UNIVERSIDAD CARLOS III DE MADRID working papers

Working Paper 12-23
Statistics and Econometrics Series 17
October 2012

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

MORE IS NOT ALWAYS BETTER: BACK TO THE KALMAN FILTER IN DYNAMIC FACTOR MODELS

P. Poncela and E. Ruiz*

Abstract

In the context of dynamic factor models (DFM), it is known that, if the cross-sectional and time dimensions tend to infinity, the Kalman filter yields consistent smoothed estimates of the underlying factors. When looking at asymptotic properties, the cross-sectional dimension needs to increase for the filter or stochastic error uncertainty to decrease while the time dimension needs to increase for the parameter uncertainty to decrease. In this paper, assuming that the model specification is known, we separate the finite sample contribution of each of both uncertainties to the total uncertainty associated with the estimation of the underlying factors. Assuming that the parameters are known, we show that, as far as the serial dependence of the idiosyncratic noises is not very persistent and regardless of whether their contemporaneous correlations are weak or strong, the filter un-certainty is a non-increasing function of the cross-sectional dimension. Furthermore, in situations of empirical interest, if the cross-sectional dimension is beyond a relatively small number, the filter uncertainty only decreases marginally. Assuming weak contemporaneous correlations among the serially uncorrelated idiosyncratic noises, we prove the consistency not only of smooth but also of real time filtered estimates of the underlying factors in a simple case, extending the results to non-stationary DFM. In practice, the model parameters are un-known and have to be estimated, adding further uncertainty to the estimated factors. We use simulations to measure this uncertainty in finite samples and show that, for the sample sizes usually encountered in practice when DFM are fitted to macroeconomic variables, the contribution of the parameter uncertainty can represent a large percentage of the total uncertainty involved in factor extraction. All results are illustrated estimating common factors of simulated time series.

Keywords: Common factors, Cross-sectional dimension, Filter uncertainty, Parameter uncertainty, Steady-state.

*Poncela, Pilar. Departamento de Análisis Económico: Economía Cuantitativa, Universidad Autónoma de Madrid, 28049 Cantoblanco (Madrid), email: pilar.poncela@uam.es
Ruiz, Esther. Departamento de Estadística and Instituto Flores de Lemus, Universidad Carlos III de Madrid, C/ Madrid 126, 28903 Getafe (Madrid), e-mail: ortega@est-econ.uc3m.es

More is not always better: back to the Kalman filter in Dynamic Factor Models

Pilar Poncela
Universidad Autónoma de Madrid
pilar.poncela@uam.es

Esther Ruiz
Universidad Carlos III de Madrid
ortega@est-econ.uc3m.es

October 14, 2012

Abstract

In the context of dynamic factor models (DFM), it is known that, if the cross-sectional and time dimensions tend to infinity, the Kalman filter yields consistent smoothed estimates of the underlying factors. When looking at asymptotic properties, the cross-sectional dimension needs to increase for the filter or stochastic error uncertainty to decrease while the time dimension needs to increase for the parameter uncertainty to decrease. In this paper, assuming that the model specification is known, we separate the finite sample contribution of each of both uncertainties to the total uncertainty associated with the estimation of the underlying factors. Assuming that the parameters are known, we show that, as far as the serial dependence of the idiosyncratic noises is not very persistent and regardless of whether their contemporaneous correlations are weak or strong, the filter uncertainty is a non-increasing function of the cross-sectional dimension. Furthermore, in situations of empirical interest, if the cross-sectional dimension is beyond a relatively small number, the filter uncertainty only decreases marginally. Assuming weak contemporaneous correlations among the serially uncorrelated idiosyncratic noises, we prove the consistency not only of smooth but also of real time filtered estimates of the underlying factors in a simple case, extending the results to non-stationary DFM. In practice, the model parameters are unknown and have to be estimated, adding further uncertainty to the estimated factors. We use simulations to measure this uncertainty in finite samples and show that, for the sample sizes usually encountered in practice when DFM are fitted to macroeconomic variables, the contribution of the parameter uncertainty can represent a large percentage of the total uncertainty involved in factor extraction. All results are illustrated estimating common factors of simulated time series.

1 Introduction

Dynamic factor models (DFM), originally introduced by Geweke (1977) and Sargent and Sims (1977), are designed to reduce the dimensionality of large systems of multivariate time series by assuming that there is a small number of underlying states common to the variables in the system. The main information contained in the variables is, consequently, summarized by the underlying states or factors. DFM have been implemented with many different goals; see Stock and Watson (2011) for a recent survey. In some cases, the factors have a direct economic interpretation as in Stock and Watson (2010) who fit a DFM to obtain common and regional factors of the building permits of new residential units. Another popular implementation of DFM consists on entering the estimated factors into fairly simple regression models to predict key macroeconomic variables as the Gross Domestic Product (GDP) or inflation; see Stock and Watson (2006) and Eickmeier and Ziegler (2008). Dynamic factor models have also been implemented to obtain business cycle indicators with the business cycle generally represented by the common component of the series; see Artis *et al.* (2004), Arouba *et al.* (2009), Altissimo *et al.* (2010) and Camacho and Perez-Quiros (2010), among others. DFM also have a long tradition in the context of financial variables. In this context, Ross (1976) proposes the first theoretically grounded multifactor model in asset pricing; see Chamberlain and Rothschild (1983). DFM have also been fitted to represent comovements among the volatilities of different financial assets as in Diebold and Nerlove (1989) and Harvey *et al.* (1994) who propose incorporating common factors in multivariate GARCH and Stochastic Volatility models, respectively. There are also financial applications to find common factors in interest rates of different maturities; see, for example, Moon and Perron (2007) and Jungbacker *et al.* (2009). Recently, Timmermann (2008) implements DFM to extract factors in order to forecast returns. Finally, the recent field of macro-finance has also relied on the estimation of factors from bond yields in order to improve the performance of small-sample structural macroeconomic models or from macroeconomic factors to estimate the term structure; see, for example, Ang and Piazzesi (2003), Forni *et al.* (2003) and Koopman and van der Wel (2011).

Looking at the wide range of applications described above, it seems obvious that DFM play a central role in modern econometrics. Whatever the application, the underlying factors are unobserved and need to be estimated. Assuming that their number is known, the original procedure to extract them from a set of observed variables was based on using the Kalman filter and smoothing algorithms after expressing the DFM as a state space model. These algorithms require knowledge of the model parameters which, in practice, were estimated by Maximum Likelihood (ML) maximizing the one-step-ahead decomposition of the log-Gaussian likelihood; see Engle and Watson (1981) and Watson and Engle (1983) for some early references. However, the maximization of the log-likelihood entails nonlinear optimization which restricts the number of parameters that can be estimated and, consequently, the number of series that can be handled when estimating the underlying factors. Recently, Jungbacker and Koopman (2008) have proposed computationally efficient procedures for the ML estimation of the parameters and factors based on univariate Kalman filter and smoothing (KFS) methods. Alternatively, the EM algorithm allows to maximize the likelihood function of very large DFM; see Shumway and Stoffer (1982) and Watson and Engle (1983) who propose the use of the EM algorithm in the context of state-space models. Other alternative procedures that allow to deal with the large systems of variables often available in practice without imposing restrictions on the specification of the idiosyncratic noises, have been proposed; see, for example, Kapetanios and Marcelino (2009) who propose a computationally heavy subspace algorithm that allows the factors to be estimated without specifying and identifying the full state-space model. These computational difficulties have limited the implementation of KFS procedures to extract the factors in favour of Principal Components (PC) procedures that require weak cross-correlations of the idiosyncratic noises; see, for example, Stock and Watson (2002) and Forni *et al.* (2000, 2005). Due to its wide popularity, PC has been extended in several directions. For example, Breitung and Tenhofen (2011) and Choi (2012) propose using Generalized Least Squares (GLS) procedures to estimate the factors in the presence of heteroscedastic noises while Bai and Ng (2004) and Lam *et al.* (2011) propose procedures to deal with nonstationary systems. Finally, Doz *et al.* (2011) propose combining the Kalman filter and PC approaches by first extracting the common factors by PC, then estimating the DFM parameters by Ordinary Least Squares (OLS) implemented to the extracted factors and, finally, using the smoothing algorithm of the Kalman filter with the estimated parameters to extract the factors.

In spite of the computational burden involved in the estimation of the DFM parameters, the Kalman filter has several advantages when implemented to extract the factors. First, it allows to handle data irregularities, as mixed frequencies or missing data. Second, it can be implemented in real time as individual data are released; see Stock and Watson (2011). Third, it provides a framework for incorporating restrictions derived from economic theory; see Doz *et al.* (in press). Fourth, the Kalman filter is more efficient than PC for a flexible range of specifications that include non-stationary DFM and idiosyncratic noises with strong cross-correlations. Finally, it allows obtaining uncertainty measures associated with the estimated factors when the cross-sectional dimension is finite. In contrast, PC procedures allow to handle systems with a large number of variables and do not rely on parametric assumptions on the dynamics of the factors and idiosyncratic noises. However, it is important to point out that the finite sample performance of the PC estimator is poor when the explanatory power of the factors does not strongly dominate the explanatory power of the idiosyncratic noises; see Onatski (in press). Further, Stock and Watson (2002) also find deterioration in PC performance when the idiosyncratic noises have strong serial correlations and to a less extend when there is heteroscedasticity or when the serial correlation of the factors is large.

Regardless of the procedure used to extract the common factors, the literature dealing with their statistical properties has focused on asymptotic results obtained when both the cross-sectional and time dimensions of the system tend to infinity. The temporal dimension needs to go to infinity for the estimates of the unknown parameters or coefficients of the model to converge to their population counterparts while the cross-sectional dimension needs to go to infinity for the uncertainty associated with the extraction procedure itself to decrease towards zero. In the context of PC, Stock and Watson (2002) show that the estimated factors are consistent while Bai (2003) derives their asymptotic Mean Square Error (MSE) which could be used as an approximation of the finite sample MSE. More recently, Choi (2012) derives the asymptotic distribution of a Generalized Principal Component estimator with smaller asymptotic variance. When the factors are extracted using the Kalman filter, Doz *et al.* (2011, in press) show that the smoothed estimates are consistent in stationary DFM when the true parameters are substituted by either OLS or QML estimates.

Given that both PC and the Kalman filter factor estimates are consistent as the cross-sectional dimension increases, including all available variables in the system seems to be a natural choice. Consequently, there is an increasing literature using DFM that incorporates a large number of variables; see, among many others, Stock and Watson (2002) fitting a DFM to a system consisting of 132 variables, Forni *et al.* (2003) who incorporate 447 variables, Amengual and Watson (2007) with 124 variables, Eickmeier (2009) with 173 variables, Altissimo *et al.* (2010) with 145 variables or Gupta and Kabundy (2011) with 267 variables. In contrast, several authors argue that, in real life problems, introducing many variables is not always a good strategy. For example, Bai and Ng (2002) and Watson (2003) show that the predictive precision of the common factors extracted using PC implemented to real and simulated data, does not increase when increasing the cross-sectional dimension beyond 40 or 50 variables, respectively. Later, Boivin and Ng (2006) point out that, in the context of predicting US GDP, if by adding an extra variable we are not adding information about the factor but rather simply extra cross-sectional correlation among the idiosyncratic disturbances, then the estimated factors deteriorate and their predictive precision is not necessarily increased. More recently, Caggiano *et al.* (2011) conclude that between 12 and 20 variables are enough to obtain the best performance when predicting euro area GDP using extracted factors; see also Banbura and Runstler (2011). Finally, Bai and Ng (2008) propose selecting the variables before performing PC and conclude that by doing so, the predictive performance of the estimated factors can increase with respect to that obtained when all available variables are included. Therefore, the debate about whether it is best to include all available variables or to select an appropriate subset when estimating unobserved factors using DFM is still open.

The objective of this paper is to analyze the uncertainty associated with the estimated common factors when the cross-sectional and temporal dimensions are finite. As mentioned above, in the context of PC, only asymptotic MSEs of the extracted factors are available. Instead, in this paper, we focus on the properties of the KFS procedures which allow us to obtain finite sample MSEs. When, as usual, the Kalman filter is run with the unknown parameters substituted by consistent estimates, the MSEs of the estimated factors have two sources of uncertainty, one stemming from the filtering process itself, also known in the literature as stochastic error uncertainty, and the other from the estimated parameters. In order to separate both sources of uncertainty, we analyze first the MSE of the estimated factors when the parameters are known and, then, the MSE when they are estimated using a consistent estimator. Furthermore, by focusing on the Kalman filter, we also contribute to the literature by considering a wide range of specifications of the factors and idiosyncratic noises and analyzing how different characteristics of the DFM affect the finite sample MSE. In particular, we consider idiosyncratic noises with weak and strong contemporaneous correlations and stationary and non-stationary specifications. Finally, we obtain expressions of the underlying uncertainty associated not only with smoothed but also with real time one-step-ahead and filtered estimates of the factors. We show that, as far as the idiosyncratic noises are serially uncorrelated and regardless of whether their contemporaneous correlations are weak or strong, the filter uncertainty is a non-decreasing function of the cross-sectional dimension. Furthermore, in situations of empirical interest, if the cross-sectional dimension is beyond a relatively small number, the filter uncertainty only decreases marginally. However, the limiting behavior of the MSE depends on the properties of the covariances of the idiosyncratic noises and the weights of the factors. Weak cross-correlations together with pervasive factors is a sufficient condition for the uncertainty of the filtered and smoothed factor estimates to converge to zero with the cross-sectional dimension. If this is the case, the factors can be consistently estimated using the Kalman filter even if the system contains a relatively small number of variables.

In practice, the model parameters are unknown and have to be estimated, adding further uncertainty to the estimated factors. However, the MSE obtained from the Kalman filter equations implemented with estimated parameters do not incorporate this further uncertainty and, consequently, subestimate the true uncertainty associated with the estimated factors; see, for example, Quenneville and Singh (2000) and Rodríguez and Ruiz (2012). In this paper, we measure the contribution of the parameter uncertainty in the total uncertainty when both the cross-sectional and temporal dimensions are finite. For this goal, we carry out Monte Carlo experiments incorporating the parameter uncertainty into the Kalman filter MSE using the proposal of Delle Monache and Harvey (2011). We show that, in cases of empirical interest, the parameter uncertainty could represent a large percentage of the total uncertainty associated with the estimation of the underlying factors. Furthermore, even for relatively large sample sizes, the parameter and total uncertainties could increase with respect to the cross-sectional dimension.

The rest of the paper is organized as follows. In section 2, we analyze the properties of the KFS estimators of the underlying factors in approximated DFM with known parameters. Section 3 is devoted to the DFM model with estimated parameters. Using simulations, we measure the uncertainty added to the estimation of the underlying factors in strict DFM when the parameters are estimated by ML with the likelihood maximized using the EM algorithm. Section 4 illustrates the results estimating the underlying factors of simulated data. Finally, section 5 concludes the paper.

2 Kalman filter with known parameters

In this section, we describe the DFM considered in this paper and derive the steady-state Kalman filter MSE when the cross-sectional dimension is finite and the model parameters and, consequently, the number of factors are known. Therefore, the focus is on the filter or stochastic error uncertainty. We consider alternative specifications of the underlying factors and of the idiosyncratic noises, including stationary and non-stationary models and weakly and strongly correlated idiosyncratic noises. We start by considering the simplest case when the number of factors is just one and then generalize the results to models with more than one factor.

