Armah, Nii Ayi; Swanson, Norman R.

Working Paper

Seeing inside the black box: Using diffusion index methodology to construct factor proxies in largescale macroeconomic time series environments

Working Paper, No. 2011-05

Provided in Cooperation with:
Department of Economics, Rutgers University

Suggested Citation: Armah, Nii Ayi; Swanson, Norman R. (2011) : Seeing inside the black box: Using diffusion index methodology to construct factor proxies in largescale macroeconomic time series environments, Working Paper, No. 2011-05, Rutgers University, Department of Economics, New Brunswick, NJ

This Version is available at:
https://hdl.handle.net/10419/59456

Seeing Inside the Black Box: Using Diffusion Index Methodology to Construct Factor Proxies in Largescale Macroeconomic Time Series Environments

Nii Ayi Armah* and Norman R. Swanson†

*Rutgers University
†Rutgers University and the Federal Reserve Bank of Philadelphia

February 2007; this revision July 2008

Abstract

In economics, common factors are often assumed to underlie the co-movements of a set of macroeconomic variables. For this reason, many authors have used estimated factors in the construction of prediction models. In this paper, we begin by surveying the extant literature on diffusion indexes. We then outline a number of approaches to the selection of factor proxies (observed variables that proxy unobserved estimated factors) using the statistics developed in Bai and Ng (2006a,b). Our approach to factor proxy selection is examined via a small Monte Carlo experiment, where evidence supporting our proposed methodology is presented, and via a large set of prediction experiments using the panel dataset of Stock and Watson (2005). One of our main empirical findings is that our "smoothed" approaches to factor proxy selection appear to yield predictions that are often superior not only to a benchmark factor model, but also to simple linear time series models which are generally difficult to beat in forecasting competitions. In some sense, by using our approach to predictive factor proxy selection, one is able to open up the "black box" often associated with factor analysis, and to identify actual variables that can serve as primitive building blocks for (prediction) models of a host of macroeconomic variables, and that can also serve are policy instruments, for example. Our findings suggest that important observable variables include: various S&P500 variables, including stock price indices and dividend series; a 1-year Treasury bond rate; various housing activity variables; industrial production; and exchange rates.

JEL classification: C22, C33, C51.
Keywords: diffusion index, factor, forecast, macroeconometrics, parameter estimation error, proxy.

---

*,† Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA (armah@econ.rutgers.edu and nswanson@econ.rutgers.edu). The authors wish to thank the editors, Esfandiar Maasoumi and Marcelo Medeiros, as well as two anonymous referees for providing numerous useful suggestions on earlier versions of this paper. Additionally, a great many thanks are owed to Mark Watson for making the data used in this paper available for public consumption. Thanks are also owed to Valentina Corradi, Roger Klein, Serena Ng, and participants of the conference on "Empirical Methods in Macroeconomics and Finance" at Bocconi University in October 2003, where many interesting discussions led to the idea for this paper. Swanson thanks the Rutgers University Research Council for financial support. The views expressed in this paper are solely our own, and do not necessarily reflect the official positions or policies of the Federal Reserve Bank of Philadelphia or the Federal Reserve System.

1 Introduction

The idea that individual economic variables can be forecast with some precision by refining the information from a large panel of data into a small set of estimated factors (predictors) is intriguing. It suggests that there is a small set of crucial latent factors which generate the co-movements in a large set of macroeconomic variables. This idea is consistent, for example, with the notion that a small set of underlying shocks are responsible for the dynamic behavior implicit in dynamic stochastic general equilibrium models. The practice of using observable economic variables to proxy the latent factors is espoused on the Federal Reserve Bank of New York's website: *"In formulating the nation's monetary policy, the Federal Reserve considers a number of factors, including the economic and financial indicators which follow, as well as the anecdotal reports compiled in the Beige Book. Real Gross Domestic Product (GDP); Consumer Price Index (CPI); Nonfarm Payroll Employment Housing Starts; Industrial Production/Capacity Utilization; Retail Sales; Business Sales and Inventories; Advance Durable Goods Shipments, New Orders and Unfilled Orders; Lightweight Vehicle Sales; Yield on 10-year Treasury Bond; S&P 500 Stock Index; M2"* (see *http://www.newyorkfed.org/education/bythe.html).* The recent literature on factor (diffusion index) models is rich and diverse. A very few of the most important papers include: Bai (2003), Bai and Ng (2002, 2005, 2006a,b,c,d), Boivin and Ng (2005), Connor and Korajczyk (1993), Ding and Hwang (1999), Forni, Hallin, Lippi, and Reichlin (2000, 2005), Forni and Reichlin (1996, 1998), Geweke (1977), Rapach and Strauss (2007), and Stock and Watson (1996, 1998, 1999, 2002a,b, 2004a,b, 2005).

In this paper, our purpose is twofold: first, we provide a review of the extant literature, with careful emphasis on the implementation of factor estimation and prediction using the methods of Bai and Ng as well as Stock and Watson. We then outline a simple methodology for the construction of factor proxies for use in prediction models, where our proxies are observable economic variables. In this sense, we attempt to look inside the "black box", in the sense that our proxy factors are observable and hence have clear economic meaning, while factors in general are often hard to interpret economically (see below for further discussion). As a case in point, policy makers use individual observable variables as policy instruments, for example. Our factor proxies might thus be used for policy, while estimated unobserved factors are not as obviously used in policy applications. In this sense, our main contribution is to add to the broad literature on prediction using factor

models. The methodology that we outline is very straightforward, and is based upon application of the $A(j)$ and $M(j)$ statistics developed in Bai and Ng (2006a,b). An ancillary purpose in this paper is to note that in some cases factor proxies defined as observable variables may actually perform as well as estimates of unobserved factors based on standard factor analysis. This is rather an interesting finding, suggesting for example that factor analysis should be applied with caution, particularly in cases where parameter estimation error implicit to factor construction may be great.

Following the approach of Stock and Watson (2002a,b), diffusion index forecasts involve a two-step procedure. First, the method of principal components is used to estimate the factors from a large panel of possible predictors, X. Second, the estimated factors are used to forecast the variable of interest, y_{t+1}. Stock and Watson (2002a) demonstrate that diffusion index forecasts yield encouraging results. Bai and Ng (2006a), however, point out that the regressors (factors) in the diffusion index model are estimated, hence substantially increasing the forecast error variance. In a related paper, Bai and Ng (2006b) examine whether observable economic variables can serve as proxies for the underlying unobserved factors. In particular, they use the $A(j)$ and $M(j)$ statistics to determine whether a group of observed variables yields precisely the same information as that contained in the latent factors. Stock and Watson (2002a) have also attempted to link the factors to observed variables. Thus, in some sense, Bai and Ng, as well as Stock and Watson, have already looked inside the "black box". Our approach is to take their argument one step further, and to argue that if observable economic variables are indeed good proxies of the unobserved factors, then these proxies can be used in place of the factors in the diffusion index model for prediction. Once the set of factor proxies is fixed, we effectively eliminate the incremental increase in forecast error variance (i.e., uncertainty) associated with the use of estimated factors. Along these lines, we consider "smoothed" versions of the $A(j)$ and $M(j)$ statistics that pre-select a set of factor proxies prior to the ex-ante construction of a sequence of predictions. It is worth noting that by replacing the estimated factors with observed variables, we are trading off the above variety of uncertainty with "variable selection uncertainty". Our empirical results suggest that there are cases in macro forecasting where the trade-off is worthwhile.

In a Monte Carlo experiment, we show that the $A(j)$ and $M(j)$ statistics can be used to construct prediction models that compare perform quite favorably when compared against standard factor model predictions. We additionally carry out a large variety of prediction experiments using the macroeconomic dataset of Stock and Watson (2005). In these experiments, we predict a number of

price and income variables, including industrial production, real personal income less transfers, real manufacturing and trade sales, the number of employees on non-agricultural payrolls, the consumer price index, the personal consumption expenditure implicit price deflator, and the producer price index for finished goods. Using recursively estimated models, we construct $h = 1, 3, 12$, and 24 step ahead forecasts. We show that the $A(j)$ and $M(j)$ statistics appear to offer an interesting means by which factor proxies for later use in prediction models can be chosen. Indeed, our "smoothed" approaches to factor proxy selection appear to yield predictions that are often mean square forecast error "superior" not only relative to a benchmark factor model, but also to simple linear time series models which are often difficult to beat in forecasting competitions. Furthermore, our methods based on the use of the $A(j)$ statistic appear to perform better than those based on the $M(j)$ statistic. Finally, we provide evidence that: (i) versions of our factor proxy selection method that use only a single factor proxy are preferred to those based on the use of $\widehat{k}$ proxies, where $\widehat{k}$ is a consistent estimate of the true number of factors; and (ii) while our "smoothed" proxy selection method is clearly superior for $h = 1, 3$, and 12, the method breaks down at the longest forecast horizon that we consider (i.e., $h = 24$). For the longest horizons, the estimated factor approach to prediction (e.g., that used by Stock and Watson (2002a,b)) dominates.

By using our approach to predictive factor proxy selection, we believe that we are able to "open up" the "black box" often associated with factor analysis, at least to a certain extent, and to identify actual variables that can serve as primitive building blocks for (prediction) models of a host of macroeconomic variables. Our empirical analysis suggests that important underlying observable variables, in the sense that they are good proxies for latent factors, include: the S&P500 price index and dividend series; the 1-year Treasury bond rate; various housing activity variables; industrial production; and an exchange rate.

The rest of the paper is organized as follows. In Section 2 we review the diffusion index literature, with some focus on the methods that are used in our Monte Carlo and empirical experiments. In Section 3 we discuss the use of factor proxies, including a discussion of the Bai and Ng (2006a,b) tests, and a discussion of the methodological approach to the construction and use of factor proxies for prediction. Section 4 contains a summary of the empirical methodology used in the paper, and Section 5 summarizes the data used. In Section 6, the results of a small Monte Carlo experiment studying the finite sample properties of the Bai and Ng (2006a,b) tests are presented, and in Section 7 we summarize our empirical findings. Finally, in Section 8 we briefly discuss the most

recent advances in the diffusion index methodology; and concluding remarks are gathered in Section 9.

2 Review: Diffusion Index Models and the Principle Components Approach to Estimation

2.1 The diffusion index model

Following Stock and Watson (2002a,b), let y_{t+1}be the series we wish to forecast and X_t be an N-dimensional vector of predictor variables, for $t = 1, \ldots, T$. Assume that (y_{t+1}, X_t) has a dynamic factor model representation with $\overline{r}$ common dynamic factors, f_t. Hence, f_t is an $\overline{r} \times 1$ vector. The dynamic factor model is written as:

$$y_{t+h} = \alpha(L) f_t + \beta' W_t + \varepsilon_{t+h} \tag{1}$$

and

$$x_{it} = \lambda_i(L) f_t + e_{it}, \tag{2}$$

for $i = 1, 2, \ldots, N$, where W_t is an $l \times 1$ vector of other observable variables with $l << N$, such as contemporaneous and lagged values of y_t; $h > 0$ is the lead time between information available and the dependent variable; x_{it} is a single datum for a particular predictor variable; e_{it} is the idiosyncratic shock component of x_{it}; and $\alpha(L)$ and $\lambda_i(L)$ are lag polynomials in nonnegative powers of L. In general, dynamic factor models can be transformed into static factor models. In Stock and Watson (2002a), the lag polynomials $\alpha(L)$ and $\lambda_i(L)$ are modeled as $\alpha(L) = \sum_{j=0}^{q} \alpha_j L^j$ and $\lambda_i(L) = \sum_{j=0}^{q} \lambda_{ij} L^j$. The finite order of the lag polynomials allows us to rewrite (1) and (2) as:

$$y_{t+h} = \alpha' F_t + \beta' W_t + \varepsilon_{t+h} \tag{3}$$

and

$$x_{it} = \Lambda_i' F_t + e_{it}, \tag{4}$$

where $F_t = (f_t', \ldots, f_{t-q}')'$ is an $r \times 1$ vector, with $r = (q+1)\overline{r}$ and α is an $r \times 1$ vector. Here, r is the number of static factors (i.e., the number of elements in F_t). Additionally, $\Lambda_i = (\lambda_{i0}', \ldots, \lambda_{iq}')'$

is a vector of factor loadings on the r static factors, where λ_{ij} is an $\overline{r} \times 1$ vector for $j = 0, \ldots, q$ and $\beta = (\beta_1, \ldots, \beta_l)'$. Alternatively, from (2), the dynamic factor model can be represented as:

$$\begin{aligned} x_{it} &= \lambda_{i0}' f_t + \lambda_{i1}' f_{t-1} + \ldots + \lambda_{iq}' f_{t-q} + e_{it} \\ &= \lambda_i'(L) f_t + e_{it} \end{aligned} \tag{5}$$

and:

$$\lambda_i(L) = \lambda_{i0} + \lambda_{i1} L^1 + \ldots + \lambda_{iq} L^q.$$

For complete details, see Bai and Ng (2005). Now, (5) can be written in the static form (4) where F_t and Λ_i are defined as above. The static factor model refers to the contemporaneous relationship between x_{it} and F_t. One major advantage of the static representation of the dynamic factor model is it enables us to use principal components to estimate the factors. This involves estimating F_t using an eigenvalue-eigenvector decomposition of the sample covariance matrix of the data. It is worth noting that the use of principal components to estimate the factors cannot be done with infinitely distributed lags of the factors (see Stock and Watson (2002a)). Ding and Hwang (1999), Forni et al. (2000), Stock and Watson (2002b), Bai and Ng (2002) and Bai (2003) showed that the space spanned by both the static and dynamic factors can be consistently estimated when N and T are both large. For forecasting purposes, little is gained from a clear distinction between the static and the dynamic factors. However, many economic analyses hinge on the ability to isolate the primitive shocks or the number of dynamic factors (see Bai and Ng (2007)). Boivin and Ng (2005) also compare alternative factor based forecast methodologies, and conclude that when the dynamic structure is unknown and the model is characterized by complex dynamics, the approach of Stock and Watson performs favorably. If the idiosyncratic errors, $e_t = (e_{1t}, \ldots, e_{Nt})'$, are cross-sectionally independent and i.i.d. over time, then (4) is the classical factor analysis model. It is important at this juncture to note that the factor model does not generally require the idiosyncratic errors to be cross-sectionally independent (see e.g., Bai and Ng (2002)). This is a crucial departure, as it ensures that we can assume the existence of an "approximate" rather than "strict" factor model. (Moreover, the idiosyncratic errors are restricted to be "weakly" correlated, roughly speaking, as the basic structure of the factor model requires the factors to account for the "bulk" of the co-movement across variables). Of final note, it should be mentioned that Geweke (1977) and Sargent and Sims (1977) were among the first to extend the classical factor analysis model to dynamic models.

