 13: EGARCH-VT asymmetric t parameters	
d_1	4.20
d_2	0.08
C	0.5196
A	0.1137
B	1.0030
–A/B	–0.1133
Maximised log-likelihood	–3270.59
N. estimated parameters	2
AICc	6545.17
BIC	6556.72

Figure 9: QQ-plot of EGARCH-VT shocks against the asymmetric t distribution

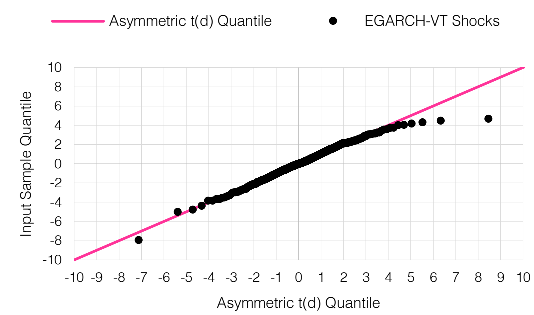


Table 14: GJR-VT parameter estimation	
ω (omega)	1.07E-05
α (alpha)	0.1955
β (beta)	0.8022
θ (theta)	–0.1194
Persistence	0.9380
Maximised log-likelihood	9461.39
R-squared	8.72%
N. estimated parameters	3
Likelihood ratio test, 99%	33.70 (6.63)
Improves on GARCH-VT?	YES
AICc	–18916.79
BIC	–18899.46

Table 15: GJR-VT shocks statistics	
Mean	–0.0341
Variance	0.9762
Standard deviation	0.9880
Skewness	0.0055
Excess kurtosis	3.1453

fit than its symmetric counterpart (likelihood ratio: 13.52; AICc: –11.52; BIC: –5.74).

Overall, we do not believe EGARCH-VT to be a very good model for TOP-20 Index.

GJR-VT

GJR[d] is an asymmetric GARCH model completely defined by the following parameters: ω, α, β (as in GARCH), and θ. The latter is the leverage parameter and, contrary to the EGARCH case:

$\theta < 0$: presence of negative leverage

$\theta = 0$: symmetric response to news

$\theta > 0$: presence of leverage

As before, we find the optimal values for α, β, and θ via maximum-likelihood estimation, then derive ω (Table 14). The following restrictions apply (on top of the usual ones): $\alpha + \theta \geq 0$, $\alpha + \beta + 0.5\,\theta < 1$ (persistence). $\alpha = 0.1955$, $\beta = 0.8022$, $\theta = -0.1194$, and $\omega =$ 1.07E-05. A negative value for θ implies the presence of negative leverage, as expected. The regression R^2 is two percentage points higher than that of GARCH, although lower than in the EGARCH case. Yet, the likelihood ratio test, as well as the information criteria, show that GJR is a much nicer fit to the data. The higher moments of the shocks, for example, are greatly reduced: in absolute value, skewness is down 81% while excess kurtosis 14%, compared to their previous values in GARCH (Table 15).
The distribution of shocks is fit to both the symmetric (Table 16) and the asymmetric (Table 17) t densities. In the first case the fit is unsatisfactory, with shocks in the left tail visibly offset (Figure 10) – we tried to fit an asymmetric model to a symmetric distribution. In the second, however, the fit is excellent, with shocks in the left tail almost in a straight line (Figure 11).

Table 16: GJR-VT standardised t parameters	
Degrees of freedom	4.09
Maximised log-likelihood	–3224.31
N. estimated parameters	1
AICc	6450.62
BIC	6456.39

Figure 10: QQ-plot of GJR-VT shocks against the standardised t distribution

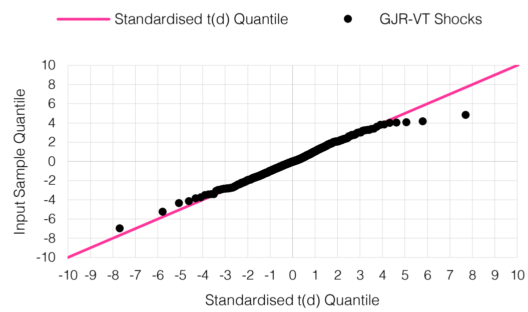
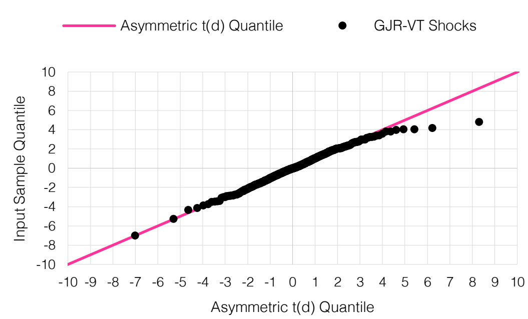


