THE BUSINESS SCHOOL
FOR FINANCIAL MARKETS

The University of Reading

Unifying Volatility Models

Carol Alexander

(Co-authored works with E. Lazar and L. Nogueira)

ISMA Centre, University of Reading

Email: c.alexander@ismacentre.rdg.ac.uk

Pdfs: www.ismacentre.rdg.ac.uk/alexander

Newton Institute, Cambridge, May 2005

I

Stochastic Local Volatility

Model Acronyms

- BS: Black-Scholes
- SV: Stochastic Volatility
- LV: Local Volatility
- SD: Sticky Delta
- ST: Sticky Tree
- SIV: Scale Invariant Volatility
- SLV: Stochastic Local Volatility
- SABR: Stochastic Alpha-Beta-Rho
- CEV: Constant Elasticity of Variance
- MM: Market Model of implied volatility

Notation

Claim Characteristics :

k: claim strike

T: maturity of claim

Model Price of Claim :

$$f_{k,T} = f_{k,T}\left(s_0, \sigma_0, \boldsymbol{\lambda}_0\right)$$

s_0: underlying price at time t_0

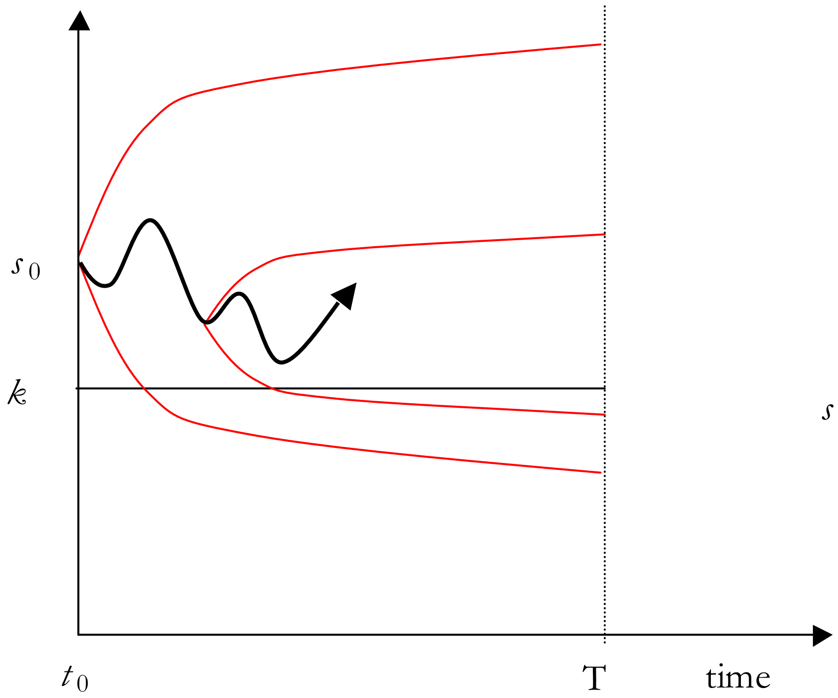
σ_0: instantaneous volatility at time t_0

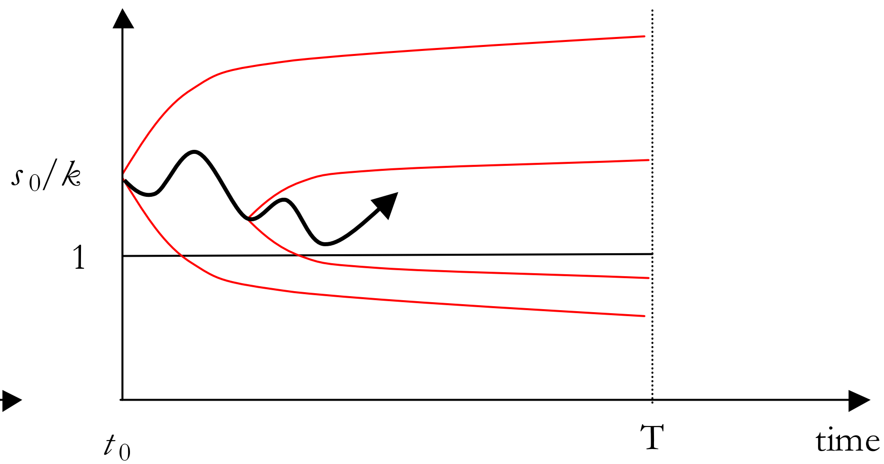
$\boldsymbol{\lambda}_0$: vector of volatility parameters at time t_0

t_0: calibration time

Scale Invariant Volatility (SIV)

$$f_{1,T}\left(\frac{s_0}{k},\sigma_0,\boldsymbol{\lambda}_0\right)=\frac{1}{k}f_{k,T}\left(s_0,\sigma_0,\boldsymbol{\lambda}_0\right)$$





Which Models are Scale Invariant?

- Most **stochastic volatility** models: even if price and/or volatility driven by Levy process
- **Local volatility** models: but *only* sticky delta models, not sticky tree models
- All BS **mixture models**, e.g.:
  – Merton's (1976) jump diffusion
  – Brigo-Mercurio's (2001) normal mixture diffusion

Example: Stochastic Volatility (SV)

- Heston (1993) **is** SIV:

$$\frac{dS}{S} = (r-q)\,dt + \sigma(t)\,dB$$

$$d(\sigma^2) = \left((1-\varsigma)\varrho\xi + \sqrt{\psi^2 - \varsigma(1-\varsigma)\xi^2}\right)(\omega - \sigma^2)\,dt + \xi\sigma\,dZ$$

$$< dB, dZ > = \varrho\,dt$$

- SABR Hagan *et al.* (2002) **is not** SIV (unless $\beta = 0$):

$$\frac{dS}{S} = (r-q)\,dt + \alpha S^{\beta}\,dB$$

$$\frac{d\alpha}{\alpha} = \zeta\,dZ$$

$$< dB, dZ > = \varrho\,dt$$

Example: Local Volatility (LV)

- Dupire (1993), Derman and Kani (1994), Rubinstein (1994):

$$\frac{dS}{S} = (r-q)\,dt + \sigma(t,S)\,dB$$

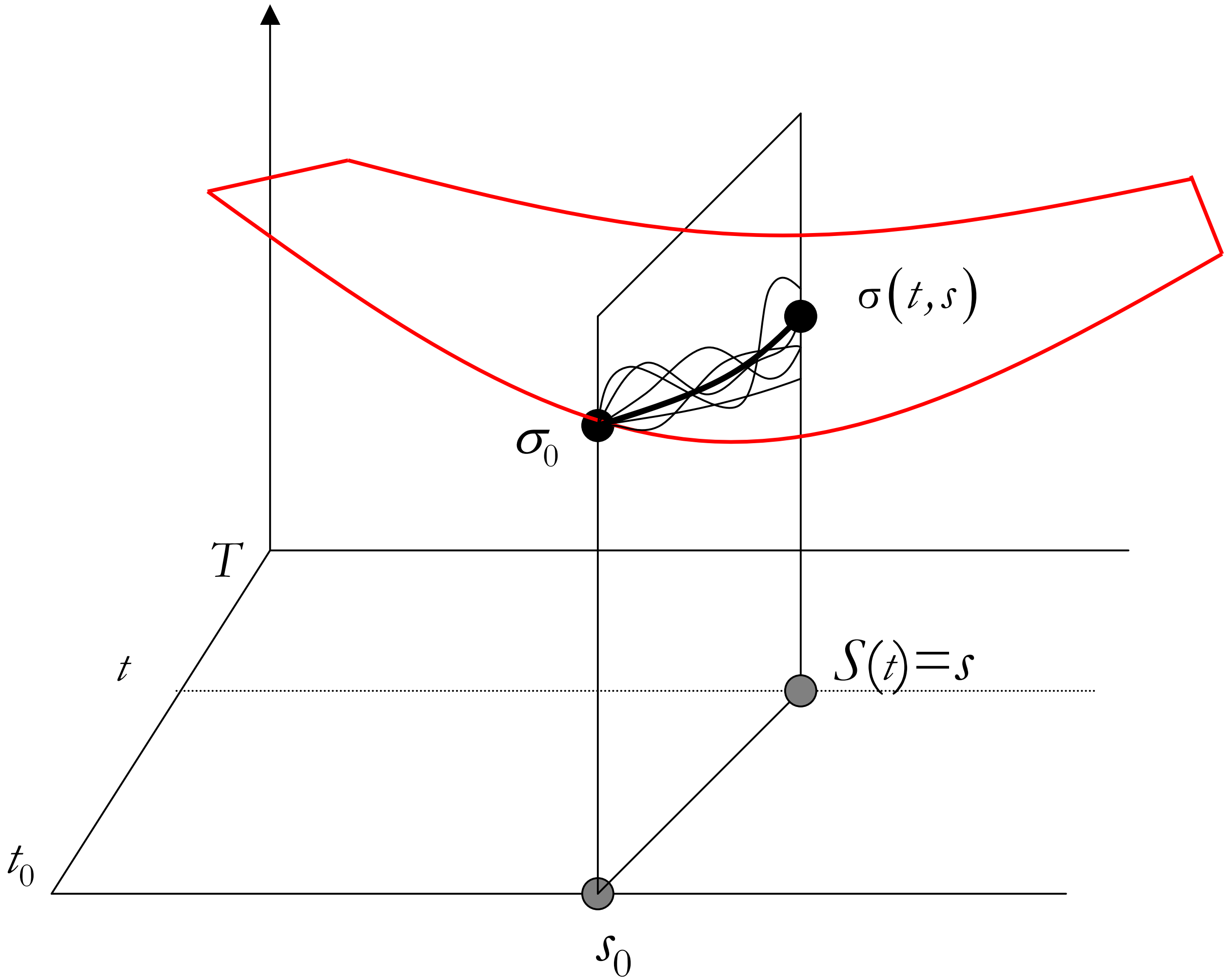
$$\sigma^2(t,s)\Big|_{t=T,s=k} = \frac{2\left(\frac{\partial f}{\partial T} + (r-q)k\frac{\partial f}{\partial k} + q f\right)}{k^2 \frac{\partial^2 f}{\partial k^2}}$$

