July 2021

Working Paper

019.2021

Permanent-Transitory decomposition of cointegrated time series via Dynamic Factor Models, with an application to commodity prices

Chiara Casoli, Riccardo (Jack) Lucchetti

Permanent-Transitory decomposition of cointegrated time series via Dynamic Factor Models, with an application to commodity prices

By Chiara Casoli, Fondazione Eni Enrico Mattei
Riccardo (Jack) Lucchetti, Università Politecnica delle Marche

Summary

In this article, we propose a cointegration-based Permanent-Transitory decomposition for non-stationary Dynamic Factor Models. Our methodology exploits the cointegration relations among the observable variables and assumes they are driven by a common and an idiosyncratic component. The common component is further split into a long-term non-stationary part and a short-term stationary one. A Monte Carlo experiment shows that taking into account the cointegration structure in the DFM leads to a much better reconstruction of the space spanned by the factors, with respect to the most standard technique of applying a factor model in differenced systems. Finally, an application of our procedure to a set of different commodity prices allows to analyse the comovement among different markets. We find that commodity prices move together due to *long-term* common forces and that the trend for most primary good prices is declining, whereas metals and energy ones exhibit an upward or at least stable pattern since the 2000s.

Keywords: Cointegration, Dynamic Factor Models, P-T decomposition, Commodity prices co-movement

JEL Classification: C32, C38, Q02

Address for correspondence:
Chiara Casoli
Econometrics of the Energy Transition Programme
Fondazione Eni Enrico Mattei
Corso Magenta 63
20123 Milano
Italy
E-mail: chiara.casoli@feem.it

Permanent-Transitory decomposition of cointegrated time series via Dynamic Factor Models, with an application to commodity prices*

Chiara Casoli† Riccardo (Jack) Lucchetti‡

July 9, 2021

Abstract

In this article, we propose a cointegration-based Permanent-Transitory decomposition for non-stationary Dynamic Factor Models. Our methodology exploits the cointegration relations among the observable variables and assumes they are driven by a common and an idiosyncratic component. The common component is further split into a long-term non-stationary part and a short-term stationary one. A Monte Carlo experiment shows that taking into account the cointegration structure in the DFM leads to a much better reconstruction of the space spanned by the factors, with respect to the most standard technique of applying a factor model in differenced systems. Finally, an application of our procedure to a set of different commodity prices allows to analyse the comovement among different markets. We find that commodity prices move together due to *long-term* common forces and that the trend for most primary good prices is declining, whereas metals and energy ones exhibit an upward or at least stable pattern since the 2000s.

Keywords: Cointegration, Dynamic Factor Models, P-T decomposition, Commodity prices co-movement.

JEL codes: C32, C38, Q02.

1 Introduction

Dynamic Factor Models (DFMs) are an increasingly popular tool for summarising information of a large number of time series into a smaller number of factors. Essentially, the vector of variables is split into a common component, capturing the joint movement of all the observable series, and an idiosyncratic component, which is variable-specific. Although DFMs are now a standard tool in applied macroeconomics and finance (extensive surveys can be found in Bai and Ng (2008); Stock and Watson (2011, 2016); Doz and Fuleky (2020)), it is only recently that the issues of non-stationarity and cointegration have begun to receive systematic attention in the literature.

*We wish to thank Matteo Barigozzi, Tomás del Barrio Castro, Stefano Fachin, Carlo Favero, Søren Johansen and Marco Lippi for their useful comments. Needless to say, none of them bear any responsibility for any errors, which are all ours.

†Fondazione Eni Enrico Mattei, Milano (Italy), chiara.casoli@feem.it

‡Univesità Politecnica delle Marche, Ancona (Italy), r.lucchetti@univpm.it

The most obvious way of dealing with $I(1)$ systems is to difference the whole set of variables and to estimate the model in first differences. Eventually, as proposed in Bai and Ng (2004), $I(1)$ common factors can be recovered by integration of the differenced extracted factors; for this reason, this routine is known as "differencing and recumulating" (DR). However, if variables of the system are in fact cointegrated, taking first differences drops the long-term information possibly contained in the data.

Therefore, other approaches avoid the differentiation step and deal directly with the original data in levels. Bai (2004) proposes a Principal Components (PC) estimation procedure for $I(1)$ systems on levels, thus allowing for the direct estimation of non-stationary common factors. A possible drawback of this procedure is that the idiosyncratic component is assumed to be stationary, which is tantamount to saying that all the non-stationarity of the system is captured by the common component. This is a very strong assumption, since in many examples the variable-specific components can reasonably thought to be non-stationary. Barigozzi et al. (2020) stress that in standard datasets used within DFM literature the idiosyncratic term is most likely $I(1)$.

On the contrary, by using the DR approach non-stationarity is not necessarily captured by the common component, but may be specific to individual series. Corona et al. (2020) extend the hybrid method proposed by Doz et al. (2011) to non-stationary cases and compare the performance of the PC methodology for factors extracted using non-stationary system in levels with those obtained with the DR approach.

Barigozzi et al. (2016, 2020, 2021) make important contributions by extending the DFM framework including cointegration, where factors are assumed to admit a VECM representation and idiosyncratic components are allowed either to be $I(0)$ or $I(1)$. Corona et al. (2020) point out that the estimators proposed by Bai and Ng (2004); Barigozzi et al. (2016) are asymptotically equivalent, with some finite sample differences if deterministic trends are included in the model. Finally, Barigozzi and Luciani (2019) propose a DFM in which a Trend-Cycle decomposition is performed from the extracted factors. Specifically, they first estimate a non-stationary DFM by Quasi-Maximum Likelihood through the EM algorithm, and then they assume factors, which are cointegrated, are driven by a non-stationary long-term component and a $I(0)$ short-term component.

In this paper, we develop a method similar in spirit to Barigozzi and Luciani (2019); contrary to previous proposals, however, we include the cointegration structure of data into a Dynamic Factor Model by using the fact that in many cases of practical interest some information on the cointegration properties of the observable variables is available *prior* to setting up the DFM. In these cases, it is possible to transform the observables as a first step via a Permanent-Transitory (P-T) decomposition and then operate on the transformed variables. As the Monte Carlo evidence presented in Section 3 suggests, this leads to sizeable improvements in reconstructing the factor space compared to the DR approach.

The rest of the paper has the following structure: Section 2 provides a quick and general introduction to DFMs, mainly to establish notation, followed by a description of the P-T decomposition we use. Some Monte Carlo evidence is provided in Section 3; finally, in Section 4, we apply our procedure to investigate the common movement of commodity prices and find that the series are, as a rule, mainly driven by their long-term common components; support for the Prebisch-Singer hypothesis is found for primary commodities. Section 5 concludes the paper.

2 Econometric methods

2.1 Dynamic factor models

Dynamic factor models were introduced by Geweke (1977); Sargent and Sims (1977), but widespread adoption in the empirical literature has started since the beginning of this century. The general setup of DFMs that we consider here can be described by the following two equations:

$$Y_t = \Lambda_0 f_t + \Lambda_1 f_{t-1} + ... + \Lambda_s f_{t-s} + \varepsilon_t, \tag{1}$$

$$f_t = A_1 f_{t-1} + A_2 f_{t-2} + ... + A_p f_{t-p} + u_t., \tag{2}$$

where:

- Y_t is a $n \times 1$ vector of time series observable variables;
- f_t is a $q \times 1$ vector of the common latent dynamic factors;
- ε_t is a $n \times 1$ vector containing the idiosyncratic terms;
- u_t is the vector of the dynamic factors shocks; ε_t and u_t are independent;
- $\Lambda_0, ..., \Lambda_s$ are the $n \times q$ factor loading matrices;
- $A_1, ..., A_p$ are the $q \times q$ matrices containing the VAR parameters of the unobserved factors.

In the jargon of state-space models, Equation (1) is the *observation equation* and describes Y_t as the sum of a common component and an idiosyncratic component, which is variable-specific. In the special case $s = 0$, the DFM is called *static*. Note, however that by defining

$$F_t = [f'_t, f'_{t-1}, \ldots, f'_{t-s}] \qquad \mathbf{\Lambda} = [\Lambda_0, \Lambda_1, \ldots, \Lambda_s]$$

the dynamic model also admits a static representation as

$$Y_t = \mathbf{\Lambda} F_t + \varepsilon_t.$$

Equation (2) is the *state equation* and expresses the dynamics of the q latent common factors.

The models we consider here belong to a class known as *Dynamic Approximate Factor Models*, because no strict assumptions are made on the distribution of the idiosyncratic term ε_t, which is allowed to have some mild form of correlation: they are assumed to be uncorrelated with the factors at all leads and lags but are allowed to be correlated either serially and contemporaneously. Moreover, we take the scalar s to be finite, which sets this model apart from the so-called *Generalised DFMs*, that are typically handled via spectral methods (see for example Forni et al. (2000, 2005, 2015, 2017); Hallin and Liška (2007)).

For the estimation of Equations (1) and (2) several techniques have been proposed, under the assumption of stationarity. In the static case ($s = 0$) a simple and computationally convenient method is PC, which is the standard choice in most empirical applications. As mentioned, the DR procedure circumvents the problem of nonstationarity of Y_t by taking first differences. In this way, factors are estimated in first differences and then recovered by cumulation of $\Delta\hat{F}_t$. Bai and Ng (2004) demonstrate that the first difference of factors and the model parameters can be estimated consistently.

Other estimation techniques have also been proposed, mainly with the aim of achieving higher efficiency than PC, notably Doz et al. (2011, 2012). Generally, these methods involve Maximum Likelihood estimation, achieved by the Kalman filtering; the advantage is that the same state-space setup can be used for retrieving estimates of the factors f_t by the smoothing algorithm that are in some cases significantly more efficient than PC (see *eg* Lucchetti and Venetis, 2020); however, the increase in computational complexity can sometimes be considerable. In this paper, we will use both approaches: we use PC for the simulation study in Section 3 and ML estimation via the EM algorithm in Section 4.

2.2 A Permanent-Transitory decomposition

We assume that the vector of observable variables has the persistence feature of a multivariate $I(1)$ process:

$$Y_t = \mathcal{T}_t + \mathcal{C}_t + \xi_t,$$

in which the common component is the sum of $\mathcal{T}_t + \mathcal{C}_t$, where the trend component $\mathcal{T}_t$ is $I(1)$ while the cycle component $\mathcal{C}_t$ is stationary; both are assumed to have a factor structure, that is, to be driven by a small number of shocks q; the idiosyncratic component is given by ξ_t. Note that ξ_t may be $I(1)$. The possibility to include non-stationarity in the idiosyncratic term is not trivial, as it implies that it is feasible to allow for some non-stationarity to be variable-specific, rather than assuming that all of the $I(1)$ component is captured by the common movement. A similar set up is allowed in Barigozzi et al. (2020, 2021); Barigozzi and Luciani (2019). Assuming that Y_t can be represented as a VAR process of finite order p and that the cointegration rank is $0 \le r < n$, the VECM representation exists, and is given by:

$$\Delta Y_t = \alpha\beta' Y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + \epsilon_t. \tag{3}$$

For given values of the parameters, a P-T decomposition can be achieved by re-expressing the original observable variables Y_t as an invertible linear transformation. The most popular one is the so-called Gonzalo-Granger decomposition (Gonzalo and Granger, 1995), in which Y_t is pre-multiplied by $[\beta \quad \alpha_\perp]'$;[1] in this paper, however, we use the decomposition put forward in Kasa (1992), which is based on $\beta_\perp$. We do so for three reasons: first, it is possible (albeit unlikely) that $[\beta \quad \alpha_\perp]'$ has not full rank, and our algorithm would break down. Second, it is often the case that β can be considered (at least partly) known a *a priori* on the grounds on theoretical arguments; similar considerations hold much more seldom for α. Finally, even when β is estimated, its convergence rate is $O(T^{-1})$, while α (and hence $\alpha_\perp$) is only $\sqrt{T}$-consistent.

