Flexible Stochastic Volatility Structures for High Frequency Financial Data

Dorit FELDMANN

Wolfgang HÄRDLE

Christian HAFNER

Marc HOFFMANN

Oleg LEPSKI

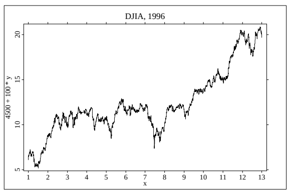
Alexander TSYBAKOV

Sonderforschungsbereich 373

Humboldt–Universität zu Berlin, Germany

and

University of Paris, France



SFB

Why Stochastic Volatility?

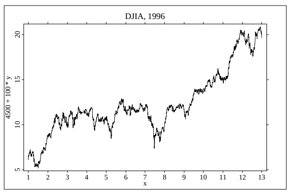
Volatility is an important financial decision variable

It can be used for

1. Prediction intervals for prices or indices
2. Risk Evaluation, Value at Risk (VaR)
3. Portfolio Selection
4. Option Pricing

Econometric Specification of volatility

1. constant (e.g. Black and Scholes, 1973)
2. time–varying
   (a) ARCH (Engle, 1982)
   (b) Stochastic Volatility (Taylor, 1986)



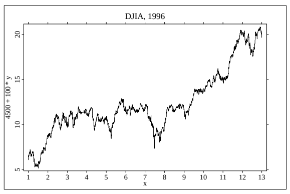
Test of the Standard Parametric SV Model

S_t: underlying asset price at time t, $t = 1, \ldots, n$

Returns $y_t = \log S_t / S_{t-1}$

standard Stochastic Volatility model, Taylor (1986)

$$
\begin{aligned}
y_t &= \exp(h_t/2)\xi_t^* \qquad (1) \\
h_t &= \vartheta h_{t-1} + \varepsilon_t \qquad (2)
\end{aligned}
$$



$$\xi_t = \log(\xi_t^*)^2 - \omega \text{ with } \omega = \mathsf{E}[\log(\xi_t^*)^2]$$

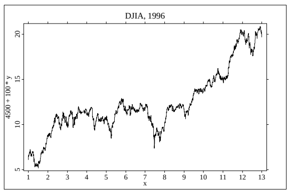
$$z_t = \log y_t^2$$

linear state space form:

$$\begin{aligned} z_t &= \omega + h_t + \xi_t \qquad (3) \\ h_t &= \vartheta h_{t-1} + \varepsilon_t \qquad (4) \end{aligned}$$

$$h_t = m\,(h_{t-1}) + \varepsilon_t$$

$$\mathbf{H}_0: \quad m(x) = \vartheta x, \vartheta \in [a, b] \subset (-1, 1) \qquad (5)$$

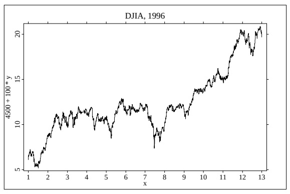


Assumptions:

1. $(\xi_t, t = 1, \ldots, n)$ and $(\varepsilon_t, t = 1, \ldots, n)$ are stochastically independent
2. Up to fourth moments are finite and constant

Define

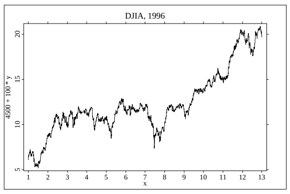
$$
\begin{aligned}
M(\vartheta) &= \mathsf{E}[\varepsilon_1 + \xi_2 - \vartheta \xi_1]^2 \\
B(\vartheta) &= \mathsf{E}[\varepsilon_1 + \xi_2 - \vartheta \xi_1]^4 \\
&+ \mathsf{E}[(\varepsilon_1 + \xi_2 - \vartheta \xi_1)^2 (\varepsilon_2 + \xi_3 - \vartheta \xi_2)^2]^2 \\
&- 3M(\vartheta)^2
\end{aligned}
$$



$$\bar{\vartheta}_n = \frac{\sum_{t=3}^{n} z_t z_{t-2}}{\sum_{t=3}^{n} z_{t-1} z_{t-2}} \tag{6}$$

