Modeling Financial Security Returns and Option Pricing with Time-Changed Lévy Processes

Liuren Wu

Zicklin School of Business, Baruch College
Overview

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

- **Generality:**
  - Lévy processes can generate pretty much any return innovation distribution.
  - Applying stochastic time changes on Lévy processes randomizes the innovation distribution over time \(\Rightarrow\) stochastic volatility, correlation, skewness, ....

- **Explicit economic mapping** by modeling returns with several time-changed Lévy components (versus models with hidden state vectors):
  - Each Lévy component captures shocks from one economic source.
  - Time change captures the time-varying intensity of its impact.

\(\Rightarrow\) makes model design more intuitive, parsimonious, and sensible.

- **Tractability:** A model is tractable for option pricing if we have under the risk-neutral measure \(\mathbb{Q}\):
  - tractable characteristic exponent for the Lévy components.
  - tractable Laplace transform for the activity rates underlying the time change.

\(\Rightarrow\) any combinations of the two generate tractable return dynamics.
Outline

I. Modeling return innovation using Lévy processes
II. Capturing stochastic volatility via stochastic time changes
III. Model design and assembly
IV. Option pricing
V. Model estimation
VI. Conclusion
Lévy processes and innovation distributions

- In discrete time, we can assume an arbitrary distribution for the return innovation: \( R_{t+1} = \mu_t + \sigma_t \varepsilon_{t+1} \).

- In continuous time, until recently, we had modeled return innovation using
  - either a Brownian motion (Black-Scholes)
  - or a compound Poisson process with normal jump size (Merton).

  \[ \Rightarrow \text{The return innovation distribution is either normal or mixture of normals.} \]

- Lévy processes greatly expand our continuous-time choice of iid return innovation distributions via the Lévy triplet \((\mu, \sigma, \pi(x))\). \((\pi(x)\text{—Lévy density})\).

- The Lévy-Khintchine Theorem:
  \[
  \phi_{X_t}(u) \equiv \mathbb{E}[e^{iuX_t}] = e^{-t\psi(u)},
  \]
  \[
  \psi(u) = -iu\mu + \frac{1}{2}u^2\sigma^2 + \int_{\mathbb{R}_0} \left(1 - e^{iux} + iux1_{|x|<1}\right) \pi(x)dx,
  \]
  distribution \leftrightarrow \text{characteristic exponent} \psi(u) \leftrightarrow \text{Lévy triplet} \ (\mu, \sigma, \pi(x))

- The model is “tractable” if the integral in \( \psi(u) \) can be carried out explicitly.
Tractable examples

• Brownian motion ($\mu t + \sigma W_t$): *Continuous movements; normal shocks.*

• Merton’s compound Poisson jumps: *Large but rare events, corporate default.*

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

• Dampened power law (CGMY, Wu):

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

– *Finite activity* when $\alpha < 0$: \( \int_{\mathbb{R}} \pi(x) dx < \infty \). Large but rare events.

– *Infinite activity* when $\alpha \geq 0$: Both small and large jumps. Jump frequency increase 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:
  \[
  \psi(u) = \lambda \left( 1 - e^{iu\mu - \frac{1}{2}u^2v} \right).
  \]

- Dampened power law: \( \psi(u) = -\Gamma(-\alpha)\lambda \left[ (\beta_+ - iu)^\alpha - \beta_+^\alpha + (\beta_- + iu)^\alpha - \beta_-^\alpha \right] - iuC(h) \) for \( \alpha \neq 0, \alpha \neq 1. \)
  
  - When \( \alpha \to 2 \), smooth transition to diffusion (quadratic function of \( u \)).
  
  - When \( \alpha = 0 \) (Variance-gamma by Madan et al):
    \[
    \psi(u) = \lambda \ln \left( 1 - iu/\beta_+ \right) \left( 1 + iu/\beta_- \right) = \lambda \left( \ln(\beta_+ - iu) - \ln \beta + \ln(\beta_- + iu) - \ln \beta_- \right).
    \]

  - When \( \alpha = 1 \) (exponentially dampened Cauchy, Wu 2006):
    \[
    \psi(u) = -\lambda \left( (\beta_+ - iu) \ln(\beta_+ - iu)/\beta_+ + \lambda (\beta_- + iu) \ln(\beta_- + iu)/\beta_- \right) - iuC(h).
    \]
Other Lévy examples

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

- All tractable in terms of the characteristic exponents $\psi(u)$.

