



Modeling the Implied Volatility Surface

Jim Gatheral
Global Derivatives and Risk Management 2003
Barcelona
May 22, 2003

This presentation represents only the personal opinions of the author and not those of Merrill Lynch, its subsidiaries or affiliates.

Outline of this talk

- A compound Poisson model of stock trading
- Empirical verification of modeling assumptions
- Stochastic volatility
- Empirical dynamics of SPX and VIX
- Dynamics of the implied volatility skew
- Which stochastic volatility model?
- Do stochastic volatility models fit option prices?
- Jumps
- The impact of large option trades

Stock trading as a compound Poisson process

- Consider a random time change from conventional calendar time to trading time such that the rate of arrival of stock trades in a given (transformed) time interval is a constant λ .
- Intuitively, relative to real time, trading time flows faster when there is more activity in the stock and more slowly when there is less activity.
- Suppose that the (random) size n of a trade is independent of the level of activity in the stock.
- Assume further that each trade impacts the mid-log-price of the stock by an amount proportional to \sqrt{n} .
 - This is a standard assumption in the market microstructure literature
- Then the change in log-mid-price over some time interval is given by

$$\Delta x = \sum_{i=1}^N \text{sgn}(n_i) \sqrt{n_i}$$

- Note that both the number of trades N and the size of each trade n_i in a given time interval Δt are random.

Volatility and volume: a relationship

- The variance of this random sum of random variables is given by

$$\begin{aligned} \text{Var} [\Delta x] &= E [N] \text{Var} [\mathbf{a} \sqrt{n_i}] + \text{Var} [N] E [\mathbf{a} \sqrt{n_i}]^2 \\ &= \mathbf{a}^2 E [N] E [n_i] \end{aligned}$$

- Rewriting this in terms of volatility, we obtain

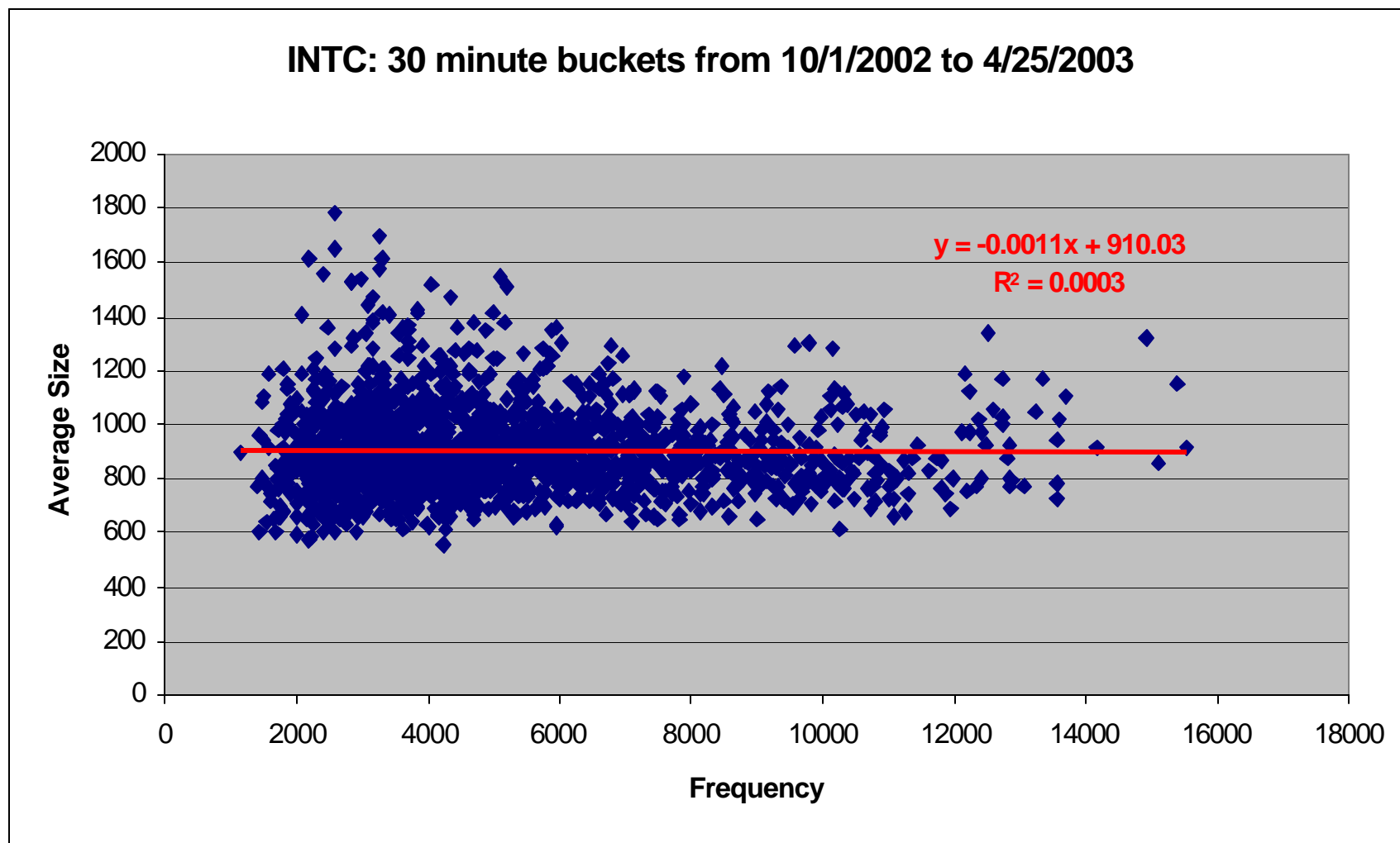
$$\text{Var} [\Delta x] = \mathbf{s}^2 \Delta t = \mathbf{a}^2 E [N] E [n_i]$$

- But $E [N] E [n_i]$ is just the expectation of the volume over the time interval Δt . Let $E [N] = \mathbf{I} \Delta t$ where \mathbf{I} is the trading rate.
- The factor Δt cancels and transforming back to real time, we see that variance is directly proportional to volume in this simple model.
- Moreover, the distribution of returns *in trading time* is approximately Gaussian for large Δt .

The \sqrt{n} relationship

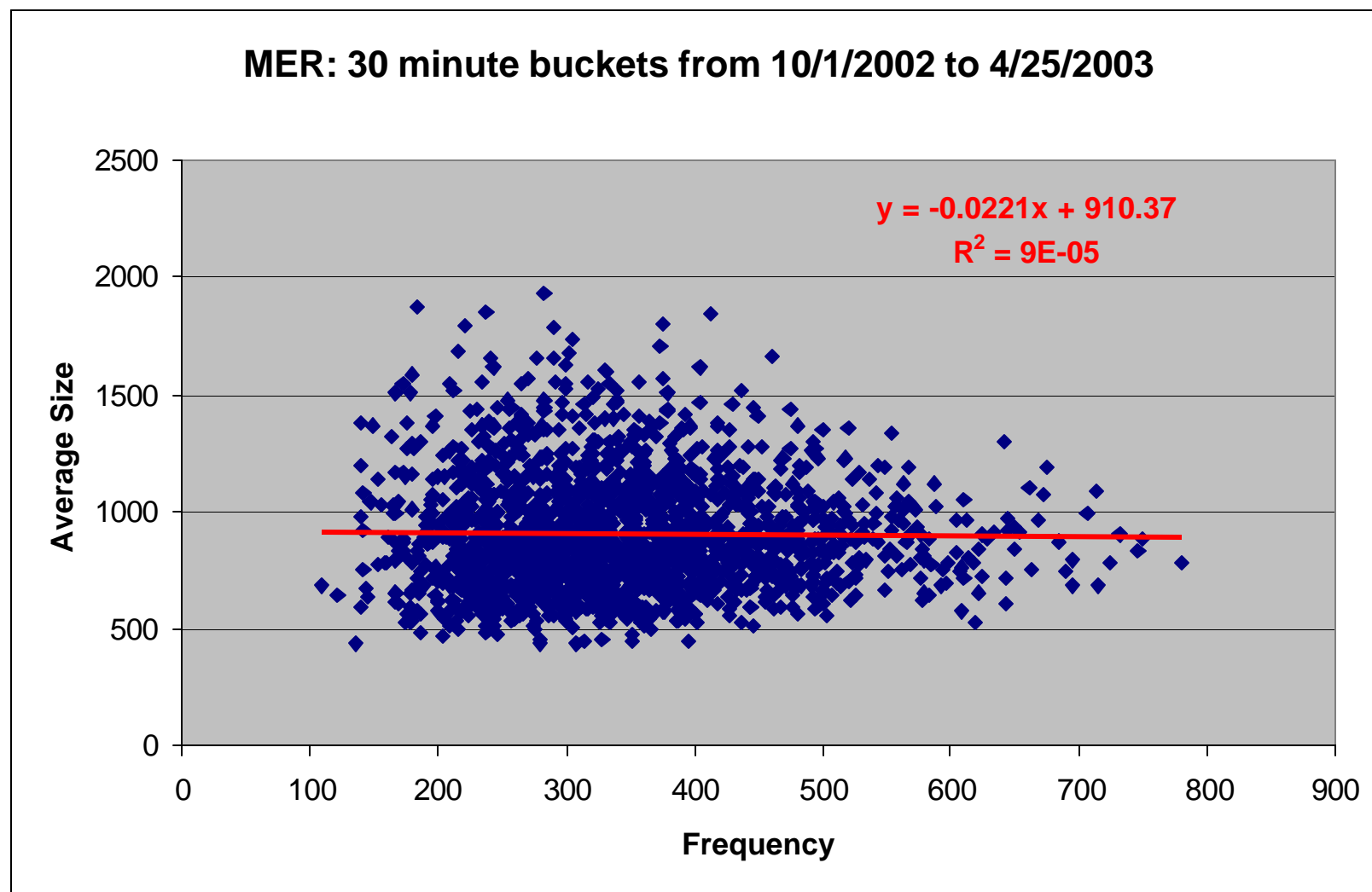
- A key assumption in our simple model is that market impact is proportional to the square root of the trade size \sqrt{n} . The following argument shows why this is plausible:
 - A market maker requires an excess return proportional to the risk of holding inventory.
 - Risk is proportional to $\sigma\sqrt{T}$ where T is the holding period.
 - The holding period should be proportional to the size of the position.
 - So the required excess return must be proportional to \sqrt{n} .

