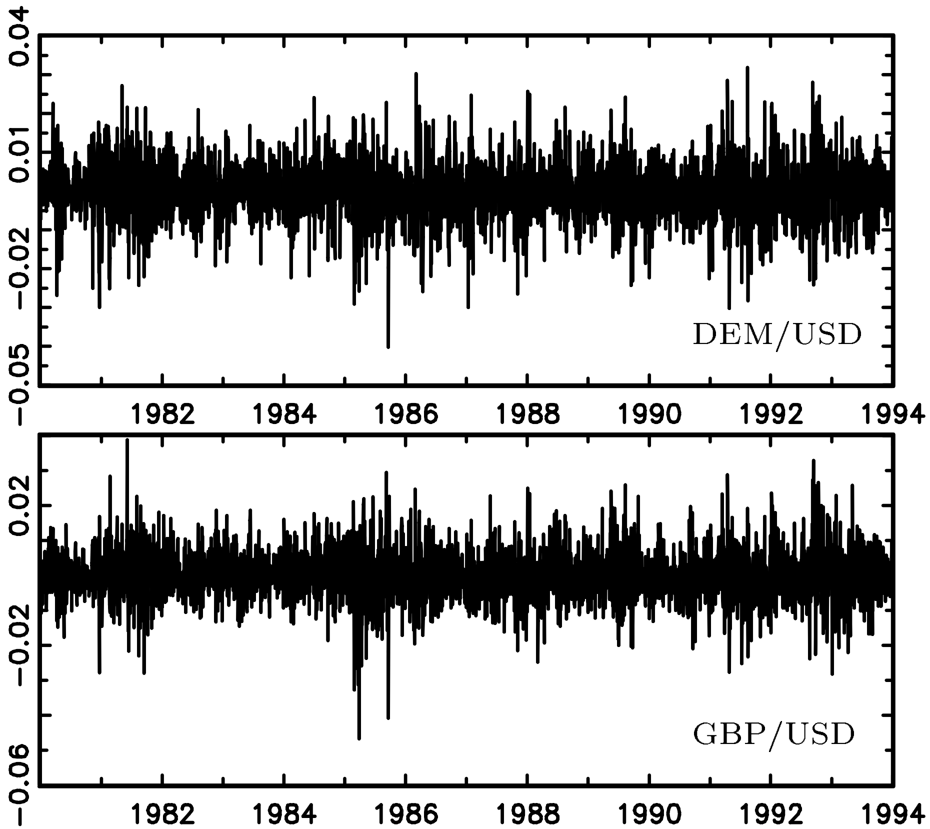
Time Inhomogeneous Multiple Volatility Modelling

Wolfgang HÄRDLE[1]
Helmut HERWARTZ[1]
Vladimir SPOKOINY[2]

[1] Center for Applied Statistics and Economics (CASE)
[2] Weierstrass Institute for Applied Analysis and Stochastics


0.04
0.01
−0.02
−0.05
DEM/USD
1982
1984
1986
1988
1990
1992
1994
0.02
−0.02
−0.06
GBP/USD
1982
1984
1986
1988
1990
1992
1994

Motivation

- Time varying volatility (Engle, 1982)
  - Interest rate parities
  - No arbitrage condition
  - Impact of global news (Engle, Ito and Lin, 1990)
  - Impact of market news (Braun, Nelson and Sunier, 1995)
- Time varying covariances
  - CAPM (Bollerslev, Engle and Wooldridge, 1988)
  - Backtesting

Parametric benchmark model

- Time homogeneous multivariate GARCH model

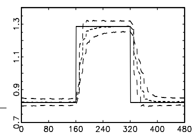
$$R_t = \Sigma_t^{1/2} \varepsilon_t$$

ε_t , $t \geq 1$, iid standard Gaussian in $\mathbb{R}^d$, $\Sigma_t \sim \mathcal{F}_{t-1}$

- vec model

$$\text{vech}(\Sigma_t) = c + \sum_{i=1}^{q} A_i \text{vech}(R_{t-i} R_{t-i}^\top) + \sum_{i=1}^{p} G_i \text{vech}(\Sigma_{t-i})$$

  - large parameter space
  - difficult to estimate (iterative numerical estimation)



Example

vec model: 21 parameters for $p = q = 1$, $d = 2$

$$
\begin{aligned}
\begin{bmatrix} \sigma_{11,t} \\ \sigma_{12,t} \\ \sigma_{22,t} \end{bmatrix}
&= \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}
+ \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}
\begin{bmatrix} r^2_{1,t-1} \\ r_{1,t-1} r_{2,t-1} \\ r^2_{2,t-1} \end{bmatrix} \\
&+ \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{bmatrix}
\begin{bmatrix} \sigma_{11,t-1} \\ \sigma_{12,t-1} \\ \sigma_{22,t-1} \end{bmatrix}
\end{aligned}
$$

BEKK model (Baba, Engle, Kraft and Kroner, 1992)

$$\Sigma_t = C_0^\top C_0 + \sum_{k=1}^{K}\sum_{i=1}^{q} A_{ki}^\top R_{t-i} R_{t-i}^\top A_{ki} + \sum_{k=1}^{K}\sum_{i=1}^{p} G_{ki}^\top \Sigma_{t-i} G_{ki}$$

Example

only 11 parameters for $K = p = q = 1$, $d = 2$

$$\begin{bmatrix} \sigma_{11,t} & \sigma_{12,t} \\ \sigma_{21,t} & \sigma_{22,t} \end{bmatrix} = \begin{bmatrix} \tilde{c}_{11} & \tilde{c}_{12} \\ \tilde{c}_{12} & \tilde{c}_{22} \end{bmatrix}$$

$$+ \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^\top \begin{bmatrix} r_{1,t-1}^2 & r_{1,t-1} r_{2,t-1} \\ r_{1,t-1} r_{2,t-1} & r_{2,t-1}^2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$+ \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}^\top \begin{bmatrix} \sigma_{11,t-1} & \sigma_{12,t-1} \\ \sigma_{21,t-1} & \sigma_{22,t-1} \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$$

An inhomogeneous model

- Time homogeneity

$$\Sigma_t \equiv \Sigma,\ t \leq T,\ \tilde{\Sigma} = \frac{1}{T}\sum_{t=1}^{T} R_t R_t^\top$$