- LV is SIV $\Leftrightarrow$ $\sigma(t,S)$ is a function of time and S/S_0 only.
- However many parametric deterministic functions $\sigma(t,S)$ have the *sticky delta* property that implied volatility is a function of delta and residual time to maturity only:

$$\theta_{k,T}(t,S) = \theta\left(\frac{S}{k}, T-t\right)$$

- 'Sticky Delta' (SD) models **are** SIV, 'Sticky Tree' (ST) models **are not** SIV.

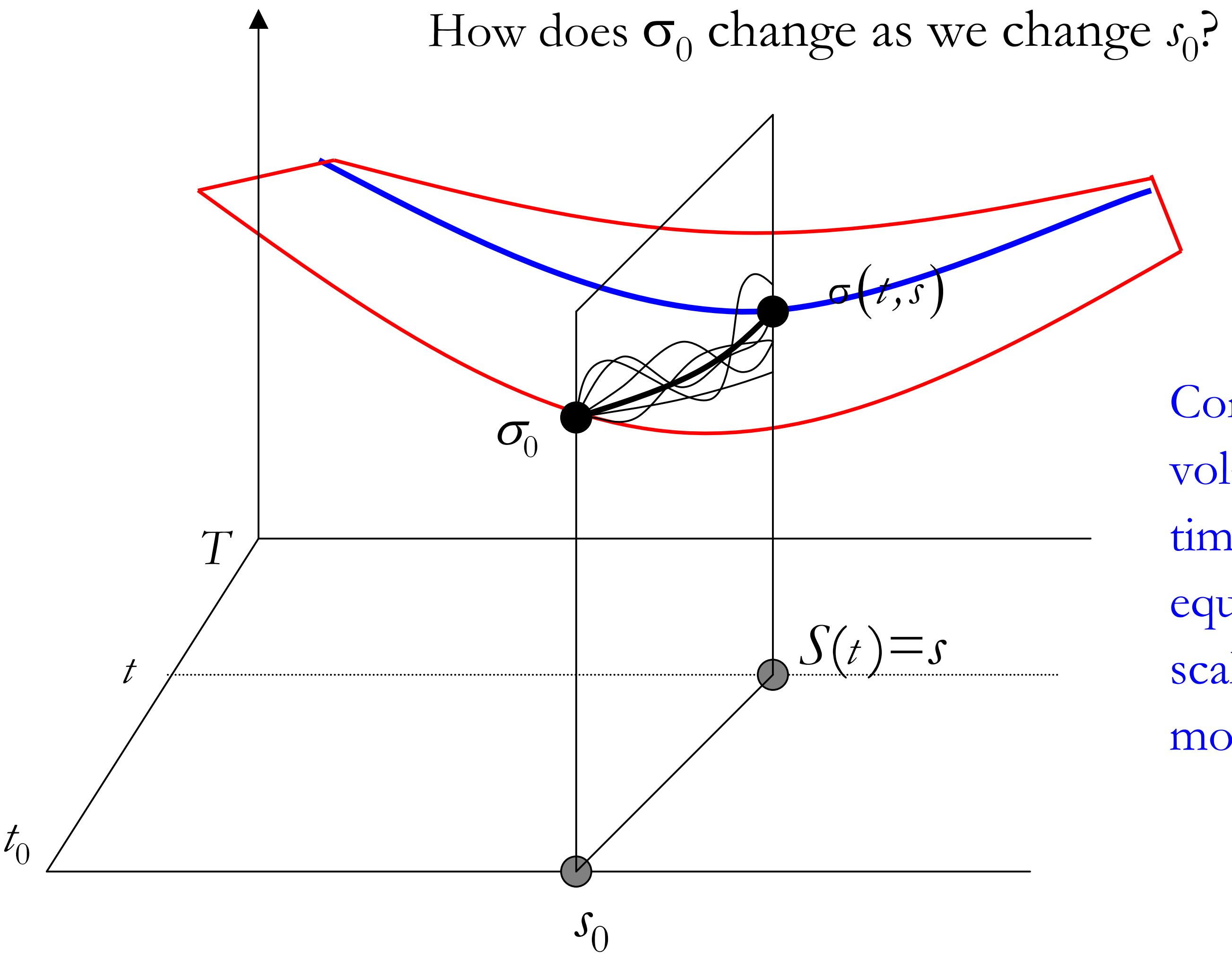
Relationship Between SV & LV



Equivalent
Local
Volatility
Surface

$$\sigma^2(t,s) = \mathrm{E}\left[\sigma^2(t) \,|\, S(t) = s\right]$$

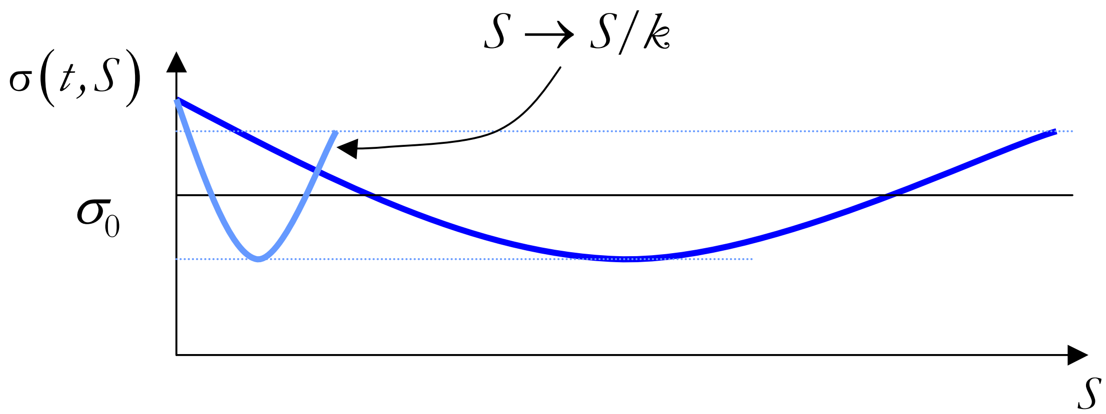
Scale Invariance of Local Volatility



Consider the local volatility curve at time t that is equivalent to a scale invariant SV model:

Independence of σ_0 and S_0

$\sigma_0^2 = E\left[\sigma^2(t)\right]$ is the **un**conditional expectation at t_0



- The local volatility depends only on the **ratio** S/S_0 (and time)
- Hence when we change the calibration value of S, the local volatility surface **moves as shown**
- Hence σ_0 is **independent** of S_0

Stochastic Local Volatility (SLV)

- Dupire (1996), Kani, Derman & Kamal (1997)
- Local variance calibrated at time $t_0 < t \equiv$ expectation of a *stochastic* instantaneous variance:

$$\sigma_L^2(t,\ s) = E^Q[\sigma^2(t,S,\mathbf{x}) \,|\, S(t)=s,\Im_0] = \int_{\Omega_t} \sigma^2(t,S,\mathbf{x})\, h_t(\mathbf{x} \,|\, S(t)=s,\Im_0)\, d\mathbf{x}$$

where $\mathbf{x} = \mathbf{x}(t, S)$ is a vector of all sources of uncertainty that influence the instantaneous volatility process.

Parametric Stochastic Local Volatility

- Alexander & Nogueira (2004) studies the general parametric form:

$$\frac{dS}{S} = (r-q)\,dt + \sigma(t,S,\boldsymbol{\lambda})\,dB$$

$$d\lambda_i = \mu_i(t,\boldsymbol{\lambda})\,dt + \zeta_i(t,\boldsymbol{\lambda})\,dZ_i$$

$$< dZ_i, dB > = \varrho_{iS}\,dt \qquad < dZ_i, dZ_j > = \varrho_{ij}\,dt$$

- Usual regularity conditions for drift and volatility in the parameter diffusions
- Strict no-arbitrage conditions

Example: CEV-SLV & SABR Model

- An example of a SLV model is the CEV-SLV:

$$\frac{dS}{S} = (r - q)\,dt + \alpha S^{\beta} dB$$

$$\frac{d\alpha}{\alpha} = \zeta dZ$$

$$< dB, dZ > = \varrho dt$$

- **<u>Same parameterization</u>** as SABR model!

