Multivariate modelling of long memory processes with common components

Claudio Morana
University of Piemonte Orientale,
International Centre for Economic Research (ICER),
and Michigan State University

Contributions of the paper

A 2-step approach to the modelling of common long memory

dynamics in multivariate time series is proposed.

- The first step is concerned with a Persistent-Non Persistent decomposition or a Permanent-Transitory decomposition of the univariate processes.

- The second step is concerned with the modelling of common persistent dynamics within the principal components analysis framework.

Relevant literature

So far few papers in the literature have been concerned with the

signal-noise decomposition in the framework of long memory processes, as well as with the estimation of common factors for fractionally cointegrated processes.

Univariate framework:

- Harvey (1998): Wiener-Kolmogorov filter
- Arino and Marmol (2005): Beveridge Nelson decomposition
- Beltratti and Morana (2006): semiparametric filter

Multivariate framework:

- Morana (2002): WKF and common factors
- Morana (2004): Kasa decomposition
- Morana (2006): Gonzalo-Granger decomposition

Differently from previous contributions to the literature, the proposed approach is suitable of implementation also for the case of large data sets, both in terms of temporal and cross-sectional dimensions, not requiring neither the estimation of the fractional cointegration space nor the maximization of a frequency domain likelihood function.

Econometric methodology
P–P-NP and P-T decompositions

Consider the long memory process $\{y_t\}_{t=0}^{T-1}$ ($I(d),\ 0 < d < 1$).

Stationary long memory case ($0 < d < 0.5$):

Permanent-Persistent-Non Persistent (P-P-NP) decomposition

$$y_t = \mu + P_t + NP_t, \qquad \#$$

μ : unconditional mean of the series (permanent component, long-run forecast (component) for the series y_t, since $\lim_{s\to\infty} E_{t+s} y_t = \mu$, given that $\lim_{s\to\infty} E_{t+s} NP_t = 0$ and, for $d < 0.5$, $\lim_{s\to\infty} E_{t+s} P_t = 0$).

P_t is the persistent component ($I(d)$, $0 < d < 0.5$), or medium-run component, since for a sufficiently long, but finite, forecast horizon $\lim_{s \to k < \infty} E_{t+s}(y_t - \mu) = \lim_{s \to k < \infty} E_{t+s} P_t$, since $\lim_{s \to k < \infty} E_{t+s} NP_t = 0$.

NP_t is the non persistent component ($I(0)$), or short-run component.

Non stationary long memory case ($I(d)$, $0.5 \leq d < 1$):

Permanent-Transitory (P-T) decomposition

$$y_t = P_t + T_t, \qquad \#$$

P_t: permanent component ($I(d)$, $0.5 \leq d < 1$) or long-run forecast (component) for the series y_t, since $\lim_{s\to\infty} E_{t+s} y_t = P_t$ and $\lim_{s\to\infty} E_{t+s} T_t = 0$.

T_t: transitory component ($I(0)$), or short-run component (SRC).

Estimation

permanent component ($I(d)$, $0 < d < 0.5$): sample mean estimator;

persistent ($I(d)$, $0 < d < 0.5$)/permanent ($I(d)$, $0.5 \leq d < 1$) component P_t can be estimated as follows:

1) compute the discrete Fourier transform of the demeaned y_t process ($y_{M,t}$)

$$\tilde{y}_t = \frac{1}{T} \sum_{k=0}^{T-1} y_{M,t} e^{i2\pi k/T}. \qquad \#$$

2) discard the portion of the transformed process corresponding to the non persistent/transitory component

$$\tilde{y}_t^* = \begin{cases} \tilde{y}_t & 0 \le t \le H \\ 0 & t > H \end{cases}.$$

3) estimate the P_t component by applying the inverse discrete Fourier transform to $\tilde{y}_t^*$

$$\hat{P}_t = \frac{1}{T} \sum_{k=0}^{T-1} \tilde{y}_t^* e^{-i2\pi k/T}.$$

#

Determination of the trimming frequency $2\pi H/T$:

1) estimate the degree of fractional integration ($\hat{d}_y$) of the process y_t;

2) compute the candidate persistent/permanent processes by allowing H to vary, i.e. $H = \{3, 4, \ldots T - 1\}$, and in correspondence of each value of H evaluate the degree of persistence of the reconstructed persistent/permanent ($\hat{d}_{s,H}$) and non persistent/transitory ($\hat{d}_{n,H}$) components.