Following Bai and Ng (2002), let $\underset{(T\times 1)}{\underline{X}_i}$ be a $T \times 1$ vector of observations for the ith variable. For a given cross-section i, we have $\underset{(T\times 1)}{\underline{X}_i} = \underset{(T\times r)(r\times 1)}{F^0}\ \Lambda_i + \underset{(T\times 1)}{\underline{e}_i}$ where $\underline{X}_i = (X_{i1}, \ldots, X_{iT})'$, $F^0 = (F_1, \ldots, F_T)'$ and $\underline{e}_i = (e_{i1}, \ldots, e_{iT})'$. The whole panel of data $X = (\underline{X}_1, \ldots, \underline{X}_N)$ can consequently be represented as $\underset{(T\times N)}{X} = \underset{(T\times r)(r\times N)}{F^0}\ \Lambda' + \underset{(T\times N)}{e}$, where $\Lambda = (\Lambda_1, \ldots, \Lambda_N)'$ and $e = (\underline{e}_1, \ldots, \underline{e}_N)$. Connor and Korajczyk (1986, 1988, 1993) (1996, 1998) and Forni, Hallin, Lippi and Reichlin (2000) Stock and Watson (2002b) We will also assume $\{F_t\}$ and $\{e_{it}\}$ are two groups of mutually independent stochastic variables. Furthermore, it is well known that for $\Lambda F_t = \Lambda Q Q^{-1} F_t$, a normalization is needed in order to uniquely define the factors, where Q is a nonsingular matrix. Now, assuming that $(\Lambda'\Lambda/N) \to I_r$, we restrict Q to be orthonormal, for example. This assumption, together with others noted in Stock and Watson (2002b), enables us to identify the factors up to a change of sign and consistently estimate them up to an orthonormal transformation. Forecasts of y_{T+h} based on (3) and (4) involve a two step procedure because both the regressors and coefficients in the forecasting equations are unknown. The data sample $\{X_t\}_{t=1}^T$ are first used to estimate the factors, $\{\widetilde{F}_t\}_{t=1}^T$ by means of principal components. With the estimated factors in hand, we obtain the estimators $\widehat{\alpha}$ and $\widehat{\beta}$ by regressing y_{t+h} onto $\widetilde{F}_t$ and the observable variables in W_t. Of note is that if $\sqrt{T}/N \to 0$, then the generated regressor problem does not arise, in the sense that least squares estimates of $\widehat{\alpha}$ and $\widehat{\beta}$ are $\sqrt{T}$ consistent and asymptotically normal (see Bai and Ng (2005)).

2.2 Common factor estimation using principal components

The problem of obtaining the necessary estimates in (4) would be simplified if we knew F^0. Then Λ_i could be estimated via least squares by setting $\{x_{it}\}_{t=1}^T$ to be the dependent variable and $\{F_t\}_{t=1}^T$ to be the explanatory variable. On the other hand, if Λ were known, F_t could be estimated by regressing $\{x_{it}\}_{i=1}^N$ on $\{\Lambda_i\}_{i=1}^N$. Since the common factors are not observed, in the regression analysis of (4), we replace F_t by $\widetilde{F}_t$, estimates that span the same space as F_t when $N, T \to \infty$. Estimation of these common factors from large panel data sets of macroeconomic variables can be carried out using principal component analysis. We refer the reader to Stock and Watson (1998, 2002a, 2002b, 2004a, 2004b) and Bai and Ng (2002) for a detailed explanation of this procedure, and to Connor and Korajczyk (1986, 1988, 1993), Forni and Reichlin (1996, 1998) and Forni, Hallin, Lippi and Reichlin (2000) for further detailed discussion of diffusion models, in general.

As noted earlier F_t and λ_i are not separately identified, but rather identifiable only up to a

square matrix. Stock and Watson (1998) further demonstrate that when principal components is used, the factors remain consistent even when there is some time variation in Λ and small amounts of data contamination, so long as the number of variables in the panel data set or the number of predictors is very large (i.e., $N >> T$). In this paper, we only give an outline of how principal component analysis is carried out, with particular emphasis on those features of the analysis that allow us to carry out our prediction experiments using the $A(j)$ and $M(j)$ statistics of Bai and Ng (2006b).

Let k ($k < \min\{N,T\}$) be an arbitrary number of factors, Λ^k be the $N \times k$ matrix of factor loadings, $(\Lambda_1^k, \ldots, \Lambda_N^k)'$, and F^k be a $T \times k$ matrix of factors $(F_1^k, \ldots, F_T^k)'$. From (4), estimates of Λ_i^k and F_t^k are obtained by solving the optimization problem:

$$V(k) = \min_{\Lambda^k, F^k} (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \Lambda_i^{k\prime} F_t^k)^2 \tag{6}$$

Let $\widetilde{F}^k$ and $\widetilde{\Lambda}^k$ be the minimizers of equation (6). Since Λ^k and F^k are not separately identifiable, if $N > T$, a computationally expedient approach would be to concentrate out $\widetilde{\Lambda}^k$ and minimize (6) subject to the normalization $F^{k\prime}F^k/T = I_k$. Minimizing (6) is equivalent to maximizing $tr[F^{k\prime}(XX')F^k]$. This optimization is solved by setting $\widetilde{F}^k$ to be the matrix of the k eigenvectors of XX' that correspond to the k largest eigenvalues of XX'. Note that $tr[\cdot]$ represents the matrix trace. The superscript in Λ^k and F^k signifies the use of k factors in the estimation and the fact that the estimates will depend on k. Let $\widetilde{D}$ be a $k \times k$ diagonal matrix consisting of the k largest eigenvalues of XX'. The estimated factor matrix, denoted by $\widetilde{F}^k$, is $\sqrt{T}$ times the eigenvectors corresponding to the k largest eigenvalues of the $T \times T$ matrix XX'. Given $\widetilde{F}^k$ and the normalization $F^{k\prime}F^k/T = I_k$, $\widetilde{\Lambda}^{k\prime} = (\widetilde{F}^{k\prime}\widetilde{F}^k)^{-1}\widetilde{F}^{k\prime}X = \widetilde{F}^{k\prime}X/T$ is the corresponding factor loadings matrix.

The solution to the optimization problem in (6) is not unique. If $N < T$, it becomes computationally advantageous to concentrate out $\overline{F}^k$ and minimize (6) subject to $\overline{\Lambda}^{k\prime}\overline{\Lambda}^k/N = I_k$. This minimization is the same as maximizing $tr[\Lambda^{k\prime}X'X\Lambda^k]$, the solution of which is to set $\overline{\Lambda}^k$ equal to the eigenvectors of the $N \times N$ matrix $X'X$ that correspond to its k largest eigenvalues. One can consequently estimate the factors as $\overline{F}^k = X'\overline{\Lambda}^k/N$. $\widetilde{F}^k$ and $\overline{F}^k$ span the same column spaces, hence for forecasting purposes, they can be used interchangeably depending on which one is more computationally efficient. Given $\widetilde{F}^k$ and $\widetilde{\Lambda}^k$, let $\widehat{V}(k) = (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \widetilde{\Lambda}_i^{k\prime}\widetilde{F}_t^k)^2$ be the sum of squared residuals from regressions of X_i on the k factors, $\forall i$. A penalty function for over fitting,

$g(N,T)$, is chosen such that the loss function

$$IC(k) = \log(\widehat{V}(k)) + kg(N,T) \tag{7}$$

can consistently estimate r. Let $kmax$ be a bounded integer such that $r \leq k\max$. Bai and Ng (2002) propose three versions of the penalty function $g(N,T)$. Namely $g_1(N,T) = \left(\frac{N+T}{NT}\right)\log\left(\frac{NT}{N+T}\right)$, $g_2(N,T) = \left(\frac{N+T}{NT}\right)\log C_{NT}^2$, and $g_3(N,T) = \left(\frac{\log(C_{NT}^2)}{C_{NT}^2}\right)$, all of which lead to consistent estimation of r. In our empirical and Monte Carlo experiments, we use $g_2(N,T)$. Of note is that we tried the other penalty functions above, and our results were qualitatively the same. However, Bai and Ng (2002), as well as others, have shown that in certain contexts, results are sensitive to the choice of penalty function. Hence, (7) becomes:

$$IC(k) = \log(\widehat{V}(k)) + k(\frac{N+T}{NT})\log C_{NT}^2$$

where $C_{NT} = \min\{\sqrt{N}, \sqrt{T}\}$. The consistent estimate of the true number of factors is then:

$$\widehat{k} = \arg \min_{0 \leq k \leq k\max} IC(k), \tag{8}$$

and $\lim_{N,T\to\infty}\text{Prob}[\widehat{k} = r] = 1$ if $g(N,T) \to 0$ and $C_{NT}^2 \cdot g(N,T) \to \infty$ as $N,T \to \infty$, as shown in Bai and Ng (2002).

3 Using Proxies In Place of Factors for Prediction

3.1 Prediction using factors

Reconsider the general equation (3), $y_{t+h} = \alpha' F_t + \beta' W_t + \varepsilon_{t+h}$. As mentioned above, and shown in Stock and Watson (2002b) and Bai and Ng (2005), under a set of moment conditions on (ε, e, F^0) and an asymptotic rank condition on Λ, if the space spanned by F_t can be consistently estimated, then $\sqrt{T}$ consistent estimates of α and β are obtainable. Under a similar set of conditions, it is also possible to obtain $\min[\sqrt{N}, \sqrt{T}]$ consistent forecasts if $\sqrt{T/N} \to 0$ as $N,T \to \infty$. Let $z_t = (F_t', W_t')'$; $E(\varepsilon_{t+h}|y_t, z_t, y_{t-1}, z_{t-1}, \dots) = 0$, for any $h > 0$; and let z_t and ε_t be independent of the idiosyncratic errors e_{is}, $\forall i, s$. If F_t is observable and α and β are known, based on the above assumption that the mean of ε_{t+h} conditional on past information is zero, the conditional mean and minimum mean square error forecast of y_{T+h} is given by:

$$y_{T+h|T} = E(y_{T+h}|z_T, z_{T-1}, \dots) = \alpha' F_T + \beta' W_T \equiv \delta' z_T$$

Such a prediction is not feasible, however, since α, β and F_t are all unobserved. The feasible prediction that replaces the unknown objects by their estimates is:

$$\widehat{y}_{T+h|T} = \widehat{\alpha}'\widetilde{F}_T + \widehat{\beta}'W_T = \widehat{\delta}'\widehat{z}_T, \tag{9}$$

where $\widehat{z}_t = (\widetilde{F}_t', W_t')'$. Here, $\widehat{\alpha}$ and $\widehat{\beta}$ are the least squares estimates obtained from regressing y_{t+h} on $\widetilde{F}_t$ and W_t, $t = 1, \ldots, T-h$. We suppress the k superscript on $\widetilde{F}_t^k$ because we assume we have consistently estimated the number of factors underlying the dataset. The factors, F_t, are estimated from x_{it} by the method of principal components, as discussed above. As the objective is to forecast y_{T+h}, a crucial aspect of our analysis is the distribution of the forecast error. As explained in detail in Bai and Ng (2006a), since $y_{T+h} = y_{T+h|T} + \varepsilon_{T+h}$, it follows that the forecast error is:

$$\widehat{\varepsilon}_{T+h} \equiv \widehat{y}_{T+h|T} - y_{T+h} = (\widehat{y}_{T+h|T} - y_{T+h|T}) - \varepsilon_{T+h}$$

