Neglecting Parameter Changes in Autoregressive Models

Eric Hillebrand
Department of Economics
Louisiana State University

Financial Mathematics Seminar
North Carolina State University

October 1, 2004

Autoregressive Structures in Econometrics

- ARMA:

$$y_t = c + \sum_{j=1}^{p} \phi_j y_{t-j} + \sum_{j=1}^{q} \psi_j \varepsilon_{t-j} + \varepsilon_t,$$

ε_t white noise.

- GARCH:

$$r_t := \log S_t - \log S_{t-1} = \mu(t) + \varepsilon_t,$$
$$\varepsilon_t|_{\mathcal{F}_{t-1}} \sim \mathcal{N}(0, h_t)$$
$$h_t = \omega + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}.$$

- VAR(p):

$$y_t = c + \sum_{j=1}^{p} \Phi_j y_{t-j} + \varepsilon_t,$$

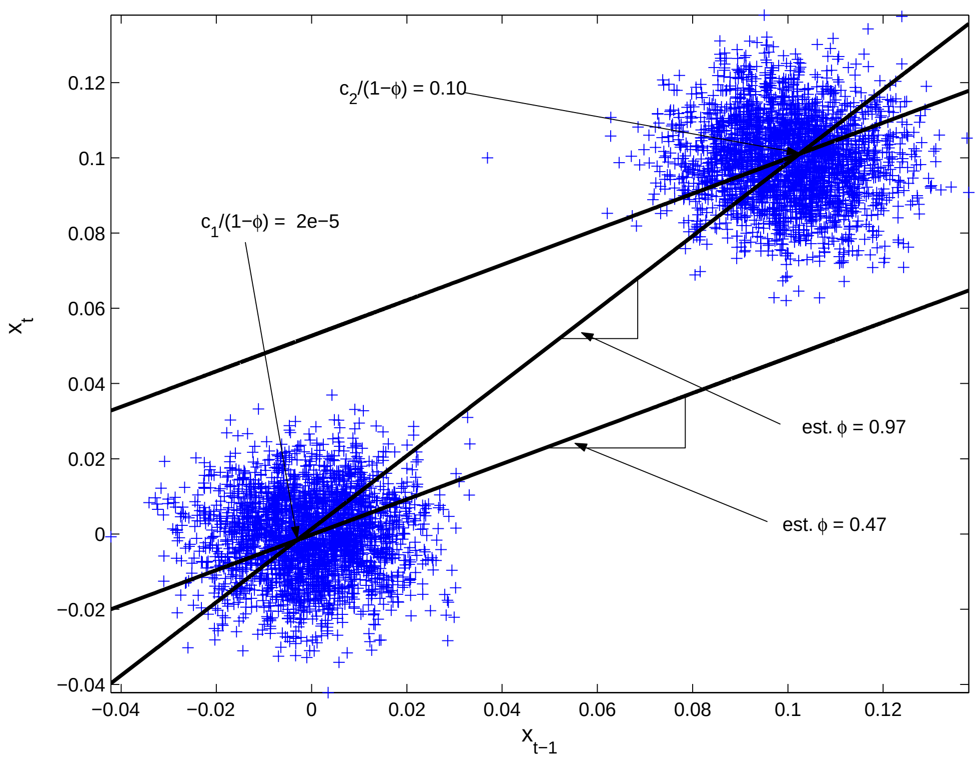
$y_t \in \mathbb{R}^n$, $\Phi_j \in \mathbb{R}^{n \times n}$, ε n-dim white noise.

Example 1: AR(1)

