# Beyond stochastic volatility and jumps in returns and volatility

Garland Durham and Yangho Park
Leeds School of Business
University of Colorado



## Background

- Given a financial asset ("stock"), we **observe**
  - the price of the **asset**, and
  - prices of **options** on it (various times to maturity and moneyness)
- We would like to have **models** that describes these data:
  - **physical measure** describes **asset price dynamics**
  - **risk neutral measure** describes **option prices**
- Given physical and risk-neutral models
  - we should be able to learn something about **risk premia** by looking at **differences** between them.



## Background — continued

- **Physical measure:** Daily prices on the underlying typically involve **several thousand** observations. Sufficient to fit a basic model with
  - shape of returns distribution (conditional on volatility)
  - time-varying volatility
  - possibly other features
- **Risk-neutral measure:** Options provide **far more information**. A typical data set may involve **hundreds of thousands** of option prices.
  - options with varying **moneyness** provide information on **shape** of returns distribution
  - options with varying **time to maturity** provide information on expectations at **various horizons**

  Fitting these requires the model to have **time-varying features** to fit
  - implied variance, skewness, kurtosis, ...
  - at various horizons

  on a **daily basis** (not just on average).



## Background — continued

While this seems like a fairly **straightforward problem**, it turns out to be **difficult**.

- The fundamental problem here is one that many people would like to have:
  - We have **too much** data!
  - Easy to **invalidate** almost any model you might imagine.
- Also, option prices provide only **indirect** information about states.
  - Models are written in terms of **one-day ahead** densities, but options are at **multi-day** horizons.
  - Need to **invert** risk-neutral measure.

## Background — continued

The goals are:

- Find models that **combine** the information from these two sources
- Assess the extent to which the **dynamics** implied by **option prices** are consistent with the dynamics of the **underlying asset price**.



## Background — continued

Modelling framework typically uses **affine-jump models** with a single volatility factor.

**Issues:**

- Only **one factor** (volatility).
  - Can fit option-implied volatility at a single horizon.
  - Cannot fit option-implied skewness, kurtosis ... at any horizons.
- **Jump specification** is too restrictive to fit shapes of distributions
  - Mixtures of normals?
- **Affine model** does not fit the data. Adding **jumps** doesn't fit the problem.
  - But give easy option-pricing formulae...
- How to get good approximations to true **continuous-time model** (theoretical basis)
  - How good is the **Euler scheme** approximation?
  - How much **better** can we do at reasonable computational cost (cpu and programmer)?



## Background — continued

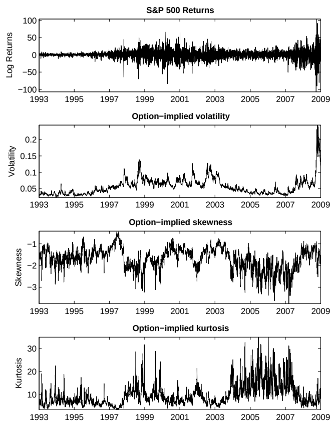
Our **research agenda** is directed toward addressing these problems.

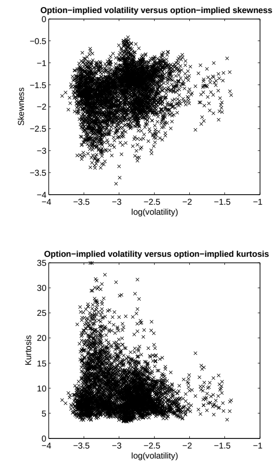
- We have developed a good collection of **tools and techniques**, and accumulated lots of **results**.
- Are now beginning to **write** some of this up.
- The paper of interest here is a part of this **broad agenda**.



## This paper...

- Financial asset returns are well-known to exhibit **stochastic volatility**
- But, there is also strong evidence in favor of **time-varying shape** in return distributions (at least under the risk-neutral measure):
  - skewness
  - kurtosis
- This can be seen from variation across time in shape of Black-Scholes **implied volatility smiles**.






## Goals

Address the following issues:

- Is there also evidence of **time-varying shape** of return distributions under the **physical measure**?
- How should we **model** the factor(s) responsible for time-varying skewness?
- Does time-variation in shape of physical measure have **explanatory power** for variation in risk-neutral measure?
  - Or, alternatively, is variation in risk-neutral measure due largely (or entirely) to changes in **risk premia**? (e.g., due to **supply and demand** for options, independent of expectations for dynamics of underlying.)
- Develop **techniques** to analyze multi-factor, non-affine models
  - fit
  - assess

**Note:** This paper is part of a long-term research agenda designed to address related issues.



## Shortcomings in existing models — part 1

Existing models used to fit physical and risk-neutral measures simultaneously typically include only a **single state variable**, volatility.

