Long Memory through Marginalization

Hidden & Ignored Cross-Section Dependence

Guillaume Chevillon
ESSEC Business School (CREAR) and CREST

joint with

Alain Hecq — Maastricht University
and
Sébastien Laurent — Université Aix-Marseille & GREQAM

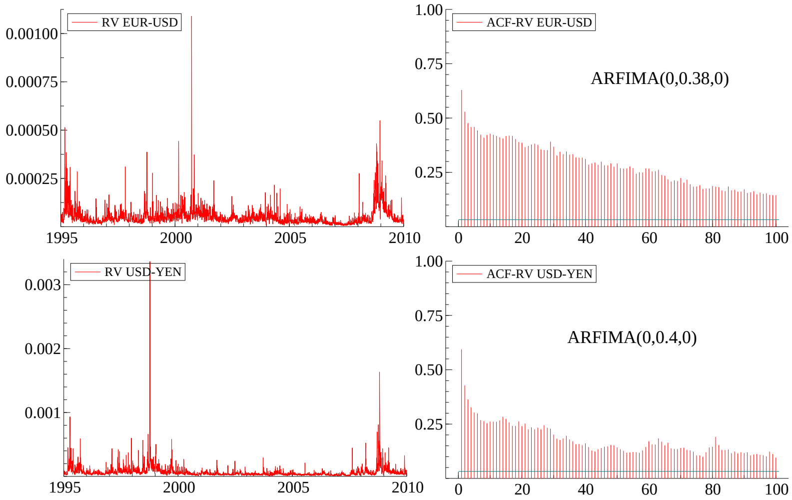
15th Oxmetrics User Conference,
Cass Business School, London
September 5, 2014

"There is an emerging consensus in empirical finance that realized volatility series typically display long range dependence with a memory parameter (d) around 0.4 (Andersen et al., JASA, 2001)."

Lieberman and Phillips (2008, ER).

Some Empirical Evidence

Exchange rates



Daily Realized Volatility, Autocorrelations and fitted ARFIMA$(0, d, 0)$

Overview

- Empirical evidence:
  - volatilities are persistent, exhibiting long memory, ACF decay slowly (hyperbolically rather than exponentially)
  - all seem to present the same degree of memory $d = 0.4$ when analyzed individually (Andersen et al's magic number).
- This paper shows: **large n multivariate** (Big Data) framework may generate (identical) long memory in all the univariate (**marginalized**) representations.
  - **beware** small (hidden/ignored) cross-section dependences (Manski, Pesaran, Hendry & Pretis...)
- **Tool:** From cross-section to time series dependence in a Final Equation Representation in VAR(1) context
- Potential applications to realized volatility, prices, river volume.... (Cox and Townsend, 1947)

Outline

- Preliminary: Long Memory and Final Equation Representation
- Our argument
- Monte Carlo and stylized empirical facts

Long Memory

- Many commonly used definitions of short memory or weak dependence imply the property that
$$\text{var}\left(T^{-1/2}S_T\right) \to c \in (0, \infty)$$
where $S_T = \sum_{t=1}^{T} z_t$.
- Any process that does not satisfy this property may be said to exhibit long memory (Diebold & Inoue, 2001).
- For a covariance stationary process, Beran (1994) shows the equivalence between
$$\begin{cases} \text{var}\left(T^{-1/2}S_T\right) \sim c_v T^{2d}, & \text{as } T \to \infty, \\ \rho_z(k) \sim c_\rho k^{2d-1}, & \text{as } k \to \infty, \\ f_z(\omega) \sim c_f |\omega|^{-2d}, & \text{as } \omega \to 0, \end{cases} \tag{1}$$
where $d \in (-1/2, 1/2)$, $d \neq 0$, $c_v, c_\rho, c_f > 0$, $\rho_z(k) = \text{Corr}[z_t, z_{t+k}]$ and $f_z(\omega)$ is spectral density.
- Fractionally integrated processes exhibit unit roots.

