Economics
The Open-Access, Open-Assessment E-Journal

Discussion Paper
No. 2020-7 | March 30, 2020 | http://www.economics-ejournal.org/economics/discussionpapers/2020-7

A comment on the dynamic factor model with dynamic factors

Pilar Poncela and Esther Ruiz

Abstract

In this paper, the authors comment on the Monte Carlo results of the paper by Lucchetti and Veneti (A replication of "A quasi-maximum likelihood approach for large, approximate dynamic factor models" (Review of Economics and Statistics), 2020)) that studies and compares the performance of the Kalman Filter and Smoothing (KFS) and Principal Components (PC) factor extraction procedures in the context of Dynamic Factor Models (DFMs). The new Monte Carlo results of Lucchetti and Veneti (2020) refer to a DFM in which the relation between the factors and the variables in the system is not only contemporaneous but also lagged. The authors´ main point is that, in this context, the model specification, which is assumed to be known in Lucchetti and Veneti (2020), is important for the properties of the estimated factors. Furthermore, estimation of the parameters is also problematic in some cases.

JEL C15, C32, C55, C87
Keywords Dynamic factor models; EM algorithm; Kalman filter; principal components

Authors
Pilar Poncela, Universidad Autónoma de Madrid, Spain
Esther Ruiz, Universidad Carlos III de Madrid, Spain, Ortega@est-econ.uc3m.es

Citation Pilar Poncela and Esther Ruiz (2020). A Comment on the Dynamic Factor Model with Dynamic Factors. Economics Discussion Papers, No 2020-7, Kiel Institute for the World Economy. http://www.economics-journal.org/economics/discussionpapers/2020-7

Received March 27, 2020 **Accepted as Economics Discussion Paper** March 30, 2020 **Published** March 30, 2020

1 Introduction

Dynamic Factor Models (DFMs) are a very powerful and popular tool to reduce the dimension of large systems of economic and financial variables by assuming that their dynamic dependence relies on a relatively small number of underlying unobserved common factors. New results pushing forward the frontiers of knowledge about the properties of DFMs, as those in the Discussion Paper by Lucchetti and Venetis (2020) (LV20), published in *Economics*, are always welcome. In particular, LV20 investigate the properties of factors extracted using Kalman Filter and Smoothing (KFS) algorithms. First, they replicate the Monte Carlo experiments carried out by Doz, Giannone and Reichlin (2012) in which the Data Generating Process (DGP) is a static DFM (S-DFM) with the variables in the system contemporaneously related with the factors, whose number is assumed to be known. Second, LV20 extend the simulations to "*more realistic DGPs, more interesting for potential users*". With this purpose, they consider DGPs in which the factors are related with the variables in the system not only contemporaneously but also with lags. Many authors refer to this model as the dynamic version of the DFM (D-DFM). In their Monte Carlo experiments, LV20 compare the performance of the factors extracted using KFS procedures with that of the factors extracted using procedures based on Principal Components (PC). LV20 conclude that Maximum Likelihood (ML) based on KFS factor extraction is often the dominant method and that the persistence characteristics of the system under analysis play a crucial role. Furthermore, they also conclude that a correct specification of the underlying dynamics is of paramount importance.

In this comment, we put forward some limitations of the analysis carried out by LV20 with the aim of giving directions for future research that could help moving forward our knowledge about DFMs.

2 The model

To clarify our discussion, we follow LV20 and define the DFM as follows

$$x_t = \Lambda_0 f_t + \Lambda_1 f_{t-1} + ... + \Lambda_s f_{t-s} + e_t \quad (1)$$

$$f_t = A_1 f_{t-1} + A_2 f_{t-2} + ... + A_p f_{t-p} + u_t \quad (2)$$

$$e_t = D e_{t-1} + v_t \quad (3)$$

where x_t is the $N \times 1$ vector of observations at time $t = 1, ..., T$, which is assumed to be stationary, f_t is the $q \times 1$ vector of unobserved factors at time t and e_t is the corresponding $N \times 1$ vector of idiosyncratic components, which are assumed to be weakly cross-sectional and serially correlated.

