Continuous Time Analysis of Fleeting Discrete Price Moves

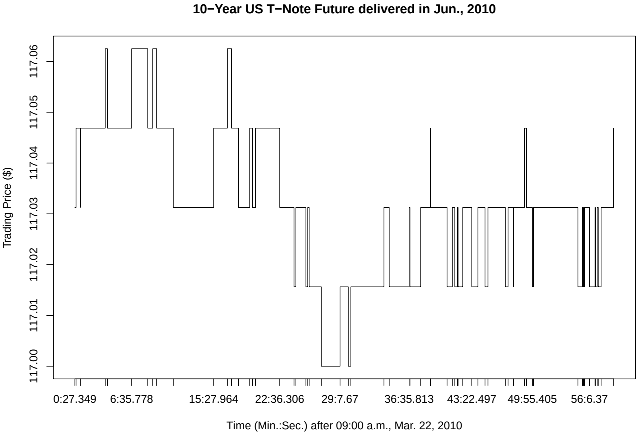
Neil Shephard and Justin J. Yang

Economics & Statistics Departments,
Harvard University

December 2014

The data in our hand: price process

Low latency data, recorded to 1/1000 of a second.



Data is always discrete. Academic TAQ will mislead you, aggregates to 1 second.

Contribution: modelling the data in our hand

- Continuous time model for prices (or best bid). Like the data:
  - prices are discrete (tick structure);
  - prices change in continuous time;
  - a high proportion of price changes are reversed in a fraction of a second.
- Model is analytically tractable: role of the calendar time is explicit.
- Formulated in terms of a "price impact curve".
- Price is càdlàg, piecewise constant semimartingale with finite activity, finite variation and no Brownian motion component.
- For futures data sets: describes the observed dynamics of price changes over three diferent orders of time
  - 0.1 seconds, 1 seconds, 10 seconds and 1 minute.

Why study?

- Trading at high frequency (prediction and control)
  - Minimising trading costs for fundamental trader (your pension)
  - Statistical arbitrage
  - Risk management
- Information extraction from high frequency data
  - Time-varying vol and correlation
  - Skews and statistical leverage
- For policy
  - Advantages/disadvantages of multiple exchanges (fragmentation/competition), dark pools, etc
- Regulation (e.g. auction each second, not continuously?)
- Forensic finance
  - Does some trading systems create a false market.

Existing analysis

- Much econometrics focused on very short-term predictive models
  - next trade or quota update
  - change in price or time between event
  - reviews in Engle (2000), Engle & Russell (2010) and Hautsch (2012)
- Relatively little about discreteness
  - Rydberg & Shephard (2003), Russell & Engle (2006), Liesenfeld, Nolte, & Pohmeier (2006), Large (2011), Oomen (2005), Oomen (2006) and Griffin & Oomen (2008). Early work includes Harris (1990), Gottlieb & Kalay (1985), Ball, Torous, & Tschoegl (1985) and Ball (1988)
- Rounding, rounding plus measurement error
  - Hasbrouck (1999), Rosenbaum (2009), Delattre & Jacod (1997), Jacod (1996) and Li & Mykland (2014).

Similar themes: discreteness in continuous time

Small literature on moves in our direction.

- Barndorff-Nielsen (BN), Pollard & Shephard (2012). Lévy process.
  - Difference of two subordinators (non-negative Lévy processes). e.g. count up moves, modelled as Poisson process. Likewise downs. Difference is price and is Skellam Lévy process. Extends to non-single tick markets.
- Bacry, Delattre, Hoffman & Muzy (2013a,b)
  - For single tick markets: extend Lévy process to difference of two Hawkes processes (up and down moving counting processes).
- Fodra and Pham (2013a,b).

Work draws on math in

- Deeper parts of the math used here draws on
  - Barndorff-Nielsen (BN), Pollard & Shephard (2012). Lévy process.
  - Barndorff-Nielsen, Lunde, Shephard & Veraart (2014). Stationary model.
    - Related to Wolpert & Taqqu (2008) and Wolpert & Brown (2012)
    - Related to $M/G/\infty$ queues, e.g. Lindley (1956), Reynolds (1968) and Bartlett (1978, Ch. 6.31)
    - Related to mixed moving average models of Surgailis, Rosinski, Mandrekar, and Cambanis (1993).
- Related to discrete time integer valued processes.

