Beyond Black-Scholes: Option Pricing with Time-Changed Lévy Processes

Liuren Wu

Zicklin School of Business, Baruch College

Options Markets
Overview

Key advantages of modeling returns with time-changed Lévy processes:

- **Generality:**
  - Lévy processes can generate almost any return innovation distributions.
  - Applying stochastic time changes on Lévy processes randomizes the innovation distribution over time ⇒ stochastic volatility, correlation, skewness, ....

- **Explicit economic mapping:**
  - Each Lévy component ↔ shocks from one economic source.
  - Time change captures the time-varying intensity of its impact.

⇒ makes model design more intuitive, parsimonious, and sensible.

- **Tractability:** A model is tractable for option pricing if we have
  - tractable characteristic exponent for the Lévy components.
  - tractable Laplace transform for the activity rates.

⇒ any combinations of the two generate tractable return dynamics.
Lévy processes

- A Lévy process is a continuous-time process that generates stationary, independent increments ...

- Think of return innovations \((\varepsilon)\) in discrete time: \(R_{t+1} = \mu_t + \sigma_t \varepsilon_{t+1}\).
  - Normal return innovation — diffusion
  - Non-normal return innovation — jumps

- Classic Lévy specifications in finance:
  - Brownian motion (Black-Scholes, Merton)
  - Compound Poisson process with normal jump size (Merton)

  \(\Rightarrow\) The return innovation distribution is either normal or mixture of normals.
Lévy characteristics

- Lévy processes greatly expand our continuous-time choices of iid return innovation distributions via the Lévy triplet \((\mu, \sigma, \pi(x))\). (\(\pi(x)\)–Lévy density).

- The Lévy-Khintchine Theorem:

\[
\begin{align*}
\phi_{X_t}(u) & \equiv \mathbb{E}[e^{iuX_t}] = e^{-t\psi(u)}, \\
\psi(u) & = -iu\mu + \frac{1}{2}u^2\sigma^2 + \int_{\mathbb{R}_0} \left(1 - e^{iux} + iux1_{|x|<1}\right) \pi(x)dx,
\end{align*}
\]

Innovation distribution

\(\leftrightarrow\) characteristic exponent \(\psi(u)\)

\(\leftrightarrow\) Lévy triplet \((\mu, \sigma, \pi(x))\)

- Constraint: \(\int_0^1 x^2\pi(x)dx < \infty\).
- “Tractable:” if the integral can be carried out explicitly.
Tractable examples

- Brownian motion \((\mu t + \sigma W_t)\): normal shocks.
- Compound Poisson jumps: Large but rare events.

\[
\pi(x) = \lambda \frac{1}{\sqrt{2\pi v_J}} \exp \left( -\frac{(x - \mu_J)^2}{2v_J} \right).
\]

- Dampened power law (DPL):

\[
\pi(x) = \begin{cases} 
\lambda \exp (-\beta_+ x) x^{-\alpha-1}, & x > 0, \\
\lambda \exp (-\beta_- |x|) |x|^{-\alpha-1}, & x < 0,
\end{cases} \quad \lambda, \beta_\pm > 0, \quad \alpha \in [-1, 2)
\]

- **Finite activity** when \(\alpha < 0\): \(\int_{\mathbb{R}} \pi(x) dx < \infty\). Compound Poisson. Large and rare events.
- **Infinite activity** when \(\alpha \geq 0\): Both small and large jumps. Jump frequency increases with declining jump size, and approaches infinity as \(x \to 0\).
- **Infinite variation** when \(\alpha \geq 1\): many small jumps.

*Market movements of all magnitudes, from small movements to market crashes.*
Analytical characteristic exponents

- **Diffusion**: \( \psi(u) = -iu\mu + \frac{1}{2}u^2\sigma^2. \)

- **Merton’s compound Poisson jumps**: 
  \[
  \psi(u) = \lambda \left( 1 - e^{iu\mu J - \frac{1}{2}u^2V_J} \right).
  \]

- **Dampened power law**: (for \( \alpha \neq 0, 1 \))
  \[
  \psi(u) = -\lambda \Gamma(-\alpha) \left[ (\beta_+ - iu)^\alpha - \beta_+^\alpha + (\beta_- + iu)^\alpha - \beta_-^\alpha \right] - iuC(h)
  \]

  - When \( \alpha \to 2 \), smooth transition to diffusion (quadratic function of \( u \)).
Other Lévy examples

- Other examples:
  - The normal inverse Gaussian (NIG) process of Barndorff-Nielsen (1998)
  - The generalized hyperbolic process (Eberlein, Keller, Prause (1998))
  - The Meixner process (Schoutens (2003))
  - ...