Consider the data-generating process

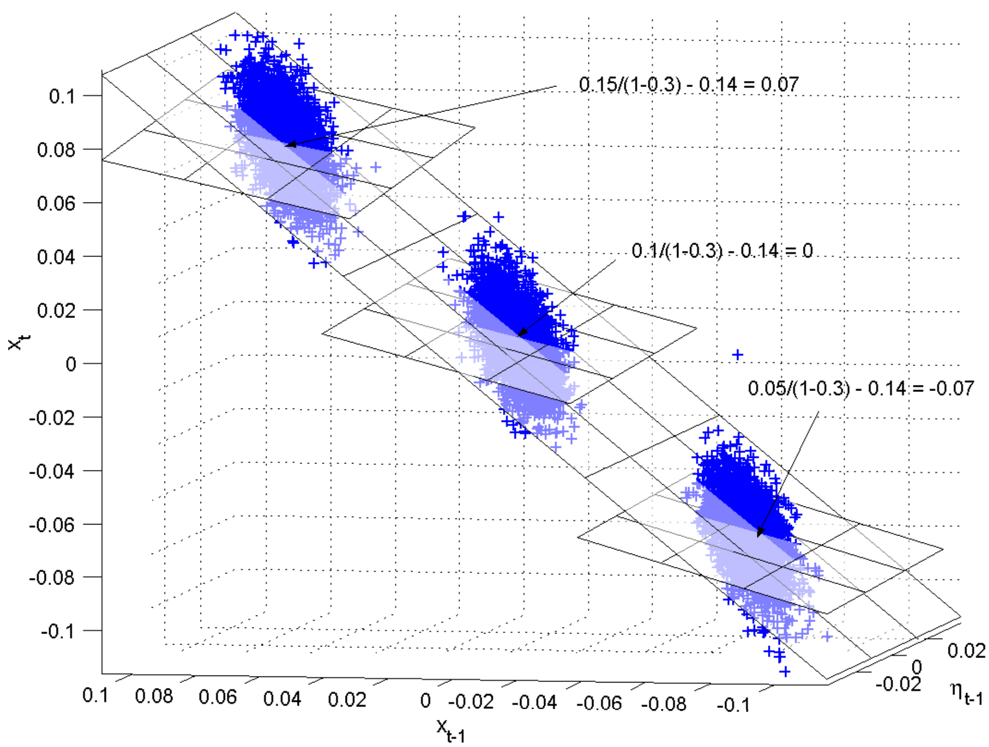
$$x_t = \begin{cases} 1\text{e-}5 + 0.50 x_{t-1} + \varepsilon_t, & t \in \{1, \ldots, N_1\} \\ 5\text{e-}2 + 0.50 x_{t-1} + \varepsilon_t, & t \in \{N_1 + 1, \ldots, N\} \end{cases}$$

ε_t white noise, $N = 5000$, $N_1 = 2500$.



Example 2: ARMA(1,1)

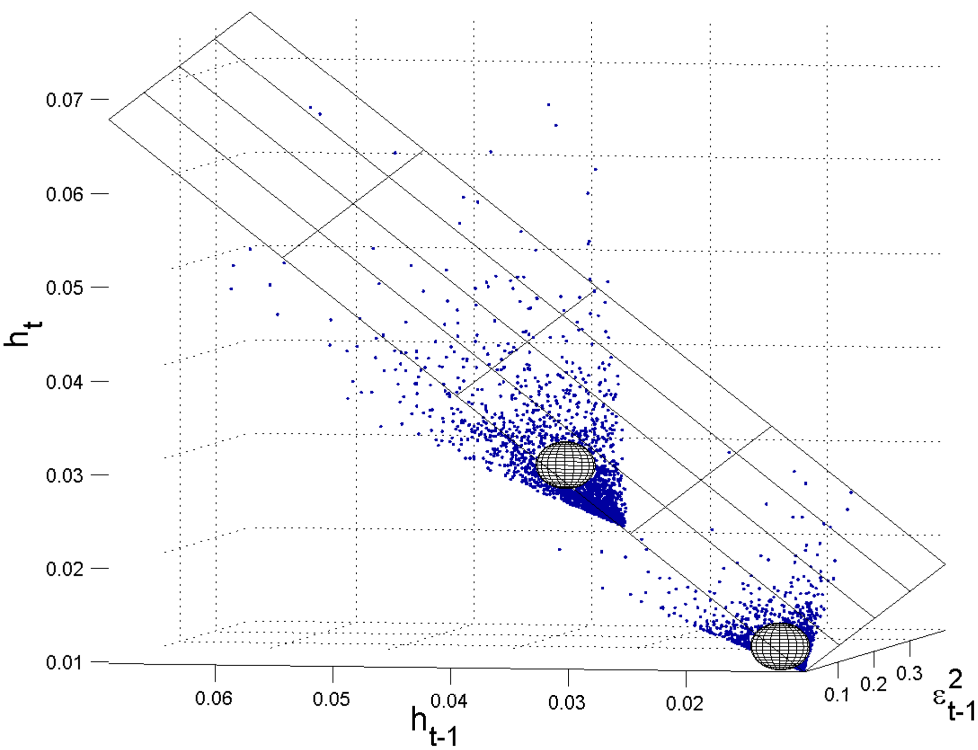
$$x_t - 0.3x_{t-1} = \begin{cases} 0.05 + \epsilon_t + 0.1\epsilon_{t-1} & \text{for } t \in \{1, \ldots, 3000\} \\ 0.10 + \epsilon_t + 0.1\epsilon_{t-1} & \text{for } t \in \{3001, \ldots, 6000\} \\ 0.15 + \epsilon_t + 0.1\epsilon_{t-1} & \text{for } t \in \{6001, \ldots, 9000\}, \end{cases}$$



