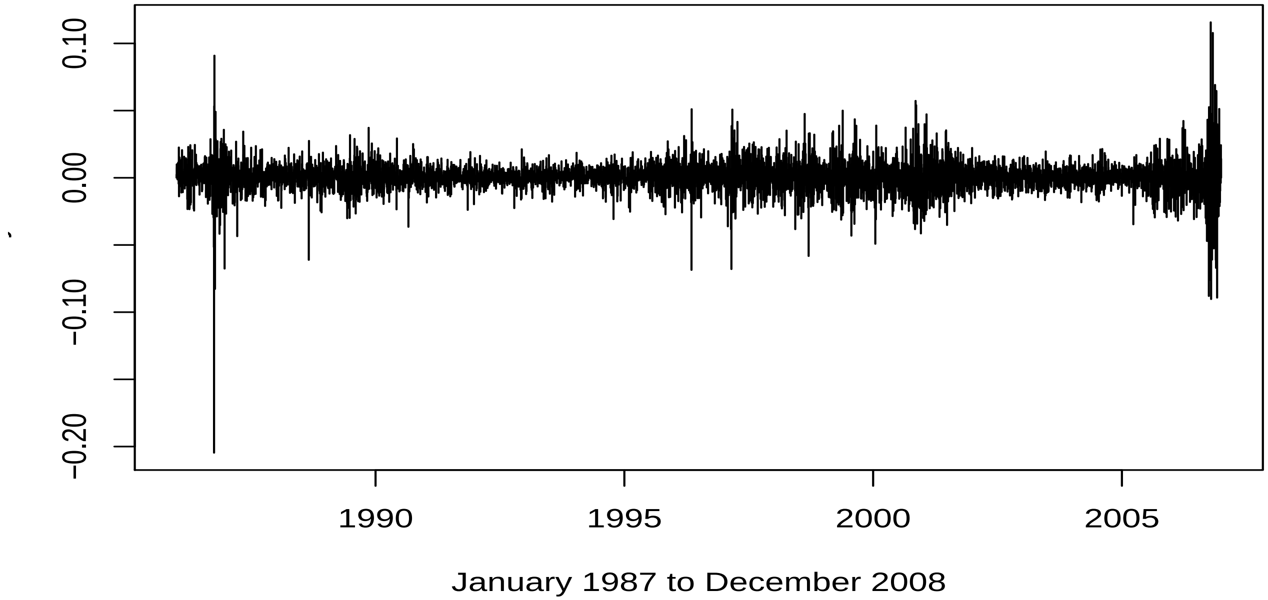
The Contribution of Jumps to the Volatility of Asset Prices

James E. Gentle

Johan Bjursell

George Mason University
Contact: jgentle@gmu.edu

The Obligatory Plot: S&P 500 Rate of Return That Includes October 19, 1987



Stylized Properties of Financial Data

Some of the characteristics of financial data that make it interesting and challenging to analyze are the following.

- Aggregational normality.
- Quasi long range dependence.
- Seasonality (sometimes).
- Asymmetry in rates of return. Rates of return are slightly negatively skewed.
- Asymmetry in lagged correlations. Coarse volatility predicts fine volatility better than the other way around.

Stylized Properties of Financial Data (Continued)

These are the real biggies.

- **Heavy tails**. The frequency distribution of rates of return decrease more slowly than $\exp(-x^2)$.
  - Can an outlier-generating distribution model it well?

- The **volatility** is **not constant**.
  - Also, the volatility clusters.

So how to model this?

The Problem with Models

We don't care whether the models are correct.

We want models that are useful.

Models that we can fit to data.

Models that give us good predictions.

Models should match observational data.

Time in Models of Financial Data

Time in financial models can be treated either as a continuous or as a discrete variable.

In any real analysis, ultimately, of course, it is discrete; the difference is in the basic types of the models.

Most standard time series models, which usually derive from a basic ARIMA approach, treat time as a discrete variable.

Other models treat financial data as discrete points in a continuous diffusion process. This approach allows us to study data at any frequency. In active markets, data are generated at a frequency arbitrarily high.

Diffusion Models

A good general model of a random walk is a Brownian motion, or a Wiener process; hence, we may write the model as a drift and diffusion,

$$\frac{\mathrm{d}S(t)}{S(t)} = \mu(S(t),t)\mathrm{d}t + \sigma(S(t),t)\mathrm{d}B,$$

where $\mathrm{d}B$ is a Brownian motion, which has a standard deviation of 1. The standard deviation of the rate of change is therefore $\sigma(S(t),t)$.

In this equation both the mean drift and the standard deviation may depend on both the magnitude of the price $S(t)$ and also on the time t itself.

Simple Diffusion Model

We often assume that $\mu(\cdot)$ and $\sigma(\cdot)$ do not depend on the value of the state and that they are constant in time, we have the model

$$\frac{\mathrm{d}S(t)}{S(t)} = \mu \mathrm{d}t + \sigma \mathrm{d}B.$$

A "geometric Brownian motion".

In this form,
μ is called the drift and the diffusion component and
σ is called the volatility.

Geometric Brownian Motion

We note that as a model for the rate of return, $\mathrm{d}S(t)/S(t)$ geometric Brownian motion is similar to other common statistical models:

$$\frac{\mathrm{d}S(t)}{S(t)} = \mu\mathrm{d}t + \sigma\mathrm{d}B(t)$$

or

$$\text{response} = \text{systematic component} + \text{random error.}$$

Also, note that without the stochastic component, the differential equation has the simple solution

$$S(t) = c\mathrm{e}^{\mu t},$$

from which we get the formula for continuous compounding for a rate μ.

Whatever .... Brownian Motion Doesn't Fit

Either

- the Gaussian distribution doesn't work

or

- the volatility is not constant

or both.

More Realistic Models

There are various approaches for modifying the simple models.