The factors and idiosyncratic components are mutually uncorrelated for all lags and leads. The noise v_t is assumed to be normal with zero mean and covariance matrix $\mathcal{T}$, defined in LV20.

The first important issue about the specification of the D-DFM in equations (1)-(3) is related with the identifying restrictions; see, for example, Trenkler and Weber (2016) and the references therein for problems related with the identification of unobserved component models and Bai and Ng (2013) for identification issues related to PC factor extraction. The identification issues could be very relevant when estimating the model parameters. In the context of the D-DFM there is an even more important identification issue related with identifying simultaneously the lag order of the VAR model for the factors, p, and the number of lags, s, in equation (1). Defining $\Lambda(L) = \Lambda_0 + \Lambda_1 L + \Lambda_s L^s$ and $A(L) = I_q - A_1 L - A_2 L^2 - A_p L^p$, the D-DFM can be written as follows

$$x_t = \Pi(L) u_t + e_t \tag{4}$$

where $\Pi(L) = \Lambda(L) A(L)^{-1}$. The model in equation (4) is known as Generalized DFM and the parameters of the infinite lag polynomial $\Pi(L)$ matrix cannot be estimated by ML or QML. Several authors propose estimating them using frequency-domain procedures; see Forni, Hallin, Lippi and Reichlin (2000, 2004, 2005). The identification problem appears because from $\Pi(L)$ it is not possible to recover the polynomials $\Lambda(L)$ and $A(L)$ in a unique form without imposing restrictions; see, for example, the discussion in Lütkepohl (2006).

The second comment about the DFM in equations (1)-(3) is that it should be made crystal clear that its specification is known, in the sense that q, s and p are assumed to be known. Knowledge of these quantities is related with the last identification issue mentioned above. However a "*potential user*" should start from scratch by determining these quantities before she can use KFS to extract the factors. A correct determination of these quantities could be crucial for the good properties of KFS being, at the same time, a very difficult task (mainly when the idiosyncratic components are serially and cross-sectionally correlated). Although LV20 conclude that a correct specification of the underlying dynamics is of paramount importance, they do not challenge the KFS factor extraction when the relevant quantities, p, q and s, are unknown.

3 Factor extraction (smoothing), parameter estimation and model specification

There are two main types of procedures for extracting factors in the context of DFMs. First, one can use procedures based on PC that, although simple computationally, are not efficient as they do not take into account the temporal dependence of the factors. However, this is also the main advantage of PC factor extraction as it is a non-parametric procedure that does not require any assumption about this temporal dependence. Alternatively, methods based on the KFS open the way to ML estimation and, consequently, if the model is correctly specified, they are efficient. Furthermore, KFS procedures allow to easily deal with several data irregularities as, for example, missing observations, mixed frequencies and aggregation constraints. However, the correct specification of the model, and therefore, the efficiency of KFS factor extraction, involves choosing the number of factors and the number of lags in the corresponding VAR model as well as the temporal relationship between the variables in the system and the factors. In summary, you need to determine q, p and s. Furthermore, one needs to assume a particular specification for the temporal and cross-sectional dependences of the idiosyncratic noises (although this seems to be less important in practice).

If the specification and the model parameters were known, the DFM in equations (1)-(3) can be written as a state space model and the KFS algorithms can be used to extract the factors and to construct confidence intervals. As an illustration, we simulate two systems by the same one-single factor S-DFM considered by LV20 and Doz, Giannone and Reichlin (2012). The first system is simulated with $N = 10$ and $T = 50$ while the second one is simulated with $N = T = 100$. The only difference between our DGP and that in the above mentioned papers is that, in this comment, following Breitung and Tenhofen (2011) and Poncela and Ruiz (2015), we randomly draw the elements of the factor loading matrix, Λ_0, from a uniform $[0, 1]$ distribution instead of from a standard normal variable.[1] The top left panel of Figure 1 plots the true simulated factor together with the estimated factor obtained using KFS and the corresponding 95% confidence intervals.[2] The top right panel of Figure 1 plots the same quantities when the factor is extracted by minimizing the same criterium as in the PC estimator but assuming that the loadings are known, i.e. by Ordinary Least Squares (OLS). The same quantities for $N = T = 100$ are plotted in the bottom panels of Figure 1. Looking at the results for $N = 10$, we can observe that, not only the KFS point estimates of the factor are closer to the true factor than PC estimates, but also that their confidence intervals are thinner (and still contain the true factor with the desired nominal coverage). Obviously, when $N = 100$, regardless of whether the factors are extracted using KFS or PC, the extracted and true factors are much closer to each other and the confidence intervals are very tiny. However, we can still observe that the confidence intervals of the PC factors are slightly wider than those of the KFS factors. This is an effect of the lack of efficiency of PC that is not using information about the (very strong) factor dependence. If the autoregressive parameter of the factor were smaller than 0.9, the differences between KFS and PC are expected to be also smaller.

