FENS
Warszawa 25-27.XI.2010
5 Ogólnopolskie Sympozjum
Fizyka w Ekonomii i Naukach Społecznych

# Prices and Volumes on the Stock Market

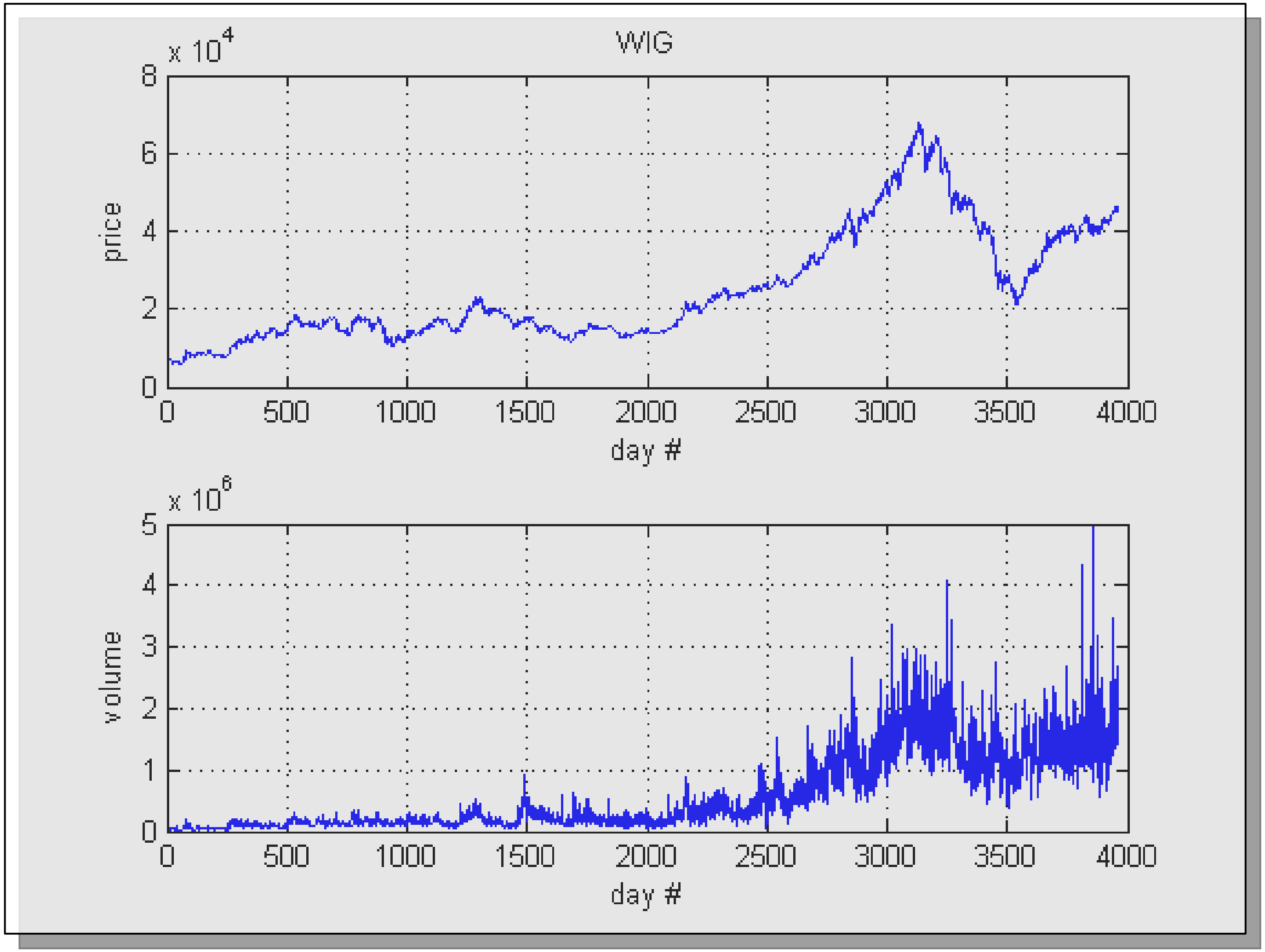
Krzysztof Karpio
Piotr Łukasiewicz
Arkadiusz Orłowski

Warszawa, 25-27 listopada 2010

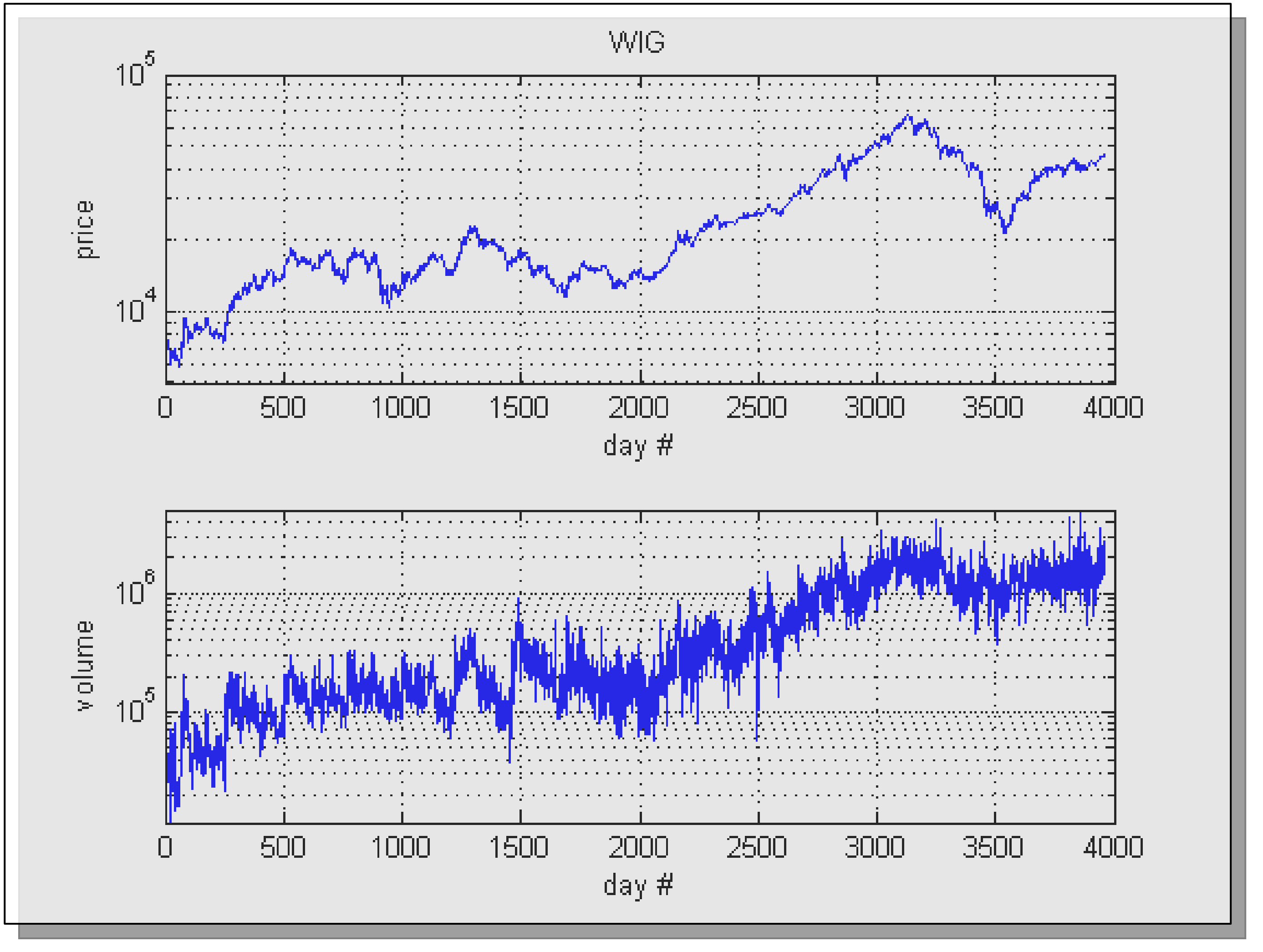
# Data selection

- Warsaw Stock Exchange
- WIG index
  - assets most important for investors
  - different sectors
  - content: every 3 months (volume, capitalization)
- 1995-01-02 to 2010-10-21
  - >= 27 assets
  - 3961 trading days
  - ~ 15 years.

# WIG: prices and volumes



# WIG: prices and volumes



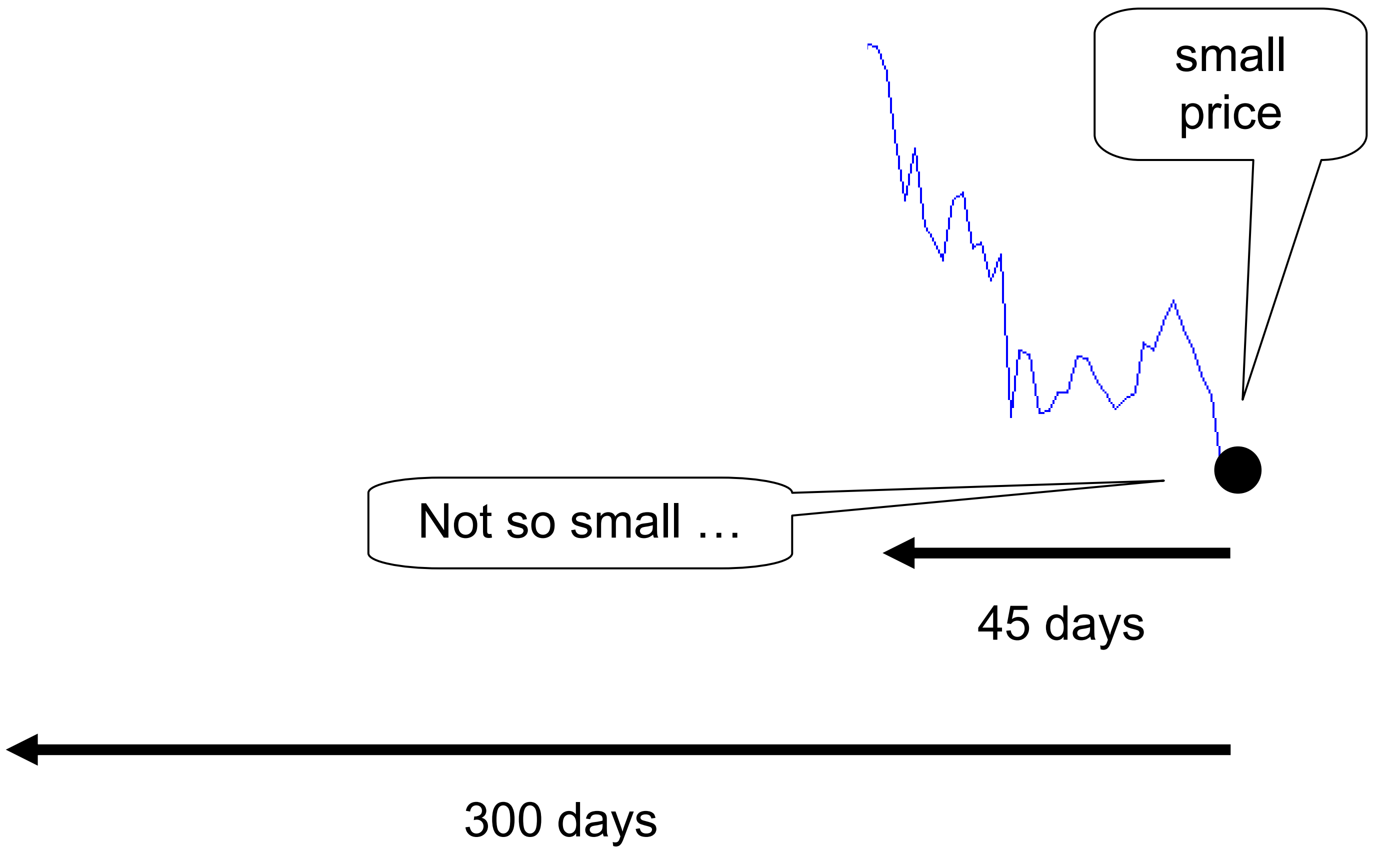
# WIG: prices and volumes


- Prices
  - 3-7 times larger
- Volumes
  - 45 times larger!
  - greater fluctuations than for prices
  - some spikes
- Prices & volumes seem to go together.

# Questions

- How to compare prices/volumes to themselves?
  - 15 years of history
- How to define a big/small?
  - differences: few orders of magnitude.

# Current value vs history



# Current value vs history

If you can not look forward ….

… just look backward!

How far?

We don't know. Lets look *k* - days.

# Variables: current value vs average

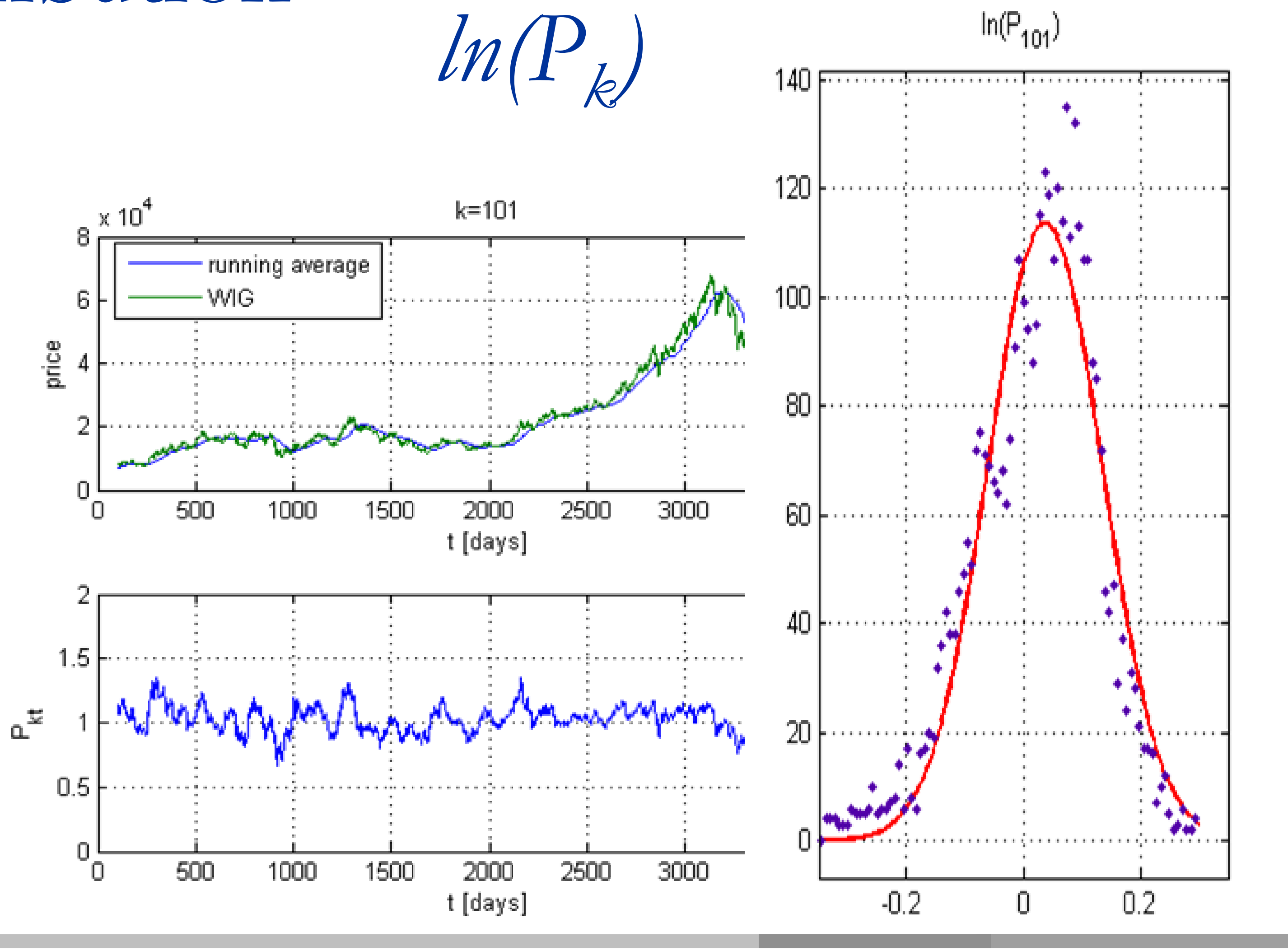
$$P_{k\,t} = \frac{p_t}{\frac{1}{k}(p_{t-1} + p_{t-2} + \ldots + p_{t-k})}$$

$$V_{k\,t} = \frac{v_t}{\frac{1}{k}(v_{t-1} + v_{t-2} + \ldots + v_{t-k})}$$

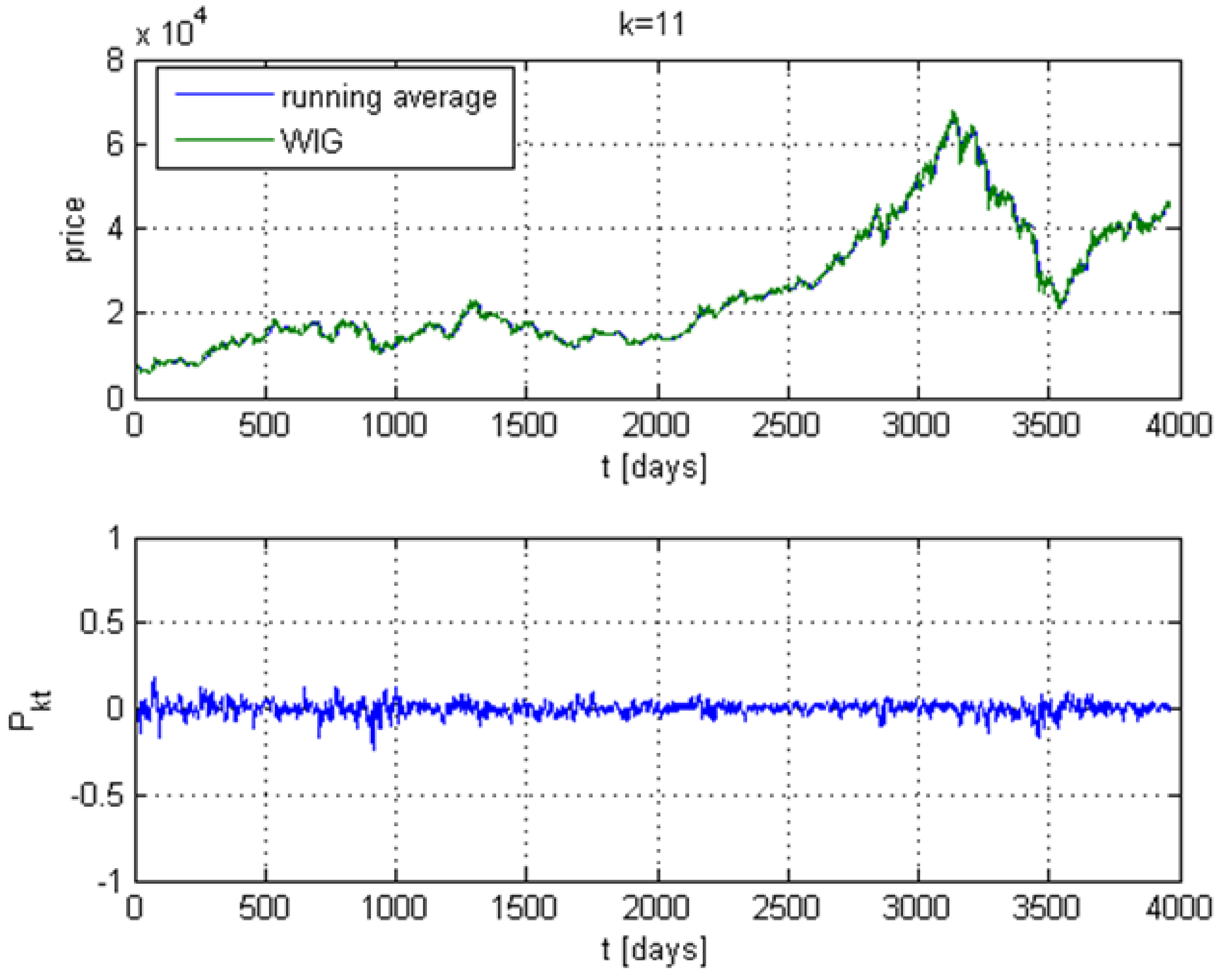
- $p_t$, $v_t$ - daily closure price, trading volume
- $k$ - investor memory ($k$<=401).