[1] We use the $\perp$ footer to indicate the orthogonal complement of a matrix, as is customary in the cointegration literature. If β is $n \times r$ with rank r, then $\beta_\perp$ is an $n \times (n-r)$ matrix such that $\beta'\beta_\perp = 0$.

The resulting decomposition is:[2]

$$G(L)Y_t = \left[\begin{array}{c} \beta' \\ \beta_\perp'(1-L) \end{array} \right] Y_t = \left[\begin{array}{c} z_t \\ \Delta m_t \end{array} \right] = W_t. \quad (4)$$

By hypothesis, $W_t \sim I(0)$. Note that we can define the inverse of the $G(L)$ filter as

$$G(L)^{-1} = \left[\begin{array}{cc} \beta(\beta'\beta)^{-1} & \beta_\perp(\beta_\perp{}'\beta_\perp)^{-1}\frac{1}{(1-L)} \end{array} \right]$$

where $\frac{1}{(1-L)}$ is the cumulation operator, and therefore write

$$Y_t = G(L)^{-1} W_t.$$

The representation linking observables to factors may be written as

$$W_t = \Lambda^*(L) f_t^* + e_t,$$

where $\Lambda^*(L)$ is a matrix polynomial of order s.

The space spanned by the factors f_t can be estimated in several ways (see Section 2.1). In the simulation experiment presented in Section 3, we will use the PC methods on the grounds of computational convenience and to conform to the method that is most widely used among practitioners; however, we also considered quasi-ML estimation of factors along the lines of Doz et al. (2012).

By partitioning the loading matrix Λ in an appropriate way

$$\left[\begin{array}{c} z_t \\ \Delta m_t \end{array} \right] = \left[\begin{array}{c} \Lambda_z \\ \Lambda_\Delta \end{array} \right] f_t + e_t$$

we have

$$Y_t = \left[\begin{array}{cc} \beta(\beta'\beta)^{-1} & \beta_\perp(\beta_\perp'\beta_\perp)^{-1}\frac{1}{(1-L)} \end{array} \right] \left\{ \left[\begin{array}{c} \Lambda_z \\ \Lambda_\Delta \end{array} \right] f_t + e_t \right\} = \mathcal{T}_t + \mathcal{C}_t + \xi_t$$

where

$$\begin{array}{rcl} \mathcal{C}_t & = & \beta(\beta'\beta)^{-1}\Lambda_z f_t \\ \mathcal{T}_t & = & \beta_\perp(\beta_\perp'\beta_\perp)^{-1}\Lambda_\Delta f_t^c \\ \xi_t & = & G(L)^{-1} e_t \end{array}$$

and f_t^c is the cumulation of f_t, that is, $f_t = \Delta f_t^c$. The $I(1)$ process f_t^c is a q-variate process whose first difference is the vector of $I(0)$ factors. Note that the DR approach is a special case of the decomposition outlined here, in which the cointegration rank is 0 and $\mathcal{C}_t = 0$. Our procedure makes it possible to decompose the common component in long- and short-term components in a very natural way.

Finally, an important point to stress here is that the β matrix used in Equation (4) does not necessarily have to span the full cointegration space of the system. The important requisite for the decomposition (4) to yield a vector of $I(0)$ variables is that

[2]Interestingly, a closely related idea is hinted at in (Bai and Ng, 2004, p. 1131), where the transformation is applied to the factors instead of the observables.

the β matrix spans a subset of the true cointegration space. Therefore, underestimating the actual cointegration rank is a lesser evil than overestimating it. If some of the cointegration vectors are left out of β, they will be present in the space spanned by its orthogonal complement $\beta_\perp$, for which the DR approach remains unchanged.

Moreover, in practical situations it is often the case that the cointegration rank (and sometimes the cointegration matrix) can be assumed to be known *a priori* on the grounds of economic theory (*eg* interest rates). That said, the numerical experiment we perform in Section 3 analyses both cases.

2.3 Cointegration analysis by blocks

The procedure described in Subsection 2.2 takes the cointegration structure of data into account and incorporates it into the DFM, so it is feasible when the cointegration matrix β is either known or it can be estimated consistently, which is hardly ever the case in practice.

For very large systems of variables, however, standard cointegration analysis becomes unfeasible; apart from the practical difficulty of setting up a VAR system when the number of variables n is even moderately large, it is well known that inference may be quite unreliable in finite samples. Cavaliere et al. (2012) suggest a comprehensive solution for the bootstrap implementation of the rank test, with Onatski and Wang (2018) providing a theoretical analysis of the relevant inferential issues. Similarly, Cavaliere et al. (2015) explore the properties of bootstrap-corrected hypothesis tests on the cointegration parameters.

However, for medium-sized problems, it is often possible to split the vector of observable variables into blocks and assume that cointegration only occurs within blocks, and not between. This is exactly the case we are studying. This makes it possible not only to properly include the information coming from the cointegration relations in the model, but also to disentangle the long-term common movement from the short-term one, captured by the Permanent and the Transitory components, respectively. Typically, the grouping of the observed series into blocks follows from *a priori* information; for example, in macroeconomic applications it may be perfectly legitimate to assume that a vector of interest rates contains only one common trend, with the obvious consequences on the cointegration rank; additionally, some of the cointegration vectors may be fixed *a priori* so as to imply that spreads are stationary. Similar considerations, with the necessary adaptations, may apply to subsets of macroeconomic series such as sectoral industrial production indices. In the empirical application contained in Section 4, we divide commodity prices into blocks on the basis of "natural" groupings (food, energy, etc.).

In formulae, we assume that Y_t can be divided into B different blocks (not necessarily of the same size):

$$Y_t' = \left[Y_{1,t}', \quad Y_{2,t}', \ldots Y_{B,t}'\right];$$

cointegration analysis may be then performed within each block. This would be equivalent to estimating the cointegration matrices through B different "partial systems" (see Johansen, 1992). The estimated system-wide β matrix would therefore be obtained by stacking diagonally the per-block cointegration matrices:

$$\beta = \begin{bmatrix} \beta_1 & & & \\ & \beta_2 & & \\ & & \ddots & \\ & & & \beta_B \end{bmatrix}$$

Of course, the overall estimated rank equals $\sum_{b=1}^{B} r_b$.

Bearing in mind that the cointegration rank should not be overestimated, but an underestimate would not undermine the procedure we are putting forward, rank determination can be done via the usual Johansen procedure, possibly with additional caveats in order to mistakenly identify spurious cointegration relationships. For example, in the simulation study in Section 3 and the empirical analysis in Section 4, we use as an estimate of the cointegration rank the smallest integer that leads to rejecting the Johansen trace test a 1% level instead of the more customary 5% level. Alternatively, Bartlett-type (see Johansen, 2000) or bootstrap correction (Cavaliere et al., 2012) may be used.

2.4 The workflow

In practice, a DFM model is applied to data that have been centred and standardised, so that the workflow goes as follows:

1. Estimate the matrix β on the original data Y_t (possibly, by blocks) and form its orthogonal complement $\beta_\perp$;

2. compute the Kasa-decomposed vector W_t as in Equation (4);

3. compute the vector of standard deviations σ so that $Z_t = \langle\sigma\rangle^{-1}\left[W_t - \bar{W}\right]$, where the notation $\langle x\rangle$ indicates a diagonal matrix that has x on its diagonal; note that $\langle\sigma\rangle$ can be written as

$$\langle\sigma\rangle = \begin{bmatrix} \langle\sigma_z\rangle & 0 \\ 0 & \langle\sigma_\Delta\rangle \end{bmatrix}$$

   in standard notation (again, see equation (4));

4. compute the factors in the DFM

$$Z_t = \Lambda f_t + e_t;$$

   and partition the loading matrix Λ as

$$\Lambda = \begin{bmatrix} \Lambda_z \\ \Lambda_\Delta, \end{bmatrix}$$

   where Λ_z has r rows and Λ_Δ has $(n-r)$;

5. recover the permament and transitory component of the factor structure as

$$\mathcal{T}_t = \beta_\perp(\beta_\perp'\beta_\perp)^{-1}\langle\sigma_\Delta\rangle\Lambda_\Delta f_t^c \quad (5)$$

$$\mathcal{C}_t = \beta(\beta'\beta)^{-1}\langle\sigma_z\rangle\Lambda_z f_t. \quad (6)$$

An alternative possibility to step 5 is to recover $\hat{z}_t$ and $\Delta\hat{m}_t$ via OLS projections from the extracted factors. At this point, obtaining the Permanent and Transitory components is straightforward:

$$\begin{bmatrix} \hat{Y}_t \\ \hat{Y}_t^c \end{bmatrix} = \begin{bmatrix} \beta & \beta_\perp \end{bmatrix}^{-1} \begin{bmatrix} \hat{z}_t \\ \hat{m}_t^c \end{bmatrix}.$$

3 Simulation results

In this Section, we present the result of a simulation study aimed at assessing the possible gain from explicitly considering cointegration among the observables along the lines described in Section 2. In order to do so, we generate several DGPs like (3) with the following structure:

1. The number of observed variables n ranges from 32 to 128;
2. the sample size T is either 200 or 400;
3. the VAR order is always 1;
4. the VAR innovations are generated with a factor structure

   $$\epsilon_t = \Lambda f_t + u_t$$

   where the number of system-wide factors q ranges from 1 to 8; Λ is generated as a conformable matrix of uniforms and u_t is a multivariate white noise; the standard deviation of each element of u_t is 0.1;
5. each factor is an independent AR(1) process $f_{i,t} = \phi f_{i,t} + \varepsilon_{i,t}$, with ϕ ranging from 0.1 to 0.9;
6. the observables can be split into B blocks, that are known *a priori*, inside each of which there is a certain number of cointegrating relationships;
7. for each block, the cointegration matrix is generated as

   $$\beta_b = \begin{bmatrix} I \\ \tilde{\beta}_b \end{bmatrix},$$

   where the elements of $\tilde{\beta}_b$ are independent standard normal pseudo random variates;
8. no cointegration occurs between blocks.

Table 1 contains a summary of the principal features of the seven experiments we ran. For each experiment, each block contains the same number of observables and has the same cointegration rank, so for example in the case $n = 64, B = 4, r = 4$ we have 4 blocks of 16 series each, and the cointegration rank within each block is 4. In other words, Y_t is a 64-variate $I(1)$ process with cointegration rank $r = 16$.

We measure the ability of the DFM to reconstruct the factor space via the trace statistic

$$H = \text{tr}\left[\frac{F_t' P_f F_t}{F_t' F_t}\right]$$

Table 1: Simulation design

n	32	32	64	64	64	64	128
B	4	4	4	4	8	8	8
r	2	4	2	4	2	4	4

where F_t are the simulated factors and P_f is the projection matrix for the estimated factors. This is a standard tool in the DFM literature and has been used, among others, in Doz et al. (2012).

For each DGP, we generate 400 realisations and estimate for each block the corresponding set of cointegrating vectors via Johansen's ML procedure, with a lag length chosen by minimising the Hannan-Quinn information criterion. Then, after performing the variable transformation (4), a factor model is estimated; to save CPU time, estimation of the factor models is performed via PC, using the estimated factors to compute the trace statistic that we call H_r. We then repeat the DFM estimation stage on the purely differenced system (that is, the rotated system with $r = 0$) and compute the trace statistic H_d. The tables presented in this Section and in Appendix A refer to the difference $\nabla = H_r - H_d$. Table 2 contains the results.

In order to assess the consequences of small-sample issues when estimating the cointegration rank for each block, we repeat the whole experiment by estimating the cointegration rank via the Johansen (1991) trace test at a 1% level (see discussion at Subsection 2.3) and report the results in Table 3. Additional descriptive statistics on the ∇ index for both scenarios are reported in Appendix A.