- We can drop pure Gaussianity.

  Replace with a heavier-tailed distribution.

  Add a mixture.

- Retain Gaussianity, but drop constant volatility.

  Introduce coupled diffusion.

  Use a discrete-time autoregressive model with generalized autoregressive conditional heteroskedasticity (GARCH).

or both ... or all.

Fixing the Models: Coupled Diffusion

A simple model in which the volatility is a function of a separate mean-reverting Ornstein-Uhlenbeck process:

$$
\begin{aligned}
\mathrm{d}X_t &= \mu X_t \mathrm{d}t + \sigma_t X_t \mathrm{d}B_t, \\
\mathrm{d}\sigma_t^2 &= f(Y_t), \\
\mathrm{d}Y_t &= \alpha(\mu_Y - Y_t)\mathrm{d}t + \beta \mathrm{d}\tilde{B}_t,
\end{aligned}
$$

where α and β are constants and $\tilde{B}_t$ is a linear combination of B_t and an independent Brownian motion.

The function f can incorporate various degrees of complexity, including the simple identity function.

These coupled equations provide a better match for observed stock and option prices.

(Well, of course, increased degrees of freedom always yield better fits.)

Fixing the Models: Adding Jumps to Diffusion

A very realistic modification of the model is to assume that Z has a superimposed jump or shock on its $\mathrm{N}(0, 1)$ distribution.

A useful model is

$$\mathrm{d}X_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t + \kappa_t \mathrm{d}q_t(\lambda_t),$$

where $\mathrm{d}q_t(\lambda_t)$ is a Poisson process with intensity λ_t; that is, $\Pr(\mathrm{d}q_t = 1) = \lambda_t \mathrm{d}t$.

We may use a constant value for κ_t or else some reasonable distribution, say a normal with a negative mean.

A preponderance of bear jumps, that is, instances in which $\kappa_t < 0$, of course, would decrease the fair price of a call option from the Black–Scholes price and would increase the fair price of a put option.

Diffusion with Jumps

How simple can we realistically be?

$$\mathrm{d}X_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t + \kappa_t \mathrm{d}q_t(\lambda_t).$$

Can $\mu_t = \mu$?

It may not matter too much.

Can $\sigma_t = \sigma$?

No. Must we model σ_t separately, or can we tie it to $\kappa_t \mathrm{d}q_t(\lambda_t)$?

Can $\mathrm{d}W_t$ be Gaussian?

Maybe with enough modifications on σ_t.

How simple can we realistically be? (continued)

$$\mathrm{d}X_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t + \kappa_t \mathrm{d}q_t(\lambda_t).$$

How carefully must we deal with κ_t?

We cannot have $\kappa_t = \kappa$, but we must realistically limit the degrees of freedom.

What else must we do about exogenous information?
We must realistically limit the degrees of freedom!

Diffusion with Jumps

The question is how far we can go with

$$\mathrm{d}X_t = \mu \mathrm{d}t + \sigma \mathrm{d}W_t + \kappa_t \mathrm{d}q_t(\lambda_t).$$

The modeling questions involve inference on κ_t and $\mathrm{d}q_t(\lambda_t)$.

Keep $\mathrm{d}q_t(\lambda_t)$ Poisson.

Can $\lambda_t = \lambda$? For now, yes.

How complicated must κ_t be? Probably skewed. Probably heavy-tailed, but can that be accounted for by the Poisson process itself?

If we can make any progress on $\kappa_t \mathrm{d}q_t$, maybe then we can go back and consider $\sigma_s = f(\kappa_t \mathrm{d}q_t)$ for $s > t$.

Diffusion with Jumps

We are currently trying to see how far we can go with

$$\mathrm{d}X_t = \mu \mathrm{d}t + \sigma \mathrm{d}W_t + \kappa_t \mathrm{d}q_t,$$

with constant λ, and some simple distribution on κ_t.

The simple questions are:

Is there a simple useful test of the existence of jumps and/or can we estimate the occurrence of a jump and its magnitude?

An Exploratory Investigation

In real data, can we identify jumps?

If we identify them, do we believe they exists?

Background

Theoretical framework: Barndorff-Nielsen and Shephard (2004, 2006), Huang and Tauchen (2005)

Market microstructure noise: Zhang, Mykland, and Aït-Sahalia (2005), Bandi and Russell (2006) and Andersen et al (2007)

Application:

- Equity, foreign exchange, and T-bond markets: Mahue and McCurdy (2004) Andersen, Bollerslev, and Diebold (2007), Jiang, Lo, and Verdelhan (2008)

- Energy futures markets: Linn and Zhu (2004), Pindyck (2004), Ates and Wang (2007), Mu (2007), and Wang, Wu and Yang (2008)

Energy Futures Prices

We study the diffusion with jumps model for the energy futures market.

The models for commodity prices should be similar to those for stock prices, but the exogenous effects are often more understandable.

There is an explainable seasonality.

The Data

Intraday data series for three contracts from the U.S. energy futures markets: Crude Oil, Heating Oil and Natural Gas.

Traded on the New York Mercantile Exchange (NYMEX).

The data for crude oil and heating oil range from January 1, 1990 to December 31, 2007; the natural gas contract span from January 1, 1993 to January 31, 2008.

The datasets include date, time and price per transaction.

We use data from nearby contract months; switch the first deferred to become the nearby contract during the delivery month based on trading frequency.

