NBER WORKING PAPER SERIES

DIFFUSION INDEXES

James H. Stock
Mark W. Watson

Working Paper 6702
http://www.nber.org/papers/w6702

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 1998

The authors thank Frank Diebold, Gregory Chow, Lucrezia Reichlin, and Charles Whiteman for helpful discussions and\or comments on the original July 1997 draft of this paper, and Lewis Chan and Alexey Onatski for skilled research assistance. This research was supported in part by National Science Foundation grants no. SBR-9409629 and SBR-9730489 through the NBER. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

On the Dynamics of Trade Reform
James H. Stock and Mark W. Watson
NBER Working Paper No. 6702
August 1998
JEL No. C32, E37

ABSTRACT

This paper considers forecasting a single time series using more predictors than there are time series observations. The approach is to construct a relatively few indexes, akin to diffusion indexes, which are weighted averages of the predictors, using an approximate dynamic factor model. Estimation is discussed for balanced and unbalanced panels. The estimated dynamic factors are (uniformly) consistent, even in the presence of time varying parameters and/or data contamination, and forecasts based on the estimated factors are efficient. In an application to forecasting U.S. inflation and industrial production using 224 monthly time series, these forecasts outperform various state-of-the-art benchmark models.

James H. Stock
Kennedy School of Government
Harvard University
Cambridge, MA 02138
and NBER
James_Stock@harvard.edu

Mark W. Watson
Woodrow Wilson School
Princeton University
Princeton, NJ 08544-1013
and NBER
mwatson@princeton.edu

1. Introduction

Recent advances in information technology now 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 output or inflation. Time series models currently used for macroeconomic forecasting, however, incorporate only a handful of series: vector autoregressions, for example, typically contain a half-dozen to one dozen variables, rarely more. Although thousands of time series are available in real time, a theoretical framework for using these data for time series forecasting remains undeveloped.

This paper addresses the problem of forecasting a single time series using a very large number of predictors, potentially many more predictors than dates at which the time series are observed. Our approach is motivated by the diffusion indexes developed by business cycle analysts at the National Bureau of Economic Research (NBER). These indexes are averages of contemporaneous values of a large number of time series; a classic use of a diffusion index is to measure whether a recession or expansion is widespread throughout the economy. Because it is an average of many variables, a diffusion index summarizes the information in a large number of economic time series. In constructing diffusion indexes, NBER business cycle analysts exercised expert judgment to identify the series and the weight placed on each series in the index.

Section 2 provides a probability model in which diffusion indexes are interpreted as estimates of the unobserved factors in a dynamic factor model, and discusses the estimation of these factors. This dynamic factor model has several important features. First, because empirical evidence suggests that time variation in macroeconomic relations is widespread (e.g. Stock and Watson [1996]), the factor loadings are permitted to evolve over time. Second, the factor structure is approximate, in the sense that the idiosyncratic errors can be correlated across series. Third, the model is nonparametric, in the sense that the correlation structures and distributions of the idiosyncratic terms and the factors, and the precise lag structure by which the factors enter, are not specified parametrically. Fourth, a practical concern when working with a large number of time series is that a large break or outlier arising from a data entry error or a redefinition might go undetected, and this possibility is introduced into the analysis. Fifth, because economic time series are typically available over different spans, the model and estimation procedures are developed for the cases of both a balanced and unbalanced panel.

This factor-based approach to forecasting can be contrasted to conventional regression-based model selection. With a very large number of predictor variables, it is computationally infeasible to enumerate and to estimate all possible models up to a given order. Although this computational problem can be ameliorated by making informed choices about the models to be estimated, more fundamentally the model selection approach is prone to producing particularly poor out of sample forecasts because of fitting so many models. In contrast, the dynamic factor model places very strong restrictions on the joint behavior of the predictors that permits extreme parameter reduction, so that for forecasting purposes the very many predictors can be replaced by a handful of factors.

Asymptotic results are presented in Section 3. The asymptotic framework is motivated by the application to macroeconomic forecasting. Because the number of time series (N) far exceeds the number of observation dates (T), N and T are modeled as tending to infinity, but $T/N \rightarrow 0$. Because macroeconomic theory does not clearly suggest finitely many factors, the number of factors (r) is treated as tending to infinity, but much more slowly than T. Because r is not known, the number of estimated factors (k) is not assumed to equal the number of true factors. In this framework, it is shown that, if $k \geq r$, the estimated factors are uniformly consistent (they span the space of the true factors, uniformly in the time index). Given this result and some additional conditions, it is then shown that, if $k \geq r$, an information criterion will consistently estimate the number of factors entering the forecasting equation for the variable of interest, and the resulting forecasts are as efficient asymptotically as if the true factors were observed. These theoretical predictions are examined in and supported by a Monte Carlo experiment reported in section 4.

In section 5, these methods are used to produce monthly forecasts of the twelve-month growth of industrial production (IP) and the twelve-month growth of the consumer price index (CPI) in the United States. The full data set spans 1959:1-1997:9. Factors are extracted and forecasts are made for a balanced panel of 170 time series and an unbalanced panel of 224 time series. These diffusion index forecasts perform well in a simulated real-time forecasting comparison with several state of the art benchmark multivariate models.

This research is related to two bodies of literature. The first is a relatively small literature in which dynamic factor models have been applied to macroeconomic data. Geweke (1977) and Sims and Sargent (1977) analyzed these models in the frequency domain for a small number of variables. Engle and Watson (1981), Sargent (1989), and Stock and Watson (1991) estimated small-N parametric time domain dynamic factor models by maximum likelihood. Quah and Sargent (1993) used the EM algorithm to extend this approach to a moderate number of series ($N=60$). The second related literature is the large body of work that uses approximate factor structures to study asset prices. Contributions include Chamberlain and Rothschild (1983), Connor and Korajczyk (1986, 1988, 1993), Mei (1993), Schneewwiss and Mathes (1995), Bekker et. al. (1996), Geweke and Zhou (1996), and Zhou (1997); also see the survey in Campbell, Lo and McKinley (1996, chapter 6)).

The work in these literatures most closely related to the present paper is by Connor and Korajczyk (1986, 1988, 1993) and Forni and Reichlin (1996, 1997, 1998); both consider the determination of the number of factors and their estimation in large systems. Working within a static approximate factor model that allows some cross-sectional dependence among the idiosyncratic errors, Connor and Korajczyk (1986, 1988, 1993) show that factors estimated by principal components are consistent (at a given date) as $N \rightarrow \infty$ with T fixed. They apply their methods to evaluating the arbitrage pricing theory of asset prices. Forni and Reichlin (1998), working with a dynamic factor model with mutually uncorrelated idiosynchratic errors, show that cross sectional averages consistently estimate a certain scalar linear combination of the factors. They use this insight in Forni and Reichlin (1996, 1998) to motivate heuristically a dynamic principal components procedure for estimating the vector common factors and for studying the dynamic properties of the factors. Forni and Reichlin (1997) suggest an alternative estimator, which they motivate by dynamic principal components, although no proofs of consistency of the estimated factors are provided.[1] In the first applications of this large cross section approach to macroeconomic data, they apply their methods to large regional and sectoral data sets, for example Forni and Reichlin (1998) analyze productivity and output for 450 U.S. industries.