We also illustrate the results in the context of D-DFMs by simulating a system by the same

[1] See our comment below about simulating the weights from a standard normal distribution.

[2] The KFS is run as if the DFM were exact although it is not.

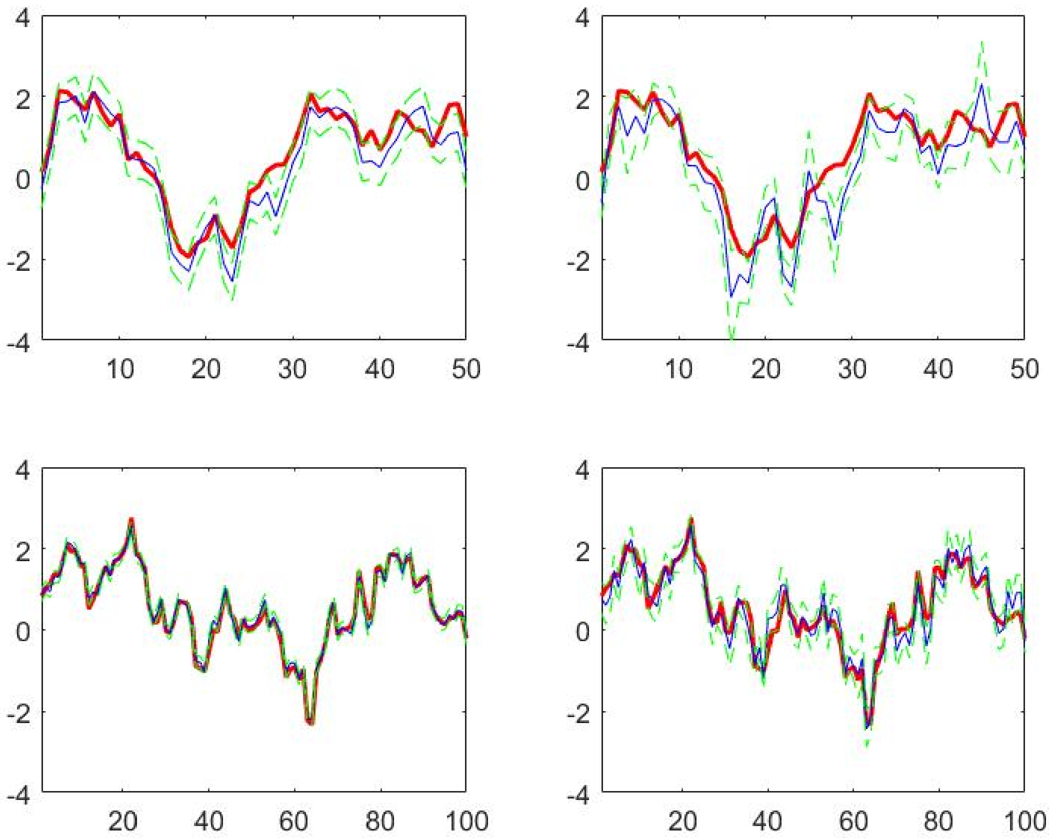

Figure 1: True factor (thick red line) simulated by a S-DFM with a unique factor. The dimensions are $N = 10$ and $T = 50$ (top panels) and $N = T = 100$ (bottom panels). The estimated factors (thin blue lines) together with 95% confidence bounds (discontinuous green lines) are obtained assuming known parameters by KFS (left panel) and OLS (right panel).

