**Macroeconomic Forecasting Using Diffusion Indexes**

James H. Stock

Kennedy School of Government, Harvard University
and the NBER

and

Mark W. Watson*

Woodrow Wilson School, Princeton University
and the NBER

April 2000
(Revised Junes 2001)

*The material in this paper originally appeared in our paper titled "Diffusion Indexes" (NBER Working Paper #6702, August 1998, revised June 1999), which this paper and Stock and Watson (2000) supercede. The authors thank Michael Boldin, Frank Diebold, Gregory Chow, Andrew Harvey, Lucrezia Reichlin, Ken Wallis, Charles Whiteman and several referees for helpful discussions and/or comments, and Lewis Chan and Alexei Onatski for skilled research assistance. This research was supported in part by National Science Foundation grants no. SBR-9409629 and SBR-9730489.

**Macroeconomic Forecasting Using Diffusion Indexes**

April 2000
(Revised June 2001)

## Abstract

This paper studies forecasting a macroeconomic time series variable using a large number of predictors. The predictors are summarized using a small number of indexes constructed by principal component analysis. An approximate dynamic factor model serves as the statistical framework for the estimation of the indexes and construction of the forecasts. The method is used to construct 6, 12, and 24 month ahead forecasts for 8 monthly U.S. macroeconomic time series using 215 predictors in simulated real time from 1970 through 1998. Over this sample period these new forecasts outperform univariate autoregressions, small vector autoregressions and leading indicator models.

Key Words: factor models, forecasting, principal components

## 1. Introduction

Recent advances in information technology make it possible to access in real time, at a reasonable cost, literally thousands of economic time series for major developed economies. This raises the prospect of a new frontier in macroeconomic forecasting, in which a very large number of time series are used to forecast a few key economic quantities such as aggregate production or inflation. Time series models currently used for macroeconomic forecasting, however, incorporate only a few series: vector autoregressions, for example, typically contain fewer than ten variables. While variable selection procedures can be used to choose a small subset of predictors from a large set of potentially useful variables, the performance of these methods ultimately rests on the few variables that are chosen. For example, real economic activity is often used to predict inflation (the so-called Philips curve), but is the unemployment rate, the rate of capacity utilization, or the GDP-gap the best measure of real activity for this purpose? An alternative to selecting a few predictors is to pool the information in all of the candidate predictors, averaging away idiosynchratic variation in the individual series. In this paper, we use an approximate factor model for this purpose. The premise is that, for forecasting puposes, the information in the large number of predictors can be replaced by a handful of estimated factors.

This idea has a long tradition in macroeconomics. For example, the notion of a common business cycle underlies the classic work of Burns and Mitchell (1947) and the indexes of leading and coincident indicators originally developed at the National Bureau of Economic Research. This notion was formally modeled by Sargent and Sims (1977) in their dynamic generalization of the classic factor analysis model. Versions of their model have been used by several researchers to study dynamic covariation among sets of variables (Geweke (1977), Singleton (1980), Engle and Watson (1981), Stock and Watson (1989,1991), Quah and Sargent (1993), and Forni and Reichlin (1996,1998)). Modern dynamic general equilibrium macroeconomic models often postulate that a small set of driving variables are responsible for variation in macro time series, and these variables can be viewed as a set of common factors. While the previous empirical research has focused on estimating indexes of covariation, this paper uses the estimated factors for prediction.

The approximate dynamic factor model, which relates the variable to be forecast, $y_{t+1}$, to a set of predictors collected in the vector $X_t$, is presented in section 2. Forecasting is carried out in a two-step process: first the factors are estimated (by principal components) using $X_t$, then these estimated factors are used to forecast $y_{t+1}$. Focusing on the forecasts implied by the factors rather than on the factors themselves permits sidestepping the difficult problem of identification (or rotation) inherent in factor models. One interpretation of the estimated factors is in terms of "diffusion" indexes developed by NBER business cycle analysts to measure common movement in a set of macroeconomic variables, and accordingly we call the estimated factors "diffusion indexes."

The performance of the diffusion index (DI) forecasts is examined in sections 3 and 4. The experiment reported in these sections simulates real time forecasting over 1970-1998 of eight U.S. macroeconomic variables, four measures each of real economic activity and of price inflation. The DI forecasts are constructed at horizons of 6, 12 and 24 months using as many as 215 predictor series. These forecasts are compared to several conventional benchmarks: univariate autogressions, small vector autoregressions, leading indicator models, and, for inflation, unemployment-based Phillips curve models. Generally speaking, the diffusion index forecasts based on a small number of factors (in most cases, one or two) are found to perform well, with relative performance improving as the horizon increases. The improvement over the benchmark forecasts can be dramatic, in several cases producing simulated out of sample mean square forecast errors that are one-third less than those of the benchmark models.

## 2. Econometric Framework

### 2.1. An approximate dynamic factor model

We begin with a discussion of the statistical model that motivates the DI forecasts. Let $y_{t+1}$ denote the scalar series to be forecast and let $X_t$ be a N-dimensional multiple time series of predictor variables, observed for $t=1,\ldots,T$, where $y_t$ and $X_t$ are both taken to have mean zero. (The different time subscripts used for y and X emphasize the forecasting relationship.) We suppose that $(X_t,y_{t+1})$ admit a dynamic factor model representation with $\bar{r}$ common dynamic factors $f_t$,

$$y_{t+1} = \beta(L)f_t + \gamma(L)y_t + \epsilon_{t+1} \tag{2.1}$$

$$X_{it} = \lambda_i(L)f_t + e_{it}, \tag{2.2}$$

for $i=1,\ldots,N$, where $e_t = (e_{1t},\ldots,e_{Nt})'$ is the $N\times 1$ idiosyncratic disturbance and $\lambda_i(L)$ and $\beta(L)$ are lag polynomials in non-negative powers of L. It is assumed that $E(\epsilon_{t+1} \mid f_t,y_t,X_t,f_{t-1},y_{t-1},X_{t-1},\ldots) = 0$. Thus, if $\{f_t\}$, $\beta(L)$, and $\gamma(L)$ were known, the minimum mean square error forecast of $y_{T+1}$ would be $\beta(L)f_T + \gamma(L)y_T$.

We make two important modifications to (2.1) and (2.2). First, the lag polynomials $\lambda_i(L)$, $\beta(L)$ and $\gamma(L)$ are modeled as having finite orders of at most q, so $\lambda_i(L)=\Sigma_{j=0}^{q}\lambda_{ij}L^j$ and $\beta(L)=\Sigma_{j=0}^{q}\beta_jL^j$. The finite lag assumption permits rewriting (2.1) and (2.2) as,

$$y_{t+1} = \beta'F_t + \gamma(L)y_t + \epsilon_{t+1} \tag{2.3}$$

$$X_t = \Lambda F_t + e_t \tag{2.4}$$

where $F_t = (f_t',\ldots,f_{t-q}')'$ is $r\times 1$, where $r\leq(q+1)\bar{r}$, the i-th row of $\Lambda$ in (2.4) is $(\lambda_{i0},\ldots,\lambda_{iq})$, and $\beta=(\beta_0,\ldots,\beta_q)'$. The main advantage of this static representation of the dynamic factor model is that the factors can be estimated using principal components. This comes at a cost, because the assumption is inconsistent with infinite distributed lags of the factors. Whether this cost is large is ultimately an empirical question, addressed here by studying whether (2.3)-(2.4) can be used to produce accurate forecasts.

Second, our empirical application focuses on h step ahead forecasts. At least two approaches to multistep forecasting are possible. One is to develop a vector time series model for $F_t$, to estimate this using the estimated factors, and to roll the $(y_t, F_t)$ model forward, but this entails estimating a large number of parameters which could erode forecast performance. Another approach is to recognize that the ensuing multistep forecasts would be linear in $F_t$ and $y_t$ (and lags), and to use an h-step ahead projection to construct the forecasts directly. We adopt the latter approach, and the resulting multi-step ahead version of (2.3) is,

$$y^h_{t+h} = \alpha_h + \beta_h(L)F_t + \gamma_h(L)y_t + \epsilon^h_{t+h} \tag{2.5}$$

where $y^h_{t+h}$ is the h-step ahead variable to be forecast, the constant term is introduced explicitly, and the subscripts reflect the dependence of the projection on the horizon.

### 2.2. Estimation and forecasting

Because $\{F_t\}$, $\alpha_h$, $\beta_h(L)$, and $\gamma_h(L)$ are unknown, forecasts of $y_{T+h}$ based on (2.4) and (2.5) are constructed using a two-step procedure. First, the sample data $\{X_t\}_{t=1}^{T}$ are used to estimate a time series of factors (the diffusion indexes), $\{\hat{F}_t\}_{t=1}^{T}$. Second, the estimators $\hat{\alpha}_h$, $\hat{\beta}_h(L)$ and $\hat{\gamma}_h(L)$ are obtained by regressing $y_{t+1}$ onto a constant, $\hat{F}_t$ and $y_t$ (and lags). The forecast of $y^h_{T+h}$ is then formed as $\hat{\alpha}_h + \hat{\beta}_h(L)\hat{F}_T + \hat{\gamma}_h(L)y_T$.

Stock and Watson (1998) develop theoretical results for this two-step procedure applied to (2.3) and (2.4). The factors are estimated by principal components because these estimators are readily calculated even for very large N and because principal components can be generalized to handle data irregularities as discussed below. Under a set of moment conditions for ($\epsilon$,e,F) and an asymptotic rank condition on $\Lambda$, the feasible forecast is asymptotically first-order efficient in the sense that its mean square forecast error (MSE) approaches the MSE of the optimal infeasible forecast as $N,T\rightarrow\infty$ where $N=O(T^{\rho})$ for any $\rho>1$. This result suggests that feasible forecasts are likely to be nearly optimal when N and T are large, regardless of the ratio of N to T. The assumptions in Stock and Watson (1998) are similar to assumptions made in the literature on approximate factor models (Chamberlain and Rothschild (1983), Connor and Korajczyk (1986), (1988) and (1993)), generalized to allow for serial correlation. A related dynamic generalization and estimation (but not forecasting) results are discussed in Forni, Hallin, Lippi and Reichlin (2000). Stock and Watson (1998) also show that the principal components remain consistent when there is some time variation in $\Lambda$ and small amounts of data contamination, as long as the number of predictors is very large, $N>>T$.

### 2.3. Data irregularities and computational issues

In our data set, some series contain missing observations and/or are available over a diminished time span. Although our data are all monthly, further complications would arise in applications in which mixed sampling frequencies are used, e.g. monthly and quarterly. In these cases standard principal components analysis does not apply. However, the EM algorithm can be used to estimate the factors by solving a suitable minizimation problem iteratively. Details are given in appendix A.

Although the components of $X_t$ typically will be distinct time series, $X_t$ could contain multiple lags of one or more series. Because the estimated factors $F_t$ could include lags of the dynamic factors $f_t$, estimation of $F_t$ might be enhanced by augmenting a vector of distinct time series with its lags. This will be referred to below as "stacking" $X_t$ with its lags, in which case the principal components of the stacked data vector are computed.

## 3. The Data and Forecasting Experimental Design

### 3.1. Forecasting models and data

The forecasting experiment simulates real-time forecasting for eight major monthly macroeconomic variables for the United States. The complete data set spans 1959:1 - 1998:12. Four of these eight variables are the measures of real economic activity used to construct the Index of Coincident Economic Indicators maintained by The Conference Board (formerly by the U.S. Department of Commerce): total industrial production (ip); real personal income less transfers (gmyxpq); real manufacturing and trade sales (msmtq); and the number of employees on nonagricultural payrolls (lpnag) (additional details are given in Appendix B, which lists series by the mnemonics given here in parenthesis). The remaining four series are price indexes: the consumer price index (punew); the personal consumption expenditure implicit price deflator (gmdc); the CPI less food and energy (puxx); and the producer price index for finished goods (pwfsa). These series and the predictor series were taken from the May 1999 release of the DRI/McGraw Hill Basic Economics database (formerly Citibase). In general these series represent the fully revised historical series available as of May 1999, and in this regard the forecasting results will differ from results that would be calculated using real-time data.

For each series, several forecasting models are compared at the 6, 12, and 24 month forecasting horizons: DI forecasts based on estimated factors; a benchmark univariate autoregression; and benchmark multivariate models. For both the real and price series, one of the benchmark multivariate models is a trivariate vector autoregression, and a second is based on leading economic indicators. As a further comparison, inflation forecasts are also computed using an unemployment-based Phillips curve.