Relative to this literature, the current paper makes four main methodological contributions, which are motivated by our focus on real-time economic forecasting. First, consistency of the estimated factors is shown when the factor loadings are time varying and the number of factors tends to infinity. Second, the estimated factors are shown to be uniformly (in the time subscript) consistent, not just for a given date, and rate of consistency is given. Third, estimation methods that are computationally feasible for large N are presented for both balanced and unbalanced panels. Fourth, results are given on the use of information criteria to select the number of factors for forecasting. The empirical contribution of this paper is to demonstrate the potential for substantial improvements in macroeconomic time series forecasts using these methods.

2. The Model and Estimator

2.1. The model

Let y_t be a scalar time series variable and let X_t be a N-dimensional multiple time series variable. Throughout, y_t is taken to be the variable to be forecast while X_t is the vector time series variable that contain useful information for forecasting y_{t+1}. It is assumed that X_t can be represented by the factor structure,

$$(2.1) \qquad X_t = \Lambda_t F_t + e_t$$

where F_t is the $r \times 1$ common factor and e_t is the $N \times 1$ idiosyncratic disturbance. The idiosyncratic disturbances are in general correlated across series and over time; specific assumptions used for the asymptotic analysis are given in section 3.

Our main objective is to estimate $E(y_{t+1} | X_t)$. We model y_{t+1} as,

$$(2.2) \qquad y_{t+1} = \beta_t{}' F_t + \epsilon_{t+1}$$

where $E(\epsilon_{t+1} | X_t, y_t, \beta_t, X_{t-1}, y_{t-1}, \beta_{t-1} \ldots) = 0$. This embodies three assumptions: that $E(y_{t+1} | X_t, y_t, \beta_t, X_{t-1}, y_{t-1}, \beta_{t-1} \ldots)$ depends F_t but not otherwise on X_t; that lags of F_t do not enter (2.2); and that lags of y_t do not enter (2.2). This first assumption is the key assumption that permits the dimension reduction necessary for handling very large X_t. The second assumption is not restrictive, because F_t in (2.2) can be reinterpreted as including lags without changing any essential argument. The third assumption also is not restrictive in the sense that y_{t+1} can be reinterpreted as a quasidifference so that lagged values of y_t can be incorporated into the model.

The factor loadings Λ_t $(N \times r)$ and the coefficients β_t $(r \times 1)$ vary over time according to,

$$\Lambda_t = \Lambda_{t-1} + h\zeta_t, \tag{2.3}$$

$$\beta_t = \beta_{t-1} + \eta_t \tag{2.4}$$

where h is a diagonal $N \times N$ scaling matrix and η_t and ζ_t are, respectively, $r \times 1$ and $N \times r$ stochastic disturbances. Specific assumptions about h are stated in section 3.

Depending on what further assumptions are made concerning the disturbances and the factor loading matrices, this model contains several important special cases. One is the static factor model in which the factor loadings are constant (so $\Lambda_t = \Lambda_0$), e_t is serially uncorrelated, F_t and $\{e_{jt}\}$ are mutually uncorrelated and are i.i.d.. If e_{it} and e_{jt} are independent for $i \neq j$, the model is referred to as an exact static factor model. If the idiosyncratic disturbances are weakly correlated across series, the model is an approximate (static) factor model (cf. Chamberlain and Rothschild (1983) and Connor and Korajczyk (1986, 1993)).

Another important special case of (2.1)-(2.4) is the dynamic factor model without time variation, as has been studied by, among others, Geweke (1977), Sargent and Sims (1977), Engle and Watson (1981), Sargent (1989), Stock and Watson (1991), and Quah and Sargent (1993). In the standard dynamic factor model, dynamics are introduced in three ways: the factors are assumed to evolve according to a time series process; the idiosyncratic error terms are serially correlated; and the factors can enter with lags (or, in general, leads).

By suitable redefinition of the factors and the idiosyncratic disturbances, the dynamic factor model can be rewritten in the form (2.1) with Λ_t constant. To see this, let Z_t denote a $n \times 1$ vector of time series variables which are assumed to satisfy the dynamic factor model,

$$Z_{it} = \alpha_i(L)f_t + v_{it} \tag{2.5}$$

(2.6) $g_i(L)v_{it} = \eta_{it}$, η_{it} i.i.d. $N(0,\phi_i^2)$

for $i=1,\dots,n$, where f_t is a vector of factors and L is the lag operator. In the econometric literature using dynamic factor models, $\{\nu_t\}$ and $\{\eta_{it}\}$, $i=1,\dots,n$ are taken to be mutually independent. Let $\alpha_i(L)$ have order q and let $g_i(L)$ be a finite order lag polynomial with roots outside the unit circle. (Typically normality of these disturbances is further assumed to motivate using the Kalman filter to compute the maximum likelihood estimates of the factors.) The model is completed by making an additional assumption specifying the stochastic process followed by the factors, such as a Gaussian vector autoregression, where the factors are distributed independently of $\{v_{it}\}$.

There are at least two ways to rewrite this model in the form (2.1) with time invariant parameters. The first is to let $X_t = Z_t$, $F_t = (f_t', f_{t-1}', \dots f_{t-q}')'$, $\Lambda = (\alpha_0\ \alpha_1\ \dots\ \alpha_q)$, and $e_t = v_t$. With these definitions, (2.5) and (2.6) are equivalent to (2.1), where the idiosyncratic errors e_t are serially correlated, F_t has dimension $r=\dim(f_t)\times(q+1)$, and $\Lambda_t=\Lambda$. In this representation, the factors F_t are dynamically singular in the sense that the spectral density matrix of F_t has rank $\dim(f_t)$.

