Variance derivatives and estimating realised variance from high-frequency data

John Crosby

UBS, London
and Centre for Economic and Financial Studies,
Department of Economics,
Glasgow University

Presentation at
the "Quantitative Trading and Asset Management" conference
at Columbia University
on Friday, 19th November 2010

File date 19.38 on 11th November 2010

Acknowledgements

- This is based on joint work with Mark Davis at Imperial College London.
- We thank Peter Carr, Floyd Hanson, Roger Lee, Aleksandar Mijatović and Vimal Raval.

- We denote the initial time (today) by $t_0 \equiv 0$. We consider a stock whose price, at time t, is $S(t)$. We consider a time interval $[t_0, T]$ which is partitioned into N time periods (not necessarily equal in length) whose end-points are t_j, $j = 1, 2, ..., N$, where $0 \equiv t_0 < t_1 < ... < t_{j-1} < t_j < ... < t_N \equiv T$.
- What difference does it make if realised variance is measured by log changes squared (i.e. $\sum_{i=1}^{N}(\log(S(t_i)/S(t_{i-1})))^2$) or by proportional differences squared (i.e. $\sum_{i=1}^{N}((S(t_i)/S(t_{i-1})) - 1)^2$)?
- What impact does monitoring frequency (i.e. the value of N above) have on the measurement of realised variance?
- What impact do jumps in the underlying stock price have on the measurement of realised variance?
- Building on Broadie and Jain (2008), Carr and Lee (2009) and Hong (2004), we will try to answer these questions.

- Our results have two important applications:
- 1./ The pricing (under an equivalent martingale measure (EMM) $\mathbb{Q}$) of variance swaps which pay $\sum_{i=1}^{N}(\log(S(t_i)/S(t_{i-1})))^2$ (which is how the payoffs are usually defined in practice) and of proportional variance swaps which pay $\sum_{i=1}^{N}((S(t_i)/S(t_{i-1})) - 1)^2$ at maturity T. In particular, we consider the case when N is infinite (continuously monitored) and the case when N is finite (discretely monitored - as they must always be in practice).
- 2./ Given observations of $S(t_i)$ for times t_i, $i = 1, 2, ..., N$ (from historical data under the real-world physical measure $\mathbb{P}$), what can we say about the process which generated this data? We are thinking, in particular, of high-frequency data (at least several, perhaps, a few hundred observations per day).

- For the first two-thirds of my talk, I will focus on variance swaps and model stock price dynamics under an EMM $\mathbb{Q}$.
- Nearly all papers on variance swaps have focussed on the log-contract replication approach (eg. Neuberger (1990), Dupire (1993), Derman et al. (1999)).
- However, there is a completely different approach (see Hong (2004) and Broadie and Jain (2008)) which utilises characteristic functions. We build upon this approach. However, firstly, we discuss the assumed stock price dynamics.

- We construct the stock price process by assuming that the log of the stock price is a time-changed Lévy process (allows a very generic process which includes (nearly) all models seen in the literature).
- We have a Lévy process (eg Brownian motion, Kou (2002) jump-diffusion, Variance Gamma or CGMY) denoted by X_t, satisfying $X_{t_0} = 0$. We assume that we mean-correct X_t so that $\exp(X_t)$ is a (non-constant) martingale (under $\mathbb{Q}$) - with respect to the natural filtration generated by X_t i.e. that $\mathbb{E}^{\mathbb{Q}}_{t_0}[\exp(X_t)] = \exp(X_{t_0}) = 1$ for all $t \geq t_0$.
- Lévy-Khinchin formula implies we can write the (mean-corrected) characteristic exponent $\overline{\psi}_X(z)$ (defined via $\mathbb{E}^{\mathbb{Q}}_{t_0}[\exp(iuX_t)] \equiv \exp(-(t-t_0)\overline{\psi}_X(u))$) in the form:

$$-\overline{\psi}_X(z) = -\frac{1}{2}\sigma^2(z^2 + iz) + \int_{-\infty}^{\infty} (\exp(izx) - 1 - iz(\exp(x) - 1))\nu(dx).$$

  For future reference, $'$ denotes differentiation i.e. $\overline{\psi}'_X(z) \equiv \partial\overline{\psi}_X(z)/\partial z$, $\overline{\psi}''_X(z) \equiv \partial^2\overline{\psi}_X(z)/\partial z^2$ and $\overline{\psi}'''_X(z) \equiv \partial^3\overline{\psi}_X(z)/\partial z^3$.
- For the case of Brownian motion, "$X_t = -\frac{1}{2}\sigma^2 t + \sigma W(t)$ where $W(t)$ is standard (driftless) Brownian motion".

- We assume that we have a non-decreasing, continuous time-change process denoted by Y_t. We normalise so that $Y_{t_0} = t_0 \equiv 0$.
- In general, Y_t may be correlated with X_t.
- Our assumption, for example, allows Y_t to be of the form $Y_t = \int_{t_0}^{t} y_s ds$ where the activity rate y_t (which must be non-negative) follows, for example, a Heston (1993) square-root process, a non-Gaussian OU process (Barndorff-Nielsen and Shephard (2001)) or it could follow the Heston (1993) plus jumps process of Duffie et al. (2000). In the latter two cases, y_t is discontinuous but Y_t is always continuous.
- (The time-change will allow us to model stochastic volatility / leverage / volatility clustering type effects).

- We time-change the Lévy process X_t by Y_t to get a process X_{Y_t}, with $X_{Y_{t_0}} = 0$.
- The stock price $S(t)$, at time t, is assumed to have the following dynamics (under $\mathbb{Q}$):

$$S(t) = S(t_0)\exp\left(\int_{t_0}^{t} (r(s) - q(s))ds + X_{Y_t}\right).$$

- Here, $r(t)$ is the risk-free interest-rate and $q(t)$ is the dividend yield (assumed finite and deterministic), at time t.
- To lighten notation, I will henceforth write equations as if $r(t) - q(t) \equiv 0$ for all t (or equivalently work with forward or future prices - the paper considers the general case). Hence, $S(t) = S(t_0)\exp(X_{Y_t})$.

