AWALEE NOTES

Dynamic of agricultural commodity: A focus on extreme soybeans price variation via Extreme Value Theory

Study carried out by the Quantitative Practice,
special thanks to Xiaoying Huang

awalee
FINANCIAL CONSULTING

JUNE 2018

Tables of content

1. INTRODUCTION

The commodity markets develop very extensively through the derivatives, such as futures and options contracts, Trackers index and other OTC derivatives. Among those, agricultural commodities occupy a special role in the economy. They are specified by the fact that agricultural commodity prices are highly dependent on the environment and result in social consequence. The weak elasticity of demand, climate risks, exchange rate risks in global markets, make agricultural commodity prices highly volatile. The agricultural commodities have risen since early-2000 and reached the peak in mi-2008 and the second half of 2012. The recent food price spikes are explained by many factors, noting macroeconomic development, which affects the structure changes, new biofuel demands, record oil prices etc. In the last few years, the fundamentals do not seem enough to explain these price movements.

In the policy and academic debate, the phenomenon labelled as 'financialization' of commodity market is blamed for the high and volatile prices or even extreme variation. As other commodity markets, energy, precious metal markets, agricultural commodities become one of the assets that have been widely used for the portfolio diversification in recent years. The introduction of new financial derivatives promotes not only the hedging activities for producers and consumers, but also makes the risk management more complicated. With the derivatives exchanged in the financial markets, agricultural commodities are reacting to the news fastly. The extreme price evolution or the tail part of both sides of price returns has impact on the risk management decision.

In the modelling or the risk management of agricultural commodities, advanced models are developed to capture the new features of commodities prices. To capture the fat tail and high volatility, relative volatility model GARCH is implemented. More specifically, in the stochastic process, the integration of jump and stochastic volatility models tries to tackle the fat tail distribution. The extreme variation or the tail in the distribution becomes a focus of recent risk management. These extreme price changes expose the producers and consumers to significant risks. Among those methods, the extreme value theory is useful to improve the examination of extreme quantiles and focuses on the tail of the price distribution.

The risk measures, Value-at-Risk, Expected-shortfall are all concerned on the tail of returns distributions. As a result, a usual application of extreme value theory in the finance is to compute these risk measures and to provide an effective risk management. This note discusses the extreme variation of agricultural prices via Extreme Value Theory.

2. A BRIEF RELATIVE LITERATURE

Regarding the features of agricultural commodity markets, time series models and stochastic models have been developed to stimulate the agricultural dynamics. These models allow considering the mean-reversion, skewness right, excess kurtosis, and discontinuity etc., all of which are often observed from agricultural commodity prices.

In the reaction to the imbalance between supply and demand, the prices revert to an equilibrium level, and during this process storage plays an important role. The mean-reversion is more likely to occur in agricultural commodities markets than other commodities. The Ornstein-Uhlenbeck model by Ornstein and Uhlenbeck (1930) captures the mean-reversion. In most of its extended stochastic models, convenience yield is introduced and is defined as the revenue by conceiving the stocks. The convenience yield is considered as an indicator of storage.

In commodity market, volatility can be modelled by general volatility models such as ARCH by Engle (1982). Those volatility models are extended and focus on the volatility behaviour. The application of those models can be referred to Beck (2001).

The observation of discontinuity of prices returns (jumps/spikes) exists in different kinds of commodities. The most widely discussed is the spikes in electricity prices. In agricultural commodities markets, the jumps in price distribution are found. The traditional Merton jump model is applied in the pricing. One can also refer to Huang (2017) for the application of a Double-exponential Jump model in wheat prices.

The modelling methods cited above for agricultural commodities are usually applied on the pricing of agricultural derivatives or on the computation of risk measures. Different from the pricing derivatives, the interests in more accurate risk management need a superior quality of worst-case scenario evaluation, for which EVT has its specific focus. The extreme value theory permits to examine the stochastic behaviour of the extreme value in a single process. EVT is usually applied in other financial assets prices tail modelling.

This note introduces the fundamentals of agricultural commodities markets and discusses the implication of EVT in the risk management of agricultural markets. Section 3 introduces the extreme value theory. Section 4 shows the empirical application of EVT in risk management of soybean markets. Section 5 concludes.