- Bottom line:
  - All tractable in terms of analytical characteristic exponents $\psi(u)$.
  - We can use FFT to generate the density function of the innovation (for model estimation).
  - We can also use FFT to compute option values.
Clark (1973): If one runs a Brownian motion on a business clock, the resulting process matches financial time series better.

The possibility that business clock may not move while calendar time marches forward is important ...

- If the clock is a standard Poisson process
  ⇒ The resulting process is a compound Poisson process with normal jump sizes.
- If the clock is a compound Poisson process with exponentially distributed jump size
  ⇒ DPL with $\alpha = -1$
- If the clock is a gamma process
  ⇒ DPL with $\alpha = 0$.
- If the clock is continuous
  ⇒ a continuous process.
General evidence on Lévy return innovations

- **Credit risk:** *(compound)* Poisson process
  - The whole intensity-based credit modeling literature...

- **Market risk:** Infinite-activity jumps
  - Evidence from stock returns (CGMY (2002)): The $\alpha$ estimates for DPL on most stock return series are greater than zero.
  - Evidence from options: Models with infinite-activity return innovations price equity index options better (Carr and Wu (2003), Huang and Wu (2004))
Implied volatility smiles & skews on a stock

Moneyness = $\ln(\frac{K}{F})$

$\sigma \sqrt{\tau}$

Implied Volatility

Short-term smile

Long-term skew

Maturities: 32, 95, 186, 368, 732

Liuren Wu

Time-Changed Lévy Processes

Options Markets
Implied volatility skews on a stock index (SPX)

More skews than smiles

Maturities: 32  60  151  242  333  704

Liuren Wu
Time-Changed Lévy Processes
Options Markets
Average implied volatility smiles on currencies

Maturities: 1m (solid), 3m (dashed), 1y (dash-dotted)
The role of jumps at very short maturities

- Implied volatility smiles (skews) ↔ non-normality (asymmetry) for the risk-neutral return distribution.

\[ IV(d) \approx ATMV \left( 1 + \frac{\text{Skew.}}{6} d + \frac{\text{Kurt.}}{24} d^2 \right), \quad d = \frac{\ln K/F}{\sigma \sqrt{\tau}} \]

- Two mechanisms to generate return non-normality:
  - Use Lévy jumps to generate non-normality for the innovation distribution.
  - Use stochastic volatility to generate non-normality through mixing over multiple periods.

- Over very short maturities (1 period), *only jumps contribute to return non-normalities.*
Time decay of short-term OTM options

- As option maturity ↓ zero, OTM option value ↓ zero.
- The speed of decay is exponential $O(e^{-c/T})$ under pure diffusion, but linear $O(T)$ in the presence of jumps.
- Term decay plot: $\ln(\text{OTM}/T) \sim \ln(T)$ at fixed $K$:

  In the presence of jumps, the Black-Scholes implied volatility for OTM options $\text{IV}(\tau, K)$ explodes as $\tau \downarrow 0$. 
(II) The impacts of jumps at very long horizons

- Central limit theorem (CLT): As option maturity increases, the smile should flatten.
- Evidence: The skew does not flatten, but steepens!
- FMLS: Maximum negatively skewed $\alpha$-stable process.
  - Return variance is infinite. Hence, CLT does not apply.
  - Down jumps only. $\Rightarrow$ Option has finite value.
- *But CLT seems to hold fine statistically:*

![Skewness on S&P 500 Index Return](image1)

![Kurtosis on S&P 500 Index Return](image2)
Reconcile $\mathbb{P}$ with $\mathbb{Q}$ via DPL jumps