Known Sources of Long Memory

- Long memory is commonly observed in economics and finance but its origin is unclear.
- Müller & Watson (2008) show the difficulty in discriminating between the various models of low frequency variation.
- Known explanations:
  - aggregation across heterogeneous series, frequencies or economic agents (Granger 1980, Lieberman & Phillips, 2008, Abadir & Talmain, 2002);
  - nonlinearity (Davidson & Sibbertsen 2005, Miller & Park, 2010)
  - structural change (Diebold & Inoue, 2001, Perron and Qu, 2006).
  - learning (bounded rationality) in forward looking models of expectations (Chevillon & Mavroeidis, 2013)
  - Network effects (Schennach, 2013).
  - **our contribution:** memory as an artefact of univariate modelling and marginalizing small & numerous cross section correlations

Final Equation Representation

- Consider an n-vector $\mathbf{x}_t$ admitting an invertible VAR(1) representation

$$(\mathbf{I}_n - \mathbf{A}_n L)\,\mathbf{x}_t = \epsilon_t \Leftrightarrow \mathbf{x}_t = (\mathbf{I}_n - \mathbf{A}_n L)^{-1}\,\epsilon_t$$

- Where the inverse satisfies

$$(\mathbf{I}_n - \mathbf{A}_n L)^{-1} = \frac{1}{\det(\mathbf{I}_n - \mathbf{A}_n L)} (\widetilde{\mathbf{I}_n - \mathbf{A}_n L})$$

  with $(\widetilde{\mathbf{I}_n - \mathbf{A}_n L})$ the adjugate matrix.
- yielding a priori an ARMA$(n, n-1)$

$$\det(\mathbf{I}_n - \mathbf{A}_n L)\,\mathbf{x}_t = (\widetilde{\mathbf{I}_n - \mathbf{A}_n L})\epsilon_t$$

  with **common** AR **polynomial for all variables.**
- The spectral density of element $\{x_{1t}\}$ is

$$f(\omega) = \sum_{j=1}^{n} \left| \frac{\left(\widetilde{\mathbf{I}_n - \mathbf{A}_n e^{-i\omega}}\right)_{1j}}{\det\left(\mathbf{I}_n - \mathbf{A}_n e^{-i\omega}\right)} \right|^2$$

Example of FER

Consider the trivariate VAR(1)

$$\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = \begin{bmatrix} a & b & 0 \\ b & a & b \\ 0 & b & a \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^x \\ \varepsilon_t^y \\ \varepsilon_t^z \end{bmatrix}$$

with FER $A(L)\mathbf{x}_t = \mathbf{B}(L)\epsilon_t$ with

$$A(L) = (1-aL)\left[1-\left(aL+\sqrt{2}b\right)L\right]\left[1-\left(aL-\sqrt{2}b\right)L\right]$$

$$\mathbf{B}(L) = \begin{bmatrix} (1-aL)^2 & bL(1-aL) & b^2L^2 \\ bL(1-aL) & (1-aL)^2 & -(1-aL)^2 \\ b^2L^2 & bL(1-aL) & (1-aL)^2-b^2L^2 \end{bmatrix}$$

hence all univariate elements follow and ARMA$(3,2)$ which simplifies e.g. if $b=0$. In the context of BEKK, see Hecq, Laurent & Palm (2012).

Outline

- Preliminary
- Our argument
- Monte Carlo and stylized empirical facts

Analytic Framework

Assume $\mathbf{A}_n$ circulant matrix

$$\mathbf{A}_n = \begin{bmatrix} a_0^{(n)} & a_1^{(n)} & \cdots & a_{n-1}^{(n)} \\ a_{n-1}^{(n)} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_1^{(n)} \\ a_1^{(n)} & \cdots & a_{n-1}^{(n)} & a_0^{(n)} \end{bmatrix}$$