3) The optimal trimming frequency can then be determined by selecting H in such a way that

$$\hat{d}_{s,H} \simeq \hat{d}_y \text{ and } d_{n,H} \simeq 0.$$

P-P-NP and P-T decomposition for long memory processes with common persistent components

Consider the vector of n $I(d)$ fractionally cointegrated pure long memory processes $\mathbf{y}_t$ (Engle and Granger (1987) definition)

stationary long memory case

$$\mathbf{y}_t = \hat{\boldsymbol{\mu}} + \hat{\mathbf{P}}_t + \hat{\mathbf{N}}\mathbf{P}_t \qquad \#$$

non stationary long memory case

$$\mathbf{y}_t = \hat{\mathbf{P}}_t + \hat{\mathbf{T}}_t \qquad \#$$

$\hat{\mu}$: is the $n \times 1$ vector of estimated unconditional mean components,

$\hat{\mathbf{P}}_t$: is the $n \times 1$ vector of estimated long memory components ($I(d)$, $0 < d < 1$),

$\hat{\mathbf{N}}\mathbf{P}_t$ or $\hat{\mathbf{T}}_t$: is the $n \times 1$ vector of estimated short memory components (or less persistent components $I(b)$, $b < d$).

Estimation

1) Compute the permanent/persistent components ($\hat{\mathbf{P}}_t$) for the individual series (first step);

2) the s common long memory factors can be obtained by

means of principal components analysis, applied to the estimated persistent/permanent processes (second step).

The decomposition can be written as

$$\mathbf{y}_t = \hat{\boldsymbol{\mu}} + \hat{\boldsymbol{\Theta}}\mathbf{f}_t + \mathbf{N}\mathbf{P}_t^*, \qquad \#$$

or

$$\mathbf{y}_t = \hat{\boldsymbol{\Theta}}\mathbf{f}_t + \mathbf{T}_t^*, \qquad \#$$

$\hat{\boldsymbol{\Theta}} = \hat{\mathbf{B}}\hat{\boldsymbol{\Lambda}}_p^{1/2}$: $n \times s$ common long memory factor loading matrix;

$\hat{\boldsymbol{\Lambda}}_p$: estimated diagonal matrix of the non zero eigenvalues of $\hat{\boldsymbol{\Sigma}}_p$ (rank $s < n$),

$\hat{\mathbf{B}}$: estimated matrix of the associated orthogonal eigenvectors

$\mathbf{f}_t = \hat{\boldsymbol{\Lambda}}_p^{-1/2}\hat{\mathbf{B}}'\hat{\mathbf{P}}_t$: estimated $s \times 1$ vector of the standardized ($\hat{\boldsymbol{\Sigma}}_p = \mathbf{I}_s$) estimated principal components or common persistent/permanent factors.

$\mathbf{NP}_t^* = \hat{\mathbf{N}}\mathbf{P}_t + \hat{\boldsymbol{\varepsilon}}_{p,t}$ and $\mathbf{T}_t^* = \hat{\mathbf{T}}_t + \hat{\boldsymbol{\varepsilon}}_{p,t}$

$\hat{\boldsymbol{\varepsilon}}_{p,t}$: $n \times 1$ vector of estimated idiosyncratic components from $\hat{\mathbf{P}}_t = \hat{\boldsymbol{\Theta}}\mathbf{f}_t + \hat{\boldsymbol{\varepsilon}}_{p,t}$.