$$\hat{\phi} \approx 1$$

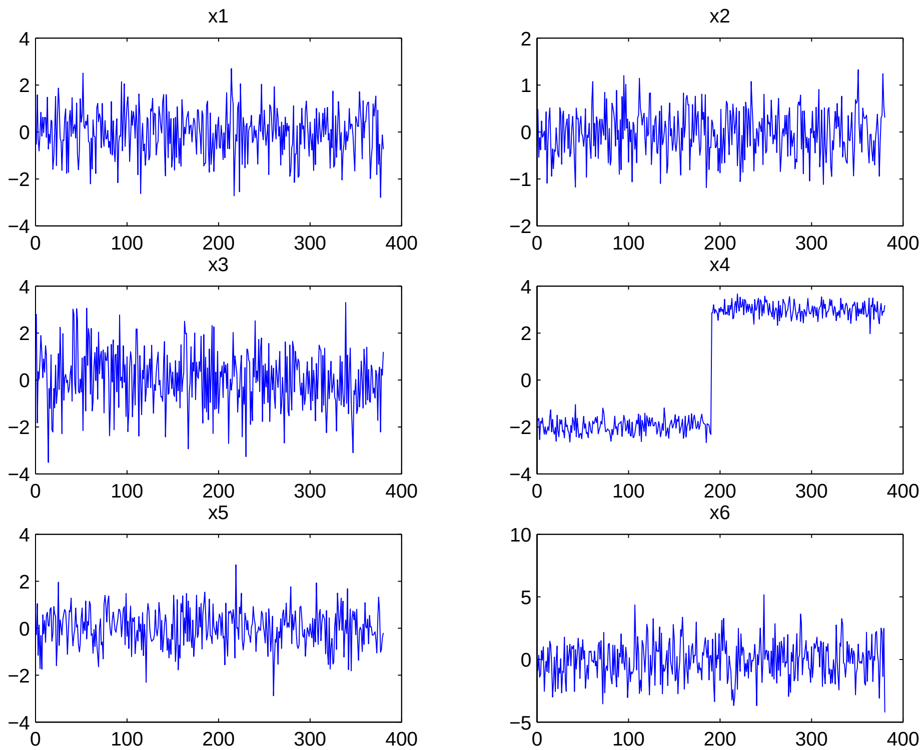
Example 3: GARCH(1,1)

$$h_t = \begin{cases} 5\text{e-}3 + 0.10\varepsilon_{t-1}^2 + 0.50h_{t-1} & \text{for } t \in \{1,\dots,2100\} \\ 0.0125 + 0.10\varepsilon_{t-1}^2 + 0.50h_{t-1} & \text{for } t \in \{2101,\dots,4200\} \end{cases}$$



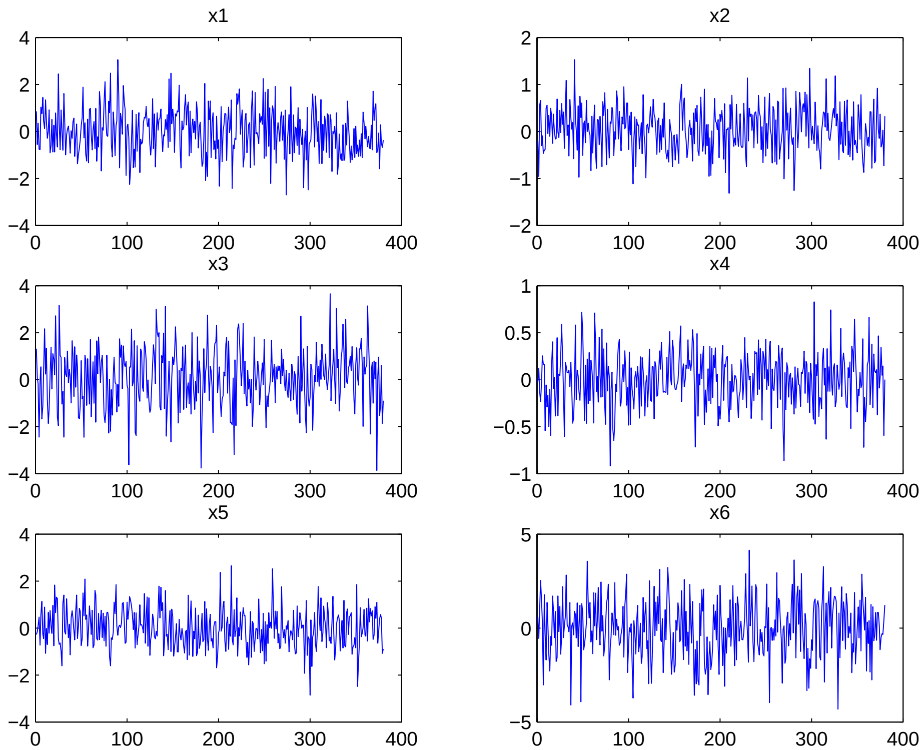
$$\hat{\alpha} + \hat{\beta} \approx 1$$