3. EXTREME VALUE THEORY

The first method of applying EVT is the known as the Block Maxima (Minima) approach, which is based on the extreme value distributions of the Gumbel, Frechet or Weibull distributions. This method fixes a block of maxima or minima in a time series, which is independent and identical distributed. The second method of applying EVT is the Peak Over Threshold (POT) approach. The POT approach fixes a threshold over which the extreme value is fixed to a Generalized Pareto Distribution (GPD). This approach is more efficient as it does not require large data sets as the Black Maxima (Minima) approach. In this note, the second method is applied, because of the advantage that the POT approach uses available data more efficiently.

3.1. Excess distribution

(Balkema and de Hann 1974 and Pickands, 1975) For a large class of underlying distributions F, the excess distribution function F_u can be approximated by GPD for an increasing large enough threshold u.

$$F_u(y) \approx G_{\varepsilon,\sigma}(y), u \to \infty$$

Where, $G_{(\varepsilon,\sigma)}$ (y) is the Generalized Pareto Distribution (GPD):

$$G_{\varepsilon,\sigma}(y) = \begin{cases} (1 + \frac{\epsilon}{\sigma} y)^{\frac{1}{\epsilon}}, \text{ if } \epsilon \neq 0 \\ 1 - e^{\frac{y}{\sigma}}, \text{ if } \epsilon = 0 \end{cases}$$

In this equation, the ε is the shape parameter and determines the heaviness of the distribution tail. And the σ is the scale parameter of GDP.
And for $y \in [0, (\chi_F - u)], if \epsilon < 0$ *and* $y \in [0, -\frac{\epsilon}{\sigma}], if \epsilon < 0$

$F_u(y)$ is the conditional excess distribution of a threshold u and defined as:

$$F_u(y) = P(X - u \leq |X > u)) = \frac{F(y+u) - F(u)}{1 - F(u)} = \frac{F(x) - F(u)}{1 - F(u))}$$
for $0 \leq y \leq \chi_F - u$

With the density function F of random variable X. $\chi_F < \infty$ is the right endpoint of the support of the unknown population distribution F of X.

3.2 Maximization of likelihood function

The parameters are computed by maximizing the likelihood function. The log-likelihood function of n observations derived from the above function is:

$$L(\epsilon, \sigma|y) = \begin{cases} -nlog\sigma - (\frac{1}{\epsilon} + 1) \sum_{i=1}^{n} log(1 + \frac{\epsilon}{\sigma} y_i), \text{ if } \epsilon \neq 0 \\ -nlog\sigma \ - \frac{1}{\sigma} \sum_{i=1}^{n} y_i, \text{ if } \epsilon = 0 \end{cases}$$

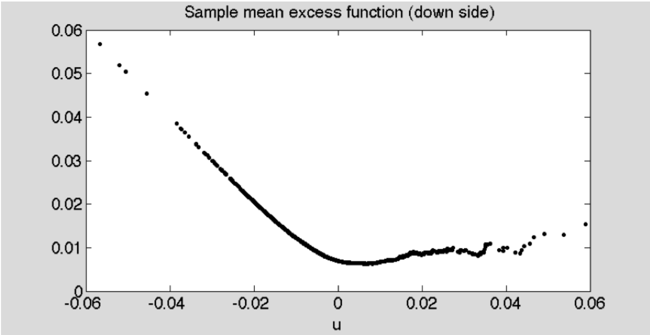
For applying the POT approach, a proper threshold is needed to fix the GPD to the soybean data. There is a trade-off between the bias and variance of the maximum likelihood estimates of the parameters. EVT needs a large value of u to minimize the bias so that the conditional excess distribution function satisfies the convergence of the GPD. However, with a larger threshold value u, there are fewer observations in the tail, as a result the estimation efficiency declines. Graphical tool can be used to choose the threshold, such as the Hill Graph, mean excess graph (Picklands Balkema-In Haan). In this note, mean excess graph and an intuitive criteria will be applied in the choosing of threshold u. Various threshold values are applied to estimate the likelihood function and the threshold value is kept with which the parameters estimates are stable.

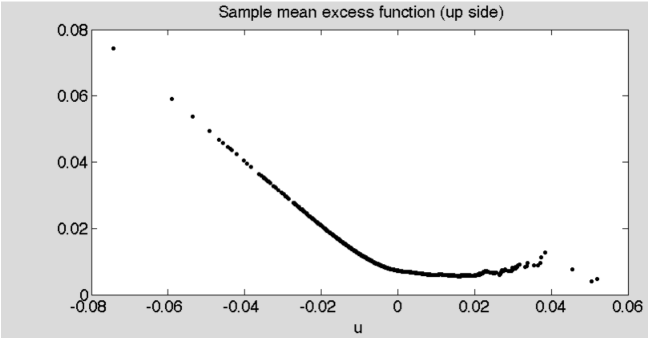
3.3 Sample mean excess plot

In order to chose an appropriate threshold u, the sample mean excess plot shows the couple of points: $(u, e_u(u))$.

