INTERNATIONAL CENTRE FOR ECONOMIC RESEARCH

ICER

WORKING PAPER SERIES

Claudio Morana

MULTIVARIATE MODELLING OF LONG MEMORY PROCESSES WITH COMMON COMPONENTS

Working Paper No. 40/2006

Multivariate modelling of long memory processes with common components

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

December 2006

Abstract

In the paper a new approach to the modelling of common components in long memory processes is introduced. The approach is based on a two-step procedure relying on Fourier transform methods (first step) and principal components analysis (second step), which, differently from previous contributions to the literature, allows the modelling of large data sets, both in terms of temporal and cross-sectional dimensions. Monte Carlo evidence, supporting the two-step estimation procedure, is also provided, as well as an empirical application to real data.

Keywords: long memory; common long memory factor model; permanent-persistent-non persistent decomposition; permanent-transitory decomposition.

JEL classification: C32.

Address for correspondence: Claudio Morana, University of Piemonte Orientale, Faculty of Economics and Quantitative Methods, Via Perrone 18, 28100, Novara, Italy; morana@eco.unipmn.it. The author is grateful to two anonymous referees for constructive comments.

1 Introduction

Recent contributions to the literature have considered the estimation of single or common long memory components for time series observed with or without measurement error/noise.[1] Among the seminal contributions, Harvey (1998) has proposed the use of the Wiener-Kolmogorov filtering approach, generalized to the multivariate common long memory components case by Morana (2002). Beltratti and Morana (2006) have also proposed a semiparametric version of the Harvey (1998) filter, showing a similar performance of the parametric version in Monte Carlo simulations, with the advantage of not requiring the specification of the exact parametric structure of the process under investigation. Moreover, Arino and Marmol (2004) have generalized the Beveridge-Nelson (1981) permanent-transitory decomposition to the case of univariate non stationary long memory processes, while in Morana (2004) and in Morana (2006) two approaches for the estimation of the common long memory components in fractionally cointegrated processes[2], based on the Kasa (1992) and the Gonzalo and Granger (1995) decompositions, respectively, have been proposed. The advantage of these latter two approaches, relatively to those existing in the literature, is that the decomposition is carried out without requiring the numerical maximization of the frequency domain likelihood function, which can become computationally difficult as the number of processes increases.

In the paper a new permanent (P)-persistent (P)-non persistent (NP) decomposition (P-P-NP) or permanent (P)-transitory (T) decomposition (P-T) for multivariate long memory processes is introduced. The approach is based on a two-step procedure, requiring the computation of the P-P-NP or P-T decomposition of each individual series by means of Fourier transform (FT) based filtering (first step), and of the common long memory factors by means of principal components analysis (PCA) (second step). Relatively to previous contributions, the proposed approach has the advantage of being computationally fast and easily implementable also when the temporal or cross-sectional dimension is very large. Moreover, both a statistical and economic interpretation in terms of long term, medium term and short term forecasts/components, respectively, can be given to the estimated components. While recent works of Bai (2003, 2004) and Bai and Ng (2004) have justified the use of the PCA estimator also for strongly dependent processes, in this paper Monte Carlo evidence supporting the use of the PCA estimator also for the case of long memory processes is provided. Monte Carlo evidence

[1] See Baillie (1996) and Robinson (2003) for an introduction to long memory processes.

[2] See Engle and Granger (1987), Robinson and Yajima (2002) and Marinucci and Robinson (2001) for details about the concept of fractional cointegration.

supporting the proposed two-step estimation procedure is provided as well. Finally, an empirical application showing how the two-step approach can be implemented is provided.

After this introduction, the paper is organized as follows. In Sections 2 and 3 the econometric methodology is introduced and the Monte Carlo evidence discussed; in Section 4 the empirical application is presented, while in Section 5 conclusions are drawn.

2 Econometric methodology

2.1 P–P-NP and P-T decompositions

Consider the long memory process $\{y_t\}_{t=0}^{T-1}$ ($I(d)$, $0 < d < 1$). For the stationary long memory case ($0 < d < 0.5$) the following Permanent-Persistent-Non Persistent (P-P-NP) decomposition holds

$$y_t = \mu + P_t + NP_t, \tag{1}$$

where μ is the unconditional mean of the series, i.e. the permanent component, P_t is the persistent component ($I(d)$, $0 < d < 0.5$) and NP_t is the non persistent component ($I(0)$[3]). The unconditional mean component μ can be interpreted as the long-run forecast 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 $d < 0.5$, $\lim_{s\to\infty} E_{t+s}P_t = 0$.[4] The mean component is denoted as the long-run component (LRC). The persistent component can be interpreted as the medium-run component (MRC), 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 $\lim_{s\to k<\infty} E_{t+s}NP_t = 0$. The non persistent component can finally be interpreted as the short-run component (SRC).

On the other hand, for the non stationary long memory case ($I(d)$, $0.5 \leq d < 1$) the following Permanent-Transitory (P-T) decomposition holds

$$y_t = P_t + T_t, \tag{2}$$

where P_t is the permanent component ($I(d)$, $0.5 \leq d < 1$) and T_t is the transitory component ($I(0)$). In this setting the component P_t can be interpreted as the long-run forecast for the series y_t, since $\lim_{s\to\infty} E_{t+s}y_t = P_t$

[3] An I(0) process is a stochastic process which does not require (integer/fractional) differencing to become weakly stationary. It is characterized by finite long-run variance, or, equivalently, finite spectral density at the zero frequency.

[4] This follows from the fact that, under covariance stationarity, the mean reversion property implies that the long-run forecast converges to the unconditional mean.

and $\lim_{s \to \infty} E_{t+s} T_t = 0$.[5] Moreover, the permanent component is denoted as the long-run component (LRC), while the transitory component as the short-run component (SRC).

2.1.1 Estimation

While the estimation of the permanent component in the stationary long memory case can be performed by means of the sample mean estimator[6], the estimation of the persistent ($I(d)$, $0 < d < 0.5$)/permanent ($I(d)$, $0.5 \leq d < 1$) component P_t can be performed as follows. Firstly, the discrete Fourier transform of the demeaned y_t process ($y_{M,t}$), is computed

$$\tilde{y}_t = \frac{1}{T} \sum_{k=0}^{T-1} y_{M,t} e^{i2\pi k/T}. \tag{3}$$

Then, the portion of the transformed process corresponding to the non persistent/transitory component is discarded by setting

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

Finally, the P_t component is estimated by applying the inverse discrete Fourier transform to $\tilde{y}_t^*$, yielding

$$\hat{P}_t = \frac{1}{T} \sum_{k=0}^{T-1} \tilde{y}_t^* e^{-i2\pi k/T}.^7 \tag{4}$$

[5] The $0.5 \leq d < 1$ case is often referred as the non stationary, yet mean reverting, long memory case. The mean reversion property depends on the fact that the effects of shocks eventually dye out, i.e, provided $d < 1$, the sequence of impulse response weights converges to zero asymptotically. It is the dissipation of shocks that allows one to denote these non covariance stationary processes as “mean reverting” (Robinson, 2003; p.20). Yet this characterization is not fully accepted in the literature. For instance, Phillips and Xiao (1999, p.34) have argued against the latter interpretation, due to the lack of covariance stationarity. The author is grateful to a referee for having raised this issue and a very useful discussion on this point.

[6] Results of Beran (1994, ch.8) point that the sample mean estimator is unbiased and efficient also in the stationary long memory case ($0 < d < 0.5$). In the Gaussian case the estimator is also the maximum likelihood estimator and therefore it is optimal relative to the class of both linear and non linear estimators. The sample mean estimator is also consistent, albeit the rate of convergence in the stationary long memory case is slower than for the i.i.d. case: T^{-d} rather than $T^{-0.5}$. Efficient and robust estimators of the location parameter belonging to the class of M-estimators have also been proposed in Beran (1994, ch.8). It is found that the loss of efficiency for this latter class of consistent estimators, relatively to the sample mean estimator, is small and null in the case of Gaussian long memory processes.