# $P_k$ distribution

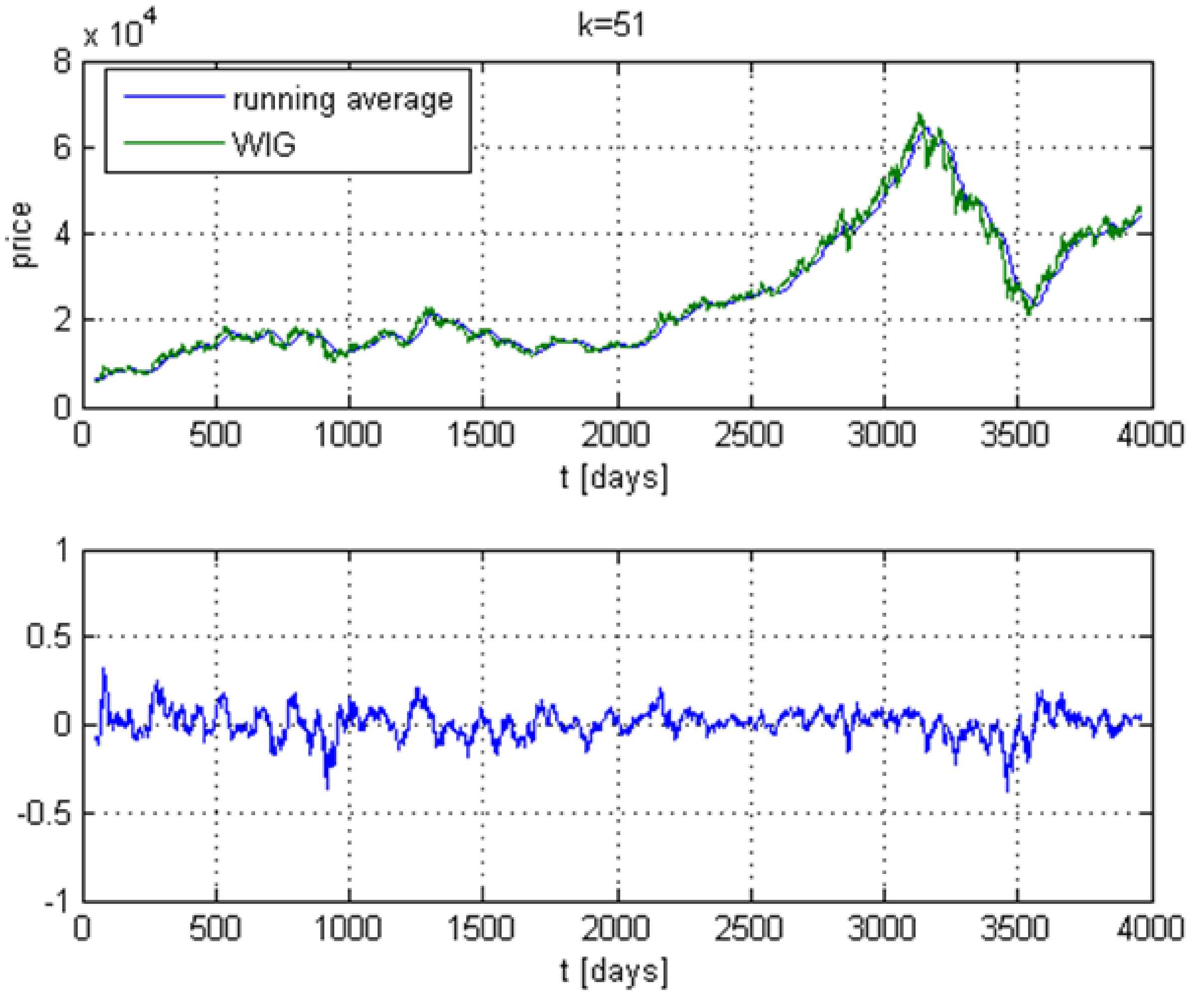
$ln(P_k)$



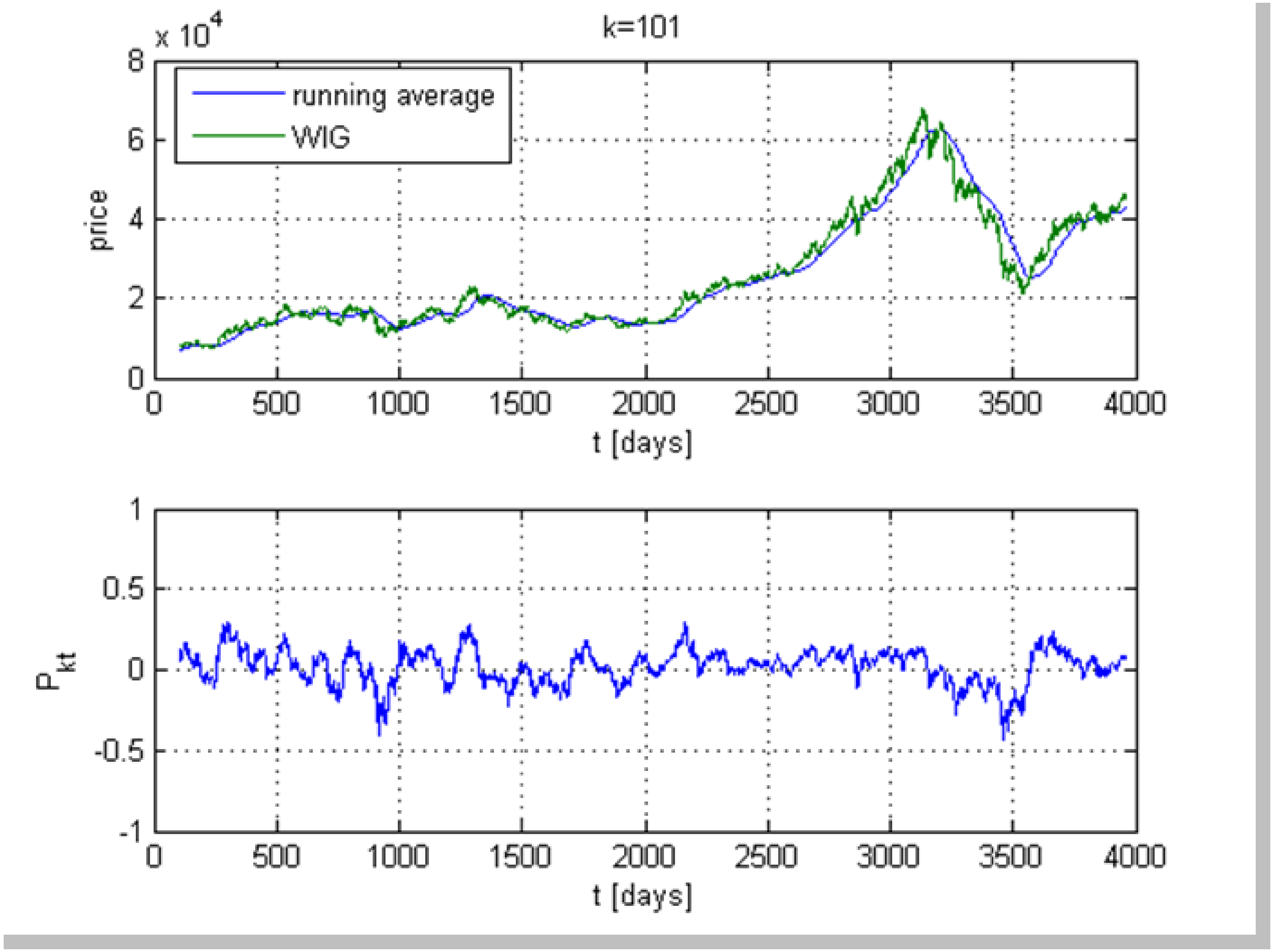
# $\ln(P_{11t})$



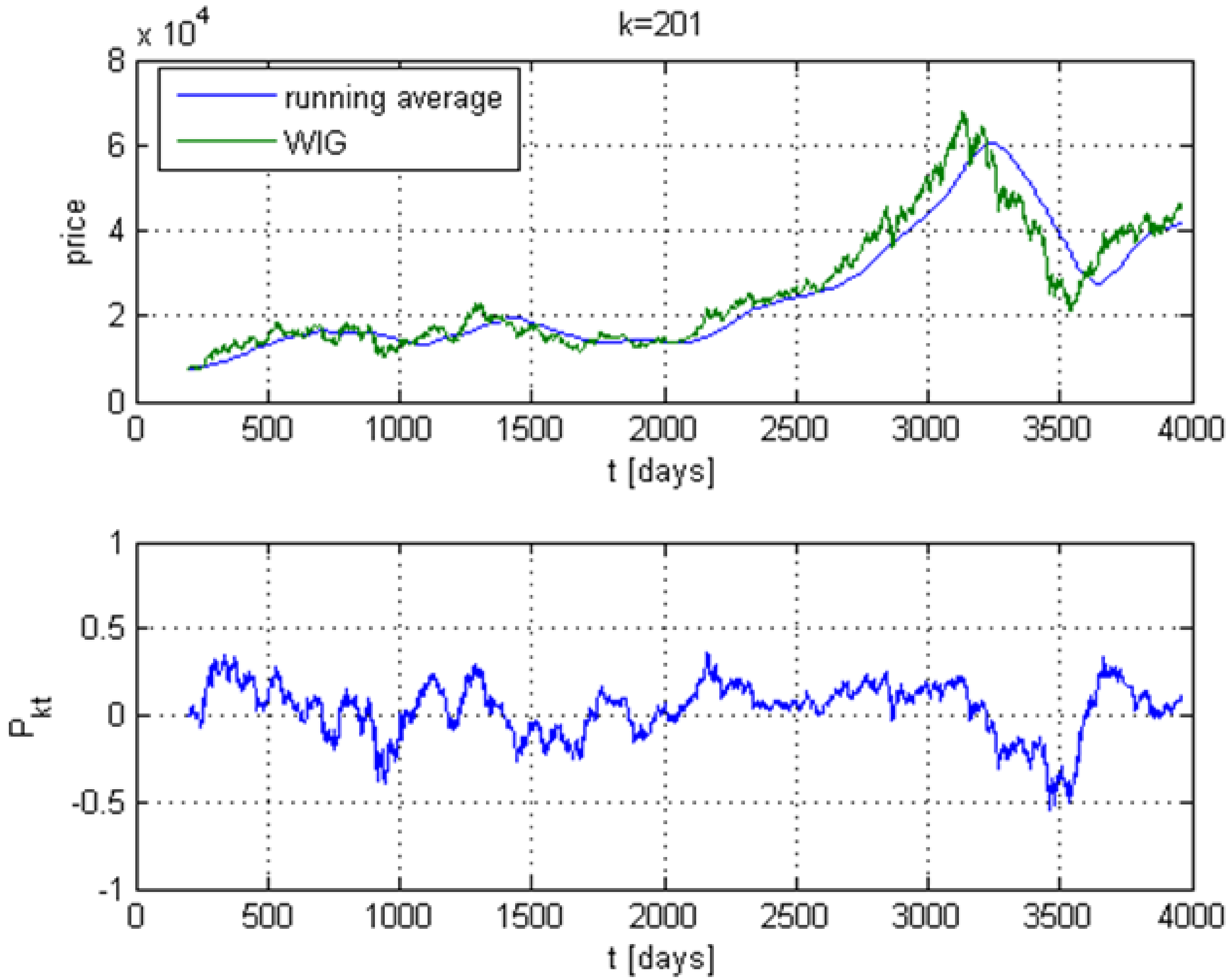
# $\ln(P_{51t})$



# $\ln(P_{101t})$

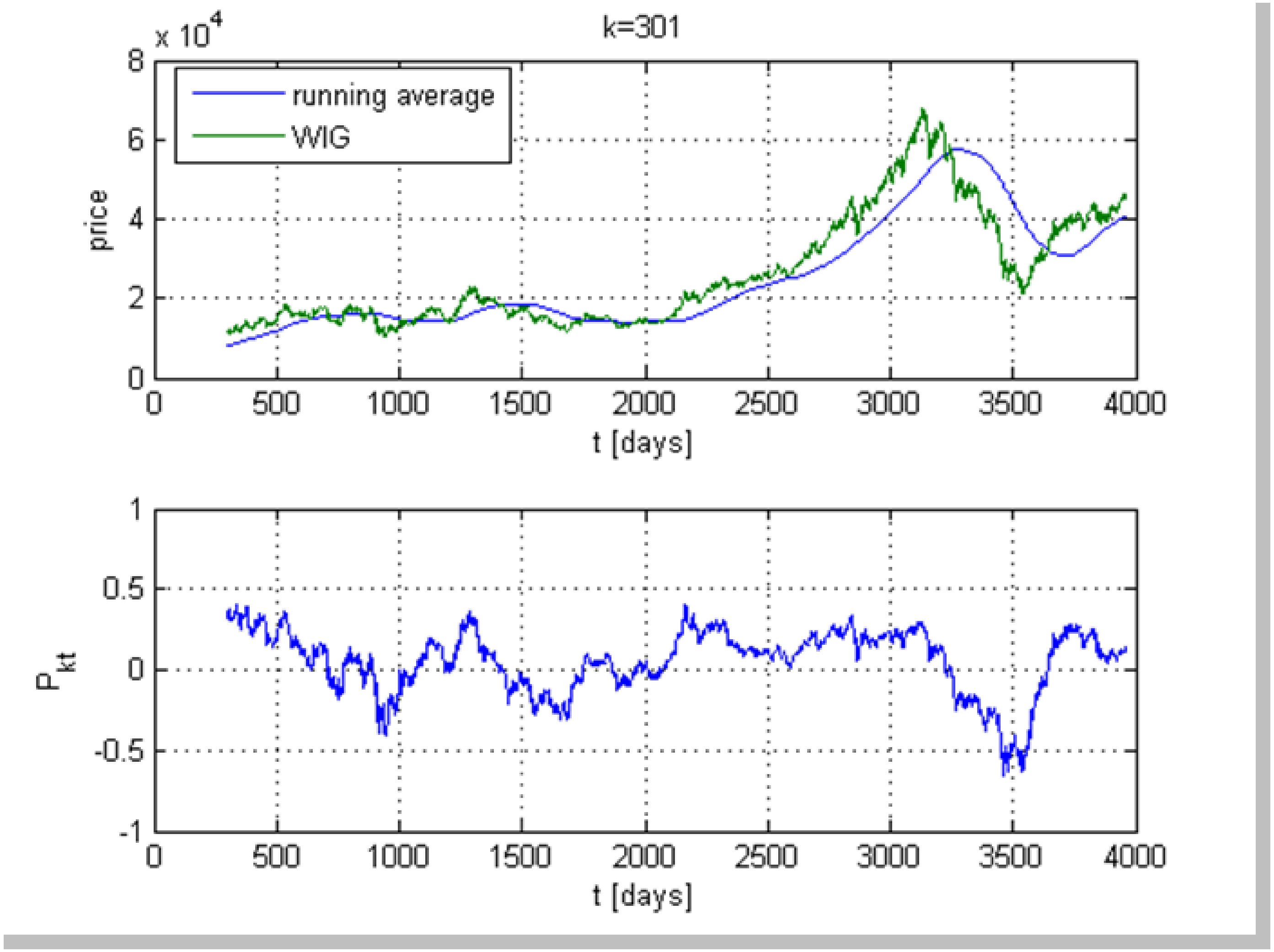


# $\ln(P_{201t})$

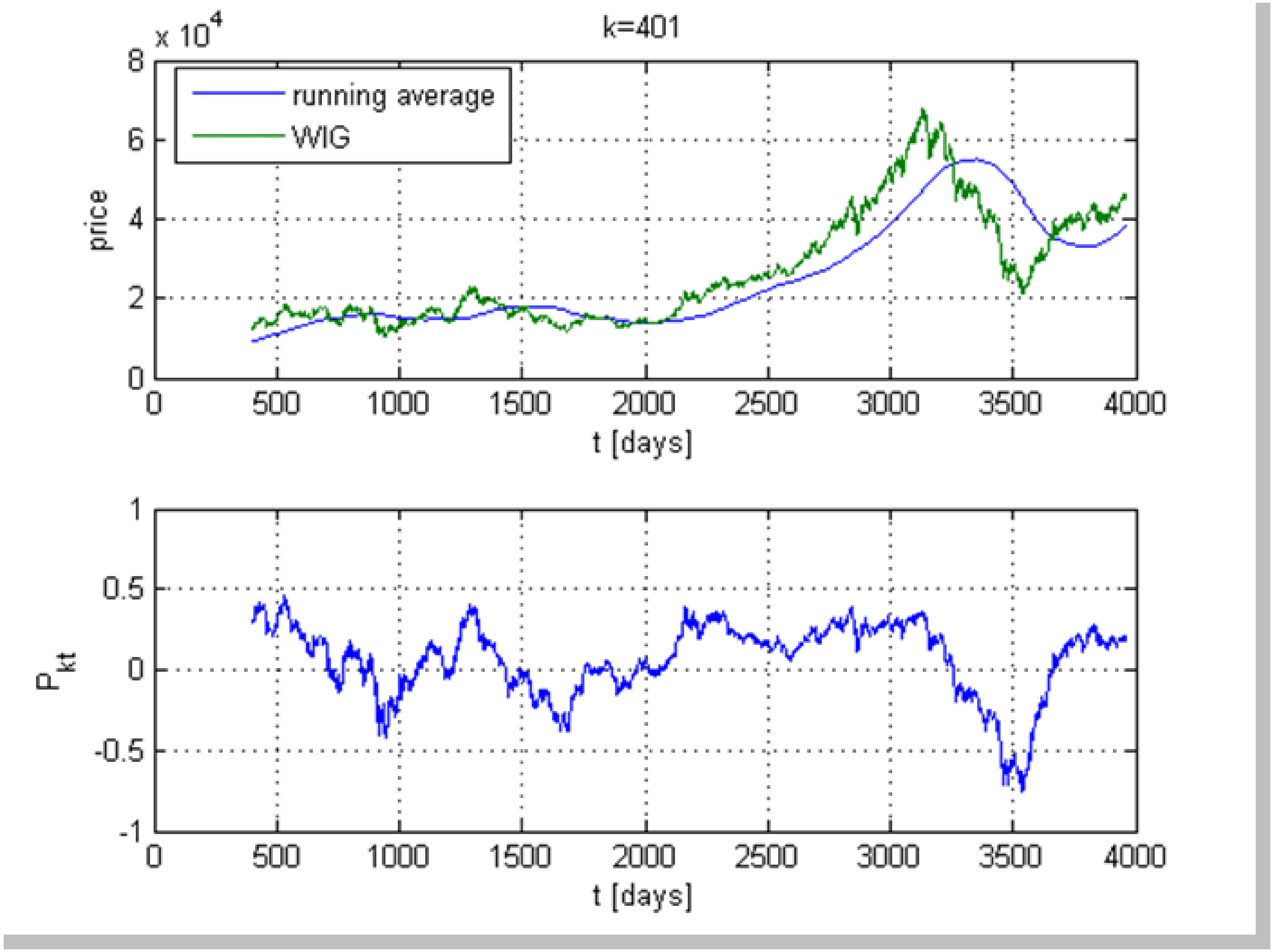


x 10^4
k=201
running average
WIG
price
0
2
4
6
8
0
500
1000
1500
2000
2500
3000
3500
4000
t [days]
P_kt
1
0.5
0
-0.5
-1
0
500
1000
1500
2000
2500
3000
3500
4000
t [days]