Example 4: VAR, simulated



$$\max \left\{ \lambda_i \text{ Eigenvalues of } \hat{\Phi} \right\} = 0.98$$

No Jump

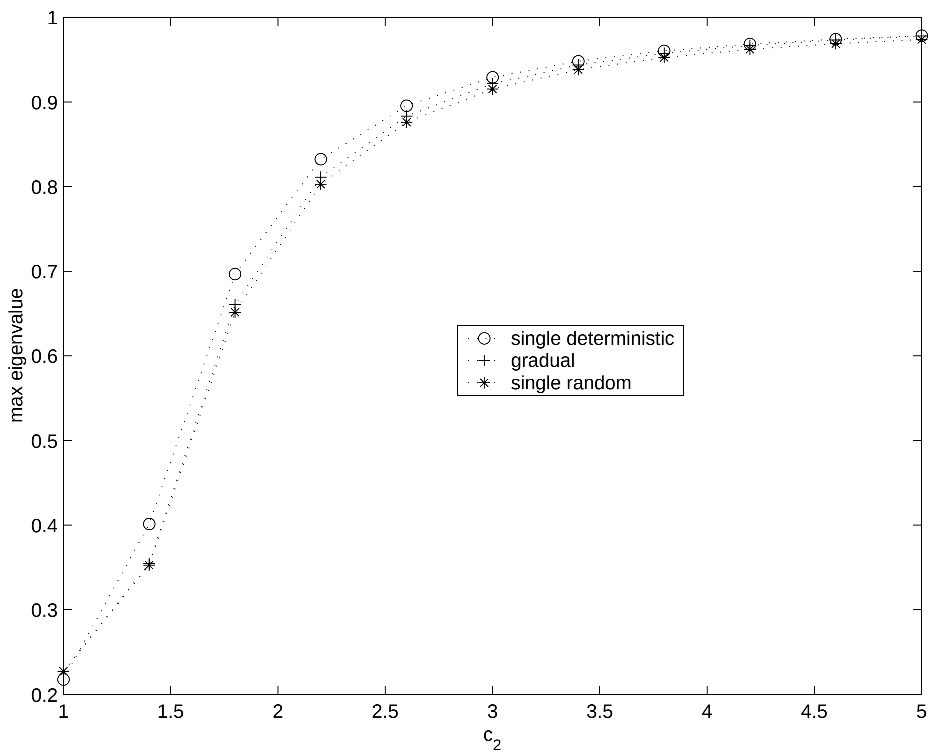


$$\max\left\{\lambda_i \text{ Eigenvalues of } \hat{\Phi}\right\} = 0.14$$

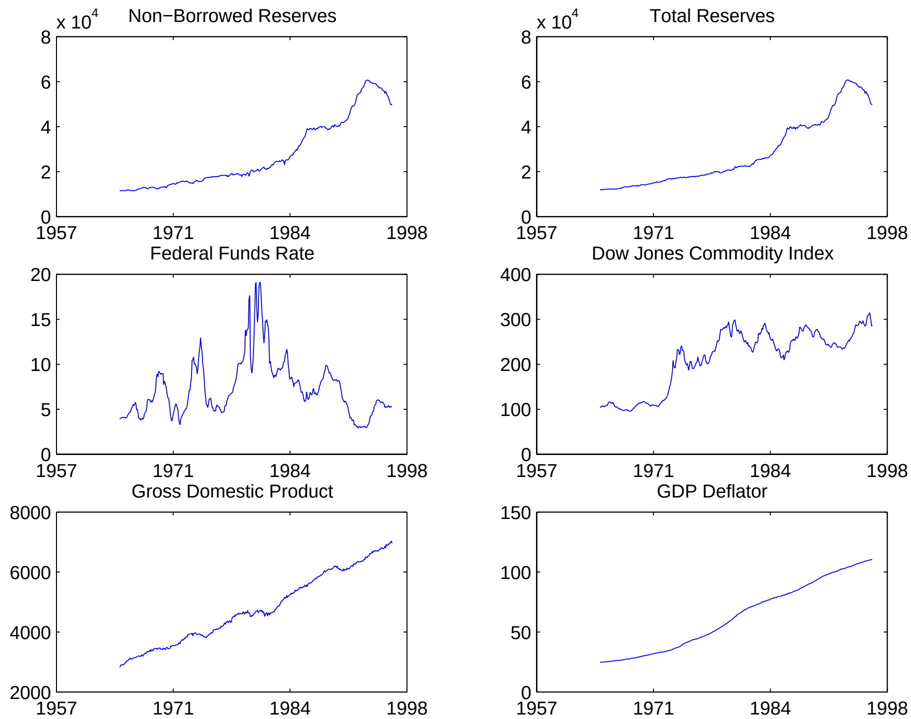
Gradual Changes, Random Changes

$$y_t = c + 0.10y_{t-1} + \varepsilon_t, \; T = 400$$