Observations

Energy futures prices are very volatile and often exhibit jumps (price spikes)

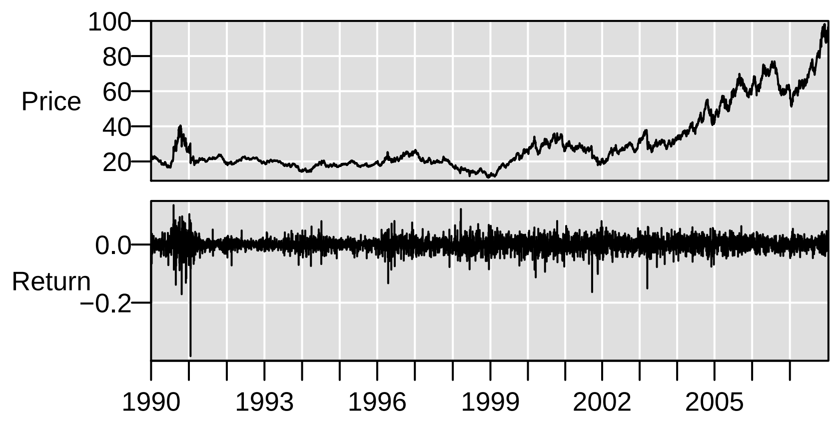
Violates the assumption of a continuous sample path diffusion process for the asset price behavior

Consequently, there are strong interests in identifying the jump process and its relative contribution to the total volatility

The volatility behavior is a central topic in option pricing, risk management and asset allocation strategies

Example: Crude Oil Futures Prices

January 1, 1990 to December 31, 2007



Objectives

This study examines the volatility and jump dynamics in the U.S. energy futures markets.

The main purposes:

- Document the realized volatility in energy futures prices
- Identify significant jump components and estimate the relative contribution of jumps to price volatility
- Investigate seasonal and intraday patterns in the smooth and discontinuous components

more

- Test whether extreme cold weather and low inventory announcements are associated with significant jumps

- Examine whether including jump components as an explanatory variable improve the modeling of the realized variance

Methodology

Consider the stochastic differential equation,

$$\mathrm{d}p(t) = \mu(t)\mathrm{d}t + \sigma(t)\mathrm{d}W(t) + \kappa(t)\mathrm{d}q(t), \qquad (1)$$

where $p(t)$ denotes the logarithmic asset price at time t, $W(t)$ a Wiener process; $\kappa(t)$ the size of a jump, and $q(t)$ a counting process.

We will use the decomposition developed by Barndorff-Nielsen and Shephard (2004, 2006) for identifying days with a discontinuous jump in the price process.

Methodology

The intraday return, $r_{t,j} = p_{t,j\Delta} - p_{t,(j-1)\Delta}$ where Δ is the sampling interval

Realized and bipower variation:

$$\mathrm{RV}_t = \sum_{j=1}^{m_t} r_{t,j}^2 \xrightarrow{p} \int_{t-1}^{t} \sigma_s^2 \mathrm{d}s + \int_{t-1}^{t} \kappa_s^2 \mathrm{d}q_s \quad (2)$$

$$\mathrm{BV}_t = \frac{\pi}{2} \sum_{j=2}^{m_t} |r_{t,j}||r_{t,j-1}| \xrightarrow{p} \int_{t-1}^{t} \sigma_s^2 \mathrm{d}s, \quad (3)$$

as $\Delta \rightarrow 0$. Hence, the differences can be used to measure the variation due to the jump component,

$$\mathrm{RV}_t(\Delta) - \mathrm{BV}_t(\Delta) \xrightarrow{p} \int_{t-1}^{t} \kappa_s^2 \mathrm{d}q_s. \quad (4)$$

Methodology: Test Statistic

A statistic for testing the null hypothesis for no jumps,

$$Z_t = \frac{\mathrm{RJ}_t}{\sqrt{(\nu_{bb} - \nu_{qq})\frac{1}{m_t}\max\left\{1, \frac{\mathrm{TP}_t}{\mathrm{BV}_t^2}\right\}}}, \tag{5}$$

where ν_{bb} and ν_{qq} are constants, TP_t is an estimator $\int \sigma_u^4 \mathrm{d}u$ and,

$$\mathrm{RJ}_t = \frac{\mathrm{RV}_t - \mathrm{BV}_t}{\mathrm{RV}_t}. \tag{6}$$

Methodology: Decomposing the Total Variation

Variance due to the jump component:

$$J_t = \max[\mathrm{RV}_t - \mathrm{BV}_t, 0]. \tag{7}$$

Variance due to the significant jump component:

$$J_{t,\alpha} = (\mathrm{RV}_t - \mathrm{BV}_t)\, \mathrm{I}_{\left(Z \geq \Phi^{-1}_{1-\alpha}\right)}, \tag{8}$$

where Φ is the cumulative distribution function of the standard normal distribution; α is the significance level; and $\mathrm{I}_{\left(Z > \Phi^{-1}_{1-\alpha}\right)}$ is the indicator function which is equal to one if the test rejects the null hypothesis and zero otherwise.

Variance due to the smooth components:

$$C_{t,\alpha} = \mathrm{I}_{\left(\mathrm{Z}_t \leq \Phi^{-1}_{1-\alpha}\right)} \mathrm{RV}_t + \mathrm{I}_{\left(\mathrm{Z}_t > \Phi^{-1}_{1-\alpha}\right)} \mathrm{BV}_t. \tag{9}$$

Methodology: Market Microstructure Noise

Observed prices are noisy due to price formation under specific trade mechanisms and rules, e.g., the discrete price grid, minimum tick size and bid-ask spread.

The realized (RV) and bipower (BV) variations are dominated by the noise variance at high sampling rates.

Methods to reduce the bias:

- Sample at low frequencies (5 to 30-minute sampling)
- Staggered returns (Anderson, Bollerslev and Diebold, 2007)
- Optimal sampling rate (Bandi and Russell, 2005)

Methodology: Market Microstructure Noise

Staggered returns

- Bipower definition:

$$\text{BV}_{i,t} = \frac{\pi}{2}\frac{m_t}{m_t - 1}\sum_{j=2+i}^{m} |r_{t,j}||r_{t,j-(1+i)}|$$

where i is the offset.