- Local homogeneity

$$\Sigma_t = \Sigma_I,\ t \in I,\ I = [\tau - m, \tau[$$

Then

$$\tilde{\Sigma}_I = \frac{1}{|I|}\sum_{t \in I} R_t R_t^\top$$

- Interval I is unknown

Adaptive Modelling

- Data driven procedure to estimate I by $\hat{I}$
- local estimator of Σ_t

$$\hat{\Sigma}_t = \frac{1}{|\hat{I}|} \sum_{t \in \hat{I}} R_t R_t^\top$$

- Lepski (1990), Lepski and Spokoiny (1997), Spokoiny (1998)
- Compare adaptive procedure (local homogeneous) with global parametric benchmark

Outline of the talk

1. Motivation ✓
2. Adaptive modelling
3. Implementation
4. Monte Carlo results
5. Empirical Illustrations
6. Conclusions

Adaptive Modelling

- Semiparametric adaptive model:

$$\hat{\Sigma}_t = \frac{1}{|\hat{I}|} \sum_{t \in \hat{I}} R_t R_t^\top$$

- Issues:
  - Dimension reduction (for choice of $\hat{I}$)
  - Power transformation (for interval of homogeneity)
  - Global parameter selection (for fine tuning)

Dimension Reduction

- Finite family of unit vectors

$$\mathcal{W} = \{ w_1, \ldots, w_r \},\ w_i \in \mathbb{R}^d,\ r \le d$$

- Then

$$\begin{aligned} \mathsf{E}\left(|w^\top R_t|^2 \middle| \mathcal{F}_{t-1} \right) &= \mathsf{E}\left(w^\top R_t R_t^\top w \middle| \mathcal{F}_{t-1} \right) \\ &= w^\top \Sigma_t w \\ &= \sigma_{t,w}^2 \end{aligned}$$

Power Transformation

- Theoretical moments

$$
\begin{aligned}
\mathsf{E}\left(\left|w^\top R_t\right|^\gamma \left|\mathcal{F}_{t-1}\right.\right) &= C_\gamma \sigma_{t,w}^\gamma \\
\mathsf{E}\left(\left|w^\top R_t\right|^\gamma - C_\gamma \sigma_{t,w}^\gamma \left|\mathcal{F}_{t-1}\right.\right)^2 &= \sigma_{t,w}^{2\gamma} D_\gamma^2
\end{aligned}
$$

ξ standard Gaussian rv, $C_\gamma = \mathsf{E}\,|\xi|^\gamma$ and $D_\gamma^2 = \mathsf{Var}\,|\xi|^\gamma$

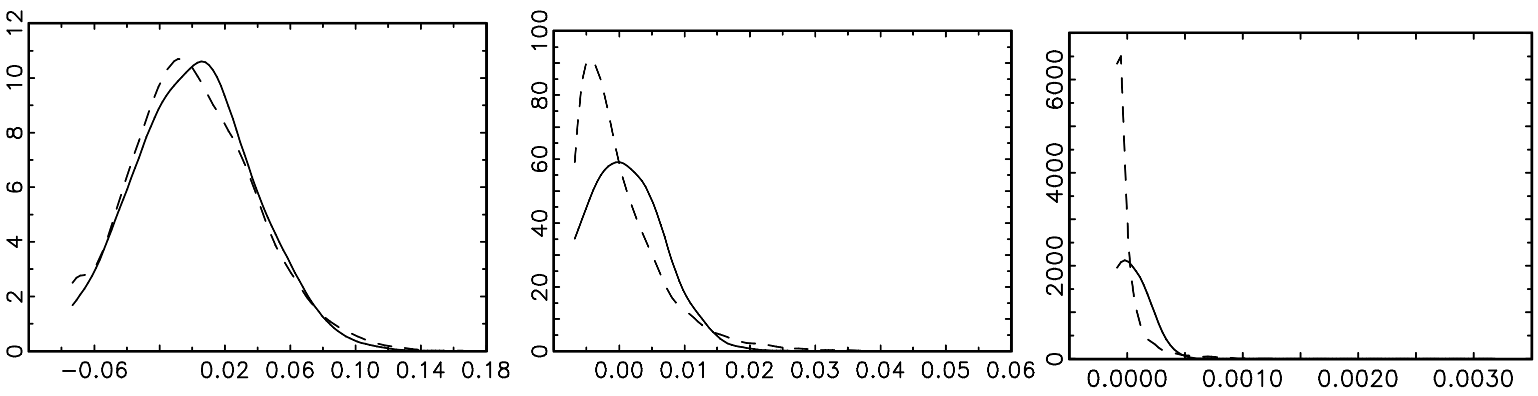
- Thus

$$
\begin{aligned}
|w^\top R_t|^\gamma &= C_\gamma \sigma_{t,w}^\gamma + D_\gamma \sigma_{t,w}^\gamma \zeta_{t,w} \\
\zeta_{t,w}|\mathcal{F}_{t-1} &\sim \left(|\xi|^\gamma - C_\gamma\right)/D_\gamma
\end{aligned}
$$

- Regression model

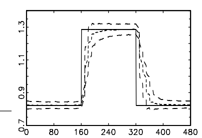
$$
\begin{aligned}
Y_{t,w} &= |w^\top R_t|^\gamma \\
&= \theta_{t,w} + s_\gamma \theta_{t,w} \zeta_{t,w}
\end{aligned}
$$

Choosing γ



$\gamma = 0.5$ $\gamma = 1$ $\gamma = 2$

Density estimates for transformed bivariate FX-returns $|w^\top R_t|^\gamma$ (dashed) and for random sample from a normal distribution (solid) (w: first eigenvector of unconditional covariance matrix).



Estimating θ (I known)

- OLS-estimator

$$
\begin{aligned}
\tilde{\theta}_{I,w} &= \frac{1}{|I|}\sum_{t\in I} Y_{t,w} \\
&= \frac{1}{|I|}\sum_{t\in I} \theta_{t,w} + \frac{s_\gamma}{|I|}\sum_{t\in I} \theta_{t,w}\zeta_{t,w}
\end{aligned}
$$

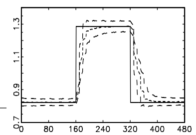
$s_\gamma = D_\gamma / C_\gamma$

- Under local homogeneity

$$
\theta_{t,w} \equiv \theta_{I,w} = C_\gamma \left(w^\top \Sigma_I w\right)^{\gamma/2},\ t \in I
$$

Hence,

$$
\begin{aligned}
\mathsf{E}\,\tilde{\theta}_{I,w} &= \theta_{I,w} \\
\mathsf{Var}\,\tilde{\theta}_{I,w} &= \frac{s_\gamma^2 \theta_{I,w}^2}{|I|} = v_{I,w}^2
\end{aligned}
$$



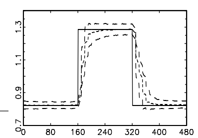
Accuracy of $\tilde{\theta}_{I,w}$ Variability of $\theta_{t,w}$ within $I = [\tau - m, \tau[$:

$$\Delta^2_{I,w} = |I|^{-1} \sum_{t \in I} (\theta_{t,w} - \theta_{\tau,w})^2$$

Then,

$$\mathsf{P}\left(|\tilde{\theta}_{I,w} - \theta_{\tau,w}| > \Delta_{I,w} + \lambda v_{I,w} \right) \leq 2 \exp\left(-\frac{\lambda^2}{2a_\gamma} \right)$$

for some fixed a_γ and $\lambda > 0$ governing the impact of estimation uncertainty. (The smaller λ the higher the probability of rejecting homogeneity.)



