Modeling Financial Security Returns Using Lévy Processes

Liuren Wu
Baruch College

http://faculty.baruch.cuny.edu/lwu/
Modeling return innovation using Lévy processes
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 return innovation distribution over time ⇒ stochastic volatility, and higher moments.

- **Explicit economic mapping** by modeling returns with several time-changed Lévy components (versus models with hidden state vectors):
  - Each Lévy component captures shocks from each economic source.
  - Time change captures the time-varying intensity of its impact.
  ⇒ makes model design more intuitive, parsimonious, and to the point.

- **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 transforms for the activity rates underlying the time change.
  ⇒ any combinations of the two generate tractable return dynamics.
Modeling return innovations

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

- Until recently, we had modeled return innovation using
  - either a Brownian motion (Black-Scholes)
  - or a compound Poisson process with normal jump size (Merton).
  \[ \Rightarrow \text{The return innovation distribution is either normal or mixture of normals.} \]

- For every return innovation distribution assumption, we can in principle find a Lévy triplet \((\mu, \sigma, \pi(x))\) that generates such a distribution.

- The Lévy-Khintchine Theorem:

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

  distribution \(\leftrightarrow\) characteristic exponent \(\psi(u)\) \(\leftrightarrow\) Lévy triplet \((\mu, \sigma, \pi(x))\)

- Constraint on the Lévy density \(\pi(x)\):
  \[
  \int_0^1 x^2\pi(x)dx < \infty \text{ (finite quadratic variation), a necessary and sufficient condition for } X_t \text{ to be a semimartingale, and hence a Lévy process.} \]
Generalized Fourier transforms

- We expand the characteristic coefficient $u$ to the subset $(\mathcal{D})$ of the complex plane ($\mathbb{C}$) where the characteristic exponent $\psi(u)$ is well-defined.

$$
\psi(u) = -iu\mu + \frac{1}{2}u^2\sigma^2 + \int_{\mathbb{R}_0} \left(1 - e^{iu\pi} + iux1_{|x|<1}\right)\pi(x)dx, \quad u \in \mathcal{D} \in \mathbb{C}. \quad (1)
$$

$\phi(u)$ under the generalized $u$ is called the generalized Fourier transform.

- It is often convenient to define the extended version of the cumulant exponent,

$$
\varphi_x(s) \equiv \frac{1}{t} \ln \mathbb{E}[e^{sX_t}] = s\mu + \frac{1}{2}s^2\sigma^2 + \int_{\mathbb{R}_0} (e^{sx} - 1 - sx1_{|x|<1})\pi(x)dx, \quad s \in \mathcal{D} \in \mathbb{C}.
$$

- When both are well defined, the characteristic exponent and the cumulant exponent are linked by

$$
\varphi(s) = -\psi(-is).
$$

- We regard a Lévy process “tractable” as long as we can derive the characteristic (cumulant) exponent analytically, i.e., we can carry out the integral in (1) explicitly.
Example: Diffusion

- Arithmetic Brownian motion \((\mu t + \sigma W_t)\).

- Application: Continuous movements; normally distributed shocks.

- Lévy triplet: \((\mu, \sigma^2, 0)\).

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

- Cumulant exponent: \(\varphi(s) = s\mu + \frac{1}{2}s^2\sigma^2\).

- Probability density: \(f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\).

- Literature:
  - The Black-Scholes (Merton, Garman, Kohlhagen...) model:

    Risk-neutral measure \(\mathbb{Q}\):  
    \[
    \ln S_t / S_0 = (r - q)t + \sigma W_t - \frac{1}{2}\sigma^2.
    \]

    Statistical measure \(\mathbb{P}\):  
    \[
    \ln S_t / S_0 = \mu t + \sigma W_t - \frac{1}{2}\sigma^2.
    \]

  - Bachelier model:  
    \[
    S_t = S_0 + \mu t + \sigma W_t.
    \]
Example: Compound Poisson jumps

... with normal-distributed jump size

- Jumps arrive via a Poisson process with mean arrival rate $\lambda$.
- Upon a jump arrival, the jump size is normally distributed $(\mu_J, \nu_J)$.
- Application: Large but rare events, such as corporate default.
- Lévy triplet: $(0, 0, \pi(x))$ where

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

(2)

- Characteristic/cumulant exponents

$$\psi(u) = \lambda \left( 1 - e^{iu\mu_J - \frac{1}{2}u^2\nu_J} \right), \quad \varphi(s) = \lambda \left( e^{s\mu_J + \frac{1}{2}s^2\nu_J} - 1 \right).$$

No truncation term necessary for compound Poisson jumps.