- Different no-arbitrage conditions

Equivalence with Market Model (MM)

- Schönbucher (1999): 'market model' of implied volatilities

$$d\theta_{k,T} = \xi_{k,T}dt + \psi_{k,T}dB + \sum_{j=1}^{n} \zeta_{j,k,T}\, dZ_j$$

- This has the same implied volatility dynamics as the general SLV model

- Hence the SLV and MM models have the same **<u>same claim prices and hedge ratios</u>**

- But MM has <u>many</u> option specific parameters

II

Hedging:

Theoretical and Empirical Results

Hedging with SIV Models

Proposition :

When a model has been calibrated to the market implied volatilty surface:

$$f_{k,T}\left(S_0,\sigma_0,\boldsymbol{\lambda}_0\right)=k\,f_{1,T}\left(\frac{S_0}{k},\sigma_0,\boldsymbol{\lambda}_0\right)$$

$$\Updownarrow$$

$$\theta_{k,T}\left(S_0,\sigma_0,\boldsymbol{\lambda}_0\right)=\theta_{1,T}\left(\frac{S_0}{k},\sigma_0,\boldsymbol{\lambda}_0\right)$$

$\theta_{k,T}$: model implied volatility, calibrated at time t_0

S_0: underlying price at the calibration time

Proof

$$f_{k,T}\left(S_0,\sigma_0,\boldsymbol{\lambda}_0\right)=k\,f_{1,T}\left(\frac{S_0}{k},\sigma_0,\boldsymbol{\lambda}_0\right)$$

$\Updownarrow$ ← Euler's Theorem

$$f_{k,T}\left(S_0,\sigma_0,\boldsymbol{\lambda}_0\right)=S_0\frac{\partial f_{k,T}}{\partial S_0}+k\frac{\partial f_{k,T}}{\partial k}$$

$\Updownarrow$ ← Differentiate $f_{k,T}\equiv f_{k,T}^{BS}$ $w.r.t\ S_0$ and k

$$\frac{\partial\theta_{k,T}}{\partial S_0}=-\frac{k}{S_0}\frac{\partial\theta_{k,T}}{\partial k}$$

$\Updownarrow$ ←

$\Uparrow$: Trivial

$\Downarrow : f_{1,T}^{BS}\left(\frac{S_0}{k},\theta_{k,T}\right)\equiv f_{1,T}\left(\frac{S_0}{k},\sigma_0,\boldsymbol{\lambda}_0\right)$

see Fouque et al. (2000).

$$\theta_{k,T}=\theta_{1,T}\left(\frac{S_0}{k},\sigma_0,\boldsymbol{\lambda}_0\right)$$

SIV Partial Price Sensitivities are <u>Model Free</u>

$$f_{k,T}(S_0,\sigma_0,\boldsymbol{\lambda}_0) = S_0 \frac{\partial f_{k,T}}{\partial S_0} + k \frac{\partial f_{k,T}}{\partial k}$$

$$\Downarrow$$

$$\frac{\partial f_{k,T}}{\partial S_0} = S_0^{-1}\left(f_{k,T} - k\frac{\partial f_{k,T}}{\partial k}\right) \text{ and } \frac{\partial^2 f_{k,T}}{\partial S_0^2} = \left(\frac{k}{S_0}\right)^2 \frac{\partial^2 f_{k,T}}{\partial k^2}$$

- Partial price sensitivities obtained <u>**directly from market data**</u> !
- Hence the only differences in <u>**partial**</u> price sensitivities, between all SIV models, is in their fit to the smile !!!!

Relationship with BS Delta

- We can relate the first (and second) order partial price sensitivities to the BS delta (and gamma):
  - Calibration to smile means setting $f_{k,T} \equiv f_{k,T}^{BS}$
  - Differentiating this: $$\frac{\partial f_{k,T}}{\partial S_0} = \frac{\partial f_{k,T}^{BS}}{\partial S_0} + \frac{\partial f_{k,T}^{BS}}{\partial \theta_{k,T}} \frac{\partial \theta_{k,T}}{\partial S_0}$$
  - But in any SIV model $$\frac{\partial \theta_{k,T}}{\partial S_0} = -\frac{k}{S_0} \frac{\partial \theta_{k,T}}{\partial k}$$
  - Hence:

$$\frac{\partial f_{k,T}}{\partial S_0} = \delta_{k,T}^{BS} - \nu_{k,T}^{BS} \frac{k}{S_0} \frac{\partial \theta_{k,T}}{\partial k}$$ ← Slope of smile

Models with Model Free Delta

- In deterministic volatility models, delta ≡ partial price sensitivity
- This holds for SD and deterministic mixture models
- But these models are also SIV
- Hence their delta (and gamma) are model free.
- Conclusions:
  1. All of the SD and deterministic mixture models have the **same** delta (and gamma)
  2. This delta is **inappropriate** in most equity markets

Sticky Tree (ST) Delta

- Sticky tree models are **not** scale invariant, indeed

$$f_{k,T}^{ST}\left(S_0,\sigma_0,\boldsymbol{\lambda}_0\right)=S_0\frac{\partial f_{k,T}^{ST}}{\partial S_0}+k\frac{\partial f_{k,T}^{ST}}{\partial k}-S_0\frac{\partial f_{k,T}^{ST}}{\partial \sigma_0}\frac{\partial \sigma_0}{\partial S_0}$$

$$\Downarrow$$

$$\delta_{k,T}^{ST}=\delta_{k,T}^{BS}-\underbrace{\nu_{k,T}^{BS}\frac{k}{S_0}}_{\text{Model Free}}\frac{\partial \theta_{k,T}}{\partial k}+\underbrace{\frac{\partial f_{k,T}^{ST}}{\partial \sigma_0}\frac{\partial \sigma_0}{\partial S_0}}_{<0\text{ in Equity}}$$

Model Free   < 0 in Equity

SV Delta

- At time $t > t_0$: $\dfrac{df_{k,T}^{SV}(t,S,\sigma)}{dS} = \dfrac{\partial f_{k,T}^{SV}}{\partial S} + \dfrac{\partial f_{k,T}^{SV}}{\partial \sigma}\dfrac{\partial \sigma}{\partial S}$

- We define delta at the calibration time t_0 as:

$$\delta_{k,T}^{SV} = \lim_{t \downarrow t_0} \frac{df_{k,T}^{SV}(t,S,\sigma)}{dS} = \frac{\partial f_{k,T}^{SV}}{\partial S_0} + \left.\frac{\partial f_{k,T}^{SV}}{\partial \sigma}\right|_{t=t_0} \lim_{t \downarrow t_0} \frac{\partial \sigma}{\partial S}$$

- **Example**: Heston model

$$\delta_{k,T}^{SV} = S_0^{-1}\Bigg(\underbrace{f_{k,T}^{SV} - k\frac{\partial f_{k,T}^{SV}}{\partial k}}_{\text{Model Free}} + \underbrace{\xi\varrho\frac{\partial f_{k,T}^{SV}}{\partial \sigma_0^2}}_{<\,0\text{ in Equity}}\Bigg)$$

ξ: vol of vol

ϱ: price-vol correlation

SLV Delta

- Delta is the standard local volatility delta, but *adjusted* for stochastic movements in local volatility surface:

$$\delta_{k,T}^{SLV} = \delta_{k,T}^{LV} + \sum_i \frac{\beta_i \varrho_{i,S}}{\sigma S_0} \frac{\partial f_{k,T}^{LV}}{\partial \lambda_i}$$