# $\ln(P_{301t})$

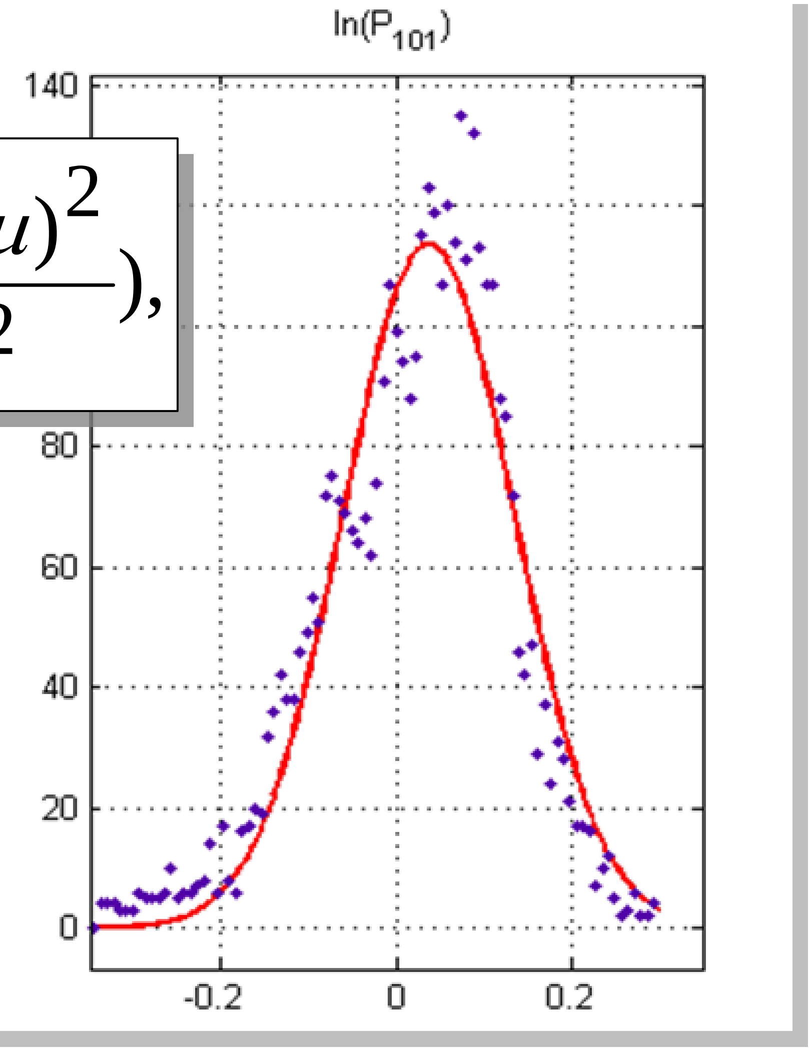


# $\ln(P_{401t})$



# $ln(P_k)$ Gaussian fit

$$f(x) = a\exp\left(-\frac{(x-\mu)^2}{\sigma^2}\right),$$



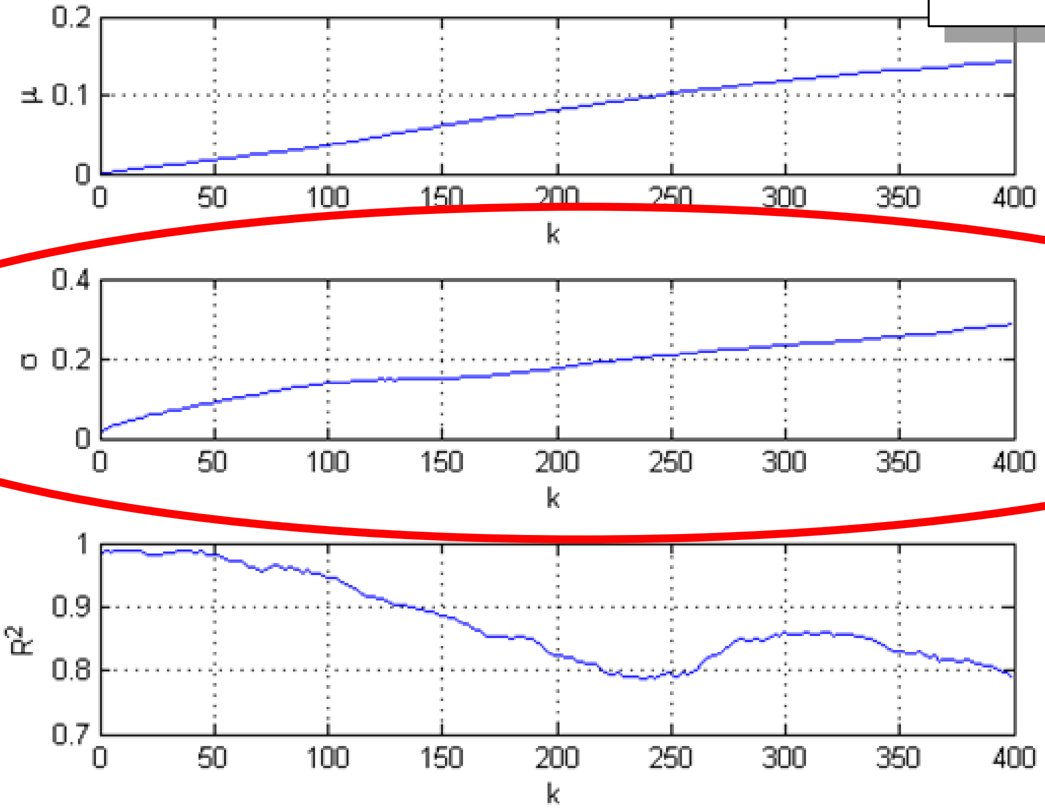

μ = 0.037 (0.032 ; 0.0425)

σ = 0.138 (0.131 ; 0.1455)

$R^2$ = 0.93

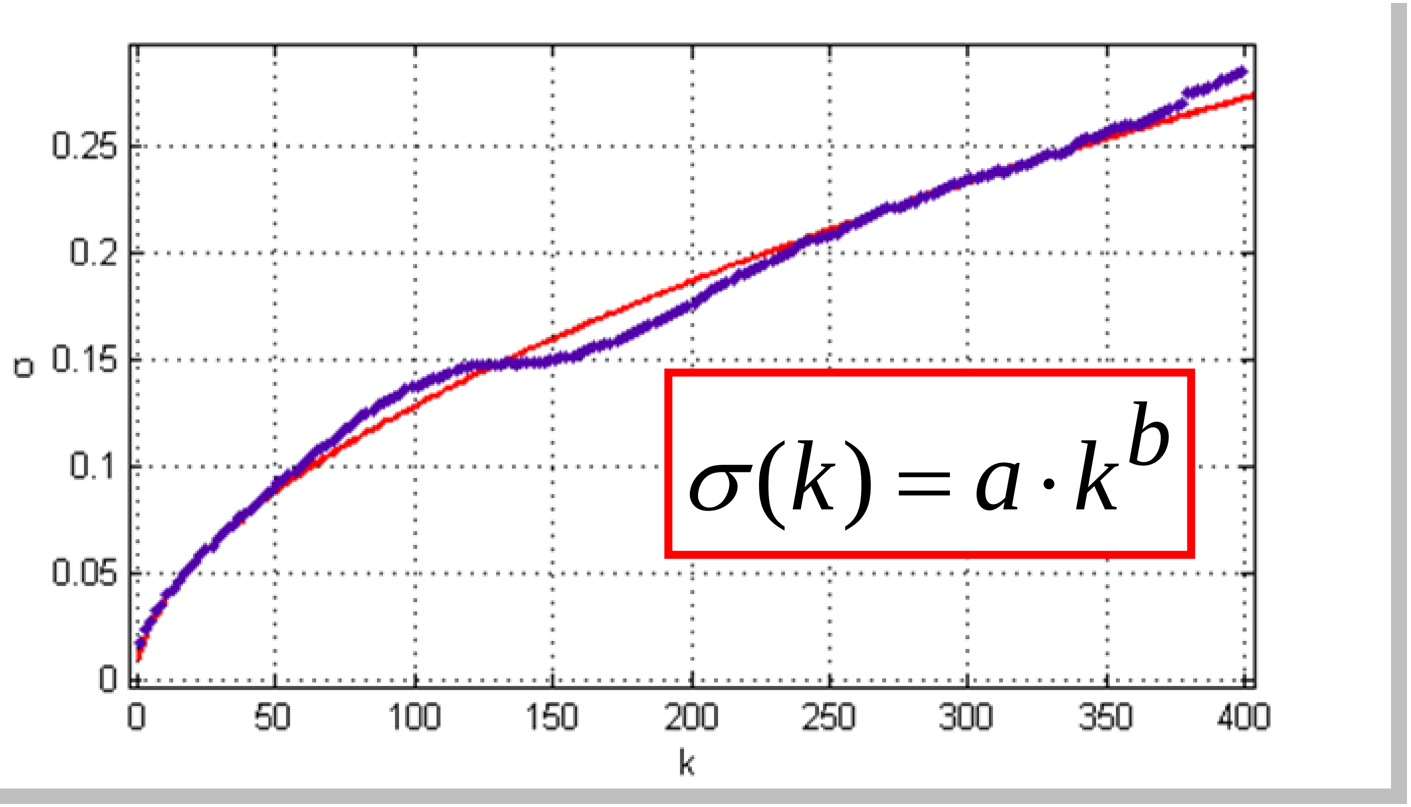
# $\ln(P_k)$ Gaussian fit parameters

$$f(x) = a\exp\left(-\frac{(x-\mu)^2}{\sigma^2}\right),$$



# Autocorrelations

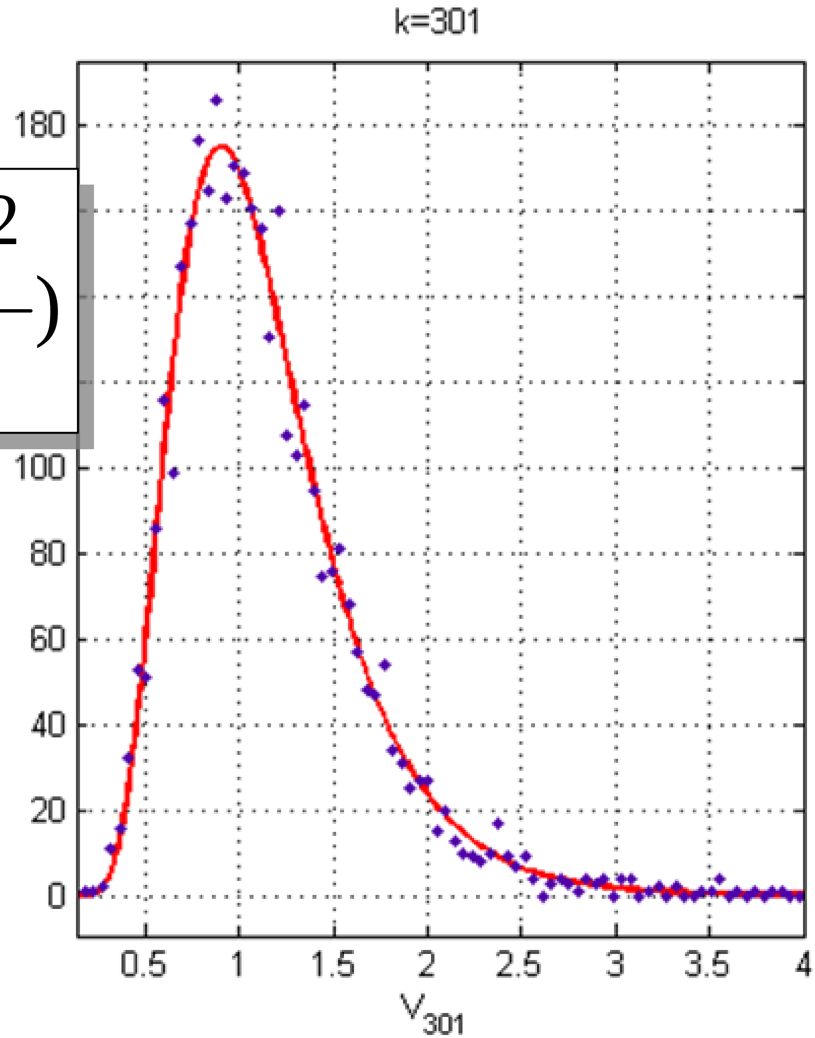
- No autocorrelations -> $\sigma(k) \propto \sqrt{k}$
- $R^2$ = 0.
- $b$ = 0.5