Table 2: Simulation results: empirical frequency of $\nabla > 0$ (rank assumed known)

			q	1			3			5			8		
			ϕ	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
n	B	r													
32	4	2	$T = 200$	0.70	0.80	0.83	0.89	0.92	0.86	0.98	0.97	0.92	1.00	0.98	0.44
			$T = 400$	0.68	0.84	0.87	0.91	0.94	0.89	0.99	0.98	0.93	1.00	1.00	0.54
		4	$T = 200$	0.80	0.89	0.93	0.99	0.99	0.95	1.00	1.00	0.90	1.00	1.00	0.66
			$T = 400$	0.82	0.87	0.89	0.99	1.00	0.96	1.00	1.00	0.96	1.00	1.00	0.83
64	4	2	$T = 200$	0.66	0.85	0.85	0.84	0.89	0.85	0.95	0.97	0.84	0.98	0.94	0.74
			$T = 400$	0.65	0.84	0.88	0.86	0.92	0.85	0.97	0.95	0.87	1.00	0.96	0.77
		4	$T = 200$	0.74	0.90	0.93	0.95	0.99	0.93	1.00	1.00	0.96	1.00	1.00	0.97
			$T = 400$	0.73	0.93	0.95	0.96	0.97	0.91	1.00	1.00	0.95	1.00	1.00	0.98
	8	2	$T = 200$	0.68	0.90	0.94	0.96	0.97	0.90	1.00	1.00	0.96	1.00	1.00	0.66
			$T = 400$	0.69	0.87	0.92	0.96	0.98	0.92	1.00	1.00	0.97	1.00	1.00	0.68
		4	$T = 200$	0.83	0.89	0.94	1.00	1.00	0.97	1.00	1.00	0.92	1.00	1.00	0.71
			$T = 400$	0.81	0.93	0.94	1.00	1.00	0.98	1.00	1.00	0.97	1.00	1.00	0.90
128	8	4	$T = 200$	0.76	0.93	0.97	0.99	1.00	0.95	1.00	1.00	0.98	1.00	1.00	1.00
			$T = 400$	0.75	0.92	0.96	1.00	1.00	0.95	1.00	1.00	0.98	1.00	1.00	1.00

As can be seen, injecting information on the cointegration structure of the data improves the ability of the model to reconstruct the factor space almost uniformly: in several cases, there were no replications at all when $\nabla < 0$; in other words, the DFM on the transformed model (4) yielded a better reconstruction of the factor space than the DFM on the purely differenced data. The one exceptional case was the small-sample ($T = 200$) case with 32 observable series and 8 very persistent factors; we conjecture that in this particular case the poor performance of our proposed procedure is caused by the imprecision in estimating the cointegrating vectors.

Table 3: Simulation results: empirical frequency of $\nabla > 0$ (rank estimated)

			q	1			3			5			8		
			ϕ	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
n	B	r													
32	4	2	$T=200$	0.74	0.80	0.83	0.88	0.92	0.87	0.96	0.94	0.55	0.90	0.85	0.06
			$T=400$	0.68	0.84	0.87	0.90	0.94	0.89	0.98	0.96	0.75	0.95	0.94	0.26
		4	$T=200$	0.78	0.87	0.93	0.97	0.99	0.89	1.00	0.99	0.81	1.00	0.99	0.73
			$T=400$	0.82	0.86	0.89	0.99	0.99	0.93	0.99	0.98	0.90	1.00	1.00	0.77
64	4	2	$T=200$	0.70	0.84	0.90	0.83	0.89	0.85	0.94	0.96	0.82	0.95	0.92	0.05
			$T=400$	0.65	0.86	0.86	0.85	0.92	0.85	0.97	0.98	0.87	0.96	0.92	0.32
		4	$T=200$	0.77	0.86	0.92	0.94	0.98	0.90	1.00	1.00	0.90	1.00	0.99	0.05
			$T=400$	0.74	0.92	0.95	0.98	0.97	0.90	1.00	1.00	0.97	1.00	1.00	0.43
	8	2	$T=200$	0.68	0.90	0.94	0.95	0.98	0.88	0.99	0.98	0.54	0.98	0.92	0.05
			$T=400$	0.73	0.86	0.92	0.95	0.98	0.92	1.00	0.97	0.74	0.99	0.97	0.26
		4	$T=200$	0.82	0.88	0.92	0.99	0.99	0.90	1.00	1.00	0.86	1.00	1.00	0.73
			$T=400$	0.81	0.90	0.94	1.00	1.00	0.96	1.00	0.99	0.89	1.00	1.00	0.83
128	8	4	$T=200$	0.76	0.89	0.96	0.98	1.00	0.87	1.00	1.00	0.90	1.00	1.00	0.05
			$T=400$	0.78	0.92	0.96	1.00	1.00	0.94	1.00	1.00	0.95	1.00	1.00	0.45

When the cointegration rank is estimated rather than known, differences are qualitatively minor, and mostly appear in situations, such as $n = 32, q \geq 5$, when the number of factors is not much smaller than the number of observables and the cross-sectional information available to reconstruct the factor structure is less abundant.

4 Empirical analysis of commodity prices

We now employ the methodology proposed here to analyse the comovement of different commodity prices. Commodity prices provide a perfect example of how our procedure works with real data, since there is a large amount of empirical research demonstrating that there is common movement among different kinds of commodity markets (see for example Pindyck and Rotemberg (1990); Byrne et al. (2013); Delle Chiaie et al. (2017); Alquist et al. (2020)). Many empirical investigations agree on the increased relative importance of the common movement starting from the mid-2000s (Vansteenkiste, 2009; Poncela et al., 2014; Delle Chiaie et al., 2017).

By using our proposed procedure, the question of whether the co-movement is originating from short-run or long-run forces can be given, at least in principle, an empirical answer. Even though results on the identification and assessment of relative importance of co-movement drivers are mixed, the common component is often summarised by a single global factor. According to Byrne et al. (2013), the most relevant driver of co-movement is the interest rate; Vansteenkiste (2009) cite oil price, US dollar exchange rate, interest rate but also, recently increasing in importance, global demand; Delle Chiaie et al. (2017); Alquist et al. (2020) conclude that the single found factor is closely related to fluctuations in global economic activity. As pointed out in Baumeister et al. (2020), the idea behind this link is that demand-induced fluctuations in economic activity cause common movements of prices in the same direction, whereas idiosyncratic shocks reflect supply-side behaviour specific to single commodity markets.

The demand-induced common movements reflecting the global economic activity are of course thought to be more relevant on the medium/long-run. Poncela et al. (2014), instead, focus on short-term fluctuations, suggesting that co-movement can originate because of speculative or financial causes; thay find that uncertainty has an important role in determining short-run common movement.

Further, the debate on the existence of a commodity prices declining trend, originally stated by Prebisch (1962); Singer (1975) and thus known as the *Prebisch and Singer hypothesis* (PHS), is still unresoved. According to this thesis, the prices of primary goods are expected to decline with respect to manufactured goods over the long-term, but empirical evidence on the existence of such a downward trend is controversial (see Harvey et al. (2010)). The rapid and huge rise in commodity prices that occurred during the mid-2000s led to the general worry among observers that this rise was the consequence of a paradigm change: specifically, an increase in the global demand for commodities mainly driven by China. By allowing for a P-T decomposition we are able to analyze not only the comovement of different prices, but, as already mentioned, also the long-term dynamics of the series.

Last but not least, commodity markets offer a very natural example of variables which can easily be split into cointegration blocks: by assuming that cointegration is present among similar commodities (*eg*, belonging to the same market), but not among different kinds, we are simply grouping prices following the economic theory. Interestingly, also Delle Chiaie et al. (2017), even if within a different framework, consider different commodity market blocks for the analysis of co-movement.

The rest of this Section will describe the data and the main results obtained by using our procedure.

4.1 Data

For our analysis, we consider the prices of different kinds of commodities, including energy, metals, food (which include different kinds of goods, too: livestock products, crop commodities, beverages, etc.) and other agricultural commodities. Specifically, we use a set of 37 monthly commodity spot prices provided by the IMF primary commodities database, covering a range from January 1980 to July 2020.[3]

Prices are expressed as indices (January 2000=100) in order to get rid of different units of measure and deflated using the Consumer Price Index provided by the FED to obtain real prices. Note that by considering deflated commodity prices, we are implicitly analysing the relative prices of commodities; therefore, the Prebisch-Singer hypothesis, which refers to the ratio between primary and manufactured goods, is equivalent to that of a long-term declining trend in real commodity prices. Therefore, the basic series we are using are log real price indices. Unit root tests confirm the non-stationarity hypothesis for most series.[4]

[3]The complete list of the used prices could be found in Appendix C.

[4]Alongside the classic ADF test (Dickey and Fuller, 1979), we also used the PP (Phillips and Perron, 1988) and the KPSS (Kwiatkowski et al., 1992) tests; tests for first differences prices reject non-stationarity in all cases.

4.2 Results

In the first step of our analysis, we determine the number of cointegration relationships and, thus, of common trends. To do so, we split the log commodity prices in six blocks, each representing a different category: these are "base metals", "precious metals", "energy", "livestock", "raw materials" and "food". Table 30, reported in Appendix C, shows the six blocks with the included commodity prices. Results of cointegration analysis are summarised in Table 4; they are obtained with the Johansen (1991) trace test for each block, with a rejection level set to $\gamma = 0.01$ instead of the usual 5% level for the reason explained in Section 2; we always used an unrestricted constant as the deterministic component. The lag length is chosen via the Hannan-Quinn criterion on the unrestricted VAR.

Table 4: Cointegration analysis by blocks

	VAR length	Cointegration rank	Common trends
Base metals	2	1	5
Precious metals	2	1	2
Energy	2	3	1
Livestock	2	2	4
Raw materials	2	1	7
Food	2	9	1
Total		17	20

We end up with a total of 17 cointegration relationships and 20 common trends. Note that at least one cointegration relationship is found in each block.

Subsequently, we compute the Kasa-decomposed series via equation (4), and estimation of the DFM is carried out on the centred/standardised series. We set $q = 2$, $s = 0$, $p = 2$, with q determined following the information criteria proposed in Bai and Ng (2007). Figure 1 reports the two estimated $I(0)$ factors. As previously pointed out, for the empirical analysis we estimate the DFM by ML via the EM algorithm; this procedure is computationally more demanding, but evidence presented in Doz et al. (2012) and Lucchetti and Venetis (2020) suggests that this method offers a better reconstruction of the original factor space than other techniques. The two factors capture the general common movement of all the 37 transformed series and, interestingly, look rather time-persistent.

