UNIVERSITÀ COMMERCIALE LUIGI BOCCONI

PHD IN ECONOMICS
CICLO XX

BOOTSTRAP METHODS FOR DYNAMIC FACTOR MODELS

MARIO PORQUEDDU
Matr. 1017037

COMMISSION:
PROF MASSIMILIANO MARCELLINO, EUI
PROF FRANCESCO CORIELLI, UNIVERSITA' BOCCONI
PROF LUCA SALA, UNIVERSITA' BOCCONI

Contents

Preface

Dynamic Factor Models (DFM) have been proposed as an useful way to summarize the information contained in large datasets with a few latent factors.

A simple way to represent DFM is to view them as in Stock Watson (2005), as a restricted vector autoregressive model where the time series set includes observable variables and a few (latent) static factors (which can be consistently estimated with principal components from a large dataset with T observations and N series). These particular VAR models have been called Factor Augmented VARs (FAVARs).

Inferential theory for these models is reported in a series of papers published in Econometrica by Jushan Bai and Serena Ng, where they show that the factor loadings are $\sqrt{T}$ consistent and asymptotically normal if $\sqrt{T}/N \to 0$ also when the factors are generated by a VAR process. The same theoretical results have been extended to applications like diffusion index forecasts in Bai-Ng (2006a) and to tests for the null hypothesis that an observed economic series is the unobserved factor (Bai-Ng(2006b)).

In this thesis we look at these problems and applications trying to find to what extent there are asymptotic theory-based tests failures , and to what extent bootstrap methods can help us in fixing them. The main result is that asymptotic theory confidence intervals are heavily influenced by data generating process parameters and by small sample biases, especially for highly nonlinear statistics. Bootstrap methods are designed for dealing with these problems, but there are so many techniques for resampling and providing bootstrap confidence intervals, that each application or model may require a different strategy. These methods are very helpful when we deal with biased estimates, because they provide a relatively simple way to estimate the bias, also if bias correction maybe counterproductive (as an example with the common component). The block bootstrap and simple percentile intervals work well when we deal with the common component and if we want to give an economic interpretation to our latent factors. Semiparametric bootstrap are more helpful when we deal with complicated nonlinear functions, as in the second paper related to impulse responses.

Acknowledgements

I am grateful to Professor Massimiliano Marcellino for helpful supervision and unbelivable patience.

I am indebted to Alessia Paccagnini, Sara Pinoli, Paolo Bianchi, Paolo Di Giannatale, Claudia Foroni, Linlin Niu, Elena Besedina and my family for their warm support.

Chapter 1

Diffusion Index Forecasts: Asymptotic Theory and Bootstrap Methods

1.1 Introduction

Dynamic Factor Models (DFM) have been proposed as an useful way to summarize the information contained in large datasets with a few latent factors.

A simple way to represent DFM is to view them as in Stock Watson (2005), as a restricted vector autoregressive model where the time series set includes observable variables and a few (latent) static factors (which can be consistently estimated with principal components from a large dataset with T observations and N series). These particular VAR models have been called Factor Augmented VARs (FAVARs).

Inferential theory for these models is reported in a series of papers published in Econometrica by Jushan Bai and Serena Ng, where they show that the factor loadings are $\sqrt{T}$ consistent and asymptotically normal if $\sqrt{T}/N \rightarrow 0$ also when the factors are generated by a VAR process. The same theoretical results have been extended to applications like diffusion index forecasts in Bai-Ng (2006a) and to tests for the null hypothesis that an observed economic series is the unobserved factor (Bai-Ng(2006b)).

Stock Watson (1999) show that a few static factors provide more precise US inflation forecasts Other papers which apply this technique to US data are

Stock Watson (2002b), Bernanke Boivin(2003) and Boivin Ng(2003); for other countries see Forni, Hallin, Lippi and Reichlin (2003b, Euro Area), Brisson, Campbell, Galbraith (2002, Canada) and Artis, Banejee and Marcellino (2001, UK).

In this paper we show that if we look at a static factor model with a more complicated dynamic structure than the one reported in the previous literature, the forecast confidence intervals obtained according to asymptotic theory are too small in small samples. We want to check if bootstrap methods can provide confidence intervals with a rejection probability close to the nominal one and with a reasonable length.

This paper has six sections, including this introduction.

In the second section we present the Dynamic Factor Model, as presented in Bai (2003) and we review estimation issues and asymptotic theory. In the third section we present different bootstrap methods which depend on how we create artificial samples and on how we build confidence intervals.

The fourth section discusses a DFM model with two common shocks and four static factors and compares confidence intervals for the common part based on asymptotic theory and bootstrap methods in terms of coverage probability and interval length . The fifth section presents a simulation related to diffusion index forecasts similar to the one proposed in Bai-Ng (2006), and in the sixth section we illustrate the methods with an empirical example related to forecasting of industrial production and inflation.

1.2 The model

The dynamic factor model expresses each series as a function of contemporaneous and lagged dynamic factors and idiosyncratic error terms:

$$X_{ti} = f_t\bar{\lambda}_i(L) + e_{ti}, i = 1, ., n \tag{1.1}$$

$$\bar{\lambda}_i(L) = \sum_{j=0}^{p} \bar{\lambda}_{i,j} L^j \tag{1.2}$$

where e_{ti} is the idiosyncratic error while f_t is a $1 \times q$ vector of covariance stationary unobserved dynamic factors and $\bar{\lambda}_i(L)$ is a $q \times 1$ vector of lag polynomials. Factors and idiosyncratic terms are assumed to be uncorrelated with each other. Moreover:

$$E(e_{ti}e_{sj}) = 0 \text{ for all i,j,t,s, if } i \neq j \tag{1.3}$$

This last strong assumption characterizes the "Exact Dynamic Factor Model " (DFM) proposed by Geweke (1977) and Sargent and Sims (1978). Chamberlain and Rothschild (1983) introduced the Approximate Dynamic Factor Model (ADFM) which allows for a limited amount of correlation between different idiosyncratic terms; they assume that the maximum eigenvalue of the matrix $\Omega = E\left(e_t e_t'\right)$ is bounded.

The $1 \times r$ vector $F_t = \begin{bmatrix} f_t & f_{t-1} & ... & f_{t-p} \end{bmatrix}$ includes r contemporaneous and lagged factors. The maximum of r is equal to $q(p+1)$ if all factors and lagged factors are loaded with nonzero loadings. λ_i includes all stacked factor loadings, so that the dynamic factor model in static form can be written as:

$$X_{ti} = F_t \lambda_i + e_{ti} \tag{1.4}$$

In this paper we will focus our attention on the asymptotic theory proposed in Bai (2003), which can be applied to a more general approximate factor model which allows for limited time series and cross section dependence in the idiosyncratic components. His results can be applied to our exact dynamic factor model.

Following his notation the DFM can be summarized as:

$$\underset{TxN}{X} = \underset{Txr}{F} * \underset{rxN}{\Lambda} + \underset{TxN}{e} \tag{1.5}$$

Small scale DFM can be estimated by Gaussian Maximum Likelihood using the Kalman filter (Engle and Watson (1981), Stock and Watson (1989,1991), Sargent (1989), and Quah and Sargent (1993)). This estimation strategy is in the literature considered unfeasible for large panels, but in a recent paper Doz, Giannone and Reichlin show that a quasi-maximum likelihood estimation method can be applied to this case. We follow the results in Chamberlain and Rothschild (1983) who show that the principal components method is equivalent to ML when N goes to infinity.

The principal component method minimizes the function $(NT)^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N}[X_{ti}-F_t\lambda_i]^2$ that, under the normalization F'F/T=I_r, corresponds to maximizing $tr(F'(XX')F)$.

$\hat{F}$ will be equal to $\sqrt{T}$ times the eigenvectors corresponding to the first r eigenvalues of XX', while $\hat{\Lambda}=\left(\hat{F}'\hat{F}\right)^{-1}\hat{F}'X=\hat{F}'X/T$.

In his paper Bai provides asymptotic distributions for $\hat{F}$ under assumptions which allow for a VAR structure of F, factor loadings independently distributed with respect to factors and idiosyncratic shocks, limited time series and cross section dependence in the idiosyncratic components, weak dependence between factors and idiosyncratic shocks, plus other additional technical assumptions.

He shows that under these assumptions, as N,T$\rightarrow\infty$ if $\sqrt{N}/T\rightarrow 0$, for each t up to an invertible matrix H:

$$\begin{aligned}\sqrt{N}\left(\hat{F}_t-H'F_t\right) &= V_{NT}^{-1}\left(\frac{\hat{F}'F}{T}\right)\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\lambda_i e_{ti}+o_p(1)\underset{d}{\rightarrow} \\ & N(0,V^{-1}Q\Gamma_t Q'V^{-1}) \end{aligned} \tag{1.6}$$

Where V_{NT} is a diagonal matrix with the first r eigenvalues of XX'/NT in decreasing order,

V are the first r eigenvalues of $\Sigma_\Lambda^{1/2}\Sigma_F\Sigma_\Lambda^{1/2}$ and $Q=V^{1/2}\Upsilon'\Sigma_\Lambda^{-1/2}$, where Υ is the matrix of the r eigenvectors associated to V, with $\Upsilon'\Upsilon=I$.

Moreover, $\Gamma_t=\lim_{N\rightarrow\infty}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda_j'E\left(e_{ti}e_{tj}\right)$, $\Sigma_F=p\lim\frac{1}{T}\sum_{i=1}^{T}F_iF_i$, and $||\Lambda'\Lambda/N-\Sigma_\Lambda||\rightarrow 0$

He suggests a consistent estimator for this covariance matrix under the assumption of cross-section independence for the idiosyncratic shocks:

$$\hat{\Pi}_t = V_{NT}^{-1}\left(\frac{\hat{F}'\hat{F}}{T}\right)\left(\frac{1}{N}\sum_{i=1}^{N}\hat{e}_{ti}^2\hat{\lambda}_i\hat{\lambda}_i'\right)\left(\frac{\hat{F}'\hat{F}}{T}\right)V_{NT}^{-1} \tag{1.7}$$

He derives under the same assumptions the asymptotic distribution of the loadings (if $\sqrt{T}/N \to 0$) :

$$\begin{aligned}\sqrt{T}\left(\hat{\lambda}_i - H^{-1}\lambda_i\right) &= V_{NT}^{-1}\left(\frac{\hat{F}'F}{T}\right)\left(\frac{\Lambda'\Lambda}{N}\right)\frac{1}{\sqrt{T}}\sum_{i=1}^{N}F_t e_{ti} + o_p(1) \underset{d}{\to} \\ & N(0,(Q')^{-1}\Phi_i Q^{-1})\end{aligned} \tag{1.8}$$

Where $\Phi_i = p\lim_{T\to\infty}\frac{1}{T}\sum_{s=1}^{T}\sum_{t=1}^{T}E\left[F_t F_t' e_{si}e_{ti}\right].$

He suggests a Newey and West (1987) HAC estimator, constructed with the estimated series:

$$\hat{\Theta}_i = D_{0,i} + \sum_{v=1}^{q}\left(1 - \frac{v}{q+1}\right)\left(D_{vi} + D_{vi}'\right) \tag{1.9}$$

$$D_{vi} = \frac{1}{T}\sum_{t=v+1}^{q}\hat{F}_t\hat{e}_{ti}\hat{e}_{t-v,i}\hat{F}_{t-v} \tag{1.10}$$

In the same paper, under the assumption of no cross section correlation between idiosyncratic shocks the asymptotic distribution (with convergence rate $\delta_{NT} = \min(\sqrt{T},\sqrt{N})$ of the common component $C_{ti} = F_t\lambda_i$ is:

$$\frac{\delta_{NT}\left(\hat{C}_{ti} - C_{ti}\right)}{\sqrt{\frac{\delta_{NT}^2}{N}V_{ti} + \frac{\delta_{NT}^2}{T}W_{ti}}} \underset{d}{\to} N(0,1) \tag{1.11}$$

$$V_{ti} = \lambda_i'\Sigma_\Lambda^{-1}\Gamma_t\Sigma_\Lambda^{-1}\lambda_i \tag{1.12}$$

$$W_{ti} = F_t'\Sigma_F^{-1}\Phi_i\Sigma_F^{-1}F_t \tag{1.13}$$

These last two quantities can be consistently estimated as:

$$\hat{V}_{ti} = \hat{\lambda}_i' \left(\frac{\hat{\Lambda}'\hat{\Lambda}}{N}\right)^{-1} \left(\frac{1}{N}\sum_{i=1}^{N} \hat{e}_{ti}^2 \hat{\lambda}_i \hat{\lambda}_i'\right) \left(\frac{\hat{\Lambda}'\hat{\Lambda}}{N}\right)^{-1} \hat{\lambda}_i \tag{1.14}$$

$$\hat{W}_{ti} = \hat{F}_t' \hat{\Theta}_i \hat{F}_t \tag{1.15}$$

In his paper he showed that these asymptotic estimators work well in a Montecarlo experiment with q=1 and a White Noise factor.

In this literature one empirical issue is the choice of the number of factors. In another paper (Bai-Ng (2002)) propose information criteria for the choice of r and they show with Montecarlo simulations that these criteria allow us to select an appropriate number of factors when the number of observations is higher than 40. Stock & Watson (2005) show with another Montecarlo simulation which mimics their sample's structure that the most reliable is the criterion $IC_{p2}(k)$:

$$IC_{p2}(k) = \ln\left(V\left(k, \hat{F}^k\right)\right) + k\left(\frac{n+T}{nT}\right)\ln\left(\min\left(n, T\right)\right) \tag{1.16}$$

where $V\left(k, \hat{F}^k\right) = \min_{\Lambda} \frac{1}{NT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(\bar{X}_{ti} - \Lambda_i^k \hat{F}_t^k\right)^2$ is the sum of squared residuals from the time series regressions of $\bar{X}_{ti}$on $\hat{F}_t^k$.

Other criteria are proposed in Hallin & Liska (2007) and Amengual & Watson (2007).

In this paper, for the sake of simplicity, we assume to know the true number of factors, but studying the effects of its information-based determination on the small-sample properties of factor estimators is an interesting topic for future research..

1.3 The Bootstrap

Bootstrap methods were introduced in Efron (1979) as a way to obtain standard errors for complicated statistics without relying on asymptotic theory.

If we have a sample X which is randomly drawn from an unknown probability distribution F and we want to create a confidence interval for the parameter of interest θ we need an unbiased estimator $\hat{\theta} = t(X)$ and its standard error $s.e._F\left(\hat{\theta}\right) = \sqrt{\int \left(t(X) - E_F\left(t\left(X\right)\right)\right)^2 dF}$.

While for the mean of an iid sample it is easy to derive the standard error, this is much harder for more complicated parameters, especially in the presence of correlated random variables.

The bootstrap substitutes the unknown F with the empirical distribution F^*, the one which assigns probability 1/N to the N observations of the available sample. In practice B simulated samples are drawn from the empirical distribution, and the estimated variance is the variance of the bootstrap estimates across the simulated samples.

In simple linear models with exogenous regressors artificial samples are created by drawing without replacement from the estimated residuals, but this procedure depends on the assumption that the residuals are iid; if the residuals are autocorrelated this Bootstrap procedure does not provide consistent estimates.