- Lit: Merton (1976) jump-diffusion, Lévy triplet $(\mu, \sigma^2, \pi(x))$ and with $\pi(x)$ as in (2).
- Special case: Upon arrival, stock price jumps to zero (log return jumps to $-\infty$).
  Derive $\pi(x), \psi(u), \varphi(s)$.
- Think of other jump size distributions, work out their $(\pi(x), \psi(u), \varphi(s))$. 

Example: Dampened Power Law (DPL)

- The Lévy density follows exponentially dampened power law:

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

- **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 many small jumps. (Brownian motion)

- \(\alpha \leq 2\) is needed to guarantee *finite quadratic variation*.

*Market movements of all magnitudes, from small movements to market crashes.*
The Lévy density follows exponentially dampened power law:

\[
\pi(x) = \begin{cases} 
\lambda \beta_+ \exp(-\beta_+ x) x^{-\alpha-1}, & x > 0, \\
\lambda \beta_- \exp(-\beta_- |x|) |x|^{-\alpha-1}, & x < 0,
\end{cases} \quad \lambda, \beta_+, \beta_- > 0, \alpha \leq 2. \quad (3)
\]

- When \(\beta_\pm = 0\) and \(\alpha > 0\), the power law specification in (3) describes the \(\alpha\)-stable motion, which generates the \(\alpha\)-stable distribution (Fama, Mandelbrot).

- Applying measure changes defined by exponential martingales on \(\alpha\)-stable motion generates DPL.

- When \(\alpha = -1\), no power, only exponential function. It describes the double-exponential model of Kou (2005). Compound Poisson with double-exponential jump specification.

- \(\alpha = 0\) corresponds to the variance gamma model (VG) of Madan et al.

- The specification in (3) originates from CGMY (2002), who consider it as an extension of VG.
Characteristic exponents for DPL

- The general case ($\alpha \neq 0, \alpha \neq 1$):

  \[
  \psi(u) = -\Gamma(-\alpha)\lambda \left[ (\beta_+ - iu)^\alpha - \beta_+^\alpha + (\beta_- + iu)^\alpha - \beta_-^\alpha \right] - iuC(h),
  \]

  \[
  \varphi(s) = \Gamma(-\alpha)\lambda \left[ (\beta_+ - s)^\alpha - \beta_+^\alpha + (\beta_- + s)^\alpha - \beta_-^\alpha \right] + sC(h),
  \]

  where $C(h)$ is a linear term induced by the truncation function, which is needed for infinite variation jumps (when $\alpha \geq 1$).

  - When $\alpha \to 2$, smooth transition to diffusion (quadratic function of $u$).
  - When $\alpha = 0$ (VG):

    \[
    \psi(u) = \lambda \ln \left(1 - iu/\beta_+\right) \left(1 + iu/\beta_-\right) = \lambda \left(\ln(\beta_+ - iu) - \ln \beta + \ln(\beta_- + iu) - \ln \beta_-\right)
    \]

  - When $\alpha = 1$ (exponentially dampened Cauchy):

    \[
    \psi(u) = -\lambda \left( (\beta_+ - iu) \ln (\beta_+ - iu) / \beta_+ + \lambda (\beta_- + iu) \ln (\beta_- + iu) / \beta_- \right) - iuC(h).
    \]

- Read Carr, Madan, Chang (98) for VG, read Wu (2006) for DPL. Derive the cumulant exponents for DPL, including $C(h)$, under the truncation $x^1_{|x|\leq 1}$. 
Other Lévy examples for return innovations

- 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 $\psi(u)$.

