Stochastic volatility: Modelling the latent process empirically.

Erik Bølviken
Institute of Mathematics, University of Oslo &
The Norwegian Computing Centre

Financial variable examined:

- **Log-return** of stock indexes
- **Definition:**

  The logarithm of **relative** price change

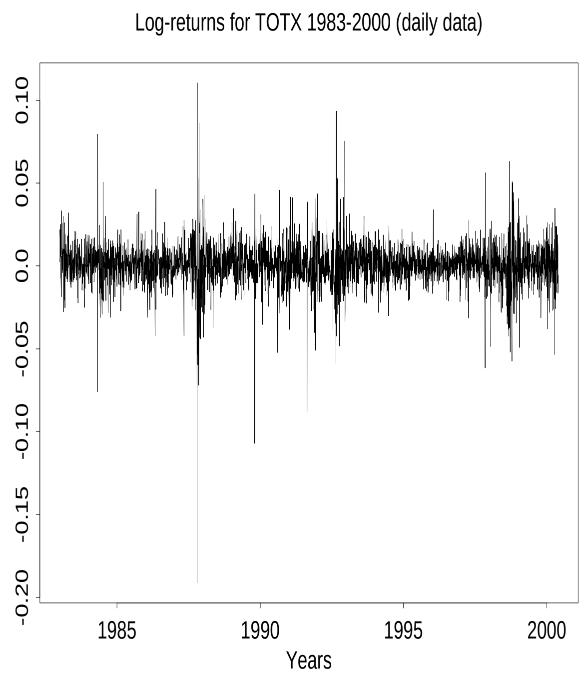
Volatility:

- The same as standard deviation.

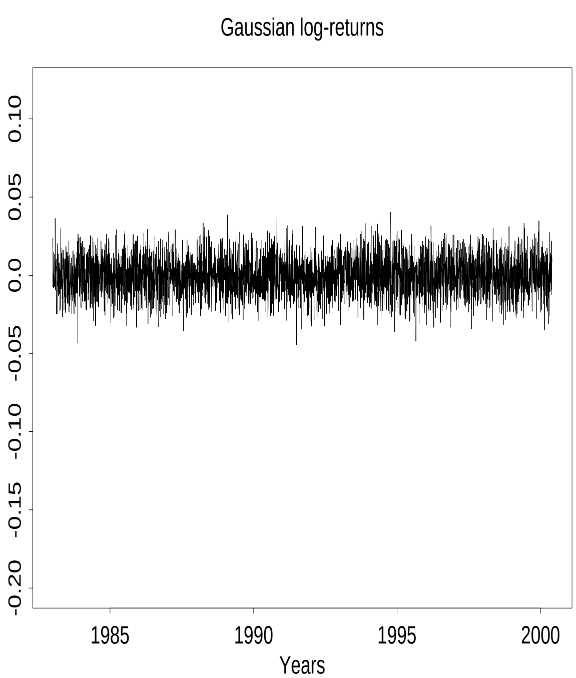
What is stochastic volatility?

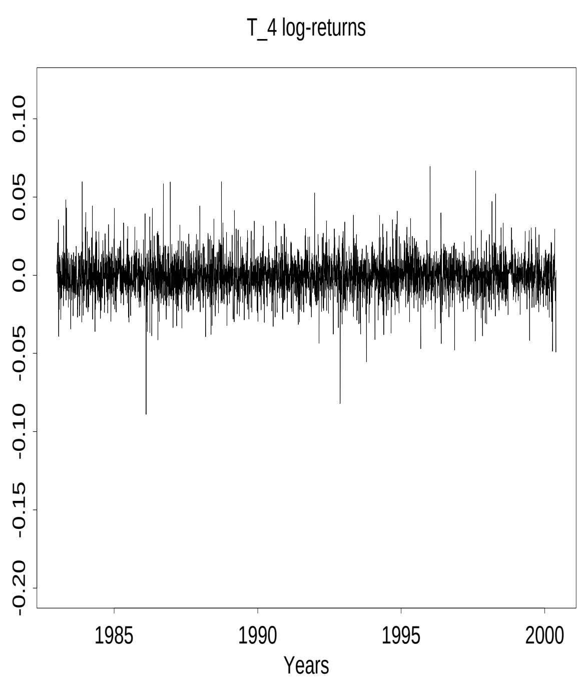
Phenomena such as this:

The **real** stock index of Oslo (Norway):



Simulations from fitted models **ignoring** stochastic volatility:





Why is stochastic volatility of importance?