- Given value of state variable, **everything** about volatility surface is known:
  - Level
  - Slope
  - Curvature
  - Term structure
- But, empirically, changes in these features are **not perfectly correlated**.
- Compare this to models of **term structure** of interest rates, which typically include at least **three states** (level, slope, curvature).

**Note:** Can add in as many different jumps and/or risk premia as you want, basic problem remains...

## Shortcomings in existing models — part 2

Existing work almost always uses **affine models**:

- Closed form option pricing formulas
- But, do these models fit the data?

**Note:** Again, one can add in as many different jumps and/or risk premia as you want, basic problem remains...



## Model — starting point

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and information filtration $\{\mathcal{F}_t\}$, the ex-dividend stock price, $x_t$, is assumed to evolve as

$$dx_t/x_t = \left[\mu - \overline{\mu}_{1Jt}\lambda\right]dt + \exp(v_t/2)dW_{1t} + (e^{J_{1t}} - 1)dN_t$$
$$dv_t = \left[k(\overline{v} - v_t) - \overline{\mu}_{2Jt}\lambda\right]dt + \sigma(s_t)dW_{2t} + J_{2t}dN_t$$

$W_{1t}$ and $W_{2t}$ are standard Brownian motions with correlation $\rho$. $N_t$ is a Poisson process with intensity $\lambda$.

**Note:** that the model includes jumps in both returns and volatility. Let $\overline{\mu}_{1Jt} = E(e^{J_{1t}} - 1)$ and $\overline{\mu}_{2Jt} = E(J_{2t})$ denote the mean jump sizes.



## Model — discussion

Suppose that one wanted to think about **extensions** of this model with potential to explain **variation in shape** of implied volatility surface:

- Stochastic **volatility of volatility**
- Stochastic **leverage effect** (correlation in return and volatility innovations)
- Time-varying **jump dynamics**
  - affects shape of return distributions, but primarily at short horizons
- Additional **volatility factor**
  - primarily affects term structure

We focus on stochastic **volatility of volatility** and **leverage effect**.



## Model — proposed extension

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and information filtration $\{\mathcal{F}_t\}$, the ex-dividend stock price, $x_t$, is assumed to evolve as

$$dx_t/x_t = \left[\mu - \overline{\mu}_{1Jt}\lambda_1\right]dt + \exp(v_t/2)dW_{1t} + (e^{J_{1t}} - 1)dN_{1t}$$
$$dv_t = \left[k(\overline{v} - v_t) - \overline{\mu}_{2Jt}\lambda_1\right]dt + \sigma(s_t)dW_{2t} + J_{2t}dN_{1t}$$
$$ds_t = (1 - 2s_t)dN_{2t}$$

where

- $v_t$ and $s_t$ are the volatility state and the regime state, respectively (the regime state is either 0 or 1).
- $W_{1t}$ and $W_{2t}$ are standard Brownian motions with regime-dependent correlation $\rho(s_t)$.
- $N_{1t}$ and $N_{2t}$ are Poisson processes with intensity $\lambda_1$ and $\lambda_2(s_t)$, respectively.

**Notes:**

- The model allows for regime-switching in **volatility of volatility** and **leverage effect**.
- Regime dependence of $\lambda_2$ lets regimes differ in **persistence**.
- Could also allow for regime switching in **jump dynamics** (we looked at such models but did not find them to be useful... ).
- Still only include **single volatility factor**. Make no effort to capture term structure effects.

## Why regime switching?

- Regime switching is the **simplest possible framework** allowing for variation in parameters of interest.
- It also turns out to be **sufficient** to capture features of interest ....
- We are not arguing that this should be taken too literally. Could also look at models with **continuous state spaces**. But,...
  - would have to fully specify dynamics, interactions, and risk-premia
  - analysis is less transparent



## Model — log returns (physical)

It is often useful to transform the model into log prices, $y_t = \log(x_t)$.

Under the $P$-measure

$$\begin{aligned} dy_t &= \left[\mu - \overline{\mu}_{1Jt}\lambda_1 - \frac{1}{2}\exp(v_t)\right]dt + \exp(v_t/2)dW_{1t} + J_{1t}dN_{1t} \\ dv_t &= \left[k(\overline{v} - v_t) - \overline{\mu}_{2Jt}\lambda_1\right]dt + \sigma(s_t)dW_{2t} + J_{2t}dN_{1t} \\ ds_t &= (1 - 2s_t)dN_{2t}. \end{aligned}$$

where everything else is as before.