- Drive their cumulant exponents, characteristic exponents.
- Code the characteristic function.
- Use FFT to generate the density function.
Generate distributions from CF

Given the characteristic function (CF) $\phi(u)$ for a process, we can obtain its probability density function via the following inversion:

$$p(x) = \frac{1}{2} \int_{-\infty}^{\infty} e^{-iux} \phi(u) du = \frac{1}{\pi} \int_{0}^{\infty} e^{-iux} \phi(u) du.$$  

It can be cast into the Fast Fourier Transform (FFT) format:

$$p(x_j) \approx \hat{p}(x_j) = \frac{1}{\pi} \sum_{n=0}^{N-1} e^{-iux_n} \phi(u_n) \Delta u = \sum_{n=0}^{N-1} f_n e^{-j\frac{2\pi n}{N} i}, \quad j = 0, 1, ..., N - 1,$$

with $f_n = \frac{1}{\pi} e^{ib\eta n} \eta \phi(\eta n)$, $u_n = \eta n$, $x_j = -b + \lambda j$, $\eta \lambda = 2\pi/N$.

Consider fractional FFT for separation of $\eta$ and $\lambda$.

The cumulative distribution function can be obtained by

$$P(x) = \frac{1}{2} + \frac{1}{2\pi} \int_{0}^{\infty} \frac{e^{iux} \phi(-u) - e^{-iux} \phi(u)}{iu} du.$$  

Use quadrature methods to approximate the integral.
Evidence on Lévy return innovations

■ Credit risk: (compound) Poisson process
  ◆ The whole “reduced-form” credit modeling literature...
  ◆ Carr and Wu: The impact of corporate default on stock price: Poisson arrival with jump to zero.
  ◆ Carr and Wu: The impact of sovereign default on currency price: Poisson arrival with random downside jump size distribution.

■ Market risk:
  ◆ Evidence from stock returns (CGMY (2002)): The $\alpha$ estimates 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))

■ 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)).
  ◆ Unresolved: Does SPX return contain a diffusion component?
Capturing stochastic volatility via time changes
Modeling stochastic volatility

- In discrete time, we can write the return as $R_{t+1} = \mu_t + \sigma_t \varepsilon_{t+1}$
  - $\varepsilon_{t+1}$ 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...

- In continuous time, how do we model stochastic 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).
  - 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, mean arrival rate of the Poisson process are both proportional to time.*
Randomize the time

- Review the Lévy-Khintchine Theorem:

\[
\phi(u) \equiv \mathbb{E}[e^{iuX_t}] = e^{-t\psi(u)},
\]

\[
\psi(u) = -iu\mu + \frac{1}{2}u^2\sigma^2 \text{ for diffusion with drift } \mu \text{ and variance } \sigma^2,
\]

\[
\psi(u) = \lambda \left(1 - e^{iu\mu J - \frac{1}{2}u^2v_J} \right) \text{ for Merton's compound Poisson jump.}
\]

- 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.
Examples

- If \( X_t \) is characterized by \((\mu, 0, 0)\): \( X_t = \mu t \), randomizing time is equivalent to randomizing \( \mu t \).

\[
X_{T_t} = \mu T_t = \mu \int_0^t v_s \, ds = \int_0^t \mu_s \, ds,
\]

with \( \mu_t = \mu v_t \). Hence, the activity rate \( v_t \) is the instantaneous drift up to a scale.

- If \( X_t \) is characterized by \((0, \sigma^2, 0)\): \( X_t = \sigma W_t \), randomizing time is equivalent to randomizing the variance \( \sigma_t^2 \).

\[
X_{T_t} = \sigma W_{T_t} = \mu W \int_0^t v_s \, ds = \int_0^t \sigma \sqrt{v_s} \, dW_s = \int_0^t \sigma_s \, dW_s,
\]

with \( \sigma_t^2 = \sigma^2 v_t \). Hence, \( v_t \) is the instantaneous variance up to a scale.