2.1 Description of the dynamic factor model

In this subsection, we consider a DFM with a single factor; see Engle and Watson (1981), Stock and Watson (1991), Arouba *et al.* (2009) and Camacho and Perez-Quiros (2010) for empirical applications with just one factor. The underlying factor, F_t, is given by the following AR(1) model

$$F_t = \phi F_{t-1} + \eta_t, \tag{1}$$

where ϕ is the autoregressive parameter such that $-1 < \phi \leq 1$. The disturbance, η_t, is a Gaussian white noise process with variance σ_η^2. When $|\phi| < 1$, the factor is a zero mean, stationary process. This is the most popular DFM; see Stock and Watson (2011). On the other hand, if $\phi = 1$, the common factor is non-stationary and represents a stochastic level; see, for example, Peña and Poncela (2004, 2006), Moon and Perron (2007), Eickmeier (2009) and Lam *et al.* (2011) for non-stationary factors.

Alternatively, the factor can be assumed to have a fixed finite variance and, in this case, it is specified as follows

$$F_t = \phi F_{t-1} + (1 - \phi^2)^{1/2} \eta_t; \tag{2}$$

see, for example, Harvey and Streibel (1998) for a specification of underlying economic cycles using the specification in (2) and Onatski (in press) who imposes this restriction when estimating the factors.

It is important to emphasize that when approaching the unit root, the dynamic behavior of the factor is different depending on whether it is specified as in equations (1) or (2). When the factor is specified as in equation (1), its variability increases as ϕ approaches 1. In the limit, when $\phi = 1$, the factor is non-stationary. However, in (2), the variability of the factor is fixed and, consequently, as ϕ increases, the variance of the noise associated with the factor is smaller. In the limit, when $\phi = 1$, the variance is zero and, consequently, the factor is deterministic and observable (and estimated without filter error). This difference is going to have implications when computing the MSE of the factor estimators.

Dynamic factor models assume that the factor, F_t, is unobserved and affects the evolution of an indeterminate number of variables, denoted by y_{it}. Consider that the first N of these variables are observed by the econometrician and that the following DFM is used to estimate the factor

$$Y_t^{(N)} = \mu^{(N)} + P^{(N)} F_t + \varepsilon_t^{(N)} \tag{3}$$

where $Y_t^{(N)} = (y_{1t}, ..., y_{Nt})'$ and $\varepsilon_t^{(N)}$ is a $N \times 1$ vector of idiosyncratic noises which follows the following VAR(1) process

$$\varepsilon_t^{(N)} = \Gamma^{(N)} \varepsilon_{t-1}^{(N)} + a_t^{(N)} \tag{4}$$

where $a_t^{(N)}$ is a Gaussian white noise vector with finite and positive definite covariance matrix $\Sigma_a^{(N)}$. The idiosyncratic noises, $\varepsilon_t^{(N)}$, are independently distributed of $\eta_{t-\tau}$ for all leads and lags. The vector of constants is $\mu^{(N)} = (\mu_1, ..., \mu_N)'$ and $P^{(N)} = (p_1, ..., p_N)'$ is the factor loading vector. When the factor is stationary, the VAR(1) process in (4) can be either stationary or not; see, for example, Bai and Ng (2004) for DFM with unit roots both in the factors and the idiosyncratic components and Eickmeier (2009) and Chmelarova and Nath (2010) for empirical applications with nonstationary idiosyncratic noises to output and prices in the Euro area countries and US prices respectively. However, when the factor is non-stationary, we assume that the idiosyncratic noises are either stationary or they have at most $N-1$ unit roots. In this case, the DFM is observable and the Kalman filter reaches the steady-state; see Harvey (1989). Note that, if the factor is non-stationary and the idiosyncratic noises are stationary, the presence of a common factor means that the series are cointegrated; see, for instance, Escribano and Peña (1994).

There are several particular cases of the DFM in equations (3) and (4) that have attracted quite a lot of attention in the related literature. When $\Gamma^{(N)} = 0$ and $\Sigma_a^{(N)}$ is diagonal, the idiosyncratic noises are contemporaneously and serially independent. In this case, the DFM is known as strict; see Breitung and Eickemeir (2006). When there is serial correlation with $\Gamma^{(N)}$

being diagonal, the model is known as exact; see, Doz *et al.* (2011, in press). Chamberlain and Rothschild (1983) introduce the term "approximate factor structure" in static factor models where the idiosyncratic components do not need to have a diagonal variance-covariance matrix. Bai and Ng (2002), Stock and Watson (2002) and Forni *et al.* (2000, 2005) generalize the approximate factor model to the dynamic case, allowing for weak cross-correlation.

The DFM in equations (3) and (4), with F_t defined either as in (1) or (2), is conditionally Gaussian. Consequently, when the idiosyncratic noises are serially uncorrelated, the Kalman filter and smoothing algorithms provide Minimum MSE estimates of the underlying factors. On the other hand, when the idiosyncratic noises are serially correlated, the DFM can be reformulated in two alternative ways to preserve the optimal properties of the Kalman filter and smoother. First, considering that the factor is defined as in (1), it is possible to express the DFM in state space form as follows:

$$Y_t^{(N)} = \mu^{(N)*} + \Gamma^{(N)} Y_{t-1}^{(N)} + \begin{bmatrix} P^{(N)} & -\Gamma^{(N)} P^{(N)} \end{bmatrix} \begin{bmatrix} F_t \\ F_{t-1} \end{bmatrix} + a_t^{(N)}$$

$$\begin{bmatrix} F_t \\ F_{t-1} \end{bmatrix} = \begin{bmatrix} \phi & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_{t-1} \\ F_{t-2} \end{bmatrix} + \begin{bmatrix} \eta_t \\ 0 \end{bmatrix}, \qquad (5)$$

where $\mu^{(N)*} = \mu^{(N)}(I - \Gamma^{(N)})$; see Reis and Watson (2010) and Jungbacker *et al.* (2011) for implementations of the model in (5). The DFM in (5) can be also adapted to the factor specification in (2). Alternatively, one can deal with the autocorrelation of the idiosyncratic noises by augmenting the state vector by $\varepsilon_t^{(N)}$; see, for example, Jungbacker *et al.* (2011). Although both formulations lead to the same results when the initialization issues are properly accounted for, the latter enlarges the dimension of the state vector which can become too large when N is large. Consequently, when necessary, we consider the reformulation in (5) when dealing with DFM with known parameters.

It is well known that, in conditionally Gaussian models, as the DFM considered in this paper, the estimates of the underlying factor provided by the Kalman filter and smoothing algorithms are given by the corresponding conditional means. Denoting by $f_{t|\tau}$ the estimate of F_t obtained with the information available up to time τ, if the model parameters are known, the filter delivers

$$f_{t|\tau} = E\left[F_t|Y_1^{(N)}, ..., Y_\tau^{(N)}\right]$$

where $\tau = t-1$, for one-step-ahead estimates, $\tau = t$ for filtered estimates and $\tau = T$ for smoothed factor estimates. Therefore, by construction, regardless of the cross-sectional dimension, the filter delivers unbiased estimates of the factor. Note that, we cannot talk about misspecification when some of the variables affected by the underlying factor are not included in the model implemented for factor extraction. However, including more variables implies more information to estimate the factor and, consequently, under mild conditions, the MSEs of the factor estimates are expected to be non-increasing functions of the cross-sectional dimension. Next, we derive the MSE of $f_{t|\tau}$ as a function of the cross-sectional dimension and analyze whether this is the case. Because the filter is run in two different state space models depending on whether the idiosyncratic noises are or not serially correlated, we consider separately both cases.

2.2 Serially uncorrelated idiosyncratic noises

In this subsection, we analyze how the MSEs of $f_{t|\tau}$ depend on N in the DFM with serially uncorrelated idiosyncratic noises, i.e. when $\Gamma^{(N)} = 0$ in equation (4). Given that the system matrices are time-invariant, the Kalman filter reaches the steady-state in which the MSE of the one-step-ahead and filtered estimates are constant; see Harvey (1989). Note that when dealing with smoothed estimates, their MSEs are also constant in the middle of the sample. From now on, we focus on steady-state MSE.

Consider first, the steady-state MSE of the one-step-ahead estimates of the underlying factor, denoted by $V(N)$, which is obtained after solving the following Riccati equation

$$V(N) = \phi^2 \left[V(N) - V(N)P^{(N)\prime} \left(P^{(N)} V(N) P^{(N)\prime} + \Sigma_\varepsilon^{(N)} \right)^{-1} P^{(N)} V(N) \right] + k\sigma_\eta^2 \quad (6)$$

where $\Sigma_\varepsilon^{(N)} = \Sigma_a^{(N)}$ and $k = 1$ when the factor is specified as in equation (1) while $k = (1 - \phi^2)$ if it is defined as in (2). Note that, as in any time-invariant model, the filter uncertainty of one-step-ahead estimated factors is independent of the particular data available; see Harvey (1989). However, it depends on the dynamic properties of the factors through ϕ and $k\sigma_\eta^2$, on the variances and contemporaneous correlations between the idiosyncratic noises through $\Sigma_\varepsilon^{(N)}$ and on the factor loadings that appear in the vector $P^{(N)}$. The cross-sectional dimension affects the steady-state MSE through these last two terms.

The following lemma establishes the solution of the Riccati equation.

Lemma 1. Given the DFM in (3) with the factor defined either as in expressions (1) or (2), and $\varepsilon_t^{(N)}$ being a serially uncorrelated vector process with contemporaneous covariance matrix given by $\Sigma_\varepsilon^{(N)}$, non necessarily diagonal, the one-step-ahead steady-state MSE is given by the solution of the Riccati equation in (6) which is given by

$$V(N) = \frac{k\sigma_\eta^2 Q(N) - 1 + \phi^2 + \sqrt{\left(k\sigma_\eta^2 Q(N) - 1 + \phi^2\right)^2 + 4k\sigma_\eta^2 Q(N)}}{2Q(N)} \quad (7)$$

where $Q(N) = P^{(N)\prime}\Omega^{(N)}P^{(N)}$ with $\Omega^{(N)} = \left(\Sigma_\varepsilon^{(N)}\right)^{-1}$. Furthermore, $V(N+1) = V(N)$ when adding an additional variable, if i) $\phi = 0$, or ii) F_t is given by (2) with $\phi = 1$, or iii) $p_{N+1} = 0$ and its corresponding idiosyncratic noise, ε_{N+1}, is not correlated with any of the other N variables previously included in the system. Otherwise, $V(N+1) < V(N)$.

Proof. See the Appendix. □

Lemma 1 establishes that, when the factor is either deterministic or white noise, the MSE of $f_{t|t-1}$ does not depend on the cross-sectional dimension, being zero in the first case and σ_η^2 in the second. In all other cases, adding a new variable to the system never decreases the precision in the estimation of the underlying factor. It is important to remark that this result is satisfied regardless of whether the cross-correlations between the idiosyncratic noises are weak or strong or whether the factor is stationary or not. Furthermore, when adding an additional variable with zero weight, if this variable is correlated with the variables already included in the model, the steady-state MSE also decreases. This result somehow contradicts the conclusion of Boivin and Ng (2006) about the deterioration of factor estimates when adding an extra variable which is not adding information about the factor but rather simply extra cross-sectional correlation among the idiosyncratic disturbances. We can guess that their conclusion could be attributed to the estimation method. Finally, note that, in the particular case of a strict DFM, in which there is not contemporaneous correlation among the idiosyncratic noises, i.e. $\Sigma_\varepsilon^{(N)}$ is diagonal, $Q(N) = \sum_{i=1}^{N} q_i$ with $q_i = \frac{p_i^2}{\sigma_i^2}$.

As we mentioned above, for a given cross-sectional dimension, N, the steady-state one-step-ahead MSE also depends on the dynamics of the underlying factor. First, it is obvious that for fixed ϕ, $V(N)$ always increases with σ_η^2. On the other hand, when the underlying factor is defined as in (1) and σ_η^2 is fixed, the precision of $f_{t|t-1}$ decreases as ϕ increases. This could be expected as the variability of the factor, is larger as it approaches a random walk. Therefore, as ϕ is larger, more variables are needed as to estimate the underlying factor with a given precision. If the underlying factor is defined as in (2), it is obvious that $V(N)$ decreases as ϕ increases given that the factor is closer to be deterministic.

Consider now the limiting behaviour of $V(N)$ when N tends to infinity. There are situations in which, by definition, there is a finite number of variables, N^*, that depend on the factor; see, for example, the factor model in Chamberlain and Rothschild (1983). In these cases, according to (7), the minimum MSE is given by

$$V(N^*) = 0.5\left(k\sigma_\eta^2 - 1 + \phi^2\right)\left(1 - \sqrt{1 + \frac{4k\sigma_\eta^2\overline{Q}}{\left(k\sigma_\eta^2\overline{Q} - 1 + \phi^2\right)^2}}\right), \quad (8)$$

where $\overline{Q} = Q(N^*)$. The MSE in (8) is always larger than $k\sigma_\eta^2$ for finite $\overline{Q}$.

Alternatively, there are applications in which there are potentially infinity variables that depend on the factor. In this case, the limiting behavior of $V(N)$ depends on whether the contemporaneous correlations between the idiosyncratic noises are weak or strong. Note that the steady-state MSE of the one-step-ahead estimates of the underlying factor, in expression (7), depends on the cross-sectional dimension, N, through the term $Q(N)$. Therefore, the limiting behavior of $V(N)$ depends on the limiting behavior of $Q(N)$. From expression (7), it is straightforward to show that, if $Q(N)$ converges to infinity with N, then the steady-state MSEs of the one-step-ahead factor estimates tend to $k\sigma_\eta^2$, the variance of the noise in the factor equation. Notice that this result could be expected as the one-step-ahead predictions of the underlying factor always involve the uncertainty associated with η_t. However, when $Q(N)$ converges to a constant, $\overline{Q}$, then the limit of the MSE is given by expression (8).

The following lemma establishes the conditions for $Q(N)$ to diverge and, consequently, the steady-state MSE of $f_{t|t-1}$ to converge to its minimum $k\sigma_\eta^2$.