- We now define, for all $t \geq t_0$:

$$\Xi_t(u) \equiv \exp(iuX_{Y_t} + Y_t\overline{\psi}_X(u)).$$

Since the mean-corrected characteristic exponent $\overline{\psi}_X(u)$ is defined via: $\mathbb{E}^{\mathbb{Q}}_{t_0}[\exp(iuX_t)] = \exp(-(t-t_0)\overline{\psi}_X(u))$, then $\exp(iuX_t + (t-t_0)\overline{\psi}_X(u))$ is a martingale, under $\mathbb{Q}$, with respect to the natural filtration generated by X_t.

- By a "randomising time" (Optional Stopping Theorem) argument, for any u, $\Xi_t(u)$ is a martingale, under $\mathbb{Q}$, with respect to the filtration generated by $\mathcal{F}_t \equiv \{X_t \cup Y_t\}$.

- In particular,

$$\mathbb{E}^{\mathbb{Q}}_{t_{j-1}}[\frac{\Xi_{t_j}(u)}{\Xi_{t_{j-1}}(u)}] = \mathbb{E}^{\mathbb{Q}}_{t_{j-1}}[\exp(iu(X_{Y_{t_j}} - X_{Y_{t_{j-1}}}) + (Y_{t_j} - Y_{t_{j-1}})\overline{\psi}_X(u))] = 1.$$

- We now introduce what we call the joint extended characteristic function $\Phi(z;j)$, which we define, for each j, $j = 1, \ldots, N$, by:

$$\begin{aligned}\Phi(z;j) &\equiv \mathbb{E}^{\mathbb{Q}}_{t_0}[\exp(iz\log\frac{S(t_j)}{S(t_{j-1})})] = \mathbb{E}^{\mathbb{Q}}_{t_0}[\exp(iz(X_{Y_{t_j}} - X_{Y_{t_{j-1}}}))] \\ &= \mathbb{E}^{\mathbb{Q}}_{t_0}[\exp(iz(X_{Y_{t_j}} - X_{Y_{t_{j-1}}}) + (Y_{t_j} - Y_{t_{j-1}})\overline{\psi}_X(z))\exp(-(Y_{t_j} - Y_{t_{j-1}})\overline{\psi}_X(z))] \\ &= \mathbb{E}^{\mathbb{Q}}_{t_0}[\mathbb{E}^{\mathbb{Q}}_{t_{j-1}}[\frac{\Xi_{t_j}(z)}{\Xi_{t_{j-1}}(z)}\exp(-(Y_{t_j} - Y_{t_{j-1}})\overline{\psi}_X(z))]].\end{aligned}$$

- (Note as an aside, $\Phi(z;j)$ is "a kind of forward characteristic function". One can compute $\Phi(z;j)$, for cases of interest, via conditioning arguments and by using results in Carr and Wu (2004) and Duffie et al. (2000), so we will say nothing more about this.)

- We note that the joint extended characteristic function $\Phi(z;j)$ allows us to immediately evaluate the price of a discretely monitored proportional variance swap. We let $iz = 2$ in the equation for $\Phi(z;j)$, then sum over j and simplify.
- $\Rightarrow$: The price $\mathrm{PVS}(t_0, T, N)$, at time t_0, of a (discretely monitored) proportional variance swap (paying $\sum_{i=1}^{N}((S(t_i)/S(t_{i-1})) - 1)^2$ at time T) is:

$$\mathrm{PVS}(t_0, T, N) = P(t_0, T)\Big(\sum_{j=1}^{N}\big(\Phi(-2i;j) - 1\big)\Big).$$

  Here, $P(t_0, T)$ is the price of a zero-coupon bond, at time t_0, that matures at time T.
- We will examine the limit as $N \to \infty$ of this equation later.

- Now we differentiate $\Phi(z;j)$ with respect to z and divide by i:

$$\begin{aligned}
\frac{1}{i}\frac{\partial\Phi(z;j)}{\partial z} &= \mathbb{E}^{\mathbb{Q}}_{t_0}[\log\frac{S(t_j)}{S(t_{j-1})}\exp(iz\log\frac{S(t_j)}{S(t_{j-1})})] \\
&= \mathbb{E}^{\mathbb{Q}}_{t_0}[\mathbb{E}^{\mathbb{Q}}_{t_{j-1}}[\frac{\Xi_{t_j}(z)}{\Xi_{t_{j-1}}(z)}\exp(-(Y_{t_j}-Y_{t_{j-1}})\overline{\psi}_X(z)) \\
&\quad \Big(\varpi^{(j)}(iz)+\big((X_{Y_{t_j}}-X_{Y_{t_{j-1}}})-i\overline{\psi}'_X(z)(Y_{t_j}-Y_{t_{j-1}})\big)\Big)]], \qquad \text{where} \\
\varpi^{(j)}(iz) &\equiv i\overline{\psi}'_X(z)(Y_{t_j}-Y_{t_{j-1}}).
\end{aligned}$$

- It is now straightforward to value log-forward-contracts (paying $\log(S(t_N)/S(t_0))$ at time T). We set $iz=0$, then we sum from $j=1$ to N and then simplify. The price $\mathrm{LFC}(t_0,T)$, at time t_0, of a log-forward-contract is:

$$\mathrm{LFC}(t_0,T) = P(t_0,T)i\overline{\psi}'_X(0)\mathbb{E}^{\mathbb{Q}}_{t_0}[Y_T-Y_{t_0}] \equiv P(t_0,T)m_X\mathbb{E}^{\mathbb{Q}}_{t_0}[Y_T-Y_{t_0}].$$

Note m_X defined by $m_X \equiv i\overline{\psi}'_X(0)$ is real.

