Federal Reserve Bank of New York
Staff Reports

Model Selection Criteria for Factor-Augmented Regressions

Jan J. J. Groen
George Kapetanios

Staff Report no. 363
February 2009
Revised May 2012

Model Selection Criteria for Factor-Augmented Regressions
Jan J. J. Groen and George Kapetanios
Federal Reserve Bank of New York Staff Reports, no. 363
February 2009; revised May 2012
JEL classification: C22, C52, E37

Abstract

Factor-augmented regressions are often used as a parsimonious way of modeling a variable using information from a large data set, through a few factors estimated from this data set. But how does one determine which factors are relevant for such a regression? Existing work has focused on criteria that can consistently estimate the appropriate number of factors in a large-dimensional panel of explanatory variables. However, these are not necessarily all relevant for modeling a specific dependent variable within a factor-augmented regression. This paper develops a number of theoretical conditions selection criteria have to fulfill in order to provide a consistent estimate of the factor dimension that is relevant for such a regression. Our framework takes into account factor estimation error, and it does not depend on a specific factor estimation methodology. Our conditions indicate that standard model selection criteria, such as BIC, are not consistent for factor-augmented regressions, but they can be once we modify these such that the corresponding penalty function for dimensionality also penalizes factor estimation error. We show through Monte Carlo and empirical applications that these modified information criteria are useful in determining appropriate factor-augmented regressions.

Key words: factor models, information criteria, macroeconomic forecasting

---

Groen: Federal Reserve Bank of New York (e-mail: jan.groen@ny.frb.org).
Kapetanios: Queen Mary University of London (e-mail: g.kapetanios@qmul.ac.uk). The authors thank two anonymous referees as well as Anindya Banerjee (the editor) for helpful comments. They also thank Craig Kennedy for excellent research assistance.

1 Introduction

When forecasting an economic variable, it is often necessary to incorporate information from a large set of potential explanatory variables into the forecasting model. Most traditional macroeconomic prediction approaches, however, are unable to deal with this, either because it is inefficient or downright impossible to incorporate a large number of variables in a single forecasting model and estimate it using standard econometric techniques. As an alternative approach to this problem factor-augmented regressions have gained a prominent place. A seminal application is Stock and Watson (2002b), where a limited number of principal components extracted from a large data set are added to a standard linear regression model which then is used to forecast key macroeconomic variables. Stock and Watson (2002a) and Bai (2003) formalized the underlying asymptotic theory, which allows the use of principal components in very large data sets to identify the common factors in such a data set.

Dynamic factor research in econometrics has spend substantial effort on developing tests and selection criteria aimed at determining that number of factors that describes best the dynamics in a large data set of explanatory variables. A well-known contribution is Bai and Ng (2002), who derive a range of consistent information criteria that can be used to identify the common factor space underlying a large panel of predictor series. While the number of factors selected in such a way provides an upper bound for the number of factors that should enter the forecasting regression for a particular variable, there is no *a priori* reason to suppose that all factors should enter this regression. Therefore, it is of importance that a form of factor selection is carried out that is tailored at determining a factor-based forecasting model for a specific variable. This problem has received far less attention in the literature than the aforementioned issue of determining the number of factors that best explains the dynamics in large data sets of explanatory variables.

One further important reason for considering this problem has to do with the well known evidence (see, e.g., Kapetanios (2010)) that determining the number of factors in large datasets is a difficult undertaking. As a result, the performance of existing methods suffer considerably under a variety of circumstances. On the other hand, determining the identity and number of variables in a regression, through information criteria, is a well understood problem. Further, such information criteria have desirable properties both asymptotically and in finite samples. Therefore, it seems reasonable to try and use such criteria for the problem at hand, even if all factors in a large dataset appear in the regression under consideration.

Intuitively, since the aim is to specify a regression model for a single variable, standard information criteria may be considered useful in selecting the optimal number of factors for a particular forecasting regression. However, factor variables are not observed and as a result this estimation error may matter and make standard information criteria invalid. Stock and Watson (1998) make this point and propose a selection criterion that takes into account this estimation error. However, their criteria do not take into account the sharper asymptotic analysis of Bai and Ng (2006) and, therefore, the form of the penalty term they propose and the conditions under which it is valid, can be improved upon. Building on Bai and Ng (2006), Bai and Ng (2009) propose a final prediction error (FPE) criterion in which an extra penalty term is added to proxy for the effect of factor estimation error on the forecasting regression. Optimizing this FPE will yield the number of factors that asymptotically minimizes the prediction error, but it does not necessarily provide an asymptotically consistent estimate of the number of factors present in the regression of interest. Also, the finite sample performance of this FPE criterion depends on the choice of a consistent estimator of the factor estimation error variance. Alternatively, one can follow Bai and Ng (2008) and select a subgroup of predictors from the overall macro panel with the best fit for the target variable, based on some threshold rule, and subsequently apply principal components on these 'targeted predictors' in order to get the most relevant factors for forecasting. In this paper, we rather focus, like Stock and Watson (1998) and Bai and Ng (2009), on the construction of appropriate selection criteria that can provide the econometrician with the optimal factor-augmented regression.

We propose a number of novel insights with respect to this issue of determining the relevant factors for a specific factor-augmented regression. Firstly, we show that standard information criteria are inconsistent estimators of the true dimension of the relevant factor space, in particular when the time series dimension of the underlying panel of predictor variables grows slower than its cross-section dimension. Next, we suggest alternative criteria that are consistent estimators in all cases - essentially we build on existing consistent information criteria for time series analysis and modify them to take into account the effects of factor estimation error. Further, we generalize our analysis to factor estimation methods other than principal components. Both Monte Carlo and empirical exercises show the relevance and added value of our proposed framework.

The paper is structured as follows. In Section 2, we present our setup and theoretical results. Section 3 reports on a detailed Monte Carlo study of our new selection criteria in comparison with existing ones. Section 4 presents an empirical forecasting application and Section 5 concludes.

2 Theory

We focus on a single variable y_t that we wish to model using an N-dimensional set of variables x_t and the latter is assumed to have a factor structure. In particular, we posit the following model for x_t:

$$x_t = \Lambda f_t + u_t, \quad t = 1, \dots, T \tag{1}$$

where $f_t = (f_{1t}, ..., f_{rt})'$ is an $r \times 1$ vector of factor variables such that $r << N$ and u_t is an $N \times 1$ vector of zero-mean errors. The factors f_t are not observed and need to be estimated from the $N \times 1$ data vector, x_t. Let the forecasting equation for y_t, be specified as

$$y_t = f_t^{0\prime} \beta + e_t \tag{2}$$

where f_t^0 is an $r_0 \times 1$ vector of factor variables that is possibly a subset of f_t, i.e., $1 \leq r_0 \leq r$. Finally, e_t is an error term with finite variance σ^2. The aim of our work is to provide information criteria for selecting the appropriate set of factors that should be entered in (2). There has been a considerable amount of work on determining r, which is the true number of factors needed to explain x_t (see, e.g., Bai and Ng (2002)). Our focus is different in the sense that not all factors underlying x_t may be relevant for modeling y_t. It is clear that standard information criteria may be of use in specifying (2), but care needs to be taken given that f_t are not observed and must be estimated from x_t.

Now let us consider the class of information criteria (IC) given by

$$IC = \frac{T}{2} \ln\{\hat{\sigma}_{\hat{e}}^2\} + C_{T,N} \tag{3}$$

where $\hat{\sigma}_{\hat{e}}^2$ denotes the estimated residual variance from the regression

$$y_t = \hat{f}_t' \beta + \hat{e}_t \tag{4}$$

and $\hat{f}_t$ denotes some subset of the estimated factor set obtained by applying principal components (PC) to (1). We further specify that $C_{T,N} = i\tilde{C}_{T,N}$ where i denotes the dimension of the candidate set of factors to be entered in (2) and $\tilde{C}_{T,N}$ denotes a penalty term that depends solely on T and N. This class includes all popular IC such as the Akaike (1974) IC (AIC), the Bayesian IC (BIC - see Schwarz (1978)) and the IC proposed by Hannan and Quinn (1979) (HQIC). It is important to realise that our search is not just over the number of factors to include in the forecasting regression, but most importantly over the identity of factors. To clarify this we discuss the search space in more detail. Let

$$\hat{\mathcal{F}} = \left\{ \left\{ \hat{f}_t^{(1)} \right\}_{t=1}^{T}, \left\{ \hat{f}_t^{(2)} \right\}_{t=1}^{T} \dots, \left\{ \hat{f}_t^{(r)} \right\}_{t=1}^{T} \right\}$$

denote the set of estimated factor variables over which the information criterion search is carried over, where $\hat{F}^{(i)} = \left(\hat{f}_1^{(i)}, \hat{f}_2^{(i)}, ..., \hat{f}_T^{(i)} \right)'$ and $\hat{f}_t^{(i)}$ denotes the i-th candidate vector of estimated factors at time t. $s \leq 2^r$, since there are 2^r distinct forecasting models that can be constructed from the set of r estimated factors. It is possible that either all 2^r combinations are considered or if the number of combinations is too large a subset of them is considered.[1] The cardinality of the subset is denoted by s. Further, let

$$\mathcal{F} = \left\{ \left\{ f_t^{(1)} \right\}_{t=1}^{\infty}, \left\{ f_t^{(2)} \right\}_{t=1}^{\infty} \dots, \left\{ f_t^{(s)} \right\}_{t=1}^{\infty} \right\}$$

where $f_t^{(i)}$denotes the probability limit of $\hat{f}_t^{(i)}$ as $N, T \to \infty$ for all $i = 1, ..., s$ and $t = 1, ..., t$. Also, denote the vector of true factors entering (2) at time t, by f_t^0. Finally, denote the information criterion associated with the candidate set of factors $\left\{ \hat{f}_t^{(i)} \right\}_{t=1}^{T}$ by $\widehat{IC}^{(i)}$ and the dimension of the factor vector $\hat{f}_t^{(i)}$ by $r^{(i)}$. Then, we have the following theorem concerning the consistency of factor selection using IC of the above form.

Theorem 1 *Let Assumptions A-E of Bai and Ng (2006) hold. Let* $\tilde{C}_{T,N} = o(T)$ *and* $\lim_{N,T\to\infty} T^{-1} \min(N, T)\tilde{C}_{T,N} = \infty$. *Then, for all* $i, j = 1, ..., s$, *we have: (i) If there exists a matrix* A *such that* $f_t^0 = Af_t^{(i)}$, $\forall t$, *but no such matrix exists for* $f_t^{(j)}$, *then*

$$\lim_{N,T\to\infty} \Pr \left(\widehat{IC}^{(i)} < \widehat{IC}^{(j)} \right) = 1. \tag{5}$$

(ii) If there exist matrices $A^{(i)}$ *and* $A^{(j)}$ *such that* $f_t^0 = A^{(i)} f_t^{(i)}$, *and* $f_t^0 = A^{(j)} f_t^{(j)}$, $\forall t$, *and* $r^{(i)} < r^{(j)}$ *then (5) holds.*

Proof: *See Appendix A for details on the proof of this theorem.*

Remark 1 *Theorem 1 is a factor selection consistency result. The Theorem gives two results. The first is that any set of estimated factors whose probability limits span the true factors entering*

[1] Usually, empirical researchers extract a small number of factors from large datasets allowing the consideration of all possible combinations. If r is large, say above 20, the calculation of information criterion values for all combinations becomes impractical. Then, one may wish to use non-standard search methods that minimise information criteria without considering all combinations. A discussion of such methods is provided in Kapetanios (2007).

(2), will be chosen over any set of estimated factors that do not span the true factors as long as the penalty term is $o(T)$. The second and nonstandard result is that if two sets of estimated factors both span the set of true factors the one with the smaller dimension will be chosen, as long as the penalty term is of higher order than $T^{-1}\min(N,T)$. Note that f_t^0 may contain all factors that explain the whole data set x_t but this is by no means the only possible case, implying that the problem of determining the factors entering (2) is distinct from the problem of determining the factors entering (1). As such this problem is worthy of consideration on its own merit. Further, it is worth noting that many authors have tried to provide some further identification of the factors. For example, data sets comprising of domestic or international data have been used to construct separate national and international factors such as in Monacelli and Sala (2009). Deciding which factors, from such different sets of factors, can be included in a factor-augmented regression cannot be done using criteria that operate on the large data sets. Our proposed criteria are the only means of solving such problems.

Theorem 1 relates explicitly to factor estimates obtained by static PC and although this is the most widely used method for estimating factors, there exist a variety of other estimation methods. For example, we have dynamic principal components as suggested in, e.g., Forni *et al.* (2000), there are methods based on estimation of state space factor models (Doz *et al.* (2006) and Kapetanios and Marcellino (2009)) or one can follow Groen and Kapetanios (2008) and use partial least squares to directly estimate the factors relevant for a specific dependent variable. These methods may have different consistency rates both for the factor estimates and the coefficients entering (2). It is therefore useful to generalize our consistency result to cover cases where factors are estimated by some other method. As we wish our result to be general we make the following high level assumption where $\hat{\sigma}_e^2$ denotes the residual variance from (2):

Assumption 1 *$\hat{\sigma}_e^2 - \hat{\sigma}_{\hat{e}}^2 = O_p(q_{NT}^{-1})$ where $q_{NT} \to 0$.*

Theorem 2 spells out the generalization of the consistency result in Theorem 1.

Theorem 2 *Let Assumption 1 and Assumption A of Bai and Ng (2006) hold. Further, assume that e_t in (2) has finite variance and satisfies a law of large numbers. Let $\tilde{C}_{T,N} = o(T)$ and $\lim_{N,T\to\infty} T^{-1} q_{NT}\tilde{C}_{T,N} = \infty$. Then, Theorem 1 holds.*

Proof: *The proof of Theorem 2 is straightforward and given in Appendix A.*

Note that Remark 1 holds for Theorem 2 as well.