For two reasons:

- Of interest **in itself**
- Influences evaluations of risk:
  - Tail measures like **VaR** especially sensitive

Purpose of talk:

- Describe the volatility process using

  a **weak** mathematical model

  and **plenty** of historical data

Outline of talk

Main themes:

- Introduction (completed)

- Technical material

  Mathematical model (not **parametric** like GARCH)

  Estimation: Through **pseudo-likelihood**

  Can it be done? Testing on **simulated** data

- Examination of index series

- Concluding remarks

Mathematical model

Notation:

- Period: k, time resolution: **Day, week, month**
- Log-return: y_k
- **State** of the market: s_k (**unobserved**)

Model:

- $$y_k = \overbrace{\underset{\substack{\uparrow \\ \text{parameter}}}{\sigma} \exp(\underset{\substack{\uparrow \\ \text{parameter}}}{\alpha} s_k)}^{\text{volatility}} \varepsilon_k$$

- s_k **stationary** process, called **latent** or **regime**

  responsible for volatility **fluctuations**

  - assumed **gaussian**

- ε_k independent random terms

  with no relation to s_k

Problem raised:

- Underlying model for s_k?

Statistical estimation: Method

Target:

- Autocorrelation function of s_k,

  defined as $\text{cor}(s_k, s_{k+l})$, $\quad l = \text{lag}$

  $\uparrow$

  correlation same for all k

Estimation:

- Trough a **pseudo**-likelihood criterion

  as explained in the paper

- Conditions too **weak** for ordinary likelihood

- Technicalities:

  A lot of numerical integration

  Numerical optimization

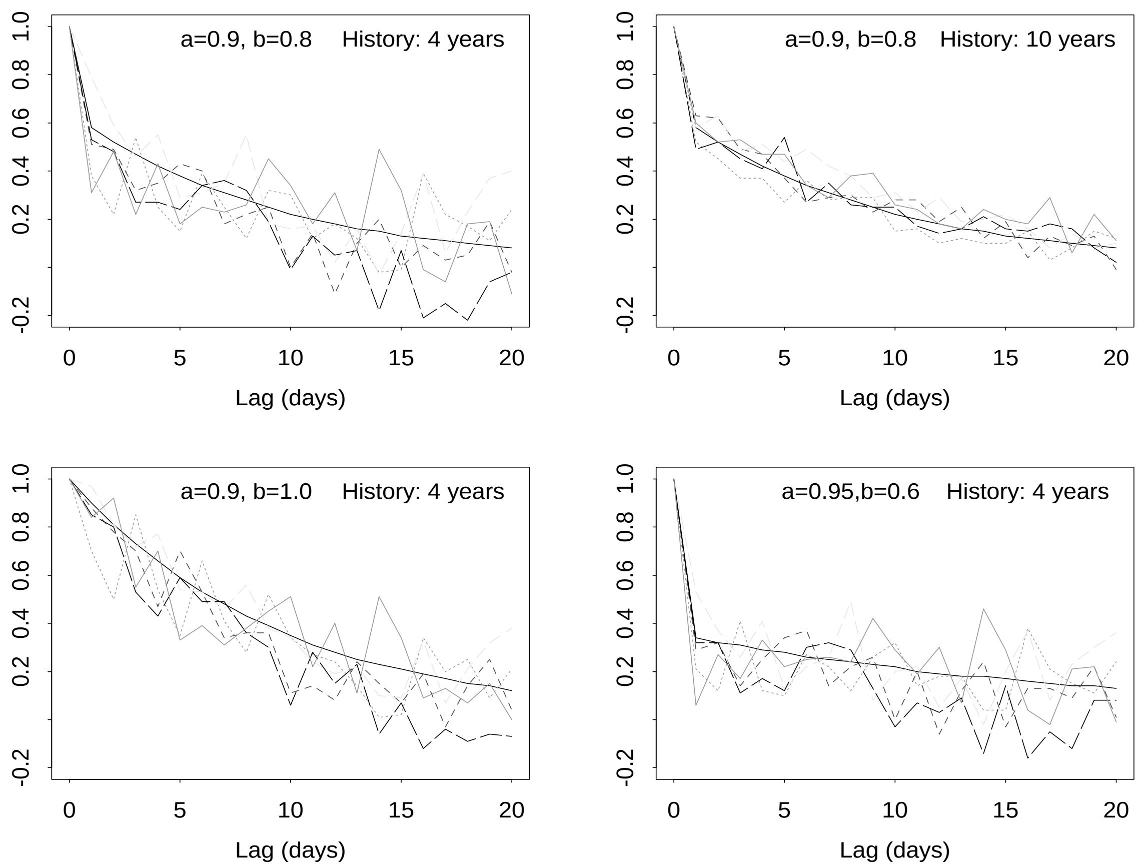
Simulations: How long must the series be?