- Model return innovations under $\mathbb{P}$ by DPL:

$$
\pi(x) = \begin{cases} 
\lambda \exp(-\beta_+ x) x^{-\alpha-1}, & x > 0, \\
\lambda \exp(-\beta_- |x|) |x|^{-\alpha-1}, & x < 0.
\end{cases}
$$

All return moments are finite with $\beta_\pm > 0$. **CLT applies.**

- Market price of jump risk ($\gamma$): \[ \frac{d\mathbb{Q}}{d\mathbb{P}} \bigg|_t = \mathcal{E}(-\gamma X) \]

- The return innovation process remains DPL under $\mathbb{Q}$:

$$
\pi(x) = \begin{cases} 
\lambda \exp(-(\beta_+ + \gamma) x) x^{-\alpha-1}, & x > 0, \\
\lambda \exp(-(\beta_- - \gamma) |x|) |x|^{-\alpha-1}, & x < 0.
\end{cases}
$$

- To break CLT under $\mathbb{Q}$, set $\gamma = \beta_-$ so that $\beta^-_Q = 0$.

- Reconciling $\mathbb{P}$ with $\mathbb{Q}$: *Investors charge maximum allowed market price on down jumps.*
(III) Default risk & long-term implied vol skew

- When a company defaults, its stock value jumps to zero.
- It generates a steep skew in long-term stock options.
- Evidence: Stock option implied volatility skews are correlated with credit default swap (CDS) spreads written on the same company.
Three Lévy jump components

I. Market risk (FMLS under $\mathbb{Q}$, DPL under $\mathbb{P}$)

II. Idiosyncratic risk (DPL under both $\mathbb{P}$ and $\mathbb{Q}$)

III. Default risk (Compound Poisson jumps).

Remarks:

- Identify market risk from SPX or QQQQ options.
- Identify the credit risk component from the credit default swap (CDS) market.

Currency options:

- Model currency returns as the difference of two log pricing kernels (market risks).
- Default risk also shows up in FX for low-rating economies.
Beyond Lévy processes

- Lévy processes can be used to generate different iid return innovation distributions.

- Yet, return distribution is not iid. It varies stochastically over time.

- We need to go beyond Lévy processes to capture the stochastic nature of the return distribution.

- Applying separate stochastic time changes to different Lévy components generates
  - separate stochastic responses to each economic shock.
  - stochastic volatility, skewness, ...
Capturing stochastic volatility via time changes

- Discrete-time analog again: $R_{t+1} = \mu_t + \sigma_t \varepsilon_{t+1}$
  
  - $\varepsilon_{t+1}$ is an iid return innovation, with an arbitrary distribution assumption $\leftrightarrow$ Lévy process.
  
  - $\sigma_t$ is the conditional volatility, $\mu_t$ is the conditional mean return, both of which can be time-varying, stochastic...
  
  - If the return innovation is modeled by a general Lévy process, it is tractable to randomize the time, or something proportional to time.

Variance of a Brownian motion, intensity of a Poisson process are both proportional to time.
Randomize the time

- The drift $\mu$, the diffusion variance $\sigma^2$, and the Poisson arrival rate $\lambda$ are all proportional to time $t$.

- We may as well randomize time $t \rightarrow T_t$ instead of $(\mu, \sigma^2, \lambda)$, for the same result.

- We define $T_t \equiv \int_0^t \nu_s \, ds$ as the (stochastic) time change, with $\nu_t$ being the instantaneous activity rate.

  - Depending on the Lévy specification, it has the same meaning (up to a scale) as a randomized version of the instantaneous drift, instantaneous variance, or instantaneous arrival rate.
Applying separate time changes

... to different Lévy components

- Consider a Lévy process $X_t \sim (\mu, \sigma^2, \lambda p(x))$.
  - If we apply random time change to $X_t \rightarrow X_{T_t}$ with $T_t = \int_0^t \nu_s ds$, it is equivalent to assuming that $(\mu_t, \sigma_t^2, \lambda_t)$ are all time varying, but they are all proportional to one common source of variation $\nu_t$.
  - Suppose we want $(\mu_t, \sigma_t^2, \lambda_t)$ to vary separately, then we need to apply separate time changes to the three Lévy components.
    - Decompose $X_t$ into three Lévy processes: $X_t^1 \sim (\mu, 0, 0)$, $X_t^2 \sim (0, \sigma^2, 0)$, and $X_t^3 \sim (0, 0, \lambda p(x))$, and then apply separate time changes to the three Lévy processes.