At this point, each commodity price is decomposed into a common component and an idiosyncratic one. To do so, we recover $\hat{Y}_t$ and $\hat{Y}_t^c$ as explained at the end of Section 2.4. Figure 2 shows as an example the performed decomposition for the log-price of wheat in its three components. The common ones are denoted by `l_wheat_perm` and `l_wheat_trans` and correpsond to the $\mathcal{T}_t$ and $\mathcal{C}_t$ components, respectively; the idiosyncratic component is shown on the bottom-right panel. The applied decomposition shows that the Permanent and Idiosyncratic components are the most important in explaining the dynamics of the wheat log-price, whereas the Transitory component is rather marginal. This means that both the common and the wheat-specific movements are determinant, but, crucially, the common movement is a *long-term* one. This result is

Figure 1: Extracted factors, Doz et al. (2012) EM algorithm

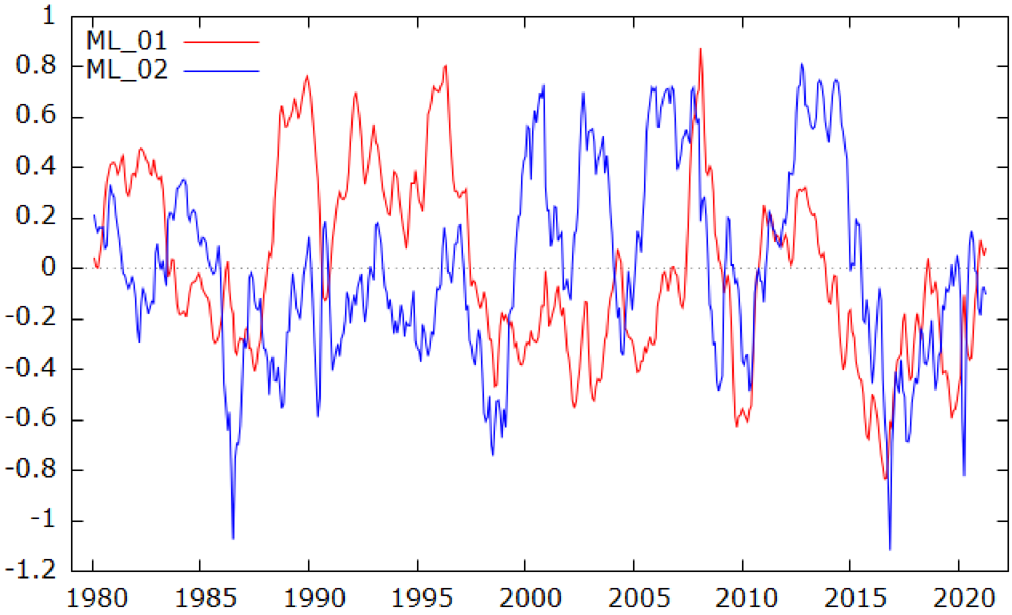
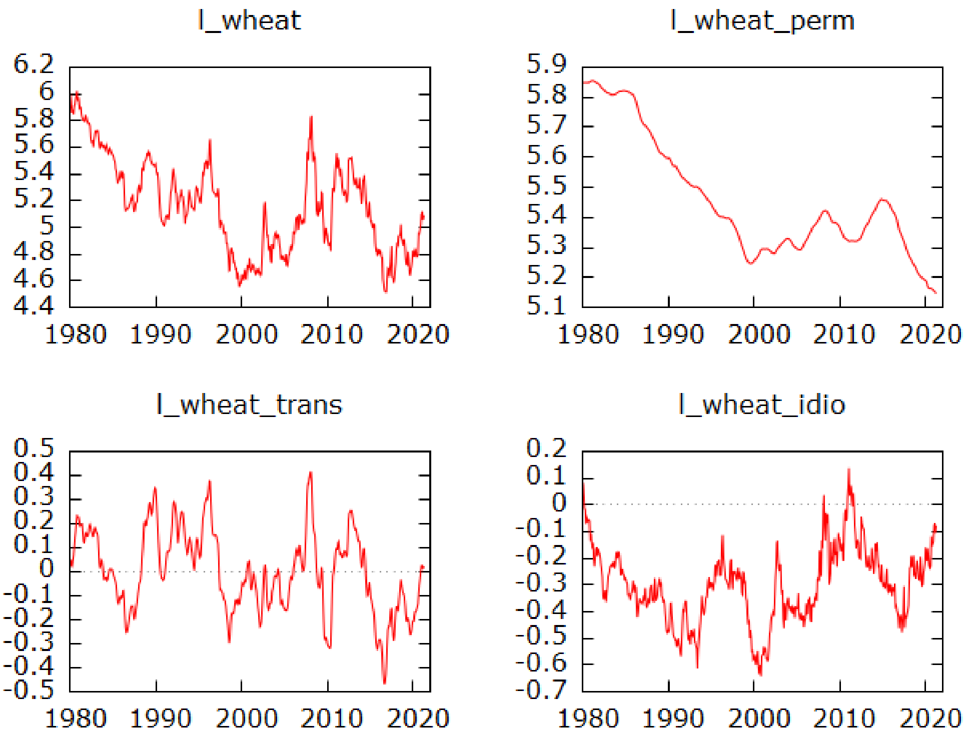


Figure 2: Decomposition of the wheat log-price in its three components



found for all the other log-prices, suggesting that, overall, the short-term comovement can be considered negligible, and that the prices of different commodities move together specifically over the long run.

Figures 3, 5, 7, 9, 11 and 13, contained in Appendix B, display all the log commodity prices with their corresponding Permanent components, divided by blocks.

From our results, it appears that the co-movement of commodity prices is mostly driven by long-run forces, while evidence of an important short-term impact is scant. Focusing on the Permanent component, which is obtained as a combination of the within-block cointegration structure and the global common factors, evidence suggests that the behaviour is different if we consider primary commodity prices or energy and metal ones. The trend for the energy block shows a declining pattern and then an upward one since the beginning of the century, whereas for metals this happens only for some prices; the others exhibit a trend stabilisation since the 2000s. For the other blocks, instead, overall the trend is declining over the entire time span, or at least, it remains stable, suggesting there is no room for the break of the PSH.

5 Conclusions

The possibility of including cointegration in DFMs is a long-standing issue in the literature. Especially for very large systems of variables, performing standard cointegration analysis is a daunting task.

However, in some cases the vector of observable variables can be split into blocks on the basis of prior information; this makes it possible to estimate cointegration vectors for blocks of moderate dimension. On this basis, we propose to estimate a DFM from a vector of transformed variables, instead than the original observed ones; to be specific, we apply the decomposition suggested by Kasa (1992) and separate the stationary part from the non-stationary one. Subsequently, we take first differences of the $I(1)$ part of the system and perform the factor extraction. We analyse the advantages that stem from using the cointegration matrix for re-expressing the original variables in a way that is more suitable for factor analysis, that is, avoiding differentiation whenever possible to make the series $I(0)$.

A simulation experiment indicates that the ability of a DFM to reconstruct the factors space is improved by taking cointegration into account, with respect to the routine followed by many practitioners of estimating the DFM on first differences.

We also apply our proposed method to analyse comovement of different commodity prices, divided by blocks. We find that the common movement is mainly driven by long-run forces, whereas the transitory common fluctuations are of rather marginal importance for all commodity prices. Being able to decompose the common component into a long- and a short-run component makes it also possible to provide interesting empirical evidence on the Prebish and Singer hypothesis, according to which primary commodity prices are expected to decline with respect to prices for manufactured goods. For the food, livestock and agricultural raw materials prices, we find support for an overall declining trend. On the contrary, metals and energy commodities, which where characterised by a declining trend until the beginning of this century, are now exhibiting an upward, or at least stable, pattern.

References

Alquist, R., S. Bhattarai, and O. Coibion (2020): "Commodity-price comovement and global economic activity," *Journal of Monetary Economics*, 112, 41–56.

Bai, J. (2004): "Estimating cross-section common stochastic trends in nonstationary panel data," *Journal of Econometrics*, 122, 137–183.

Bai, J. and S. Ng (2004): "A PANIC attack on unit roots and cointegration," *Econometrica*, 72, 1127–1177.

——— (2007): "Determining the number of primitive shocks in factor models," *Journal of Business & Economic Statistics*, 25, 52–60.

——— (2008): "Large Dimensional Factor Analysis," *Foundations and Trends (R) in Econometrics*, 3, 89–163.

Barigozzi, M., M. Lippi, and M. Luciani (2016): "Non-stationary dynamic factor models for large datasets," *arXiv preprint arXiv:1602.02398*.

——— (2020): "Cointegration and error correction mechanisms for singular stochastic vectors," *Econometrics*, 8, 3.

——— (2021): "Large-dimensional Dynamic Factor Models: Estimation of Impulse–Response Functions with $I(1)$ cointegrated factors," *Journal of Econometrics*, 221, 455–482.

Barigozzi, M. and M. Luciani (2019): "Measuring the output gap using large datasets," Tech. rep., Technical report.

Baumeister, C., D. Korobilis, and T. K. Lee (2020): "Energy markets and global economic conditions," *The Review of Economics and Statistics*, 1–45.

Byrne, J. P., G. Fazio, and N. Fiess (2013): "Primary commodity prices: Comovements, common factors and fundamentals," *Journal of Development Economics*, 101, 16–26.

Cavaliere, G., H. B. Nielsen, and A. Rahbek (2015): "Bootstrap testing of hypotheses on co-integration relations in vector autoregressive models," *Econometrica*, 83, 813–831.

Cavaliere, G., A. Rahbek, and A. R. Taylor (2012): "Bootstrap determination of the co-integration rank in vector autoregressive models," *Econometrica*, 80, 1721–1740.

Corona, F., P. Poncela, and E. Ruiz (2020): "Estimating non-stationary common factors: implications for risk sharing," *Computational Economics*, 55, 37–60.

Delle Chiaie, S., L. Ferrara, and D. Giannone (2017): "Common factors of commodity prices," *La Défense-EconomiX*.

Dickey, D. A. and W. A. Fuller (1979): "Distribution of the estimators for autoregressive time series with a unit root," *Journal of the American statistical association*, 74, 427–431.

Doz, C. and P. Fuleky (2020): "Dynamic factor models," in *Macroeconomic forecasting in the era of big data*, ed. by P. Fuleky, Springer, 27–64.

Doz, C., D. Giannone, and L. Reichlin (2011): "A two-step estimator for large approximate dynamic factor models based on Kalman filtering," *Journal of Econometrics*, 164, 188–205.

——— (2012): "A quasi–maximum likelihood approach for large, approximate dynamic factor models," *Review of economics and statistics*, 94, 1014–1024.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000): "The generalized dynamic-factor model: Identification and estimation," *Review of Economics and statistics*, 82, 540–554.

——— (2005): "The generalized dynamic factor model: one-sided estimation and forecasting," *Journal of the American Statistical Association*, 100, 830–840.

Forni, M., M. Hallin, M. Lippi, and P. Zaffaroni (2015): "Dynamic factor models with infinite-dimensional factor spaces: One-sided representations," *Journal of econometrics*, 185, 359–371.

——— (2017): "Dynamic factor models with infinite-dimensional factor space: Asymptotic analysis," *Journal of Econometrics*, 199, 74–92.

Geweke, J. (1977): "The dynamic factor analysis of economic time series," *Latent variables in socio-economic models.*

Gonzalo, J. and C. Granger (1995): "Estimation of common long-memory components in cointegrated systems," *Journal of Business & Economic Statistics*, 13, 27–35.

Hallin, M. and R. Liška (2007): "Determining the number of factors in the general dynamic factor model," *Journal of the American Statistical Association*, 102, 603–617.

Harvey, D. I., N. M. Kellard, J. B. Madsen, and M. E. Wohar (2010): "The Prebisch-Singer hypothesis: four centuries of evidence," *The review of Economics and Statistics*, 92, 367–377.

Johansen, S. (1991): "Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models," *Econometrica: journal of the Econometric Society*, 1551–1580.

——— (1992): "Cointegration in partial systems and the efficiency of single-equation analysis," *Journal of econometrics*, 52, 389–402.

——— (2000): "A Bartlett correction factor for tests on the cointegrating relations," *Econometric Theory*, 16, 740–778.

Kasa, K. (1992): "Common stochastic trends in international stock markets," *Journal of monetary Economics*, 29, 95–124.

Kwiatkowski, D., P. C. Phillips, P. Schmidt, and Y. Shin (1992): "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?" *Journal of econometrics*, 54, 159–178.

Lucchetti, R. and I. A. Venetis (2020): "A replication of "A quasi-maximum likelihood approach for large, approximate dynamic factor models"(Review of Economics and Statistics, 2012)," *Economics*, 14.

Onatski, A. and C. Wang (2018): "Alternative asymptotics for cointegration tests in large vars," *Econometrica*, 86, 1465–1478.

Phillips, P. C. and P. Perron (1988): "Testing for a unit root in time series regression," *Biometrika*, 75, 335–346.

Pindyck, R. S. and J. J. Rotemberg (1990): "The Excess Co-Movement of Commodity Prices," *The Economic Journal*, 100, 1173–1189.

Poncela, P., E. Senra, and L. P. Sierra (2014): "Common dynamics of nonenergy commodity prices and their relation to uncertainty," *Applied Economics*, 46, 3724–3735.

Prebisch, R. (1962): "The economic development of Latin America and its principal problems," *Economic Bulletin for Latin America*.

Sargent, T. J. and C. A. Sims (1977): "Business cycle modeling without pretending to have too much a priori economic theory," *New methods in business cycle research*, 1, 145–168.