Core model: Poisson random measure

The basic framework:
(i) events arriving in continuous time,
(ii) some events have fleeting impact, some permanent
(iii) events of variable size and direction.

- Minimal mathematical core: 3 dim *Poisson random measure* N with intensity measure

$$\mathbb{E}\{N(\mathrm{d}y, \mathrm{d}x, \mathrm{d}s)\} = \nu(\mathrm{d}y)\mathrm{d}x\mathrm{d}s.$$

  $\nu(\mathrm{d}y)$ is a Lévy measure.

  - s is time (with arrivals randomly scattered on $\mathbb{R}$) and
  - x is a random height (uniformly scattered over $[0,1]$): random source for the degree of fleetingness of the event and
  - y marks the variable size and direction of the integer events.

Zero chance two points with common height or time.

Lévy basis

- The Lévy basis records the value of the y variable at each point in time s (which lives on $\mathbb{R}$) and height x (which lives on $[0,1]$). It is given by

$$L(\mathrm{d}x, \mathrm{d}s) = \int_{-\infty}^{\infty} yN(\mathrm{d}y, \mathrm{d}x, \mathrm{d}s), \qquad (x,s) \in [0,1] \times \mathbb{R}.$$

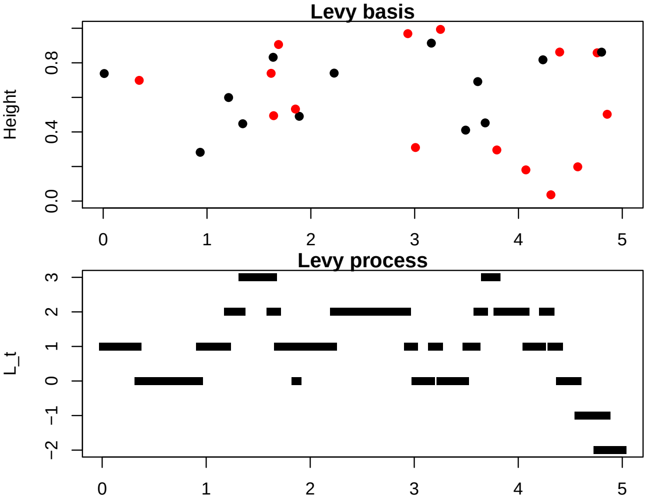
- The Lévy process

$$\begin{aligned} L_t &= \int_0^t \int_0^1 L(\mathrm{d}x, \mathrm{d}s) & (1) \\ &= L(D_t), \qquad D_t \triangleq [0,1] \times (0,t]. & (2) \end{aligned}$$

Here D_t is a rectangle which grows with t. Thus the Lévy process counts up all the points in the Lévy basis with heights under 1 and from time 0 to time t.

Example: Skellam basis and Skellam process

$L(\mathrm{d}x, \mathrm{d}s)$, s time, x height. Black: $+1$, Red: -1.



Drivers of price process

- Lévy process: apply increasing rectangle D_t to Lévy basis L

$$L_t = L(D_t), \qquad D_t \triangleq [0,1] \times (0,t].$$

- Need fleeting component too. Multiple shapes
  - Drag through time a fixed shape

$$A \subseteq [b,1] \times (-\infty, 0]\,, \quad \text{where} \quad b \in [0,1]$$
$$A_t \triangleq A + (0,t)\,.$$

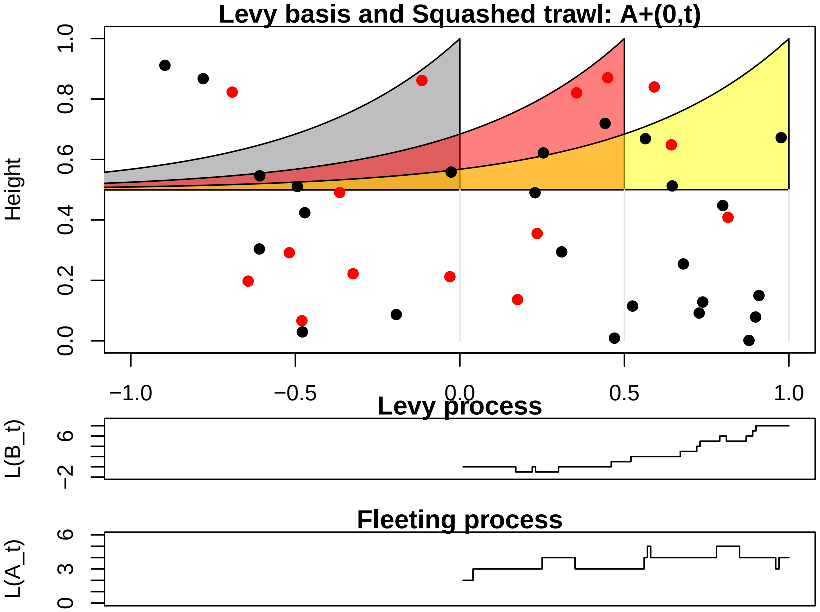
  - Build an increasing rectangle $B_t \triangleq [0,b) \times (0,t]$.
- Union of two shapes

