Jestr

Journal of Engineering Science and Technology Review 8 (1) (2015) 12 -15
Special Issue on Econophysics

JOURNAL OF
Engineering Science and Technology Review

www.jestr.org

Conference Article

Price Formation Modelling by Continuous-Time Random Walk: An Empirical Study

Frédéric Délèze and Sergey Osmekhin

Department of Finance and Statistics, Hanken School of Economics, P.O. Box 479, FI-00101, Helsinki, Finland

Abstract

Markovian and non-Markovian models are presented to model the futures market price formation. We show that the waiting-time and the survival probabilities have a significant impact on the price dynamics. This study tests analytical solutions and present numerical results for the probability density function of the continuous-time random walk using tick-by-tick quotes prices for the DAX 30 index futures.

Keywords: econophysics, continuous-time random walk, DAX futures, non-Markovian model, price dynamics

1. Introduction

Continuous-time random walk (CTRW) is an extension of the classical random walk model initiated by Montroll and Weiss in 1965 [1]. CTRW was introduced as a theoretical approach to describe the diffusion process in solid-state physics, where the waiting-time between two sequential space jumps of a moving particle is modelled stochastically. It models the dynamics of the probability density function of observing a particle in the space point x at time t. Similar processes take place in financial markets, where the time between transactions is stochastic and where trades induce price jumps [2]. Nowadays the CTRW framework is widely used in finance to predict and analyse the price behaviour of stock and derivatives [3,4,5] by calculating the probability density function (pdf) p of finding a certain price at a given time t.

Two main forms of CTRW have been recently described to model the price of financial assets [3,4,6]: a Markovian (memoryless) and a non-Markovian model. The present study tests the two models. It compares analytical solutions with numerical results for the price probability density function using tick-by-tick mid-quote prices for the German DAX 30 futures. We find a significant difference between the solutions of the Markovian and non-Markovian equations. In addition we show that market liquidity has a meaningful impact on future market price formation.

2. Continuous-time Random Walk

We follow the same approach as Scalas, et al [4,6,7]. Let $x(t)$ be the log-price of asset S at time t. The time between two transactions, called waiting-time is $\tau_i = t_{i+1} - t_i$. The log-return is $\xi_i = x(t_{i+1}) - x(t_i)$. The joint probability of returns and waiting-time is defined as $\varphi(\xi, \tau)$. The two marginal distributions, $\psi(\tau) = \int_{-\infty}^{\infty} \varphi(\xi,\tau) d\xi$ and $\lambda(\xi) = \int_0^{\infty} \varphi(\xi,\tau) d\tau$, represent the waiting-time and asset return probability density functions, respectively. We call $p(x,t)$ the probability of finding a lot price x at time t. The Laplace transform of $f(t)$ is denoted by

$$\tilde{f}(s) = \int_0^{\infty} e^{-st} f(t) dt.$$

The non-Markovian form of the CTRW solves the following master equation:

$$\int_0^t \phi(t-t') \frac{\partial}{\partial t'} p(x,t') dt' = -p(x,t) + \int_{-\infty}^{\infty} \lambda(x - x') p(x',t) dx' \quad (1)$$

where the kernel $\phi(t)$ is defined through its Laplace transform $\tilde{\phi}(s) = \frac{s\,\tilde{\psi}(s)}{1-\tilde{\psi}(s)}$. As $\tilde{\phi}(s) = 1$ the master equation for the CTRW becomes Markovian:

$$\frac{\partial}{\partial t} p(x,t) = -p(x,t) + \int_{-\infty}^{\infty} \lambda(x-x') p(x',t) dx' \quad (2)$$

The CTRW Markovian equation describes the standard dynamic to model the price of financial instruments. The model can be seen as a generalisation of the geometric Brownian motion as it uses the asset return distribution as the unique driver to model the price fluctuation of an asset over time.

