Optimal Hedging of Variance Derivatives

John Crosby

Presentation at the Bachelier Finance Society World Congress in Toronto, Canada, June 2010

File date 20.07 on 12th June 2010

Acknowledgements p 2/25

- We are indebted to Roger Lee for spotting an error in the first version of the paper on which this talk is based.
- We express our thanks to Peter Carr, Mark Davis, Aleksandar Mijatović and Vimal Raval.



- A number of investment banks are reputed to have lost significant amounts of money on their variance derivatives books in Autumn 2008 as stock indices moved by seven per cent or more in a day.
- This makes very pertinent the question of how to optimally hedge variance derivatives.
- Nearly all papers on variance swaps have focussed on the log-contract replication approach (eg. Neuberger (1990), (1994), (1996), Dupire (1993), Demeterfi et al. (1999), Carr and Lee (2008)).
- This approach works by noting that, under the assumption that the stock price process has continuous sample paths, the payoff of a (continuously monitored) variance swap can be perfectly hedged by a static position of being long 2 log-forward-contracts and by a dynamic position of being short 2/F(t,T) units of forward contracts on the stock, where $F(t,T) \equiv F(t)$ is the forward stock price, at time t, to time T.
- \bullet We will henceforth refer to this approach as the "standard 2+2 log-contract replication" approach.



- Main assumption: Continuous sample paths i.e. no jumps.
- But every empirical study (even before Autumn 2008) shows that stocks and stock indices exhibit jumps in their dynamics and that jumps are necessary to fit implied vol. surfaces, etc.
- The "standard 2 + 2 log-contract replication" approach is often described as model-independent (which is true in some ways), but actually it assumes away that which is empirically most important (i.e. jumps).
- Can we do better? This is the subject of my talk.
- Actually, we can do **much** better and our results have a considerable degree of robustness to model (mis-)specification.



- We define the initial time (today) by $t_0 \equiv 0$ and denote calendar time by $t, t \geq t_0$. Consider a market, which we assume to be free of arbitrage. There is a stock whose forward price, at time t, to time T, is F(t, T). We assume that interest-rates and dividend yields are deterministic and finite.
- The absence of arbitrage guarantees the existence of a risk-neutral equivalent martingale measure. However, as we will utilise Lévy processes, the market is incomplete and, hence, the risk-neutral equivalent martingale measure is not unique. We will assume that one such measure \mathbb{Q} has been fixed on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq t_0}, \mathbb{Q})$. We denote by $\mathbb{E}_t^{\mathbb{Q}}$ the conditional expectation, under \mathbb{Q} , at time t.
- We construct the stock price process by assuming that the log of the stock price is a time-change of (possibly, multiple) Lévy processes.



- We have K independent Lévy processes denoted by $X_t^{(k)}$, for k = 1, 2, ..., K, satisfying $X_{t_0}^{(k)} = 0$. We assume, for each k = 1, 2, ..., K, that we mean-correct $X_t^{(k)}$ so that $\exp(X_t^{(k)})$ is a (non-constant) martingale with respect to the natural filtration generated by $X_t^{(k)}$ i.e. that $\mathbb{E}_{t_0}^{\mathbb{Q}}[\exp(X_t^{(k)})] = \exp(X_{t_0}^{(k)}) = 1$ for all $t \ge t_0$.
- Lévy-Khinchin formula, for each k, implies we can write the (mean-corrected) characteristic exponent $\overline{\psi}_X^{(k)}(z)$ in the form:

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

For future reference, we define, for each k, the deterministic quantity:

$$m_X^{(k)}(iz) \equiv i \overline{\psi}_X^{(k)'}(z),$$

where ' denotes differentiation i.e. $\overline{\psi}_X^{(k)'}(z) \equiv \partial \overline{\psi}_X^{(k)}(z) / \partial z$, $\overline{\psi}_X^{(k)''}(z) \equiv \partial^2 \overline{\psi}_X^{(k)}(z) / \partial z^2$, and further, for $n \geq 3$, $\overline{\psi}_X^{(k),(n)}(z) \equiv \partial^n \overline{\psi}_X^{(k)}(z) / \partial z^n$.



- We assume that we have K (possibly, deterministic) non-decreasing, continuous time-change processes denoted by $Y_t^{(k)}$, for each k = 1, 2, ..., K. We normalise so that $Y_{t_0}^{(k)} = t_0 \equiv 0$.
- In general, $Y_t^{(k)}$ may be correlated with $X_t^{(\ell)}$, for any $\ell = 1, 2, \ldots, K$.
- Our assumption, for example, allows $Y_t^{(k)}$ to be of the form $Y_t^{(k)} = \int_{t_0}^t y_s^{(k)} ds$ where the activity rate $y_t^{(k)}$ (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, $y_t^{(k)}$ is discontinuous but $Y_t^{(k)}$ is always continuous.
- We will sometimes refer to an important special case when the time-change processes are "common" i.e. $Y_t^{(k)} = Y_t$, say, for all k.