Our focus is on multistep ahead prediction, and most of the forecasting regressions are projections of a h-step ahead variable $y^h_{t+h}$ onto t-dated predictors, sometimes including lagged transformed values $y_t$ of the variable of interest. The real variables are modeled as being I(1) in logarithms. Because all four real variables are treated identically, consider industrial production, for which

$$(3.1)\qquad y^h_{t+h} = (1200/h)\ln(IP_{t+h}/IP_t) \text{ and } y_t = 1200\ln(IP_t/IP_{t-1}).$$

The price indexes are modeled as being I(2) in logarithms. The I(2) specification is consistent with standard Phillips curve equations and is a good description of the series over much of the sample period. However, I(1) specifications also provide adequate descriptions of the data, particularly in the early part of the sample. Stock and Watson (1999) find little difference in I(1) and I(2) factor model forecasts for these prices over the sample period studied here, so for the sake of brevity we limit our analysis here to the I(2) specification. Accordingly, for the CPI (and similarly for the other price series),

$$(3.2)\qquad y^h_{t+h} = (1200/h)\ln(CPI_{t+h}/CPI_t)\text{-}1200\ln(CPI_t/CPI_{t-1}) \text{ and } y_t = 1200\Delta\ln(CPI_t/CPI_{t-1}).$$

*Diffusion Index forecasts*. Following (2.5), the most general DI forecasting function is,

$$(3.3)\qquad \hat{y}^h_{T+h|T} = \hat{\alpha}_h + \Sigma^m_{j=1}\hat{\beta}_{hj}'\hat{F}_{T-j+1} + \Sigma^p_{j=1}\hat{\gamma}_{hj}y_{T-j+1},$$

where $\hat{F}_t$ is the vector of k estimated factors. Results for three variants of (3.3) are reported. The first, denoted in the tables by "DI-AR,Lag", includes lags of the factors and lags of $y_t$, with k and lag orders m and p estimated by BIC, with $1 \le k \le 4$, $1 \le m \le 3$, and $0 \le p \le 6$. Thus the smallest candidate model that BIC can choose here includes only a single contemporaneous factor and excludes $y_t$. The second, denoted "DI-AR", includes contemporaneous $\hat{F}_t$, that is,

$m=1$, and k and p are chosen by BIC with $1 \leq k \leq 12$ and $0 \leq p \leq 6$. The third, denoted "DI", includes only contemporaneous $\hat{F}_t$, so $p=0$, $m=1$, and k is chosen by BIC, $1 \leq k \leq 12$.

The full data set used to estimate the factors contains 215 monthly time series for the U.S. from 1959:1-1998:12. The series were selected judgmentally to represent 14 main categories of macroeconomic time series: real output and income; employment and hours; real retail, manufacturing and trade sales; consumption; housing starts and sales; real inventories and inventory-sales ratios; orders and unfilled orders; stock prices; exchange rates; interest rates; money and credit quantity aggregates; price indexes; average hourly earnings; and miscellaneous. The list of series is given in Appendix B, and is similar to lists that we have used elsewhere (Stock and Watson [1996, 1999]). These series were taken from a somewhat longer list, from which we eliminated series with gross problems such as redefinitions. However no further pruning was performed.

The theory outlined in sections 2 assumes that $X_t$ is I(0), so these 215 series were subjected to three preliminary steps: possible transformation by taking logarithms, possible first differencing, and screening for outliers. The decision to take logarithms or to first difference the series was made judgmentally after preliminary data analysis, including inspection of the data and unit root tests. In general, logarithms were taken for all nonnegative series that were not already in rates or percentage units. Most series were first differenced. A code summarizing these transformations is given for each series in Appendix B. After these transformations, all series were further standardized to have sample mean zero and unit sample variance. Finally, the transformed data were screened automatically for outliers (generally taken to be coding errors or exceptional events such as labor strikes), and observations exceeding ten times the interquartile range from the median were replaced by missing values.

Using this transformed and screened dataset, three sets of empirical factors were constructed. The first was computed using principal components from the subset of 149 variables available for the full sample period (the balanced panel). The second set of factors was computed using the nonbalanced panel of all 215 series using the methods of Appendix A. The third set of factors was computed by stacking the 149 variables in the balanced panel with their first lags, so the augmented data vector has dimension 298. Empirical factors were then estimated by the principal components of the stacked data as discussed in section 2.

*Autoregressive forecast*. The autoregressive forecast is a univariate forecast based on (3.3) where the terms involving $\hat{F}$ are excluded. The lag order p was selected recursively by BIC with $0 \leq p \leq 6$, where $p=0$ indicates that $y_t$ and its lags are excluded.

*Vector autoregressive forecast*. The first multivariate benchmark model is a VAR with p lags each of three variables. One version of the VAR used $p=4$ lags, and another version selected p recursively by BIC. The fixed-lag VARs performed somewhat better than the BIC selected lag lengths (which often set $p=1$), and we report results for the fixed lag specifications in the results below. The variables in the VAR are a measure of the monthly growth in real activity, the change in monthly inflation, and the change in the 90 day U.S. Treasury bill rate. When used to forecast the real series, the relevant real activity variable was used and the inflation measure was CPI inflation. For forecasting inflation, the relevant price series was used and the real activity measure was industrial production. Multistep forecasts were computed by iterating the VAR forward. This contrasts to the AR forecasts, which were computed by h-step ahead projection rather than iteration.

*Multivariate leading indicator forecasts*. The leading indicator forecasts have the form,

$$\hat{y}^h_{T+h|T} = \hat{\delta}_{h0} + \Sigma^m_{j=1}\hat{\delta}_{hi}'W_{T-j+1} + \Sigma^p_{j=1}\hat{\gamma}_{hj}y_{T-j+1}, \tag{3.4}$$

where $W_t$ is a vector of leading indicators that have been featured in the literature and/or in real-time forecasting applications and $\hat{\delta}_{h0}$, etc. are OLS coefficient estimates.

For the real variables, $W_t$ consists of eleven leading indicators that we have used for real-time monthly forecasting in experimental leading and recession indicators; see Stock and Watson (1989). (The list used here consists of the leading indicators used to produce the XRI and the XRI-2, which are released monthly (and documented) at the web site http://www.nber.org.) Five of these leading indicators are also used in the factor estimation step in the diffusion index forecasts. These are: average weekly hours of production workers in manufacturing (lphrm); the capacity utilization rate in manufacturing (ipxmca); housing starts (building permits) (hsbr); the index of help-wanted advertising in newspapers (lhel); and the interest rate on 10-year U.S. Treasury bonds (fygt10). The remaining six leading indicators are: the interest rate spread between 3-month U.S. Treasury bills and 3-month commercial paper; the spread between 10-year and 1-year U.S. Treasury bonds; the number of people working part-time in nonagricultural industries because of slack work; real manufacturers' unfilled orders in durable goods industries; a trade-weighted index of nominal exchange rates between the U.S. and the U.K., West Germany, France, Italy, and Japan; and the National Association of Purchasing Managers' index of vendor performance (the percent of companies reporting slower deliveries).

For the inflation forecasts, eight leading indicators are used. These variables were chosen because of their good individual performance in previous inflation forecasting exercises. In particular these variables performed well in at least one of the historical episodes considered in Staiger, Stock and Watson (1997) (also see Stock and Watson [1999]). Seven of these variables are also used in the factor estimation step in the diffusion index forecasts: the total unemployment rate (lhur); real manufacturing and trade sales (msmtq); housing starts (hsbr); new orders in durable goods industries (mdoq); the nominal M1 money supply (fm1); the federal funds overnight interest rate (fyff); and the interest rate spread between 1-year U.S. Treasury bonds and the federal funds rate (sfygt1). The remaining variable is the trade-weighted exchange rate listed in the previous paragraph.

In all cases, the leading indicators were transformed so that $W_t$ is I(0). This entailed taking logarithms of variables not already in rates, and differencing all variables except the interest rate spreads, housing starts, the index of vendor performance, and the help wanted index.

For each variable to be forecast, p and m in (3.4) were determined by recursive BIC with $1 \leq m \leq 4$ and $0 \leq p \leq 6$, so 28 possible models were compared in each time period.

*Phillips curve forecasts*. The unemployment-based Phillips curve is considered by many to have been a reliable method for forecasting inflation over this period, cf. Gordon (1982) and, more recently, the Congressional Budget Office (1996), Fuhrer (1995), Gordon (1997), Staiger, Stock and Watson (1997), and Tootel (1994). The Phillips curve inflation forecasts considered here have the form (3.4), where $W_t$ consists of: the unemployment rate (LHUR) and m-1 of its lags; the relative price of food and energy (current and one lagged value only); and Gordon's (1982) variable that controls for the imposition and removal of the Nixon wage and price controls. The wage and price control variable is introduced for forecasts made in 1971:7+h, before which it produces singular regressions. The lag lengths m and p were chosen by recursive BIC, where $1 \leq m \leq 6$ and $0 \leq p \leq 6$.

### 3.2. Simulated real-time experimental design

Estimation and forecasting was conducted to simulate real-time forecasting. This entailed fully recursive parameter estimation, factor estimation, model selection, etc. The first simulated out of sample forecast was made in 1970:1. To construct this forecast, the data were screened for outliers and standardized, the parameters and factors were estimated, and the models were selected, using only data available from 1959:1 through 1970:1 (the first date for the regressions was 1960:1, with earlier observations used for initial conditions as needed). Thus regressions (3.3) and (3.4) were run for t=1960:1,...,1970:1-h, then the values of the regressors at t=1970:1 were used to forecast $y^h_{1970:1+h}$. All parameters, factors, etc. were then reestimated, information criteria were recomputed, and models were selected using data from 1959:1 through 1970:2, and forecasts from these models were then computed for $y^h_{1970:2+h}$. The final simulated out of sample forecast was made in 1998:12-h for $y^h_{1998:12}$.

## 4. Empirical Results

### 4.1. Forecasting results

The results for the real variables are reported in detail in table 1 for 12-month ahead forecasts, and summaries for 6- and 24-month ahead forecasts are reported in table 2. Two sets of statistics are reported. The first is the mean squared error (MSE) of the candidate forecasting model, computed relative to the MSE of the univariate autoregressive forecast (so the AR forecast has a relative MSE of 1.00). For example, the simulated out of sample MSE of the leading indicator (LI) forecast of industrial production is 86% that of the AR forecast at the 12 month horizon. Autocorrelation consistent standard errors for these relative MSEs, calculated following West (1996), are reported in parentheses. The second set of statistics are the coefficient on the candidate forecast from the forecast combining regression,

$$y^h_{t+h} = \alpha\hat{y}^h_{t+h|t} + (1-\alpha)\hat{y}^{h,AR}_{t+h|t} + u^h_{t+h} \tag{4.1}$$

where $\hat{y}^h_{t+h|t}$ is the candidate h-step ahead forecast and $\hat{y}^{h,AR}_{t+h|t}$ is the benchmark h-step ahead AR forecast. HAC standard errors for $\alpha$ are reported in parentheses. For example, $\alpha$ is estimated to be .57 when the candidate forecast is the leading indicator forecast at the 12 month horizon, with a standard error of .13, so the hypothesis that the weight on the leading indicator forecast is zero ($\alpha$=0) is rejected at the 5% level, but so is the hypothesis that the leading indicator forecast receives unit weight.

We now turn to the results for the real variables. First consider the DI forecasts with factors estimated using the full data set (the unbalanced panel). These forecasts with BIC factor selection generally improve substantially over the benchmark univariate and multivariate forecasts. The DI-AR,Lag model, which allows recursive BIC selection across own lags and lags of the factors, outperforms all three benchmark models in 10 of the 12 variable/horizon combinations, the exceptions being 6- and 12-month ahead forecasts of employment. In most cases the performance of the simpler DI forecasts, which exclude lags of $\hat{F}_t$ and $y_t$, is comparable to or even better than that of the DI-AR,Lag forecasts. This is rather surprising, because it implies that essentially all the predictable dynamics of these series are accounted for by the estimated factors. In some cases, the improvement over the benchmark forecasts are quite substantial, for example, for industrial production at the 12 month horizon the DI-AR,Lag forecast has a forecast error variance 57% that of the AR model and two-thirds that of the leading indicator model. The relative improvements are more modest at the 6 month horizon. At the 24 month horizon, the multivariate benchmark forecasts break down and perform worse than the univariate forecast, however the DI-AR,Lag, DI-AR, and DI forecasts continue to outperform the AR benchmark very substantially.

The performance of comparable models is usually better when the empirical factors from the full data set are used, relative to those from the balanced panel subset. Performance is not improved by using empirical factors from augmenting the balanced panel with its first lag; for these real series, doing so does comparably, or somewhat worse, than using the empirical factors from the unstacked balanced panel.

Inspection of the final panels of tables 1 and 2 reveals a striking finding: simply using DI or DI-AR forecasts with two factors captures most of the forecasting improvement. In most cases, incorporating BIC factor and lag order selection provides little or no improvement over just using two factors, with no lags of the factors and no lagged dependent variables.

The results for the price series are given in tables 3 and 4. There are three notable differences in these results, relative to those for the real variables. First, the DI-AR,Lag forecasts outperform all the benchmark forecasts less often, in only 6 of the 12 variable/horizon combinations. Second, including lagged inflation dramatically improves the forecasts, and without this the DI forecasts are actually worse than the autoregressive forecasts. Third, other factor forecasts generally outperform the DI-AR,Lag forecasts. Notably, the full data set DI-AR forecast with k=1 (and no lagged factors) outperforms all the benchmarks in 11 of 12 cases, and typically improves upon the DI-AR lag. Thus most of the forecasting gains seem to come from using a single factor.

As with the real variables, forecasts based on the stacked data perform less well than those based on the unstacked data. While the full data set forecasts are typically better than the balanced panel subset forecasts for the 6 and 12 month horizons, at the 24 month horizon the balanced panel forecasts slightly outperform the full data set forecasts.

Additional analysis of factor-based forecasts of CPI and consumption deflator inflation, and additional comparisons of these forecasts to other Phillips-curve forecasts and to forecasts based on other leading indicators, are contained in Stock and Watson (1999). Three findings from that study are worth noting here. First, the DI-AR and DI-AR,Lag forecasts are found to perform well relative to a large number of additional multivariate benchmarks. Second, the forecasts reported here can be further improved upon using a single-factor forecast, where the factor is computed from a set of variables that measure only real economic activity. Forecasts based on this real economic activity factor have MSEs approximately 10% less than the best forecasts reported in table 3. Finally, similar rankings of methods are obtained using I(1) forecasting models, rather than the I(2) models used here, that is, when first rather than second differences of log prices are used for the forecasting equation and factor estimation.