Estimating I

- **Basic idea** Suppose a family $\mathcal{I}$ of interval-candidates I,
  $I = [\tau - m, \tau[$, $m \in \mathbb{N}$, is given. Under homogeneity of I each subinterval $J \subset I$ is also homogeneous and $\theta_{I,w} = \theta_{J,w}$. Note, $J \in \mathcal{J}(I)$, $\mathcal{J}$ is a set of subintervals.

Estimating I (2)

- **Step1:** Select (initial interval estimate) I out of $\mathcal{I}$

  Step2: Test homogeneity within I

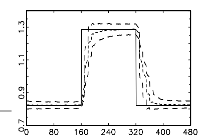
  If for given global parameters λ and μ there is $J \in \mathcal{J}(I)$ and $w \in \mathcal{W}$ such that

$$\left|\tilde{\theta}_{I,w} - \tilde{\theta}_{J,w}\right| > \lambda\, \tilde{v}_{J,w} + \mu\, \tilde{v}_{I,w}.$$

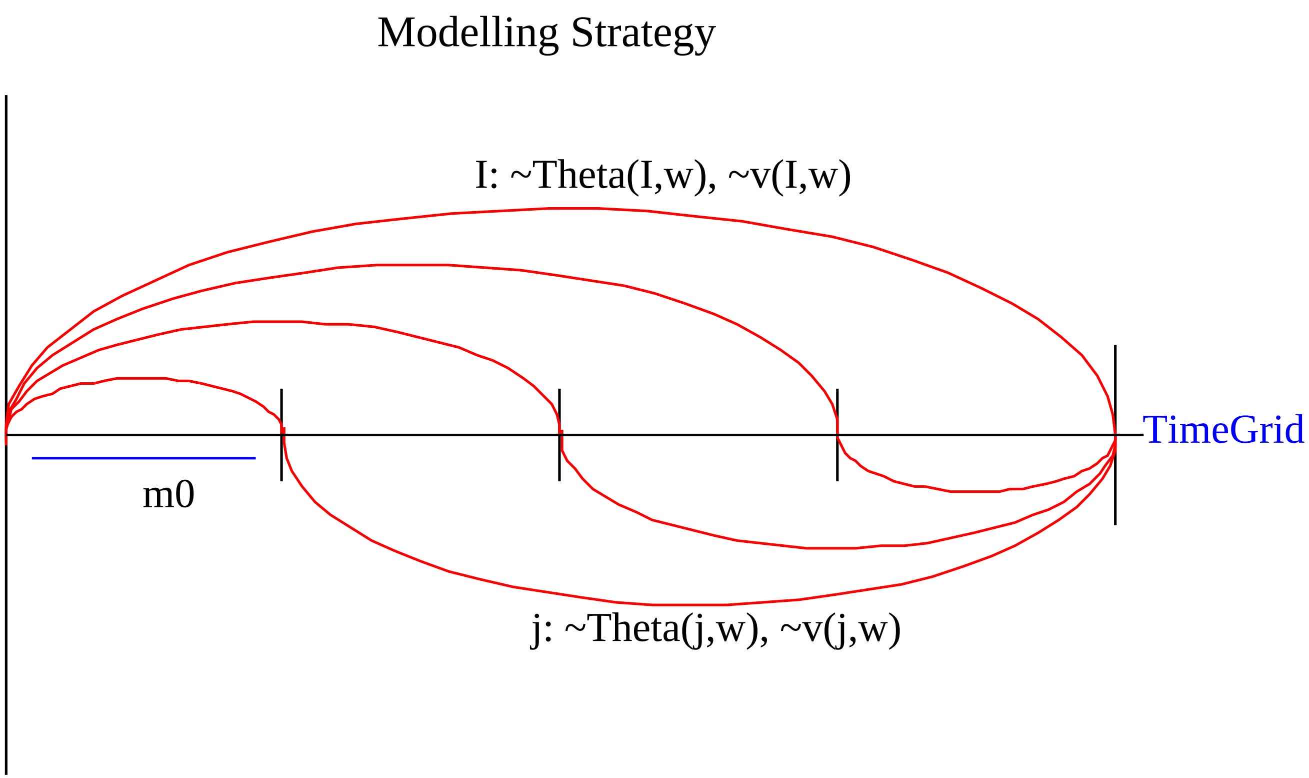
  reject homogeneity

  else replace I by large interval from I and interate until rejection of homogeneity

  Step3: $\hat{I} =$ "latest" non rejected interval I, XFGlochom



Estimating I (3)



Given $I \in \mathcal{I}$ and $m_0 \in \mathbb{N}$, $\tilde{\theta}_{I,w}$ and $\tilde{v}_{j,w}$ are estimated for each $j \in \mathcal{J}(I)$.

Estimating I (4)

Impact of sensitivity parameters λ and μ

- λ and μ large
  - conservative algorithm, slow reaction
  - larger bias of estimator
  - intervals of homogeneity possibly too large
- λ and μ small
  - frequent rejection of homogeneity
  - high variability of estimator
  - intervals of homogeneity possibly too small

Implementation

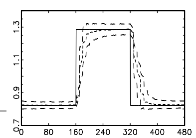
- Choice of the family $\mathcal{W}$ by Principal Components
- Choice of the sets $\mathcal{I}$ and $\mathcal{J}(I)$

$$
\begin{aligned}
\mathcal{I} &= \{I_k = [t_k, \tau[: t_k < \tau,\ t_k = m_0 k,\ k \in \mathbb{N}\} \\
\mathcal{J}(I_k) &= \{J = [t_{k'}, \tau[\text{ or } J = [t_k, t_{k'}[: k < k' < k^*\}.
\end{aligned}
$$

- Data-driven choice of λ and μ
  - Cross validation

$$
(\hat{\lambda}, \hat{\mu}) = \inf_{\lambda, \mu} \sum_{w \in \mathcal{W}} \sum_{t=t_0}^{T} \left(Y_{t,w} - \hat{\theta}_t^{(\lambda,\mu)} \right)^2
$$