$$C_t \triangleq A_t \cup B_t, \quad A_t \cap B_t = \varnothing.$$

- Price

$$P_t = V_0 + L(C_t) = V_0 + L(A_t) + L(B_t).$$

- Lévy process $L(B_t)$ independent of fleeting $L(A_t)$.
  $leb(C_t) = leb(A) + tb$.



Levy basis and Squashed trawl: A+(0,t)
Height
0.0
0.2
0.4
0.6
0.8
1.0
−1.0
−0.5
0.0
0.5
1.0
Levy process
L(B_t)
−2
6
Fleeting process
L(A_t)
0
3
6

Trawl function

- $A_t \triangleq A + (0, t)$,
- Shape A?
- Curve denoted d. Called a "trawl function".

$$A \triangleq \{(x, s) : s \leq 0, b \leq x < d(s)\}. \tag{3}$$

- Here makes sense for d to be monotonic.
- Write

$$G(s) = 1 - d(-s), \quad s \geq 0.$$

- Lifetime of j-th arrival is

$$G^{-1}(U_j), \quad U_j \overset{iid}{\sim} U(0, 1).$$

- $G^{-1}(U_j) = \infty$ permanent
- $G^{-1}(U_j) < \infty$ temporary
- Hence trawl function parameterises a "price impact curve".

Stochastic analysis

Like the data:

- Price is càdlàg,
- Piecewise constant semimartingale with
  - finite activity (so the Blumenthal-Getoor index is always zero)
  - finite variation and
- No Brownian motion component.

Jump probabilities of prices

For this model

$$\mathbb{P}\left(\Delta P_t = y | \Delta P_t \neq 0\right) = \frac{\nu\left(y\right) + \nu\left(-y\right)\left(1-b\right)}{\left(2-b\right)\left\|\nu\right\|}. \tag{4}$$

Price moves

$$\begin{aligned} P_t &= V_0 + L(C_t) = V_0 + L(A_t) + L(B_t), \quad t \geq 0, \\ P_t - P_0 &= L(C_t) - L(C_0), \quad t > 0. \end{aligned}$$

Theorem

Let $A \backslash B$ be set subtraction (all of set A except those also in B). Then

$$P_t - P_0 = L\left(C_t \backslash C_0\right) - L(C_0 \backslash C_t),$$

where $L\left(C_t \backslash C_s\right)$ is independent of $L(C_s \backslash C_t)$. Characteristic function of returns is

$$\begin{aligned} M\left(\theta \ddagger P_t - P_0\right) &\triangleq \log \mathbb{E}\left\{e^{i\theta(P_t - P_0)}\right\}, \quad i \triangleq \sqrt{-1}, \\ &= btM\left(\theta \ddagger L_1\right) + leb(A_t \backslash A)\left\{M\left(\theta \ddagger L_1\right) + M\left(-\theta \ddagger L_1\right)\right\}, \end{aligned}$$

where L_t is the corresponding Lévy process.

Thm continued: Furthermore, if the j-th cumulant of L_1 exists and is written as $\kappa_j\left(L_1\right)$, then

$$\begin{aligned}
\kappa_j(P_t - P_0) &= bt\kappa_j(L_1), \quad j = 1, 3, 5, ... \\
\kappa_j(P_t - P_0) &= \left\{bt + 2leb(A_t \backslash A)\right\}\kappa_j(L_1), \quad j = 2, 4, 6, ...
\end{aligned}$$

Notice that $C_t \backslash C_0$ has the physical interpretation of arrivals since time 0.