- In practice, we can use one Lévy process to model one source of economic shock, and use separate time changes on different Lévy processes to capture the intensity variation of different economic shocks.
Example: Return on a stock

- Model the return on a stock to reflect shocks from two sources:
  - **Credit risk**: In case of corporate default, the stock price falls to zero. Model the impact as a Poisson Lévy jump process with log return jumps to negative infinity upon jump arrival.
  - **Market risk**: Daily market movements (small or large). Model the impact as a diffusion or infinite-activity (infinite variation) Lévy jump process or both.

- Apply separate time changes to the two Lévy components to capture (1) the intensity variation of corporate default, (2) the market risk (volatility) variation.

- Key: *Each component has a specific economic purpose.*

Example: Return on an exchange rate

- Exchange rate reflects the interaction between two economic forces.
- Use two Lévy processes to model the two economic forces separately.
- Consider a negatively skewed distribution (downside jumps) from each economic source (crash-o-phobia from both sides). Use the difference to model the currency return between the two economies.
- Apply separate time changes to the two Lévy processes to capture the strength variation of the two economic forces.
  - Stochastic time changes on the two negatively skewed Lévy processes generate both stochastic volatility and stochastic skew.
- Key: *Each component has its specific economic purpose.*

Carr and Wu (JFE, forthcoming), “Stochastic Skew in Currency Options”
Exchange rates and pricing kernels

- Exchange rate reflects the interaction between two economic forces.
- The economic meaning becomes clearer if we model the pricing kernel of each economy.

  - Let $m_{0,t}^{US}$ and $m_{0,t}^{JP}$ denote the pricing kernels of the US and Japan. Then the dollar price of yen $S_t$ is given by

    $$\ln \frac{S_t}{S_0} = \ln m_{0,t}^{JP} - \ln m_{0,t}^{US}.$$  

  - If we model the negative of the logarithm of each pricing kernel ($-\ln m_{0,t}^j$) as a time-changed Levy process, $X_{T_t}^j$ ($j = US, JP$) with negative skewness. Then, $\ln \frac{S_t}{S_0} = \ln m_{0,t}^{JP} - \ln m_{0,t}^{US} = X_{T_t}^{US} - X_{T_t}^{JP}$.

  - Consistent and simultaneous modeling of all currency pairs (not limited to 2 economies).

Model assembly

- Start with the risk-neutral (\( \mathbb{Q} \)) process — That’s where tractability is needed the most dearly.
  - Identify the economic sources (\( X^k_t \) for \( k = 1, \cdots, K \))
  - Decide whether to apply separate time changes: \( X^k_t \rightarrow X^k_{T_t} \)
  - Adjust to guarantee the martingale condition: \( \mathbb{E}^\mathbb{Q}[S_t/S_0] = e^{(r-q)t} \).

\[
\ln S_t/S_0 = (r - q)t + \sum_{k=1}^{K} \left( b^k X^k_{T_t} - \varphi_x^k (b^k) T^k_t \right),
\]

- \( \mathbb{E}^\mathbb{Q}[e^{bX_t}] = e^{\varphi(b)t} \). Hence, \( \mathbb{E}^\mathbb{Q}[e^{bX_t - \varphi(b)t}] = 1, \mathbb{E}^\mathbb{Q}[e^{bX_{T_t} - \varphi(b)T_t}] = 1 \).

- Example: A CAPM model:
\[
\ln S^j_t/S^j_0 = (r - q)t + \left( \beta^j X^m_{T_t} - \varphi_x^m (\beta^j) T^m_t \right) + \left( X^j_{T_t} - \varphi_x^j (1) T^j_t \right).
\]
  - Estimate \( \beta \) and market prices of return and volatility risk using index and single name options.
  - Cross-sectional analysis of the estimates.
Market prices and statistics dynamics

- Since we can always use Euler approximation for model estimation, tractability requirement is not as strong for the statistical dynamics.

- We can specify pretty much any forms for the market prices subject to (i) technical conditions, (ii) economic sensibility, and (iii) identification concerns.