Independence of trade size and trade frequency*



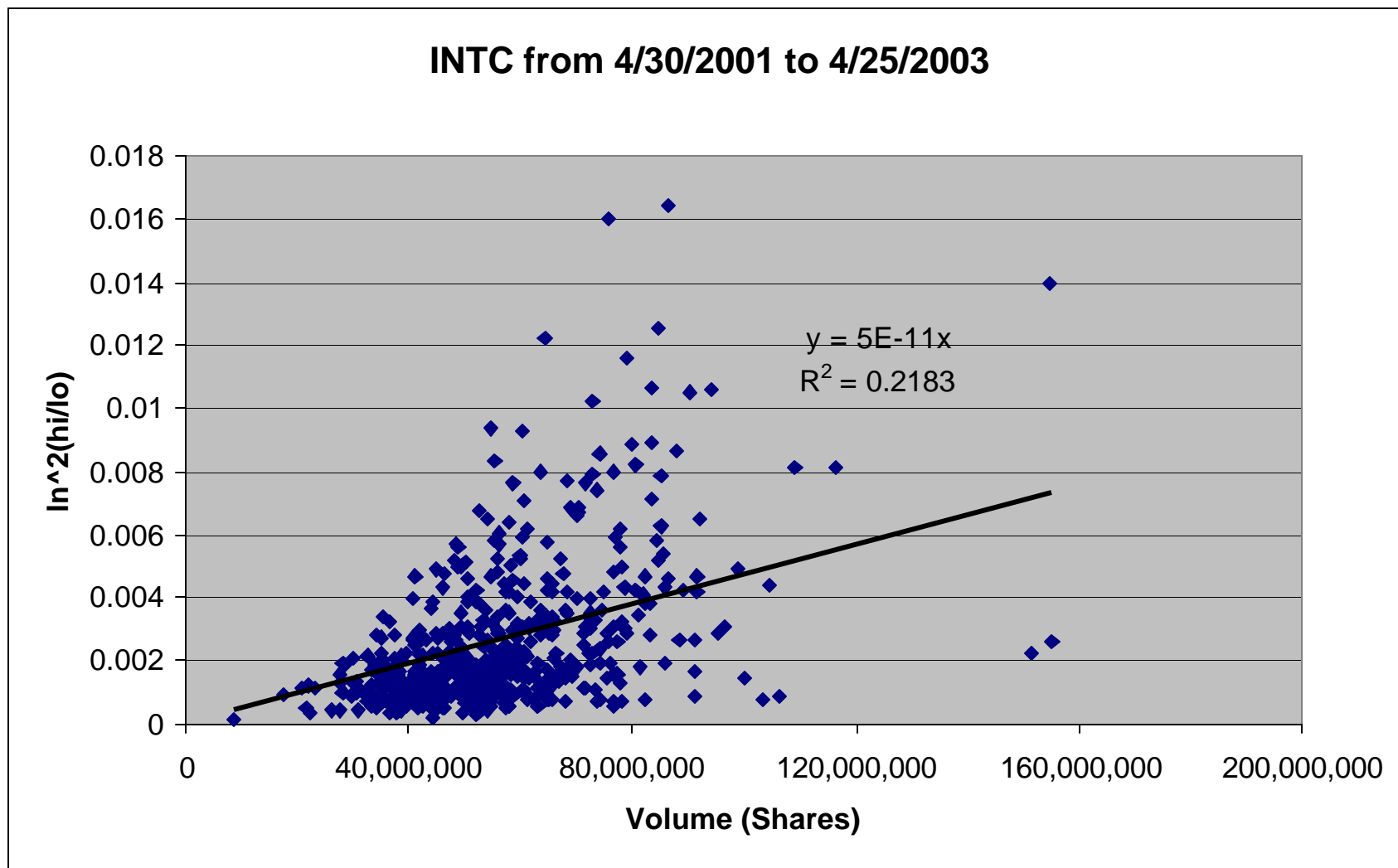
* Excluding trades over 10,000 shares

Independence of trade size and trade frequency*

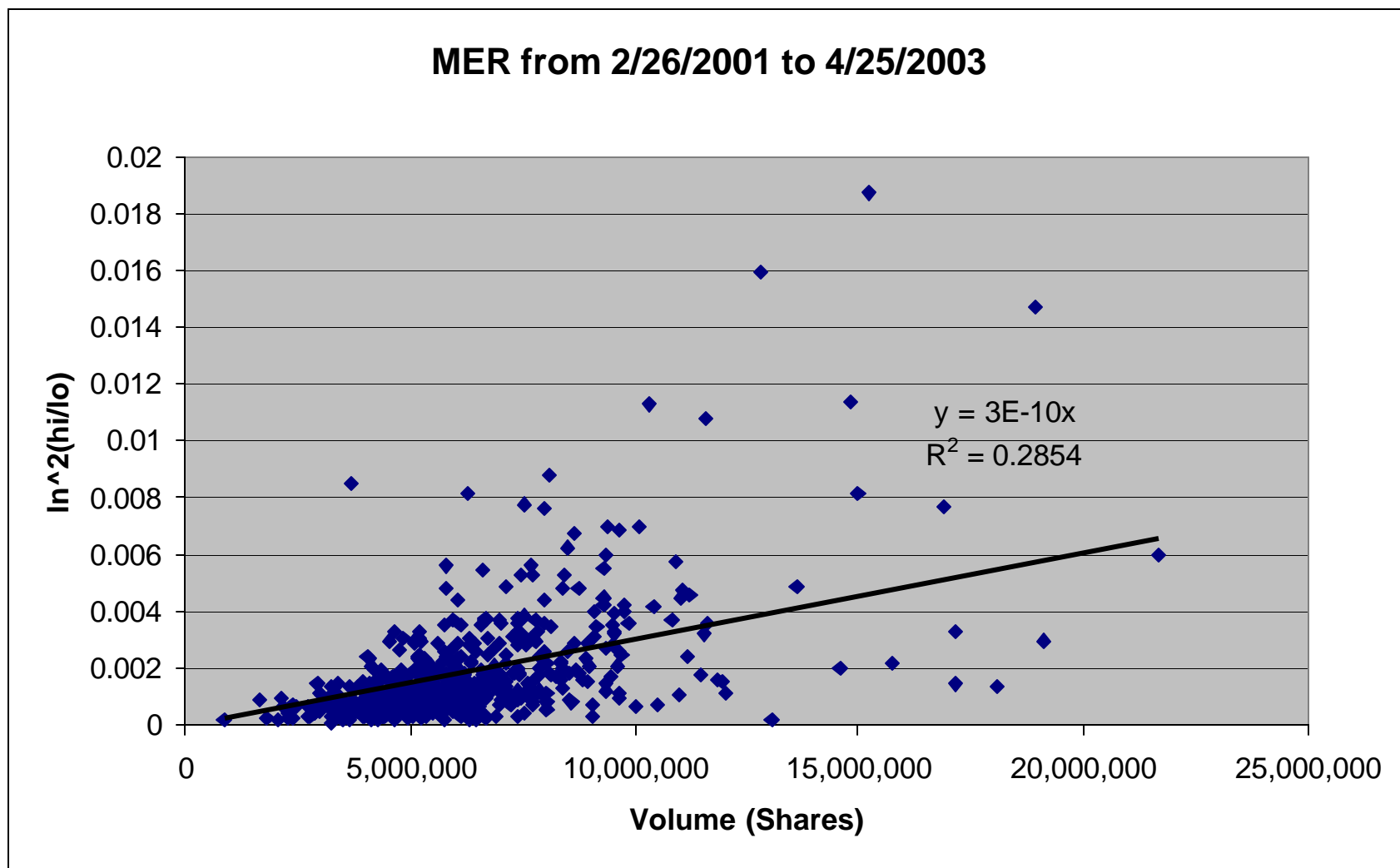


* Excluding trades over 10,000 shares

Empirical variance vs volume



Empirical variance vs volume



Implications of our simple model

- Our simple but realistic model has the following modeling implications
 - Log returns are roughly Gaussian with constant variance *in trading time* (sometimes call intrinsic time) defined in terms of transaction volume
 - Trading time is the inverse of variance
 - When we transform from trading time to real time, variance appears to be random
- It is natural to model the log stock price as a diffusion process subordinated to another random process which is really trading volume in our model - stochastic volatility
- What form should the volatility process take?

Empirical observations

- In the underlying index return data we observe
 - Clustering: large moves follow large moves, small moves follow small moves
 - Mean reversion of volatility
 - Negative correlation between implied volatility moves and index returns
- In the implied volatility data we observe
 - The volatility envelope: short-dated volatilities move more than long-dated ones
 - A pronounced skew with downside implied volatilities significantly higher than upside implied volatilities.
 - Implied volatilities move more when implied volatility is high.

A generic stochastic volatility model

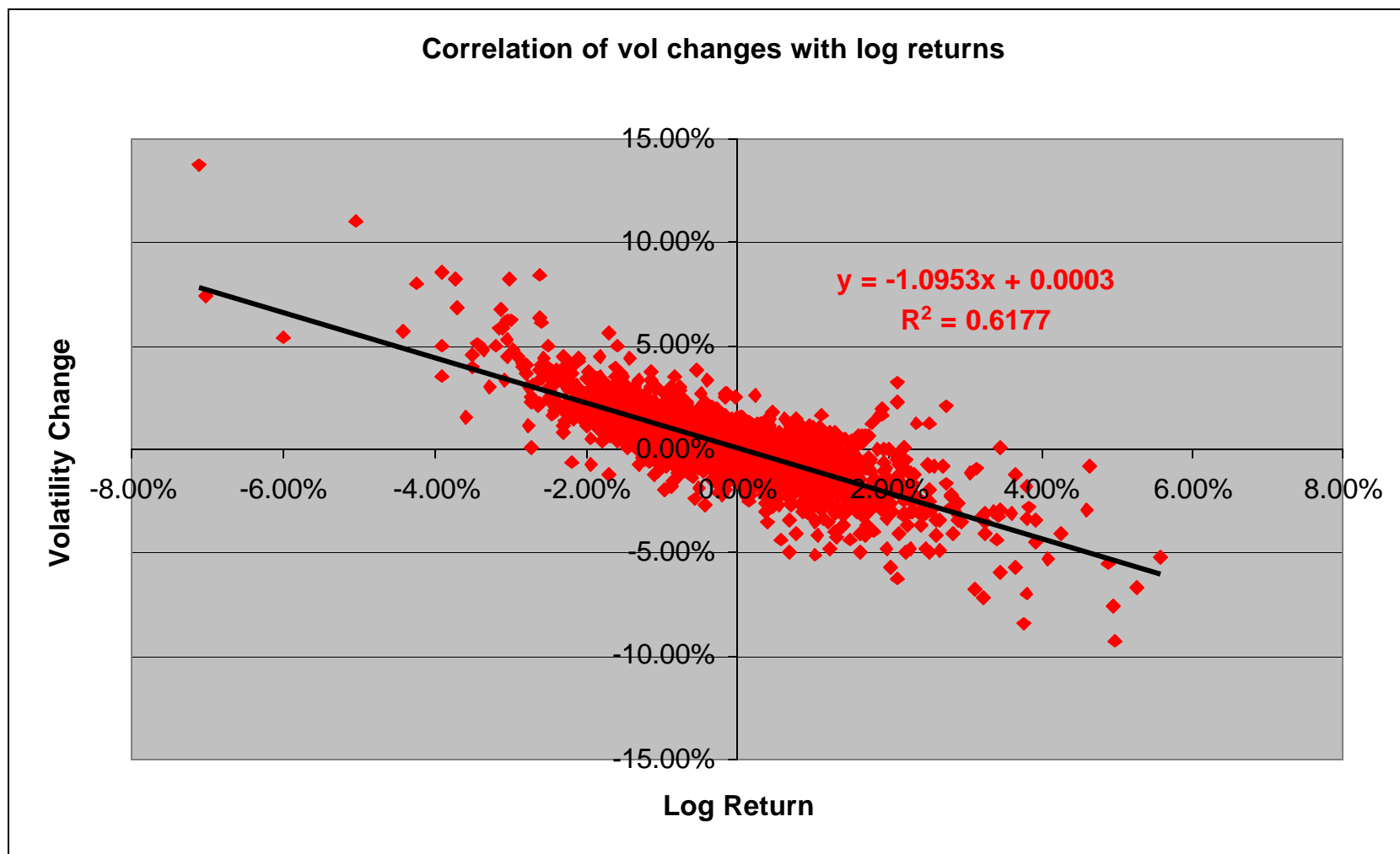
- We are now in a position to write down a generic stochastic volatility model consistent with our observations. Let x denote the log stock price and v denote its variance. Then

$$dx = \mathbf{m}dt + \sqrt{v} dZ_1$$
$$dv = \mathbf{a}(v) + \mathbf{h}v^b dZ_2$$

with $\langle dZ_1, dZ_2 \rangle = \mathbf{r} dt$.