Another way to rewrite (2.5) in static form is to create an augmented, or stacked, vector of data, where Z_t is augmented by lags. Specifically, let $X_t^\dagger = (Z_t', Z_{t-1}', \dots, Z_{t-p+1}')'$, $F_t^\dagger = (f_t', f_{t-1}', \dots f_{t-q-p+1}')'$, $e_t^\dagger = (v_t', v_{t-1}', \dots, v_{t-p+1}')'$, and let $\Lambda^\dagger$ be the $nq\times(p+q)\dim(f_t)$ matrix, partitioned into $n\times\dim(f_t)$ blocks defined by,

$$\Lambda^\dagger = \begin{bmatrix} A_0 & A_1 & \cdots\cdots & A_p & 0 & \cdots & 0 \\ 0 & A_0 & \cdots\cdots & A_{p-1} & A_p & \cdots & 0 \\ \vdots & & & & & & \vdots \\ 0 & \cdots & 0 & A_0 & & \cdots\cdots\cdots & A_p \end{bmatrix}$$

where $A_j = (\alpha'_{1j}, \alpha'_{2j}, \ldots, \alpha'_{nj})'$, where $\alpha_i(L) = \Sigma_{j=0}^{p}\alpha_{ij}L^j$, where α_{ij} is $1\times\dim(f_t)$. Thus, (2.5)-(2.6) can be rewritten in static form (2.1):

$$X_t^\dagger = \Lambda^\dagger F_t^\dagger + e_t^\dagger \tag{2.7}$$

where $N=\dim(X_t^\dagger)=np$ and $r=\dim(F_t^\dagger)=(p+q)\times\dim(f_t)$. As in the first representation, the factors $F_t^\dagger$ are dynamically singular. Operationally, these two representation suggest different estimation strategies, the first by extracting dynamic factors using contemporaneous values of X_t only, the second by using lags of X_t as well. A potential advantage of the second representation is that additional indicators (lagged values of X_t) are introduced for the estimation of f_t, which could improve finite sample performance.

These representations exploit the fact that $\alpha_i(L)$ has finite order. If it has infinite order then these static representations have infinitely many factors. Parametric time domain dynamic factor models explored to date in the literature assume finite order $\alpha_i(L)$ (cf. Sargent (1989), Stock and Watson (1991), and Quah and Sargent (1993)). Whether this is a problem is an empirical issue.

It is convenient at this point to introduce some additional notation. Let X_{it} denote the observation on variable i at time t and let the $T\times 1$ vector $\underline{X}_i = (X_{i1}, X_{i2},\ldots,X_{iT})'$. Also let F_{it} be the observation on the ith factor at time t, let the $T\times r$ matrix $F=(F_1, F_2,\ldots,F_T)'$, let $P_F = F(F'F)^{-1}F'$, and let $\Lambda_t = (\lambda_{1t}, \lambda_{2t},\ldots,\lambda_{Nt})'$, where λ_{it} is the $r\times 1$ vector of factor loadings on variable i at time t. Let F^0 denote the true value of F. Let ζ'_{it} denote the i'th row of ζ_t. Throughout, c and d denote generic finite positive constants. For any matrix M, its (i,j) element is M_{ij} and its norm is $\|M\| = \{tr(M'M)\}^{1/2}$. For a real symmetic matrix M, mineval(M) and maxeval(M) denote its minimum and maximum eigenvalue.

Finally, to address the problem of missing data in an unbalanced panel, let I_{it} be a nonrandom indicator function, where $I_{it}=1$ if the i^{th} variable is observed at date t, and $I_{it}=0$ otherwise.

2.2. Estimation

There are three challenges in the estimation of the factors. First, the number of parameters is large; with N=500 and k=15, for example, the initial factor loading matrix Λ_0 has 7500 elements. Second, if both F_t and Λ_t are treated as stochastic, then the model is a bilinear form in random variables. Third, in our application we must handle an unbalanced panel, since different macroeconomic time series are available over different periods of time.

The standard method of estimation of dynamic factor models is by maximum likelihood using the Kalman filter. Application of the Kalman filter to dynamic factor models can be justified by assuming that the factor loading matrices are constant and by making suitable parametric assumptions on the disturbances (mutual independence, Gaussianity, and a parametric serial correlation structure); then the Gaussian likelihood can be evaluated using the Kalman filter and the likelihood can be maximized accordingly. This has been implemented in low dimensional systems (e.g. Engle and Watson (1981), Sargent (1989), Stock and Watson (1991)) and in higher dimensional systems (N=60) where the maximization is done using a modification of this appproach based on the EM algorithm (Quah and Sargent (1993)). Although the Kalman filter is easily modified for missing data, nonlinear filters are needed to compute the likelihood when F_t and Λ_t are both random. Moreover, likelihood maximization when N is very large does not seem promising from a computational perspective.

We therefore take a different approach and estimate the dynamic factor model in its static (or stacked) form. The approach here is quasi-MLE, in the sense that the estimator is motivated by making strong parametric assumptions, but the consistency of the estimated factors is shown under weaker nonparametric assumptions given in section 3. To motivate the estimation strategy, we suppose that $h=0$ so $\Lambda_t=\Lambda_0$, and e_{it} is i.i.d. $N(0,\sigma_e^2)$ and independent across series. We also diverge from the treatment of F_t in dynamic factor models, in which F_t is modeled as obeying a stochastic process, and instead treat $\{F_t\}$ as a $T\times r$ dimensional unknown nonrandom parameter to be estimated. With this notation and under these restrictive assumptions, the maximum likelihood estimator for (Λ_0, F) solves the nonlinear least squares problem with the objective function,

$$V_{NT}(F,\Lambda_0) = (NT)^{-1}\Sigma_{i=1}^{N}\Sigma_{t=1}^{T} l_{it}(X_{it}-\lambda_{i0}'F_t)^2. \tag{2.8}$$

Let $(\tilde{F}, \tilde{\Lambda}_0)$ denote the minimizers of $V_{NT}(F,\Lambda_0)$. These satisfy the first order conditions,

$$\tilde{\lambda}_{i0} = (\Sigma_{t=1}^{T} l_{it}\tilde{F}_t\tilde{F}_t')^{-1}(\Sigma_{t=1}^{T} l_{it}\tilde{F}_tX_{it}), \tag{2.9a}$$

$$\tilde{F}_t = (\Sigma_{i=1}^{N} l_{it}\tilde{\lambda}_{i0}\tilde{\lambda}_{i0}')^{-1}(\Sigma_{i=1}^{N} l_{it}\tilde{\lambda}_{i0}X_{it}). \tag{2.9b}$$

The estimator $\hat{F}$ for which we provide results is the numerator matrix of $\tilde{F}_t$:

$$\hat{F}_t = \Sigma_{i=1}^{N} l_{it}\tilde{\lambda}_{i0}X_{it}/\Sigma_{i=1}^{N} l_{it} \tag{2.10}$$

Efficient computation of $(\tilde{F}, \tilde{\Lambda})$ depends on whether the panel is balanced. If the panel is balanced, then the parameters can be estimated by solving either of two eigenvalue problems. The first eigenvalue problem obtains by subsituting (2.9a) into (2.8) with $l_{it}=1$ to yield the concentrated objective function, $V_{NT}(\tilde{\Lambda},F) = (NT)^{-1}\Sigma_{i=1}^{N}\underline{X}_i'\underline{X}_i-(NT)^{-1}\Sigma_{i=1}^{N}\underline{X}_i'P_F\underline{X}_i$. Because of the normalization $F'F/T=I_k$, minimizing $V_{NT}(\tilde{\Lambda},F)$ is equivalent to maximizing $\mathrm{tr}\{F'(N^{-1}\Sigma_{i=1}^{N}\underline{X}_i\underline{X}_i')F\}$, which is solved by choosing F as the eigenvectors corresponding to the k largest eigenvalues of the T×T matrix $N^{-1}\Sigma_{i=1}^{N}\underline{X}_i\underline{X}_i'$. This is the computational strategy used by Connor and Korajczyk (1986, 1993).