D-DFM considered by LV20 with $p = 1$, $s = 1$ and $q = 1$. As above, the only difference with the DGP considered by LV20 is that we generate the factor loadings from a uniform $[0, 1]$ instead of using a standard normal. Figure 2 plots the same quantities described above when the factor is extracted using KFS. Note that comparing these plots with those in the first column of Figure 1, we cannot observe any appreciable difference. If the specification and the parameters are known, it is not relevant whether the model is static or dynamic when implementing KFS. Note that PC factor extraction in the D-DFM is not well solved in the literature. Although LV20 try different alternatives, we do not explore them further in this comment.

The illustration above is not realistic because, in practice, the model parameters are unknown and should be estimated. If the idiosyncratic components are serial and cross-sectionally uncorrelated, the DFM parameters can be estimated by ML regardless of the cross-sectional dimension, N. However, if the idiosyncratic components are serially and cross-sectionally correlated and N

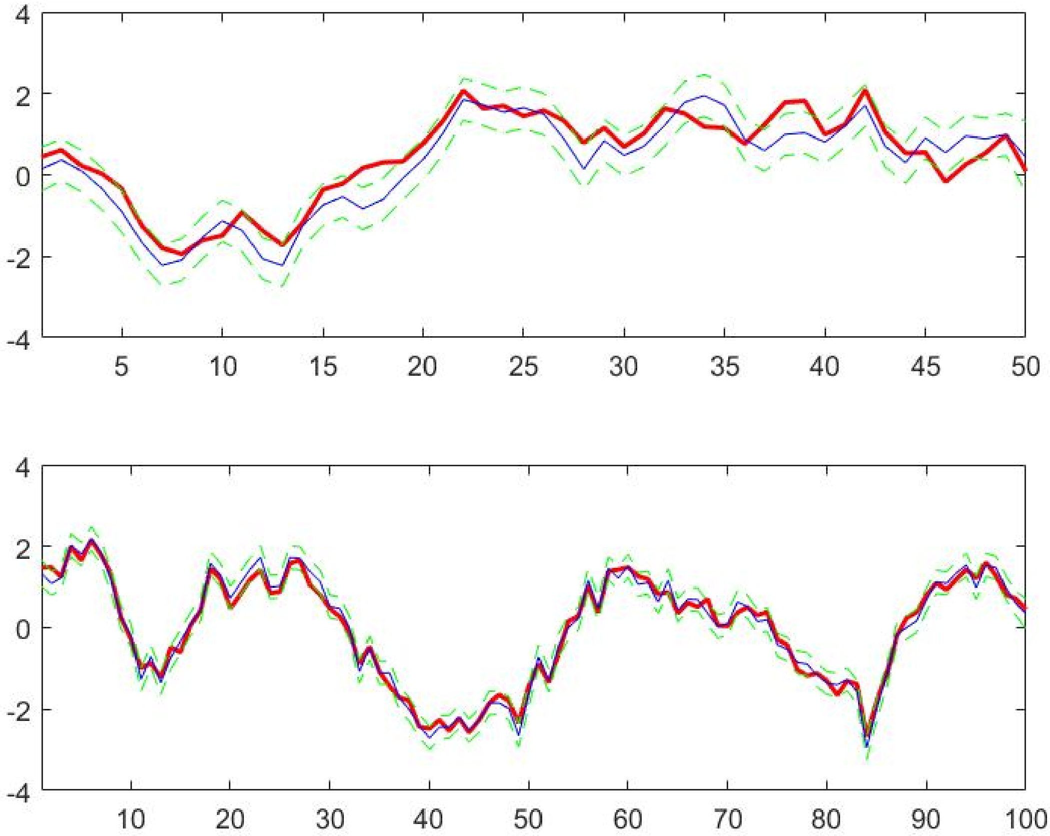

Figure 2: True factor (thick red line) simulated by a D-DFM with a unique factor ($q = 1$) and one lag ($s = 1$). The idiosyncratic noises are heteroscedastic and the autoregressive parameter of the factor is $\phi = 0.9$. The dimensions are $N = 10$ and $T = 50$ (top panel) and $N = T = 100$ (bottom panel). The estimated factors (thin blue lines) together with 95% confidence bounds (discontinuous green lines) are extracted, assuming known parameters, by KFS.