  - Change point model



Change point model - Type I error

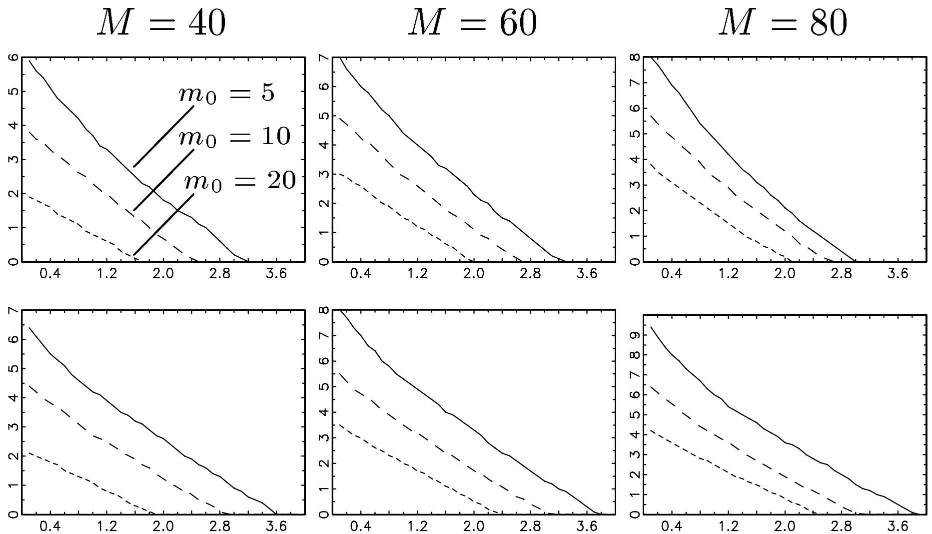
- Determine for a homogeneous interval I of length M λ and $\mu = \mu(\lambda)$ such that the probability of rejecting I is at most α.

$$\mathsf{P}\left(I \text{ is rejected}\right) \leq \sum_{J \in \mathcal{J}(I)} \sum_{w \in \mathcal{W}} 2 \exp\left(-\frac{\lambda^2}{2a_\gamma(1+\lambda s_\gamma |J|^{-1/2})^2}\right) \leq \alpha$$

  for λ not too small.

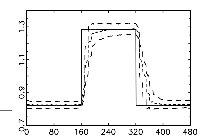
- Estimating $\mu = \mu(\lambda)$ by simulation.
  - $\alpha = 0.05$, $T = M + 1$
  - DGP is bivariate standard Gaussian
  - $\mathcal{W}_1$ ($\mathcal{W}_2$) contains first (both) eigenvector of $\tilde{\Sigma}$

Choosing λ and μ



$$\mu = \mu(\lambda)$$

Upper (lower) panels $\mathcal{W} = \mathcal{W}_1$ ($\mathcal{W} = \mathcal{W}_2$).



Change point model - Type II error

Let $\Sigma_t = \Sigma$ before change-point T_{cp} and $\Sigma_t = \Sigma'$ after it, and let $b_w = |\theta'_w - \theta_w|/\theta_w$. Thus, $|\mathbb{I}| = [T_{cp}, \tau[$ is an ideal choice $(m' = |\mathbb{I}|)$. $I = [\tau - m' - m, \tau[$. Then

$$\mathrm{P}\left(I \text{ is not rejected}\right) \leq 4 \exp\left(-\frac{\lambda^2}{2a_\gamma}\right)$$

if

$$\delta = \frac{\lambda s_\gamma}{\sqrt{\min\{m, m'\}}},\ 1 - \delta - \frac{\mu}{\lambda\sqrt{2}}\delta(1+\delta) > 0$$

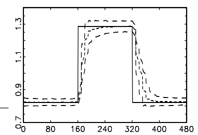
and

$$b_w \geq \frac{\delta + \delta(1+\delta) + \frac{\mu}{\lambda\sqrt{2}}\delta(1+\delta)}{1 - \delta - \frac{\mu}{\lambda\sqrt{2}}\delta(1+\delta)}.$$