- **small p**

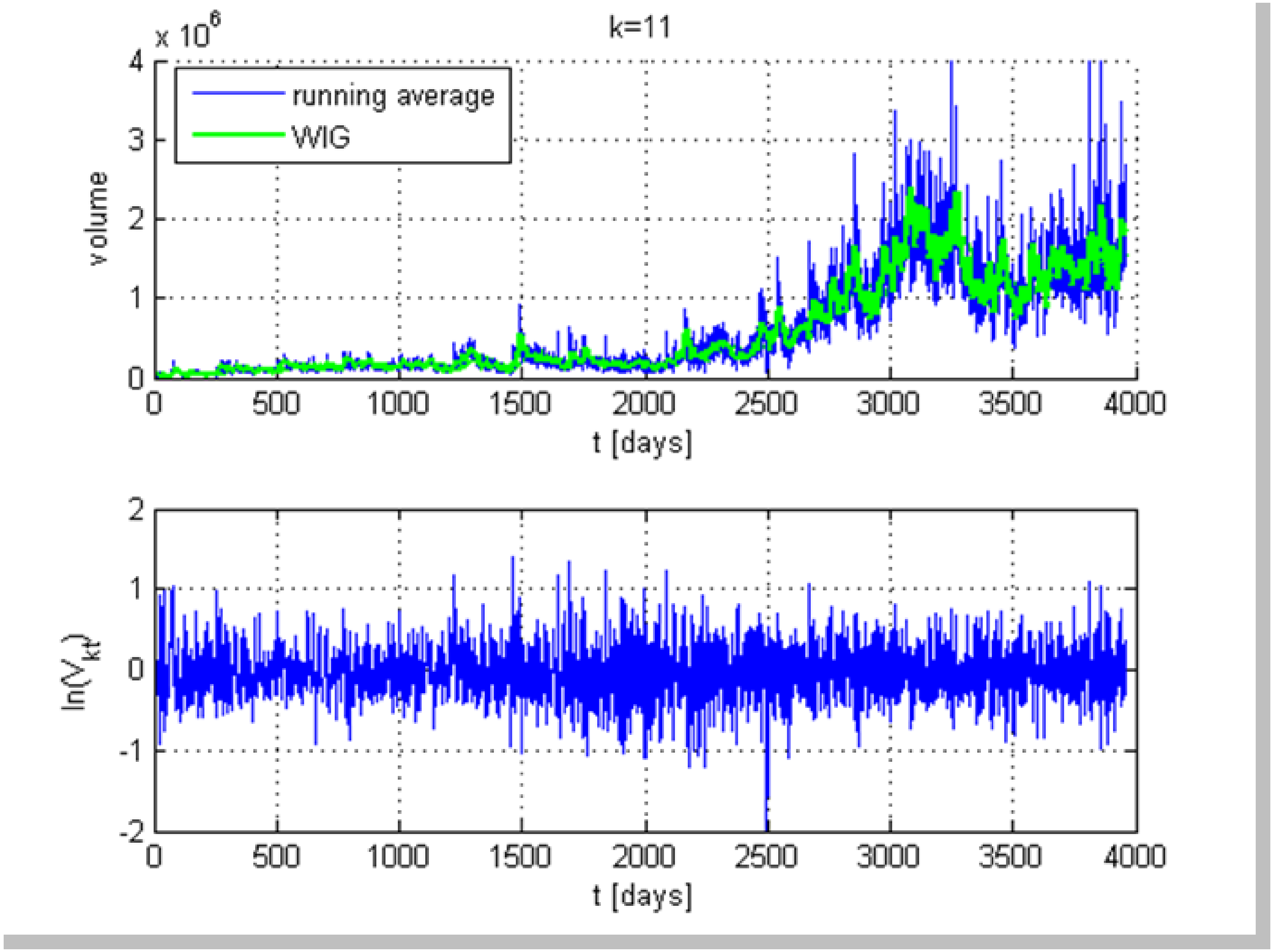
# $V_k$ distribution

$$f(x) = \frac{a}{x}\exp\left(-\frac{(\ln(x)-\mu)^2}{\sigma^2}\right)$$

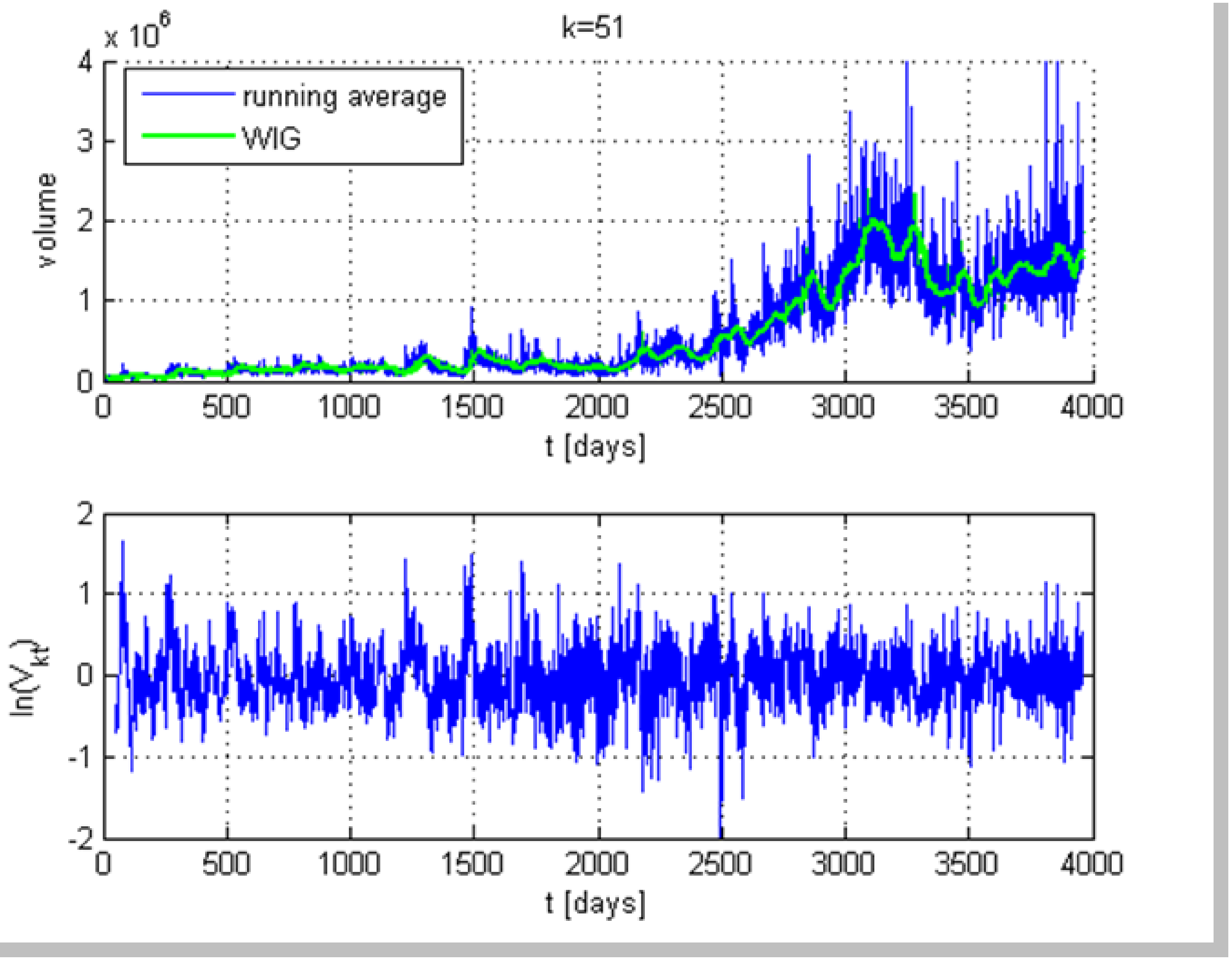



$\mu$ = 0.053 (0.042 ; 0.065)
$\sigma$ = 0.564 (0.550 ; 0.579)
$R^2$ =0.988

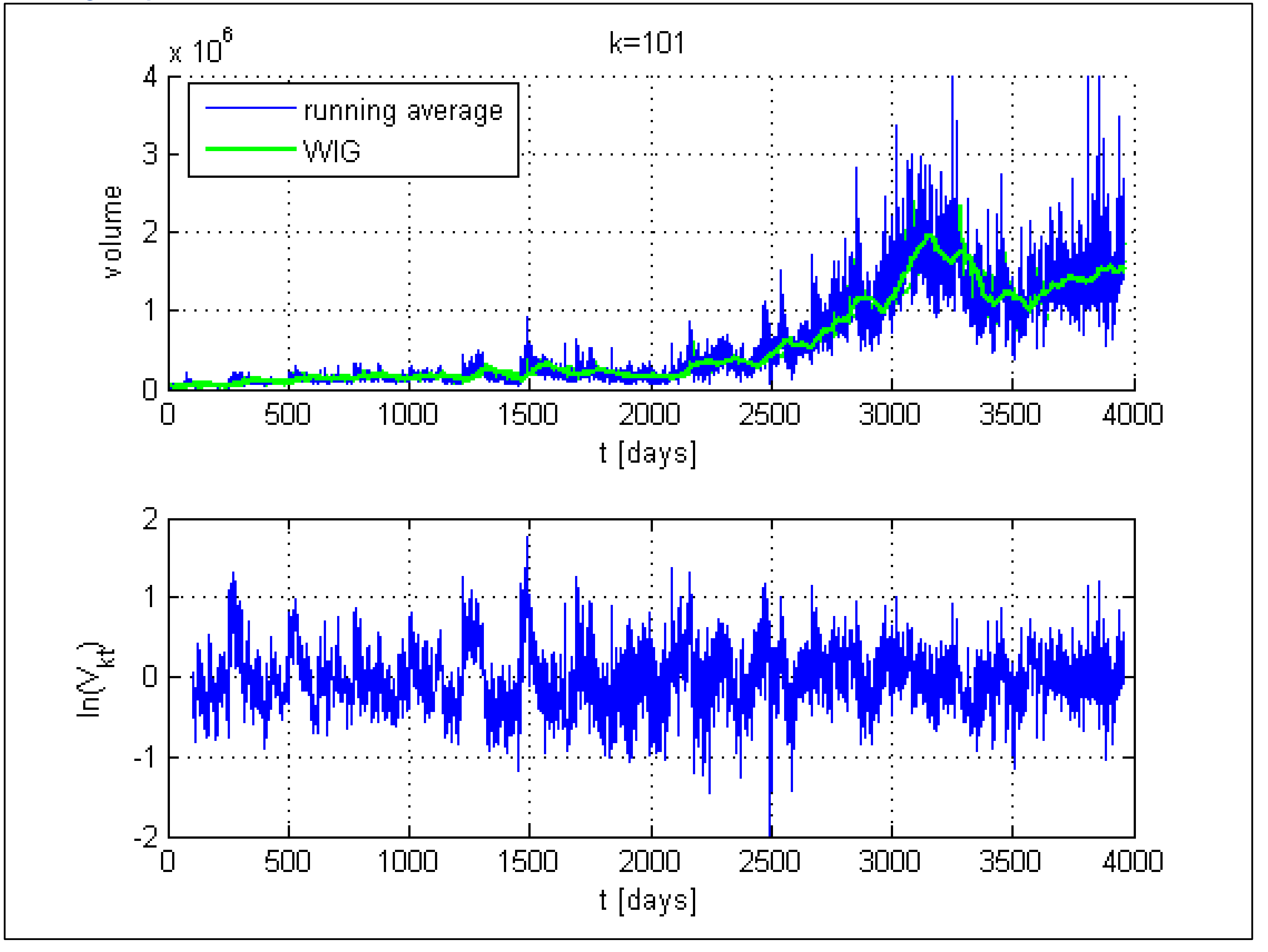
# $\ln(V_{11t})$



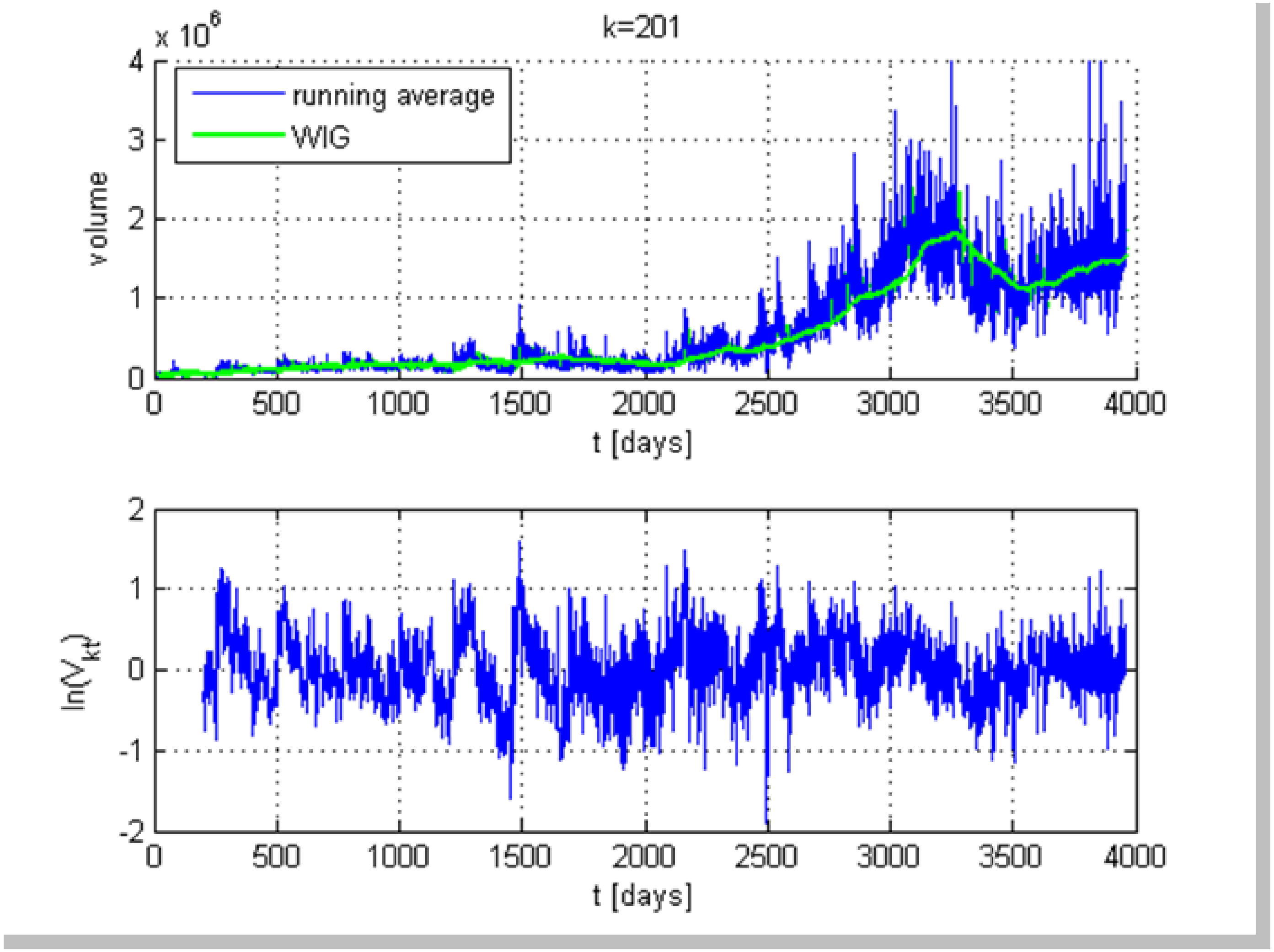
# $\ln(V_{51t})$



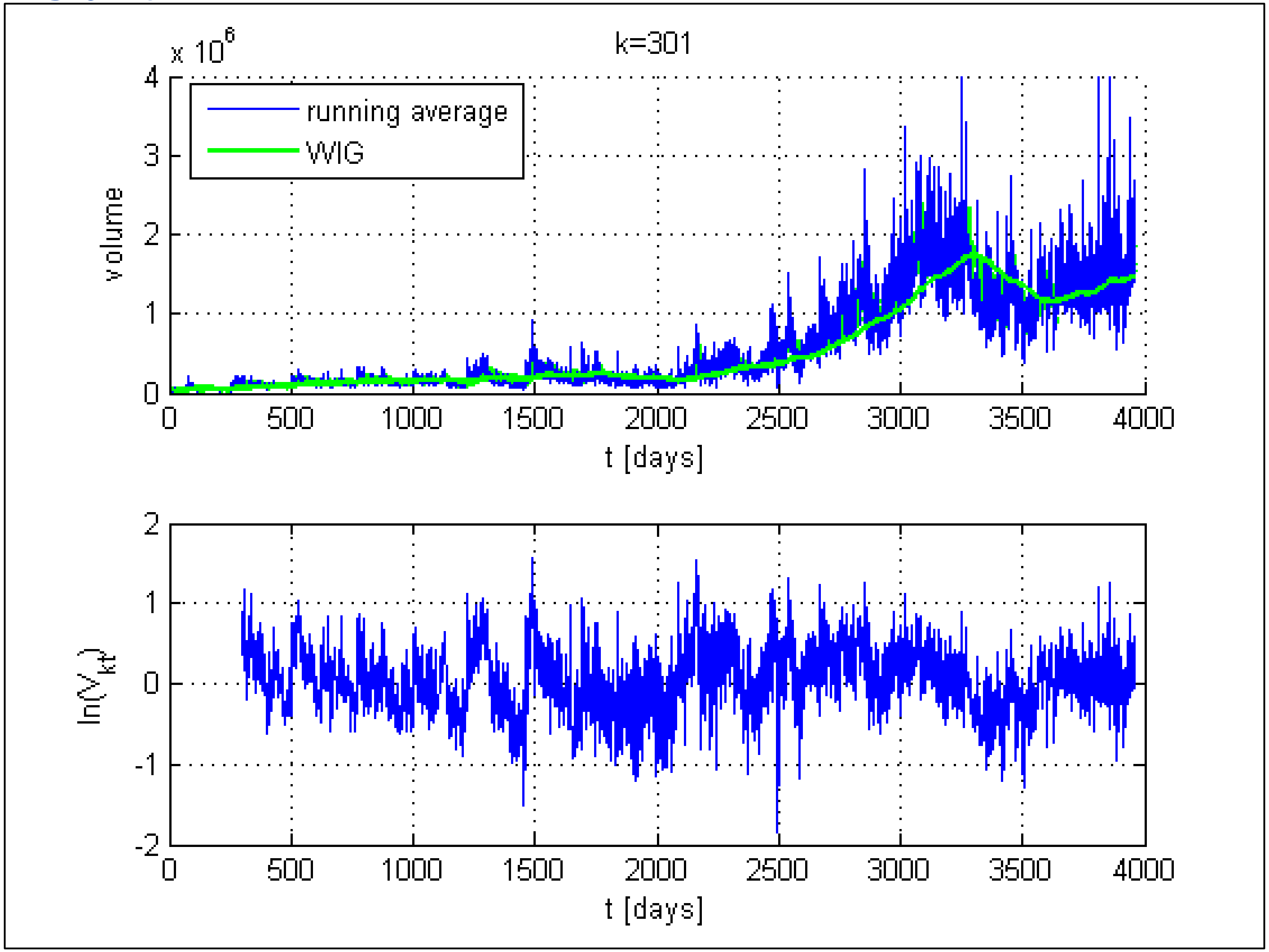
# $\ln(V_{101t})$



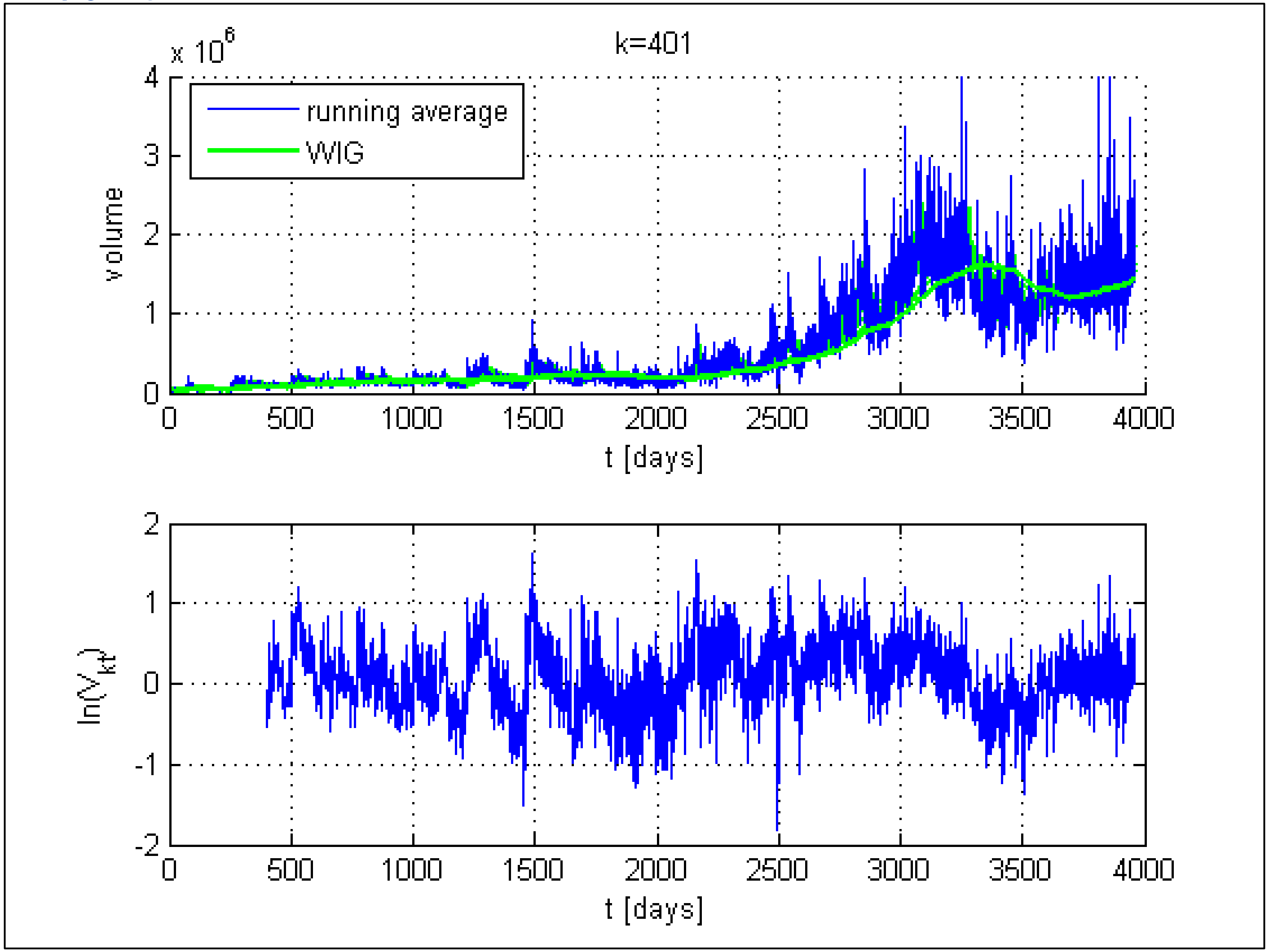
# $\ln(V_{201t})$



# $\ln(V_{301t})$

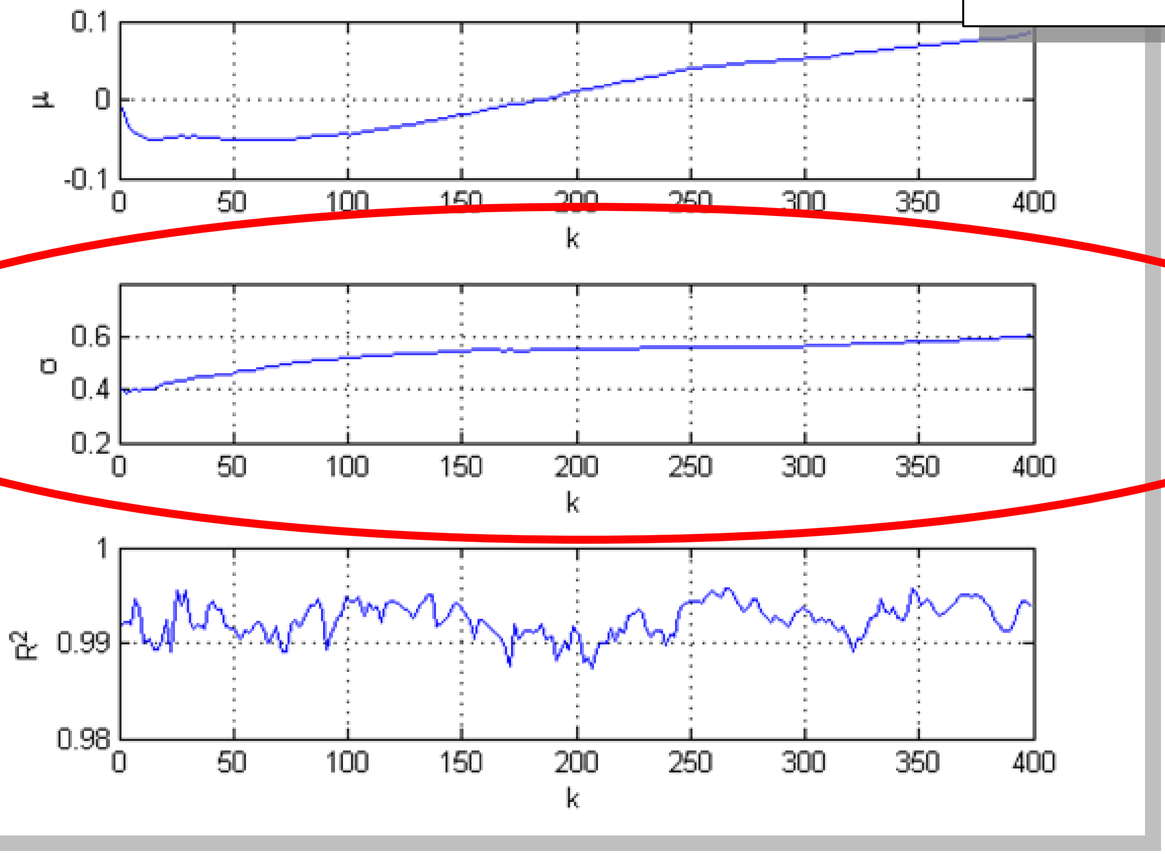


# $\ln(V_{401t})$



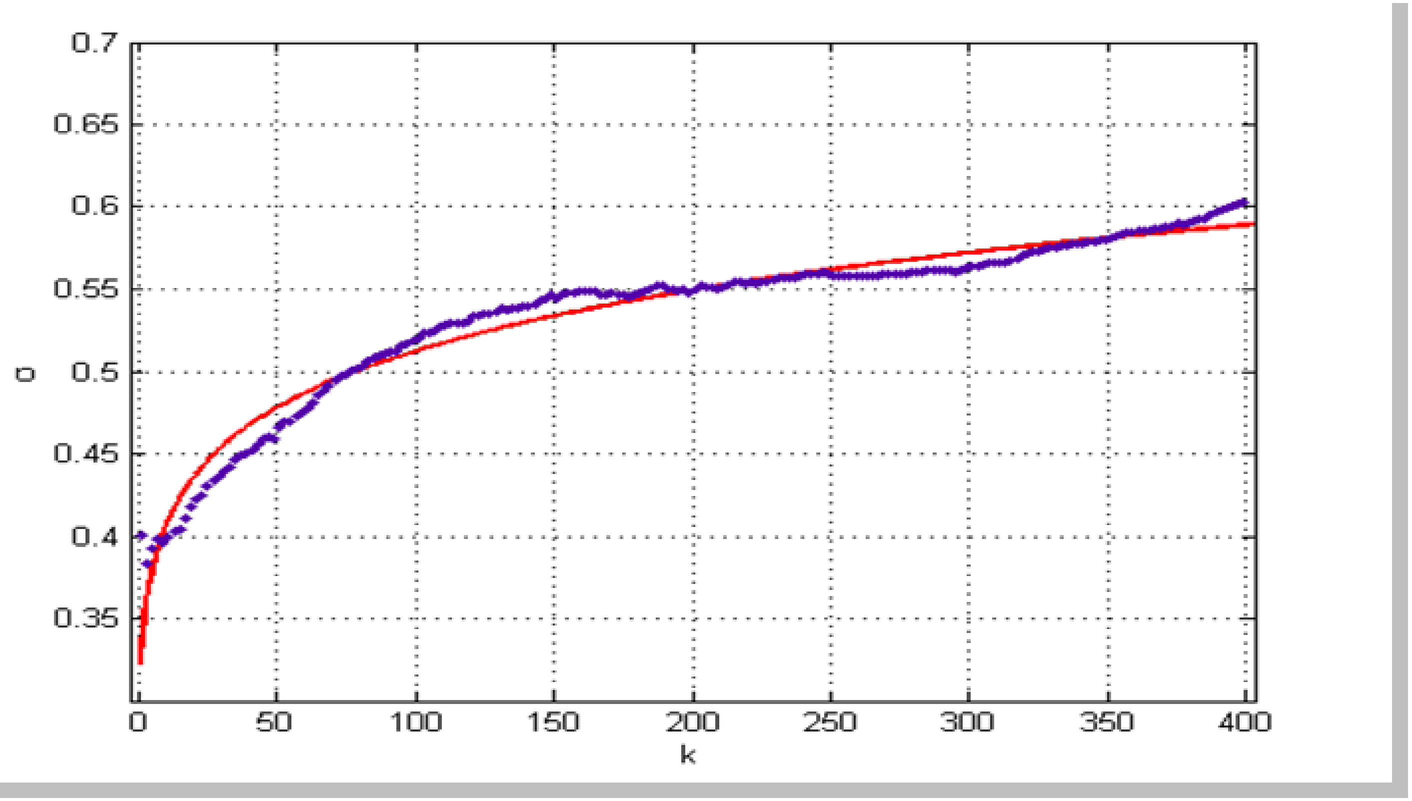
# $ln(V_k)$ - Gaussian fit parameters

$$f(x) = a\exp\left(-\frac{(x-\mu)^2}{\sigma^2}\right),$$



# Autocorrelations

- No autocorrelations -> $\sigma(k) \propto \sqrt{k}$
- $R^2$ = 0.
- $b$ = 0.1




- **big ne**

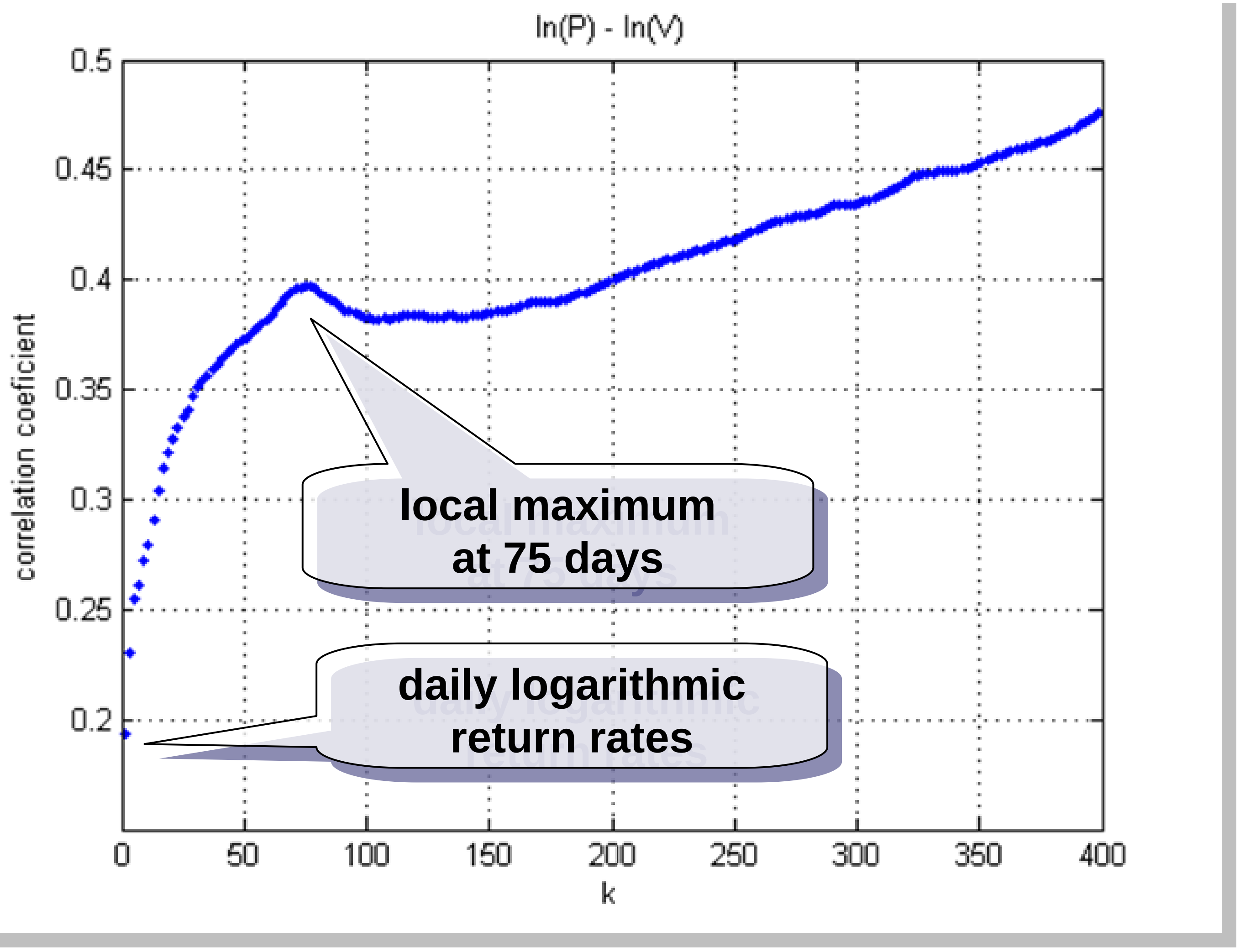