The second eigenvalue problem obtains by substituting (2.9b) into (2.8) to yield the concentrated objective function $V_{NT}(\Lambda,\tilde{F})$, which is minimized by the k eigenvectors corresponding to the k largest eigenvalues of the N×N matrix $T^{-1}\Sigma_{t=1}^{T}X_tX_t'$. These eigenvectors are the first k principal components of X_t.

A different approach must be used in an unbalanced panel. In principal it is possible to iterate on the first order conditions (2.9), subject to the normalization condition $\tilde{F}'\tilde{F}/T=I_k$. This is, however, computationally burdensome for large N. In the unbalanced panel we therefore minimize V_{NT} using the EM algorithm. Continue to assume that $\Lambda_t=\Lambda_0$. Let X^*_{it} denote the latent value of X_{it}, so $X^*_{it} = \lambda_{i0}'F_t + e_{it}$, and $X_{it} = X^*_{it}$ if $I_{it}=1$ and X_{it} is unobserved otherwise, and let $V^*_{NT}(F,\Lambda_0)$ denote the "complete-data" likelihood, $V^*_{NT}(F,\Lambda_0) = -(NT)^{-1}\Sigma_{i=1}^{N}\Sigma_{t=1}^{T}(X^*_{it}-\lambda_{i0}'F_t)^2$. The EM algorithm proceeds by iteratively maximizing the expected complete-data likelihood, $Q_{NT}(F^{(i)},\Lambda_0^{(i)}) = E[V^*_{NT}(F^{(i)},\Lambda_0^{(i)})|X,F^{(i-1)},\Lambda_0^{(i-1)}]$, where $F^{(i)}$ and $\Lambda_0^{(i)}$ respectively denote the i^{th} iterates of F and Λ_0. Under the assumption that e_{it} is i.i.d. $N(0,\sigma^2)$, this has the simple form, $Q_{NT}(F^{(i)},\Lambda_0^{(i)}) = -(NT)^{-1}\Sigma_{i=1}^{N}\Sigma_{t=1}^{T}(\hat{X}^{*(i-1)}_{it}-\lambda_{i0}^{(i)\prime}F_t^{(i)})^2$ (plus terms that do not depend on Λ_0 or F), where $\hat{X}^{*(i-1)}_{it} = E[X^*_{it}|X,F^{(i-1)},\Lambda_0^{(i-1)}] = X_{it}$ if $I_{it}=1$ and $= \lambda_{i0}^{(i-1)\prime}\hat{F}_t^{(i-1)}$ if $I_{it}=0$. The arguments about concentrating the likelihood above apply, so $F^{(i)}$ are computed as the eigenvalues of $N^{-1}\Sigma_{i=1}^{N}\hat{\underline{X}}_i^{*(i-1)\prime}\hat{\underline{X}}_i^{*(i-1)}$, where $\hat{\underline{X}}_i^{*(i-1)}$ is defined analogously to $\underline{X}_i$ above except using $\hat{X}^{*(i-1)}_{it}$ rather than X_{it} as needed for estimates of the missing observations. The unbalanced panel quasi-MLEs are obtained by iterating this process to convergence. Note also that this approach extends to other data irregularities, in particular situations with mixed sampling frequencies, for example some variables might be observed monthly while others are observed either as end-of-quarter values (stocks) or as quarterly averages (flows).

3. Asymptotics

3.1. Asymptotic framework

We wish to apply the model and estimation methods of the previous section to empirical settings with the following features: the number of series is very large, perhaps in the thousands; the number of time periods is large, but less than the number of series, in the range of T = 100-400 for monthly data; the number of factors is far less than the number of time series, for example $r=10$ or 15; and the researcher does not know the number of factors, so $k \neq r$ in general. In addition, two types of parameter instability (instability in the factor loading matrices) are of particular concern: drifts in the parameters resulting from the ongoing evolution of economic relations, and gross breaks in the parameters resulting from series redefinitions or data entry errors. Empirical evidence in the literature suggests that the former type of instability is widely present in U.S. macroeconomic time series, but that magnitude of parameter drift is fairly small. One would hope that careful attention to data would keep the second type of instability to a minimum, but realistically when hundreds or thousands of time series are used some such gross errors might go undetected.

The asymptotic nesting adopted here is designed to capture these features. Specifically, both N and T are taken to tend to infinity, but $T/N \rightarrow 0$. Also, the number of true and estimated factors are assumed to tend to infinity, but at the same slow rate (both as lnT); because they are sequences, they will be denoted r_T and k_T. These assumptions are summarized in

Condition R (rates)

$T \rightarrow \infty$, $\ln(N)/\ln(T) \rightarrow \rho > 2$, $1 \leq k_T \leq \bar{k}\ln T$,, $1 \leq r_T \leq \bar{r}\ln T$, and $k_T/r_T \rightarrow \mu > 0$.

The possibility of two types of parameter instability is addressed by modeling h as a sequence of random matrices h_T that satisfy,

Condition TV (time varying factor loadings)

$h_T = \text{diag}(h_{1T},\dots,h_{NT})$, where h_{iT} is i.i.d., h_T is independent of $(\epsilon_t, e_t, \eta_t, \zeta_t\}$, and $T\kappa_{4T} = O(1)$, where $\kappa_{pT} \equiv (Eh_{iT}^p)^{1/p}$.

Two concerns about time varying parameters were outlined in the introduction: moderate parameter drift because of structural change for many series, and large occasional jumps because of redefinitions or coding errors for a few series. Condition TV handles these problems. Consider the following example. Suppose a fraction π of the seies are subject to a redefinition error at date t*, so that for these series $\Delta\Lambda_t=a$ if $t=t^*$ and $=0$ otherwise. The remaining $1-\pi$ series experience moderate parameter drift of the form $h_{it}=b/T$ (so the full-sample parameter drift is $O(T^{-1/2})$, the same order as conventional sampling uncertainty were F_t observed)[2]. Then $T\kappa_{qT} \rightarrow [a^qT^{q-1}\pi+b^q(1-\pi)]^{1/q}$, so $T\kappa_{4t}=O(1)$ if $\pi=O(1/T^3)$. If $\rho=3$ in condition R, this corresponds to a constant fraction of the series having redefinition contamination and the rest having moderate parameter drift.

The next condition restricts Λ_0.