Long run returns

- If $\kappa_2(L_1) < \infty$, then

$$t^{-1/2}\left(P_t - P_0 - bt\kappa_1(L_1)\right) \xrightarrow{\mathcal{L}} N\left(0, b\kappa_2(L_1)\right) \quad \text{as } t \to \infty.$$

- This is the obvious result that the fleeting returns have no impact in the long run and that the non-Gaussian becomes irrelevant.

Theorem

Assume that $\kappa_2(L_1) < \infty$*. Then the gross returns have the autocorrelation structure, for some sampling interval* $\delta > 0$ *and* $k = 1, 2, ...$

$$
\begin{aligned}
\gamma_k &\triangleq \operatorname{Cov}\left(\left(P_{(k+1)\delta} - P_{k\delta}\right), \left(P_\delta - P_0\right)\right) \\
&= \left(leb(A_{(k+1)\delta} \backslash A) - 2leb(A_{k\delta} \backslash A) + leb(A_{(k-1)\delta} \backslash A)\right) \kappa_2(L_1), \\
\rho_k &\triangleq \operatorname{Cor}\left(\left(P_{(k+1)\delta} - P_{k\delta}\right), \left(P_\delta - P_0\right)\right) \\
&= \frac{leb(A_{(k+1)\delta} \backslash A) - 2leb(A_{k\delta} \backslash A) + leb(A_{(k-1)\delta} \backslash A)}{b\delta + 2leb(A_\delta \backslash A)}.
\end{aligned}
$$

Corollary

$\rho_k \leq 0$ *for all* $k = 1, 2, \ldots$. *This inequality becomes strict when* d *is strictly increasing (i.e.* $d(s_1) < d(s_2)$ *for all* $s_1 < s_2 \leq 0$*).*

g-variation

Quadratic variation plays a large role in modern financial econometrics (e.g. ABDL (01), BNS (02)). Extensions to power variation were rationalised in econometrics by BNS(04,06). More general functions were introduced by BN, Graversen, Jacod, S (06a,b). Here

- Recall the Lévy basis is

$$L(\mathrm{d}x, \mathrm{d}s) = \int_{-\infty}^{\infty} yN(\mathrm{d}y, \mathrm{d}x, \mathrm{d}s), \qquad (x, s) \in [0, 1] \times \mathbb{R}.$$

- Define the g-Lévy basis as

$$\Sigma(\mathrm{d}x, \mathrm{d}s; g) = \int_{-\infty}^{\infty} g(y)N(\mathrm{d}y, \mathrm{d}x, \mathrm{d}s),$$

  with mean measure

$$\mu(\mathrm{d}x, \mathrm{d}s; g) = \mathrm{d}x\mathrm{d}s \int_{-\infty}^{\infty} g(y)\nu(\mathrm{d}y),$$

  assuming $\int_{-\infty}^{\infty} g(y)\nu(\mathrm{d}y) < \infty$.

g-variation

- Then the unnormalised g-variation is

$$\{P;g\}_t = \lim_{\delta \to 0} \sum_{k=1}^{t/\delta} g\left(P_{k\delta} - P_{(k-1)\delta}\right) = \sum_{0<s\leq t} g(\Delta P_s).$$

  This is always finite.
- Quadratic case: many econometric researchers in effect assume a priori that this is infinity. This does not match the data or the predictions from our model.

$$\{P; g\}_t = \Sigma(B_t; g) + Z_t(g), \quad B_t \triangleq [0, b) \times (0, t],$$

where the impact of the fleeting events is $Z_t(g)$

$$Z_t(g) = \Sigma(H_t; g) + \Sigma(G_t; g), \quad H_t \triangleq [b, 1] \times (0, t], \quad G_t \triangleq (H_t \cup A) \setminus A_t.$$

Further,

$$\mathrm{E}\left[\{P; g\}_t\right] = (2 - b)t \int_{-\infty}^{\infty} g(y)\nu(\mathrm{d}y) = \mathrm{E}\left[\{P; 1\}_t\right] \frac{\int_{-\infty}^{\infty} g(y)\nu(\mathrm{d}y)}{\|\nu\|}.$$

Example: realized variance

- The realized variance (ABDL (01), BNS(02)) is

$$RV^{(n)} \triangleq \sum_{k=1}^{n} (P_{k\delta_n} - P_{(k-1)\delta_n})^2, \qquad \delta_n \triangleq \frac{T}{n}.$$