Stock price process p 8/25

- For each k = 1, 2, ..., K, we time-change the Lévy process $X_t^{(k)}$ by $Y_t^{(k)}$ to get a process $X_{Y_t^{(k)}}^{(k)}$ which we henceforth denote by $X_{Y_t}^{(k)}$, with $X_{Y_{t_0}}^{(k)} = 0$.
- The forward stock price F(t, T), at time t, to time T, is assumed to have the following dynamics:

$$F(t,T) = F(t_0,T) \exp(\sum_{k=1}^{K} X_{Y_t}^{(k)}).$$

Note that F(t,T) is a martingale, under \mathbb{Q} , in the enlarged filtration generated by $\{X_t^{(1)} \cup X_t^{(2)} \cup \ldots \cup X_t^{(K)} \cup Y_t^{(1)} \cup Y_t^{(2)} \cup \ldots \cup Y_t^{(K)}\}.$



- We have already seen that with continuous sample paths, the price of a continuously monitored variance swap equals minus two times the price of a log-forward-contract.
- In general, i.e. with jumps, the price $VS(t_0, T)$, at time t_0 , of (the floating leg of) a continuously monitored variance swap maturing at time T equals $-Q_X$ times the price $LFC(t_0, T)$, at time t_0 , of a log-forward-contract paying $\log(F(T, T)/F(t_0, T))$ at time T, where

$$-Q_X \equiv \frac{VS(t_0, T)}{LFC(t_0, T)}$$

$$\equiv \frac{\sum_{k=1}^{K} \overline{\psi}_X^{(k)''}(0) \mathbb{E}_{t_0}^{\mathbb{Q}}[Y_T^{(k)} - Y_{t_0}^{(k)}]}{\sum_{k=1}^{K} m_X^{(k)}(0) \mathbb{E}_{t_0}^{\mathbb{Q}}[Y_T^{(k)} - Y_{t_0}^{(k)}]}.$$
 Note that $Q_X > 0$, since $m_X^{(k)} < 0$, for all k .

Proof: Peter Carr's talk in 25 minutes time, Carr and Lee (2009), Crosby et al. (2010).

- In particular, there is no up-front cost of entering into a position of being long the floating leg of one variance swap and being long Q_X log-forward-contracts.
- When the time-change processes are "common" i.e. $Y_t^{(k)} = Y_t$, say, for all k, then Q_X does not depend on Y_t in any way.

Self-financing trading strategy

- Basic idea: We construct a self-financing trading strategy as follows: We commence the strategy at time $t_0 \equiv 0$. At each time $t \in [t_0, T]$, we hold a long position in one variance swap and in Θ_t^{LFC} log-forward-contracts. Additionally, we trade dynamically in the underlying stock. Specifically, for all $t \in [t_0, T]$, we hold a short position in $\Delta_t \equiv \phi_t/F(t-, T)$ units of forward contracts on the stock.
- We compute the variance, under \mathbb{Q} , of the time T P+L of the self-financing trading strategy i.e. the variance of the residual hedging error.
- It is a non-negative quadratic function of Θ_t^{LFC} and ϕ_t . Minimise by differentiating w.r.t. portfolio weight and setting the resulting equation to zero.
- Can do this analytically (does not need Monte Carlo see paper for full details)

$p \ 10/25$

Self-financing trading strategy 2 p 11/25

- For simplicity, I'll just write the equations with deterministic time-changes.
- The P+L of the self-financing strategy, at time T, is:

$$\begin{split} \epsilon(T) &\equiv \int_{t_0}^T \sum_{k=1}^K y_u^{(k)} \int_{-\infty}^\infty x^2 \left(\mu^{(k)}(dx) - \nu^{(k)}(dx) \right) du \\ &+ \int_{t_0}^T \Theta_u^{\text{LFC}} \sum_{k=1}^K y_u^{(k)} \left(\sigma^{(k)} dW_u^{(k)} + \int_{-\infty}^\infty x \left(\mu^{(k)}(dx) - \nu^{(k)}(dx) \right) du \right) \\ &- \int_{t_0}^T \Delta_u F(u-,T) \sum_{k=1}^K y_u^{(k)} \left(\sigma^{(k)} dW_u^{(k)} + \int_{-\infty}^\infty (\exp(x) - 1) \left(\mu^{(k)}(dx) - \nu^{(k)}(dx) \right) du \right) \\ &= \int_{t_0}^T \sum_{k=1}^K y_u^{(k)} \int_{-\infty}^\infty \left(x^2 + \Theta_u^{\text{LFC}} x - \phi_u(\exp(x) - 1) \right) \left(\mu^{(k)}(dx) - \nu^{(k)}(dx) \right) du \\ &+ \int_{t_0}^T \left(\Theta_u^{\text{LFC}} - \phi_u \right) \sum_{k=1}^K y_u^{(k)} \sigma^{(k)} dW_u^{(k)}, \text{ using } \phi_t \equiv \Delta_t F(t-,T). \end{split}$$