Theoretical results of Bai (2003, 2004) and Bai and Ng (2004) have justified the use of the PCA estimator also for strongly dependent processes. Yet, proof of consistency and asymptotic normality for long memory processes are still missing.

Monte Carlo results

P-NP and P-T decomposition

$$
\begin{aligned}
y_t &= P_t + NP_t \\
(1-L)^d P_t &= \varepsilon_t \\
&\varepsilon_t \sim n.i.d.(0,1) \\
&NP_t = \eta_t \text{ or } T_t = \eta_t \\
&\eta_t \sim n.i.d.(0,\sigma^2) \\
Cov(\varepsilon_t, \eta_s) &= 0 \quad t-s = 0,1,\ldots
\end{aligned}
$$

with $d = \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$,
$\sigma^2 = \{2, 1.5, 1, 0.5, 0.25, 0.125\}$, $t = 1,\ldots,T$, $T = 100, 1000$,
$b = \{0.1, 0.2, 0.3, -0.3, -0.2, -0.1\}$, 500 replications.

Common long memory factor estimation

$$
\begin{aligned}
\mathbf{y}_t &= \mathbf{\Phi f}_t + \mathbf{v}_t \\
& \mathbf{v}_t \sim n.i.d.(\mathbf{0}, \mathbf{\Sigma}) \\
\mathbf{\Lambda f}_t &= \boldsymbol{\varepsilon}_t \\
& \boldsymbol{\varepsilon}_t \sim n.i.d.(\mathbf{0}, \mathbf{I}), \\
Cov(\varepsilon_{it}, \eta_{js}) &= 0 \quad t - s = 0, 1, \ldots \quad \forall i, j
\end{aligned}
$$

$\mathbf{y}_t$, $t = 1, \ldots, T$, $T = \{100, 1000\}$, is a $n \times 1$ vector of generated processes with $n = \{2, 3, 4, 6, 8\}$, $\mathbf{\Lambda} = diag((1-L)^d, \ldots, (1-L)^d)_k$, $k = \{1, 2\}$, $\mathbf{\Phi}_{n \times k} = \begin{bmatrix} 1 & \ldots & 1 \end{bmatrix}'$ in the single factor case for $n = 2, 3, 4$,

$\boldsymbol{\Phi}_{n\times k} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \end{bmatrix}'$ in the two-factor case for $n = 4$, $\boldsymbol{\Phi}_{n\times k} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 & 0.5 \end{bmatrix}'$ in the two-factor case for $n = 6$, and $\boldsymbol{\Phi}_{n\times k} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0.5 & 0.5 \end{bmatrix}'$ in the two-factor case for $n = 8$,
$d = \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$,
$\Sigma = diag(\sigma^2, \ldots, \sigma^2)_n$, $\sigma^2 = \{2, 1.5, 1, 0.5, 0.25, 0.125\}$, 500 replications.

Results

The Theil inequality coefficient (IC) and the correlation coefficient (ρ) have been employed in the evaluation

$$IC = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(x_t^* - \hat{x}_t)^2}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}x_t^{*2}} + \sqrt{\frac{1}{T}\sum_{t=1}^{T}\hat{x}_t^2}},$$

$$\rho = \frac{Cov(x_t^*, \hat{x}_t)}{\sqrt{Var(x_t^*)Var(\hat{x}_t)}},$$

where x_t^* is the true unobserved component and $\hat{x}_t$ is its estimated counterpart.

Overall, the findings suggest that both the P-NP and P-T decomposition and the PCA approach may be successfully employed in the case of long memory processes.

P-NP and P-T decomposition

In general the performance of the filtering approach is satisfactory,

- increasing with the degree of persistence of the series and the sample size
- decreasing with the inverse signal to noise ratio.