It is easy to see that within the context of regression (2) modified versions of the BIC and HQIC given by

$$\begin{aligned} BICM &= \frac{T}{2}\ln\{\hat{\sigma}_{\hat{e}}^2\} + i\ln(T)\left(1+\frac{T}{N}\right), \\ HQICM &= \frac{T}{2}\ln\{\hat{\sigma}_{\hat{e}}^2\} + 2i\ln\ln(T)\left(1+\frac{T}{N}\right), \end{aligned} \tag{6}$$

with $1 \le i \le r$, fulfill the conditions of Theorem 1. Unlike the case where criteria are used to determine the number of factors in a large dataset where both N and T are large, the context here is that of a forecasting regression where N and T have very different functions.

This is a forecasting regression with a few regressors which happen to be unobserved factors, and it is because of this that the observed N is relevant and therefore our criteria in (6) are asymmetric in T and N. Suitably defined penalty terms, that are expressed in terms of q_{NT}, can be straightforwardly specified to produce information criteria that satisfy the conditions of Theorem 2. We do not report these as our main focus, in the rest of the paper, is criteria that relate to Theorem 1.

Of course, other variables can enter the regression and one can also envisage other types of selection. The most obvious one is lag selection where lags of y_t or possibly other variables enter the regression and

the number of lags needs to be selected. Given that the conditions of the Theorems 1 and 2 imply that the relevant IC will be consistent also for lag selection, it is clear that such joint searches are feasible. Therefore, we can modify regression (2) such that

$$y_t = z_t'\gamma + f_t^{0'}\beta + e_t \tag{7}$$

with z_t is a $k \times 1$ vector of non-generated regressors and γ is the corresponding parameter vector; z_t can contain an intercept, lags of y_t and so on. The following versions of the modified criteria in (6) are valid for regression (7) under the framework spelled out in Theorems 1 and 2:

$$\begin{aligned} BICM &= \frac{T}{2}\ln\{\hat{\sigma}_{\hat{e}}^2\} + k\ln(T) + i\ln(T)\left(1 + \frac{T}{N}\right), \\ HQICM &= \frac{T}{2}\ln\{\hat{\sigma}_{\hat{e}}^2\} + 2k\ln\ln(T) + 2i\ln\ln(T)\left(1 + \frac{T}{N}\right). \end{aligned} \tag{8}$$

Hence, searching for the optimal values of the modified ICs in (8) will provide the econometrician with a consistent, simultaneous, estimate of the optimal values of k and i in regression (7).

3 Monte Carlo Analysis

In this section we carry out a Monte Carlo study of the new selection criteria for factor-augmented regressions suggested in Section 2. The set-up of the Monte Carlo experiments are spelled out in Section 3.1. In the experiments we compare our suggested criteria with existing ones and the results of this comparison are reported in Section 3.2.

3.1 Set-Up

Our Monte Carlo experiments are based on the following data generating processes (DGPs):

$$y_t = \alpha' x_t + \epsilon_t, \quad t = 1, ..., T, \tag{9}$$

$$x_t = c_r \Lambda' f_t + u_t, \quad c_r = \frac{1}{\sqrt{r}}, \tag{10}$$

$$\epsilon_t = \sqrt{cN}\varepsilon_t. \tag{11}$$

In (9)-(11), $x_t = (x_{1,t} \cdots x_{N,t})'$ is the $N \times 1$ vector of explanatory variables with corresponding $N \times 1$ vector of regression parameters $\alpha = (\alpha_1 \cdots \alpha_N)'$. The term ϵ_t is a zero-mean disturbance term, which we discuss in more detail below. The explanatory variables x_t in (10) are generated by r factors with a $N \times 1$ vector of zero-mean disturbances $u_t = (u_{1,t} \cdots u_{N,t})'$ and a $r \times N$ matrix of factor loadings $\Lambda = (\lambda_1 \cdots \lambda_N)$ that corresponds with the $r \times 1$ vector of factors $f_t = (f_{1,t} \cdots f_{r,t})'$ with $\lambda_i = (\lambda_{i,1} \cdots \lambda_{i,r})'$.

The individual regression coefficients in (10) are drawn from a normal distribution: $\alpha_i \sim \text{iid}N(1,1)$, whereas the disturbances for the N explanatory variables u_t in (11) are generated from a multivariate normal distribution that allows for some degree of cross-sectional dependence. To achieve this, spatial dependence is incorporated by setting $u_t = R\vartheta_t$, where $\vartheta_t = \sqrt{\sigma_\vartheta^2}\eta_t; \quad \eta_t \sim IIDN(0,I)$, $\sigma_\vartheta^2 = N/tr(RR')$, and $R = (I - dS)^{-1}$. We set $d = 0.2$ and define the $N \times N$ matrix

$$S = \begin{bmatrix} 0 & 1 & 0 & & 0 \\ 1/2 & 0 & 1/2 & & \\ \vdots & & \ddots & \ddots & \ddots \\ & & & & 1/2 \\ 0 & 0 & & 1 & 0 \end{bmatrix}. \tag{12}$$

The S matrix (12) implies that the elements of u_t have unit variance but are weakly cross-sectionally dependent.

The individual factor loadings in Λ in (11) are determined as $\lambda_{i,j} \sim \text{iid } \sqrt{12}U(0,1)$ and the r factors are each generated as $f_{j,t} \sim \text{iid}N(0,1)$. Note that the above implies that $\lambda_{i,j}$ has unit variance. Within this context it is important to ensure that the factor loadings have non zero mean as a zero mean compounded by random loadings implies a much weaker factor structure than otherwise would be the case. For example, a factor model with zero mean random loadings implies zero average correlations across $x_{i,t}$. Next, c_r is given by $\frac{1}{\sqrt{r}}$, and is chosen so to ensure that the R^2 of the y_t equation is constant as r increases. Finally, concerning the choice of r in our simulation experiments, we consider $r = 1, 2, 4, 6$.

The regression model for y_t in (9) has the alternative representation given by

$$y_t = c_r \alpha' \Lambda' f_t + \left(\alpha' u_t + \epsilon_t\right), \quad t = 1, ..., T \tag{13}$$

which warrants a factor-augmented type of regression model to explain the dynamics in y_t in a parsimonious manner. In each artificial sample generated from (9)-(11), we therefore select the optimal number of factors in a regression of y_t on factors that are estimated by applying PC on x_t. We do this for different selection criteria, both the ones we considered in Section 2, in particular (6) given the set-up in (9)-(11), as well as standard information criteria and the FPE criterion suggested in Bai and Ng (2009). Crucial for the Monte Carlo study is controlling the fit of the y_t regression equation. This is done by calibrating the variance of the disturbance term ϵ_t through c, where $\varepsilon_t \sim \text{iid}N(0,1)$. The calculation of this fit requires care for a variety of reasons. Firstly, we note that the relevant fit is not the one pertaining to (9) but to (13) since our regressors are the estimated factors and not x_t. Secondly, the fact the we use random loadings complicates matters. Our measure of fit, that resembles the usual R^2, is given by

$$1 - \frac{E\left(\alpha_i\right)^2 E\left(u_{i,t}\right)^2 + cE\left(\epsilon_t\right)^2}{E\left(\alpha_i\right)^2 E\left(u_{i,t}\right)^2 + cE\left(\epsilon_t\right)^2 + E\left(\alpha_i\right)^2 E\left(\lambda_{i,j}\right)^2 E\left(f_{j,t}\right)^2} \tag{14}$$

where the term $E\left(\alpha_i\right)^2 E\left(\lambda_{i,j}\right)^2 E\left(f_{j,t}\right)^2$ proxies the explained sum of squares, whereas $E\left(\alpha_i\right)^2 E\left(u_{i,t}\right)^2 + cE\left(\epsilon_t\right)^2$ proxies the residual sum of squares. The measure in (14) varies from 0 to 1, and setting $c = 0.545,\ 1,\ 3.5$ and 11 gives fit measures equal to $0.66,\ 0.50,\ 0.40$ and 0.20, respectively.

We generate data through (9) for N, $T = 20, 30, 50, 100, 200, 400$. The Monte Carlo experiments are based on 1000 replications and in each replication we set for each selection criterion the estimate of r equal to the number that minimizes the criterion across potential numbers of factors that range from 0 up to 8, inclusive. We report for each selection criterion under consideration the average number of factors selected across the replications, where the best performing criterion should on average select a number of factors that is close to the assumed factor order r in a particular Monte Carlo experiment. Note that unlike much previous work, we allow for the possibility that no factors should enter the regression and so we potentially allow for the value of $r = 0$ in the selection criterion search.

An important question relates to the space over which we search. Our theory focuses on the general question of which factors to use in the regression model for y_t. The main theoretical problem associated with existing information criteria is that they will choose redundant factors, as is clear from the proof of our Theorems. So, in general, they will over-parameterise the forecasting regression. The true factors are orthonormal in our Monte Carlo study and we impose the same restriction on the estimated factors. Therefore, selecting the appropriate factors in the above experiments boils down to selecting the optimal number of factors. However, we will also consider a related but distinct Monte Carlo setup where a subset of the factors enter the forecasting regression and we wish to determine which factors do so. The setup is the same as that given above apart from replacing (9) with

$$y_t = \alpha_1 f_{1t} + \epsilon_t, \quad t = 1, ..., T \tag{15}$$

where $\epsilon_t = \sqrt{0.545}\varepsilon_t$, implying an asymptotic fit (through (14)) of 0.66, α_1 and ε_t are specified as before, $r = 4$, and we consider all possible 16 combinations of the first four dominant principal components in our IC search. In this variant of our Monte Carlo set-up, the estimated factors are normalised so that they converge to the true factors as T and N increase. We evaluate the criteria by defining the true set of factors entering the forecasting regression as the 4×1 vector $\mathcal{I}^0 = (1, 0, 0, 0)'$ and similarly defining the vector of factors selected using some information criterion, which we denote by $\widehat{\mathcal{I}}$. Then, a measure of the performance of a certain information criterion is given by the mean of the squared deviations (MSD) of the IC-based selection vector of factors relative to the true selection vector of factors $\mathcal{I}^0$:

$$\text{MSD} = \frac{1}{B}\sum_{i=1}^{B}(\widehat{\mathcal{I}}_i - \mathcal{I}^0)'(\widehat{\mathcal{I}}_i - \mathcal{I}^0) \tag{16}$$

where B is the number of Monte Carlo replications, which we set, as before, to 1000.

3.2 Results

Apart from the modified information criteria (6) in Section 2, which are relevant given DGP (9) for the Monte Carlo experiments, we analyze in our study also the performance of standard information criteria. As AIC is known to be inconsistent, we will in particular focus on the performance of the BIC and HQIC criteria. Given that the extra penalty term in our criteria will be most relevant for cases were $T \geq N$, we expect specially for those cases to observe large differences for the criteria in (6) *vis-à-vis* BIC and HQIC.

Our modified information criteria are not the first set of selection criteria that are specifically developed to determine the dimensions of factor-augmented regressions, albeit that ours are the first *consistent* criteria to be proposed for this purpose. Bai and Ng (2009) suggest a forecast prediction error (FPE) criterion that in the limit minimizes the mean squared prediction error of a factor-augmented regression. This FPE criterion essentially entails adding a cross-sectional penalty factor, which depends on an estimate of the factor covariance matrix, to a standard information criterion. So using BIC the FPE for regression model (7) becomes

$$FPE = \ln\{\hat{\sigma}^2_{\hat{e}}\} + (k+i)\left(\frac{\ln(T)}{T}\right) + c_i\left(\frac{\ln(N)}{N}\right) \tag{17}$$

with

$$c_i = \frac{\hat{\beta}'\Sigma_i\hat{\beta}}{\hat{\sigma}^2_{\hat{e}}}$$

where Σ_r is a consistent estimator of the covariance matrix of the i factors included in (7).[2] We use this FPE criterion as a third alternative, next to BIC and HQIC, for our BICM and HQICM criteria, which we implement by setting $k = 0$ in (17).

The results of the experiments based on (9)-(11) are reported in Tables 1-3. When we first focus on the results for the standard information criteria in the upper panels of Tables 1 and 2, it becomes clear that these criteria overestimate the number of factors in a number of cases by a considerable margin. Particularly when the time series dimension T is larger than the cross-section dimension N of the regressor variable vector x_t, indicating the potential severity of the impact of factor estimation error variance in that case. Especially for cases when the true r is small, the results for both large T and N in Tables 1 and 2 suggest that BIC and HQIC are not able to provide consistent estimates of the optimal number of factors that underlie a factor-augmented regression. On the other hand our modified criteria, see the lower panels of Tables 1 and 2, seem to be behaving consistently, where the differences relative to the standard criteria are the most pronounced for smaller true numbers of factors and $T > N$. As in the standard time series case with non-generated regressors for BIC, our modified BIC criterion as a slight tendency to underestimate the true r which is much less so in case of our modification of the HQIC criterion. Not surprisingly, when the calibrated fit of (13) decreases in Tables 1-**??** all the corresponding criteria perform poorly and structurally underestimate the true factor order.

Next, the performance of the FPE criterion (17) is remarkably bad relative to both standard and modified BIC and HQIC criteria, as can be observed from Table 3, where in a significant number of replications it selects no factors at all across a variety of configurations. Finally, we consider the Monte Carlo experiment based on (15) and (10)-(11) where we wish to determine which factors enter the forecasting regression. We report the relative MSD for the standard *versus* modified BIC and HQIC criteria, so (16) of the modified criterion divided by that of the standard criterion, in Table 4. Again, we see that the modified criteria obtain relative MSDs that are considerably below one, in the majority of cases, indicating their superior performance.

[2] There are a variety of estimators possible for Σ_r, and the choice of such an estimator impacts the finite sample behavior of (17). We choose to apply a HAC-consistent covariance matrix estimator on the r (i.e. the total number factors driving the dynamics in x_t) estimated factors to proxy Σ_r - this makes sense as each factor is a linear combination of the individual predictor series whose dynamics is not explicitly modeled. Also, we use a HAC estimator for $\hat{\sigma}^2_{\hat{e}}$ in c_r of (17) when $h > 1$. In particular, we found that both in the Monte Carlo and the empirical applications using the Den Haan and Levin (1997) VAR-HAC estimator based on BIC lag selection resulted in the most accurate performance of the FPE criterion.