Table 17: GJR-VT asymmetric t parameters	
d_1	4.05
d_2	0.08
C	0.5273
A	0.1117
B	1.0031
–A/B	–0.1114
Maximised log-likelihood	–3217.67
N. estimated parameters	2
AICc	6439.34
BIC	6450.89

Figure 11: QQ-plot of GJR-VT shocks against the asymmetric t distribution



Model selection

We go back to the model selection criteria to look for the best model, among GARCH-VT, EGARCH-VT, and GJR-VT for TOP-20 Index. We recall we would like such a model to display the highest likelihood ratio, as well as the lowest values for the information criteria.

Conditional volatility model – In terms of likelihood ratios, GARCH < EGARCH < GJR. In terms of both AICc and BIC, the same ranking applies. Winner: GJR.

Distribution model – In terms of likelihood ratios, EGARCH < GARCH < GJR. Also, symmetric t GJR < asymmetric t GJR. In terms of both AICc and BIC, the same ranking applies. Winner: asymmetric t GJR (Table 18).

News impact curve

We previously stated that the negative leverage effect is among the most important anomalies of TOP-20 Index, and we verified its existence through the optimal values of the leverage parameters: $\psi = 0.0667$ (positive) and $\theta = -0.1194$ (negative). A visual representation of this interesting phenomenon is the news impact curve by Engle and Ng[b]. The curve analyses the impact that news, proxied by shocks (standardised residuals), have on conditional variance the next day. It is a parabola centered at the origin, and with two branches whose steepness determines the intensity of the reaction to such news: the steeper the branches, the higher the predicted value of tomorrow's conditional variance. The news impact curve for GARCH is symmetric because the model equally reacts to positive and negative news: what really impacts on volatility is the magnitude, not the sign of the previous day's shock (Figure 12). As a consequence, GARCH cannot capture the positive skewness and the negative leverage of TOP-20 Index returns. By contrast, the curves for EGARCH and GJR are asymmetric, with a steeper (GJR) or much steeper (EGARCH) right branch: what influences volatility are not only the size but also the sign

Table 18: Model selection chart

Conditional volatility model						
	LR		AICc		BIC	
GARCH-VT						
EGARCH-VT						
GJR-VT	✓		✓		✓	
Distribution model						
	LR		AICc		BIC	
	S	A	S	A	S	A
GARCH-VT						
EGARCH-VT						
GJR-VT		✓		✓		✓

Figure 12: News impact curve

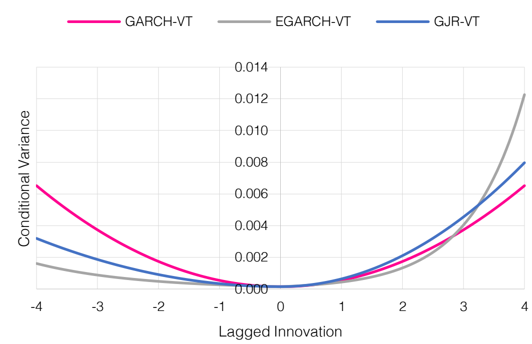


Table 19: News impact curve intercept

GARCH-VT	0.000145
EGARCH-VT	0.000144
GJR-VT	0.000149

of the previous day's shock, and large, positive shocks have a much bigger impact. Depending on the values of the parameters, the curves are rotated, translated and expanded. Rotation is dictated by the leverage parameters ψ and θ, translation depends on ω, β, γ, and sample variance, and expansion is mainly determined by α and γ. Compared to the GARCH curve, the asymmetric ones are very mildly translated (Table 19), rotated to the left (due to the negative leverage effect), and expanded. In particular, the EGARCH curve develops in an exponential fashion, a behaviour most visible in the right branch of the parabola.