Lemma 2. Let $g_N^2 = \sigma_{N|N-1}^{-2}\left(p_N - P^{(N-1)\prime}\Omega^{(N-1)}\Sigma_{N,N-1}\right)^2$ where $\sigma_{N|N-1}^2 = \sigma_N^2 - \Sigma_{N,N-1}'\Omega^{(N)}\Sigma_{N,N-1}$ is the variance of ε_N conditional on ε_i for $i = 1, ..., N-1$ and $\Sigma_{N,N-1}$ is the $N-1$ vector containing the covariances between ε_N and ε_i for $i = 1, ..., N-1$ and let $Q(N)$ be defined as in expression (7). If $\lim\limits_{N\to\infty} \frac{g_{N+1}^2}{g_N^2} = l$ exists and $l > 1$, then

$$\lim_{N\to\infty} Q(N) = \infty.$$

Proof. In the proof of lemma 1, we show that $Q(N+1) = Q(N) + g_N^2$. Therefore, $Q(N)$ is a series of positive terms and the result of the lemma is obtained as a direct consequence of the D'Alembert criterion; see, for example, Piskunov (1969). □

Lemma 2 has several important implications for the empirical implementation of DFM. First, it is important to note that the usual assumption in the large DFM literature about the idiosyncratic noises being weakly correlated and the factors being pervasive, i.e., their cumulative loadings on N cross-sectional variables rising proportional to N, is a sufficient condition for $Q(N)$ to go to infinity with N. In this case, when $N \to \infty$, $\frac{1}{N}\sum_{i=1}^{N} p_i^2$ is asymptotically larger than the maximum eigenvalue of $\Sigma_\varepsilon^{(N)}$ and the explanatory content of the factor strongly dominates the explanatory content of the idiosyncratic noises; see, for example, Onatski (in press). Second, note that, in any case, the condition in lemma 2 for $Q(N)$ to go to infinity is sufficient but not necessary. If $\lim\limits_{N\to\infty} \frac{g_{N+1}^2}{g_N^2}$ does not exist or is equal to 1, the series $Q(N)$ can either converge or diverge when the number of variables increase, and it is necessary to use alternative criteria to solve the problem. Third, from an empirical point of view, we expect the variables being introduced in the model according to a criterion that implies some kind of ordering. Consider, for example, situations in which the variables are introduced according to their explanatory content with respect to the factor so that the less pervasive variables are those introduced later in the model. In this case, we expect the conditions in lemma 2 for $Q(N)$ to go to infinity with N not to be satisfied and the MSE is never as small as $k\sigma_\eta^2$.

Next, we illustrate the results of lemmas 1 and 2 focusing on the issue of how many variables should be included in the system used to extract the factors and the role of the contemporaneous correlations on the steady-state MSE of $f_{t|t-1}$. For this purpose, we first consider a strict DFM with the factor specified as in equation (1) with $q_i = 1$, $i = 1, ..., N$ $\sigma_\eta^2 = 1$ and $\phi = 0.4$, 0.8 and 1; see Camacho *et al.* (2012) for the empirical adequacy of these parameter values. In this case, the minimum MSE of the one-step-ahead estimates of the factor is 1. Using expression (7), we can see that, if $\phi = 0.4$,

$N = 7$ variables are needed to obtain a MSE which is 2% larger than the minimum attainable when $N = \infty$. However, if $\phi = 0.8$, $N = 30$ variables are needed for a similar precision. Finally, when $\phi = 1$, $N = 50$ variables are required for the MSE to be a 2% larger than the minimum. Therefore, if the factor is specified as in equation (1), more variables are needed to extract it with a given precision as ϕ approaches the unity. This result is illustrated in the top left panel of figure 1 that displays the steady-state MSE in (7) as a function of the cross-sectional dimension, N. We can observe that the larger is ϕ, the larger the steady-state MSEs of $f_{t|t-1}$ and a larger number of variables is required to converge. However, even when $\phi = 1$, if the number of variables in the system is around 30, adding additional variables only decreases the MSE marginally. It is surprising to observe that this number of variables is similar to those found by Bai and Ng (2002), Watson (2003) Caggiano *et al.* (2011) and Banbura and Runstler (2011) in the context of real and simulated data when the factor is estimated by PC. On the other hand, when the autoregressive dependence of the underlying factor is small, $\phi = 0.4$, the MSEs are very close to $\sigma_\eta^2 = 1$ regardless of the number of variables included in the system. In this case, using the information of approximately 5 variables, the steady-state MSE is already quite close to its minimum. For comparison purposes, the steady-state MSE of the one-step-ahead estimates of F_t defined as in equation (2), are plotted in the right column of figure 1 for the same parameters as above. We can observe an apparently different picture but with similar conclusions about the number of variables. In this case, it is important to remember that the variability of the factor decreases as ϕ increases. Therefore, the steady-state MSEs are smaller as ϕ increases and their limit is $(1 - \phi^2)$ when the cross-sectional dimension tends to infinity. When $\phi = 1$, the factor is deterministic and constant and, consequently, the MSE is obviously zero regardless of the cross-sectional dimension, N; see figure 1 and lemma 1. Once more, figure 1 illustrates that this limit is reached for a relatively small number of time series.

Figure 1 about here

Next, we analyze the effects of the presence of contemporaneous correlations on the previous conclusions. For this purpose, we consider a DFM with contemporaneously correlated idiosyncratic noises. The factor is defined as in (1) with the values of the parameters ϕ and σ_η^2 as above. The relative weights, q_i, are also the same as above. The idiosyncratic noises have weak cross-correlations with covariance matrix, $\Sigma_\varepsilon^{(N)} = \Sigma_a^{(N)}$, given by the following Toeplitz matrix

$$\Sigma_\varepsilon^{(N)} = \begin{bmatrix} 1 & b & \cdots & b^{N-1} \\ b & 1 & \cdots & b^{N-2} \\ \vdots & \vdots & \ddots & \vdots \\ b^{N-1} & b^{N-2} & \cdots & 1 \end{bmatrix}. \tag{9}$$

where $b = 0.5$. Therefore, the ijth element of the correlation matrix of the idiosyncratic noises is given by $0.5^{(j-i)}$ which implies that the correlation between any two noises is weaker as they are further apart in the vector. The left column of figure 2 plots the corresponding steady-state MSE. Comparing the left columns of figures 1 and 2, we can observe that the steady-state MSE are approximately the same although, when there is cross-correlation, they decay slightly slower than in the DFM with uncorrelated idiosyncratic noises. Therefore, adding weak cross-correlations only increases the uncertainty marginally and the main conclusions are the same.

Finally, we consider a DFM with strong cross-correlations among the idiosyncratic noises. The same DFM is again considered but now all out-diagonal elements of $\Sigma_\varepsilon^{(N)}$ are equal to 0.5. The right column of figure 2 plots the corresponding steady-state MSE in equation (7) as a function of N. We can observe that, as established in lemma 2, the MSEs do not converge to σ_η^2. This result is in concordance with the results in Onatski (in press) for PC factor extraction when the explanatory power of the factors does not strongly dominate the explanatory power of the idiosyncratic noises.

Figure 2 about here

Now, we consider the properties of real time steady-state MSE of the filtered estimates which are established in the following lemma.

Lemma 3. Given the DFM in (3) with the factor defined either as in expressions (1) or (2), and $\varepsilon_t^{(N)}$ being a serially uncorrelated vector process with contemporaneous covariance matrix given by $\Sigma_\varepsilon^{(N)}$, non necessarily diagonal, the filtered steady-state MSE is given by

$$W(N) = \frac{V(N)}{1 + V(N)Q(N)}. \tag{10}$$

Furthermore, when adding an additional variable, $W(N+1) = W(N)$ if i) $\phi = 0$, or ii) F_t is given by (2) with $\phi = 1$, or iii) $p_{N+1} = 0$ and its corresponding idiosyncratic noise, ε_{N+1}, is not correlated with any of the other N variables previously included in the system. Otherwise, $W(N+1) < W(N)$.

Proof. See the Appendix. $\square$

It is obvious that when the factor is defined as in (2) and $\phi = 1$, the steady-state MSE of the updated estimates is trivially equal to zero. Furthermore, when the factor is white noise or the added variables do not incorporate new information in the system, the real time filtered MSEs are also constant. In all other cases, the filtered uncertainty always decreases as more variables are used to estimate the underlying factor. Finally, note that the asymptotic behavior of $W(N)$ depends on the convergence of $Q(N)$. If $Q(N)$ tends to ∞ with N, $W(N)$ converges to zero and the filtered estimates of the underlying factor are consistent when the cross-sectional dimension tends to infinity. Therefore, the conditions for the consistency of the filtered estimates are the same as those established in lemma 2 for $Q(N)$ to diverge.

As an illustration of the performance of the filtered MSE, we consider, once more, the same DFM considered above. The second row of figure 1 plots the steady-state MSE of $f_{t|t}$ as a function of N in the strict DFM when the factor is specified as in (1) in the left column while the second column plots the same quantities when the factor is specified as in (2). This figure shows that, except when the factor is defined as in (2) and $\phi = 1$, the MSEs are very similar for all values of ϕ considered. The steady-state MSEs decrease very slowly for N larger than 20. Furthermore, when N is larger than 30, the MSEs are approximately equal to zero. For instance, if the factor is defined as in (1) and $\phi = 1$, the reduction in the MSE is approximately 25% when going from $N = 2$ to 20 variables while it is only about 3% when going from $N = 20$ to 30 variables. On the other hand, the second row of figure 2 plots the steady-state filtered MSE when the factor is defined as in (1) and the idiosyncratic noises have weak contemporaneous correlations in the left column and when they have strong correlations in the right column. We can see that by adding weak cross-correlation, the filtered MSEs increase with respect to those in the strict DFM. The MSEs still converge to zero with the cross-sectional dimension but the rate is slower. However, if the cross-correlations are strong, then the filtered MSEs do not converge to zero. It is important to note that in this latter case, having around 10 variables in the system already generates filtered MSEs which are very close to the minimum.

Finally, we analyze the properties of the MSE of the smoothed estimator of the underlying factor, $f_{t|T}$. Using the results in Harvey (1989), we can see that if the Kalman filter is in its steady-state, the smoother filter also has constant MSE in the middle of the sample, which is given by

$$S(N) = \frac{V(N)\left(1 + V(N)Q(N) - \phi^2\right)}{\left(1 + V(N)Q(N)\right)^2 - \phi^2}. \tag{11}$$

The following lemma establishes the non-increasing property of the steady-state MSE of $f_{t|T}$.

Lemma 4. Consider the DFM in (3) with the factor defined either as in expressions (1) or (2), and $\varepsilon_t^{(N)}$ being a serially uncorrelated vector process with contemporaneous covariance matrix given by $\Sigma_\varepsilon^{(N)}$. When adding an additional variable $S(N+1) = S(N)$ if: i) $\phi = 0$, or ii) F_t is given by (2) and $\phi = 1$, or iii) $p_{N+1} = 0$ and its corresponding idiosyncratic noise, ε_{N+1}, is not correlated with any of the other N variables previously included in the system. Otherwise, $S(N+1) = S(N)$.

Proof. See the Appendix. □

Note that if the factor is defined as in (2) with $\phi = 1$, then $V(N) = 0$ and, consequently, according to equation (11), the steady state MSEs of the smoothed estimates are also equal to zero. From expression (11), it is also straightforward to see that, if $Q(N)$ tends infinity with N, then $S(N)$ converges to zero as N goes to infinity. Therefore, the conditions in lemma 2 are sufficient for the smoothed estimates of the underlying factor to be consistent. This result can be compared with Doz *et al.* (2011) who prove the consistency of Kalman filter smoothed estimates assuming a more restrictive stationary DFM.

The results are finally illustrated for the same DFM considered above. The third row of figure 1 displays the MSE of $f_{t|T}$ as a function of N for the strict DFM while the third row of figure 2 plots the same quantities for the DFM with contemporaneously correlated idiosyncratic noises. The plots and conclusions are very similar to those obtained for the filtered estimates. There are no big gains in increasing the number of variables in the system beyond 30 variables.

2.3 Serially correlated idiosyncratic noises

As mentioned above, if the idiosyncratic noises are serially correlated, the Kalman filter is optimal when implemented in the state-space model in (5) in which the state vector is not scalar as it contains both F_t and F_{t-1}. In this case, the corresponding Riccati equation does not have a closed-form solution as a function of the parameters of the model; see Lancaster and Rodman (1995) and Rojas (2011) for solutions of the Riccati equation. In this subsection, we analyze the effects of the serial correlation of the idiosyncratic noises on the steady-state MSE by running the prediction equations of the Kalman filter until the steady-state is reached for several particular DFM. Note that, given the state vector in model (5), the Kalman filter prediction equations deliver both one-step-ahead and filtered MSE without requiring the simulation of the series. We also obtain the value of the smoothed MSE in the middle of the sample where they are constant.

The particular DFM considered include both stationary and non-stationary factors. The factor is specified as in equation (1) with $\sigma_\eta^2 = 1$. First, we consider a stationary factor with $\phi = 0.8$ and second a non-stationary factor with $\phi = 1$. The idiosyncratic noises are assumed to be contemporaneously uncorrelated, i.e. $\Sigma_a = I^{(N)}$, with the autoregressive matrix in equation (5) given by $\Gamma^{(N)} = \rho I^{(N)}$, where $I^{(N)}$ is the order N identity matrix and $\rho = 0$, 0.5 and 0.9. Note that when $\rho = 0$, the same strict DFM considered above is obtained for comparative purposes. As mentioned above, the existence of the steady-state requires the DFM to be observable. Consequently, the case when both the underlying factor and the idiosyncratic noises are random walks has been ruled out. Furthermore, we should note that, when both ρ and ϕ are close to one, the steady-state is only reached after a very large number of steps of the Kalman filter.

The left column of figure 3 plots the steady-state MSE as a function of the cross-sectional dimension for one-step-ahead, filtered and smoothed estimates of the underlying factor for the DFM with a stationary factor, i.e. when $\phi = 0.8$, while the right column plots the same quantities when the factor is non-stationary, i.e. $\phi = 1$. We can observe that, regardless of whether the factor is stationary or not, the steady-state MSE of $f_{t|\tau}$, $\tau = t-1$, t and T, are very similar when the idiosyncratic noises are serially uncorrelated and when they have moderate temporal dependences, i.e. when $\rho = 0.5$. The conclusions about the number of variables needed in order to have estimates of the underlying factor with a precision close to the maximum are very similar to those obtained in the previous section. However, when the idiosyncratic noises are very persistent, i.e. $\rho = 0.9$, and, consequently, they are close to be non-stationary, we observe much larger MSE for each cross-sectional dimension. Furthermore, the convergence of the MSE towards their minimum is slower than when the idiosyncratic errors are moderately autocorrelated. Therefore, to have the same precision in the estimation of the factor, it is necessary to introduce in the model a larger number of variables. This result could be expected given that, if both the factor and the idiosyncratic noises are highly persistent, it could be difficult for the filter to distinguish between them.

Figure 3 about here

Summarizing, it seems that unless the idiosyncratic noises are highly persistent, the previous conclusions about the number of variables to be included in the system to estimate the underlying factors are maintained in the presence of serially correlated idiosyncratic noises.