In interpreting these results, it should be stressed that the multivariate leading indicator models are sophisticated forecasting tools that provide a stiff benchmark against which to judge the diffusion index forecasts. In our judgment, the performance of the leading indicator models reported here overstates their true potential out of sample performance, because the lists of leading indicators used to construct the forecasts were chosen by model selection methods based on their forecasting performance over the past two decades, as discussed in section 3. In this light, we consider the performance of the various diffusion index models to be particularly encouraging.

### 4.2. Empirical factors

Because the factors are identified only up to a $k \times k$ matrix, detailed discussion of the individual factors is unwarranted. Nevertheless the finding that good forecasts can be made with only one or two factors suggests briefly characterizing the first few factors.

Figure 1 therefore displays the $R^2$s of the regressions of the 215 individual time series against each of the first six empirical factors from the balanced panel subset, estimated over the full sample period. These $R^2$s are plotted as bar charts with one chart for each factor. (The series are grouped by category and ordered numerically using the ordering in the appendix.) Broadly speaking, the first factor loads primarily on output and employment; the second factor on interest rate spreads, unemployment rates, and capacity utilization rates; the third, on interest rates; the fourth, on stock returns; the fifth, on inflation; and the sixth, on housing starts. Taken together, these six factors account for 39% of the variance of the 215 monthly time series in the full data set, as measured by the trace-$R^2$; the first twelve factors together account for 53% of the variance of these series. (The contributions to the trace-$R^2$ by the first six factors are, respectively: 0.137, 0.085, 0.048, 0.040, 0.034, and 0.041, for a total of 0.385.)

## 5. Discussion and Conclusions

We find two features of the empirical results surprising and intriguing. First, only six factors account for much of the variance of our 215 time series. One interpretation of this result is that there are only a few important sources of macroeconomic variability. Second, just a few factors are needed to forecast real activity and the most accurate forecasts of inflation use lags of inflation together with a single factor. This suggests that a very small state vector may be necessary for forecasting macroeconomic time series.

These results raise several issues for future empirical and theoretical research. We mention five here. First, classical diffusion indexes are computed using nonlinear transformations of the data, but our indexes are linear functions of the data. This raises the possibility that there are further forecasting gains that can be realized using a nonlinear version of the dynamic factor model. Second, the results reported here rely on monthly data, but data from other sampling frequencies (weekly, quarterly, etc.) may improve the forecasts. A computational algorithm for estimating the factors with mixed frequency data is outlined in Appendix A. Third, we have considered only U.S. data, and it would be useful to study the relative forecasting performance of these methods for other countries. Fourth, the estimated factors that we used here were based on simple estimators and it would be useful to study other estimators designed to exploit the heteroskedasticity and serial correlation in the data to improve efficiency. Finally, our results are based on 215 time series chosen judgementally from the large number of available macroeconomic time series. Would there be additional improvements if we were to use 500 series, or much of loss to restricting ourselves to only 100 series? Alternatively, the problem of systematically selecting many series from very many series is a difficult problem that requires further research.

## Appendix A: EM Estimation with an Unbalanced Panel and Data Irregularities

In practice, when N is large one encounters various data irregularities, including occasionally missing observations, unbalanced panels, and mixed frequency (e.g., monthly and quarterly) data. In this case, a modification of standard principal component estimation is necessary. To motivate the modification, consider the least squares estimators of $\Lambda$ and $F_t$ from (2.4) from a balanced panel. The objective function is

$$V(F,\Lambda) = \Sigma_{i=1}^{N} \Sigma_{t=1}^{T} (X_{it} - \lambda_i' F_t)^2. \tag{A.1}$$

where $\lambda_i$ is the i'th row of $\Lambda$. (A.1) can be minimized by the usual eigenvalue calculations and $\hat{F}_t$ are the principal components of $X_t$.

When the panel is unbalanced, least squares estimators of $F_t$ can be calculated from the objective function

$$V^{\dagger}(F,\Lambda) = \Sigma_{i=1}^{N} \Sigma_{t=1}^{T} I_{it}(X_{it} - \lambda_i' F_t)^2. \tag{A.2}$$

where $I_{it}=1$ if $X_{it}$ is available and 0 otherwise. Minimization of (A.2) requires iterative methods. This appendix summarizes an iterative method based on the EM algorithm that has proven to be easy and effective.

To motivate this EM algorithm, notice that $V(F,\Lambda)$ is proportional to the log-likelihood under the assumption that $X_{it}$ are iid $N(\lambda_i' F_t, 1)$, in which case the least squares estimators are the Gaussian MLEs. Because $V^{\dagger}$ is just a "missing data" version of V and because minimization of V is computationally simple, a simple EM algorithm can be constructed to minimize $V^{\dagger}$.

The j-th iteration of the algorithm is defined as follows. Let $\hat{\Lambda}$ and $\hat{F}$ denote estimates of $\Lambda$ and F constructed from the (j-1)'st iteration, and let

(A.3) $$Q(X^\dagger,\hat{F},\hat{\Lambda},F,\Lambda) = E_{\hat{F},\hat{\Lambda}}[V(F,\Lambda)|X^\dagger]$$

where $X^\dagger$ denotes the full set of observed data and $E_{\hat{F},\hat{\Lambda}}[V(F,\Lambda)|X^\dagger]$ is the expected value of the "complete data" log-likelihood $V(F,\Lambda)$, evaluated using the conditional density of $X|X^\dagger$ evaluated at $\hat{F}$ and $\hat{\Lambda}$. The estimates of F and $\Lambda$ at iteration j solve $\mathrm{Min}_{F,\Lambda}Q(X^\dagger,\hat{F},\hat{\Lambda},F,\Lambda)$.

To carry out the calculations, note that

(A.4) $$Q(X^\dagger,\hat{F},\hat{\Lambda},F,\Lambda)=\Sigma_i\Sigma_t\{E_{\hat{F},\hat{\Lambda}}(X_{it}^2|X^\dagger)+(\lambda_i'F_t)^2-2\hat{X}_{it}(\lambda_i'F_t)\}$$

where $\hat{X}_{it}=E_{\hat{F},\hat{\Lambda}}(X_{it}|X^\dagger)$. The first term on the right hand side of (A.4) does not depend on F or $\Lambda$, and so for purposes of minimization it can be replaced by $\Sigma_i\Sigma_t\hat{X}_{it}^{\;2}$. This implies that the values of F and $\Lambda$ that minimize (A.4) can be calculated as the minimizers of $\hat{V}(F,\Lambda) = \Sigma_i\Sigma_t(\hat{X}_{it}-\lambda_i'F_t)^2$. At the jth step, this reduces to the usual principal component eigenvalue calculation where the missing data are replaced by their expectation conditional on the observed data and using the parameter values from the previous iteration. If the full data set contains a subset that constitutes a balanced panel, then starting values for $\hat{F}$ in the EM iteration can be obtained using estimates from the balanced panel subset.

We now provide some additional details on the calculation of $\hat{X}_{it}$ for some important special cases. Let $\underline{X}_i = (X_{i1},\ldots,X_{iT})'$, and let $\underline{X}_i^\dagger$ be the vector of observations on the i-th variable. Suppose that $\underline{X}_i^\dagger=A_i\underline{X}_i$ for some known matrix $A_i$, as can be done in the cases of missing values and temporal aggregation, for example. Then $E(\underline{X}_i|X^\dagger) = E(\underline{X}_i|\underline{X}_i^\dagger) = F\lambda_i+A_i'(A_iA_i')^-(\underline{X}_i^\dagger-A_iF\lambda_i)$, where $(A_iA_i')^-$ is the generalized inverse of $A_iA_i'$. The particulars of these calculations are now presented for some important special cases. In the first four special cases discussed below, this level of generality is unnecessary and the formula for $\hat{X}_{it}$ follows quite simply from the nature of the data irregularity.

*A. Missing observations.* Suppose some observations on $X_{it}$ are missing. Then, during iteration j, the elements of the estimated balanced panel are constructed as $\hat{X}_{it}=X_{it}$ if $X_{it}$ observed, and $\hat{X}_{it}=\hat{\lambda}_i'\hat{F}_t$ otherwise. The estimate of F is then updated by computing the eigenvectors corresponding to the largest r eigenvalues of $N^{-1}\sum_i \underline{\hat{X}}_i\underline{\hat{X}}_i'$ where $\underline{\hat{X}}_i=(\hat{X}_{i1}, \hat{X}_{i2},\ldots,\hat{X}_{iT})'$. The estimate of $\Lambda$ is updated by the OLS regression of $\hat{X}$ onto this updated estimate of F.

*B. Mixed monthly/quarterly data - I(0) stock variables.* A series that is observed quarterly and is a stock variable would be the point-in-time level of a variable at the end of the quarter, say the level of inventories at the end of the quarter. If this series is I(0) then it is handled as in case A, that is, it is treated as a monthly series with missing observations in the first and second months of the quarter.

*C. Mixed monthly/quarterly data - I(0) flow variables.* A quarterly flow variable is the average (or sum) of unobserved monthly values. If this series is I(0), it can be treated as follows. The unobserved monthly series, $X_{it}$, is measured only as the time aggregate $X^q_{it}$ where $X^q_{it}=(1/3)(X_{i,t-2}+X_{i,t-1}+X_{it})$ for t=3,6,9,12,..., and $X^q_{it}$ is missing for all other values of t. In this case estimation procedes as in case A, but with $\hat{X}_{it}=\hat{\lambda}_i'\hat{F}_t + \hat{e}_{it}$, where $\hat{e}_{it}=X^q_{i\tau}-\hat{\lambda}_i'(\hat{F}_{\tau-2}+\hat{F}_{\tau-1}+\hat{F}_{\tau})/3$, where $\tau=3$ when t=1,2,3, $\tau=6$, when t=4,5,6, etc.

*D. Mixed monthly/quarterly data - I(1) stock variables.* Suppose that underlying monthly data are I(1) and let $X^q_{it}$ denote the quarterly first difference stock variable, assumed to be measured in the third month of every quarter, and let $X_{it}$ denote the monthly first difference of the variable. Then $X^q_{it}=(X_{i,t-2}+X_{i,t-1}+X_{it})$ for t=3,6,9,12,..., and $X^q_{it}$ is missing for all other values of t. In this case estimation procedes as in case A, but with $\hat{X}_{it}=\hat{\lambda}_i'\hat{F}_t+(1/3)\hat{e}_{it}$, where $\hat{e}_{it}=X^q_{i\tau}-\hat{\lambda}_i'(\hat{F}_{\tau-2}+\hat{F}_{\tau-1}+\hat{F}_{\tau})$, where $\tau=3$ when t=1,2,3, $\tau=6$, when t=4,5,6, etc.

*E. Mixed monthly/quarterly data - I(1) flow variables.*

Construction of $\hat{X}_{it}$ is more difficult here than in the earlier cases. Here the general regression formula given above can be implemented after specifying $\underline{X}_i^\dagger$ and $A_i$. Let the quarterly first differences be denoted by $X^q_{it}$, which is assumed to be observed at the end of every quarter. The vector of observations is then $\underline{X}_i^\dagger=(X^q_{i3}, X^q_{i6},\ldots,X^q_{i\tau})'$, where $\tau$ denotes the month of the last quarterly observation. If the underlying quarterly data are averages of monthly series, and if the monthly first differences are denoted by $X_{it}$, then $X^q_{it}=(1/3)(X_{i,t}+2X_{i,t-1}+3X_{it-2}+2X_{it-3}+X_{it-4})$ for t=3,6,9,12,..., and this implicitly defines the rows of $A_i$. Then the estimate of $\underline{X}_i$ is given by $\hat{\underline{X}}_i=F\lambda_i+A_i'(A_iA_i')^{-1}(\underline{X}_i^\dagger-A_iF\lambda_i)$.