Under hypothesis $\mathbf{H}_0$, $\bar{\vartheta}_n$ is $\sqrt{n}$-consistent for ϑ

$$\hat{\vartheta}_n = \begin{cases} \bar{\vartheta}_n, & \text{if} \quad a \le \bar{\vartheta}_n \le b; \\ a \ , & \text{if} \quad \bar{\vartheta}_n < a; \\ b \ , & \text{if} \quad \bar{\vartheta}_n > b; \end{cases}$$



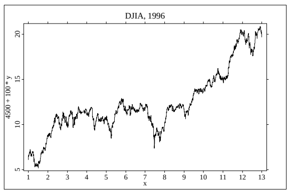
Test statistic

$$T_n = \frac{1}{\sqrt{nB_n}} \sum_{t=1}^{n} \left\{ \left(z_t - \hat{\vartheta}_n z_{t-1} \right)^2 - M_n \right\}. \qquad (7)$$

Fix some $0 < \alpha < 1$ and set

$$\Delta_n = \begin{cases} 0, & \text{if } \; T_n \leq \sqrt{2 \ln \frac{1}{\alpha}} \\ 1, & \text{if } \; T_n > \sqrt{2 \ln \frac{1}{\alpha}} \end{cases}$$

The hypothesis is accepted if $\Delta_n = 0$ and it is rejected if $\Delta_n = 1$.