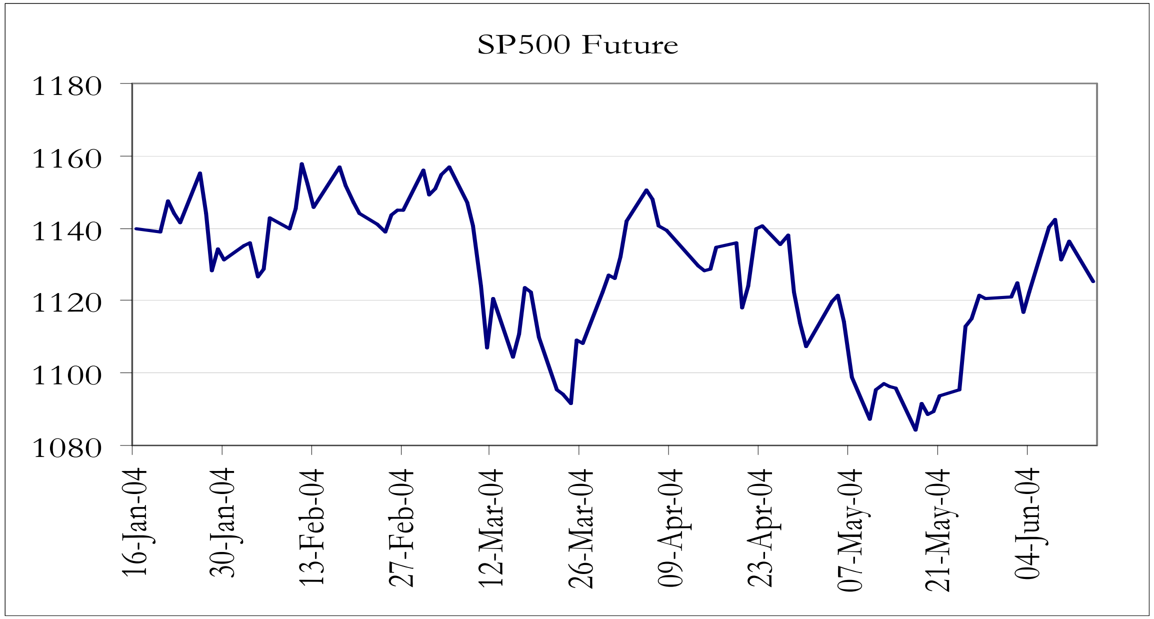
- Similar adjustments for gamma and theta
- Bias and efficiency? Both δ^{LV} and δ^{SLV} delta hedged portfolios have zero expectation but δ^{SLV} hedged portfolio has smaller hedging error variance.
- Hence SLV model may be useful for hedging

Summary of Deltas

Model	Model Delta	BS Delta – Model Delta
Sticky Delta & Mixtures	$S_0^{-1}\left(f_{k,T} - k\frac{\partial f_{k,T}}{\partial k}\right)$ **Model Free**	$\nu_{k,T}^{BS}\frac{k}{S_0}\frac{\partial \theta_{k,T}}{\partial k}$
Sticky Tree	$S_0^{-1}\left(f_{k,T} - k\frac{\partial f_{k,T}}{\partial k}\right) + \frac{\partial f_{k,T}}{\partial \sigma_0}\frac{\partial \sigma_0}{\partial s_0}$	$\nu_{k,T}^{BS}\frac{k}{S_0}\frac{\partial \theta_{k,T}}{\partial k} - \frac{\partial f_{k,T}}{\partial \sigma_0}\frac{\partial \sigma_0}{\partial s_0}$
Stochastic Volatility	$S_0^{-1}\left(f_{k,T} - k\frac{\partial f_{k,T}}{\partial k}\right) + \left.\frac{\partial f_{k,T}}{\partial \sigma}\right\rvert_{t=t_0} \lim_{t\downarrow t_0}\frac{\partial \sigma}{\partial S}$	$\nu_{k,T}^{BS}\frac{k}{S_0}\frac{\partial \theta_{k,T}}{\partial k} - \left.\frac{\partial f_{k,T}}{\partial \sigma}\right\rvert_{t=t_0} \lim_{t\downarrow t_0}\frac{\partial \sigma}{\partial S}$
Stochastic Local Volatility: Sticky Delta	$S_0^{-1}\left(f_{k,T} - k\frac{\partial f_{k,T}}{\partial k}\right) + \sum_i \frac{\zeta_i \varrho_{i,S}}{\sigma S_0}\frac{\partial f_{k,T}}{\partial \lambda_i}$	$\nu_{k,T}^{BS}\frac{k}{s_0}\frac{\partial \theta_{k,T}}{\partial k} - \sum_i \frac{\zeta_i \varrho_{i,S}}{\sigma S_0}\frac{\partial f_{k,T}}{\partial \lambda_i}$
Stochastic Local Volatility: Sticky Tree	$S_0^{-1}\left(f_{k,T} - k\frac{\partial f_{k,T}}{\partial k}\right) + \frac{\partial f_{k,T}}{\partial \sigma_0}\frac{\partial \sigma_0}{\partial S_0} + \sum_i \frac{\zeta_i \varrho_{i,S}}{\sigma S_0}\frac{\partial f_{k,T}}{\partial \lambda_i}$	$\nu_{k,T}^{BS}\frac{k}{S_0}\frac{\partial \theta_{k,T}}{\partial k} - \frac{\partial f_{k,T}}{\partial \sigma_0}\frac{\partial \sigma_0}{\partial s_0} - \sum_i \frac{\zeta_i \varrho_{i,S}}{\sigma S_0}\frac{\partial f_{k,T}}{\partial \lambda_i}$

Empirical Results

- Data on June 2004 European options on SP500:
  - Daily close prices from 02 Jan 2004 to 15 June 2004 (111 business days)
  - Strikes from 1005 to 1200 (i.e. 34 different strikes)
  - All strikes within ±10% of index level used for daily calibration



Hedging Race

7 Models

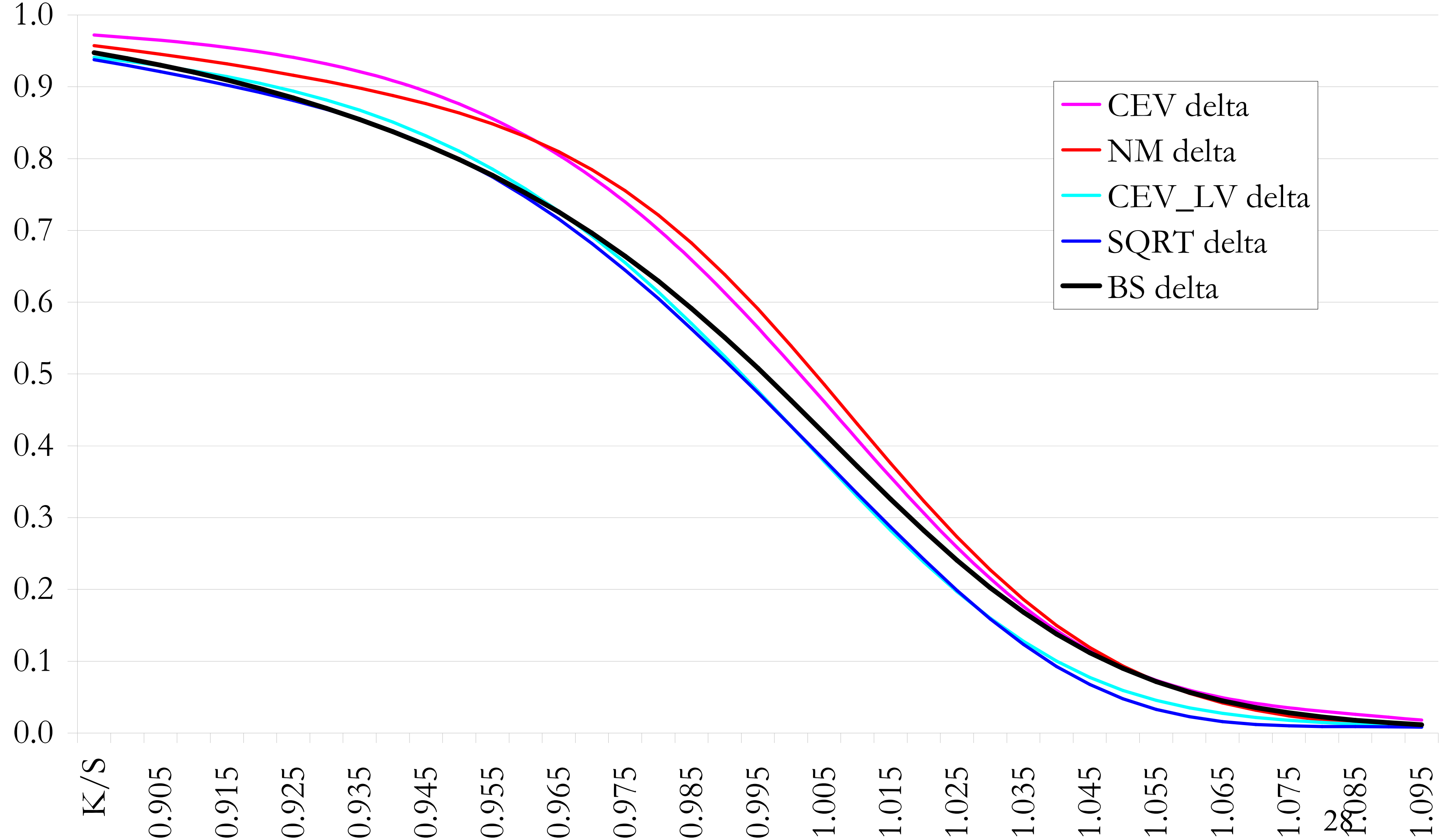
1. Black & Scholes (BS)
2. Sticky Delta (CEV)
3. Normal Mixture (NM)
4. Sticky Tree (CEV_LV)
5. Heston (SQRT)
6. SABR (CEV_SLV)
7. NM with stochastic parameters: (NM_SLV)