## Appendix B: Data Description

This appendix lists the time series used to construct the diffusion index forecasts discussed in section 5. The format is: series number; series mnemonic; data span used; transformation code; and brief series description. The transformation codes are: 1 = no transformation; 2 = first difference; 4 = logarithm; 5 = first difference of logarithms; 6 = second difference of logarithms. An asterisk after the date denotes a series that was included in the unbalanced panel but not the balanced panel, either because of missing data or because of gross outliers which were treated as missing data. The series were either taken directly from the DRI-McGraw Hill Basic Economics database, in which case the original mnemonics are used, or they were produced by authors' calculations based on data from that database, in which case the authors calculations and original DRI/McGraw series mnemonics are summarized in the data description field. The following abbreviations appear in the data definitions: SA = seasonally adjusted; NSA = not seasonally adjusted; SAAR = seasonally adjusted at an annual rate; FRB = Federal Reserve Board; AC = Authors calculations

Real output and income (Out)

| | | | | |
|---|---|---|---|---|
| 1. | ip | 1959:01-1998:12 | 5 | industrial production: total index (1992=100,sa) |
| 2. | ipp | 1959:01-1998:12 | 5 | industrial production: products, total (1992=100,sa) |
| 3. | ipf | 1959:01-1998:12 | 5 | industrial production: final products (1992=100,sa) |
| 4. | ipc | 1959:01-1998:12 | 5 | industrial production: consumer goods (1992=100,sa) |
| 5. | ipcd | 1959:01-1998:12 | 5 | industrial production: durable consumer goods (1992=100,sa) |
| 6. | ipcn | 1959:01-1998:12 | 5 | industrial production: nondurable condsumer goods (1992=100,sa) |
| 7. | ipe | 1959:01-1998:12 | 5 | industrial production: business equipment (1992=100,sa) |
| 8. | ipi | 1959:01-1998:12 | 5 | industrial production: intermediate products (1992=100,sa) |
| 9. | ipm | 1959:01-1998:12 | 5 | industrial production: materials (1992=100,sa) |
| 10. | ipmd | 1959:01-1998:12* | 5 | industrial production: durable goods materials (1992=100,sa) |
| 11. | ipmnd | 1959:01-1998:12 | 5 | industrial production: nondurable goods materials (1992=100,sa) |
| 12. | ipmfg | 1959:01-1998:12 | 5 | industrial production: manufacturing (1992=100,sa) |
| 13. | ipd | 1959:01-1998:12 | 5 | industrial production: durable manufacturing (1992=100,sa) |

| | | | | |
|---|---|---|---|---|
| 14. | ipn | 1959:01-1998:12 | 5 | industrial production: nondurable manufacturing (1992=100,sa) |
| 15. | ipmin | 1959:01-1998:12 | 5 | industrial production: mining (1992=100,sa) |
| 16. | iput | 1959:01-1998:12 | 5 | industrial production: utilities (1992-=100,sa) |
| 17. | ipx | 1967:01-1998:12* | 1 | capacity util rate: total industry (% of capacity,sa)(frb) |
| 18. | ipxmca | 1959:01-1998:12 | 1 | capacity util rate: manufacturing,total(% of capacity,sa)(frb) |
| 19. | ipxdca | 1967:01-1998:12* | 1 | capacity util rate: durable mfg (% of capacity,sa)(frb) |
| 20. | ipxnca | 1967:01-1998:12* | 1 | capacity util rate: nondurable mfg (% of capacity,sa)(frb) |
| 21. | ipxmin | 1967:01-1998:12* | 1 | capacity util rate: mining (% of capacity,sa)(frb) |
| 22. | ipxut | 1967:01-1998:12* | 1 | capacity util rate: utilities (% of capacity,sa)(frb) |
| 23. | pmi | 1959:01-1998:12 | 1 | purchasing managers' index (sa) |
| 24. | pmp | 1959:01-1998:12 | 1 | NAPM production index (percent) |
| 25. | gmpyq | 1959:01-1998:12* | 5 | personal income (chained) (series #52) (bil 92$,saar) |
| 26. | gmyxpq | 1959:01-1998:12 | 5 | personal income less transfer payments (chained) (#51) (bil 92$,saar) |

Employment and hours (EMP)

| | | | | |
|---|---|---|---|---|
| 27. | lhel | 1959:01-1998:12 | 5 | index of help-wanted advertising in newspapers (1967=100;sa) |
| 28. | lhelx | 1959:01-1998:12 | 4 | employment: ratio; help-wanted ads:no. unemployed clf |
| 29. | lhem | 1959:01-1998:12 | 5 | civilian labor force: employed, total (thous.,sa) |
| 30. | lhnag | 1959:01-1998:12 | 5 | civilian labor force: employed, nonagric.industries (thous.,sa) |
| 31. | lhur | 1959:01-1998:12 | 1 | unemployment rate: all workers, 16 years & over (%,sa) |
| 32. | lhu680 | 1959:01-1998:12 | 1 | unemploy.by duration: average(mean)duration in weeks (sa) |
| 33. | lhu5 | 1959:01-1998:12 | 1 | unemploy.by duration: persons unempl.less than 5 wks (thous.,sa) |
| 34. | lhu14 | 1959:01-1998:12 | 1 | unemploy.by duration: persons unempl.5 to 14 wks (thous.,sa) |
| 35. | lhu15 | 1959:01-1998:12 | 1 | unemploy.by duration: persons unempl.15 wks + (thous.,sa) |
| 36. | lhu26 | 1959:01-1998:12 | 1 | unemploy.by duration: persons unempl.15 to 26 wks (thous.,sa) |
| 37. | lpnag | 1959:01-1998:12 | 5 | employees on nonag. payrolls: total (thous.,sa) |
| 38. | lp | 1959:01-1998:12 | 5 | employees on nonag payrolls: total, private (thous,sa) |
| 39. | lpgd | 1959:01-1998:12 | 5 | employees on nonag. payrolls: goods-producing (thous.,sa) |
| 40. | lpmi | 1959:01-1998:12* | 5 | employees on nonag. payrolls: mining (thous.,sa) |
| 41. | lpcc | 1959:01-1998:12 | 5 | employees on nonag. payrolls: contract construction (thous.,sa) |
| 42. | lpem | 1959:01-1998:12 | 5 | employees on nonag. payrolls: manufacturing (thous.,sa) |
| 43. | lped | 1959:01-1998:12 | 5 | employees on nonag. payrolls: durable goods (thous.,sa) |
| 44. | lpen | 1959:01-1998:12 | 5 | employees on nonag. payrolls: nondurable goods (thous.,sa) |
| 45. | lpsp | 1959:01-1998:12 | 5 | employees on nonag. payrolls: service-producing (thous.,sa) |
| 46. | lptu | 1959:01-1998:12* | 5 | employees on nonag. payrolls: trans. & public utilities (thous.,sa) |
| 47. | lpt | 1959:01-1998:12 | 5 | employees on nonag. payrolls: wholesale & retail trade (thous.,sa) |
| 48. | lpfr | 1959:01-1998:12 | 5 | employees on nonag. payrolls: finance,insur.&real estate (thous.,sa |
| 49. | lps | 1959:01-1998:12 | 5 | employees on nonag. payrolls: services (thous.,sa) |
| 50. | lpgov | 1959:01-1998:12 | 5 | employees on nonag. payrolls: government (thous.,sa) |
| 51. | lw | 1964:01-1998:12* | 2 | avg. weekly hrs. of prod. wkrs.: total private (sa) |
| 52. | lphrm | 1959:01-1998:12 | 1 | avg. weekly hrs. of production wkrs.: manufacturing (sa) |
| 53. | lpmosa | 1959:01-1998:12 | 1 | avg. weekly hrs. of prod. wkrs.: mfg.,overtime hrs. (sa) |

| | | | | |
|---|---|---|---|---|
| 54. | pmemp | 1959:01-1998:12 | 1 | NAPM employment index (percent) |

Real retail, manufacturing and trade sales (RTS)

| | | | | |
|---|---|---|---|---|
| 55. | msmtq | 1959:01-1998:12 | 5 | manufacturing & trade: total (mil of chained 1992 dollars)(sa) |
| 56. | msmq | 1959:01-1998:12 | 5 | manufacturing & trade:manufacturing;total(mil of chained 1992 dollars)(sa) |
| 57. | msdq | 1959:01-1998:12 | 5 | manufacturing & trade:mfg; durable goods (mil of chained 1992 dollars)(sa) |
| 58. | msnq | 1959:01-1998:12 | 5 | manufact. & trade:mfg;nondurable goods (mil of chained 1992 dollars)(sa) |
| 59. | wtq | 1959:01-1998:12 | 5 | merchant wholesalers: total (mil of chained 1992 dollars)(sa) |
| 60. | wtdq | 1959:01-1998:12 | 5 | merchant wholesalers:durable goods total (mil of chained 1992 dollars)(sa) |
| 61. | wtnq | 1959:01-1998:12 | 5 | merchant wholesalers:nondurable goods (mil of chained 1992 dollars)(sa) |
| 62. | rtq | 1959:01-1998:12 | 5 | retail trade: total (mil of chained 1992 dollars)(sa) |
| 63. | rtnq | 1959:01-1998:12 | 5 | retail trade:nondurable goods (mil of 1992 dollars)(sa) |

Consumption (PCE)

| | | | | |
|---|---|---|---|---|
| 64. | gmcq | 1959:01-1998:12 | 5 | personal consumption expend (chained)-total (bil 92$,saar) |
| 65. | gmcdq | 1959:01-1998:12 | 5 | personal consumption expend (chained)-total durables (bil 92$,saar) |
| 66. | gmcnq | 1959:01-1998:12 | 5 | personal consumption expend (chained)-nondurables (bil 92$,saar) |
| 67. | gmcsq | 1959:01-1998:12 | 5 | personal consumption expend (chained)-services (bil 92$,saar) |
| 68. | gmcanq | 1959:01-1998:12 | 5 | personal cons expend (chained)-new cars (bil 92$,saar) |

Housing starts and sales (HSS)

| | | | | |
|---|---|---|---|---|
| 69. | hsfr | 1959:01-1998:12 | 4 | housing starts:nonfarm(1947-58);total farm&nonfarm(1959-)(thous.,sa |
| 70. | hsne | 1959:01-1998:12 | 4 | housing starts:northeast (thous.u.)s.a. |
| 71. | hsmw | 1959:01-1998:12 | 4 | housing starts:midwest(thous.u.)s.a. |
| 72. | hssou | 1959:01-1998:12 | 4 | housing starts:south (thous.u.)s.a. |
| 73. | hswst | 1959:01-1998:12 | 4 | housing starts:west (thous.u.)s.a. |
| 74. | hsbr | 1959:01-1998:12 | 4 | housing authorized: total new priv housing units (thous.,saar) |
| 75. | hsbne | 1960:01-1998:12* | 4 | houses authorized by build. permits:northeast(thou.u.)s.a |
| 76. | hsbmw | 1960:01-1998:12* | 4 | houses authorized by build. permits:midwest(thou.u.)s.a. |
| 77. | hsbsou | 1960:01-1998:12* | 4 | houses authorized by build. permits:south(thou.u.)s.a. |
| 78. | hsbwst | 1960:01-1998:12* | 4 | houses authorized by build. permits:west(thou.u.)s.a. |
| 79. | hns | 1963:01-1998:12* | 4 | new 1-family houses sold during month (thous,saar) |
| 80. | hnsne | 1973:01-1998:12* | 4 | one-family houses sold:northeast(thou.u.,s.a.) |
| 81. | hnsmw | 1973:01-1998:12* | 4 | one-family houses sold:midwest(thou.u.,s.a.) |
| 82. | hnssou | 1973:01-1998:12* | 4 | one-family houses sold:south(thou.u.,s.a.) |
| 83. | hnswst | 1973:01-1998:12* | 4 | one-family houses sold:west(thou.u.,s.a.) |
| 84. | hnr | 1963:01-1998:12* | 4 | new 1-family houses, month's supply @ current sales rate(ratio) |
| 85. | hniv | 1963:01-1998:12* | 4 | new 1-family houses for sale at end of month (thous,sa) |
| 86. | hmob | 1959:01-1998:12 | 4 | mobile homes: manufacturers' shipments (thous.of units,saar) |
| 87. | contc | 1964:01-1998:12* | 4 | construct.put in place:total priv & public 1987$(mil$,saar) |
| 88. | conpc | 1964:01-1998:12* | 4 | construct.put in place:total private 1987$(mil$,saar) |
| 89. | conqc | 1964:01-1998:12* | 4 | construct.put in place:public construction 87$(mil$,saar) |

| | | | | |
|---|---|---|---|---|
| 90. | condo9 | 1959:01-1998:10* | 4 | construct.contracts: comm'l & indus.bldgs(mil.sq.ft.floor sp.;sa) |

Real inventories and inventory-sales ratios (Inv)

| | | | | |
|---|---|---|---|---|
| 91. | ivmtq | 1959:01-1998:12 | 5 | manufacturing & trade inventories:total (mil of chained 1992)(sa) |
| 92. | ivmfgq | 1959:01-1998:12 | 5 | inventories, business, mfg (mil of chained 1992 dollars, sa) |
| 93. | ivmfdq | 1959:01-1998:12 | 5 | inventories, business durables (mil of chained 1992 dollars, sa) |
| 94. | ivmfnq | 1959:01-1998:12 | 5 | inventories, business, nondurables (mil of chained 1992 dollars, sa) |
| 95. | ivwrq | 1959:01-1998:12 | 5 | manufacturing & trade inv:merchant wholesalers (mil of chained 1992 dollars)(s |
| 96. | ivrrq | 1959:01-1998:12 | 5 | manufacturing & trade inv:retail trade (mil of chained 1992 dollars)(sa) |
| 97. | ivsrq | 1959:01-1998:12 | 2 | ratio for mfg & trade: inventory/sales (chained 1992 dollars, sa) |
| 98. | ivsrmq | 1959:01-1998:12 | 2 | ratio for mfg & trade:mfg;inventory/sales (87$)(s.a.) |
| 99. | ivsrwq | 1959:01-1998:12 | 2 | ratio for mfg & trade:wholesaler;inventory/sales(87$)(s.a.) |
| 100. | ivsrrq | 1959:01-1998:12 | 2 | ratio for mfg & trade:retail trade;inventory/sales(87$)(s.a.) |
| 101. | pmnv | 1959:01-1998:12 | 1 | napm inventories index (percent) |