The Figure 1 illustrates the modelled solution of the Markovian master equation. At time *t=0*, the probability density function $p(x,t)$ is a delta Dirac function because the

* E-mail address: frederik.deleze@hanken.fi
ISSN: 1791-2377

current price is known. The uncertainty then increases with time and broadens $p(x,t)$. The skewness of the distribution $\lambda(x)$, modelled in the figure with an exponential distribution, orientates $p(x,t)$. The cross-sectional view at a given point in time *t>0* is the distribution of asset returns. Thus, $p(x, t = 200\text{ sec})$ reflects the expected distribution of the log-price *x* 200 seconds ahead.

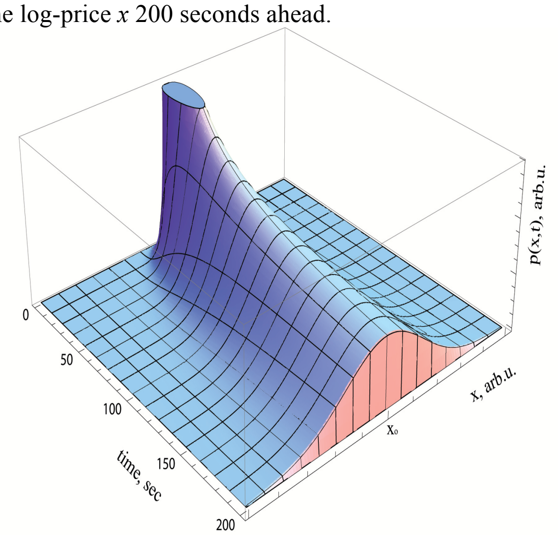

Fig. 1. Modelled probability density function of finding a log-price x at time t, p(x,t) for the Markovian master equation (1)

Asset returns are known to be leptokurtic [8] and the assumption of independent and identically distributed equity returns underestimates the real probability of extreme events [9]. The general framework of CTRW supports fat-tailed distributions.

The non-Markovian CTRW is an extension of the Markovian CTRW, where both the time between transactions, called waiting-time, and the asset returns are modelled stochastically. The waiting-time distribution reflects the market liquidity. A transaction in a very illiquid market, i.e. when the waiting-time is abnormally long, translates into abrupt price changes while a transaction in a very liquid period has very little impact on price [10]. As the waiting-time distribution conveys relevant information about price formation we can expect the non-Markovian approach to outperform the memoryless model.

3. Intraday Returns

The high-frequency probability density function of asset return $\lambda(\xi)$ needs to be evaluated to describe the dynamics of the Markovian master equation and the non-Markovian space and time convolutions stochastic differential equation. To reflect the fat-tailness of asset returns [11,12], the probability density function of returns $\lambda(\xi)$ of each model is approximated with Kou's double exponential jump-diffusion model [13]:

$$dS(t)/S(t-) = \mu dt + \sigma dW(t) + d\left(\sum_{i=1}^{N(t)} (V_i - 1)\right) \quad (3)$$

where *W(t)* is a Wiener process, σ the volatility and *N(t)* is a Poisson process with rate ν. With high-frequency data, the drift is irrelevant and we set $\mu = 0$. $\{V_i\}$ is a sequence of positive random variables such that $Y \equiv \log(V)$ is defined as an asymmetric double exponential distribution:

$$f_Y(y) = q\eta_1 e^{-\eta_1 y} 1_{\{y \geq 0\}} + (1-q)\eta_2 e^{\eta_2 y} 1_{\{y<0\}} \quad (4)$$

The parameters q represents the probability of positive jumps. η_1 and η_2 control the decrease of the distribution tails of positive, respectively negative jumps. For small Δt, the solution of the stochastic differential equation can be approximated as $(S(t+\Delta t) - S(t))/S(t) = \mu\Delta t + \sigma\sqrt{\Delta t}\,\epsilon + B \cdot Y$ with $\epsilon \sim N(0,1)$ and $B \sim \text{Binomial}(n, q)$. Hence, with probability q the return ξ jumps up to $\xi^+$and with probability *(1-q)* ξ jumps down to $\xi^-$where ξ^+ and ξ^- follow an exponential random variable with mean $1/\eta_1$ and $1/\eta_2$.