One-month ahead forecast

We are now going to provide a one-month ahead forecast of the distribution of TOP-20 Index returns using the models discussed in this report. We pair the GJR-VT conditional volatility model with both filtered historical simulation and Monte Carlo asymmetric t.
We generate 100,000 scenarios, and make no assumptions regarding the future level of volatility (volatility multiplier: 1.0). The current level, as of February 24, 2017, is σ = 1.37% (σ^2 = 0.00188).

Filtered historical simulation, GJR-VT

FHS, which we introduced in part one of this risk report, combines a model-based method of conditional variance with a model-free method of distribution [2]. Here, we assume that conditional variance is correctly specified by GJR-VT, and that the historical density of TOP-20 Index returns is a valid proxy for the future. The steps of the process are only mildly changed with respect to those in part one:

1. We create 21 series of 100,000 discrete RV in [1, sample size]. Each series represents a day, each RV one of the daily returns in the sample;
2. From the set of GJR-VT shocks, we extract the ones whose indices correspond to the discrete numbers included in the first series;
3. We multiply each shock by the level of volatility predicted for the day, to obtain a series of simulated returns;
4. We calculate the vector of GJR-VT conditional variances (with α, β, ω, and θ already estimated) using the simulated squared returns and the predicted level of volatility;

Table 23: FHS investment statistics, %	
No. of simulations	100,000
Average return	–0.9442
Volatility	5.9571
Skewness	0.4811
Excess Kurtosis	3.8157
Neg./pos. return frequency	59.16; 40.84
Average if neg./pos.	–4.64; 4.41
Freq. returns beyond ±10	4.83; 3.77
Average if beyond ±10	–13.19; 14.52

Figure 18: Distribution of FHS GJR-VT monthly returns

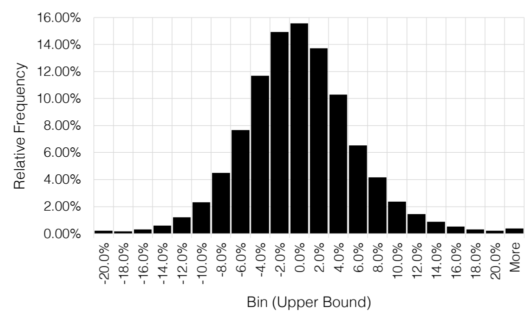
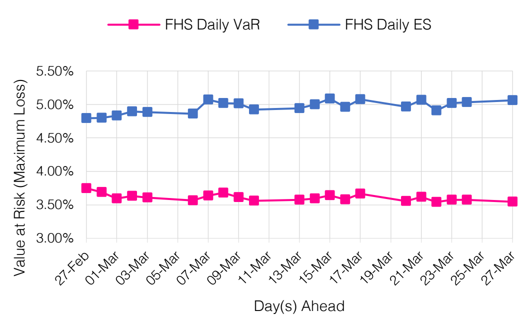


Table 24: FHS monthly VaR and ES	
Value at Risk, 99%	14.78%
Expected Shortfall, 99%	18.54%
Portfolio Value	MNT 1,000,000
MNT Value at Risk	MNT 137,399.58
MNT Expected Shortfall	MNT 169,211.38

Figure 19: FHS GJR-VT term structures of VaR and ES



5. We repeat each step 21 times, then aggregate the results to obtain a series of monthly returns.

6. We compute monthly Value at Risk and Expected Shortfall, together with their term structures.

The model captures several stylised facts of TOP-20 Index: the slightly negative mean return (–0.94%), a level of volatility in line with current market figures (~6%), positive skewness (0.48), and positive excess kurtosis (3.82). Of negative leverage, it considers the higher frequency of losses (60-40), and the larger average extreme gain (~1.5% more, in absolute value, than the average extreme loss), and does not consider the higher frequency of extreme gains (here, ~1% lower than that of extreme losses), and the smaller average loss (here, ~0.25% higher than the average gain). One month ahead Value at Risk, the potential loss that would only be exceeded 1% of the times in the period, is forecast at 14.78% (Table 24), between 3.55% and 3.75% per day (Figure 19). MNT VaR is around MNT 137,400 for each million invested in the portfolio (~EUR 52.60). One month ahead Expected Shortfall, the average loss given that the actual return is worse than the VaR, is forecast at 18.54%, between 4.80% and 5.10% per day. MNT ES is around MNT 169,200 per million invested, (~EUR 64.70).