Table 1: Monte Carlo results for standard and modified BIC

		$r=1$					$r=2$					$r=4$					$r=6$				
		Standard BIC																			
Fit	N / T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
	30	3.35	3.54	4.52	5.65	6.50	3.59	3.81	4.72	5.62	6.53	4.17	4.55	5.31	6.10	6.79	4.53	4.88	5.78	6.42	6.93
	50	2.53	2.60	3.18	4.46	5.48	3.10	3.16	3.74	4.63	5.66	3.78	4.06	4.56	5.34	6.00	4.08	4.44	5.27	6.04	6.65
0.66	100	2.06	2.00	2.28	2.85	3.90	2.56	2.57	2.81	3.35	4.20	3.44	3.62	4.01	4.51	5.25	3.91	4.29	4.97	5.58	6.25
	200	1.78	1.58	1.52	1.94	2.48	2.29	2.26	2.26	2.63	3.08	3.36	3.41	3.59	4.07	4.54	3.91	4.17	4.88	5.44	5.98
	400	1.49	1.38	1.33	1.43	1.68	2.14	2.06	2.05	2.16	2.38	3.38	3.42	3.55	3.85	4.09	4.15	4.25	4.84	5.26	5.71
	30	2.77	2.89	3.81	4.80	5.99	3.20	3.37	3.98	4.85	6.07	3.62	3.82	4.65	5.62	6.33	3.73	4.35	5.15	5.98	6.60
	50	2.22	2.27	2.76	3.62	4.77	2.63	2.75	3.09	3.98	5.02	3.36	3.56	3.94	4.86	5.58	3.54	4.01	4.79	5.50	6.38
0.50	100	1.84	1.72	1.84	2.39	3.26	2.42	2.23	2.38	2.84	3.60	3.04	3.21	3.59	4.20	4.75	3.39	3.68	4.45	5.20	5.97
	200	1.64	1.44	1.45	1.64	2.13	2.13	2.05	2.09	2.21	2.66	3.12	3.14	3.40	3.81	4.19	3.50	3.80	4.46	5.04	5.62
	400	1.42	1.35	1.27	1.25	1.44	2.08	1.96	1.92	2.02	2.18	3.24	3.16	3.36	3.60	3.85	3.74	3.89	4.45	5.07	5.53
	30	1.74	1.66	1.94	2.47	3.69	1.94	2.12	2.34	2.74	3.94	2.33	2.49	2.96	3.66	4.77	2.46	2.53	3.33	4.26	5.38
	50	1.66	1.45	1.50	1.86	2.59	1.90	1.82	1.92	2.36	2.91	2.39	2.44	2.71	3.28	4.13	2.23	2.57	3.24	4.06	5.05
0.40	100	1.44	1.37	1.28	1.42	1.67	1.86	1.73	1.78	1.96	2.23	2.31	2.26	2.67	3.09	3.60	2.30	2.47	2.98	3.87	4.77
	200	1.42	1.24	1.21	1.19	1.29	1.80	1.69	1.73	1.84	1.95	2.29	2.39	2.68	2.97	3.40	2.36	2.57	3.12	3.96	4.78
	400	1.36	1.17	1.13	1.11	1.10	1.86	1.74	1.71	1.80	1.82	2.55	2.41	2.67	2.95	3.32	2.65	2.65	3.26	3.93	4.67
	30	1.35	1.25	1.25	1.35	1.62	1.58	1.47	1.54	1.68	2.12	1.69	1.51	1.85	2.23	2.91	1.64	1.59	1.84	2.45	3.27
	50	1.34	1.18	1.20	1.17	1.33	1.53	1.45	1.50	1.63	1.77	1.63	1.63	1.76	2.20	2.58	1.67	1.49	1.83	2.25	3.09
0.20	100	1.32	1.18	1.10	1.10	1.12	1.54	1.44	1.48	1.57	1.66	1.63	1.65	1.80	2.10	2.59	1.65	1.57	1.76	2.26	3.04
	200	1.24	1.11	1.10	1.09	1.06	1.60	1.47	1.46	1.55	1.62	1.68	1.61	1.81	2.13	2.58	1.72	1.62	1.74	2.35	3.15
	400	1.30	1.12	1.09	1.05	1.03	1.70	1.52	1.50	1.56	1.63	1.84	1.76	1.87	2.23	2.63	1.80	1.69	1.79	2.29	3.21
		Modified BIC																			
	30	1.42	1.18	1.08	1.03	1.01	1.73	1.60	1.52	1.48	1.43	2.06	2.00	1.89	1.78	1.72	2.19	1.89	1.86	1.68	1.56
	50	1.43	1.22	1.13	1.06	1.02	1.96	1.75	1.64	1.60	1.52	2.48	2.38	2.29	2.25	2.13	2.57	2.37	2.41	2.25	2.15
0.66	100	1.50	1.35	1.17	1.08	1.03	2.03	1.88	1.76	1.72	1.62	2.83	2.72	2.79	2.81	2.77	3.13	3.06	3.24	3.44	3.48
	200	1.55	1.28	1.15	1.13	1.04	2.00	1.93	1.84	1.79	1.76	3.08	3.01	3.08	3.14	3.19	3.50	3.65	4.03	4.29	4.46
	400	1.38	1.25	1.14	1.13	1.07	2.05	1.97	1.88	1.83	1.87	3.27	3.25	3.28	3.44	3.45	3.96	3.98	4.44	4.74	5.01
	30	1.24	1.14	1.07	1.02	1.01	1.61	1.49	1.38	1.40	1.35	1.83	1.67	1.64	1.53	1.47	1.82	1.70	1.50	1.44	1.29
	50	1.28	1.17	1.08	1.03	1.01	1.73	1.61	1.52	1.51	1.46	2.17	2.10	1.93	2.02	1.95	2.12	2.03	1.99	1.92	1.80
0.50	100	1.31	1.19	1.09	1.04	1.02	1.92	1.73	1.66	1.65	1.61	2.50	2.45	2.48	2.56	2.60	2.62	2.58	2.86	2.81	3.02
	200	1.42	1.22	1.10	1.06	1.05	1.95	1.82	1.76	1.70	1.69	2.81	2.71	2.84	2.99	3.03	3.06	3.21	3.53	3.75	3.99
	400	1.35	1.21	1.13	1.08	1.04	2.01	1.85	1.75	1.78	1.80	3.11	3.01	3.06	3.28	3.33	3.53	3.57	4.02	4.36	4.75
	30	1.08	1.03	1.00	1.00	1.00	1.26	1.23	1.21	1.15	1.18	1.33	1.27	1.18	1.16	1.12	1.29	1.20	1.12	1.10	1.06
	50	1.11	1.05	1.01	1.00	1.00	1.39	1.32	1.30	1.34	1.28	1.52	1.47	1.35	1.35	1.29	1.46	1.36	1.27	1.20	1.19
0.40	100	1.18	1.09	1.02	1.01	1.00	1.55	1.48	1.41	1.43	1.45	1.83	1.77	1.77	1.70	1.78	1.79	1.63	1.62	1.64	1.66
	200	1.26	1.13	1.04	1.02	1.01	1.63	1.55	1.53	1.55	1.58	2.04	2.06	2.16	2.20	2.30	2.05	2.07	2.18	2.42	2.55
	400	1.27	1.12	1.05	1.03	1.01	1.81	1.66	1.61	1.65	1.63	2.43	2.23	2.40	2.50	2.82	2.45	2.37	2.69	3.06	3.52
	30	1.04	1.00	1.00	1.00	1.00	1.12	1.10	1.05	1.04	1.03	1.11	1.05	1.03	1.02	1.01	1.07	1.06	1.01	1.01	1.00
	50	1.08	1.01	1.00	1.00	1.00	1.21	1.14	1.11	1.09	1.10	1.18	1.14	1.07	1.05	1.04	1.19	1.09	1.05	1.03	1.02
0.20	100	1.14	1.04	1.00	1.00	1.00	1.36	1.24	1.20	1.22	1.18	1.33	1.25	1.23	1.19	1.17	1.32	1.21	1.14	1.11	1.09
	200	1.13	1.05	1.02	1.00	1.00	1.43	1.36	1.33	1.33	1.33	1.50	1.39	1.43	1.40	1.47	1.50	1.33	1.26	1.33	1.35
	400	1.23	1.07	1.04	1.01	1.00	1.61	1.45	1.42	1.42	1.46	1.70	1.65	1.63	1.78	1.89	1.67	1.52	1.51	1.59	1.75

Notes: The entries are the average number of factors selected based on the BIC criterion that is minimized for a range of 0 to 8 factors across 1,000 Monte Carlo replications. The variables are generated through the DGPs in (9)-(11) for different values of the time series dimension of all series, T, and the cross-section dimension, N, of series in x_t. We also impose different levels for the asymptotic fit of the prediction regression through (14), symbolized by the column titled 'Fit'.

Table 2: Monte Carlo results for standard and modified HQIC

		$r = 1$					$r = 2$					$r = 4$					$r = 6$				
		Standard HQIC																			
Fit	N / T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
	30	4.62	4.92	5.76	6.64	7.13	4.76	5.08	5.89	6.51	7.15	5.21	5.46	6.16	6.81	7.24	5.44	5.84	6.54	7.01	7.32
	50	3.85	4.03	4.74	5.95	6.54	4.21	4.41	5.09	5.89	6.68	4.94	5.13	5.55	6.20	6.78	5.19	5.46	6.12	6.67	7.07
0.66	100	3.12	3.08	3.65	4.49	5.53	3.58	3.53	4.09	4.77	5.64	4.41	4.43	4.90	5.59	6.22	4.91	5.15	5.74	6.29	6.75
	200	2.71	2.50	2.59	3.37	4.36	3.19	3.04	3.25	3.78	4.69	4.20	4.26	4.43	4.88	5.47	4.84	5.11	5.59	6.08	6.49
	400	2.16	2.11	2.16	2.43	3.15	2.91	2.75	2.70	3.04	3.62	4.03	4.02	4.17	4.61	4.99	4.92	5.09	5.52	5.83	6.25
	30	3.96	4.14	5.25	6.13	6.83	4.23	4.53	5.28	6.05	6.93	4.78	4.93	5.69	6.41	6.99	5.01	5.31	6.08	6.71	7.12
	50	3.29	3.43	4.18	5.18	6.10	3.74	3.95	4.36	5.29	6.27	4.42	4.53	4.99	5.87	6.45	4.62	5.00	5.68	6.30	6.96
0.50	100	2.80	2.76	3.04	4.04	5.00	3.29	3.00	3.46	4.25	5.13	3.97	4.09	4.47	5.14	5.71	4.39	4.69	5.36	6.08	6.51
	200	2.48	2.13	2.48	2.76	3.73	2.92	2.69	2.81	3.21	4.06	3.92	3.94	4.17	4.67	5.13	4.47	4.66	5.24	5.81	6.23
	400	2.15	1.90	1.90	2.02	2.56	2.79	2.58	2.54	2.70	3.19	3.89	3.81	3.89	4.24	4.64	4.47	4.80	5.23	5.71	6.06
	30	2.65	2.61	3.08	4.14	5.53	2.86	2.92	3.47	4.34	5.58	3.34	3.44	4.07	4.95	5.92	3.47	3.61	4.46	5.45	6.31
	50	2.41	2.17	2.47	3.14	4.43	2.81	2.63	2.80	3.57	4.43	3.23	3.27	3.69	4.41	5.28	3.31	3.52	4.36	5.16	5.95
0.40	100	2.09	1.98	1.97	2.26	3.07	2.58	2.31	2.40	2.84	3.57	3.11	3.01	3.43	3.97	4.60	3.31	3.52	4.08	4.87	5.67
	200	2.02	1.83	1.62	1.80	2.12	2.39	2.23	2.24	2.42	2.73	3.07	3.12	3.36	3.67	4.12	3.33	3.61	4.08	4.91	5.55
	400	1.90	1.65	1.48	1.45	1.53	2.42	2.15	2.03	2.15	2.29	3.27	3.13	3.20	3.56	3.83	3.58	3.61	4.26	4.84	5.38
	30	2.02	1.81	1.80	2.06	2.93	2.21	1.98	2.19	2.48	3.33	2.42	2.24	2.61	3.10	4.04	2.39	2.30	2.81	3.59	4.63
	50	2.06	1.62	1.59	1.73	2.29	2.15	1.90	1.98	2.21	2.55	2.31	2.27	2.42	2.96	3.52	2.40	2.17	2.65	3.30	4.36
0.20	100	1.87	1.61	1.38	1.38	1.56	2.19	1.90	1.84	1.95	2.19	2.29	2.20	2.42	2.83	3.38	2.43	2.24	2.65	3.29	4.04
	200	1.76	1.44	1.34	1.33	1.39	2.13	1.85	1.75	1.85	2.02	2.32	2.27	2.36	2.83	3.19	2.43	2.26	2.57	3.40	4.24
	400	1.80	1.43	1.35	1.22	1.27	2.17	1.91	1.76	1.85	1.92	2.47	2.38	2.43	2.76	3.19	2.58	2.33	2.66	3.39	4.29
		Modified HQIC																			
	30	2.07	1.57	1.35	1.22	1.13	2.43	2.08	1.84	1.71	1.62	2.97	2.76	2.49	2.44	2.33	3.15	2.79	2.78	2.63	2.56
	50	2.07	1.72	1.50	1.27	1.15	2.64	2.29	2.03	1.86	1.74	3.31	3.13	2.95	2.88	2.82	3.63	3.43	3.53	3.37	3.44
0.66	100	2.21	1.89	1.57	1.36	1.25	2.74	2.48	2.19	2.05	1.89	3.62	3.51	3.43	3.40	3.26	4.14	4.12	4.30	4.39	4.56
	200	2.27	1.85	1.53	1.51	1.34	2.79	2.54	2.26	2.16	2.08	3.84	3.71	3.60	3.67	3.67	4.48	4.53	4.88	5.11	5.18
	400	2.01	1.79	1.59	1.50	1.40	2.70	2.43	2.30	2.22	2.14	3.90	3.77	3.78	3.93	3.91	4.76	4.89	5.14	5.32	5.56
	30	1.71	1.41	1.21	1.11	1.06	2.11	1.84	1.61	1.58	1.51	2.52	2.17	2.13	2.07	2.00	2.59	2.42	2.23	2.09	1.99
	50	1.83	1.50	1.28	1.13	1.06	2.30	2.03	1.79	1.69	1.63	2.92	2.75	2.45	2.66	2.57	3.03	2.96	2.90	2.90	2.92
0.50	100	1.99	1.61	1.38	1.24	1.13	2.56	2.15	1.99	1.85	1.80	3.18	3.08	3.10	3.14	3.09	3.61	3.54	3.86	3.94	4.10
	200	2.07	1.66	1.46	1.33	1.22	2.57	2.24	2.10	1.92	1.90	3.59	3.43	3.40	3.46	3.45	4.05	4.14	4.48	4.66	4.86
	400	1.94	1.71	1.46	1.30	1.26	2.61	2.33	2.13	2.05	2.01	3.75	3.61	3.58	3.65	3.68	4.32	4.55	4.81	5.15	5.33
	30	1.25	1.08	1.03	1.01	1.00	1.43	1.36	1.35	1.27	1.30	1.68	1.52	1.45	1.41	1.38	1.70	1.44	1.34	1.31	1.24
	50	1.42	1.15	1.06	1.02	1.01	1.68	1.51	1.45	1.46	1.40	2.07	1.83	1.71	1.73	1.72	1.96	1.82	1.68	1.61	1.59
0.40	100	1.55	1.33	1.11	1.05	1.02	1.95	1.69	1.57	1.57	1.58	2.45	2.20	2.29	2.26	2.36	2.45	2.36	2.31	2.46	2.60
	200	1.70	1.40	1.21	1.10	1.04	2.08	1.83	1.73	1.70	1.70	2.71	2.65	2.70	2.71	2.83	2.91	2.90	3.14	3.44	3.66
	400	1.75	1.41	1.27	1.14	1.06	2.24	1.97	1.82	1.82	1.75	3.05	2.83	2.87	3.04	3.19	3.37	3.33	3.73	4.04	4.39
	30	1.12	1.02	1.00	1.00	1.00	1.23	1.18	1.12	1.10	1.09	1.28	1.12	1.08	1.08	1.06	1.24	1.13	1.05	1.04	1.03
	50	1.22	1.06	1.01	1.00	1.00	1.39	1.25	1.22	1.21	1.18	1.43	1.30	1.19	1.19	1.17	1.47	1.22	1.16	1.12	1.08
0.20	100	1.40	1.15	1.03	1.01	1.00	1.62	1.42	1.34	1.35	1.31	1.71	1.58	1.51	1.50	1.48	1.72	1.51	1.41	1.36	1.35
	200	1.51	1.19	1.10	1.04	1.00	1.85	1.58	1.46	1.45	1.45	2.00	1.83	1.81	1.86	1.94	2.13	1.82	1.75	1.96	1.98
	400	1.65	1.32	1.18	1.07	1.01	2.03	1.73	1.59	1.57	1.57	2.32	2.14	2.10	2.30	2.42	2.38	2.08	2.21	2.44	2.77

