INTERNATIONAL JOURNAL OF ECONOMICS AND FINANCE STUDIES
Vol 4, No 2, 2012 ISSN: 1309-8055 (Online)

STEEL PRICE MODELLING WITH LEVY PROCESS

Emre Kahraman
Türk Ekonomi Bankası (TEB) A.Ş.
Director / Risk Capital Markets Department
emre.kahraman@teb.com.tr

Gazanfer Unal
Yeditepe University Financial Economics
Coordinator / Financial Economics Graduate and PhD Programs
gunal@yeditepe.edu.tr

─Abstract ─
The aim of this study is to model steel price returns by Lévy process. The daily LME Steel Billets Spot Prices between 04.01. 2010 and 31.10.2011 are analyzed and AR[1] ~ GARCH[1,1] discrete model is found to be the best candidate taking all indicators into account. Then the continuous analogue of the discrete model is derived from the discrete model parameters. During the overall study, time (pathwise), distributional and spectral analysis performed. Finally, it is shown that the volatility simulated from both discrete and continuous models shows similar volatility patterns. The results of the study could be utilized to predict the behavior of future steel prices' moves. In addition, the finding could be a good reference specialist and researchers who are interested in steel market.

Key Words: *Steel Modelling, ARMA, GARCH, COGARCH, Lévy Processes, Discrete Time Models, Continuous Time Models, Stochastic Modelling*

JEL Classification: C01, C51

1. INTRODUCTION

The volatility in the commodity market increases the importance of modeling studies in those area and the models which have a success in forecasting the commodity prices receive great attention in the last decade. Geman (2005:4) analyzed commodity market and relationships between commodities in his book. In addition to general commodity market attraction, the steel prices are becoming more and more important nowadays and they directly or indirectly affect the economy in world. If the historical trends are analyzed, high volatilities, upward/downward jumps and drifts could be easily observed. This means that there is no equilibrium in the steel market. Although the researches and analyses on the behavior of steel prices together with modeling studies are increasing, it is also a reality that steel price modelling is scarce and it could be defined as quite new subject in modeling. Those models are really important especially in hedging and risk management purposes as well as in trading.

To be able understand the characteristic of steel market, the essential criteria is to figure out the stochastic models of steel prices. Since the confidence intervals of the models could change with time, the accuracy of the models can be improved with variance of error models. It comes with heteroskedasticity concept in error terms. If the conditional heteroskedasticity of steel prices could be captured with stochastic volatility models (GARCH), then the accuracy of the model could be improved easily. It is also time to mention the importance of time horizon selection in the analysis. It plays a vital role in the appropriate model selection procedure and should be taken into account during all studies.

In financial econometrics, most of the volatility models are in discrete time, namely GARCH models. Those discrete time models have been widely used in various modeling studies to be able to capture the characteristics of financial data. Nelson (1990:7) and Duan (1997:3) studies are only some of those. They tried to model the financial data characteristics by GARCH diffusion approximations. Because of the fact that continuous time models allow closed form solutions, they have advantages with respect to discrete ones and those advantages are tried to be utilized with several studies carried out with continuous models. Klüppelberg (2004:5) worked on a new continuous time GARCH model, namely COGARCH by adapting the single noise process idea.

2. METHODOLOGY

First of all, the stationarity of the data has to be check before starting the discrete modelling part. The trend analysis, Augmented Dickey-Fuller test, autocorrelation and partial autocorrelation functions are utilized to be able to identify whether the data is stationary or not. If the data is not stationary, then either difference or logarithmic difference is applied to make it stationary.

The next step is to execute discrete modeling via “Hannan-Rissanen” algorithm. In this approach, the results of the algorithm are checked with AIC and BIC values to find the best candidate model. The best model should has the lowest

AIC and BIC values. The resulted model is controlled for both stationarity and autocorrelations left in the residuals. The ARCH effects in the residuals will be eliminated with ARCH/GARCH modeling by selecting the best GARCH model. Finally, the non-negativity constraints of GARCH process and the covariance stationary condition are controlled before moving to the continuous modeling.

In the continuous side, discrete model parameters are used to find the parameters for COGARCH. The simulations are carried out for discrete and continuous models and results are compared with each others. Overall, the study is covering not only time analysis, but also distributional and spectral analysis on the data.