2.4 Generalization to more than one factor

The results above have been obtained assuming that there is a unique common factor in the system. However, in practice, a larger number of common underlying factors could be expected in large systems. The natural question to ask is whether the conclusions are still the same when we need to estimate more than one factor. As mentioned above, we do not consider the model uncertainty and, therefore, we assume that the number of factors is known and given by r. The vector of factors is then given by

$$F_t = \Phi F_{t-1} + \eta_t \tag{12}$$

where Φ is an $r \times r$ diagonal matrix containing the autoregressive parameters and η_t is an $r \times 1$ Gaussian vector with diagonal covariance matrix Σ_η. The variables in the system are related with the underlying factors through equation (3) where $P^{(N)}$ is now an $N \times r$ matrix of factor loadings. To simplify the discussion, we consider a DFM with serially uncorrelated idiosyncratic noises with contemporaneous covariance matrix $\Sigma_\varepsilon^{(N)}$. In this case, the Riccati equation is given by

$$V(N) = \Phi \left[V(N) - V(N)P^{(N)\prime} \left(P^{(N)}V(N)P^{(N)\prime} + \Sigma_{\varepsilon}^{(N)} \right)^{-1} P^{(N)}V(N) \right] \Phi' + \Sigma_{\eta}. \tag{13}$$

As commented above, given that the state vector is not scalar, the Riccati equation in (13) does not need to have a closed-form solution in terms of the parameters of the model. Therefore, once more, in this subsection, we obtain the steady-state MSE associated with the estimates of the vector of underlying factors by running the prediction equations of the Kalman filter until the steady-state is reached. The particular DFM considered is defined as in 12 with

$$\Phi = \begin{bmatrix} 1 & 0 \\ 0 & 0.4 \end{bmatrix}$$

and $\Sigma_{\eta} = I^{(2)}$. The weights are given by $p_{i1} = 1$, $i = 1, ..., N$, while $p_{12} = 0$ and the remaining weights of the second factor, p_{i2}, $i = 2, ..., N$, have been randomly generated from a uniform $[0, 1]$ distribution. Figure 4 plots the steady-state MSE of one-step-ahead (first column), filtered (second column) and smoothed (third column) factor estimates for the first factor in the first row and for the second in the second row. Finally, the third row of figure 4 plots the corresponding covariances delivered by the filter. Once more, we can observe that the MSE of filtered and smoothed estimates are very similar. We can also observe that the absolute covariances between the estimated factors decrease with the cross-sectional dimension, N. From figure 4, it is also clear that the one-step-ahead estimates of both factors have very similar steady-state MSE. In any case, the MSE of the second factor decrease very quickly with the first variables added to the system but then, after having around 30 variables, the decrease is rather slow.

Figure 4 about here

Therefore, it seems that the presence of more than one factor, require, in general, more variables to be estimated with a given precision. This could be due to the correlation between the estimated factors.

3 Estimation of parameters

In the previous section, we assume known parameters and analyze how the filter uncertainty depends on the cross-sectional dimension. However, in practice, when implementing KFS methods, the parameters are unknown and are usually substituted by consistent estimates. In this case, the total uncertainty associated with the estimation of the underlying factors has two components, one related with the stochastic error uncertainty, considered in the previous section, and another with the parameter uncertainty. Note that the MSEs delivered by the filter run with estimated parameters, which are usually reported in practice by many authors, subestimate the true uncertainty as they do not incorporate the additional uncertainty due to the parameter estimation; see, for example, Quenneville and Singh (2000) and Rodríguez and Ruiz (2012) who quantify this uncertainty as a 5% of the total uncertainty in a univariate non-stationary one factor model with $T = 100$. In this section, we measure the additional uncertainty attributable to parameter estimation and its relation with the cross-sectional and time dimensions. Note that the parameter uncertainty is expected to increase with the cross-sectional dimension as more parameters need to be estimated when adding additional variables to the system. On the other hand, if the parameter estimator is consistent, increasing the temporal dimension decreases the parameter uncertainty which disappears in the limit. Because of its popularity and given that the model considered in this paper is assumed to be conditionally Gaussian, we focus on the ML estimator of the parameters with the log-likelihood maximized using the EM algorithm. This algorithm has the attractiveness of being derivative free and only requires one pass of the smoother in each iteration. Therefore, it is computationally convenient when dealing with the estimation of the parameters in large DFM.

In order to measure the total MSE associated with one-step-ahead estimates of the factors when the Kalman filter is implemented with estimated parameters, we treat the estimated model as if it were misspecified and use the results in Delle Monache and Harvey (2011) who establish a general framework to compute the MSE in misspecified state-space models. Consider the strict DFM with r factors given by equations (3) and (12) and denote by $\Theta^{(N)}$ the vector of parameters to be estimated and by $\widehat{\Theta}^{(N,T)}$, the vector of the corresponding ML estimates. Also denote by $\widehat{K}_t$ the filter gain, which is given by $\widehat{K}_t = \widehat{\Phi}\widehat{V}_{t|t-1}\widehat{P}^{(N)\prime}\left[\widehat{P}^{(N)}\widehat{V}_{t|t-1}\widehat{P}^{(N)\prime}+\widehat{\Sigma}_{\varepsilon}^{(N)}\right]^{-1}$ and by $\widehat{V}_{t+1|t}$ the one-step-ahead MSE matrices delivered by the filter run with estimated parameters, which are given by

$$\widehat{V}_{t+1|t} = \widehat{\Phi}\left[\widehat{V}_{t|t-1} - \widehat{V}_{t|t-1}\widehat{P}^{(N)\prime}\left(\widehat{\Phi}\widehat{V}_{t|t-1}\widehat{\Phi}' + \widehat{\Sigma}_{\varepsilon}\right)^{-1}\widehat{P}^{(N)}\widehat{V}_{t|t-1}\right]\widehat{\Phi}' + \widehat{\Sigma}_{\eta}. \quad (14)$$

Note that although both, $\widehat{K}_t$ and $\widehat{V}_{t+1|t}$, depend on the cross-sectional and temporal dimensions, we do not made this dependence explicit to simplify the notation. Then, according to the formulae derived by Delle Monache and Harvey (2011), the true MSE of $\widehat{f}_{t|t-1}$, the one-step-ahead estimates of the underlying factors delivered by the Kalman filter with the true parameters substituted by estimated parameters, is given by

$$\begin{aligned}
V_{t+1|t} \quad = \quad & \left[\widehat{\Phi} - \widehat{K}_t\widehat{P}^{(N)}\right]V_{t|t-1}\left[\widehat{\Phi} - \widehat{K}_t\widehat{P}^{(N)}\right]' + \\
& \left[\left(\Phi - \widehat{\Phi}\right) - \widehat{K}_t\left(P^{(N)} - \widehat{P}^{(N)}\right)\right]X_t\left[\left(\Phi - \widehat{\Phi}\right) - \widehat{K}_t\left(P^{(N)} - \widehat{P}^{(N)}\right)\right]' + \\
& \widehat{K}_t\Sigma_{\varepsilon}^{(N)}\widehat{K}_t{}' + \Sigma_{\eta} + \left[\left(\Phi - \widehat{\Phi}\right) - \widehat{K}_t\left(P^{(N)} - \widehat{P}^{(N)}\right)\right]C_{t|t-1}\left[\widehat{\Phi} - \widehat{K}_t\widehat{P}^{(N)}\right]' + \\
& \left[\widehat{\Phi} - \widehat{K}_t\widehat{P}^{(N)}\right]C'_{t|t-1}\left[\left(\Phi - \widehat{\Phi}\right) - \widehat{K}_t\left(P^{(N)} - \widehat{P}^{(N)}\right)\right]' \qquad (15)
\end{aligned}$$

where

$$X_{t+1} = \Phi X_t\Phi' + \Sigma_{\eta}$$

and

$$C_{t+1|t} = \Phi C_{t|t-1}\left[\widehat{\Phi} - \widehat{K}_t\widehat{P}^{(N)}\right]' + \Phi X_t\left[\left(\Phi - \widehat{\Phi}\right) - \widehat{K}_t\left(P^{(N)} - \widehat{P}^{(N)}\right)\right]' + \Sigma_{\eta}$$

with $X_0 = C_0 = V_{1/0} = 0$. Delle Monache and Harvey (2011) show that the true MSE in expression (15) has a steady-state.

The MSE of the filtered estimates of the underlying factors obtained when the filter is run with estimated parameters, $\widehat{f}_{t|t}$, can also be derived from the results in Delle Monache and Harvey (2011) as follows

$$\begin{aligned} W_{t|t} &= V_{t|t-1} + \widehat{M}_t G_t \widehat{M}_t' - V_{t|t-1} \widehat{P}^{(N)\prime} \widehat{M}_t' - \widehat{M}_t P^{(N)} V_{t|t-1}' - \\ & \quad C_{t|t-1}' \left(P^{(N)} - \widehat{P}^{(N)}\right)' \widehat{M}_t' - \widehat{M}_t \left(P^{(N)} - \widehat{P}^{(N)}\right) C_{t|t-1}, \quad (16) \end{aligned}$$

where

$$\widehat{M}_t = \widehat{V}_{t|t-1} \widehat{P}^{(N)\prime} \left[\widehat{P}^{(N)} \widehat{V}_{t|t-1} \widehat{P}^{(N)\prime} + \widehat{\Sigma}_{\varepsilon}^{(N)}\right]^{-1},$$

and

$$\begin{aligned} G_t &= \widehat{P}^{(N)} V_{t|t-1} \widehat{P}^{(N)\prime} + \Sigma_{\varepsilon}^{(N)} + \left(P^{(N)} - \widehat{P}^{(N)}\right) X_t \left(P^{(N)} - \widehat{P}^{(N)}\right)' + \\ & \quad \widehat{P}^{(N)} C_{t|t-1}' \left(P^{(N)} - \widehat{P}^{(N)}\right)' + \left(P^{(N)} - \widehat{P}^{(N)}\right) C_{t|t.-1} \widehat{P}^{(N)\prime}. \end{aligned}$$

Delle Monache and Harvey (2011) do not provide results for the true MSE of the fixed interval smoothed estimates considered in this paper. However, we have seen in the previous section that the MSE of filtered and smoothed estimates are nearly indistinguishable. Therefore, we expect the MSE of the smoothed estimates of the factors obtained when the filter is run with estimated parameters to be similar to that of the filtered estimates and only consider the latter in this section.

The "true" steady-state MSEs in equations (15) and (16) depend on the parameter estimates obtained in a particular data set. Consequently, we carry out Monte Carlo experiments to measure the uncertainty associated with $\widehat{f}_{t|t-1}$ and $\widehat{f}_{t|t}$ in the context of a strict DFM. For a given specification, we generate $R = 500$ replicates of sizes $T = 100$ and 200 and for each replicate, we obtain the ML estimates of the parameters using the EM algorithm and compute the true MSE of the one-step-ahead and filtered estimates of the underlying factors, using expressions (15) and (16) until they reach the steady-state. Then we average the steady-state MSE through all replicates. The resulting averages are denoted as $\overline{V}(N,T)$ and $\overline{W}(N,T)$, respectively.

As an illustration, we consider a DFM with two factors defined as in equation (11) with

$$\Phi = \begin{bmatrix} 0.8 & 0 \\ 0 & 0.4 \end{bmatrix}$$

and $\Sigma_\eta = I^{(2)}$. The factor loadings are given by $p_{i1} = 1$, $i = 1, ..., N$, $p_{21} = 0$, p_{2i} are randomly generated from a uniform $[0, 1]$ distribution for $i = 2, ..., N$ and $\Sigma_\varepsilon^{(N)} = I^{(N)}$. The usual identifying restrictions are imposed before estimation. First, as the model is stationary, $Y_t^{(N)}$ is assumed to be zero mean and, consequently, all series in $Y_t^{(N)}$ are centered previous to their analysis. Finally, we restrict $\Sigma_\eta = I^{(2)}$ and $p_{11} > 0$ and $p_{12} = 0$; see Harvey (1989). Note that the positivity of the loading parameter of the first variable in the system is needed to identify the sign of the factor. We also compute the steady-state MSE associated with the stochastic error as given in equation (13). In order to save space, we only report the results related with filtered estimates of the factors. The first row of the left column of figure 5 plots the total steady-state MSE of the filtered estimates of the first factor together with the corresponding stochastic error MSE when $T = 100$. The second row plots the same quantities for the second factor. Finally, the third row plots the corresponding covariances. The right column of figure 5 plots the same quantities when $T = 200$. We can observe that, obviously, the total MSEs are always larger than the MSEs obtained when the filter is run with known parameters. Furthermore, for a given sample size, T, while the filter MSE approaches zero for relatively small cross-sectional dimensions, the total uncertainty has a U shape. As more variables are introduced in the system, more parameters need to be estimated and, consequently, the total uncertainty could even increase with N. The difference between the total and the filter uncertainty is relatively small when the number of variables in the system is small but increases with the cross-sectional dimension. When we are dealing with a system with $N = 60$ and $T = 100$, we can observe that most of the uncertainty can be attributed to the parameter estimation. When $T = 200$, we can observe that the total MSEs of $\widehat{f}_{t|t}$ are obviously smaller for each N. This is obviously expected given that the ML estimator of Θ is consistent, $\underset{T\rightarrow\infty}{P\lim}\left(\widehat{\Theta}\right)=\Theta$ and, consequently, the limit when T tends to infinity of the MSE in (16) is given by the MSE in (13). However, the parameter uncertainty is still important when $N=60$.

Figure 5 about here

To assess the importance of the parameter uncertainty when extracting the factors by running the Kalman filter with estimated parameters, the left column of figure 6 plots the relative difference between the average total MSE, $\overline{V}(N,T)$ and the MSE attributable to the filter uncertainty, $V(N)$, i.e. the percentage of the total uncertainty that can be attributed to the parameter uncertainty for one-step-ahead (left column) and filtered (right column) estimates of the factors, when $T=100$. The right column of figure 6 plots the same quantities when $T=200$. We can observe that, for one-step-ahead estimates of the first factor, this percentage is a quadratic function of the cross-sectional dimension. It is minimum when the number of variables in the system is between 20 and 30 variables when the parameter uncertainty represents around a 4.5% of the total uncertainty. However, if there is a small number of variables in the system, it can represent around 8% and when $N=60$, it represents around a 6%. When looking at the results for the second factor, we can see that the percentage decreases with N, being around 12% when there are few variables in the system and being as small as 2% when the number of variables is 60. Nevertheless, the percentages of the filtered estimates are completely different. Regardless of whether we look at the results for the first or second factor, we can observe that the percentage of the parameter uncertainty over the total uncertainty is minimum when the number of variables is around 10. In this case, the percentage is around 1%. However, as the number of variables increases, this percentage also increases and could be as large as 55% in the first factor or 30% in the second. Furthermore, the percentage of the total MSE that can be attributed to the parameter estimation is now smaller.

Figure 6 about here

Another important issue, obtained as a subproduct of the analysis carried out in this section, is a measure of the bias of the MSE computed using the estimated parameters without taking into account the additional uncertainty associated with the estimation. With this purpose, we also compute the steady-state MSEs delivered by the filter with estimated parameters by computing recursively (14) until it reaches the steady-state. Then, we compute the Monte Carlo averages of the steady-states for all replicates, denoted by $\widehat{V}(N,T)$ and $\widehat{W}(N,T)$ for one-step-ahead and filtered estimates, respectively. These averages are plotted in figure 5 for $T = 100$ and 200. We can observe that these MSE are clearly smaller to the total MSE and closer to the MSE computed when the parameters are known. The biases are larger the smaller T and the larger N. The left column of figure 7 plots $(\widehat{V}(N,T) - \overline{V}(N,T))/\overline{V}(N,T)$, the relative baises of the MSE delivered by the Kalman filter with estimated parameters for one-step-ahead. The same quantities for filtered estimates are plotted in the right column of figure 7. Note that when $T = 100$ and the number of the variables in the system is large, the biases could be as large as 30%. Therefore, the parameter uncertainty should be taken into account if we want to have realistic measures of the uncertainty of the underlying factors close to the true ones. The parameter uncertainty can be an important issue when estimating large systems and should not be ignored. Bootstrap procedures as those proposed by Rodriguez and Ruiz (2012) could be implemented to incorporate this uncertainty.