Condition FL (factor loadings)

$|\lambda_{i0,m}| \leq \bar{\lambda} < \infty$, $i=1,\dots,N$, $m=1,\dots,r_T$; $r_T\text{mineval}(\Lambda_0'\Lambda_0/N) \geq d > 0$; $\text{tr}(\Lambda_0'\Lambda_0/N) \leq c < \infty$; and there is a sequence of positive semidefinite $r_T \times r_T$ matrices D such that $\|\Lambda_0'\Lambda_0/N\text{-}D\| \rightarrow 0$.

The condition that $\text{tr}(\Lambda_0'\Lambda_0/N) \leq c$ ensures that the expected contribution of the factors to the variance of X_t is finite. On the other hand, the eigenvalue assumption ensures that the contribution of each factor to the variance of the multiple time series X_t be sufficiently large. For example, suppose that Λ_0 is such that the factor F_{jt} loads onto (say) N_{jT} variables, that the corresponding coefficient is unity, and that each variable has a single factor. Then $\Lambda_0'\Lambda_0/N$ is diagonal with eigenvalues N_{jT}/N and $tr(\Lambda_0'\Lambda_0/N)=1$. Evidently the eigenvalue condition is satisfied if $N_{jT}/(N/r_T)$ converges to a nonzero constant, $j=1,\ldots,r_T$. On the other hand, if $N_{jT}/(N/r_T)\to 0$ for some j, this condition fails. In this example, the number of variables upon which the factors load must be the same order of magnitude.

3.2 Moment conditions

The final condition concerns the moments of the various stochastic terms in (2.1) and (2.4). Let $\Gamma_{ij,\ell m}(u)$ denote the (ℓ,m) element of $\Gamma_{ij}(u)$, etc.

Condition M (moments and dependence)

The random variables $\{e_t, \zeta_t, F_t\}$ satisfy:

(a) (i) $Ee_{it} = 0$, $E(e_t'e_{t+u}/N) = \gamma(u)$, and $\Sigma_{u=-\infty}^{\infty} |\gamma(u)| < \infty$,

(ii) $Ee_{it}e_{jt} = \tau_{ij}$, where $\lim_{N\to\infty} N^{-1}\Sigma_{i=1}^{N}\Sigma_{j=1}^{N}|\tau_{ij}| < \infty$,

(iii) $\sup_{i,t}Ee_{it}^4 < \infty$ and $\lim_{N\to\infty}\sup_{s,t}N^{-1}\Sigma_{i=1}^{N}\Sigma_{j=1}^{N}|cov(e_{is}e_{it},e_{js}e_{jt})| < \infty$.

(b) (i) $E\zeta_{it,m} = 0$, $E\zeta_{it}\zeta_{jt+u}' = \Gamma_{ij}(u)$, and $\Sigma_{u=-\infty}^{\infty}\sup_{i,j,l,m}|\Gamma_{ij,lm}(u)| < \infty$

(ii) $\lim_{N\to\infty}\sup_m N^{-1}\Sigma_{i=1}^{N}\Sigma_{j=1}^{N}\Sigma_{u=-\infty}^{\infty}|\Gamma_{ij,mm}(u)| < \infty$,

(iii) $\sup_{i,s,m}E\zeta_{is,m}^4 < \infty$ and

$\lim_{N\to\infty}\sup_{\ell,m}N^{-1}\Sigma_{i=1}^{N}\Sigma_{j=1}^{N}\sup_{t,u_1,u_2,u_3}|cov(\zeta_{it,\ell}\zeta_{it+u_1,m},\zeta_{jt+u_2,\ell}\zeta_{jt+u_3,m})| < \infty$.

(c) (i) $E\zeta_{it}e_{jt+u} = \Psi_{ij}(u)$ and $\sup_i\Sigma_{u=-\infty}^{\infty}\sup_m|\Psi_{ii,m}(u)| < \infty$,

(ii) $\sup_m N^{-1}\Sigma_{i=1}^{N}\Sigma_{j=1}^{N}\sup_{t,u,v}|cov(e_{it}\zeta_{it+u,m},e_{jt}\zeta_{jt+v,m})| < \infty$.

(d) (i) $|F_{it}^0| \le \bar{F} < \infty$, $i=1,\ldots,r_T$, $t=1,\ldots,T$,

(ii) $EF_t^0F_t^{0\prime} = \Sigma_{F,T}$, where $0<d\le \text{mineval}(\Sigma_{F,T})\le c<\infty$.

(iii) $\sup_{\ell,m,t} \sum_{u=-\infty}^{\infty} |\mathrm{cov}(F^0_{\ell t}F^0_{mt}, F^0_{\ell t+u}F^0_{mt+u})| < \infty$.

These assumptions limit the dependence across series and over time of these disturbances. It should be emphasized that the various disturbances are not assumed to be mutually independent. No restriction is made on the dependence between F_t and the errors (e_t, ζ_t). Also, e_t and ζ_t can be dependent, even across series, subject to condition (c).

Condition M is all satisfied in the leading case of an exact time invariant factor model, in which e_{it} and F_{it} are i.i.d. and mutually independent and $\zeta_t=0$. However, they allow for more temporal and cross-series dependence than in the time invariant factor model and in this sense accomodate an approximate factor structure.

Condition M is also satisfied when (2.1) is the static representation of a parametric dynamic factor model with constant factor loadings as discussed in section 2.1. Condition M(a) holds by (2.6), the independence of v_{it} across series, and the assumed stationarity of v_t. Conditions M(b) and M(c) are not relevant because the coefficients are time invariant. Condition M(d)(ii) is satisfied by the factor model as written. Conditions M(d)(i) and M(d)(iii) represent additional conditions.[3] Finally, r_T, which is $(q+1)\dim(f_t)$ in the dynamic factor model, satisfies condition R if $\dim(f_t) = O(\ln T)$.

Because F^0_t is assumed bounded by condition M(d)(i), this assumption is also made for its estimator: $\tilde{F}_t$ is assumed to satisfy $|\tilde{F}_{it}| \leq \bar{F} < \infty$, $i=1,\ldots,r_T$, $t=1,\ldots,T$. Because F is identified only up to a nonsingular $k\times k$ transformation, without loss of generality the additional normalization $\tilde{F}'\tilde{F}/T = I_{k_T}$ is imposed.

3.3. Results

The results are all developed for the case of a balanced panel. This is done primarily to streamline the notation and calculations; extension of the results to the unbalanced panel is left to future research.

Our first theoretical result is that the estimator $\hat{F}_t$ given in (2.10) is uniformly (in t) consistent for a linear combination of the true factors F_T, at the rate δ_{NT}.

Theorem 1. Let X_t and Λ_t obey (2.1) and (2.3). Suppose that conditions R, FL, and M hold, and let $h=h_T$, where h_T satisfies condition TV.

(a) Then, $\delta_{NT}\sup_t \|\hat{F}_t-H_{NT}F_t\| \xrightarrow{p} 0$, where $\delta_{NT} = T^b$ for any $b<\min[\frac{1}{2}\rho-1,1]$, and H_{NT} is not a function of (i,t).