- We differentiate again with respect to z and again divide by i:

$$
\begin{aligned}
-\frac{\partial^2\Phi(z;j)}{\partial z^2} &= \mathbb{E}_{t_0}^{\mathbb{Q}}[(\log \frac{S(t_j)}{S(t_{j-1})})^2 \exp(iz\log \frac{S(t_j)}{S(t_{j-1})})] \\
&= \mathbb{E}_{t_0}^{\mathbb{Q}}[\mathbb{E}_{t_{j-1}}^{\mathbb{Q}}[\frac{\Xi_{t_j}(z)}{\Xi_{t_{j-1}}(z)} \exp(-(Y_{t_j} - Y_{t_{j-1}})\overline{\psi}_X(z)) \\
&\quad \Big(\varpi^{(j)\,2}(iz) \\
&\quad +\Big\{2\varpi^{(j)}(iz)\left((X_{Y_{t_j}} - X_{Y_{t_{j-1}}}) - i\overline{\psi}'_X(z)(Y_{t_j} - Y_{t_{j-1}})\right)\Big\} \\
&\quad +\left((X_{Y_{t_j}} - X_{Y_{t_{j-1}}}) - i\overline{\psi}'_X(z)(Y_{t_j} - Y_{t_{j-1}})\right)^2 - \overline{\psi}''_X(z)(Y_{t_j} - Y_{t_{j-1}}) \\
&\quad +\overline{\psi}''_X(z)(Y_{t_j} - Y_{t_{j-1}})\Big)]].
\end{aligned}
$$

.

- The price, at time t_0, of a variance swap $\mathrm{VS}(t_0,T,N)$ can be obtained by setting $iz=0$, summing from $j=1$ to N and simplifying: The price $\mathrm{VS}(t_0,T,N)$ is:

$$\begin{aligned} &\mathrm{VS}(t_0,T,N) \\ =\ & P(t_0,T)\mathbb{E}^{\mathbb{Q}}_{t_0}[\sum_{j=1}^{N}\mathbb{E}^{\mathbb{Q}}_{t_{j-1}}[\varpi^{(j)\,2}(0)]] \\ +\ & P(t_0,T)\mathbb{E}^{\mathbb{Q}}_{t_0}[\sum_{j=1}^{N}\mathbb{E}^{\mathbb{Q}}_{t_{j-1}}[2m_X(Y_{t_j}-Y_{t_{j-1}})\big((X_{Y_{t_j}}-X_{Y_{t_{j-1}}})-m_X(Y_{t_j}-Y_{t_{j-1}})\big)]] \\ +\ & P(t_0,T)\overline{\psi}''_X(0)\mathbb{E}^{\mathbb{Q}}_{t_0}[\sum_{j=1}^{N}(Y_{t_j}-Y_{t_{j-1}})]. \end{aligned} \tag{1}$$

- Note that $\varpi^{(j)}(0)$ is the drift of log of the stock price (over the time interval t_{j-1} to t_j) (it is real and for Brownian motion and a deterministic time-change it is "$(r-q-\frac{1}{2}\sigma^2)(t_j-t_{j-1})$").
- Here $m_X \equiv i\overline{\psi}'_X(0)$ (note m_X is real and for Browian motion it is "$-\frac{1}{2}\sigma^2$").
- Lets look at each of the three lines of equation (1) in turn.

- Again, $\mathrm{VS}(t_0, T, N)$

$$
\begin{aligned}
&= P(t_0,T)\mathbb{E}^{\mathbb{Q}}_{t_0}[\sum_{j=1}^{N}\mathbb{E}^{\mathbb{Q}}_{t_{j-1}}[\varpi^{(j)\,2}(0)]] \\
&+ P(t_0,T)\mathbb{E}^{\mathbb{Q}}_{t_0}[\sum_{j=1}^{N}\mathbb{E}^{\mathbb{Q}}_{t_{j-1}}[2m_X(Y_{t_j}-Y_{t_{j-1}})\big((X_{Y_{t_j}}-X_{Y_{t_{j-1}}})-m_X(Y_{t_j}-Y_{t_{j-1}})\big)]] \\
&+ P(t_0,T)\overline{\psi}''_X(0)\mathbb{E}^{\mathbb{Q}}_{t_0}[\sum_{j=1}^{N}(Y_{t_j}-Y_{t_{j-1}})].
\end{aligned}
$$

- Note that, with a deterministic time-change, $\varpi^{(j)\,2}(0)$ is $O(1/N^2)$. Broadie and Jain (2008) show that it is $O(1/N^2)$ if the activity-rate of the time-change is Heston (1993). In the paper, we show that it is $O(1/N^2)$ for "almost any" time-change.
- Hence the first line is $O(1/N)$ and $\to 0$ as $N \to \infty$.
- Since $\varpi^{(j)}(0)$ is real, $\varpi^{(j)\,2}(0)$ is definitely non-negative and zero only if the drift of the log of the stock price is identically equal to zero.

- Again, the second line is:

$$P(t_0,T)\mathbb{E}^{\mathbb{Q}}_{t_0}[\sum_{j=1}^{N}\mathbb{E}^{\mathbb{Q}}_{t_{j-1}}[2m_X(Y_{t_j}-Y_{t_{j-1}})\big((X_{Y_{t_j}}-X_{Y_{t_{j-1}}})-m_X(Y_{t_j}-Y_{t_{j-1}})\big)]].$$

- Note $\mathbb{E}^{\mathbb{Q}}_{t_{j-1}}[(X_{Y_{t_j}}-X_{Y_{t_{j-1}}})-m_X(Y_{t_j}-Y_{t_{j-1}})]\equiv 0$ (by construction it is a martingale eg the whole term is standard Brownian motion).
- Therefore, if X_t and Y_t are independent, the second line is identically equal to zero.
- m_X is always negative (eg for Browian motion it is "$-\frac{1}{2}\sigma^2$"). Therefore, if X_t and Y_t are negatively correlated, the second term is positive.
- Results in Broadie and Jain (2008) show, for Heston (1993) that the (absolute value of the) second line is $O(1/N)$. In the paper, we show that it is $O(1/N)$ for any Lévy process and "almost any" time-change.