Variance of residual hedging error	$p \ 12/25$
------------------------------------	-------------

• From Ito's isometry formula, the variance, under \mathbb{Q} , of the time T P+L of the self-financing strategy is:

$$\begin{split} Var_{t_0}^{\mathbb{Q}}[\epsilon(T)] &\equiv \sum_{k=1}^{K} \mathbb{E}_{t_0}^{\mathbb{Q}}[\Big(\int_{t_0}^{T} y_u^{(k)} (\Theta_u^{\text{LFC}} - \phi_u)^2 \sigma^{(k)\,2} du\Big) \\ &+ \Big(\int_{t_0}^{T} y_u^{(k)} \int_{-\infty}^{\infty} \left(x^2 + \Theta_u^{\text{LFC}} x - (\phi_u(\exp(x) - 1))\right)^2 \nu^{(k)}(dx) du\Big)] \\ &= \sum_{k=1}^{K} \mathbb{E}_{t_0}^{\mathbb{Q}}[\int_{t_0}^{T} y_u^{(k)} \Big(\phi_u^2(-\overline{\psi}_X^{(k)}(-2i)) \\ &- 2\phi_u \Big(\Theta_u^{\text{LFC}}(m_X^{(k)}(1) - m_X^{(k)}(0)) + (\overline{\psi}_X^{(k)''}(-i) - \overline{\psi}_X^{(k)''}(0))\Big) \\ &- \overline{\psi}_X^{(k),(4)}(0) - 2\Theta_u^{\text{LFC}} i \overline{\psi}_X^{(k),(3)}(0) + \Theta_u^{\text{LFC}\,2} \overline{\psi}_X^{(k)''}(0)\Big) du]. \end{split}$$

This is a non-negative quadratic function of ϕ_u and Θ_u^{LFC} . We minimise by differentiating with respect to ϕ_u and Θ_u^{LFC} and setting to zero.

Hedging strategies p 13/25

- We hedge a (static) long position in one variance swap.
- We consider two types of hedging strategy labelled A and B.
- The first type of hedging strategy (labelled hedging strategy A) consists of a static position in $\Theta_t^{\text{LFC}} = Q_X$ log-forward-contracts and a dynamic short position in $\Delta_t \equiv \phi_t/F(t-,T)$ forward contracts on the underlying stock. The static position Q_X is motivated by slide "Important result" (but is not necessarily optimal).
- The second type of hedging strategy (labelled hedging strategy B) consists of a, possibly, dynamic position in Θ_t^{LFC} log-forward-contracts and a dynamic short position in $\Delta_t \equiv \phi_t/F(t-,T)$ forward contracts on the underlying stock.

Hedging strategy A p 14/25

• For hedging strategy A, we have $\Theta_t^{\text{LFC}} = Q_X$ (by design) and we find that the optimal value $\hat{\Delta}_t$ which minimises the variance is:

$$\hat{\Delta}_{t} \equiv \frac{\hat{\phi}_{t}}{F(t-,T)}, \text{ where}$$

$$\hat{\phi}_{t} = \frac{\sum_{k=1}^{K} y_{t}^{(k)}(\overline{\psi}_{X}^{(k)''}(-i) - \overline{\psi}_{X}^{(k)''}(0)) + Q_{X} \sum_{k=1}^{K} y_{t}^{(k)}(m_{X}^{(k)}(1) - m_{X}^{(k)}(0))}{-\sum_{k=1}^{K} y_{t}^{(k)} \overline{\psi}_{X}^{(k)}(-2i)}$$

• Sanity check: For Brownian motion with volatility σ (wlog), $\overline{\psi}_X^{(k)}(z) = \frac{1}{2}\sigma^2(z^2 + iz)$, $m_X^{(k)}(0) = -\frac{1}{2}\sigma^2$, $\overline{\psi}_X^{(k)''}(0) = \sigma^2$. Implies: $Q_X = -\frac{\sigma^2}{-\frac{1}{2}\sigma^2} = 2$.