Nelson (1990:7) worked on GARCH diffusion approximation with a different way. In the model, there are two different and independent Brownian motions which drive the diffusion, although the process is driven by a single noise sequence. G_t and volatility process σ_t^2 can be represented as:

$$dG_t = \sigma_t dB_t^{(1)} \qquad t \geq 0 \qquad [1]$$

$$\sigma_t^2 = (\beta - \eta\sigma_t^2)dt + \varphi\sigma_t^2 dB_t^{(2)} \qquad t \geq 0 \qquad [2]$$

where β > 0, η ≥ 0, and φ ≥ 0 are constants.

Klüppelberg (2004:5) and Ross (2008:8 / 2009:9) together with all others model COGARCH with a direct analogue of GARCH. The model is based on Lévy process and the model construction is done by taking limits of an explicit representation of the discrete time GARCH process.

The COGARCH process $(G_t)_{t\geq 0}$ is defined in terms of its stochastic differential dG, such that,

$$dG_t = \sigma_t + dL_t \qquad t \geq 0 \qquad [3]$$

$$d\sigma_t^2 = (\beta - \eta\sigma_{t-}^2)dt + \varphi\sigma_{t-}^2 d[L,L]_t^{(d)} \qquad t > 0 \qquad [4]$$

where β > 0, η ≥ 0, and φ ≥ 0 are constants.

$[L,L]_t^{(d)}$ is a quadratic variation process of L (Lévy process) which is defined as;

$$[L,L]_t^{(d)} = \sum_{0<s<t} (\Delta L_s)^2 = \sum_{i=1}^{Nt} V_i \qquad \text{where } \Delta L_t = L_t - L_{t-1} \text{ for t} \geq 0 \qquad [5]$$

The process G jumps at the same time as L (Lévy process) does, and has jump sizes;

$$\Delta G = \sigma_t \Delta L_t \qquad \text{t} \geq 0 \qquad [6]$$

Deriving a recursive and deterministic approximation for the volatilities at the jump times, Klüppelberg (2004:5) shows;

$$\sigma_i^2 = \sigma_{i-1}^2 - \beta + \eta \int_0^t \sigma_s^2 ds + \varphi \sum_{0<s\leq t} \sigma_s^2 (\Delta L_t)^2 \qquad [7]$$

Since σ_s is latent ΔL_s is usually not observable, hence using Euler approximation for the integral we get;

$$\int_0^t \sigma_s^2 ds \approx \sigma_{t-1}^2 \qquad [8]$$

$$\sum_{0<s\leq t} \sigma_s^2 (\Delta L_t)^2 \approx (G_t - G_{t-1})^2 \qquad [9]$$

Therefore, for the volatility estimation, we end up with;

$$\sigma_i^2 = \beta + (1-\eta)\sigma_{i-1}^2 + \varphi(G_t - G_{t-1})^2 \qquad [10]$$

The bivariate process $(\sigma_t, G_t)_{t\geq 0}$ is Markovian. If $(\sigma_t^2)_{t\geq 0}$ is the stationary version of the process with $\sigma_0^2 = \sigma_\infty^2$, then $(G_t)_{t\geq 0}$ is a process with stationary increments.

3. DATA ANALYSIS

3.1. Analysis in Time Domain

In this study, LME Steel Billets Spot Prices (in US dollars) have been analyzed. The analysis has been performed by using the daily close data over the period from January 4, 2010 to October 31, 2011. The data contains 462 observations and the graphical representation is given on Figure-1.

Figure-1: LME Steel Billets Spot Prices / Time Series Plot

Source: Bloomberg / LMFMDY Commodity

It is visually clear from the graph that the price series have a trend and are mean non-stationary. The data is checked for a trend and the result shows that there is a trend which means it is not stationary. [-9673.2+0.251505 t] It is also checked by Augmented Dickey-Fuller test. Test results confirm that the series is not stationary. (Table-1)

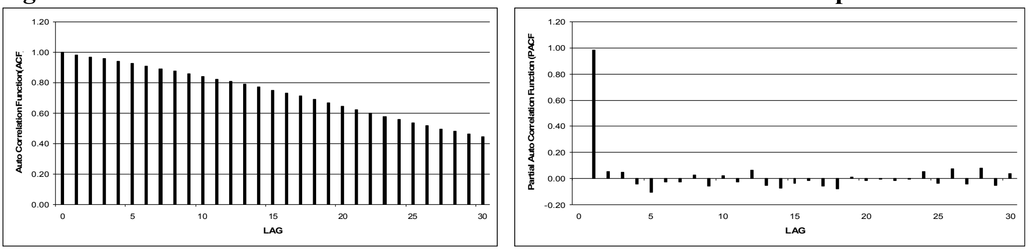
Table 1: Unit Root Test for LME Steel Billets Spot Prices