- Again, VS(t_0, T, N)

$$
\begin{aligned}
&= P(t_0,T)\mathbb{E}^{\mathbb{Q}}_{t_0}[\sum_{j=1}^{N}\mathbb{E}^{\mathbb{Q}}_{t_{j-1}}[\varpi^{(j)\,2}(0)]] \\
&+ P(t_0,T)\mathbb{E}^{\mathbb{Q}}_{t_0}[\sum_{j=1}^{N}\mathbb{E}^{\mathbb{Q}}_{t_{j-1}}[2m_X(Y_{t_j}-Y_{t_{j-1}})\big((X_{Y_{t_j}}-X_{Y_{t_{j-1}}})-m_X(Y_{t_j}-Y_{t_{j-1}})\big)]] \\
&+ P(t_0,T)\overline{\psi}''_X(0)\mathbb{E}^{\mathbb{Q}}_{t_0}[Y_T-Y_{t_0}].
\end{aligned}
$$

- The term $\mathbb{E}^{\mathbb{Q}}_{t_0}[\sum_{j=1}^{N}(Y_{t_j}-Y_{t_{j-1}})] = \mathbb{E}^{\mathbb{Q}}_{t_0}[Y_T-Y_{t_0}]$ due to a telescoping sum.
- The third line is the price of the continuously monitored version of the variance swap.

- Again, $\mathrm{VS}(t_0, T, N)$

$$
\begin{aligned}
&= P(t_0,T)\mathbb{E}^{\mathbb{Q}}_{t_0}[\sum_{j=1}^{N}\mathbb{E}^{\mathbb{Q}}_{t_{j-1}}[\varpi^{(j)\,2}(0)]] \\
&+ P(t_0,T)\mathbb{E}^{\mathbb{Q}}_{t_0}[\sum_{j=1}^{N}\mathbb{E}^{\mathbb{Q}}_{t_{j-1}}[2m_X(Y_{t_j}-Y_{t_{j-1}})\big((X_{Y_{t_j}}-X_{Y_{t_{j-1}}})-m_X(Y_{t_j}-Y_{t_{j-1}})\big)]] \\
&+ P(t_0,T)\overline{\psi}''_X(0)\mathbb{E}^{\mathbb{Q}}_{t_0}[Y_T-Y_{t_0}].
\end{aligned}
$$

- The price of a (discretely monitored) variance swap is the sum of three terms: A non-negative "drift-related" term, a "covariance" term which is non-negative (respectively, zero) if $\mathrm{Correl}(X_t, Y_t)$ is negative (respectively, zero) and the price of the continuously monitored version of the variance swap.
- In particular, if the "covariance" term is non-positive, a discretely monitored variance swap is always worth than its continuously monitored counterpart.
- Convergence is always $O(1/N)$.

- We saw earlier that the price $\mathrm{PVS}(t_0, T, N)$, at time t_0, of a (discretely monitored) proportional variance swap (paying $\sum_{i=1}^{N}((S(t_i)/S(t_{i-1})) - 1)^2$ at time T) is:

$$\mathrm{PVS}(t_0, T, N) = P(t_0, T)\Big(\sum_{j=1}^{N}\big(\Phi(-2i; j) - 1\big)\Big).$$

- Hence:

$$\begin{aligned}\lim_{N\to\infty}\mathrm{PVS}(t_0, T, N) &= \lim_{N\to\infty} P(t_0, T)\Big(\sum_{j=1}^{N}\big(\Phi(-2i; j) - 1\big)\Big)\\ &= P(t_0, T)\lim_{N\to\infty}\sum_{j=1}^{N}\mathbb{E}_{t_0}^{\mathbb{Q}}[\frac{\Xi_{t_j}(-2i)}{\Xi_{t_{j-1}}(-2i)}(\exp(-(Y_{t_j} - Y_{t_{j-1}})\overline{\psi}_X(-2i)) - 1)]\\ &= -P(t_0, T)\overline{\psi}_X(-2i)\mathbb{E}_{t_0}^{\mathbb{Q}}[Y_T - Y_{t_0}] + O(1/N).\end{aligned}$$

- Hence, the price of the continuously monitored version of the proportional variance swap is $-P(t_0, T)\overline{\psi}_X(-2i)\mathbb{E}_{t_0}^{\mathbb{Q}}[Y_T - Y_{t_0}]$.
- Convergence is also $O(1/N)$.

- From the previous slide,

$$\begin{aligned}\mathrm{PVS}(t_0,T,N) &= P(t_0,T)\Big(\sum_{j=1}^{N}\big(\Phi(-2i;j)-1\big)\Big) \text{ with}\\ \Phi(-2i;j) &= \mathbb{E}^{\mathbb{Q}}_{t_0}\big[\frac{\Xi_{t_j}(-2i)}{\Xi_{t_{j-1}}(-2i)}\exp(-(Y_{t_j}-Y_{t_{j-1}})\overline{\psi}_X(-2i))\big].\end{aligned}$$

  Hence, it is clear (since $\overline{\psi}^{(k)}_X(-2i)<0$ eg. for Brownian motion $\overline{\psi}^{(k)}_X(-2i)=-\sigma^2$) that when X_t and Y_t are positively correlated then the price of a discretely monitored proportional variance swap is higher than the price of the same discretely monitored proportional variance swap under the assumption that they are independent (the opposite way round to a variance swap).

- Under the assumption of independence, a discretely monitored proportional variance swap is always worth at least as much as an otherwise identical continuously monitored proportional variance swap (the same way round as a variance swap).