With $e_u(u) = \frac{\sum_{i=k}^{n} (X_i^n - u)}{n - k + 1}, k = min(i|X_i^n > u)$

And $n - k + 1$ is number of observation that exceeds the threshold u.
The sample mean excess function should be linear. As a result, the value of threshold is located at the end of the linear curve in the sample mean excess plot.
The following graph shows the sample excess plot for the soybean futures returns. As we are working on the absolute value of two extreme observations, according to the graph, for the downside tail, -0.015 and for the up side tail, 0.02 would be applied to select the tail observations to compute the extreme value distribution.





4. TAIL BEHAVIOUS ANALYSIS OF SOYBEAN MARKET VIA EVT

The focus of this note is the nearest soybean futures contracts traded in the Chicago Mercantile Exchange (CME). The data range is from 04/01/1999 to 04/01/2017. The data returns are calculated using the first two contracts to eliminate the artificial jump during the changes of contracts.

4.1. Primary observation

In the following graph, soybean prices and returns are presented by red and blue curves respectively. The price of soybean increases since 2008. During 2009 to 2012, the prices attend to about 35 Dollars per ton. For single year 2010, the price dropped 15 dollars from the beginning to the mid-2010 and climbed 20 dollars until the end-2010. It should be noted that this period coincides with the high volatility period by looking at the variation of prices returns, which is usually noted as volatility clustering. The return could increase by 4% or even decrease 6% during one day.

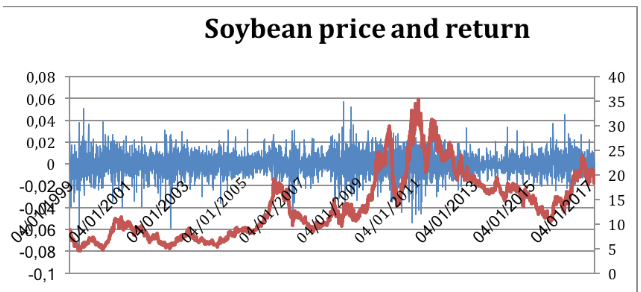


Figure 1: Price and Return of Soybean market

Normal distribution is a common hypothesis on the return evolution. A QQ-plot between sample data and normal distribution is drawn in the following graph. The fat-fail feature is observed in the last few quantiles. The tail of both upside and downside are visibly deviated far from the normal distribution. The normal distribution fits well in the centre, but neither downside-tail nor upside-tail fits with the normal distribution. For an investor being short on soybean futures, the risk of loss consists on the positive tail. As a result, in the following tail behaviour study, both position and negative extreme variation will be discussed.

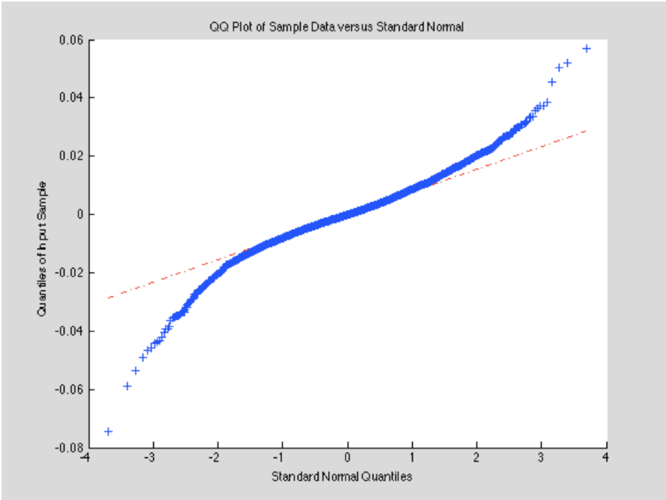


Figure 2: QQ plot of sample data with standard normal distribution

4.2. EVT Parameter estimates of Soybean returns

According to sample mean excess plot, the threshold is chosen with about 90% quantile for the right tail and 10% for the left right. For every estimation, the threshold moves to a more extreme value by 1 basis point. The final parameter and threshold is kept when the value of parameter is stable.