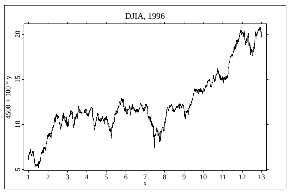
The Test on a Workbench

1000 replications of the series

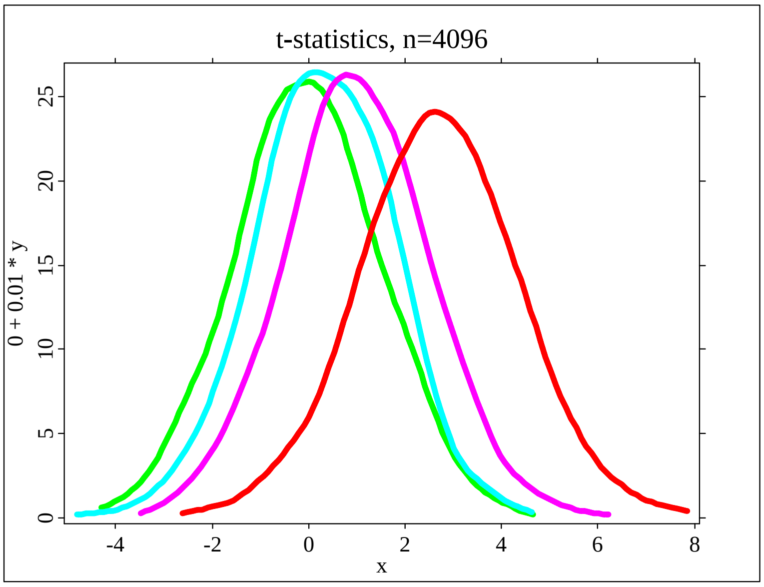
$$z_t = h_t + \xi_t, \tag{8}$$

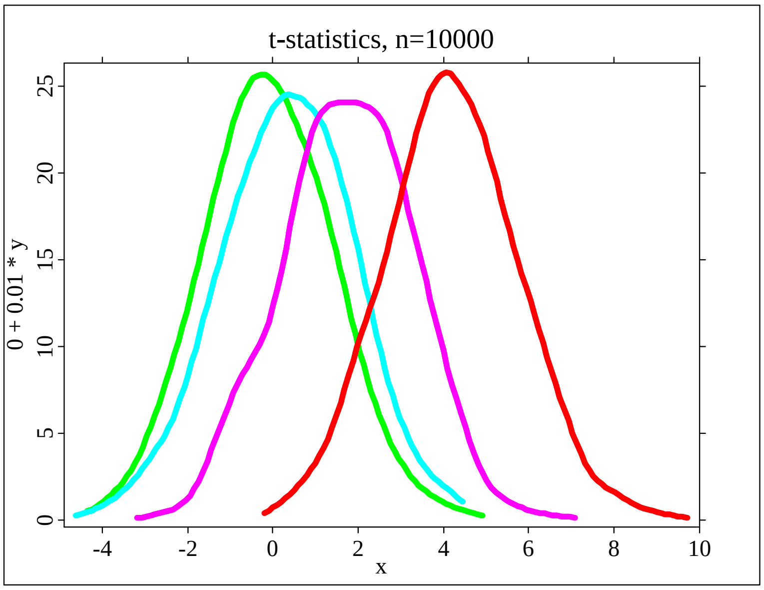
$$h_t = \vartheta h_{t-1} + \frac{\lambda}{n^{1/4}} \sin(2\pi h_{t-1}) + \varepsilon_t \tag{9}$$

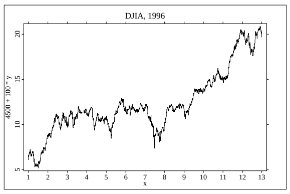
$$n = 10000, 6561, 4096$$
$$\vartheta = 0.95$$



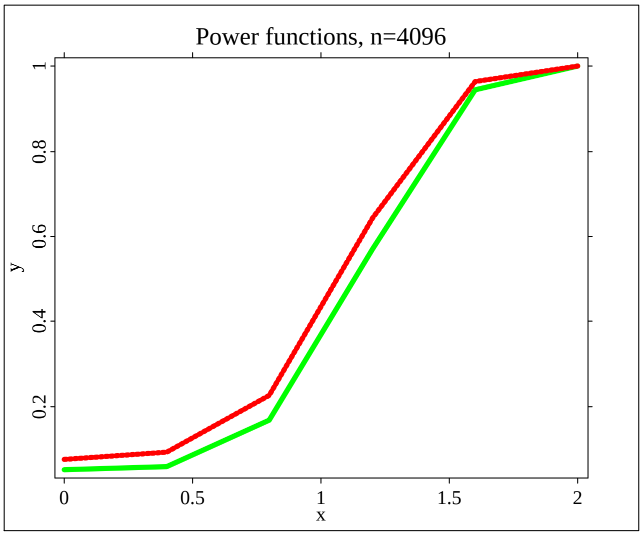
n	10000	6561	4096
λ	0.000	0.000	0.000
mean	-0.048	0.019	0.014
std.dev.	1.502	1.551	1.490
λ	0.500	0.450	0.400
mean	0.383	0.365	0.357
std.dev.	1.553	1.569	1.418
λ	1.000	0.900	0.800
mean	1.801	1.418	1.052
std.dev.	1.521	1.520	1.492
λ	1.500	1.350	1.200
mean	4.251	3.445	2.747
std.dev.	1.545	1.664	1.643
λ	2.000	1.800	1.600
mean	8.486	6.868	5.377
std.dev.	1.903	1.864	1.823
λ	2.500	2.250	2.000
mean	15.406	12.445	9.881
std.dev.	3.306	3.974	3.332