(b) If in addition the normalization $\tilde{F}'\tilde{F}/T=I_{k_T}$ is adopted, $\|H_{NT}-H\| \xrightarrow{p} 0$, where $H = R\Sigma_{F,T}D$, where R is a nonrandom $k_T\times r_T$ matrix with row rank of $\min(k_T,r_T)$.

All proofs are given in Appendix A.

Several remarks are in order. First, the consistency of the estimated factors is obtained by averaging over a very large number of cross sectional observations, relative to the number of time series observations. Technically this is reflected in the condition that $\rho>2$ (so $T=o(N^{\frac{1}{2}})$), and in the exponent b being an increasing function of ρ. In contrast to conventional factor model estimation approaches which require a large number of time series observations and a small number of variables, a larger N relative to T improves the asymptotic performance in the sense that consistency is achieved at a faster rate.

Second, for the interpretation of the matrices H and R it is useful to consider separately the three cases of $k_T<r_T$, $k_T=r_T$, and $k_T>r_T$. When $k_T<r_T$, the rows of R span the space of the first k_T eigenvectors of the positive definite $r_T\times r_T$ matrix $A = \Sigma^{\frac{1}{2}}_{F,T}D\Sigma^{\frac{1}{2}\prime}_{F,T}$. When $k_T=r_T$, R is a full rank square matrix with $R'R=I_{k_T}$ so asymptotically $\hat{F}_t$ equals F^0_t up to the full rank transformation matrix H. When $k_T>r_T$, the row rank of R and thus H is only r_T, so $\hat{F}_t$ contains k_T-r_T redundant estimates of the factors that are just linear combinations of the r_T elements of $\hat{F}_t$.[4]

The second theorem addresses the use of the estimated factors to forecast y_{t+1}. Intuitively, the uniform consistency of $\hat{F}_t$ suggests that these estimated factors can in effect be treated as the true factors for the purposes of forecasting y_{t+1}. Two complications arise however. The first is that, if the parameters β_t in (2.2) evolve over time and $var(\Delta\beta_t)=O(1)$, then β_t is not consistently estimable even were F^0 known. Here, we provide formal results only for the case of no time variation in the forecasting equation, i.e. $\beta_t=\beta$.

The second complication arises when the true number of factors is unknown, as is the case in practice. We therefore consider the problem of the estimation of the number of factors, q, that enter the forecasting equation using an information criterion. The information criterion is of the form,

$$(3.1) \qquad IC_q = \ln(\hat{\sigma}^2_\epsilon(q)) + g(T)q$$

where $\hat{\sigma}^2_\epsilon(q) = SSR(q)/T$, where SSR(q) is the sum of squared residuals from estimation of (2.2) by OLS using q estimated factors. The function g(T) is the penalty function, for example $g(T)=\ln T/T$ for the Bayes Information Criterion (BIC). The information criterion estimate of r, $\hat{r}$, solves $\min_{1\leq q\leq k} IC_q$.

The following theorem provides sufficient conditions for forecasts based on $\hat{F}$ to be uniformly consistent for forecasts based on F^0.

Theorem 2. Suppose that the conditions of theorem 1 hold but in addition $k_T=k$ and $r_T=r$ are fixed and $\sup_t E\epsilon_t^4 < \infty$.

(a) If $k \geq r$ then $\hat{\sigma}^2_\epsilon(k) \xrightarrow{p} \sigma^2_\epsilon$.

(b) Let $\hat{r}$ be the estimate of r produced by an information criterion with $g(T)\to 0$ and $\delta_{NT}g(T)\to\infty$, where $k\geq r$. Then $\Pr(\hat{r}=r)\to 1$ and $\hat{\sigma}^2_\epsilon(\hat{r}) \xrightarrow{p} \sigma^2_\epsilon$.

In (2.2), the efficient forecast of y_{t+1} given past (y_t, X_t) (and given $\beta_t=\beta$) is $\beta' F_t^0$. Theorem 2 implies that this efficient forecast can be achieved (in a mean-square sense) asymptotically even if the factors, and indeed the number of factors, are unknown. Theorem 2(a) states that forecast efficiency can be achieved even if "too many" factors are estimated, and that the overestimation introduces no additional error asymptotically. In practice, however, one might be concerned about the effect of estimating more coefficients than are needed, so it might be desirable to use an information criterion to reduce the number of factors. Theorem 2(b) provides conditions under which doing so produces an efficient forecast and moreover provides a consistent estimate of the number of factors.

The conditions on g(T) in theorem 2 differ from the usual conditions to justify information criteria. With observable regressors, model selection by information criteria in the stationary case generally is consistent if $Tg(T)\rightarrow\infty$ and $g(T)\rightarrow 0$, which are satisfied by the BIC but not the AIC. However, neither the AIC nor the BIC satisfy $\delta_{NT} g(T)\rightarrow\infty$. A penalty function which does satisfy this condition is,

$$g(T) = \omega \ln T/\delta_{NT} \tag{3.2}$$

where δ_{NT} is given in theorem 1, that is, $\delta_{NT} = \min(N^{1/2}/T^{1+\epsilon}, T^{1-\epsilon})$, where ϵ is a small positive constant, and where ω is a positive constant. If for example $N=T^3$ (which satisfies condition R), then $\delta_{NT} = T^{1/2-\epsilon}$, so that $g(T)=\omega \ln T/T^{1/2-\epsilon}$, which is a larger penalty than the BIC asymptotically. The constant ω is indeterminant in the theorem so a suitable choice of ω in practice is one topic to be investigated in the Monte Carlo study in the next section.[5]

4. Monte Carlo Analysis

A Monte Carlo experiment was performed to study the finite sample performance of this factor extraction procedure and its application to forecasting. This experiment has three objectives. The first is to verify numerically the predictions of theorem 1, which in particular entails ascertaining the extent to which the estimated factors are close to the true factors in finite samples for various values of T, N, r and k. The second objective is to quantify the increase in forecast error that arises from using the estimated factors rather than the true factors, assuming that r is known. The final objective is to quantify the additional forecasting error introduced when the true number of factors is unknown so the number of factors is selected by an information criterion, as studied in theorem 2.

The experimental design is the parametric dynamic factor model that, in its most general form, allows for time varying factor loadings, an autoregressive factor, and idiosyncratic terms that are serially correlated and correlated across series. All the results here are for a balanced panel. The design is,

(4.1) $$X_{it} = \Sigma_{j=0}^{q} \lambda_{ijt}' F_{t-j} + e_{it}$$

(4.2) $$F_t = \alpha F_{t-1} + u_t$$

(4.3) $$(1-aL)e_{it} = (1+b^2)v_{it} + bv_{i+1,t} + bv_{i-1,t}$$

(4.4) $$\lambda_{ijt} = \lambda_{ijt-1} + (h/T)\zeta_{ijt}$$

where $i=1,\ldots,N$ and $t=1,\ldots,T$, F_t and λ_{ijt} are $r\times 1$, $\{e_{it}, v_{it}, \zeta_{ijt}\}$ are i.i.d. N(0,1), u_t is i.i.d. $N(0,I_r)$, and $\{u_t\}$ is independent of $\{e_{it}, v_{it}, \zeta_{ijt}\}$. As discussed in Section 2.3, because this is a dynamic factor model, in the static form the true number of factors is $(q+1)r$. The time variation here is a special case of the heterogeneous time variation allowed in the theoretical work in which $h_{iT} = h/T$ with probability one.