- Tripower definition:

$$\text{TP}_t = m_t \mu_{4/3}^{-3} \frac{m_t}{m_t - 2(1+i)} \sum_{j=1+2(1+i)}^{m_t} \prod_{k=0}^{2} |r_{t_{j-k(1+i)}}|^{4/3}.$$

Methodology: Market Microstructure Noise

Optimal Sampling

- Minimizing MSE,

$$\mathsf{E}\left(\sum_{j=1}^{m_t} r_{t_i}^2 - \int_{t-1}^{t} \sigma_s^2 \mathrm{d}s\right)^2 = 2\frac{1}{m_t}\mathsf{TP}_t + m_t\beta + {m_t}^2\alpha + \gamma,$$

where,

$$\begin{aligned} \alpha &= \left(\mathsf{E}\left(\eta_t^2\right)\right)^2, \\ \beta &= 2\mathsf{E}\left(\eta_t^4\right) - 3\left(\mathsf{E}\left(\eta_t^2\right)\right)^2, \\ \gamma &= 4\mathsf{E}\left(\eta_t^2\right)\int_{t-1}^{t} \sigma_s^2 \mathrm{d}s - \mathsf{E}\left(\eta_t^4\right) - 2\left(\mathsf{E}\left(\eta_t^2\right)\right)^2. \end{aligned}$$

The Datasets Again

Intraday data series for three contracts from the U.S. energy futures markets:

A: Crude Oil

B: Heating Oil

C: Natural Gas

Traded on the New York Mercantile Exchange (NYMEX).

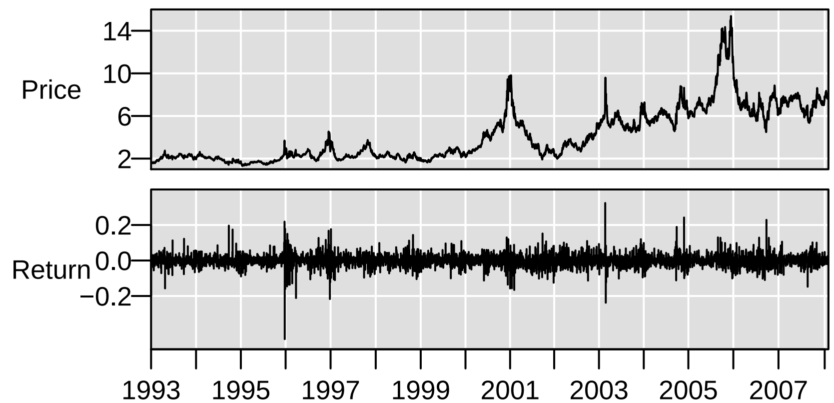
We analyzed these in a variety of ways.

In the following, I present a sampling of the analyses, some for one dataset and others for other datasets.

Panel C: Henry Hub Natural Gas Futures
Trading Unit
10,000 million British thermal units (mmBtu).
Price Quotation
U.S. dollars and cents per barrel
Trading Hours
Open outcry trading is conducted from 9:00 AM until 2:30 PM.
Trading Months
The current year plus the next twelve years through December 2020.
Minimum Price Flucuation
\$0.001 (0.1) per mmBtu (\$10.00 per contract).
Maximum Daily Price Flucuation
\$3.00 per barrel (\$30,000 per contract) for all months.
Last Trading Day
Trading terminates three business days prior to the first calendar day of the delivery month.
Settlement Type
Physical
Delivery
The Sabine Pipe Line Co. Henry Hub in Louisiana. Seller is responsible for the movement of the gas through the Hub; the buyer, from the Hub. The Hub fee will be paid by seller.
Trading Symbol
NG

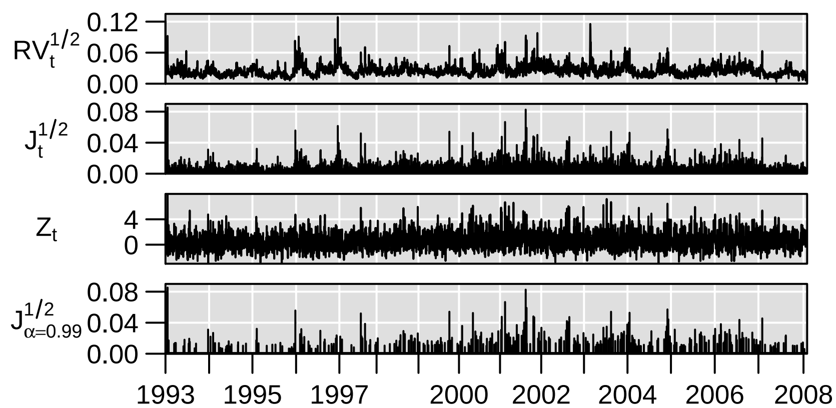
Data - Natural Gas

January 1, 1993 to January 31, 2008



Empirical Results

Time-series plots of the total volatility ($\mathrm{RV}_t^{1/2}$), jump component ($\mathrm{J}_t^{1/2}$), jump statistic (Z_t), and significant jumps ($J_{t,\alpha=0.99}^{1/2}$).