[7] For the non stationary long memory case the standard FT formula requires a cor-

A two-step procedure for the determination of the trimming frequency $2\pi H/T$ can be followed. Once the degree of fractional integration $(\hat{d}_y)$ of the process y_t has been determined, candidate persistent/permanent processes are computed by allowing H to vary, i.e. $H = \{3, 4, ...T-1\}$, computing, in correspondence of each value of H, the degree of persistence of the reconstructed persistent/permanent $(\hat{d}_{s,H})$ and non persistent/transitory $(\hat{d}_{n,H})$ components. The optimal trimming frequency can then be determined by selecting H in such a way that

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

i.e. as the frequency at which the reconstructed persistent/permanent component has a degree of persistence not statistically different from the one of the actual process and the reconstructed non persistent/transitory component has a degree of persistence not statistically different from zero. In the case more than a frequency satisfied this latter criterion, the farthest one from the zero frequency may be selected in order not to disregard signal potentially belonging to the persistent/permanent component. On the other hand, if a smoothed version of the unobserved component is sought, or when the non persistent/transitory component may still be an integrated process $I(b)$, $0 < b < d$, although of lower order than the actual series, the optimal trimming frequency may be set equal to the closest one to the zero frequency, still satisfying the above persistence requirements. Once the permanent and persistent components are available, the $\hat{N}P_t$ and $\hat{T}_t$ components can be obtained as $y_{M,t} - \hat{P}_t$. Monte Carlo results, providing full support to the proposed methodology, are reported in the next section.[8]

rection. In fact, from formula [59] in Phillips (1999), for the process $(1-L)^d x_t = u_t$, $t = 1, ..., T$, $0.5 < d < 1$ it is found $w_x(\lambda_s) \simeq (1-e^{i\lambda_s})^{-d} w_u(\lambda_s) - \frac{e^{i\lambda_s}}{1-e^{i\lambda_s}} \frac{x_T}{\sqrt{2\pi T}}$. See also Knsch (1986) on FT of stationary long memory processes, and Pollock (2005) for additional details on FT based signal extraction methods. Finally, see Engle (1974) for a seminal contribution to FT based filtering and estimation.

[8]There are two other non stationary cases of interest, which are not directly covered by the proposed approach. The first one is the non mean reverting fractional integration case, i.e. the $d \geq 1$ case. The second one is the deterministic non stationary case. Yet, the proposed approach could be still employed to handle such cases. The non mean reverting case would require integer differencing to be applied to the series before FT filtering, and then the integration of the estimated signal. On the other hand, the deterministic non stationary case would require the filtering out of the deterministic break process before FT filtering is performed on the series, and then its addition to the estimated signal obtained from the break-free process.

2.2 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 $\mathbf{y}_t$. The definition of fractional cointegration employed is the one of Engle and Granger (1987), according to which the vector process of order n is fractionally cointegrated of order d and b, i.e. $CI(d,b)$, if there exist up to $n-1$ linear combinations of the series characterized by an order of integration equal to b, with $b < d$. Then, a permanent-persistent-non persistent (P-P-NP) decomposition or a permanent-transitory (P-T) decomposition for the n series can be carried out by means of a two-step procedure based on Fourier transform and principal components analysis, and written as

$$\mathbf{y}_t = \hat{\boldsymbol{\mu}} + \hat{\mathbf{P}}_t + \hat{\mathbf{N}}\mathbf{P}_t \tag{5}$$

for the stationary long memory case, or

$$\mathbf{y}_t = \hat{\mathbf{P}}_t + \hat{\mathbf{T}}_t \tag{6}$$

for the non stationary long memory case, where $\hat{\mu}$ 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 ($I(d)$, $0 < d < 1$), and $\hat{\mathbf{N}}\mathbf{P}_t$ or $\hat{\mathbf{T}}_t$ is the $n \times 1$ vector of estimated short memory components (or less persistent components $I(b)$, $b < d$).

Once an estimate of the individual permanent/persistent components is available ($\hat{\mathbf{P}}_t$), i.e. the FT based first step has been carried out, the s common long memory factors can be obtained by means of principal components analysis, applied to the estimated persistent/permanent processes (second step).[9]

The decomposition can be written as

$$\mathbf{y}_t = \hat{\boldsymbol{\mu}} + \hat{\boldsymbol{\Theta}}\mathbf{f}_t + \mathbf{NP}_t^*, \tag{7}$$

or

$$\mathbf{y}_t = \hat{\boldsymbol{\Theta}}\mathbf{f}_t + \mathbf{T}_t^*, \tag{8}$$

where $\hat{\boldsymbol{\Theta}} = \hat{\mathbf{B}}\hat{\boldsymbol{\Lambda}}_p^{1/2}$ is the estimated $n \times s$ common long memory factor loading matrix, $\hat{\boldsymbol{\Lambda}}_p$ is the estimated diagonal matrix of the non zero eigenvalues of the estimated reduced rank variance-covariance matrix of the persistent/permanent processes $\hat{\boldsymbol{\Sigma}}_p$ (rank $s < n$), $\hat{\mathbf{B}}$ is the estimated matrix of

[9] Bierens (2000) has proposed a similar approach for the estimation of the common non linear deterministic components from a set of estimated individual non linear deterministic components. Yet, the discussion of this approach cannot be found in the published version of his paper.

the associated orthogonal eigenvectors, and $\mathbf{f}_t = \hat{\mathbf{\Lambda}}_p^{-1/2}\hat{\mathbf{B}}'\hat{\mathbf{P}}_t$ is the estimated $s \times 1$ vector of the standardized ($\hat{\mathbf{\Sigma}}_p = \mathbf{I}_s$) estimated principal components or common persistent/permanent processes. Finally, $\mathbf{NP}_t^* = \hat{\mathbf{N}}\mathbf{P}_t + \hat{\boldsymbol{\varepsilon}}_{p,t}$ and $\mathbf{T}_t^* = \hat{\mathbf{T}}_t + \hat{\boldsymbol{\varepsilon}}_{p,t}$, where $\hat{\boldsymbol{\varepsilon}}_{p,t}$ is an $n \times 1$ vector of estimated idiosyncratic components from $\hat{\mathbf{P}}_t = \hat{\mathbf{\Theta}}\mathbf{f}_t + \hat{\boldsymbol{\varepsilon}}_{p,t}$.[10]

To date no theoretical results concerning the asymptotic properties of the (time domain) PCA estimator for long memory processes have been provided in the literature. Yet, recent theoretical developments of Bai (2003, 2004) and Bai and Ng (2004) have justified the use of the PCA estimator also for dependent processes. In particular, Bai (2004) has considered the generalization of PCA to the case in which the series are weakly dependent processes, establishing consistency and asymptotic normality when both the unobserved factors and idiosyncratic components show limited serial correlation, also allowing for heteroskedasticity in both the time and cross section dimension in the idiosyncratic components. In Bai (2003) consistency and asymptotic normality has been derived for the case of I(1) unobserved factors and I(0) idiosyncratic components, also in the presence of heteroskedasticity in both the time and cross section dimension in the idiosyncratic components. Finally, Bai and Ng (2004) have established consistency also for the case of I(1) idiosyncratic components. As pointed out by Bai and Ng (2004), consistent estimation should also be achieved by PCA in the intermediate case represented by long memory processes. Monte Carlo evidence that the performance of the principal components approach is indeed not affected by the presence of long memory, being also robust to the presence of moderate noise, is provided in the section below.

[10] The Engle and Granger (1987) definition of fractional cointegration is not the only one. Generalizations considered in Marinucci and Robinson (2001) and Robinson and Yajima (2002) allow for the case of subsets of variables showing different orders of integration. If some of the higher order variables cointegrate within their own subset, the consequent order reduction may allow fractional cointegration to concern variables belonging to different subsets. The proposed approach allows to handle this latter case, requiring the PCA analysis to be carried out considering subsets of variables characterized by the same order of integration. For instance, common factors would be first extracted within the block of higher order variables and the corresponding cointegration relationships estimated. Then, the lower order estimated cointegrating errors would be included, as variables, in the block composed of the other series of the same integration order and the PCA analysis carried out on this latter block of variables.

3 Monte Carlo results

Two Monte Carlo experiments have been carried out to evaluate the performance of the decomposition methodology, as well as the PCA approach to common long memory factor estimation.

3.1 P-NP and P-T decomposition

The following data generation process has been assumed for the series

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

with $d = \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$, $\sigma^2 = \{2, 1.5, 1, 0.5, 0.25, 0.125\}$, $t = 1, ..., T$, $T = 100, 1000$. The number of replications has been set to 500 for each case. In the assessment, the degree of persistence of both the actual series and the persistent component have been estimated[11], so that the evaluation concerns the feasible version of the proposed methodology. The performance of the approach has also been evaluated assuming different degrees of bias in the estimation of the fractional differencing operator, i.e. the size of the bias has been set to $b = \{0.1, 0.2, 0.3, -0.3, -0.2, -0.1\}$.