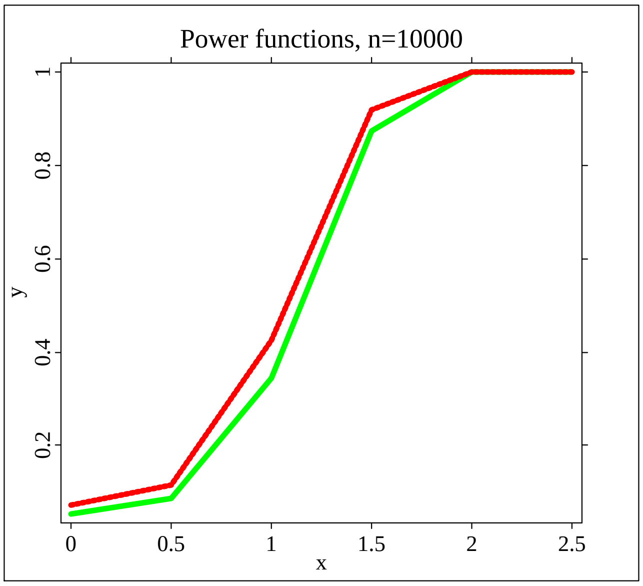

t-statistics, n=4096
0 + 0.01 * y
0
5
10
15
20
25
-4
-2
0
2
4
6
8
x

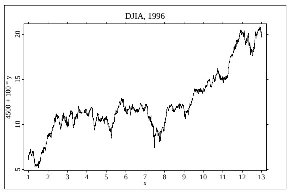

t-statistics, n=10000
0 + 0.01 * y
0
5
10
15
20
25
-4
-2
0
2
4
6
8
10
x


DJIA, 1996

SFB


Power functions, n=4096
y
0.2
0.4
0.6
0.8
1
x
0
0.5
1
1.5
2


Power functions, n=10000
y
0.2
0.4
0.6
0.8
1
x
0
0.5
1
1.5
2
2.5


DJIA, 1996

SFB

The HFDF96 data set

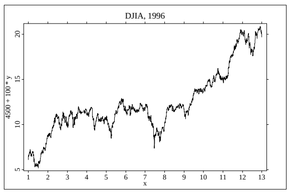
half hourly sampled values of 1996

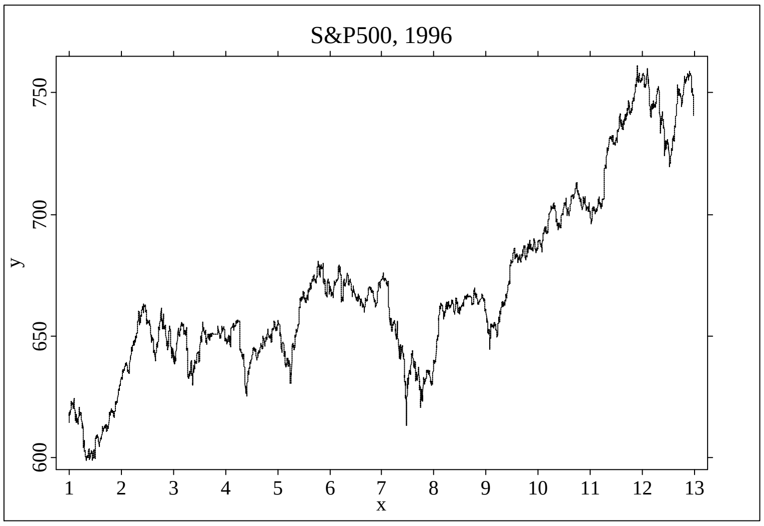
3 data sets:

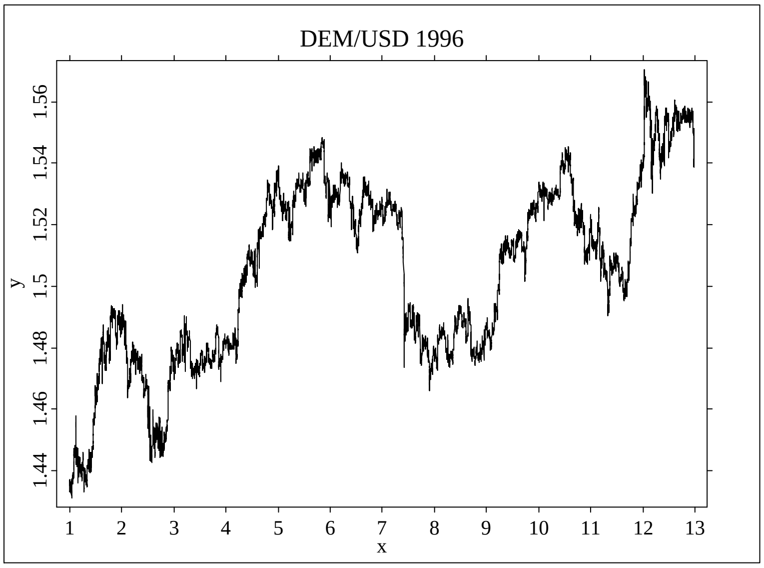
1. Dow Jones Industrial Average, $n = 3680$
2. Standard & Poors 500, $n = 3680$
3. DEM/USD exchange rate, $n = 14234$