Jump in c at $T/2$.
$c_1 = 1$ first segment, c_2 second segment



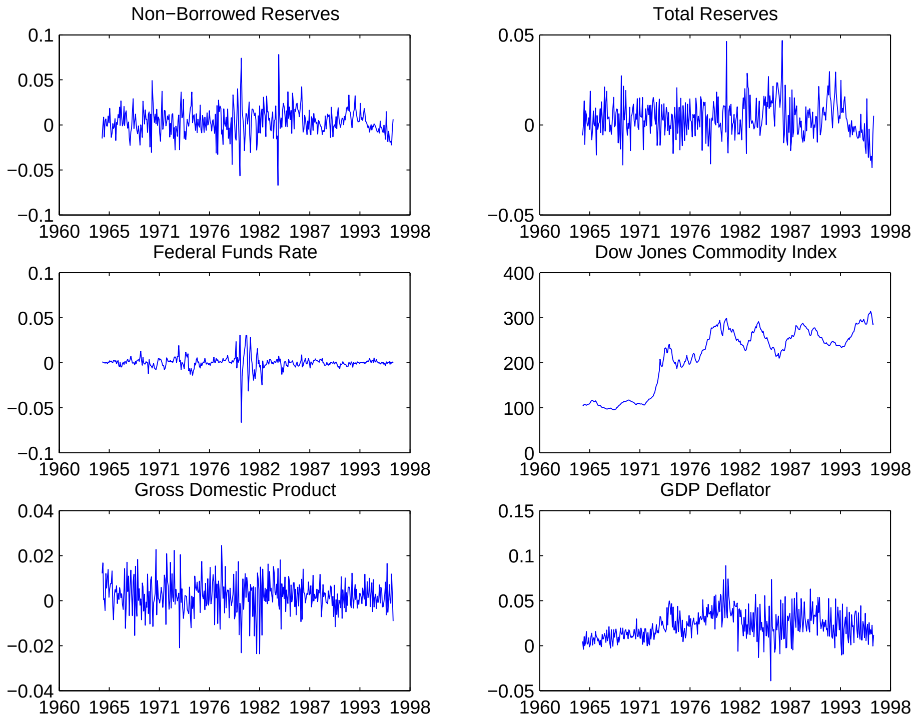
Example 5: VAR, real



$$\max\left\{\lambda_i \text{ Eigenvalues of } \sum \Phi_j\right\} = 1$$

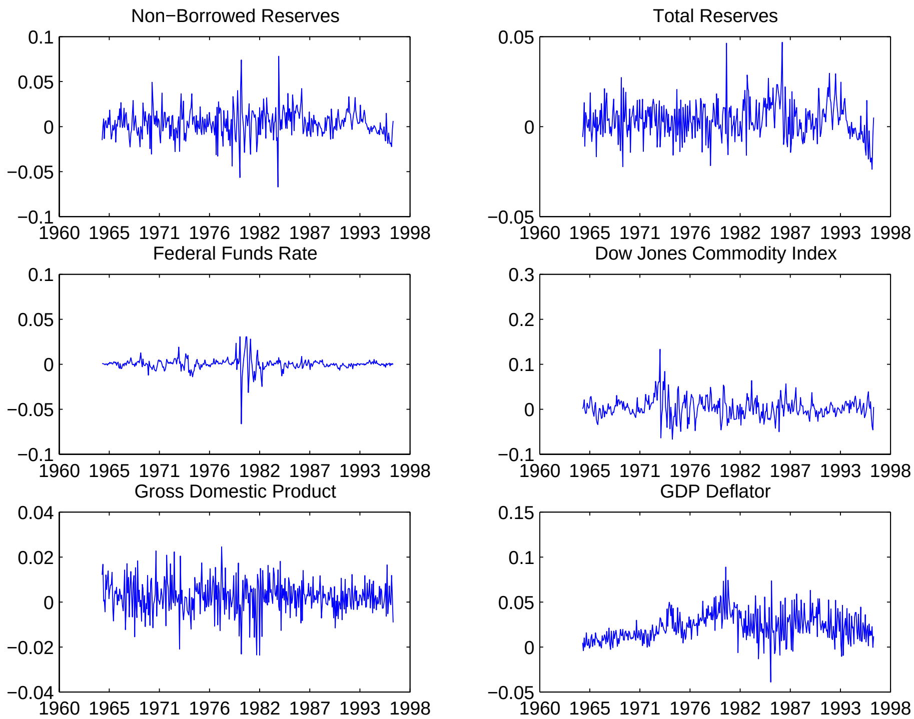
(Bernanke and Mihov: Measuring Monetary Policy. Quarterly Journal of Economics. Aug 1998. 869–902.)