Notes: See the notes for Table 1, albeit that factor selection is now based on the HQIC criterion.

Table 3: Monte Carlo results for the Bai-Ng FPE criterion

		$r=1$					$r=2$					$r=4$					$r=6$				
Fit	N / T	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400	30	50	100	200	400
	30	0.57	0.32	0.26	0.23	0.16	0.59	0.36	0.30	0.28	0.24	0.67	0.46	0.38	0.32	0.30	0.71	0.46	0.39	0.36	0.33
	50	0.52	0.29	0.18	0.15	0.14	0.63	0.41	0.29	0.29	0.24	0.70	0.41	0.35	0.32	0.25	0.71	0.43	0.36	0.31	0.29
0.66	100	0.62	0.31	0.16	0.08	0.07	0.58	0.38	0.28	0.20	0.18	0.67	0.44	0.36	0.27	0.26	0.75	0.50	0.33	0.30	0.24
	200	0.53	0.22	0.11	0.04	0.02	0.71	0.39	0.28	0.17	0.16	0.72	0.47	0.30	0.26	0.22	0.85	0.47	0.37	0.30	0.24
	400	0.57	0.24	0.06	0.02	0.01	0.68	0.41	0.29	0.21	0.16	0.76	0.57	0.35	0.26	0.23	0.82	0.50	0.37	0.29	0.22
	30	0.73	0.46	0.37	0.32	0.34	0.71	0.53	0.44	0.42	0.38	0.84	0.55	0.50	0.47	0.42	0.86	0.60	0.49	0.43	0.45
	50	0.81	0.45	0.33	0.30	0.27	0.84	0.56	0.44	0.38	0.37	0.83	0.57	0.48	0.47	0.40	0.83	0.63	0.53	0.47	0.44
0.50	100	0.70	0.37	0.24	0.22	0.19	0.73	0.50	0.42	0.35	0.31	0.89	0.59	0.45	0.44	0.43	0.95	0.60	0.52	0.49	0.45
	200	0.83	0.40	0.22	0.14	0.13	0.81	0.54	0.41	0.32	0.30	0.95	0.61	0.50	0.44	0.38	1.09	0.68	0.54	0.49	0.45
	400	0.85	0.31	0.23	0.10	0.06	0.82	0.55	0.46	0.32	0.29	0.92	0.70	0.52	0.46	0.42	1.02	0.67	0.60	0.53	0.47
	30	1.25	0.98	0.92	0.87	0.91	1.32	1.04	0.96	0.92	0.93	1.34	1.08	1.00	1.00	0.97	1.35	1.04	1.00	1.00	0.99
	50	1.29	1.00	0.96	0.93	0.96	1.31	1.10	0.98	1.00	0.98	1.38	1.10	1.04	1.01	1.01	1.39	1.10	1.03	1.00	1.01
0.40	100	1.42	1.03	0.96	0.97	0.99	1.37	1.05	1.00	1.00	1.00	1.44	1.12	1.03	1.01	1.00	1.53	1.10	1.02	1.05	1.04
	200	1.47	1.04	1.01	1.04	0.99	1.39	1.05	1.05	1.06	1.02	1.52	1.14	1.07	1.04	1.02	1.61	1.16	1.05	1.06	1.02
	400	1.52	1.07	1.07	1.02	1.02	1.49	1.08	1.07	1.02	1.03	1.53	1.13	1.06	1.04	1.03	1.67	1.14	1.07	1.07	1.02
	30	1.48	1.13	1.06	1.08	1.05	1.49	1.21	1.15	1.13	1.18	1.52	1.19	1.16	1.14	1.26	1.55	1.16	1.13	1.13	1.20
	50	1.50	1.19	1.05	1.10	1.04	1.61	1.18	1.10	1.12	1.15	1.60	1.23	1.08	1.16	1.12	1.62	1.21	1.11	1.11	1.12
0.20	100	1.58	1.25	1.10	1.05	1.02	1.54	1.19	1.11	1.07	1.06	1.69	1.30	1.11	1.09	1.07	1.75	1.22	1.09	1.08	1.04
	200	1.73	1.23	1.07	1.04	1.06	1.75	1.19	1.10	1.04	1.04	1.70	1.26	1.11	1.05	1.02	1.81	1.23	1.12	1.05	1.05
	400	1.71	1.16	1.11	1.04	1.02	1.69	1.24	1.06	1.02	1.02	1.77	1.27	1.09	1.04	1.04	1.89	1.24	1.11	1.04	1.01

Notes: See the notes for Table 1, albeit that factor selection is now based on the Bai and Ng (2009) FPE criterion - see (17).

Table 4: Monte Carlo results for selecting which factors enter (2)

	BIC, modified relative to standard					HQIC, modified relative to standard				
N / T	30	50	100	200	400	30	50	100	200	400
30	0.753	0.803	1.384	1.687	2.885	0.566	0.467	0.638	0.809	1.117
50	0.683	0.745	0.918	1.411	2.212	0.568	0.497	0.504	0.587	0.738
100	0.737	0.675	0.792	0.919	1.230	0.707	0.563	0.410	0.395	0.320
200	0.854	0.789	0.730	0.802	1.025	0.859	0.777	0.546	0.434	0.305
400	0.891	0.920	0.811	0.679	0.711	0.940	0.901	0.723	0.491	0.391

Notes: The entries are the ratio of the MSD statistic (16) for the modified BIC (HQIC) criterion relative to the standard BIC (HQIC) criterion, where each are minimized for 2^4 potential combinations of the first 4 factors extracted from x_t across 1,000 Monte Carlo replications. The variables are generated through the DGPs in (15) and (10)-(11) that imposes an asymptotic fit (through (14)) of 0.66, for different values of the time series dimension of all series, T, and the cross-section dimension, N, of series in x_t.

4 Empirical Application

The purpose of this section is to assess the performance of our proposed framework when applied on real world data, in particular by assessing its impact on the out-of-sample forecasting performance of factor-augmented regressions. We summarize the set-up of our application in Section 4.1 and discuss the results in Section 4.2.

4.1 Set-Up

We focus in Section 4.2 on the performance of direct forecasts from factor-augmented regressions for a number of macroeconomic variables. It is standard practice in the macroeconomic forecasting literature to use as forecasting benchmarks for factor-augmented regressions an autoregressive (AR) model and the unconditional mean. The AR benchmark model in the context of direct forecasting can be writing as

$$\Delta y_{t+h,t} = \alpha_h + \sum_{j=1}^{\hat{p}} \rho_{h,j} \Delta y_{t-j+1,t-j} + \epsilon_{t+h,t}, \quad t = 1, \ldots, T \tag{18}$$

with $\Delta y_{t+h,t} = y_{t+h} - y_t$ for $h > 0$ and $\Delta y_{t-j+1,t-j} = y_{t-j+1} - y_{t-j}$ for $j = 1, \ldots, p$. The number of lagged first differences $\hat{p}$ in (18) is determined by sequentially applying the standard BIC starting with a maximum lag order of $p = p_{max}$ down to $p = 0$. The unconditional mean benchmark is simply

$$\Delta y_{t+h,t} = \alpha_h + \epsilon_{t+h,t}, \tag{19}$$

which implies a random walk (RW) forecast for the level of the forecast variable y_t. The assessment of the forecasting performance relative to pure AR-based and random walk-based forecasts is based on the square root of the mean of the squared forecast errors (RMSE). In Section 4.2 we will report ratios of the RMSE of factor-augmented regressions *vis-à-vis* the RMSE based on either (18) or (19). Obviously, superior out-of-sample performance of a factor-augmented regression relative to these benchmarks is indicated by a RMSE ratio smaller than one and *vice versa*.

Following Stock and Watson (2002b) we take our $T \times N$ matrix of N predictor variables $X = (X_1' \cdots X_T')'$ and normalize these such that they are in zero-mean and unity variance space, which results in the $T \times N$ matrix $\tilde{X}$. We then compute the r^{max} eigenvectors of the $N \times N$ matrix $\tilde{X}'\tilde{X}$ that correspond to the first r^{max} largest eigenvalues of that matrix and post-multiplying $\tilde{X}$ with these eigenvectors results in the estimated factors $\hat{f}_{1,t}, \ldots, \hat{f}_{r^{max},t}$ that potentially can be used in our factor-augmented regressions. These adhere to the following specification:

$$\Delta y_{t+h,t} = \alpha_h + \sum_{i=1}^{r^{max}} \delta_{h,i}\beta_{h,i}\hat{f}_{i,t} + \sum_{j=1}^{\hat{p}} \rho_{h,j}\Delta y_{t-j+1,t-j} + \epsilon_{t+h,t}, \tag{20}$$

where $\delta_{h,i} = 1$ if factor i is in the optimal subset of the r^{max} factors otherwise $\delta_{h,i} = 0$.

It is, of course, the aim of this exercise to evaluate the finite sample performance of different selection criteria that can be used to determine the optimal dimensions of a factor-augmented regression like (20). Given (20), we search across the range $p = 0, \ldots, p^{max}$ as well as the $2^{r^{max}}$ possible factor subsets ($(p^{max} + 1) \times 2^{r^{max}}$ potential lag order-factor subset combinations), and select the optimal lag order-factor subset combination that minimizes either the BICM or the HQICM criterion outlined in (8). In addition, we do similar searches using the standard BIC and HQIC criteria as well as the Bai-Ng FPE criterion (17). In the end this results in five different versions of (20) for each forecast horizon that we will assess relative to our two benchmark models. The forecasting models will be updated based on a fixed window of w observations:

1. First forecast for all h is generated on t_0.

2. Extract r^{max} principal components from the N predictor variables over the sample $t = t_0 - w + 1, \ldots, t_0 - h$.

3. Determine for each h over the sample $t = t_0 - w + 1, \ldots, t_0 - h$ the optimal lag order and the optimal subset of factors for (20) for each of our five criteria: BICM, HQICM, BIC, HQIC and Bai-Ng FPE across the lag order range $p = 0, \ldots, p^{max}$ and the $2^{r^{max}}$ possible factor subsets. In a similar vein, determine also the optimal lag order for the AR benchmark based on BIC.

4. Given the outcome of step 3, estimate (18), (19) and versions of (20) that are based on BICM, HQICM, BIC, HQIC and Bai-Ng FPE selection over the sample $t = t_0 - w + 1, \ldots, t_0 - h$ for each h.

5. Extract r^{max} principal components from the N predictor variables N over the sample $t = t_0 - w + 1, \ldots, t_0$.

6. Generate for h the forecast $\Delta\hat{y}_{t+h,t}$ using the estimated dimensions from step 3 and the parameter estimates from step 4 as well as, in case of (20), the common factors from step 5.