## Model — log returns (risk-neutral)

Under the $Q$-measure,

$$\begin{aligned} dy_t &= \left[r_t - q_t - \overline{\mu}_{1Jt}\lambda_1 - \frac{1}{2}\exp(v_t)\right]dt + \exp(v_t/2)dW_{1t}^{\mathbb{Q}} + J_{1t}^{\mathbb{Q}}dN_{1t}^{\mathbb{Q}} \\ dv_t &= \left[k(\overline{v} - v_t) - \eta(s_t)v_t - \overline{\mu}_{2Jt}\lambda_1\right]dt + \sigma(s_t)dW_{2t}^{\mathbb{Q}} + J_{2t}^{\mathbb{Q}}dN_{1t}^{\mathbb{Q}} \\ ds_t &= (1 - 2s_t)dN_{2t}^{\mathbb{Q}} \end{aligned}$$

where $r_t$ and $q_t$ denote the risk-free rate and the dividend rate, respectively.



## Model — Jump forms

**Unscaled jump model (UJ):** jump innovations are identically distributed across time,

$$\begin{aligned} J_{1t} &\sim N\left(\mu_{1J}, \sigma_{1J}^2\right) \\ J_{2t} &\sim N\left(\mu_{2J}, \sigma_{2J}^2\right) \\ \mathrm{corr}(J_{1t}, J_{2t}) &= \rho_J. \end{aligned}$$

(This form has been commonly used in the existing literature.)

**Scaled jump model (SJ):** jumps scale in proportion to the volatility of the diffusion component of the process.

$$\begin{aligned} J_{1t}/\exp(v_t/2) &\sim N\left(\mu_{1J}, \sigma_{1J}^2\right) \\ J_{2t}/\sigma(s_t) &\sim N\left(\mu_{2J}, \sigma_{2J}^2\right) \\ \mathrm{corr}(J_{1t}, J_{2t}) &= \rho_J. \end{aligned}$$

By generating larger jumps when volatility is higher, the SJ model is potentially capable of providing more realistic dynamics.

## Model — discussion of risk-premia

- Bias between risk neutral and physical volatility. Can be be explained by
  - volatility risk premium, or
  - jump risk premium.
- In order to disentangle these requires information about
  - Term structure (of volatility), or
  - Shape of return distributions.
- We make no attempt to separately identify jump risk premium and volatility risk premium.
  - Account for bias by volatility risk premium. Absorbs potential jump risk premium.
  - See Pan (JFE, 2002) for additional details...

**Note:** But we do allow for different risk premia depending on regime state.



## Study design

**Data:**

- Use daily observations of **SPX index** and **option prices**.
- Compute daily time-series for
  - option-implied **volatility**
  - option-implied **skewness**
  - option-implied **kurtosis**

**Step 1:**

- Fit models using **only** information from **returns** and **implied volatility**
  - Maximum likelihood estimation
  - Volatility state backed out from option-implied volatility
  - Implied skewness and kurtosis are **withheld** from estimation
- Model comparisons (**likelihood-based**)
- Look at diagnostics to assess model fit (based on idea of **generalized residuals**).



## Study design — continued

**Step 2:**

- Use nonlinear filter to back out **implied regime states**.
- **Regress** option-implied skewness and kurtosis on states (and some control variables) to see if the models have **explanatory power**
  - i.e., do changes in characteristics of physical dynamics help explain changes in shape of **implied volatility smile**.

**Note:** This is a **meaningful diagnostic** since skewness/kurtosis are not used in estimation...



## Questions

A few **totally random** questions which I have **never seen** before taken from workshop participants chosen entirely at **random**...

## Question 1

You have all these great option price data which are hugely informative about skewness/kurtosis. Why don't you use them in fitting the `$#&!!@` model?



## Question 1 — discussion

- Given observations on e.g. implied **volatility and** implied **skewness**:
  - Essentially **any** two-factor model can fit both **exactly** on a day-by-day basis.
  - Even if the model is **badly misspecified**.
- But unclear if implied dynamics are actually present under physical model or just **artifacts** of forcing the model to fit the shape of the implied volatility smile.
- Differences between risk-neutral and physical measures are typically attributed to **risk premia**.
- But, if the model is **misspecified**, the risk premia "discovered" in this way will also be just artifacts.



## Question 1 — discussion (continued)

To **illustrate**:

- Suppose (purely hypothetically) that time-variation in the shape of return distributions is due to stochastic **leverage effect**.
- But, now suppose that we used observations of option-implied **volatility** and **skewness** to fit (incorrectly) a model with time-varying **jump dynamics** (say, jump intensity) but not leverage effect.
- One would find that the model could **perfectly match** the observed values of option-implied volatility and skewness. (If the only tool you have is a hammer, everything looks like a nail...).
- One would find:
  - Strong evidence of time-varying jump intensity in the risk-neutral measure.
  - But, no evidence of time-varying jump intensity in the physical measure.

  So the effect would be attributed entirely to time-varying risk premia.
- The exercise would
  - generate results that are **false**
  - errors that are **difficult to diagnose**.