Summary so far

- We have explicit expressions for the prices of variance swaps and proportional variance swaps (both discretely monitored and continuously monitored). Discretely monitored prices tend to their continuously monitored counterparts as $O(1/N)$ (for both variance swaps and proportional variance swaps).
- In the paper, we prove $O(1/N)$ convergence is also true for discontinuous time-changes.
- In the paper, we prove $O(1/N)$ convergence is also true for gamma swaps, self-quantoed variance swaps and skewness swaps.
- The prices of continuously monitored variance swaps and proportional variance swaps (and also gamma swaps and skewness swaps) do **NOT** depend upon $\mathrm{Correl}(X_t, Y_t)$.
- Can easily see dependence of discretely monitored versions of these swaps on $\mathrm{Correl}(X_t, Y_t)$.
- In particular,

$$\mathrm{VS}(t_0, T, N) \geq \mathrm{VS}(t_0, T, \infty) \text{ provided } \mathrm{Correl}(X_t, Y_t) \leq 0,$$

(and a non-positive correlation seems most likely in practice).

- The price of a continuously monitored proportional variance swap is:
  $\mathrm{PVS}(t_0, T, \infty) = -P(t_0, T)\overline{\psi}_X(-2i)\mathbb{E}^{\mathbb{Q}}_{t_0}[Y_T - Y_{t_0}]$.
- The price of a continuously monitored variance swap is:
  $\mathrm{VS}(t_0, T, \infty) = P(t_0, T)\overline{\psi}''_X(0)\mathbb{E}^{\mathbb{Q}}_{t_0}[Y_T - Y_{t_0}]$.
- The price of a log-forward-contract is:
  $\mathrm{LFC}(t_0, T) = P(t_0, T)i\overline{\psi}'_X(0)\mathbb{E}^{\mathbb{Q}}_{t_0}[Y_T - Y_{t_0}] \equiv P(t_0, T)m_X\mathbb{E}^{\mathbb{Q}}_{t_0}[Y_T - Y_{t_0}]$.
- Hence:
  $$\frac{\mathrm{VS}(t_0, T, \infty)}{\mathrm{LFC}(t_0, T)} = \frac{\overline{\psi}''_X(0)}{m_X}, \qquad \frac{\mathrm{PVS}(t_0, T, \infty)}{\mathrm{LFC}(t_0, T)} = \frac{-\overline{\psi}_X(-2i)}{m_X}.$$
  Carr and Lee (2009) have already proven the left-hand-side equation (i.e. for variance swaps (VS)) by a different method. In the paper, we show similar analogous results, not only for proportional variance swaps, but also for other types of variance derivatives.
- Hence, given vanilla prices, can price variance swaps and proportional variance swaps independent of any assumption on Y_t (and therefore robust to model (mis-)specification).

- For the case, when X_t is Brownian motion with volatility σ:
  We have: $\overline{\psi}_X(z) = \sigma^2(z^2 + iz)/2$, $m_X = -\sigma^2/2$, $\overline{\psi}''_X(0) = \sigma^2$, $\overline{\psi}''_X(-i) = \sigma^2$, $\overline{\psi}'''_X(0) = 0$ and $\overline{\psi}_X(-2i) = -\sigma^2$.

$$\frac{\mathrm{VS}(t_0, T, \infty)}{-\mathrm{LFC}(t_0, T)} = 2, \qquad \frac{\mathrm{PVS}(t_0, T, \infty)}{-\mathrm{LFC}(t_0, T)} = 2.$$

- The left-hand-side equation restates Neuberger (1990), Dupire (1993) and Derman et al. (1999):
  The price of a variance swap equals (minus) two times the price of a log-forward-contract (with the assumption of continuous sample paths (i.e. the log of the stock price is time-changed Brownian motion)).

- The right-hand-side equation says that it makes **no difference** if realised variance is measured by log changes squared (i.e. $\sum_{i=1}^{N}(\log(S(t_i)/S(t_{i-1})))^2)$ or by proportional differences squared (i.e. $\sum_{i=1}^{N}((S(t_i)/S(t_{i-1})) - 1)^2)$ **when there are no jumps (i.e. continuous sample paths) and when $N = \infty$ (i.e. continuously monitored)**.

- For the case, when X_t is a compound Poisson process with a fixed jump amplitude a (and with no diffusion component), then we have:

$$\begin{aligned}\frac{\mathrm{VS}(t_0,T,\infty)}{-\mathrm{LFC}(t_0,T)} &= \frac{a^2}{(\exp(a)-1-a)} \approx 2\left(1-\frac{a}{3}\right),\\ \frac{\mathrm{PVS}(t_0,T,\infty)}{-\mathrm{LFC}(t_0,T)} &= \frac{(\exp(a)-1)^2}{(\exp(a)-1-a)} \approx 2\left(1+\frac{2a}{3}\right),\end{aligned}$$