Monte Carlo simulation, GJR-VT asymmetric Student's t distribution

Monte Carlo generates new, hypothetical returns from a predetermined density calibrated on sample data. We use the asymmetric t distribution, and assume that the conditional variance is correctly specified by GJR-VT. Due to the asymmetric nature of the density, we need to vary the simulation steps. Using the inverse transform method [3]:

1. We create 21 sets of 100,000 uniform

Table 25: Monte Carlo investment statistics, %	
No. of simulations	100,000
Average return	0.0096
Volatility	6.3085
Skewness	0.9768
Excess Kurtosis	16.6698
Neg./pos. return frequency	53.15; 46.85
Average if neg./pos.	–4.24; 4.83
Freq. returns beyond ±10	3.35; 5.11
Average if beyond ±10	–13.39; 15.70

Figure 20: Distribution of Monte Carlo GJR-VT asymmetric t monthly returns

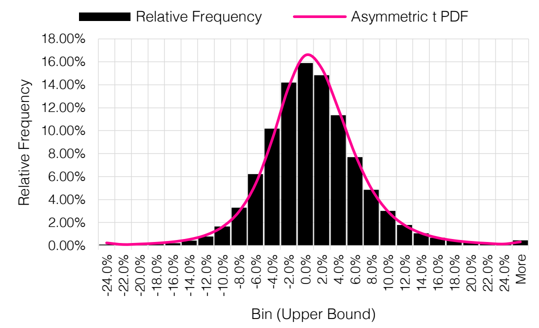
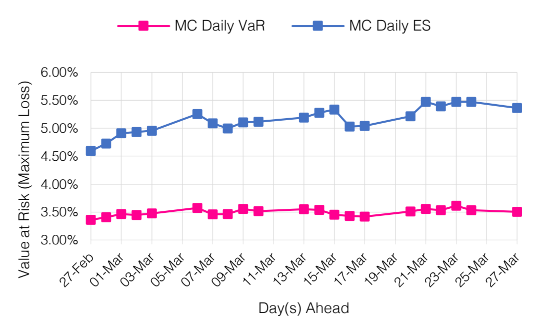


Table 26: Monte Carlo monthly VaR and ES	
Value at Risk, 99%	13.64%
Expected Shortfall, 99%	17.91%
Portfolio Value	MNT 1,000,000
MNT Value at Risk	MNT 127.467.84
MNT Expected Shortfall	MNT 163,961.31

Figure 21: MCS GJR-VT asymmetric t term structures of VaR and ES



RV – variables which randomly take any of the values in [0,1] with equal probability. We interpret these RV as the quantiles of a cumulative distribution function;

2. We transform the data into asymmetric t quantiles, with different formulae depending on whether they fall below or above $(1 – d_2)/2$, with $d_2 = 0.08$ so that the threshold value is 0.4606. These quantiles are equivalent to asymmetric t shocks;
3. We calculate the vector of returns one day ahead by multiplying the first set of shocks with the most recent GJR-VT figure;
4. We build the vector of GJR-VT conditional variances using the squared simulated returns and the predicted level of volatility;
5. We repeat each step 21 times, then aggregate the figures to obtain a set of monthly TOP-20 Index returns;
6. We compute monthly VaR and ES, together with their term structures.

With the exception of the negative mean return, which is nevertheless too noisy to be correctly estimated, the model preserves all the stylised facts of TOP-20 Index: a level of volatility in line with current market figures (~6.3%), positive skewness (0.98), positive excess kurtosis (16.67), and negative leverage, which is visible in the higher frequency of losses (53-47), the larger average gain (~0.6% more, in absolute value, than the average loss), the higher frequency of extreme gains (~1.75% higher than that of extreme losses), and the larger average extreme gain (~2.3% more than the average extreme loss) (Table 25, Figure 20). One month ahead Value at Risk is forecast at 13.64% (Table 26), between 3.40% and 3.60% per day (Figure 21). MNT VaR is about MNT 127,500 per million invested (~EUR 48.80).