If $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$, then:

$$\widehat{\varepsilon}_{T+h} \sim N(0, \sigma_\varepsilon^2 + var(\widehat{y}_{T+h|T})) \tag{10}$$

where

$$var(\widehat{y}_{T+h|T}) = \frac{1}{T}\widehat{z}_T' Avar(\widehat{\delta})\widehat{z}_T + \frac{1}{N}\widehat{\alpha}' Avar(\widetilde{F}_T)\widehat{\alpha}. \tag{11}$$

Here, $var(\widehat{y}_{T+h|T})$ reflects both parameter uncertainty and regressor uncertainty. In large samples, $var(\widehat{\varepsilon}_{T+h})$ is dominated by σ_ε^2. If we ignore $var(\widehat{y}_{T+h|T})$, σ_ε^2 alone will under-estimate the true forecast uncertainty for finite T and N. Let us now assume for a moment that F_t is observable. The feasible prediction of y_{T+h} would then be $\overline{y}_{T+h|T} = \overline{\alpha}'F_T + \overline{\beta}'W_T = \overline{\delta}'z_T$, where $\overline{\alpha}$ and $\overline{\beta}$ are the least squares estimates obtained from regressing y_{t+h} on F_t and W_t. Once again, since $y_{T+h} = y_{T+h|T} + \varepsilon_{T+h}$, the forecast error is:

$$\overline{\varepsilon}_{T+h} = \overline{y}_{T+h|T} - y_{T+h} = (\overline{y}_{T+h|T} - y_{T+h|T}) - \varepsilon_{T+h}$$

If $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$, then

$$\overline{\varepsilon}_{T+h} \sim N(0, \sigma_\varepsilon^2 + var(\overline{y}_{T+h|T})), \tag{12}$$

where

$$var(\overline{y}_{T+h|T}) = \frac{1}{T}z_T' Avar(\overline{\delta})z_T. \tag{13}$$

Thus, and as discussed by Bai and Ng (2006a), when comparing $var(\overline{y}_{T+h|T})$ with $var(\widehat{y}_{T+h|T})$, it is clear that estimating the factors increases the forecast error variance, $var(\widehat{y}_{T+h|T})$, by $\frac{1}{N}\widehat{\alpha}' Avar(\widetilde{F}_T)\widehat{\alpha}$. Of course, if we could observe the factors instead of estimating them, we would reduce the forecast error variance from (10) to (12). In finite samples, this may yield important prediction error variance reduction. It is for this reason that we consider replacing the factors in (9) with observable variables that closely proxy the factors. The approach taken in order to do this involves implementing a "first stage" factor analysis in which proxies are formed using the $A(j)$ and $M(j)$ statistics of Bai and Ng (2006b). In a "second stage", the observable proxies are used in the construction of a prediction model. In this way, all estimation error associated with the factor analysis and proxy selection is essentially "hidden" in the first stage, and does not directly manifest itself in the "second stage" prediction models and prediction errors. Put another way, we are trading-off "estimated factor uncertainty" for "variable selection uncertainty" (see introduction for further discussion). Of course, issues related to "pre-testing" and sequential testing bias still arise. Nevertheless, in our prediction experiments we attempt to quantify through finite sample experiments the potential gains to using the "proxy" approach.

3.2 Using the $A(j)$ and $M(j)$ tests of Bai and Ng (2006b) to uncover factor proxies

For a detailed theoretical discussion of the results presented in this subsection, the reader is referred to Bai and Ng (2006b). Here, we draw heavily on aspects of that paper that are relevant to our empirical implementation. Note that while Bai and Ng (2006b) suggest using the $A(j)$ and $M(j)$ statistics to assess whether key business cycle indicators approximate the latent factors, we use the $A(j)$ and $M(j)$ statistics to select factor proxies for subsequent use in prediction models. The $A(j)$ statistic depends on the actual size of a t-test. The $M(j)$ test is based on a measure of the distance between observed variables and factor estimates thereof.

Suppose we observe G', a $(T \times m)$ matrix of observable economic variables that could potentially proxy the latent factors (i.e., G is an $m \times T$ matrix). At any given time t, any of the m elements of G_t $(m \times 1)$ will be a good proxy if it is a linear combination of the $r \times 1$ latent factors, F_t. Let G_{jt} be an element of the m vector G_t. The null hypothesis is that G_{jt} is an exact proxy, or more precisely, $\exists\ \theta_j$ $(r \times 1)$ such that $G_{jt} = \theta_j' F_t$. In order to implement all of the methods, consider the

regression $G_{jt} = \gamma_j' \widetilde{F}_t + \rho_t$. Let $\widehat{\gamma}_j$ be the least squares estimate of γ_j and let $\widehat{G}_{jt} = \widehat{\gamma}_j' \widetilde{F}_t$. The test is carried out by constructing the following t-statistic:

$$\tau_t(j) = \frac{(\widehat{G}_{jt} - G_{jt})}{(\widehat{var}(\widehat{G}_{jt}))^{1/2}} \tag{14}$$

where

$$\begin{aligned} \widehat{var}(\widehat{G}_{jt}) &= \frac{1}{N}\widehat{\gamma}_j' \widetilde{D}^{-1} \left(\frac{\widetilde{F}'\widetilde{F}}{T}\right) \widetilde{\Gamma}_t \left(\frac{\widetilde{F}'\widetilde{F}}{T}\right) \widetilde{D}^{-1}\widehat{\gamma}_j \\ &= \frac{1}{N}\widehat{\gamma}_j' \widetilde{D}^{-1} \widetilde{\Gamma}_t \widetilde{D}^{-1}\widehat{\gamma}_j, \end{aligned} \tag{15}$$

and $\widetilde{\Gamma}_t$ is defined below. The last step above is due to the normalization that $\widetilde{F}'\widetilde{F}/T = I_{\widehat{k}}$. Once again, $\widetilde{D}$ is a $k \times k$ diagonal matrix consisting of the k largest eigenvalues of XX'. Given the null hypothesis that $G_{jt} = \theta_j' F_t$ and that $\widehat{G}_{jt}$ converges to G_{jt} at rate $\sqrt{N}$, Bai and Ng (2006b) show that the limiting distribution of $\sqrt{N}(\widehat{G}_{jt} - G_{jt})$ is asymptotically normal and hence $\tau_t(j)$ has a standard normal limiting distribution. Consistent choices for the the $\widehat{k} \times \widehat{k}$ matrix $\widetilde{\Gamma}_t$ include the following:

$$\widetilde{\Gamma}_t^1 = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\widetilde{\Lambda}_i\widetilde{\Lambda}_j' \frac{1}{T}\sum_{t=1}^{T}\widetilde{e}_{it}\widetilde{e}_{jt} \ , \ \forall \ t, \tag{16}$$

$$\widetilde{\Gamma}_t^2 = \frac{1}{N}\sum_{i=1}^{N}\widetilde{e}_{it}^2\widetilde{\Lambda}_i\widetilde{\Lambda}_i', \tag{17}$$

and

$$\widetilde{\Gamma}^3 = \widehat{\sigma}_e^2 \frac{\widetilde{\Lambda}'\widetilde{\Lambda}}{N}, \tag{18}$$

where $\widehat{\sigma}_e^2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\widetilde{e}_{it}^2$, $\widetilde{e}_{it} = x_{it} - \widetilde{\Lambda}_i'\widetilde{F}_t$ and $\frac{n}{\min[N,T]} \to 0$ as $N, T \to \infty$. In our Monte Carlo simulation and our empirical analysis, we choose $n = \min\{\sqrt{N}, \sqrt{T}\}$. Equation (16) allows cross-section correlation but assumes time-series stationarity of e_{it}. This covariance estimator is a HAC type estimator because it is robust to cross-correlation (see Bai and Ng (2006a) for complete details). Equation (17) allows for time-series heteroskedasticity, but assumes no cross-sectional correlation of e_{it}. Equation (18) assumes no cross-sectional correlation and constant variance, $\forall i$ and $\forall t$. For small cross-sectional correlation in e_{it}, Bai and Ng (2006a) found that constraining the correlations to be zero could sometimes be desirable. In this regard, they make the point that (17) and (18) are useful even if residual cross-correlation is genuinely present.

As mentioned earlier, $\tau_t(j)$ in (14) has a standard normal limiting distribution. Let Φ_ξ^τ be the ξ percentage point of the limiting distribution of $\tau_t(j)$. The hypothesis test based on the t-statistic in (14) enables us to determine whether an observed value of a candidate variable is a good proxy at a specific time t. For our purposes however, given information up to time T, whatever methods or procedures we use to select the proxies ought to select whole time series G_j, for which G_{jt} satisfies the null hypothesis, $\forall t$. In this regard, our first proxy selection method is based upon the following statistic:

$$A(j) = \frac{1}{T}\sum_{t=1}^{T} 1(|\tau_t(j)| > \Phi_\xi^\tau). \tag{19}$$

The $A(j)$ statistic is the actual size of the test (i.e., the probability of Type I error given the sample size). Since $\tau_t(j)$ is asymptotically standard normal and the test is a two-tailed test, the actual size, $A(j)$, of the t-test should converge to the nominal size (the desired significance level is 2ξ) as $T \to \infty$. This means that if a candidate variable is a good proxy of the underlying factors of a data set, the $A(j)$ statistic calculated from its sample time series should approach 2ξ as the sample size increases. This is the basis on which we use the $A(j)$ statistic to select proxies. It should be noted that the $A(j)$ statistic does not constitute a test in the strict sense since we do not compare a test statistic to a critical value to determine whether or not to reject a null hypothesis. Rather, this procedure gives a ranking of the proxies with the best proxy having an $A(j)$ statistic value closest to 2ξ. In our implementation, the candidate set of proxies, G', is the same as the the panel data set X from which we estimate the factors. Given the choice of the significance level 2ξ, the $A(j)$ statistic incorporates some degree of robustness by allowing G_{jt} to deviate from $\widehat{G}_{jt}$ for a specified number of time points.

The second method for selecting the proxies considers the statistic:

$$M(j) = \max_{1\le t\le T} |\tau_t(j)|, \tag{20}$$

which is based on a measure of how far the $\widehat{G}_{jt}$ curve is from G_{jt}. If e_{it} is serially uncorrelated, then:

$$P(M(j) \le x) \approx [2\Phi(x) - 1]^T, \tag{21}$$

where $\Phi(x)$ is the cdf of a standard normal random variable. From (20) and (21), we can perform a test to determine whether the time series of a candidate variable is a good proxy for the latent

factors. For instance, suppose we are given a significance level 2ξ and a sample of size T from a particular candidate variable, G_j. From the right hand side of (21), we can calculate the corresponding critical value, x, for the test. For the same sample, we calculate $M(j)$ from (20) and conclude that G_j is a good proxy if $M(j) \leq x$, and a bad proxy otherwise. The test based on the $M(j)$ statistic is thus stronger than the selection method based on the $A(j)$ statistic, as the $M(j)$ test gives a sharp decision rule. However, the $M(j)$ test has at least one disadvantage. It requires e_{it} to be serially uncorrelated. We ignore this requirement in our experimental analysis. It should be noted that x increases with the sample size, T. Depending on the nature of the observed sample, this fact could either preserve or reduce the power of the $M(j)$ test.[1]

The proxies selected depend on the structure of the $\widehat{k} \times \widehat{k}$ matrix $\widetilde{\Gamma}_t$ that we use in (15). For a given proxy selection method, if the choice of $\widetilde{\Gamma}_t^1, \widetilde{\Gamma}_t^2, \widetilde{\Gamma}^3$ used in calculating (15) all produce the same proxies, it could mean that the respective assumptions associated with the use of $\widetilde{\Gamma}_t^1, \widetilde{\Gamma}_t^2, \widetilde{\Gamma}^3$ might not be very relevant, empirically. We found no gains, in our experimental set-up, to using $\widetilde{\Gamma}_t^1$ and $\widetilde{\Gamma}^3$, and hence all reported results are for the case where we use $\widetilde{\Gamma}_t^2$

Finally, Shanken (1992) points out that it is theoretically crucial for the observed selected proxies to span the same space as the r latent factors, as discussed above. We nevertheless consider versions of the above methods where the number of factors is greater than the number of proxies, given the principle of parsimony in forecasting.