• Further, $\overline{\psi}_X^{(k)}(-2i) = -\sigma^2$, which implies: $\hat{\phi}_t = 2$.

- This agrees with standard results i.e. the standard 2 + 2 log-contract replication approach naturally appears as a special case of our analysis. Further, for this special case, substituting back in, the variance $Var_{t_0}^{\mathbb{Q}}[\epsilon(T)]$ is identically equal to zero.
- \Rightarrow perfect hedge.

Hedging strategy A: Special case of a compound Poisson process

- For hedging strategy A, we also consider the special case of a compound Poisson process with a fixed jump amplitude a_1 (with no diffusion component). Substituting in the relevant characteristic function, we find:
- $Q_X = a_1^2 / (\exp(a_1) 1 a_1).$
- $\hat{\phi}_t = a_1^2 / (\exp(a_1) 1 a_1).$
- Further, for this special case, substituting back in, the variance is identically equal to zero.
- \Rightarrow perfect hedge.
- In the limit that $a_1 \rightarrow 0$, we find:

$$\hat{\phi}_t = Q_X = \frac{a_1^2}{(\exp(a_1) - 1 - a_1)} \approx \frac{2}{(1 + (a_1/3))}.$$

We see that when a_1 is small but positive, $\hat{\phi}_t = Q_X$ is just below two and when a_1 is small but negative, $\hat{\phi}_t = Q_X$ is just above two. In either case, as $a_1 \to 0$, $\phi_t \to 2$, which is the same as the case of Brownian motion.

Hedging strategy B p 16/25

- For hedging strategy B, we optimise over Θ_t^{LFC} (the position in log-forward-contracts) and over ϕ_t (recall $\Delta_t \equiv \phi_t / F(t-,T)$ is the position in forward contracts on the underlying stock).
- It turns out that, in the special case that the time-changes are "common" (ie the same for all Lévy processes in the sense that $Y_t^{(k)} = Y_t$, say, for all k, which must be true if K = 1), then the position in log-forward-contracts is constant in time i.e. it is a static buy-and-hold position (which is important as dynamic positions would incur significant transactions costs).
- Further, in this special case, ϕ_t and Θ_t^{LFC} do not depend upon the time-change process in any way \Rightarrow considerable degree of robustness to model (mis-)specification.

Numerical examples p 17/25

- We now consider some numerical examples which compare three possible hedging strategies.
- The first hedging strategy is the standard 2 + 2 log-contract replication approach (sets $\phi_t = 2$, $\Theta_t^{\text{LFC}} = 2$).
- The second and third are hedging strategies A and B respectively which we described earlier.
- In the examples, we always work with a "common" time-change. Hence, we have constant values of ϕ_t , Θ_t^{LFC} (to repeat, this has the additional benefit of robustness to transactions costs).
- We consider the hedging of a long position in one variance swap with maturity T = 0.5.
- We consider two sets of numerical results each with six different sets of process parameters. The first uses combinations of a Brownian motion and upto three compound Poisson processes with fixed jump amplitudes together with a deterministic time-change. The second uses stochastically time-changed CGMY processes (there are more results in the paper).



• Table 1.

We consider six combinations (labelled params 1 to params 6) of a Brownian motion and upto three compound Poisson processes, with intensity rates λ_1 , λ_2 and λ_3 and with fixed jump amplitudes a_1 , a_2 and a_3 . We assume a common, deterministic time-change (not necessarily of the form $Y_t^{(k)} = t$).

								Skewness	
	λ_1	a_1	λ_2	a_2	λ_3	a_3	Vol	swap price	Q_X
params 1	1.00000000	-0.2	0	0	0	0	0.15	-0.00400	2.0846708
params 2	1.53186275	-0.2	0.76593137	0.04	0	0	0	-0.00610	2.1320914
params 3	0.98039216	-0.2	0.49019608	0.04	0	0	0.15	-0.00391	2.0825752
params 4	1.50240385	-0.2	0.75120192	0.04	0.75120192	-0.04	0	-0.00601	2.1299626
params 5	0.96153846	-0.2	0.48076923	0.04	0.48076923	-0.04	0.15	-0.00385	2.0812748
params 6	0.54086538	-0.2	0.27043269	0.04	0.27043269	-0.04	0.2	-0.00216	2.0449185

For all parameters, the (annualised) variance swap rate expressed as a volatility is 0.25. All values of $Var_{t_0}^{\mathbb{Q}}[\epsilon(T)]$ in the table below are multiplied by 1,000,000 to improve readability.



Notice we get perfect hedges in some special cases.