7. Repeat for $t_0 + 1, \ldots, T - h$ for each h.

The choice of using rolling data windows in updating our forecasting models serves two purposes. Firstly, it allows the models to cope better with the unavoidable instabilities in the underlying data. Also, it greatly simplifies inference on the significance of any observed outperformance of the benchmark models by a factor-augmented model. In order to be able to do that we employ the Diebold and Mariano (1995)-West (1996) test statistic (hereafter DMW statistic) for the null hypothesis of equal forecasting performance *vis-à-vis* the alternative hypothesis that a factor-augmented model has a lower MSE than the benchmark model:

$$z_{\mathrm{MSE}} = \sqrt{T - t_0 - h}\left(\frac{\mathrm{MSE_B} - \mathrm{MSE_F}}{\sqrt{Var(u_{t+h} - (\mathrm{MSE_B} - \mathrm{MSE_F}))}}\right) \tag{21}$$

with B = AR or RW, and

$$u_{t+h} = e^2_{\mathrm{B},s,s+h} - e^2_{\mathrm{F}s,s+h}; \quad s = t_0, \ldots, T - h$$

In (21) $\mathrm{MSE_B}$ and $\mathrm{MSE_F}$ are the MSE corresponding to the benchmark prediction, based on either (18) or (19), and the factor-augmented regression respectively, u_{t+h} is the difference in the squared prediction error from the benchmark and factor-augmented based forecasts, and $Var(u_{t+h} - (\mathrm{MSE_B} - \mathrm{MSE_F}))$ is an estimate of the variance of the demeaned u_{t+h}'s. It is shown in Giacomini and White (2006, Theorem 4) that the DMW statistic (21) has a standard normal limiting distribution when rolling windows are used, as long as a HAC estimator is used for $Var(u_{t+h} - (\mathrm{MSE_B} - \mathrm{MSE_F}))$ at each horizon h. We will use the Den Haan and Levin (1997) VAR-HAC estimator based on BIC lag selection to approximate $Var(u_{t+h} - (\mathrm{MSE_B} - \mathrm{MSE_F}))$ in (21).

4.2 Empirical Results

We base our empirical exercise on a large panel of monthly macroeconomic, financial and survey-based indicator variables for the United States, which is similar to that used Stock and Watson (2007) but updated by us up to mid-2008. This panel consists of 108 monthly series, which before transformation span a sample starting in January 1959 and ending in July 2008. It spans real variables (sectoral industrial production, employment, subcomponents of unemployment and hours worked), nominal variables (subcomponents of consumer price index, producer price indexes, deflators, wages, money and credit aggregates), asset prices (interest rates, stock prices and exchange rates) and surveys. Of these 108 series, we use 106 as predictor variables that are transformed such that they are $I(0)$. In general this means that the real variables are expressed in log first differences and we simply use first differences of series expressed in rates, such as interest rates and unemployment series; see Appendix B for more details. We transform the nominal variables into first differences of annual growth rates in order to guarantee that the dynamic properties of these transformed series are comparable to those of the rest of the predictor variable panel, as for example motivated in D'Agostino and Giannone (2006, Appendix B). Hence, after transforming the predictor variables we end up with an effective span of the data that starts in February 1960 (i.e. 1960.2) and ends in July 2008 (i.e. 2008.07).

The aforementioned panel is used to forecast inflation based on the U.S. personal consumption expenditures (PCE) price index as well as the federal funds rate – see Table 5 for an overview of the appropriate transformation of each forecast variable – and we deliberately keep these two variables separate from the panel of predictor series. The federal funds rate is determined by the Federal Reserve Board, which sets the target for the federal funds rate by taking into account both a range of nominal and real developments, so factor methods could potentially be very

Table 5: Transformation of the forecast variables

Y_t	$\Delta y_{t,t-1}$	$\Delta y_{t+h,t}$
PCE index	$\Delta \ln Y_{t,t-12} - \Delta \ln Y_{t-1,t-13}$	$\Delta \ln Y_{t+h,t+h-12} - \Delta \ln Y_{t,t-12}$
Federal Funds rate	$\Delta Y_{t,t-1}$	$\Delta Y_{t+h,t}$

Notes: The table illustrates the transformation of a forecast variable Y_t, indicated in the first column, for use in the prediction regression (20).

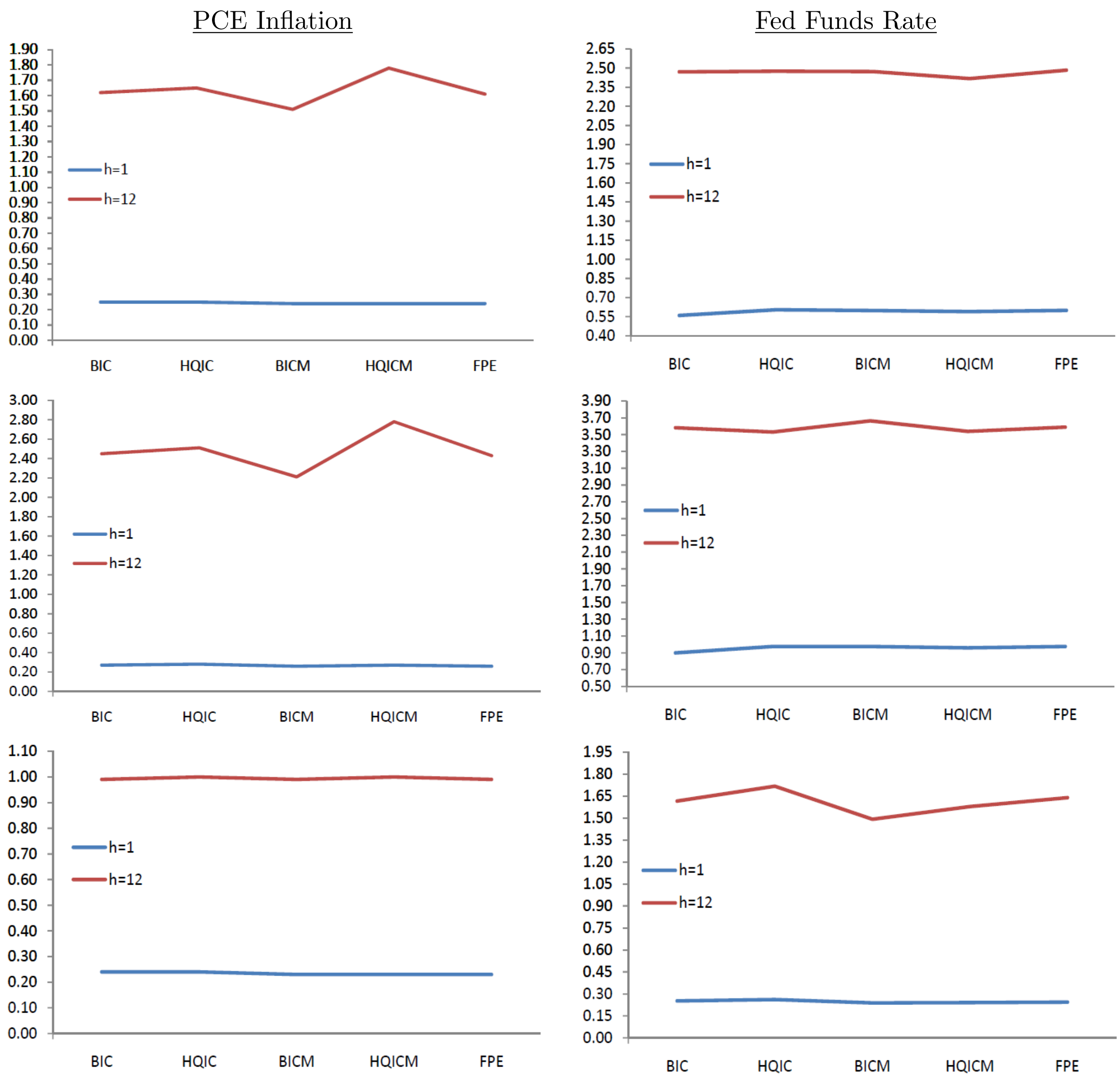
useful in predicting this variable. Inflation based on the PCE price index is of interest, as the expenditure weights of the individual consumption goods in this price index vary from period to period and it is a chain-linked index. As such, one would expect that PCE inflation better reflects the effects of substitution across goods by consumers when relative prices change than other inflation measures, such as CPI inflation. Also, the Federal Reserve Board has made it clear that it views PCE inflation as its primary measure of inflation.[3]

As described in the previous subsection, the forecasting models are updated based on a fixed window of data of size w and all forecasts are direct forecasts for 2 horizons (in months): $h = 1$ and $h = 12$, which are horizons commonly analyzed in the literature. We set $w = 156$, which corresponds with 13 years worth of data. This window size was chosen as a compromise between having a data set with a large enough time series relative to cross-sectional size of the predictor variable panel as well as the potential lag order in (20), and having a not 'too long' data window so that our models can adapt when there are instabilities in the data. In each update we determine five versions of the factor-augmented regression (20) using our modified information criteria in (8), BIC, HQIC and the Bai-Ng FPE measure (17), using a lag order upper bound of 12 lags and 8 principal components extracted from our panel. For each criterion we simultaneously select the optimal lag order as well as the optimal subset of factors across 13×2^8 potential lag order-factor subset combinations such that a particular criterion is minimized. In case of the AR benchmark (18) we select that lag order from $p = 0, \ldots, 12$ that minimizes the BIC criterion for (18). The forecast evaluation spans three samples: January 1973 - July 2008, January 1973 - December 1984 and January 1985 - July 2008. The latter two sub-samples split the first sample in two around the start of the 'Great Moderation': e.g. Kim and Nelson (1999), McConnell and Perez-Quiros (2000) and Sensier and van Dijk (2004) all find evidence for a downward, exogenous, shift in the volatility of a large number of U.S. macroeconomic time series around 1985.

Let us now turn to the out-of-sample forecasting results for both PCE inflation and the federal funds rate. Before we discuss the out-of-sample comparison relative to naive benchmark models, it may be worthwhile to firstly report the RMSE estimates themselves for the different factor-augmented regressions. We plot these RMSE estimates in Figure 1 across both the full 1973-2008 evaluation sample as well as the two sub-samples (1973-1984 and 1985-2008, respectively). One striking observation in this figure is that for both PCE inflation and the fed funds rate at both $h = 1$ and $h = 12$ the RMSE's are higher for the 1973-1984 evaluation sample than for the

[3] See 'Monetary Policy Report to the Congress', February 2000, Federal Reserve Board.

Figure 1: Out-of-sample RMSE estimates



Notes: In this figure we plot the out-of-sample RMSE, which are reported in percentage points, for both PCE inflation and the Federal Funds rate at horizons $h = 1$ and $h = 12$ based on factor- and lag order selection using, respectively, BIC, HQIC, BICM, HQICM (see (8)) and FPE (see (17)) criteria. The first row plots the RMSE estimates for the full January 1973 - July 2008 evaluation sample, the second row for the January 1973 - December 1984 sub-sample, and the last row for the January 1985 - July 2008 sub-sample.

later 1985-2008 sample. This is, of course, a testimony to the earlier cited finding from the empirical macroeconomics literature that there has been a downward shift in the volatility of a range of macroeconomic series around 1985. The resulting lower RMSE estimates would suggest an improved prediction performance for the factor-augmented regressions, albeit that a lower volatility environment could also imply that naive but parsimonious autoregressive or random walk specifications are very hard to beat by our factor-augmented models. Another observation from Figure 1 is that it seems at first sight that the differences in the RMSE estimates across BIC-, HQIC-, BICM-, HQICM-, and FPE-based factor-augmented regressions are often not that big, especially at $h = 1$. Of course, the only way to really assess the validity of these observations is to compare the out-of-sample performance of our range of factor-augmented regressions relative to parsimonious time series models, taking into account the variability of the different RMSE estimates, and we do that in the following.

In Table 6 we summarise the out-of-sample comparison of the factor-augmented models relative to (18) and (19) for PCE inflation and the Federal funds rate at one-month ($h = 1$) and one-year ($h = 12$) horizons. Both the point estimates, the RMSE ratios, as well as test statistics for the null that the benchmark models cannot be outperformed by the factor-augmented regressions are reported. In case of PCE inflation the factor-augmented model selection strategy based on our BICM criterion results in the best performing inflation forecast, and in case of the 1-month horizon BICM's outperformance of the benchmark forecasts is statistically significant in most cases. Our BICM and HQICM criteria applied to a factor-augmented regression appears to be the most useful in forecasting future changes in the federal funds rate at the 1-year horizon, especially for factor-augmented models based on the BICM measure. Only for the 1972-1984 sub-sample it appears that all the employed strategies for factor selection in factor-augmented regressions are very effective in outperforming the benchmark models for fed fund rate forecasts at $h = 12$, particularly relative to the AR-based benchmark predictions.

The empirical results in this section confirm our earlier insights from theory and Monte Carlo experiments. By taking into account factor estimation error when selecting the dimensions of a factor-augmented regression, which perform at least as well as, and often better than, factor-augmented regressions whose dimensions are determined through standard model selection criteria.

5 Conclusions

Factor-augmented regressions are often used for macroeconomic forecasting and analysis as a parsimonious way of basing the forecast or the analysis on information from a large number of variables. This paper is focused on the issue of how to determine which factors that are relevant for such a factor-augmented regression, whereas existing work has been more focused on criteria that can consistently estimate the appropriate number of factors that drive the dynamics in a large-dimensional panel of explanatory variables. However, in the latter case the resulting number of factors are not necessarily all relevant for modeling a specific dependent variable within a factor-augmented regression. Further, determining the number of factors in large datasets is a very difficult task, and, as of yet, no satisfactory method that works in the majority of possible modeling scenarios seems to have been developed.