3.3 Smoothed $A(j)$ and $M(j)$ tests for selecting factor proxies

The $A(j)$ and $M(j)$ statistics discussed above may yield a different set of proxies at each point in time when used to construct a sequence of recursive forecasts. Namely, if the information set used in the parameterization of a prediction model is updated prior to the construction of each new forecast for some sequence of E ex ante predictions, then the "first stage" factor analysis discussed above may yield a sequence of E different vectors of factor proxies. Thus, in addition to the $A(j)$ and $M(j)$ proxy selection methods, we also consider a version of these methods where the sample period in our empirical analysis is broken into three subsamples (R_1, R_2, and E, such that $T = R_1 + R_2 + E$). The first subsample is used to estimate proxies. Thereafter, one observation from R_2 is added, and this new larger sample is used to recursively select a second set of factor proxies. This is

[1] Note that we also considered the confidence interval approach of Bai and Ng (2006); but it did not perform better than the above methods.

continued until the second subsample is exhausted, yielding a sequence of R_2 different vectors of factor proxies. Individual proxies are then ranked according to their selection frequency, and those occurring the most frequently are selected and fixed for further use in constructing E ex ante predictions. As some of our models (such as the autoregressive model) select the number of lags and re-estimate all parameters prior to the formation of each new prediction, this smoothed approach is at a disadvantage, in the sense that it is static (i.e., the set of proxies is fixed throughout the forecast experiment). However, loading parameters for the proxies are still re-estimated prior to the formation of each new recursive prediction. Of course, the potential advantage to this approach is that noise across the proxy selection process is suppressed.

4 Empirical Methodology

In order to assess the performance of factor proxy based prediction models, we focus our attention on direct multistep-ahead predictions. Forecasts are generated as h-step ahead predictions of y_t, say. Namely, we predict $y_{t+h} = \log\left(\frac{Y_{t+h}}{Y_{t+h-1}}\right)$, where Y_t is the variable of interest.[2] Our approach is to compare the performance of factor based predictions with a host of proxy based predictions as well as various "strawman" predictions. For the "strawman" forecast models, we use an autoregressive ($AR(p)$) model (with lags selected using the Schwarz Information Criterion (SIC)) and a random walk model. The "strawman" models are included because they serve as parsimonious benchmarks that are often difficult to outperform. In Table 1, we provide the specifications and brief descriptions of all of the forecast models examined.

We consider two classes of proxy forecasts models. The first class of models, which we call "ordinary" proxy forecast models, include Model 4 - Model 7. With these models, proxies are re-selected recursively, prior to the construction of each h-step ahead prediction. Let $\{A(j)\}_{j=1}^{m}$, be a set of $A(j)$ statistics calculated for each candidate proxy variable j. As suggested above, in this particular paper, we set $m = N$; but this need not always be the case. Define:

$$S^A = \{G^A_{j1}, \ldots, G^A_{j\widehat{k}}\} \tag{22}$$

[2]While cummulative changes are very useful in prediction contexts, we predict the growth rate from one period to the next, $y_{t+h} = log(Y_{t+h}/Y_{t+h-1})$ instead of the cumulative change, $y_{t+h} = log(Y_{t+h}/Y_t)$. Our approach is in accord with the Federal Reserve Economic Database (FRED), where the same period on period growth rates are reported. We have experimented also with cummulative growth rates, with similar empirical findings to those reported here.

where $\widehat{k} \leq m$ and $|A(j_1) - 2\xi| \leq |A(j_2) - 2\xi| \leq \ldots \leq |A\left(j_{\widehat{k}}\right) - 2\xi|$.Here, S^A is the set of $\widehat{k}$ proxy variables selected via implementation of the $A(j)$ test. Further, define $G^A_{j_1}$ as the "best" possible proxy as determined by the $A(j)$ while $G^A_{j_2}$ is the next "best" proxy, and so on. Recall that G_j is an observable time series variable, such as the CPI or the Federal Funds Rate. Turning next to proxies selected via implementation of the $M(j)$ test, define:

$$S^M = \{G_j \in G \mid M(j) \leq x\}, \quad j = 1, \ldots, m.$$

Here, S^M is a set of proxies selected by the $M(j)$ test. The number of proxy variables selected at each recursive stage is indeterminate. Furthermore, the selected proxies are not ranked. For Model *6*, where the $M(j)$ test is used to select a single proxy, our approach is to select the proxy in the set S^M that is associated with the smallest value of $M(j)$.

The second class of models, which we call "smoothed" proxy forecast models, are discussed in Section 3.3, and include Model *8* - Model *15*. The proxies used in these models are still based on implementing the $A(j)$ and $M(j)$ statistics as discussed above. The factors and the proxies are estimated recursively, just as in Models *1, 4-7,* but this is done starting with R_1 observations and ending with $R_1 + R_2$ observations. The "smoothed" proxies are selected as the $\widehat{k}$ proxies that are "most frequently" picked by the $A(j)$ and $M(j)$ tests. Thereafter, all proxies are fixed, although their "weights" in the prediction models are still re-estimated recursively, prior to the construction of each of the E ex-ante forecasts. To differentiate between proxies picked using the "ordinary" and "smoothed" versions of the tests, we define S^{A*} and S^{M*} to be the "smoothed" versions of S^A and S^M. The ex-ante prediction period, E, is the same for all models in our empirical experiments.

In order to evaluate forecast performance, we compare mean squared forecast errors (MSFEs) defined as $\frac{1}{E}\sum_{t=R-h+1}^{T-h}\left(\widehat{y}_{t+h} - y_{t+h}\right)^2$, where $R = R_1 + R_2$. We also carry out Diebold and Mariano (DM: 1995) predictive accuracy tests. Let $\{\widehat{y}_{1,t}\}_{t=R-h+1}^{T-h}$ and $\{\widehat{y}_{2,t}\}_{t=R-h+1}^{T-h}$ be two forecasts of the time series $\{y_t\}_{t=R-h+1}^{T-h}$. The "benchmark" is Model *1* (i.e., the factor model), and is used to generate $\{\widehat{y}_{1,t}\}_{t=R-h+1}^{T-h}$, while Models *2-15* are used to generate $\{\widehat{y}_{2,t}\}_{t=R-h+1}^{T-h}$. Since the "benchmark" contains estimated factors and the alternative models contain no estimated factors, the "benchmark" and alternative models are non-nested. The corresponding out-of-sample forecast errors are $\{\widehat{\varepsilon}_{1,t}\}_{t=R-h+1}^{T-h}$ and $\{\widehat{\varepsilon}_{2,t}\}_{t=R-h+1}^{T-h}$. The null hypothesis of equal forecast accuracy for two forecasts is given by $H_0 : E[\widehat{\varepsilon}^2_{1,t}] = E[\widehat{\varepsilon}^2_{2,t}]$ or $H_0 : E[\widehat{d}_t] = 0$,where $\widehat{d}_t = \widehat{\varepsilon}^2_{1,t} - \widehat{\varepsilon}^2_{2,t}$ is the loss differential series. The DM test statistic is $DM = E^{-1/2}\frac{\overline{d}}{\sqrt{\widehat{\sigma}^2_{\overline{d}}}}$, where $\overline{d} = \frac{1}{E}\sum_{t=R-h+1}^{T-h}\widehat{d}_t$, and $\widehat{\sigma}^2_{\overline{d}}$ is a HAC

standard error for $\widehat{d}_t$. Since the forecast models are non-nested, and assuming that parameter estimation error vanishes, the DM test statistic has a $N(0,1)$ limiting distribution. Finally, given this setup, a negative DM t-stat indicates that the factor model yields a lower point MSFE. For further discussion of parameter estimation error and nestedness issues in the context of predictive accuracy tests, the reader is referred to Corradi and Swanson (2002, 2006a, 2006b).

5 Data

The dataset used to estimate the factors is the same as that used in Stock and Watson (2005), which can be obtained at *http://www.princeton.edu/˜mwatson*. This dataset contains 132 monthly time series for the United States for the entire period from 1960:1 to 2003:12, hence $N = 132$ and $T = 528$ observations. The series were selected to represent the following categories of macroeconomic time series: real output and income; employment, manufacturing and trade sales; consumption; housing starts and sales; real inventories and inventory-sales ratios; orders and unfilled orders; stock price indices; exchange rates; interest rate spreads; money and credit quantity aggregates; and price indexes. Most of the series were taken from the Global Insights Basic Economic Database or The Conference Board's Indicators Database (TCB). Others were calculated by Stock and Watson with information from either Global Insights or TCB or both. The theory outlined assumes that the panel dataset used to estimate the factors is $I(0)$. To achieve this, some of the 132 series were subjected to transformations by taking logarithms and/or first differencing. In general, logarithms were taken for all nonnegative series that were not already in rates or percentage units (see Stock and Watson (2002a,2005) for complete details). After these transformations were carried out, all series were further standardized to have sample mean zero and unit sample variance. Using the transformed data set, denoted above by X, the factors are estimated by the method of principal components. As mentioned earlier, in our implementation, the set of candidate proxies for the factors G', will be the same as X. Although this need not be the case, it is done mainly because X represents the biggest set of (standardized and stationary) observable time series variables available to us. We perform real-time forecasts of 7 of the 8 major monthly macroeconomic time series studied in Stock and Watson (2002a). The four real variables we concentrate on are total industrial production (IP), real personal income less transfers, real manufacturing and trade sales and the number of employees on nonagricultural payrolls. The three price series considered are the consumer price index (CPI), the

personal consumption expenditure implicit price deflator (PCED) and the producer price index for finished goods (PPI). All of these variables are expressed in log-differences (i.e., monthly growth rates).[3]

6 Monte Carlo Experiment

Table 2 contains the results from a small Monte Carlo experiment used to assess the finite sample forecast performance of the $A(j)$ and $M(j)$ tests. In the empirical panel dataset spanning 1960:1 to 2003:12 discussed in Section 5 above, $\widehat{k} = 13$ factors are consistently estimated using the methodology of Bai and Ng (2002). For this reason, we assume there are 13 factors underlying our simulated dataset and set $r = 13$. The simulated dataset also has the same dimensions as the empirical dataset discussed in the next section. Hence, we set $N = 132$ and $T = 528$. Each of the thirteen latent factors is generated as

$$F_{kt} = \nu_k F_{kt-1} + u_{kt}, \tag{23}$$

where $0.6 \leq \nu_k \leq 0.8$, $u_{kt} \sim N(0,1)$, and u_{kt} is uncorrelated with u_{jt}, for $k \neq j$, $k, j = 1, \ldots, r$. $F_t = (F_{1t}, \ldots, F_{rt})'$, $\Lambda_{ik} \sim N(0,1)$, and e_{it} is uncorrelated with e_{jt}, for $i \neq j$, $i, j = 1, \ldots, N$, $t = 1, \ldots, T$. Following Bai and Ng (2002), the simulated panel dataset is generated as

$$x_{it} = \Lambda_i' F_t + \sqrt{\eta} e_{it}, \tag{24}$$

for $i = 1, \ldots, 119$, where η is a measure of the variance of the idiosyncratic errors, e_{it}, relative to the common component, $\Lambda_i' F_t$. More specifically, for $i = 1, \ldots, 39$, we make the idiosyncratic errors homoskedastic and set $e_{it} \sim N(0,1)$. We introduce heteroskedasticity into the variables for which $i = 40, \ldots, 79$ and let

$$e_{it} = \begin{cases} e_{it}^1 & \text{if } t \text{ is even} \\ e_{it}^1 + e_{it}^2 & \text{if } t \text{ is odd} \end{cases} \tag{25}$$

[3]Note that Stock and Watson (1999, 2002a) model some of our price variables as $I(2)$ in logarithms. However, they find little discrepancy in performance under $I(1)$ and $I(2)$ assumptions for factor forecasts of our three target price variables. For this reason, we limit our analysis by assuming that our price variables as well as other variables in X are $I(1)$ in logarithms (see Section 10 for further details). In all other respects, our dataset is the same as that used by Stock and Watson (2005).

where e_{it}^1 and e_{it}^2 are independent $N(0,1)$ (see Bai and Ng (2002)). For $i = 80, \ldots, 119$, the idiosyncratic component of (24) is generated as an $MA(1)$ process such that

$$e_{it} = 0.6e_{it-1} + e_{it}^3 \tag{26}$$

and $e_{it}^3 \sim N(0,1)$. Define a variable to be a good proxy if it is a linear combination of the underlying latent factors (see Bai and Ng (2006b) for complete details). Thus, for $i = 120, \ldots, 132$, the proxy variables are generated as

$$x_{it} = \Lambda_i' F_t \tag{27}$$

Since the generated factors in (23) are assumed to be latent, they are not wholly included in the simulated panel dataset. The above setup ensures that four separate DGPs generate a total of 132 simulated variables. Ex ante forecasts are constructed for four variables. In Table 2, the target variables labelled "Homoskedastic", "Heteroskedastic" and "$MA(1)$" are all generated from (24). However, the corresponding idiosyncratic errors are specified by i.i.d. $N(0,1)$, (25) and (26) respectively. Let $x_t^p = (x_{120t}, \ldots, x_{132t})$, $e_{it}^4 \sim N(0,1)$ and $\Omega_{il} \sim N(0,1)$ for $l = 1, \ldots, 13$, then the target variable labelled "Proxy (Homoskd.)" is generated by

$$y_p = \Omega_i' x_t^p + e_{it}^4 \tag{28}$$

The difference between (28) and (24) is that in (28), x_t^p are observed and can be selected by the $A(j)$ or $M(j)$ tests as regressors in a forecast model for the respective target variable. Of course, there is still no guarantee that they will be selected; rather this is the only case where the true regressor variables are actually in the panel dataset and can be selected. This is an important case, and defines the case we are most interested in. On the contrary in (24), F_t are not observed and can consequently not be selected as predictors in a forecasting exercise. For each of the four target variables in Table 2 and Table 3, the last third of simulated values are recursively forecasted. Since there are 528 data points across time in our setup, we effectively use observations from $t = 352$ to $t = 528$ to evaluate forecast performance via examination of the mean squared forecast error (MSFE). Prior observations are used to estimate the forecast models. In strict recursive fashion, all models, factors, number of factors, k and proxies are re-estimated and re-selected for each constructed forecast. Forecasting at time $t+1$, the panel dataset from $1, \ldots, t$ is also standardized to have mean zero and unit variance before the factors are recursively estimated and proxies selected.