Experimental conditions:

- **Four** and **ten** years of **daily** data
- Realistic parameters

Autocorrelation functions reconstructed:

- **Solid** lines: The truth*
- **Dashed/dotted** lines: Attempted recontructions



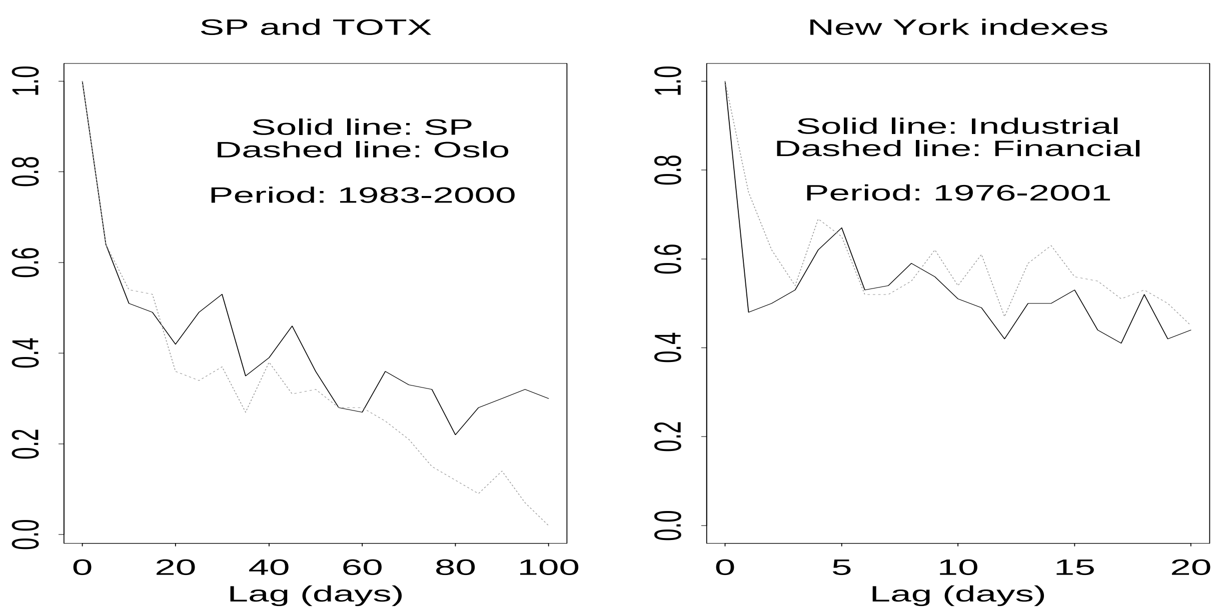
*The parameters a and b are explained in the paper

Example 1: Financial communities of different size

Notation and facts:

- **SP**: Standard & Poor 500 index
- **TOTX**: Index of the stock exchange of Oslo (Norway)
- **Daily** data 1983-2000

Estimated autocorrelation functions



Remarks:

- **Slow** decay with the time lag
- **Faster** for the small unit (Oslo)
- Interpretation as model: Later