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.183019	0.2129
Test critical values:	1% level	-3.444311	
	5% level	-2.86759	
	10% level	-2.570055	

Source: Own Study

Finally, autocorrelation and partial correlation function graphs have been given on Figure-2. While autocorrelation values decrease slowly, partial autocorrelation values sharply converge to almost zero levels. The decrease in autocorrelation could be thought as the fact that random shocks to the system dissipate with time. It could be concluded that LME Steel Billets Spot Prices have a trend and so are mean non-stationary. Finally, considering the slow decrease in autocorrelation values, it could be concluded that there is a long memory structure in the data.

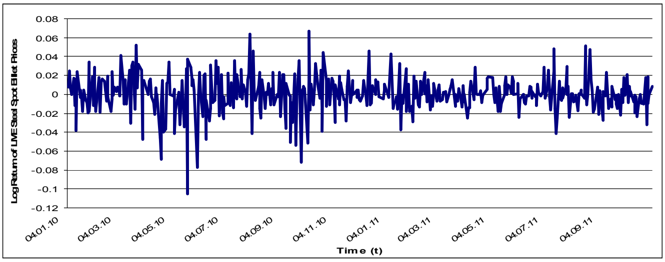
Figure-2: Autocorrelation & Partial Autocorrelation Functions of LME Spot Prices



Source: Own Study

The next step is to make the data stationary. To achieve mean stationary, difference of the series could be used. But, if the series show non-linear trend, the differencing creates non-stationary variance. So, to achieve both mean and variance stationary, first the logarithm should be taken and then difference of the series, which is the logarithmic return. The resulted series is given on Figure-3.

Figure-3: Log Return of LME Steel Billets Spot Prices



Source: Own Study

The logarithmic return of steel price series seems to have no trend. The next step is to check that whether it is a non-stationary or not. Augmented Dickey-Fuller test results confirm that the series is not stationary. (Table-2)

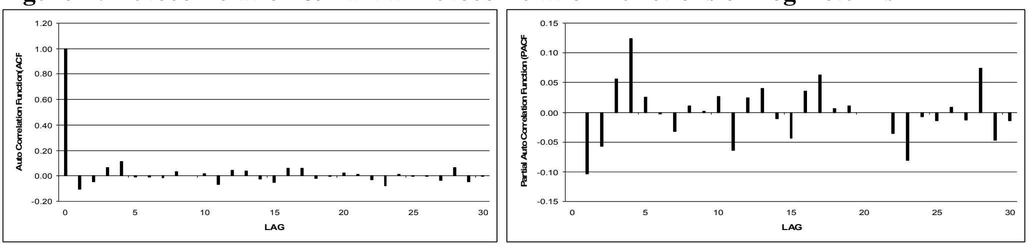
Table 2: Unit Root Test for Log Returns

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-23.75465	0.0000
Test critical values:	1% level	-3.444342	
	5% level	-2.867603	
	10% level	-2.570063	

Source: Own Study

Finally, the autocorrelation and partial autocorrelation results are obtained. (Figure-4) It seems that there is relationship in an order of one.

Figure-4: Autocorrelation & Partial Autocorrelation Functions of Log Returns



Source: Own Study

The next step is to apply "Hannan-Rissanen" algorithm to decide the mean equation. The idea behind the procedure is first to fit an AR model to the data in order to obtain the estimates of the noise or innovation. When this estimated noise is used in place of the true noise, it enables us to estimate ARMA parameters using the less expensive method of least squares regression. The orders are determined within the procedure itself using an information criterion. It gives AR[1] as the best fit. The results for the first eight models are given at Table-3.

Table 3: Hannan-Rissanen Results and AIC&BIC Values

Ranking	Model	Model Parameters	AIC	BIC
1	AR [1]	[{-0.102947},0.000368701]	-7.8922	-7.8721
2	MA [1]	[{-0.118533},0.000368849]	-7.8918	-7.8717
3	AR [2]	[{-0.109846,-0.0564174},0.000366953]	-7.8836	-7.8435
4	ARMA [1,1]	[{0.289847},{-0.408522},0.000367832]	-7.8812	-7.8411
5	MA [2]	[{-0.117819,-0.0336272},0.000369023]	-7.8780	-7.8378
6	MA [4]	[{-0.113389,-0.0375167,0.0863235,0.113773},0.000359727]	-7.8768	-7.7966
7	AR [4]	[{-0.115142,-0.0447406,0.0698673,0.125534},0.000361382]	-7.8722	-7.7920
8	AR [3]	[{-0.108007,-0.0507689,0.0568196},0.000366382]	-7.8718	-7.8116