Stationarity - The ADF test

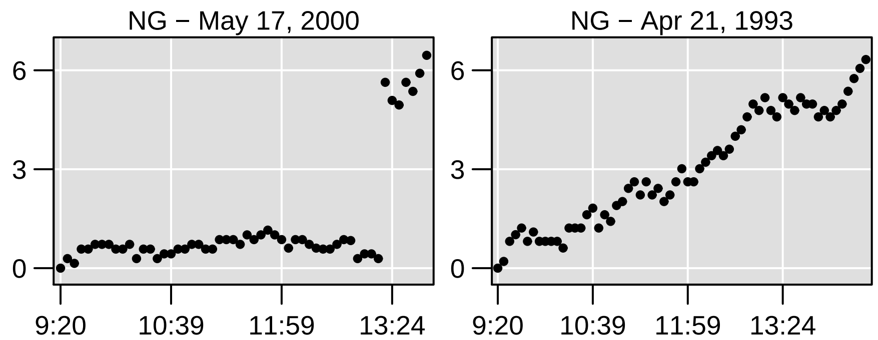
Augmented Dickey-Fuller test of the realized variance, RV. Null Hypothesis: RV has a unit root.

		t-Statistic	Prob.
	Panel A: Crude Oil		
Augmented Dickey-Fuller test statisic		−16.03719	0.0000
Test critical values:	1% level	−3.431619	
	5% level	−2.861986	
	10% level	−2.567050	
	Panel B: Heating Oil		
Augmented Dickey-Fuller test statisic		−15.35070	0.0000
Test critical values:	1% level	−3.431639	
	5% level	−2.861995	
	10% level	−2.567055	
	Panel C: Natural Gas		
Augmented Dickey-Fuller test statisic		−10.90829	0.0000
Test critical values:	1% level	−3.431949	
	5% level	−2.862131	
	10% level	−2.567128	

Summary Statistics

	RV_t	$RV_t^{1/2}$	$\log(RV_t)$	J_t	$J_t^{1/2}$	$\log(J_t+1)$
Panel A: Crude Oil						
Mean	0.0003	0.0164	−8.3774	0.0000	0.0033	0.0000
Std Dev	0.0007	0.0072	0.7718	0.0003	0.0045	0.0003
Skewness	44.3013	4.8741	0.0393	59.8240	7.3385	59.7116
Kurtosis	2534.1413	88.6175	1.0611	3835.8889	175.3440	3825.6311
Min	0.0000	0.0030	−11.6462	0.0000	0.0000	0.0000
Max	0.0381	0.1953	−3.2664	0.0188	0.1370	0.0186
LB_{10}	968	10926	16947	91	283	93
Panel C: Natural Gas						
Mean	0.0007	0.0248	−7.5419	0.0001	0.0061	0.0001
Std Dev	0.0008	0.0105	0.7556	0.0003	0.0075	0.0003
Skewness	6.7288	2.1632	0.2184	11.8350	2.7404	11.8041
Kurtosis	81.8544	10.0056	0.6582	207.6906	14.6202	206.6025
Min	0.0000	0.0038	−11.1209	0.0000	0.0000	0.0000
Max	0.0165	0.1286	−4.1015	0.0073	0.0852	0.0072
LB_{10}	2912	6184	8503	194	231	194

Significant Jumps - An Example



Significant Jumps - Yearly Estimates

Natural Gas:

	No. Days	No. Jumps	Prop	RJ on Jump Days (%) Min	Mean	Median	Max
1993	250	14	0.056	25.84	40.42	32.69	85.34
1994	248	18	0.073	25.18	35.10	35.27	54.62
1995	250	15	0.060	24.99	35.80	31.31	75.23
1996	248	24	0.097	26.62	36.54	33.81	61.08
1997	213	14	0.066	28.01	34.95	31.27	73.60
1998	240	20	0.083	26.47	39.19	35.91	78.50
1999	232	24	0.103	25.32	32.93	31.64	55.07
2000	235	25	0.106	25.99	44.06	38.52	87.47
2001	236	42	0.178	23.56	43.20	37.07	85.92
2002	245	23	0.094	25.67	42.78	41.03	72.97
2003	249	36	0.145	25.89	36.06	32.10	77.15
2004	249	32	0.129	25.56	40.22	36.31	69.19
2005	251	27	0.108	26.20	38.31	33.92	68.96
2006	250	30	0.120	25.47	39.24	35.86	62.96
2007	258	22	0.085	22.73	31.26	30.59	52.13

Significant Jumps - Yearly Estimates

Crude Oil:

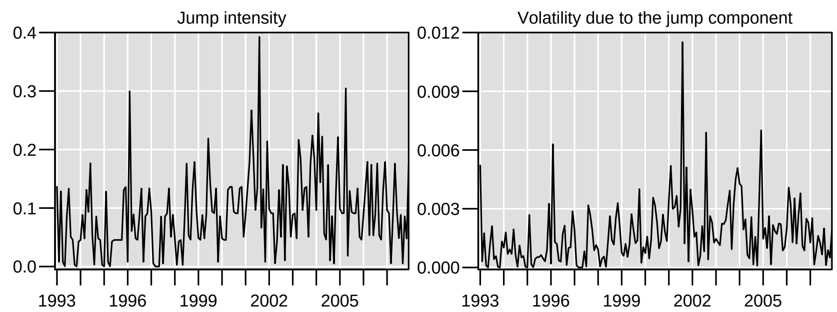
				RJ on Jump Days (%)			
	No. Days	No. Jumps	Prop	Min	Mean	Median	Max
1993	250	21	0.084	22.61	29.51	28.92	40.14
1994	250	26	0.104	23.41	33.75	31.78	45.92
1995	250	17	0.068	22.58	34.93	34.73	62.32
1996	251	22	0.088	23.41	32.93	31.67	62.08
1997	251	22	0.088	23.03	32.24	31.54	47.48
1998	251	16	0.064	25.26	30.43	30.03	38.68
1999	250	17	0.068	22.92	30.59	29.12	41.82
2000	249	11	0.044	23.32	30.60	30.57	44.61
2001	246	6	0.024	24.99	29.42	27.88	36.76
2002	250	11	0.044	26.48	34.09	33.55	49.11
2003	250	19	0.076	27.86	39.28	35.60	58.77
2004	249	15	0.060	25.61	34.82	35.64	47.22
2005	251	17	0.068	24.97	32.37	29.98	44.20
2006	250	10	0.040	28.10	34.31	29.96	48.73
2007	258	32	0.124	22.61	31.14	29.40	64.18