## Question 1 — discussion (continued)

Our approach is to **withhold** information on option-implied skewness when fitting the model (step 1).

We can then **test** whether **forecasts** from our models are consistent with what is **observed** (step 2).

## Question 2

Do you do the standard sort of option pricing exercise? Do your models lead to improved option pricing performance?



## Question 2 — discussion

We do not make any effort to demonstrate potential improvements in fitting observed option prices.

- To do this, one would want to use the full panel of observed option prices to back out implied states.
- Model can fit, e.g., both implied volatility and skewness on a day-by-day basis.
- Fit to option prices improves correspondingly.
- But, this is not the point of this paper...



## Summary of results — part 1

- Including **jumps in returns** and **volatility** is important (as is already known).
- But, allowing for time-variation (**regime switching**) in stochastic volatility of volatility and leverage effect are also important.
- The best model includes regime-switching in both.
  - improvement in log likelihood is around **118 points** relative to model without regime switching.
  - likelihood ratio test indicates a $p$-value of around $10^{-48}$. (Typically considered to be **significant**...).
  - other **diagnostics** are also improved.



## Summary of results — part 2

- **Regressions** show strong evidence of **explanatory power**.
- Best model includes states from **both** regime-switching in **volatility of volatility** and **leverage effect**.
  - **Slope coefficients** are **large** and in **expected directions** (high volatility of volatility and strong leverage effect are each associated with more skewed and leptokurtotic return distributions under risk-neutral measure).
  - Vol of vol and leverage effect states each provide **independent** sources of information.
  - $R^2$ is **over 32%** (compared to 9% for control variables alone).
  - $t$-statistics are greater than 11 (in absolute value), corresponding to $p$-values of around $10^{-27}$.
  - Results are both **economically** and **statistically significant**.

## Conclusions

- We do indeed find evidence of characteristics associated with time-varying shape of returns distribution under **physical measure**.
- These characteristics do have strong explanatory power for time-varying shape of returns under **risk-neutral measure**.
- We can **reject** the idea that time-variation in shape of Black-Scholes implied volatility smile is due entirely to **changes in risk-premia**.



## Directions for future work

- Reverse process: is information in shape under risk-neutral measure useful for understanding dynamics under physical measure?
- Fit option prices using full information.
- Multi-factor models to fit term-structure and slope/curvature of smile.
  - Need at least three states to capture basic features.

**Notes:**

- All of this using log volatility (or CEV) models (since affine models don't fit data).
- Mixture of normals rather than jumps?



## Details

- Option-implied volatility, skewness and kurtosis
- Backing out volatility state
- Backing out regime state
- Maximum-likelihood estimation
- Model comparison (likelihood-based)
- Generalized residuals
- Diagnostics
- Regressions



## Option-implied volatility, skewness and kurtosis

- Given a panel of option-prices across moneyness (with fixed time to maturity) it should be possible to extract the full **risk-neutral return density** (matching time to maturity).
- We use approach of Bakshi, Kapadia and Madan (2003) to compute option-implied volatility, skewness and kurtosis for **log returns**.

  E.g., for **volatility**:

$$V(t,T) = \int_S^\infty \frac{2(1-\log(K/S))}{K^2} C(t,T;K)\,dK + \int_0^S \frac{2(1-\log(K/S))}{K^2} P(t,T;K)\,dK$$

- Use **linear interpolation** based on two closest times to maturity greater than 8 days to get **constant 30-day maturity** series.

## Backing out volatility

Given a model, candidate parameter vector (risk-neutral), regime state, and **observed option-implied volatility** we need to **back out the volatility state**.

This is a simple **scalar mapping**. Just need to evaluate the mapping at a few points and use standard techniques for **functional approximation** (curve fitting).

**Procedure**

- Let $\widehat{SV}_1, \ldots, \widehat{SV}_G$ be a grid of possible values for **spot volatility**.
- For each, compute $\widehat{IV}_1, \ldots, \widehat{IV}_G$, the corresponding value for **implied volatility** (at 30 days under risk neutral measure). Use **brute force** Monte Carlo.
- Now, we have a collection of pairs $\{(IV_g, SV_g\}_{g=1}^{G}$. It is just a matter of fitting a curve.

**Notes:**

- Many possible curve-fitting/interpolation schemes are possible. We find a simple global (cubic) polynomial scheme to work well.
- That there are two possible sources of approximation error. One should check that both are negligible (they are).
  - Monte Carlo error
  - interpolation error



## Figure — IV to SV mapping



## Backing out regime state

Given an observed value of $IV_t$, we now have **two possible values** for $SV_t$ (and $v_t$), one for each regime state.