In the literature there are basically four bootstrap methods for autoregressive models.

Parametric (Montecarlo) methods rely on the estimation of the residual's ARMA structure and on assumptions about their distribution. As an example in Forni, Giannone, Lippi and Reichlin, (2005) they estimate the model:

$$X_{it} = B_i\left(L\right)\eta_t + u_{it} = \chi_{it} + u_{it} \tag{1.17}$$

where η_t are the common shocks with unit variance and mutually uncorrelated and u_{it} are idiosyncratic shocks with an AR structure:

$$a_i(L)u_{it} = \sigma_i v_{it} \tag{1.18}$$

The common shocks and the iid residuals v_{it} are drawn from a N(0,1) distribution in order to obtain the simulated series.

Semiparametric methods (sieve bootstrap) are similarly based on the estimation of the autoregressive structure, but the samples are drawn from the residuals' empirical distribution

Nonparametric methods do not assume a parametric AR structure, but recognize the presence of autocorrelated residuals by resampling blocks of residuals (the block bootstrap). An important issue with block bootstrap is the estimation of the optimal block length (or the optimal average length), a choice that depends on the implicit loss function. There are different types of block bootstrap, as an example Politis-White (2004) report two different types: the circular bootstrap (CB) and the stationary bootstrap (SB)

The Circular Bootstrap assumes that the block length is the same for each block, while the Stationary Bootstrap assumes that the length changes randomly at each draw; in their paper they provide an automatic procedure which selects the mean block length which minimizes the mean squared error of the asymptotic variance of the autocorrelated series. This optimal length depends on the persistence of the series: if we have high persistence, the effect of a shock lasts longer and we need to draw a longer block in order to preserve the series' autocorrelation.

Another strategy is to work in the frequency domain, where it is possible to find iid variable which can be exploited in bootstrapping. As an example, Ramos (1984) proposed to estimate the spectral density function of a covariance stationary series, generate pseudo- Fourier coefficients and pseudo-data from these pseudo-Fourier coefficients. This is possible because for each frequency the ratio between the fourier coefficients and the square root of the spectral density function is asymptotically distributed as a N(0,1/2).

There is a trade-off between non-parametric and parametric methods: the former are more likely to encompass the true model, while the latter are more precise. As an example, Berkowitz and Kilian compare the confidence intervals of the T-Bill rate impulse response to a standard deviation shock. They show with a Montecarlo simulation where the DGP is a ARMA(2,4), that the parametric bootstrap has an higher coverage probability if confronted to the block bootstrap and Ramos' bootstrap. We think that in our model robustness is a more important issue than precision.

One very important issue which will especially encounter in the second paper is estimation bias.

As an example in small samples VAR parameters are systematically biased: Kilian (1998) shows that the small sample distribution of the impulse responses are not Gaussian, and that simple naive quantile confidence intervals have a lower coverage than expected.

In other words, if we simply look at the $\alpha/2$ and 1-$\alpha/2$ quantiles of the B simulated $\theta^B_{i,j}$ (the vector of VAR parameters):

$$Prob\left(\theta_i \in \left[\theta^B_{i,B\alpha/2}, \theta^B_{i,B(1-\alpha/2)}\right]\right) < 1-\alpha \qquad (1.19)$$

For this reason Kilian (1998) proposed a simulation-based correction, where in the case of a stationary VAR the bootstrap replications are drawn after correcting for the bias $\hat{Bias} = \sum_{j=1}^{B} \theta_{i,j}^{B}/B - \hat{\theta}_i$.

But correcting for small sample estimation bias is not enough for testing and reporting confidence intervals with the appropriate coverage.

In order to have exact tests with the help of simulations, the statistic of interest should be pivotal: it should not depend on unknown parameters.

Tests based on bootstrap replications are very often asymptotically pivotal: the larger is the sample size, the lower is the dependence of the coverage probability from unknown parameters.

In MacKinnon (2002) there are various examples where he shows a few asymptotically pivotal bootstrap tests. He shows confidence interval coverage probability as a function of the true parameter for a given number of observations t and that when t increases it becomes flat at a faster rate than tests based on asymptotic theory. This is not always the case, since each specific model require an extensive montecarlo evaluation in order to understand if a reliable bootstrap procedure for providing accurate confidence interval exists.

Given the number of parameters it is a quite impossible task to provide an extensive evaluation of asymptotic pivotalness when we deal with the DFM. We will only see an example in the second paper, when we will look at the AR(1) parameter in a very simple DFM.

There is not just one way to build bootstrap-based confidence intervals. The simple percentile confidence interval now coexist with more complicated intervals which have been developed when it became clear that estimation bias and asymptotic pivotalness were two important issues . Benkowitz-Lutkepohl-Wolters (2001) propose the following methods:

1. The percentile confidence interval: $\text{CI}(\alpha) = \left[\theta_{\alpha/2}^{Boot}, \;\; \theta_{(1-\alpha/2)}^{Boot}\right]$.
2. The Hall(1992) percentile confidence interval:

   $$\text{CI}_H(\alpha) = \left[2\hat{\theta} - \theta_{(1-\alpha/2)}^{Boot}, \;\; 2\hat{\theta} - \theta_{\alpha/2}^{Boot}\right]$$
3. The studentized percentile confidence interval:

   $$\text{CI}_{SH}(\alpha) = \left[\hat{\theta} - t_{(1-\alpha/2)}^{Boot2}\sqrt{Var(\hat{\theta})}, \;\; \hat{\theta} - t_{\alpha/2}^{Boot2}\sqrt{Var(\hat{\theta})}\right]$$

The percentile confidence interval is the most naive: it is obtained directly by sorting the parameters estimated from each artificial sample. These intervals have not the desired coverage probability when the estimator is biased, since this confidence interval derives from this assumption. For this reason it does not work when we deal with biased AR model parameter estimators.

The Hall percentile confidence interval takes account of this bias, since it is based on the assumption that the difference between the estimate $\hat{\theta}$ and the average estimate across simulated samples $\bar{\theta}^{Boot}$ is equivalent to the estimator bias.

In order to understand the logic behind this confidence interval, let start by assuming that we have a biased estimator such that

$$\hat{\theta} \sim N(\theta + Bias, \sigma^2_{\hat{\theta}}) \tag{1.20}$$

Given the bootstrap principle we can also assume that the estimator obtained from the bootstrap sample is biased and normally distributed:

$$\hat{\theta}^B \sim N(\hat{\theta} + Bias, \sigma^2_{\hat{\theta}^B}) \tag{1.21}$$

From the empirical distribution we can estimate the bias as the difference between the average bootstrap estimate and the original estimate:

$$\hat{Bias} = \bar{\theta}^{Boot} - \hat{\theta} = \frac{\sum_{i=1}^{B} \theta_i^{Boot}}{B} - \hat{\theta} \tag{1.22}$$

Under these assumptions:

$$\Pr\left(z_{\alpha/2} \leq \frac{\hat{\theta} - \theta - \hat{Bias}}{\sigma_{\hat{\theta}}} \leq z_{1-\alpha/2} \right) = 1 - \alpha \tag{1.23}$$

And the bias-corrected confidence interval becomes:

$$\left[\hat{\theta} - \hat{Bias} - z_{1-\alpha/2}\sigma_{\hat{\theta}} \quad \hat{\theta} - \hat{Bias} - z_{\alpha/2}\sigma_{\hat{\theta}} \right] \tag{1.24}$$

Under this normal approximation z_α is the 100α-th percentile of a gaussian distribution while the parameter variance can be estimated as the standard deviation of the bootstrap parameters:

$$\hat{\sigma}_{\hat{\theta}} = \frac{1}{B-1}\sum_{i=1}^{B}\left(\theta_i^{Boot} - \bar{\theta}^{Boot}\right)^2 \tag{1.25}$$

If we relax the normality assumption we can approximate the z distribution by studentizing the empirical distribution. In practice it means that: z_α is the $100\alpha th$ percentile of

$$t_i^{Boot} = \frac{(\theta_i^{Boot} - \bar{\theta}^{Boot})}{\sigma_{\hat{\theta}^B}}. \tag{1.26}$$

Under the assumption that the bootstrap estimator standard deviation $\sigma_{\hat{\theta}^B}$ and estimator standard deviation $\sigma_{\hat{\theta}}$ are equal we can derive the Hall Percentile interval as:

$$\begin{aligned} \hat{\theta} - \hat{Bias} - z_{1-\alpha/2}\sigma_{\hat{\theta}} &= \hat{\theta} - \left(\bar{\theta}^{Boot} - \hat{\theta}\right) - \hat{\sigma}_{\hat{\theta}}\frac{(\theta_{(1-\alpha/2)}^{Boot} - \bar{\theta}^{Boot})}{\sigma_{\hat{\theta}^B}} = \\ &= 2\hat{\theta} - \bar{\theta}^{Boot} - \theta_{(1-\alpha/2)}^{Boot} + \bar{\theta}^{Boot} = 2\hat{\theta} - \theta_{(1-\alpha/2)}^{Boot} \end{aligned}$$

This method and the studentized confidence interval derives by the observation in the bootstrap literature that by studentizing it is possible to obtain a better approximation of the bootstrap to the real distribution.

The studentized percentile confidence interval requires much more computation because it is a double bootstrap. B_2 artificial samples are simulated starting from the bootstrapped parameters θ_i^{Boot}, in order to estimate $Var(\theta_i^{Boot})$.

The confidence interval has the following structure:

$$CI_{SH} = \left[\hat{\theta} - t_{(1-\alpha/2)}^{Boot2}\sqrt{Var(\hat{\theta})},\quad \hat{\theta} - t_{\alpha/2}^{Boot2}\sqrt{Var(\hat{\theta})}\right] \tag{1.27}$$

Where $Var(\hat{\theta})$ is the estimated variance of the bootstrapped parameters θ_i^{Boot}.

$$Var(\hat{\theta}) = \frac{1}{B-1}\sum_{i=1}^{B}\left(\theta_i^{Boot} - \bar{\theta}^{Boot}\right)^2 \tag{1.28}$$

The confidence interval's critical values come from the studentized bootstrap parameters; they require a second draw of B_2 artificial samples, in order to estimate $Var(\theta_i^{Boot})$.

$$t_i^{Boot2} = \left(\theta_i^{Boot} - \hat{\theta}\right) / \sqrt{Var(\theta_i^{Boot})} \tag{1.29}$$

$$Var(\theta_i^{Boot}) = \frac{1}{B_2 - 1} \sum_{i=1}^{B} \left(\theta_i^{Boot2} - \bar{\theta}_i^{Boot2}\right)^2 \tag{1.30}$$

In the bootstrap literature there is also another percentile confidence interval: the Bias Corrected percentile confidence interval (BC, see DeCiccio Efron (1996)).

Let us assume that it exists a monotone increasing transformation of our parameter of interest $\phi = m(\theta)$ such that:

$$\hat{\phi} = m(\hat{\theta}) \sim N(\phi - z_0, 1) \tag{1.31}$$

Under this assumption (which is weaker than in the previous case where $\theta = m(\theta)$) there is another monotone transformation $\xi = M(\theta)$ such that $\hat{\xi} = M(\hat{\theta}) = \xi + W$, with W having the same distribution for any ξ and the αth percentile of $\hat{\xi}$ is:

$$\hat{\xi}(\alpha) = \xi - W(1 - \alpha) \tag{1.32}$$

In practice if we define $\Phi(.)$ as the cumulative distribution function of a N(0,1) variable (where $z^{(\alpha)} = \Phi^{-1}(\alpha)$, so that $z^{(0.025)} = 1.96$) and $G(c)$ is the cumulative distribution empirical function, we have that:

$$G(c) = \frac{I(\theta_i^{Boot} < c)}{B} \tag{1.33}$$

$$\hat{z}_0 = \Phi^{-1}\left(G(\hat{\theta})\right) \tag{1.34}$$

$$\theta_\alpha^{BC} = G^{-1}(\Phi(\ 2\hat{z}_0 + z^{(\alpha)})) \tag{1.35}$$

And the bias corrected percentile interval with significance level α is:

$$BC(\alpha) = \left[\theta_{\alpha/2}^{BC}, \quad \theta_{(1-\alpha/2)}^{BC}\right]$$

This last interval is equivalent to the simple percentile one if there is no bias, since if $\hat{z}_0 = 0$, $\Phi(\; z^{(\alpha/2)}) = \alpha/2$ and $G^{-1}(\alpha/2) = \theta_{\alpha/2}^{Boot}$.

Concluding this section we have seen five different ways to construct confidence intervals: Efron simple percentile intervals, Hall percentile, Bias corrected percentile, normal approximation (bias corrected) and studentized percentile. Different models may require or not bias corrected intervals, computation-intensive studentized percentile intervals or simple normal approximations.

In the next sections we will evaluate these different methods by simulating a dynamic factor model and looking at the common component and at forecast intervals.

1.4 Our Montecarlo simulations

In this section we want to examine the small sample properties of confidence intervals based on asymptotic theory in a model more complex than Bai's (2003) and see if we can obtain better results with the bootstrap. While in his model there is only an iid dynamic factor, in our model there are q=2 dynamic factors, modelled as independent AR(2) processes (h=2); these factors are loaded as an MA(1) (p=1). This specification is closer to what we observe empirically; as an example Stock Watson (2005) model a panel of 132 time series of the US economy as a DFM with 9 static factors and 7 common shocks.

Hence, our DGP is the following :

$$X_{t,i} = \left(\begin{bmatrix} \lambda_{i1,0} & \lambda_{i2,0} & \lambda_{i1,1} & \lambda_{i2,1} \end{bmatrix} \begin{bmatrix} f_{1,t} \\ f_{2,t} \\ f_{1,t-1} \\ f_{2,t-1} \end{bmatrix} \right)' + \sigma_i e_{t,i} \tag{1.36}$$

$$\begin{bmatrix} f_{1,t} \\ f_{2,t} \end{bmatrix} = \begin{bmatrix} \Gamma_{11} & 0 \\ 0 & \Gamma_{21} \end{bmatrix} \begin{bmatrix} f_{1,t-1} \\ f_{2,t-1} \end{bmatrix} + \tag{1.37}$$

$$\begin{bmatrix} \Gamma_{12} & 0 \\ 0 & \Gamma_{22} \end{bmatrix} \begin{bmatrix} f_{1,t-2} \\ f_{2,t-2} \end{bmatrix} + \begin{bmatrix} \sigma_{u1} & 0 \\ 0 & \sigma_{u2} \end{bmatrix} \begin{bmatrix} u_{t,1} \\ u_{t,2} \end{bmatrix} \tag{1.38}$$

In this model we assume that $e_{t,i}$, $u_{t,1}$ and $u_{t,2}$ are iid random variables with unit variance. If we define the static factor F_t as the r=q(p+1)=4 vector:

$$F_t = \begin{bmatrix} f_{1,t} \\ f_{2,t} \\ f_{1,t-1} \\ f_{2,t-1} \end{bmatrix}' \tag{1.39}$$

we can rewrite the model in static form as:

$$X_{ti} = F_t \lambda_i + \sigma_i e_{ti} \tag{1.40}$$

$$F_t = \begin{bmatrix} f_{1,t} \\ f_{2,t} \\ f_{1,t-1} \\ f_{2,t-1} \end{bmatrix}' = \left(\begin{bmatrix} \Gamma_{11} & 0 & \Gamma_{12} & 0 \\ 0 & \Gamma_{21} & 0 & \Gamma_{22} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} f_{1,t-1} \\ f_{2,t-1} \\ f_{1,t-2} \\ f_{2,t-2} \end{bmatrix} \right)' +$$

$$\left(\begin{bmatrix} \sigma_{u1} & 0 \\ 0 & \sigma_{u2} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} \right)' = F_{t-1}\Gamma + u_t \Omega_u \tag{1.41}$$

As an example let's assume that our n series are driven by two common shocks, and that the q=2 orthogonal factors have the following AR structure, where the second shock is more persistent:

$$(1-0.1L)(1-0.7L)f_{1t} = u_{t,1} \tag{1.42}$$

$$(1-0.3L)(1-0.9L)f_{2t} = u_{t,2} \tag{1.43}$$

We assume that each series is also a function of lagged factors (p=1). Factor loadings assume value 0 with probability 50% and a value in the intervals [-2,-1] or [1, 2] with probability 25% in both cases.