Trends in Daily Time Series

	$RV_t^{1/2}$	$C_{t,\alpha}^{1/2}$	$J_{t,\alpha}^{1/2}$	$RJ_t^{1/2}$
	Panel A: Crude Oil			
Intercept	16356.0	16106.0	767.9	38141.0
	(152.79)	(159.28)	(14.22)	(18.18)
Trend	0.383	0.398	−0.044	−6.207
	(4.61)	(5.07)	(−1.05)	(−3.82)
R Adj	0.00	0.01	0.00	0.00
F Stat	21.26	25.70	1.10	14.56
	Panel B: Heating Oil			
Intercept	16646.0	16316.0	982.0	56079.0
	(173.84)	(179.79)	(17.39)	(25.13)
Trend	0.485	0.470	0.043	−3.057
	(6.44)	(6.59)	(0.97)	(−1.74)
R Adj	0.01	0.01	0.00	0.00
F Stat	41.51	43.39	0.95	3.04
	Panel C: Natural Gas			
Intercept	24830.0	24131.0	1840.7	67217.0
	(143.45)	(148.29)	(16.84)	(24.98)
Trend	0.326	0.198	0.352	10.307
	(1.98)	(1.28)	(3.38)	(4.03)
R Adj	0.00	0.00	0.00	0.00
F Stat	3.91	1.64	11.44	16.20

Trends in Intensity and Size

Exponentially smoothed time series of the monthly jump intensity and variance



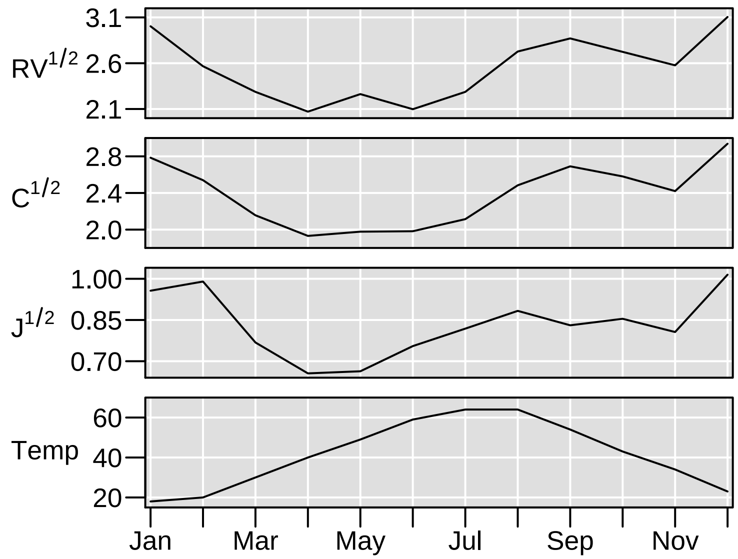
Summary Statistics - Significant Jumps

α	0.500	0.950	0.990	0.999	0.9999
		Panel A: Crude Oil			
No. Jumps	2665	700	326	127	50
Proportion	0.59	0.16	0.07	0.03	0.01
Mean ($J_{t,\alpha}^{1/2}$)	0.0056	0.0092	0.0107	0.0133	0.0138
Std Dev	0.0047	0.0068	0.0089	0.0130	0.0071
$\text{LB}_{10}, J_{t,\alpha}^{1/2}$	91	70	58	57	0
		Panel B: Heating Oil			
No. Jumps	2843	823	392	152	77
Proportion	0.64	0.18	0.09	0.03	0.02
Mean ($J_{t,\alpha}^{1/2}$)	0.0059	0.0094	0.0114	0.0136	0.0151
Std Dev	0.0043	0.0055	0.0069	0.0088	0.0110
$\text{LB}_{10}, J_{t,\alpha}^{1/2}$	137	102	97	40	0
		Panel C: Natural Gas			
No. Jumps	2369	710	367	161	87
Proportion	0.64	0.19	0.10	0.04	0.02
Mean ($J_{t,\alpha}^{1/2}$)	0.0094	0.0157	0.0186	0.0243	0.0286
Std Dev	0.0074	0.0098	0.0114	0.0138	0.0149
$\text{LB}_{10}, J_{t,\alpha}^{1/2}$	194	165	127	131	19

Summary Statistics - Signed Jumps

Contract	N	Mean	Median	StdDev	Max	Min
			Positive Jumps			
Crude Oil	146	0.011	0.009	0.012	0.137	0.003
Heating Oil	154	0.011	0.010	0.008	0.102	0.005
Natural Gas	143	0.018	0.014	0.012	0.083	0.006
			Negative Jumps			
Crude Oil	177	0.010	0.009	0.005	0.031	0.003
Heating Oil	232	0.011	0.010	0.005	0.036	0.004
Natural Gas	219	0.019	0.016	0.011	0.085	0.006