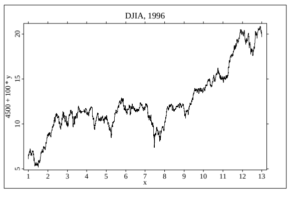
For the stock indices: intervals corresponding to non–trading hours at the NYSE are skipped

For DEM/USD: intervals for which the time of the last quote was more than half an hour behind are skipped




S&P500, 1996
y
x
600
650
700
750
1
2
3
4
5
6
7
8
9
10
11
12
13


DEM/USD 1996
y
x
1.44
1.46
1.48
1.5
1.52
1.54
1.56
1
2
3
4
5
6
7
8
9
10
11
12
13

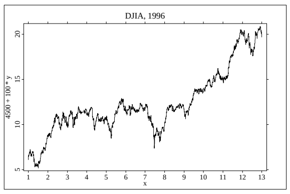

DJIA, 1996

SFB

$$\hat{\vartheta} = 0.9004 \text{ for DEM/USD}$$
$$\hat{\vartheta} = 1.0153 \text{ for DJIA}$$
$$\hat{\vartheta} = 0.9241 \text{ for S\&P500}$$

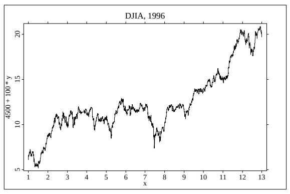
Estimate the unknown ω, B_n and M_n by taking means of the corresponding moments over the entire sample

For T_n, consider k subsamples of size n/k



Entries: Number of rejections for k subsamples

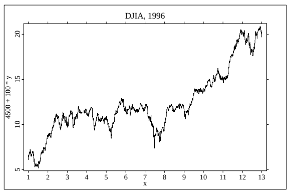
	DEM/USD		DJIA		S&P500	
k	α=0.05	α=0.1	0.05	0.1	0.05	0.1
2	0	0	0	0	0	0
3	0	0	0	0	1	1
4	0	0	1	1	0	0
5	1	1	1	1	0	0
6	0	0	1	1	2	2
10	1	1	1	1	0	2
12	1	1	1	1	1	2
20	2	3	1	1	1	3
50	5	7	2	3	4	6
100	12	15	10	13	9	9



DAX Data

Source: TP C1; daily stock prices, 1974-1996

Stock	$\hat{\vartheta}$	$\alpha = 0.05$	$\alpha = 0.1$
Allianz	0.9680	2	3
BASF	0.9423	6	6
Bayer	0.9890	3	5
BMW	0.9723	2	3
Commerzbank	0.9713	5	5
Daimler	0.9691	2	4
Deutsche Bank	0.9803	2	3
Dresdner Bank	0.9483	6	8
Hoechst	0.9455	2	4
Karstadt	0.9943	2	7
MAN	0.9402	2	5
Mannesmann	0.9683	5	6
Preussag	0.9967	2	3
RWE	0.9811	3	3
Siemens	1.0019	1	3
Thyssen	0.9741	4	6
Volkswagen	0.9992	4	5



Conclusions and Outlook

- Hypothesis could not be rejected for any data set
- No saturation or nonlinearity effect
- Extended SV models to cope with leverage effect