2 Strategies

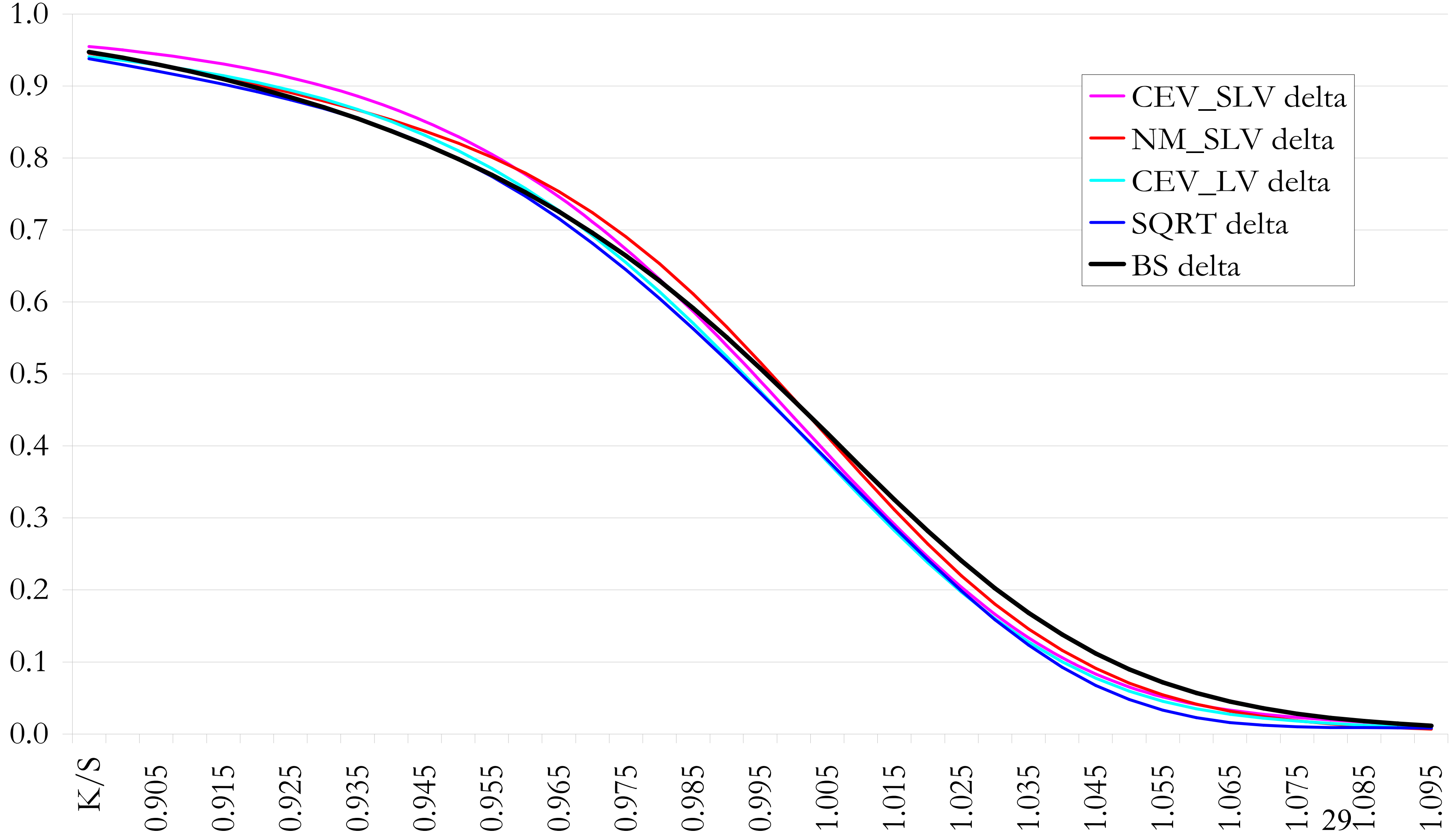
1. Delta Neutral: Short one call on each strike and delta hedge with underlying
2. Delta-Gamma Neutral: Buy 1125 strike to zero gamma then delta hedge

Rebalance daily from 16 Jan 2004 onwards (100 business days) assuming zero transactions costs

Comparison of Delta



Comparison of Delta



Hedged Portfolio P&L Stats

	Average	Std.Dev.	Skewness	XS Kurtosis
Delta Hedge				
CEV_ST	0.1462	0.5847	-0.3424	0.7820
SQRT	0.1370	0.6103	-0.5705	1.6737
CEV_SLV	0.1393	0.6280	-0.5701	1.6041
NM_SLV	0.1399	0.6329	-0.5224	1.2351
BS	0.1401	0.7451	-0.7029	2.0370
CEV	0.1367	1.1035	-0.6525	1.6691
SI	0.1446	1.1448	-0.5395	1.3708
NM	0.1373	1.1788	-0.5928	1.4834
Delta-Gamma Hedge				
BS	-0.0014	0.2612	-0.4353	2.5297
CEV_ST	0.0098	0.2691	-0.0291	3.0850
NM_SLV	0.0182	0.2773	0.1673	3.2069
SQRT	0.0111	0.2789	0.1929	3.6019
CEV_SLV	0.0154	0.2832	0.1304	4.1414
CEV	0.0382	0.4102	0.2501	4.0340
SI	0.0370	0.4205	0.1438	3.7003
NM	0.0428	0.4548	0.0208	4.0123

III

GARCH from Discrete to Continuous Time

GARCH(1,1)

Bollerslev (1986)

Generalized Engle's (1982) ARCH model to GARCH(1,1):

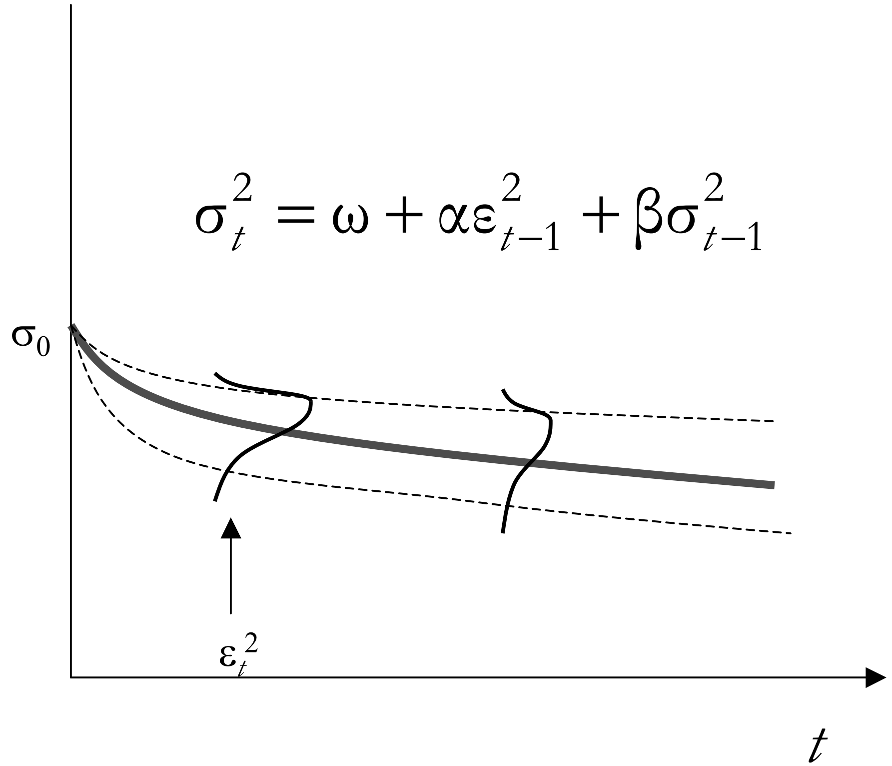
$$ln\left(S_t / S_{t-1}\right) = \mu + \varepsilon_t$$

$$\varepsilon_t \left| I_{t-1} \sim N(0, \sigma_t^2)\right.$$

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

$$\omega > 0, \alpha, \beta \geq 0, \alpha + \beta \leq 1$$

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$



GARCH Option Pricing

- Duan (1995), Siu *et al.* (2005) → Discrete time, numerical methods, Esscher transform
- Heston *&* Nandi (2000) → Closed form, continuous limit is Heston SV model with perfect price-volatility correlation

Continuous time limit of GARCH(1,1) → conflicting results

- Nelson (1990) → **Diffusion** limit model
- Corradi (2000) → **Deterministic** limit model

} **Which?**

GARCH Diffusion

Nelson (1990)

Continuous limit of GARCH(1,1)