Seasonal Effects: Natural gas



Seasonal Effects: Natural Gas

	$RV_t^{1/2}$	$C_t^{1/2}$	$J_t^{1/2}$	RJ_t	Stock	Temp
Intercept	0.0224	0.0216	0.0020	0.0735	22.1869	64.1667
	(38.82)	(39.98)	(5.20)	(7.81)	(56.85)	(74.77)
D1	0.0073	0.0073	−0.0001	−0.0084	−0.3722	−45.4444
	(9.04)	(9.65)	(−0.27)	(−0.64)	(−0.67)	(−37.44)
D2	0.0049	0.0046	0.0009	0.0004	−6.2694	−42.8333
	(5.89)	(5.88)	(1.56)	(0.03)	(−11.18)	(−35.29)
D3	0.0008	0.0011	−0.0007	−0.0229	−9.8154	−34.8137
	(0.97)	(1.49)	(−1.36)	(−1.72)	(−17.92)	(−28.27)
D4	−0.0017	−0.0014	−0.0006	−0.0028	−10.0588	−24.9314
	(−2.03)	(−1.74)	(−1.15)	(−0.20)	(−18.15)	(−20.25)
D5	−0.0022	−0.0019	−0.0006	−0.0097	−7.4043	−15.5556
	(−2.69)	(−2.53)	(−1.22)	(−0.74)	(−13.47)	(−12.82)
D6	−0.0005	−0.0003	−0.0003	0.0006	−3.4215	−5.4444
	(−0.60)	(−0.39)	(−0.59)	(0.04)	(−6.20)	(−4.49)
D8	0.0029	0.0025	0.0007	0.0030	2.5947	−0.9444
	(3.66)	(3.35)	(1.32)	(0.23)	(4.74)	(−0.78)
D9	0.0039	0.0043	−0.0009	−0.0131	5.5917	−9.5556
	(4.79)	(5.63)	(−1.62)	(−0.98)	(10.09)	(−7.87)
D10	0.0032	0.0033	−0.0001	−0.0106	8.2283	−20.5556
	(3.97)	(4.37)	(−0.11)	(−0.81)	(14.67)	(−16.94)
D11	0.0024	0.0026	−0.0003	−0.0104	8.8009	−31.2222
	(2.93)	(3.37)	(−0.49)	(−0.78)	(15.62)	(−25.72)
D12	0.0080	0.0078	0.0004	−0.0016	5.3676	−41.3333
	(9.78)	(10.26)	(0.76)	(−0.12)	(9.69)	(−34.06)
R Adj	0.09	0.09	0.01	0.00	0.81	0.95
F Stat	32.71	34.61	2.15	0.65	297.89	370.35

Intraday Pattern: Storage Reports

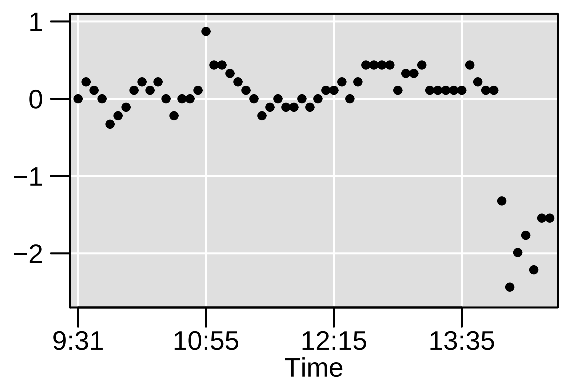
The primary announcements with implications for the energy markets are released by the Energy Information Administration (EIA), DOE.

- Weekly natural gas storage report:
  - 2:00 PM on Wednesdays, 2000 - 2002
  - 10:30 AM on Thursdays, 2002 - present
- Weekly petroleum status reports:
  - 10:35 AM on Wednesdays, small version
  - 1:00 PM on Wednesdays, full version

Intraday Patterns

"Gas and Crude Oil Prices Dive As U.S. Inventory Levels Surge" (Los Angeles Times on May 3, 2001)

Cumulative intraday returns in natural gas futures on May 2, 2001 (EIA released the storage report at 2 PM)



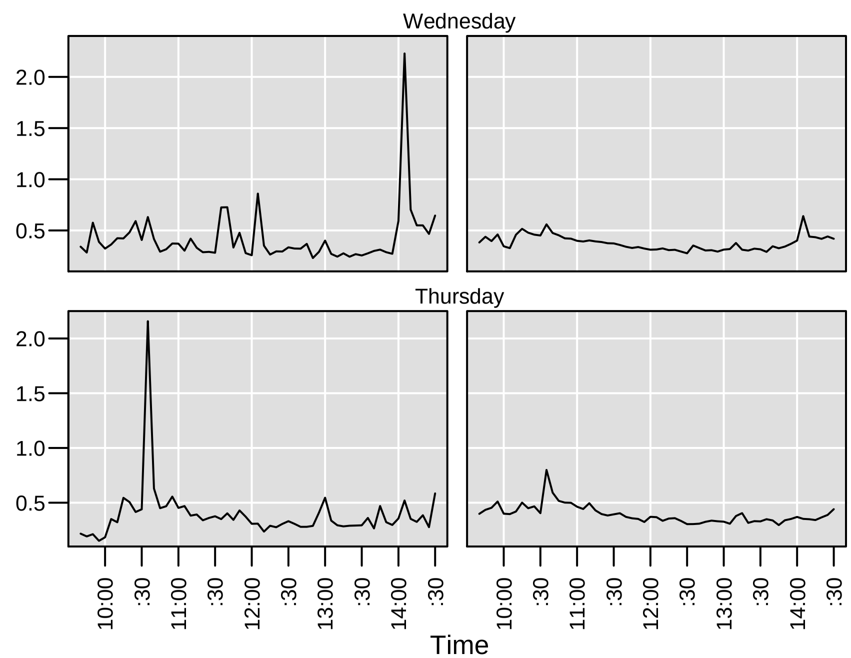
Intraday: Natural Gas - Thursday

	RV
10:00 - 10:05	-0.010
	(-0.150)
10:10 - 10:15	0.094
	(2.430)
10:15 - 10:20	0.047
	(1.210)
10:20 - 10:25	0.047
	(1.200)
10:25 - 10:30	0.004
	(0.110)
10:30 - 10:35	**1.111**
	(28.570)
10:35 - 10:40	0.194
	(5.000)
10:40 - 10:45	0.093
	(2.390)
10:45 - 10:50	0.082
	(2.110)
10:50 - 10:55	0.095
	(2.440)
10:55 - 11:00	0.049
	(1.270)