Singer, H. W. (1975): "The distribution of gains between investing and borrowing countries," in *The strategy of international development*, Springer, 43–57.

Stock, J. H. and M. Watson (2011): "Dynamic factor models," *Oxford handbook on economic forecasting*.

Stock, J. H. and M. W. Watson (2016): "Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics," in *Handbook of macroeconomics*, Elsevier, vol. 2, 415–525.

Vansteenkiste, I. (2009): "How important are common factors in driving non-fuel commodity prices? A dynamic factor analysis," Working Paper Series 1072, European Central Bank.

A Simulation results

Table 5: Simulation results, rank assumed known, $q = 1$, $\phi = 0.1$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.070	0.695	0.171	0.008	8.215	0.000
			$T = 400$	0.065	0.678	0.157	0.006	8.198	0.000
		4	$T = 200$	0.196	0.800	0.289	0.040	13.549	0.000
			$T = 400$	0.217	0.815	0.316	0.038	13.692	0.000
64	4	2	$T = 200$	0.020	0.658	0.077	0.002	5.293	0.000
			$T = 400$	0.034	0.648	0.115	0.002	5.873	0.000
		4	$T = 200$	0.052	0.743	0.131	0.007	7.955	0.000
			$T = 400$	0.042	0.725	0.123	0.006	6.758	0.000
	8	2	$T = 200$	0.042	0.680	0.132	0.004	6.353	0.000
			$T = 400$	0.040	0.693	0.135	0.003	5.887	0.000
		4	$T = 200$	0.162	0.825	0.277	0.028	11.697	0.000
			$T = 400$	0.203	0.808	0.320	0.021	12.664	0.000
128	8	4	$T = 200$	0.022	0.760	0.069	0.005	6.399	0.000
			$T = 400$	0.023	0.750	0.084	0.004	5.527	0.000

Table 6: Simulation results, rank assumed known, $q = 1$, $\phi = 0.5$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.114	0.800	0.159	0.072	14.327	0.000
			$T = 400$	0.131	0.838	0.173	0.075	15.162	0.000
		4	$T = 200$	0.266	0.890	0.253	0.207	21.049	0.000
			$T = 400$	0.264	0.870	0.258	0.190	20.460	0.000
64	4	2	$T = 200$	0.060	0.845	0.074	0.043	16.046	0.000
			$T = 400$	0.063	0.843	0.080	0.040	15.694	0.000
		4	$T = 200$	0.114	0.898	0.115	0.082	19.795	0.000
			$T = 400$	0.122	0.928	0.111	0.103	21.975	0.000
	8	2	$T = 200$	0.112	0.900	0.117	0.086	19.148	0.000
			$T = 400$	0.109	0.868	0.131	0.071	16.610	0.000
		4	$T = 200$	0.231	0.885	0.222	0.181	20.799	0.000
			$T = 400$	0.257	0.928	0.227	0.214	22.675	0.000
128	8	4	$T = 200$	0.091	0.925	0.096	0.066	18.907	0.000
			$T = 400$	0.105	0.920	0.096	0.082	21.712	0.000

Table 7: Simulation results, rank assumed known, $q = 1$, $\phi = 0.9$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.057	0.833	0.080	0.046	14.312	0.000
			$T = 400$	0.053	0.868	0.069	0.045	15.528	0.000
		4	$T = 200$	0.136	0.933	0.123	0.113	21.960	0.000
			$T = 400$	0.129	0.888	0.142	0.110	18.095	0.000
64	4	2	$T = 200$	0.030	0.850	0.032	0.027	18.385	0.000
			$T = 400$	0.029	0.880	0.032	0.027	18.430	0.000
		4	$T = 200$	0.062	0.925	0.058	0.054	21.305	0.000
			$T = 400$	0.070	0.953	0.067	0.060	20.767	0.000
	8	2	$T = 200$	0.057	0.943	0.046	0.051	24.401	0.000
			$T = 400$	0.062	0.918	0.079	0.052	15.800	0.000
		4	$T = 200$	0.120	0.943	0.098	0.107	24.357	0.000
			$T = 400$	0.128	0.943	0.107	0.114	23.889	0.000
128	8	4	$T = 200$	0.064	0.965	0.050	0.058	25.654	0.000
			$T = 400$	0.066	0.958	0.053	0.064	25.210	0.000

Table 8: Simulation results, rank assumed known, $q = 3$, $\phi = 0.1$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.070	0.888	0.089	0.032	15.611	0.000
			$T = 400$	0.075	0.908	0.096	0.032	15.647	0.000
		4	$T = 200$	0.159	0.985	0.110	0.144	28.949	0.000
			$T = 400$	0.147	0.993	0.109	0.124	26.842	0.000
64	4	2	$T = 200$	0.042	0.840	0.073	0.011	11.421	0.000
			$T = 400$	0.045	0.858	0.074	0.013	12.124	0.000
		4	$T = 200$	0.104	0.948	0.113	0.045	18.438	0.000
			$T = 400$	0.093	0.963	0.110	0.034	16.882	0.000
	8	2	$T = 200$	0.128	0.963	0.118	0.080	21.692	0.000
			$T = 400$	0.117	0.963	0.115	0.067	20.375	0.000
		4	$T = 200$	0.202	1.000	0.120	0.245	33.701	0.000
			$T = 400$	0.223	0.998	0.117	0.273	38.264	0.000
128	8	4	$T = 200$	0.182	0.990	0.130	0.205	28.038	0.000
			$T = 400$	0.184	0.998	0.132	0.219	27.829	0.000

Table 9: Simulation results, rank assumed known, $q = 3$, $\phi = 0.5$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.075	0.923	0.074	0.054	20.316	0.000
			$T = 400$	0.083	0.943	0.075	0.061	22.148	0.000
		4	$T = 200$	0.164	0.993	0.092	0.169	35.568	0.000
			$T = 400$	0.159	0.995	0.094	0.158	33.850	0.000
64	4	2	$T = 200$	0.046	0.885	0.055	0.026	16.511	0.000
			$T = 400$	0.049	0.918	0.059	0.027	16.466	0.000
		4	$T = 200$	0.109	0.985	0.086	0.087	25.302	0.000
			$T = 400$	0.101	0.970	0.083	0.078	24.566	0.000
	8	2	$T = 200$	0.112	0.973	0.089	0.089	25.225	0.000
			$T = 400$	0.104	0.978	0.081	0.088	25.487	0.000
		4	$T = 200$	0.203	0.995	0.093	0.219	43.742	0.000
			$T = 400$	0.201	1.000	0.089	0.220	45.150	0.000
128	8	4	$T = 200$	0.148	0.995	0.085	0.157	34.652	0.000
			$T = 400$	0.155	0.995	0.085	0.175	36.533	0.000

Table 10: Simulation results, rank assumed known, $q = 3$, $\phi = 0.9$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.065	0.855	0.081	0.057	16.094	0.000
			$T = 400$	0.079	0.888	0.077	0.067	20.269	0.000
		4	$T = 200$	0.132	0.953	0.093	0.125	28.227	0.000
			$T = 400$	0.140	0.963	0.095	0.137	29.429	0.000
64	4	2	$T = 200$	0.048	0.848	0.067	0.038	14.356	0.000
			$T = 400$	0.048	0.845	0.065	0.039	14.595	0.000
		4	$T = 200$	0.083	0.933	0.070	0.073	23.813	0.000
			$T = 400$	0.092	0.910	0.084	0.078	21.815	0.000
	8	2	$T = 200$	0.073	0.903	0.070	0.061	21.023	0.000
			$T = 400$	0.080	0.915	0.074	0.068	21.541	0.000
		4	$T = 200$	0.137	0.973	0.087	0.127	31.566	0.000
			$T = 400$	0.141	0.975	0.086	0.126	32.847	0.000
128	8	4	$T = 200$	0.086	0.950	0.065	0.075	26.453	0.000
			$T = 400$	0.086	0.948	0.060	0.079	28.684	0.000

Table 11: Simulation results, rank assumed known, $q = 5$, $\phi = 0.1$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.137	0.980	0.070	0.146	39.055	0.000
			$T = 400$	0.133	0.990	0.070	0.138	37.986	0.000
		4	$T = 200$	0.213	1.000	0.070	0.206	60.953	0.000
			$T = 400$	0.208	1.000	0.074	0.202	55.932	0.000
64	4	2	$T = 200$	0.106	0.948	0.076	0.098	27.660	0.000
			$T = 400$	0.107	0.968	0.072	0.106	29.610	0.000
		4	$T = 200$	0.198	1.000	0.073	0.192	54.487	0.000
			$T = 400$	0.193	0.998	0.075	0.190	51.344	0.000
	8	2	$T = 200$	0.164	0.998	0.075	0.175	44.010	0.000
			$T = 400$	0.159	0.998	0.074	0.172	42.770	0.000
		4	$T = 200$	0.214	1.000	0.080	0.210	53.302	0.000
			$T = 400$	0.222	1.000	0.079	0.215	56.037	0.000
128	8	4	$T = 200$	0.190	1.000	0.082	0.196	46.463	0.000
			$T = 400$	0.181	1.000	0.083	0.195	43.727	0.000

Table 12: Simulation results, rank assumed known, $q = 5$, $\phi = 0.5$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.123	0.970	0.070	0.124	35.448	0.000
			$T = 400$	0.126	0.983	0.065	0.125	38.535	0.000
		4	$T = 200$	0.205	1.000	0.072	0.200	57.074	0.000
			$T = 400$	0.209	0.998	0.072	0.202	57.979	0.000
64	4	2	$T = 200$	0.094	0.965	0.061	0.089	30.784	0.000
			$T = 400$	0.091	0.953	0.057	0.087	31.820	0.000
		4	$T = 200$	0.173	1.000	0.064	0.171	53.922	0.000
			$T = 400$	0.173	1.000	0.063	0.166	54.783	0.000
	8	2	$T = 200$	0.146	1.000	0.057	0.147	50.975	0.000
			$T = 400$	0.146	1.000	0.063	0.146	46.723	0.000
		4	$T = 200$	0.216	1.000	0.070	0.210	61.897	0.000
			$T = 400$	0.215	1.000	0.067	0.214	64.237	0.000
128	8	4	$T = 200$	0.177	1.000	0.064	0.170	55.483	0.000
			$T = 400$	0.178	1.000	0.062	0.172	57.581	0.000

Table 13: Simulation results, rank assumed known, $q = 5$, $\phi = 0.9$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.060	0.923	0.050	0.053	24.120	0.000
			$T = 400$	0.066	0.925	0.052	0.061	25.332	0.000
		4	$T = 200$	0.086	0.900	0.067	0.082	25.620	0.000
			$T = 400$	0.122	0.955	0.073	0.123	33.749	0.000
64	4	2	$T = 200$	0.031	0.838	0.037	0.025	16.824	0.000
			$T = 400$	0.036	0.868	0.043	0.031	16.941	0.000
		4	$T = 200$	0.084	0.960	0.053	0.076	31.740	0.000
			$T = 400$	0.091	0.953	0.059	0.085	31.211	0.000
	8	2	$T = 200$	0.081	0.955	0.054	0.073	29.960	0.000
			$T = 400$	0.089	0.968	0.056	0.082	31.966	0.000
		4	$T = 200$	0.097	0.923	0.069	0.097	28.040	0.000
			$T = 400$	0.127	0.965	0.066	0.132	38.652	0.000
128	8	4	$T = 200$	0.088	0.980	0.046	0.084	38.700	0.000
			$T = 400$	0.097	0.978	0.053	0.093	36.450	0.000