- Assume that $\kappa_2(L_1) < \infty$. Then

$$\mathbb{E}\left(RV^{(n)}\right) = \left(b + 2\frac{leb\left(A_{\delta_n} \backslash A\right)}{\delta_n}\right) T\kappa_2\left(L_1\right) + b^2 T\delta_n \kappa_1^2\left(L_1\right).$$

- For $n = 1$, as $T \to \infty$,

$$\mathbb{E}\left(RV^{(1)}\right) = \left(b + 2\frac{leb\left(A_T \backslash A\right)}{T}\right) T\kappa_2\left(L_1\right) + b^2 T^2 \kappa_1^2\left(L_1\right) \approx \mathbb{E}\left(L\left(B_T\right)^2\right.$$

- For $n \to \infty$ and a fixed T,

$$\lim_{n\to\infty} \mathbb{E}\left(RV^{(n)}\right) = (2-b)\, T\kappa_2\left(L_1\right).$$

- The QV is highly distorted by the fleeting component.

Model of trawl function

Example

A class of squashed monotonic trawls is the *superposition model*

$$d(s) = b + (1-b)\int_0^\infty e^{\lambda s}\pi(\mathrm{d}\lambda), \quad s \le 0, \tag{5}$$

where π is an arbitrary probability measure on $(0,\infty)$. Then

$$leb(A) = (1-b)\int_0^\infty \frac{1}{\lambda}\pi(\mathrm{d}\lambda), \qquad leb(A_t \backslash A) = (1-b)\int_0^\infty \frac{1-e^{-t\lambda}}{\lambda}\pi(\mathrm{d}\lambda).$$

Special cases

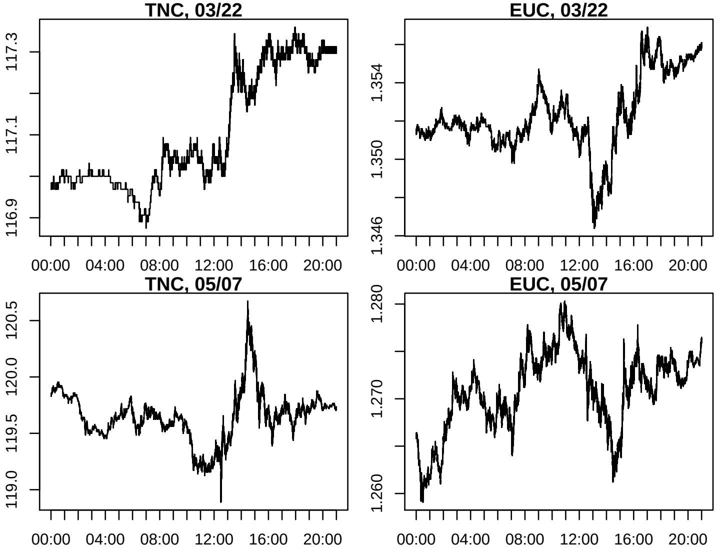
- single atom (exponential trawl function)
- gamma (allow long memory for some parameters)
- GIG (which includes inverse gamma)

Empirical analysis using moment based estimates

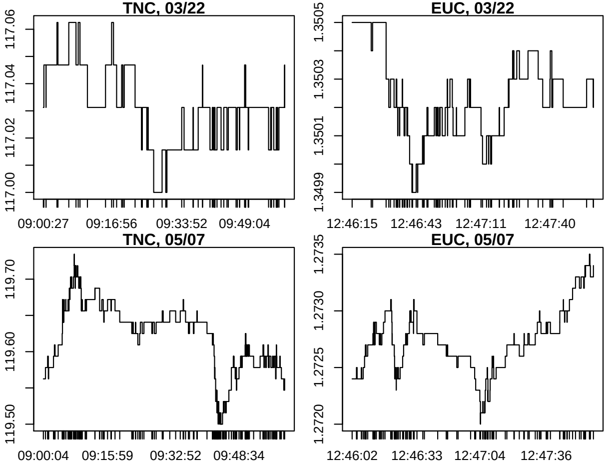
In this Subsection, we employ the moment-based estimation for empirical analysis. Four data set are studied here:

1. the Ten-Year US Treasury Note future contract delivered in June 2010 (TNC1006) during March 22, 2010;
2. the International Monetary Market (IMM) Euro-Dollar Foreign Exchange (EUC1006) future contract during March 22, 2010;
3. TNC1006 during May 7, 2010;
4. EUC1006 during May 7, 2010. Figure From now on, we will no longer mention the delivery date of each data set and the year 2010.