- We can use FFT to generate the density function of the innovation (for model estimation).
Evidence on Lévy return innovations

- Credit risk: \textit{(compound) Poisson process}
  - The whole intensity-based credit modeling literature... \textit{Wait for Ken’s talk}

- Market risk: \textit{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))
  - Li and Yu (2005): Infinite-activity jumps cannot be approximated by finite-activity jumps.

- The role of diffusion (in the presence of infinite-variation jumps)
  - Not big, difficult to identify (CGMY (2002), Carr and Wu (2003a,b)).
  - Generate correlations with diffusive activity rates (Huang and Wu (2004)).
  - The diffusion component ($\sigma^2$) is identifiable even in presence of infinite-variation jumps (Aït-Sahalia (2004), Aït-Sahalia & Jacod 2005).
II. Capturing stochastic volatility via time changes

- Discrete-time analog again: $R_{t+1} = \mu_t + \sigma_t \epsilon_{t+1}$
  - $\epsilon_{t+1}$ is an iid return innovation, with an arbitrary distribution assumption ↔ Lévy process.
  - $\sigma_t$ is the conditional volatility, $\mu_t$ is the conditional mean return, both of which can be time-varying, stochastic...

- In continuous time, how do we model stochastic mean/volatility tractably?
  - If the return innovation is modeled by a Brownian motion, we can let the instantaneous variance to be stochastic and tractable, not volatility (Heston(1993), Bates (1996)).
  - If the return innovation is modeled by a compound Poisson process, we can let the Poisson arrival rate to be stochastic, not the mean jump size, jump distribution variance (Bates(2000), Pan(2002)).
  - 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

• Review the Lévy-Khintchine Theorem:

\[
\begin{align*}
\phi(u) & \equiv \mathbb{E} \left[ e^{iuX_t} \right] = e^{-t\psi(u)}, \\
\psi(u) & = -iu\mu + \frac{1}{2}u^2\sigma^2 \text{ for diffusion with drift } \mu \text{ and variance } \sigma^2, \\
\psi(u) & = \lambda \left( 1 - e^{iu\mu - \frac{1}{2}u^2v} \right) \text{ for Merton’s compound Poisson jump.}
\end{align*}
\]

• 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 \( t \to T_t \) instead of \( (\mu, \sigma^2, \lambda) \), for the same result.

• We define \( T_t \equiv \int_0^t v_s \, ds \) as the (stochastic) time change, with \( v_t \) 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 $(\mu_t, \sigma^2_t, \lambda_t)$ are all time varying, but they are all proportional to one common source of variation $\nu_t$.
  
  – Suppose we want $(\mu_t, \sigma^2_t, \lambda_t)$ 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 $X_t^1 \sim (0, 0, \lambda p(x))$, 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.

 grief (Economic Shock From Source i)}^{\text{Stochastic impacts}}.
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

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

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

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

Stochastic volatility/skew in currency options

II. Capturing stochastic volatility via time changes

Liuren Wu
Stochastic volatility/skew in SPX options

Model: Separate time changes to a diffusion and a negatively-skewed jump component. (Huang and Wu, JF 2004)
Tractable activity rate dynamics

- Most works use the square root process (affine):

  \[ dv_t = \kappa(\theta - v_t)dt + \sigma_v\sqrt{v_t}dW_t. \]

  - Including jumps with constant (Eraker, Johannes, Polson (2003)) and proportional (Wu(2005)) arrival rates.
  - Multi-dimensional extensions, with interactions (Duffie, Pan, Singleton (2000))

- *Quadratic* models: \( v_t \) is a quadratic function of an Ornstein-Uhlenbeck process (Santa-Clara and Yan (2005)), multivariate versions (Leippold and Wu (2004)).

- *Wishart* models: Buraschi, Cielask, Porchia, Tebaldi, Trojani,...

- *3/2* models (Heston, Lewis):

  \[ dv_t = \kappa v_t(\theta - v_t)dt + \sigma_v\sqrt{v_t^3}dW_t. \]

Much evidence favoring 3/2 over 1/2 in one-factor diffusion setting:

III. Model assembly

- Start with the risk-neutral ($\mathcal{Q}$) 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: $X^k_t \rightarrow X^k_{\tau_t}$
  - Adjust to guarantee the martingale condition: $\mathbb{E}_\mathcal{Q}[S_t/S_0] = e^{(r-q)t}$.