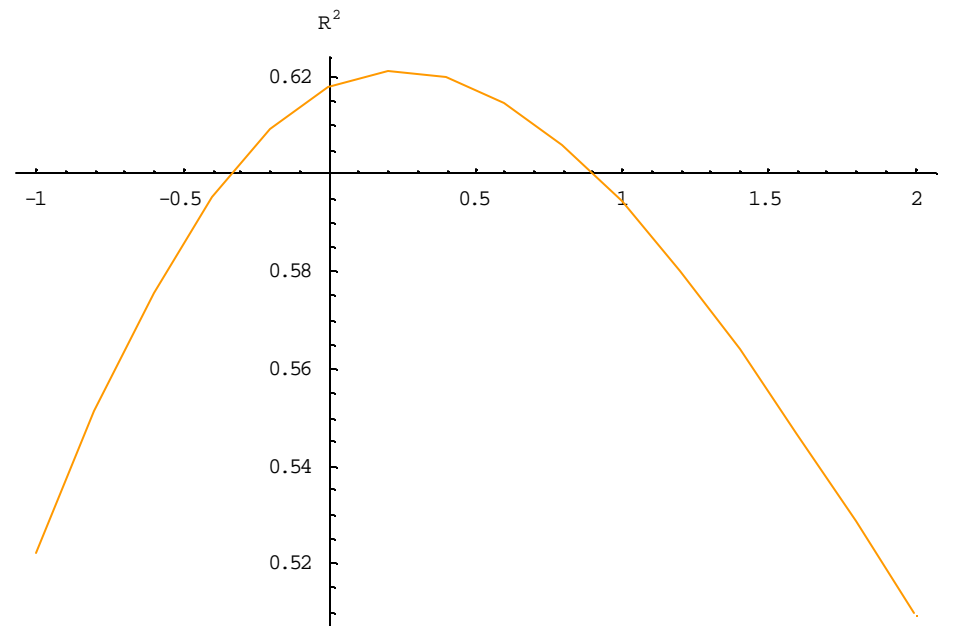
- $\mathbf{a}(v)$ is a mean-reversion term, \mathbf{r} is the correlation between volatility moves and stock price moves and \mathbf{h} is called “volatility of volatility”.
- $b = \frac{1}{2}$ gives the Heston model
 $b=1$ gives Wiggins' lognormal model
 $b = \frac{3}{2}$ gives Lewis's $\frac{3}{2}$ model
- To estimate \mathbf{b} we analyze historical SPX and VIX data.

SPX from 1/1/1990 to 4/25/2003



Regressing VIX vs SPX

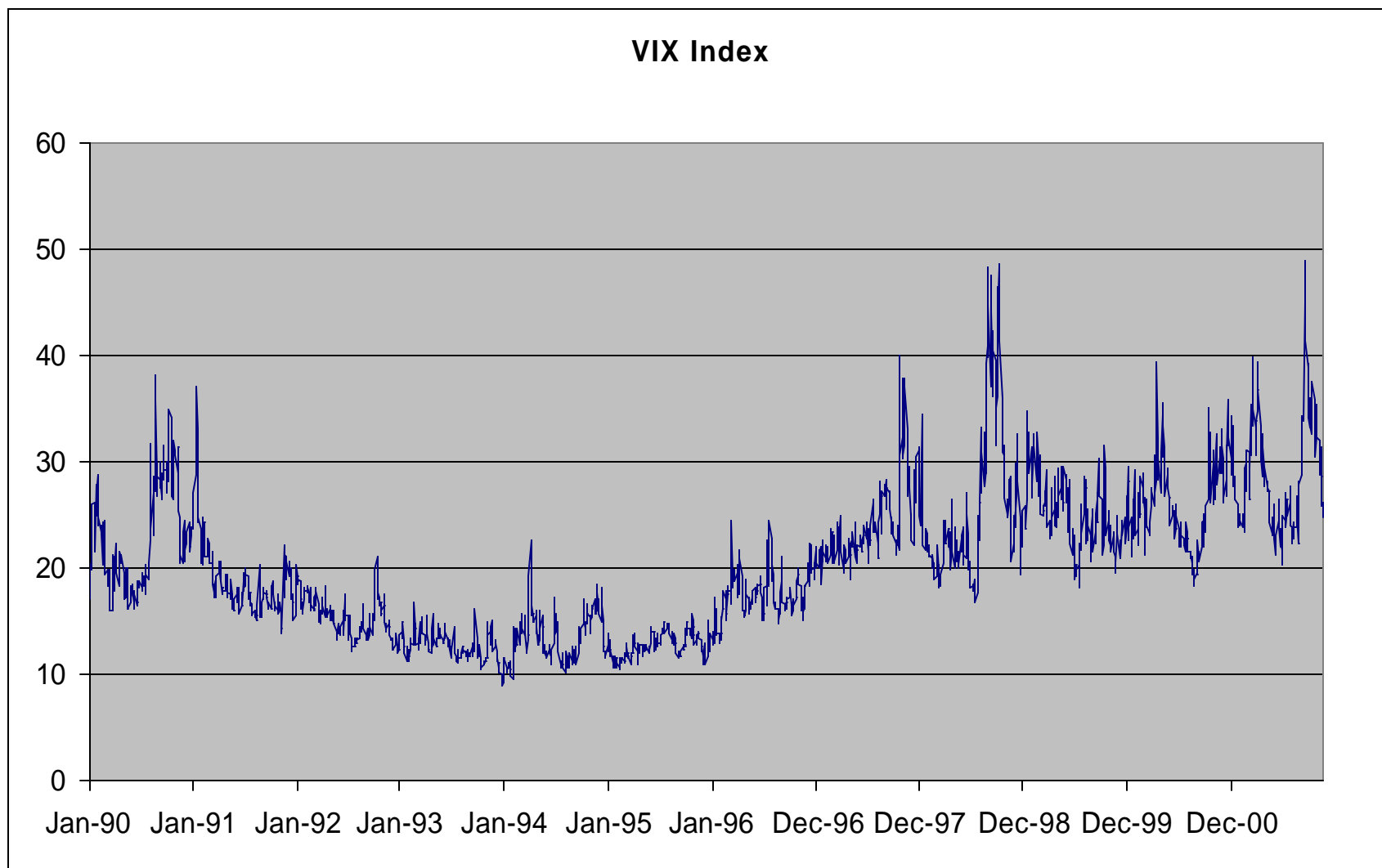
- We regress $VIX^a \Delta VIX$ vs $\Delta \ln(\text{SPX})$ for different values of a .
- Fit quality is plotted below:



$a = 0$ is lognormal
 $a = -1$ is 3/2 model
 $a = 1$ is Heston

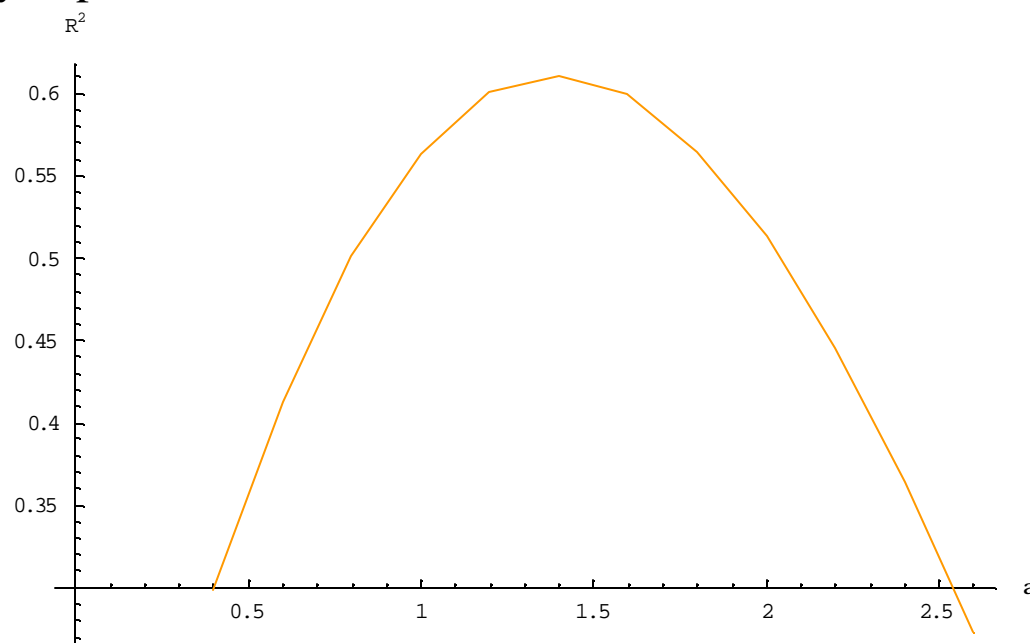
- We note that R^2 is maximized at roughly $a = 0$ which corresponds to a lognormal process for volatility.
- At $a = 0$, $R^2 \approx 0.62$ so $r \approx -0.79$. Then, from the slope of the regression, $\Delta v \approx h v \sqrt{\Delta t} Z$ with $h \approx 2.8$.