In order to make the experiment credible, the model used for the factor forecasts is Model *1* and those used for the "$A(j)$" and "$M(j)$"proxy forecasts are Model *5* and Model *7,* respectively (see Table 1 for the specification of these models). From prior work, Model *5* and Model *7* performed worst among all the alternative proxy forecast models specified in Table 1, and hence our setup is as "tough as possible" on our approach. It is left to future research to establish whether other model specifications discussed in this paper that perform better in our empirical experiments also perform better in Monte Carlo simulation experiments.

We perform the same forecast evaluation exercise for a subsets of $N = 40$ and $N =$132. The 40 variable subsets are randomly selected from the original 132 simulated variables under the constraint that at least 2 and at most 7 proxies as defined by (27) are selected. Forecast horizons of $h = 1, 12$ are considered. The entire monte carlo experiment is conducted for 250 iterations and at each iteration, for $N = 40$ and $N = 132$, we calculate the MSFE from 176 recursive forecasts for $t = 352, \ldots, 528$.

The numerical entries in Table 2 represent the fraction of times (out of 250 Monte Carlo iterations) that the proxy forecasts have a lower MSFE than the factor forecasts. Regardless of the number of variables in the panel dataset or the forecast horizon, the proxy forecasts outperformed the factor forecasts about 50% of the time in almost all cases. This is significant as it demonstrates that the worst performing proxy forecast models equally match the factor forecasts. Under $h = 1$, entries of 0.720, 0.795, 0.880 and 0.895 for "Proxy (Homoskd.)" indicate that the proxy forecasts strongly outperform the factor forecasts. This particular outcome is as might be expected, given that this is the case where the $A(j)$ and $M(j)$ test statistics are afforded the possibility of selecting the truly correct elements of x_t^p used to generate y_p in (28) and suggests that our approach is working as desired. However, under $h = 12$ for the same target variable "Proxy (Homoskd.)", the proxy forecasts perform just as well as the factor forecasts. One explanation for this result might be that as the forecast horizon gets longer, the informational content in the proxies deteriorates faster relative to that of the factors.

The entries in Table 3 not in parenthesis represent the mean of the various MSFEs across Monte Carlo iterations. The standard deviations of the MSFEs are reported in parentheses. From Table 3, proxy forecasts constructed from the $A(j)$ or $M(j)$ statistic marginally outperform the factor forecasts most of the time in terms of the mean of the MSFEs. However, the equal performance of the factor and proxy forecasts in Table 2 is demonstrated in Table 3 by the fact that the mean of

the proxy MSFEs is generally only slightly less than the mean of the factor MSFEs.

Overall, these results are interesting, and suggest that our prediction/proxy approach outperforms a standard factor approach in favorable cases, and perform equally as well in non-favorable cases.

7 Empirical Findings

In this section, we discuss the results of a series of prediction experiments using the dataset discussed above, and applying the models outlined in Table 1 to construct sequences of recursive ex-ante h-step ahead predictions. The dataset consists of 132 variables (see Section 5), and data are available for the period 1960:1-2003:12. Furthermore, predictions are constructed for the period 1989:5-2003:12. Please see Section 4 for complete details concerning the strategy used to specify and estimate the prediction models prior to forecast construction. For models in which proxies were selected using the $M(j)$ and $A(j)$ tests, we set $2\xi = 0.05$. Hence we carry out the tests at a 5% significance level. We include 1 autoregressive lag in most of the models because the importance of autoregressive lags in prediction is well established. Furthermore, adding autoregressive terms of the target variable to the basic factor model is a good way to give the factor model a fair chance to "win" our forecasting competition.

Results of our empirical experiments are gathered in Table 4 (frequency of selected factor proxies), Table 5 (CPI, PCED, and PPI forecasting competition results), and Table 6 (Industrial Production, Personal Income; Nonagricultural Employment, Manufacturing and Trade Sales). In Table 4, selection frequencies are reported, while in Tables 5-6 MSFEs and DM test statistics are reported. The MSFE values reported for CPI, PCED and Nonagricultural Employment are multiplied by 100,000 and those reported for Producer Price Index, Industrial Production, Manufacturing and Trade Sales and Personal Income are multiplied by 10,000. For the benchmark Model *1* (i.e., the factor model), the only tabular entry for all forecast horizons is the MSFE. With all of the other models (i.e., our alternative models), there are two entries: The top entry is the MSFE and the bottom entry in parenthesis is the DM t-statistic. As mentioned earlier, a positive DM statistic value indicates that the alternative model has a MSFE that is lower than the benchmark, while a negative statistic value indicates the reverse. Entries in bold signify instances where the alternative model outperforms the factor model as determined by a point MSFE comparison. Boxed MSFE

entries represent the lowest MSFE value among all the models for a particular forecast horizon. DM statistic entries with a $*$ indicate instances where the respective alternative model significantly outperforms the factor model at a 10% significance level, whereas for entries with a † sign, the factor model significantly outperforms the alternative model at a 10% significance level. We now provide a number of conclusions based on the tables.

Upon inspection of Table 5, it is clear that the benchmark factor model (i.e., Model *1)* significantly outperforms most of the alternative models in the forecast of CPI and PCED. This point is supported by the overwhelming number of DM test rejections in Panels A and B of Table 5. While the benchmark still yields the lower MSFE in many pairwise comparisons when examining PPI results (see Panel C of the table), the DM test null of equal predictive accuracy is not frequently rejected.

A key exception to the above conclusion that the benchmark model yields superior predictions is in the case of Models *12-15*. From Table 1, recall that these are autoregressive models with exogenous variables (ARX). The lags of the ARX models are selected by the SIC and the exogenous variables are based on smoothed versions of the $A(j)$ and $M(j)$ tests. For $h = 1, 3, 12$, these models not only frequently yield lower point MSFEs than the benchmark, but the difference in performance is often significant. Across all 3 panels and 3 forecast horizons (i.e., 9 variable/horizon combinations), it is interesting to note that one or many of Models 12-15 are "MSFE-best" 7 times. Furthermore, of these 7 "wins" it is Model 12 that yields the lowest MSFE in 4 instances. Thus, we have direct evidence that the parsimonious single proxy smoothed $A(j)$ model fares very well when compared not only to the benchmark, but also to other models which yield lower MSFEs than the benchmark. This suggests that while the factor approach is very useful, often beating the pure autoregressive and other linear models when used for predicting price variables, a parsimonious version of the smoothed $A(j)$ factor proxy approach performs the best, overall. Thus, as pointed out by Bai and Ng (2006c), parsimony is still important. This is even true in the context of ordinary proxy models (Models *4-7*), as choosing one proxy rather than $\widehat{k}$ proxies often yields the lowest MSFE model.

Interestingly, in Table *5*, the above conclusions hold for $h = 1, 3, 12$ and not for $h = 24$. Indeed, it appears that all models perform quite poorly for $h = 24$, with the notable exception of the benchmark, which clearly outperforms virtually all competitors in all price variable cases when $h = 24$. Thus, at the longest forecast horizons, we have evidence that our simple factor proxy

approaches are not faring well at all.

Turning now to Table 6, the above conclusions still hold, with the exception that many other alternative models, and not just Models 12-15, are point MSFE "better" than the benchmark. Summarizing the results in Table 6, the benchmark model does yield the lowest MSFE for 3 of the 4 variables when $h = 1$ and for 1 variable when $h = 3$, although the DM test null is not rejected in any of these cases. Furthermore, for all remaining horizon/variable combinations, the benchmark does not yield the lowest MSFE. Indeed, in all but one of these other cases, factor proxy approaches yield the lowest MSFE (the sole exception is a random walk "win" for Manufacturing and Trade Sales when $h = 3$).

Given the above results, it is of interest to tabulate which factor proxies were used in our prediction experiments. This is done in Table 4, where factor proxies that are (most frequently) selected using the $A(j)$ and $M(j)$ test and the frequencies with which they are selected are reported. The second column under "Trans" indicates the data transformation that was performed to induce data stationarity. As is evident, S&P's Common Stock Price Index, Industrials; S&P's Common Stock Price Index, Composite; Dividend Yields, a 1-Year Bond Rate; and Housing Starts are the five most common proxies selected by both $A(j)$ and $M(j)$. Structural change could account for some of the proxies being selected less frequently than the five above proxies. Clearly, the importance of proxies may in some cases depend on the period in history represented by the data. However, it is interesting that a variety of factor proxies are "picked" across our entire ex-ante prediction period.

The diagrams in Panel 1- 3 of Figure 1 are time series plots of the first three estimated factors (i.e., the most important factors for explaining the variability in our panel dataset). Panels 4-6 are time series plots of the three most frequently selected proxies based on use of the $A(j)$ and $M(j)$ test statistics. The S&P Common Stock Price Index does proxy the estimated factors to some extent, although the relatively high level of noise in the S&P variable does appear to obscure this fact to a certain degree. The Housing Starts, Nonfarm variable (which has less noise - see Panel 6) better illustrates the close relationship between the estimated factors and selected proxies. Results in Table 4 indicate that almost all three proxies in Figure 1 are selected 100% of the time by both the $A(j)$ and $M(j)$ statistics although the $M(j)$ test has more power than the $A(j)$ test. The lone exception to this is the Housing Starts, Nonfarm variable which is selected 95% of the time by the $M(j)$ test. This suggests how strongly the three variables proxy the underlying factors. In addition, one gets a "sense" of the robustness of the $A(j)$ and $M(j)$ test statistics in consistently

selecting good proxies, since the uderlying factors are re-estimated at each recursive iteration.

In closing, we note that factor proxies appear useful for prediction. Additionally, since factors are unobserved, analyzing and studying them on their own can be quite difficult. For instance, in our context is not clear how relevant it is to study the evolution of the individual factors over time because prior to each new prediction, the factors are re-estimated. Creating a clearly defined historical path for a factor is consequently complicated. The ability to proxy the unobserved factors with observed variables enables us to identify actual variables that can serve as primitive building blocks for (prediction) models of a host of macroeconomic variables.

8 Recent Advances in the Construction of Diffusion Indices

In this section, we briefly highlight some of the most recent work relating to diffusion index (factor) models. Many of these ideas could potentially be applied to the issues discussed in this paper, although we leave that to future research. Some of the concerns raised in this paper such as the use of the same factors and consequently the same proxies to forecast *any* variable are addressed in a number of the papers. For example, Bai and Ng (2006c) offer two refinements to the method of factor forecasting. The current framework is confined to a linear relation between the predictors and the forecasted series. Bai and Ng (2006c) propose a more flexible structure. Their so-called squared principal components approach allows the relationship between the predictors and the factors to be non-linear. They use a non-linear "link" function that involves expanding the set of predictors to include non-linear functions of the observed variables. In this regard, (4) can be modified as follows:

$$h(x_{it}) = \vartheta_i' J_t + e_{it},$$

where $h(\cdot)$ is a non-linear link function, J_t are the common factors, and ϑ_i is the vector of factor loadings. The second order factor model is consequently:

$$x_t^* = \Omega J_t + e_t \tag{29}$$

where $x_t^* = \{x_{it}, x_{it}^2\}\forall i$ is an $N^* \times 1$ vector and $N^* = 2N$. Estimation of J_t from (29) is done using the usual method of principal components. The forecasting equation in (9) remains linear regardless of the form of $h(\cdot)$. The second refinement proposed by Bai and Ng (2006c) takes explicit account of the fact that the ultimate aim is to forecast a specific time series variable, say y_t. The

authors propose using principal components analysis with a "targeted" subset of the predictors in X, which have been tested to have predictive power for y. This implies that the set of predictors used to extract the factors change with y, the targeted forecast variable. "Hard" and so-called "soft" thresholding is used to determine which subset of X the factors are to be extracted from. Under "hard" thresholding, a test with a sharp decision rule determines which variables are "in" or "out". With "soft" thresholding, the top variables are kept in the subset of predictors used to extract the factors. The ordering of the predictors is based on the particular soft-thresholding rule. The "soft" thresholding approach is thus related to our "smoothed" test statistic approach to factor proxy selection.