Need to compute

$$p_t^j = p(s_t = j|\mathcal{F}_t) = p(v_t = v_t^j|\mathcal{F}_t), \qquad j = 1, 2$$

where $v_t^j$ denotes the volatility state corresponding to regime $j$.

**Filter** can be constructed recursively using **standard techniques**.

**Procedure**

- Let $p_t = (p_t^0, p_t^1)'$ for each $t = 0, \ldots, n$.
- Initialize by setting $p_0^j$ equal to the marginal probability of state $j$ $(j = 0, 1)$.
- Now, suppose that $p_t$ is known. The problem is to compute $p_{t+1}$. This is given by (for $j = 0, 1$)

$$p_{t+1}^j = \frac{\sum_{i=0}^{1} p(y_{t+1}, v_{t+1}^j|y_t, v_t^i, s_t = i) \cdot p(s_{t+1} = j|s_t = i) \cdot p_t^i}{\sum_{k=0}^{1}\sum_{i=0}^{1} p(y_{t+1}, v_{t+1}^k|y_t, v_t^i, s_t = i) \cdot p(s_{t+1} = k|s_t = i) \cdot p_t^i}$$

**Note:** Sometimes speak of **filtered regime state**: the expected value of $s_t$ conditional on information available at time $t$,

$$\hat{s}_t = E_t(s_t) = p_t^0 \cdot 0 + p_t^1 \cdot 1 = p_t^1.$$



## Maximum likelihood estimation

- Having backed out volatility states and computed filtered regime state probabilities, computing the **log likelihood** is straightforward:

$$\log L(\{y_t\}_{t=1}^n, \{IV_t\}_{t=1}^n; \theta) \approx \sum_{t=1}^{n-1}\sum_{i=0}^{1}\sum_{j=0}^{1}\Big[\log p(y_{t+1}, v_{t+1}^j|y_t, v_t^i, s_t = i) + \log p(s_{t+1} = j|s_t = i) + \log p(s_t = i) + \log J_{t+1}^j\Big], \tag{1}$$

  where $J_{t+1}^j = \left|dv_{t+1}^j / dIV_{t+1}\right|$ is the Jacobian corresponding to regime state $j$.

- Parameter estimates are obtained by numerical optimization,

$$\widehat{\theta} = \arg\max \log L(\{y_t\}_{t=1}^n, \{IV_t\}_{t=1}^n; \theta).$$

  We use a BHHH optimizer, but the criterion function is well-behaved and nearly any optimizer will work fine.