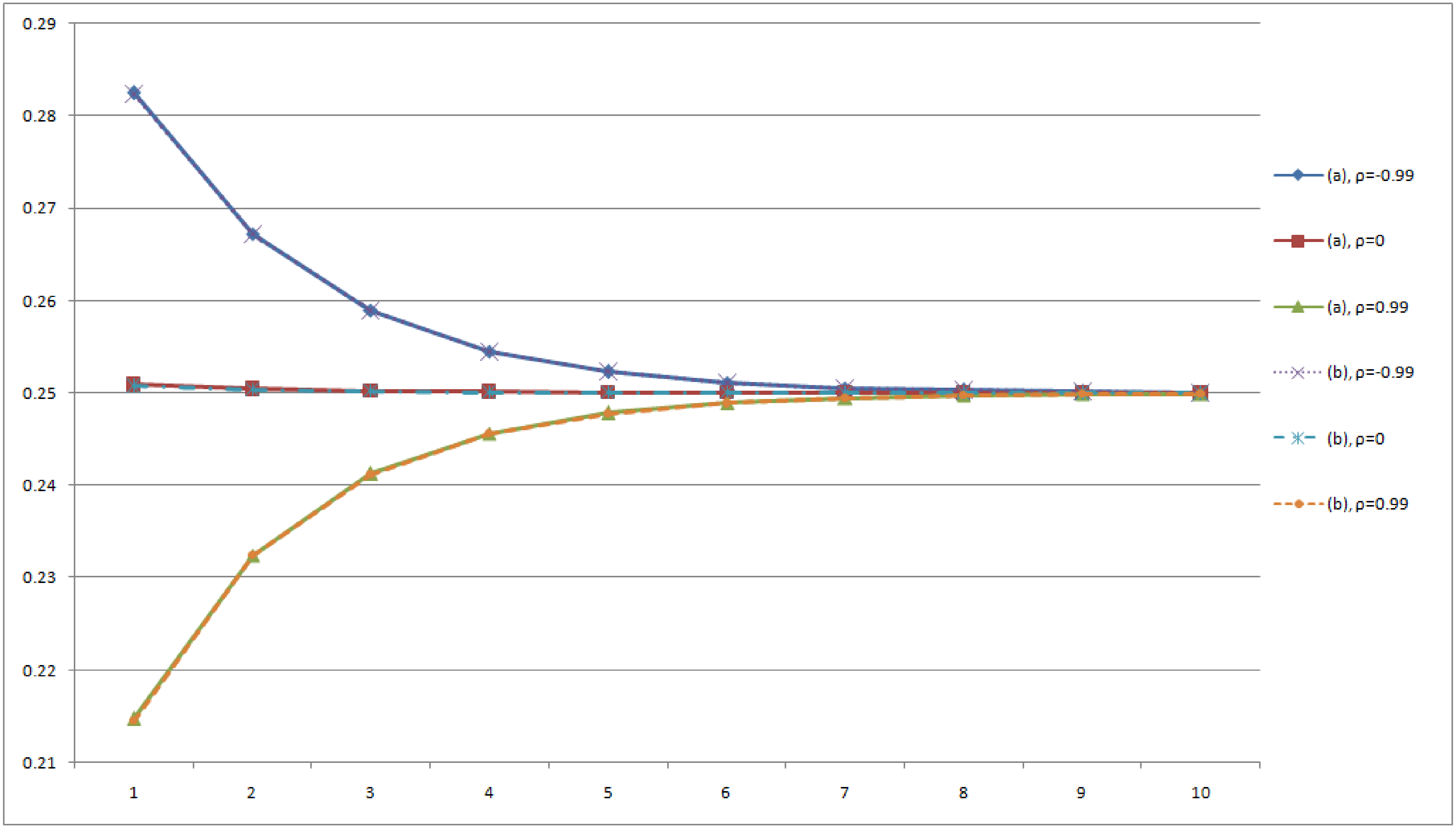
  where, in each part, the first term is exact and the second term is the expansion of the first term to leading order when $|a|$ is small.

- $\Rightarrow$: The prices of variance swaps and proportional variance swaps have the **opposite sensitivities** to jumps (and the impact will be larger in magnitude (perhaps, twice as large) for proportional variance swaps).

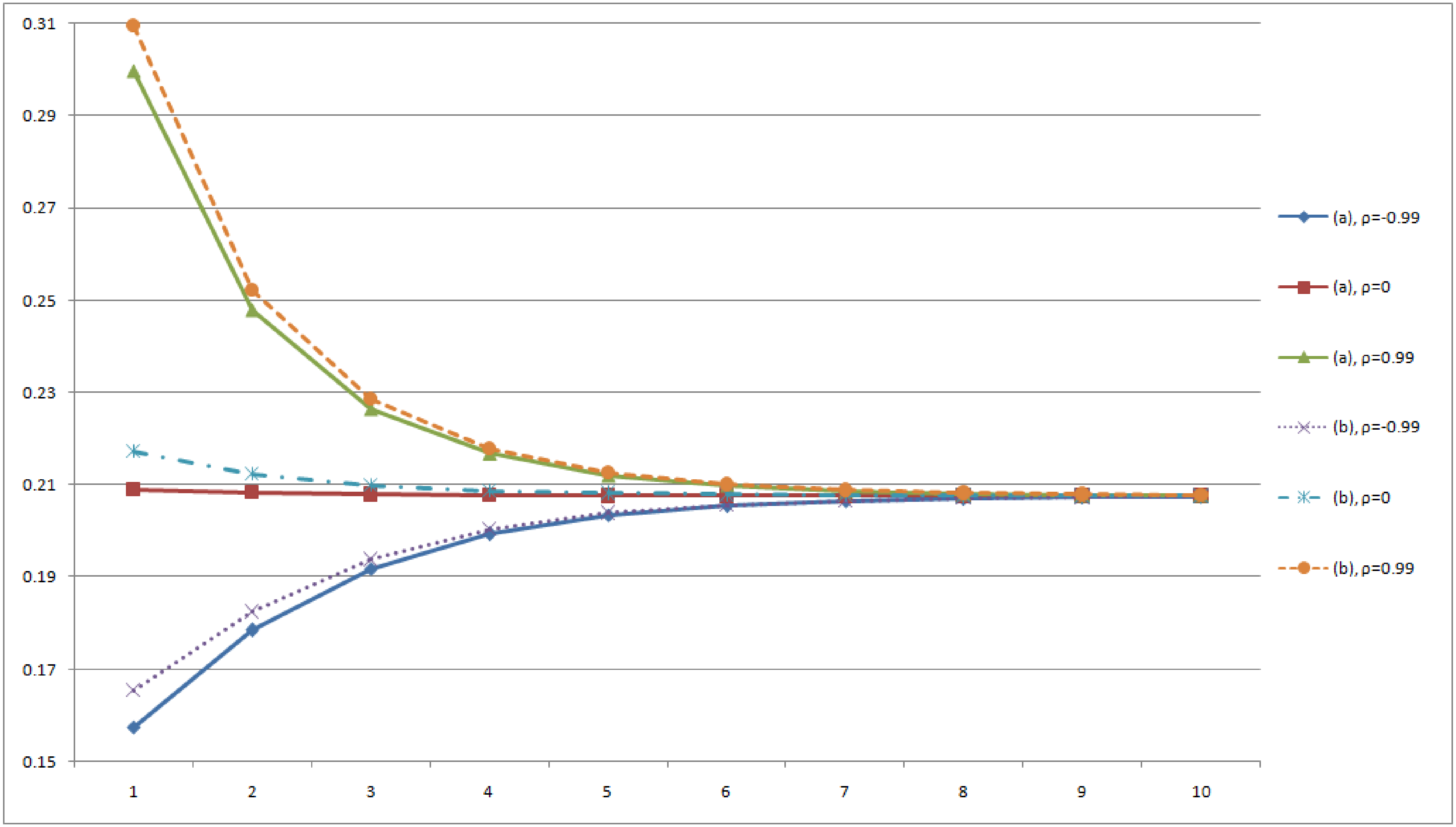
- The right-hand-side equation suggests that it will make a **big difference** if realised variance is measured by log changes squared (i.e. $\sum_{i=1}^N (\log(S(t_i)/S(t_{i-1})))^2$) or by proportional differences squared (i.e. $\sum_{i=1}^N ((S(t_i)/S(t_{i-1}))-1)^2$) **when there are (large) jumps**.