# Autocorrelations: prices vs volumes

- Prices
- Small positive
- Trend to keep changes
- Prices are unlimited

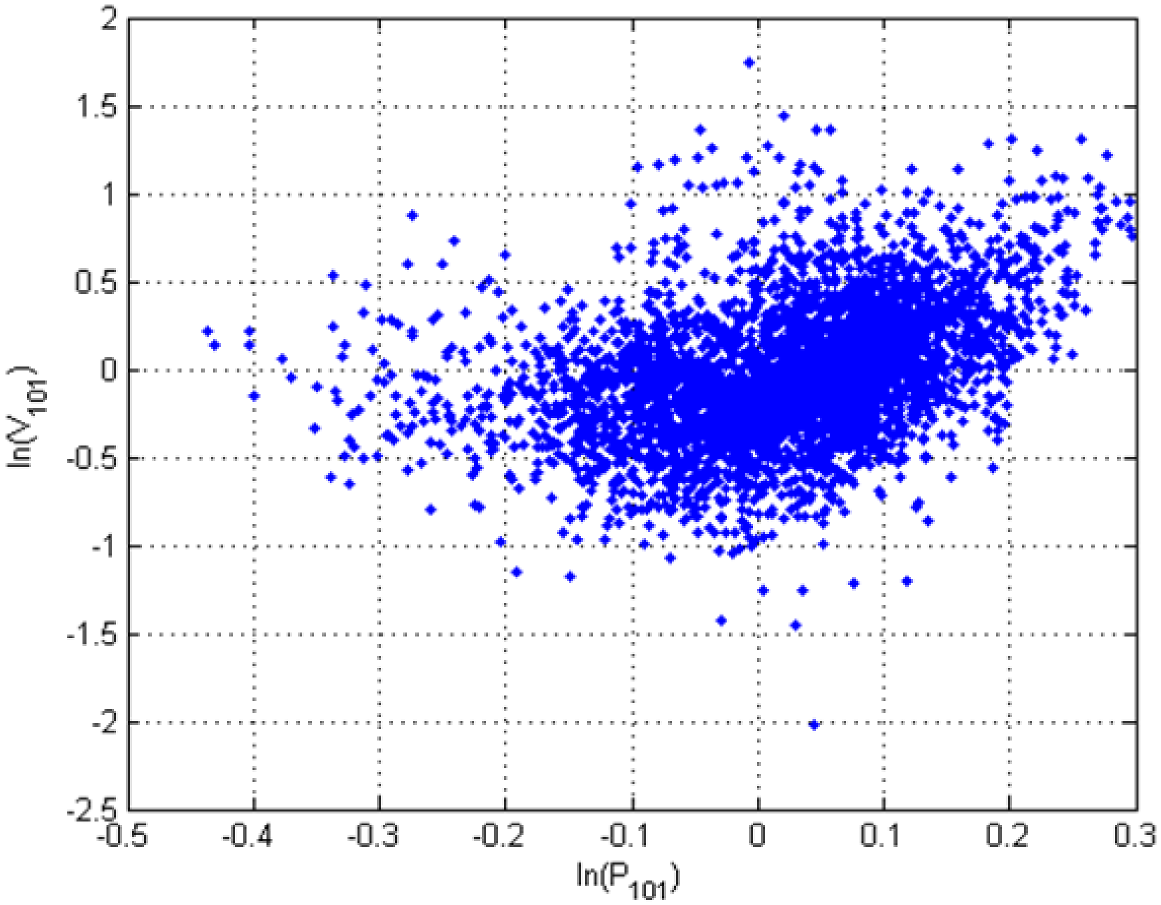
- Volumes
- Big negative
- Counteraction against changes
- Volume is limited

- Demand is restrained by the price growth
- Supppply is restrained by the price descrease.

# Correlations $\ln(P_k)$ vs $\ln(V_k)$



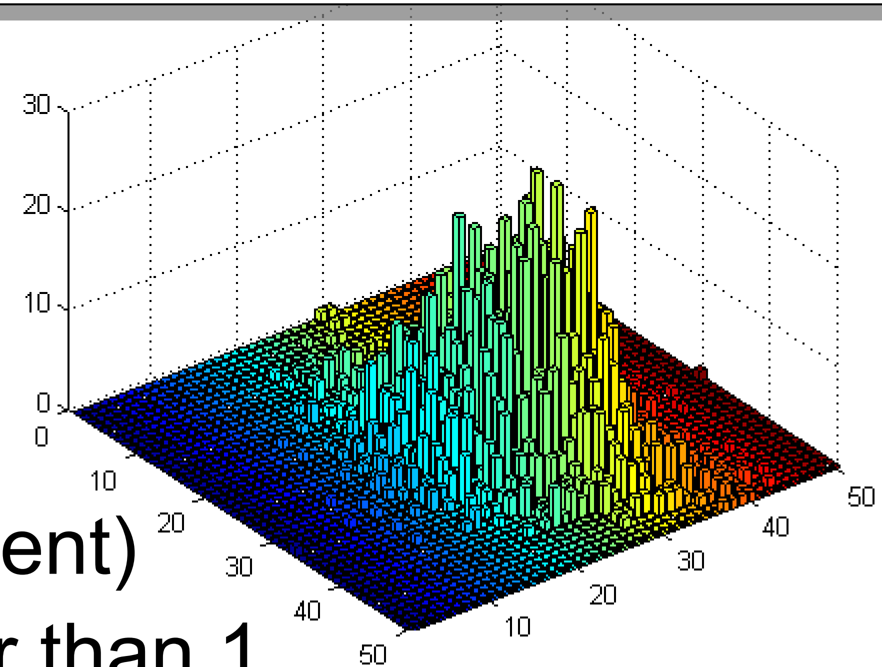
# Correlation plot, k=101



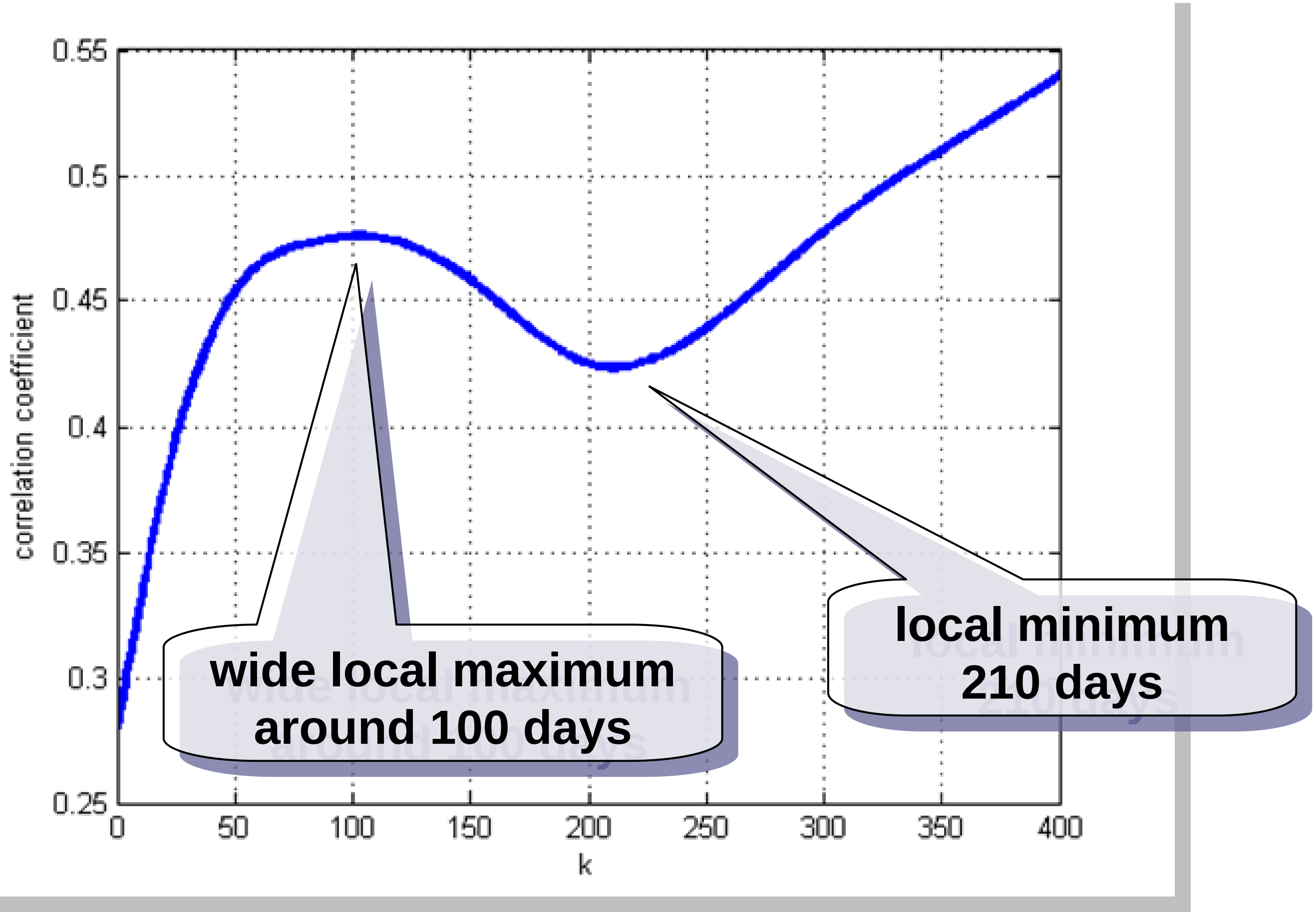
# Parametrization of correlation plots

$$f(p,v) = \frac{A}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(\frac{-1}{2(1-\rho^2)}\left[\frac{(p-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(p-\mu_1)(v-\mu_2)}{\sigma_1\sigma_2} + \frac{(v-\mu_2)^2}{\sigma_2^2}\right]\right),$$

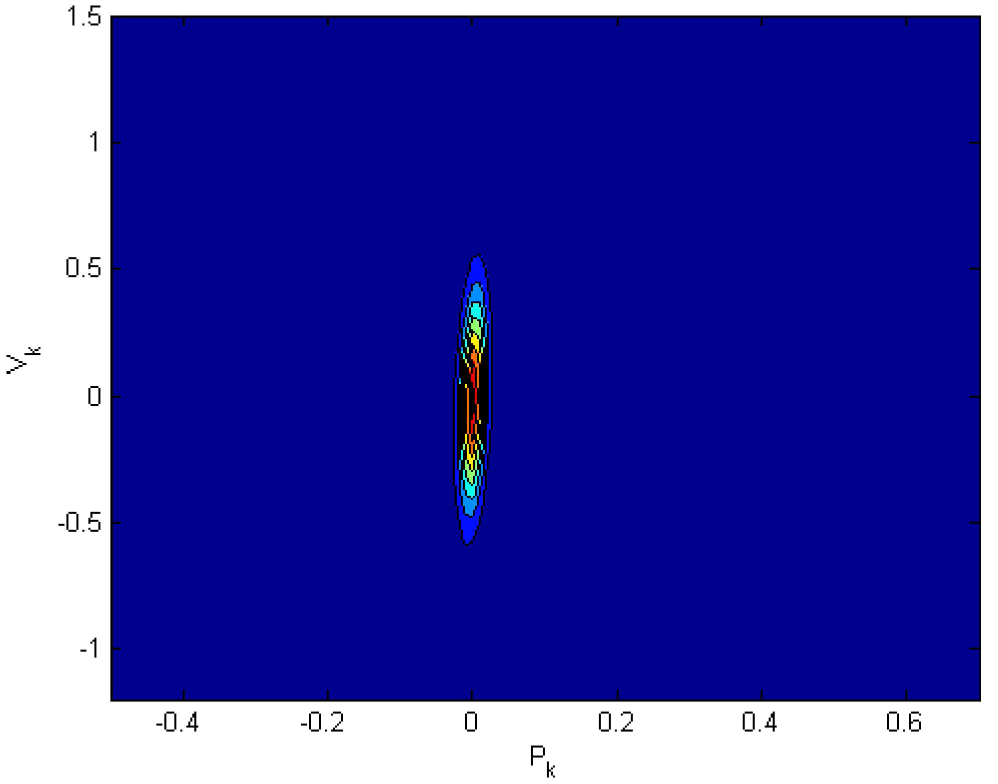
- 2-dim Gauss
- Weighted $\chi^2$
- error = sqrt(bin content)
- $\chi^2$/ndf slightly bigger than 1