is very large, ML is not feasible due to the extremely large number of parameters that need to be estimated. In a very important contribution, Doz, Giannone and Reichlin (2012) show that procedures based on KFS in S-DFMs are still efficient even if they are implemented assuming that the idiosyncratic components were serially and cross-sectionally uncorrelated when they are not. Doz, Giannone and Reichlin (2012) propose estimating the parameters of the S-DFM by the EM algorithm using as starting values the parameters obtained using the PC factors. This is a QML estimator.

Next, we illustrate the KFS estimation of the factors based on the same S-DFM considered above when the parameters are estimated by the QML estimator proposed by Doz, Giannone and Reichlin (2012).[3] As before, the left panels of Figure 3 plot the true and smoothed factors

[3]Note that this is the case considered in the simulations by LV20 with the model assumed to be known and

together with their corresponding 95% bounds when $N = 10$ and $T = 50$ (top panel) and when $N = T = 100$ (bottom panel). The corresponding factors extracted using PC have also been ploted in the right panels of Figure 3. It is important to note that the identifying restrictions are different when the factors are extracted by KFS or by PC. The factors are estimated up to a rotation which, in this case, given that $q = 1$ and $s = 0$, is a change of scale. Consequently, in order to check whether the confidence intervals contain the true factors, we follow Poncela and Ruiz (2016) and rotate the estimated factors to be in the same scale of the true factors as follows

$$\hat{f}^{*} = \hat{f}\left[\left(\hat{\Lambda}'\hat{\Lambda}\right)^{-1}\hat{\Lambda}'\Lambda\right]^{-1}. \tag{5}$$

The Mean Square Errors (MSEs), needed to construct the intervals, should be accordingly transformed. Looking at Figure 3, we can observe that, for this particular realization when $N = T = 100$, the conclusions are very similar to those obtained from Figure 1 with known parameters. In this case, the role of parameter uncertainty is very mild. Even when $N = 10$ and $T = 50$, the estimated KFS factor is very similar to that extracted with known parameters although the confidence intervals are slighly thiner (they do not incorporate parameter uncertainty). The effect of parameter estimation on PC factors seems to be very mild. The only apparent effect on this particular realization is a slight increase of the length of confidence intervals.

However, as mentioned above, in practice, one needs to decide about the specification of the DFM and this decision is crucial when considering the properties of the factors extracted using KFS. To illustrate this point, we consider again the same basic systems described above simulated by the S-DFM with $N = 10$ and $T = 50$ and with $N = T = 100$. The first step of any empirical analysis is to determine the number of common factors, q. Setting the maximum number of common factors to 5 (50) and applying the test by Onatski (2010), the number of common factors is $\hat{q} = 3$ ($\hat{q} = 5$) when $N = 10$ ($N = 100$). To shed some light about the number of common factors, Figure 4, which plots the corresponding scree plots, illustrates that, even if the true underlying model only has one common factor, the tests to determine q are missleading. Obviously, this misspecification will have consequences on the estimation of the parameters and the quality of the factor estimates. Finally, note that, if the model is dynamic instead of static, the identification is further complicated with serious consequences over the properties of the extracted factors.

This example just illustrates the difficuty involved in the identification, even in the context of this very simple S-DFM, when the idiosyncratic errors are temporal and cross-sectionally correlated. Identification in the context of the D-DFM is even more challenging. This is an issue that deserves further research.

---

the parameters estimated by QML.