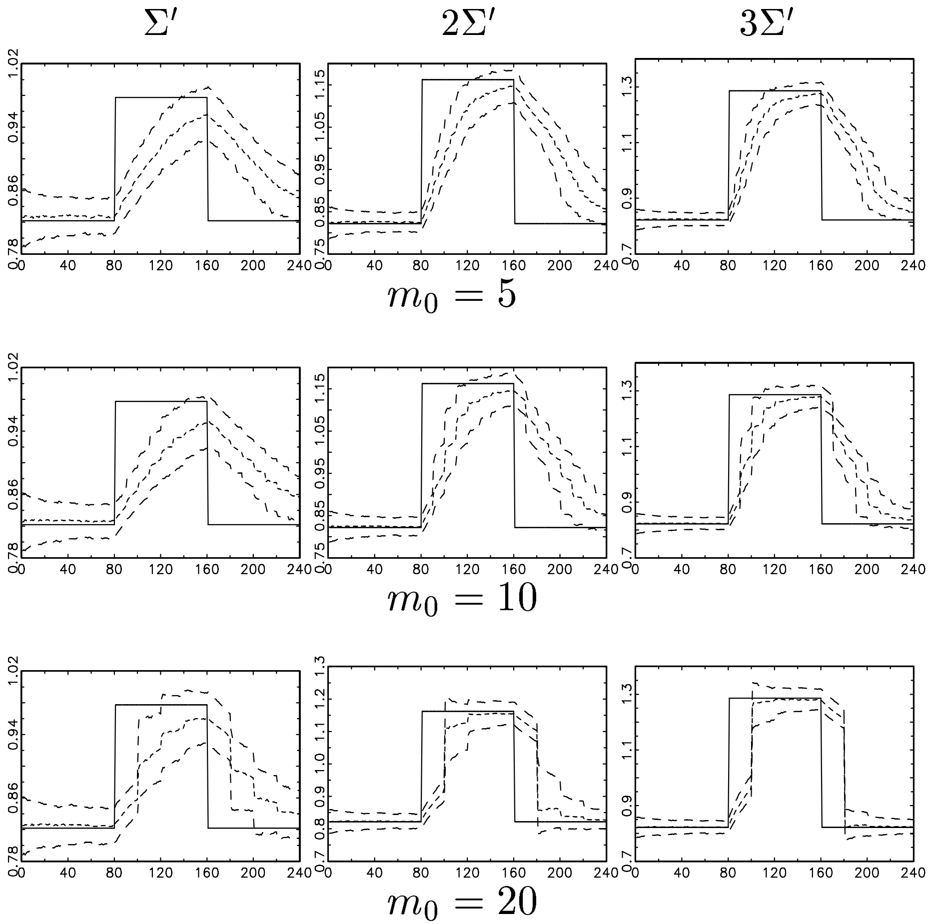
for some $w \in \mathcal{W}$

Monte Carlo Investigation

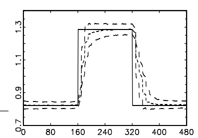
- Sample size $7M$, $M = 40, 60, 80$
- Estimation period: $t = M + 1, \ldots, 7M$
- Change point model with two changes ($3M$ an $5M$)
- DGP is bivariate standard Gaussian for $t \in [1, 3M[$ and $t \in [5M, 7M]$
- $\Sigma'_t = \Sigma'$, $2\Sigma'$ and $3\Sigma'$ for $t \in [3M, 5M[$, $\Sigma' = \begin{pmatrix} 2 & 0 \\ 0 & .5 \end{pmatrix}$.
- $m_0 = 5, 10, 20$, $\mathcal{W} = \mathcal{W}_2$
- 1000 replications



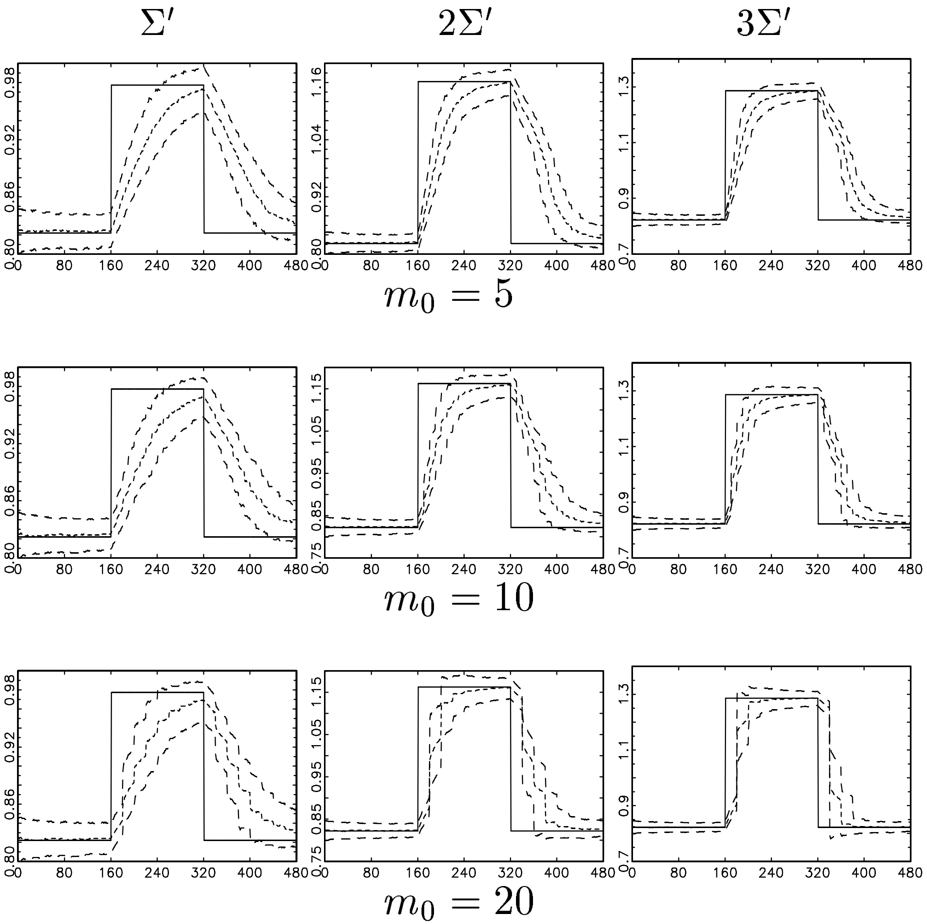
Monte Carlo results ($M = 40$)



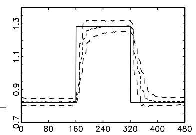
True quantities θ_{t,w_1} and median and interquartile range of $\hat{\theta}_{t,w_1}$.



Monte Carlo results ($M = 80$)



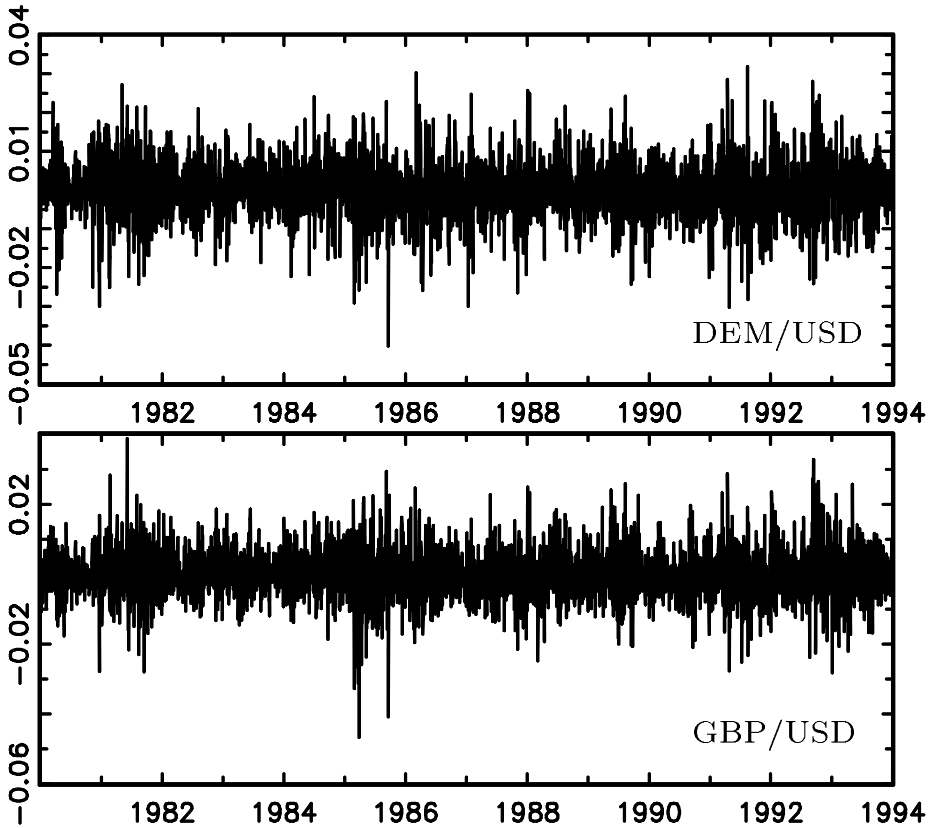
True quantities θ_{t,w_1} and median and interquartile range of $\hat{\theta}_{t,w_1}$.