- Simple/parsimonious specification: *Constant* market prices of return and vol risks \((\gamma_k, \gamma_{kv})\)

\[
\mathcal{M}_t = e^{-rt} \prod_{k=1}^{K} \exp \left( -\frac{\gamma_k X_{T_t}^k}{\gamma_k} + \varphi_{x^k} \left( -\frac{\gamma_k}{\gamma_k} \right) T_t^k - \frac{\gamma_{kv} X_{T_t}^{kv}}{\gamma_{kv}} + \varphi_{x^{kv}} \left( -\frac{\gamma_{kv}}{\gamma_{kv}} \right) T_t^k \right) \cdot \zeta,
\]

  - \(\sigma W_t \rightarrow \) constant drift adjustment \(\eta = \gamma \sigma^2\).

  - Pure jump Lévy process \(\rightarrow\) \(\pi^P(x) = e^{\gamma x} \pi^Q(x)\), drift adjustment: \(\eta = \varphi^P_j(1) - \varphi^Q_j(1) = \varphi^Q_j(1 + \gamma) - \varphi^Q_j(\gamma) - \varphi^Q_j(1)\).

  - Time change: instantaneous risk premium \((\eta \nu_t)\) proportional to the risk level \(\nu_t\).
Option pricing

- To compute the time-0 price of a European option price with maturity at $t$, we first compute the Fourier transform of the log return $\ln S_t / S_0$. Then we compute option value via Fourier inversions.

- The Fourier transform of a time-changed Lévy process:

$$
\phi_Y(u) \equiv E^Q \left[ e^{iuX_t} \right] = E^M \left[ e^{-\psi_x(u)T_t} \right], \quad u \in D \in \mathbb{C},
$$

- Tractability of the transform $\phi(u)$ depends on the tractability of (i) $\psi_x(u)$, and (ii) the Laplace transform of $T_t$ under $M$.

- Tractable $\psi_x(u)$ comes from the Lévy specification: diffusion, compound Poisson, DPL, NIG,...

- Tractable Laplace comes from activity rate dynamics: affine, quadratic, 3/2.

- The two $(X, T_t)$ can be chosen separately as building blocks, for different purposes.
Fourier inversion for a cumulative distribution

Example: a European call: \( C(k) = C(K, t)/S_0 = e^{-rt}E_0^Q [(e^{s_t} - e^k)1_{s_t \geq k}] \).

1. Treat \( C(x) = C(k = -x) = e^{-rt}E_0^Q [(e^{s_t} - e^{-x})1_{s_t \leq x}] \) analogous to a cumulative distribution.

- The option transform:

\[
\chi_c(z) \equiv \int_{-\infty}^{\infty} e^{izk} dC(x) = e^{-rt} \frac{\phi_s(-i - z)}{1 - iz}, \quad z \in \mathbb{R}.
\]

- The inversion is analogous to that for a cumulative distribution:

\[
C(x) = \frac{1}{2} \chi_c(0) + \frac{1}{2\pi} \int_{0}^{\infty} \frac{e^{izx} \chi_c(-z) - e^{-izx} \chi_c(z)}{iz} dz.
\]

with \( \chi_c(0) = e^{-qt} \).

- Use quadrature methods for the numerical integration.
Fourier inversion for a probability density

II. Treat $C(k)$ analogous to a *probability density function*.

- The option transform:

$$
\chi_p(z) \equiv \int_{-\infty}^{\infty} e^{izk} C(k) dk = e^{-rt} \frac{\phi_s(z - i)}{(iz)(iz + 1)}
$$

with $z = z_r - iz_i$, $z_i \in \mathcal{D} \subseteq \mathbb{R}^+$ for the option transform to be well defined.

- The inversion is analogous to that for a probability density:

$$
C(k) = \frac{1}{2} \int_{-iz_i - \infty}^{-iz_i + \infty} e^{-izk} \chi_p(z) dz = \frac{e^{-z_i k}}{\pi} \int_0^{\infty} e^{-izr} \chi_p(z_r - iz_i) dz_r.
$$

- The numerical integration can be cast into an FFT to improve the computational speed. Obtain options across all strikes simultaneously.