Correlation coefficient

$\sigma \backslash d$	0	0.2	0.4	0.6	0.8
2	0.343	0.381	0.419	0.728	0.765
1.5	0.410	0.453	0.492	0.780	0.813
1	0.509	0.552	0.590	0.842	0.866
0.5	0.676	0.713	0.741	0.914	0.928
0.25	0.806	0.832	0.850	0.954	0.962
0.125	0.892	0.908	0.918	0.976	0.980

Theil Inequality coefficient

$\sigma \backslash d$	0	0.2	0.4	0.6	0.8
2	0.511	0.486	0.462	0.282	0.258
1.5	0.468	0.442	0.419	0.249	0.228
1	0.409	0.384	0.361	0.207	0.189
0.5	0.313	0.290	0.273	0.150	0.137
0.25	0.232	0.215	0.201	0.108	0.098
0.125	0.169	0.155	0.146	0.077	0.071

PCA approach: in general the performance is satisfactory; independently of the number of factors, the performance of the approach is

- negatively affected by the presence of noise,
- positively affected by the degree of persistence of the series and the temporal and crossectional dimensions of the sample.

Correlation coefficient, $n = 2$, $k = 1$

$\sigma \backslash d$	0	0.2	0.4	0.6	0.8
2	0.512	0.531	0.596	0.766	0.925
1.5	0.585	0.602	0.661	0.814	0.943
1	0.677	0.694	0.746	0.867	0.961
0.5	0.809	0.819	0.854	0.929	0.980
0.25	0.894	0.901	0.921	0.963	0.990
0.125	0.944	0.948	0.959	0.981	0.995

Theil Inequality coefficient, $n = 2$, $k = 1$

$\sigma \backslash d$	0	0.2	0.4	0.6	0.8
2	0.484	0.473	0.437	0.339	0.234
1.5	0.443	0.433	0.399	0.309	0.220
1	0.390	0.380	0.350	0.275	0.206
0.5	0.312	0.306	0.283	0.231	0.190
0.25	0.256	0.251	0.237	0.204	0.181
0.125	0.219	0.216	0.207	0.189	0.176

Correlation coefficient, $n = 4$, $k = 1$

$\sigma\backslash d$	0	0.2	0.4	0.6	0.8
2	0.686	0.696	0.761	0.918	0.982
1.5	0.745	0.754	0.809	0.938	0.987
1	0.813	0.821	0.864	0.957	0.991
0.5	0.897	0.901	0.927	0.978	0.995
0.25	0.946	0.948	0.962	0.989	0.998
0.125	0.972	0.974	0.981	0.994	0.999

Theil Inequality Coefficient, $n = 4$, $k = 1$

$\sigma \backslash d$	0	0.2	0.4	0.6	0.8
2	0.493	0.488	0.456	0.377	0.343
1.5	0.465	0.460	0.432	0.367	0.341
1	0.430	0.427	0.405	0.357	0.338
0.5	0.388	0.386	0.373	0.345	0.336
0.25	0.363	0.362	0.354	0.339	0.335
0.125	0.349	0.348	0.344	0.336	0.334

An application to interest rate volatility

Daily realized volatility processes for

- the overnight interest rate;
- the one-week, two-week,
- one-month, three-month, six-month, twelve-month

EONIA swap rates, obtained from intra-daily observations, sampled at the 5-minute frequency. The latter have been computed as averages of real-time, bid-ask quotes taken from REUTERS screens. The sample investigated is from 28/11/2000 through 22/04/2005, for a total of 91,392 usable observations, i.e. 952 days, with 95 5-minutes observations each (from 9 a.m. to 5 p.m.).