Historical SPX implied volatility



Empirical dynamics of VIX

- We regress the 20-day moving average std of ΔVIX vs VIX^a for different values of a .
- Fit quality is plotted below:



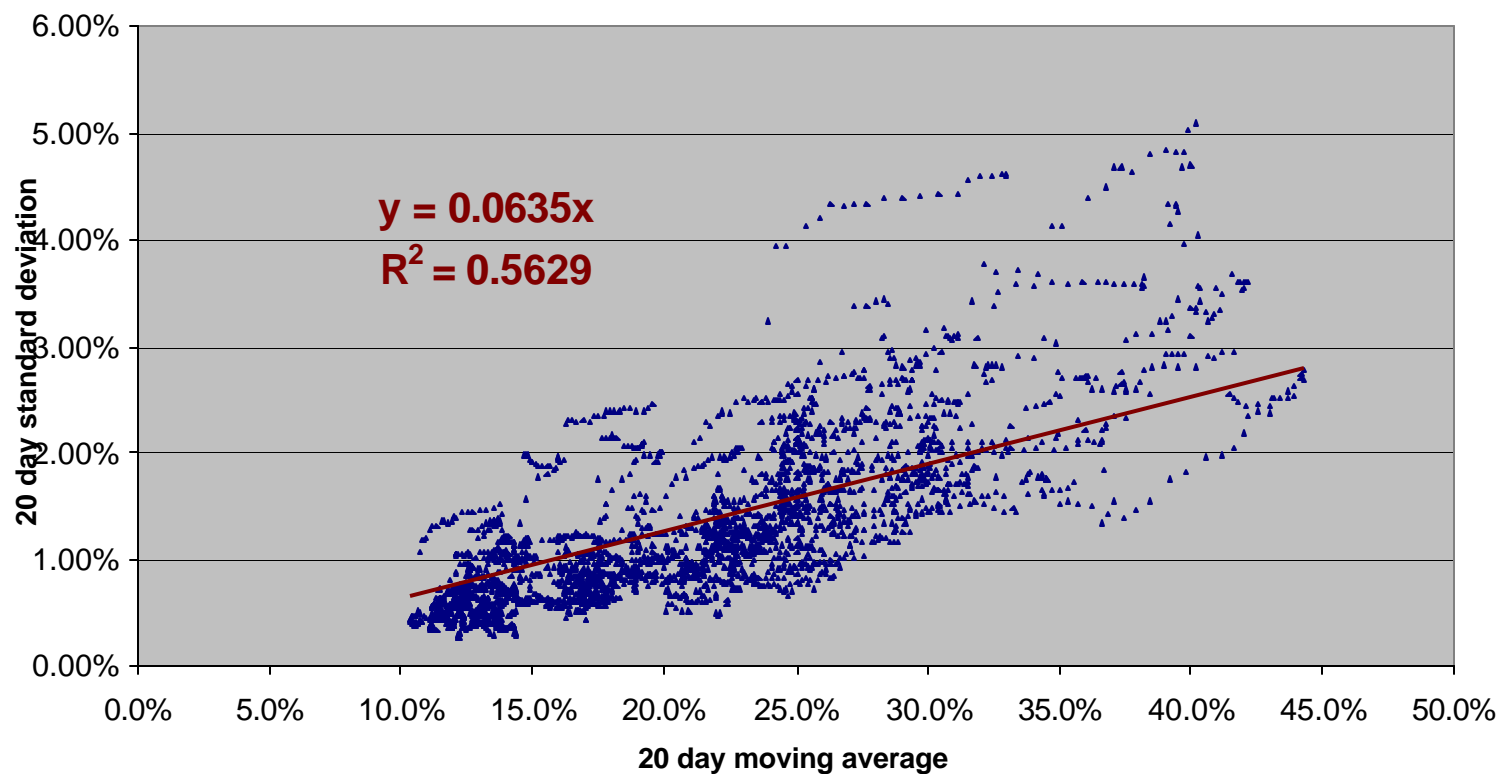
$a = 1$ is lognormal

$a = 2$ is 3/2 model

$a = 0$ is Heston

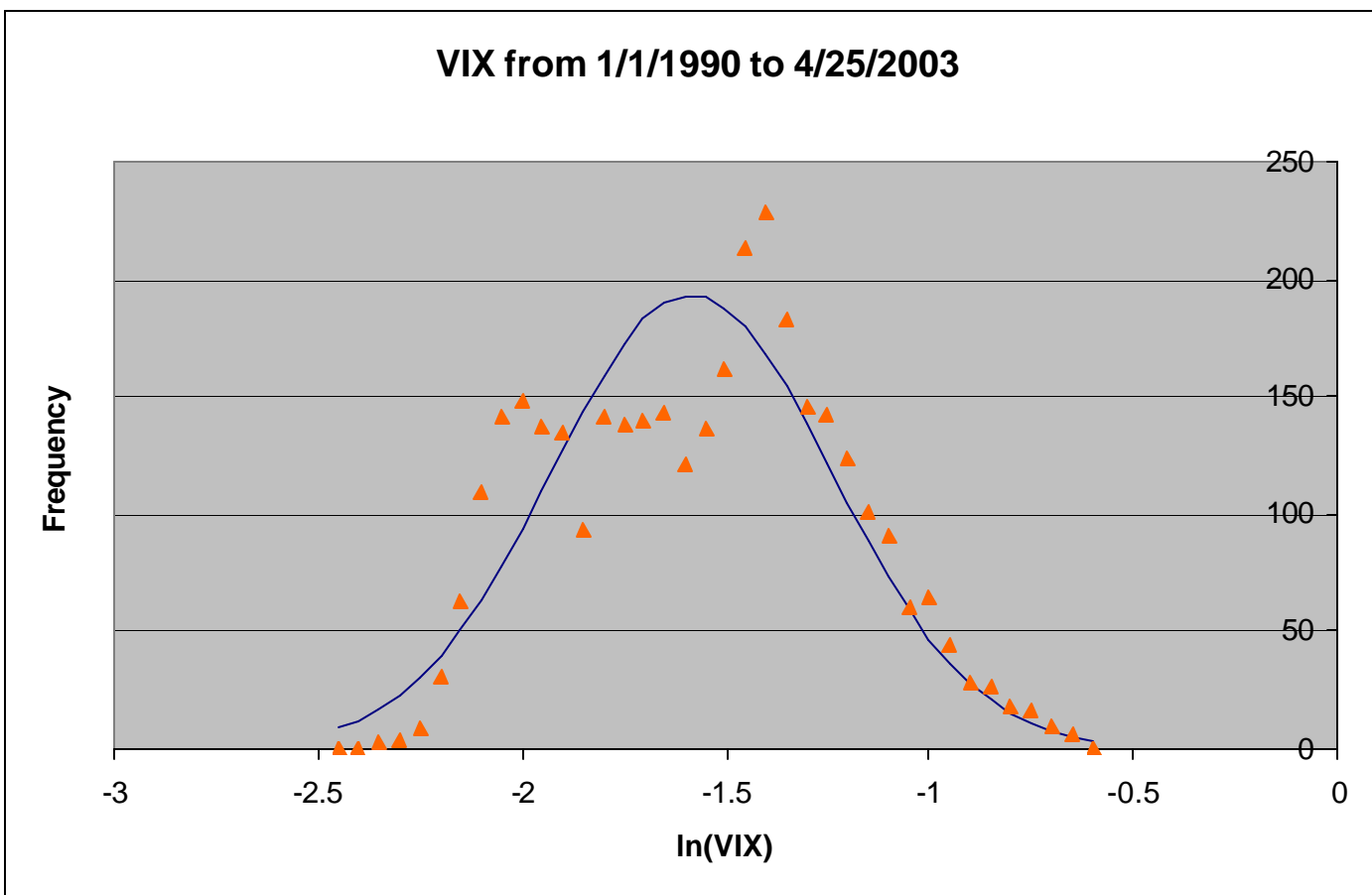
- We note that R^2 is maximized at roughly $a = 1.4$. This would put the volatility dynamics between lognormal and 3/2 with $b = (a + 1)/2 \approx 1.2$
- Imposing lognormality, the slope of the regression should give us volatility of volatility directly.

Regression of VIX volatility vs VIX level



- The regression gives $\Delta s \approx \frac{1}{2} h s \sqrt{\Delta t} \approx 0.0635 s$. Then, $h \approx 0.0635 \times 16 \times 2 \approx 2.0$

Unconditional distribution of VIX vs lognormal



- Consistent with lognormal volatility dynamics!

Dynamics of the volatility skew

- Our empirical investigations so far indicate that the appropriate specification of stochastic volatility is lognormal with volatility of volatility h somewhere between 2.0 and 3.0.
- This choice of specification has implications for the dynamics of the implied volatility skew. We can check to see whether or not the volatility skew behaves consistently with this specification.

Vol term structure and skew under stochastic volatility

- All stochastic volatility models generate volatility surfaces with approximately the same shape
- The Heston model $dv = -l(v - \bar{v})dt + h\sqrt{v}dZ$ has an implied volatility term structure that looks to leading order like

$$s_{BS}(x, T)^2 \approx \bar{v} + (v - \bar{v}) \frac{(1 - e^{-lT})}{lT}$$

It's easy to see that this shape should not depend very much on the particular choice of model.

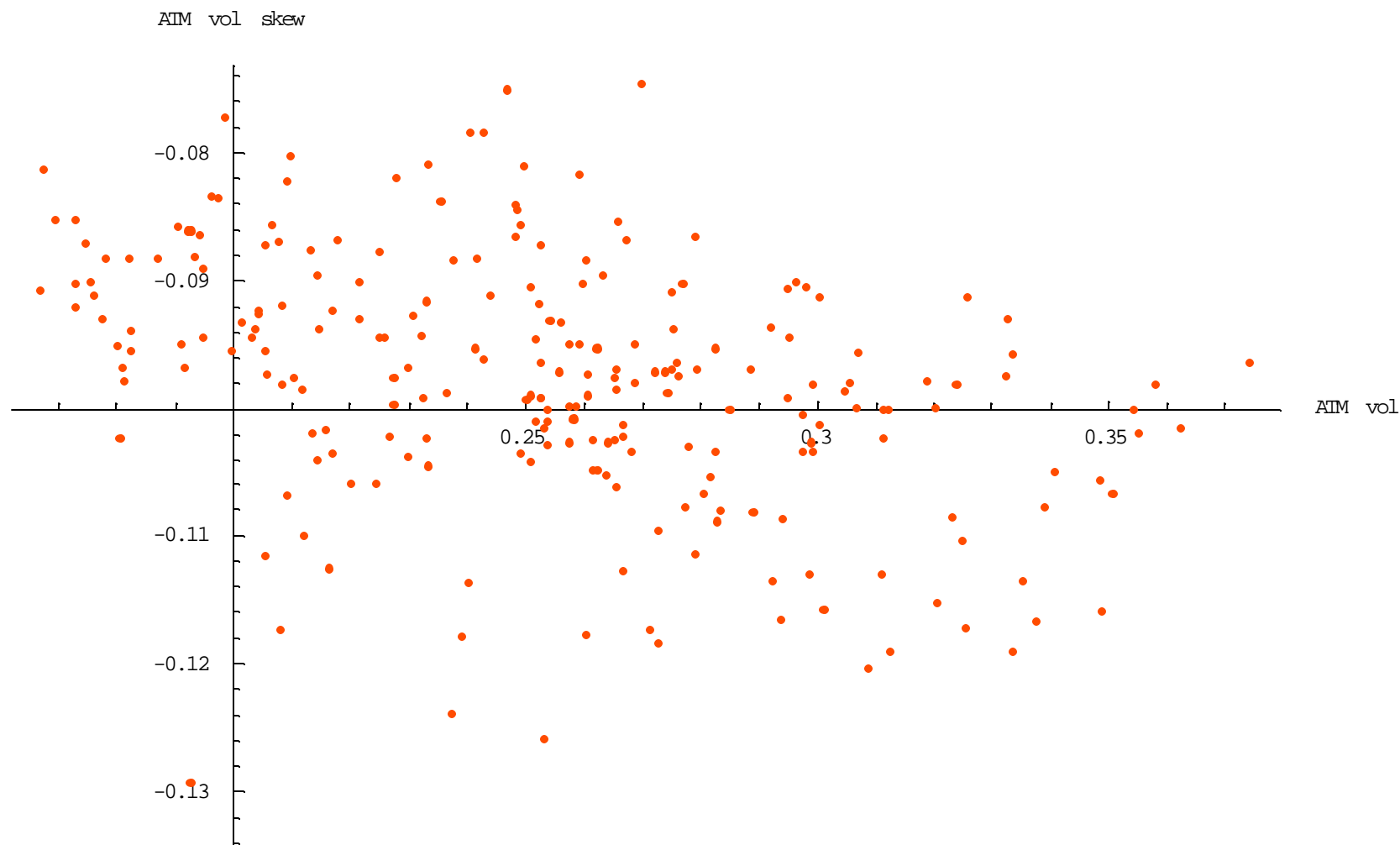
- Also, Gatheral (2002) shows that the term structure of the at-the-money volatility skew has the following approximate behavior for all stochastic volatility models of the form $dv = -l(v - \bar{v})dt + hv^b dZ$

$$\left. \frac{\partial}{\partial x} s_{BS}(x, T)^2 \right|_{x=0} \approx \frac{rhv^{b-1/2}}{l'T} \left\{ 1 - \frac{(1 - e^{-l'T})}{l'T} \right\}$$

$$\text{with } l' = l - rhv^{b-1/2} / 2$$

- So we can estimate b by regressing volatility skew against volatility level.