In this way the simulated dataset is driven by 4 static factors, and when we apply principal components also the third and the fourth factor explain a fair share of the total variance. Otherwise Bai-Ng (2002) criteria would select no more than 2 static factors.

Before dealing with forecasting as a first exercise we show the rejection probability over 1000 simulated artificial samples for the common component.

From the second section we know that asymptotically the common part is distributed as a normal distribution, where the asymptotic covariance matrix takes into account both the uncertainty in estimating the factors and in estimating the factor loadings:

$$\frac{\left(\hat{C}_{ti} - C_{ti}\right)}{\sqrt{\frac{1}{N}\hat{V}_{ti} + \frac{1}{T}\hat{W}_{ti}}} \underset{d}{\rightarrow} N(0,1) \tag{1.44}$$

$$\hat{V}_{ti} = \hat{\lambda}_i' \left(\frac{\hat{\Lambda}'\hat{\Lambda}}{N}\right)^{-1} \left(\frac{1}{N}\sum_{i=1}^{N} \hat{e}_{ti}^2 \hat{\lambda}_i \hat{\lambda}_i'\right) \left(\frac{\hat{\Lambda}'\hat{\Lambda}}{N}\right)^{-1} \hat{\lambda}_i \tag{1.45}$$

$$\hat{W}_{ti} = \hat{F}_t' \hat{\Theta}_i \hat{F}_t \tag{1.46}$$

Given the definition of the rejection probability as the probability that the true parameter is not inside the confidence interval at α% confidence level, we calculated it as the average over n series, t observations and m=1000 artificial samples of an indicator function which assumes value 1 when the studentized common part value is larger in absolute value of the 100(1-$\alpha/2$)th percentile of the normal distribution:

$$RP(t,n) = \frac{\left(\sum_{k=1}^{m} \sum_{i=1}^{t} \sum_{j=1}^{n} I \left(\left| \frac{(\hat{C}_{ij}^k - C_{ij}^k)}{\sqrt{\frac{1}{n}\hat{V}_{ij}^k + \frac{1}{t}\hat{W}_{ij}^k}} \right| > z_{(1-a/2)} \right) \right)}{tnm} \tag{1.47}$$

Table(1) shows our results when the nominal confidence level is α =5% ($z_{(1-a/2)} = 1.96$); the rejection probability for the confidence intervals based on asymptotic theory is close to the nominal one only when we have 100 series and 200 observations. However, it is very high (around 20%) when we deal with samples which are small either in the N or in the T dimensions.

Rejection Probabilities given asymptotic confidence intervals				
Common Part	n=15	n=25	n=50	n=100
t=25	0.24261	0.22161	0.24968	0.28158
t=50	0.23228	0.20729	0.24414	0.29693
t=100	0.20791	0.16677	0.19406	0.24662
t=200	0.18632	0.10382	0.072631	0.060652

Table 1:Common Part asymptotic confidence intervals.
Rejection Probabilities average over 1000 simulations, n series
and t observations at the 5% significance level

This is the main reason why we want to evaluate whether applying bootstrap methods we can obtain an appropriate coverage probability also in small samples.

In particular, we propose a bootstrap method based on the block bootstrap. We choose this method in order to avoid the bias related to the estimation of the static factors' VAR structure and because it is consistent with the model structure.

When we estimate the static factors with principal components we have (r+n) independent autocorrelated stationary series: r estimated factors and n idiosyncratic errors.

$$\underset{TxN}{X} = \underset{Txr}{F} * \underset{rxN}{\Lambda} + \underset{TxN}{e} \tag{1.48}$$

We can resample these series with the stationary bootstrap proposed in Politis-Romano. Their procedure requires two steps:

1. Draw from a geometric distribution the length of the ith block; the average length depends on the time series' autocorrelation.

2. Draw with repetition the ith block with the length extracted in the first step.

These two steps must be iterated untill we obtain a sample with the required length. In this way we obtain B=1000 bootstrap samples and by re-estimating the common part we have an empirical distribution of t*n parameters. Table(2) shows the rejection probability average over parameters and m=1000 artificial samples when confidence intervals are constructed with Efron Percentile, Hall percentile, Normal approximation and bias corrected percentile interval.

Common part: CI Rejection probabilities				
Efron	n=15	n=25	n=50	n=100
t=25	0.050917	0.05344	0.057778	0.060725
t=50	0.049469	0.050751	0.055296	0.058186
t=100	0.047135	0.047843	0.051659	0.054565
t=200	0.045709	0.046891	0.049427	0.052969
Hall	n=15	n=25	n=50	n=100
t=25	0.19613	0.17963	0.16213	0.15145
t=50	0.15188	0.13549	0.11803	0.10753
t=100	0.11603	0.10173	0.082943	0.073741
t=200	0.097058	0.080697	0.059366	0.05368
Normal Approximation	n=15	n=25	n=50	n=100
t=25	0.19622	0.17502	0.16005	0.14965
t=50	0.14867	0.13443	0.11700	0.10842
t=100	0.11517	0.10196	0.08371	0.07413
t=200	0.09889	0.08272	0.06221	0.05545
BC Perc	n=15	n=25	n=50	n=100
t=25	0.20615	0.18663	0.17338	0.16454
t=50	0.15803	0.14385	0.12884	0.11970
t=100	0.12172	0.10793	0.09103	0.08181
t=200	0.10089	0.08469	0.06410	0.05640

Table 2 Rejection Probabilities average for common part over m=1000 simulations and t*n parameters at the 5% significance level Bootstrap samples are created with an algorithm based on the Stationary Block Bootstrap, with B=1000 replications

We can observe in Table(2) that in every bootstrap confidence interval the rejection probability is closer to the nominal one (5%) if confronted with the asymptotic confidence interval.

In particular Efron's simple percentile interval is in practice always equal to 5%. The other three converge rapidly to the nominal level as the number of observations and series increase.

The big difference with the simple percentile interval is related to bias correction: as explained in section 3 Efron's percentile interval is simply obtained by sorting the common part estimated in each artificial sample. The other confidence intervals take account of bias correction, measured as the difference between the average bootstrapped parameter value and the estimated one.

The bootstrap average of the j-th observation common component for the ith series $\bar{C}_{ji}^{BOOT}$ is always close to 0, which is the theoretical expected common component value. In practice bias correction gives too much weight to short term movements of the common component from its expected value when we deal with small samples.

In Figure(1) we show as an example four different confidence intervals and the actual common part of one of the artificial series when we have t=25 observations and n=15 series.

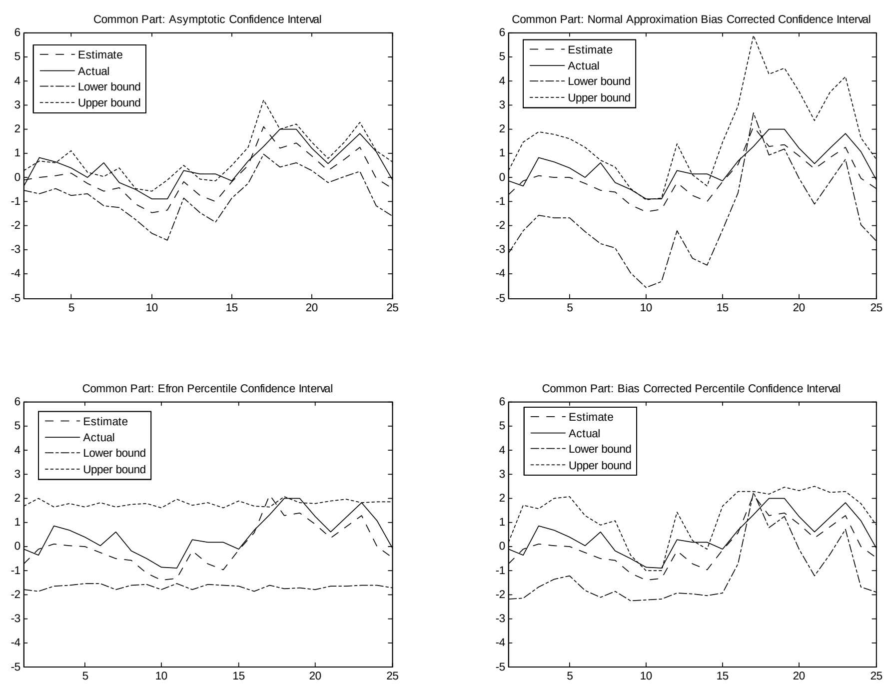


Figure 1: Common Part Asymptotic vs Bootstrap Confidence intervals when t=25 and n=15.

If we look at the length of these different confidence intervals, we observe that confidence intervals based on asymptotic theory are in general smaller than the confidence intervals based on the bootstrap. This finding is confirmed in table 3, which shows the average length across observations, series and artificial samples, calculated as the difference between the upper bound and the lower bound of a 5% level confidence interval.

$$AL(t,n) = \frac{\left(\sum_{k=1}^{m}\sum_{i=1}^{t}\sum_{j=1}^{n}(UB_{ij,k} - LB_{ij,k})\right)}{tnm}$$

Common Part: CI Average Length				
Asymptotic	n=15	n=25	n=50	n=100
t=25	1.3608	1.3302	1.1712	1.0592
t=50	1.1509	1.1172	0.92262	0.78831
t=100	0.98759	0.95376	0.73077	0.58621
t=200	0.86358	0.84108	0.61846	0.46453

Efron	n=15	n=25	n=50	n=100
t=25	3.4292	3.2634	3.1709	3.0871
t=50	3.5236	3.2591	3.1452	3.0975
t=100	3.6066	3.5659	3.3651	3.1688
t=200	3.6272	3.505	3.2875	3.2061

Normal Approximation	n=15	n=25	n=50	n=100
t=25	3.5475	3.3883	3.2815	3.2240
t=50	3.5699	3.4241	3.2978	3.2293
t=100	3.6026	3.4558	3.3161	3.2367
t=200	3.6233	3.4521	3.3010	3.2188

BC Perc	n=15	n=25	n=50	n=100
t=25	2.5662	2.4932	2.4241	2.3888
t=50	2.6009	2.5357	2.4411	2.3889
t=100	2.6428	2.5811	2.4709	2.4087
t=200	2.7002	2.6277	2.5033	2.4223

Table 3: Confidence interval average length for common part over m=1000 simulations and t*n parameters at the 5% significance level given asymptotic and bootstrap methods Bootstrap samples are created with an algorithm based on the Stationary Block Bootstrap, with B=1000 replications

Summarizing these results we have seen that with an appropriate bootstrap algorithm we can obtain confidence intervals which do not underestimate the variance of the common component if our goal is to get a rejection probability close to the nominal one, but this may lead us to report too large confidence intervals.

Bearing in mind what we have seen for the common component, we can now focus on forecasting. We look at the Montecarlo simulations reported in Bai-Ng (2005), where they simulate for each artificial sample, and for h=4 steps ahead, the following series y:

$$y_{t+h} = 1 + f_{1t} + f_{2t} + f_{1t-1} + f_{2t-1} + \varepsilon_{t+h} \tag{1.49}$$

where ε_{t+h} is an iid N(0,1)

If we assume that the factors are observable, we can estimate this equation with OLS and derive the optimal forecast and the asymptotic forecast error variance, as reported below.

$$\begin{aligned} y_{t+i} &= \begin{bmatrix} \hat{c} & \hat{\lambda}_{y1,0} & \hat{\lambda}_{y2,0} & \hat{\lambda}_{y1,1} & \hat{\lambda}_{y2,1} \end{bmatrix} \begin{bmatrix} 1 \\ f_{1,t} \\ f_{2,t} \\ f_{1,t-1} \\ f_{2,t-1} \end{bmatrix} + \hat{\varepsilon}_{y,t+h} = && (1.50) \\ &= \hat{\Lambda}'_y z_t + \hat{\varepsilon}_{y,t+h} = \hat{y}_{t+h|t} + \hat{\varepsilon}_{y,t+h} && (1.51) \\ AsyVar(y_{t+h} - \hat{y}_{t+h|t}) &= \hat{\sigma}^2_{\varepsilon} + z'_t AsyVar(\hat{\Lambda}'_y) z_t / T && (1.52) \end{aligned}$$

In their paper they showed under the regularity conditions that we reported in the second section that if the factors have been estimated with principal components, the asymptotic variance should take account of this unobservable factors uncertainty:

$$\begin{aligned} AsyVar(y_{t+h} - \hat{y}_{t+h|t}) &= \hat{\sigma}^2_{\varepsilon} + \hat{z}'_t AsyVar(\hat{\Lambda}'_y) \hat{z}_t / T + (1/N) * \\ & \begin{bmatrix} \hat{\lambda}_{y1,0} & \hat{\lambda}_{y2,0} & \hat{\lambda}_{y1,1} & \hat{\lambda}_{y2,1} \end{bmatrix} AsyVar\left(\hat{F}_t\right) \begin{bmatrix} \hat{\lambda}_{y1,0} \\ \hat{\lambda}_{y2,0} \\ \hat{\lambda}_{y1,1} \\ \hat{\lambda}_{y2,1} \end{bmatrix} && (1.53) \end{aligned}$$

With these results in mind we can construct a 95% confidence interval for y_{t+h} :

$$\left[\hat{y}_{t+h|t} \mp 1.96\sqrt{AsyVar(y_{t+h} - \hat{y}_{t+h|t})}\right] \tag{1.54}$$

If we want to show a confidence interval based on bootstrap methods we must simulate the forecast error distribution. If the factors are known this confidence interval can be estimated in three steps:

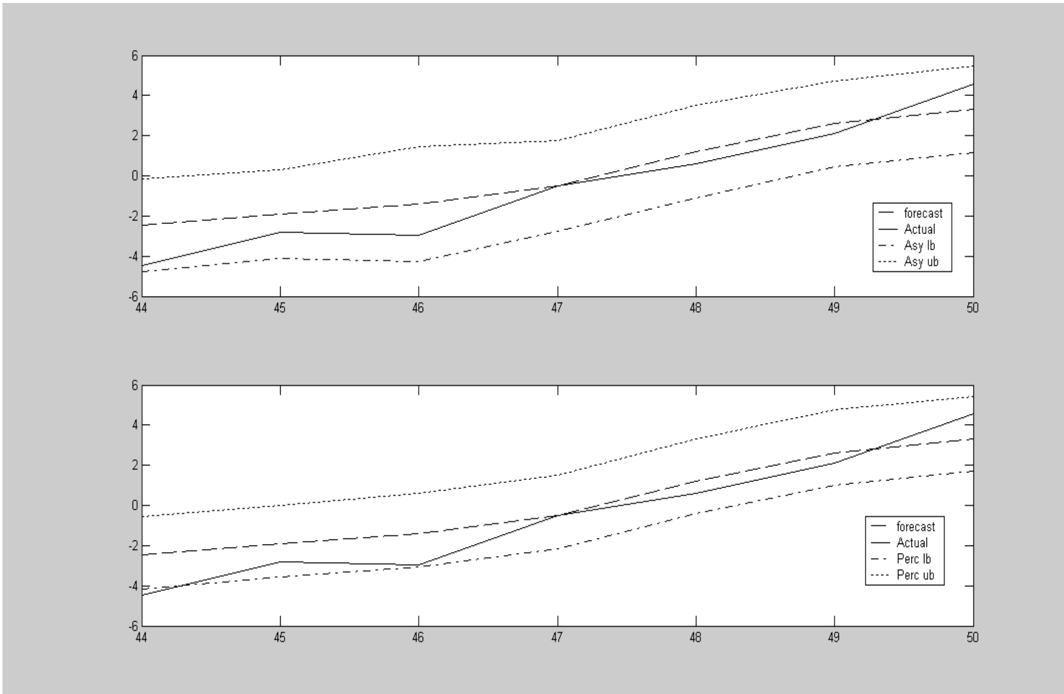
1. Estimate $\hat{\varepsilon}_{y,t+i}$ with the first t observations, generate $\mathrm{B}_1 = 100$ artificial samples $\varepsilon^j_{y,t+i}$, with the block bootstrap and simulate B_1 series y^j_{t+i}.
2. For each artificial sample y^j_{t+i} estimate the forecasting errors $\hat{\varepsilon}^j_{y,t+i}$, and from its empirical distribution resample $\mathrm{B}_2 = 100$ forecasting errors $\varepsilon^{j,k}_{y,t+i}$.
3. Sort $\varepsilon^{j,k}_{y,t+i}$,save the $\alpha/2$ and $(1-\alpha/2)$ percentiles and create the percentile confidence interval for y_{t+i} :

$$\left[\hat{y}_{t+i|t} + \varepsilon^{j,k}_{y,t+i}(\alpha/2),\ \hat{y}_{t+i|t} + \varepsilon^{j,k}_{y,t+i}(1-\alpha/2)\right]$$

If the factors are unknown this algorithm requires also in step 1 to draw B_1 artificial samples from the empirical distribution of the r estimated static factors $\hat{F}$ and the n idiosyncratic shocks $\hat{e}$ via the block bootstrap. In this way we can simulate B_1 artificial samples $X^{(j)}$, and reestimate the factors $\hat{F}^{(j)}$ which we will use in the second step for obtaining $\hat{\varepsilon}^j_{y,t+h} = y^j_{t+h} - \hat{\Lambda}'_y \begin{bmatrix} 1 \\ \hat{F}^{(j)}_t \end{bmatrix}$.

As before we evaluated these confidence intervals considering samples with t observations and n series. In our experiment the forecast sample corresponds to the last (0.2t-h) observations. As an example if we have t=50 observations and we forecast h=4 steps ahead, we use the first 0.8t=40 observations of X for estimating the factors and observations j=(5,6,..44) of the forecasted variable y in order to estimate the direct forecast equation. This means that we evaluate the forecast intervals only for the last 7 observations.

Figure(2) shows this example connected with an artificial series when we have only 50 observations and 50 series in X. In practice there is not a big difference between the two confidence intervals.



Figure(2): An example of Asymptotic and Percentile Forecast intervals when t=50, n=50

Table 4 shows the rejection probability average over m=1000 simulation and (0.2t-4) forecast sample observations of asymptotic theory vs bootstrap percentile confidence intervals. Surprisingly we see that when confidence intervals are based on asymptotic theory the rejection probability is closer to the nominal one than what we obtained with our bootstrap algorithm. This is associated, as it can be seen in Table 5 with slightly larger confidence intervals. One possible reason is that the bootstrap algorithm that we proposed has not captured all the uncertainty. If for example we resample idiosyncratic shocks conditionally on a first draw of static factors, we obtain percentile bootstrap confidence intervals with rejection probability close to 0 but their average length doubles.

Rejection Probabilities for Forecast CI

Asymptotic Theory

	n=15	n=25	n=50	n=100
t=25	0.1375	0.1425	0.1525	0.15875
t=50	0.069643	0.072143	0.0775	0.085
t=100	0.045882	0.046618	0.053382	0.055294
t=200	0.038514	0.03973	0.042838	0.046892

Bootstrap Percentile

	n=15	n=25	n=50	n=100
t=25	0.3375	0.32625	0.3375	0.35
t=50	0.14357	0.14929	0.14643	0.14071
t=100	0.083382	0.084265	0.081618	0.082794
t=200	0.061216	0.061419	0.060203	0.063378

Table 4: Rejection Probabilities average over 1000 simulations and forecast sample, with h=4

Forecast CI average length

Asymptotic Theory

	n=15	n=25	n=50	n=100
t=25	4.6751	4.4094	4.2854	4.138
t=50	4.6179	4.4135	4.2254	4.0954
t=100	4.6876	4.4553	4.2592	4.1227
t=200	4.6843	4.4422	4.2357	4.1067

Bootstrap Percentile

	n=15	n=25	n=50	n=100
t=25	2.9563	2.8346	2.7666	2.7083
t=50	3.7599	3.6176	3.5323	3.4779
t=100	4.1254	3.9524	3.8506	3.7778
t=200	4.2478	4.0767	3.9684	3.8953

Table 5: CI average length over 1000 simulations and forecast sample, with h=4

Summarizing the results of this Montecarlo simulation we can observe that when we deal with parameters which are a simple linear combination of the DGP parameters asymptotic theory confidence intervals work well when we deal with samples which have a reasonable length. This is not true when our goal is to provide confidence intervals for the common part in small samples: in this case bootstrap methods perform far better, but we should bear in mind that any particular statistical model requires a different bootstrap algorithm. When we deal with the common part bias correcting is counterproductive. With forecasting there is probably a problem related with the algorithm that we used, because it fails in providing an appropriate level of uncertainty.

We should take account of two important issues: the desirability of a correct but large confidence interval and the trade off between a correct confidence interval and the computational time needed. Both arguments are in favour of asymptotic theory confidence intervals.

1.5 Empirical Application: Bai Ng 2005

In this part we present two empirical examples: we will forecast the growth rate of industrial production and of inflation, as in Bai-Ng (2005) and Stock & Watson (2002a).

In this way we will show the equivalence of the two method when we deal with large samples.

We use a balanced panel with 160 series out of 215 series in the original paper in order to have a balanced sample with 481 observations in the period 1959:2-1999:2.

We consider the one-year ahead growth rate which is defined as:

$$y_{t+12}^{12} = 100(\ln(IP_{t+12}) - \ln(IP_t)) \tag{1.55}$$

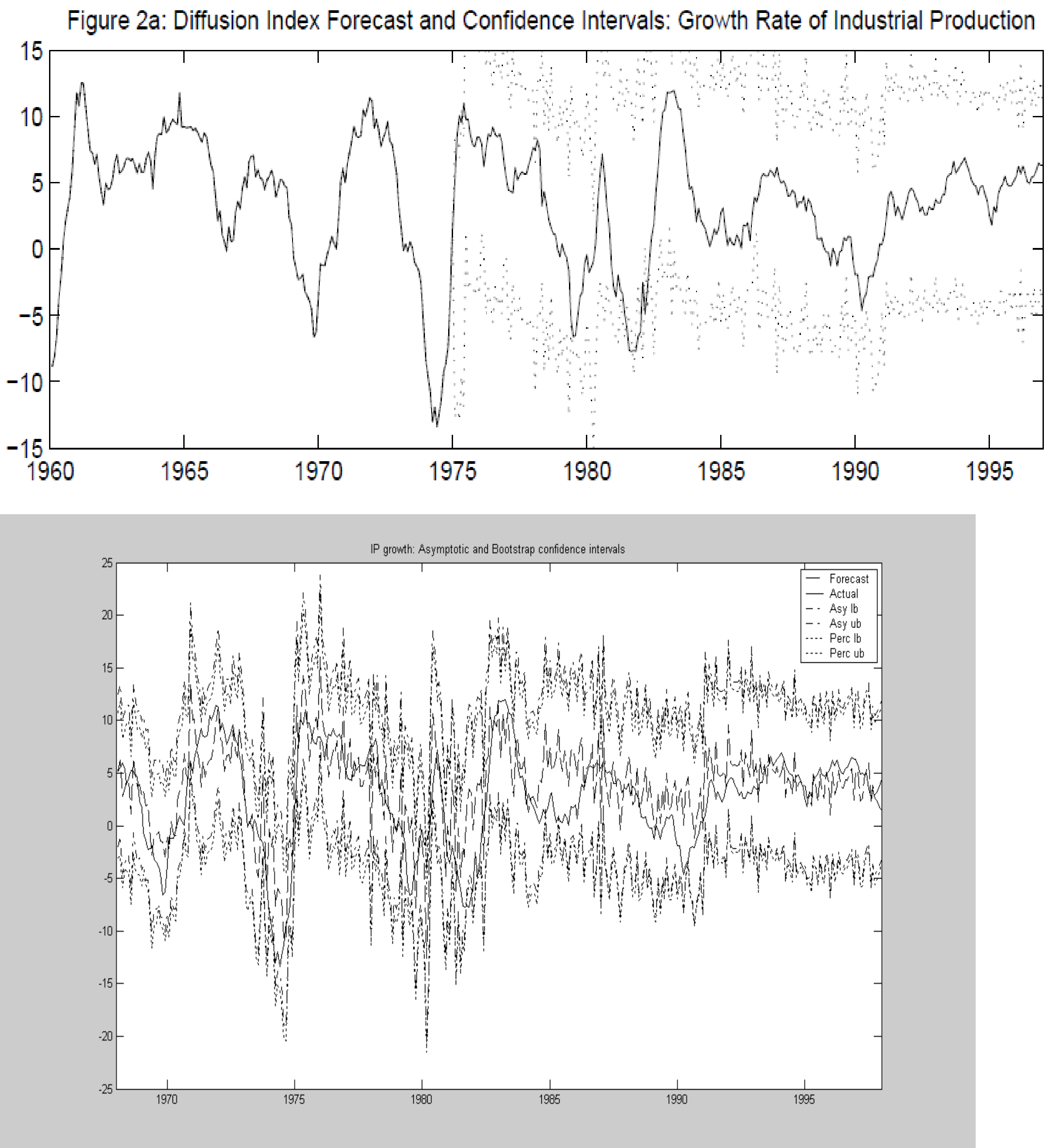
And given the definition of the annualized industrial production growth rate as:

$$y_t = 1200(\ln(IP_t) - \ln(IP_{t-1})) \tag{1.56}$$

We have the following forecast based on the Diffusion Index Autoregressive (DI-AR) model:

$$\hat{y}_{t+12|t}^{12} = \hat{c}_{12} + \hat{\Lambda}_{12}\hat{F}_t + \sum_{j=1}^{p} \hat{\gamma}_{12,j} y_{t-j+1} \tag{1.57}$$

For what concerns our example if we consider the BIC criterion we have p=1, and two specifications: the AR specification without factors, and the DI-AR specification which includes nine factors estimated recursively.



Figure(3): Confront with the asymptotic CI in the original paper and between the two methods

Figure(3) shows that, as expected given our simulation results with large N and large T, asymptotic and bootstrap confidence intervals are very similar. We report also figure (2a) in

Bai-Ng (2005) paper for a comparison.

As a second empirical example we consider one-year ahead inflation forecasting, which is defined as $\pi_t^{12} = 100 \ln(CPI_{t+12}/CPI_t)$. Since the Consumer Price Index turns out to be I(2), we consider the model .

$$\hat{y}_{t+12|t}^{12} = \hat{c}_{12} + \hat{\Lambda}_{12}\hat{F}_t + \sum_{j=1}^{p} \hat{\gamma}_{12,j} y_{t-j+1} \tag{1.58}$$

$$y_{t+12}^{12} = \pi_t^{12} - \pi_t^1 = 100 \ln(CPI_{t+12}/CPI_t) - 1200 \ln(CPI_t/CPI_{t-1}) \tag{1.59}$$

$$y_t = 1200\Delta \ln(CPI_t/CPI_{t-1}) \tag{1.60}$$

The DI-AR specification has 6 lags and provides the forecasting confidence interval reported in figure (4). Figure(4) shows also what is reported in Bai-Ng (2005).Also in this case both confidence intervals are very similar.