The three coefficients q, η_1 and η_2 can either be estimated by maximum likelihood method described in [14] or, alternatively, the jumps can be detected by performing the Lee-Mykland test [15]. The detection technique consists of disentangling price jumps due to the Wiener process from the pure jump components by computing the realised bi-power variation on an optimal window size. q is estimated as the ratio of all positive jumps to all detected jumps, $\hat{\eta}_1 = 1/E[|\text{jump}||\text{jump size} > 0]$ and $\hat{\eta}_2 = 1/E[|\text{jump}||\text{jump size} < 0]$.

Computing the realised volatility $\hat{\sigma} = \sum_{i=1}^{n} \left(\xi_{t_{i+1}} - \xi_{t_i}\right)^2$with high-frequency data introduces a bias proportional to the number of observations n as shown in [16]. To circumvent the problem, we follow the approach suggested in [17] and estimate the volatility with a Two Scales Realised Volatility (TSRV). The observed log-price at time *t*, ξ_t, is noisy due to imperfections of the trading process. Let us define $\xi_t = X_t + \epsilon_t$ with X_t the unobserved efficient log-price at time *t* and ϵ_t the noise level. A generalisation of the TSRV for the continuous quadratic variation $< X, X >_T = \int_0^T \sigma_t^2 dt$is defined as

$$\widehat{< X, X >}_T = [\xi, \xi]_T^{(K)} - \frac{\bar{n}_K}{\bar{n}_J} [\xi, \xi]_T^{(J)}, \; 1 \leq J < K \leq n \quad (5)$$

for the unobserved efficient log-price X from which the average lag j realised volatility is given by $[\xi, \xi]_T^{(J)} = \frac{1}{J}\sum_{i=0}^{n-J} \left(\xi_{t_{i+J}} - \xi_{t_i}\right)^2$ where $\bar{n}_K = (n - K + 1)/K$, $\bar{n}_J = (n - J + 1)/J$.

4. Relationship between Markovian and Non-Markovian Approaches

The intertrade duration at tick-by-tick level follows a mixture of compound Poisson processes $\psi(\tau) = \sum_{i=1}^{T} a_i \mu_i e^{-\mu_i T}$ as described in [18] where $\{a_i\}_{i=1}^{T}$is a set of weights reflecting the intraday activity observed on the market. Hence, for each trading period, the waiting-time probability density function $\psi(\tau)$ can be modelled as an exponential distribution of parameters $\theta e^{-\theta\tau}, \tau \geq 0$.

Modelling $\psi(\tau)$ as an exponential distribution simplifies the discretisation of the non-Markovian stochastic differential equation. Indeed, the Laplace transform of the kernel $\phi(t)$ is constant $\tilde{\phi}(t) = \frac{t\,\tilde{\psi}(t)}{1-\tilde{\psi}(t)} = \frac{t\frac{\theta}{t+\theta}}{1-\frac{\theta}{t+\theta}} = \theta$ and therefore, $\phi(t) = \theta\delta(t)$ where $\delta(t)$ is the Dirac function at t.

The discretisation of the non-Markovian master equation (1) shows that $p(x, t+1)$, the probability density function of finding a log-price x at a future time *t+1*, is the sum of the survival probability up to time t times $p(x,t)$ with the solution of the Markovian stochastic density function (2):

$$\underbrace{p(x,t+1)}_{\text{non-Markovian pdf}} = \underbrace{\left(1-\frac{1}{\theta}\right)}_{survival\ prob.\ \Psi} p(x,t) + \frac{1}{\theta}\underbrace{p(x,t+1)}_{\text{Markovian pdf}} \quad (6)$$

5. Empirical data

The analytical solutions of the two forms of CTRW are illustrated with the DAX30 futures. The index contains the 30 biggest German stocks by capitalisation. DAX30 futures are traded from 07:50 until 22:00 on Eurex. We use level-1 tick-by-tick quote data provided by tickmarketdata.com [19] spanning the period from 20 December 2010 until 23 December 2011. The dataset contains the first expiry of the March, June, September and December contracts. Figure 2 shows the four intraday trading periods observed for the DAX30 futures. Before the opening of the regional stock market at 09:00 CET, the trading activity is very low. In the second phase, before the NYSE and NASDAQ open at 15:30 CET, the activity increases but the time between transactions remains higher than between 15:30 and 17:30 CET when both the US and European stock markets are open. Finally, the trading activity declines again after the closure of the European stock markets.