3.2 Common long memory factor estimation

The following data generation process has been assumed for the series

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

where $\mathbf{y}_t$, $t = 1, ..., T$, $T = \{100, 1000\}$, is a $n \times 1$ vector of generated processes with $n = \{2, 3, 4, 6, 8\}$, $\mathbf{\Lambda}$ is a $k \times k$ diagonal matrix $\mathbf{\Lambda} = diag((1-$

[11] For its good properties and robustness to bandwidth selection, the estimator proposed by Robinson (1998) has been employed.

$L)^d, ..., (1-L)^d)$, $\mathbf{f}_t$ is a $k \times 1$ vector of generated common order (d) long memory processes with $k = \{1, 2\}$, $\mathbf{\Phi}$ is a $n \times k$ factor loading matrix , defined as $\mathbf{\Phi} = \begin{bmatrix} 1 & ... & 1 \end{bmatrix}'$ in the single factor case for $n = 2, 3, 4$, $\mathbf{\Phi} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \end{bmatrix}'$ in the two-factor case for $n = 4$, $\mathbf{\Phi} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 & 0.5 \end{bmatrix}'$ in the two-factor case for $n = 6$, and $\mathbf{\Phi} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0.5 & 0.5 \end{bmatrix}'$ in the two-factor case for $n = 8$, $d = \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$, $\mathbf{v}_t$ is a $n \times 1$ vector of generated idiosyncratic Gaussian white noise processes, Σ is a $n \times n$ diagonal matrix, with $\Sigma = diag(\sigma^2, ..., \sigma^2)$, $\sigma^2 = \{2, 1.5, 1, 0.5, 0.25, 0.125\}$. The number of replications has been set to 500 for each case. The robustness of the approach to wrongly neglecting the presence of one factor, in the estimation of the other factor, has been evaluated as well.

3.3 Results

The performance of the estimators has been assessed with reference to the ability of recovering the unobserved components P_t, and f_t, respectively. The Theil inequality coefficient (IC) and the correlation coefficient (ρ) have been employed in the evaluation

$$\begin{aligned} IC &= \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(x_t^* - \hat{x}_t)^2}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}x_t^{*2}} + \sqrt{\frac{1}{T}\sum_{t=1}^{T}\hat{x}_t^2}}, \\ \rho &= \frac{Cov(x_t^*, \hat{x}_t)}{\sqrt{Var(x_t^*)Var(\hat{x}_t)}}, \end{aligned}$$

where x_t^* is the true unobserved component and $\hat{x}_t$ is its estimated counterpart. The above statistics have been computed for each Monte Carlo replication and then averaged. For the two-factor case, averages of the statistics obtained for the two estimated factors are reported.

The results of the Monte Carlo exercise for the P-NP and P-T decompositions and the principal components (PCA) factor estimation approach are reported in Tables 1-9. As is shown in Table 1, the performance of the P-NP and P-T filtering approaches is satisfactory, both in terms of correlation coefficient and inequality coefficient. In general the performance of the filtering approach tends to increase with the degree of persistence of the series and the sample size, and to decrease with the inverse signal to noise ratio. As far as the effects of a change in the degree of persistence is concerned, they are particularly noticeable when the non stationary and stationary long memory cases are compared, for large inverse signal to noise ratios. For instance, in the worst case ($T = 100$, $\sigma^2 = 2$) an increase in the correlation coefficient of up 200%, from 0.36 to 0.78, may be noted as the persistence parameter increases from 0 to 0.9 (from 0.343 to 0.973 in the large sample case). Less dramatic improvements can be noted as the inverse signal to noise ratio falls, bottoming at about 10% for the best case scenario ($\sigma^2 = 0.125$), where the correlation coefficient is never below 0.87 (0.89 in the large sample case). The increase in performance due to the increase of the sample size is also noticeable, particularly for the non stationary case and when the inverse signal to noise ratio is large, with an increase in the correlation coefficient up to about 20%, i.e. from 0.78 to 0.98 in the best case scenario ($d = 0.9$, $\sigma^2 = 2$).

Similar findings hold when the inequality coefficient is employed in the assessment. For instance, even in the worst case scenario values larger than 0.50 are found only in 8% of the cases ($T = 100$, $\sigma^2 = 2$, $d < 0.5$), the IC statistic being below 0.30 in 50% of the cases and below 0.40 in about 70% of the cases. Even better results can be found for the large sample case, with the IC statistic taking values larger than 0.50 only in the 5% of the cases, and below 0.30 and 0.40 in 60% and 72% of the cases, respectively.

Moreover, as shown in Table 2, estimating the fractional differencing parameter with moderate positive bias only has a small negative impact on the ability of the estimator to recover the actual process, particularly when the inverse signal to noise is large. In terms of correlation coefficient, up to a 20% loss in performance can be found in the worst case ($\sigma = 2$) for the large sample case, while an up to 10% loss in performance can be found in the small sample case. Yet, the findings are not univocal since, in some cases, also an improvement in performance is associated with the postive bias, particularly for the $d = 0.4$ case. Coherent with the findings for the correlation coefficient, no major differences can also be found for the IC statistic.[12] Therefore, it can be concluded that the proposed P-NP filtering approach shows a satisfactory performance, being robust to the presence of noise, sample size, degree of persistence, and biased estimation of the degree of persistence.

As far as the PCA approach is concerned, as shown in Tables 3-9, independently of the number of factors, the performance of the approach is negatively affected by the presence of noise, but positively affected by the degree of persistence of the series, and the temporal and crossectional dimensions of the sample. For instance, for the single factor case increasing

[12]Results for the negative bias case are qualitatively similar and are not included for reasons of space. They are however available upon request from the author.

the crossectional dimension of one unit, i.e. from $n = 2$ to $n = 3$, leads to an increase in the correlation coefficient of up to 20%, while increasing it of a further unit, i.e. from $n = 3$ to $n = 4$, yields an additional improvement of up to 10%. Yet, the improvement in performance, due to increasing the crossectional dimension of the sample, tends to be less noticeable as the degree of persistence increases. Differently, the improvement in the correlation coefficient due to increasing the temporal dimension of the sample tends to increase with the degree of persistence, with the extension of the sample size from 100 to 1000 observations leading to an increase of up to 30% in the correlation coefficient between the actual and generated factors. As far as the effects of observational noise are concerned, independently of the crossectional dimension of the sample, the negative impact tends to become stronger as the temporal dimension of the sample lowers as well as the degree of persistence decreases. For instance, in the worst case ($n = 2$, $d = 0.2$) increasing the inverse signal to noise ratio from 0.125 to 2 yields more than a halving in the correlation coefficient. Finally, as far as the effects of the degree of persistence on the performance of the approach are concerned, independently of the size dimensions of the sample, an increase in the fractional differencing parameter leads to an improvement in the correlation coefficient between the actual and estimated factors, with the improvement been particularly noticeable when the stationary long memory case is compareed with the non stationary long memory one. The best cases for the $n = 2$, $n = 3$, and $n = 4$ cases point to an increase of up to 100%, 70%, and 50%, respectively. Yet, the performance of the approach is always very satisfactory. Even in the worst case, i.e. $n = 2$, $d = 0.2$, $T = 100$, $\sigma^2 = 2$, the correlation coefficient between the actual and estimated factor is close to 0.40, while, when the effects of the observational noise are negligible ($\sigma^2 = 0.125$), the correlation coefficient, for the same case, is close to 0.90. Yet, even in the presence of observational noise ($\sigma^2 = 2$), the correlation coefficient is larger than 0.60 for the non stationary case ($n = 2$, $d > 0.5$, $T = 100$). Similar conclusions holds when the inequality coefficient is employed, with the coefficient exceeding the 0.5 value only in 8% of cases in the worst scenario ($n = 2$, $T = 100$). Still in the worst scenario the Theil inequality coefficient takes values below 0.3 in 50% of the cases and below 0.40 in 75% of the cases.

As shown in Tables 7, 8, and 9, similar results hold for the two-factor case as well, although in general the performance of the PCA approach tends to be lower than for the single factor case. This finding is a consequence of the averaging of the results for the two factors, since, as far as the first factor alone is concerned, the performance of the method is actually superior to what found for the single factor case, due to the large number of relevant series involved.[13]

Moreover, by comparing Tables 6 and 3 it is possible to notice that the performance of the approach does not seem to be affected by the inclusion of information (additional variables) non relevant for the estimation of the factor of interest. In fact, figures for the two cases are always very close, without evidence of a clear ranking pattern. This finding can also be interpreted as poiting that the performance of the approach is not negatively affected by neglecting a second factor when estimating the one of interest, i.e. the first one. This result is fully coherent with the way the PCA estimator extracts the common factors, with the factor explaining the higher proportion of total variance being not affected, by construction, by the other factors which follows in estimation.