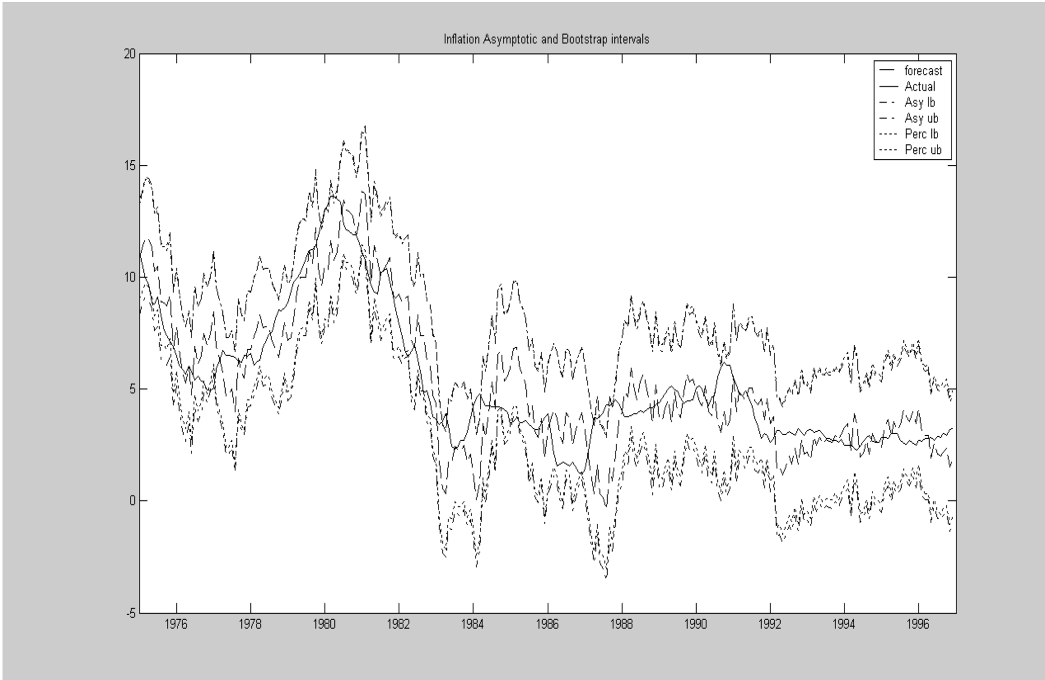
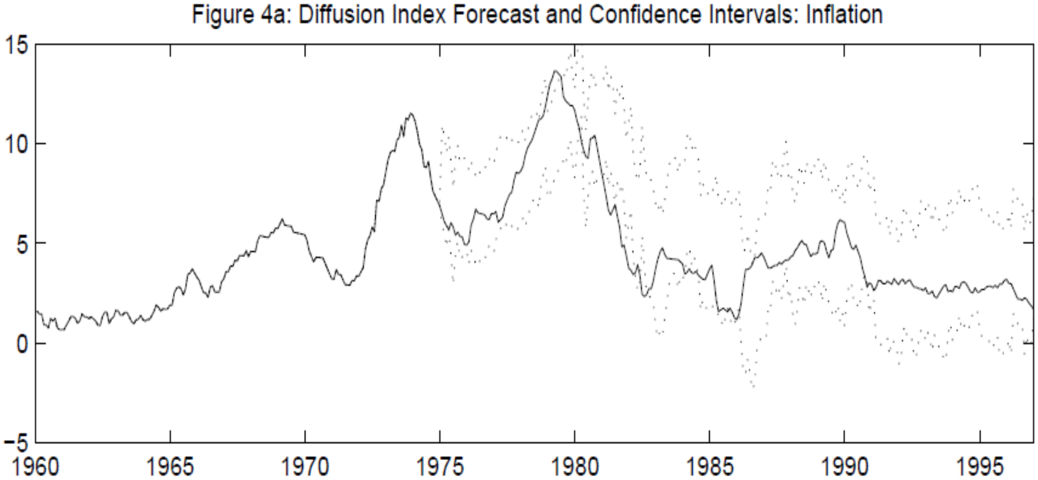


Figure 4a: Diffusion Index Forecast and Confidence Intervals: Inflation



Figure(4)Confront with the asymptotic CI in the original paper and between the two methods

1.6 Conclusion

In this first chapter we have presented the Dynamic Factor Model in static form and the asymptotic theory developed by Bai and Ng. In the third section we have seen that there is not a single way to apply bootstrap methods and that each model and each problem requires a different procedure. This is the main goal of this thesis; in this chapter we looked at the common part and at diffusion index forecasts. We have seen for the common part that a simple percentile interval work well for any (t,n) combination, also if in this case we risk to obtain it at the cost of having a naive interval. Moreover we have seen that bias correction can be counterproductive when the bootstrap mean is already close to the theoretical mean: the common component has mean zero and bias-correcting intervals are relatively large because they give too much weight to short term shocks.

On the other hand our bootstrap algorithm does not work when we deal with forecasting:

any application of the bootstrap requires a careful montecarlo experiment before we can propose a suitable algorithm for dealing with our problem.

Chapter 2

Bootstrap Methods for DFM: confidence intervals for impulse responses

2.1 Introduction

Dynamic Factor Models (DFM) have been proposed as an useful way to summarize the information contained in large datasets with a few latent factors.

Different techniques have been proposed in order to obtain forecasts, identify common factors and the effects of specific shocks in a DFM framework.

Inferential theory based on asymptotic theory for these models is reported in a series of papers published in Econometrica by Jushan Bai and Serena Ng, where for example they show that the FAVAR parameters are $\sqrt{T}$ consistent and asymptotically normal if $\sqrt{T}/N \rightarrow 0$. They also propose in their papers confidence intervals for factors and diffusion index forecasts, but they do not provide confidence intervals for impulse responses function. Only recently Forni and Gambetti (2008) showed confidence intervals for impulse responses which are based on the block bootstrap.

In the first paper we dedicated a section to the asymptotic theory proposed by Jushan Bai and Serena Ng and a section to bootstrap methods. We confronted confidence intervals for diffusion index forecasts and the common component. In this paper we want to show how bootstrap techniques can be applied to the task of providing confidence intervals for nonlinear functions of the DGP parameters: impulse responses.

After this introduction this second paper has two other sections, plus the conclusion.

The second section applies asymptotic theory and bootstrap methods to the task of providing impulse responses confidence intervals for a DFM with a single common shock when there are no identification problems (apart to the sign of the impact effect). The third section presents and explain how with a rotation matrix we can rotate the estimated impulse responses in order to track the actual impulse responses. The conclusion summarizes the results and report preliminary results for the complicated model: the block bootstrap is probably a good choice for our task.

2.2 Confidence intervals for impulse responses when identification is not a problem

In this section we evaluate the performance of asymptotic methods in providing reliable confidence intervals for impulse responses when we have a simple Dynamic Factor Model. Before doing it we will exploit this simple model for showing something that was hard to see in the preceding chapter: how confidence intervals based on asymptotic theory are heavily influenced by key DGP parameters.

The model is an exact dynamic factor model with one single common factor (which is modeled as a AR(1)), up to 200 observations and to 100 series.

Loadings, idiosyncratic and common shocks are extracted from independent N(0,1) distributions, so by construction the common component explains a share of the total variance for each series which is equal to $\lambda_i^2/(1+\lambda_i^2)$.

We can represent this model with two equations:

$$X_{it} = \lambda_i f_t + e_{it} \quad (2.1)$$

$$f_t = \alpha f_{t-1} + u_t \quad (2.2)$$

Where $\alpha = 0.9$, $\sigma_u = \sqrt{(1-\alpha^2)}$ and $\sigma_f = 1$.

In Bai (2003) and in the first paper we have seen that confidence intervals based on asymptotic theory have good coverage probability, if we have a reasonable large number of observations and series, but we have not seen how these results are influenced by important parameters, as it is α in our example. This is a typical situation where we can appreciate the asymptotic pivotalness of confidence intervals based on the bootstrap.

As a first step we revisit the Montecarlo simulations proposed in Bai(2003). In order to show factor and factor loadings confidence intervals, he observes that the estimated factor is a linear combination of the true factor. For this reason after regressing the true factor on the estimated one:

$$f_t = \hat{f}_t\hat{\beta} + \hat{\varepsilon}_t \tag{2.3}$$

The estimated asymptotic confidence interval for $\sqrt{N}f_t$ (when the significance level is 5%) becomes:

$$\sqrt{N}\hat{f}_t\hat{\beta} \pm 1.96\hat{\beta}\sqrt{\hat{\Pi}_t} \tag{2.4}$$

If we estimate the common factor with principal components and look at the estimated asymptotic variance of the factors, we see in this simple case that only when the number of series and observations relatively high, the probability that the real factor is not inside the proposed confidence interval is closer to 5% regardless of the AR(1) parameters

This is shown in Table(1), which is the result of a simulation with 2000 artificial samples, with 100 series and 200 observations which share the same loadings and the AR(1) parameter α.

Rejection Probabilities for factors asymptotic CI

α=0	n=15	n=25	n=50	n=100
t=25	0.11456	0.09244	0.07524	0.07428
t=50	0.09628	0.07630	0.06222	0.06148
t=100	0.08801	0.06944	0.05609	0.05601
t=200	0.08640	0.06869	0.05657	0.05435

α=0.5	n=15	n=25	n=50	n=100
t=25	0.08789	0.07737	0.07286	0.07183
t=50	0.07460	0.06680	0.06097	0.05951
t=100	0.06976	0.06084	0.05809	0.05511
t=200	0.06943	0.05991	0.05501	0.05460

α=0.9	n=15	n=25	n=50	n=100
t=25	0.25476	0.19048	0.15704	0.13172
t=50	0.14954	0.11136	0.09388	0.08168
t=100	0.11064	0.08209	0.06846	0.05977
t=200	0.08777	0.06801	0.05793	0.05328

Table 1: Rejection Probabilities average over 1000 simulations,
for factor asymptotic confidence intervals.
The factor is identified with a simple regression and generated
with an AR(1) process with parameter α
The significance level is 5%

This table shows a significant increase of the rejection probabilities when $\alpha = 0.9$.

We can observe the same pattern when looking at the rejection probability of the confidence intervals for the factor loadings.

Considering that in order to apply principal components we need to normalize the series by dividing for the standard deviation of the single series, the asymptotic confidence interval for $\sqrt{T}\lambda_i$ (when the significance level is 5%) will be:

$$\sqrt{T}\frac{\hat{\sigma}_{x_i}\hat{\lambda}_i}{\hat{\beta}} \pm 1.96\frac{\hat{\sigma}_{x_i}}{\hat{\beta}}\sqrt{\hat{\Theta}_i} \tag{2.5}$$

Also in this case the rejection probability is heavily influenced by the parameter α, as it is shown in Table(2).

Rejection Probabilities for factor loadings asymptotic CI

α=0	n=15	n=25	n=50	n=100
t=25	0.16687	0.12832	0.10566	0.09921
t=50	0.13120	0.08656	0.06270	0.05773
t=100	0.10853	0.04740	0.02170	0.01705
t=200	0.01887	0.00000	0.00000	0.00000

α=0.5	n=15	n=25	n=50	n=100
t=25	0.11257	0.10126	0.09780	0.09251
t=50	0.09067	0.07091	0.06460	0.06014
t=100	0.06971	0.04457	0.03151	0.02750
t=200	0.02752	0.00537	0.00129	0.00071

α=0.9	n=15	n=25	n=50	n=100
t=25	0.21793	0.14212	0.10482	0.08734
t=50	0.18300	0.11236	0.08180	0.06872
t=100	0.19733	0.10400	0.06876	0.05641
t=200	0.28993	0.12016	0.06154	0.04773

Table 2: Rejection Probabilities average over 1000 simulations,
for factor loadings asymptotic confidence intervals.
The factor is identified with a simple regression and generated
with an AR(1) process with parameter α
The significance level is 5%

We obtain even worse results in terms of rejection probabilities when we look at the asymptotic confidence interval for the common part $C_{ti} = f_t \lambda_i'$, which is:

$$\hat{\sigma}_{x_i} \hat{C}_{it} \pm 1.96 \hat{\sigma}_{x_i} \sqrt{\frac{\hat{V}_{it}}{N} + \frac{\hat{W}_{it}}{T}} \tag{2.6}$$

The rejection probability is close to the theoretical one only when t=200, n=100 and $\alpha = 0.9$, as it can be seen in the table(3) for the normalized residuals $\hat{\varepsilon}_{\hat{c}_{it}}$.

$$\hat{\varepsilon}_{\hat{c}_{it}} = \frac{\left(\hat{\sigma}_{x_i} \hat{C}_{it} - C_{it}\right)}{\hat{\sigma}_{x_i} \sqrt{\frac{\hat{V}_{it}}{N} + \frac{\hat{W}_{it}}{T}}} \tag{2.7}$$

Rejection Probabilities for common part asymptotic CI

α=0	n=15	n=25	n=50	n=100
t=25	0.12273	0.12186	0.15339	0.17222
t=50	0.08530	0.08005	0.09480	0.10608
t=100	0.06097	0.05170	0.05278	0.05430
t=200	0.04265	0.03340	0.02830	0.02330

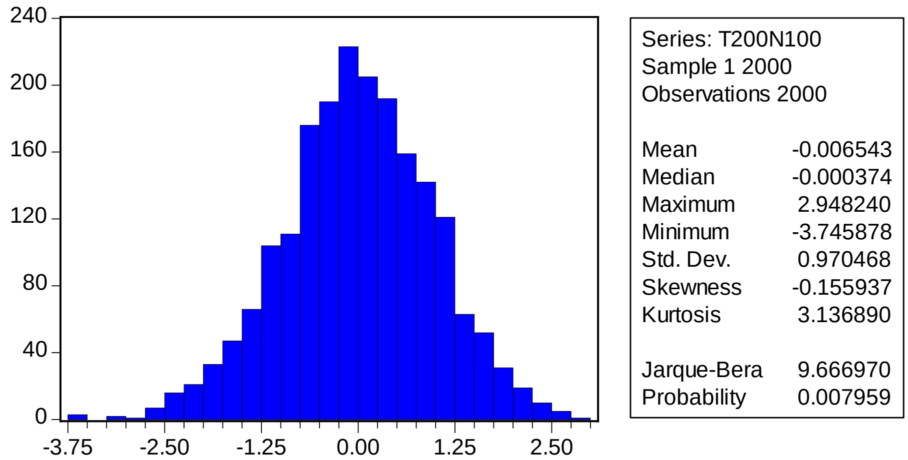
α=0.5	n=15	n=25	n=50	n=100
t=25	0.20069	0.22268	0.25806	0.29441
t=50	0.14009	0.15876	0.18817	0.22939
t=100	0.08379	0.08741	0.10100	0.12476
t=200	0.04896	0.03979	0.03298	0.02855

α=0.9	n=15	n=25	n=50	n=100
t=25	0.26459	0.29996	0.35263	0.41412
t=50	0.19159	0.23359	0.29764	0.37696
t=100	0.12854	0.15292	0.20780	0.30106
t=200	0.06341	0.05542	0.05036	0.04860

Table 3: Rejection Probabilities average over 1000 simulations,
for common part asymptotic confidence intervals.
The factor is identified with a simple regression and generated
with an AR(1) process with parameter α
The significance level is 5%

These normalized residuals where analyzed in Bai(2003), where he reported mean and standard deviations for these residuals, but he did not report rejection probability tables.

As an example in our simulations only when we have 2000 artificial samples and with $\alpha = 0.9$, we obtain a slightly biased distribution for $\hat{\varepsilon}_{\hat{c}_{50,100}}$., with standard deviation close to 1, as it can be seen in Figure(1).



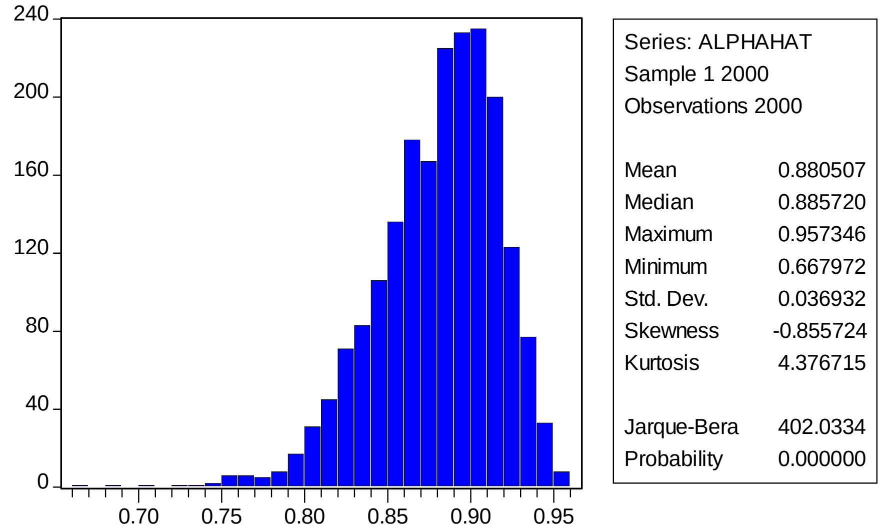
Figure(1)

Summarizing not only we find the same small sample rejection probability problems which we found in the first paper, but we have also seen that asymptotic confidence intervals are influenced by important parameters like α.