- Use fractional FFT to separate the choice of strike grids from the integration grids (Chourdakis (2005)).
Estimating the statistical dynamics

- Constructing likelihood of the Lévy return innovation based on Fourier inversion of the characteristic function. (CGMY (2002), Wu (2006))

- Euler approximation in the presence of complicated drift functions.

- Maximum likelihood with particle filtering in the presence of time changes and hence unobservable activity rates (Javaheri (2005)).

- MCMC Bayesian estimation (ErakerJohannesPolson (2003))

- Constructing variance swap rates from options and realized variance from high-frequency returns to make activity rates more observable. (Wu(2005))

  - **Future research**: Use more cross sections to estimate time-series dynamics (esp. return innovation distributions).
    - Combine index with single names.
    - Variance swap rates (interest rates) across different maturities.
Estimating the risk-neutral dynamics

- Nonlinear weighted least square to fit Lévy models to option prices. Daily calibration (Bakshi, Cao, Chen (1997), Carr and Wu (2003))

- Sometimes separate calibration per maturity is needed for a simple Lévy model.
  - Lévy processes with finite variance implies that non-normality dies away quickly with time aggregation.
  - Implied volatility smile/smirk flattens out at long maturities.
  - Separate calibration is necessary to capture smiles at long maturities.

- Adding a persistent stochastic volatility process (time change) helps improve the fitting along the maturity dimension.
  - Daily calibration: activity rates and model parameters are treated the same as free parameters.
  - Dynamic consistent estimation: Parameters are fixed, only activity rates are allowed to vary over time.
Dynamically consistent estimation

- Nested nonlinear least square (Huang and Wu (2004)).
- Cast the model into state-space form and use MLE.
  - Define state propagation equation based on the $\mathbb{P}$-dynamics of the activity rates. (Need to specify market price on activity rates, but not on return risks).
  - Define the measurement equation based on option prices (out-of-money values, weighted by vega,...)
  - Use an extended version of Kalman filter (EKF, UKF, PKF) to predict/filter the distribution of the states and measurements.
  - Define the likelihood function based on forecasting errors on the measurement equations.
  - Estimate model parameters by maximizing the likelihood.
Static v. dynamic consistency

- **Static cross-sectional consistency**: Option values across different strikes/maturities are generated from the same model (same parameters) at a point in time.

- **Dynamic consistency**: Option values over time are also generated from the same no-arbitrage model (same parameters).

While most academic & practitioners appreciate the importance of being both cross-sectionally and dynamically consistent, it can be difficult to achieve while generating good pricing performance. So it comes to compromises.

- **Market makers**:
  - Achieving static consistency is sufficient.
  - Matching market is important to provide two-sided quotes.

- **Long-term convergence traders**:
  - Pricing errors represent trading opportunities.
  - Dynamic consistency is important for long-term trading.
Joint estimation of $P$ and $Q$ dynamics

- Pan (2002): GMM. Choosing moment conditions becomes increasing difficult with increasing number of parameters.
- Eraker (2004): Bayesian with MCMC. Choose 2-3 options per day. Throw away lots of cross-sectional ($Q$) information.
- Bakshi & Wu (2005): MLE with filtering
  - Cast activity rate $P$-dynamics into state equation, cast option prices into measurement equation.
  - Use UKF to filter out the mean and covariance of the states and measurement.
  - Construct the likelihood function of options based on forecasting errors (from UKF) on the measurement equations.
  - Given the filtered activity rates, construct the conditional likelihood on the returns by Fast Fourier inversion of the conditional characteristic function.
Concluding remarks

- Modeling security returns with (time-changed) Lévy processes enjoys three key virtues:
  
  1. **Generality**: Lévy process can be made to capture any return innovation distribution; applying time changes can make this distribution vary stochastic over time.
  
  2. **Explicit economic mapping**: Each Lévy component captures shocks from one economic source. Time changes capture the relative variation of the intensities of these impacts.
  
  3. **Tractability**: Combining any tractable Lévy process (with tractable $\psi(u)$) with any tractable activity rate dynamics (with a tractable Laplace) generates a tractable Fourier transform for the time changed Lévy process. The two specifications are separate.

- It is a nice place to start with for generating security return dynamics that are parsimonious, tractable, economically sensible, and statistically performing well.