Intraday: Natural Gas - Thursdays

	RV - Jump	RV - No Jump
10:00 - 10:05	0.020	−0.019
	(0.060)	(−0.380)
10:10 - 10:15	0.194	0.075
	(1.210)	(2.340)
10:15 - 10:20	0.152	0.026
	(0.950)	(0.820)
10:20 - 10:25	0.069	0.042
	(0.430)	(1.300)
10:25 - 10:30	0.090	−0.013
	(0.560)	(−0.400)
10:30 - 10:35	**4.553**	**0.463**
	(28.320)	**(14.490)**
10:35 - 10:40	0.293	0.175
	(1.830)	(5.470)
10:40 - 10:45	0.099	0.091
	(0.620)	(2.850)
10:45 - 10:50	0.115	0.075
	(0.720)	(2.340)
10:50 - 10:55	0.206	0.073
	(1.280)	(2.280)
10:55 - 11:00	0.101	0.039
	(0.630)	(1.210)

Intraday: Natural Gas



Modeling Realized Variation with Jumps

We examine whether including the jump component as an explanatory variable improves the modeling of the realized variance.

A number of studies in various markets have found long-memory dependencies by employing ARCH, stochastic volatility and ARFIMA (autoregressive fractionally-integrated moving average) models.

Andersen et al. (2007) model and forecast the realized volatility by augmenting the HAR-RV (heterogeneous AR realized volatility) model (Corsi (2004), Müller et al. (1997))

Modeling Realized Variation with Jumps

Autocorrelation in natural gas:

	RV_t	$C_{t,\alpha}$	$J_{t,\alpha}$
1	0.440	0.495	-0.00405
2	0.316	0.362	-0.00979
3	0.285	0.338	-0.01074
4	0.308	0.339	0.02936
5	0.307	0.296	0.16478
6	0.210	0.241	-0.00715
7	0.198	0.230	0.00219
8	0.191	0.215	-0.01399
9	0.219	0.219	0.07325
10	0.243	0.239	0.02225
11	0.202	0.231	-0.01017
12	0.200	0.225	-0.01090
13	0.198	0.220	-0.00423
14	0.215	0.214	0.07761
15	0.231	0.250	0.00737
16	0.185	0.211	-0.00376

Modeling Realized Variation with Jumps

HAR-RV model:

$$\text{RV}_{t,t+h} = \beta_0 + \beta_D \text{RV}_t + \beta_W \text{RV}_{t-5,t} + \beta_M \text{RV}_{t-22,t} + \epsilon_{t,t+h}.$$

HAR-RV-J model

$$\text{RV}_{t,t+h} = \beta_0 + \beta_D \text{RV}_t + \beta_W \text{RV}_{t-5,t} + \beta_M \text{RV}_{t-22,t} + \beta_J J_t + \epsilon_{t,t+h}.$$

HAR-RV: Natural Gas

	$RV_{t,t+h}$			$(RV_{t,t+h})^{1/2}$		
h	1	5	22	1	5	22
			Panel C: Natural Gas			
β_0	0.000	0.000	0.000	0.004	0.006	0.011
	(6.520)	(5.550)	(6.770)	(8.130)	(7.090)	(8.180)
β_D	0.230	0.140	0.058	0.227	0.168	0.090
	(2.900)	(2.900)	(3.630)	(6.100)	(6.370)	(7.120)
β_W	0.327	0.232	0.231	0.379	0.315	0.329
	(5.840)	(3.330)	(3.150)	(11.000)	(8.090)	(5.860)
β_M	0.244	0.343	0.223	0.216	0.300	0.177
	(4.530)	(7.740)	(3.150)	(5.700)	(7.320)	(2.320)
AdjR^2	0.27	0.37	0.31	0.40	0.50	0.40
F	449.44	704.31	547.10	788.89	1201.91	799.25

HAR-RV-J: Natural Gas

	$RV_{t,t+h}$			$(RV_{t,t+h})^{1/2}$		
h	1	5	22	1	5	22
			Panel C: Natural Gas			
β_0	0.000	0.000	0.000	0.004	0.005	0.011
	(6.140)	(5.390)	(6.650)	(7.930)	(6.930)	(8.110)
β_D	0.290	0.187	0.084	0.283	0.202	0.103
	(2.750)	(3.540)	(5.310)	(7.700)	(7.740)	(6.830)
β_W	0.301	0.211	0.220	0.356	0.301	0.324
	(4.540)	(2.980)	(3.030)	(10.710)	(7.840)	(5.840)
β_M	0.244	0.343	0.223	0.215	0.299	0.177
	(4.800)	(7.990)	(3.140)	(5.810)	(7.370)	(2.320)
β_J	−0.224	−0.175	−0.097	−0.109	−0.067	−0.025
	(−0.950)	(−2.120)	(−2.200)	(−3.100)	(−2.620)	(−1.330)
AdjR^2 J	0.27	0.37	0.32	0.40	0.50	0.40

Summary and Conclusions

- Natural gas futures are more volatile than crude oil and heating oil.

- There are significant jump components in all three series. The relative contribution (RJ) ranges from:
  - 22.7% to 87.5% in natural gas
  - 22.6% to 62.2% in crude oil,
  - 22.8% to 73.6% in heating oil.

Summary and Conclusions

- We demonstrate that the continuous diffusion process model for energy futures is violated in practice. (Surprise!)
- The volatility due to the jump component is less persistent than the continuous in energy futures markets.
- Significant jumps occur more often during the winter and at the release time of energy reports.
- Separating the continuous and jump components improves the modeling of the realized variance.

Summary and Conclusions

- Implications for market participants in energy futures markets:
  - Option pricing
  - Hedging
  - Risk management

Further Work

This was a preliminary study with real data. There are many more interesting questions in the analysis of jumps.

Rate of convergence of the asymptotic test statistics.

Other methods of inference about jumps.

Sampling frequency.

With some understanding of the historical price behavior, we can build simulation models to study these issues.

(If you understand something, you can simulate it.)