- We now consider some numerical examples.
- We consider variance swaps and proportional variance swaps, with maturity $T = 0.5$, and with N (equally-spaced) monitoring times where $N = 2^{(J-1)}$, for $J = 1, 2, ..., 10$.
- We consider a generalised CGMY process (with a diffusion component) time-changed by a Heston (1993) activity rate (parameters from calibration to the market prices of vanilla options on the S & P500 stock index).
- To see effect of drift and correlation:
- We consider two possible choices, labelled (a) and (b) for the values of the interest-rate $r(t)$ and the dividend yield $q(t)$. In the first choice (a), $r(t) = 0$, $q(t) = 0$, for all t. In the second choice (b), $r(t) = 0.065$, $q(t) = 0.015$, for all t.
- We consider three different combinations for the correlation ρ between the activity rate and the diffusion component of the CGMY process: Namely, $\rho = -0.99$, $\rho = 0$ and $\rho = 0.99$.

Variance swap rates (expressed in volatility terms as a decimal)



Proportional variance swap rates (expressed in volatility terms as a decimal)



Estimating parameters from historical data

- Now lets consider the problem of estimating process parameters from historical data.
- Either assume that we have structure-preserving risk-premia which means we have time-changed Lévy process dynamics under the real-world physical measure $\mathbb{P}$ and under $\mathbb{Q}$ (with, in general, different parameters).
- Or simply regard estimating process parameters as a seperate problem.
- Either way, we assume henceforth time-changed Lévy process dynamics under $\mathbb{P}$.

- Suppose we are given stock prices $S(t_j)$ for times t_j, for $j = 1, 2, ..., N$ where N is large and is of the form $N = LM$, for integers L and M.
- Let us identify L and M as follows: L is the total number of days on which we observe the stock prices and on each day we observe M prices (not necessarily at equal intervals).
- The quadratic variation $QV(\ell)$ of log of the stock price over the period from time $t_{(\ell-1)M}$ to time $t_{\ell M}$ (i.e. on the ℓ^{th} day) is defined as:

$$QV(\ell) \equiv \lim_{\hat{N}\to\infty} \sum_{n=1}^{\hat{N}} (\log(S(u_n)/S(u_{n-1})))^2,$$

for any sequence of partitions $t_{(\ell-1)M} \equiv u_0 < u_1 < u_2 < ... < u_{\hat{N}-1} < u_{\hat{N}} \equiv t_{\ell M}$ with $\sup\{u_n - u_{n-1}\} \to 0$.

- Note that on the ℓ^{th} day, for each $\ell = 1, 2, ..., L$, we can compute the realised variance $\tilde{RV}(\ell, M)$ via:

$$\tilde{RV}(\ell, M) \equiv \sum_{m=1}^{M} (\log(S(t_{(\ell-1)M+m})/S(t_{(\ell-1)M+m-1})))^2.$$

- This (discrete) realised variance $\tilde{RV}(\ell, M)$ is clearly a discrete approximation to the quadratic variation $QV(\ell)$.
- There is a central limit theorem type result (Barndorff-Nielsen and Shephard (2004)) that says that $\tilde{RV}(\ell, M)$, for each $\ell = 1, ..., L$, are (approximately) multi-variate normal provided M is not too small (say, $M \geq 15$).
- Recall M is the number of observations per day of the stock price.

- High-frequency data uses values of M that are large, for example, M equal to 288 (every 5 minutes, 24 hours in a working day - Barndorff-Nielsen and Shephard (2004)) or sampling every 60 seconds ($M = 480$ for 8 hour working day) or every 10 seconds ($M = 2880$ for 8 hour working day) - Barndorff-Nielsen, Hanson, Lunde and Shephard (2008).
- Good point: Large M seems to use more data - therefore better estimates??
- Bad point: Concern that market microstructure effects - eg minimum tick-size, indicative prices or actual transactions prices, illiquidity - distort the estimates if M is too large.