In this paper we are mainly interested in the small-sample distribution of the impulse responses of the single series to the common factor. In this simple example, the impulse responses of each series to a unitary shock to the single common shock depends only on the parameters λ_i, α and h.

$$\phi_i(h) \equiv \frac{\partial X_{i,t+h}}{\partial u_t} = \lambda_i \alpha^h \tag{2.8}$$

We start with $\alpha = 0.9$. There is a small downward bias for the OLS estimator of α, as it can be seen in the figure(2), which shows the histogram of 2000 Montecarlo simulations for $\hat{\alpha}_{ols}$ estimated with the actual factors.:



Figure(2)

The shape of the histogram for this AR parameter when we don't have the actual factor is very similar; what changes is an higher bias, which is only marginally influenced by the number of series, as it can be seen in table 4.

Mean Estimated Alpha with PC				
b=2000	n=15	n=25	n=50	n=100
t=25	0.675235	0.684257	0.69086	0.694214
t=50	0.781532	0.788258	0.793147	0.795325
t=100	0.837005	0.842287	0.845663	0.847282
t=200	0.862025	0.866396	0.869092	0.870171

Table 4
Mean estimated alpha over 2000 simulations when the factor is estimated with n series and t observations

The dynamics of the impulse responses depend on the parameter α while the impact effect depends on the parameter λ_i. Given this relation between factor loadings and impulse responses, in order to observe the coverage probability of such estimator, we simply propose for the h-steps ahead impulse responses the following estimator:

$$\hat{\phi}_i(h) = \frac{\hat{\sigma}_{x_i}\hat{\lambda}_i}{\hat{\beta}}\hat{\alpha}^h \tag{2.9}$$

With the help of the delta method, we can derive the asymptotic variance of the impulse responses (where $\hat{\sigma}_{x_i}$ and $\hat{\beta}$ are considered constant since they are simply the parameters which allow us to span the real λ_i).

$$\sqrt{T}\left(\hat{\phi}_i(h) - \phi_i(h)\right) \underset{d}{\rightarrow} N\left(0, Var(\hat{\phi}_i(h))\right) \tag{2.10}$$

$$\begin{aligned}
Var(\hat{\phi}_i(h)) &= Var(\frac{\hat{\sigma}_{x_i}\hat{\lambda}_i}{\hat{\beta}}\hat{\alpha}^h) = \begin{bmatrix} \frac{\partial\hat{\phi}_i(h)}{\partial\hat{\lambda}_i} & \frac{\partial\hat{\phi}_i(h)}{\partial\hat{\alpha}} \end{bmatrix} \begin{bmatrix} \hat{\Theta}_i & 0 \\ 0 & \hat{\Theta}_\alpha \end{bmatrix} \begin{bmatrix} \frac{\partial\hat{\phi}_i(h)}{\partial\hat{\lambda}_i} \\ \frac{\partial\hat{\phi}_i(h)}{\partial\hat{\alpha}} \end{bmatrix} = \\
&= \begin{bmatrix} \frac{\hat{\sigma}_{x_i i}}{\hat{\beta}}\hat{\alpha}^h & \frac{\hat{\sigma}_{x_i}\hat{\lambda}_i}{\hat{\beta}}h\alpha^{h-1} \end{bmatrix} \begin{bmatrix} \hat{\Theta}_i & 0 \\ 0 & \hat{\Theta}_\alpha \end{bmatrix} \begin{bmatrix} \frac{\hat{\sigma}_{x_i i}}{\hat{\beta}}\alpha^h \\ \frac{\hat{\sigma}_{x_i}\hat{\lambda}_i}{\hat{\beta}}h\alpha^{h-1} \end{bmatrix} = \\
&= \hat{\alpha}^{2h}\left(\frac{\hat{\sigma}_{x_i i}}{\hat{\beta}}\right)^2\hat{\Theta}_i + (\frac{\hat{\sigma}_{x_i}\hat{\lambda}_i}{\hat{\beta}}h\hat{\alpha}^{h-1})^2\hat{\Theta}_\alpha
\end{aligned} \tag{2.11}$$

The asymptotic variance of α (defined as $\hat{\Theta}_\alpha$) can be estimated with Newey-West method in much the same way as for λ_i ; these two parameters are not correlated because the common shock is not correlated with the idiosyncratic shocks. An asymptotic confidence interval for $\sqrt{T}\phi_i(h)$ is:

$$\sqrt{T}\frac{\hat{\sigma}_{x_i}\hat{\lambda}_i}{\hat{\beta}}\hat{\alpha}^h \pm 1.96\sqrt{Var(\hat{\phi}_i(h))} \tag{2.12}$$

These confidence intervals have good coverage, provided that $\hat{\alpha}$ is not too far from the real α; table 5 shows the average coverage probability across series and simulations, when $\alpha = 0.9$.

Rejection probability for Impulse Responses				
1 step ahead				
b=2000	n=15	n=25	n=50	n=100
t=25	0.225033	0.21942	0.20115	0.15571
t=50	0.1385	0.13272	0.12278	0.09882
t=100	0.087067	0.08408	0.07801	0.06687
t=200	0.062267	0.06086	0.05748	0.0477
12 steps ahead				
b=2000	n=15	n=25	n=50	n=100
t=25	0.720333	0.70784	0.69322	0.656985
t=50	0.550767	0.5349	0.51969	0.48297
t=100	0.3993	0.37294	0.34649	0.3177
t=200	0.304533	0.27252	0.25093	0.228895

Table 5
Mean rejection probability over 2000 simulations
for different n and t (when a=0.9)
(Probability that the real impulse response to the
common shock is not inside the confidence interval)

Table(5) shows that when there is a large bias $Bias(\hat{\alpha}) = E(\hat{\alpha}) - \alpha$ also such nonlinear functions of α as the impulse responses are influenced in terms of rejection probability.

When the actual parameter is low also the bias is low; this is confirmed by the rejection probabilities when $\alpha = 0.5$ reported in table 5b; we can see that they are significantly different from the ones reported in table 5.

Rejection probability for Impulse Responses				
1 step ahead				
b=2000	n=15	n=25	n=50	n=100
t=25	0.1333	0.13688	0.14696	0.13715
t=50	0.089833	0.09162	0.09387	0.086975
t=100	0.061	0.06048	0.05928	0.04985
t=200	0.053467	0.05224	0.04715	0.029675
12 steps ahead				
b=2000	n=15	n=25	n=50	n=100
t=25	0.520833	0.52262	0.51305	0.50447
t=50	0.422333	0.4237	0.40879	0.40319
t=100	0.342	0.33232	0.32395	0.31356
t=200	0.272333	0.2557	0.25514	0.24419

Table 5b
Mean rejection probability over 2000 simulations
for different n and t (when a=0.5)
(Probability that the real impulse response to the
common shock is not inside the confidence interval)

This difference means that the studentized statistic $\frac{\sqrt{T}\left(\hat{\phi}_i(h)-\phi_i(h)\right)}{\sqrt{Var(\hat{\phi}_i(h))}}$, which is asymptotically distributed as a N(0,1) is not a pivotal statistic, since it depends on α.

Summarizing we have two main problems: in small samples: $\hat{\alpha}$ is biased and the statistics of interest are not pivotal, especially the impulse responses. We also should consider that these statistics are a nonlinear function of the basic parameters, for this reason there is a huge difference between the impact impulse response and the impulse response 12 steps ahead.

In order to solve the first problem we can apply the Kilian (1998) bootstrap-after-bootstrap correction, which in our case can be implemented as follows:

1. Estimate $\hat{\alpha}$, $\hat{\lambda}_i$, $\hat{e}_i$ and $\hat{u}$ with principal components and OLS regressions, generate B_1 artificial samples $X^{(j)}$ by resampling from the residuals and for each artificial sample estimate $\hat{\alpha}^{(j)}$.

2. Estimate the bias for α as $\Psi(\alpha) = \sum_{j=1}^{B_1} \hat{\alpha}^{(j)}/B_1 - \hat{\alpha}$, and generate B_2 new samples with the parameter $\tilde{\alpha}$=$\hat{\alpha}-$ $\Psi(\alpha)$, obtaining B_2 new estimates $\tilde{\alpha}^{(j)}$ and $\tilde{\lambda}_i^{(j)}$. If $\tilde{\alpha} \geq 1$, shrink iteratively the bias $\Psi(\alpha)$, until $\tilde{\alpha} < 1$.

3. Finally, evaluate the distribution of the statistic of interest (as an example the impulse responses) as a function of $\tilde{\alpha}^{(j)*}$ =($\tilde{\alpha}^{(j)} - \Psi(\alpha)$)), $\tilde{\lambda}_i^{(j)}$ and $\hat{\sigma}_{x_i}$

   This correction works well for the estimation of α, if the number of observations is higher than 25 (see table 6).

Mean bias-corrected α				
b=100	n=15	n=25	n=50	n=100
t=25	0.769322	0.781846	0.781781	0.793754
t=50	0.853963	0.856962	0.858212	0.85818
t=100	0.893894	0.894326	0.893664	0.896168
t=200	0.89857	0.900077	0.900205	0.901392

Table 6
Mean Bias-corrected α over 100 simulations
(Actual α=0.9)

In Kilian (1999) it is shown that simple percentile intervals based on bias-corrected parameters perform better than more sophisticated studentized percentile intervals in small samples and for nonlinear statistics. Table 7 shows similar results in our case for different α, with 100 artificial samples for each α and when we have 100 observations and 100 series. In this case B_1=100 and B_2=100.

Bootstrap-after-bootstrap for different α			
Actual α	0.1	0.5	0.9
Mean estimated α	0.092663	0.479481	0.842317
Mean bias-corrected α	0.108467	0.509152	0.896168
Rejection probability for α	0.04	0.06	0.06
Rejection probability for φ(1)	0.0277	0.0278	0.0283
Rejection probability for φ(12)	0.0217	0.0217	0.028

Table 7
Mean estimated α and rejection probabilities related to 3 simulations with different alpha, 100 artificial samples, 100 series, 100 observations and 100 bootstrap replications φ(h) is the impulse response h-steps ahead of the single φ series to the common shock.

The rejection probabilities are on average over 100 series and 100 replications, and they measure how many times the real α or the real $\phi_i(h)$ are not inside the 95% percentile bootstrap intervals. Table 7 shows clearly that these bootstrap-after-bootstrap confidence intervals are more reliable than the ones based on asymptotic theory, also if we observe that the rejection probabilities are too low, meaning that the simple percentile intervals used here are too large.

2.3 Rotation matrices

In this section we come back to the DGP that we proposed in the first paper: we have now two common shocks modelled as independent AR(2) processes (h=2); the two factors are loaded as an MA(1) (p=1).

$$X_{it} = \begin{bmatrix} \lambda_{i1,0} & \lambda_{i2,0} & \lambda_{i1,1} & \lambda_{i2,1} \end{bmatrix} \begin{bmatrix} f_{1,t} \\ f_{2,t} \\ f_{1,t-1} \\ f_{2,t-1} \end{bmatrix} + \sigma_i e_{i,t} \tag{2.13}$$

$$\begin{bmatrix} f_{1,t} \\ f_{2,t} \end{bmatrix} = \begin{bmatrix} \Gamma_{11} & 0 \\ 0 & \Gamma_{21} \end{bmatrix} \begin{bmatrix} f_{1,t-1} \\ f_{2,t-1} \end{bmatrix} + \tag{2.14}$$

$$\begin{bmatrix} \Gamma_{12} & 0 \\ 0 & \Gamma_{22} \end{bmatrix} \begin{bmatrix} f_{1,t-2} \\ f_{2,t-2} \end{bmatrix} + \begin{bmatrix} \sigma_{u1} & 0 \\ 0 & \sigma_{u2} \end{bmatrix} \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} \tag{2.15}$$

In this model we assume that $e_{i,t}$, $u_{1,t}$ and $u_{2,t}$ are iid random variables with unit variance. If we define the static factor F_t as the q(p+1) vector:

$$F_t = \begin{bmatrix} f_{1,t} \\ f_{2,t} \\ f_{1,t-1} \\ f_{2,t-1} \end{bmatrix} \tag{2.16}$$

we can rewrite the model in static form as:

$$X_{it} = \lambda_i F_t + \sigma_i e_{i,t} \tag{2.17}$$

$$F_t = \begin{bmatrix} f_{1,t} \\ f_{2,t} \\ f_{1,t-1} \\ f_{2,t-1} \end{bmatrix} = \begin{bmatrix} \Gamma_{11} & 0 & \Gamma_{12} & 0 \\ 0 & \Gamma_{21} & 0 & \Gamma_{22} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} f_{1,t-1} \\ f_{2,t-1} \\ f_{1,t-2} \\ f_{2,t-2} \end{bmatrix} + \begin{bmatrix} \sigma_{u1} & 0 \\ 0 & \sigma_{u2} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} \tag{2.18}$$

$$= \Gamma F_{t-1} + \Omega_u u_t \tag{2.19}$$

The moving average representation of this model as a function of common and idiosyncratic shocks is:

$$X_{it} = \lambda_i \sum_{j=0}^{\infty} \Gamma^j \Omega_u u_{t-j} + \sigma_i e_{i,t}$$

And the impulse responses to the common shocks are:

$$\frac{\partial X_{it+j}}{\partial u_t} = \lambda_i \Gamma^j \Omega_u$$

This static model is observationally equivalent to a model where static factors are rotated by a full rank matrix H, as for example:

$$\tilde{F}_t = HF_t = H\Gamma H^{-1} HF_{t-1} + H\Omega_u u_t = \tilde{\Gamma}\tilde{F}_{t-1} + \tilde{v}_t \tag{2.20}$$

$$X_{it} = \lambda_i H^{-1} HF_t + \sigma_i e_{i,t} = \tilde{\lambda}_i \tilde{F}_t + \sigma_i e_{i,t} \tag{2.21}$$

So when we apply principal components to the real variables, we obtain an estimate of $\tilde{F}_t$, $\tilde{\Gamma}$, $\tilde{v}_t, \tilde{\lambda}_i$

If we want to estimate the impulse responses we must take account of the fact that the (p+1)q static factors are driven by q common shocks. The relation between the reduced form common shocks $\tilde{v}_t$ and the real common shocks is the following:

$$\tilde{v}_t = H\Omega_u u_t \Leftrightarrow \Sigma_{\tilde{v}} \equiv E(\tilde{v}_t \tilde{v}_t') = \tag{2.22}$$

$$= H\Omega_u E\left(u_t u_t'\right) \Omega_u' H' \tag{2.23}$$

$$= H\Omega_u \Omega_u' H' \tag{2.24}$$

But we can also find orthogonal shocks $\tilde{u}_t = R(\theta)' u_t$, where $R(\theta)R(\theta)' = I_q = R(\theta)'R(\theta)$, such that $\tilde{v}_t = \tilde{H}\tilde{\Omega}_u \tilde{u}_t$.

$R(\theta)$ is a q-by-q rotation matrix which depends on the q(q-1)/2-by-1 vector of parameters θ.

In general $R(\theta)$ is the product of q(q-1)/2 matrices Q_j, where j corresponds to the indexation of the q(q-1)/2 combinations (without repetition) of the integers 1,2,..q. As an example if q=7, we have 7(7-1)/2=21 combinations: (1,2) (1,3) (1,4),... and so on until (6,7).