Orders and unfilled orders (Ord)

| | | | | |
|---|---|---|---|---|
| 102. | pmno | 1959:01-1998:12 | 1 | napm new orders index (percent) |
| 103. | pmdel | 1959:01-1998:12 | 1 | napm vendor deliveries index (percent) |
| 104. | mocmq | 1959:01-1998:12 | 5 | new orders (net)-consumer goods & materials, 1992 dollars (bci) |
| 105. | mdoq | 1959:01-1998:12 | 5 | new orders, durable goods industries, 1992 dollars (bci) |
| 106. | msondq | 1959:01-1998:12 | 5 | new orders, nondefense capital goods, in 1992 dollars (bci) |
| 107. | mo | 1959:01-1998:12 | 5 | mfg new orders: all manufacturing industries, total (mil$,sa) |
| 108. | mowu | 1959:01-1998:12 | 5 | mfg new orders: mfg industries with unfilled orders(mil$,sa) |
| 109. | mdo | 1959:01-1998:12 | 5 | mfg new orders: durable goods industries, total (mil$,sa) |
| 110. | mduwu | 1959:01-1998:12 | 5 | mfg new orders:durable goods indust with unfilled orders(mil$,sa) |
| 111. | mno | 1959:01-1998:12 | 5 | mfg new orders: nondurable goods industries, total (mil$,sa) |
| 112. | mnou | 1959:01-1998:12 | 5 | mfg new orders: nondurable gds ind.with unfilled orders(mil$,sa) |
| 113. | mu | 1959:01-1998:12 | 5 | mfg unfilled orders: all manufacturing industries, total (mil$,sa) |
| 114. | mdu | 1959:01-1998:12 | 5 | mfg unfilled orders: durable goods industries, total (mil$,sa) |
| 115. | mnu | 1959:01-1998:12 | 5 | mfg unfilled orders: nondurable goods industries, total (mil$,sa) |
| 116. | mpcon | 1959:01-1998:12 | 5 | contracts & orders for plant & equipment (bil$,sa) |
| 117. | mpconq | 1959:01-1998:12 | 5 | contracts & orders for plant & equipment in 1992 dollars (bci) |

Stock prices (SPr)

| | | | | |
|---|---|---|---|---|
| 118. | fsncom | 1959:01-1998:12 | 5 | NYSE common stock price index: composite (12/31/65=50) |
| 119. | fsnin | 1966:01-1998:12* | 5 | NYSE common stock price index: industrial (12/31/65=50) |
| 120. | fsntr | 1966:01-1998:12* | 5 | NYSE common stock price index: transportation (12/31/65=50) |
| 121. | fsnut | 1966:01-1998:12* | 5 | NYSE common stock price index: utility (12/31/65=50) |
| 122. | fsnfi | 1966:01-1998:12* | 5 | NYSE common stock price index: finance (12/31/65=50) |
| 123. | fspcom | 1959:01-1998:12 | 5 | S&P's common stock price index: composite (1941-43=10) |
| 124. | fspin | 1959:01-1998:12 | 5 | S&P's common stock price index: industrials (1941-43=10) |
| 125. | fspcap | 1959:01-1998:12 | 5 | S&P's common stock price index: capital goods (1941-43=10) |

| | | | | |
|---|---|---|---|---|
| 126. | fsptr | 1970:01-1998:12* | 5 | S&P's common stock price index: transportation (1970=10) |
| 127. | fsput | 1959:01-1998:12 | 5 | S&P's common stock price index: utilities (1941-43=10) |
| 128. | fspfi | 1970:01-1998:12* | 5 | S&P's common stock price index: financial (1970=10) |
| 129. | fsdxp | 1959:01-1998:12 | 1 | S&P's composite common stock: dividend yield (% per annum) |
| 130. | fspxe | 1959:01-1998:12 | 1 | S&P's composite common stock: price-earnings ratio (%,nsa) |
| 131. | fsnvv3 | 1974:01-1997:07* | 5 | NYSE mkt composition:reptd share vol by size,5000+ shrs,% |

Exchange rates (EXR)

| | | | | |
|---|---|---|---|---|
| 132. | exrus | 1959:01-1998:12 | 5 | United States effective exchange rate (merm)(index no.) |
| 133. | exrger | 1959:01-1998:12 | 5 | foreign exchange rate: Germany (deutsche mark per U.S.$) |
| 134. | exrsw | 1959:01-1998:12 | 5 | foreign exchange rate: Switzerland (swiss franc per U.S.$) |
| 135. | exrjan | 1959:01-1998:12 | 5 | foreign exchange rate: Japan (yen per U.S.$) |
| 136. | exruk | 1959:01-1998:12* | 5 | foreign exchange rate: United Kingdom (cents per pound) |
| 137. | exrcan | 1959:01-1998:12 | 5 | foreign exchange rate: Canada (canadian $ per U.S.$) |

Interest rates (Int)

| | | | | |
|---|---|---|---|---|
| 138. | fyff | 1959:01-1998:12* | 2 | interest rate: federal funds (effective) (% per annum,nsa) |
| 139. | fycp90 | 1959:01-1998:12* | 2 | interest rate: 90 day commercial paper, (ac) (% per ann,nsa) |
| 140. | fygm3 | 1959:01-1998:12* | 2 | interest rate: U.S.treasury bills,sec mkt,3-mo.(% per ann,nsa) |
| 141. | fygm6 | 1959:01-1998:12* | 2 | interest rate: U.S.treasury bills,sec mkt,6-mo.(% per ann,nsa) |
| 142. | fygtl | 1959:01-1998:12* | 2 | interest rate: U.S.treasury const maturities,1-yr.(% per ann,nsa) |
| 143. | fygt5 | 1959:01-1998:12 | 2 | interest rate: U.S.treasury const maturities,5-yr.(% per ann,nsa) |
| 144. | fygt10 | 1959:01-1998:12 | 2 | interest rate: U.S.treasury const maturities,10-yr.(% per ann,nsa) |
| 145. | fyaaac | 1959:01-1998:12 | 2 | bond yield: moody's aaa corporate (% per annum) |
| 146. | fybaac | 1959:01-1998:12 | 2 | bond yield: moody's baa corporate (% per annum) |
| 147. | fwafit | 1973:01-1994:04* | 1 | weighted avg foreign interest rate(%,sa) |
| 148. | fyfha | 1959:01-1998:12 | 2 | secondary market yields on fha mortgages (% per annum) |
| 149. | sfycp | 1959:01-1998:12 | 1 | spread fycp - fyff |
| 150. | sfygm3 | 1959:01-1998:12 | 1 | spread fygm3 - fyff |
| 151. | sfygm6 | 1959:01-1998:12 | 1 | spread fygm6 - fyff |
| 152. | sfygtl | 1959:01-1998:12 | 1 | spread fygtl - fyff |
| 153. | sfygt5 | 1959:01-1998:12 | 1 | spread fygt5 - fyff |
| 154. | sfygt10 | 1959:01-1998:12 | 1 | spread fygt10 - fyff |
| 155. | sfyaaac | 1959:01-1998:12 | 1 | spread fyaaac - fyff |
| 156. | sfybaac | 1959:01-1998:12 | 1 | spread fybaac - fyff |
| 157. | sfyfha | 1959:01-1998:12 | 1 | spread fyfha - fyff |

Money and credit quantity aggregates (Mon)

| | | | | |
|---|---|---|---|---|
| 158. | fm1 | 1959:01-1998:12 | 6 | money stock: m1(curr,trav.cks,dem dep,other ck'able dep)(bil$,sa) |
| 159. | fm2 | 1959:01-1998:12 | 6 | money stock:m2(m1+o'nite rps,euro$,g/p&b/d mmmfs&sav&sm time dep(bil$, |
| 160. | fm3 | 1959:01-1998:12 | 6 | money stock: m3(m2+lg time dep,term rp's&inst only mmmfs)(bil$,sa) |
| 161. | fml | 1959:01-1998:09* | 6 | money stock:l(m3 + other liquid assets) (bil$,sa) |

| | | | | |
|---|---|---|---|---|
| 162. | fm2dq | 1959:01-1998:12 | 5 | money supply-m2 in 1992 dollars (bci) |
| 163. | fmfba | 1959:01-1998:12 | 6 | monetary base, adj for reserve requirement changes(mil$,sa) |
| 164. | fmrra | 1959:01-1998:12 | 6 | depository inst reserves:total,adj for reserve req chgs(mil$,sa) |
| 165. | fmrnbc | 1959:01-1998:12 | 6 | depository inst reserves:nonborrow+ext cr,adj res req cgs(mil$,sa) |
| 166. | fcls | 1973:01-1998:12* | 5 | loans & sec @ all coml banks: total (bils,sa) |
| 167. | fcsgv | 1973:01-1998:12* | 5 | loans & sec @ all coml banks: U.S.govt securities (bil$,sa) |
| 168. | fclre | 1973:01-1998:12* | 5 | loans & sec @ all coml banks: real estate loans (bil$,sa) |
| 169. | fclin | 1973:01-1998:12* | 5 | loans & sec @ all coml banks: loans to individuals (bil$,sa) |
| 170. | fclnbf | 1973:01-1994:01* | 5 | loans & sec @ all coml banks: loans to nonbank fin inst(bil$,sa) |
| 171. | fclnq | 1959:01-1998:12* | 5 | commercial & industrial loans oustanding in 1992 dollars (bci) |
| 172. | fclbmc | 1959:01-1998:12* | 1 | wkly rp lg com'l banks:net change com'l & indus loans(bil$,saar) |
| 173. | cci30m | 1959:01-1995:09* | 1 | consumer instal.loans: delinquency rate,30 days & over, (%,sa) |
| 174. | ccint | 1975:01-1995:09* | 1 | net change in consumer instal cr: total (mil$,sa) |
| 175. | ccinv | 1975:01-1995:09* | 1 | net change in consumer instal cr: automobile (mil$,sa) |
| 176. | ccinrv | 1980:01-1995:09* | 1 | net change in consumer instal cr: revolving(mil$,sa) |

Price indexes (Pri)

| | | | | |
|---|---|---|---|---|
| 177. | pmcp | 1959:01-1998:12 | 1 | napm commodity prices index (percent) |
| 178. | pwfsa | 1959:01-1998:12 | 6 | producer price index: finished goods (82=100,sa) |
| 179. | pwfcsa | 1959:01-1998:12 | 6 | producer price index:finished consumer goods (82=100,sa) |
| 180. | pwimsa | 1959:01-1998:12* | 6 | producer price index:intermed mat.supplies & components(82=100,sa) |
| 181. | pwcmsa | 1959:01-1998:12* | 6 | producer price index:crude materials (82=100,sa) |
| 182. | pwfxsa | 1967:01-1998:12* | 6 | producer price index: finished goods,excl. foods (82=100,sa) |
| 183. | pw160a | 1974:01-1998:12* | 6 | producer price index: crude materials less energy (82=100,sa) |
| 184. | pw150a | 1974:01-1998:12* | 6 | producer price index: crude nonfood mat less energy (82=100,sa) |
| 185. | psm99q | 1959:01-1998:12 | 6 | index of sensitive materials prices (1990=100)(bci-99a) |
| 186. | punew | 1959:01-1998:12 | 6 | cpi-u: all items (82-84=100,sa) |
| 187. | pu81 | 1967:01-1998:12* | 6 | cpi-u: food & beverages (82-84=100,sa) |
| 188. | puh | 1967:01-1998:12* | 6 | cpi-u: housing (82-84=100,sa) |
| 189. | pu83 | 1959:01-1998:12 | 6 | cpi-u: apparel & upkeep (82-84=100,sa) |
| 190. | pu84 | 1959:01-1998:12 | 6 | cpi-u: transportation (82-84=100,sa) |
| 191. | pu85 | 1959:01-1998:12 | 6 | cpi-u: medical care (82-84=100,sa) |
| 192. | puc | 1959:01-1998:12 | 6 | cpi-u: commodities (82-84=100,sa) |
| 193. | pucd | 1959:01-1998:12 | 6 | cpi-u: durables (82-84=100,sa) |
| 194. | pus | 1959:01-1998:12 | 6 | cpi-u: services (82-84=100,sa) |
| 195. | puxf | 1959:01-1998:12 | 6 | cpi-u: all items less food (82-84=100,sa) |
| 196. | puxhs | 1959:01-1998:12 | 6 | cpi-u: all items less shelter (82-84=100,sa) |
| 197. | puxm | 1959:01-1998:12 | 6 | cpi-u: all items less midical care (82-84=100,sa) |
| 198. | pcgold | 1975:01-1998:12* | 6 | commodities price:gold,london noon fix,avg of daily rate,$ per oz |
| 199. | gmdc | 1959:01-1998:12 | 6 | pce,impl pr defl:pce (1987=100) |
| 200. | gmdcd | 1959:01-1998:12 | 6 | pce,impl pr defl:pce; durables (1987=100) |
| 201. | gmdcn | 1959:01-1998:12 | 6 | pce,impl pr defl:pce; nondurables (1987=100) |

| | | | | |
|---|---|---|---|---|
| 202. | gmdcs | 1959:01-1998:12 | 6 | pce,impl pr defl:pce; services (1987=100) |

Average hourly earnings (AHE)

| | | | | |
|---|---|---|---|---|
| 203. | leh | 1964:01-1998:12* | 6 | avg hr earnings of prod wkrs: total private nonagric ($,sa) |
| 204. | lehcc | 1959:01-1998:12 | 6 | avg hr earnings of constr wkrs: construction ($,sa) |
| 205. | lehm | 1959:01-1998:12 | 6 | avg hr earnings of prod wkrs: manufacturing ($,sa) |
| 206. | lehtu | 1964:01-1998:12* | 6 | avg hr earnings of nonsupv wkrs: trans & public util($,sa) |
| 207. | lehtt | 1964:01-1998:12* | 6 | avg hr earnings of prod wkrs:wholesale & retail trade(sa) |
| 208. | lehfr | 1964:01-1998:12* | 6 | avg hr earnings of nonsupv wkrs: finance,insur,real est($,sa) |
| 209. | lehs | 1964:01-1998:12* | 6 | avg hr earnings of nonsupv wkrs: services ($,sa) |

Miscellaneous (Oth)

| | | | | |
|---|---|---|---|---|
| 210. | fste | 1986:01-1998:12* | 5 | U.S.mdse exports: total exports(f.a.s. value)(mil.$,s.a.) |
| 211. | fstm | 1986:01-1998:12* | 5 | U.S.mdse imports: general imports(c.i.f. value)(mil.$,s.a.) |
| 212. | ftmd | 1986:01-1998:12* | 5 | U.S.mdse imports: general imports (customs value)(mil$,s.a.) |
| 213. | fstb | 1986:01-1998:12* | 2 | U.S.mdse trade balance:exports less imports(fas/cif)(mil$,s.a.) |
| 214. | ftb | 1986:01-1998:12* | 2 | U.S.mdse trade balance:exp.(fas) less imp.(custom)(mil$,s.a.) |
| 215. | hhsntn | 1959:01-1998:12 | 1 | u. of mich. index of consumer expectations(bcd-83) |

## References

Burns, Arthur F. and Mitchell, Wesley C., *Measuring Business Cycles*, New York: National Bureau of Economic Research, 1947.