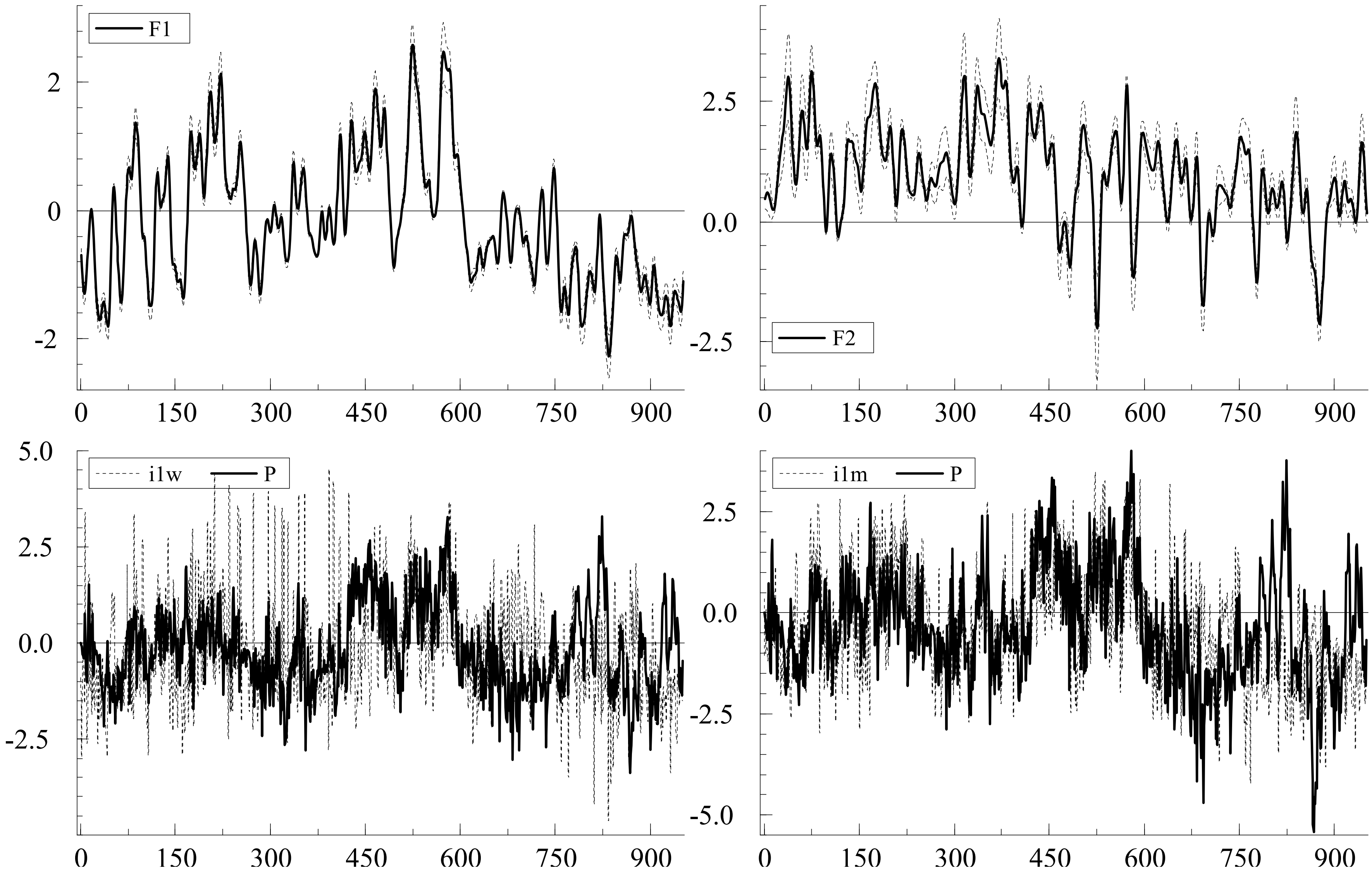
Results

Moderate degree of long memory for all the variables, with median estimates in the range 0.22-0.35, and an average value equal to 0.296 (0.032).

The null of equality of the fractional differencing parameters is not rejected at the 1% significance level in all cases.

Six cointegrating vectors can be found for the eight realized volatility series, at the 1% significance level, with the two implied common long memory factors explaining 100% of total variance at the selected bandwidth (2 ordinates).

According to the PCA results, two factors account for about 75% of total variance, supporting the results of cointegration analysis.