From above the results it can then be concluded that also the overall performance of the PCA approach is very satisfactory, being robust to the presence of noise, degree of persistence and temporal/crossectional sample size.

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

4 An application to interest rate volatility

In the application daily realized volatility[14] processes for the overnight interest rate, and the one-week, two-week, one-month, three-month, six-month and twelve-month EONIA swap rates have been employed. The realized volatility processes have been 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 107,291 5-minute observations, excluding weekends and holidays. Excluding thin trading days (days in which prices did not change) 91,392 usable observations were left, i.e. 952 days, with 95 5-minutes observations each (from 9 a.m. to 5 p.m.).

The implementation of the approach requires at the first stage the estimation of the degree of persistence of each series. In order to achieve robust conclusions on the order of integration of the series, both parametric and

[13] For reasons of space detailed results are not reported. They are however available from the author upon request.

[14] See Andersen et al. (2001) and Barndorff-Nielsen and Shephard (2002).

semiparametric estimators have been employed. The semiparametric estimators employed belong to two classes, i.e. log periodogram estimation (GPH, Geweke and Porter-Hudak, 1983), and local Whittle estimation (Künsch, 1987; Robinson,1995), considering, in addition to the original contributions, also recent extensions.[15] Following Taqqu and Teverovsky (1998), in all the cases the final estimates have been obtained as averages over the stable region closest to the zero frequency. On the other hand, parametric estimation of the fractional differencing operator has been performed by means of exact maximum likelihood (Sowell, 1992), modified profile likelihood (An and Bloomfield, 1993) and non linear least squares estimation (Beran, 1995).[16] An overall estimate of the fractional differencing parameter for each series has then been computed as the median of the estimates obtained by means of the various estimators. The evidence points to a 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).[17] Since the test for the equality of the fractional differencing parameters (Robinson and Yajima, 2002; Morana, 2006) does not allow to reject the null of equality of the fractional differencing parameters at the 1% significance level in all cases, the average estimate obtained for the eight series (0.296 (0.032)) has then been employed in the fractional cointegration analysis. According to the results of the Robinson and Yajima (2002) fractional cointegrating rank test, 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)[18].

The filtering out of the non persistent components has been carried out by

[15]Recent extensions to these two classes of estimators, aiming to improve the performance of the estimators in the presence of short memory dynamics (Andrews and Guggenberger (2003) bias reduced log periodogram estimator, Andrews and Sun (2004) biased reduced local Whittle estimator, Shimotsu and Phillips (2005) pooled log periodogram estimator, Moulines and Soulier (1999) broad band log periodogram estimator), observational noise (Sun and Phillips (2003) non linear log periodogram estimator), non stationarity (Shimotsu and Phillips (2004) exact local Whittle estimator), have been considered. On the basis of Monte Carlo results available in the literature, also the estimator of Robinson (1998), which has proved to be robust to both bandwidth selection and the presence of nonstationarity, has been employed.

[16]The ARFIMA(0,d,0) model was selected for all of the series using the BIC criterion.

[17]The estimated fractional differencing parameters are equal to 0.217(0.032), 0.268(0.031), 0.333(0.030), 0.352(0.029), 0.332(0.031), 0.279(0.035), 0.292(0.032), and 0.292(0.032), for the overnight, the one-week, two-week, one-month, three-month, six-month, nine-month and one-year rates, respectively.

[18]The non zero eigenvalues of the spectral matrix at the zero frequency are equal to 1.531 and 0.175. Hence, the proportion of variance explained by the two common long memory factors is 89.8% and 10.2%, respectively.

applying the FT to the demeaned individual series, trimming high frequency ordinates, and then applying inverse FT to the trimmed transformed series. Optimal trimming has been implemented as described in the methodological section, i.e. the optimal trimming frequency has been selected as closest one to the origin at which the decomposition objective is achieved.[19] The estimation of the common long memory factors has then been carried out by applying principal components analysis to the estimated persistent components. On the basis of the results of the cointegration analysis, it is expected that two common long memory factors explain the bulk of persistent fluctuations. Indeed, according to the PCA results, two factors account for about 75% of total variance, supporting the results of cointegration analysis. The estimated factors, with 95% confidence bounds, and the persistent components for the one week and one month log standard deviation processes, are reported in Figure 1. As shown by the estimated factor loading matrix[20], 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.[21]

[19] A smoothed persistent component characterized by a degree of persistence non statistically different from the one determined by the semiparametric analysis ($d = 0.296$ (0.032)), and a non persistent component characterized by a degree of persistence non statistically different from zero. The optimal trimming ordinates range between 5 (the overnight rate) and 121 (the one month rate).

[20] The estimated factor loadings are as follows. First factor: 0.127(0.056), 0.618(0.059), 0.646(0.027), 0.965(0.008), 0.533(0.062), 0.402(0.016), 0.317(0.036), 0.246(0.028), 0.635(0.029) for the overnight, one-week, two-week, one-month, three-month, six-month, nine-month, one-year rates, respectively. Second factor: -0.232(0.066), -0.292(0.054), -0.295(0.006), 0.119(0.007), 0.066(0.031), 0.284(0.034), 0.369(0.032), 0.078(0.026), 0.126(0.007) for the overnight, one-week, two-week, one-month, three-month, six-month, nine-month, one-year rates, respectively. Standard errors have been computed from the cross sectional distributions of the parameters obtained allowing the persistent components to be estimated using several optimal trimming bandwidths, falling in the range 5-121, i.e. the range matching the one found relevant for the actual series.

[21] At issue is the control exercised by the ECB on the shortest end on the yield curve. Low volatility should be expected in the case of successful control. For reason of space detailed results for the empirical application have not been reported. They are however available from the author upon request.

5 Conclusions

In the paper a new approach to multivariate modelling of common long memory components has been introduced. 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. Finally, an empirical application, showing that the proposed approach can be easily implemented using real data, has been provided.

References

[1] An, S and P. Bloomfield, 1993, Cox and Reid's Modification in Regressions Models with Correlated Errors, mimeo, North Carolina State University, Department of Economics.

[2] Andersen, T.G., T. Bollerslev, F.X. Diebold and P. Labys, 2001, The Distribution of Realized Exchange Rate Volatility, Journal of the American Statistical Association, 96, 42-55.

[3] Andrews, D.W.K. and P. Guggenberger, 2003, A Bias-Reduced Log Periodogram Regression Estimator for the Long Memory Parameter, Econometrica, 71, 675-712.

[4] Andrews, D.W.K. and Y. Sun, 2004, Adaptive Local Polynomial Whittle Estimation of Long Range Dependence, Econometrica, 72(2), 569-614.

[5] Arino, M.A. and F. Marmol, 2004, A Permanent-Transitory Decomposition for ARFIMA Processes, Journal of Statistical Planning and Inference, 124, 87-97.

[6] Bai, J., 2004, Estimating Cross-Section Common Stochastic Trends in Nonstationary Panel Data, Journal of Econometrics, 122, 137-38.

[7] Bai, J., 2003, Inferential Theory for Factor Models of Large Dimensions, Econometrica, 71(1), 135-171,

[8] Bai, J. and S. Ng, 2004, A Panick Attack on Unit Roots and Cointegration, Econometrica, 72(4), 1127-1177.

[9] Baillie, R.T., 1996, Long Memory Processes and Fractional Integration in Econometrics, Journal of Econometrics, 73, 5-59.

[10] Barndorff-Nielsen, O. and N. Shephard, 2002, Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models, Journal of the Royal Statistical Society, Series B, 64, 253-80.

[11] Beltratti, A. and C. Morana, 2006, Breaks and Persistence: Macroeconomic Causes of Stock Market Volatility, Journal of Econometrics, 131, 151-177.

[12] Beran, J., 1994, Statistics for Long Memory Processes, Chapman and Hall, New York.

[13] Beran, J., 1995, Maximum Likelihood Estimation of the Differencing Parameter for Invertible Short and Long Memory Autoregressive Integrated Moving Average Models, Journal of the Royal Statistical Society 57, 659-672.

[14] Beveridge, S. and C.R. Nelson, 1981, A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the "Business Cycle", Journal of Monetary Economics 7, 151-174.

[15] Bierens, H.J., 2000, Non Parametric Nonlinear Cotrending Analysis, with an Application to Interest and Inflation in the United States, Journal of Business and Economic Statistics, 18(3), 323-37.

[16] Engle, R.F., 1974, Band Spectrum Regression, 15(1), 1-11.

[17] Engle, R.F. and C.W.J. Granger, 1987, Co-integration and Error Correction Representation, Estimation, and Testing, Econometrica, 55, 251-76.