Chamberlain, G. and M. Rothschild (1983), "Arbitrage Factor Stucture, and Mean-Variance Analysis of Large Asset Markets," *Econometrica*, Vol. 51, No. 5.

Congressional Budget Office (1996), *The Economic and Budget Outlook: Fiscal Years 1997-2006*, May.

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

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

Connor, G. and R.A. Korajczyk (1993), "A Test for the Number of Factors in an Approximate Factor Model," *Journal of Finance*, Vol. XLVIII, No. 4.

Engle, R.F., and M.W. Watson (1981), "A One-Factor Multivariate Time Series Model of Metropolitan Wage Rates," *Journal of the American Statistical Association 76*, no. 376, 774-781.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000), "The Generalized Dynamnic Factor Model: Identification and Estimation," *The Review of Economics and Statistics*, Vol. 82, No. 4, November, 540-552.

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

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

Fuhrer, Jeffrey C. (1995), "The Phillips Curve is Alive and Well," *New England Economic Review of the Federal Reserve Bank of Boston*, March/April 1995, 41-56.

Geweke, J. (1977), "The Dynamic Factor Analysis of Economic Time Series," in D.J. Aigner and A.S. Goldberger (eds.), *Latent Variables in Socio-Economic Models*, North-Holland, Amsterdam, Ch. 19.

Gordon, Robert J. (1982), "Price Inertia and Ineffectiveness in the United States," *Journal of Political Economy 90*, 1087-1117.

Gordon, Robert J. (1997), "The Time-Varying NAIRU and its Implications for Economic Policy," *Journal of Economic Perspectives*, 11-32.

Quah, D. and T.J. Sargent (1993), "A Dynamic Index Model for Large Cross Sections," ch. 7 in J.H. Stock and M.W. Watson (eds.), *Business Cycles, Indicators, and Forecasting*, University of Chicago Press for the NBER, 1993, 285-306.

Sargent, T.J, and C.A. Sims (1977), "Business Cycle Modeling without Pretending to have Too

Much a-priori Economic Theory," in C. Sims et al. (eds.), *New Methods in Business Cycle Research*, Minneapolis: Federal Reserve Bank of Minneapolis.

Singleton, Kenneth .J. (1980), "A Latent Time Series Model of the Cyclical Behavior of Interest Rates," *International Economic Review*, 21, 559-575.

Staiger, D., J.H. Stock and M.W. Watson (1997), "The NAIRU, Unemployment, and Monetary Policy," *Journal of Economic Perspectives* 11 (Winter), 33-51.

Stock, J.H. and M.W. Watson (1989), "New Indexes of Coincident and Leading Economic Indicators," *NBER Macroeconomics Annual*, 351-393.

Stock, J.H. and M.W. Watson (1991), "A Probability Model of the Coincident Economic Indicators," in K. Lahiri and G.H. Moore (eds.), *Leading Economic Indicatorss: New Approaches and Forecasting Records*, ch. 4. New York: Cambridge University Press, 63-85.

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

Stock, J.H. and M.W. Watson (1998), "Diffusion Indexes," NBER Working Paper No. #6702.

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

Stock, J.H. and M.W. Watson (2000), "Forecasting Using Principal Components from a Large

Number of Predictors," manuscript.

Tootell, Geoffrey M.B. (1994), "Restructuring, the NAIRU, and the Phillips Curve," *New England Economic Review of the Federal Reserve Bank of Boston*, September/October, 31-44.

West, K.D. (1996), "Asymptotic Inference about Predictive Ability," *Econometrica*, 64, 1067-1084.

**Table 1**
**Simulated out-of-sample forecasting results: Real variables, 12 month horizon**

| Forecast Method | industrial production Rel. MSE | $\hat{\alpha}$ | personal income Rel. MSE | $\hat{\alpha}$ | mfg & trade sales Rel. MSE | $\hat{\alpha}$ | nonag. employment Rel. MSE | $\hat{\alpha}$ |
|---|---|---|---|---|---|---|---|---|
| *Benchmark models* | | | | | | | | |
| AR | 1.00 | | 1.00 | | 1.00 | | 1.00 | |
| LI | 0.86 (0.27) | 0.57 (0.13) | 0.97 (0.21) | 0.52 (0.15) | 0.82 (0.25) | 0.63 (0.17) | 0.89 (0.23) | 0.56 (0.14) |
| VAR | 0.97 (0.07) | 0.75 (0.68) | 0.98 (0.05) | 0.68 (0.34) | 0.98 (0.04) | 0.73 (0.58) | 1.05 (0.09) | 0.22 (0.41) |
| *Full data set (N=215)* | | | | | | | | |
| DI-AR, Lag | 0.57 (0.27) | 0.76 (0.13) | 0.77 (0.14) | 0.76 (0.13) | 0.48 (0.25) | 0.99 (0.15) | 0.91 (0.13) | 0.63 (0.18) |
| DI-AR | 0.63 (0.25) | 0.71 (0.12) | 0.86 (0.16) | 0.61 (0.12) | 0.57 (0.24) | 0.84 (0.18) | 0.99 (0.31) | 0.51 (0.20) |
| DI | 0.52 (0.26) | 0.88 (0.17) | 0.86 (0.16) | 0.61 (0.12) | 0.56 (0.23) | 0.94 (0.20) | 0.92 (0.26) | 0.55 (0.20) |
| *Balanced panel (N=149)* | | | | | | | | |
| DI-AR, Lag | 0.67 (0.25) | 0.70 (0.13) | 0.82 (0.15) | 0.70 (0.13) | 0.56 (0.23) | 0.91 (0.16) | 0.88 (0.14) | 0.68 (0.18) |
| DI-AR | 0.67 (0.25) | 0.70 (0.12) | 0.92 (0.14) | 0.57 (0.12) | 0.61 (0.23) | 0.80 (0.17) | 0.88 (0.22) | 0.58 (0.17) |
| DI | 0.59 (0.25) | 0.81 (0.17) | 0.92 (0.14) | 0.57 (0.12) | 0.57 (0.23) | 0.91 (0.18) | 0.84 (0.21) | 0.62 (0.16) |
| *Stacked balance panel* | | | | | | | | |
| DI-AR | 0.65 (0.25) | 0.70 (0.12) | 0.93 (0.15) | 0.56 (0.12) | 0.61 (0.22) | 0.89 (0.19) | 1.02 (0.30) | 0.49 (0.14) |
| DI | 0.62 (0.25) | 0.81 (0.18) | 0.93 (0.15) | 0.56 (0.12) | 0.66 (0.21) | 0.85 (0.20) | 0.95 (0.24) | 0.53 (0.14) |
| *Full data set; m=1, p=BIC, k fixed* | | | | | | | | |
| DI-AR, k=1 | 1.06 (0.11) | 0.27 (0.34) | 1.03 (0.08) | 0.34 (0.41) | 0.98 (0.06) | 0.63 (0.46) | 1.01 (0.09) | 0.49 (0.24) |
| DI-AR, k=2 | 0.63 (0.25) | 0.76 (0.14) | 0.78 (0.14) | 0.77 (0.14) | 0.53 (0.24) | 0.93 (0.15) | 0.77 (0.13) | 0.82 (0.15) |
| DI-AR, k=3 | 0.56 (0.26) | 0.84 (0.14) | 0.77 (0.15) | 0.77 (0.13) | 0.52 (0.23) | 0.99 (0.16) | 0.84 (0.14) | 0.75 (0.20) |
| DI-AR, k=4 | 0.54 (0.26) | 0.85 (0.14) | 0.76 (0.15) | 0.78 (0.14) | 0.51 (0.23) | 1.01 (0.16) | 0.83 (0.15) | 0.73 (0.19) |
| *Full data set; m=1, p=0, k fixed* | | | | | | | | |
| DI, k=1 | 1.03 (0.07) | 0.30 (0.49) | 1.01 (0.09) | 0.46 (0.34) | 0.98 (0.05) | 0.67 (0.49) | 1.01 (0.09) | 0.48 (0.24) |
| DI, k=2 | 0.55 (0.25) | 0.89 (0.15) | 0.78 (0.14) | 0.76 (0.13) | 0.57 (0.24) | 0.95 (0.17) | 0.78 (0.13) | 0.83 (0.16) |
| DI, k=3 | 0.51 (0.25) | 1.00 (0.16) | 0.77 (0.15) | 0.77 (0.13) | 0.60 (0.21) | 1.02 (0.19) | 0.84 (0.14) | 0.76 (0.19) |
| DI, k=4 | 0.49 (0.25) | 1.00 (0.16) | 0.76 (0.15) | 0.78 (0.14) | 0.59 (0.22) | 1.03 (0.20) | 0.82 (0.15) | 0.75 (0.18) |
| RMSE, AR Model | 0.049 | | 0.027 | | 0.045 | | 0.017 | |

Notes to table 1: For each variable/forecast method combination, the first entry is the ratio of the MSE of the forecast made by the method for that row, to the MSE of a univariate autoregressive forecast with lag length selected by BIC ("AR" in the table), and the HAC standard error of that ratio appears next in parentheses. The second pair of entries are the estimated forecast combining coefficient $\hat{\alpha}$ from regression (4.1) and its HAC standard error. All forecasts are simulated out of sample. The LI (leading indicator), Phillips curve (for inflation series), DI, DI-AR, and DI-AR,Lag forecasts were computed with BIC lag and/or variable selection, see the text for details. The method for computing the factors (full data set, balanced panel, stacked balanced panel) are indicated in italics above the associated panel of results. The final line presents the root MSE for the AR model in native (decimal growth rate) units at an annual rate.

**Table 2**

**Simulated out-of-sample forecasting results: Real variables, 6 and 24 month horizons**

| Forecast Method | industrial production Rel. MSE | industrial production $\hat{\alpha}$ | personal income Rel. MSE | personal income $\hat{\alpha}$ | mfg & trade sales Rel. MSE | mfg & trade sales $\hat{\alpha}$ | nonag. employment Rel. MSE | nonag. employment $\hat{\alpha}$ |
|---|---|---|---|---|---|---|---|---|
| **A. Horizon = 6 months** | | | | | | | | |
| *Benchmark models* | | | | | | | | |
| AR | 1.00 | | 1.00 | | 1.00 | | 1.00 | |
| LI | 0.70 (0.25) | 0.68 (0.13) | 0.83 (0.15) | 0.64 (0.11) | 0.77 (0.19) | 0.68 (0.14) | 0.75 (0.19) | 0.67 (0.12) |
| VAR | 1.01 (0.05) | 0.43 (0.39) | 0.99 (0.03) | 0.63 (0.43) | 0.99 (0.04) | 0.64 (0.45) | 1.06 (0.07) | 0.12 (0.34) |
| *Full data set (N=215)* | | | | | | | | |
| DI-AR, Lag | 0.69 (0.25) | 0.69 (0.14) | 0.77 (0.12) | 0.86 (0.15) | 0.63 (0.18) | 0.89 (0.17) | 0.94 (0.16) | 0.56 (0.18) |
| DI-AR | 0.77 (0.30) | 0.62 (0.16) | 0.81 (0.16) | 0.66 (0.13) | 0.70 (0.20) | 0.76 (0.17) | 1.02 (0.32) | 0.49 (0.19) |
| DI | 0.74 (0.25) | 0.68 (0.17) | 0.81 (0.16) | 0.65 (0.13) | 0.67 (0.20) | 0.79 (0.18) | 0.96 (0.28) | 0.52 (0.19) |
| *Balanced panel (N=149)* | | | | | | | | |
| DI-AR, Lag | 0.73 (0.25) | 0.68 (0.16) | 0.79 (0.13) | 0.78 (0.13) | 0.66 (0.17) | 0.87 (0.17) | 0.93 (0.17) | 0.58 (0.21) |
| DI-AR | 0.78 (0.28) | 0.62 (0.16) | 0.81 (0.15) | 0.66 (0.11) | 0.76 (0.19) | 0.70 (0.17) | 0.97 (0.28) | 0.52 (0.19) |
| DI | 0.73 (0.24) | 0.69 (0.15) | 0.81 (0.15) | 0.66 (0.11) | 0.68 (0.19) | 0.81 (0.17) | 0.95 (0.26) | 0.53 (0.18) |
| *Full data set; m=1, p=BIC, k fixed* | | | | | | | | |
| DI-AR, k=1 | 0.97 (0.15) | 0.58 (0.33) | 0.91 (0.07) | 0.80 (0.23) | 0.99 (0.11) | 0.52 (0.29) | 0.94 (0.12) | 0.60 (0.19) |
| DI-AR, k=2 | 0.67 (0.22) | 0.77 (0.15) | 0.76 (0.11) | 0.90 (0.14) | 0.64 (0.18) | 0.86 (0.16) | 0.84 (0.13) | 0.71 (0.16) |
| DI-AR, k=3 | 0.64 (0.23) | 0.81 (0.15) | 0.75 (0.12) | 0.89 (0.14) | 0.64 (0.18) | 0.88 (0.17) | 0.88 (0.14) | 0.66 (0.17) |
| DI-AR, k=4 | 0.64 (0.23) | 0.80 (0.15) | 0.74 (0.13) | 0.87 (0.14) | 0.63 (0.18) | 0.87 (0.15) | 0.91 (0.16) | 0.60 (0.18) |
| RMSE, AR Model | 0.030 | | 0.016 | | 0.028 | | 0.008 | |