- If \( X_t \) is characterized by \((0, 0, \lambda p(x))\), where \( p(x) \) denotes the jump size distribution conditional on one jump occurring, randomizing time is equivalent to randomizing the mean arrival rate \( \lambda_t = \lambda v_t \).
The Heston model under measure $Q$:

$$dS_t/S_t = (r - q)dt + \sqrt{v_t}dW_t, \quad dv_t = \kappa (\theta - v_t) dt + \sigma_v \sqrt{v_t} dW_t^v.$$  

- In log return (Ito’s lemma):
  $$d\ln S_t = (r - q)dt + \sqrt{v_t}dW_t - \frac{1}{2}v_t dt.$$  
- In integrated return form,
  $$\ln S_t/S_0 = (r-q)t + \int_0^t \sqrt{v_s}dW_s - \frac{1}{2} \int_0^t v_s ds, v_t = v_0 + \kappa \theta t - \int_0^t v_s ds + \int_0^t \sigma_v \sqrt{v_s} dW_s.$$  
- Think of $\mathcal{T}_t = \int_0^t v_s ds$ as a time change,
  $$\ln S_t/S_0 = (r-q)t + W_{\mathcal{T}_t} - \frac{1}{2} \mathcal{T}_t, \quad v_t = v_0 + \kappa \theta t - \mathcal{T}_t + \sigma_v W_{\mathcal{T}_t}^v.$$  

Some equalities (in distribution):

$$W_{\sigma^2 t} = \sigma W_t = \sigma \sqrt{t} W_1,$$

$$W_{\mathcal{T}_t} = W_{\int_0^t v_s ds} = \int_0^t \sqrt{v_s} dW_s,$$

$$dW_{\mathcal{T}_t} = \sqrt{v_s} dW_s.$$
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 \( X_t \), 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 \( v_t \).
  - 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 shocks.
Example: Return on a stock

- Model the return on a stock as reflecting shocks from two sources
  - credit risk: Corporate default leads to 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 a stock

The dynamics under $\mathbb{Q}$:

$$dS_t / S_t = (r(t) - q(t) + \lambda(t)) \, dt + \sqrt{v(t)} \, dW^s_t,$$
conditional on no default, $S_t = 0$ otherwise.

$$dv(t) = (\theta_v - \kappa_v v(t)) \, dt + \sigma_v \sqrt{v(t)} \, dW^v_t,$$

$$\lambda(t) = \xi v(t) + z(t),$$

$$dz(t) = (\theta_z - \kappa_z z(t) - \kappa_{zv} v(t)) \, dt + \sigma_z \sqrt{z(t)} \, dW^z_t,$$

$$\rho^{sv} \equiv \mathbb{E} [dW^s dW^v] / dt < 0, \quad \rho^{sz} \equiv \mathbb{E} [dW^s dW^z] = 0, \quad \rho^{zv} \equiv \mathbb{E} [dW^z dW^v] = 0.$$

- Default arrival $\lambda(t)$ is \textit{correlated} with the diffusion variance through $\xi$, but also has \textit{independent} variation through $z_t$.

- $\rho^{sv}$ captures the leverage effect.

- Time change notation:

$$T^v_t = \int_0^t v(s) \, ds, \quad T^z_t = \int_0^t z(s) \, ds, \quad T^\lambda_t = \int_0^t \lambda(s) \, ds = \xi T^v_t + T^z_t.$$

$$\ln S_t / S_0 = (r(0, t) - q(0, t)) \, t + T^\lambda_t + W^s_{T^v_t} - \frac{1}{2} T^v_t,$$ conditional on no default.
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.
Example: Return on an exchange rate

The dynamics under $\mathcal{Q}$:

$$\ln S_t/S_0 = (r_d(0, t) - r_f(0, t))t + (X^L_{T_t} - \varphi_X(1)T^L_t) + (X^R_{T_t} - \varphi_X(1)T^R_t).$$

- $X^L = \sigma^L W^L_t + J_-$ contains a diffusion and a downward jump.
- $X^R = \sigma^R W^R_t + J_+$ contains a component and a upward jump.
- $T^L_t$ and $T^R_t$ are independent, and driven by the following activity rate process,

$$dv^j_t = \kappa^j (1 - v^j_t)dt + \sigma^j_r \sqrt{v^j_t}dW^{jv}, \quad j = R, L, \quad \mathbb{E}[dW^R_t dW^L_t] = 0.$$

- “Leverage effect” on the two economies:

$$\mathbb{E}[dW^R_t dW^R_t] = \rho^R dt > 0, \quad \mathbb{E}[dW^L_t dW^L_t] = \rho^L dt < 0, \quad \mathbb{E}[dW^R_t dW^L_t] = 0.$$