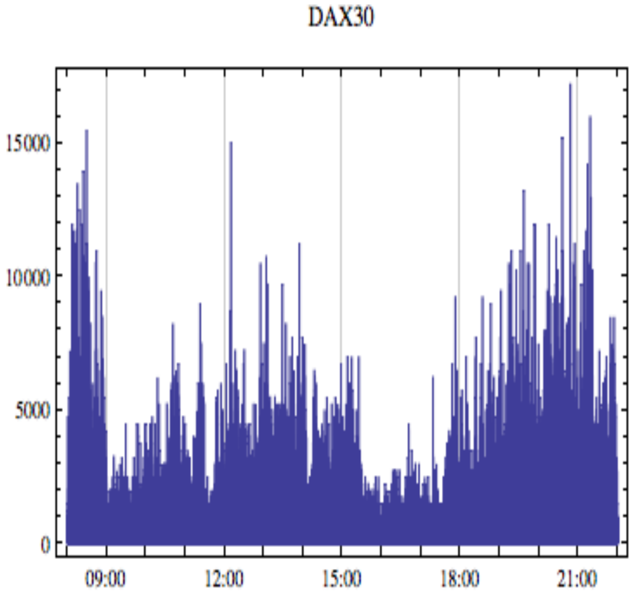

Fig. 2. Time between transactions τ on 10 January 10 for DAX30 March 11 expiry.

Sudden and large bursts in market liquidity, mainly driven by macroeconomic news announcements [20], cause jumps in the survival probability. Table 1 presents the estimation of the parameters of the Kou model (3) and (4) and TSRV volatility (5) used by the two approaches. They are computed for the first expiry of each future spanning the period from 20 December 2010 until 23 December 2011 for both liquid and illiquid periods, i.e. 15:30 – 17:30 and 08:00 – 09:00. The probability of positive jump q and the magnitude of jumps $\hat{\nu}$ are significantly higher when the market is illiquid.

Table 1. Parameters estimation for the first expiry of the DAX30 futures

	Time 08:00 - 09:00				Time 15:30 - 17:30			
DAX30	Mar-11	Jun-11	Sep-11	Dec-11	Mar-11	Jun-11	Sep-11	Dec-11
$\hat{\nu}$	0.99	1.22	2.53	3.13	0.98	1.19	2.18	2.63
$\hat{\eta}_1$	1.27	0.81	0.55	0.32	1.13	0.72	0.50	0.36
$\hat{\eta}_2$	0.86	0.84	0.32	0.32	0.92	0.98	0.44	0.40
$\hat{\sigma}$	0.09	0.17	0.29	0.30	0.09	0.10	0.24	0.21
q	46%	54%	45%	58%	50%	46%	39%	54%

Figure 2 compares the Markovian and the non-Markovian approaches for the DAX 30 June 11 expiry futures. The three upper pictures illustrate the cross-sectional view at the future time step *t=4*. The expectation and variance of the difference between the two approaches show that the uncertainty is larger when the liquidity is low. In liquid markets, buy and sell orders are more frequent and the execution time is shorter. Illiquidity increases the probability of partially filled and missed orders. The non-Markovian approach includes the survival probability density function as a liquidity measure (6). The expected difference between the Markovian and non-Markovian master equations when the price is kept unchanged, $E[p_2(x(t)=x_0,t)] - E[p_1(x(t)=x_0,t)]$ depends on market liquidity. The difference reaches the levels of 0.02% (or 1.5 basis points) and 0.01% or (0.75 basis point) for the DAX30 futures at the liquid and illiquid periods, respectively. Such price levels are substantial and significantly impact the profitability of high-frequency trading strategies. As depicted in the lower graph, the distribution of the expected difference exhibits fatter tails in illiquid than in liquid markets, indicating an enhanced risk of high price movements when the market is not active. Table 2 summarises the results of Welch's t-test [21]. The null hypothesis that the mean of the two approaches is equal is rejected at 1% confidence level.