SPX 3-month ATM volatility skew vs ATM 3m volatility



Interpreting the regression of skew vs volatility

- Recall that if the variance satisfies the SDE

$$dv \sim v^b dZ$$

at-the-money variance skew should satisfy

$$\left. \frac{\partial v}{\partial k} \right|_{k=0} \propto v^{b-1/2}$$

and at-the-money volatility skew should satisfy

$$\left. \frac{\partial \mathbf{s}_{BS}}{\partial k} \right|_{k=0} \propto \mathbf{s}_{BS}^{2b-2}$$

- The graph shows volatility skew to be roughly independent of volatility level so $b \approx 1$ again consistent with lognormal volatility dynamics.

Implied vs empirical stochastic volatility parameters

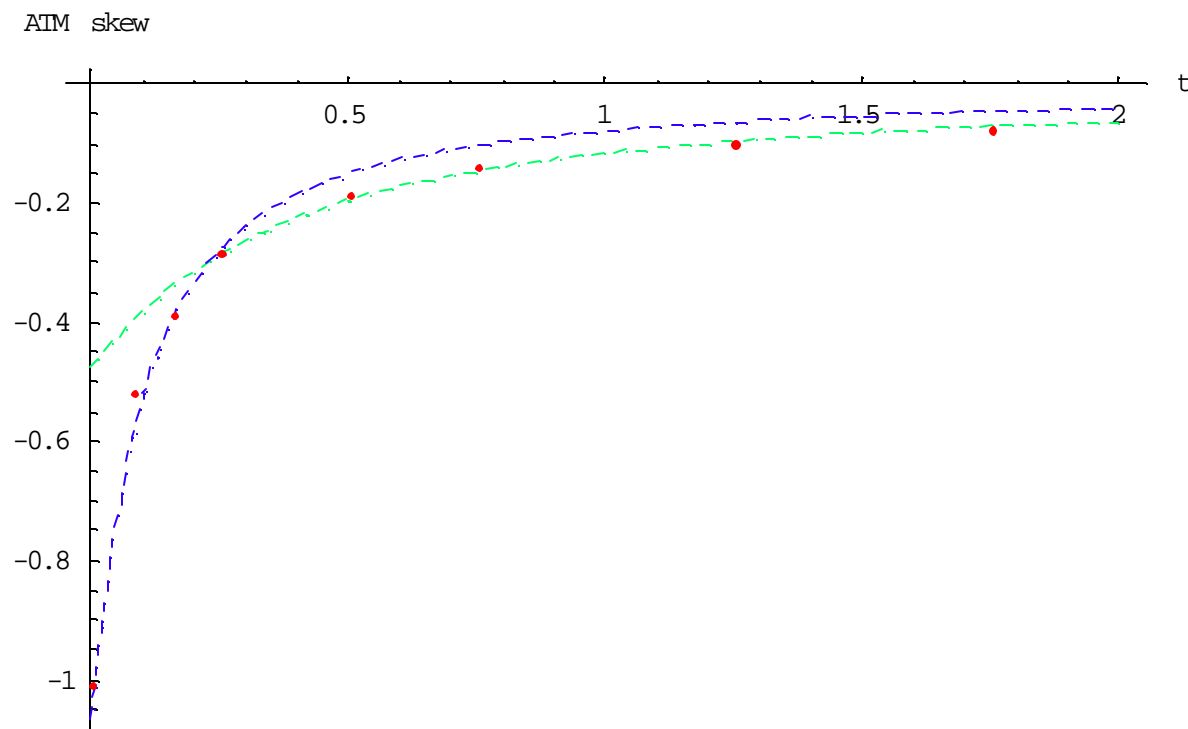
- It is often claimed that stochastic volatility parameters obtained by fitting stochastic volatility models to option prices are inconsistent with historical parameters. Specifically, the implied volatility of volatility is often thought to be extreme.
- The typical volatility of volatility returned from a fit of the Heston model to option prices is $\mathbf{h}_H \approx 1$.
- The relationship between lognormal vol of vol and Heston vol of vol should be $\mathbf{h}_H \approx \mathbf{h} \times \mathbf{s}_{BS}$. With an implied volatility level of around 30%, $\mathbf{h}_H \approx 1$ would give $\mathbf{h} \approx 3.3$ which should be compared with $\mathbf{h} \in (2.0, 3.0)$ from our empirical analysis.

Do stochastic volatility models fit option prices?

- Once again, we note that the shape of the implied volatility surface generated by a stochastic volatility model does not strongly depend on the particular choice of model.
- Given this observation, do stochastic volatility models fit the implied volatility surface? The answer is “more or less”. Moreover, fitted parameters are reasonably stable over time.
- For very short expirations however, stochastic volatility models certainly don't fit as the next slide will demonstrate.

Short Expirations

- Here's a graph of the SPX volatility skew on 17-Sep-02 (just before expiration) and various possible fits of the volatility skew formula:



- We see that the form of the fitting function is too rigid to fit the observed skews.
- Jumps could explain the short-dated skew!

More reasons to add jumps

- The statistical (historical) distribution of stock returns and the option implied distribution have quite different shapes.
- The size of the volatility of volatility parameter estimated from fits of stochastic volatility models to option prices is too high to be consistent with empirical observations

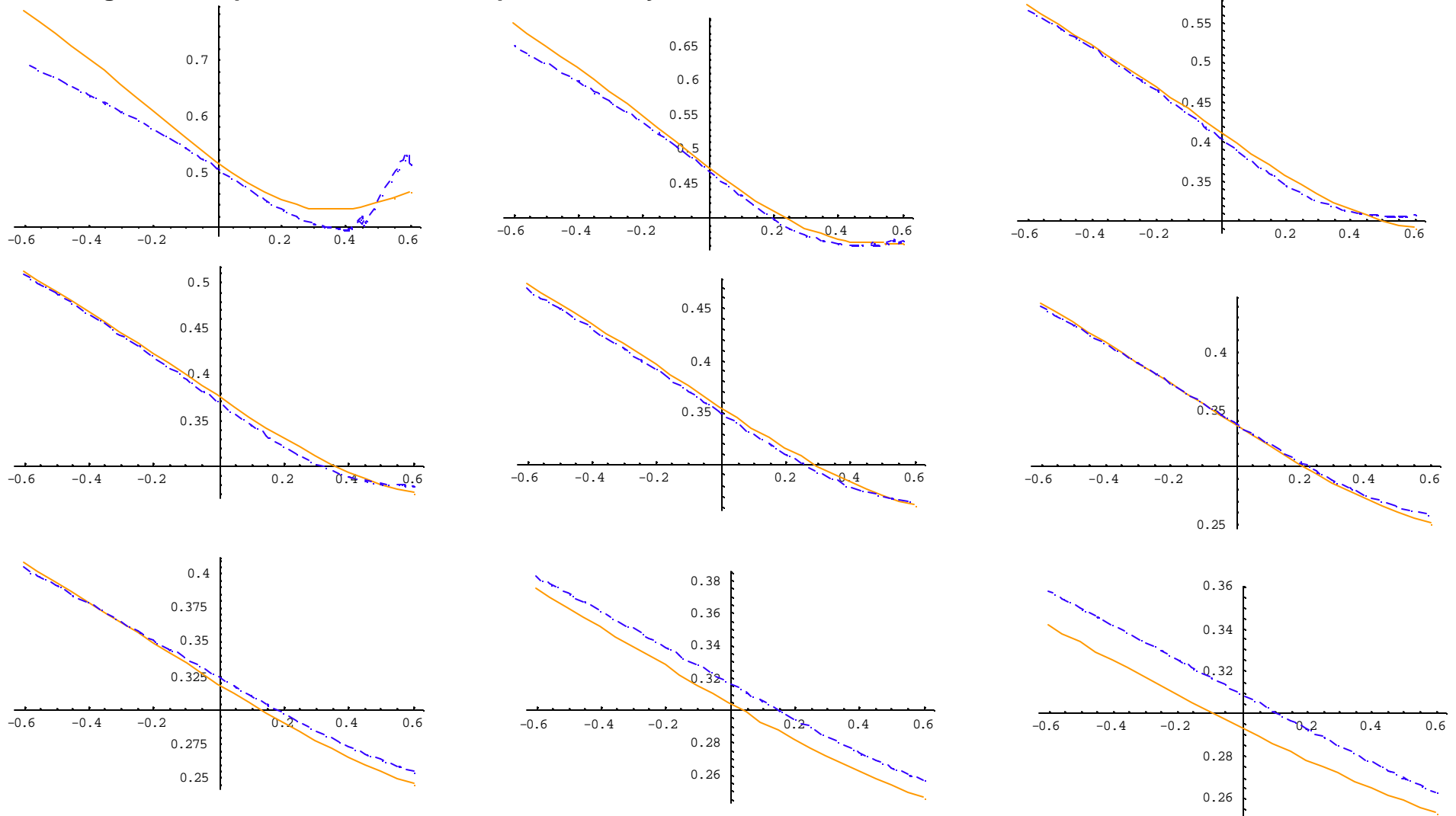
Fitting STOXX50 volatilities using Heston and SVJ

- Heston and SVJ models were fitted to SX5E implied volatility data from 18-Oct-2002.
- Heston and SVJ parameter definitions are as follows:
 - v instantaneous variance
 - $vbar$ long-term mean variance
 - κ rate of mean reversion of variance
 - h volatility of volatility
 - ρ correlation between stock and volatility changes
 - λ rate of arrival of jumps
 - a mean log-jump size
 - d standard deviation of log-jump size