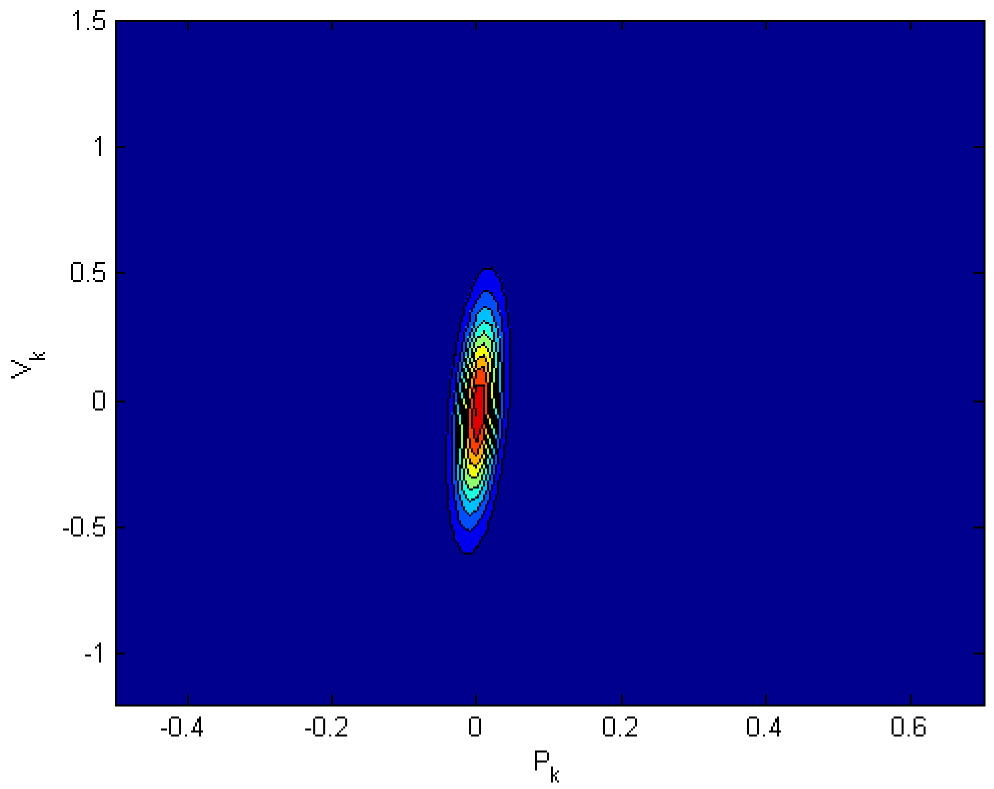
# Correlations $\ln(P_k)$ vs $\ln(V_k)$



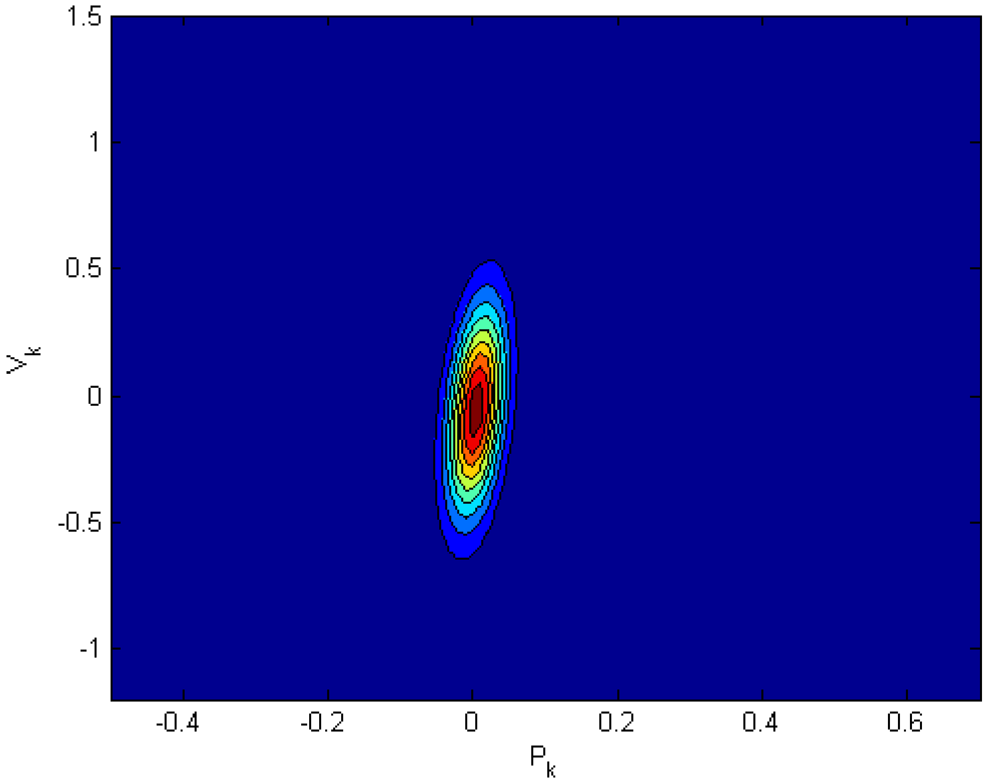
# $\ln(P_k)$ - $\ln(V_k)$, k=1



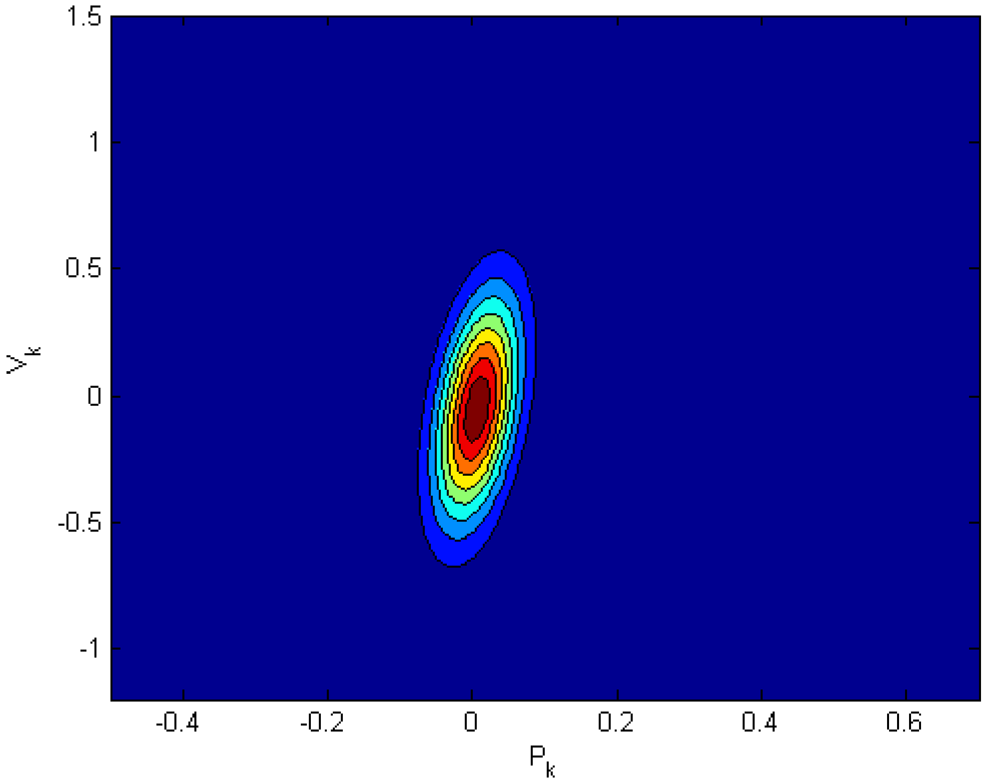
# ln($P_k$) - ln($V_k$), k=5



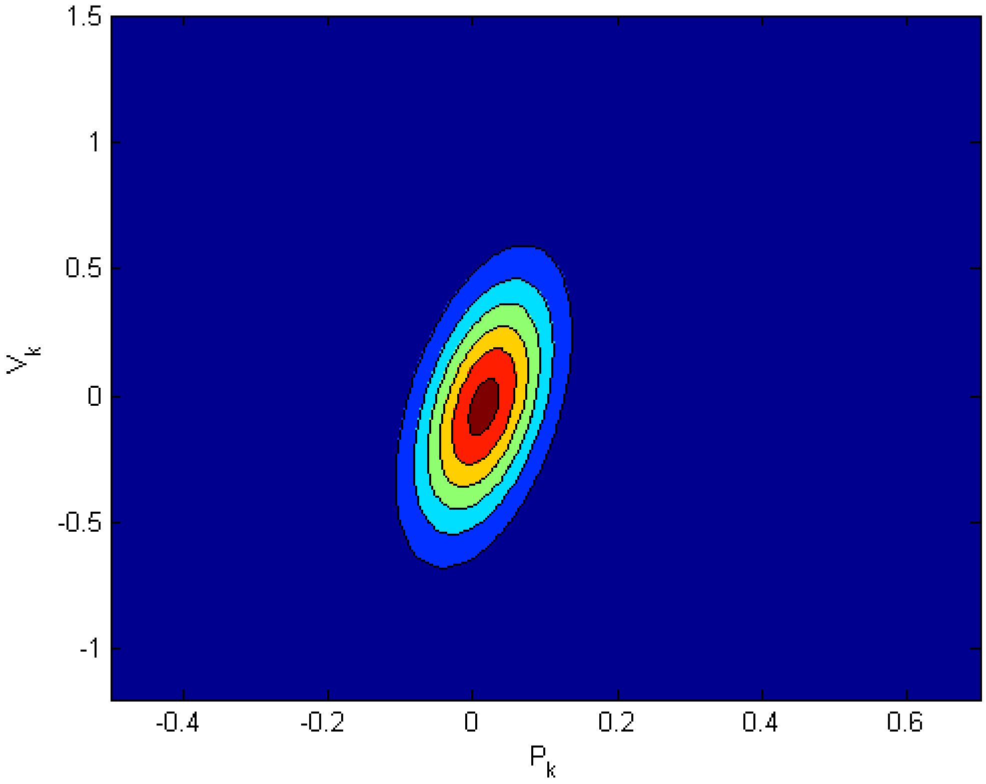
# $\ln(P_k)$ - $\ln(V_k)$, k=11



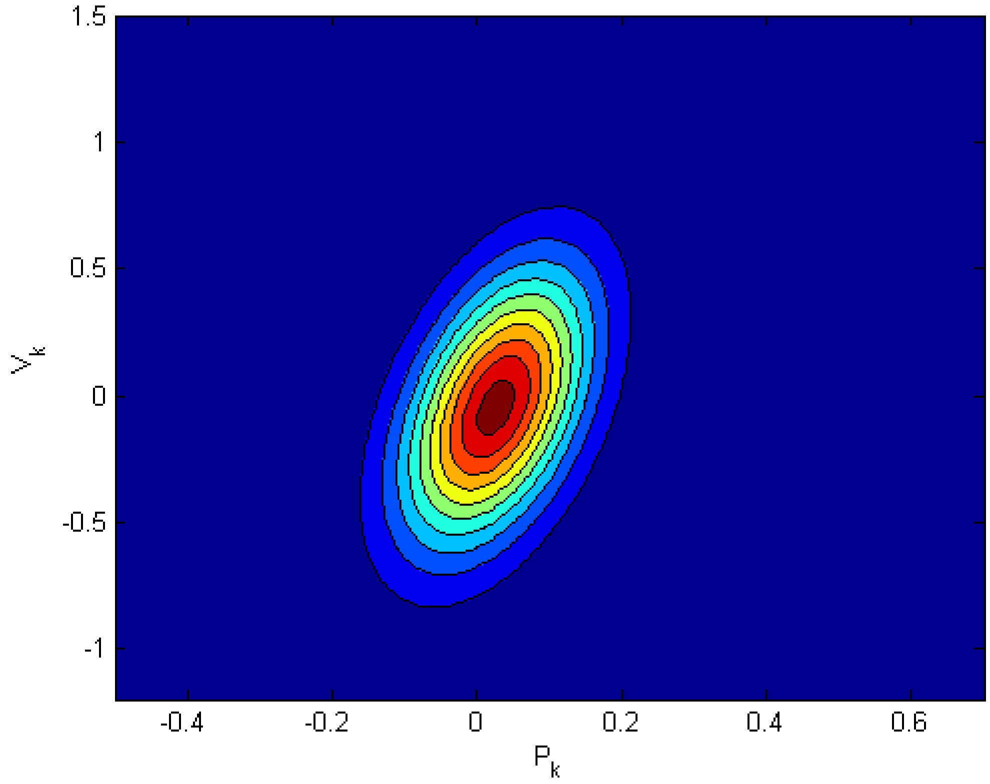
# $\ln(P_k)$ - $\ln(V_k)$, k=21



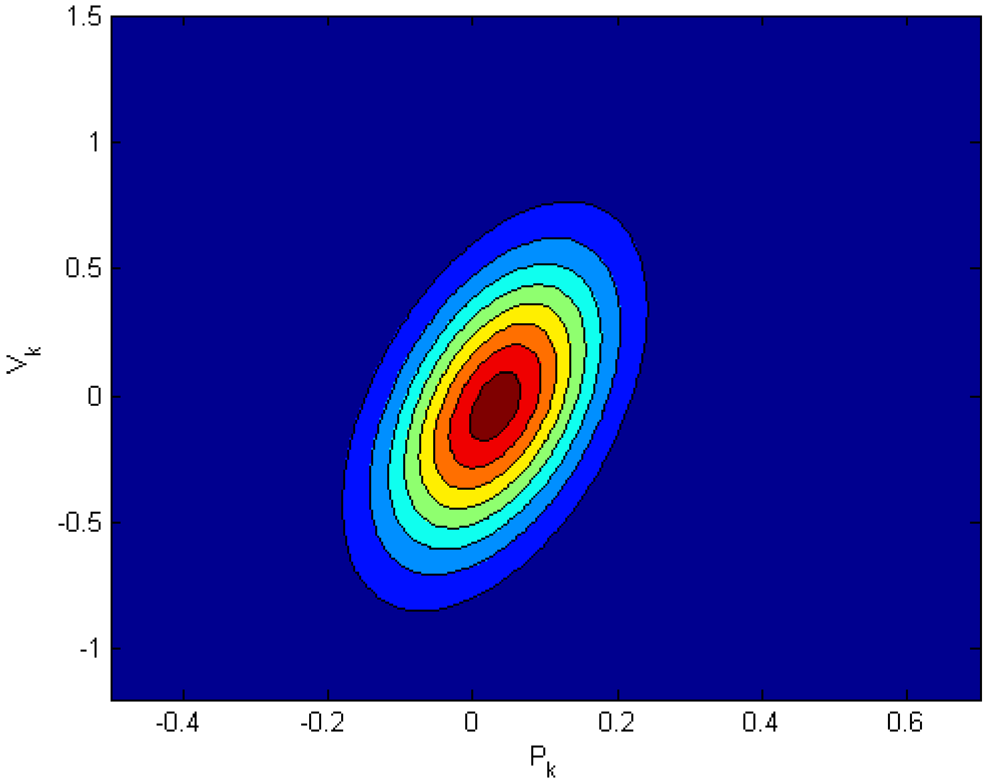
# $\ln(P_k)$ - $\ln(V_k)$, k=51



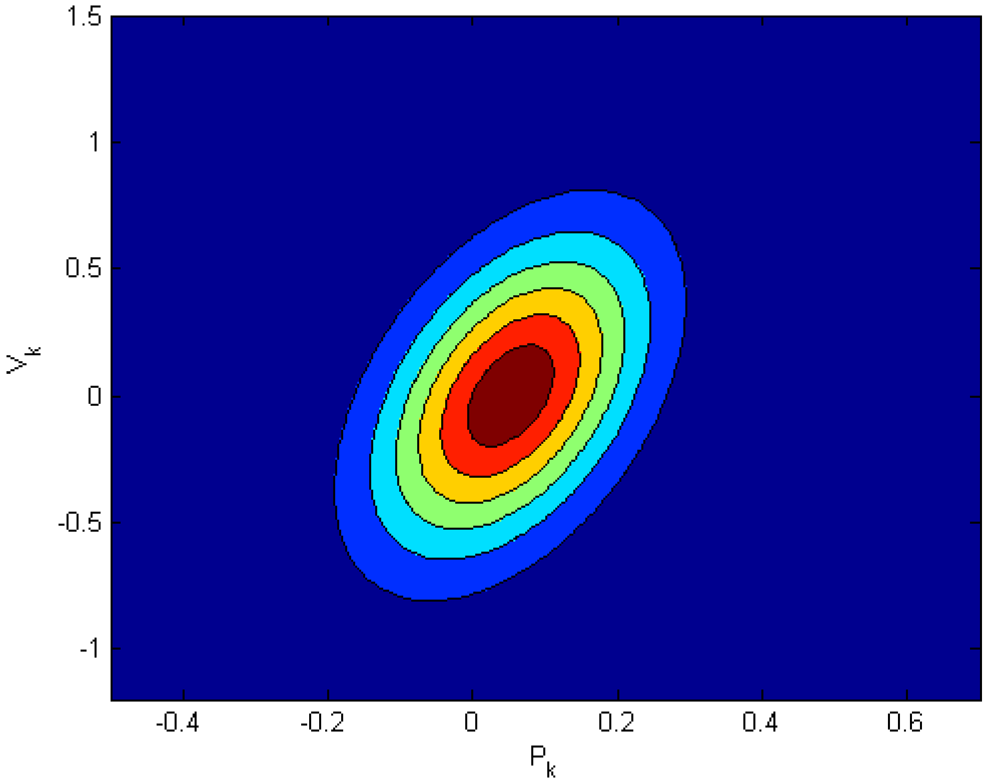
# $\ln(P_k) - \ln(V_k)$, k=81



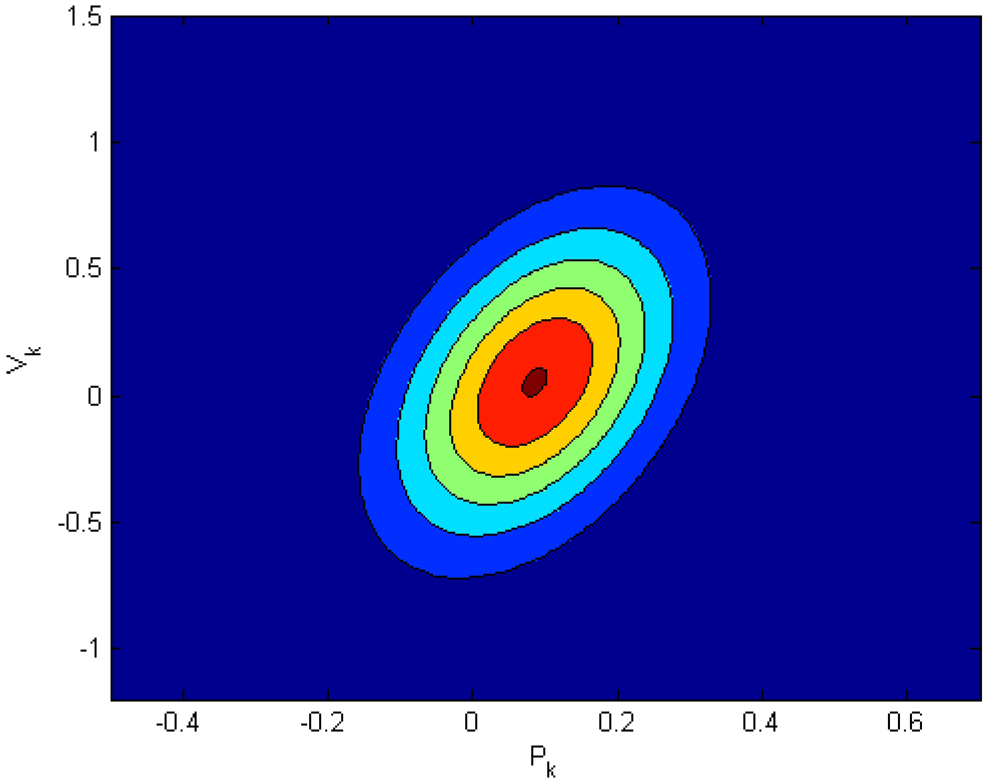
# $\ln(P_k)$ - $\ln(V_k)$, k=101



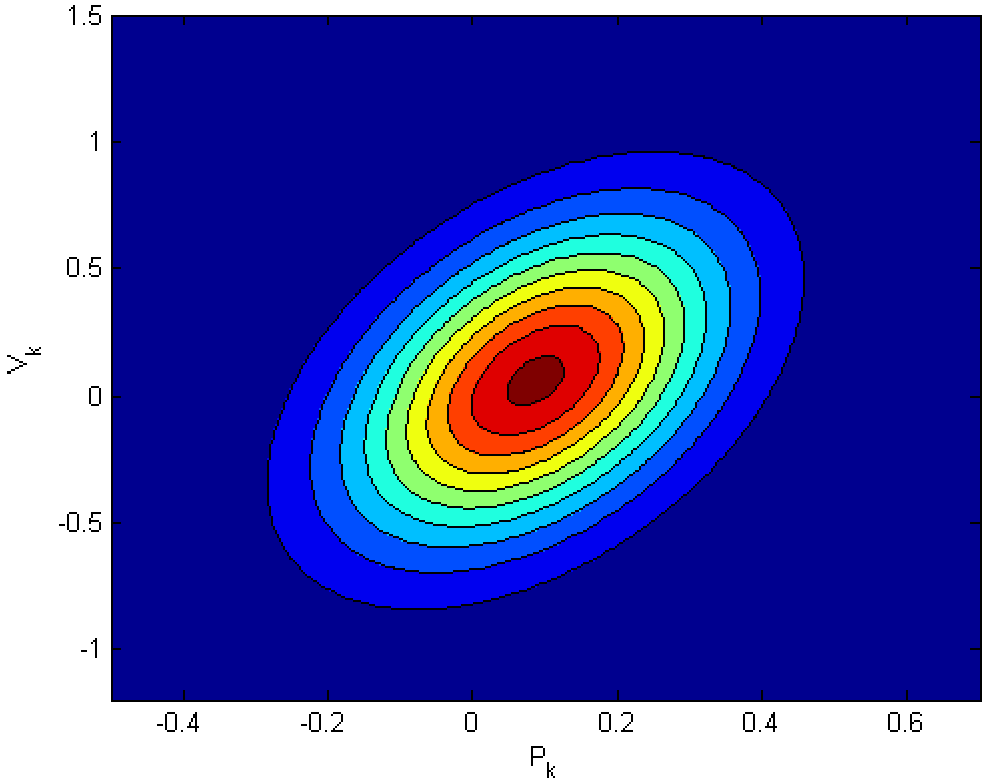