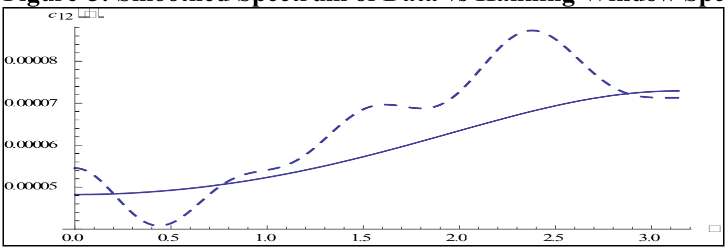
Source: Own Study

Moreover, Stationary-Q test result for AR[1] model states that it is stationary. Finally, the Portmanteau test is used to see whether there is any autocorrelation is left. AR[1] Portmanteau statistics for the first 35 autocorrelations is 30.53, while 95% confidence level Chi-Square Distribution result is 48.60. In other word, there is not enough evidence to state that there is autocorrelation left in the residuals.

3.2. Analysis in Frequency Domain

The aim in this section is to analyze the data in frequency domain. It is also called the time series analysis in Fourier space. It enables us to work with the same data in different representation and all should give the same result. When the spectrum analysis has been carried out, it could be observed that while the paths are same for both the smoothed spectrum of the data and the spectrum of AR[1] estimates, there are some noise due the GARCH effect. (Figure-5)

Figure-5: Smoothed Spectrum of Data vs Hanning Window Spectrum of AR[1]



Source: Own Study

3.3. GARCH Modelling

The data has to be checked for GARCH modeling. In other word, it is controlled to be able to see that is there any GARCH effect of not. ARCH LM test result shows that there is GARCH effect in the series which is to be modeled. (LM Statistic: 6.884 & 95% confidence level Chi-Square Distribution result: 3.841)

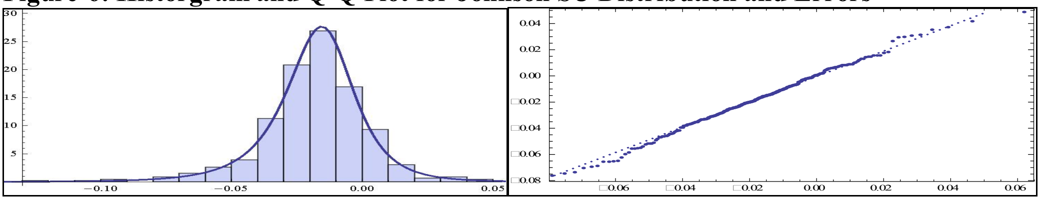
GARCH effect is analyzed and the results for different models indicated that GARCH[1,1] is the best fit via lowest AIC value. The parameters for constant, ARCH[1] and GARCH[1] coefficients are 0.000007970611852, 0.07482224309 and 0.9049436507 respectively. The estimated GARCH obeys the non-negativity constraints of GARCH process, since all coefficients are positive. The model also satisfies the covariance stationary condition that sum of coefficients is less than 1.

$$X_t = 0.000368701 - 0.102947 X_{t-1} + \varepsilon_t$$
$$\sigma_t^2 = 0.000007970611852 + 0.07482224309 \varepsilon_{t-1} + 0.9049436507 \sigma_{t-1}^2$$

3.4. Distribution Analysis on Errors

The aim of this section is to find the best distribution which fit to GARCH model error series. Normal, Johnson SU, Weibull, Gumbel and Cauchy distributions are being tested for distribution fitting. Histograms, probability plots, Q-Q plots and cumulative distribution functions are checked together with different statistical tests including Anderson-Darling, Cramér-von Mises, Kolmogorov-Smirnov and Kulper. As a result, Johnson SU is found to be the best fit. (Figure-6)

Figure-6: Historgram and Q-Q Plot for Johnson SU Distribution and Errors



Source: Own Study

3.5. Continuous Modelling

When the discrete time GARCH[1,1] model had been estimated, the continuous time COGARCH[1,1] model can be found from the discrete model parameters.