Stationary Series except Commodity Index



$$\max \left\{ \lambda_i \text{ Eigenvalues of } \sum \Phi_j \right\} = 0.995$$

Stationary Series



$$\max\left\{\lambda_i \text{ Eigenvalues of } \sum \Phi_j\right\} = 0.402$$

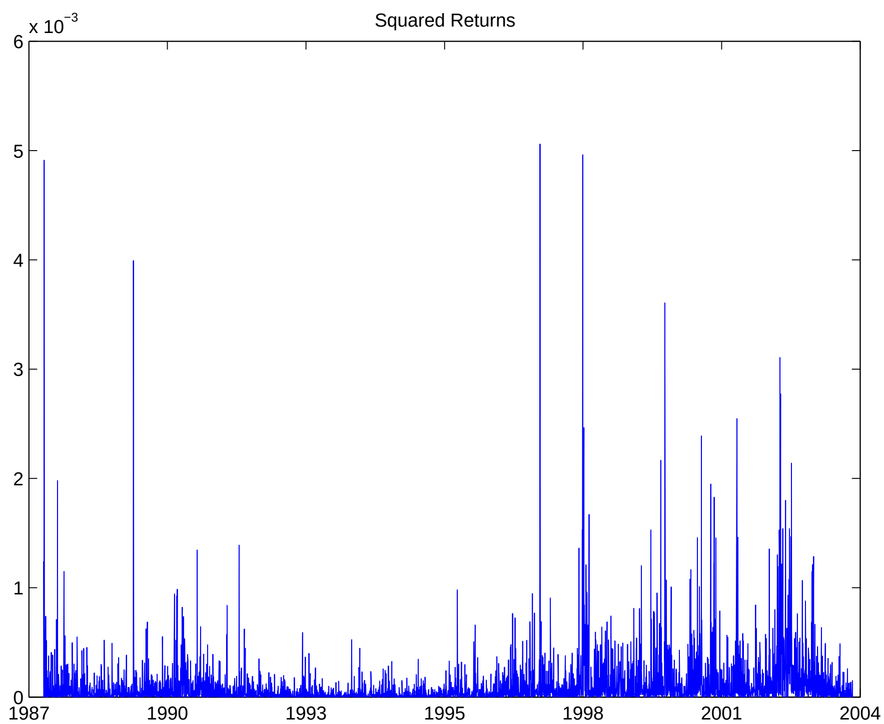
Example 6: S&P 500 Index

4-Jan-1988 through 31-Dec-2003

GARCH(1,1):

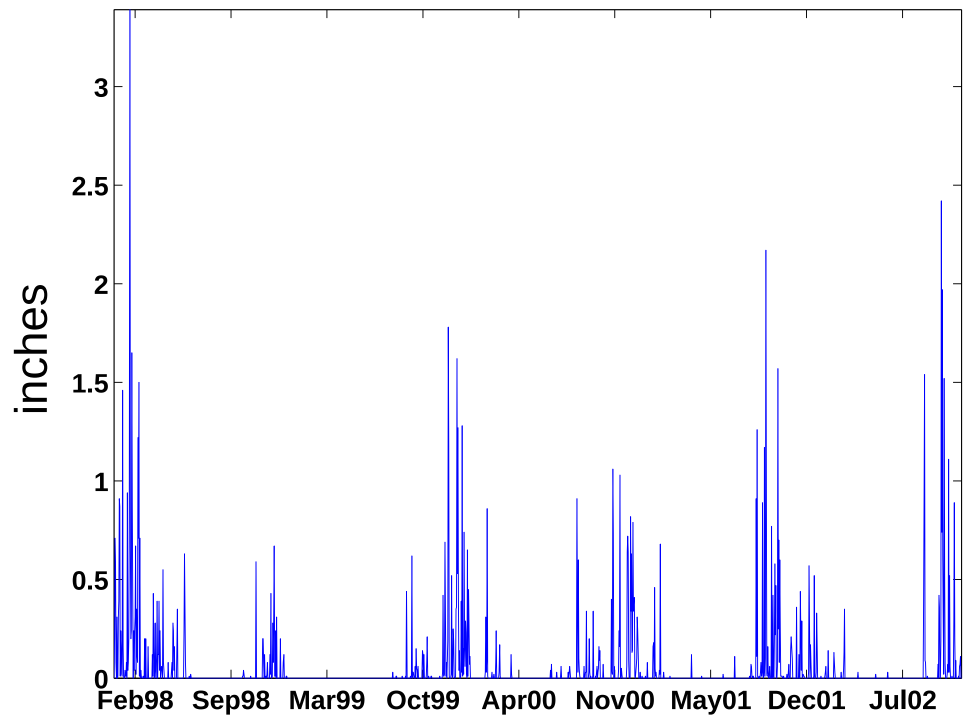
$$h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1},$$

persistence parameter $\lambda := \alpha + \beta$.



$$\lambda = 0.0447 + 0.9506 = 0.9953$$

Example 7: San Francisco Precipitation Jan 1, 1998 through Jan 31, 2003



GARCH(1,1):

$$\lambda = 0.196 + 0.803 = 0.999$$

Earlier Studies

AR Models, Dickey-Fuller Tests:

- Perron (1989)
- Perron (1990)
- Hendry and Neale (1991)

GARCH:

- Diebold (1986)
- Lamoureux and Lastrapes (1990)
- Hamilton and Susmel (1994)
- Mikosch and Starica (2000)