Figure 7 about here

4 An illustration with simulated data

To analyze the practical implications of our results when dealing with real data, the performance of the one-step-ahead estimates of the underlying factor is illustrated with simulated data when $T = 100$ and $N = 2$ or $N = 20$. We simulate a non-stationary factor by model (1) with $\phi = 1$ and $\sigma_\eta^2 = 1$ which is plotted in figure 8. Then, 30 time series are simulated by a strict DFM with relative loadings $q_i = 1$, $i = 1, ..., 20$, and the corresponding idiosyncratic noises being mutually independent Gaussian white noises. The Kalman filter is first run with the first two simulated variables and then with all 30 variables. The left panels of figure 8 plot the one-step-ahead estimates of the factor together with the corresponding 95% prediction intervals obtained using the MSE delivered by the Kalman filter with known parameters and $N = 2$ (top panel) and $N = 20$ (low panel) variables are used in the estimation. Notice that when the cross-sectional dimension increases from $N = 2$ to $N = 20$, the estimated factor is much closer to the simulated factor and the intervals are much narrower. When $N = 20$, the intervals are very narrow following closely the evolution of the simulated underlying factor.

The right panels of figure 8 also plot the simulated factor together with the corresponding one-step-ahead estimates obtained when the Kalman filter is implemented with the true parameters substituted by their ML estimates. Once more, the top panel plots the estimates obtained when $N = 2$ while the low panel plots the estimates when $N = 20$. In each of these two panels, we have plotted two different series of 95% prediction intervals for the underlying factor. First, we have plotted the intervals obtained using the MSE directly delivered by the Kalman filter with estimated parameters. The second intervals have been obtained using the "true" MSE that incorporate the parameter uncertainty which have been computed using the formulae in Delle Monache and Harvey (2011). We can observe that the Kalman filter intervals are narrower than the true with the result that the simulated factor is lying out of the intervals in too many moments of time. Therefore, it seems clear that, in practice, we need to take into account the parameter uncertainty when computing intervals for simulated factors.

Finally figure 8 also illustrates that when $N = 20$, the intervals are already very narrow so there is no substantial gains in precision using more variables to extract the underlying factor.

Figure 8 about here

5 Conclusions

In this paper, we contribute to the issue about the finite sample uncertainty associated with the extraction of unobserved factors in DFM in the context of KFS procedures. We also extend available consistency results to real time filtered estimates of the factors. Assuming that the model specification is known, if the Kalman filter is implemented, as usual, substituting the unknown parameters by consistent estimates, the total MSE can be decomposed into the part attributable to the filter uncertainty and that attributable to the parameter uncertainty. When looking at the former component of the MSE, we show that, regardless of whether the idiosyncratic noises are weakly or strongly correlated, a relatively small number of variables, typically around 30, is enough to estimate the factors with an uncertainty close to its potential minimum. However, a larger number of variables could be needed if the idiosyncratic noises are highly persistent. Furthermore, when looking at the parameter uncertainty, we show that it can represent a large percentage of the total uncertainty. For a given temporal dimension, the parameter uncertainty and the total uncertainty can even be an increasing function of the cross-sectional dimension. Therefore, it seems that there is no point in including a huge number of variables when estimating underlying factor. This result suggests that it could be worth to go back to the Kalman filter improving the efficiency in the estimation of unobserved factors with respect to PC whose main advantage is being able to deal with very large DFM,

In this paper, the results about the parameter uncertainty have been obtained using the ML estimator based on the EM algorithm. However, other alternative estimators, as those proposed by Jungbaker and Koopman (2008) or Doz *et al.* (2011, in press), could be considered. The results could then be compared with those of PC. This issue is left for further research.

The Kalman filter is based on a parametric specification of the dynamics of the factors and idiosyncratic components. Obviously, in practice these dynamics could be misspecified. Therefore, analyzing the effects of misspecification on the Kalman filter MSE is an interesting topic that we left for further research. Other sources of model misspecification are related with the number of underlying factors and the error distribution. Comparing the performance of Kalman filter and PC estimators in the context of both well specified and misspecified models is also in our research agenda.

Appendix

Proof of lemma 1: Resolution of the Riccati equation for the steady-state one-step-ahead MSE and its non-increasing property.

Using the expression of the inverse of the sum of two matrices, we obtain the following result

$$\begin{aligned}&\left(P^{(N)}V(N)P^{(N)\prime}+\Sigma_{\varepsilon}^{(N)}\right)^{-1}=\\&\Omega^{(N)}-\Omega^{(N)}P^{(N)\prime}\left(P^{(N)\prime}\Omega^{(N)}P^{(N)\prime}+\left(V(N)\right)^{-1}\right)^{-1}P^{(N)\prime}\Omega^{(N)}\end{aligned} \tag{17}$$

where $\Omega^{(N)}=\left(\Sigma_{\varepsilon}^{(N)}\right)^{-1}$.

Introducing (17) into the expression of the Riccati equation in (6), we obtain the following expression

$$\begin{aligned}&V(N)=k\sigma_{\eta}^{2}+\phi^{2}V(N)-\\&\phi^{2}V(N)P^{(N)\prime}\left(\Omega^{(N)}-\Omega^{(N)}P^{(N)}\left(Q^{(N)}+\left(V(N)\right)^{-1}\right)^{-1}P^{(N)\prime}\Omega^{(N)}\right)P^{(N)}V(N)\end{aligned} \tag{18}$$

which can be rewritten as

$$\left(1-\phi^{2}\right)V(N)+\phi^{2}V^{2}(N)\left(Q^{(N)}-\frac{\left(Q^{(N)}\right)^{2}}{Q^{(N)}+\left(V(N)\right)^{-1}}\right)=k\sigma_{\eta}^{2}. \tag{19}$$

After some straightforward algebra, the following equation is obtained

$$Q^{(N)}V^{2}(N)-\left(k\sigma_{\eta}^{2}Q^{(N)}-1+\phi^{2}\right)V(N)-k\sigma_{\eta}^{2}=0. \tag{20}$$

Taking the positive solution of the 2nd order equation (20), we obtain expression (7) for the steady-state MSE.

To prove that $V(N)$ is a non increasing function of N, subtract expression (20) evaluated at $N+1$ from the same expression evaluated at N,

$$V(N)^2 Q^{(N)} - V(N+1)^2 Q^{(N+1)} + V(N)\left(-k\sigma_\eta^2 Q(N) + 1 - \phi^2\right) - V(N+1)\left(-k\sigma_\eta^2 Q^{(N+1)} + 1 - \phi^2\right) = 0. \tag{21}$$

First, we need to prove that $Q^{(N+1)} \geq Q^{(N)}$. Consider the following partition of the covariance matrix of the idiosyncratic noises when the cross-sectional dimension is $N+1$

$$\Sigma_\varepsilon^{(N+1)} = \begin{bmatrix} \Sigma_\varepsilon^{(N)} & \Sigma_{N,N+1} \\ \Sigma'_{N,N+1} & \sigma_{N+1}^2 \end{bmatrix},$$

where σ_{N+1}^2 is the variance of ε_{N+1} and $\Sigma_{N,N+1}$ is an $N \times 1$ vector that collects the covariances between ε_i, $i = 1, ..., N$ and ε_{N+1}. Using the formula for the inverse of the partitioned matrix, we obtain the following expression for $\Omega^{(N+1)} = \left(\Sigma_\varepsilon^{(N+1)}\right)^{-1}$

$$\Omega^{(N+1)} = \begin{bmatrix} B & -\frac{1}{\sigma_{N+1}^2} B \Sigma_{N,N+1} \\ -\frac{1}{\sigma_{N+1}^2} \Sigma'_{N,N+1} B & \frac{1}{\sigma_{N+1}^2} + \frac{1}{\sigma_{N+1}^4} \Sigma'_{N,N+1} B \Sigma_{N,N+1} \end{bmatrix} \tag{22}$$

where $B = \left(\Sigma_\varepsilon^{(N)} - \frac{\Sigma_{N,N+1}\Sigma'_{N,N+1}}{\sigma_{N+1}^2}\right)^{-1}$. Applying the formula of the inverse of the sum of two matrices, B can be rewritten as

$$\begin{aligned} B &= \Omega^{(N)} + \left(1 - \frac{1}{\sigma_{N+1}^2} \Sigma'_{N,N+1} \Omega^{(N)} \Sigma_{N,N+1}\right)^{-1} \frac{1}{\sigma_{N+1}^2} \Omega^{(N)} \Sigma_{N,N+1} \Sigma'_{N,N+1} \Omega^{(N)} \\ &= \Omega^{(N)} + \frac{1}{\sigma_{N+1|N}^2} \Omega^{(N)} \Sigma_{N,N+1} \Sigma'_{N,N+1} \Omega^{(N)}, \end{aligned} \tag{23}$$

where $\sigma_{N+1|N}^2 = \sigma_{N+1}^2 - \Sigma'_{N,N+1} \Omega^{(N)} \Sigma_{N,N+1}$ is the variance of ε_{N+1} conditional on ε_i, $i = 1, ..., N$. Finally, let the factor loading vector be partitioned as

$P^{(N+1)\prime} = \left[\begin{array}{cc} P^{(N)\prime} & p_{N+1} \end{array}\right]$. Considering the previous partitions and using the results in (22) and (23), it is possible to write $Q^{(N+1)}$ as

$$
\begin{aligned}
&Q(N+1) = P^{(N+1)\prime}\Omega^{(N+1)}P^{(N+1)} = \\
&= P^{(N)\prime}\left(\Omega^{(N)} + \tfrac{1}{\sigma^2_{N+1|N}}\Omega^{(N)}\Sigma_{N,N+1}\Sigma'_{N,N+1}\Omega^{(N)}\right)P^{(N)} \\
&+p^2_{N+1}\left(\tfrac{1}{\sigma^2_{N+1}} + \tfrac{1}{\sigma^4_{N+1}}\Sigma'_{N,N+1}\left(\Omega^{(N)} + \tfrac{1}{\sigma^2_{N+1|N}}\Omega^{(N)}\Sigma_{N,N+1}\Sigma'_{N,N+1}\Omega^{(N)}\right)\Sigma_{N,N+1}\right) \\
&-2\tfrac{p_{N+1}}{\sigma^2_{N+1}}P^{(N)\prime}\left(\Omega^{(N)} + \tfrac{1}{\sigma^2_{N+1|N}}\Omega^{(N)}\Sigma_{N,N+1}\Sigma'_{N,N+1}\Omega^{(N)}\right)\Sigma_{N,N+1} \\
&= P^{(N)\prime}\Omega^{(N)}P^{(N)} + \tfrac{1}{\sigma^2_{N+1|N}}\left(P^{(N)\prime}\Omega^{(N)}\Sigma_{N,N+1}\right)^2 + \tfrac{p^2_{N+1}}{\sigma^2_{N+1}} \\
&+\tfrac{p^2_{N+1}}{\sigma^4_{N+1}}\Sigma'_{N,N+1}\Omega^{(N)}\Sigma_{N,N+1}\left(1 + \tfrac{1}{\sigma^2_{N+1|N}}\Sigma'_{N,N+1}\Omega^{(N)}\Sigma_{N,N+1}\right) \\
&-2\tfrac{p_{N+1}}{\sigma^2_{N+1}}P^{(N)\prime}\Omega^{(N)}\Sigma_{N,N+1}\left(1 + \tfrac{1}{\sigma^2_{N+1|N}}\Sigma'_{N,N+1}\Omega^{(N)}\Sigma_{N,N+1}\right).
\end{aligned}
\tag{24}
$$

Using that $1 + \frac{1}{\sigma^2_{N+1|N}}\Sigma'_{N,N+1}\Omega^{(N)}\Sigma_{N,N+1} = \frac{\sigma^2_{N+1}}{\sigma^2_{N+1|N}}$, in (24), it follows that

$$
\begin{aligned}
Q(N+1) &= Q^{(N)} + \frac{1}{\sigma^2_{N+1|N}}\left(P^{(N)\prime}\Omega^{(N)}\Sigma_{N,N+1}\right)^2 \\
&\quad + \frac{p^2_{N+1}}{\sigma^2_{N+1}} + \frac{p^2_{N+1}}{\sigma^2_{N+1}}\frac{1}{\sigma^2_{N+1|N}}\Sigma'_{N,N+1}\Omega^{(N)}\Sigma_{N,N+1} \\
&\quad -2\frac{p_{N+1}}{\sigma^2_{N+1|N}}P^{(N)\prime}\Omega^{(N)}\Sigma_{N,N+1} \\
&= Q^{(N)} + g^2_{N+1}
\end{aligned}
\tag{25}
$$

where $g^2_{N+1} = \frac{1}{\sigma^2_{N+1|N}}\left(p_{N+1} - P^{(N)\prime}\Omega^{(N)}\Sigma_{N,N+1}\right)^2$. Consequently, $Q^{(N+1)} \geq Q^{(N)}$ regardless of the properties of the covariances between the idiosyncratic noises. Note that if ε_{N+1} is uncorrelated with ε_i, $i = 1, ..., N$, then $g^2_{N+1} = q^2_{N+1} = \frac{p^2_{N+1}}{\sigma^2_{N+1}}$. If further, the variable y_{N+1} is non-informative so that $p^2_{N+1} = 0$, or $\sigma^2_{N+1} = \infty$, then $g^2_{N+1} = 0$ and, consequently, $Q^{(N+1)} = Q^{(N)}$. However, if only $p_{N+1}^2 = 0$, then $g_{N+1}^2 = \frac{1}{\sigma^2_{N+1|N}} \left(P^{(N)\prime} \Omega^{(N)} \Sigma_{N,N+1} \right)^2 > 0$ and $Q^{(N+1)} > Q^{(N)}$.

To prove that $V(N)$ is non-increasing, consider again equation (21) and substitute $Q^{(N+1)}$ by $Q^{(N)} + g_{N+1}^2$, obtaining the following expression

$$
\begin{aligned}
&\left(V(N)^2 - V(N+1)^2\right) Q^{(N)} - V(N+1)^2 g_{N+1}^2 + \\
&V(N)\left(-\sigma_\eta^2 k Q^{(N)} + 1 - \phi^2\right) - V(N+1)\left(-\sigma_\eta^2 k Q^{(N)} + 1 - \phi^2\right) + \\
&V(N+1) k \sigma_\eta^2 g_{N+1}^2 \\
&= \left(V(N) - V(N+1)\right)\left(V(N) + V(N+1)\right) Q^{(N)} + \\
&\left(V(N) - V(N+1)\right)\left(-\sigma_\eta^2 k Q^{(N)} + 1 - \phi^2\right) + \\
&V(N+1)\left(k\sigma_\eta^2 - V(N+1)\right) g_{N+1}^2 \\
&= 0.
\end{aligned}
$$

Rearranging terms,

$$
\begin{aligned}
&V(N+1)\left(V(N+1) - k\sigma_\eta^2\right) g_{N+1}^2 \\
&= \left(V(N) - V(N+1)\right)\left[\left(1 - \phi^2\right) + \left(V(N) + V(N+1) - k\sigma_\eta^2\right) Q^{(N)}\right].
\end{aligned}
\tag{26}
$$

Next, we show that if $\phi > 0$, then $V(N+1) > k\sigma_\eta^2$. Consider the expression of $V(N)$ in (7). Then,

$$
\begin{aligned}
V(N+1) &> \frac{k\sigma_\eta^2 Q^{(N+1)} - 1 + \sqrt{\left(k\sigma_\eta^2 Q^{(N+1)} - 1\right)^2 + 4k\sigma_\eta^2 Q^{(N+1)}}}{2Q^{(N+1)}} \\
&= \frac{k\sigma_\eta^2 Q^{(N+1)} - 1 + k\sigma_\eta^2 Q^{(N+1)} + 1}{2Q^{(N+1)}} \\
&= k\sigma_\eta^2.
\end{aligned}
$$

Therefore, the left hand side of expression (26) is always positive if $g_{N+1}^2 > 0$. On the other hand, $\left[\left(1 - \phi^2\right) + \left(V(N) + V(N+1) - k\sigma_\eta^2\right) Q^{(N)}\right] > 0$ and consequently, $V(N) > V(N+1)$. Finally, if $g_{N+1}^2 = 0$, then $V(N) = V(N+1)$.