The parameters of continuous COGARCH [1,1] model in terms of discrete time GARCH[1,1] model can be written as:

$$\beta = \beta \qquad \eta = -\ln(\delta) \qquad \varphi = \lambda / \delta$$

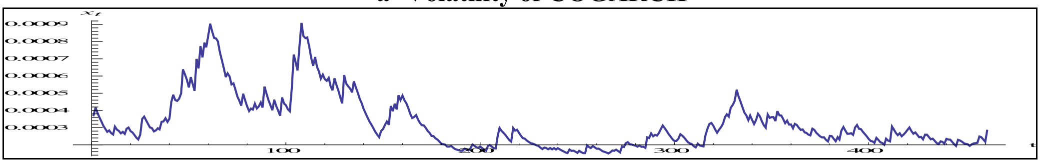
Therefore, the parameters of COGARCH [1,1] model can be calculated as:

$$\beta = 0.000007970611852$$
$$\eta = -\ln(0.9049436507) = 0.0998826016406098$$
$$\varphi = \lambda / \delta = 0.07482224309/0.9049436507 = 0.0826816598272421$$

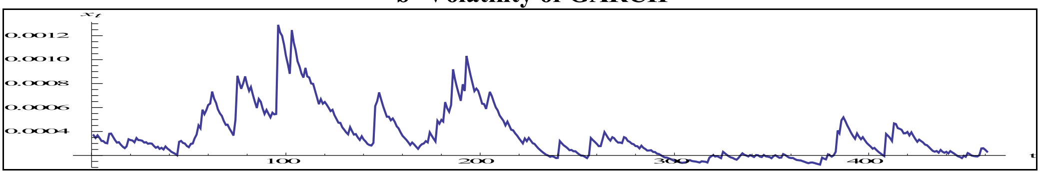
After the parameter estimation, the simulation has been carried out by using the numerical solutions for G_t and σ_t^2 . The Lévy process driven by Johnson-SU process, which is found to be the best fit, is utilized. Figure-7 shows the discrete GARCH model and continuous GARCH (COGARCH) model. It could be easily realized that both models are showing similar pattern in time.

Figure-7: Volatility Graphs of COGARCH & GARCH & Log Returns

a- Volatility of COGARCH



b- Volatility of GARCH



Source: Own Study

4. CONCLUSION

In this study, log returns of daily LME Steel Billets Spot Prices between 04.01. 2010 and 31.10.2011 have been modeled with AR[1] ~ GARCH[1,1] discrete model which is the best candidate. The discrete model parameters are used to construct the continuous COGARCH[1,1] analogue. Then, the simulated volatility results of both discrete and continuous models are compared with each other. It was shown that both models follow the similar patterns especially in the jumps.

BIBLIOGRAPHY

Barndorff-Nielsen, Ole E., Neil Shepherd (2001), "Non-Gaussian Ornstein-Uhlenbeck Based Models and some of their Use in Financial Economics (with discussion)", Journal of Royal Statistics Society Series B, Vol. 63, pp.167-241.

Christian Kleiber and Samuel Kotz (2003), "Statistical Size Distributions in Economics and Actuarial Sciences", Wiley Series in Probability and Statistics.

Duan, Jin Chuan (1997), "Augmented GARCH (p,q) Process and Its Diffusion Limit", Journal of Econometrics, Vol. 79, pp.97-127.

Geman Hélyette (2005), "Commodities and Commodity Derivatives: Modelling and Pricing for Agriculturals, Metals and Energy", Wiley Finance.

Klüppelberg Claudia, Alexander Lindner and Ross Maller (2004), "A Continuous Time GARCH Process Driven by a Lévy Process: Stationary and Second Order Behaviour", Journal of Applied Probability, Vol. 41, pp.601-622.

Klüppelberg Claudia, Alexander Lindner and Ross Maller (2006), "Continuous Time Volatility Modelling: COGARCH versus Ornstein-Uhlenbeck Models",(in: Yuri Kabanov, Robert Lipster and Jordan Stoyanov-Eds, From Stochastic Calculus to Mathematical Finance), Springer:Berlin, pp.393-419.

Nelson, Daniels B. (1990), "ARCH Models as Diffusion Approximation", Journal of Econometrics, Vol. 45, pp.7-38.

Ross A. M., Gernot M. and Alex S. (2008), "GARCH Modelling in Continuous Time for Irregular Spaced Time Series Data", Bernoulli, Vol. 14, pp.519-542.

Ross A. Maller, Gernot Müller and Alex Szimayer (2009), "Ornstein-Uhlenbeck Processes and Extensions", Handbook of Financial Time Series.

Tim Bollerslev (1986), "Generalized Autoregressive Conditionally Heteroscedasticity", Journal of Econometrics, Vol. 31, pp.307-327.