Factor estimation error is an important issue in determining the dimensions of a factor-augmented regression, particularly when the time series dimension of the underlying panel of predictor series is larger than the cross-section dimension. We develop a number of theoretical conditions selection criteria have to fulfill in order to estimate the factor subset that is relevant for such a regression in a consistent manner. The framework does not hinge on a particular factor estimation methodology. Based on this framework it is clear that standard model selection criteria like AIC, BIC and HQIC do not necessarily provide consistent estimates of the dimensions of a factor-augmented regression. As a consequence, we suggest selection criteria that do fulfill the conditions set by our theoretical framework and thus are consistent. Our criteria essentially take standard consistent information criteria that are commonly used in time series econometrics and modify these such that the corresponding penalty function for dimensionality also penalizes factor estimation error. We show through Monte Carlo applications and empirically, through forecast evaluations for PCE inflation and the fed funds rate, that our model selection criteria are useful in determining the dimensions of factor-augmented regressions.

References

Akaike, H., 1974, A New Look at the Statistical Model Identification, *IEEE Transactions on Automatic Control* **19**, 716–723.

Bai, J., 2003, Inferential Theory for Factor Models of Large Dimensions, *Econometrica* **71**, 135–173.

Bai, J. and S. Ng, 2002, Determining the Number of Factors in Approximate Factor Models, *Econometrica* **70**, 191–221.

Bai, J. and S. Ng, 2006, Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-Augmented Regressions, *Econometrica* **74**, 1133–1150.

Bai, J. and S. Ng, 2008, Forecasting Economic Time Series Using Targeted Predictors, *Journal of Econometrics* **146**, 304–317.

Bai, J. and S. Ng, 2009, Boosting Diffusion Indices, *Journal of Applied Econometrics* **24**, 607–629.

D'Agostino, A. and D. Giannone, 2006, Comparing Alternative Predictors Based on Large-Panel Factor Models, *Working Paper 680*, European Central Bank.

Den Haan, W. J. and A. Levin, 1997, A Practitioner's Guide to Robust Covariance Matrix Estimation, *Handbook of Statistics*, Vol. 25, pp. 291–341.

Diebold, F. X. and R. S. Mariano, 1995, Comparing Predictive Accuracy, *Journal of Business & Economic Statistics* **13**, 253–263.

Doz, C., D. Giannone and L. Reichlin, 2006, A Quasi Maximum Likelihood Approach for Large Approximate Dynamic Factor Models, *Discussion Paper 5724*, Centre for Economic Policy Research.

Forni, M., M. Hallin, M. Lippi and L. Reichlin, 2000, The Generalized Dynamic Factor Model: Identification and Estimation, *Review of Economics and Statistics* **82**, 540–554.

Giacomini, R. and H. White, 2006, Tests of Conditional Predictive Ability, *Econometrica* **74**, 1545–1578.

Groen, J. J. J. and G. Kapetanios, 2008, Revisiting Useful Approaches to Data-Rich Macroeconomic Forecasting, *Staff Reports 327*, Federal Reserve Bank of New York.

Hannan, E. J. and B. G. Quinn, 1979, The Determination of the Order of an Autoregression, *Journal of the Royal Statistical Society, Series B* **41**, 190–195.

Kapetanios, G., 2007, Variable selection in regression models using nonstandard optimisation of information criteria, *Computational Statistics and Data Analysis* **52**, 4–15.

Kapetanios, G., 2010, A Testing Procedure for Determining the Number of Factors in Approximate Factor Models With Large Datasets, *Journal of Business & Economic Statistics* . *Forthcoming.* Online at http://pubs.amstat.org/doi/pdf/10.1198/jbes.2009.07239.

Kapetanios, G. and M. Marcellino, 2009, A Parametric Estimation Method for Dynamic Factor Models of Large Dimensions, *Journal of Time Series Analysis* **30**, 208–238.

Kim, C-J. and C. R. Nelson, 1999, Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle, *Review of Economics and Statistics* **81**, 608–616.

McConnell, M. and G. Perez-Quiros, 2000, Output Fluctuations in the United States: What has Changed Since the Early 1980's?, *American Economic Review* **90**, 1464–1476.

Monacelli, T. and L. Sala, 2009, The International Dimension: Evidence from Disaggregated Consumer Price Data, *Journal of Money, Credit, and Banking* **41**, 101–120.

Schwarz, G., 1978, Estimating the Dimension of a Model, *Annals of Statistics* **6**, 461–464.

Sensier, M. and D. J. C. van Dijk, 2004, Testing for Volatility Changes in U.S. Macroeconomic Time Series, *Review of Economics and Statistics* **86**, 833–839.

Stock, J. H. and M. W. Watson, 1998, Diffusion Indexes, *Working Paper 6702*, National Bureau for Economic Research.

Stock, J. H. and M. W. Watson, 2002a, Forecasting Using Principal Components from a Large Number of Predictors, *Journal of the American Statistical Association* **97**, 1167–1179.

Stock, J. H. and M. W. Watson, 2002b, Macroeconomic Forecasting Using Diffusion Indexes, *Journal of Business & Economic Statistics* **20**, 147–162.

Stock, J. H. and M. W. Watson, 2007, Forecasting in Dynamic Factor Models Subject to Structural Instability, *mimeo*, Harvard University and Princeton University.

West, K. D., 1996, Asymptotic Inference About Predictive Ability, *Econometrica* **64**, 1067–1084.

Table 6: Out-of-sample forecasting results

h	*BIC*	*HQIC*	*BICM*	*HQICM*	*FPE*	*BIC*	*HQIC*	*BICM*	*HQICM*	*FPE*
			PCE Inflation					Federal Funds Rate		
			January 1973 - July 2008					January 1973 - July 2008		
			Benchmark: RW					*Benchmark: RW*		
1	0.960	0.967	**0.914*****	0.941*	0.927***	**0.898**	0.946	0.936	0.959	0.960
	(-1.129)	(-0.863)	(-2.720)	(-1.606)	(-2.358)	(-1.053)	(-0.527)	(-0.979)	(-0.583)	(-0.530)
12	1.076	1.095	**1.011**	1.199	1.050	1.018	1.030	**0.983**	0.992	1.003
	(1.287)	(1.657)	(0.234)	(1.200)	(1.140)	(0.375)	(0.667)	(-0.351)	(-0.164)	(0.043)
			Benchmark: AR					*Benchmark: AR*		
1	1.032	1.040	**0.983***	1.012	0.999	0.920	0.930	**0.883**	0.943	0.944
	(1.288)	(1.485)	(-1.295)	(0.453)	(-0.270)	(-0.645)	(-0.421)	(-0.723)	(-0.531)	(-0.406)
12	1.052	1.070	**0.989**	1.172	1.026	0.916	0.927	**0.884**	0.893	0.902
	(1.058)	(1.452)	(-0.227)	(1.136)	(0.610)	(-1.022)	(-0.872)	(-1.110)	(-0.977)	(-0.893)
			January 1985 - July 2008					January 1985 - July 2008		
			Benchmark: RW					*Benchmark: RW*		
1	0.999	0.982	**0.955****	0.970	0.977	1.175	1.212	**1.027**	1.052	1.198
	(-0.027)	(-0.399)	(-1.854)	(-0.678)	(-1.008)	(1.173)	(1.327)	(0.315)	(0.571)	(1.162)
12	1.157	1.179	**1.138**	1.161	1.145	1.058	1.136	**1.012**	1.038	1.040
	(1.328)	(2.078)	(1.336)	(1.612)	(1.864)	(0.580)	(1.740)	(0.112)	(0.502)	(0.426)
			Benchmark: AR					*Benchmark: AR*		
1	1.016	0.999	**0.971***	0.986	0.993	1.362	1.405	**1.191**	1.219	1.389
	(0.468)	(-0.036)	(-1.515)	(-0.389)	(-0.404)	(2.074)	(2.165)	(1.923)	(2.103)	(1.870)
12	1.048	1.068	**1.032**	1.052	1.040	1.031	1.107	**0.987**	1.012	1.014
	(0.790)	(1.246)	(0.671)	(1.094)	(0.681)	(0.381)	(1.603)	(-0.144)	(0.186)	(0.181)
			January 1973 - December 1984					January 1973 - December 1984		
			Benchmark: RW					*Benchmark: RW*		
1	0.912*	0.949	**0.862****	0.906**	0.865**	**0.869***	0.923	0.918*	0.949	0.937
	(-1.419)	(-0.768)	(-2.205)	(-1.715)	(-2.142)	(-1.557)	(-0.969)	(-1.289)	(-0.847)	(-1.173)
12	1.037	1.058	**0.957**	1.197	1.006	0.985	0.978	0.958	**0.952**	0.969
	(0.512)	(0.802)	(-0.760)	(0.907)	(0.092)	(-0.249)	(-0.449)	(-0.931)	(-0.964)	(-0.483)
			Benchmark: AR					*Benchmark: AR*		
1	1.058	1.101	**1.000**	1.051	1.003	0.896	0.891	**0.844**	0.922	0.910
	(1.480)	(2.314)	(0.050)	(1.280)	(0.181)	(-0.613)	(-0.908)	(-1.136)	(-0.558)	(-1.000)
12	1.041	1.062	**0.960**	1.201	1.010	0.862**	0.856**	0.839**	**0.834****	0.848
	(0.652)	(0.700)	(-0.563)	(0.982)	(0.174)	(-1.851)	(-1.866)	(-2.292)	(-2.280)	(-1.229)

Notes: The table reports the ratio of the RMSE of a version of (20) *vis-à-vis* the random walk model (19) or the autoregressive model (18) for PCE inflation and fed funds rate (see Table 5)) at each horizon h (in months). In parentheses we report the DMW statistic (21) for the null of equal forecasting performance relative to the alternative that a factor-augmented regression has a lower MSE. A $^{*(**)[***]}$ indicates a rejection of the null at a *one-sided* 10% (5%) [1%] significance level based on a standard normal distribution. For (20) the optimal lags are picked from a range between 0 and 12, whereas an upper bound of 8 is used when selecting the number of factors. Columns *BIC*, *HQIC*, *BICM*, *HQICM* and *FPE* report the results when the optimal number of lags and factors are chosen to minimize, respectively, the BIC, Hannan-Quinn IC, the modified BIC and Hannan-Quinn IC measures, see (8), and the Bai-Ng FPE criterion (17). The method that performs relatively best *vis-à-vis* the benchmark is highlighted in **bold**.

Appendix

A Proofs

Proof of Theorem 1

Let $\hat{\mathbf{F}} = (\hat{f}_1, \ldots, \hat{f}_T)'$ and $\mathbf{F} = (f_1, \ldots, f_T)'$ where $\hat{f}_t$ denotes a generic set of estimated factors and f_t denote its probability limit. From now on when the matrices $\mathbf{F}$ and $\mathbf{M} = I - \mathbf{F}'(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'$ have superscript (i), they are constructed using $f_t^{(i)}$. When the coefficient vector β has superscript (i) then it refers to a model using $f_t^{(i)}$. Also, $\hat{\beta}$ indicate estimated parameters in a model with estimated factors and $\tilde{\beta}$ indicate estimated parameters in a model with true unobserved factors. Throughout a 0 superscript denotes use of the true set of factors. We denote the penalty term for $f_t^{(i)}$ by $C_{T,N}^{(i)} = r^{(i)}\tilde{C}_{T,N}^{(i)}$, with $\tilde{C}_{T,N}^{(i)}$ solely depending on T and N. Finally, collect the error terms of (2) in $\mathbf{e} = (e_1, ..., e_T)'$. The feasible information criterion takes the following form

$$\widehat{IC}(\beta, C_{T,N}) = \frac{T}{2}\ln\{\frac{1}{T}(\mathbf{y} - \hat{\mathbf{F}}\beta)'(\mathbf{y} - \hat{\mathbf{F}}\beta)\} + C_{T,N} \tag{A.1}$$

At first, we consider the case where $\mathbf{M}^{(i)}\mathbf{F}^0 = 0$ and $\mathbf{M}^{(j)}\mathbf{F}^0 \neq 0$. We wish to show that

$$\lim_{T\to\infty} P\{\widehat{IC}(\hat{\beta}^{(j)}, C_{T,N}^{(j)}) - \widehat{IC}(\hat{\beta}^{(i)}, C_{T,N}^{(i)}) < 0\} = 0 \tag{A.2}$$

It is straightforward to show, using (A.1), that (A.2) becomes

$$\lim_{N,T\to\infty} P\left\{\begin{array}{c} \frac{T}{2}\ln\left[\frac{\frac{1}{T}\mathbf{y}'\mathbf{M}^{(j)}\mathbf{y}}{\frac{1}{T}\mathbf{y}'\mathbf{M}^{(i)}\mathbf{y}}\right] + \frac{T}{2}\ln\left[\frac{\frac{1}{T}\left(\mathbf{y}-\hat{\mathbf{F}}^{(j)}\hat{\beta}^{(j)}\right)'\left(\mathbf{y}-\hat{\mathbf{F}}^{(j)}\hat{\beta}^{(j)}\right)}{\frac{1}{T}\left(\mathbf{y}-\mathbf{F}^{(j)}\tilde{\beta}^{(j)}\right)'\left(\mathbf{y}-\mathbf{F}^{(j)}\tilde{\beta}^{(j)}\right)}\right] - \\ \frac{T}{2}\ln\left[\frac{\frac{1}{T}\left(\mathbf{y}-\hat{\mathbf{F}}^{(i)}\hat{\beta}^{(i)}\right)'\left(\mathbf{y}-\hat{\mathbf{F}}^{(i)}\hat{\beta}^{(i)}\right)}{\frac{1}{T}\left(\mathbf{y}-\mathbf{F}^{(i)}\tilde{\beta}^{(i)}\right)'\left(\mathbf{y}-\mathbf{F}^{(i)}\tilde{\beta}^{(i)}\right)}\right] < C_{T,N}^{(i)} - C_{T,N}^{(j)} \end{array}\right\} = 0. \tag{A.3}$$