All of these four data sets use all the trading activity from 00:00 to 21:00.



Large time scale: trace plots look like a continuous diffusion process. At a much smaller time scale (within one hour).



Discreteness of the data becomes distinctive.

See several multiple ticks jump in the two EUC data sets.

Summary statistics

Table summarizes some basic features of these four data set.

Contract, Day	Tick Size ($)	# Jumps	SD. (in ticks):	Min.	Max.
TNC, 03/22	1/64	3,249	1.000	−1	1
EUC, 03/22	0.0001	13,943	1.012	−2	3
TNC, 05/07	1/64	12,849	1.035	−13	15
EUC, 05/07	0.0001	55,379	1.077	−13	15

Estimation: exponential trawl

Here $d(s) = b + (1-b)\exp(\lambda s), \nu^+ \triangleq \sum_{y=1}^{\infty} \nu(y)$ and $\nu^- \triangleq \sum_{y=-\infty}^{-1} \nu(y)$.

- Core results:

Contract, Day	b	ν^+	ν^-	λ
TNC, 03/22	0.396	0.014	0.013	0.68
EUC, 03/22	0.654	0.069	0.068	2.47
TNC, 05/07	0.574	0.059	0.060	3.88
EUC, 05/07	0.694	0.282	0.279	4.03

- TNC, 03/22 estimate 40% of price moves are permanent.

Estimation: sup GIG trawl

- Core results: Then

$$\pi\left(\mathrm{d}\lambda\right)=\frac{\left(\gamma/\delta\right)^{\nu}}{2K_{\nu}\left(\gamma\delta\right)}\lambda^{\nu-1}e^{-\left(\gamma^{2}\lambda+\delta^{2}\lambda^{-1}\right)/2},\qquad \gamma,\delta>0,\ \nu\in\mathbb{R},$$

where $K_{\nu}\left(x\right)$ is the *modified Bessel function of the second kind*. Implies

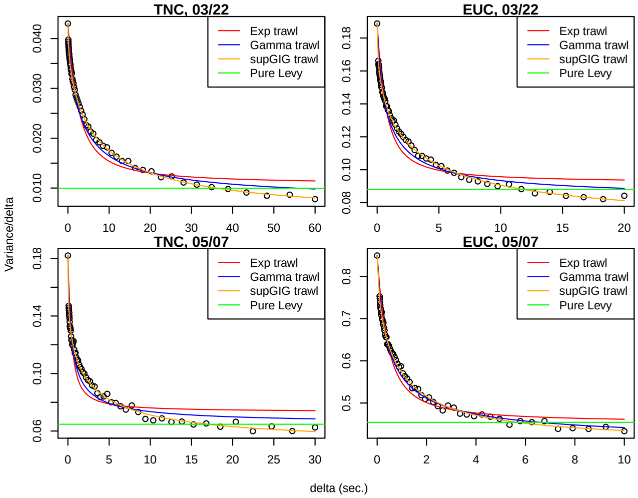
$$d\left(s\right)=\left(1-\frac{2s}{\gamma^{2}}\right)^{-\nu/2}\frac{K_{\nu}\left(\gamma\delta\sqrt{1-2s/\gamma^{2}}\right)}{K_{\nu}\left(\gamma\delta\right)}.$$

Contract, Day	b	ν^{+}	ν^{-}	γ	δ	ν
TNC, 03/22	0.186	0.013	0.011	0	0.453	-0.604
EUC, 03/22	0.528	0.063	0.062	0	0.604	-0.453
TNC, 05/07	0.440	0.054	0.055	0	0.583	-0.332
EUC, 05/07	0.648	0.272	0.269	0	1.525	-0.741

GIG is collapsing to inverse gamma trawel.