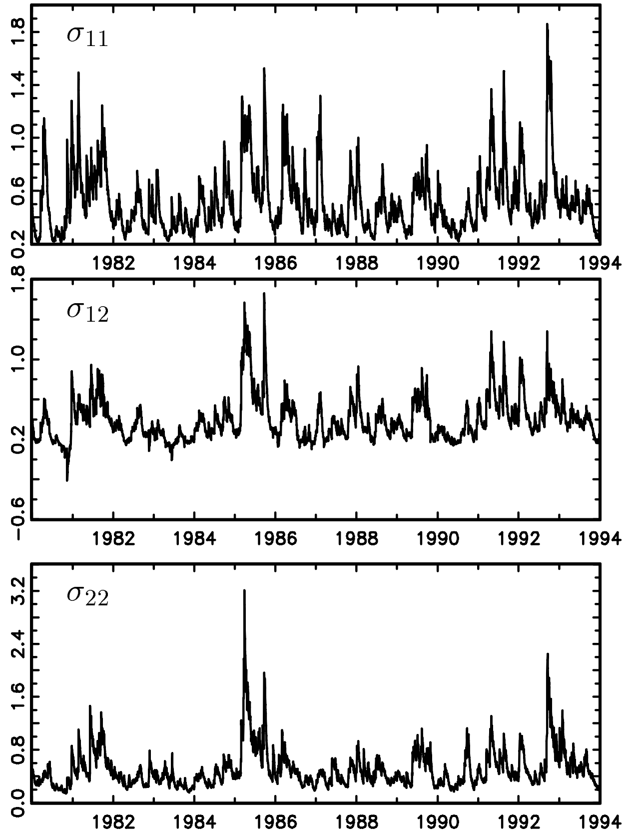
Application 1

- Bivariate FX-rate series DEM/USD and GBP/USD
- December, 31, 1979 to April, 1, 1994
- Daily quotes
- 3720 observations
- Parametric and adaptive approach
- Estimated implied cross moments

FX-rate returns XFGmvol01

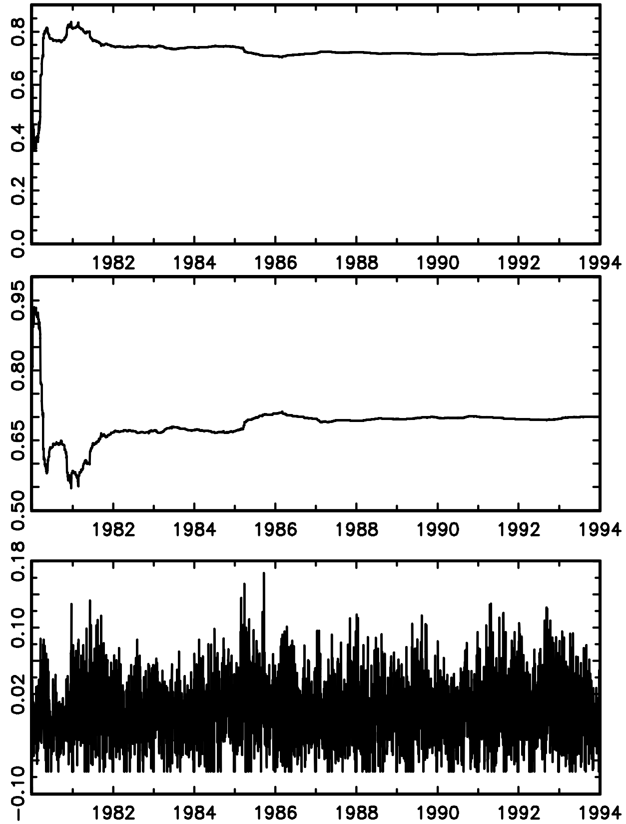


Parametric volatility estimates



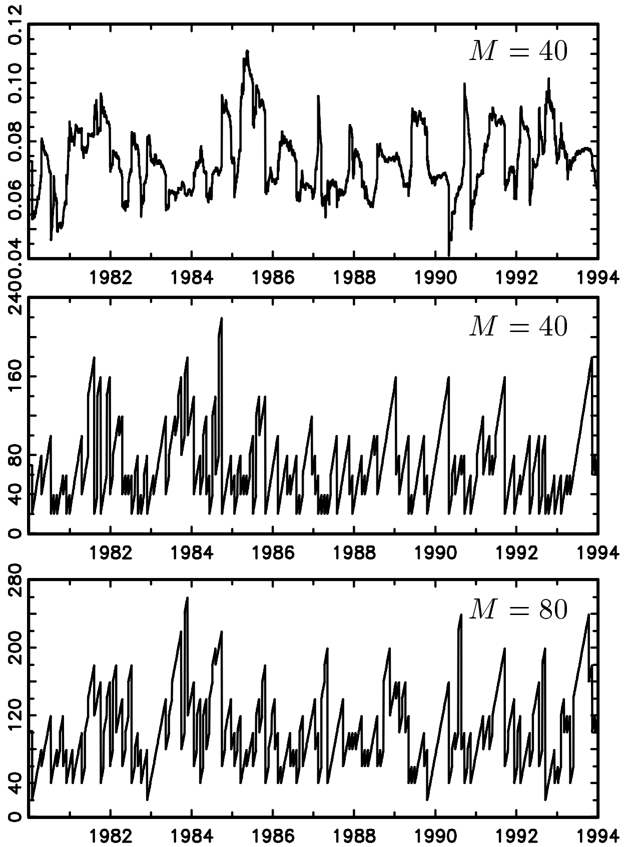
Elements of $10^4\hat{\Sigma}_t$ implied by BEKK estimate. XFGmvol02