Table 14: Simulation results, rank assumed known, $q = 8$, $\phi = 0.1$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.110	1.000	0.051	0.107	43.459	0.000
			$T = 400$	0.116	0.995	0.048	0.113	48.933	0.000
		4	$T = 200$	0.229	1.000	0.054	0.228	84.162	0.000
			$T = 400$	0.244	1.000	0.051	0.240	94.992	0.000
64	4	2	$T = 200$	0.086	0.975	0.053	0.082	32.774	0.000
			$T = 400$	0.090	0.995	0.049	0.088	36.564	0.000
		4	$T = 200$	0.214	1.000	0.060	0.211	70.884	0.000
			$T = 400$	0.229	1.000	0.064	0.225	70.983	0.000
	8	2	$T = 200$	0.208	1.000	0.056	0.212	74.688	0.000
			$T = 400$	0.213	1.000	0.061	0.214	70.262	0.000
		4	$T = 200$	0.317	1.000	0.051	0.314	123.960	0.000
			$T = 400$	0.322	1.000	0.054	0.319	118.960	0.000
128	8	4	$T = 200$	0.308	1.000	0.060	0.310	103.280	0.000
			$T = 400$	0.317	1.000	0.061	0.323	103.730	0.000

Table 15: Simulation results, rank assumed known, $q = 8$, $\phi = 0.5$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.098	0.983	0.044	0.096	44.470	0.000
			$T = 400$	0.107	0.995	0.046	0.104	46.918	0.000
		4	$T = 200$	0.223	1.000	0.052	0.223	85.349	0.000
			$T = 400$	0.238	1.000	0.049	0.240	97.209	0.000
64	4	2	$T = 200$	0.061	0.940	0.040	0.060	30.090	0.000
			$T = 400$	0.064	0.958	0.041	0.062	31.584	0.000
		4	$T = 200$	0.179	1.000	0.056	0.177	63.782	0.000
			$T = 400$	0.189	1.000	0.053	0.186	70.646	0.000
	8	2	$T = 200$	0.164	1.000	0.050	0.161	65.090	0.000
			$T = 400$	0.178	1.000	0.050	0.176	70.820	0.000
		4	$T = 200$	0.283	1.000	0.050	0.280	113.030	0.000
			$T = 400$	0.301	1.000	0.050	0.301	120.700	0.000
128	8	4	$T = 200$	0.248	1.000	0.048	0.248	103.700	0.000
			$T = 400$	0.263	1.000	0.048	0.260	109.110	0.000

Table 16: Simulation results, rank assumed known, $q = 8$, $\phi = 0.9$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	-0.007	0.440	0.043	-0.006	-3.336	1.000
			$T = 400$	0.004	0.538	0.046	0.005	1.782	0.037
		4	$T = 200$	0.020	0.663	0.054	0.024	7.454	0.000
			$T = 400$	0.049	0.830	0.054	0.048	18.153	0.000
64	4	2	$T = 200$	0.014	0.738	0.027	0.014	10.129	0.000
			$T = 400$	0.016	0.765	0.029	0.017	11.128	0.000
		4	$T = 200$	0.058	0.968	0.033	0.057	35.541	0.000
			$T = 400$	0.067	0.980	0.034	0.064	39.453	0.000
	8	2	$T = 200$	0.013	0.655	0.037	0.015	7.231	0.000
			$T = 400$	0.020	0.683	0.044	0.022	8.967	0.000
		4	$T = 200$	0.026	0.705	0.051	0.023	10.161	0.000
			$T = 400$	0.068	0.898	0.054	0.069	25.239	0.000
128	8	4	$T = 200$	0.082	0.998	0.031	0.076	53.545	0.000
			$T = 400$	0.098	1.000	0.035	0.093	56.254	0.000

Table 17: Simulation results, rank estimated, $q = 1$, $\phi = 0.1$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.076	0.738	0.176	0.009	8.654	0.000
			$T = 400$	0.067	0.678	0.159	0.007	8.439	0.000
		4	$T = 200$	0.192	0.783	0.284	0.041	13.508	0.000
			$T = 400$	0.217	0.823	0.310	0.042	13.994	0.000
64	4	2	$T = 200$	0.047	0.695	0.124	0.005	7.618	0.000
			$T = 400$	0.033	0.653	0.107	0.003	6.151	0.000
		4	$T = 200$	0.078	0.770	0.173	0.011	9.066	0.000
			$T = 400$	0.044	0.738	0.131	0.006	6.686	0.000
	8	2	$T = 200$	0.046	0.678	0.143	0.004	6.452	0.000
			$T = 400$	0.042	0.725	0.143	0.003	5.922	0.000
		4	$T = 200$	0.158	0.818	0.275	0.023	11.536	0.000
			$T = 400$	0.203	0.810	0.317	0.024	12.794	0.000
128	8	4	$T = 200$	0.035	0.758	0.101	0.006	7.014	0.000
			$T = 400$	0.027	0.775	0.092	0.004	5.908	0.000

Table 18: Simulation results, rank estimated, $q = 1$, $\phi = 0.5$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.118	0.800	0.162	0.072	14.552	0.000
			$T = 400$	0.133	0.835	0.176	0.073	15.116	0.000
		4	$T = 200$	0.264	0.865	0.253	0.215	20.840	0.000
			$T = 400$	0.257	0.863	0.253	0.185	20.334	0.000
64	4	2	$T = 200$	0.090	0.835	0.110	0.062	16.262	0.000
			$T = 400$	0.071	0.855	0.091	0.046	15.577	0.000
		4	$T = 200$	0.141	0.863	0.141	0.106	20.075	0.000
			$T = 400$	0.127	0.923	0.114	0.111	22.168	0.000
	8	2	$T = 200$	0.118	0.898	0.120	0.089	19.628	0.000
			$T = 400$	0.112	0.858	0.135	0.071	16.635	0.000
		4	$T = 200$	0.227	0.883	0.220	0.158	20.546	0.000
			$T = 400$	0.249	0.903	0.225	0.195	22.109	0.000
128	8	4	$T = 200$	0.100	0.890	0.114	0.067	17.443	0.000
			$T = 400$	0.109	0.915	0.099	0.088	21.900	0.000

Table 19: Simulation results, rank estimated, $q = 1$, $\phi = 0.9$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.059	0.833	0.086	0.047	13.774	0.000
			$T = 400$	0.056	0.870	0.068	0.044	16.375	0.000
		4	$T = 200$	0.134	0.925	0.124	0.111	21.601	0.000
			$T = 400$	0.127	0.890	0.141	0.110	17.948	0.000
64	4	2	$T = 200$	0.047	0.895	0.043	0.044	21.918	0.000
			$T = 400$	0.031	0.863	0.033	0.028	18.640	0.000
		4	$T = 200$	0.074	0.915	0.079	0.067	18.780	0.000
			$T = 400$	0.071	0.953	0.068	0.062	20.627	0.000
	8	2	$T = 200$	0.060	0.943	0.050	0.053	23.802	0.000
			$T = 400$	0.057	0.918	0.069	0.050	16.545	0.000
		4	$T = 200$	0.114	0.920	0.100	0.106	22.661	0.000
			$T = 400$	0.123	0.938	0.108	0.113	22.801	0.000
128	8	4	$T = 200$	0.076	0.958	0.068	0.068	22.459	0.000
			$T = 400$	0.065	0.960	0.053	0.061	24.593	0.000

Table 20: Simulation results, rank estimated, $q = 3$, $\phi = 0.1$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.070	0.878	0.092	0.033	15.086	0.000
			$T = 400$	0.074	0.895	0.098	0.028	15.076	0.000
		4	$T = 200$	0.155	0.973	0.111	0.144	28.016	0.000
			$T = 400$	0.147	0.990	0.110	0.126	26.865	0.000
64	4	2	$T = 200$	0.049	0.830	0.079	0.014	12.297	0.000
			$T = 400$	0.046	0.848	0.075	0.013	12.162	0.000
		4	$T = 200$	0.110	0.938	0.117	0.048	18.679	0.000
			$T = 400$	0.105	0.975	0.118	0.041	17.831	0.000
	8	2	$T = 200$	0.125	0.948	0.118	0.076	21.201	0.000
			$T = 400$	0.120	0.950	0.118	0.066	20.328	0.000
		4	$T = 200$	0.197	0.993	0.121	0.228	32.422	0.000
			$T = 400$	0.221	1.000	0.117	0.272	37.734	0.000
128	8	4	$T = 200$	0.176	0.983	0.129	0.207	27.363	0.000
			$T = 400$	0.180	0.995	0.133	0.225	27.131	0.000

Table 21: Simulation results, rank estimated, $q = 3$, $\phi = 0.5$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.080	0.920	0.080	0.056	20.000	0.000
			$T = 400$	0.084	0.940	0.077	0.060	21.854	0.000
		4	$T = 200$	0.162	0.990	0.095	0.166	34.007	0.000
			$T = 400$	0.155	0.990	0.095	0.150	32.826	0.000
64	4	2	$T = 200$	0.056	0.893	0.068	0.031	16.691	0.000
			$T = 400$	0.048	0.915	0.056	0.029	17.051	0.000
		4	$T = 200$	0.109	0.978	0.086	0.087	25.393	0.000
			$T = 400$	0.103	0.968	0.084	0.082	24.669	0.000
	8	2	$T = 200$	0.113	0.975	0.088	0.095	25.638	0.000
			$T = 400$	0.104	0.980	0.081	0.085	25.605	0.000
		4	$T = 200$	0.198	0.993	0.092	0.213	43.037	0.000
			$T = 400$	0.197	1.000	0.090	0.216	43.892	0.000
128	8	4	$T = 200$	0.159	0.998	0.089	0.175	35.588	0.000
			$T = 400$	0.155	0.995	0.089	0.163	34.793	0.000

Table 22: Simulation results, rank estimated, $q = 3$, $\phi = 0.9$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.075	0.868	0.088	0.062	17.036	0.000
			$T = 400$	0.077	0.888	0.082	0.068	18.661	0.000
		4	$T = 200$	0.110	0.893	0.103	0.106	21.304	0.000
			$T = 400$	0.132	0.933	0.102	0.128	25.827	0.000
64	4	2	$T = 200$	0.062	0.845	0.075	0.051	16.504	0.000
			$T = 400$	0.053	0.853	0.064	0.044	16.574	0.000
		4	$T = 200$	0.083	0.900	0.080	0.075	20.973	0.000
			$T = 400$	0.085	0.903	0.074	0.076	23.065	0.000
	8	2	$T = 200$	0.071	0.875	0.077	0.061	18.393	0.000
			$T = 400$	0.085	0.915	0.073	0.075	23.183	0.000
		4	$T = 200$	0.111	0.895	0.101	0.106	22.013	0.000
			$T = 400$	0.130	0.960	0.093	0.122	27.980	0.000
128	8	4	$T = 200$	0.079	0.870	0.071	0.081	22.517	0.000
			$T = 400$	0.092	0.940	0.069	0.086	26.731	0.000

Table 23: Simulation results, rank estimated, $q = 5$, $\phi = 0.1$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.130	0.963	0.074	0.139	35.330	0.000
			$T = 400$	0.128	0.975	0.072	0.134	35.266	0.000
		4	$T = 200$	0.199	0.998	0.074	0.199	53.870	0.000
			$T = 400$	0.197	0.993	0.079	0.195	49.639	0.000
64	4	2	$T = 200$	0.112	0.938	0.078	0.111	28.591	0.000
			$T = 400$	0.115	0.970	0.066	0.117	34.894	0.000
		4	$T = 200$	0.200	0.998	0.073	0.193	54.907	0.000
			$T = 400$	0.192	0.998	0.074	0.189	51.891	0.000
	8	2	$T = 200$	0.155	0.993	0.076	0.163	40.716	0.000
			$T = 400$	0.149	0.995	0.076	0.166	39.112	0.000
		4	$T = 200$	0.198	1.000	0.086	0.203	46.153	0.000
			$T = 400$	0.209	1.000	0.086	0.210	48.547	0.000
128	8	4	$T = 200$	0.202	1.000	0.079	0.201	51.414	0.000
			$T = 400$	0.195	1.000	0.076	0.196	51.467	0.000