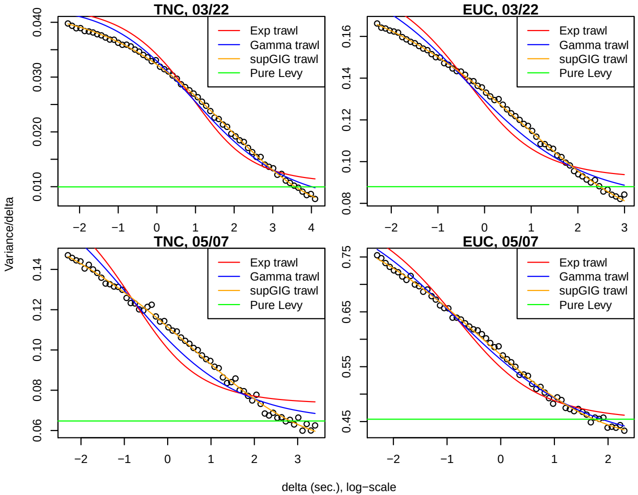
Scaled variogram

$$\frac{\sigma_\delta^2}{\delta} = \frac{1}{\delta}\mathrm{Var}\left(P_{t+\delta} - P_t\right).$$



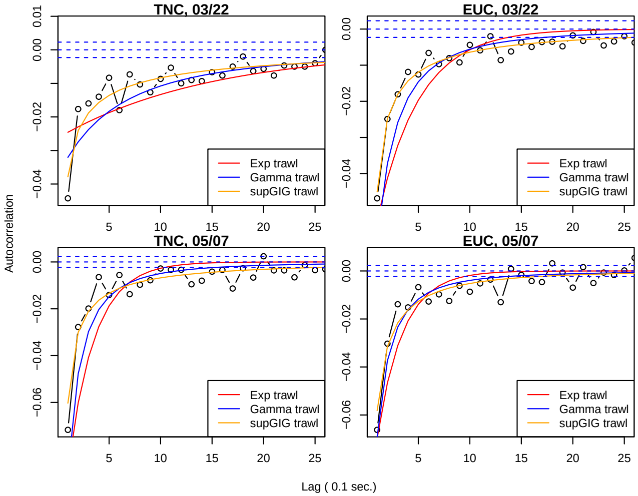
Scaled variogram — log time

$$\sigma_\delta^2/\delta = \frac{1}{\delta}\mathrm{Var}\left(P_{t+\delta} - P_t\right).$$



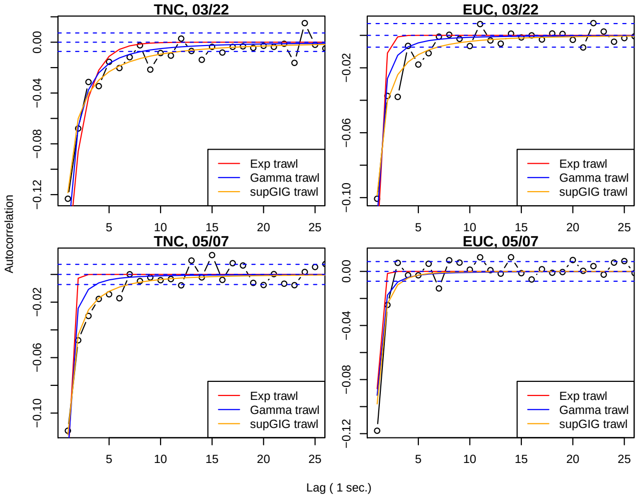
Autocorrelation at 0.1sec lags

$$\mathrm{Cor}\left\{\left(P_{(k+1)\delta} - P_{k\delta}\right), \left(P_{\delta} - P_{0}\right)\right\}$$



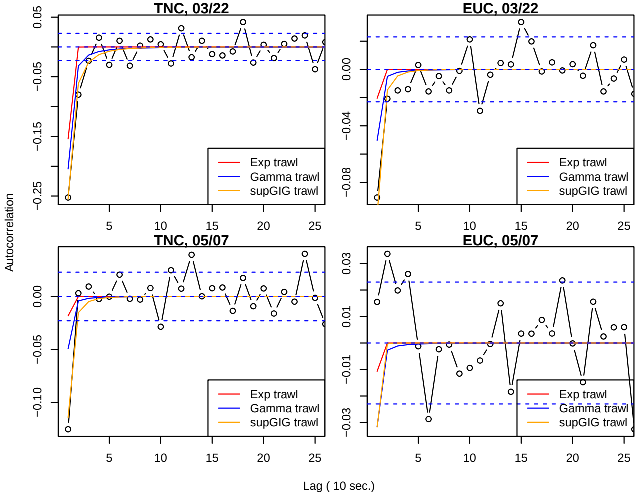
Autocorrelation at 1 second lags

$$\mathrm{Cor}\left\{\left(P_{(k+1)\delta} - P_{k\delta}\right), \left(P_{\delta} - P_0\right)\right\}$$