Proof of lemma 3: Non-increasing property of steady-state MSE of filtered estimates.

The steady-state MSE of the filtered estimates of the underlying factors are given by

$$W(N) = V(N) - V(N)P^{(N)\prime}\left[P^{(N)}V(N)P^{(N)\prime} + \Sigma_{\varepsilon}^{(N)}\right]^{-1} P^{(N)}V(N).$$

Using the well-known formula for the inverse of the sum of matrices given by Rao (1973), expression (10) of the steady-state MSE of $f_{t|t}$ is directly obtained. In order to prove that it is non decreasing, we have to show that

$$\begin{aligned}&\frac{V(N)}{1+V(N)Q^{(N)}} - \frac{V(N+1)}{1+V(N+1)Q^{(N+1)}} = \\ &\frac{V(N)\left[1+V(N+1)Q^{(N+1)}\right] - V(N+1)\left[1+V(N)Q^{(N)}\right]}{[1+V(N)Q^{(N)}]\,[1+V(N+1)Q^{(N+1)}]} \geq 0.\end{aligned}$$

Since the denominator is positive, the proof is reduced to show that the numerator is also positive. After some straightforward algebra

$$\begin{aligned}V(N)\left[1+V(N+1)Q(N+1)\right] - V(N+1)\left[1+V(N)Q^{(N)}\right] &= \\ V(N) - V(N+1) + V(N)V(N+1)\left[Q^{(N+1)} - Q^{(N)}\right] &\geq 0.\end{aligned}$$

Lemma 1 establishes that if $g_{N+1}^2 = 0$, then $Q^{(N+1)} = Q^{(N)}$ and $V(N) = V(N+1)$. Otherwise, $Q^{(N+1)} > Q^{(N)}$ and $V(N) > V(N+1)$, so the inequality is already proved.

Proof of lemma 4: Non-increasing property of steady-state MSE of smoothed estimates.

As regards the variance of the smoothed factor, $S(N)$, we can show that under the same hypothesis as lemma 1, it is also a decreasing function of the number of series N. We have to prove that $S(N) \geq S(N+1)$, where $S(N)$ is given in (11)

$$\frac{V(N)\left(1+V(N)Q^{(N)} - \phi^2\right)}{\left(1+V(N)Q^{(N)}\right)^2 - \phi^2} \geq \frac{V(N+1)\left(1+V(N+1)Q^{(N+1)} - \phi^2\right)}{\left(1+V(N+1)Q^{(N+1)}\right)^2 - \phi^2}.$$

Notice that since $V(N) \geq V(N+1)$, then

$$\begin{aligned}\frac{V(N)\left(1+V(N+1)Q^{(N+1)}-\phi^2\right)}{\left(1+V(N+1)Q^{(N+1)}\right)^2-\phi^2} &\geq \frac{V(N+1)\left(1+V(N+1)Q^{(N+1)}-\phi^2\right)}{\left(1+V(N+1)Q^{(N+1)}\right)^2-\phi^2} = \\ &= S(N+1).\end{aligned}$$

Therefore, if we can prove that

$$S(N) = \frac{V(N)\left(1+V(N)Q^{(N)}-\phi^2\right)}{\left(1+V(N)Q^{(N)}\right)^2-\phi^2} \geq \frac{V(N)\left(1+V(N+1)Q^{(N+1)}-\phi^2\right)}{\left(1+V(N+1)Q^{(N+1)}\right)^2-\phi^2} \tag{27}$$

then it will be already proved that $S(N) \geq S(N+1)$. To prove (27), it suffices to prove that

$$\begin{aligned}&\left(1+V(N)Q^{(N)}-\phi^2\right)\left(\left(1+V(N+1)Q^{(N+1)}\right)^2-\phi^2\right)\\ &-\left(1+V(N+1)Q^{(N+1)}-\phi^2\right)\left(\left(1+V(N)Q^{(N)}\right)^2-\phi^2\right) \geq 0,\end{aligned}$$

since the denominators are positive and $V(N)$ can be simplified from both sides of the inequality (27).

After some straightforward algebra

$$\begin{aligned}&\left[(1-\phi^2)+V(N)Q(N)\right]\left[1-\phi^2+V^2(N+1)Q^2(N+1)+2V(N+1)Q(N+1)\right]\\ &-\left[(1-\phi^2)+V(N+1)Q(N+1)\right]\left[1-\phi^2+V^2(N)Q^2(N)+2V(N)Q(N)\right]\\ = &\left[V(N)Q(N)-V(N+1)Q(N+1)\right](1-\phi^2)+\\ &\left[V^2(N+1)Q^2(N+1)-V^2(N)Q^2(N)\right](1-\phi^2)\\ &+V(N)V(N+1)Q(N)Q(N+1)\left[V(N+1)Q(N+1)-V(N)Q(N)\right]\\ &+2(1-\phi^2)\left[V(N+1)Q(N+1)-V(N)Q(N)\right]\\ = &\left[(1-\phi^2)+V(N)V(N+1)Q(N)Q(N+1)\right]\left[V(N+1)Q(N+1)-V(N)Q(N)\right]\\ &+\left[V(N+1)Q(N+1)-V(N)Q(N)\right]\left[V(N+1)Q(N+1)+V(N)Q(N)\right](1-\phi^2)\\ = &\left[(1-\phi^2)\left(1+\left(V(N+1)Q(N+1)+V(N)Q(N)\right)\right)+V(N)V(N+1)Q(N)Q(N+1)\right]\\ &\times\left[V(N+1)Q(N+1)-V(N)Q(N)\right].\end{aligned}$$

Since $1-\phi^2 \geq 0$ and both $V(N)$ and $Q(N)$ are positive for all N, it suffices to prove that $V(N+1)Q(N+1) \geq V(N)Q(N)$. By (7)

$$V(N)Q(N) = \frac{k\sigma_\eta^2 Q^{(N)} - 1 + \phi^2 + \sqrt{\left(k\sigma_\eta^2 Q^{(N)} - 1 + \phi^2\right)^2 + 4k\sigma_\eta^2 Q^{(N)}}}{2}.$$

Then, taking also into account that $Q^{(N+1)} = Q^{(N)} + g_{N+1}^2$

$$\begin{aligned}
& V(N+1)Q(N+1) - V(N)Q(N) \\
= \; & \frac{k\sigma_\eta^2 Q^{(N+1)}}{2} + \frac{\sqrt{\left(k\sigma_\eta^2 Q^{(N+1)} - 1 + \phi^2\right)^2 + 4k\sigma_\eta^2 Q^{(N+1)}}}{2} \\
& - \frac{k\sigma_\eta^2 Q^{(N)}}{2} - \frac{\sqrt{\left(k\sigma_\eta^2 Q^{(N)} - 1 + \phi^2\right)^2 + 4k\sigma_\eta^2 Q^{(N)}}}{2} \\
= \; & \frac{k\sigma_\eta^2 g_{N+1}^2}{2} + \frac{\sqrt{\left(k\sigma_\eta^2 Q^{(N+1)} - 1 + \phi^2\right)^2 + 4k\sigma_\eta^2 Q^{(N+1)}}}{2} - \\
& \frac{\sqrt{\left(k\sigma_\eta^2 Q^{(N)} - 1 + \phi^2\right)^2 + 4k\sigma_\eta^2 Q^{(N)}}}{2}.
\end{aligned}$$

Since $\frac{k\sigma_\eta^2 g_{N+1}^2}{2}$ is positive, we have to prove that

$$\sqrt{\left(k\sigma_\eta^2 Q^{(N+1)} - 1 + \phi^2\right)^2 + 4k\sigma_\eta^2 Q^{(N+1)}} \geq \sqrt{\left(k\sigma_\eta^2 Q^{(N)} - 1 + \phi^2\right)^2 + 4k\sigma_\eta^2 Q^{(N)}}$$

or taking squares

$$\left(k\sigma_\eta^2 Q^{(N+1)} - 1 + \phi^2\right)^2 + 4k\sigma_\eta^2 Q^{(N+1)} \geq \left(k\sigma_\eta^2 Q^{(N)} - 1 + \phi^2\right)^2 + 4k\sigma_\eta^2 Q^{(N)}.$$

After some straightforward algebra and taking into account that $Q^{(N+1)} = Q^{(N)} + g_{N+1}^2$

$$\begin{aligned}
& \left(k\sigma_\eta^2 \left(Q^{(N)} + g_{N+1}^2\right) - \left(1-\phi^2\right)\right)^2 + 4k\sigma_\eta^2 \left(Q^{(N)} + g_{N+1}^2\right) - \\
& \left(k\sigma_\eta^2 Q^{(N)} - \left(1-\phi^2\right)\right)^2 + 4k\sigma_\eta^2 Q^{(N)} \\
= \; & k^2\sigma_\eta^4 g_{N+1}^4 + 2k\sigma_\eta^2 Q^{(N)} g_{N+1}^2 - 2k\sigma_\eta^2 g_{N+1}^2 \left(1-\phi^2\right) + 4k\sigma_\eta^2 g_{N+1}^2 \\
= \; & k^2\sigma_\eta^4 g_{N+1}^4 + 2k\sigma_\eta^2 Q^{(N)} g_{N+1}^2 + 2(1+\phi^2)k\sigma_\eta^2 g_{N+1}^2 \geq 0
\end{aligned}$$

which is positive since it is the sum of three positive terms. Therefore the result is proved and $S(N) \geq S(N+1)$.

Acknowledgements

This paper was presented at the Workshop in Honour of the 65th Birthday of Andrew Harvey held in the Oxford Man Institute on June 2012. We are very grateful to all participants. Financial support from the Spanish Government projects ECO2009-10287 and ECO2012-32854 is acknowledged by the first author while the second author acknowledges support from projects ECO2009-08100 and ECO2012-32401. We also appreciate very useful comments from the participants in the Computational and Financial Econometrics Conference held in London in December 2011, the Nonlinear Dynamics and Econometrics Conference held in Istanbul in April 2012 and the II_t Workshop on Time Series Econometrics held in Zaragoza in May 2012. We very specially appreciate comments from Davide Delle Monache and Gabriel Perez-Quiros. Obviously, any remaining errors are our responsibility.

References

[1] Altissimo, F., Cristadoro, R., Forni, M., Lippi, M., and Veronese, G. (2010). 'New Eurocoin: Tracking economic growth in real time', *Review of Economics and Statistics,* 92:1024-1034.

[2] Amengual, D., and Watson, M.W. (2007). 'Consistent estimation of the number of dynamic factors in a large N and T panel', *Journal of Business and Economic Statistics,* 25(1):91-96.

[3] Ang, A., and Piazzesi, M. (2003). 'A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables', *Journal of Monetary Economics,* 50(4):745-787.

[4] Artis, M., Marcelino, M. and Proietti, T. (2004). 'Dating business cycle: a methodological contribution with an application to the euro area', *Oxford Bulletin of Economics and Statistics,* 66(4):537-565.

[5] Aruoba, S.B., Diebold, F.X., and Scotti, C. (2009). 'Real-Time Measurement of Business Conditions', *Journal of Business & Economic Statistics,* 27:417-427.

[6] Bai, J. (2003). 'Inferential theory for factor models of large dimensions', *Econometrica,* 71(1):135-171.

[7] Bai, J., and Ng, S. (2002). 'Determining the number of factors in approximate factor models', *Econometrica,* 70(1):191-221.

[8] Bai, J., and Ng, S. (2004). 'A PANIC attack on unit roots and cointegration', *Econometrica,* 72(4):1127-1177.

[9] Bai, J., and Ng, S. (2008). 'Forecasting economic time series using targeted predictors', *Journal of Econometrics*, 146:304-3.

[10] Banbura, M., and Runstler, G. (2011). 'A look into the factor model black box: publication lags and the role of hard and soft data in forecasting GDP', *International Journal of Forecasting,* 27(2):333-346.

[11] Boivin, J., and Ng, S. (2006). 'Are more data always better for factor analysis?', *Journal of Econometrics*, 132:169-194.

[12] Breitung, J., and Eickemeir, S. (2006). 'Dynamic Factor Models', in O. Hübler, and J. Frohn (eds.), *Modern Econometric Analysis*, Location: Springer, Chapter 3.

[13] Breigtung, J., and Tenhofen, J. (2011). 'GLS estimation of dynamic factor models', *Journal of the American Statistical Association,* 106(495):1150-1166.

[14] Caggiano, G., Kapetanios G., and Labhard, V. (2011). 'Are more data always better for factor analysis? Results from the euro area, the six largest euro area countries and the UK', *Journal of Forecasting,* 30:736-752.

[15] Camacho, M., and Perez-Quiros, G. (2010). 'Introducing the Euro-STING: Short Term INdicator of Euro Area Growth', *Journal of Applied Econometrics*, 25(4):663-694.

[16] Camacho, M., Perez-Quiros G., and Poncela, P. (2012). 'Markov switching dynamic factor models in real time', WP 1205 (Bank of Spain, Madrid).

[17] Chamberlain, G., and Rothschild, M. (1983). 'Arbitrage, factor structure and mean-variance analysis in large asset markets', *Econometrica*, 51:1305-1324.

[18] Chmelarova, V., and Nath, H.K. (2010). 'Relative price convergence among US cities: Does the choice of numeraire city matter?', *Journal of Macroeconomics,* 32:405-414.

[19] Choi, I. (2012). 'Efficient estimation of factor models', *Econometric Theory,* 28:274-308.

[20] Delle Monache, D., and Harvey, A.C. (2011). 'On the effect of misspecification in models for extracting trends and cycles', Manuscript.

[21] Diebold, F.X., and Nerlove, M. (1989). 'The dynamics of exchange rate volatility. A multivariate latent factor ARCH model', *Journal of Applied Econometrics,* 4(1):1-21.

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

[23] Doz, C., Giannone, D., and Reichlin, L. (in press). 'A Quasi Maximum Likelihood approach for large approximate dynamic factor models', *Review of Economics and Statistics.*

[24] Eickmeier, S. (2009). 'Comovements and heterogeneity in the euro area analyzed in a non-stationary dynamic factor model', *Journal of Applied Econometrics*, 24(6):933-959.

[25] Eickmeier, S., and Ziegler, C. (2008). 'How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach', *Journal of Forecasting,* 27:237-265.

[26] Engle, R.F., and Watson, M.W. (1981). 'A one-factor multivariate time series model of metropolitan wage rates', *Journal of the American Statistical Association,* 76:774-781.

[27] Escribano, A., and Peña, D. (1994). 'Cointegration and common factors', *Journal of Time Series Analysis*, 15:577-586.