Heston Fits

$\lambda \textcircled{R} 3.9, r \textcircled{R} -.7, h \textcircled{R} 1.15, v \textcircled{R} 0.54^2, vbar \textcircled{R} .31^2$

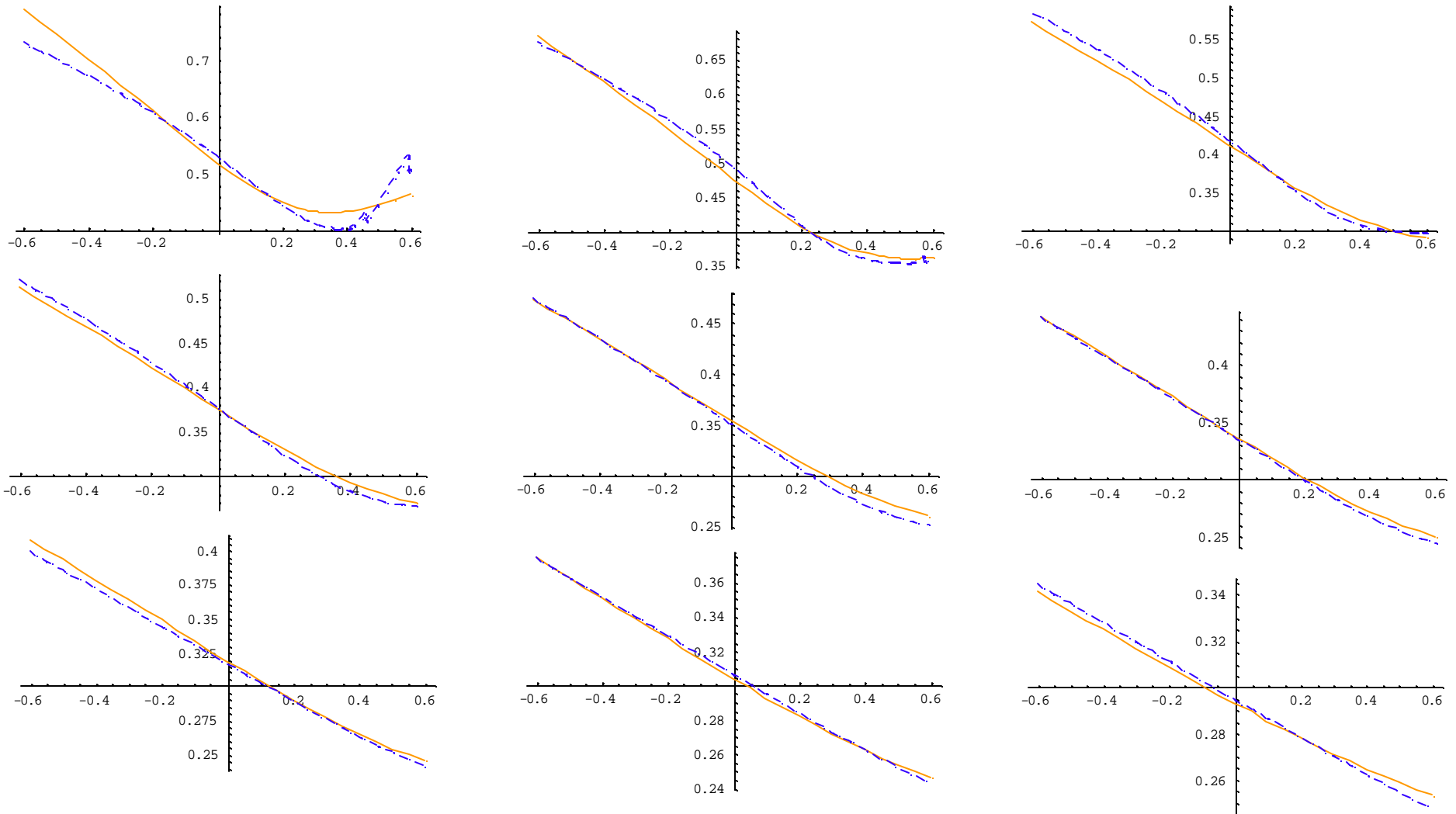
The orange line is a parametric fit to the implied volatility data; the dashed blue line is the Heston fit.



SVJ Fits

$l \text{ ® } 3.5, r \text{ ® } -.75, h \text{ ® } 1.15, v \text{ ® } 0.55^2, \text{vbar} \text{ ® } .25^2, l_J \text{ ® } 1, a \text{ ® } -.1, d \text{ ® } .1$

The orange line is a parametric fit to the implied volatility data; the dashed blue line is the SVJ fit.



Observations on the fits

- We note that the Heston model cannot fit volatility levels at all expirations simultaneously nor can it fit the skew at the front.
- On the other hand, the SVJ model allows us to fit the short term skew and also more or less all of the at-the-money volatility levels.
- Both fits return volatility-of-volatility parameters that are larger than seem reasonable.

Volatility skew from the characteristic function

- We can compute the volatility skew directly if we know the characteristic function $\mathbf{f}_T(u) = \mathbb{E}[e^{iu x_T}]$.
- The volatility skew is given by the formula (Gatheral (2002)):

$$\left. \frac{\partial \mathbf{s}_{BS}}{\partial k} \right|_{k=0} = -e^{-\mathbf{s}_{BS}^2 T / 8} \sqrt{\frac{2}{p}} \frac{1}{\sqrt{T}} \int_0^\infty du \frac{u \operatorname{Im}[\mathbf{f}_T(u - i/2)]}{u^2 + 1/4}$$

- The Heston characteristic function is given by

$$\mathbf{f}_T^{SV}(u) = \exp\{C(u, T) \bar{v} + D(u, T) v\}$$

where $C(u, T)$ and $D(u, T)$ are the familiar Heston coefficients with parameters $\mathbf{l}, \mathbf{h}, \mathbf{r}$.

- Adding a jump in the stock price (SVJ) gives the characteristic function

$$\mathbf{f}_T^{SVJ}(u) = \mathbf{f}_T^{SV}(u) \mathbf{f}_T^J(u)$$

$$\text{with } \mathbf{f}_T^J(u) = \exp\left\{-iu \mathbf{l}_J T \left(e^{\mathbf{a} + \mathbf{d}^2 / 2} - 1\right) + \mathbf{l}_J T \left(e^{i \mathbf{a} u - u^2 \mathbf{d}^2 / 2} - 1\right)\right\}$$

SVJJ

- Adding a simultaneous jump in the volatility gives the characteristic function (Matytsin (2000)):

$$\mathbf{f}_T^{SVJ}(u) = \mathbf{f}_T^{SV}(u) \mathbf{f}_T^J(u) \mathbf{f}_T^{J_v}(u)$$

with $\mathbf{f}_T^{J_v}(u) = \exp\left\{\bar{v} I_J T \left(e^{i a u - u^2 d^2 / 2} I(u) - 1\right)\right\}$

where $I(u) = \frac{1}{T} \int_0^T dt e^{\mathbf{g}_v D(u, T)} = -\frac{2 \mathbf{g}_v}{p_+ p_-} \int_0^{-\mathbf{g}_v D(u, T)} \frac{e^{-z} dz}{(1 + z / p_+)(1 + z / p_-)}$

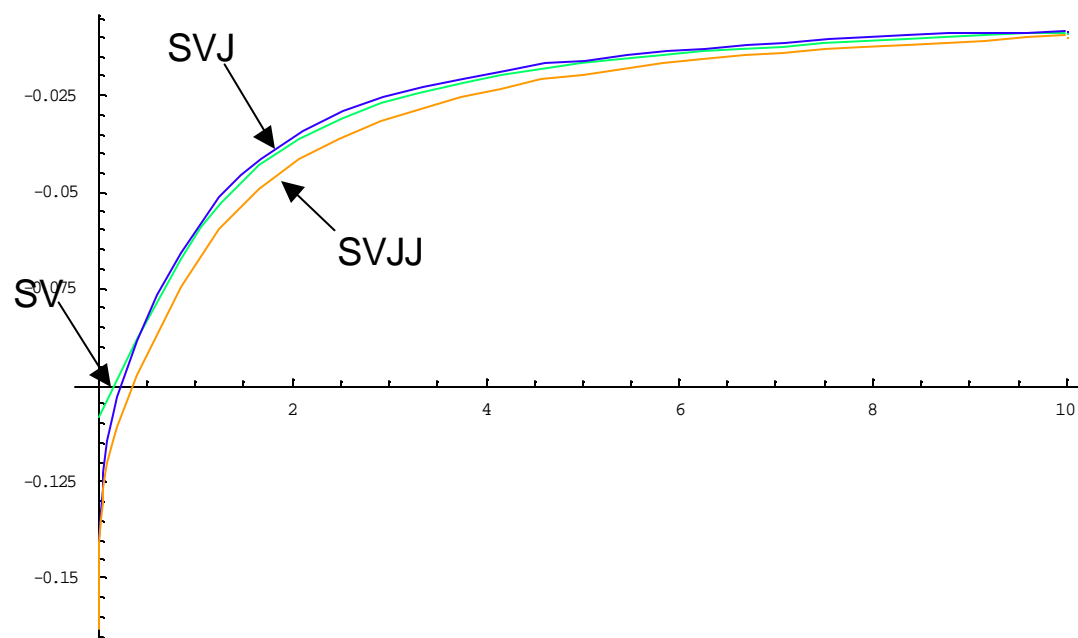
with $p_{\pm} = \frac{\mathbf{g}_v}{h^2} \{b - \mathbf{r} h i u \pm d\}$ (usual Heston notation)