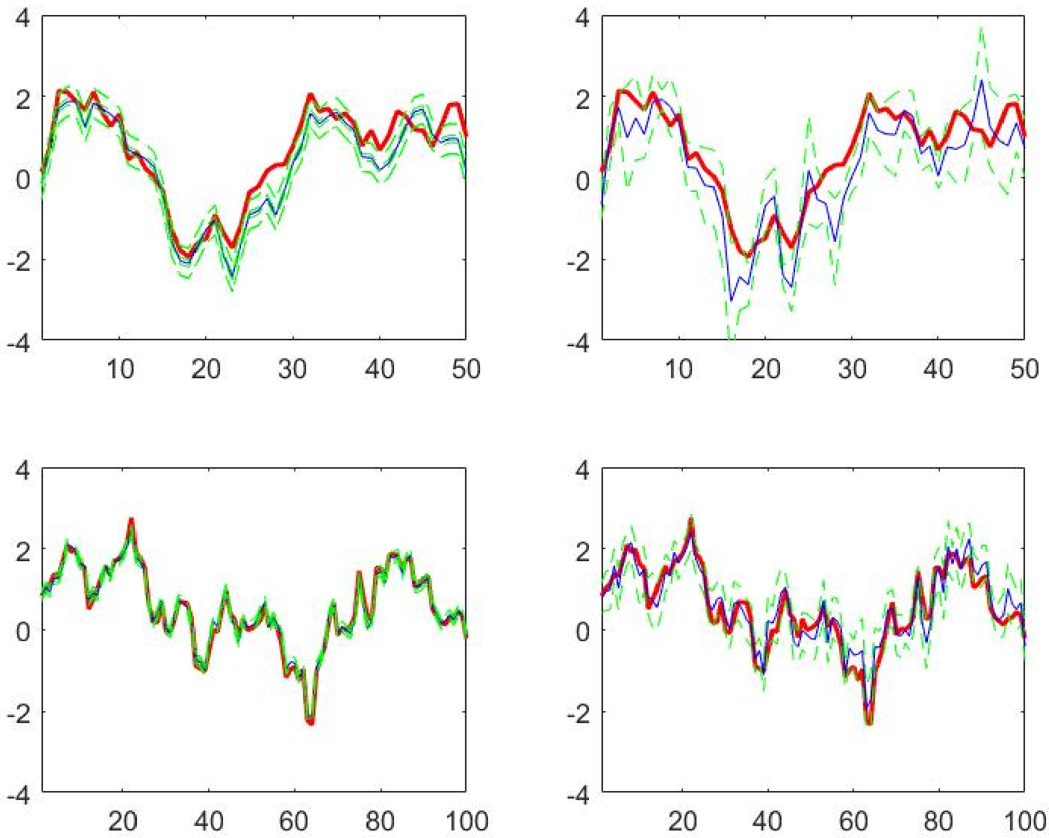

Figure 3: True factor (thick red line) simulated by a static DFM with a unique factor. The idiosyncratic noises are heteroscedastic and the autoregressive parameter of the factor is $\phi = 0.9$. The dimensions are $N = 10$ and $T = 50$ (top panels) and $N = T = 100$ (bottom panels). The estimated factors (thin blue lines) together with 95% confidence bounds (discontinuous green lines) are extracted, after estimating the parameters assuming that the model is known, by KFS (left panel) and PC (right panel).

4 The Monte Carlo design

The first comment to make about the design of the Monte Carlo experiments in LV20 is that there is not any mention to the distribution of the errors u_t. We assume that, as often done in the related literature and, in particular, in the Monte Carlo experiments in Doz, Giannone and Reichlin (2012), LV20 assume them to be standard normal. Although normality could be a good approximation to start with, it could also be useful to investigate the performance of KFS under other distributions of the factors and/or idiosyncratic components; see Barigozzi and Luciani (2019) who carry out Monte Carlo experiments assuming that the innovations are Student-t. Our guess is that, while point estimates of the factors may not be severely affected by non-normal errors, interval estimates could be. Therefore, analyzing also the performance of