Preliminary analysis of FX-returns



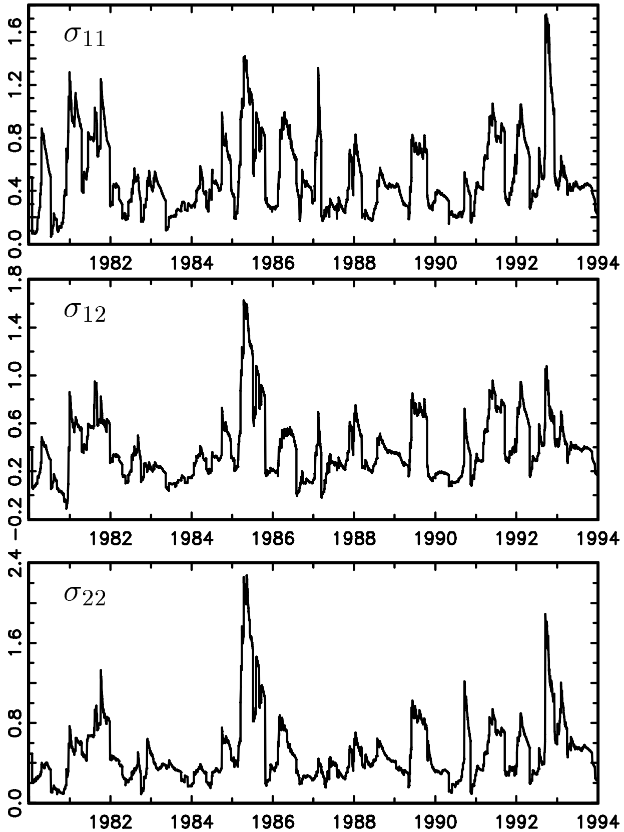
Elements of recursive eigenvectors (w_1) and centered process Y_{t,w_1}.

Adaptive estimates (FX-return modelling)



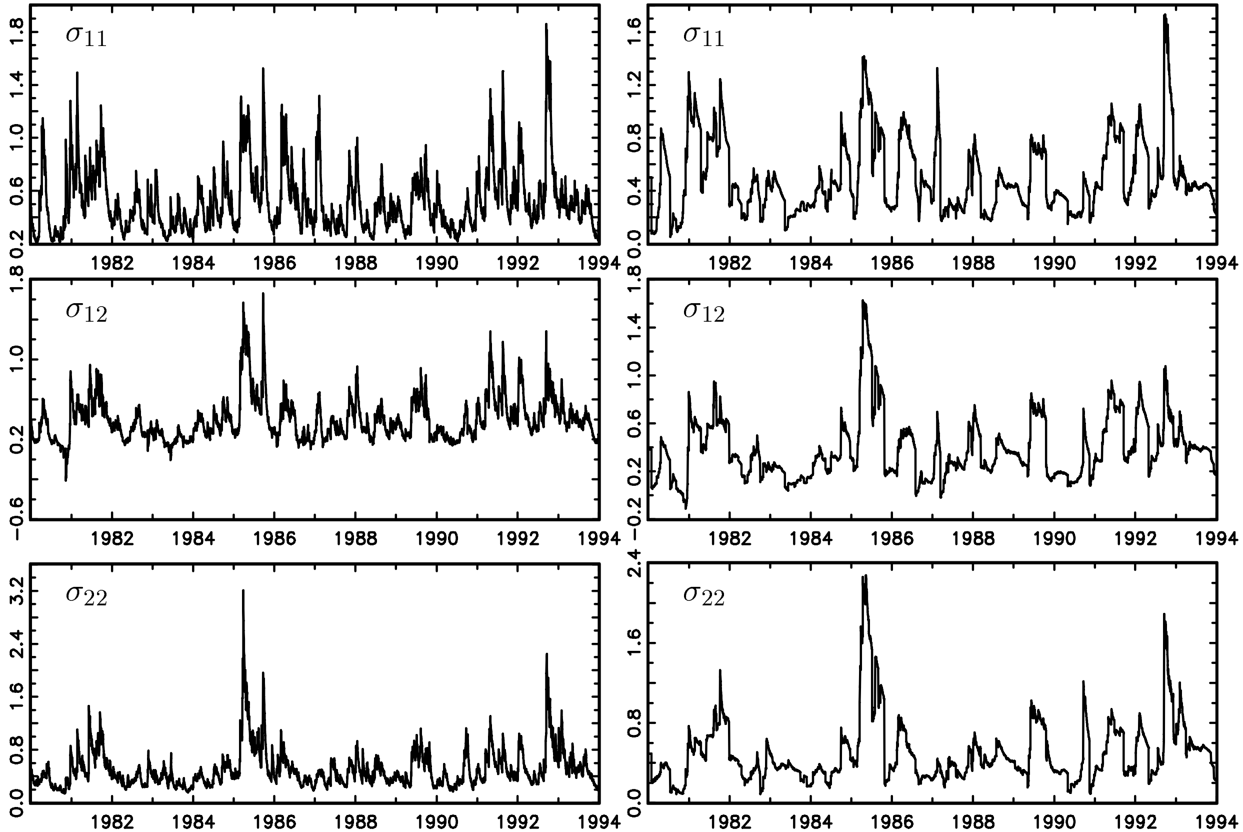
Smoothed process Y_{t,w_1} and estimated lengths of homogeneous intervals for $M = 40$ (medium) and $M = 80$ (lower panel), $m_0 = 20$.

Adaptive volatility estimates



$$M = 40,\ m_0 = 20,\ \mathcal{W} = \mathcal{W}_2,\ \lambda = 1.5,\ \mu = 0.4.$$

Volatility estimates



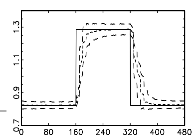
Parametric (BEKK) and adaptive estimates $10^4\hat{\Sigma}_t$

($M = 40$, $m_0 = 20$, $\mathcal{W} = \mathcal{W}_2$, $\lambda = 1.5$, $\mu = 0.4$).

Implied moments

	$M=40$		$M=80$		BEKK
	$m_0=5$	$m_0=20$	$m_0=5$	$m_0=20$	
λ	0.1	1.5	0.1	0.5	
$\mu(\lambda)$	6.4	0.4	9.4	3.4	
$\hat{\varepsilon}_{1t}\hat{\varepsilon}_{2t}$	0.02	-0.00	0.02	0.00	-0.01
	$(.021)$	$(.024)$	$(.023)$	$(.027)$	$(.026)$
$\hat{\varepsilon}^1_{1t}\hat{\varepsilon}^2_{2t}$	-0.08	-0.16	-0.11	-0.20	-0.21
	$(.054)$	$(.085)$	$(.078)$	$(.130)$	$(.107)$
$\hat{\varepsilon}^2_{1t}\hat{\varepsilon}^1_{2t}$	0.13	0.19	0.16	0.24	0.24
	$(.057)$	$(.087)$	$(.077)$	$(.130)$	$(.113)$
$\hat{\varepsilon}^2_{1t}\hat{\varepsilon}^2_{2t}$	1.63	2.17	1.99	2.73	2.56
	$(.180)$	$(.401)$	$(.329)$	$(.839)$	$(.600)$