Long Memory:

- Granger and Hyung (1999)
- Granger and Teräsvirta (2001)
- Diebold and Inoue (2001)

VAR(1)

AR(1):

$$y_t = c + \phi y_{t-1} + \varepsilon_t;$$

$$y_t = c\frac{1-\phi^t}{1-\phi} + \sum_{j=0}^{t-1} \phi^j \varepsilon_{t-j} + \phi^t y_0.$$

VAR(1):

$$x_t = c + \Phi x_{t-1} + \varepsilon_t;$$

$$x_t = (I-\Phi)^{-1}(I-\Phi^t)c + \sum_{j=0}^{t-1} \Phi^j \varepsilon_{t-j} + \Phi^t x_0$$

k segments:

$$x_t^{(i)} = c_i + \Phi_i x_{t-1}^{(i)} + \varepsilon_t,$$

$$t = T_{i-1}{+}1, \ldots, T_i;\ i = 1, \ldots, k;\ T_0 = 0,\ T_k = T$$

$$\mu_i := (I-\Phi_i)^{-1} c_i.$$

$$\mathbb{E}_i(\cdot) = \mathbb{E}(\cdot | x_{T_{i-1}})$$

$$\begin{aligned} \mathbb{E}_i x_t &= (I - \Phi_i)^{-1}(I - \Phi_i^{t-T_{i-1}})c_i \\ &\quad + \mathbb{E}_i[\sum_{j=0}^{t-T_{i-1}-1} \Phi_i^j \varepsilon_{t-j}] + \Phi_i^{t-T_{i-1}} x_{T_{i-1}} \\ &= \mu_i + o(1), \end{aligned}$$

Auxiliary Result:

$$\begin{aligned} \bar{x} = \frac{T_1 - T_0}{T}\mu_1 + \frac{T_2 - T_1}{T}\mu_2 + \ldots + \frac{T_k - T_{k-1}}{T}\mu_k \\ + o(1). \end{aligned}$$

Estimated Model:

$$x_t = \hat{c} + \hat{\Phi} x_{t-1} + \varepsilon_t$$

Subtract sample mean $\bar{x} = \hat{c} + \hat{\Phi}\bar{x}$, take $\mathbb{E}_i$:

$$\mathbb{E}_i(x_t - \bar{x}) = \mathbb{E}_i(\hat{\Phi}(x_{t-1} - \bar{x})).$$

Assumption:

$$\mathrm{Var}\hat{\Phi} = o(1)$$

Then,

$$\mathbb{E}_i(x_t - \bar{x}) = \mathbb{E}_i\hat{\Phi}\,\mathbb{E}_i(x_{t-1} - \bar{x}) + o(1).$$

Use the result $\mathbb{E}_i x_t = \mu_i + o(1)$ and Auxiliary Result:

$$\left[\mu_i - \frac{1}{T}\sum_{j=1}^{k}(T_j - T_{j-1})\mu_j\right]$$
$$= \mathbb{E}_i\hat{\Phi}\left[\mu_i - \frac{1}{T}\sum_{j=1}^{k}(T_j - T_{j-1})\mu_j\right] + o(1)$$

Theorem 1. *If a first-order vector autoregressive model is estimated on a time series that underwent* $(k-1)$ *parameter changes, the matrix* $\mathbb{E}_i(\hat{\Phi}) = \mathbb{E}(\hat{\Phi}|x_{T_{i-1}})$ *has a unit eigenvalue with corresponding eigenvector*

$$e_i := \mu_i - 1/T\sum_{j=1}^{k}(T_j - T_{j-1})\mu_j.$$

This holds true in all segments $i = 1, 2, \ldots, k$, *up to terms that vanish with growing segment sizes* $T_i - T_{i-1}$.

VAR(p)

$$\tilde{x}_t = \tilde{c} + \Phi_1 \tilde{x}_{t-1} + \Phi_2 \tilde{x}_{t-2} + \ldots + \Phi_p \tilde{x}_{t-p} + \tilde{\varepsilon}_t,$$

VAR(1) representation:

$$x_t = c + M x_{t-1} + \varepsilon_t,$$

$$x_t = \begin{bmatrix} \tilde{x}_t \\ \tilde{x}_{t-1} \\ \vdots \\ \tilde{x}_{t-p+1} \end{bmatrix} \; c = \begin{bmatrix} \tilde{c} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \; M = \begin{bmatrix} \Phi_1 & \Phi_2 & \ldots & \Phi_{p-1} & \Phi_p \\ I & 0 & \ldots & 0 & 0 \\ 0 & I & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & I & 0 \end{bmatrix}$$

$$\varepsilon_t = \begin{bmatrix} \tilde{\varepsilon}_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Theorem 1 applies to

$$\mathbb{E}_i(\hat{\Phi}_1 + \hat{\Phi}_2 + \ldots + \hat{\Phi}_p).$$

ARMA(p,q)