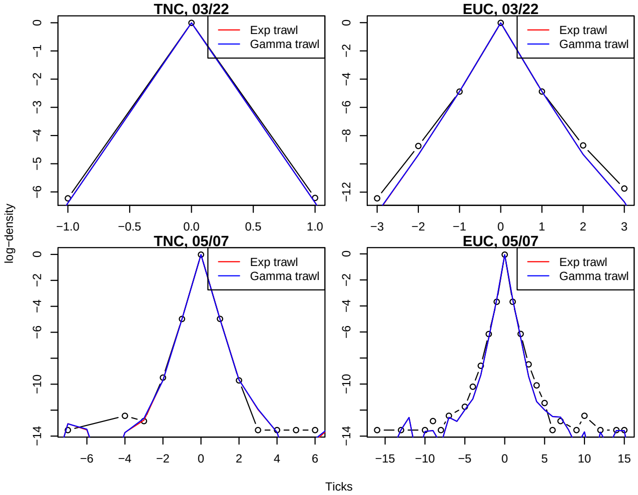
Autocorrelation at 10 second lags

$$\mathrm{Cor}\left\{\left(P_{(k+1)\delta} - P_{k\delta}\right), \left(P_{\delta} - P_{0}\right)\right\}$$



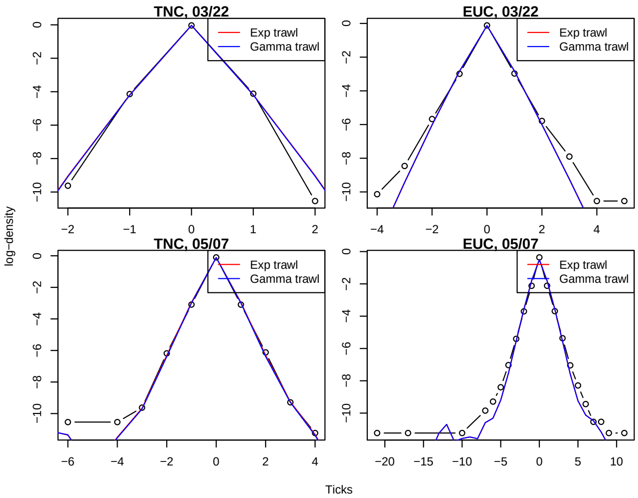
ACF is non-monotonic. Consistent with the models.

Log-probability at 0.1 seconds



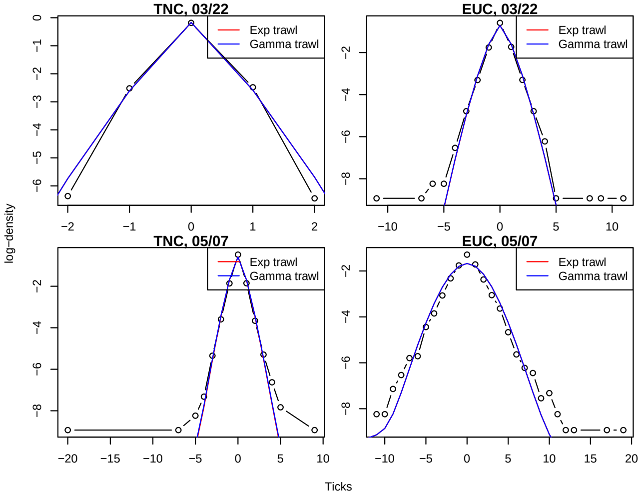
Circles: raw log histogram

Log-probability at 1 seconds



Circles: raw log histogram

Log-probability at 10 seconds



Circles: raw log histogram

Conclusions

- For high frequency data, discreteness is dominant.
  - build models for the data we have in our hand
  - model structure determined by the specifics of the problem.
- Continuous time, non-stationary discrete model
- Flexible memory, analytically tractable & easy to simulate
- Moment based estimation is easy.
- Nice cumulant functions (stochastic discount factors).
- Extensions being worked on
  - Understanding impotence of robust measures (kernels, 2 scale, preaveraging, etc) and what to do (with Mikkel Bennedsen)
  - Filtering (is a new price arrival fleeting or permanent)
  - Multi case (random delay Lévy process) to capture Epps effects.
  - Allow parameters of the model to wobble through time
    - Conditioning on other information, e.g. order book
    - Stochastic processes, e.g. stochastic volatility