# $\ln(P_k)$ - $\ln(V_k)$, k=151



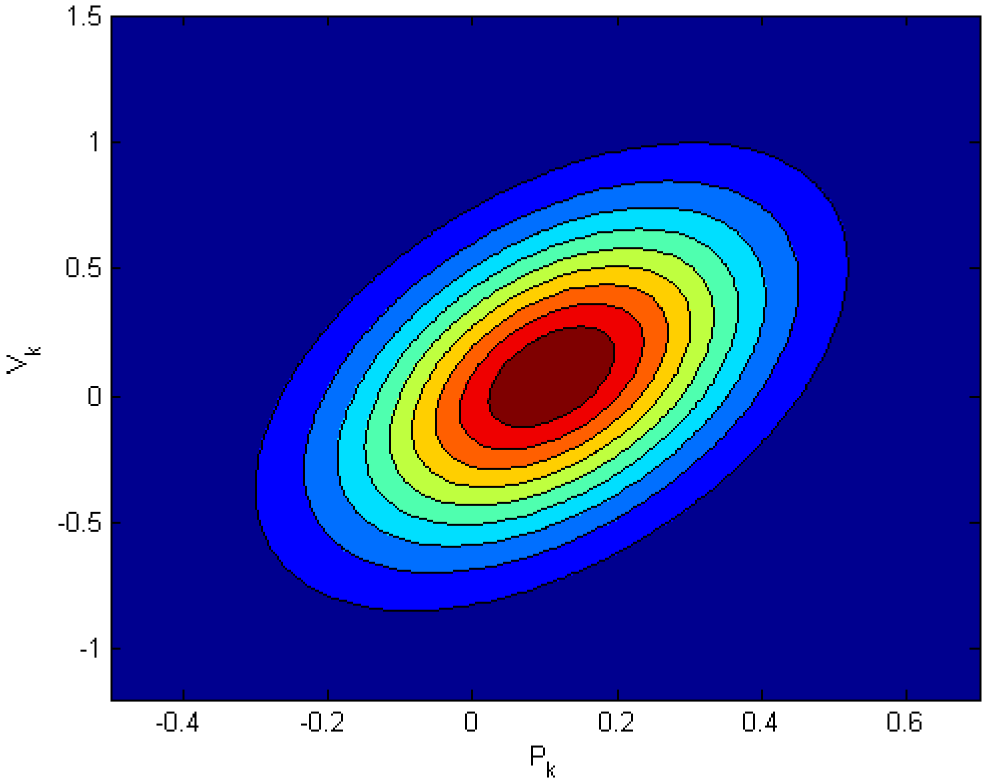
# $\ln(P_k)$ - $\ln(V_k)$, k=201



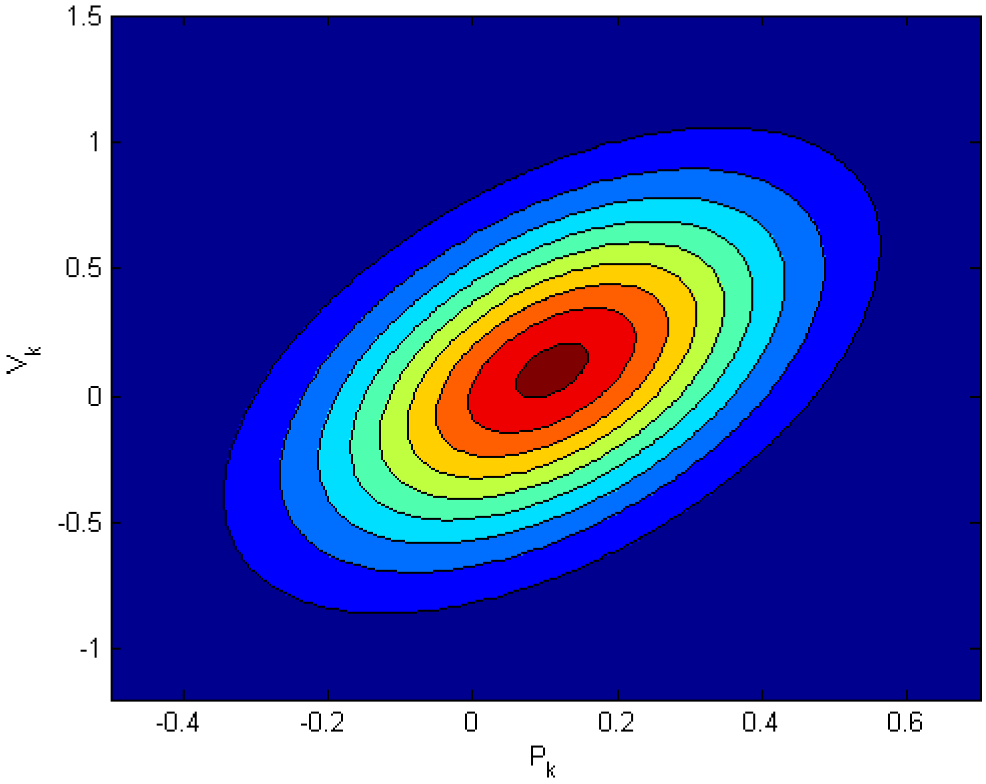
# $\ln(P_k)$ - $\ln(V_k)$, k=251



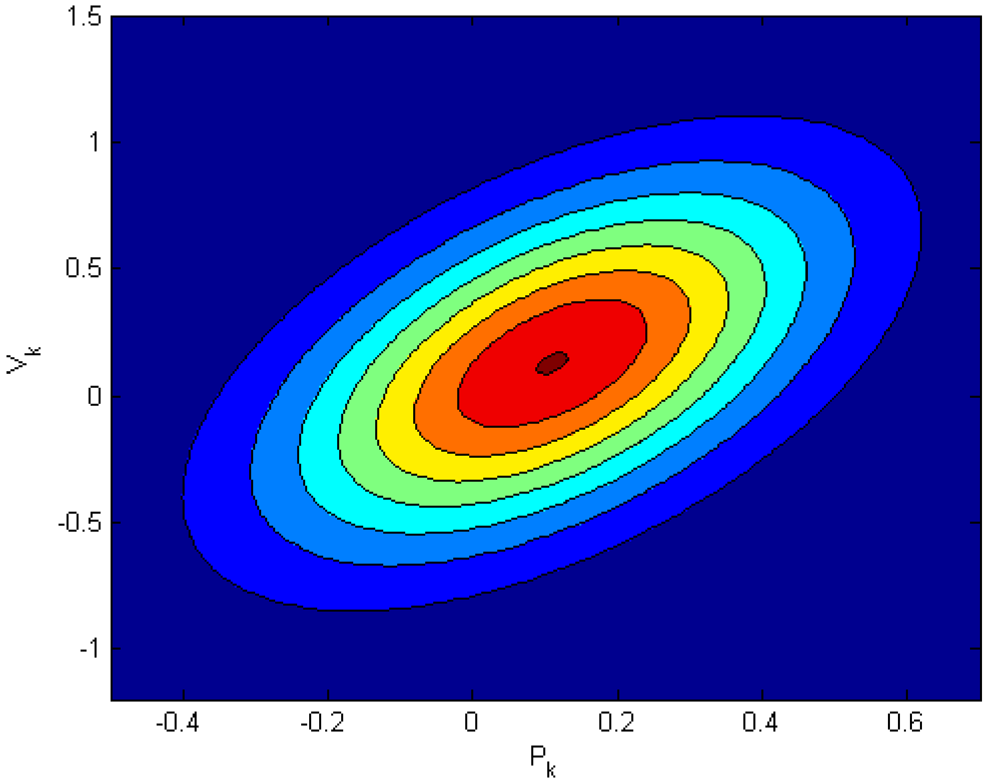
# $\ln(P_k) - \ln(V_k)$, k=301



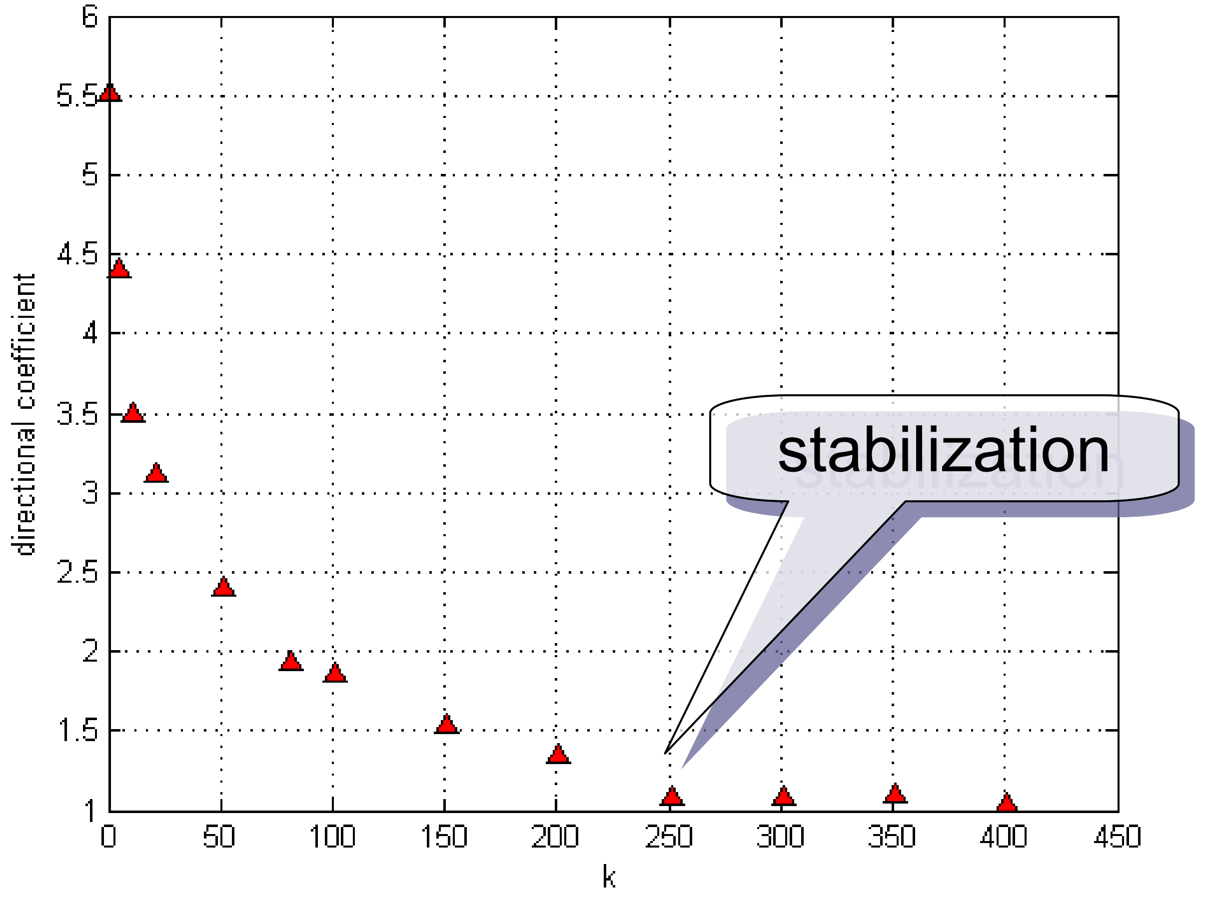
# $\ln(P_k)$ - $\ln(V_k)$, k=351



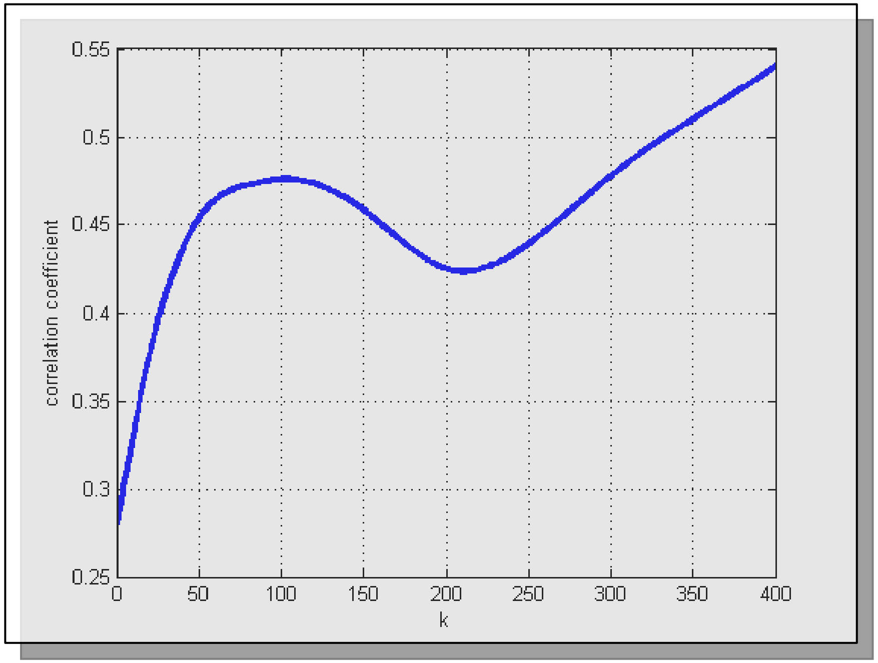
# ln($P_k$) - ln($V_k$), k=401.



# Slope coefficient vs k



# Causality




- Correlations exist
- What's the direction?
  - $\ln(P_k) \to \ln(V_k)$
  - $\ln(V_k) \to \ln(P_k)$
  - $\ln(V_k) \leftrightarrow \ln(P_k)$
- Granger test.