$$\tilde{x}_t = \tilde{c} + \phi_1 \tilde{x}_{t-1} + \ldots + \phi_p \tilde{x}_{t-p} + \epsilon_t + \psi_1 \epsilon_{t-1} + \ldots + \psi_q \epsilon_{t-q}$$

$$m = \max\{p, q\}$$
$$c = (\tilde{c}, 0, \ldots, 0)' \in \mathbb{R}^m$$
$$x_t = (\tilde{x}_t, \tilde{x}_{t-1}, \ldots, \tilde{x}_{t-m+1})'$$
$$\varepsilon_t = (\epsilon_t, \epsilon_{t-1}, \ldots, \epsilon_{t-m})' \in \mathbb{R}^{m+1}$$

$$\Phi = \begin{bmatrix} \phi_1 & \phi_2 & \ldots & \phi_{p-1} & \phi_p & [0] \\ 1 & 0 & \ldots & 0 & 0 & [0] \\ 0 & 1 & \ldots & 0 & 0 & [0] \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 & [0] \\ [0] & [0] & \ldots & [0] & [1] & [0] \end{bmatrix} \in \mathbb{R}^{m \times m}.$$

$$\Psi = \begin{bmatrix} 1 & \psi_1 & \ldots & \psi_q & [0] \\ 0 & 0 & \ldots & 0 & [0] \\ 0 & 0 & \ldots & 0 & [0] \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & [0] \\ [0] & [0] & \ldots & [0] & [0] \end{bmatrix} \in \mathbb{R}^{m+1 \times m+1},$$

Vector form:

$$x_t = c + \Phi x_{t-1} + \Psi \varepsilon_t,$$

Theorem 1 applies to

$$\mathbb{E}_i \hat{\Phi},$$

or

$$\mathbb{E}_i \sum_{j=1}^{p} \hat{\phi}_j = 1 + o(1).$$

GARCH(p,q)

$$r_t = \mathbb{E}(r_t|\mathcal{F}_{t-1}) + \epsilon_t = f(b) + \epsilon_t \; t = 1, \ldots, T,$$
$$\epsilon_t|\mathcal{F}_{t-1} \sim \mathcal{N}(0, \mathtt{h}_t),$$
$$\mathtt{h}_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \mathtt{h}_{t-i},$$

$$\epsilon_t^2 = \mathtt{h}_t \, \eta_t^2, \;\; \eta_t \sim \mathcal{N}(0,1)$$

$$m = \max\{p, q\}$$
$$h_t = (\mathtt{h}_t, \, \mathtt{h}_{t-1}, \, \ldots, \, \mathtt{h}_{t-m+1})' \in \mathbb{R}^m$$
$$\varepsilon_t^2 = (\epsilon_t^2, \, \epsilon_{t-1}^2, \ldots, \, \epsilon_{t-m+1}^2)' \in \mathbb{R}^m$$
$$\omega = (\alpha_0, \, 0, \ldots, \, 0)' \in \mathbb{R}^m$$

$$A = \begin{bmatrix} \alpha_1 \eta_{t-1}^2 & \alpha_2 \eta_{t-2}^2 & \ldots & \alpha_q \eta_{t-q}^2 & [0] \\ 0 & 0 & \ldots & 0 & [0] \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & [0] \\ [0] & [0] & \ldots & [0] & [0] \end{bmatrix} \in \mathbb{R}^{m \times m},$$

$$B = \begin{bmatrix} \beta_1 & \beta_2 & \dots & \beta_{p-1} & \beta_p & [0] \\ 1 & 0 & \dots & 0 & 0 & [0] \\ 0 & 1 & \dots & 0 & 0 & [0] \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & [0] \\ [0] & [0] & \dots & [0] & [1] & [0] \end{bmatrix} \in \mathbb{R}^{m\times m}.$$

Vector form:

$$h_t = \omega + (A+B)h_{t-1},$$

Theorem 1 applies to

$$\mathbb{E}_i(A+B),$$

or

$$\mathbb{E}_i\left(\sum_{j=1}^{q}\alpha_j + \sum_{j=1}^{p}\beta_j\right) = 1 + o(1).$$

Conclusions

- Looked at neglected parameter switches that change the in-segment means of autoregressive processes.
- In VAR:
$$\mathbb{E}_i \sum_{j=1}^{p} \hat{\Phi}_j$$
has asymptotic unit eigenvalue.
- In ARMA:
$$\mathbb{E}_i \sum_{j=1}^{p} \hat{\phi}_j = 1 + o(1).$$
- In GARCH:
$$\mathbb{E}_i \left(\sum_{j=1}^{q} \alpha_j + \sum_{j=1}^{p} \beta_j \right) = 1 + o(1).$$
- Geometric phenomenon: not confined to specific estimators. ($\mathrm{Var}\hat{\Phi} = o(1)$)
- Before specifying sample: Change-point detection.