| | No regime switching | | | RV model | | RL model | | RVL model | |
|---|---|---|---|---|---|---|---|---|---|
| | SV | SJ | UJ | SJ-RV | UJ-RV | SJ-RL | UJ-RL | SJ-RVL | UJ-RVL |
| $\mu \times 10^4$ | 4.69 | 3.70 | 3.08 | 3.03 | 2.07 | 3.42 | 3.08 | 3.00 | 2.08 |
| | (1.27) | (1.28) | (1.25) | (1.25) | (1.25) | (1.26) | (1.23) | (1.25) | (1.25) |
| $\kappa \times 10^3$ | 4.29 | 6.00 | 6.69 | 6.34 | 11.07 | 7.48 | 7.36 | 6.27 | 11.04 |
| | (1.29) | (1.35) | (1.33) | (1.32) | (1.28) | (1.35) | (1.34) | (1.32) | (1.28) |
| $\overline{v}$ | -9.71 | -10.03 | -9.70 | -10.06 | -9.63 | -10.00 | -9.80 | -9.98 | -9.63 |
| | (0.55) | (0.38) | (0.36) | (0.32) | (0.20) | (0.32) | (0.33) | (0.33) | (0.20) |
| $\eta^0 \times 10^3$ | 3.01 | 2.84 | 3.84 | 2.86 | 4.18 | 2.49 | 2.72 | 3.01 | 4.20 |
| | (0.25) | (0.36) | (0.35) | (0.42) | (0.36) | (0.47) | (0.44) | (0.37) | (0.36) |
| $\eta^1 \times 10^3$ | | | | 2.97 | 2.46 | 3.39 | 4.27 | 2.66 | 2.46 |
| | | | | (0.59) | (0.41) | (0.40) | (0.38) | (0.54) | (0.42) |
| $\sigma^0 \times 10^1$ | 1.32 | 0.91 | 1.23 | 0.84 | 1.18 | 0.99 | 1.25 | 0.85 | 1.18 |
| | (0.02) | (0.03) | (0.03) | (0.03) | (0.03) | (0.03) | (0.03) | (0.03) | (0.03) |
| $\sigma^1 \times 10^1$ | | | | 1.33 | 2.12 | | | 1.33 | 2.12 |
| | | | | (0.06) | (0.06) | | | (0.05) | (0.06) |
| $\rho^0$ | -0.74 | -0.78 | -0.82 | -0.81 | -0.91 | -0.53 | -0.61 | -0.77 | -0.91 |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.04) | (0.03) | (0.01) | (0.01) |
| $\rho^1$ | | | | | | -0.82 | -0.85 | -0.87 | -0.91 |
| | | | | | | (0.01) | (0.01) | (0.02) | (0.01) |
| $\rho_J$ | | -0.71 | -0.53 | -0.70 | -0.05 | -0.77 | -0.72 | -0.69 | -0.04 |
| | | (0.03) | (0.04) | (0.03) | (0.10) | (0.03) | (0.04) | (0.03) | (0.10) |
| $\lambda$ | | 0.47 | 0.33 | 0.55 | 1.08 | 0.30 | 0.15 | 0.46 | 1.08 |
| | | (0.06) | (0.04) | (0.09) | (0.14) | (0.04) | (0.02) | (0.07) | (0.15) |
| $\mu_{1J} \times 10^2$ | | -1.13 | 0.05 | -4.52 | 0.08 | -11.90 | -0.05 | -3.78 | 0.08 |
| | | (6.59) | (0.04) | (6.38) | (0.02) | (8.78) | (0.07) | (6.98) | (0.02) |
| $\sigma_{1J} \times 10^1$ | | 12.94 | 0.06 | 12.44 | 0.03 | 13.42 | 0.08 | 12.30 | 0.03 |
| | | (0.75) | (0.00) | (0.77) | (0.00) | (0.85) | (0.01) | (0.77) | (0.00) |
| $\mu_{2J} \times 10^1$ | | 2.79 | 0.24 | 3.39 | 0.09 | 3.59 | 0.44 | 3.04 | 0.08 |
| | | (0.69) | (0.07) | (0.66) | (0.04) | (0.95) | (0.13) | (0.70) | (0.04) |
| $\sigma_{2J}$ | | 1.54 | 0.15 | 1.15 | 0.05 | 1.72 | 0.21 | 1.25 | 0.05 |
| | | (0.07) | (0.01) | (0.07) | (0.01) | (0.09) | (0.01) | (0.07) | (0.01) |
| $\pi_0$ | | | | 0.98 | 0.98 | 0.97 | 0.96 | 0.99 | 0.98 |
| | | | | (0.01) | (0.01) | (0.01) | (0.01) | (0.00) | (0.01) |
| $\pi_1$ | | | | 0.94 | 0.94 | 0.99 | 0.98 | 0.96 | 0.94 |
| | | | | (0.01) | (0.01) | (0.00) | (0.01) | (0.01) | (0.01) |
| log(L) | 38,491 | 38,851 | 38,810 | 38,961 | 38,960 | 38,883 | 38,858 | 38,969 | 38,960 |



## Remarks

- Jumps are important (about 360 points in log likelihood).
- SJ better than UJ (about 40 points in log likelihood)
- Regime switching is also important
  - **volatility in volatility** is better than leverage effect
  - best model includes **both features**
  - improvement of around **118 points** in log likelihood relative to non-regime switching model



## Model comparison

- To the extent that models are nested, can use **likelihood ratio** tests.

  E.g., best regime-switching model rejects non-regime-switching counterpart with a $p$-value of around $10^{-48}$.

- Alternatively, one could use **information criteria** (**AIC**, **BIC**, etc). Results are **sufficiently clearcut** that it doesn't make much difference which ...



## Diagnostics — generalized residuals

- Let $\{z_t\}_{t=1}^n$ be a sequence of random vectors where $z_t$ has distribution $G_t(z|\mathcal{F}_{t-1})$.
- Let $u_t = G_t(z_t|\mathcal{F}_{t-1})$ $(t = 1, \ldots, n)$.
- If the model is correctly specified, $\{u_t\}$ should be iid uniform$(0, 1)$.
- The hypothesis that $\{G_t(z_t|\mathcal{F}_{t-1}; \theta)\}$ is the true data generating process for $\{z_t\}$ can be tested by performing diagnostics on $\{u_t\}$.
- But, it is often more useful to instead perform diagnostics on

$$\widetilde{u}_t = \Phi^{-1}(u_t), \qquad t = 1, \ldots, n \tag{2}$$

  where $\Phi$ is the standard normal distribution function.

- In this case, the transformed residuals $\{\widetilde{u}_t\}$ should be iid standard normal under the hypothesis of correct model specification.

**Note:** We will always use the transformed residuals in this paper.

## Generalized residuals — continued

- For the models in this paper, the generalized residuals are computed in a manner similar to the log likelihood (equation (1)),

$$u_{t+1} = \sum_{i=0}^{1}\sum_{j=0}^{1} P(y_{t+1}, v_{t+1}^{j} | y_t, v_t^i, s_t = i) \cdot p(s_{t+1} = j | s_t = i) \cdot p_t^i,$$

  where $P(\cdot)$ denotes a cdf.