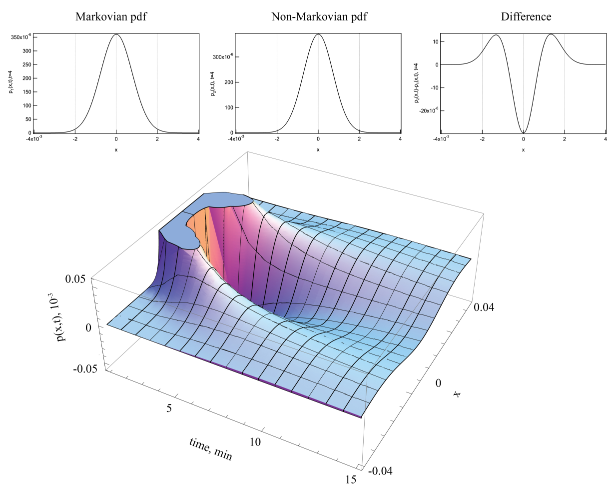

Fig. 3. On top: The graph on the left illustrates the Markovian PDE at t=4 for the DAX30 June 11 futures. The central figure depicts the non-Markovian PDE. The cross-sectional difference between the non-Markovian and the Markovian PDEs at t=4 is displayed on the right. The survival probability ψ=43%. At the bottom: 3D plot displays the differences between the two PDEs over time.

Table 2: Difference between the non-Markovian $p_2(x,t)$ and the Markovian $p_1(x,t)$ approaches for DAX JUN11 at time t=4 when the market is illiquid (08:00-09:00) and during the most liquid period (15:30 – 17:30). (*) The results of the Welch's t-test indicate that the two pdfs differ at 1% confidence level.

	Illiquid time between 08:00 and 09:00			Liquid time between 15:30 and 17:30		
$x \times 10^{-3}$	$E[p_2 - p_1]$	$Var[p_2 - p_1]$	$t - test$	$E[p_2 - p_1]$	$Var[p_2 - p_1]$	$t - test$
10	3.1×10^{-6}	6.3×10^{-12}	7.5^*	2.5×10^{-8}	2.3×10^{-14}	–
7	5.2×10^{-5}	1.7×10^{-13}	3.2^*	1.7×10^{-6}	5.6×10^{-13}	16.59^*
5	3.0×10^{-5}	5.3×10^{-11}	1.7	2.1×10^{-5}	3.3×10^{-11}	10.3^*
3	-2.9×10^{-5}	1.2×10^{-10}	−0.6	8.1×10^{-5}	5.0×10^{-11}	5.9^*
2	-6.2×10^{-5}	1.1×10^{-10}	-4.9^*	5.1×10^{-5}	4.3×10^{-10}	2.6^*
1	-8.9×10^{-5}	5.7×10^{-11}	-11.8^*	-6.3×10^{-5}	1.1×10^{-9}	-1.8^*
0	-1.0×10^{-4}	1.1×10^{-11}	-68.5^*	-1.8×10^{-4}	2.8×10^{-10}	-14.4^*
−1	-9.9×10^{-5}	8.3×10^{-11}	-18.2^*	-1.4×10^{-4}	5.6×10^{-10}	-11.8^*
−2	-8.5×10^{-5}	6.7×10^{-11}	-7.4^*	-3.2×10^{-5}	9.2×10^{-10}	−1.6
−3	-5.6×10^{-5}	1.1×10^{-10}	-3.5^*	6.1×10^{-5}	2.3×10^{-10}	2.6^*
−5	4.4×10^{-6}	7.5×10^{-11}	−0.3	4.1×10^{-5}	6.9×10^{-11}	5.1^*
−7	3.7×10^{-5}	8.6×10^{-12}	1.8^*	4.1×10^{-6}	2.7×10^{-12}	10.3^*
−10	2.9×10^{-5}	1.0×10^{-11}	3.1^*	6.3×10^{-8}	5.0×10^{-14}	16.4^*