[18] Geweke, J. and S. Porter-Hudak, 1983, The Estimation and Application of Long Memory Time Series, Journal of Time Series Analysis, 4, 221-38.

[19] Gonzalo, J. and C. Granger, 1995, Estimation of Common Long-Memory Components in Cointegrated Systems, Journal of Businness and Economic Statistics, 13(1), 27-35.

[20] Harvey, A.C., 1998, Long Memory in Stochastic Volatility, in Forecasting Volatility in Financial Markets. J Knight and S Satchell (eds), 1998, 307-320. Oxford: Butterworth-Heineman.

[21] Kasa, K., 1992, Common Stochastic Trends in International Stock Markets, Journal of Monetary Economics, 29, 95-124.

[22] Künsch, H.R., 1987, Statistical Aspects of Self Similar Processes. In Proceedings of the First World Congress of the Bernoulli Society, Y. Prohorov and V.V. Sazanov, eds., 1, 67-74, VNU Science Press, Utrecht.

[23] Künsch, H.R., 1986, Discriminating Between Monotonic Trends and Long-Range Dependence, Journal of Applied Probability, 23, 1025-30.

[24] Marinucci, D. and P.M. Robinson, 2001, Semiparametric Fractional Cointegration Analysis, Journal of Econometrics, 105, 225-47.

[25] Morana, C, 2002, Common Persistent Factors in Inflation and Excess Nominal Money Growth and a New Measure of Core Inflation, Studies in Non Linear Dynamics and Econometrics, 6(3), art.3-5.

[26] Morana, C., 2004, Frequency Domain Principal Components Estimation of Fractionally, Cointegrated Processes, Applied Economics Letters, 11, 837-42.

[27] Morana, C., 2006, A Small Scale Macroeconometric Model for Euro-12 Area , Economic Modelling, 23(3), 391-426.

[28] Moulines, E. and P. Soulier, 1999, Broadband Log-Periodogram Regression of Time Series with Long Range Dependence, The Annals of Statistics, 27(4), 1415-39.

[29] Phillips, P.C., 1999, Discrete Fourier Transforms of Fractional Processes, Yale University, mimeo.

[30] Phillips, P.C.B and Z. Xiao, 1999, A Primer on Unit Root Testing, in M. McAleer and L. Oxley, eds., *Practical Issues in Cointegration Analysis*, Blackwell Publishers.

[31] Pollock, D.S.G., 2005, Econometric Methods of Signal Extraction, mimeo, University of London.

[32] Robinson, P.M., 1994, Semiparametric Analysis of Long Memory Time Series, Annals of Statistics, 22, 515-39.

[33] Robinson, P.M., 1995, Gaussian Semiparametric Estimation of Long Range Dependence, The Annals of Statistics, 23(5), 1630-61.

[34] Robinson, P.M., 1998, Comment, Journal of Business and Economic Statistics, 16(3), 276-79.

[35] Robinson, P.M. and Y. Yajima, 2002, Determination of Cointegrating Rank in Fractional Systems, Journal of Econometrics, 106(2), 217-41.

[36] Robinson, P.M., 2003, Long Memory Time Series, in P.M Robinson, ed., *Time Series with Long Memory*, Oxford University Press.

[37] Shimotsu, K. and P.C.B. Phillips, 2002, Pooled Log Periodogram Regression, Journal of Time Series Analysis, 23(1), 57-93.

[38] Shimotsu, K. and P.C.B. Phillips, 2005, Exact Local Whittle Estimation of Fractional Integration, Annals of Statistics 33, 1890-1933.

[39] Sowell, F., 1992, Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models, Journal of Econometrics, 53, 165-88.

[40] Sun, Y. and P.C.B. Phillips, 2003, Non Linear Log-Periodogram Regression for Perturbed Fractional Processes, Journal of Econometrics, 115, 355-89.

[41] Taqqu, M.S. and V. Teverovsky, 1998, Semi-Parametric Graphical Estimation Techniques for Long Memory Data, in P.M. Robinson and M. Rosemblatt, eds., *Time Series Analysis in Memory of E.J. Hannan*, 420-32, New York: Springer Verlag.

Table 1

P-NP filtering, Monte Carlo results; sample size 1,000 observations

Correlation coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.343	0.328	0.381	0.358	0.419	0.527	0.728	0.797	0.765	0.975
1.5	0.410	0.394	0.453	0.427	0.492	0.599	0.780	0.840	0.813	0.981
1	0.509	0.494	0.552	0.529	0.590	0.689	0.842	0.887	0.866	0.987
0.5	0.676	0.661	0.713	0.691	0.741	0.816	0.914	0.939	0.928	0.993
0.25	0.806	0.796	0.832	0.816	0.850	0.898	0.954	0.968	0.962	0.997
0.125	0.892	0.885	0.908	0.898	0.918	0.945	0.976	0.983	0.980	0.998

Theil Inequality coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.511	0.521	0.486	0.501	0.462	0.398	0.282	0.238	0.258	0.080
1.5	0.468	0.478	0.442	0.457	0.419	0.357	0.249	0.209	0.228	0.070
1	0.409	0.418	0.384	0.397	0.361	0.304	0.207	0.173	0.189	0.057
0.5	0.313	0.321	0.290	0.303	0.273	0.225	0.150	0.125	0.137	0.041
0.25	0.232	0.239	0.215	0.225	0.201	0.164	0.108	0.090	0.098	0.029
0.125	0.169	0.174	0.155	0.164	0.146	0.118	0.077	0.065	0.071	0.021

P-NP filtering, Monte Carlo results; sample size 100 observations

Correlation coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.360	0.315	0.312	0.332	0.351	0.468	0.492	0.443	0.636	0.780
1.5	0.418	0.372	0.371	0.393	0.422	0.543	0.564	0.509	0.703	0.824
1	0.514	0.477	0.472	0.495	0.522	0.631	0.656	0.609	0.781	0.876
0.5	0.683	0.646	0.638	0.662	0.685	0.777	0.794	0.758	0.876	0.934
0.25	0.809	0.783	0.775	0.793	0.813	0.876	0.885	0.861	0.933	0.966
0.125	0.892	0.876	0.872	0.885	0.896	0.934	0.938	0.926	0.965	0.982

Theil Inequality coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.502	0.533	0.535	0.521	0.507	0.436	0.420	0.451	0.337	0.251
1.5	0.463	0.493	0.493	0.480	0.463	0.392	0.379	0.410	0.298	0.222
1	0.404	0.430	0.433	0.418	0.403	0.339	0.325	0.352	0.250	0.183
0.5	0308	0.332	0.335	0.322	0.307	0.253	0.241	0.264	0.183	0.132
0.25	0.230	0.249	0.252	0.241	0.228	0.183	0.176	0.194	0.132	0.094
0.125	0.169	0.182	0.185	0.176	0.167	0.132	0.127	0.140	0.094	0.068

The table reports the root mean square forecast error ($RMSFE$) and Theil inequality coefficient (IC) for the P-NP filtering approach.

Table 2
P-NP filtering, Monte Carlo results, biased estimation; sample size 1,000 observations

Correlation coefficient

	$b = 0.10$				$b = 0.20$				$b = 0.30$			
$\sigma \backslash d$	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
2	0.352	0.451	0.615	0.845	0.352	0.444	0.591	0.839	0.351	0.468	0.678	0.847
1.5	0.421	0.522	0.680	0.878	0.421	0.515	0.660	0.874	0.418	0.537	0.737	0.881
1	0.520	0.622	0.760	0.915	0.519	0.615	0.744	0.911	0.522	0.637	0.808	0.916
0.5	0.683	0.764	0.864	0.954	0.683	0.759	0.850	0.953	0.684	0.777	0.893	0.956
0.25	0.810	0.866	0.927	0.975	0.811	0.861	0.916	0.974	0.811	0.875	0.942	0.977
0.125	0.893	0.927	0.961	0.986	0.894	0.924	0.953	0.985	0.895	0.933	0.969	0.988

Theil Inequality coefficient

	$b = 0.10$				$b = 0.20$				$b = 0.30$			
$\sigma \backslash d$	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
2	0.504	0.443	0.348	0.205	0.504	0.446	0.359	0.209	0.505	0.433	0.310	0.204
1.5	0.460	0.401	0.310	0.180	0.461	0.404	0.321	0.183	0.462	0.392	0.275	0.178
1	0.401	0.343	0.262	0.149	0.402	0.347	0.271	0.152	0.401	0.335	0.231	0.148
0.5	0.307	0.259	0.191	0.108	0.308	0.261	0.201	0.110	0.307	0.251	0.168	0.106
0.25	0.228	0.189	0.138	0.080	0.228	0.192	0.147	0.081	0.228	0.183	0.122	0.077
0.125	0.167	0.137	0.099	0.060	0.167	0.140	0.109	0.060	0.166	0.132	0.088	0.056