As a reminder, the use of factor models (diffusion indices) involves a two-step approach in which the factors are first estimated from a large panel dataset. The estimated factors are then used as predictors in the forecast models. Although the estimated factors in the first stage are capable of parsimoniously capturing almost all the information in a large dataset, standard tools for specifying the forecast model in the second stage remain unsatisfactory in certain contexts. The specified prediction models are still susceptible to overfitting or underfitting, for example. In this light, Bai and Ng (2006d) suggest a stopping rule for "boosting" that prevents a model from being overfitted with estimated factors or other predictors. Boosting is a procedure that estimates the conditional mean using M stagewise regressions (Bai and Ng (2006d)). The authors also propose two ways to handle lagged predictors: a component-wise approach that treats each lag as a separate variable, and a block-wise approach that treats lags of the same variable jointly. Some important papers on boosting include Schapire (1990), Freund (1995), Friedman (2001) and Buhlmann and Hothorn (2006).

9 Concluding Remarks

Using Monte Carlo and empirical analysis, we have shown that the $A(j)$ and $M(j)$ statistics of Bai and Ng (2006b) appear to offer an interesting means by which factor proxies for later use in prediction models can be chosen. Indeed, our "smoothed" approaches to factor proxy selection appear to yield predictions that are often superior not only to a benchmark factor model, but also to simple linear time series models which have in many practical applications hitherto been found to be difficult to beat in forecasting competitions. More specifically, we find that our factor proxy

models (e.g., see Model 5 and Model 7 in Table 1) perform (slightly) better than a standard factor model (Model 1) in our Monte Carlo experiments. The implication is that a policymaker will be better served by using the proxy model. At the very least, the methodology suggested in this paper should perhaps be added to the practitioners "tool-box", and one should examine on a case-by-case basis whether or not proxy observable factors are more effective that standard factors. This is particularly relevant since, unlike the factor model which has estimated regressors, the proxy model uses observed regressors that can act as policy instruments, for example. By using our approach to predictive factor proxy selection, one is able to open up the "black box" often associated with factor analysis, to some extent. This is because one can identify actual variables that can serve as "primitive" building blocks for (prediction) models of a host of other macroeconomic variables. This approach in some cases leads to improved prediction, and may also possibly lead to improved policy analysis if used in policy related prediction modelling.

10 References

Bai, J. (2003). Inferential Theory for Factor Models of Large Dimensions. *Econometrica* 71(1), 135-172.

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

Bai, J. and Ng, S. (2006a). Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-Augmented Regressions. *Econometrica* 74(4), 1133-1150.

Bai, J. and Ng, S. (2006b). Evaluating Latent and Observed Factors in Macroeconomics and Finance. *Journal of Econometrics* 113, 507-537.

Bai, J. and Ng, S. (2006c). Forecasting Economic Time Series Using Targeted Predictors. *Working Paper*, University of Michigan, Ann-Arbor.

Bai, J. and Ng, S. (2006d). Boosting Diffusion Indices. *Working Paper*, University of Michigan, Ann-Arbor.

Bai, J. and Ng, S. (2007). Determining the Number of Primitive Shocks in Factor Models. *Journal of Business and Economic Statistics* 25, 52-60.

Battini, J. and Haldane, A. (1999). Forward-looking Rules for Monetary Policy. In Taylor, J., (ed.), *Monetary Policy Rules*. University of Chicago Press for NBER, Chicago.

Bernanke, B. and Boivin, J. (2003). Monetary Policy in a Data-Rich Environment. *Journal of Monetary Economics* 50, 525-546.

Boivin, J. and Ng, S. (2005). Understanding and Comparing Factor Based Macroeconomic Forecasts. *International Journal of Central Banking* 1(3), 117-152.

Breeden, D., Gibbons, M. and Litzenberger, R. (1989). Empirical Tests of the Consumption-Oriented CAPM. *Journal of Finance* XLIV (2), 231-262.

Buhlmann, P. and Hothorn, T. (2006). Boosting Algorithms: Regularization, Prediction and Model Fitting. *Working Paper*, ETH Zürich.

Clarida, R, Gali, G. and Gertler, M. (1999). The Science of Monetary Policy: A new Keynesian Perspective. *Journal of Economic Literature* 37, 1661-1707.

Clarida, R, Gali, G. and Gertler, M. (2000). Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory. *Quarterly Journal of Economics* 115, 147-180.

Connor, G., and Korajczyk, R. (1986). Performance Measurement With the Arbitrage Pricing Theory. *Journal of Financial Economics* 15, 373-394.

Connor, G., and Korajczyk, R. (1988). Risk and Return in an Equilibrium APT: Application of a New Test Methodology. *Journal of Financial Economics* 21, 255-289.

Connor, G., and Korajczyk, R. (1993). A Test for the Number of Factors in an Approximate Factor Model. *Journal of Finance* 48(4) 1263-1291.

Corradi, V., and Swanson, N.R. (2002). A Consistent Test for Out of Sample Nonlinear Predictive Ability. *Journal of Econometrics* 110, 353-381.

Corradi, V., and Swanson, N.R. (2006a). Bootstrap Conditional Distribution Tests in the Presence of Dynamic Misspecification. *Journal of Econometrics* 133, 779-806.

Corradi, V., and Swanson, N.R. (2006b). Predictive Density Evaluation. In Granger, C., Elliot, G. and Timmerman, A. (eds.), *Handbook of Economic Forecasting*. Elsevier, Amsterdam, 197-284.

Diebold, F. and Mariano, R. (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*. 13, 253-263.

Ding, A. and Hwang, J. (1999). Prediction Intervals, Factor Analysis Models, and High-Dimensional Empirical Linear Prediction. *Journal of the American Statistical Association* 94, 446-455.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2000). The Generalized Dynamic Factor Model: Identification and Estimation. *The Review of Economics and Statistics* 82(4), 540-552.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2005). The Generalized Dynamic Factor Model: One-Sided Estimation and Forecasting. *Journal of the American Statistical Association* 100(471), 830-840.

Forni, M.and Reichlin, L. (1996). Dynamic Common Factors in Large Cross-Sections. *Empirical Economics* 21, 27-42.

Forni, M.and Reichlin, L. (1998). Lets Get Real: A Dynamic Factor Analytical Approach to Disaggregated Business Cycle. *Review of Economic Studies* 65, 453-474.

Freund, Y. (1995). Boosting a Weak Learning Algorithm by Majority. *Information and Computation* 121, 256-285.

Friedman, J. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics* 29, 1189-1232.

Geweke, J. (1977). The Dynamic Factor Analysis of Economic Time Series. In Aigner, D. and Goldberger, A. (eds.), *Latent Variables in Socio-Economic Models*. Amsterdam.

Hayashi, F. (2000). *Econometrics*. Princeton University Press, Princeton.

Newey, W. and West, K. (1987). A Simple Positive-Definite Heteroskedasticity and Autocorelation Consistent Covariance Matrix. *Econometrica* 55, 703-708.

Rapach, D. and Strauss, J. (2007). Bagging or Combining (or Both)? An Analysis Based on Forecasting U.S. Unemployment Growth. *Econometric Reviews*, (forthcoming).

Sargent, T., and Sims, C. (1977). Business Cycle Modeling without Pretending to Have Too Much A-Priori Economic Theory. In Sims, C. et al. (eds.), *New Methods in Business Cycle Research*. Minneapolis: Federal Reserve Bank of Minneapolis.

Schapire, R. (1990). The Strength of Weak Learnability. *Machine Learning* 5, 197-227.

Shanken, J. (1992). On the Estimation of Beta-Pricing Models. *Review of Financial Studies* 5 (1), 1-33.

Stock, J. and Watson, M. (1996). Evidence on Structural Instability in Macroeconomic Time Series Relations. *Journal of Business and Economic Statistics* 14, 11-30.

Stock, J. and Watson, M. (1998). Diffusion Indexes. *Working Paper* 6702, National Bureau of Economic Research.

Stock, J. and Watson, M. (1999). Forecasting Inflation. *Journal of Monetary Economics* 44, 293-335.

Stock, J. and Watson, M. (2002a). Macroeconomic Forecasting Using Diffusion Indexes. *Journal of Business and Economic Statistics* 20, 147-161.

Stock, J. and Watson, M. (2002b). Forecasting Using Principal Components from a Large Number of Predictors. *Journal of American Statistical Association* 97, 1167-1179.

Stock, J. and Watson, M. (2004a). An Empirical Comparison of Methods for Forecasting Using Many Predictors. *Working Paper*, Princeton University.

Stock, J. and Watson, M. (2004b). Forecasting with Many Predictors. *Handbook of Forecasting*, forthcoming.

Stock, J. and Watson, M. (2005). Implications of Dynamic Factor Models for VAR Analysis. *Working Paper* 11467, National Bureau of Economic Research.

Taylor, J. (1993). Discretion Versus Policy Rules in Practice. *Carnegie Rochester Conference Series on Public Policy* 39, 195-214.

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica* 48, 817-838.

Table 1: Prediction Models Used in Empirical Experiments*

Model 1 (Factor Model): This is the standard factor forecast model: $\widehat{y}_{T+h|T} = \widehat{a}_0 + \widehat{\alpha}'\widetilde{F}_T + \widehat{\beta}y_T$

Model 2 (Autoregressive Model): This is an $AR(p)$ forecast model, with lags selected by the SIC: $\widehat{y}_{T+h|T} = \widehat{a}_0 + \sum_{j=1}^{p} \widehat{\alpha}_j y_{T-j+1}$

Model 3 (Random Walk Model): This is a random walk forecast model: $\widehat{y}_{T+h|T} = y_T$

Model 4 (Ordinary $A(j)$ - 1 Proxy Model): In this forecast model, the single "best" proxy selected by the $A(j)$ test (i.e., the proxy associated with the $A(j)$ statistic value closest to 2ξ in absolute value) is used as the only proxy regressor in the forecast model: $\widehat{y}_{T+h|T} = \widehat{a}_0 + \widehat{\alpha}G^A_{j_1T} + \widehat{\beta}y_T$

Model 5 (Ordinary $A(j)$ - $\widehat{k}$ Proxies Model): The "best" $\widehat{k}$ factor proxies selected by the $A(j)$ test are used: $\widehat{y}_{T+h|T} = \widehat{a}_0 + \widehat{\alpha}'S^A_T + \widehat{\beta}y_T$, where $S^A_T = \{G^A_{j_1T}, \ldots, G^A_{j_{\widehat{k}}T}\}$.

Model 6 (Ordinary $M(j)$ - 1 Proxy Model): In this forecast model, the single "best" factor proxy selected by the $M(j)$ test (i.e., the proxy associated with the lowest $M(j)$-statistic) is used as the only proxy regressor in the forecast model: $\widehat{y}_{T+h|T} = \widehat{a}_0 + \widehat{\alpha}G^M_{jT} + \widehat{\beta}y_T$. Since it is possible for the $M(j)$ test to select no proxies at all, should that scenario occur, the model degenerates to: $\widehat{y}_{T+h|T} = \widehat{a}_0 + \widehat{\beta}y_T$.

Model 7 (Ordinary $M(j)$ - $\widehat{k}$ Proxies Model): This forecast model is the same as Model 6, but $\widehat{k}$ factor proxies selected by the $M(j)$ test are used: $\widehat{y}_{T+h|T} = \widehat{a}_0 + \widehat{\alpha}'S^M_T + \widehat{\beta}y_T$.

Model 8 (Smoothed $A(j)$ - 1 Proxy Model): This forecast model is the same as Model 4, except that the smoothed version of the $A(j)$ test is used (see Section 3.3 for further discussion).

Model 9 (Smoothed $A(j)$ - $\widehat{k}$ Proxies Model): This forecast model is the same as Model 5, except that the smoothed version of the $A(j)$ test is used (see Section 3.3 for further discussion).

Model 10 (Smoothed $M(j)$ - 1 Proxy Model): This forecast model is the same as Model 6, except that the smoothed version of the $M(j\ )$ test is used (see Section 3.3 for further discussion).

Model 11 (Smoothed $M(j)$ - $\widehat{k}$ Proxies Model): This forecast model is the same as Model 7, except that the smoothed version of the $M(j\ )$ test is used (see Section 3.3 for further discussion).

Model 12 (Autoregressive plus Smoothed $A(j)$ - 1 Proxy Model): This forecast model is the same as Model 8, except that the lag of the autoregressive component is selected by the SIC rather than restricted to 1: $\widehat{y}_{T+h|T} = \widehat{a}_0 + \widehat{\alpha}G^{A*}_{j_1T} + \sum_{j=1}^{p_x} \widehat{\beta}_j y_{T-j+1}$.

Model 13 (Autoregressive plus Smoothed $A(j)$ - $\widehat{k}$ Proxies Model): This forecast model is the same as Model 9, except that the lag of the autoregressive component is selected by the SIC rather than restricted to 1.

Model 14 (Autoregressive plus Smoothed $M(j)$ - 1 Proxy Model) : This forecast model is the same as Model 10, except that the lag of the autoregressive component is selected by the SIC rather than restricted to 1.

Model 15 (Autoregressive plus Smoothed $M(j)$ - $\widehat{k}$ Proxies Model): This forecast model is the same as Model 8, except that the lag of the autoregressive component is selected by the SIC rather than restricted to 1.

* Note: See Sections 3.3 and 4 for further discussion of the factor proxy selection methodology used in the construction of the above models.