We first examine the first term of the LHS of the inequality within (A.3). By expanding $\mathbf{y}$ we get that

$$\frac{T}{2}\ln\left[\frac{\frac{1}{T}\mathbf{y}'\mathbf{M}^{(j)}\mathbf{y}}{\frac{1}{T}\mathbf{y}'\mathbf{M}^{(i)}\mathbf{y}}\right] = \frac{T}{2}\ln\left[\frac{\frac{1}{T}(\beta^{0\prime}\mathbf{F}^{0\prime}\mathbf{M}^{(j)}\mathbf{F}^0\beta^0 + \mathbf{e}\mathbf{M}^{(j)}\mathbf{e} + 2\mathbf{e}'\mathbf{M}^{(j)}\mathbf{F}^0\beta^0)}{\frac{1}{T}(\beta^{0\prime}\mathbf{F}^{0\prime}\mathbf{M}^{(i)}\mathbf{F}^0\beta^0 + \mathbf{e}\mathbf{M}^{(i)}\mathbf{e} + 2\mathbf{e}'\mathbf{M}^{(i)}\mathbf{F}^0\beta^0)}\right]$$

But idempotency implies positive-definiteness and as a result $\beta^{0\prime}\mathbf{F}^{0\prime}\mathbf{M}^{(j)}\mathbf{F}^0\beta^0 > 0$. Also, $\mathbf{M}^{(i)}\mathbf{F}^0 = 0$, and so $\beta^{0\prime}\mathbf{F}^{0\prime}\mathbf{M}^{(i)}\mathbf{F}^0\beta^0 = 0$. Further, by the assumed stationarity of the model, and using the assumptions of the theorem, we get

$$0 < p\lim\frac{1}{T}\left\|\mathbf{F}^{0\prime}\mathbf{F}^0\right\| < \infty,\ 0 < p\lim\frac{1}{T}\left\|\mathbf{F}^{(i)\prime}\mathbf{F}^{(i)}\right\| < \infty,$$

$$0 < p\lim\frac{1}{T}\left\|\mathbf{F}^{(j)\prime}\mathbf{F}^{(j)}\right\| < \infty,\ 0 < p\lim\frac{1}{T}\left\|\mathbf{F}^{(j)\prime}\mathbf{F}^{(i)}\right\| < \infty,$$

$$\frac{1}{T}\mathbf{F}^{0\prime}\mathbf{e} = O_p\left(T^{-1/2}\right),\ \frac{1}{T}\mathbf{F}^{(i)\prime}\mathbf{e} = O_p\left(T^{-1/2}\right),$$

and

$$\frac{1}{T}\mathbf{F}^{(j)\prime}\mathbf{e} = O_p\left(T^{-1/2}\right).$$

So,

$$\left(\frac{1}{T}\mathbf{F}^{(i)\prime}\mathbf{e}\right)^{\prime}\left(\frac{1}{T}\mathbf{F}^{(i)\prime}\mathbf{F}^{(i)}\right)^{-1}\left(\frac{1}{T}\mathbf{F}^{(i)\prime}\mathbf{e}\right)=O_p\left(T^{-1}\right), \tag{A.4}$$

and

$$\left(\frac{1}{T}\mathbf{F}^{(j)\prime}\mathbf{e}\right)^{\prime}\left(\frac{1}{T}\mathbf{F}^{(j)\prime}\mathbf{F}^{(j)}\right)^{-1}\left(\frac{1}{T}\mathbf{F}^{(j)\prime}\mathbf{e}\right)=O_p\left(T^{-1}\right). \tag{A.5}$$

Further,

$$\left(\frac{1}{T}\mathbf{F}^{(i)\prime}\mathbf{e}\right)^{\prime}\left(\frac{1}{T}\mathbf{F}^{(i)\prime}\mathbf{F}^{(i)}\right)^{-1}\left(\frac{1}{T}\mathbf{F}^{(i)\prime}\mathbf{F}^{0}\beta^{0}\right)=O_p\left(T^{-1/2}\right),$$

and

$$\left(\frac{1}{T}\mathbf{F}^{(j)\prime}\mathbf{e}\right)^{\prime}\left(\frac{1}{T}\mathbf{F}^{(j)\prime}\mathbf{F}^{(j)}\right)^{-1}\left(\frac{1}{T}\mathbf{F}^{(j)\prime}\mathbf{F}^{0}\beta^{0}\right)=O_p\left(T^{-1/2}\right).$$

Thus,

$$\frac{1}{T}\mathbf{e}^{\prime}\mathbf{M}^{(i)}\mathbf{e}=\sigma^2+O_p\left(T^{-1}\right),\ \frac{1}{T}\mathbf{e}^{\prime}\mathbf{M}^{(j)}\mathbf{e}=\sigma^2+O_p\left(T^{-1}\right), \tag{A.6}$$

$$\frac{1}{T}\mathbf{e}^{\prime}\mathbf{M}^{(j)}\mathbf{F}^{0}\beta^{0}\rightarrow_p 0,\ \text{and}\ \frac{1}{T}\mathbf{e}^{\prime}\mathbf{M}^{(i)}\mathbf{F}^{0}\beta^{0}\rightarrow_p 0.$$

As a result of all the above $\ln\left[\frac{\frac{1}{T}\mathbf{y}^{\prime}\mathbf{M}^{(j)}\mathbf{y}}{\frac{1}{T}\mathbf{y}^{\prime}\mathbf{M}^{(i)}\mathbf{y}}\right]$ is positive and $O_p(1)$ and, therefore, we have in (A.3)

$$\frac{T}{2}\ln\left[\frac{\frac{1}{T}\mathbf{y}^{\prime}\mathbf{M}^{(j)}\mathbf{y}}{\frac{1}{T}\mathbf{y}^{\prime}\mathbf{M}^{(i)}\mathbf{y}}\right]=O_p(T).$$

Next, we need to analyse $\frac{T}{2}\ln\left[\frac{\frac{1}{T}\left(\mathbf{y}-\hat{\mathbf{F}}^{(j)}\hat{\beta}^{(j)}\right)^{\prime}\left(\mathbf{y}-\hat{\mathbf{F}}^{(j)}\hat{\beta}^{(j)}\right)}{\frac{1}{T}\left(\mathbf{y}-\mathbf{F}^{(j)}\tilde{\beta}^{(j)}\right)^{\prime}\left(\mathbf{y}-\mathbf{F}^{(j)}\tilde{\beta}^{(j)}\right)}\right]$ and $\frac{T}{2}\ln\left[\frac{\frac{1}{T}\left(\mathbf{y}-\hat{\mathbf{F}}^{(i)}\hat{\beta}^{(i)}\right)^{\prime}\left(\mathbf{y}-\hat{\mathbf{F}}^{(i)}\hat{\beta}^{(i)}\right)}{\frac{1}{T}\left(\mathbf{y}-\mathbf{F}^{(i)}\tilde{\beta}^{(i)}\right)^{\prime}\left(\mathbf{y}-\mathbf{F}^{(i)}\tilde{\beta}^{(i)}\right)}\right]$ in (A.3). We have to show that

$$\frac{T}{2}\ln\left[\frac{\frac{1}{T}\left(\mathbf{y}-\hat{\mathbf{F}}^{(j)}\hat{\beta}^{(j)}\right)^{\prime}\left(\mathbf{y}-\hat{\mathbf{F}}^{(j)}\hat{\beta}^{(j)}\right)}{\frac{1}{T}\left(\mathbf{y}-\mathbf{F}^{(j)}\tilde{\beta}^{(j)}\right)^{\prime}\left(\mathbf{y}-\mathbf{F}^{(j)}\tilde{\beta}^{(j)}\right)}\right]=O_p\left(T\min(N,T)^{-1}\right)=o_p(T), \tag{A.7}$$

and

$$\frac{T}{2}\ln\left[\frac{\frac{1}{T}\left(\mathbf{y}-\hat{\mathbf{F}}^{(i)}\hat{\beta}^{(i)}\right)^{\prime}\left(\mathbf{y}-\hat{\mathbf{F}}^{(i)}\hat{\beta}^{(i)}\right)}{\frac{1}{T}\left(\mathbf{y}-\mathbf{F}^{(i)}\tilde{\beta}^{(i)}\right)^{\prime}\left(\mathbf{y}-\mathbf{F}^{(i)}\tilde{\beta}^{(i)}\right)}\right]=O_p\left(T\min(N,T)^{-1}\right)=o_p(T), \tag{A.8}$$

as long as $N\rightarrow\infty$. Since $C_{T,N}^{(i)}-C_{T,N}^{(j)}=o(T)$, (A.3) holds, proving (A.2). We will show (A.7). (A.8) follows very similarly. We start by noting Lemma A.1 of Bai and Ng (2006), which states that

$$\frac{1}{T}\sum_{t=1}^{T}(\hat{f}_t-f_t)q_t^{\prime}=O_p\left(\min(k,T)^{-1}\right) \tag{A.9}$$

as long as q_t has finite fourth moments, nonsingular covariance matrix and $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(q_t-E(q_t)\right)$ satisfies a central limit theorem. These conditions are satisfied for f_t and y_t. Using this result, we have that

$$\frac{1}{T}\mathbf{F}^{(j)}\mathbf{y}-\frac{1}{T}\hat{\mathbf{F}}^{(j)\prime}\mathbf{y}=O_p\left(\min(N,T)^{-1}\right)\ \text{and}\ \frac{1}{T}\hat{\mathbf{F}}^{(j)\prime}\hat{\mathbf{F}}^{(j)}-\frac{1}{T}\hat{\mathbf{F}}^{(j)\prime}\mathbf{F}^{(j)}=O_p\left(\min(N,T)^{-1}\right).$$

Then,

$$\frac{1}{T}\mathbf{y}'\hat{\mathbf{F}}^{(j)}\left(\hat{\mathbf{F}}^{(j)\prime}\hat{\mathbf{F}}^{(j)}\right)^{-1}\hat{\mathbf{F}}^{(j)\prime}\mathbf{y}-\frac{1}{T}\mathbf{y}'\mathbf{F}^{(j)}\left(\mathbf{F}^{(j)\prime}\mathbf{F}^{(j)}\right)^{-1}\mathbf{F}^{(j)\prime}\mathbf{y}=O_p\left(\min(N,T)^{-1}\right),$$

which implies that

$$\frac{1}{T}(\mathbf{y}-\hat{\mathbf{F}}^{(j)}\hat{\beta})'(\mathbf{y}-\hat{\mathbf{F}}^{(j)}\hat{\beta})-\frac{1}{T}(\mathbf{y}-\mathbf{F}^{(j)}\tilde{\beta})'(\mathbf{y}-\mathbf{F}^{(j)}\tilde{\beta})=O_p\left(\min(N,T)^{-1}\right),$$

which immediately implies (A.7).

Now, we want to prove that (A.2) holds, when $\mathbf{M}^{(i)}\mathbf{F}^0=\mathbf{M}^{(j)}\mathbf{F}^0=0$ and $r^{(i)}<r^{(j)}$. Then,

$$\beta^{0\prime}\mathbf{F}^{0\prime}\mathbf{M}^{(j)}\mathbf{F}^0\beta^0=\mathbf{e}'\mathbf{M}^{(j)}\mathbf{F}^0\beta^0=\beta^{0\prime}\mathbf{F}^{0\prime}\mathbf{M}^{(i)}\mathbf{F}^0\beta^0=\mathbf{e}'\mathbf{M}^{(i)}\mathbf{F}^0\beta^0=0$$

Thus,

$$\frac{T}{2}\ln\left[\frac{\frac{1}{T}\mathbf{y}'\mathbf{M}^{(j)}\mathbf{y}}{\frac{1}{T}\mathbf{y}'\mathbf{M}^{(i)}\mathbf{y}}\right]=\frac{T}{2}\ln\left[\frac{\frac{1}{T}\mathbf{e}'\mathbf{M}^{(j)}\mathbf{e}}{\frac{1}{T}\mathbf{e}'\mathbf{M}^{(i)}\mathbf{e}}\right]$$

We need to show that

$$\ln\left[\frac{\frac{1}{T}\mathbf{e}'\mathbf{M}^{(j)}\mathbf{e}}{\frac{1}{T}\mathbf{e}'\mathbf{M}^{(i)}\mathbf{e}}\right]=O_p(T^{-1}) \tag{A.10}$$

which implies

$$\frac{T}{2}\ln\left[\frac{\frac{1}{T}\mathbf{e}'\mathbf{M}^{(j)}\mathbf{e}}{\frac{1}{T}\mathbf{e}'\mathbf{M}^{(i)}\mathbf{e}}\right]=O_p(1)$$

To show (A.10), we have

$$\ln\left[\frac{\frac{1}{T}\mathbf{e}'\mathbf{M}^{(j)}\mathbf{e}}{\frac{1}{T}\mathbf{e}'\mathbf{M}^{(i)}\mathbf{e}}\right]=\frac{\frac{1}{T}\mathbf{e}'\left(\mathbf{M}^{(j)}-\mathbf{M}^{(i)}\right)\mathbf{e}}{\frac{1}{T}\mathbf{e}'\mathbf{M}^{(i)}\mathbf{e}}+o_p\left(\frac{\frac{1}{T}\mathbf{e}'\mathbf{M}^{(j)}-\mathbf{M}^{(i)}\mathbf{e}}{\frac{1}{T}\mathbf{e}'\mathbf{M}^{(i)}\mathbf{e}}\right). \tag{A.11}$$