F1
F2
i1w
P
i1m
P
2
0
-2
2.5
0.0
-2.5
5.0
-5.0
0
150
300
450
600
750
900

Persistent-Non Persistent decomposition

	d_P	d_{NP}	m	f_1	f_2
i_0	0.353 (0.071)	0.020 (0.049)	5	0.127 (0.056)	–0.232 (0.066)
i_{1w}	0.290 (0.061)	0.000 (0.022)	42	0.618 (0.059)	–0.292 (0.054)
i_{2w}	0.281 (0.059)	0.084 (0.065)	63	0.646 (0.027)	–0.295 (0.006)
i_{1m}	0.300 (0.062)	0.014 (0.093)	121	0.965 (0.008)	0.119 (0.007)

i_{3m}	0.282 (0.057)	0.076 (0.090)	73	0.533 (0.062)	0.066 (0.031)
i_{6m}	0.316 (0.067)	0.066 (0.065)	64	0.402 (0.016)	0.284 (0.034)
i_{9m}	0.290 (0.062)	0.083 (0.062)	83	0.317 (0.036)	0.369 (0.032)
i_{1y}	0.287 (0.061)	0.069 (0.091)	90	0.246 (0.028)	0.078 (0.026)
pv				0.635 (0.029)	0.126 (0.007)

The first factor affects positively all series, while the second factor affects the shorter (up to the two-week horizon) and the longer maturities (from the one-month horizon onwards) with different signs, reflecting an excess persistent volatility component in the longer maturities, relatively to the shorter ones.

While the first factor points to forward transmission of persistent volatility shocks along the term structure, the second factor could capture the reaction to the flow of news about economic conditions, to which only the longer end of the curve is likely to react, given the characteristics of the monetary policy operational framework of the European Central Bank.

Median estimated fractional differencing operators

	i_0	i_{1w}	i_{2w}	i_{1m}
d	0.217	0.268	0.333	0.352
	0.032	0.031	0.030	0.029

i_{3m}	i_{6m}	i_{9m}	i_{1y}
0.332	0.279	0.292	0.292
0.031	0.035	0.032	0.032

Test for the equality of the fractional differencing operators

	i_0	i_{1w}	i_{2w}	i_{1m}	i_{3m}	i_{6m}	i_{9m}	i_{1y}
i_0		0.178	0.003	$3E-4$	0.002	0.097	0.047	0.047
i_{1w}	0.252		0.091	0.025	0.090	0.768	0.525	0.525
i_{2w}	0.008	0.132		0.618	0.979	0.155	0.286	0.286
i_{1m}	0.002	0.048	0.649		0.593	0.048	0.108	0.108
i_{3m}	0.010	0.144	0.982	0.638		0.154	0.288	0.288
i_{6m}	0.191	0.814	0.241	0.108	0.257		0.727	0.727
i_{9m}	0.097	0.590	0.350	0.165	0.369	0.784		1.000
i_{1y}	0.097	0.590	0.350	0.165	0.369	0.784	1.000	

Fractional cointegrating rank test, all variables

$b = 2$					eig	pv
	1%	5%	10%	TV	1.531	0.898
$r = 1$	0.000	0.000	0.000	0.013	0.175	0.102
$r = 2$	0.000	0.000	0.000	0.025	0.000	0.000
$r = 3$	0.000	0.000	0.000	0.038	0.000	0.000
$r = 4$	0.000	0.000	0.000	0.050	0.000	0.000
$r = 5$	0.000	0.000	0.000	0.063	0.000	0.000
$r = 6$	0.000	0.000	0.000	0.075	0.000	0.000
$r = 7$	0.256	0.211	0.187	0.088	0.000	0.000

Conclusions

In the paper a new approach to multivariate modelling of common long memory components has been introduced.

The approach is 2-step and based on FT filtering and PCA.

Differently from previous contributions to the literature, the proposed approach is suitable of implementation also for the case of large data sets, both in terms of temporal and cross-sectional dimensions, not requiring neither the estimation of the fractional cointegration space nor the maximization of a frequency domain likelihood function.

Monte Carlo evidence strongly supports the proposed approach.