6. Conclusion

The present paper tests and compares the joint probability of finding a log-price *x* at a future time *t* for both the Markovian and non-Markovian forms of the CTRW. The non-Markovian probability density function is derived in terms of the solution of the Markovian equation where the waiting-time density function is exponentially distributed. The two models are constructed and their parameters are estimated with tick-by-tick data for the DAX 30 index futures. We find a significant difference between the two approaches. Market liquidity, reflected by the waiting-time and survival probability density functions is not constant throughout the trading day and plays a central role in the price formation at a market microstructure level. Further research is needed to test if the probability density function of the volume traded affects price formation.

References

[1] E.W. Montroll and G.H. Weiss, Random walk on lattices, II, J. Math. Phys. 6 (1965) 167-181

[2] M. Raberto, E. Scalas, F. Mainardi, Waiting-times and returns in high-frequency financial data: an empirical study, Physica A 314 (2002) 749- 755

[3] E. Scalas, The application of the continuous-time random walks in finance and economics, Physica A 362 (2006) 225-239

[4] F. Mainardi, M. Raberto, R. Gorenflo, E. Scalas, Fractional calculus and continuous-time finance II: the waiting-time distribution, Physica A 287 (2000) 468-481

[5] A. Cartea, T. Meyer-Brandis, How duration between trades of underlying securities affects option prices, Review of Finance, 14 (4) (2010) 749-785

[6] R. Gorenflo, F. Mainardi, E. Scalas, M. Raberto, Fractional calculus and continuous-time finance III: the diffusion limit, Mathematical finance, 17010180, Trends Math, Birkhäuser, Basel (2001) 171-180

[7] E. Scalas, R. Gorenflo and F. Mainardi, Fractional calculus and continuous-time finance, Physica A 284 (2000) 376-384.

[8] R.N. Mantegna, H. E. Stanley, Scaling behaviour in the dynamics of an economic index, Nature 376.6535 (1995): 46-49.

[9] D.A. Hsieh, Chaos and nonlinear dynamics: application to financial markets, Journal of Finance 46.5 (1991): 1839-1877.

[10] M. Chng, Measuring the Summary Informativeness of Orders and Trades, Review of Futures Markets, (2004): 245-281.

[11] Y. Aït-Sahalia, Disentangling diffusion from jumps, J Financ Econ, 74 (2004), 487-528

[12] D.S. Bates, Empirical option pricing: a retrospection, J Econom 116(1- 2), (2003a), 387-404

[13] S. Kou, A jump diffusion model for option pricing, Manage Sci, 48(8) (2002), 1086-1101

[14] C.A. Ramezani, Y. Zeng, Maximum likelihood of the double exponential jump-diffusion process, Annals of Finance, 3 (2007) 487-507

[15] S.S. Lee, P.A. Mykland, Jumps in Financial Markets: A New Non-parametric Test and Jump Dynamics, The Review of Financial Studies, (2007) 21 (6): 2535-2563

[16] L. Zhang, et al., A tale of two time scales: determining integrated volatility with noisy high-frequency data, Journal of the American Statistical Association 100, (2005), 1394-1411

[17] Y. Aït-Sahalia, et al., Ultra high frequency volatility estimation with dependent microstructure noise, Journal of Econometrics, (2010)

[18] E. Scalas, Mixtures of compound Poisson processes as models of tick-by-tick financial data, Chaos, Solitons and Fractals 34 (2007) 33-40

[19] Tick-by-tick prices are provided by tickmarketdata.com

[20] Andersen, Torben G., et al., Real-time price discovery in global stock, bond and foreign exchange markets, Journal of International Economics 73.2 (2007): 251-277

[21] B.L. Welch, The generalization of Student's problem when several different population variances are involved, Biometrika 34 (1947) (12): 2835