P-NP filtering, Monte Carlo results, biased estimation; sample size 100 observations

Correlation coefficient

	$b = 0.10$				$b = 0.20$				$b = 0.30$			
$\sigma \backslash d$	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
2	0.285	0.470	0.569	0.584	0.295	0.371	0.517	0.775	0.320	0.376	0.462	0.605
1.5	0.345	0.529	0.643	0.659	0.351	0.433	0.598	0.819	0.390	0.428	0.525	0.673
1	0.447	0.633	0.727	0.746	0.449	0.534	0.683	0.873	0.488	0.533	0.638	0.747
0.5	0.610	0.769	0.841	0.849	0.614	0.701	0.812	0.932	0.654	0.698	0.773	0.855
0.25	0.757	0.872	0.913	0.920	0.765	0.822	0.897	0.964	0.792	0.823	0.869	0.921
0.125	0.858	0.929	0.954	0.958	0.866	0.902	0.946	0.981	0.882	0.901	0.931	0.958

Theil Inequality coefficient

	$b = 0.10$				$b = 0.20$				$b = 0.30$			
$\sigma \backslash d$	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8
2	0.552	0.434	0.375	0.366	0.546	0.494	0.406	0.254	0.526	0.494	0.439	0.355
1.5	0.512	0.397	0.333	0.324	0.505	0.449	0.360	0.224	0.483	0.457	0.400	0.316
1	0.448	0.338	0.283	0.273	0.445	0.393	0.310	0.185	0.422	0.395	0.337	0.270
0.5	0.351	0.256	0.209	0.203	0.348	0.299	0.229	0.133	0.325	0.302	0.255	0.199
0.25	0.263	0.186	0.151	0.145	0.259	0.221	0.166	0.096	0.243	0.222	0.188	0.144
0.125	0.196	0.136	0.109	0.105	0.190	0.162	0.118	0.069	0.178	0.163	0.134	0.105

The table reports the root mean square forecast error ($RMSFE$) and Theil inequality coefficient (IC) for the P-NP filtering approach.

Table 3

PCA analysis, Monte Carlo results: $n = 2$, $k = 1$; sample size 1,000 observations

Correlation coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.512	0.479	0.531	0.582	0.596	0.734	0.766	0.940	0.925	0.985
1.5	0.585	0.551	0.602	0.651	0.661	0.787	0.814	0.954	0.943	0.989
1	0.677	0.649	0.694	0.736	0.746	0.847	0.867	0.969	0.961	0.993
0.5	0.809	0.787	0.819	0.848	0.854	0.917	0.929	0.984	0.980	0.996
0.25	0.894	0.881	0.901	0.918	0.921	0.957	0.963	0.992	0.990	0.998
0.125	0.944	0.936	0.948	0.957	0.959	0.978	0.981	0.996	0.995	0.998

Theil Inequality coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.484	0.503	0.473	0.445	0.437	0.357	0.339	0.222	0.234	0.185
1.5	0.443	0.462	0.433	0.406	0.399	0.326	0.309	0.211	0.220	0.182
1	0.390	0.407	0.380	0.356	0.350	0.288	0.275	0.199	0.206	0.179
0.5	0.312	0.326	0.306	0.287	0.283	0.240	0.231	0.186	0.190	0.175
0.25	0.256	0.265	0.251	0.239	0.237	0.209	0.204	0.179	0.181	0.173
0.125	0.219	0.225	0.216	0.209	0.207	0.192	0.189	0.175	0.176	0.172

PCA analysis, Monte Carlo results: $n = 2$, $k = 1$; sample size 100 observations

Correlation coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.462	0.465	0.409	0.553	0.483	0.643	0.723	0.691	0.704	0.898
1.5	0.543	0.539	0.485	0.628	0.558	0.707	0.777	0.750	0.764	0.922
1	0.643	0.634	0.594	0.715	0.654	0.784	0.839	0.818	0.829	0.947
0.5	0.784	0.776	0.746	0.834	0.794	0.879	0.913	0.900	0.905	0.973
0.25	0.879	0.876	0.853	0.912	0.884	0.936	0.954	0.948	0.950	0.986
0.125	0.935	0.934	0.921	0.953	0.940	0.967	0.976	0.973	0.974	0.993

Theil Inequality coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.514	0.514	0.546	0.461	0.503	0.411	0.365	0.384	0.377	0.253
1.5	0.469	0.472	0.501	0.421	0.459	0.376	0.334	0.349	0.343	0.237
1	0.413	0.417	0.442	0.370	0.405	0.330	0.294	0.307	0.302	0.217
0.5	0.331	0.333	0.353	0.297	0.322	0.267	0.244	0.253	0.248	0.196
0.25	0.268	0.270	0.285	0.244	0.264	0.226	0.212	0.217	0.216	0.184
0.125	0.227	0.228	0.238	0.212	0.223	0.202	0.193	0.196	0.195	0.178

The table reports the root mean square forecast error ($RMSFE$) and Theil inequality coefficient (IC) for the PCA approach.

Table 4

PCA analysis, Monte Carlo results: $n = 3$, $k = 1$; sample size 1,000 observations

Correlation coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.600	0.611	0.618	0.643	0.714	0.741	0.803	0.933	0.988	0.981
1.5	0.667	0.678	0.682	0.706	0.769	0.793	0.844	0.949	0.991	0.985
1	0.751	0.759	0.763	0.782	0.833	0.852	0.891	0.965	0.994	0.990
0.5	0.857	0.864	0.866	0.878	0.909	0.920	0.942	0.982	0.997	0.995
0.25	0.923	0.927	0.928	0.935	0.952	0.958	0.970	0.991	0.998	0.998
0.125	0.960	0.962	0.963	0.966	0.976	0.979	0.985	0.996	0.999	0.999

Theil Inequality coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.492	0.486	0.483	0.470	0.432	0.418	0.384	0.311	0.276	0.281
1.5	0.457	0.452	0.449	0.437	0.403	0.390	0.362	0.301	0.274	0.278
1	0.413	0.408	0.406	0.396	0.368	0.358	0.335	0.290	0.272	0.274
0.5	0.354	0.351	0.349	0.343	0.325	0.319	0.305	0.280	0.270	0.271
0.25	0.316	0.314	0.313	0.309	0.298	0.295	0.287	0.274	0.269	0.270
0.125	0.294	0.292	0.292	0.290	0.284	0.282	0.278	0.271	0.268	0.269

PCA analysis, Monte Carlo results: $n = 3$, $k = 1$; sample size 100 observations

Correlation coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.590	0.598	0.581	0.637	0.630	0.775	0.784	0.800	0.864	0.764
1.5	0.663	0.668	0.643	0.698	0.697	0.821	0.824	0.841	0.895	0.814
1	0.747	0.752	0.734	0.780	0.776	0.874	0.880	0.890	0.928	0.867
0.5	0.855	0.861	0.847	0.876	0.873	0.933	0.936	0.941	0.962	0.930
0.25	0.920	0.924	0.918	0.935	0.933	0.966	0.967	0.970	0.981	0.963
0.125	0.959	0.961	0.957	0.966	0.965	0.983	0.983	0.985	0.990	0.981

Theil Inequality coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.499	0.493	0.506	0.474	0.479	0.401	0.396	0.387	0.352	0.406
1.5	0.463	0.459	0.470	0.441	0.443	0.375	0.371	0.366	0.334	0.381
1	0.417	0.413	0.423	0.397	0.401	0.345	0.342	0.337	0.314	0.348
0.5	0.357	0.354	0.362	0.344	0.346	0.311	0.309	0.306	0.292	0.312
0.25	0.318	0.315	0.320	0.310	0.311	0.290	0.290	0.288	0.281	0.292
0.125	0.294	0.294	0.296	0.290	0.291	0.280	0.279	0.278	0.275	0.281

The table reports the root mean square forecast error ($RMSFE$) and Theil inequality coefficient (IC) for the PCA approach.