The observations of tails decrease when the absolute value of threshold rises in both directions. The following table shows the parameters estimated in both up- and down-side tails. As it is reminded in the section above, the ε describes the skews of tail distribution and σ measures the kurtosis of tail. Both parameters are estimated by maximizing the likelihood function based on the sample of tail observations.

The first results given in the following table are the parameters with the threshold chosen by sample excess mean plot in both the right and left tails distribution. Confidence intervals are also given for every parameter. Every estimate is inside the confidence interval and is thus significant. By checking simply the value, the left tail is relatively heavier than the right tail for the soybean returns. An extreme decrease of prices could be more probably happened than an extreme increase.

Threshold	ε	σ
0.015	-0.475 [-0.5384;-0.4121]	0.027 [0.0242;0.0310]
-0.02	-0.504 [-0.5830;-0.3855]	0.038 [0.0307;0.04435]

Table 1: Parameter estimates

The estimation is complemented by modifying the threshold value. From the number of observations, the down side tail has more skewness and kurtosis than the up side tail. The downside risk in soybean markets is highlighted after observing the extreme distribution parameter. Both parameters are quite stable; the parameters of table 1 are then retained.

DOWN SIDE

Threshold	ε	σ	Tail observations
0.01478	-0.341	0.020	582
0.011	-0.373	0.022	478
0.012	-0.394	0.023	419
0.013	-0.421	0.024	358
0.014	-0.447	0.026	306
0.015	-0.475	0.027	261
0.016	-0.500	0.029	227
0.017	-0.543	0.031	181
0.018	-0.576	0.033	154
0.019	-0.614	0.035	129

UP SIDE

Threshold	ε	σ	Tail observations
- 0.01	- 0.254	0.020	510
- 0.011	-0.276	0.021	430
- 0.012	-0.299	0.023	361
- 0.013	-0.321	0.025	310
- 0.014	-0.345	0.027	263
- 0.015	-0.377	0.029	215
- 0.016	-0.404	0.031	183
- 0.017	-0.436	0.033	154
- 0.018	-0.468	0.036	132
- 0.019	-0.484	0.037	122
- 0.02	-0.504	0.038	111

Table 2: Estimates with dynamic threshold

4.3. Application in VaR

As it is noted in the previous section, extreme value theory is often applied in the calculation of risk measures, which usually concentrates on the extreme variation. For this interest, Value-at-risk is calculated with EVT and alternative methods.
The estimate of the VaR for a given probability p can be obtained by the quantile function of the GPD distribution $G_{(\varepsilon,\sigma)}(y)$ and is given as :

$$VaR_p = u + \frac{\sigma}{\epsilon}\left(\frac{n}{Nn}(1-p)^{-\epsilon} - 1\right)$$

u,σ,ε are the same specified variables above.

The VaRs with both parametric method (Variance-Covariance method) and non-parametric approach (Normal distribution) are also calculated. The value p is chosen to be 0.95, 0.975 and 0.999.

The VaR under Extreme-Value Theory is the highest comparing with the other methods, Variance-covariance and Normal distribution. With a probability of 0.05 or 95% level of confidence, the expected losses will not exceed to 5.37% according to the EVT method. However, for the other two methods under normal distribution hypothesis, the loss will not exceed only around 1.6%. As it is expected, EVT captures higher tail risk. The difference is even larger with a 99.9% level of confidence.

5. CONCLUSION

This note has illustrated the application of EVT on the log-returns of soybean futures for assessing the extreme variation. Both graph tool and dynamic threshold are applied to estimate the parameters that control the shape and location of tail behaviour and to guarantee the quality of the model fit.
The evidence of fait tail is shown in both left and right tails of soybean returns. Specifically, the negative extreme variation is highlighted. The application of extreme-value theory in the calculation of Value-at-Risk has shown that EVT has captured more potential risk than normal distributed Value-at Risk. The Extreme-Value Theory could be a complement in the risk management in agricultural industry.

BIBLIOGRAPHY

Balkema, A and de Hann, L (1974). " Residual life time at great age",, Annals of Probability, 2, 792–804.