But,

$$\begin{aligned}\frac{1}{T}\mathbf{e}'\left(\mathbf{M}^{(j)}-\mathbf{M}^{(i)}\right)\mathbf{e}&=\left(\frac{1}{T}\mathbf{F}^{(j)\prime}\mathbf{e}\right)'\left(\frac{1}{T}\mathbf{F}^{(j)\prime}\mathbf{F}^{(j)}\right)^{-1}\left(\frac{1}{T}\mathbf{F}^{(j)\prime}\mathbf{e}\right)-\left(\frac{1}{T}\mathbf{F}^{(i)\prime}\mathbf{e}\right)'\left(\frac{1}{T}\mathbf{F}^{(i)\prime}\mathbf{F}^{(i)}\right)^{-1}\left(\frac{1}{T}\mathbf{F}^{(i)\prime}\mathbf{e}\right)\\&=O_p\left(T^{-1}\right),\end{aligned}$$

by (A.4) and (A.5). As a result of (A.10),

$$\lim_{T\to\infty}P\left\{\frac{T}{2}\ln\left[\frac{\frac{1}{T}\mathbf{y}'\mathbf{M}^{(j)}\mathbf{y}}{\frac{1}{T}\mathbf{y}'\mathbf{M}^{(i)}\mathbf{y}}\right]<C^0_{T,N}-C_{T,N}\right\}=0 \tag{A.12}$$

as long as $C^0_{T,N}-C_{T,N}\to-\infty$. But, (A.7) implies this is not enough for (A.2) to hold. For (A.2) to hold, and given (A.12), we need that

$$\lim_{T\to\infty}P\left\{\begin{array}{c}\frac{T}{2}\ln\left[\frac{\frac{1}{T}\left(\mathbf{y}-\hat{\mathbf{F}}^{(j)}\hat{\beta}^{(j)}\right)'\left(\mathbf{y}-\hat{\mathbf{F}}^{(j)}\hat{\beta}^{(j)}\right)}{\frac{1}{T}\left(\mathbf{y}-\mathbf{F}^{(j)}\tilde{\beta}^{(j)}\right)'\left(\mathbf{y}-\mathbf{F}^{(j)}\tilde{\beta}^{(j)}\right)}\right]+\frac{T}{2}\ln\left[\frac{\frac{1}{T}\left(\mathbf{y}-\hat{\mathbf{F}}^{(i)}\hat{\beta}^{(i)}\right)'\left(\mathbf{y}-\hat{\mathbf{F}}^{(i)}\hat{\beta}^{(i)}\right)}{\frac{1}{T}\left(\mathbf{y}-\mathbf{F}^{(i)}\tilde{\beta}^{(i)}\right)'\left(\mathbf{y}-\mathbf{F}^{(i)}\tilde{\beta}^{(i)}\right)}\right]\\<C^{(i)}_{T,N}-C^{(j)}_{T,N}\end{array}\right\}=0. \tag{A.13}$$

But (A.13) holds, if $\frac{C^{(i)}_{T,N}-C^{(j)}_{T,N}}{T\min(N,T)^{-1}}\to-\infty$. Given we assume this due to the fact that $\mathbf{F}$ has more columns than $\mathbf{F}^0$, the result is proven.

Proof of Theorem 2

The result follows immediately from the proof of Theorem 1 once we note that for any set of factors, $\mathbf{F}$, (A.7) can be replaced by

$$\frac{T}{2}\ln\left[\frac{\frac{1}{T}(\mathbf{y}-\hat{\mathbf{F}}\hat{\beta})'(\mathbf{y}-\hat{\mathbf{F}}\hat{\beta})}{\frac{1}{T}(\mathbf{y}-\mathbf{F}\tilde{\beta})'(\mathbf{y}-\mathbf{F}\tilde{\beta})}\right]=O_p\left(Tq_{NT}\right)=o_p(T). \tag{A.14}$$

Table B.1: Transformation of the predictor variables

Transformation code	Transformation X_t of raw series Y_t
1	$X_t = Y_t$
2	$X_t = \Delta Y_{t,t-1}$
3	$X_t = \Delta Y_{t,t-12} - \Delta Y_{t-1,t-13}$
4	$X_t = \ln Y_t$
5	$X_t = \Delta \ln Y_{t,t-1}$
6	$X_t = \Delta \ln Y_{t,t-12} - \Delta \ln Y_{t-1,t-13}$

B Data Set

The data set used for forecasting are the monthly series from the panel of U.S. indicator series as employed in Stock and Watson (2007), but excluding our two forecast variables: PCE inflation and the (effective) federal funds rate. In order to be sure that these predictor variables are $I(0)$, the underlying raw series need to be transformed such that this is the case; generally we employ the same transformation as Stock and Watson (2007), except for the bulk of the nominal series where we follow, e.g., D'Agostino and Giannone (2006) and use first differences of twelve-month transformations of the raw series. Table B.1 summarizes our potential transformations for the raw series.

Hence, we are using as predictor variables the following 106 series, which span the sample January 1959 - July 2008 before the appropriate transformations are applied, and we refer to Stock and Watson (2007) for more details regarding data construction and sources:

Series Y_t	**Transformation:** ***(See Table B.1)***
INDUSTRIAL PRODUCTION INDEX - PRODUCTS, TOTAL	5
INDUSTRIAL PRODUCTION INDEX - FINAL PRODUCTS	5
INDUSTRIAL PRODUCTION INDEX - CONSUMER GOODS	5
INDUSTRIAL PRODUCTION INDEX - DURABLE CONSUMER GOODS	5
INDUSTRIAL PRODUCTION INDEX - NONDURABLE CONSUMER GOODS	5
INDUSTRIAL PRODUCTION INDEX - BUSINESS EQUIPMENT	5
INDUSTRIAL PRODUCTION INDEX - MATERIALS	5
INDUSTRIAL PRODUCTION INDEX - DURABLE GOODS MATERIALS	5
INDUSTRIAL PRODUCTION INDEX - NONDURABLE GOODS MATERIALS	5
INDUSTRIAL PRODUCTION INDEX - MANUFACTURING (SIC)	5
INDUSTRIAL PRODUCTION INDEX - RESIDENTIAL UTILITIES	5
INDUSTRIAL PRODUCTION INDEX - FUELS	5
NAPM PRODUCTION INDEX (PERCENT)	1
CAPACITY UTILIZATION - MANUFACTURING (SIC)	1
AVG HRLY EARNINGS, PROD WRKRS, NONFARM - GOODS-PRODUCING	6
AVG HRLY EARNINGS, PROD WRKRS, NONFARM - CONSTRUCTION	6
AVG HRLY EARNINGS, PROD WRKRS, NONFARM - MFG	6
REAL AVG HRLY EARNINGS, PROD WRKRS, NONFARM - GOODS-PRODUCING	5
REAL AVG HRLY EARNINGS, PROD WRKRS, NONFARM - CONSTRUCTION	5
REAL AVG HRLY EARNINGS, PROD WRKRS, NONFARM - MFG	5
EMPLOYEES, NONFARM - TOTAL PRIVATE	5
EMPLOYEES, NONFARM - GOODS-PRODUCING	5
EMPLOYEES, NONFARM - MINING EMPLOYEES, NONFARM - CONSTRUCTION	5
EMPLOYEES, NONFARM - MFG	5

EMPLOYEES, NONFARM - DURABLE GOODS	5
EMPLOYEES, NONFARM - NONDURABLE GOODS	5
EMPLOYEES, NONFARM - SERVICE-PROVIDING	5
EMPLOYEES, NONFARM - TRADE, TRANSPORT, UTILITIES	5
EMPLOYEES, NONFARM - WHOLESALE TRADE	5
EMPLOYEES, NONFARM - RETAIL TRADE	5
EMPLOYEES, NONFARM - FINANCIAL ACTIVITIES	5
EMPLOYEES, NONFARM - GOVERNMENT	5
INDEX OF HELP-WANTED ADVERTISING IN NEWSPAPERS (1967=100;SA)	2
EMPLOYMENT: RATIO; HELP-WANTED ADS:NO. UNEMPLOYED CLF	2
CIVILIAN LABOR FORCE: EMPLOYED, TOTAL (THOUS.,SA)	5
CIVILIAN LABOR FORCE: EMPLOYED, NONAGRIC.INDUSTRIES (THOUS.,SA)	5
UNEMPLOY.BY DURATION: AVERAGE(MEAN)DURATION IN WEEKS (SA)	2
UNEMPLOY.BY DURATION: PERSONS UNEMPL.LESS THAN 5 WKS (THOUS.,SA)	5
UNEMPLOY.BY DURATION: PERSONS UNEMPL.5 TO 14 WKS (THOUS.,SA)	5
UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 WKS + (THOUS.,SA)	5
UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 TO 26 WKS (THOUS.,SA)	5
UNEMPLOY.BY DURATION: PERSONS UNEMPL.27 WKS + (THOUS,SA)	5
AVG WKLY HOURS, PROD WRKRS, NONFARM - GOODS-PRODUCING	1
AVG WKLY OVERTIME HOURS, PROD WRKRS, NONFARM - MFG	2
HOUSING AUTHORIZED: TOTAL NEW PRIV HOUSING UNITS (THOUS.,SAAR)	4
HOUSING STARTS:NONFARM(1947-58);TOTAL FARM&NONFARM(1959-)(THOUS.,U)SA	4
HOUSING STARTS:NORTHEAST (THOUS.U.)S.A.	4
HOUSING STARTS:MIDWEST(THOUS.U.)S.A.	4
HOUSING STARTS:SOUTH (THOUS.U.)S.A.	4
HOUSING STARTS:WEST (THOUS.U.)S.A.	4
INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,3-MO.(% PER ANN,NSA)	2
INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,6-MO.(% PER ANN,NSA)	2
INTEREST RATE: U.S.TREASURY CONST MATURITIES,1-YR.(% PER ANN,NSA)	2
INTEREST RATE: U.S.TREASURY CONST MATURITIES,5-YR.(% PER ANN,NSA)	2
INTEREST RATE: U.S.TREASURY CONST MATURITIES,10-YR.(% PER ANN,NSA)	2
BOND YIELD: MOODY'S AAA CORPORATE (% PER ANNUM)	2
BOND YIELD: MOODY'S BAA CORPORATE (% PER ANNUM)	2
INTEREST RATE SPREAD: 6-MO. TREASURY BILLS MINUS 3-MO. TREASURY BILLS	1
INTEREST RATE SPREAD: 1-YR. TREASURY BONDS MINUS 3-MO. TREASURY BILLS	1
INTEREST RATE SPREAD: 10-YR. TREASURY BONDS MINUS 3-MO. TREASURY BILLS	1
INTEREST RATE SPREAD: AAA CORPORATE MINUS 10-YR. TREASURY BONDS	1
INTEREST RATE SPREAD: BAA CORPORATE MINUS 10-YR. TREASURY BONDS	1
MONEY STOCK: M1(CURR,TRAV.CKS,DEM DEP,OTHER CK'ABLE DEP)(BIL$,SA)	6
MZM (SA) FRB St. Louis	6
MONEY STOCK:M2(M1+O'NITE RPS,EURO$,G/P&B/D MMMFS&SAV&SM TIME DEP)(BIL$,SA)	6
MONETARY BASE, ADJ FOR RESERVE REQUIREMENT CHANGES(MIL$,SA)	6
DEPOSITORY INST RESERVES:TOTAL,ADJ FOR RESERVE REQ CHGS(MIL$,SA)	6
DEPOSITORY INST RESERVES:NONBORROWED,ADJ RES REQ CHGS(MIL$,SA)	6
Commercial and Industrial Loans at All Commercial Banks (FRED) Billions $ (SA)	6
CONSUMER CREDIT OUTSTANDING - NONREVOLVING(G19)	6
Personal Consumption Expenditures, Price Index (2000=100) , SAAR	6
Personal Consumption Expenditures - Durable Goods, Price Index (2000=100), SAAR	6
Personal Consumption Expenditures - Nondurable Goods, Price Index (2000=100), SAAR	6
Personal Consumption Expenditures - Services, Price Index (2000=100) , SAAR	6
PCE Price Index Less Food and Energy (SA) Fred	6
PRODUCER PRICE INDEX: FINISHED GOODS (82=100,SA)	6
PRODUCER PRICE INDEX:FINISHED CONSUMER GOODS (82=100,SA)	6
PRODUCER PRICE INDEX:INTERMED MAT.SUPPLIES & COMPONENTS(82=100,SA)	6
PRODUCER PRICE INDEX:CRUDE MATERIALS (82=100,SA)	6
Real PRODUCER PRICE INDEX:CRUDE MATERIALS (82=100,SA)	5
SPOT MARKET PRICE INDEX:BLS & CRB: ALL COMMODITIES(1967=100)	6
Real SPOT MARKET PRICE INDEX:BLS & CRB: ALL COMMODITIES(1967=100)	5
PRODUCER PRICE INDEX: CRUDE PETROLEUM (82=100,NSA)	6
PPI Crude (Relative to Core PCE)	5
NAPM COMMODITY PRICES INDEX (PERCENT)	1
UNITED STATES;EFFECTIVE EXCHANGE RATE(MERM)(INDEX NO.)	5
FOREIGN EXCHANGE RATE: SWITZERLAND (SWISS FRANC PER U.S.$)	5
FOREIGN EXCHANGE RATE: JAPAN (YEN PER U.S.$)	5
FOREIGN EXCHANGE RATE: UNITED KINGDOM (CENTS PER POUND)	5

FOREIGN EXCHANGE RATE: CANADA (CANADIAN $ PER U.S.$)	5
S&P'S COMMON STOCK PRICE INDEX: COMPOSITE (1941-43=10)	5
S&P'S COMMON STOCK PRICE INDEX: INDUSTRIALS (1941-43=10)	5
S&P'S COMPOSITE COMMON STOCK: DIVIDEND YIELD (% PER ANNUM)	2
S&P'S COMPOSITE COMMON STOCK: PRICE-EARNINGS RATIO (%,NSA)	2
COMMON STOCK PRICES: DOW JONES INDUSTRIAL AVERAGE	5
S&P'S COMPOSITE COMMON STOCK: DIVIDEND YIELD (% PER ANNUM)	2
U. OF MICH. INDEX OF CONSUMER EXPECTATIONS(BCD-83)	2
PURCHASING MANAGERS' INDEX (SA)	1
NAPM NEW ORDERS INDEX (PERCENT)	1
NAPM VENDOR DELIVERIES INDEX (PERCENT)	1
NAPM INVENTORIES INDEX (PERCENT)	1
NEW ORDERS (NET) - CONSUMER GOODS & MATERIALS, 1996 DOLLARS (BCI)	5
NEW ORDERS, NONDEFENSE CAPITAL GOODS, IN 1996 DOLLARS (BCI)	5