One month ahead Expected Shortfall is forecast at 17.91%, between 4.70% and 5.50% per day. MNT ES is about MNT 164,000 per million (~EUR 62.80). These values are lower than those predicted by FHS because Monte Carlo asymmetric t better captures the positive skewness/negative leverage phenomena of TOP-20 Index.

Conclusion

We surveyed, and incorporated into our risk models, the most notable empirical features of TOP-20 Index: autocorrelation of log and squared returns, volatility clustering, positive skewness, negative leverage, and positive excess kurtosis.

We showed that the simultaneous presence of all these features may be beneficial for investors thanks to the greater chance of extreme gains, as opposed to extreme losses, and interpreted it as evidence of euphoria on the stock market.

Among the risk models proposed, we found GJR-VT FHS and GJR-VT asymmetric t to provide the best fit to the data, so we used both to forecast the distribution of TOP-20 returns one month from now.

Overall, we do believe TOP-20 Index to be a very nice investment opportunity at the moment.

The report is made for Standard Investment LLC by Federico M. Massari, a long distant volunteer risk analyst, using the sources provided.

---

* All data from mse.mn. We modified the value of the close price recorded on August 13, 2010 from MNT 11,145.50 to MNT 10,145.50; the previous datum was most likely an outlier resulting from transcription error.

Bibliography

[1] Jondeau E., Poon, S.-H., Rockinger, M.: *Financial Modeling Under Non-Gaussian Distributions*, 2007, Springer Finance.

[2] Christoffersen, P.F.: *Elements of Financial Risk Management*, 2nd Ed., 2012, Academic Press, Elsevier. Book and companion material.

[3] Glasserman, P.: *Monte Carlo Methods In Financial Engineering*, 2003, Springer Applications of Mathematics.

[4] Massari, F.M.: *Mongolia's TOP-20 Index Risk Analysis*, 2017, Standard Investment LLC. Available at *standardinvestment.mn/en*

[5] Massari, F.M.: *Should You Add TOP-20 To Your Asset Mix?*, 2017, Standard Investment LLC. Full report and Investors' edition available at *standardinvestment.mn/en*

[6] *Mongolia Investors Cockpit, January 2017*, Standard Investment LLC. Available at *standardinvestment.mn/en*

Additional resources

[a] Jaggia, S., Kelly-Hawke, A.: *Modelling skewness and elongation in financial returns: the case of exchange-traded funds*, 2009, Applied Finance Economics, vol. XIX, no. 16, pp. 1305–1316

[b] Engle, R.F., Ng, V.K.: *Measuring and Testing the Impact of News on Volatility*, 1993, The Journal of Finance, vol. XLVIII, no. 5, pp. 1749–1778

[c] Nelson, D.B.: *Conditional Heteroskedasticity in Asset Returns: A New Approach*, 1991, Econometrica, vol. LIX, no. 2, pp. 347–370

[d] Glosten, L.R., Jagannathan, R., Runkle, D.E.: *On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks*, 1993, The Journal of Finance, vol. XLVIII, no. 5, pp. 1779–1801

Contacts

Federico M. Massari
Long Distant Volunteer Risk Analyst
federico.massari@libero.it
Tel. +39 340 1011568

Standard Investment, LLC
Jigjidjaw St. 5/3, 1st khoroo, Chingeltei district
Ulaanbaatar, Mongolia

Postal Address:
PO Box 1487, Central Post Office
Ulaanbaatar 15160

Tel. +976 7011 4433
info@standardinvestment.mn

Disclaimer

Investors act on their own risk. This is not an offer to buy or sell or the solicitation of an offer to buy or sell any security/instrument or to participate in any particular trading strategy.

All information in this report is for general information only. The information is derived from sources which the company believes to be reliable and prepared in good faith. Standard Investment LLC makes no guarantee of accuracy, timeliness and completeness of the information. Neither Standard Investment nor its affiliates shall be liable for any damages arising out of any person's reliance upon this report.