Table 2: Monte Carlo Experiment Results*

		$h = 1$		$h = 12$	
N	Error Structure	A(j)	M(j)	A(j)	M(j)
40	Homoskedastic	0.425	0.330	0.460	0.560
40	Heteroskedastic	0.425	0.435	0.530	0.685
40	MA(1)	0.520	0.620	0.575	0.675
40	Proxy (Homoskd.)	0.720	0.795	0.390	0.450
132	Homoskedastic	0.585	0.430	0.545	0.595
132	Heteroskedastic	0.680	0.475	0.545	0.615
132	MA(1)	0.460	0.600	0.585	0.605
132	Proxy (Homoskd.)	0.880	0.895	0.585	0.620

* Notes: The numeric entries under "N" indicate the number of variables in the simulated panel dataset. Entries under "A(j)" and "M(j)" indicate the fraction of times that the alternative model (*Model 5* or *Model 7*, repsectively) has a lower MSFE than the benchmark (*Model 1*), in 250 Monte Carlo iterations. Under "Error Structure", we state the forecast "target" variable. "Homoskedastic", "Heteroskedastic" and "$MA(1)$" represent target variables for which the idiosyncratic error, e_{it}, in the DGP is an i.i.d. $N(0,1)$, a heteroskedastic, or a moving average process, respectively. For all three of these cases, the independent variables in the DGP are the latent factors. For "Proxy (Homoskd.)", the idiosyncratic error, e_{it}, is i.i.d. $N(0,1)$; and the independent variables in the DGP are potential "proxy" variables, so that the "A(j)" and "M(j)" in this case might select the "true" proxy, if they perform as desired. See Section 6 for complete details.

Table 3: Monte Carlo Experiment Descriptive Statistics*

			$h = 1$			$h = 12$	
N	Error Structure	Factor	A(j)	M(j)	Factor	A(j)	M(j)
40	Homoskedastic	52.809	53.023	53.472	60.043	60.118	60.008
		(8.882)	(8.933)	(9.219)	(12.770)	(12.826)	(12.822)
40	Heteroskedastic	49.111	49.481	49.725	56.925	56.928	56.685
		(8.950)	(8.761)	(8.783)	(14.240)	(14.226)	(14.174)
40	MA(1)	38.829	38.812	38.736	66.851	66.649	66.468
		(6.669)	(6.665)	(6.619)	(16.385)	(16.333)	(16.178)
40	Proxy (Homoskd.)	25.751	25.569	25.466	56.548	56.953	56.681
		(26.382)	(26.695)	(26.704)	(59.791)	(59.849)	(59.543)
132	Homoskedastic	49.371	48.955	50.024	62.079	61.987	61.814
		(8.567)	(8.627)	(8.680)	(13.959)	(13.609)	(13.578)
132	Heteroskedastic	44.948	44.305	45.265	57.560	57.444	57.298
		(7.369)	(7.501)	(7.834)	(12.337)	(12.245)	(12.384)
132	MA(1)	39.227	39.234	39.054	69.809	69.586	69.236
		(6.225)	(6.266)	(6.254)	(15.916)	(16.042)	(16.556)
132	Proxy (Homoskd.)	28.249	27.579	27.422	60.651	60.446	60.065
		(31.545)	(31.141)	(30.988)	(71.739)	(72.218)	(71.744)

* Notes: See notes to Table 2 above. The numerical entries not in parentheses under "Factor", "A(j)" or "M(j)" are the means of the various MSFEs calculated under the respective models, across 250 Monte Carlo iterations. The corresponding entries in parentheses are MSFE standard deviations, again calculated across all Monte Carlo iterations.

Table 4: Frequency of Selected Factor Proxies*

Selected Factor Proxy	Trans	A(j)	M(j)
fspin: S&P's Common Stock Price Index, Industrials	$\Delta \log$	1.000	1.000
fspcom: S&P's Common Stock Price Index, Composite	$\Delta \log$	1.000	1.000
fsdxp: S&P's Composite Common Stock: Dividend Yield	Δlv	1.000	
fygt1: Interest Rate: U.S. Treasury Const Maturities, 1-Yr	Δlv	1.000	
hsfr: Housing Starts, Nonfarm	log	1.000	0.949
hsbr: Housing Authorized, Total New Private Housing Units	log	0.989	0.455
ips10: Industrial Production Index, Total Index	$\Delta \log$	0.909	
exrus: United States, Effective Exchange Rate	$\Delta \log$	0.835	0.370
sfygm6: 6 month Treasury Bills - Federal Funds, spread	lv	0.813	
sfygt5: 5 yr Treasury Bond Const. Maturities - Federal Funds, spread	lv	0.750	
sfygt10: 10 yr Treasury Bond Const. Maturities - Federal Funds, spread	lv	0.659	0.420
fygm6: Interest Rate, U.S. Treasury Bills, Sec Mkt, 6-Mo.	Δlv	0.460	
a0m077: Ratio, Mfg. and Trade Inventories to Sales	Δlv	0.341	0.261

* Notes: In this table we report proxies that were frequently selected using the $A(j)$ and $M(j)$ tests, and the frequencies with which they were selected, in our recursive forecasting experiments. The second column under "Trans" indicates the data transformation that was performed to induce stationarity, lv means no transformation; the series was left at level. Δlv means first difference of the level. log means the natural log function was applied to the data. $\Delta \log$ means the series was first differenced after the natural log function was applied. Empty entries in the fourth column under $M(j)$ indicate that the respective variables were not selected at all by the $M(j)$ test.

Table 5: Predictive Performance of Various Models for Price Variables*

Forecast Horizon (h)	1	3	12	24
Panel A: CPI				
Model 1	3.496	3.464	4.299	4.089
Model 2	**3.457**	**3.330**	4.357	5.069
	(0.136)	**(0.375)**	(-0.155)	(-2.270)†
Model 3	4.785	5.270	6.347	6.129
	(-3.788)†	(-3.795)†	(-3.768)†	(-3.087)†
Model 4	3.809	4.075	4.792	5.305
	(-1.164)	(-1.873)†	(-1.336)	(-2.737)†
Model 5	4.079	4.592	5.255	5.337
	(-1.125)	(-1.775)†	(-1.650)†	(-1.878)†
Model 6	3.802	4.107	4.757	4.891
	(-1.139)	(-2.011)†	(-1.347)	(-1.770)†
Model 7	4.516	4.747	5.095	5.103
	(-1.479)	(-2.223)†	(-1.480)	(-1.600)
Model 8	3.810	4.111	4.759	4.960
	(-1.169)	(-2.048)†	(-1.382)	(-2.014)†
Model 9	3.677	3.921	4.472	4.665
	(-0.775)	(-1.798)†	(-0.618)	(-1.645)†
Model 10	3.819	4.101	4.769	5.208
	(-1.212)	(-2.040)†	(-1.304)	(-2.576)†
Model 11	3.720	4.050	4.563	4.740
	(-0.935)	(-2.022)†	(-0.881)	(-1.659)†
Model 12	**3.340**	**3.158**	**4.020**	4.448
	(0.549)	**(0.995)**	**(0.921)**	(-0.981)
Model 13	3.519	**3.296**	**4.097**	4.259
	(-0.086)	**(0.539)**	**(0.606)**	(-0.537)
Model 14	**3.486**	**3.381**	4.351	5.124
	(0.035)	**(0.232)**	(-0.145)	(-2.379)†
Model 15	**3.351**	**3.331**	**3.999**	4.297
	(0.527)	**(0.411)**	**(0.938)**	(-0.634)
Panel B: Consumption Deflator (PCE)				
Model 1	2.689	2.882	3.162	2.902
Model 2	**2.613**	**2.540**	**3.097**	3.918
	(0.245)	**(1.598)**	**(0.275)**	(-2.985)†
Model 3	4.318	3.956	4.521	4.823
	(-2.312)†	(-3.275)†	(-3.082)†	(-3.373)†
Model 4	3.561	3.214	3.608	4.114
	(-1.911)†	(-1.525)	(-1.983)†	(-3.754)†
Model 5	2.900	3.488	3.557	3.663
	(-1.106)	(-2.348)†	(-1.990)†	(-2.308)†
Model 6	3.542	3.220	3.587	3.835
	(-1.871)†	(-1.593)	(-2.118)†	(-2.933)†
Model 7	3.123	3.386	3.501	3.648
	(-1.865)†	(-2.486)†	(-1.834)†	(-2.349)†
Model 8	3.562	3.283	3.921	4.412
	(-1.910)†	(-1.847)†	(-3.021)†	(-4.066)†
Model 9	3.375	3.233	3.491	3.826
	(-1.687)†	(-1.948)†	(-1.729)†	(-2.957)†
Model 10	3.593	3.227	3.673	4.207
	(-1.887)†	(-1.614)	(-1.969)†	(-3.925)†
Model 11	3.548	3.196	3.496	3.781
	(-1.717)†	(-1.504)	(-1.769)†	(-2.905)†
Model 12	**2.619**	**2.485**	**3.118**	3.846
	(0.237)	**(2.005)***	**(0.191)**	(-2.904)†
Model 13	**2.669**	**2.554**	**2.874**	3.294
	(0.066)	**(1.669)***	**(1.360)**	(-1.562)
Model 14	**2.637**	**2.558**	**3.123**	3.978
	(0.163)	**(1.544)**	**(0.160)**	(-3.229)†
Model 15	**2.633**	**2.525**	**2.817**	3.271
	(0.175)	**(1.870)***	**(1.617)**	(-1.542)

Table 5 (cont.): Predictive Performance of Various Models for Price Variables*

Forecast Horizon (h)	1	3	12	24
Panel C: Producer Price Index (PPI)				
Model 1	2.142	2.152	2.351	2.198
Model 2	2.445	2.360	2.433	2.385
	(-1.813)†	(-1.349)	(-0.660)	(-1.232)
Model 3	3.140	4.070	3.625	3.737
	(-3.026)†	(-3.407)†	(-3.214)†	(-3.404)†
Model 4	2.201	2.413	**2.300**	2.421
	(-0.387)	(-1.424)	**(0.370)**	(-1.599)
Model 5	2.282	2.391	2.370	2.536
	(-1.143)	(-1.339)	(-0.152)	(-1.576)
Model 6	2.203	2.392	**2.256**	2.303
	(-0.402)	(-1.320)	**(0.729)**	(-0.743)
Model 7	2.332	2.480	**2.273**	2.420
	(-1.205)	(-1.828)†	**(0.632)**	(-1.110)
Model 8	2.206	2.397	**2.257**	2.332
	(-0.420)	(-1.351)	**(0.730)**	(-1.021)
Model 9	**2.115**	2.192	**2.245**	2.238
	(0.394)	(-0.769)	**(1.369)**	(-0.352)
Model 10	2.217	2.474	**2.345**	2.407
	(-0.465)	(-1.806)†	**(0.043)**	(-1.350)
Model 11	2.199	2.409	**2.200**	2.313
	(-0.385)	(-1.569)	**(1.449)**	(-0.938)
Model 12	2.396	2.299	2.356	2.332
	(-1.654)†	(-0.888)	(-0.054)	(-1.021)
Model 13	**2.115**	2.344	**2.245**	2.238
	(0.394)	(-1.512)	**(1.369)**	(-0.352)
Model 14	2.447	2.401	2.465	2.407
	(-1.784)†	(-1.558)	(-0.912)	(-1.350)
Model 15	2.406	2.387	2.383	2.313
	(-1.650)†	(-1.337)	(-0.327)	(-0.938)

* Notes: Primary entries in this table are mean square forecast errors (MSFEs) based upon recursively constructed ex ante predictions for the period 1960:01-2003:12, using Models 1-15 (see Table 1 for an explanation of the different models). Bracketed entries are MSFE type Diebold and Mariano (DM: 1995) predictive accuracy test statistics, where Model 1 is compared with each of the other models). Entries in bold indicate instances where the alternative model (i.e. each of Models 2-15) outperforms the factor model (i.e. Model 1), as indicated by both a lower MSFE and a positive DM test stistic. Boxed MSFE entries represent the lowest MSFE value amongst all models, for a particular forecast horizon, h. DM statistic entries with a $*$ sign indicate instances where the respective alternative model significantly outperforms the factor model at a 10% significance level, whereas for entries with a † sign, the factor model significantly outperforms the alternative model at a 10% significance level, under the assumption that the DM test statistic has a standard normal limiting distribution (see above for further discussion).