[28] Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2000). 'The generalized dynamic-factor model: identification and estimation', *Review of Economics and Statistics,* 82(4):540-554.

[29] Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2003). 'Do financial variables help forecasting inflation and real activity in the euro area?', *Journal of Monetary Economics,* 50(6):1243-1255.

[30] Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2005). 'The generalized dynamic factor model: one sided estimation and forecasting', *Journal of the American Statistical Association*, 100:830-840.

[31] Geweke, J. (1977). 'The dynamic factor analysis of economic time series', in D.J. Aigner, and A.S. Goldberger (eds.), *Latent Variables in Socio-Economic Models,* Amsterdam: North-Holland.

[32] Gupta, R., and Kabundi, A. (2011). 'A large factor model for forecasting macroeconomic variables in South Africa', *International Journal of Forecasting,* 27(4):1076-1088.

[33] Harvey, A.C. (1989). *Forecasting Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

[34] Harvey, A.C., and Streibel, M. (1998). 'Testing for a slowly changing level with special reference to stochastic volatility', *Journal of Econometrics* 87:167–189.

[35] Harvey, A.C., Ruiz, E., and Shephard, N.G. (1994). 'Multivariate Stochastic Variance Models', *Review of Economic Studies,* 61:247-264.

[36] Jungbaker, B., and Koopman, S.J. (2008). 'Likelihood-based analysis for dynamic factor models', Tinbergen Institute Discussion Paper, TI2008-0007/4.

[37] Jungbacker, B., Koopman, S.J., and van der Wel, M. (2009). 'Dynamic factor models with smooth loadings for analyzing the term structure of interest rates', Tinbergen Institute Discussion Paper, TI2009-041/4.

[38] Jungbacker, B., Koopman, S.J., and van der Wel, M. (2011). 'Maximum likelihood estimation for dynamic factor models with missing data', *Journal of Economic Dynamics & Control,* 35:1358-1368.

[39] Kapetanios, G., and Marcelino, M. (2009). 'A parametric estimation method for dynamic factor model of large dimensions', *Journal of Time Series Analysis,* 30(2):208-238.

[40] Koopman, S.J., and van der Wel, M. (2011). Forecasting the US term structure of interest rates using a macroeconomic smooth dynamic factor model, Tinbergen Institute Discussion Paper, TI2011-063/4.

[41] Lam, C., Yao, Q., and Bathia, N. (2011). 'Estimation of latent factors using high-dimensional time series', *Biometrika*, 98(4):901-918.

[42] Lancaster, P., and Rodman, L. (1995). *Algebraic Riccati Equations,* Oxford University Press.

[43] Moon, H.R., and Perron, B. (2007). 'An empirical analysis of nonstationarity in a panel of interest rates with factors', *Journal of Applied Econometrics,* 22:383-400.

[44] Onatski, A. (in press), 'Asymptotics of the Principal Components estimator of large factor models with weakly influential factors', *Journal of Econometrics.*

[45] Peña, D., and Poncela, P. (2004), 'Forecasting with nonstationary dynamic factor models', *Journal of Econometrics,* 119:291-321.

[46] Peña, D., and Poncela, P. (2006), 'Nonstationary dynamic factor analysis', *Journal of Statistical Planning and Inference,* 136:1237-1257.

[47] Piskunov, N. (1969). *Integral and Differential Calculus,* Mir.

[48] Quenneville, B., and Singh (2000), 'Bayesian prediction mean squared error for state space models with estimated parameters', *Journal of Time Series Analysis,* 21:219-236.

[49] Rao, C. R. (1973). *Linear Statistical Inference and Its Applications.* New York: J. Wiley.

[50] Reis, R., and Watson, M.W. (2010). 'Relative good's prices, pure inflation, and the Phillips correlation', *American Economic Journal of Macroeconomics,* 2:128-157.

[51] Rodríguez, A., and Ruiz, E. (2012). 'Bootstrap prediction mean squared errors of unobserved states based on the Kalman filter with estimated parameters', *Computational Statistics and Data Analysis,* 56:62-74.

[52] Rojas, A.J. (2011).'On the discrete-time algebraic Riccati equation and its solution in closed-form', Preprints of the 18th IFAC World Congress, Milan.

[53] Ross, S.A. (1976). 'The arbitrage theory of capital asset pricing', *Journal of Economic Theory,* 13:341-360.

[54] Sargent, T.J., and Sims, C.A. (1977). 'Business cycle modeling without pretending to have too much a priory economic theory', in C.A. Sims (ed.), *New Methods in Business Cycle Research,* Minneapolis: Federal Reserve Bank of Minneapolis.

[55] Shumway, R.H., and Stoffer, D.S. (1982). 'An approach to time series smoothing and forecasting using the EM algorithm', *Journal of Time Series Analysis,* 3:253-264.

[56] Stock, J.H., and Watson, M.W. (1991). 'A probability model of the coincident economic indicators', in K. Lahiri, and G.H. Moore (eds.), *Leading Economic Indicators: New Approaches and Forecasting Records,* Cambridge: Cambridge University Press.

[57] Stock J.H., and Watson, M.W. (2002). 'Forecasting using Principal Components from a large number of predictors', *Journal of the American Statistical Association,* 97:1167-79.

[58] Stock J.H., and Watson, M.W. (2006). 'Forecasting with many predictors', in G. Elliot, C. Granger, and A. Timmermann (eds.), *Handbook of Economic Forecasting,* vol.1, Elsevier.

[59] Stock, J.H., and Watson, M.W. (2010). 'The evolution of national and regional factors in US housing construction', in T. Bollerslev, J.R. Russell, and M.W. Watson (eds.), *Volatility and Time Series Econometrics: Essays in Honour of Robert F. Engle*, Oxford: Oxford University Press.

[60] Stock, J.H., and Watson, M.W. (2011). 'Dynamic factor models', in M.P. Clements, and D.F. Hendry (eds.), *The Oxford Handbook of Economic Forecasting,* Oxford: Oxford University Press.

[61] Timmermann, A. (2008). 'Elusive return predictability', *International Journal of Forecasting,* 24(1):1-18.

[62] Watson, M.W. (2003). 'Macroeconomic forecasting using many predictors', in M. Dewatripont, L. Hansen, and S. Turnovsky (eds.), *Advances in Econometrics, Theory and Applications: 8th World Congress of the Econometric Society,* vol. III, 87-115, Cambridge: Cambridge University Press.

[63] Watson, M.W., and Engle, R.F. (1983). 'Alternative algorithms for the estimation of dynamic factor, MIMIC and varying coefficient regression', *Journal of Econometrics,* 23:385-400.

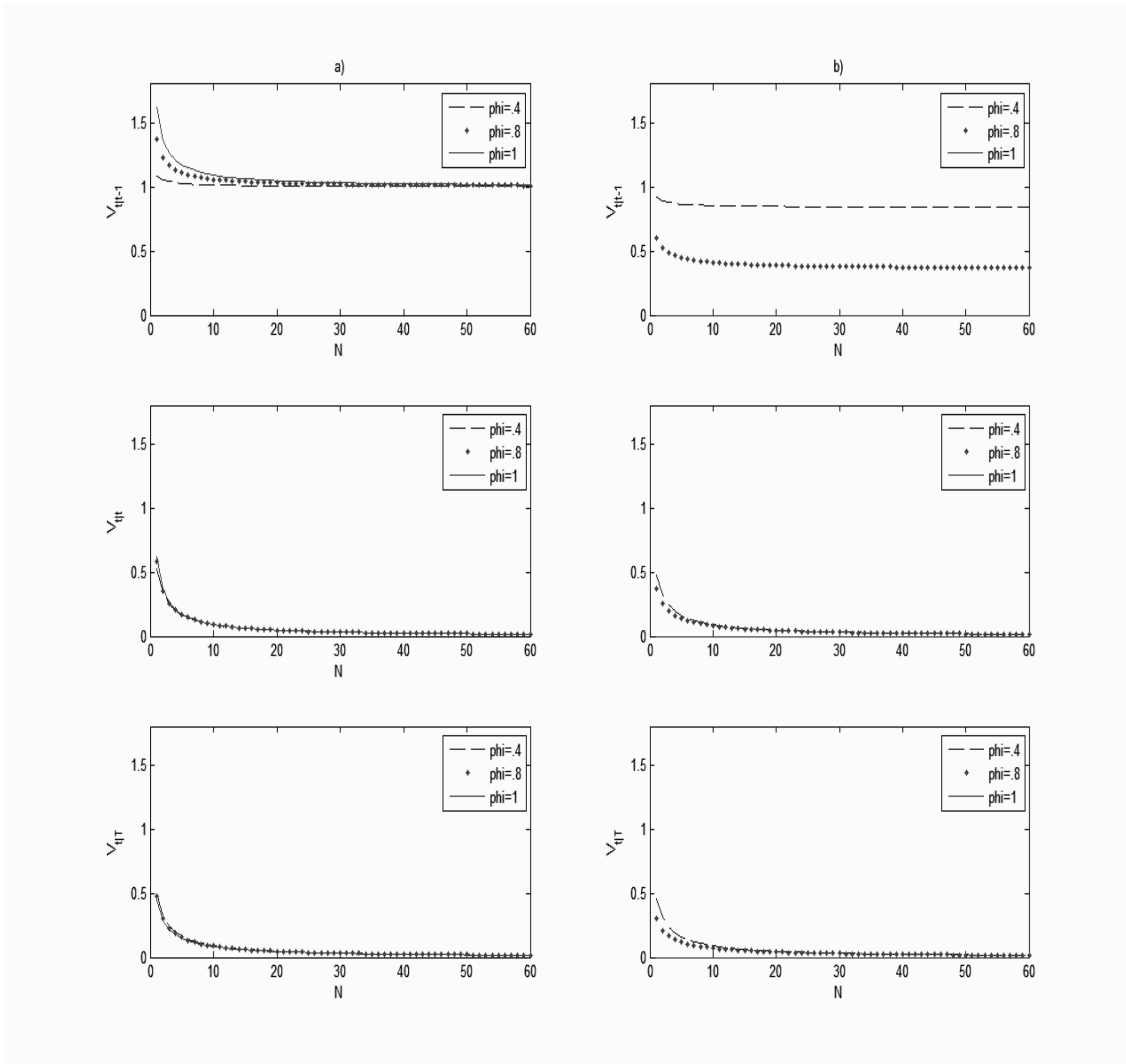

Figure 1: Steady-state MSE of one-step-ahead (first row), filtered (second row) and smoothed (third row) estimates of the underlying factor in a strict DFM with relative loadings $q_i = 1$ for different values of the autoregressive parameter, ϕ. The identifying condition is $\sigma^2_\eta = 1$ (first colum) and $\sigma^2_f = 1$ (second colum).

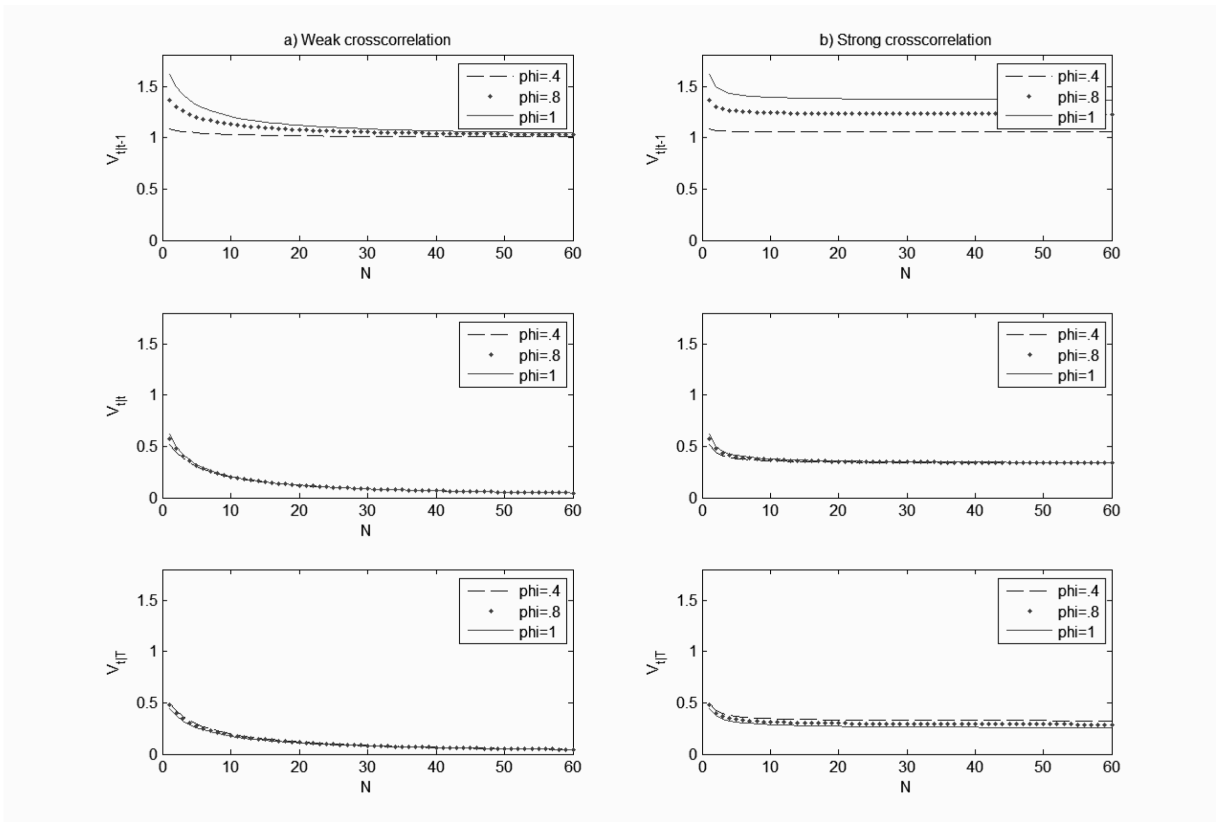

Figure 2: Steady-state MSE of one-step-ahead (first row), filtered (second row) and smoothed (thrid row) estimates of the underlying factor in a DFM with $q_i = 1$ and contemporaneously correlated idiosyncratic noises: Weak correlations (left column) and strong correlations (right column).