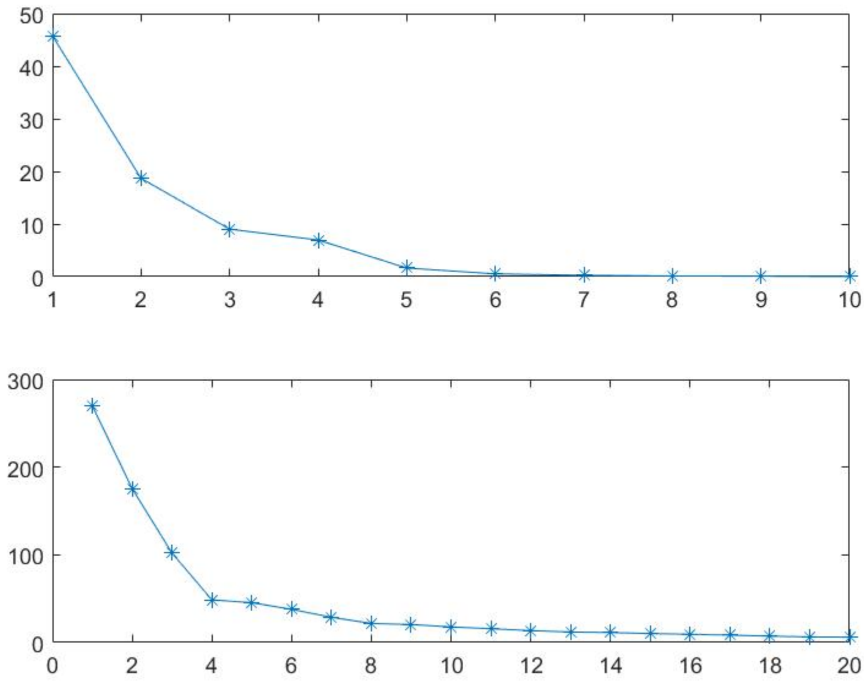

Figure 4: Scree plot of a system simulated by a S-DFM with $N = 10$ and $T = 50$ (top panel) and $N = T = 100$ (bottom panel).

confidence regions for the underlying factors could be of interest.

Second, with respect to the design of the covariance matrix, the Monte Carlo results reported by LV20 are obtained for the case in which $d_i = d, \forall i = 1, ..., N$. What is the point in changing the dynamics of the idiosyncratic components and, consequently, the definition of the matrix $\mathcal{T}$ with respect to that in Doz, Giannone and Reichlin (2012) by allowing each idiosyncratic noise to have their own autoregressive parameter if then the results are reported only for the case in which all parameters are the same?.[4]

Third, another important point about the Monte Carlo design is related with the conclusion by LV20 about the identification and parameter estimation being related with the persistence of the factors. Even though LV20 claim in the abstract that "*the persistence characteristics of the observable series play a crucial role*", they only report results for a very persistent factor, with its autoregressive parameter being 0.9. In order to analyse the role of persistence on the results, one should at least consider different levels of persistency in the factors. This is indeed a very crucial question for users of DFMs.

Fourth, although it is a common practice in the related literature, simulating the weights from

[4]It is also surprising that LV20 report 6 tables considering the same Monte Carlo design as in Doz, Giannone and Reichlin (2012) and just one table with the "extended" design.

a $N(0,1)$ does not seem to be a good idea. For the factors to be pervasive, a large enough number of weights should be different from zero. Simulating the weights from a $N(0,1)$, a large number of simulated weights could be close to zero and, consequently, it could generate weak factors; see Chudik, Pesaran and Tosetti (2011) and Onatski (2012) for the implications of weak factors.

Fifth, in their Monte Carlo experiments, LV20 drop the replicates for which $TR < 0.05$.[5] The authors are throwing away replicates for which the method is not working. Obviously, the results reported are better than they shoud be. It is important to know at least how many replicates are discarded.

Sixth, in the case of the D-DFM considered as DGP, it is not clear what the authors are estimating when implementing the estimation procedures proposed by Doz, Giannone and Reichlin (2011, 2012) that are designed for S-DFMs. In their description of these procedures, LV20 say "*estimating* $\hat{f}_t$" as if there were no lags. Does it mean that they are extracting the factors as if s were zero in equation (1)? If this is so, this is not exactly an estimation of the D-DFM. Some authors propose estimating the D-DFM by looking at its corresponding "static" version with the vector of factors being $F_t = (f_t, f_{t-1}, ..., f_{t-s})$. For example, Stock and Watson (2005) propose estimating the $q(s+1)$ vector of "static" factors, F_t by PC and then, for a given lag p, estimate the "restricted" VAR coefficients by regressing $\hat{F}_t$ onto the desired number of lags.[6] Barigozzi and Luciani (2019) also consider a D-DFM writen as a S-DFM with the vector of factors being F_t and the corresponging factor equation having a singular covariance matrix of the noise. Barigozzi and Luciani (2019) investigate the implications of estimating this model assuming that it is exact when the idiosyncratic components are serial and cross-sectionally correlated and report numerical results from an extensive Monte Carlo study.