- Existing papers have attempted to estimate the model parameters via maximum likelihood with a log-likelihood function based on multi-variate normal. For this we need to have expressions for the following quantities: $\mathbb{E}^{\mathbb{P}}_{t_0}[\tilde{RV}(\ell, M)]$, $\mathbb{V}ar^{\mathbb{P}}_{t_0}[\tilde{RV}(\ell, M)]$ and $\mathbb{C}ov^{\mathbb{P}}_{t_0}[\tilde{RV}(\ell, M), \tilde{RV}(j, M)]$ for all j and all ℓ.
- In trying to compute these quantities, existing papers seem to make at least one assumption out of the following:
  (1) Assume continuously monitored (ie actually use the expressions for $\mathbb{E}^{\mathbb{P}}_{t_0}[\tilde{RV}(\ell, \infty)]$, etc);
  (2) Ignore drift;
  (3) Assume independence between X_t and Y_t;
  (4) Assume continuous sample paths (i.e. X_t is actually Brownian motion).
- We can compute these quantities without making any of these assumptions.

- Can compute $\mathbb{E}^{\mathbb{P}}_{t_0}[\tilde{RV}(\ell, M)]$ exactly for finite M using equation (1) (the formula for the price of a discretely monitored variance swap).
- Can compute $\mathbb{V}ar^{\mathbb{P}}_{t_0}[\tilde{RV}(\ell, M)]$ via the fourth derivative of the joint extended characteristic function $\Phi(z; j)$.
- Can compute $\mathbb{C}ov^{\mathbb{P}}_{t_0}[\tilde{RV}(\ell, M), \tilde{RV}(j, M)]$ for all j and all ℓ by considering an extended characteristic function of the form $\mathbb{E}^{\mathbb{Q}}_{t_0}[\exp(iz_1 \log \frac{S(t_j)}{S(t_{j-1})} + iz_2 \log \frac{S(t_k)}{S(t_{k-1})})]$
- But how much difference does it make (compared to making the four assumptions on the last slide: (1) Continuously monitored; (2) Ignore drift; (3) Independence between X_t and Y_t; (4) Continuous sample paths)?

- Used same CGMY data as before. Values of the "C" parameters were normalised so that $\mathbb{E}^{\mathbb{P}}_{t_0}[\tilde{RV}(\ell,\infty)] = 0.25$, exactly.
- $\mathbb{E}^{\mathbb{P}}_{t_0}[\tilde{RV}(\ell, M)]$ expressed as an annualised volatility equivalent for different values of M.

$\mathbb{E}^{\mathbb{P}}_{t_0}[\tilde{RV}(\ell, M)]$	M
0.250206	1
0.250103	2
0.250051	4
0.250025	8
0.250012	16
0.250006	32
0.250002	64
0.250001	128

- Answer: It makes little difference. The theoretical value of $\mathbb{E}^{\mathbb{P}}_{t_0}[\tilde{RV}(\ell, M)]$ is very, very insensitive to M.
- We saw that the drift and the correlation between X_t and Y_t make very little difference to the price of a variance swap with, for example, daily monitoring. Its the same story with high-frequency data.
- Conclusion:
  (1) Can assume continuously monitored (ie actually use the expressions for $\mathbb{E}^{\mathbb{P}}_{t_0}[\tilde{RV}(\ell, \infty)]$, etc);
  (2) Can ignore drift;
  (3) Little to be lost (for this estimation method) by assuming independence between X_t and Y_t (because this estimation method cannot produce reliable non-zero estimates).
- If worried about market microstructure effects, one can safely use a smaller value of M (for say $M \geq 15$) - or even better (Ait-Sahalia (2005)), model market microstructure effects explicitly and find the optimal choice of M based on trading off more data against microstructure noise - not based on discrete monitoring effects.

- We **cannot** ignore jumps. We can show approximately

$$\mathbb{V}ar^{\mathbb{P}}_{t_0}[\tilde{RV}(\ell, M)] \approx \frac{-\overline{\psi}''''_X(0)}{365} + (\overline{\psi}''_X(0))^2 \mathbb{V}ar^{\mathbb{P}}_{t_0}[Y_{\ell M+m-1} - Y_{(\ell-1)M+m-1}].$$

- The second term will be very small (eg 10^{-10} or 10^{-11}) for realistic data. The first term (excess kurtosis) would be identically equal to zero for Brownian motion. In practice (based on high-frequency foreign exchange data in Barndorff-Nielsen and Shephard (2004)), the first term is of the order of one million times bigger than the second term. The Barndorff-Nielsen and Shephard data implies a value of $-\overline{\psi}''''_X(0)$ which is of the order of 0.08 to 0.8 (my CGMY data implies a value of 0.092 which is in the right ball-park).

- Generally speaking, discrete monitoring makes little difference to the prices of variance swaps and proportional variance swaps (in the paper, we show more or less the same story for self-quantoed variance swaps, gamma swaps and skewness swaps). This means they are also little affected by the value of Correl(X_t, Y_t).
- Jumps in the underlying dynamics make a lot of difference (there are more examples in the paper) - this is especially true with asymmetric jumps.
- This motivates empirical studies which try to determine how much of the negative skewness seen in stock price returns (under $\mathbb{P}$ and $\mathbb{Q}$) comes from a negatively skewed Lévy process and how much comes from a negative value of Correl(X_t, Y_t) (the maximum likelihood method outlined earlier seems incapable of doing this).
- The paper ("Variance derivatives: Pricing and convergence") on which this talk is based will soon be on my website:
  http://www.john-crosby.co.uk . (or email me - address on website).

- Broadie M. and A. Jain (2008) "The Effect of Jumps and Discrete Sampling on Volatility and Variance Swaps" *International Journal of Theoretical and Applied Finance* Vol.11 No.8. p761-797
- Carr P. and R. Lee (2009) "Pricing variance swaps on time-changed Lévy processes" Working paper
- Derman E., K. Demeterfi, M. Kamal and J. Zou (1999) "More than you ever wanted to know about volatility swaps" *Journal of Derivatives* Vol. 6 No. 4 p9-32 (also on Emanuel Derman's website at http://www.ederman.com)
- Hong G. (2004) "Forward Smile and Derivative Pricing". Unpublished seminar presentation given summer 2004 at Cambridge University. Available on the website of the Centre for Financial Research, Judge Business School, Cambridge University