The initial factor loading matrix Λ_0 is chosen as follows. Let $R_i^2 = \mathrm{var}(\Sigma_{j=0}^{q}\lambda_{ij0}'F_{t-j})/[\mathrm{var}(\Sigma_{j=0}^{q}\lambda_{ij0}'F_{t-j}) + \mathrm{var}(e_{it})]$. Then $\lambda_{ij0} = \lambda_i^*\tilde{\lambda}_{ij0}$, where $\tilde{\lambda}_{ij0}$ is i.i.d. N(0,1) and independent of $\{e_{it}, v_{it}, \zeta_{ijt}, v_t\}$, and λ_i^* is chosen so that R_i^2 has a uniform distribution on [0.1,0.8].[6] The initial values of the factor F_0 are drawn from the stationary distribution of F_t. Finally the $\{X_{it}\}$ are transformed to have sample mean zero and sample variance one (this transformation is used in the empirical work presented in the next section).

The scalar variable to be forecast obeys,

$$y_{t+1} = \Sigma_{j=0}^{q}\iota' F_{t-j} + \epsilon_{t+1} \tag{4.5}$$

where ι is a $r\times1$ vector of 1's and ϵ_{t+1} is i.i.d. N(0,1).

The factors were estimated as discussed in section 2 for the balanced panel using $\{X_{it}\}$, i=1,...,N, t=1,...,T. Estimates were based on the static framework, that is, an augmented X constructed by stacking X and its lags as discussed in section 2.5 was not used. The coefficients β in the forecasting regression were estimated by the OLS coefficients $\hat{\beta}$ in the regression of y_{t+1} on $\hat{F}_t$, t=2,...,T; in particular no lags of $\hat{F}_t$ were introduced into the forecasting equation. The out-of-sample forecast $\hat{y}_{T+1}$ was constructed as $\hat{y}_{T+1} = \hat{\beta}'\hat{F}_T$. For comparison purposes, the infeasible out-of-sample forecast $\hat{y}_{T+1}^0 = \hat{\beta}^{0\prime}F_t^0$ was also computed, where $\hat{\beta}^0$ are the OLS coefficients in the regression of y_{t+1} on F_t^0, t=2,...,T.

The free parameters to be varied in the Monte Carlo experiment are N, T, α, a, b, and h. The results are summarized by two statistics. The first is a trace R^2 of the multivariate regression of $\hat{F}$ on F_0:

$$R^2_{\hat{F},F_0} = \hat{E}\|P_{F0}\hat{F}\|^2/\hat{E}\|\hat{F}\|^2 = \hat{E}\mathrm{tr}(\hat{F}'P_{F0}\hat{F})/\hat{E}\mathrm{tr}(\hat{F}'\hat{F}), \tag{4.6}$$

where $\hat{E}$ denotes the expectation estimated by averaging the relevant statistic over the Monte Carlo repetitions. According to theorem 1, if $k \geq (q+1)r$ then $R^2_{\hat{F},F^0} \xrightarrow{p} 1$. Values of this statistic considerably less than one indicates a case in which theorem 1 provides a poor approximation to the finite sample performance of $\hat{F}$.

The second statistic measures how close the forecast based on the estimated factors is to the infeasible forecast based on the true factors:

$$(4.7) \qquad S^2_{\hat{y},\hat{y}^0} = 1 - \hat{E}(\hat{y}_{T+1} - \hat{y}^0_{T+1})^2 / \hat{E}(\hat{y}^0_{T+1})^2$$

According to theorem 2, $S^2_{\hat{y},\hat{y}^0} \xrightarrow{p} 1$ either if $k = (q+1)r$, or if $k \geq (q+1)r$ and the factors included in the forecasting regression are chosen using an information criterion that satisfies condition IC. Accordingly, results are reported for several information criteria: the AIC, the BIC, and the information criterion with the penalty function (3.2) for various choices of the scaling parameter ω.

The results are summarized in table 1. Panel A presents results for the static factor model with i.i.d. errors and factors. In panel B, this model is extended to idiosyncratic errors that are serially correlated across series. Panel C considers the dynamic factor model with serially correlated factors and lags of the factors entering X_t, and time varying factor loadings are introduced in panel D.

First consider the results for $R^2_{\hat{F},F^0}$, which checks the consistency predicted by theorem 1. In all cases, $R^2_{\hat{F},F^0}$ exceeds .8, even for T=25 and N=50. As T and N increase, this R^2 increases, for example, for T=100, N=250, r=k=5, $R^2_{\hat{F},F^0}$=.97. As predicted by the theorem, estimating k>r typically introduces little spurious noise, for example, when T=100, N=250, and r=5, increasing k from 5 to 10 decreases $R^2_{\hat{F},F^0}$ by .02. If the idiosyncratic errors are moderately serially correlated (a=.5), $R^2_{\hat{F},F_0}$ drops only slightly, although it drops further when a=.9 (although this drop is largely eliminated when T is increased). The $R^2_{\hat{F},F_0}$ is also high when the true model is dynamic but the factors are extracted from a static procedure with $k \geq r(q+1)$, although some deterioration is noticeable when the factors are highly serially correlated. The greatest deterioration of the estimates of the factors occurs when time variation in the factor weights is introduced. With large time variation (h=10), $R^2_{\hat{F},F_0}$ is between .83 and .87 for the various cases considered. In general, the results improve when T increases, with N, r, and k fixed, and when N increases, with T, r, and k fixed; for fixed T and N, results deteriorate as r increases and k=r, although they deteriorate only slightly as k increases for fixed r.

Next turn to the results for $S^2_{\hat{y},\hat{y}_0}$, which check the consistency predictions of theorem 2. The results for q=r essentially parallel the results for $R^2_{\hat{F},F_0}$, although the range of $S^2_{\hat{y},\hat{y}_0}$ values exceeds the range of $R^2_{\hat{F},F_0}$. When T=100 and N=250, $S^2_{\hat{y},\hat{y}_0}$ is generally large, typically exceeding .95 in the static models. The quality of the forecasts drops in the dynamic models and when there is time variation in the factor loadings.