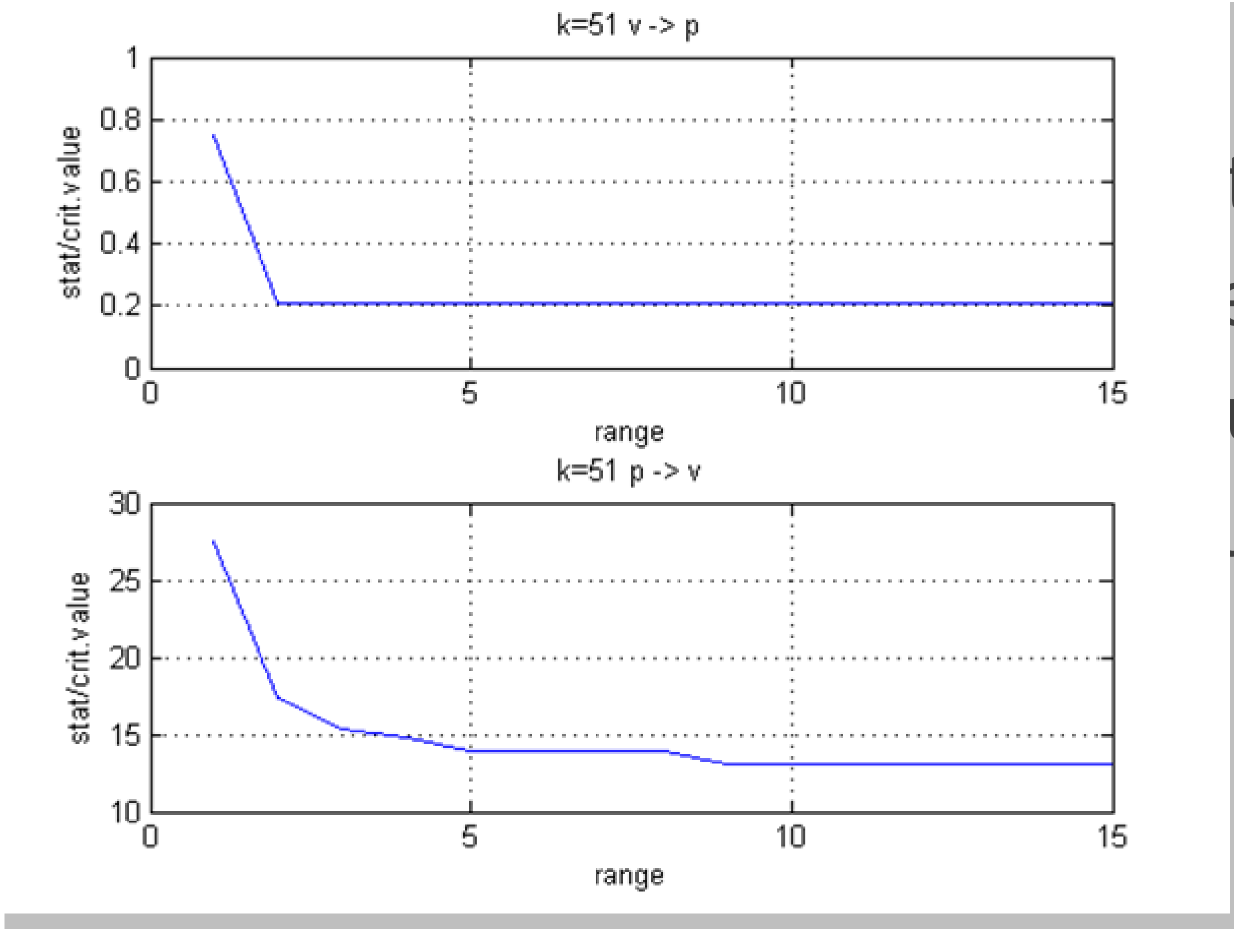
# Granger test

- What is better forecasted based on the past data?
  - prices
  - volumes

$$\ln(V_t) \;,\; \ln(P_t) = B_0 + \sum_{i=1}^{l} \alpha_i \ln(V_{t-i}) + \sum_{i=1}^{l} \beta_i \ln(P_{t-i})$$

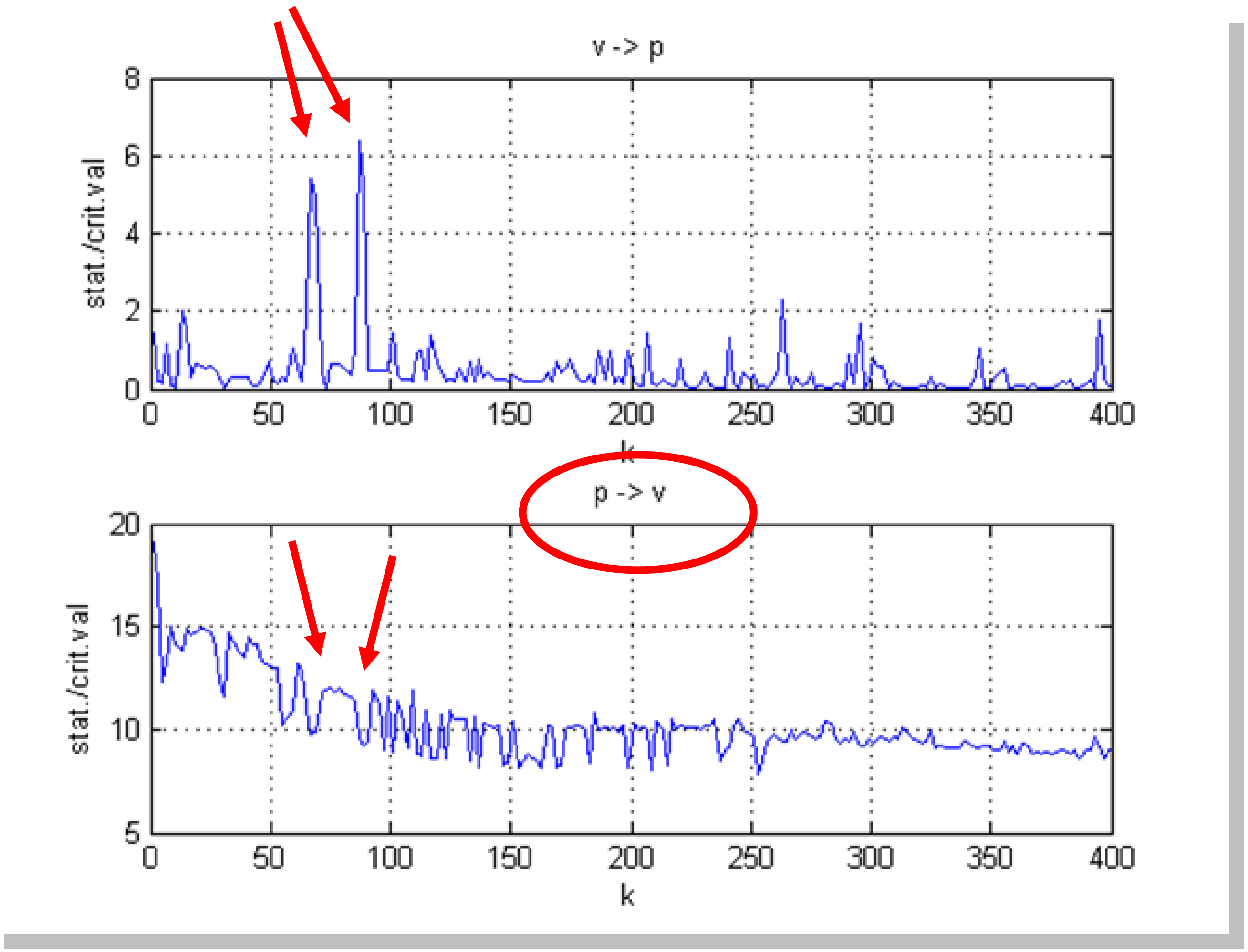
- How far to look backward? What is the range *l*.

# Granger test: range




tion
es
und 10
ure

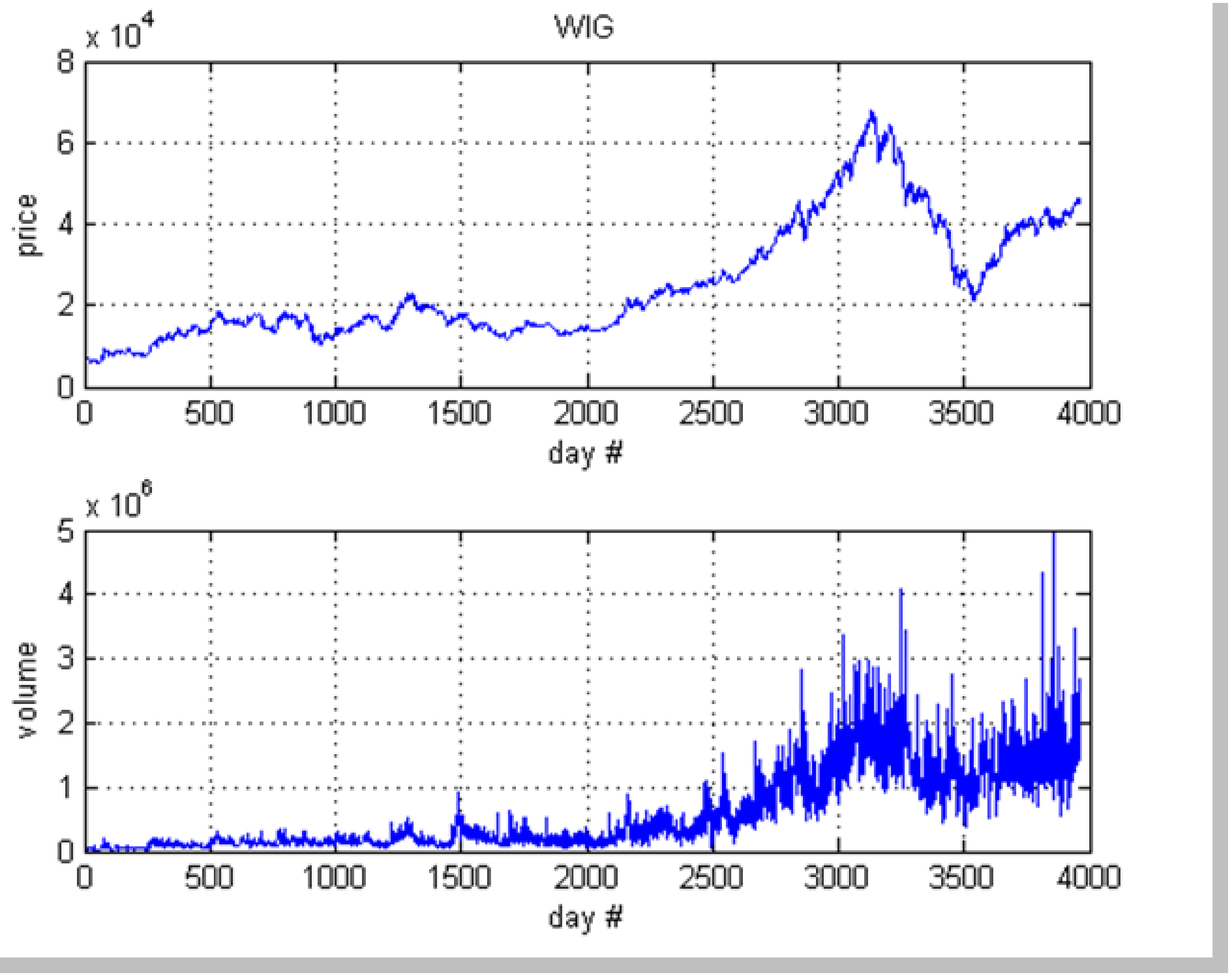
# Granger test



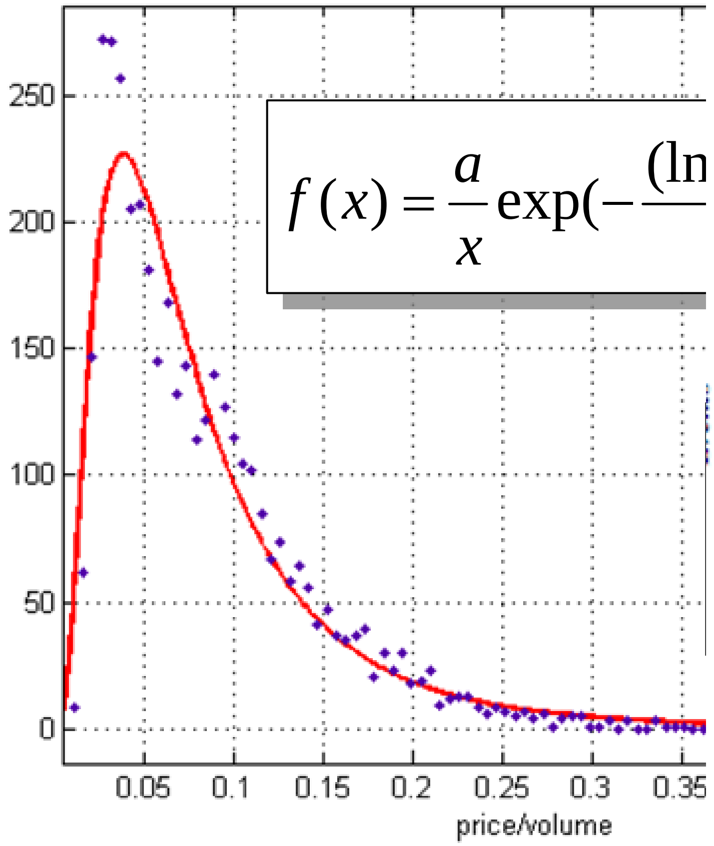
# Granger

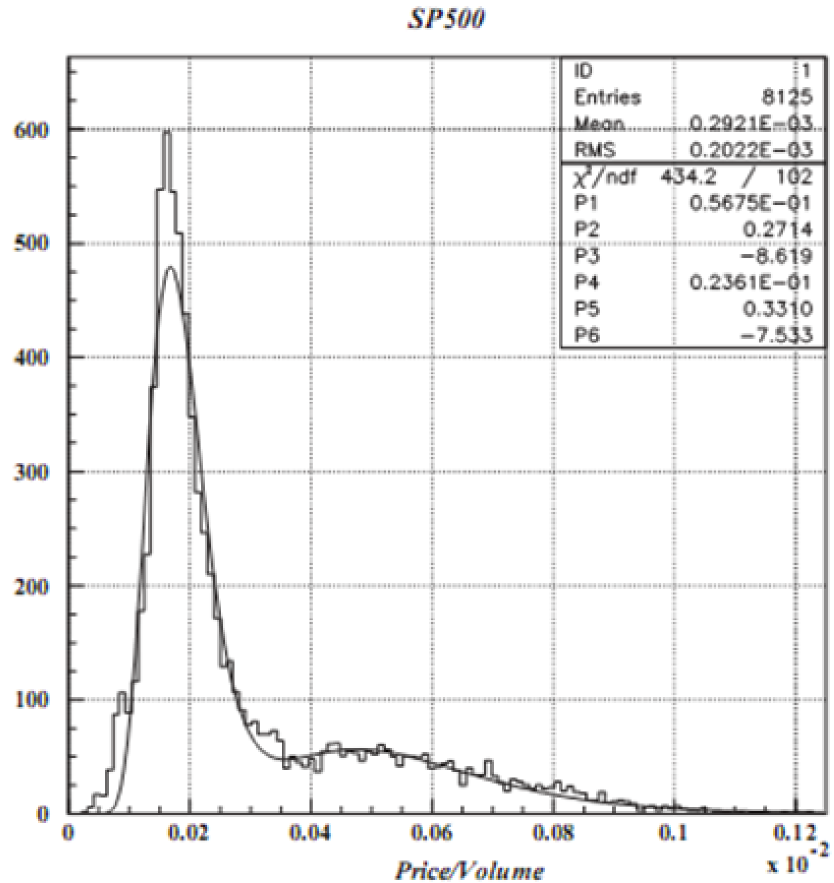
- Prices -> Volumes
- Result independent on k
- Two directional relation for some k?
- Other studies
  - Latin America emerging: vol -> prices
  - Mexico, US: prices -> vol
- Some evidence on Polish market maturity?

# Overall glance again



# Distribution of price/volume

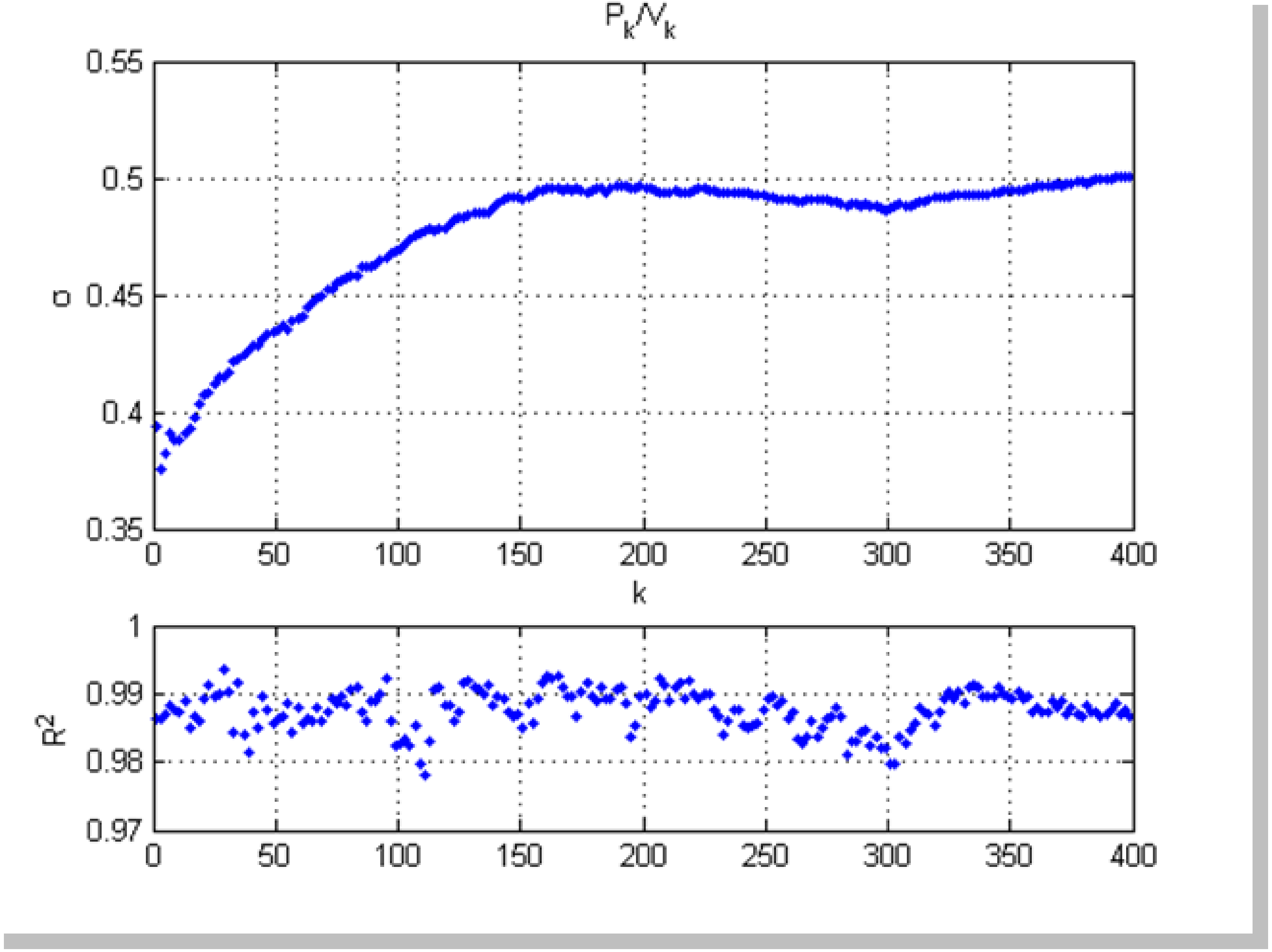


# Companies (σ)

| U.S. (NYSE) | Poland (GPW) |
|---|---|
| Solvay: 0.943 | Pekao: 1.195 |
| DaimlerChysler: 0.900 | Budimex: 3.076 |
| Microsoft: 0.673 | Rafako: 1.566 |

# $P_k/V_k$ vs k



P_k/V_k
0.55
0.5
0.45
0.4
0.35
σ
0
50
100
150
200
250
300
350
400
k
1
0.99
0.98
0.97
R²

# Final remarks

- Event prices & volumes signifitantly changed over time: $P_{kt}$ & $V_{kt}$ are stationary
- $P_{kt}$ & $V_{kt}$ are log-normal distributed
- $P_{kt}$ - small positive autocorrelations
- $V_{kt}$ - big negative autocorrelations
- Moderate correlations $P_{kt}$ - $V_{kt}$
  - local maximum (100) & minimum (210)
  - 2dim plots tend to rotate: some stabilization at 250.

# Final remarks (cont.)

- Strong Granger causality P -> V
  - similarity to mature markets
  - opposite behavious than for emerging ones
  - … at least in Latin America
- Simply P/V values conform log-normal distribution
- Width of $\ln(P_k/V_k)$ exchibites local maximum & minimum.

# Thank You!