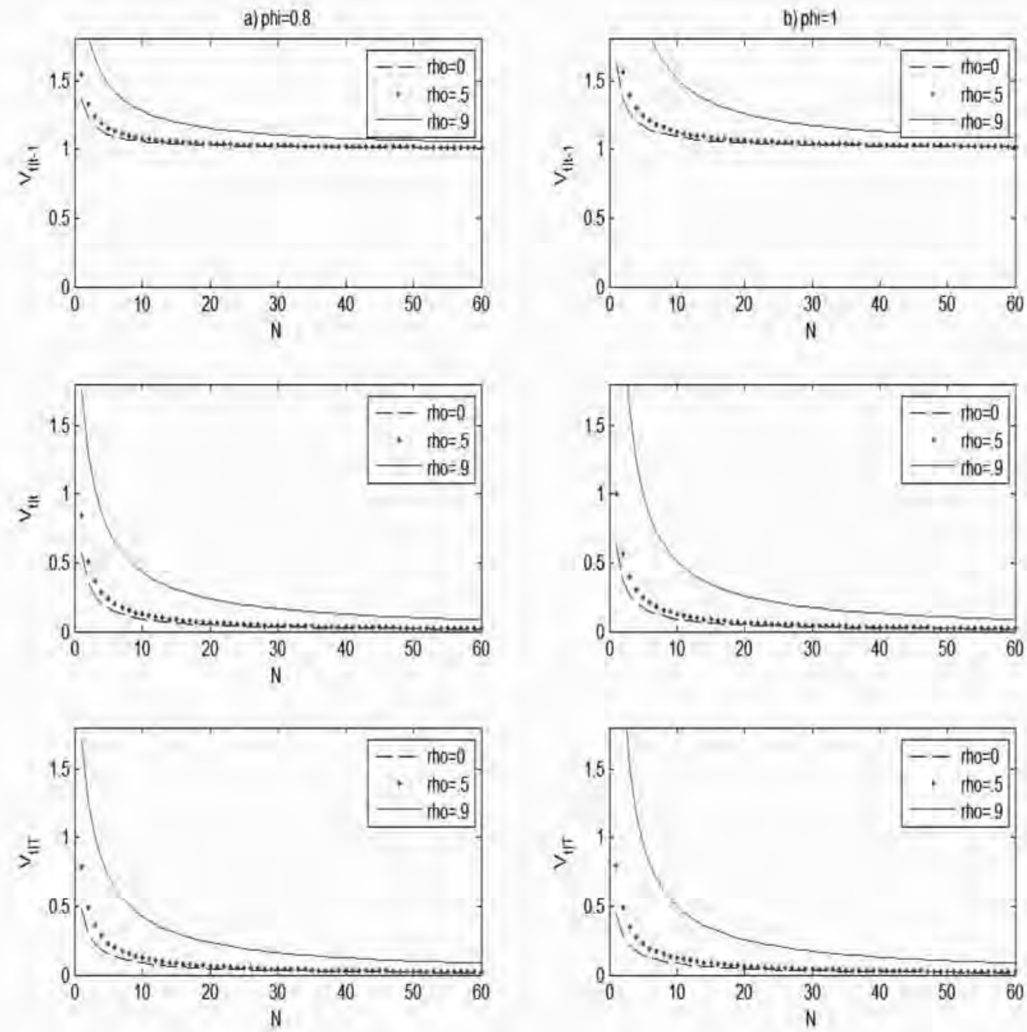

Figure 3: Steady-state MSE of one-step-ahead (first row), filtered (second row) and smoothed (third row) estimates of the underlying factor in a DFM with relative loadings $q_i = 1$ and serially correlated idiosyncratic noises with parameter ρ, for stationary (left column) and non-stationary (right column) factors.

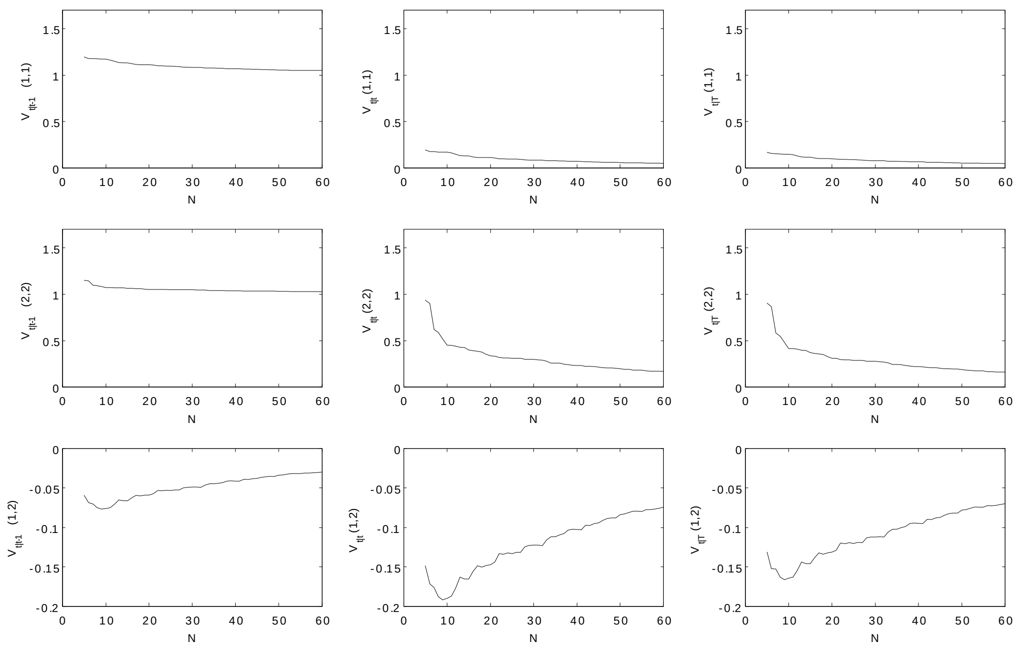

Figure 4: Steady-state MSE matrix of one-step-ahead (first column), filtered (second column) and smoothed (third column) estimates of the underlying factors in a DFM with two factors. The first two rows represent the MSE of the two factors while the third row represents the covariances.

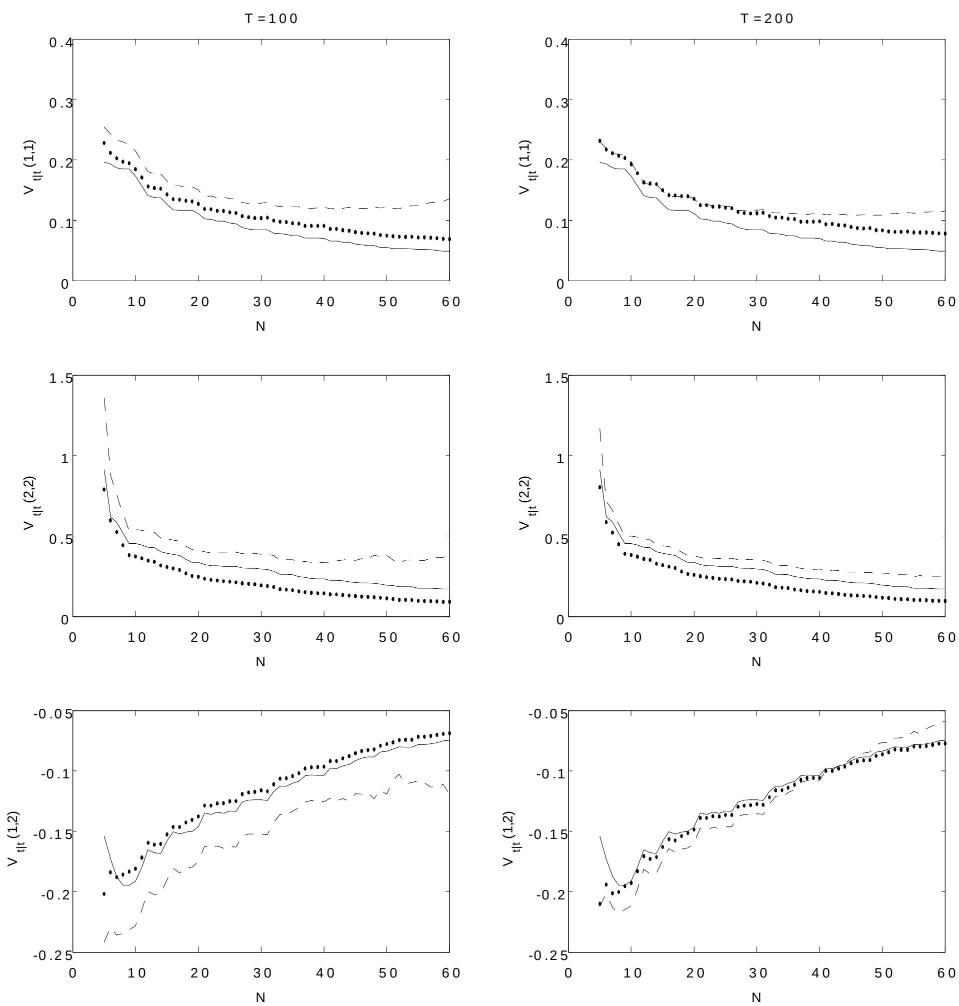

Figure 5: Steady-state total MSE (dashed lines), filter MSE (continuous lines) and estimated parameters MSE (dotted lines) matrices of filtered estimates in a strict DFM with two factors and parameters estimated by ML with $T = 100$ (left column) and $T = 200$ (second column). The first two rows represent the MSE of each of the factors while the third row represents the covariances.

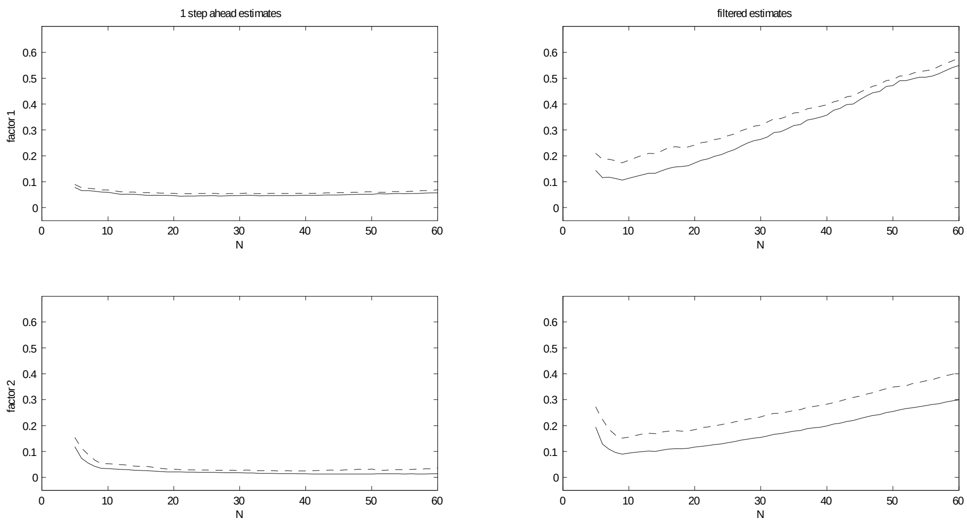

Figure 6: Percentage of the total MSE represented by the parameter uncertainty in a strict DFM with two factors for one-step-ahead (left column) and filtered (right column) estimates when the parameters are estimated with $T = 100$ (continuous lines) and $T = 200$ (dashed lines). The first row represents the results for the first factor while the second row corresponds to the second factor.

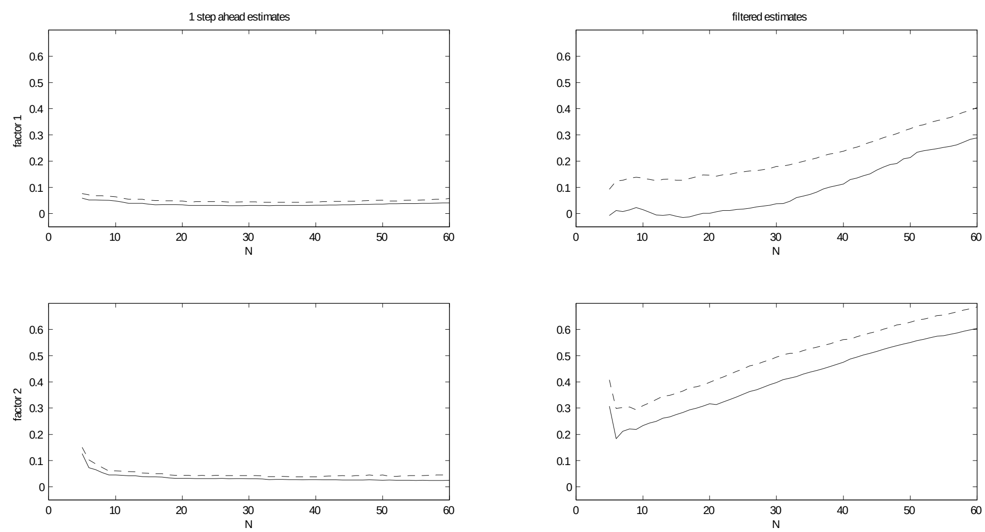

Figure 7: Relative biases in MSEs delivered by the Kalman filter with estimated parameters in a strict DFM with two factors for one-step-ahead (left column) and filtered (right column) estimates when the parameters are estimated with $T = 100$ (continuous lines) and $T = 200$ (dashed lines). The first row represents the results for the first factor while the second row corresponds to the second factor.

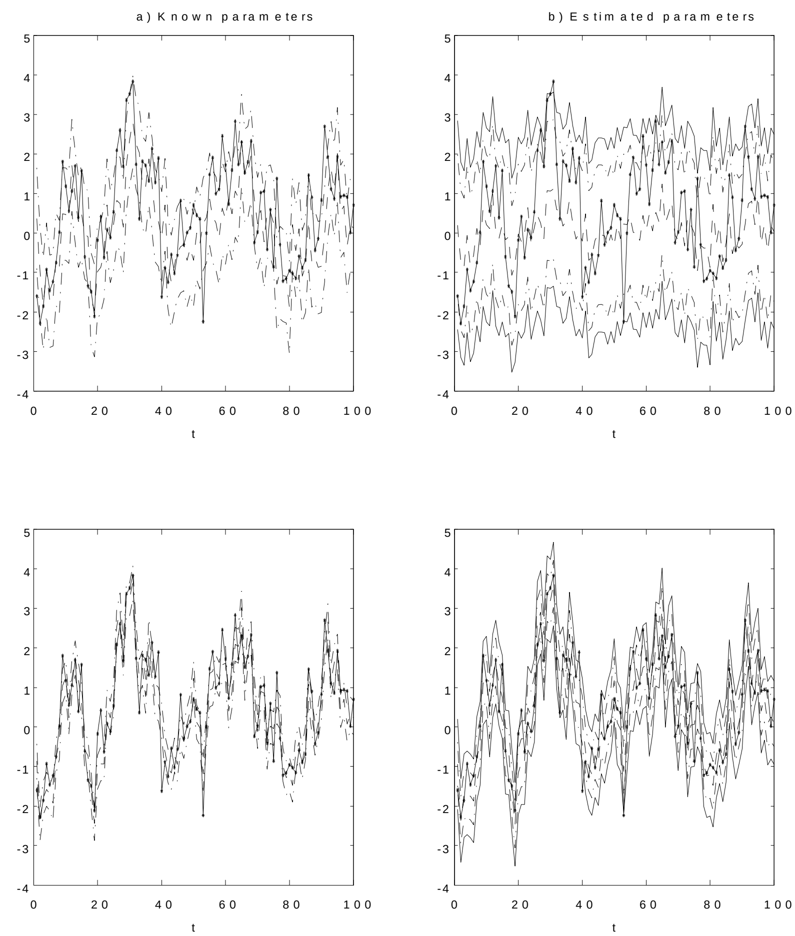

Figure 8: Underlying factor (continuous line) together with its one-step-ahead estimate (dashed line) in a strict DFM with $\phi = 1$, $\sigma_\eta^2 = 1$ and relative loadings $q_i = 1$ when $N = 2$ (top row) and $N = 20$ (low row). The left column also plots the 95% prediction intervals obtained using the MSE delivered by the Kalman filter with known parameters. The right column plots 95% prediction intervals obtained using the MSE delivered by the Kalman filter with estimated parameters and the intervals obtained using the true total uncertainty.