**Table 2, continued**

| Forecast Method | industrial production Rel. MSE | industrial production $\hat{\alpha}$ | personal income Rel. MSE | personal income $\hat{\alpha}$ | mfg & trade sales Rel. MSE | mfg & trade sales $\hat{\alpha}$ | nonag. employment Rel. MSE | nonag. employment $\hat{\alpha}$ |
|---|---|---|---|---|---|---|---|---|
| **B. Horizon = 24 months** | | | | | | | | |
| *Benchmark models* | | | | | | | | |
| AR | 1.00 | | 1.00 | | 1.00 | | 1.00 | |
| LI | 1.09 (0.28) | 0.45 (0.14) | 1.29 (0.31) | 0.30 (0.20) | 1.08 (0.21) | 0.45 (0.14) | 1.07 (0.31) | 0.47 (0.15) |
| VAR | 1.01 (0.10) | 0.44 (0.48) | 0.98 (0.06) | 0.63 (0.34) | 1.03 (0.06) | 0.13 (0.85) | 1.06 (0.13) | 0.35 (0.31) |
| *Full data set (N=215)* | | | | | | | | |
| DI-AR, Lag | 0.57 (0.24) | 0.88 (0.13) | 0.70 (0.20) | 0.94 (0.23) | 0.66 (0.18) | 0.95 (0.18) | 0.82 (0.15) | 0.88 (0.26) |
| DI-AR | 0.59 (0.25) | 0.88 (0.15) | 0.76 (0.22) | 0.80 (0.26) | 0.70 (0.20) | 0.89 (0.19) | 0.74 (0.19) | 0.97 (0.24) |
| DI | 0.55 (0.26) | 0.91 (0.14) | 0.76 (0.22) | 0.80 (0.25) | 0.70 (0.20) | 0.89 (0.19) | 0.74 (0.19) | 0.97 (0.24) |
| *Balanced panel (N=149)* | | | | | | | | |
| DI-AR, Lag | 0.57 (0.25) | 0.87 (0.14) | 0.76 (0.19) | 0.86 (0.23) | 0.64 (0.20) | 0.94 (0.18) | 0.74 (0.17) | 1.06 (0.25) |
| DI-AR | 0.58 (0.25) | 0.87 (0.14) | 0.83 (0.20) | 0.74 (0.24) | 0.67 (0.19) | 0.93 (0.18) | 0.76 (0.18) | 0.94 (0.25) |
| DI | 0.58 (0.25) | 0.87 (0.14) | 0.83 (0.20) | 0.74 (0.24) | 0.67 (0.20) | 0.94 (0.19) | 0.75 (0.18) | 0.94 (0.24) |
| *Full data set; m=1, p=BIC, k fixed* | | | | | | | | |
| DI-AR, k=1 | 1.12 (0.19) | 0.10 (0.46) | 1.07 (0.09) | -0.81 (1.00) | 0.97 (0.04) | 0.90 (0.62) | 1.03 (0.07) | 0.33 (0.46) |
| DI-AR, k=2 | 0.76 (0.19) | 0.68 (0.11) | 0.88 (0.13) | 0.68 (0.17) | 0.65 (0.20) | 0.87 (0.14) | 0.72 (0.16) | 0.99 (0.17) |
| DI-AR, k=3 | 0.58 (0.24) | 0.89 (0.13) | 0.72 (0.19) | 0.90 (0.18) | 0.70 (0.17) | 0.89 (0.14) | 0.79 (0.16) | 0.95 (0.24) |
| DI-AR, k=4 | 0.56 (0.24) | 0.90 (0.14) | 0.70 (0.20) | 0.93 (0.23) | 0.67 (0.18) | 0.95 (0.18) | 0.78 (0.16) | 0.96 (0.24) |
| RMSE, AR Model | 0.075 | | 0.046 | | 0.070 | | 0.031 | |

Notes: See the notes to table 1.

**Table 3**

**Simulated out-of-sample forecasting results: Price inflation, 12 month horizon**

| Forecast Method | CPI Rel. MSE | CPI $\hat{\alpha}$ | consumption deflator Rel. MSE | consumption deflator $\hat{\alpha}$ | CPI exc. food&energy Rel. MSE | CPI exc. food&energy $\hat{\alpha}$ | producer price index Rel. MSE | producer price index $\hat{\alpha}$ |
|---|---|---|---|---|---|---|---|---|
| *Benchmark models* | | | | | | | | |
| AR | 1.00 | | 1.00 | | 1.00 | | 1.00 | |
| LI | 0.79 (0.15) | 0.76 (0.15) | 0.95 (0.12) | 0.58 (0.17) | 1.00 (0.16) | 0.50 (0.21) | 0.82 (0.15) | 0.75 (0.19) |
| Phillips Curve | 0.82 (0.13) | 0.95 (0.20) | 0.92 (0.10) | 0.72 (0.23) | 0.79 (0.18) | 0.80 (0.22) | 0.87 (0.14) | 0.96 (0.30) |
| VAR | 0.91 (0.09) | 0.74 (0.20) | 1.02 (0.06) | 0.45 (0.20) | 0.99 (0.05) | 0.56 (0.21) | 1.29 (0.14) | 0.25 (0.12) |
| *Full data set (N=215)* | | | | | | | | |
| DI-AR, Lag | 0.72 (0.14) | 0.91 (0.14) | 0.90 (0.09) | 0.65 (0.13) | 0.84 (0.15) | 0.76 (0.20) | 0.83 (0.13) | 0.78 (0.21) |
| DI-AR | 0.71 (0.16) | 0.83 (0.13) | 0.90 (0.10) | 0.62 (0.13) | 0.85 (0.15) | 0.74 (0.20) | 0.82 (0.14) | 0.75 (0.20) |
| DI | 1.30 (0.16) | 0.34 (0.08) | 1.40 (0.16) | 0.25 (0.08) | 1.55 (0.31) | 0.24 (0.06) | 2.40 (0.88) | 0.13 (0.07) |
| *Balanced panel (N=149)* | | | | | | | | |
| DI-AR, Lag | 0.70 (0.14) | 0.94 (0.12) | 0.90 (0.08) | 0.67 (0.15) | 0.84 (0.15) | 0.77 (0.21) | 0.86 (0.11) | 0.77 (0.21) |
| DI-AR | 0.69 (0.15) | 0.88 (0.13) | 0.87 (0.10) | 0.66 (0.12) | 0.85 (0.15) | 0.73 (0.20) | 0.85 (0.14) | 0.71 (0.19) |
| DI | 1.30 (0.16) | 0.32 (0.08) | 1.34 (0.13) | 0.26 (0.09) | 1.57 (0.33) | 0.20 (0.07) | 2.44 (0.87) | 0.14 (0.06) |
| *Stacked balance panel* | | | | | | | | |
| DI-AR | 0.73 (0.15) | 0.82 (0.12) | 0.87 (0.09) | 0.65 (0.12) | 0.85 (0.15) | 0.77 (0.21) | 0.81 (0.14) | 0.75 (0.20) |
| DI | 1.54 (0.31) | 0.28 (0.08) | 1.51 (0.18) | 0.25 (0.08) | 1.55 (0.32) | 0.23 (0.06) | 3.06 (1.89) | 0.11 (0.06) |
| *Full data set; m=1, p=BIC, k fixed* | | | | | | | | |
| DI-AR, k=1 | 0.64 (0.15) | 1.14 (0.14) | 0.77 (0.12) | 0.96 (0.16) | 0.71 (0.17) | 1.25 (0.23) | 0.76 (0.16) | 0.95 (0.24) |
| DI-AR, k=2 | 0.67 (0.14) | 1.07 (0.13) | 0.83 (0.09) | 0.83 (0.14) | 0.72 (0.17) | 0.97 (0.19) | 0.77 (0.15) | 0.93 (0.23) |
| DI-AR, k=3 | 0.76 (0.13) | 0.91 (0.15) | 0.94 (0.07) | 0.61 (0.14) | 0.86 (0.14) | 0.73 (0.20) | 0.86 (0.11) | 0.78 (0.21) |
| DI-AR, k=4 | 0.74 (0.14) | 0.89 (0.15) | 0.91 (0.09) | 0.64 (0.14) | 0.87 (0.15) | 0.72 (0.21) | 0.82 (0.13) | 0.79 (0.21) |
| *Full data set; m=1, p=0, k fixed* | | | | | | | | |
| DI, k=1 | 1.60 (0.34) | 0.25 (0.07) | 1.56 (0.20) | 0.22 (0.09) | 1.55 (0.31) | 0.23 (0.06) | 2.76 (1.61) | 0.12 (0.07) |
| DI, k=2 | 1.56 (0.31) | 0.26 (0.07) | 1.58 (0.20) | 0.21 (0.08) | 1.62 (0.39) | 0.22 (0.07) | 2.72 (1.56) | 0.13 (0.07) |
| DI, k=3 | 1.57 (0.32) | 0.24 (0.08) | 1.60 (0.20) | 0.17 (0.08) | 1.69 (0.43) | 0.18 (0.07) | 2.68 (1.49) | 0.13 (0.07) |
| DI, k=4 | 1.56 (0.25) | 0.25 (0.07) | 1.56 (0.19) | 0.21 (0.08) | 1.67 (0.40) | 0.19 (0.07) | 2.55 (0.99) | 0.16 (0.06) |
| RMSE, AR Model | 0.021 | | 0.015 | | 0.019 | | 0.033 | |

Notes: see the notes to table 1.

Table 4

**Simulated out-of-sample forecasting results: Price inflation, 6 and 24 month horizons**

| Forecast Method | CPI Rel. MSE | CPI $\hat{\alpha}$ | consumption deflator Rel. MSE | consumption deflator $\hat{\alpha}$ | CPI exc. food&energy Rel. MSE | CPI exc. food&energy $\hat{\alpha}$ | producer price index Rel. MSE | producer price index $\hat{\alpha}$ |
|---|---|---|---|---|---|---|---|---|
| **A. Horizon = 6 months** | | | | | | | | |
| *Benchmark models* | | | | | | | | |
| AR | 1.00 | | 1.00 | | 1.00 | | 1.00 | |
| LI | 0.82 (0.12) | 0.78 (0.16) | 1.04 (0.09) | 0.42 (0.16) | 1.10 (0.16) | 0.32 (0.27) | 1.00 (0.09) | 0.51 (0.19) |
| Phillips Curve | 0.90 (0.11) | 0.80 (0.27) | 0.99 (0.06) | 0.54 (0.23) | 0.90 (0.11) | 0.68 (0.19) | 1.02 (0.04) | 0.34 (0.37) |
| VAR | 1.04 (0.08) | 0.41 (0.16) | 1.15 (0.07) | 0.08 (0.20) | 1.00 (0.05) | 0.50 (0.21) | 1.34 (0.16) | 0.19 (0.12) |
| *Full data set (N=215)* | | | | | | | | |
| DI-AR, Lag | 0.73 (0.14) | 1.05 (0.18) | 0.91 (0.08) | 0.71 (0.17) | 0.83 (0.13) | 0.89 (0.25) | 0.87 (0.11) | 0.87 (0.26) |
| DI-AR | 0.74 (0.14) | 1.01 (0.19) | 0.89 (0.08) | 0.79 (0.18) | 0.83 (0.13) | 0.89 (0.25) | 0.87 (0.10) | 0.87 (0.26) |
| DI | 1.57 (0.25) | 0.21 (0.08) | 1.68 (0.26) | 0.10 (0.08) | 1.74 (0.43) | 0.13 (0.07) | 2.42 (0.74) | 0.05 (0.07) |
| *Balanced panel (N=149)* | | | | | | | | |
| DI-AR, Lag | 0.79 (0.13) | 1.00 (0.22) | 0.97 (0.07) | 0.59 (0.18) | 0.85 (0.13) | 0.85 (0.25) | 0.91 (0.09) | 0.78 (0.27) |
| DI-AR | 0.78 (0.13) | 0.94 (0.21) | 0.96 (0.08) | 0.60 (0.18) | 0.85 (0.13) | 0.85 (0.25) | 0.91 (0.09) | 0.82 (0.29) |
| DI | 1.59 (0.26) | 0.19 (0.08) | 1.64 (0.21) | 0.09 (0.08) | 1.73 (0.43) | 0.13 (0.07) | 2.42 (0.70) | 0.07 (0.07) |
| *Full data set; m=1, p=BIC, k fixed* | | | | | | | | |
| DI-AR, k=1 | 0.71 (0.14) | 1.15 (0.19) | 0.85 (0.09) | 0.91 (0.20) | 0.85 (0.11) | 1.13 (0.29) | 0.85 (0.12) | 0.90 (0.26) |
| DI-AR, k=2 | 0.72 (0.14) | 1.03 (0.18) | 0.88 (0.08) | 0.78 (0.17) | 0.80 (0.13) | 1.00 (0.24) | 0.86 (0.12) | 0.86 (0.26) |
| DI-AR, k=3 | 0.76 (0.13) | 0.97 (0.18) | 0.93 (0.08) | 0.66 (0.17) | 0.86 (0.12) | 0.82 (0.25) | 0.91 (0.10) | 0.76 (0.26) |
| DI-AR, k=4 | 0.76 (0.13) | 0.96 (0.19) | 0.93 (0.08) | 0.65 (0.17) | 0.88 (0.12) | 0.79 (0.25) | 0.90 (0.10) | 0.75 (0.25) |
| RMSE, AR Model | 0.010 | | 0.007 | | 0.009 | | 0.017 | |