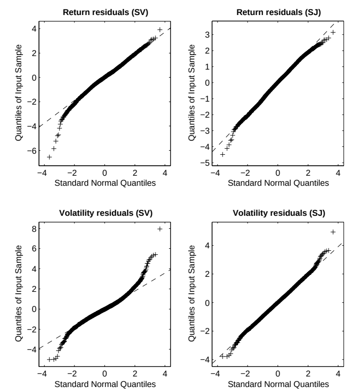
- These residuals correspond to the **joint distribution** of price and volatility innovations.

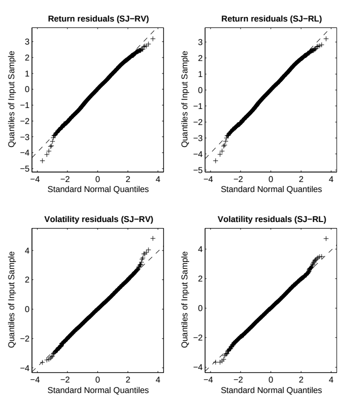
- Although one could certainly study these, we have found it more useful to study **marginal residuals** corresponding to **price** and **volatility** innovations separately,

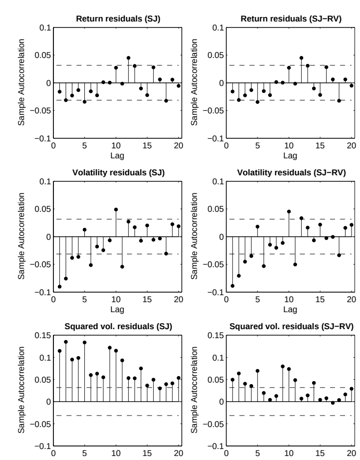
$$u_{y,t+1} = \sum_{i=0}^{1} P(y_{t+1} | y_t, v_t^i, s_t = i) \cdot p(s_t = i)$$

$$u_{v,t+1} = \sum_{i=0}^{1}\sum_{j=0}^{1} P(v_{t+1}^{j} | y_t, v_t^i, s_t = i) \cdot p(s_{t+1} = j | s_t = i) \cdot p(s_t = i).$$

- Testing can proceed using **standard time series techniques**:
  - **Normality** (QQ-plots, Jarque-Bera tests)
  - **Independence** (correlograms, Ljung-box tests)



## Table — Diagnostics

**Jarque-Bera Test**

| | No regime switching | | | RV model | | RL model | | RVL model | |
|---|---|---|---|---|---|---|---|---|---|
| | SV | SJ | UJ | SJ-RV | UJ-RV | SJ-RL | UJ-RL | SJ-RVL | UJ-RVL |
| Return | 445 | 20 | 113 | 17 | 245 | 10 | 50 | 19 | 246 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.006) | (0.000) | (0.000) | (0.000) |
| Volatility | 2,367 | 22 | 121 | 20 | 318 | 8 | 30 | 15 | 320 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.018) | (0.000) | (0.001) | (0.000) |

**Ljung-Box Test (with 20 lags)**

| | No regime switching | | | RV model | | RL model | | RVL model | |
|---|---|---|---|---|---|---|---|---|---|
| | SV | SJ | UJ | SJ-RV | UJ-RV | SJ-RL | UJ-RL | SJ-RVL | UJ-RVL |
| Return | 43 | 41 | 41 | 41 | 40 | 41 | 40 | 41 | 40 |
| | (0.002) | (0.004) | (0.004) | (0.004) | (0.005) | (0.004) | (0.005) | (0.004) | (0.005) |
| Volatility | 116 | 118 | 116 | 115 | 100 | 115 | 114 | 116 | 100 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Squared Vol. | 419 | 552 | 424 | 128 | 74 | 559 | 446 | 118 | 74 |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |



## Explanatory power for option-implied skewness and kurtosis

Look at following regressions:

$$\text{SKEW}_t = \beta_0 + \beta_{\text{RV}}\hat{s}_t^{\text{RV}} + \beta_{\text{RL}}\hat{s}_t^{\text{RL}} + \beta_{\text{VIX}}\log\text{VIX}_t + \beta_{\text{VRP}}\text{VRP}_t + \beta_{\text{JV}}\log\text{JV}_t + \varepsilon_t$$
$$\text{KURT}_t = \beta_0 + \beta_{\text{RV}}\hat{s}_t^{\text{RV}} + \beta_{\text{RL}}\hat{s}_t^{\text{RL}} + \beta_{\text{VIX}}\log\text{VIX}_t + \beta_{\text{VRP}}\text{VRP}_t + \beta_{\text{JV}}\log\text{JV}_t + \varepsilon_t$$

where

- $\hat{s}_t^{\text{RV}}$ denotes the filtered state under the models with regime switching in volatility of volatility.
- $\hat{s}_t^{\text{RL}}$ denotes the filtered state under the models with regime switching in leverage effect.
- $\text{VIX}_t$ is the VIX index, which serves as a proxy for the volatility state.
- $\text{VRP}_t$ is the variance risk premium
- $\text{JV}_t$ is past one-month jump variation, which proxies for jump risk.