Example 2: New York indexes

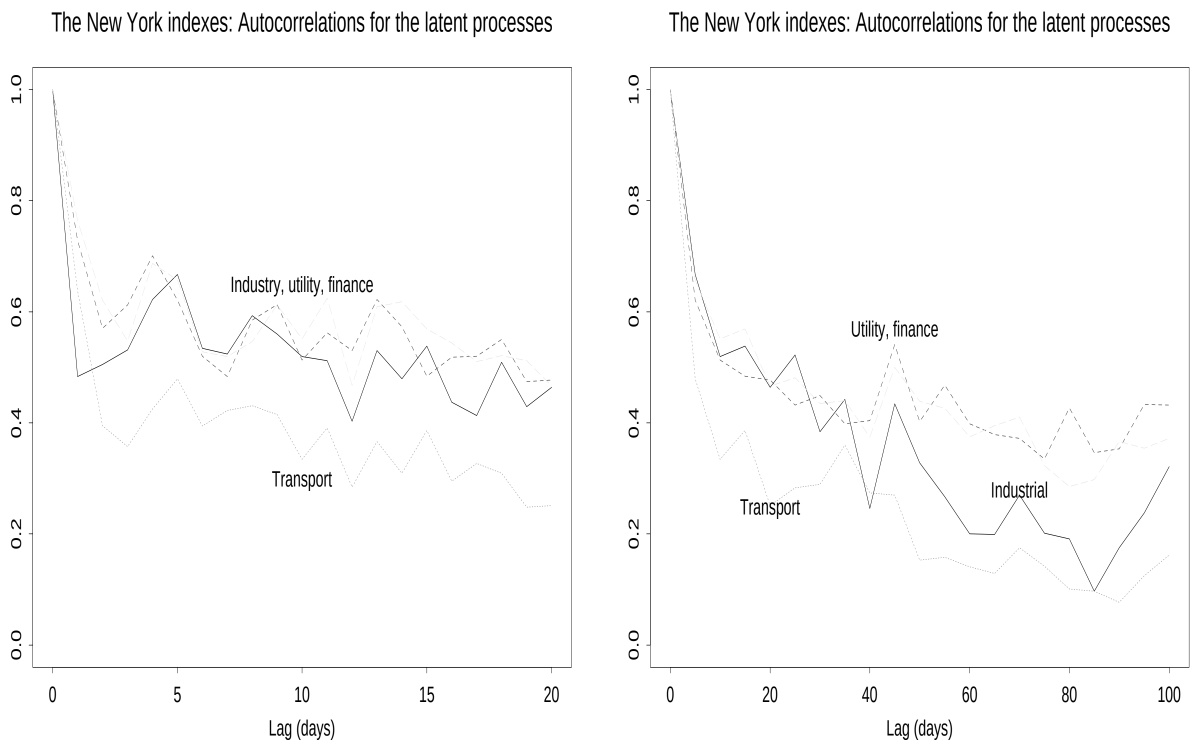
Some facts:

- The indexes examined:

  Industri, transport, uility, financial

- **Daily** data 1976-2001

Estimated autocorrelation functions



Comment:

- Left : Time lag up to **20** days
- Right : Time lag up to **100** days

Mathematical model identified

Conclusion:

- All estimated autocorrelation functions consistent with

  autoregressive, moving average (ARMA) processes of order $(1,1)$

  In mathematical form:

$$s_k = z_k + \omega_k, \qquad z_k = a z_{k-1} + \eta_k$$

ω_k: random process, independent, zero mean

a: parameter, defines **decay**

η_k: random process independent, zero mean

- Mathematical model ambiguous:

  – First form

$$y_k = \sigma \exp(\alpha s_k) \varepsilon_k$$

s_k: **non-markov** process

ε_k: **gaussian** process

  – Second form

$$y_k = \sigma \exp(\alpha z_k) \varepsilon'_k, \qquad \varepsilon'_k = \varepsilon_k \exp(\omega_k)$$

z_k: **markov** process

ε'_k: heavy-tailed, **non-gaussian**

Different latent processes

Additional problem:

- Relationship between latent processes for **different** financial variables?

Quantity sought:

- The **cross**correlation function

$$\text{cor}(s_{1k}, s_{2k+l}), \qquad l = \text{lag}$$

$\uparrow$

correlation same for all k

- for latent processes s_{1k} and s_{2k} corresponding to **different** log-returns

Method:

- Essentially as described earlier

New York indexes: *Cross*correlations latent processes

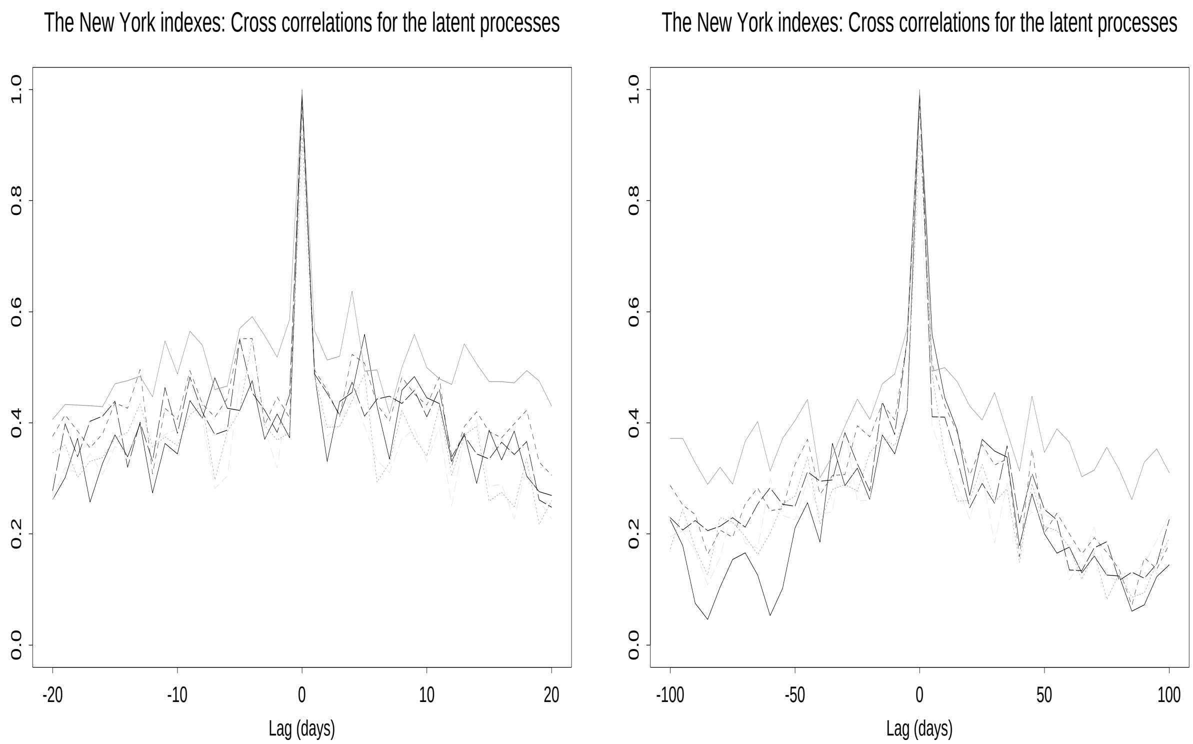
Some facts:

- The indexes examined:

  Industri, transport, uility, financial

- **Daily** data 1976-2001

Estimated *cross* correlation



Comment:

- Left : Time lag up to **20** days
- Right : Time lag up to **100** days

Latent processes for SP and Oslo indexes

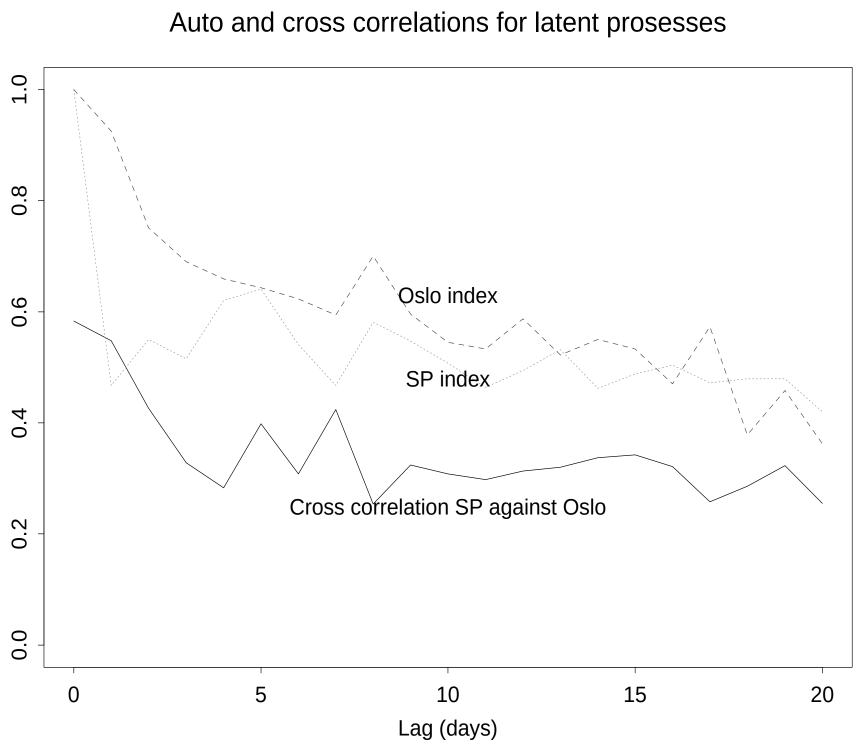
Some facts:

- The indexes examined:

  Standard&Poor 500 and Oslo stock exchange (TOTX)

- **Daily** data 1983-2000

Estimated *cross* and auto correlation



Comment:

- Reasonably parallel curves (?)
- Error (**at lag one**) for Oslo index

Suggested mathematical model

Remark:

- **Cross**correlations with losely same decay as **auto**correlations (?)
- If so, consistent with **one** latent process underlying **all**

Model in summary:

- One **single** latent process, of **Markov** type
- Non-gaussian noise

Concluding remarks

- Purpose of method presented:
  To identify model for regime (latent) process
  without parametric assumptions

- Worked well for **daily** data;
  parsimonious model for multiple series suggested

- For **Monthly** data:

  Series too short;
  Estimates too unstable

  Possible approach:
  Upscale the daily model?