Discussion of results 1

• With combinations of Brownian motions and compound Poisson processes with fixed jump amplitudes, as we increase the number of hedging instruments over which we optimise (underlying, log-forward-contracts), we increase from 1 to 2 the number of underlying stochastic processes that we can perfectly hedge against. This is highly intuitive.

p 20/25

• For example, for hedging strategy B (two instruments, i.e. underlying and log-forward-contracts), we can perfectly hedge when there are two stochastic processes (one Poisson + Brownian motion or two Poisson).



• Table 2.

We consider six combinations (labelled params 1 to params 6) of a generalised CGMY process time-changed by either a Heston (1993) activity-rate process (params 1 to 5) or a non-Gaussian OU process (the Gamma-OU process of Barndorff-Nielsen and Shephard (2001)) (params 6).

All parameters were obtained from calibrations to market prices of vanilla options on S & P 500 and are quoted from Schoutens (2003) and from Carr, Geman, Madan and Yor (2003).

Skewness

	C_{Up}	C_{Down}	G	M	$Y_{ m Up}$	$Y_{ m Down}$	Vol	swap price	Q_X
params 1	0.00740000	0.00740000	0.1025	11.394	1.6765	1.6765	0	-0.06977	2.7294158
params 2	0.16350000	0.04713705	0.6965	21.97	-3.65	1.45	0	-0.01272	2.4274086
params 3	0.35870000	0.01886762	0.4231	24.64	-4.51	1.67	0	-0.01419	2.3727413
params 4	0.40410000	0.02731716	1.64	16.91	-2.9	1.54	0	-0.00385	2.1675632
params 5	2.04400000	0.17476200	3.68	52.86	-2.12	1.22	0	-0.01054	2.1349535
params 6	0.04150000	0.04150000	3.9134	30.6322	1.3664	1.3664	0	-0.00182	2.0769284



• Table 2 continued.

The activity rate for params 1 to params 5 follows a Heston (1993) process of the form:

$dy_t = \kappa(\eta -$	- $y_t)dt$ +	$+\lambda y_t^{1/2}dz_t,$	$y_{t_0} \equiv y_0,$	with $y_0 > 0$.	
,	- >	_			

	Var swap rate (as vol)	λ	κ	η	y_0	ρ
params 1	0.232270	1.3612	0.3881	1.4012	1	0
params 2	0.179512	0.00022	8.51	0.1497	1	0
params 3	0.190740	0.0006	6.65	0.3469	1	0
params 4	0.165670	2.78E-05	4.85	0.4474	1	0
params 5	0.315297	1.7	15.91	1.3700	1	0
params 6	0.172255	0.8826	a = 0.5945	b = 0.8524	1	0

All values of $Var_{t_0}^{\mathbb{Q}}[\epsilon(T)]$ in the table below are multiplied by 100 to improve readability.



Discussion of results 2 p 24/25

- When $Q_X \gg 2$ (which implies that the \mathbb{Q} distribution of stock returns is negatively skewed which is empirically the case for equities), then $\hat{\phi}_t \gg 2$ and $\hat{\Theta}_t^{\text{LFC}} \gg 2$.
- For parameters (params 1) obtained from a calibration to market prices of options on S & P 500, the optimal hedges are **five times** greater than those implicit in the standard 2 + 2 log-contract replication approach.
- Hedging strategy B always outperforms hedging strategy A which, in turn, always outperforms the standard 2 + 2 log-contract replication approach.
- Our results show that substantial reduction in residual hedging error is possible by optimal choice of the position in log-forward-contracts and the position in forward contracts on the underlying stock.
- In the paper, we show further that further substantial reductions are possible through the use of skewness swaps.



- The standard 2 + 2 log-contract replication approach is very far from optimal in the presence of jumps.
- The good news: We can construct optimal hedges for hedging a long position in one variance swap which are of the form long $\hat{\Theta}_t^{\text{LFC}}$ log-forward-contracts and short $\hat{\phi}_t/F(t-,T)$ forward contracts on the underlying stock.
- The bad news: $\hat{\Theta}_t^{\text{LFC}}$ and $\hat{\phi}_t$ are not 2 (2 is the "small jump limit").
- For a wide class of processes (but not always), $\hat{\Theta}_t^{\text{LFC}}$ and $\hat{\phi}_t$ are independent of the time-change (\Rightarrow robust to model (mis-)specification) and constant in time (\Rightarrow robust to transactions costs) but they are highly dependent upon the skew of the Lévy process(es).
- The paper on which this talk is based ("Optimal hedging of variance derivatives") can be found on my website: http://www.john-crosby.co.uk .