Comparing ATM skews from different models

- With parameters

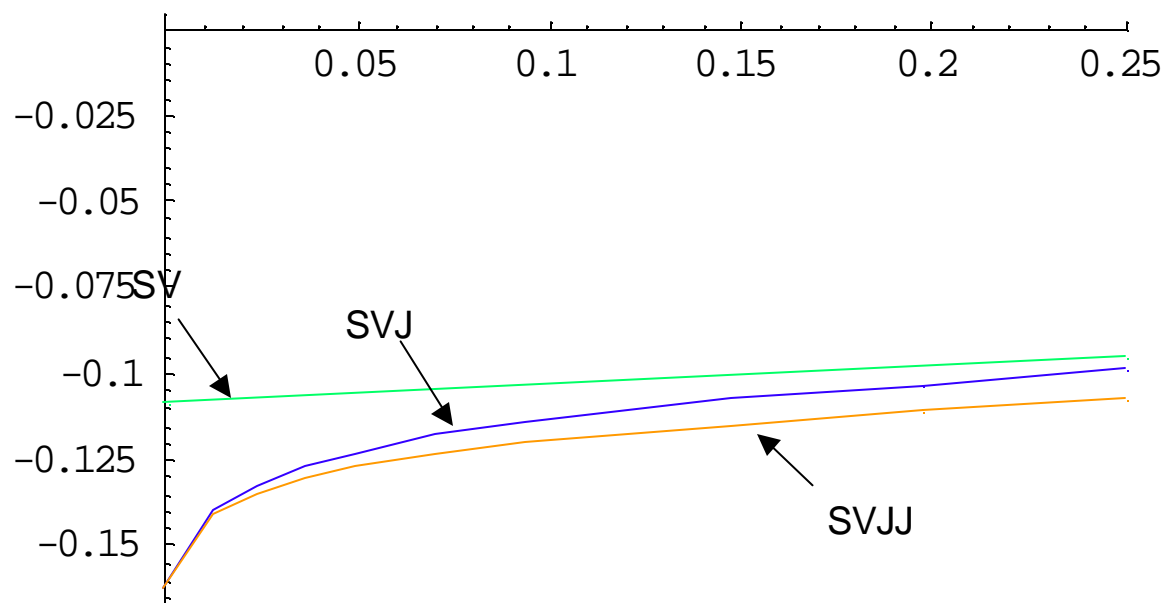
$l \approx 2.03$, $r \approx -0.57$, $h \approx 0.38$, $v \approx 0.1$, $\dot{v} \approx 0.04$, $l_J \approx 0.59$, $a \approx -0.05$, $d \approx 0.07$, $g_v \approx 0.1$

we get the following plots of ATM variance skew vs time to expiration



Short expiration detail

- SV and SVJ skews essentially differ only for very short expirations



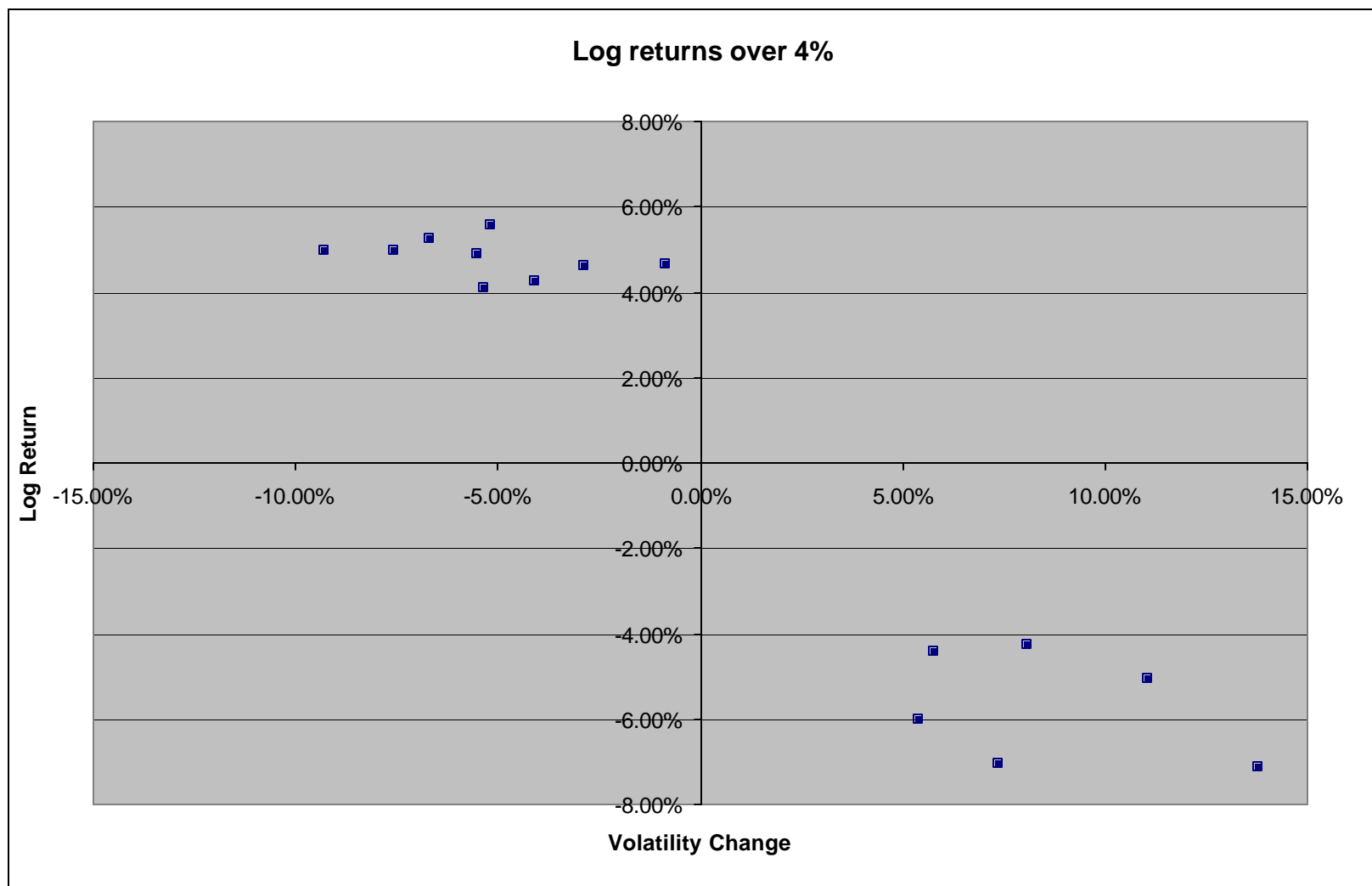
Other problems for diffusion models

- If the underlying stochastic process for the stock is a diffusion, we should be able to get from the statistical measure to the risk neutral measure using Girsanov's Theorem
 - This change of measure preserves volatility of volatility.
- However, historical volatility of volatility is significantly lower than the SV fitted parameter.
- Finally, in the few days prior to SPX expirations, out-of-the money option prices are completely inconsistent with the diffusion assumption
 - For example a 5 cent bid for a contract 10 standard deviations out-of-the-money.

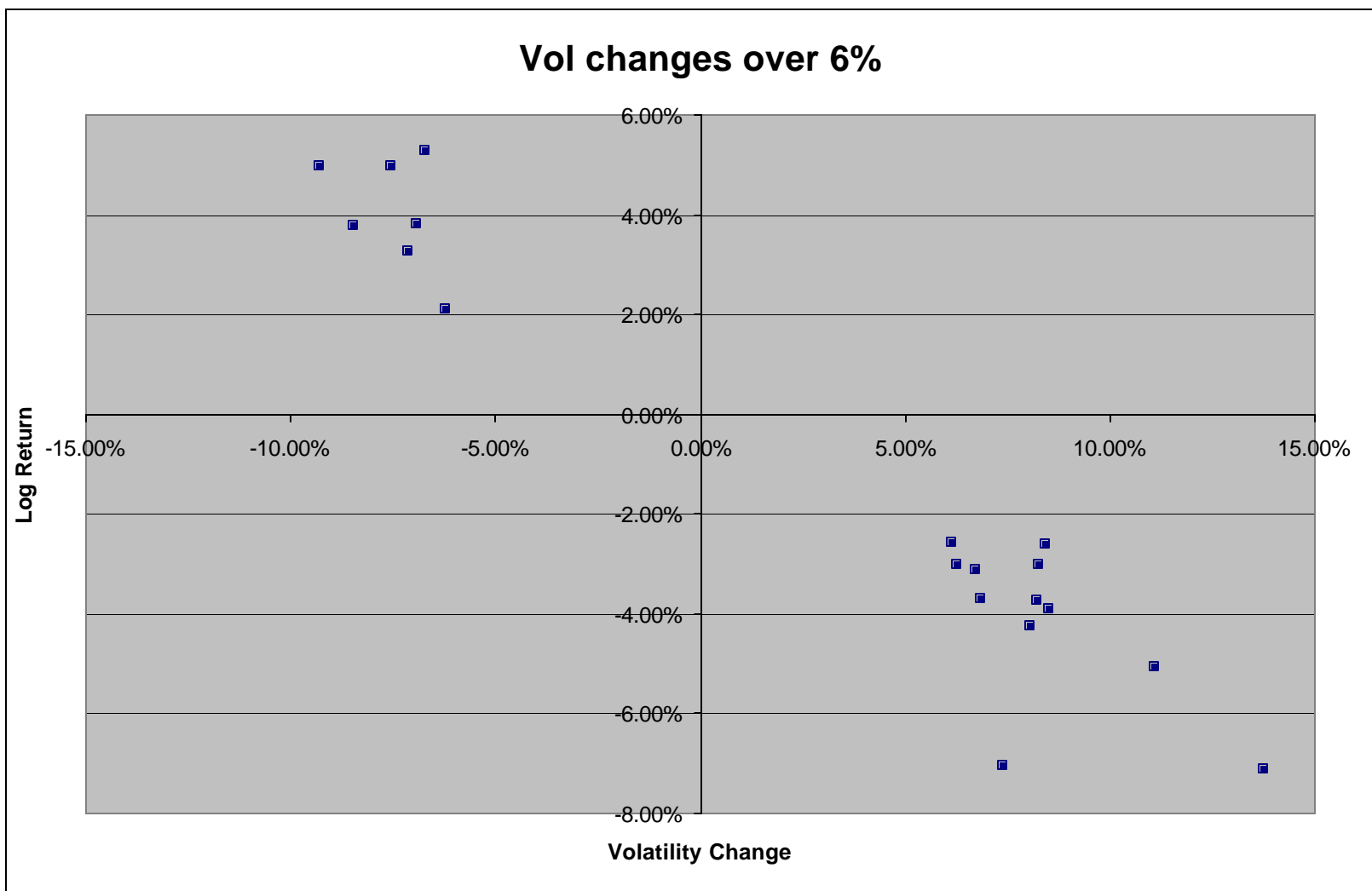
Simple jump diffusion models don't work either

- Although jumps may be necessary to explain very short dated volatility skews, introducing jumps introduces more parameters and this is not necessarily a good thing. For example, Bakshi, Cao, and Chen (1997 and 2000) find that adding jumps to the Heston model has little effect on pricing or hedging longer-dated options and actually worsens hedging performance for short expirations (probably through overfitting).
- Different authors estimate wildly different jump parameters for simple jump diffusion models.

SPX large moves from 1/1/1990 to 4/25/2003



SPX large moves from 1/1/1990 to 4/25/2003



Empirical jump observations

- A large move in the SPX index is invariably accompanied by a large move in volatility
 - Volatility changes and log returns have opposite sign
- This is consistent with clustering
 - If there is a large move, more large moves follow *i.e.* volatility must jump

- We conclude that any jumps must be double jumps!

What about pure jump models?

- Dilip Madan and co-authors have written extensively on pure jump models with stable increments
 - The latest versions of these models involve subordinating a pure jump process to the integral of a CIR process – trading time again.
- Pure jump models are more aesthetically pleasing than SVJJ
 - The split between jumps and diffusion is somewhat *ad hoc* in SVJJ
- However, large jumps in the stock price don't force an increase in implied volatilities.

The instantaneous volatility impact of option trades

- Recall our simple market price impact model

$$\Delta x = S \sqrt{x}$$

where x is the number of days' volume represented by the trade.