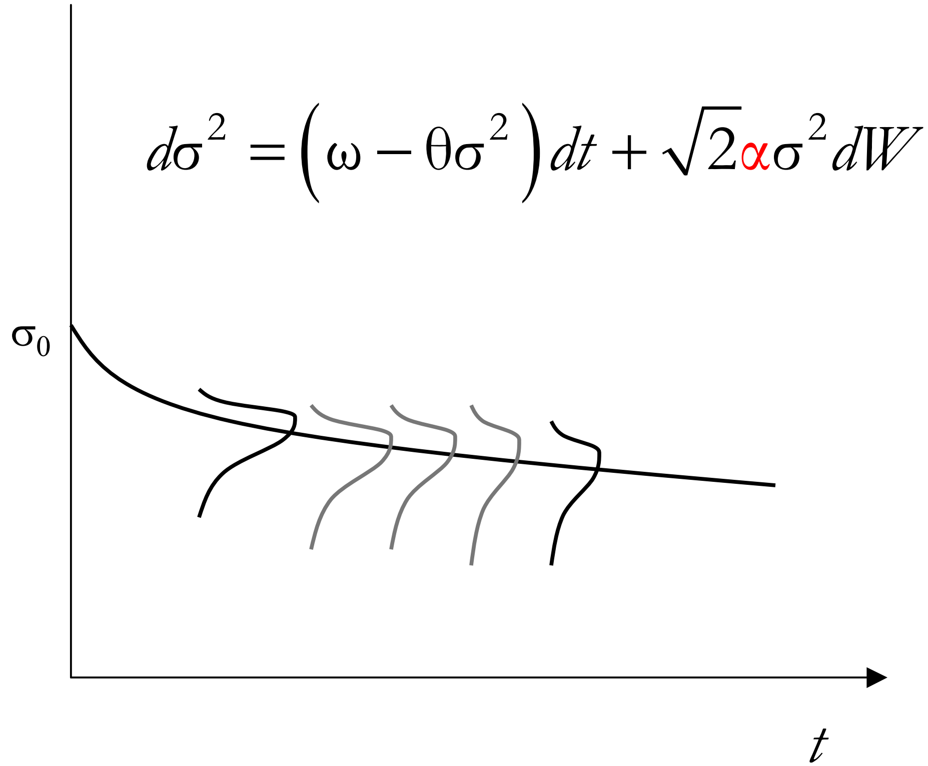
- Nelson assumed the following finite limits exist and are positive:

$$\omega = \lim_{h \downarrow 0}\left(\frac{{}_h\omega}{h}\right) \qquad \theta = \lim_{h \downarrow 0}\left(\frac{1 - {}_h\beta - {}_h\alpha}{h}\right)$$

$$\alpha^2 = \lim_{h \downarrow 0}\left(\frac{{}_h\alpha^2}{h}\right)$$

where ${}_h\omega$, ${}_h\alpha$, ${}_h\beta$ are the parameters using returns with step-length h

$$d\sigma^2 = \left(\omega - \theta\sigma^2\right)dt + \sqrt{2}\alpha\sigma^2 dW$$



Why No Diffusion?

1. GARCH has only **one source of uncertainty** in discrete time (Heston-Nandi, 2000)

2. The **discretization** of Nelson's limit model does not return the original GARCH model: it is ARV (Taylor, 1986)

3. GARCH diffusion and GARCH have **non-equivalent likelihood functions** (Wang, 2002)

4. Nelson's assumptions **cannot be extended** to other GARCH models

5. Nelson's assumptions imply that the GARCH model is **not time-aggregating** (Drost and Nijman, 1993). Only the assumptions made by Corradi (2000) are consistent with time-aggregation

6. **Related papers:** Brown *et al.* (2002) Duan *et al.* (2005)

Deterministic Limit?

Corradi (2000)

Continuous limit of GARCH(1,1)

(revisited)

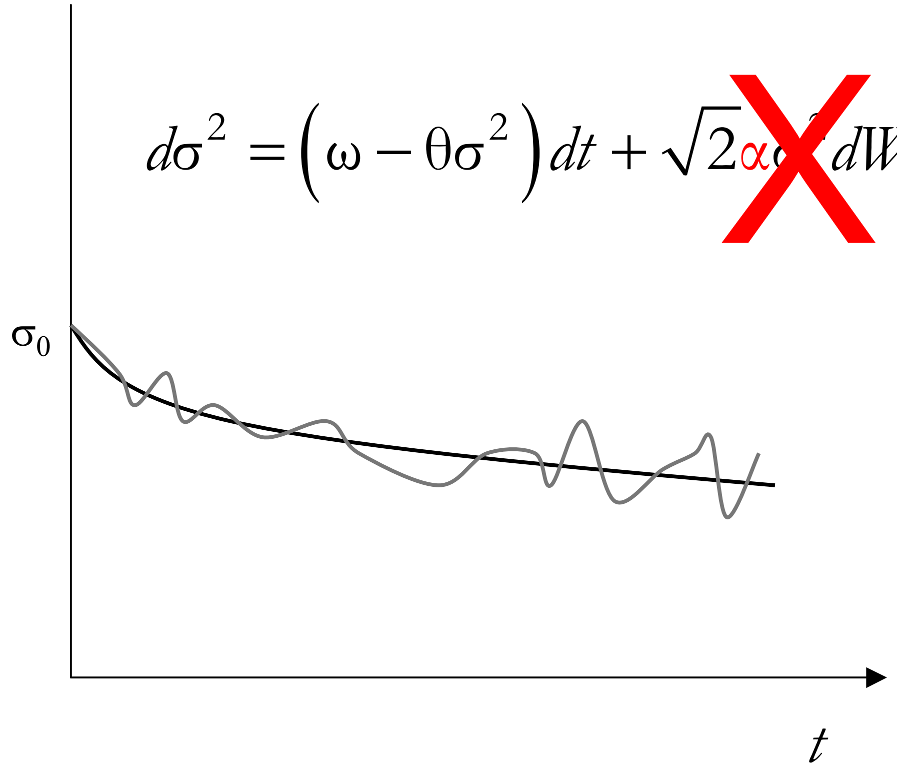
- Different assumptions:

$$\omega = \lim_{h \downarrow 0} \left(\frac{{}_h\omega}{h} \right) \qquad \psi = \lim_{h \downarrow 0} \left(\frac{1 - {}_h\beta}{h} \right)$$

$$\alpha = \lim_{h \downarrow 0} \left(\frac{{}_h\alpha}{h} \right)$$

where $\theta = \psi - \alpha$

$$d\sigma^2 = \left(\omega - \theta\sigma^2\right) dt + \sqrt{2}\alpha\sigma^2 dW$$



Multi-State GARCH

Alexander & Lazar (2004), Haas *et al.* (2004)

$$y_t = \frac{S_t - S_{t-1}}{S_{t-1}} \cong ln(S_t / S_{t-1}) \qquad y_t = \mu + \varepsilon_t$$

$$\boxed{\varepsilon_t \left| I_{t-1} \sim NM\left(\pi_1, \dots, \pi_K; \mu_1, \dots, \mu_K; \sigma_{1t}^2, \dots, \sigma_{Kt}^2\right)\right.} \quad \sum_{i=1}^{K} \pi_i = 1 \quad \sum_{i=1}^{K} \pi_i \mu_i = 0$$

GARCH(1,1): $\sigma_{it}^2 = \omega_i + \alpha_i \varepsilon_{t-1}^2 + \beta_i \sigma_{it-1}^2$

A-GARCH: $\sigma_{it}^2 = \omega_i + \alpha_i (\varepsilon_{t-1} - \lambda_i)^2 + \beta_i \sigma_{it-1}^2$

GJR: $\sigma_{it}^2 = \omega_i + \alpha_i \varepsilon_{t-1}^2 + \lambda_i d_{t-1}^- \varepsilon_{t-1}^2 + \beta_i \sigma_{it-1}^2$