Table 24: Simulation results, rank estimated, $q = 5$, $\phi = 0.5$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.113	0.940	0.074	0.115	30.465	0.000
			$T = 400$	0.118	0.960	0.072	0.119	32.886	0.000
		4	$T = 200$	0.181	0.985	0.084	0.184	42.987	0.000
			$T = 400$	0.196	0.983	0.081	0.192	48.489	0.000
64	4	2	$T = 200$	0.103	0.955	0.066	0.096	31.284	0.000
			$T = 400$	0.100	0.980	0.056	0.097	35.657	0.000
		4	$T = 200$	0.175	0.998	0.062	0.174	56.363	0.000
			$T = 400$	0.176	1.000	0.062	0.172	57.139	0.000
	8	2	$T = 200$	0.131	0.975	0.066	0.135	39.768	0.000
			$T = 400$	0.139	0.968	0.073	0.141	38.294	0.000
		4	$T = 200$	0.187	0.995	0.080	0.191	46.504	0.000
			$T = 400$	0.200	0.993	0.078	0.207	50.976	0.000
128	8	4	$T = 200$	0.187	1.000	0.062	0.181	60.465	0.000
			$T = 400$	0.179	1.000	0.061	0.171	58.370	0.000

Table 25: Simulation results, rank estimated, $q = 5$, $\phi = 0.9$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.002	0.553	0.081	0.010	0.521	0.301
			$T = 400$	0.040	0.745	0.076	0.047	10.429	0.000
		4	$T = 200$	0.066	0.810	0.076	0.070	17.406	0.000
			$T = 400$	0.096	0.895	0.082	0.099	23.245	0.000
64	4	2	$T = 200$	0.036	0.818	0.058	0.036	12.594	0.000
			$T = 400$	0.045	0.865	0.048	0.042	18.945	0.000
		4	$T = 200$	0.073	0.898	0.066	0.076	22.074	0.000
			$T = 400$	0.095	0.970	0.054	0.089	34.702	0.000
	8	2	$T = 200$	0.002	0.535	0.087	0.009	0.466	0.321
			$T = 400$	0.047	0.738	0.080	0.053	11.892	0.000
		4	$T = 200$	0.074	0.863	0.071	0.072	20.698	0.000
			$T = 400$	0.093	0.890	0.075	0.096	24.692	0.000
128	8	4	$T = 200$	0.079	0.895	0.062	0.082	25.403	0.000
			$T = 400$	0.093	0.948	0.058	0.090	32.073	0.000

Table 26: Simulation results, rank estimated, $q = 8$, $\phi = 0.1$

n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.084	0.895	0.068	0.087	24.854	0.000
			$T = 400$	0.101	0.948	0.059	0.104	34.282	0.000
		4	$T = 200$	0.215	0.998	0.061	0.215	70.206	0.000
			$T = 400$	0.231	0.998	0.064	0.231	71.868	0.000
64	4	2	$T = 200$	0.083	0.945	0.054	0.081	30.485	0.000
			$T = 400$	0.083	0.960	0.051	0.079	32.684	0.000
		4	$T = 200$	0.203	1.000	0.060	0.204	67.895	0.000
			$T = 400$	0.220	1.000	0.064	0.219	68.161	0.000
	8	2	$T = 200$	0.171	0.983	0.074	0.170	46.004	0.000
			$T = 400$	0.189	0.988	0.074	0.195	51.169	0.000
		4	$T = 200$	0.290	1.000	0.070	0.298	82.809	0.000
			$T = 400$	0.304	1.000	0.068	0.307	90.053	0.000
128	8	4	$T = 200$	0.296	1.000	0.060	0.301	98.240	0.000
			$T = 400$	0.312	1.000	0.061	0.317	101.520	0.000

Table 27: Simulation results, rank estimated, $q = 8$, $\phi = 0.5$

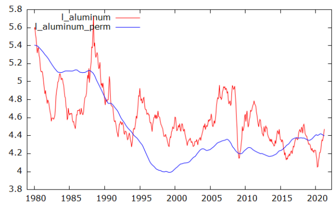
n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	0.065	0.853	0.071	0.073	18.387	0.000
			$T = 400$	0.089	0.935	0.060	0.093	29.784	0.000
		4	$T = 200$	0.193	0.993	0.063	0.196	61.783	0.000
			$T = 400$	0.222	0.998	0.065	0.230	68.637	0.000
64	4	2	$T = 200$	0.062	0.923	0.048	0.061	25.489	0.000
			$T = 400$	0.059	0.918	0.045	0.054	26.352	0.000
		4	$T = 200$	0.161	0.993	0.059	0.160	54.049	0.000
			$T = 400$	0.176	1.000	0.057	0.174	61.597	0.000
	8	2	$T = 200$	0.111	0.923	0.071	0.116	31.241	0.000
			$T = 400$	0.145	0.973	0.071	0.153	41.077	0.000
		4	$T = 200$	0.241	1.000	0.067	0.248	72.291	0.000
			$T = 400$	0.272	1.000	0.071	0.284	76.232	0.000
128	8	4	$T = 200$	0.233	1.000	0.061	0.235	76.928	0.000
			$T = 400$	0.247	1.000	0.053	0.246	93.008	0.000

Table 28: Simulation results, rank estimated, $q = 8$, $\phi = 0.9$

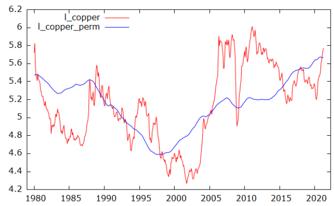
n	B	r		Mean	Freq. $\nabla > 0$	St. Dev.	Median	Test $\nabla = 0$	p-value
32	4	2	$T = 200$	-0.084	0.060	0.055	-0.085	-30.617	1.000
			$T = 400$	-0.052	0.258	0.071	-0.050	-14.683	1.000
		4	$T = 200$	0.023	0.728	0.045	0.023	10.360	0.000
			$T = 400$	0.033	0.768	0.049	0.035	13.564	0.000
64	4	2	$T = 200$	-0.128	0.045	0.072	-0.126	-35.682	1.000
			$T = 400$	-0.032	0.318	0.059	-0.025	-10.824	1.000
		4	$T = 200$	-0.105	0.050	0.061	-0.103	-34.455	1.000
			$T = 400$	-0.022	0.433	0.079	-0.008	-5.537	1.000
	8	2	$T = 200$	-0.075	0.048	0.050	-0.074	-30.157	1.000
			$T = 400$	-0.054	0.258	0.071	-0.046	-15.205	1.000
		4	$T = 200$	0.026	0.728	0.049	0.030	10.565	0.000
			$T = 400$	0.049	0.830	0.055	0.053	17.590	0.000
128	8	4	$T = 200$	-0.101	0.045	0.056	-0.102	-36.257	1.000
			$T = 400$	-0.016	0.453	0.081	-0.006	-3.962	1.000

B Commodity price data and long-term trends

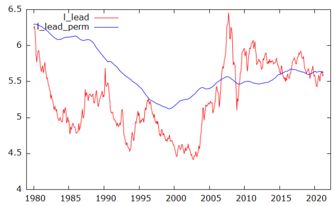
Figure 3: Base metals log prices and Permanent components



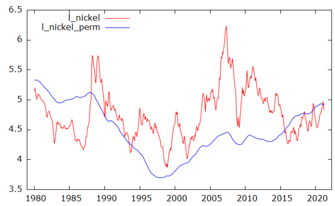
(a) Aluminium



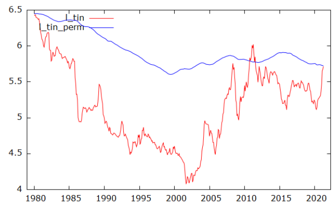
(b) Copper



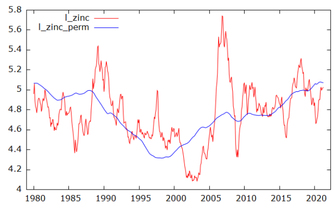
(c) Lead



(d) Nickel

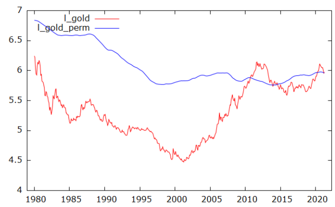


(e) Tin

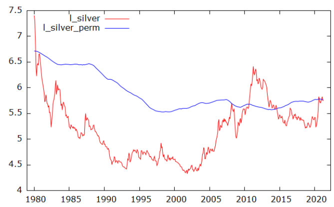


(f) Zinc

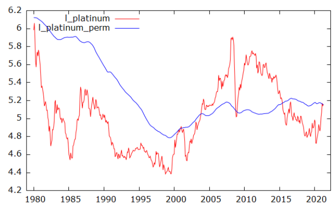
Figure 5: Precious metals log prices and Permanent components



(a) Gold

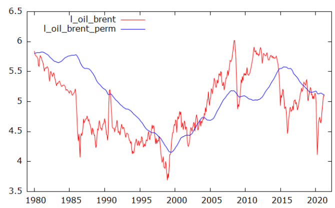


(b) Silver

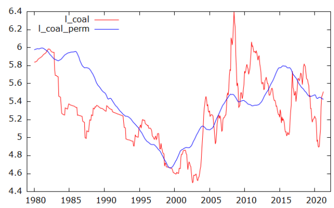


(c) Platinum

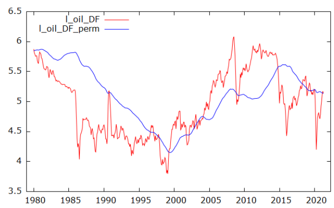
Figure 7: Energy log prices and Permanent components



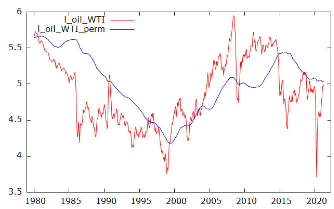
(a) Crude oil (brent)



(b) Coal

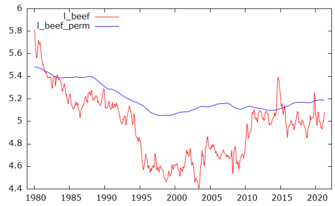


(c) Crude oil (Dubai Fateh)

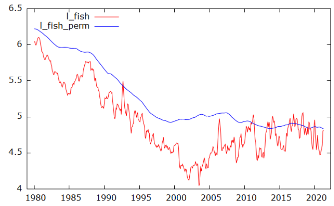


(d) Crude oil (WTI)

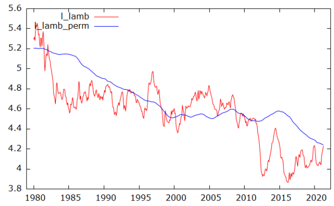
Figure 9: Livestock log prices and Permanent components



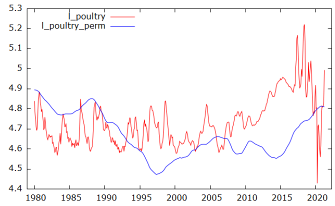
(a) Beef



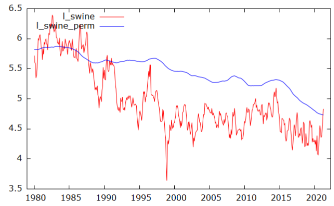
(b) Fish



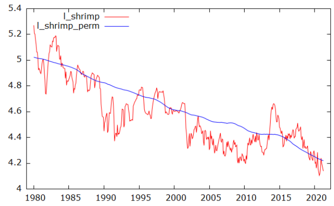
(c) Lamb



(d) Poultry

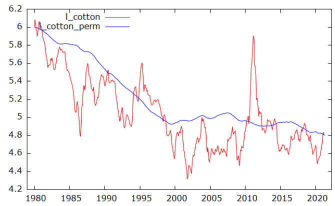


(e) Swine

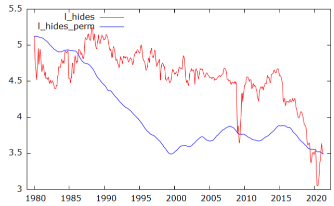


(f) Shrimp

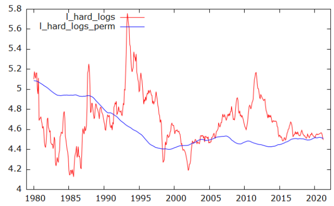
Figure 11: Raw materials log prices and Permanent components