**Table 4, continued**

| Forecast Method | CPI Rel. MSE | CPI $\hat{\alpha}$ | consumption deflator Rel. MSE | consumption deflator $\hat{\alpha}$ | CPI exc. food&energy Rel. MSE | CPI exc. food&energy $\hat{\alpha}$ | producer price index Rel. MSE | producer price index $\hat{\alpha}$ |
|---|---|---|---|---|---|---|---|---|
| **B. Horizon = 24 months** | | | | | | | | |
| *Benchmark models* | | | | | | | | |
| AR | 1.00 | | 1.00 | | 1.00 | | 1.00 | |
| LI | 0.70 (0.21) | 0.76 (0.12) | 0.70 (0.20) | 0.78 (0.11) | 0.99 (0.29) | 0.51 (0.25) | 0.65 (0.22) | 0.84 (0.19) |
| Phillips Curve | 0.84 (0.12) | 0.77 (0.08) | 0.81 (0.15) | 0.80 (0.09) | 0.72 (0.21) | 0.93 (0.19) | 0.77 (0.19) | 1.00 (0.06) |
| VAR | 0.92 (0.08) | 0.80 (0.22) | 0.98 (0.06) | 0.57 (0.18) | 1.00 (0.06) | 0.49 (0.34) | 1.18 (0.12) | 0.29 (0.10) |
| *Full data set (N=215)* | | | | | | | | |
| DI-AR, Lag | 0.74 (0.23) | 0.74 (0.18) | 0.75 (0.16) | 0.79 (0.13) | 0.92 (0.26) | 0.58 (0.28) | 0.82 (0.14) | 0.68 (0.12) |
| DI-AR | 0.75 (0.25) | 0.67 (0.16) | 0.71 (0.21) | 0.73 (0.12) | 0.96 (0.33) | 0.53 (0.27) | 0.77 (0.17) | 0.68 (0.13) |
| DI | 1.18 (0.22) | 0.40 (0.12) | 1.21 (0.18) | 0.38 (0.12) | 1.40 (0.22) | 0.30 (0.07) | 2.09 (0.72) | 0.19 (0.09) |
| *Balanced panel (N=149)* | | | | | | | | |
| DI-AR, Lag | 0.59 (0.22) | 0.95 (0.12) | 0.67 (0.18) | 0.84 (0.10) | 0.84 (0.22) | 0.69 (0.24) | 0.76 (0.14) | 0.78 (0.13) |
| DI-AR | 0.70 (0.24) | 0.72 (0.13) | 0.70 (0.20) | 0.75 (0.12) | 0.87 (0.29) | 0.61 (0.25) | 0.86 (0.15) | 0.62 (0.11) |
| DI | 1.07 (0.20) | 0.46 (0.12) | 1.08 (0.18) | 0.45 (0.12) | 1.43 (0.22) | 0.27 (0.07) | 2.10 (0.70) | 0.19 (0.08) |
| *Full data set; m=1, p=BIC, k fixed* | | | | | | | | |
| DI-AR, k=1 | 0.63 (0.20) | 1.04 (0.18) | 0.68 (0.17) | 0.97 (0.15) | 0.60 (0.25) | 1.12 (0.20) | 0.73 (0.17) | 0.93 (0.22) |
| DI-AR, k=2 | 0.61 (0.21) | 1.07 (0.17) | 0.72 (0.16) | 0.92 (0.13) | 0.64 (0.24) | 0.96 (0.17) | 0.68 (0.19) | 0.97 (0.20) |
| DI-AR, k=3 | 0.80 (0.17) | 0.82 (0.23) | 0.80 (0.12) | 0.83 (0.13) | 0.94 (0.25) | 0.56 (0.29) | 0.81 (0.11) | 0.80 (0.14) |
| DI-AR, k=4 | 0.76 (0.20) | 0.81 (0.21) | 0.74 (0.15) | 0.83 (0.14) | 0.92 (0.26) | 0.59 (0.29) | 0.78 (0.14) | 0.78 (0.14) |
| RMSE, AR Model | 0.052 | | 0.038 | | 0.046 | | 0.077 | |

Notes: See the notes to table 1.

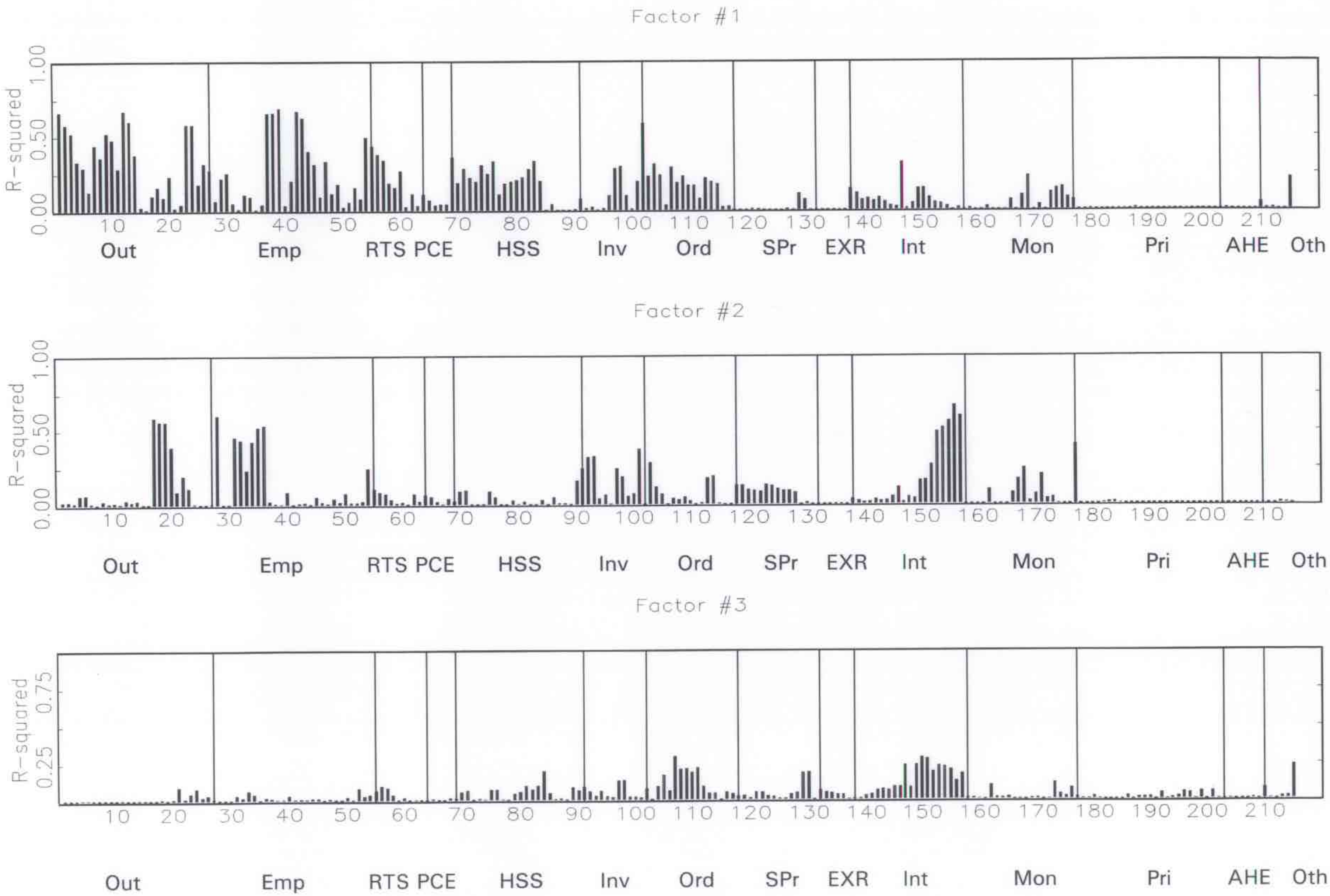


Figure 1. $R^2$s between factors and individual time series, grouped by category (see Appendix B)

Categories: Real output and income (Out); Employment and hours (Emp); Real retail, manufacturing and trade sales (RTS); Consumption (PCE); Housing starts and sales (HSS); Real inventories and inventory-sales ratios (Inv); Orders and unfilled orders (Ord); Stock prices (SPr); Exchange rates (EXR); Interest rates (Int); Money and credit quantity aggregates (Mon); Price indexes (Pri); Average hourly earnings (AHE); Miscellaneous (Oth)

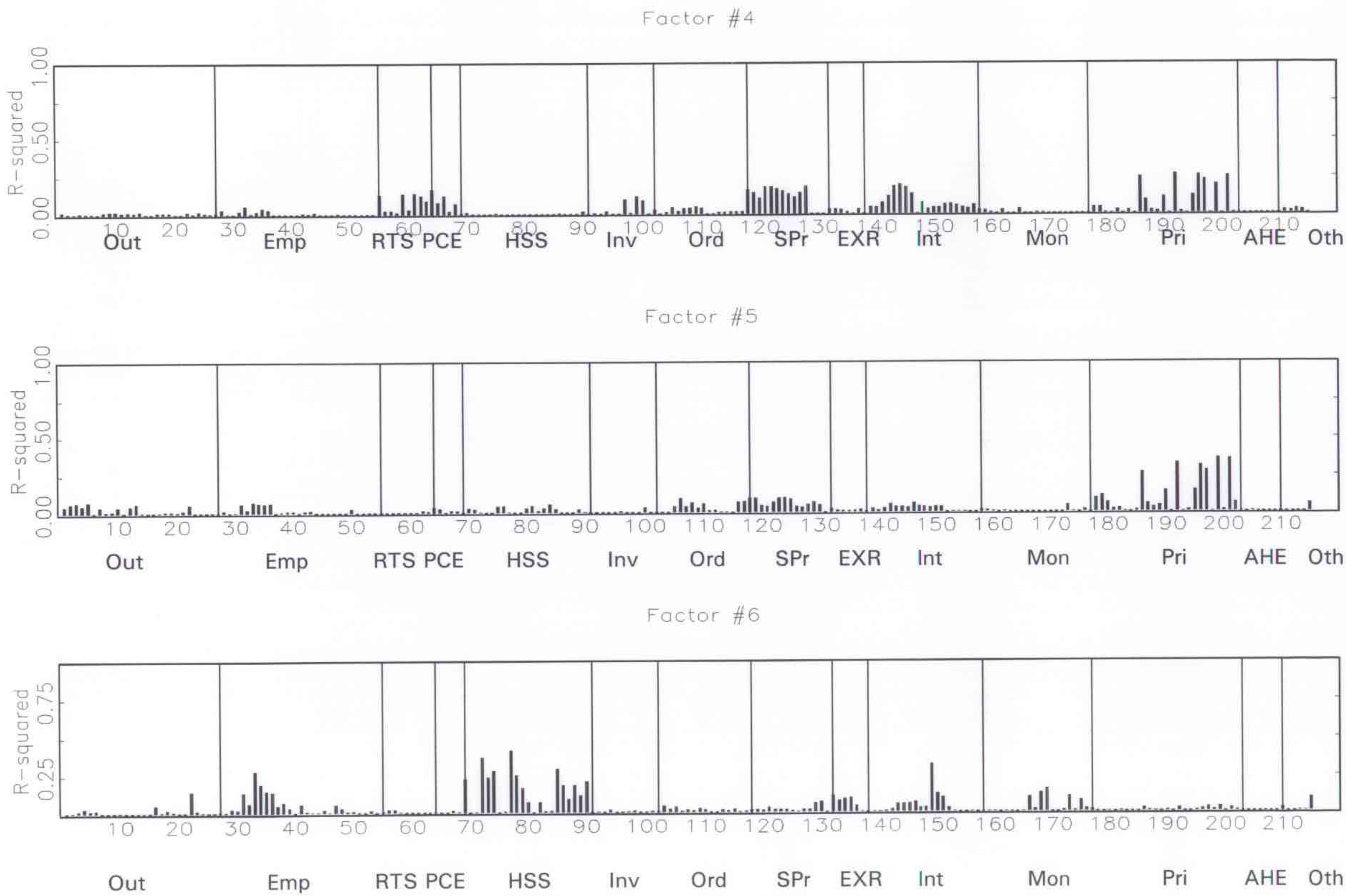


Figure 1 (continued)