- Symmetry assumption can dramatically reduces the number of parameters:

$$\sigma^L = \sigma^R, J_- = -J_+, \kappa^L = \kappa^R, \sigma^L_v = \sigma^R_v, \rho^L = -\rho^R.$$
Exchange rates and pricing kernels

- Exchange rate reflects the interaction between two economic forces.
- This economic meaning becomes clearer if we build models based on the pricing kernels of the two economies.
  - Let $S_t =$ Dollar price of yen (home currency = dollar, foreign = yen). Let $m_{0,t}^{US}$ and $m_{0,t}^{JP}$ denote the pricing kernels of the two economies. Then
    \[
    \ln S_t/S_0 = \ln m_{0,t}^{JP} - \ln m_{0,t}^{US}.
    \]
  - If we model the negative of the logarithm of each pricing kernel ($-\ln m_{0,t}^j$) as a time-changed Levy process, $X_t^j (j = US, JP)$ with negative skewness (diffusion + downside jumps, negative correlations between return and volatility). Then, \[
    \ln S_t/S_0 = \ln m_{0,t}^{JP} - \ln m_{0,t}^{US} = X_{T_t}^{US} - X_{T_t}^{JP}.
    \]
  - Then we can go back to the previous page for each of these time-changed Lévy processes.
  - No arbitrage guarantees the existence of at least one positive pricing kernel for each economy. (Cochrane, *Asset Pricing*, uses pricing kernel to discuss traditional CAPM questions)
Stochastic volatility/skew in currency options

JPYUSD

GBPUSD

Implied Volatility, %

JPYUSD

GBPUSD

RR10 and SM10, %ATMV

RR10 and SM10, %ATMV
Stochastic volatility/skew in SPX options

Model: Separate time changes to a diffusion and a negatively-skewed jump component. (Huang and Wu (2004), SV4)

- **Future research**: Joint modeling of equity indexes of two economies and the exchange rate between them.
- **Wish**: Jointly price bond, stock, and currency options.
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 (Carr and Wu).

- Quadratic models: \( v_t \) is a quadratic function of an Ornstein-Uhlenbeck process (Santa-Clara and Yan (2005)), multivariate versions (Carr and Wu (2004)).
  - Future research: Not fully explored.

- 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:
  - Future research: Need work on option pricing and implementation on 3/2.
Model assembly
Model assembly

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

\[
\ln S_t/S_0 = (r - q)t + \sum_{k=1}^K \left( b^k X_{T_t}^k - \varphi_{x,k}(b^k) T_{T_t}^k \right),
\]

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

- Examples:
  - The Black-Scholes model: \(K = 1, X_t = \sigma W_t, b = 1\):
    \[
    \ln S_t/S_0 = (r - q)t + \sigma W_t - \frac{1}{2} \sigma^2 t.
    \]
  - Heston SV model: \(\ln S_t/S_0 = (r - q)t + \sigma W_{T_t} - \frac{1}{2} \sigma^2 T_t\) with
    \[
    dv_t = \kappa(1 - v_t) dt + \sigma_v \sqrt{v_t} dW^v_t.
    \]
  - A CAPM model (with Levy return shocks and stochastic volatility):
    \[
    \ln S^j_t/S^j_0 = (r - q)t + \left( b^j X_{T_t}^m - \varphi_{x,m}(b^j) T_{T_t}^m \right) + \left( X_{T_t}^j - \varphi_{x,j}(1) T_{T_t}^j \right).
    \]
  - Future research: joint estimation on index and single name stock options.
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 \( (\gamma_k, \gamma_{kv}) \)

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

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

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

- \( X_{T_t} \) → proportional instantaneous drift adjustment \( \eta v_t \) as \( dT_t = v_t dt \).
  Example: \( \sigma dW_{T_t} = \sigma \sqrt{v_t} dW_t \rightarrow \text{drift adj} = \gamma \sigma^2 v_t \).
Arbitrarily flexible market price of risks

$(\gamma_k, \gamma_{kv})$ 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 $\gamma_v^t = \gamma_0 / v_t + \gamma_1 + \gamma_2 v_t + \cdots + \gamma_{kv} v_t^{k-1}$, and