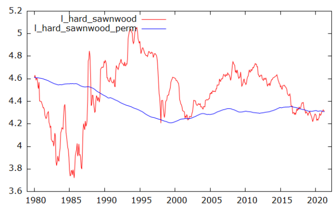
(a) Cotton



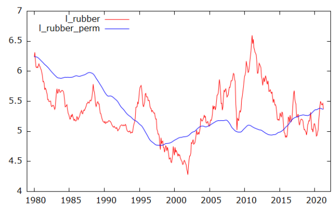
(b) Hides



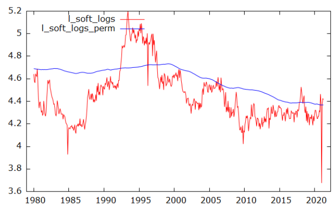
(c) Hard logs



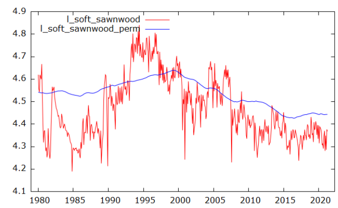
(d) Hard sawnwood



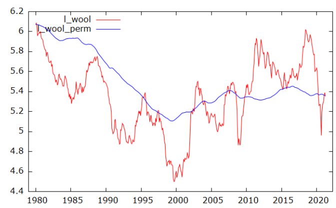
(e) Rubber



(f) Soft logs

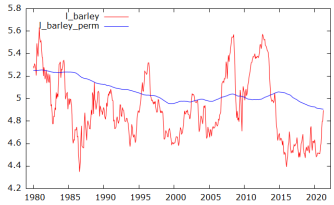


(g) Soft sawnwood

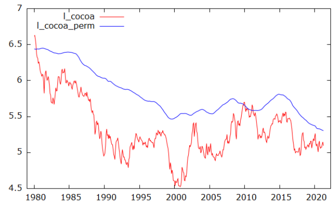


(h) Wool

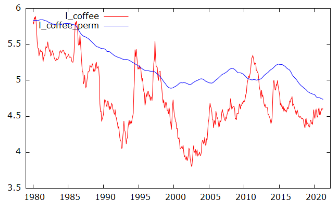
Figure 13: Food log prices and Permanent components



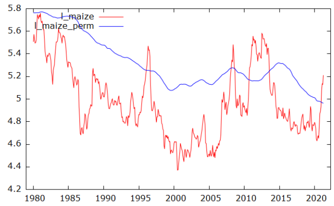
(a) Cotton



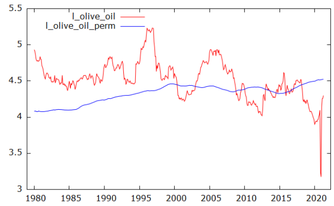
(b) Hides



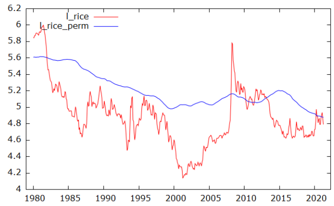
(c) Hard logs



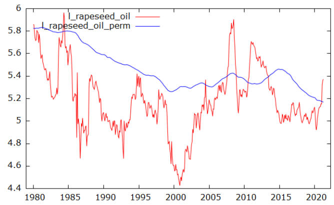
(d) Hard sawnwood



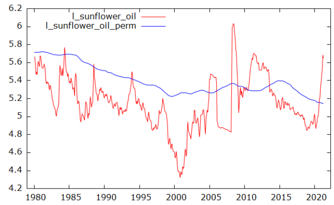
(e) Rubber



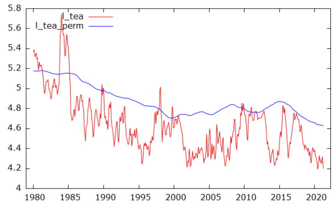
(f) Soft logs



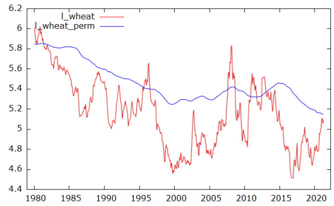
(g) Soft sawnwood



(h) Wool



(i) Soft sawnwood



(j) Wool

C Commodity price data

Table 29: Data description, IMF database of primary commodity prices

	Commodity price description
Aluminium	99.5% minimum purity, LME spot price, CIF UK ports, US$ per metric ton
Barley	Canadian no.1 Western Barley, spot price, US$ per metric ton
Beef	Australian and New Zealand 85% lean fores, CIF U.S. import price, US cents per pound
Coal	Australian thermal coal, 12,000- btu/pound, less than 1% sulfur, 14% ash, FOB Newcastle/Port Kembla, US$ per metric ton
Cocoa	Cocoa beans, International Cocoa Organization cash price, CIF US and European ports, US$ per metric ton
Coffee	Robusta, International Coffee Organization New York cash price, ex-dock New York, US cents per pound
Rapeseed oil	Crude, fob Rotterdam, US$ per metric ton
Copper	Grade A cathode, LME spot price, CIF European ports, US$ per metric ton
Cotton	Cotton Outlook 'A Index', Middling 1-3/32 inch staple, CIF Liverpool, US cents per pound
Hides	Heavy native steers, over 53 pounds, wholesale dealer's price, US, Chicago, fob Shipping Point, US cents per pound
Lamb	Frozen carcass Smithfield London, US cents per pound
Lead	99.97% pure, LME spot price, CIF European Ports, US$ per metric ton
Soft Logs	Average Export price from the U.S. for Douglas Fir, US$ per cubic meter
Hard Logs	Best quality Malaysian meranti, import price Japan, US$ per cubic meter
Maize	U.S. No.2 Yellow, FOB Gulf of Mexico, U.S. price, US$ per metric ton
Nickel	Melting grade, LME spot price, CIF European ports, US$ per metric ton
Crude oil	1) Crude Oil (petroleum), Dated Brent, light blend 38 API, fob U.K., US$ per barrel 2) Crude Oil (petroleum), Dubai Fateh Fateh 32 API, US$ per barrel 3) Crude Oil (petroleum), West Texas Intermediate 40 API, Midland Texas, US$ per barrel
Olive oil	Extra virgin less than 1% free fatty acid, ex-tanker price U.K., US$ per metric ton
Swine	51-52% lean Hogs, U.S. price, US cents per pound
Poultry	Whole bird spot price, Ready-to-cook, whole, iced, Georgia docks, US cents per pound
Rice	5 percent broken milled white rice
Rubber	Singapore Commodity Exchange, No. 3 Rubber Smoked Sheets, 1st contract, US cents per pound
Fish	Farm Bred Norwegian Salmon, export price, US$ per kilogram
Hard Sawnwood	Dark Red Meranti, select and better quality, C&F U.K port, US$ per cubic meter
Soft Sawnwood	Average export price of Douglas Fir, U.S. Price, US$ per cubic meter
Shrimps	Thailand Whiteleg Shrimp 70 Shrimps/Kg Spot Price
Sunflower oil	US export price from Gulf of Mexico, US$ per metric ton
Tea	Mombasa, Kenya, Auction Price, US cents per kilogram, From July 1998,Kenya auctions, Best Pekoe Fannings. Prior, London auctions, c.i.f. U.K. warehouses
Tin	Standard grade, LME spot price, US$ per metric ton
Wheat	No.1 Hard Red Winter, ordinary protein, Kansas City, US$ per metric ton
Wool	Coarse, 23 micron, Australian Wool Exchange spot quote, US cents per kilogram
Zinc	High grade 98% pure, US$ per metric ton
Gold	Fixing Committee of the London Bullion Market Association, London 3 PM fixed price, US$ per troy ounce
Silver	London Bullion Market Association, USD/troy ounce
Platinum	LME spot price, USD/troy ounce

Table 30: Blocks of commodity prices

Base Metals	Aluminium Copper Lead Nickel Tin Zinc
Precious Metals	Gold Silver Platinum
Energy	Crude oil (brent) Crude oil (Dubai Fateh) Crude oil (WTI) Coal
Livestock	Beef Fish Swine Poultry Lamb Shrimp
Raw Materials	Soft logs Cotton Hides Hard logs Rubber Hard sawnwood Soft sawnwood Wool
Food	Barley Cocoa Coffee Rapeseed oil Maize Olive oil Rice Sunflower oil Tea Wheat

FONDAZIONE ENI ENRICO MATTEI WORKING PAPER SERIES
"NOTE DI LAVORO"
Our Working Papers are available on the Internet at the following address:
http://www.feem.it/getpage.aspx?id=73&sez=Publications&padre=20&tab=1

"NOTE DI LAVORO" PUBLISHED IN 2021

1. 2021, Alberto Arcagni, Laura Cavalli, Marco Fattore, Partial order algorithms for the assessment of italian cities sustainability
2. 2021, Jean J. Gabszewicz, Marco A. Marini, Skerdilajda Zanaj, Random Encounters and Information Diffusion about Product Quality
3. 2021, Christian Gollier, The welfare cost of ignoring the beta
4. 2021, Richard S.J. Tol, The economic impact of weather and climate
5. 2021, Giacomo Falchetta, Nicolò Golinucci, Michel Noussan and Matteo Vincenzo Rocco, Environmental and energy implications of meat consumption pathways in sub-Saharan Africa
6. 2021, Carlo Andrea Bollino, Marzio Galeotti, On the water-energy-food nexus: Is there multivariate convergence?
7. 2021, Federica Cappelli, Gianni Guastella, Stefano Pareglio, Urban sprawl and air quality in European Cities: an empirical assessment
8. 2021, Paolo Maranzano, Joao Paulo Cerdeira Bento, Matteo Manera, The Role of Education and Income Inequality on Environmental Quality. A Panel Data Analysis of the EKC Hypothesis on OECD
9. 2021, Iwan Bos, Marco A. Marini, Riccardo D. Saulle, Myopic Oligopoly Pricing
10. 2021, Samir Cedic, Alwan Mahmoud, Matteo Manera, Gazi Salah Uddin, Information Diffusion and Spillover Dynamics in Renewable Energy Markets
11. 2021, Samir Cedic, Alwan Mahmoud, Matteo Manera, Gazi Salah Uddin, Uncertainty and Stock Returns in Energy Markets: A Quantile Regression Approach
12. 2021, Sergio Tavella, Michel Noussan, The potential role of hydrogen towards a low-carbon residential heating in Italy
13. 2021, Maryam Ahmadi, Matteo Manera, Oil Prices Shock and Economic Growth in Oil Exporting Countries
14. 2021, Antonin Pottier, Emmanuel Combet, Jean-Michel Cayla, Simona de Lauretis, Franck Nadaud, Who emits CO_2? Landscape of ecological inequalities in France from a critical perspective
15. 2021, Ville Korpela, Michele Lombardi, Riccardo D. Saulle, An Implementation Approach to Rotation Programs
16. 2021, Miguel Borrero, Santiago J. Rubio, An Adaptation-Mitigation Game: Does Adaptation Promote Participation in International Environmental Agreements?
17. 2021, Alan Finkelstein Shapiro, Gilbert E. Metcalf, The Macroeconomic Effects of a Carbon Tax to Meet the U.S. Paris Agreement Target: The Role of Firm Creation and Technology Adoption
18. 2021, Davide Bazzana, Jeremy Foltz, Ying Zhang, Impact of climate smart agriculture on food security: an agent-based analysis
19. 2021, Chiara Casoli, Riccardo (Jack) Lucchetti, Permanent-Transitory decomposition of cointegrated time series via Dynamic Factor Models, with an application to commodity prices

Fondazione Eni Enrico Mattei
Corso Magenta 63, Milano – Italia

Tel. +39 02.520.36934
Fax. +39.02.520.36946

E-mail: letter@feem.it
www.feem.it

FEEM
FONDAZIONE ENI
ENRICO MATTEI