Table 6: Predictive Performance of Various Models for Output, Employment and Sales Variables*

Forecast Horizon (h)	1	3	12	24
Panel A: Industrial Production				
Model 1	2.226	2.459	3.114	2.871
Model 2	2.471	2.490	**2.811**	**2.797**
	(-1.529)	(-0.192)	**(1.343)**	**(0.673)**
Model 3	4.267	3.931	4.541	5.528
	(-4.910)†	(-3.142)†	(-3.165)†	(-4.884)†
Model 4	2.804	2.655	**2.785**	**2.708**
	(-3.270)†	(-1.093)	**(1.436)**	**(1.417)**
Model 5	2.284	2.478	**3.100**	**2.747**
	(-0.419)	(-0.147)	**(0.081)**	**(0.560)**
Model 6	2.682	2.623	**2.795**	**2.688**
	(-2.613)†	(-1.039)	**(1.383)**	**(1.584)**
Model 7	2.678	**2.352**	**2.708**	**2.620**
	(-2.563)†	**(0.948)**	**(1.752)***	**(1.853)***
Model 8	2.719	2.652	**2.737**	**2.584**
	(-2.542)†	(-1.210)	**(1.598)**	**(2.195)***
Model 9	2.445	**2.406**	**2.912**	**2.681**
	(-1.542)	**(0.447)**	**(0.803)**	**(1.504)**
Model 10	2.666	**2.164**	**2.758**	**2.846**
	(-2.474)†	**(2.155)***	**(1.565)**	**(0.232)**
Model 11	2.512	**2.291**	**2.654**	**2.609**
	(-1.911)†	**(1.268)**	**(1.784)***	**(1.852)***
Model 12	2.594	2.615	**2.737**	**2.584**
	(-2.009)†	(-0.976)	**(1.598)**	**(2.195)***
Model 13	2.445	**2.402**	**2.912**	**2.681**
	(-1.542)	**(0.490)**	**(0.803)**	**(1.504)**
Model 14	2.453	**2.123**	**2.758**	**2.846**
	(-1.445)	**(2.445)***	**(1.565)**	**(0.232)**
Model 15	2.502	**2.240**	**2.654**	**2.609**
	(-1.840)†	**(1.608)**	**(1.784)***	**(1.852)***
Panel B: Personal Income Less Transfers				
Model 1	5.919	5.841	5.660	6.235
Model 2	7.167	6.811	**5.576**	**5.994**
	(-1.444)	(-1.522)	**(0.293)**	**(1.841)***
Model 3	15.316	12.858	6.533	10.327
	(-2.046)†	(-1.697)†	(-0.534)	(-1.459)
Model 4	6.408	6.028	**5.225**	**6.083**
	(-0.725)	(-0.927)	**(1.627)**	**(1.587)**
Model 5	6.030	6.028	**5.642**	**6.148**
	(-0.292)	(-1.125)	**(0.118)**	**(1.117)**
Model 6	6.373	5.996	**5.298**	**6.071**
	(-0.674)	(-0.790)	**(1.513)**	**(1.889)***
Model 7	6.570	6.249	**5.518**	**6.027**
	(-0.941)	(-1.328)	**(0.418)**	**(2.272)***
Model 8	6.368	5.991	**5.300**	**6.075**
	(-0.666)	(-0.764)	**(1.505)**	**(1.840)***
Model 9	6.334	6.147	5.690	**6.132**
	(-0.741)	(-2.102)†	(-0.074)	**(0.969)**
Model 10	6.569	6.077	**5.363**	**6.026**
	(-0.734)	(-0.834)	**(0.940)**	**(1.581)**
Model 11	6.336	6.057	**5.358**	**6.042**
	(-0.610)	(-0.782)	**(1.347)**	**(1.887)***
Model 12	6.766	6.674	**5.490**	**6.075**
	(-1.268)	(-1.327)	**(0.767)**	**(1.840)***
Model 13	6.659	6.791	5.920	**6.150**
	(-1.220)	(-1.589)	(-0.676)	**(1.004)**
Model 14	7.164	6.809	**5.587**	**6.007**
	(-1.440)	(-1.491)	**(0.269)**	**(1.548)**
Model 15	6.649	6.796	**5.482**	**6.042**
	(-1.022)	(-1.417)	**(0.936)**	**(1.887)***

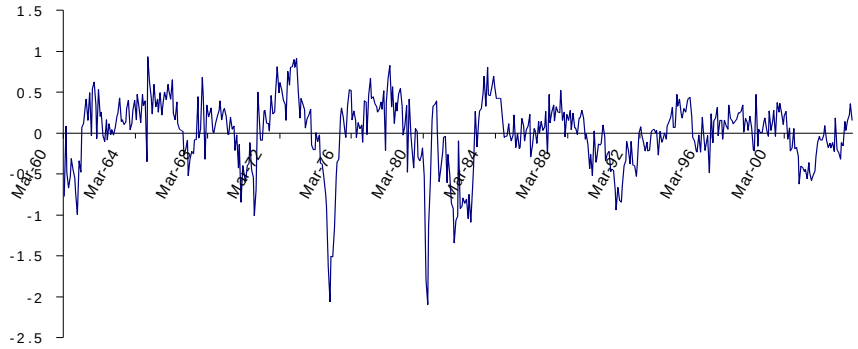
Table 6 (cont.): Predictive Performance of Various Models for Output, Employment and Sales Variables*

Forecast Horizon (h)	1	3	12	24
Panel C: Nonagricultural Employment				
Model 1	1.893	1.693	3.587	3.279
Model 2	**1.135**	**1.471**	**3.446**	3.626
	(4.013)*	**(1.323)**	**(0.561)**	(-1.836)†
Model 3	**1.655**	**1.571**	3.685	6.021
	(0.991)	**(0.542)**	(-0.239)	(-5.224)†
Model 4	2.203	2.134	3.607	3.424
	(-1.460)	(-2.614)†	(-0.079)	(-0.970)
Model 5	2.360	2.441	**3.345**	**2.726**
	(-2.191)†	(-3.580)†	**(0.977)**	**(3.068)***
Model 6	2.102	2.032	**3.566**	3.408
	(-0.982)	(-2.115)†	**(0.090)**	(-0.866)
Model 7	2.235	2.102	**3.177**	**2.992**
	(-1.323)	(-2.570)†	**(1.569)**	**(2.170)***
Model 8	2.090	2.024	**3.547**	3.426
	(-0.929)	(-2.073)†	**(0.170)**	(-0.986)
Model 9	2.223	2.219	**3.385**	**2.772**
	(-1.635)	(-3.206)†	**(0.786)**	**(2.767)***
Model 10	**1.772**	**1.632**	**3.311**	3.657
	(0.574)	**(0.333)**	**(1.066)**	(-2.064)†
Model 11	2.084	2.009	**3.029**	**2.784**
	(-0.935)	(-2.081)†	**(2.256)***	**(3.210)***
Model 12	**1.275**	1.719	**3.547**	3.426
	(3.526)*	(-0.187)	**(0.170)**	(-0.986)
Model 13	**1.327**	1.744	**3.385**	**2.772**
	(3.691)*	(-0.428)	**(0.786)**	**(2.767)***
Model 14	**1.128**	**1.406**	**3.311**	3.657
	(4.087)*	**(1.546)**	**(1.066)**	(-2.064)†
Model 15	**1.257**	1.695	**3.029**	**2.784**
	(3.825)*	(-0.015)	**(2.256)***	**(3.210)***
Panel D: Manufacturing and Trade Sales				
Model 1	7.001	8.243	8.603	8.187
Model 2	7.294	**7.729**	**8.075**	**7.920**
	(-0.639)	**(1.802)***	**(1.494)**	**(0.912)**
Model 3	21.172	12.915	15.844	18.207
	(-5.572)†	(-3.449)†	(-4.636)†	(-5.484)†
Model 4	7.811	**8.132**	**8.076**	**7.881**
	(-1.696)†	**(0.447)**	**(1.461)**	**(1.073)**
Model 5	7.885	**7.787**	**8.292**	8.425
	(-1.239)	**(2.022)***	**(0.734)**	(-0.914)
Model 6	7.541	**7.808**	**8.074**	**7.925**
	(-1.197)	**(1.895)***	**(1.451)**	**(0.915)**
Model 7	7.706	**7.890**	**8.183**	8.420
	(-1.359)	**(1.643)**	**(1.083)**	(-0.907)
Model 8	7.429	**7.795**	**8.079**	**7.926**
	(-0.959)	**(1.955)***	**(1.447)**	**(0.910)**
Model 9	7.199	**7.836**	**8.148**	**8.033**
	(-0.458)	**(1.589)**	**(1.128)**	**(0.602)**
Model 10	7.571	**7.895**	**8.091**	**7.964**
	(-1.109)	**(1.546)**	**(1.424)**	**(0.763)**
Model 11	7.465	**7.917**	**8.092**	**7.984**
	(-1.019)	**(1.585)**	**(1.237)**	**(0.687)**
Model 12	7.429	**7.795**	**8.079**	**7.926**
	(-0.959)	**(1.955)***	**(1.447)**	**(0.910)**
Model 13	7.199	**7.836**	**8.013**	**8.033**
	(-0.458)	**(1.589)**	**(1.422)**	**(0.602)**
Model 14	7.195	**7.895**	**8.091**	**7.964**
	(-0.398)	**(1.546)**	**(1.424)**	**(0.763)**
Model 15	7.465	**7.917**	**8.092**	**7.984**
	(-1.019)	**(1.585)**	**(1.237)**	**(0.687)**

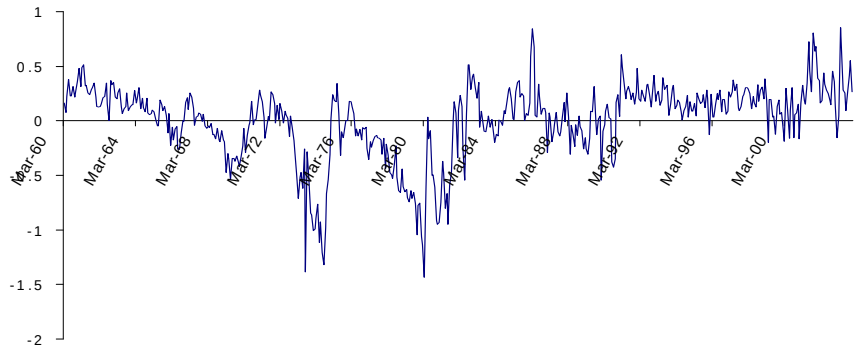
* Notes: See notes to Table 4.

Figure 1: Estimated Factors and Most Frequently Selected Factor Proxies

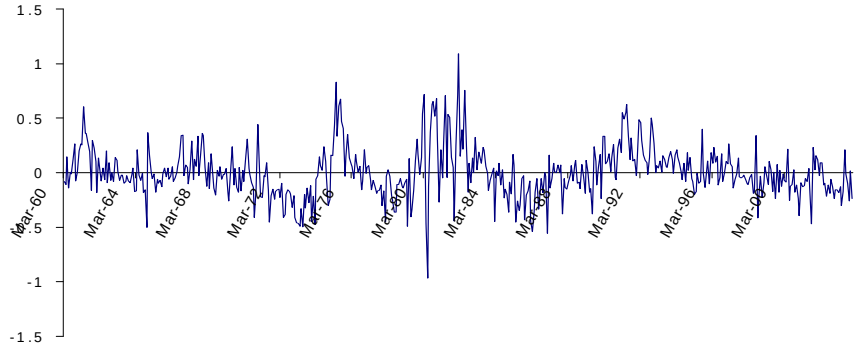
Panel 1: Estimated Factor 1



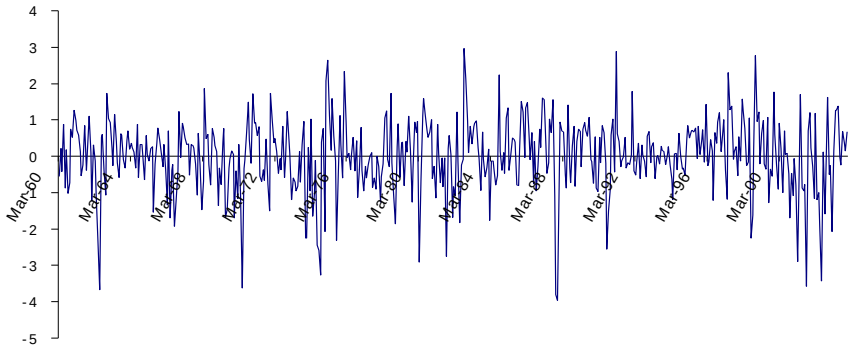
Panel 2: Estimated Factor 2



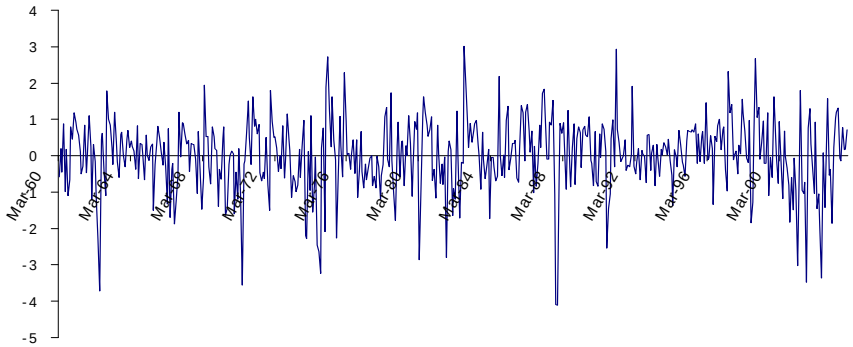
Panel 3: Estimated Factor 3



Panel 4: S&P's Common Stock Price Index, Composite



Panel 5: S&P's CommonStock Price Index, Industrials



Panel 6: Housing Starts, Nonfarm