$\mathbb{Q}$-dynamics: $dv_t = \left(a - \kappa v_t\right) dt + \sigma_v \sqrt{v_t} dW_t^v$

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

$$dv_t = \left((a + \gamma_0 \sigma^2_v) - (\kappa - \gamma_1 \sigma^2_v) v_t + \gamma_2 \sigma^2_v v_t^2 + \cdots + \gamma_{kv} \sigma^2_v v_t^k\right) dt + \sigma_v \sqrt{v_t} dW_t^v.$$  

- Note: $\gamma_v^t$ is the market price on $\sigma_v \sqrt{v_t} dW_t^v$, not on $dW_t^v$.

- Cheridito, Filipovic, Kimmel (2003), Pan& Singleton(2005): $\gamma_k = 0$ for $k \geq 2$.

- With $\gamma_0 \neq 0$, risk premium approaches a finite amount when risk $(v_t)$ goes to zero. → Economically sensible? even if it is mathematically ok (no arbitrage).
Option pricing
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 (later).

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

$$
\phi_Y(u) \equiv \mathbb{E}^Q \left[ e^{iuX_T} \right] = \mathbb{E}^M \left[ e^{-\psi_x(u)T} \right],
$$

where the new measure $\mathbb{M}$ is defined by the exponential martingale:

$$
\frac{d\mathbb{M}}{d\mathbb{Q}} \bigg|_t = \exp(iuX_T + T\psi_x(u)).
$$

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

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

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

- The two $(X, T)$ can be chosen separately as building blocks, for different purposes.
The $\mathbb{Q}$-dynamics of the log return is

$$ s_t \equiv \ln S_t / S_0 \equiv (r - q) t + \sigma W_T - \frac{1}{2} \sigma^2 T, \quad T = \int_0^t v_s ds $$

$$ dv_t = \kappa (\theta - v_t) dt + \sigma_v \sqrt{v_t} dW^v_t, \quad \mathbb{E}[dW_t dW^v_t] = \rho dt $$

For identification, we need to set either $\sigma = 1$ or $\theta = 1$.

- When $\sigma = 1$, $\theta$ becomes the log-run variance and $v_t$ the instantaneous variance rate. This is the original Heston specification.
- We instead let $\sigma$ be a free parameter and set $\theta = 1$. It is equivalent. But our specification is more convenient later...
- In time changed language, $X_t = \sigma W_t - \frac{1}{2} \sigma^2 t X_T = \sigma W_T - \frac{1}{2} \sigma^2 T$.

The Fourier transform of the log return is:

$$ \phi_s(u) \equiv \mathbb{E}^Q \left[ e^{iulnS_t / S_0} \right] = e^{iu(r-q)t} \mathbb{E}^Q \left[ e^{iuX_T} \right] = \mathbb{E}^M \left[ e^{-\psi_x(u)T_t} \right], $$

- $\psi_x(u) = \frac{1}{2} \sigma^2 (u^2 + iu)$ for $X_t = \sigma W_t - \frac{1}{2} \sigma^2 t$.
- Measure change $\frac{dM}{dQ} \bigg|_t = \exp \left( iu\sigma W_T + \frac{1}{2} u^2 \sigma^2 T \right)$. 
Heston model

- The Fourier transform of the log return is:

\[ \phi_s(u) \equiv \mathbb{E}^{Q} \left[ e^{iu \ln S_t / S_0} \right] = e^{iu(r-q)t} \mathbb{E}^{Q} \left[ e^{iu X_{\tau_t}} \right] = e^{iu(r-q)t} \mathbb{E}^{\mathbb{M}} \left[ e^{-\psi_x(u)\tau_t} \right], \]

- \( \psi_x(u) = \frac{1}{2} \sigma^2(u^2 + iu) \) for \( X_t = \sigma W_t - \frac{1}{2} \sigma^2 t. \)

- Measure change \( \frac{d\mathbb{M}}{d\mathbb{Q}} \big|_t = \exp \left( iu \sigma^2 W_{\tau_t} + \frac{1}{2} u^2 \sigma^2 \tau_t \right). \)