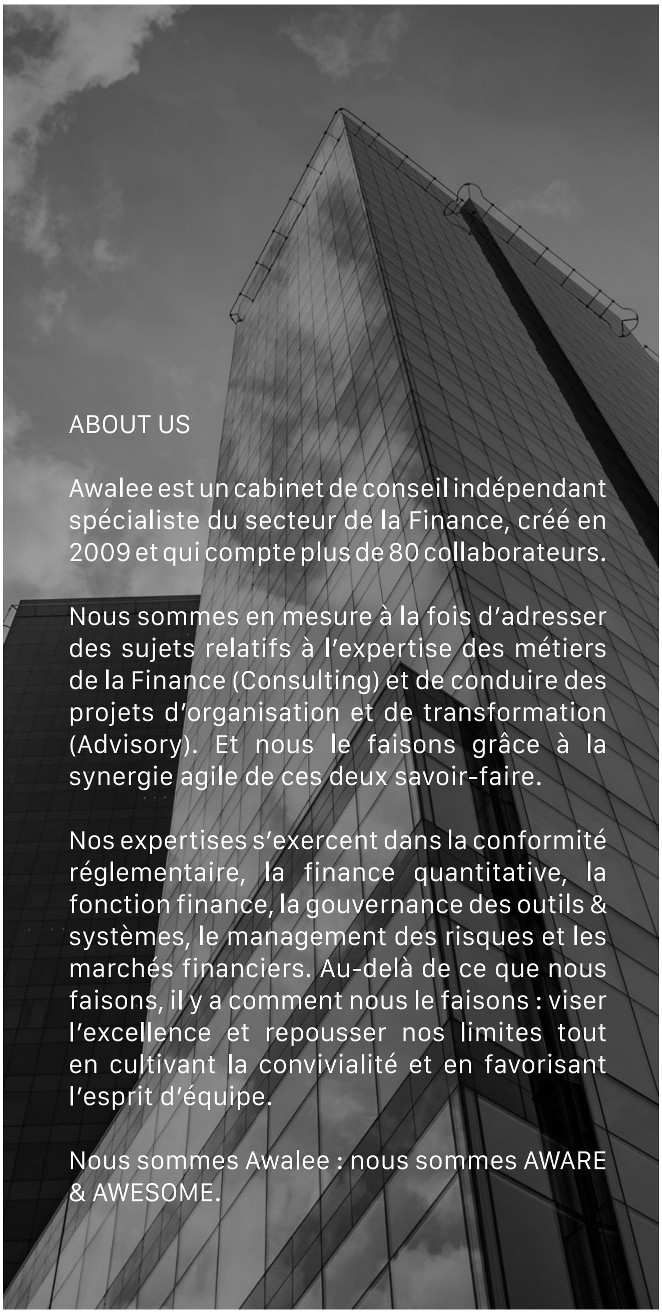
Beck, Stacie(2001). "Autoregressive conditional heteroscedasticity in commodity spot prices" Journal of Applied Econometrics, Vol. 16, Issue 2, pages 115-132.

Engle, Robert F. (1982). "Autoregressive conditional heteroscedasticity with esimates of the variance of United Kindom inflation", Journal of the Economics Society, Vol. 50, No.4, pages 987-1007.

Huang, Xiaoying (2017). "A Double-Exponential Jump model and its application to risk measure in Wheat spot market", Economics Bulletin, Vol. 37, Issue 2, pages 1298-1390.

Ornstein and Uhlenbeck (1930). "On the theory of the Bronian Motion", Physical Review, Vol. 36, Issue 5.

Pickands, J. (1975). «Statistical inference using extreme order statistics», Annals of Statistics, 3, 119–131.



ABOUT US

Awalee est un cabinet de conseil indépendant spécialiste du secteur de la Finance, créé en 2009 et qui compte plus de 80 collaborateurs.

Nous sommes en mesure à la fois d'adresser des sujets relatifs à l'expertise des métiers de la Finance (Consulting) et de conduire des projets d'organisation et de transformation (Advisory). Et nous le faisons grâce à la synergie agile de ces deux savoir-faire.

Nos expertises s'exercent dans la conformité réglementaire, la finance quantitative, la fonction finance, la gouvernance des outils & systèmes, le management des risques et les marchés financiers. Au-delà de ce que nous faisons, il y a comment nous le faisons : viser l'excellence et repousser nos limites tout en cultivant la convivialité et en favorisant l'esprit d'équipe.

Nous sommes Awalee : nous sommes AWARE & AWESOME.

Awalee consulting
59 avenue Marceau
75016 Paris

www.awaleeconsulting.com
twitter.com/awaleeconsulting
linkedin.com/awaleeconsulting