Seventh, in any case, our main concern is about the design of the Monte Carlo experiments being too restrictive as to be of real interest for users in the sense that the number of factors, q, and the lags, s and p, are assumed to be known. For the Monte Carlo experiments to be of real interest for users, one should consider uncertainty about the number of factors and their dependences.

5 Conclusions

The results in the extended Monte Carlo experiments carried out by LV20 are a step further in the knowledge about the properties of DFMs. However, a more ambitious objective could be of interest for the results to be truly useful for potential users. In particular, it is interesting

[5]It is not clear what the authors mean by TR. Is the TR of any method?

[6]Note also that LV20 mention that they extract qs instead of $q(s+1)$ factors when implementing PC.

to investigate the properties of the KFS factors when there is uncertainty about the correct specification. This is a question still open to further research.

References

[1] Bai, J. and Ng, S. (2013). Principal components estimation and identification of static factors. *Journal of Econometrics*, 176, 18-29.

[2] Barigozzi, M. and Luciani, M. (2019). Quasi Maximum likelihood estimation and inference of large approximate dynamic factor models via the EM algorithm. *Manuscript.*

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

[4] Chudik, A., Pesaran, M.H. and Tosetti, E. (2011), Weak and strong cross-section dependence and estimation of large panels, *Econometrics Journal*, 14, 45-90.

[5] 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.

[6] Doz, C., Giannone, D. and Reichlin, L. (2012). A quasi maximum likelihood approach for large approximate dynamic factor model. *Review of Economics and Statistics*, 94, 1014-1024.

[7] 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.

[8] Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2004). The generalized Dynamic factor model. Consistency and rates. *Journal of Econometrics*, 119(2), 231-255.

[9] Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2005). The generalized Dynamic factor model. One-sided estimation and forecasting. *Journal of the American Statistics Association*, 100(471), 830-840.

[10] Lucchetti, R. and Veneti, I.A. (2020). A replication of "A quasi-maximum likelihood approach for large, approximate dynamic factor models" (Review of Economics and Statistics, 2012). *Economics*, Discussion Paper No. 2020-5.

[11] Lütkepohl, H. (2006). Forecasting with VARMA models. In Elliot, G., Granger, C.W.J., Timmermann, A. (eds.), *Handbook of Economic Forecasting*, Vol. 1, 287-325.

[12] Onatski, A. (2010). Determining the number of factors in approximate from empirical distribution of eigenvalues. *Review of Economics and Statistics*, 92(4), 1004-1016.

[13] Onatski, A. (2012), Asymptotics of the principal components estimator of large factor models with weakly influential factors, *Journal of Econometrics*, 168(2), 244-258.

[14] Poncela, P.and Ruiz, E. (2015). More is not always better: back to the Kalman filter in Dynamic Factor Models, in Koopman, S.J. and Shephard, N.G. (eds.), *Unobserved Components and Time Series Econometrics*, Oxford University Press.

[15] Poncela, P. and Ruiz, E. (2016). Small versus big-data factor extraction in Dynamic Factor Models: An empirical assessment, in Hillebrand, E. and Koopman, S.J. (eds.), *Dynamic Factor Models (Advances in Econometrics)*, Vol. 35, 401-434, Emerald Group Publishing Limited.

[16] Stock, J.H. and Watson, M.W. (2005), Implications of dynamic factor models for the VAR analysis, Working paper 11467, National Bureau of Economic Research.

[17] Trenkler, C. and Weber, E. (2016), On the identification of multivariate uncorrelated unobserved components models, *Economics Letters*, 138, 15-18.

Economics
The Open-Access, Open-Assessment E-Journal

Please note:

You are most sincerely encouraged to participate in the open assessment of this discussion paper. You can do so by either recommending the paper or by posting your comments.

Please go to:

http://www.economics-ejournal.org/economics/discussionpapers/2020-20

The Editor