Define the spectral density of $\mathbf{A} = \lim_{n\to\infty} \mathbf{A}_n$, $g_{\mathbf{A}}$, with Fourier coefficients $a_k = \lim_{n\to\infty} a_k^{(n)}$:

$$g_{\mathbf{A}}(\lambda) \stackrel{def}{=} \lim_{n\to\infty} \left(a_0^{(n)} + 2 \sum_{k=1}^{n-1} a_k^{(n)} e^{k\lambda} \right); \quad a_k = \frac{1}{2\pi} \int_0^{2\pi} g_{\mathbf{A}}(\lambda) e^{-ik\lambda} d\lambda$$

Then the eigenvalues of $\mathbf{A}_n$ are given by

$$\lambda_k = g_{\mathbf{A}} \left(\frac{2\pi k}{n} \right), \quad 0 \le k < n$$

Our result: choose a parametric function g()

- Consider (a 2π-periodic version of) the low pass

$$g_{\mathbf{A}}^{(d)}(x) = 1_{\{|x|<2\pi d\}}, \quad d \in (0, 1/2),$$

the eigenvalues of $\mathbf{C}_n$ satisfy:

$$\lambda_k = g_{\mathbf{A}}^{(d)}\left(\frac{2\pi k}{n}\right) = \begin{cases} 1 & k = 0, ..., \lfloor nd \rfloor; \\ 0 & k = \lfloor nd \rfloor + 1, ..., n-1. \end{cases}$$

$\lfloor nd \rfloor$ stochastic trends, $n - \lfloor nd \rfloor$ cointegration relations.

Theorem

Any element x_t of the process $\mathbf{x}_t$, admits a spectral density $f_{n,d}$ such that as $(d, n) \to (1/2, \infty)$, for $\omega > 0$,

$$f_{n,d}(\omega) \underset{(d,n)\to\left(\frac{1}{2},\infty\right)}{\to} f_{1/2}(\omega) = \sigma_\epsilon^2 \left|1 - e^{-i\omega}\right|^{-1}$$

- $f_{1/2}(\omega)$ is the spectral density of a flicker noise ARFIMA$\left(0, \frac{1}{2}, 0\right)$.

Intuition of Proof 1/2

1. The coefficients $a_k^{(n,d)}$ of $\mathbf{A}_n$ satisfy $\lim_{n\to\infty} a_0^{(n,d)} = d$ and $a_k^{(n,d)} = O\left(n^{-1} + (d - 1/2)\right)$ so $\mathbf{A}_n \sim \frac{1}{2}\mathbf{I}_n$ as $(d, n) \to (1/2, \infty)$
2. Matrix

$$\mathbf{A}_n = \begin{bmatrix} a_0^{(n)} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{n-1} \end{bmatrix} + \mathbf{o}\left(1\right)$$

FER for the first row (x_t)

$$\det\left(\mathbf{I}_n - \mathbf{A}_n L\right) x_{1t} = \det\left(\mathbf{I}_{n-1} - \mathbf{A}_{n-1} L\right) \varepsilon_{1t} + \mathbf{o}_p\left(1\right)$$

and as $(d, n) \to \left(\frac{1}{2}, \infty\right)$

$$f_x\left(\omega\right) \sim \left|\frac{\det\left(\mathbf{I}_{n-1} - \mathbf{A}_{n-1} e^{-i\omega}\right)}{\det\left(\mathbf{I}_n - \mathbf{A}_n e^{-i\omega}\right)}\right|^2 \sigma_\varepsilon^2 \qquad (2)$$

Intuition of Proof 2/2

- **Szegö's theorem:**

$$\frac{\det\left(\mathbf{I}_{n-1}-\mathbf{A}_{n-1}e^{-i\omega}\right)}{\det\left(\mathbf{I}_{n}-\mathbf{A}_{n}e^{-i\omega}\right)}\underset{n\to\infty}{\rightarrow}\exp\left[-\frac{1}{2\pi}\int_0^{2\pi}\log\left(1-g_{\mathbf{A}}\left(\lambda\right)e^{-i\omega}\right)d\lambda\right]$$
$$=\exp\left[-\frac{1}{2\pi}\int_0^{2\pi d}\log\left(1-e^{-i\omega}\right)d\lambda\right]$$
$$=\left(1-e^{-i\omega}\right)^{-d}$$

- Remark:
  - NOT: aggregation of elements of ϵ_t with Beta-distributed weights as in Granger (1980).