**Notes:**

- Recall that we **do not use** the information from observed option-implied skewness and kurtosis when fitting the models (back out regime states using **only returns** and **volatility**.
- This allows to **test** whether the implied states have **explanatory power** for option-implied skewness and kurtosis.



## Table — skewness regressions

| Constant | | $\hat{s}_t^{\text{RV}}$ | | $\hat{s}_t^{\text{RL}}$ | | $\log\text{VIX}_t$ | | $\text{VRP}_t$ | | $\log\text{JV}_t$ | | Adj. $R^2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| coeff | t-stat | coeff | t-stat | coeff | t-stat | coeff | t-stat | coeff | t-stat | coeff | t-stat | |
| **Control variables only** | | | | | | | | | | | | |
| -2.78 | (-16.43) | | | | | 0.37 | (6.53) | | | | | 7.6% |
| -3.38 | (-12.02) | | | | | 0.60 | (5.94) | -0.10 | (-3.21) | | | 9.2% |
| -2.75 | (-13.84) | | | | | 0.36 | (4.97) | | | 0.000 | (-0.32) | 7.6% |
| -3.56 | (-9.34) | | | | | 0.67 | (4.65) | -0.11 | (-3.37) | 0.001 | (0.74) | 9.3% |
| **Control variables plus filtered states from SJ models** | | | | | | | | | | | | |
| -2.82 | (-18.07) | -0.56 | (-9.02) | | | 0.44 | (8.48) | | | | | 16.2% |
| -3.63 | (-14.57) | -0.61 | (-9.98) | | | 0.74 | (8.42) | -0.13 | (-5.01) | | | 19.0% |
| -2.76 | (-16.78) | -0.57 | (-9.06) | | | 0.41 | (7.29) | | | -0.001 | (-1.10) | 16.3% |
| -3.80 | (-12.12) | -0.61 | (-9.92) | | | 0.81 | (7.00) | -0.14 | (-5.07) | 0.001 | (1.00) | 19.1% |
| -2.82 | (-15.39) | | | -0.62 | (-7.37) | 0.54 | (7.15) | | | | | 17.5% |
| -3.25 | (-12.37) | | | -0.60 | (-7.17) | 0.69 | (7.10) | -0.07 | (-2.23) | | | 18.3% |
| -2.73 | (-11.68) | | | -0.63 | (-8.12) | 0.50 | (5.11) | | | -0.001 | (-0.80) | 17.7% |
| -3.22 | (-8.26) | | | -0.60 | (-7.89) | 0.68 | (4.43) | -0.07 | (-2.03) | 0.000 | (-0.11) | 18.3% |
| -2.87 | (-18.05) | -0.71 | (-11.53) | -0.76 | (-10.66) | 0.65 | (10.36) | | | | | 30.6% |
| -3.52 | (-16.28) | -0.74 | (-12.04) | -0.74 | (-10.77) | 0.89 | (11.23) | -0.10 | (-4.02) | | | 32.3% |
| -2.73 | (-14.90) | -0.72 | (-11.55) | -0.78 | (-11.90) | 0.59 | (8.15) | | | -0.002 | (-1.89) | 31.1% |
| -3.44 | (-11.21) | -0.74 | (-11.96) | -0.74 | (-11.69) | 0.86 | (7.28) | -0.10 | (-3.59) | -0.001 | (-0.39) | 32.3% |



## Skewness regressions — remarks

- **Control variables** have some explanatory power, but together only achieve an $R^2$ of 9.3%.
- SJ models always **outperform** UJ models.
- **Volatility of volatility** does slightly better than **leverage effect** ($R^2$ of 19% and 18.3% respectively).
- Slope coefficients are **large** and in **expected direction**.
- Best model includes **both states**, with an $R^2$ of over 32% (**leverage** state and **vol of vol** state each provide **independent sources of information**.
- Slope coefficients are around -0.7 with $t$-statistics of over 11 (in absolute value), corresponding to $p$-values of around $10^{-27}$.
- Results are **economically** and **statistically significant**.

## Regressions — kurtosis

Results for kurtosis are weaker but qualitatively similar.



## Conclusions

We have

- Demonstrated some **tools** and **techniques**.
- Obtained some results.

Lots more that can be done ...