Each matrix Q_j is a function of only one parameter $\theta j \in [0, 2\pi]$, as it can be seen below.

$$R(\theta) = \prod_{j=1}^{q(q-1)/2} Q_j \text{ where for example } Q_2 = \begin{bmatrix} \cos(\theta_2) & 0 & \sin(\theta_2) & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ -\sin(\theta_2) & 0 & \cos(\theta_2) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.25}$$

In our simple case with q=2, we need only one parameter θ and the rotation matrix will simply be:

$$R(\theta) = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \tag{2.26}$$

As an alternative we can specify a rotation matrix where the elements of the main diagonal have opposite sign, for example when q=2:

$$R(\theta) = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{bmatrix} \tag{2.27}$$

If we look at the model the matrix Ω_v should have rank q, but when we estimate a VAR(1) for the estimated static factors, $\hat{\Omega}_{\hat{v}}$ has rank q(p+1).

In order to estimate Ω_u I followed a procedure involving singular value decomposition:

1. Estimate eigenvectors and eigenvalues of $\hat{\Sigma}_{\tilde{v}}$, order eigenvalues in descending order in the diagonal matrix D, and accordingly sort the associated eigenvectors in the matrix P.

2. Apply a singular value decomposition to the first q columns of $PD^{0.5}$. We will obtain a q(p+1)-by- q(p+1) matrix U that we can interpret as H, a (p+1)*q-by-q matrix S with a structure similar to Ω_u, and a q-by-q diagonal matrix V. If we premultiply our static

factors with U', we obtain an estimate of Ω_u as $\hat{\Omega}_u = SV'$.

We can estimate the impulse responses of the single series to a generic common shock $\tilde{u}$:

$$\frac{\partial \hat{X}_{it+j}}{\partial \tilde{u}_t} = \hat{\lambda}_i U \left(U' \hat{\Gamma} U \right)^j SV' \tag{2.28}$$

In order to estimate the impulse responses to the identified shock $u_t = R(\theta)\tilde{u}_t$ it will be sufficient to post-multiply the impulse responses with the rotation matrix, where θ is chosen according to identifying restrictions.

As we have done in the first paper we assume that our n series are driven by two common shocks, and that the q=2 orthogonal factors have the following AR structure, where the second shock is more persistent:

$$(1 - 0.1L)(1 - 0.7L) f_{1t} = u_{1t} \tag{2.29}$$

$$(1 - 0.3L)(1 - 0.9L) f_{2t} = u_{2t} \tag{2.30}$$

In this DGP we impose that the first shock has a positive effect on impact, while the second shock has a negative impact effect ($\lambda_{i1,0} > 0$, $\lambda_{i2,0} < 0$ for all variables).

This rotation matrix can be useful in a Montecarlo simulation where we know the real impulse responses, because in order to show how bootstrap methods can provide good impulse responses confidence intervals we must overcome any identification problems by choosing an appropriate rotation of the shocks.

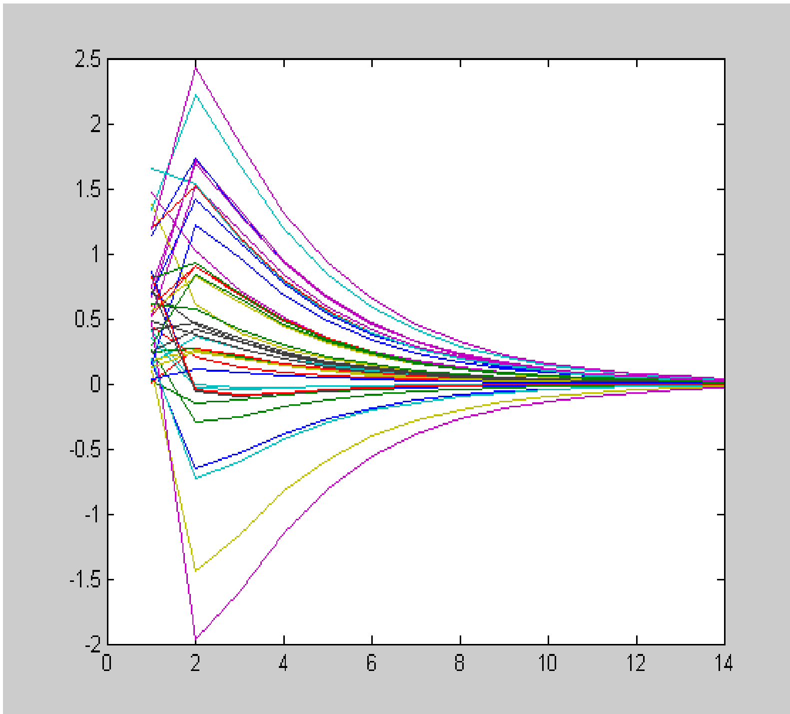
Moreover a small example shows that t these rotation matrices can be the tools we need if we want to apply for this model a sort of partial identification strategy which consists in choosing a set of $\theta \in [0, 2\pi]$ such that the impulse responses satisfy a set of sign restrictions:

1) The first shock has a positive impact effect on all variables.

2) The second shock has a negative impact effect on all variables.

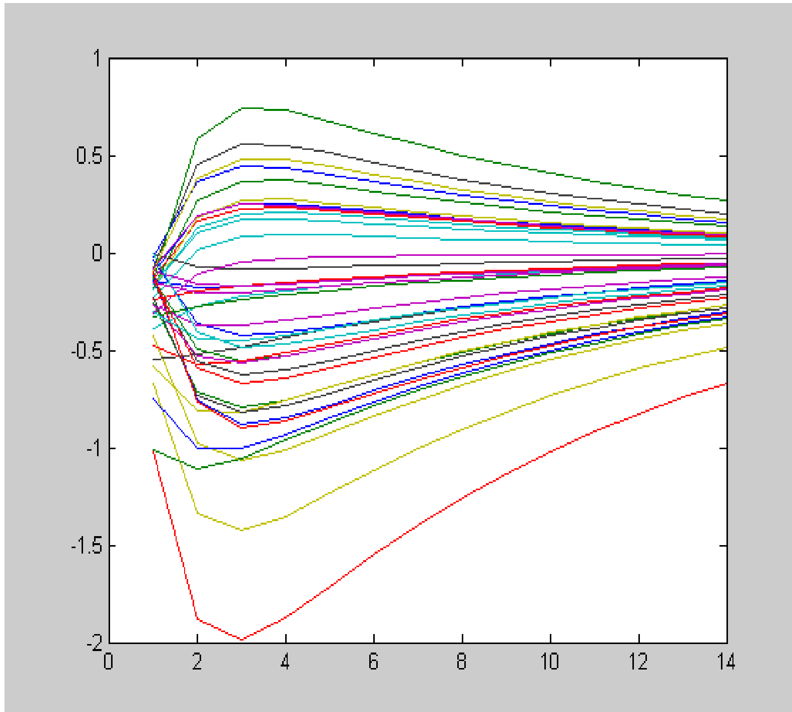
3) The effect of a shock to the first factor dies out rapidly, so we can expect that after 12 periods all impulse responses are close to zero.

As an example look at figure (3) which shows the actual impulse responses to the first shock:



Figure(3): Actual impulse responses of 100 artificial series to the first shock

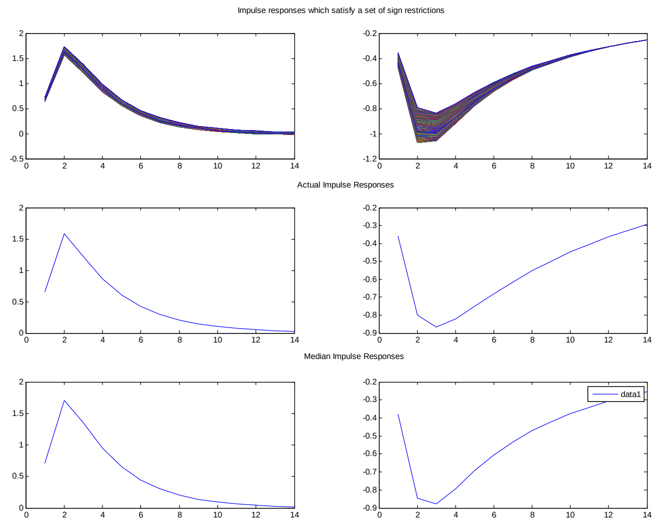
1. While the effects of the second shock is much more persistent, as it can be seen in figure (4).



Figure(4): Actual impulse responses of 100 artificial series to the second shock

If we just estimate the impulse responses without identifying the shocks, we can obtain impulse responses with opposite sign (as it was the case for this example), or maybe the second estimated shock is equal to the opposite of the real first one. In our case inside a grid of 6283 θs we found only 119 θs (equal to approximately the 2% of the grid) which satisfied the restrictions.

Figure(5) shows for each column the impulse responses of the 20th series to the two shocks. The second row shows the actual impulse responses, the first ones all the rotations which satisfy the sign restrictions while the third row shows the impulse responses related to the median θ.



Figure(5)

Figure(5)

We show this example because before analyzing any bootstrap procedure for impulse responses we must be sure to identify the right path.

2.4 Conclusions

In this paper we examined closely the issue of providing confidence intervals for impulse responses in a simple model and how to reach identification in our complicated model.

In an Montecarlo which we did not reported we simulated a DGP similar to what we have seen in the previous section: there is a positive effect of all the artificial series to a positive first shock, and a negative effect to a positive second shock (except for one of the series, where we imposed a zero impact effect to the second shock).

Also in this case with the block bootstrap: static factors and idiosyncratic shocks were resampled in order to obtain artificial samples. The rotation matrix was used in order to just identify our model: a zero restriction on the effect of the second shock to the first variable is theoretically enough for fixing the rotation parameter θ.

When using the block bootstrap we observed a persistence problem similar to what we saw in the second section with the AR(1) parameter α. While in that simple case α was a measure of the persistence for a single factor, when we deal with two dynamic factors and four static factors the persistence of a common shock can be measured either by the companion form estimated matrix $\hat{\Gamma}$ of the static factor VAR, or by the average block length of the factor that we estimate starting from the artificial samples.

If we do not take account of this aspect in a satisfactory way what we obtain are impulse responses from bootstrapped samples with a very low persistence, and which fail in containing the actual impulse response. A natural extension of this paper must deal with this problem, taking account of the fact that probably with such a complicated dynamic model we should use the parametric bootstrap instead of the block bootstrap.

Chapter 3

Bootstrap Methods for DFM: An application to latent factors

3.1 Introduction

. In this paper we want to show how simple bootstrap methods can improve the power of a test where the null hypothesis is that a given macroeconomic time series is a latent factor. These tests can be useful if we want to give an interpretation to factor forecasts and for the task of preselecting variables from our dataset when they are mainly driven by noise, following the suggestion in Boivin-Ng (2003).

3.2 Observed and latent factors

In Bai-Ng(2004) the authors propose a way to evaluate if an observable variable G_{jt} can be considered a latent factor.

They consider the factor model:

$$x_{it} = \lambda_i' F_t + e_{it} \tag{3.1}$$

And they want to test the null hypothesis that $G_{jt} = \delta_j' F_t$ for any t. In order to do it they propose to regress the observable variable G_{jt} on the factors extracted with principal components methods $\tilde{F}_t$, obtaining $\hat{G}_{jt} = \hat{\gamma}_j' \tilde{F}_t$. In this way they obtain a kind of t-statistics.

$$\tau_t(j) = \frac{\left(\hat{G}_{jt} - G_{jt}\right)}{\sqrt{\hat{Var}(\hat{G}_{jt})}} \tag{3.2}$$

$$\hat{Var}(\hat{G}_{jt}) = \frac{1}{N} \hat{\gamma}_j' \tilde{V}^{-1} \tilde{\Gamma}_t \tilde{V}^{-1} \hat{\gamma}_j \tag{3.3}$$

$\tilde{\Gamma}_t$ is a consistent estimate of $H'^{-1}\Gamma_t H^{-1}$. If the idiosyncratic errors are stationary it is still possible to obtain a consistent estimator when there is cross-section correlation between the idiosyncratic shocks. They refer to it as CS-HAC estimator, where $\frac{n}{\min[N,T]} \to 0$. In practice they average over k estimates obtained from k random samples which involve k series, where k=min$[\sqrt{N}, \sqrt{T}]$.

$$\tilde{\Gamma} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \tilde{\lambda}_i \tilde{\lambda}_j' \frac{1}{T} \sum_{t=1}^{T} \tilde{e}_{ti} \tilde{e}_{tj} \tag{3.4}$$

If the idiosyncratic shocks are not cross-sectionally correlated there is an estimator which allows also for heteroskedasticity.

$$\tilde{\Gamma}_t = \frac{1}{N} \sum_{j=1}^{N} \tilde{e}_{ti}^2 \tilde{\lambda}_i \tilde{\lambda}_i' \tag{3.5}$$

While if $E(e_{ti}^2) = \sigma_e^2$ for all t and for all i, the estimator becomes:

$$\tilde{\Gamma} = \left(\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{e}_{ti}^2 \right) \frac{\hat{\Lambda}'\hat{\Lambda}}{N} \tag{3.6}$$

Under the regularity conditions which we reported in the first paper $\tau_t(j)$ is asymptotically distributed as a N(0,1). A way to test the null hypothesis is to look at the frequency with which this statistics is higher in absolute value than the critical value of a standard N(0,1) They define with Φ_α the (1-α) percentile of this distribution (so if $\alpha = 0.025$, $\Phi_\alpha = 1.96$) and with A(j) this frequency which should tend to 2α.

$$A(j) = \frac{1}{T} \sum_{i=1}^{T} 1(|\hat{\tau}_t(j)| > \Phi_\alpha) \tag{3.7}$$

This statistic is equivalent to the rejection probability of the asymptotic confidence interval for G_{jt}, which corresponds in this case to:

$$(\hat{G}_{jt} \mp \Phi_\alpha \sqrt{Var(\hat{G}_{jt})}) \tag{3.8}$$

These tests are heavily influenced by small sample bias related to both the t and n dimension, as it,can be seen in Table (1a), which is the result of the same Montecarlo simulation reported in table(1a) in Bai-Ng (2004).