Extension to a more general framework

- Extension to a wider class of continuous densities as $(d, p, n) \rightarrow \left(\frac{1}{2}, \infty, \infty\right)$.

$$g_{p,d}(u) = \exp\left\{-\left(\Gamma_{\frac{p+1}{p}} \frac{|u|}{2\pi d}\right)^p\right\}$$

- Circulant to Toeplitz Matrices

$$\mathbf{A}_n = \begin{bmatrix} a_0^{(n)} & a_1^{(n)} & \cdots & a_{n-1}^{(n)} \\ a_{n-1}^{(n)} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_1^{(n)} \\ a_1^{(n)} & \cdots & a_{n-1}^{(n)} & a_0^{(n)} \end{bmatrix}, \quad \mathbf{T}_n = \begin{bmatrix} a_0^{(n)} & a_1^{(n)} & \cdots & a_{n-1}^{(n)} \\ a_1^{(n)} & a_0^{(n)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_1^{(n)} \\ a_{n-1}^{(n)} & \cdots & a_1^{(n)} & a_0^{(n)} \end{bmatrix},$$

as $n \rightarrow \infty$, $z \in \mathbb{C}$,

$$\det(\mathbf{I}_n - \mathbf{A}_n z) \sim \det(\mathbf{I}_n - \mathbf{T}_n z).$$

- General VAR: $\mathbf{B}_n = \mathbf{V}_n \mathbf{T}_n \mathbf{V}^{-1}$

$$(\mathbf{I}_n - \mathbf{B}_n L)\,\mathbf{x}_t = \mathbf{V}_n (\mathbf{I}_n - \mathbf{T}_n L)\,\mathbf{V}_n^{-1} \mathbf{x}_t = \epsilon_t$$

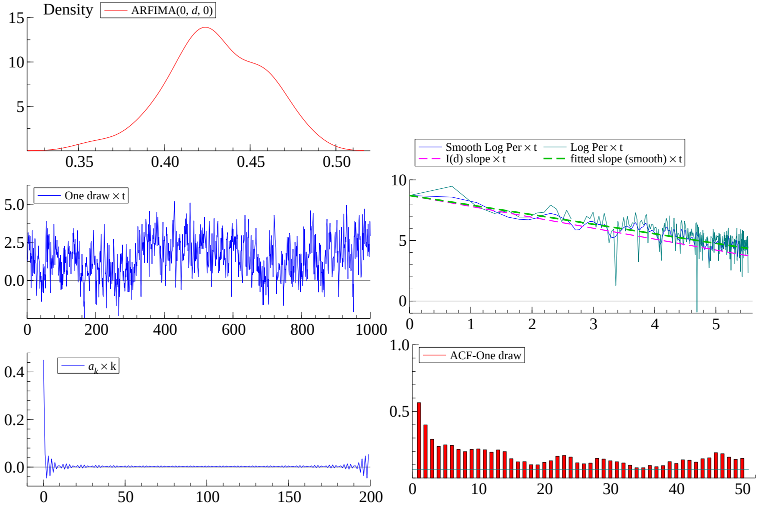
where $\det(\mathbf{I}_n - \mathbf{B}_n z) = \det(\mathbf{I}_n - \mathbf{T}_n z)$.