Table 5

PCA analysis, Monte Carlo results: $n = 4$, $k = 1$; sample size 1,000 observations

Correlation coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.686	0.678	0.696	0.700	0.761	0.837	0.918	0.965	0.982	0.991
1.5	0.745	0.738	0.754	0.757	0.809	0.873	0.938	0.974	0.987	0.993
1	0.813	0.808	0.821	0.823	0.864	0.912	0.957	0.982	0.991	0.995
0.5	0.897	0.894	0.901	0.903	0.927	0.954	0.978	0.991	0.995	0.998
0.25	0.946	0.944	0.948	0.949	0.962	0.976	0.989	0.996	0.998	0.999
0.125	0.972	0.971	0.974	0.974	0.981	0.988	0.994	0.998	0.999	0.999

Theil Inequality Coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.493	0.497	0.488	0.486	0.456	0.419	0.377	0.352	0.343	0.339
1.5	0.465	0.468	0.460	0.459	0.432	0.401	0.367	0.348	0.341	0.337
1	0.430	0.433	0.427	0.426	0.405	0.381	0.357	0.343	0.338	0.336
0.5	0.388	0.390	0.386	0.385	0.373	0.359	0.345	0.338	0.336	0.335
0.25	0.363	0.364	0.362	0.361	0.354	0.346	0.339	0.336	0.335	0.334
0.125	0.349	0.349	0.348	0.348	0.344	0.340	0.336	0.335	0.334	0.334

PCA analysis, Monte Carlo results: $n = 4$, $k = 1$; sample size 100 observations

Correlation coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.685	0.605	0.691	0.667	0.642	0.753	0.737	0.728	0.907	0.896
1.5	0.744	0.675	0.751	0.726	0.707	0.803	0.792	0.778	0.929	0.921
1	0.816	0.760	0.817	0.801	0.787	0.859	0.849	0.843	0.951	0.947
0.5	0.899	0.864	0.901	0.889	0.880	0.925	0.919	0.916	0.975	0.973
0.25	0.947	0.927	0.948	0.941	0.937	0.961	0.958	0.956	0.987	0.986
0.125	0.973	0.962	0.973	0.970	0.968	0.980	0.978	0.978	0.994	0.993

Theil Inequality coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.495	0.532	0.492	0.505	0.517	0.462	0.469	0.474	0.384	0.389
1.5	0.465	0.499	0.463	0.476	0.484	0.436	0.44	0.447	0.373	0.375
1	0.431	0.459	0.430	0.438	0.444	0.407	0.414	0.418	0.360	0.363
0.5	0.389	0.406	0.388	0.392	0.397	0.374	0.377	0.379	0.347	0.348
0.25	0.362	0.373	0.363	0.366	0.368	0.355	0.357	0.357	0.341	0.341
0.125	0.349	0.355	0.349	0.350	0.352	0.345	0.345	0.346	0.337	0.337

The table reports the root mean square forecast error ($RMSFE$) and Theil inequality coefficient (IC) for the PCA approach.

Table 6

PCA analysis, Monte Carlo results: $n = 4$, $k = 2$ (only the first factor is estimated); sample size 1,000 observations

Correlation coefficient

$\sigma \backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.525	0.533	0.511	0.553	0.624	0.705	0.800	0.930	0.951	0.958
1.5	0.595	0.603	0.584	0.621	0.690	0.761	0.842	0.945	0.957	0.965
1	0.688	0.696	0.677	0.711	0.769	0.826	0.889	0.960	0.962	0.972
0.5	0.815	0.821	0.807	0.832	0870	0.905	0.941	0.975	0.968	0.980
0.25	0.898	0.902	0.893	0.907	0.930	0.950	0.969	0.983	0.971	0.984
0.125	0.947	0.948	0.944	0.951	0.964	0.974	0.984	0.987	0.973	0.986

Theil Inequality coefficient

$\sigma \backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.477	0.472	0.485	0.461	0.420	0.375	0.317	0.191	0.254	0.218
1.5	0.437	0.431	0.444	0.422	0.383	0.342	0.291	0.179	0.250	0.212
1	0.384	0.378	0.391	0.371	0.337	0.301	0.259	0.165	0.246	0.206
0.5	0.308	0.304	0.313	0.298	0.273	0.249	0.221	0.149	0.242	0.200
0.25	0.253	0.250	0.257	0.247	0.230	0.215	0.199	0.141	0.240	0.197
0.125	0.218	0.216	0.220	0.214	0.204	0.195	0.186	0.136	0.239	0.195

PCA analysis, Monte Carlo results: $n = 4$, $k = 2$ (only the first factor is estimated); sample size 100 observations

Correlation coefficient

$\sigma \backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.422	0.555	0.452	0.436	0.586	0.589	0.616	0.727	0.831	0.892
1.5	0.500	0.627	0.527	0.518	0.659	0.663	0.686	0.775	0.869	0.914
1	0.609	0.719	0.633	0.630	0.747	0.747	0.761	0.830	0.910	0.938
0.5	0.764	0.838	0.784	0.780	0.854	0.857	0.855	0.893	0.952	0.963
0.25	0.868	0.914	0.879	0.877	0.920	0.923	0.911	0.927	0.975	0.976
0.125	0.929	0.954	0.935	0.935	0.955	0.958	0.942	0.945	0.988	0.982

Theil Inequality coefficient

$\sigma \backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.532	0.454	0.511	0.520	0.427	0.432	0.432	0.310	0.296	0.266
1.5	0.488	0.413	0.467	0.474	0.386	0.392	0.393	0.279	0.271	0.251
1	0.426	0.361	0.408	0.413	0.337	0.343	0.350	0.242	0.244	0.233
0.5	0.340	0.289	0.324	0.329	0.269	0.276	0.291	0.195	0.212	0.213
0.25	0.275	0.238	0.263	0.267	0.223	0.229	0.252	0.164	0.192	0.203
0.125	0.232	0.208	0.223	0.226	0.194	0.201	0.230	0.147	0.182	0.197

The table reports the root mean square forecast error ($RMSFE$) and Theil inequality coefficient (IC) for the PCA approach.

Table 7

PCA analysis, Monte Carlo results: $n = 4$, $k = 2$; sample size 1,000 observations

Correlation coefficient										
$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.358	0.359	0.350	0.398	0.450	0.508	0.734	0.648	0.675	0.801
1.5	0.419	0.419	0.413	0.463	0.515	0.570	0.785	0.673	0.672	0.816
1	0.507	0.509	0.503	0.555	0.602	0.652	0.845	0.701	0.677	0.833
0.5	0.655	0.655	0.651	0.701	0.738	0.775	0.915	0.735	0.692	0.847
0.25	0.781	0.779	0.778	0.816	0.842	0.867	0.955	0.754	0.699	0.856
0.125	0.872	0.870	0.870	0.894	0.911	0.925	0.977	0.765	0.697	0.862
Theil Inequality coefficient										
$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.530	0.527	0.536	0.502	0.469	0.429	0.395	0.334	0.372	0.283
1.5	0.487	0.486	0.492	0.461	0.430	0.394	0.385	0.324	0.377	0.276
1	0.432	0.432	0.434	0.406	0.378	0.347	0.378	0.314	0.376	0.269
0.5	0.350	0.347	0.352	0.327	0.306	0.284	0.391	0.302	0.370	0.265
0.25	0.285	0.284	0.286	0.267	0.256	0.239	0.397	0.296	0.367	0.261
0.125	0.341	0.238	0.248	0.228	0.220	0.210	0.427	0.293	0.369	0.258

PCA analysis, Monte Carlo results: $n = 4$, $k = 2$; sample size 100 observations

Correlation coefficient										
$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.293	0.335	0.302	0.295	0.364	0.387	0.444	0.474	0.582	0.714
1.5	0.362	0.396	0.372	0.368	0.428	0.458	0.511	0.520	0.641	0.746
1	0.471	0.482	0.471	0.475	0.519	0.548	0.596	0.573	0.715	0.785
0.5	0.641	0.633	0.642	0.647	0.654	0.696	0.706	0.643	0.821	0.828
0.25	0.780	0.758	0.776	0.785	0.759	0.806	0.772	0.686	0.895	0.853
0.125	0.871	0.849	0.867	0.876	0.830	0.879	0.819	0.711	0.939	0.866
Theil Inequality coefficient										
$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.581	0.559	0.573	0.578	0.525	0.519	0.477	0.421	0.396	0.321
1.5	0.534	0.521	0.523	0.531	0.483	0.472	0.436	0.395	0.362	0.305
1	0.467	0.463	0.465	0.467	0.423	0.415	0.385	0.369	0.312	0.287
0.5	0.376	0.363	0.367	0.370	0.340	0.326	0.329	0.338	0.258	0.269
0.25	0.303	0.290	0.293	0.290	0.289	0.269	0.298	0.319	0.224	0.258
0.125	0.251	0.247	0.246	0.248	0.258	0.235	0.277	0.308	0.203	0.253

The table reports the root mean square forecast error ($RMSFE$) and Theil inequality coefficient (IC) for the PCA approach.