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

- $\mathbb{E}_\mathcal{Q}[e^{b X_t}] = e^{\varphi(b)t}$. Hence, $\mathbb{E}_\mathcal{Q}[e^{b X_t - \varphi(b)t}] = 1$, $\mathbb{E}_\mathcal{Q}[e^{b X_{\tau_t} - \varphi(b)\tau_t}] = 1$.

- Example: A CAPM model (with Levy return shocks and stochastic volatility):

\[ \ln S^i_t/S^i_0 = (r - q)t + \left( \beta^i X^m_{\tau^m_t} - \varphi^m(\beta^i) T^m_t \right) + \left( X^i_{\tau^i_t} - \varphi^i(1) T^i_t \right). \]

- 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 \((\gamma_k, \gamma_{kv})\)

\[
M_t = e^{-rt} \prod_{k=1}^{K} \exp \left( -\gamma_k X_{\overline{r}_k} - \varphi_k (-\gamma_k) T_{\overline{r}_k} - \gamma_{kv} X_{\overline{r}_{kv}} - \varphi_{kv} (-\gamma_{kv}) T_{\overline{r}_{kv}} \right) \cdot \zeta,
\]

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

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

  - Time change: instantaneous risk premium \((\eta \nu_t)\) proportional to the risk level \(\nu_t\).
Arbitrarily flexible market price of risks (for kicks only)

$(\gamma_k, \gamma_{kv})$ are complicated functions of $(X_{T_k}^k, X_{T_k}^{kv})$ and other state variables.

- The BS model $(\sigma W_t)$ with $\gamma = \gamma_0 + \gamma_1 Z_t + \gamma_2 Z_t^2 + \gamma_3 Z_t^3$, $\mathbb{P}$-dynamics:

$$dS_t/S_t = (r - q + (\gamma_0 + \gamma_1 Z_t + \gamma_2 Z_t^2 + \gamma_3 Z_t^3)\sigma^2)dt + \sigma dW_t.$$

- $Z_t$ can be some variables that predict asset returns (dividend yields, term spreads, default spread, etc).

- The Heston model with $\gamma'_v = \gamma_0/v_t + \gamma_1 v_t + \cdots + \gamma_k v_t^{k-1}$, and $\mathbb{Q}$-dynamics: $d\nu_t = (a - \kappa \nu_t)dt + \sigma_v \sqrt{\nu_t} dW_t^\nu$

Then, its $\mathbb{P}$-dynamics

$$d\nu_t = ((a + \gamma_0 \sigma_v^2) - (\kappa - \gamma_1 \sigma_v^2)\nu_t + \gamma_2 \sigma_v^2 \nu_t^2 + \cdots + \gamma_k \sigma_v^2 \nu_t^k)dt + \sigma_v \sqrt{\nu_t} dW_t^\nu,$$

- Caveat: Economic sensibility and discipline, identification.
IV. 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:

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

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

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

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

    $$C(k) = \frac{1}{2} \int_{-iz_1}^{-iz_1+\infty} e^{-izk} \chi_p(z) dz = \frac{e^{-z_ik}}{\pi} \int_0^\infty e^{-izrk} \chi_p(z_r - iz_i) dz_r.$$ 

    - 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)).
V. 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 (Eraker Johannes Polson (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 prices is important to provide two-sided quotes.

  – **Long-term convergence traders**:
    * Pricing errors represent trading opportunities.
    * Dynamic consistency is important for long-term convergence trading.

• A well-designed model (with several time changed Lévy components) can achieve both dynamic consistency and good performance. *Fewer parameters (parsimony), more activity rates.*
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.
  - The joint log likelihood equals the sum of the log likelihood of option pricing errors and the conditional log likelihood of stock returns.
Concluding remarks

• Modeling security returns with (time-changed) Lévy processes enjoys three key virtues:

  – *Generality*: Lévy process can be made to capture any return innovation distribution; applying time changes can make this distribution vary stochastic over time.

  – *Explicit economic mapping*: Each Lévy component captures shocks from one economic source. Time changes capture the relative variation of the intensities of these impacts.

  – *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.