- \( dv_t = \kappa (1 - v_t) dt + \sigma_v \sqrt{v_t} dW_t^v \) under \( \mathbb{Q} \) becomes \( dv_t = (\kappa - (\kappa - iu \sigma \sigma_v \rho) v_t) dt + \sigma_v \sqrt{v_t} dW_t^v \) under \( \mathbb{M}. \) Let \( \kappa^\mathbb{M} = \kappa - iu \sigma \sigma_v \rho. \)

- The Laplace transform of \( \tau_t \) is exponential affine in \( v_0: \)

\[ e^{iu(r-q)t} \mathbb{E}^{\mathbb{M}} \left[ e^{-\psi_x(u)\tau_t} \right] = e^{iu(r-q)t} \mathbb{E}^{\mathbb{M}} \left[ e^{-\psi_x(u) \int_0^t v_s ds} \right] = e^{iu(r-q)t} e^{-b(t)v_0 - c(t)} \]

with \( b'(t) = \psi_x(u) - (\kappa^\mathbb{M}) b(t) - \frac{1}{2} \sigma_v^2 b(t)^2, \quad c'(t) = \kappa b(t), \) starting at \( b(0) = c(0) = 0. \) They can be solved analytically:

\[ b(t) = \frac{2 \psi_x(u) \left( 1 - e^{-\xi t} \right)}{2 \xi - (\xi - \kappa^\mathbb{M})(1 - e^{-\xi t})}, \quad c(t) = \frac{\kappa^\mathbb{M}}{\sigma_v^2} \left[ 2 \ln \left( 1 - \frac{\xi - \kappa^\mathbb{M}}{2 \xi} \left( 1 - e^{-\xi t} \right) \right) + (\xi - \kappa^\mathbb{M}) t \right], \]

with \( \xi = \sqrt{(\kappa^\mathbb{M})^2 + 2 \sigma_v^2 \psi_x(u)}. \)
Characteristic exponents of Lévy processes

- The Lévy process underlying the Heston model is pure diffusion
  \[ X_t = \sigma W_t - \frac{1}{2} \sigma^2 t. \]
  Its characteristic exponent is \( \psi_x(u) = \frac{1}{2} \sigma^2 (u^2 + iu). \)

- If we replace the diffusion with a pure jump process, \( J, X_t = J_t - \varphi_J(1)t, \) its
  characteristic exponent is \( \psi_X(u) = \psi_J(u) + iu \varphi_J(1). \)

- DPL, NIG, Merton’s compound Poisson all lead to analytical forms for \( \psi(u). \)

- For returns, we always need the concavity adjustment \( \varphi(1) \) and hence \( iu \varphi(1) \) in
  the characteristic exponent.

- Without time changes, the Fourier transform is simply
  \[ \phi_S(u) = e^{(r-q)t} e^{-t \psi_x(u)}. \]

- Code up the Fourier transform of log returns under the following Lévy models
Laplace transform of affine activity rates

Given the characteristic exponent $\psi_x(u)$, we can now focus on the Laplace transform induced by time changes,

$$\phi_s(u) = e^{iu(r-q)t} \mathbb{E}^{\mathbb{M}} \left[ e^{-\psi_x(u)T_t} \right] = e^{iu(r-q)t} \mathbb{E}^{\mathbb{M}} \left[ e^{-\psi_x(u) \int_0^t v_s ds} \right],$$

The transform is exponential affine if $v_t = c_0 + b_0^\top Z_t$, and the instantaneous drift vector $(\mu_t^Z)$, the instantaneous diffusion covariance matrix $(\Sigma_t^Z)$, and the jump arrival rate $\pi_t(x)$ of $Z_t$ are all affine in $Z_t$:

$$dZ_t^i = (A^i + B^i Z_t)dt + \left( \sqrt{\alpha_i + \beta_i^\top Z_t} \right)_{i} dW_t^{Z_i} + dJ((C^i + (D^i)^\top Z_t)\pi^i(x))$$

The Laplace transform is $\mathbb{E}^{\mathbb{M}} \left[ e^{-\psi_x(u) \int_0^t v_s ds} \right] = e^{-b(t)^\top Z_0 - c(t)}$ with

$$b'(t) = \psi(x)b_0 + B^\top b(t) - \frac{1}{2} \sum_i b(t)^\top b(t) \bigodot \beta_i - D \left( L_q(b(t)) - 1 \right);$$

$$c'(t) = \psi(u)c_0 + b(t)^\top a - b(t)^\top \alpha b(t)/2 - C \left( L_q(b(t)) - 1 \right),$$