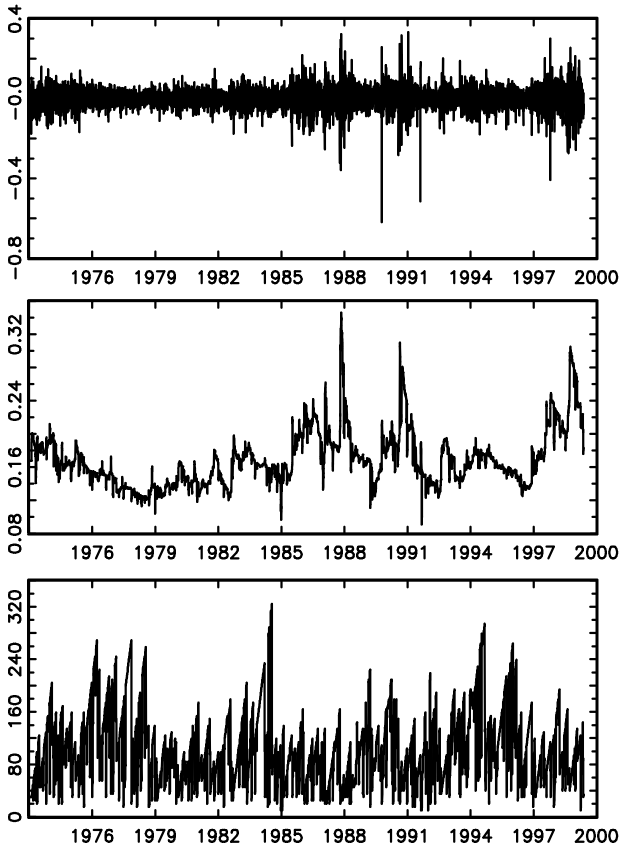
Standard errors in parentheses, $\mathcal{W}=\mathcal{W}_2$.



Application 2

- 23 German stock price series (DAX assets)
- January, 1, 1973 to May, 31, 1999
- Daily quotes
- 6891 observations

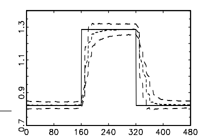
Stock return modelling



Centered process Y_{t,w_1} and adaptive estimates. Lower panel: Estimated lengths of homogeneous intervals ($M = 40$, $m_0 = 5$, $\mathcal{W} = \mathcal{W}_2$).

Rejections of implied moment conditions

$\varepsilon_i \varepsilon_j^2$		$\varepsilon_i^2 \varepsilon_j$		$\varepsilon_i^2 \varepsilon_j^2$	
i	j	i	j	i	j
BASF	BAYE	BAYE	HOEC	BASF	BAYE
BASF	HOEC	BAYE	VEBA	BASF	HOEC
BAYE	HOEC	COBA	PREU	BASF	SIEM
BMW	RWE	DEBK	HYPO	BAYE	HOEC
COBA	DRBK	SIEM	SCHE	BMW	RWE
DRBK	HYPO			COBA	DRBK
HYPO	DRBK			COBA	MAN
LUHA	MUER			COBA	SIEM
RWE	VEBA			DEBK	DRBK
SCHE	SIEM			HOEC	SCHE
				LUHA	RWE



Implied moments (I)

Stock	ε_i	ε_i^2	ε_i^3	ε_i^4
ALLI	-0.015	0.983	0.204	5.150
BASF	0.004	1.011	-0.821*	11.95
BAYE	0.009	1.000	-0.681	11.17
BMW	-0.005	1.003	0.296*	5.927
COBA	-0.004	0.963	-0.427	7.984
DEBK	-0.002	1.031	-0.241	8.666
DEGU	0.003	0.986	0.102	5.006
DRBK	-0.000	0.999	-0.275	8.073
HOEC	-0.012	1.001	-0.833*	11.19
HYPO	-0.009	1.006	-0.078	5.508
KARS	-0.009	0.989	-0.009	4.879
LUHA	0.004	1.012	0.178	5.067

Implied moments (II)

Stock	ε_i	ε_i^2	ε_i^3	ε_i^4
LIND	0.005	0.982	0.142	4.542
MAN	-0.002	0.999	0.101	4.465
MANN	-0.008	0.991	-0.151	5.560
MUER	-0.005	1.001	0.090	6.101
PREU	-0.003	1.005	0.286*	5.619
RWE	-0.000	0.985	0.043	6.842
SCHE	-0.009	1.001	0.079	5.790
SIEM	0.004	0.994	-0.172	6.766
THYS	0.003	0.995	-0.144	6.358
VEBA	0.006	1.011	-0.354	8.836
VW	0.004	0.984	-0.039	5.802

Conclusions

- Adaptive modelling is feasible even for high dimensional heteroskedastic processes (dimension reduction, power transformation)
- Selection of global parameters by means of the error probability of falsely rejecting a homogeneous interval
- Reasonable sensitivity in presence of volatility shifts

Conclusions (2)

- Moments of implied innovations are *close* to theoretical counterparts
- Good empirical performance in high dimensional system of German stock returns
- Some remaining issues: Determination of w, Application of volatility estimates

References

Bollerslev, T., Engle, R.F. & Woolridge, J.M. (1988). A Capital Asset Pricing Model with Time Varying Covariances, *Journal of Political Economy*, **96**: 116–131.

Braun, P.A, Nelson, D.B. & Sunier, A.M. (1995). Good News, Bad News, and Betas, *Journal of Finance* **50**: 1575–1603.

Engle, R.F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation, *Econometrica*, **50**: 987–1008.

Engle, R.F. , Ito, T. & Lin, W.L. (1990). Meteor Showers or Heat Waves? Heteroscedastic Intra–Daily Volatility in the Foreign Exchange Market, *Econometrica*, **58**: 525–542.

Lepski, O.V. (1990). One Problem of Adaptive Estimation in Gaussian

White Noise, *Theory Probab. Appl.*, **35**: 459–470.

Lepski, O.V. & Spokoiny, V. (1997). Optimal Pointwise Adaptive Methods in Nonparametric Estimation, *Annals of Statistics*, **25**: 2512–2546.

Mercurio, D. (2002). Locally Time Homogeneous Time Series Modeling, *Applied Finance Guide*, Springer, Berlin, Chapter 15.

Spokoiny, V. (1998). Estimation of a Function with Discontinuities via Local Polynomial Fit with an Adaptive Window Choice, *Annals of Statistics*, **26**: 1356–1378.