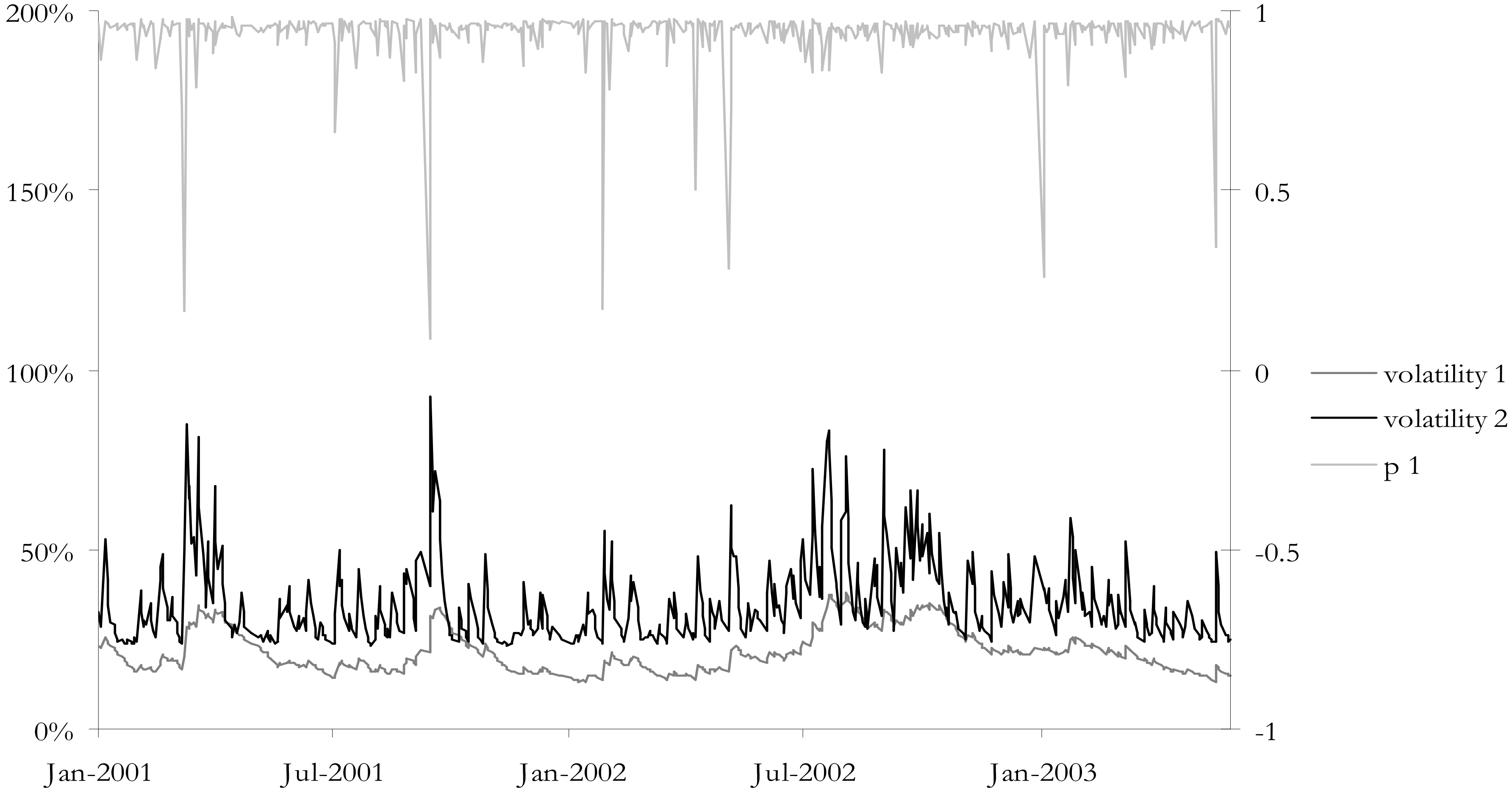
} for volatility states $i = 1, \dots, K$

+ parameter restrictions

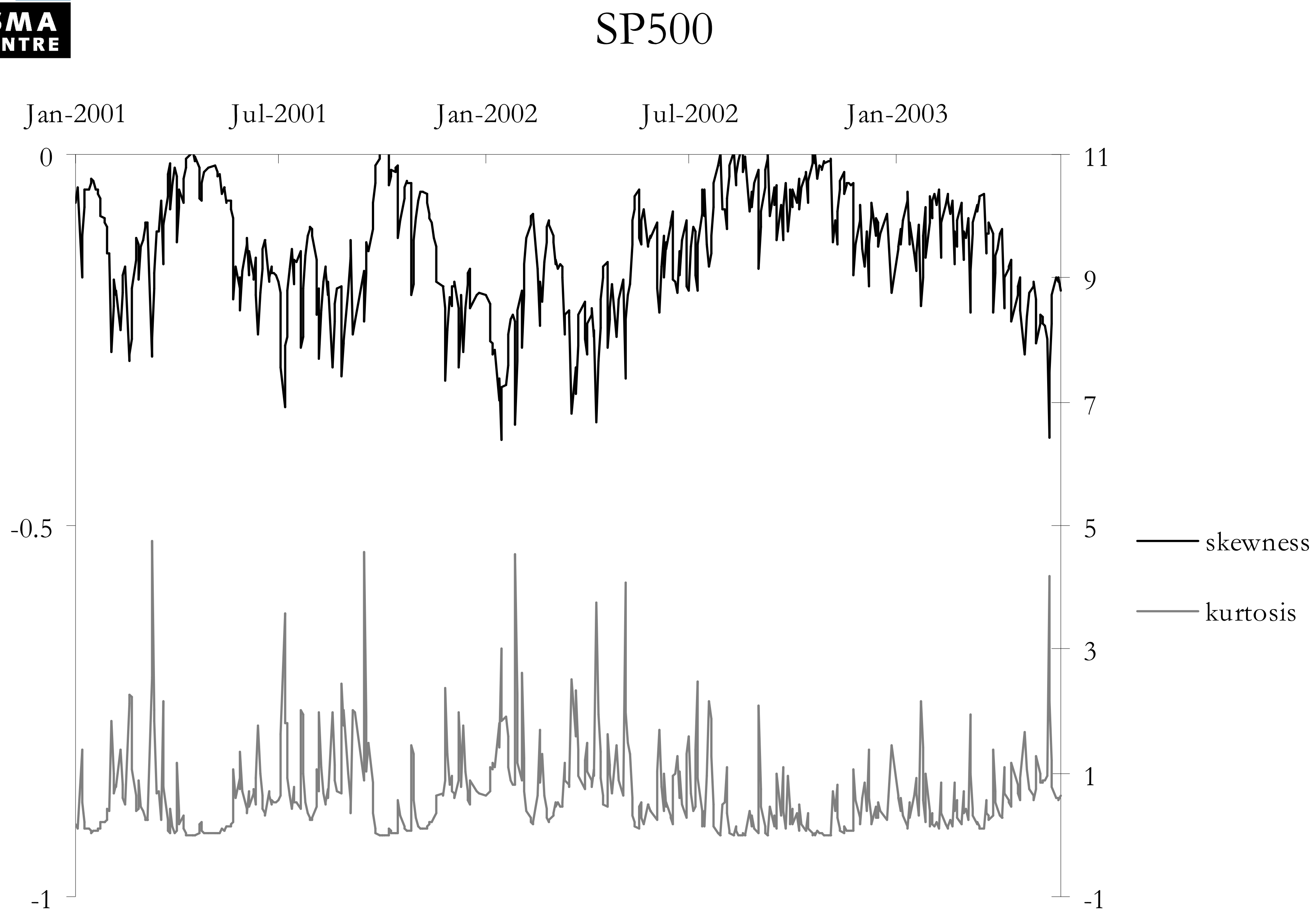
$$\sigma_t^2 = \sum_{i=1}^{K} \pi_i \sigma_{it}^2 + \sum_{i=1}^{K} \pi_i \mu_i^2$$

Volatilities and Mixing Law

SP500



Conditional Skewness and Kurtosis



Features of Multi-State GARCH

- Endogenous time varying higher moments ⇒ **endogenous** term structure in volatility skew surfaces
- Different component means ⇒ long-run **persistence** in volatility skew
- Different GARCH volatility components ⇒ regime specific **mean-reversion** of volatility
- Asymmetric GARCH components ⇒ regime specific **leverage effect** (slope of short term volatility skew)
- NM GARCH is **better fit** to historical data than single state GARCH models (Alexander & Lazar, 2004)

Volatility State Dependence

$$\mathbf{p}_t = \left(p_{1,t}, p_{2,t}, \ldots, p_{n,t} \right)' \ : p_{i,t} = \begin{cases} 1 \text{ if system is in state } i \ (s_t = i) \\ \text{otherwise} \end{cases}$$

- For continuous limit, introduce step-length h :

$$_h\mathbf{p}_{kh} = \left({}_h p_{1,hk}, {}_h p_{2,hk}, \ldots, {}_h p_{n,hk} \right)' \ : \ {}_h p_{i,kh} = \begin{cases} 1 \text{ if system is in state } i \ ({}_h s_{kh} = i) \\ \text{otherwise} \end{cases}$$

- Define K continuous volatility states:

$$\boldsymbol{\sigma}^2(t) = \left(\sigma_1^2(t), \sigma_2^2(t), \ldots, \sigma_K^2(t) \right)' \text{ with overall volatility } \sigma^2(t) = \sum_{i=1}^{K} \pi_i \sigma_i^2(t) + \sum_{i=1}^{K} \pi_i \mu_i^2$$

- Define the state indicator process:

$$\boldsymbol{p}(t) = \lim_{h \downarrow 0} {}_h\boldsymbol{p}_t \text{ where } {}_h\boldsymbol{p}_t = {}_h\boldsymbol{p}_{kh} \text{ for } t \in [\, kh, \ (k+1)h \,)$$