Heston is a one-factor diffusion special case. Carr and Wu (2005a.b.c) consider two factor examples.
Laplace transform of quadratic activity rates

- If \( v_t = c_0 + b_0^\top Z_t + Z_t^\top A_0 Z_t \), and

\[
dZ_t = -\kappa Z_t dt + dW_t^v
\]

The Laplace transform is

\[
E^M \left[ e^{-\psi(u) \int_0^t v_s ds} \right] = e^{-Z_0^\top a(t)Z_0 - b(t)^\top Z_0 - c(t)}
\]

with

\[
a'(t) = \psi(u)a_0 - a(t)\kappa - \kappa^\top a(t) - 2a(t)^2;
\]
\[
b'(t) = \psi(u)b_0 - \kappa b(t) - 2a(t)^\top b(t);
\]
\[
c'(t) = \psi c_0 + tr a(t) - b(t)^\top b(t)/2,
\]

- Too many free parameters. Model design needs to be more structured.
- Santa-Clara, Yan (2005) is a good applied example.
Laplace transform of 3/2 activity rates

- I have a small write up (derivation) and a discussion on extension on this.
- The solution involves a confluent hypergeometric function.
- I’ll leave it to you to make it fully work for option pricing.
Fourier inversion for a cumulative distribution

Example: a European call: \( C(k) = C(K, t)/S_0 = e^{-rt} \mathbb{E}_0^Q [(e^{st} - e^k)1_{s \geq k}]. \)

\[ I. \] Treat \( C(x) = C(k = -x) = e^{-rt} \mathbb{E}_0^Q [(e^{st} - e^{-x})1_{-s \leq x}] \) analogous to a cumulative distribution.

\begin{itemize}
\item The option transform:
\[ \chi_c(z) \equiv \int_{-\infty}^{\infty} e^{izx} dC(x) = e^{-rt} \phi_s (-i - z) \frac{1}{i z}, \quad z \in \mathbb{R}. \]
\end{itemize}

\begin{itemize}
\item The inversion is analogous to that for a cumulative distribution:
\[ C(x) = \frac{1}{2} \chi_c(0) + \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{izx} \chi_c(-z) - e^{-izx} \chi_c(z)}{iz} dz. \]
\end{itemize}

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

\begin{itemize}
\item The literature often treats the call as a combination of two contingent claims and does the inversion separately. \( C(x) = e^{-qt} Q_1(x) - e^{-rt} e^{-x} Q_2(x), \)
\end{itemize}

\begin{itemize}
\item Doing one inversion is faster than doing two.
\item Explore quadrature methods for the numerical integration.
\end{itemize}
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 $z = z_r - iz_i$. We need $z_i \in D \subseteq \mathbb{R}^+$ for the call option transform to be well defined.

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

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

- 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)).

*Future research*: Need more robust analysis, esp. for model estimations.
Model estimation

- Estimating the statistical dynamics
- Estimating the risk-neutral dynamics
- Dynamically consistent estimation
- Static v. dynamic consistency
- Joint estimation of $P$ and $Q$ dynamics
- Concluding remarks

Further research topics
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.
  - *Future research*: Replace UKF with PKF for non-gaussian state dynamics. Improve the speed/accuracy for model estimation.
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 $\mathbb{P}$ and $\mathbb{Q}$ 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 ($\mathbb{Q}$) information.

- Bakshi & Wu (2005): MLE with filtering
  - Cast activity rate $\mathbb{P}$-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.
  - *Future research*: Better drift identification. Model uncertainty and robust control...
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.
Further research topics

- Apply time-changed Lévy process for interest rate options, especially eurodollar futures options.
  - Consistency across caps and swaptions.

- Impose time-changed Lévy innovation to CEV specification for stock prices or interest rates.

\[
dS_t = (r - q)S_t dt + S_t^\beta dX_{T_t}.
\]

- Simulating time-changed Lévy processes for the pricing of barriers and other exotics.