Table 1a

N	T	j	A(j)	M(j)	NS(j)	CI(j)	R2	R2-	R2+	Boot-A(j)
50	50	1	0,03936	0,033	0,028264	0,96408	0,97258	0,95783	0,98733	0,0554
50	50	2	0,03794	0,03	0,027883	0,96568	0,97293	0,95837	0,9875	0,05556
50	50	3	0,18526	0,714	0,068114	0,96184	0,93646	0,90359	0,96933	0,06118
50	50	4	0,1885	0,735	0,068371	0,96356	0,93622	0,90324	0,96921	0,06114
50	50	5	0,85224	1	5,4854	0,94306	0,22716	0,13477	0,31955	0,37998
50	50	6	0,85458	1	5,053	0,93646	0,22717	0,1348	0,31953	0,38026
50	50	7	0,94926	1	354,9	0,94822	0,040253	0,019695	0,060811	0,73362
100	50	1	0,03448	0,016	0,013651	0,96906	0,98655	0,97919	0,9939	0,0529
100	50	2	0,03448	0,02	0,013666	0,96832	0,98653	0,97917	0,99389	0,05392
100	50	3	0,28426	0,97	0,052827	0,96616	0,94995	0,92365	0,97624	0,05838
100	50	4	0,28048	0,965	0,052362	0,96544	0,95036	0,92427	0,97645	0,05832
100	50	5	0,8962	1	5,6152	0,93132	0,22541	0,1341	0,31671	0,38514
100	50	6	0,8993	1	5,6508	0,93692	0,22083	0,13059	0,31107	0,38556
100	50	7	0,96438	1	187,12	0,92922	0,041874	0,020535	0,063213	0,73002
50	100	1	0,03653	0,024	0,028273	0,9651	0,97255	0,96211	0,983	0,05485
50	100	2	0,03752	0,03	0,028231	0,96443	0,97259	0,96216	0,98302	0,05469
50	100	3	0,18126	0,809	0,068636	0,9648	0,93589	0,91242	0,95937	0,05992
50	100	4	0,18121	0,815	0,06833	0,9643	0,93617	0,91279	0,95955	0,06023
50	100	5	0,85771	1	4,5255	0,9486	0,2077	0,14505	0,27035	0,38895
50	100	6	0,85737	1	4,4209	0,95086	0,20728	0,1445	0,27005	0,38675
50	100	7	0,96628	1	465,58	0,96489	0,018981	0,011821	0,02614	0,81256
200	100	1	0,0301	0,014	0,006724	0,97121	0,99332	0,99072	0,99592	0,05219
200	100	2	0,0305	0,01	0,006745	0,97074	0,9933	0,99069	0,99591	0,05215
200	100	3	0,41015	1	0,046274	0,96856	0,95581	0,93927	0,97235	0,05726
200	100	4	0,4056	1	0,046098	0,96796	0,95598	0,93949	0,97246	0,0574
200	100	5	0,92581	1	4,267	0,94208	0,21398	0,14975	0,2782	0,37962
200	100	6	0,92885	1	4,6148	0,94169	0,20635	0,144	0,2687	0,38944
200	100	7	0,98222	1	484,28	0,94748	0,01938	0,012072	0,026688	0,81331
100	200	1	0,0301	0,007	0,013554	0,97059	0,98663	0,98298	0,99029	0,05227
100	200	2	0,03042	0,007	0,013502	0,97034	0,98668	0,98304	0,99032	0,05239
100	200	3	0,27621	0,999	0,053409	0,96981	0,94933	0,936	0,96265	0,05722
100	200	4	0,27696	1	0,053421	0,96954	0,94932	0,93599	0,96265	0,057945
100	200	5	0,89876	1	4,2167	0,95576	0,20405	0,15969	0,24841	0,3838
100	200	6	0,89969	1	4,2666	0,95317	0,20233	0,15823	0,24642	0,38632
100	200	7	0,98153	1	569,36	0,95462	0,010277	0,007484	0,01307	0,86081

The first two columns show number of series and observations while j is the index of seven $G_{jt} = \delta_j' F_t + \varepsilon_{jt}$ where δ_j is a 2-by-1 vector of weights and $\varepsilon_{jt} \sim \sigma_\varepsilon(j) N(0, var(\delta_j' F_t))$.

j	1	2	3	4	5	6	7
δ_{j1}	1	1	1	1	1	1	0
δ_{j2}	1	0	1	0	1	0	0
σ_e	0	0	.2	.2	2	2	1

The first two observable variables are exact factors and the third and fourth factors are close to exact factors. Fifth and sixth observable variables are highly contaminated by idiosyncratic shocks, while the seventh is completely unrelated to the two common factors. Measures of the relative weight of common factors and noise are the Noise-Signal ratio and the R^2 which are reported in table(1).

$$NS(j) = \frac{\widehat{Var}(\hat{\varepsilon}(j))}{\widehat{Var}(\hat{G}(j))} \tag{3.9}$$

$$R^2(j) = \frac{\widehat{Var}(\hat{G}(j))}{\widehat{Var}(G(j))} \tag{3.10}$$

The table reports also the probability that, under the assumption that the idiosyncratic shocks are not correlated, the maximum absolute value of the statistic $\hat{\tau}_t(j)$ is higher than the appropriate critical value.

$$M(j) = \max_{1\leq t\leq T} |\hat{\tau}_t(j)| \tag{3.11}$$

$$P(M(j) \leq x) \approx [2\Phi(x) - 1]^T \tag{3.12}$$

We concentrate our attention on the statistic A(j). We can observe that the results are heavily influenced by the number of observations and the number of variables, while the same is not true if we use a simple Efron Percentile confidence interval. This is obtained by drawing artificial samples of factors via the block bootstrap, and assuming that the observable variable are exact factors.

3.3 Conclusion

The preliminary results show that when we deal with simple statistics, bootstrap methods can provide easy to use ways to obtain bootstrap rejection probabilities which help us in providing tests which are not affected by small sample biases. This result is heavily related with what with have seen in the first paper, where the rejection probabilities which we reported for the common part naive percentile where in practice the rejection probability of A(j).

Bibliography

[1] AMENGUAL D., 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, January 2007

[2] ARTIS M, BANERJEE A., MARCELLINO M. (2001), "Factor Forecasts for the UK"

[3] BAI J. (2003), "Inferential Theory for Factor Models of Large Dimensions", Econometrica 71:135-171.

[4] BAI J., NG S., (2002), "Determining the Number of Factors in Approximate Factor Models", Econometrica 70:191-221.

[5] BAI J., NG S., (2005), "Determing the Number of Primitive Shocks in Factor Models".

[6] BAI J., NG S., (2006,a), "Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-Augmented Regressions", Econometrica 74, No 4 (July 2006), 1133-1150.

[7] BAI J., NG S., (2006b), "Evaluating Latent and Observed Factors in Macroeconomics and Finance", Journal of Econometrics 131 (2006) 507-537.

[8] BERNANKE, B. S. & BLINDER, A. (1992) "The Federal Funds Rate and the channels of monetary transmission", American Economic Review, 82, 901-921.

[9] BERNANKE, B., BOIVIN, J.,(2003) "Monetary policy in data-rich environment", Journal of Monetary Economics 50:525-546.

[10] BERNANKE, B., BOIVIN, J., ELIASZ, P. (2005) “Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach”, Quarterly Journal of Economics, February 2005, 387-422.

[11] BERNANKE, B. S. &-MIHOV, I “Measuring Monetary Policy”, NBER working paper no 5145, June 1995

[12] BERNANKE, B. S.& MIHOV, I, “The Liquidity Effect and Long-Run Neutrality”, Carnegie-Rochester Conference Series on Public Policy 49 (1998), 149-194

[13] BERKOWITZ, J.,& KILIAN, L. "Recent Developments in Bootstrapping Time Series" (with J. Berkowitz), Econometric Reviews , 19(1), February 2000, 1-54.

[14] BLANCHARD, O.& QUAH, D., “The Dynamic Effects of Aggregate Supply and Demand Disturbances”, American Economic Review 79(4) (1989), 655-673

[15] BOIVIN, J., GIANNONI, M., (2006), “Has Monetary Policy Become More Effective?”, NBER Working Paper No 9459

[16] BOIVIN J., NG S.(2003), “Are More Data Always Better for Factor Analysis?”, NBER Working Paper No. 9829

[17] BRILLINGER D. R. (1964), “A Frequency Approach to the Techniques of Principal Components, Factor Analyses and Canonical Variates in the Case of Stationary Time Series”, Invited Paper, Royal Statistical Society Conference, Cardiff Wales.(Available at http://www.stat.berkeley.edu/users/brill/papers.html).

[18] BRILLINGER D. R. (1981), Time Series: Data Analysis and Theory, expanded edition (Holden-Day, San Francisco).

[19] BRISSON M., CAMPBELL B. & GALBRAITH J. W. (2002), "Forecasting some low-predictability time series using diffusion indices", manuscript (CIRANO).

[20] CANOVA F., DE NICOLO’ G., (2002) “Money Matters for Business Cycle Fluctuations in the G7”, Journal of Monetary Economics, 49, 1131-1159.

[21] CHRISTIANO, L. J., EICHENBAUM, M. & EVANS, C. L. (1998) “Monetary Policy Shocks: What Have We Learned and to What End?” NBER WP No. 6400.

[22] CHRISTIANO, L. J., EICHENBAUM, M. & EVANS, C. L. (1996) “The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds”, Review of Economics and Statistics, 78:16-34

[23] COGLEY, T., SARGENT, T. J., (2002) “Drift and Volatilities: Monetary Policy and Outcomes in the Post WWII U.S.”

[24] DEL NEGRO M, and SCHORFHEIDE F. (2004): " Priors from General equilibrium Models for VARs," International Economic Review, 45, 643-673.

[25] DEL NEGRO M,, SCHORFHEIDE F., SMETS F. and WOUTERS R. (2006): "On the Fit of New-Keynesisan Models," Working Paper.

[26] DICICCIO, T. J., EFRON, B (1996): "Bootstrap Confidence Intervals", Statistical Science, Vol 11, No 3, 189-228.

[27] DOZ C, GIANNONE D., REICHLIN L. ,(2006) "A Quasi Maximum Likelihood Approach for Large Approximate Dynamic Factor Models", CEPR discussion paper No 5724

[28] EFRON B. (1979) Bootstrap Methods: Another look at the jackknife, Annals of Statistics, 7, 1-26.

[29] ENGLE R. F., WATSON M. W., (1981), “A One-Factor Multivariate Time Series Model of Metropolitan Wages Rates”, Journal of the American Statistical Association , 76:774-781.

[30] FAUST, J. (1998) “The robustness of identified VAR conclusions about money”, Carnegie-Rochester Conference Series on Public Policy, 49 pp 207-244

[31] FORNI M., HALLIN M., LIPPI M., REICHLIN L., (2000), “The Generalized Factor Model: Identification and Estimation”, The Review of Economics and Statistics 82:540-554.

[32] FORNI M., HALLIN M., LIPPI M., REICHLIN L., (2003), “Do financial variables help forecasting inflation and real activity in the EURO area?", Journal of Monetary Economics, 50:1243-1255.

[33] FORNI M., GAMBETTI L., (2008), "The dynamic effects of monetary policy. A strcutural factor model approach", mimeo

[34] FORNI M., GIANNONE D., LIPPI M., REICHLIN L., (2005), “Opening the Black Box: Structural Factor Models with large cross-sections”

[35] FORNI M., HALLIN M., LIPPI M., REICHLIN L., “The Generalized Dynamic Factor Model One-Sided Estimation and Forecasting”, Journal of the American Statistical Association, Vol. 100, No. 471, September 2005 pp.830-840.

[36] GEWEKE J., (1977), “The Dynamic Factor Analyses of Economic Time Series”, in D. J. Aigner and A.S. Goldberger, eds, Latent Variables in Socio-Economic Models, (North-Holland, Amsterdam).

[37] GIANNONE D., REICHLIN L., and SALA L.(2002), "Tracking Greenspan:Systematic and Unsystematic Monetary Policy Revisited", CEPR discussion paper No 3550.

[38] HALLIN M., LISKA R.(2007) , "Determining the Number of Factors in the General Dynamic Factor Model", Journal of the American Statistical Association, 2007, vol. 102, pp 603-617

[39] KAPETANIOS G., MARCELLINO M., (2003), “A Comparison of Estimation Methods for Dynamic Factor Models of Large Dimensions”.

[40] KILIAN L. , (1998), "Small-Sample Confidence Intervals for Impulse Responses Functions", The Review of Economics and Statistics, Vol. 80, No. 2. (May 1998), pp 218-230.

[41] KILIAN L., (1999), "Finite-Sample Properties of Percentile and Percentile-t Bootstrap Confidence Intervals for Impulse Responses", The Review of Economics and Statistics, Vol. 81, No 4, (Nov. 1999), pp. 652-660.

[42] KILIAN L. (2001), "Impulse Responses Analysis in Vector Autoregressions with Unknown Lag Order", Journal of Forecasting, 20, pp 161-179 (2001)

[43] KIM C.-J., NELSON C. R., (1999), "Has the US Economy Become More Stable? A Bayesian Approach Based on a Markov Switching Model of the Business Cycle", Review of Economics and Statistics 81(4), 608-661.

[44] LEEPER, &. M. , SIMS, C. A. & ZHA, T. (1996) “What does Monetary Policy Do?” Brookings Papers on Economic Activity, serie 2, 1996, 1-63

[45] MacKINNON, J. G. (2002) "Bootstrap Inference in Econometrics" The Canadian Journal of Economics, Vol. 35, No 4. (Nov., 2002) pp 615-645.

[46] POLITIS, D. N., WHITE, H. (2004) “Automatic Block-Length Selection for the Dependent Bootstrap” Econometric Reviews Vol. 23, No 1, pp.53-70, 2004

[47] RAMOS, E. (1984), "A Bootstrap for Time Series", Qualifying paper, Department of Statistics, Harvard University.

[48] QUAH D. and SARGENT T. J. (1993), “A Dynamic Index Model for Large Cross Sections”, in J.H. Stock and M.W. Watson, eds., Business Cycles, Indicators and Forecasting (University of Chicago Press for the NBER, Chicago) Ch 7:179-204.

[49] SALA, L. (2001) “Monetary Transmission in the Euro Area: A Factor Model Approach”

[50] SARGENT, T. J. (1989), “Two Models of Measurements and the Investment Accelerator”, The Journal of Political Economy 97:251-287.

[51] STOCK, J. H., WATSON M. W. (1989), “New indexes of coincident and leading economic indicators”, NBER Macroeconomics Annual, 351-393.

[52] STOCK, J. H., WATSON M. W. (1991), “A Probability Model of the Coincident Economic Indicators” in G. Moore and K. Lahiri, eds, The Leading Economic Indicators: New Approaches and Forecasting Records (Cambridge University Press, Cambridge) 63-90.

[53] STOCK, J. H., WATSON M. W. (1999), “Forecasting Inflation”, Journal of Monetary Economics 44:293-335.

[54] STOCK, J. H., WATSON M. W. (2002a), “Macroeconomic Forecasting Using Diffusion Indexes”, Journal of Business and Economic Statistics 20:147-162.

[55] STOCK, J. H., WATSON M. W. (2002b), "Forecasting Using Principal Components from a Large Number of Predictors", Journal of the American Statistical Association, 97:1167-1179.

[56] STOCK, J. H., WATSON M. W. (2002c), "Has the Business Cycle Changed and Why?", Nber Working Paper No 9127.

[57] STOCK, J. H., WATSON M. W. (2004), "Forecasting with many predictors", Handbook of Forecasting.

[58] STOCK, J. H., WATSON M. W. (2005) "Implications of Dynamic Factor Models for VAR Analyses".

[59] STRONGIN, S. (1995) "The Identification of Monetary Policy Disturbances Explaining the Liquidity Puzzle", Journal of Monetary Economics, 35, 463-497

[60] UHLIG, H., (1997), "What Are the Effects of Monetary Policy? Results from an Agnostic Identification Procedure" CEPR Working Paper No. 2137.

[61] UHLIG, H., (2001), "Did the FED Surprise the Markets in 2001? A Case Study for VARs with Sign Restrictions" Tilburg University. Center for Economic Research Discussion paper 2001-88