Transition Matrix

- **State probabilities** $\tilde{\mathbf{p}}$ are given by the (conditional) expectation of the state indicator vector:

$$\mathbf{P}\left({}_h s_{kh} = i \mid I_{(k-1)h}\right) = E\left({}_h \mathbf{p}_{kh} \mid I_{(k-1)h}\right) = {}_h\tilde{\mathbf{p}}_{kh}, \text{ say}$$

- Introduce state dependence via **transition matrix** for step-length *h*:

$${}_h q_{ij} = \mathrm{P}\left({}_h s_{kh} = i \,\middle|\, {}_h s_{(k-1)h} = j\right); {}_h\mathbf{Q} = \left({}_h q_{ij}\right) \text{ with } \sum_{i=1}^{K} {}_h q_{ij} = 1, \quad {}_h q_{ij} \geq 0$$

- Then $\quad {}_h\tilde{\mathbf{p}}_{kh} = {}_h\mathbf{Q}\; {}_h\tilde{\mathbf{p}}_{(k-1)h}$

  and $\quad {}_h\mathbf{Q}\,\boldsymbol{\pi} = \boldsymbol{\pi}$

- Where the unconditional state probabilities are the mixing law, $\boldsymbol{\pi}$

Transition Rate Matrix

- Define limit of state probabilities: $\tilde{\mathbf{p}}(t) = \lim_{h \downarrow 0} {}_h\tilde{\mathbf{p}}_t$ where ${}_h\tilde{\mathbf{p}}_t = {}_h\tilde{\mathbf{p}}_{kh}$
- And the **transition rate matrix:** $\mathbf{\Lambda} = \lim_{h \downarrow 0} \left(\frac{{}_h\mathbf{Q} - \mathbf{I}}{h} \right) = (\lambda_{ij})$
- Then the continuous time dynamics of the state probability vector in the physical measure are given by:

$$d\,\tilde{\mathbf{p}}(t) = \mathbf{\Lambda}\ \tilde{\mathbf{p}}(t)\,dt$$

- And the number of volatility jumps follow a Poisson process:

P(*k jumps from state j in an interval of length t*) = $\frac{(\lambda_j t)^k}{k!} exp(-\lambda_j t)$

where $\lambda_j = \sum_{i \neq j} \lambda_{ij}$ is the instantaneous probability of a jump from state *j*.

GARCH Jump

- Assume: $\omega_i = \lim_{h\downarrow 0}\left(\frac{{}_h\omega_i}{h}\right)$ $\alpha_i = \lim_{h\downarrow 0}\left(\frac{{}_h\alpha_i}{h}\right)$ $\psi_i = \lim_{h\downarrow 0}\left(\frac{1-{}_h\beta_i}{h}\right)$ $\mu_i = \lim_{h\downarrow 0}\left({}_h\mu_i\right)$
- Continuous limit of Markov Switching GARCH(1,1) is:

$$d\sigma^2 = \mathbf{p}'d\boldsymbol{\sigma}^2 + \boldsymbol{\sigma}^2{}'d\mathbf{p}$$

$$d\boldsymbol{\sigma}^2 = \left(\boldsymbol{\omega} + \boldsymbol{\alpha}\sigma^2 - \boldsymbol{\psi}\otimes\boldsymbol{\sigma}^2\right)dt$$

Risk neutral measure:

$$\frac{dS}{S} = rdt + \sigma dB^*$$

$$d\tilde{\mathbf{p}} = \boldsymbol{\Lambda}^* \ \tilde{\mathbf{p}}\ dt$$

$$\boldsymbol{\Lambda}^* = \boldsymbol{\Lambda}\otimes\left(\left(\mathbf{1}+\boldsymbol{\eta}\right)\mathbf{1}'\right)$$

$\boldsymbol{\eta}$: vector of conditional state risk premia

Physical measure:

$$\frac{dS}{S} = \mathbf{p}'\boldsymbol{\mu}dt + \sigma dB$$

$$d\tilde{\mathbf{p}} = \boldsymbol{\Lambda}\ \tilde{\mathbf{p}}\ dt$$

Example: Two Volatility States

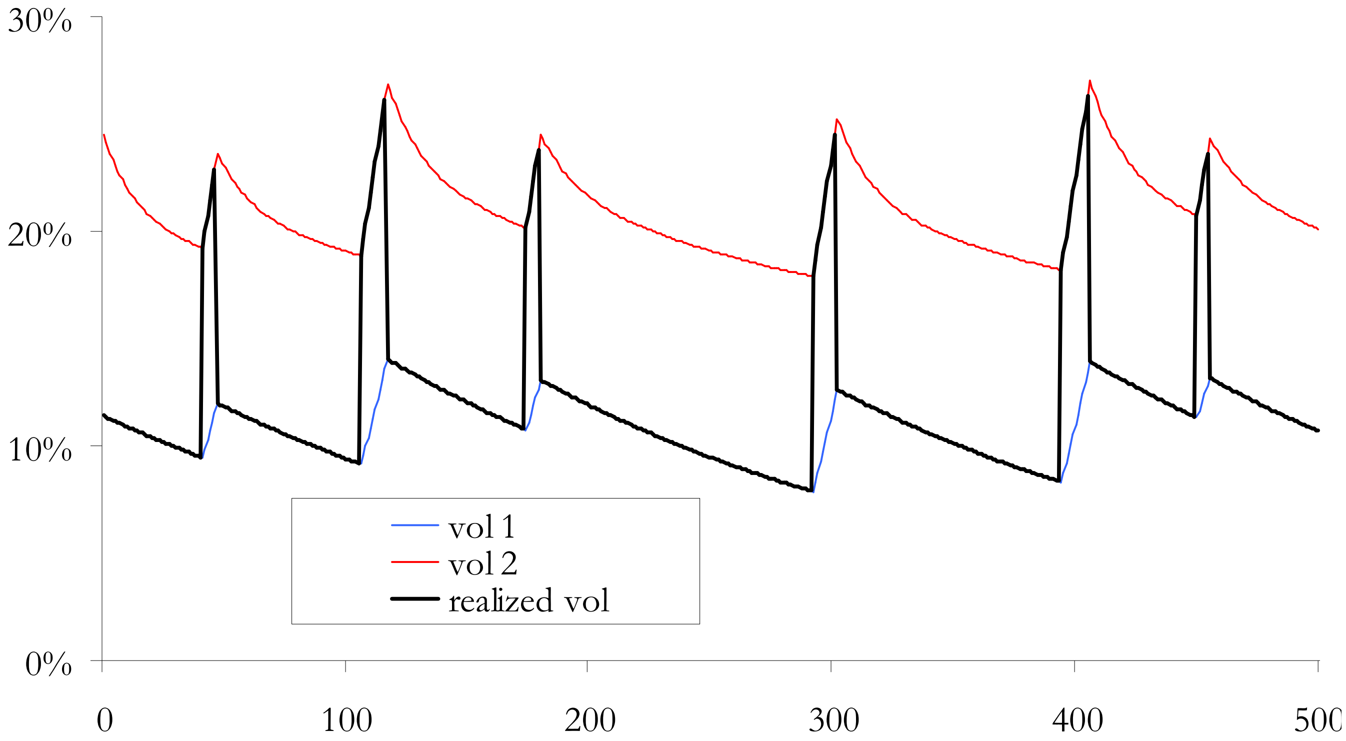
$$d\sigma^2 = p\left[\left(\omega_1 + \alpha_1\sigma^2 - \psi_1\sigma_1^2\right)dt\right] + (1-p)\left[\left(\omega_2 + \alpha_2\sigma^2 - \psi_2\sigma_2^2\right)dt\right] + \left(\sigma_1^2 - \sigma_2^2\right)dp$$

$$\mathbf{p}(t) = \left(p(t), 1 - p(t)\right)' \qquad p(t) = \begin{cases} 1 \text{ if } s(t) = 1 \\ 0 \text{ if } s(t) = 2 \end{cases}$$

$$dp(t) = \begin{cases} 1 \text{ if state changes from 2 to 1} \\ 0 \text{ if state does not change} \\ -1 \text{ if state changes from 1 to 2} \end{cases}$$

$$\sigma^2(t) = \begin{cases} \sigma_1^2(t) \text{ in state 1} \\ \sigma_2^2(t) \text{ in state 2} \end{cases}$$

Realised Volatility



Related Work

- Option pricing and hedging with Markov switching volatility processes:
  - **Naik (1993):** Theoretical results in continuous time, stochastic jump arrival rates, constant volatilities and perfectly correlated jumps in price
  - **Elliot** *et al.* **(2005):** MS version of GARCH model of Heston *&* Nandi (2000) based on Esscher transform

Acknowledgements

- **Co-authors:**
  - Emese Lazar, ISMA Centre
  - Leonardo Nogueira, ISMA Centre
- **Helpful discussions:**
  - Jacques Pézier, ISMA Centre
  - Hyungsok Ahn, Commerzbank London
  - Bruno Dupire, Bloomberg
  - Peter Carr, Bloomberg and NYU
  - Participants at Risk 2005 Europe and at Newton Institute, 2005.