Table 8

PCA analysis, Monte Carlo results: $n = 6$, $k = 2$; sample size 1,000 observations

Correlation coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.431	0.432	0.441	0.492	0.492	0.652	0.726	0.818	0.794	0.887
1.5	0.496	0.496	0.505	0.558	0.553	0.704	0.750	0.852	0.803	0.900
1	0.583	0.587	0.595	0.645	0.637	0.770	0.784	0.890	0.813	0.914
0.5	0.722	0.727	0.733	0.776	0.761	0.854	0.816	0.934	0.824	0.927
0.25	0.832	0.836	0.840	0.869	0.852	0.908	0.843	0.958	0.829	0.935
0.125	0.905	0.908	0.910	0.929	0.911	0.939	0.848	0.970	0.831	0.939

Theil Inequality coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.518	0.516	0.511	0.479	0.477	0.384	0.346	0.295	0.290	0.262
1.5	0.479	0.477	0.472	0.441	0.441	0.356	0.335	0.277	0.285	0.256
1	0.428	0.425	0.421	0.392	0.394	0.320	0.320	0.255	0.280	0.248
0.5	0.351	0.348	0.345	0.321	0.326	0.275	0.307	0.228	0.275	0.241
0.25	0.290	0.290	0.286	0.269	0.277	0.244	0.295	0.211	0.272	0.236
0.125	0.247	0.245	0.244	0.231	0.242	0.225	0.295	0.202	0.271	0.234

PCA analysis, Monte Carlo results: $n = 6$, $k = 2$; sample size 100 observations

Correlation coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.383	0.403	0.432	0.417	0.461	0.370	0.549	0.500	0.704	0.718
1.5	0.451	0.484	0.499	0.495	0.529	0.468	0.611	0.566	0.754	0.748
1	0.550	0.584	0.607	0.594	0.613	0.579	0.689	0.663	0.812	0.781
0.5	0.703	0.739	0.750	0.740	0.732	0.750	0.794	0.787	0.883	0.822
0.25	0.820	0.845	0.854	0.846	0.817	0.858	0.867	0.872	0.923	0.845
0.125	0.897	0.916	0.917	0.915	0.871	0.923	0.907	0.927	0.947	0.859

Theil Inequality coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.570	0.543	0.528	0.543	0.516	0.549	0.455	0.491	0.360	0.347
1.5	0.525	0.493	0.486	0.494	0.464	0.494	0.415	0.443	0.334	0.333
1	0.455	0.431	0.413	0.424	0.411	0.431	0.370	0.385	0.303	0.318
0.5	0.361	0.343	0.335	0.342	0.345	0.346	0.312	0.316	0.264	0.301
0.25	0.297	0.284	0.277	0.282	0.300	0.288	0.273	0.268	0.239	0.291
0.125	0.252	0.241	0.239	0.241	0.273	0.255	0.250	0.233	0.224	0.286

The table reports the root mean square forecast error ($RMSFE$) and Theil inequality coefficient (IC) for the PCA approach.

Table 9

PCA analysis, Monte Carlo results: $n = 8$, $k = 2$; sample size 1,000 observations

Correlation coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.497	0.486	0.512	0.557	0.588	0.700	0.721	0.826	0.912	0.886
1.5	0.563	0.550	0.578	0.619	0.646	0.745	0.768	0.860	0.927	0.895
1	0.650	0.639	0.663	0.699	0.723	0.800	0.824	0.897	0.944	0.904
0.5	0.780	0.768	0.790	0.813	0.826	0.870	0.891	0.939	0.960	0.914
0.25	0.872	0.864	0.878	0.893	0.897	0.914	0.930	0.961	0.969	0.918
0.125	0.930	0.925	0.933	0.941	0.939	0.938	0.951	0.973	0.973	0.921

Theil Inequality coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.511	0.518	0.502	0.478	0.458	0.393	0.384	0.325	0.273	0.281
1.5	0.474	0.481	0.465	0.444	0.425	0.368	0.358	0.304	0.263	0.276
1	0.425	0.432	0.417	0.399	0.383	0.337	0.327	0.280	0.251	0.270
0.5	0.354	0.361	0.348	0.335	0.325	0.297	0.287	0.251	0.237	0.265
0.25	0.300	0.305	0.296	0.287	0.282	0.270	0.260	0.232	0.229	0.262
0.125	0.260	0.264	0.258	0.252	0.252	0.252	0.244	0.221	0.224	0.260

PCA analysis, Monte Carlo results: $n = 8$, $k = 2$; sample size 100 observations

Correlation coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.457	0.432	0.463	0.490	0.439	0.475	0.606	0.500	0.650	0.694
1.5	0.530	0.503	0.537	0.564	0.494	0.549	0.672	0.543	0.696	0.754
1	0.636	0.609	0.637	0.656	0.565	0.637	0.749	0.595	0.747	0.823
0.5	0.775	0.753	0.770	0.784	0.657	0.757	0.848	0.661	0.821	0.904
0.25	0.870	0.854	0.865	0.871	0.713	0.834	0.913	0.703	0.869	0.950
0.125	0.930	0.919	0.924	0.924	0.474	0.882	0.951	0.729	0.896	0.974

Theil Inequality coefficient

$\sigma\backslash d$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2	0.539	0.564	0.538	0.529	0.488	0.523	0.454	0.498	0.423	0.398
1.5	0.494	0.515	0.492	0.480	0.455	0.479	0.418	0.472	0.397	0.365
1	0.433	0.452	0.434	0.425	0.418	0.428	0.374	0.443	0.367	0.327
0.5	0.357	0.372	0.360	0.353	0.370	0.364	0.316	0.413	0.327	0.276
0.25	0.301	0.312	0.305	0.302	0.342	0.321	0.275	0.395	0.301	0.242
0.125	0.260	0.270	0.265	0.266	0.325	0.294	0.245	0.385	0.285	0.218

The table reports the root mean square forecast error ($RMSFE$) and Theil inequality coefficient (IC) for the PCA approach.

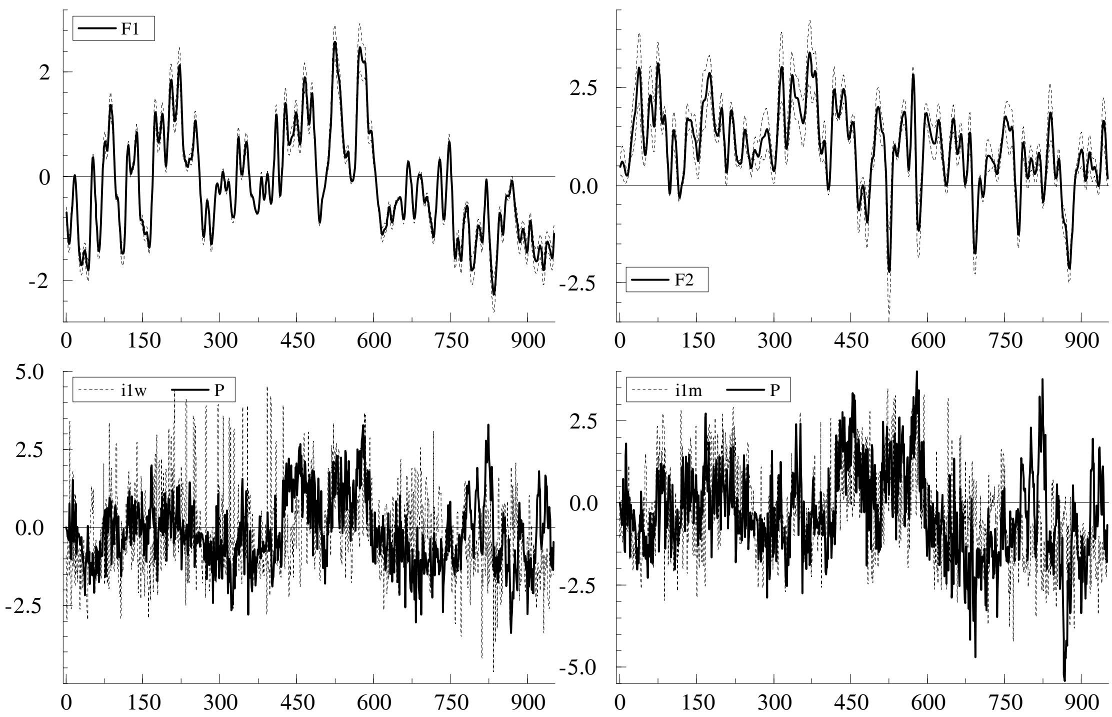

Figure 1: Persistent factors with 95% confidence bounds (top plots); selected persistent components (one week and one month rates) with actual series (bottom plots).