Outline

- Preliminary
- Our argument
- Monte Carlo and stylized empirical facts

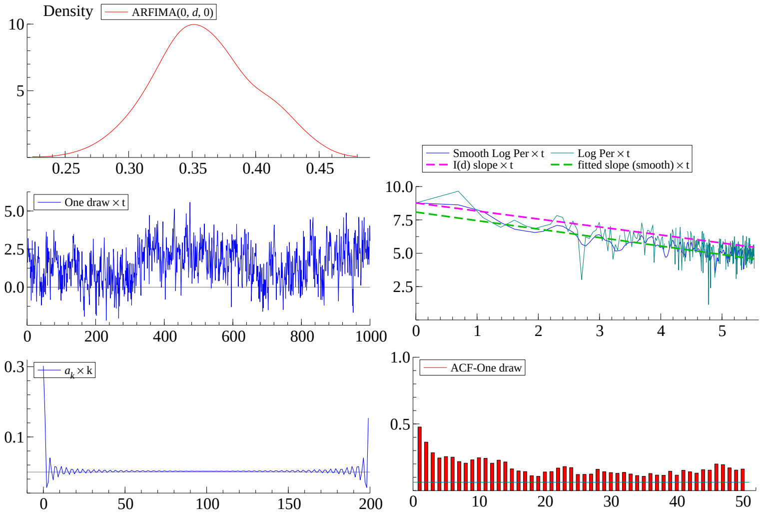
Simulation: Toeplitz matrix

Monte Carlo, $n = 200$, $T = 1000$, $d = .45$, $M = 100$



Simulation: Toeplitz matrix

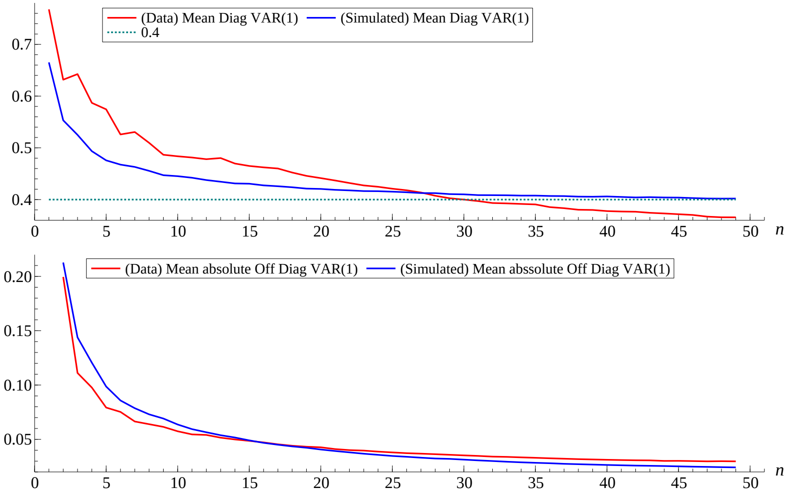
Monte Carlo, $n = 200$, $T = 1000$, $d = .30$, $M = 100$



Empirics: Stylized facts

- **Data** (provided by TickData) consists of transaction prices at the 5-minute sampling frequency for $n = 49$ large capitalization stocks from the NYSE, AMEX NASDAQ, covering the period from January 4, 1999 to December 31, 2008 (2,489 trading days). The trading session runs from 9:30 EST until 16:00 EST. Several models are estimated on daily log-returns in % (obtained by summing 5-minute log-returns) on rolling windows of 980 observations.
- Estimator of realized variation: medRV (robust to jumps)
- We Estimate a VAR(1) for data and simulated series ($d = 0.4$) and report the mean of (on/off) diagonal elements as a function of n

VAR(1) Estimation



Conclusions

- Link from cross section dependence to long memory.
- Setting
  - Large n-VAR
  - Toeplitz structure
  - small but nonzero cross correlations
  - fractional number of unit roots $\lfloor nd \rfloor$, $d \in (0, 1/2)$
- Implied univariate ARMA with AR polynomial common to all variables
  - Identical Long Memory appears in all variables as $(d, n) \to (1/2, \infty)$
- Fractional Integration should disappear in multivariate framework as n increases
- Extensions to Prices, River flow volume (Tersavirta), hidden depence (Hendry & Pretis).