- We can always either buy an option or replicate it by delta hedging until expiration: in equilibrium, implied volatilities should reflect this
- If we delta hedge, each rebalancing trade will move the price against us by

$$S \sqrt{\frac{\Delta d}{ADV}} = S \sqrt{\frac{n \Gamma \Delta S}{ADV}} = S \sqrt{x \Gamma} \sqrt{\Delta S}$$

- In the spirit of Leland (1985), we obtain the shifted expected instantaneous volatility

$$\hat{S} \approx S \left\{ 1 \pm 2 \sqrt{\frac{2}{p}} \sqrt{\frac{S \Gamma x}{\sqrt{d_t}}} \right\}$$

- What does this mean for the implied volatility?

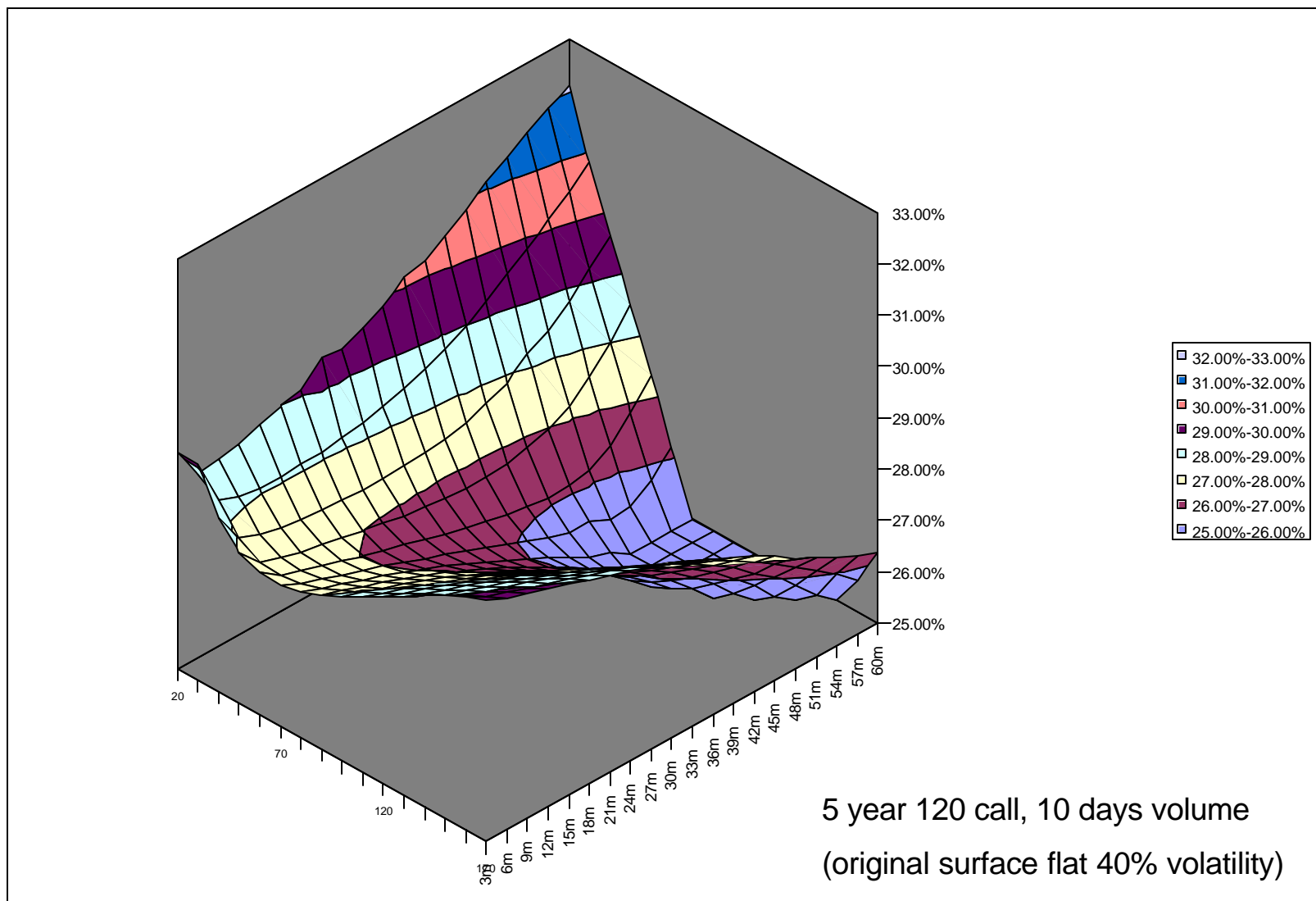
The implied volatility impact of an option trade

- To get implied volatility from instantaneous volatilities, integrate local variance along the most probable path from the current stock price today to the strike price at expiration (see Gatheral (2002))

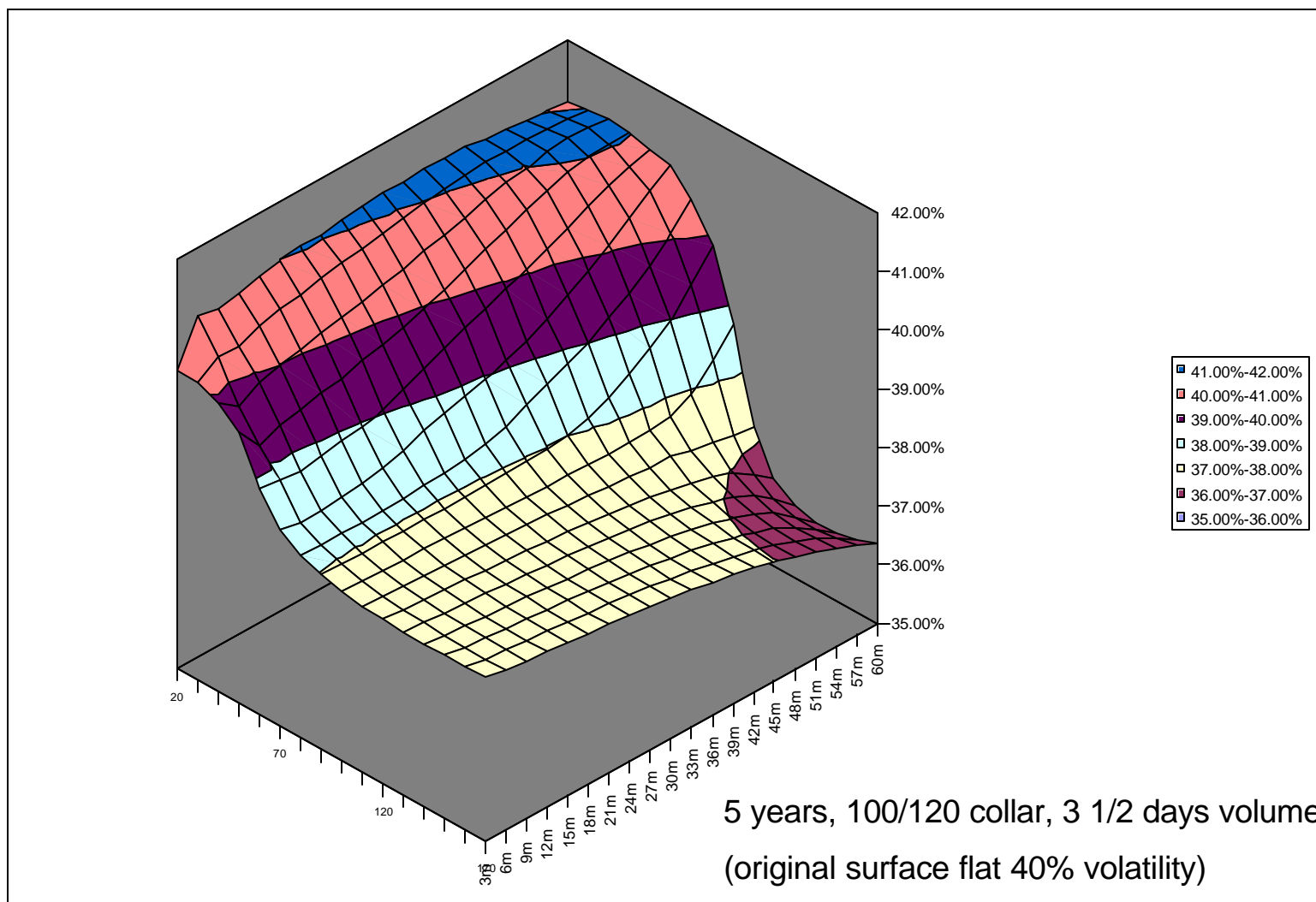
$$\sigma_{BS}^2(K, T) \approx \frac{1}{T} \int_0^T \sigma^2(\tilde{S}_t, t) dt$$

$$\text{with } \tilde{S}_t \approx S_0 \left(\frac{K}{S_0} \right)^{t/T}.$$

The volatility impact of a 5 year 120 strike call



The volatility impact of a 5 year 100/120 collar trade



Liquidity and the volatility surface

- We see that the shape of the implied volatility surface should reflect the structure of open delta-hedged option positions.
- In particular, if delta hedgers are structurally short puts and long calls, the skew will increase relative to a hypothetical market with no frictions.
 - Part of what we interpret as volatility of volatility when we fit stochastic volatility models to the market can be ascribed to liquidity effects.

How option prices reflect the behavior of stock prices

- Short-dated implied vol. more volatile than long-dated implied vol.
- Significant at-the-money skew
- Skew depends on volatility level
- Extreme short-dated implied volatility skews
- High implied volatility of volatility
- Clustering - mean reversion of volatility
- Anti-correlation of volatility moves and log returns
- Volatility of volatility increases with volatility level
- Jumps
- Stock volatility depends on the strikes and expirations of open delta-hedged options positions.

Conclusions

- Far from being *ad hoc*, stochastic volatility models are natural continuous time extensions of simple but realistic discrete-time models of stock trading.
- Stylized features of log returns can be related to empirically observed features of implied volatility surfaces.
- By carefully examining the various stylized features of option prices and log returns, we are led to reject all models except stochastic volatility with simultaneous index and volatility jumps.
- Investor risk preferences and liquidity effects also affect the observed volatility skew so SV-type models may be misspecified and fitted parameters unreasonable.

References

- Clark, Peter K., 1973, A subordinated stochastic process model with finite variance for speculative prices, *Econometrica* 41, 135-156.
- Duffie D., Pan J., and Singleton K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68, 1343-1376.
- Geman, Hélyette, and Thierry Ané, 2000, Order flow, transaction clock, and normality of asset returns, *The Journal of Finance* 55, 2259-2284.
- Gatheral, J. (2002). Case studies in financial modeling lecture notes.
http://www.math.nyu.edu/fellows_fin_math/gatheral/case_studies.html
- Heston, Steven (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies* 6, 327-343.
- Leland, Hayne (1985). Option Pricing and Replication with Transactions Costs. *Journal of Finance* 40, 1283-1301.
- Lewis, Alan R. (2000), *Option Valuation under Stochastic Volatility : with Mathematica Code*, Finance Press.
- Matytsin, Andrew (2000). Modeling volatility and volatility derivatives, Cirano Finance Day, Montreal.
<http://www.cirano.qc.ca/groupefinance/activites/fichiersfinance/matytsin.pdf>
- Wiggins J.B. (1987). Option values under stochastic volatility: theory and empirical estimates. *Journal of Financial Economics* 19, 351-372.