The results for forecasts based on the model selection criterion are generally consistent with theorem 2. Generally speaking, for T and N large, forecasts based on the BIC, AIC, or (3.2) with a=.001 perform similarly, and only slightly worse than those with q=r. However, the forecasts based on (3.2) with larger values of ω such as ω=.01 perform poorly, and when ω is further increased the forecasts deteriorate even further (these results are not shown to save space). This suggests that criteria that satisfy condition IC are unduly conservative, a possibility discussed in section 3 and a consequence of condition IC being only a sufficient condition for theorem 2.

5. Application to Forecasting U.S. Industrial Production and Inflation

This section reports the results of a simulated real-time forecasting exercise, in which forecasts based on the diffusion index approach are compared to forecasts from a variety of benchmark models.

This exercise focuses on forecasting two macroeconomic variables for the United States: real economic growth, as measured by the twelve-month growth of the index of industrial production (total) (IP), and inflation, as measured by the twelve-month growth of the consumer price index (urban, all items) (CPI). Specifically, let z_t denote either IP or the CPI in month t. In the notation of section 2, the variable to be forecast is $y_{t+1} = \ln(z_{t+12}/z_t)$. The complete data set spans 1959:1 - 1997:9.

5.1 Models

For both IP and the CPI, the diffusion index (DI) forecasts were compared to benchmark forecasts from an autoregressive model and multivariate regression-based forecasts using various leading indicators. For the CPI, as an additional comparison forecasts were also computed using models based on the Phillips curve. These various forecasting models are now described in turn.

Diffusion Index forecasts. Diffusion index forecasts were computed as outlined in section 2. Two sets of variables were used: a balanced panel of 170 monthly macroeconomic time series, 1960:1 - 1997:9, and an unbalanced panel in which these 170 series were augmented by 54 monthly series which are available for only part of this period, so that the total number of series in the unbalanced panel is 224. 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 full list is given in Appendix B. This list is similar to lists which we have used elsewhere (Stock and Watson [1996, 1998]). These series were taken from a somewhat longer list, which was scanned visually to eliminate gross problems such as series redefinitions. However no further pruning of this list was performed based on forecast performance measures. The series were taken from the February 1998 release of the DRI/McGraw Hill Basic Economics database (formerly Citibase). In general these series represent the fully revised historical series available as of February 1998.

All series on this list were subjected to two preliminary steps: possible transformation by taking logarithms, and possible first differencing. The decision to take logarithms or to first difference the series was judgmentally made. In general, logarithms were taken for all nonnegative series that were not already in rates or percentage units. In general, first differences were taken of real and nominal quantity series and of price indexes. 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.

The factors were estimated using only contemporaneous values of X_t (no stacking of lagged values of X_t). The factors were computed using the algorithms described in section 2. A total of $k=12$ factors were estimated.

In general the error term ϵ_t in (2.2) can be serially correlated. This suggests considering a variant of (2.2) in which lagged values of the dependent variable also appear as predictors. We therefore consider diffusion index forecasts of the form,

$$(5.1) \qquad \ln(z_{t+12}/z_t) = \beta_0 + \Sigma_{i=1}^{q}\beta_i\hat{F}_{it} + \Sigma_{j=0}^{p}\gamma_j\Delta\ln z_{t-j} + \tilde{\epsilon}_t$$

where $\{\hat{F}_{it}\}$ are the estimated factors. Given q and p, the coefficients of (5.1) were estimated by OLS. Four variants of (5.1) are reported, each with different treatment of q: (i) the number of factors recursively selected by BIC, $1 \leq q \leq 12$, and no autoregressive components (p=0); (ii) the number of factors recursively selected by BIC, $1 \leq q \leq 12$, and autoregressive components recursively estimated by BIC ($0 \leq p \leq 5$); (iii) a fixed number of factors and p=0; and (iv) a fixed number of factors and p selected by BIC ($0 \leq p \leq 5$).

Autoregressive forecast. The autoregressive forecast is a univariate forecast based on the model,

$$(5.2) \qquad \ln(z_{t+12}/z_t) = \mu + \Sigma_{j=0}^{p} \Delta \ln z_t + \tilde{\epsilon}_t,$$

where p is selected recursively by BIC ($0 \leq p \leq 5$).

Multivariate leading indicator forecasts. The multivariate leading indicator forecasts are of the form,

$$(5.3) \qquad \ln(z_{t+12}/z_t) = \beta_0 + \Sigma_{j=0}^{q} \Sigma_{i=1}^{m} \delta_{ij} w_{i,t-j} + \Sigma_{j=0}^{p} \gamma_j \Delta \ln z_{t-j} + \tilde{\epsilon}_t$$

where $\{w_{it}\}$ are various leading indicators that have been used elsewhere to forecast these variables.

For the IP forecasts, the set of eleven leading indicators used here are those that we have used in real time forecasting using experimental coincident, leading and recession indicators (see Stock and Watson [1989, 1991])[7]. Five of these leading indicators are also used in the factor estimation step in the the diffusion index forecasts. These are (the mnemonics under which they appear in Appendix B appear in parentheses): average weekly hours of production workers in manufacturing (lphrm); the capacity utilization rate in manufacturing (ipxmca); housing starts (building permits) (hsbp); 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 Assiation of Purchasing Managers' index of vendor performance (the percent of companies reporting slower deliveries).

For the CPI 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). Five of these variables are also used in the factor estimation step in the diffusion index forecasts: the total unemployment rate (lhur); housing starts (hsbp); new orders in durable goods industries (mdoq); the nominal M1 money supply (fm1); and the federal funds overnight interest rate (fyff). The remaining three variables are: real manufacturing and trade sales; the interest rate spread between 1-year U.S. Treasury bonds and the federal funds rate; and the trade-weighted exchange rate listed in the previous paragraph.

In all cases, the leading indicators were transformed to be approximately stationary. 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, two leading indicator forecasts were produced. To be concrete, consider forecasts of IP. The first forecast uses all eleven leading indicators (so m=11 in (5.3)), with p and q selected by recursive BIC, with $0 \leq p \leq 5$ and $0 \leq q \leq 3$. The second forecast uses recursive model selection, so that the forecast at a given date is based on a subset of the eleven leading indicators. In this case, at each date m, p and q are selected by recursive BIC, where $1 \leq m \leq 11$, $p=3$, and $q=\{1,3\}$. This entails comparing 4094 different models at any given date. This latter approach is closest to classical model selection theory, in which all models are enumerated and compared using an information criterion.

Phillips curve forecasts. The expectations-augmented Phillips curve constitutes an important tool of empirical macroeconomics and is considered by many to be a reliable tool for forecasting inflation, cf Gordon (1982) and, more recently, the Congressional Budget Office (1996), Fuhrer (1995), Gordon (1997), Staiger, Stock and Watson (1997), and Tootel (1994). For this reason, forecasts based on two variants of a Phillips curve are also included for comparison purposes. These specified the twelve-month inflation rate as the dependent variable:

$$(5.4) \qquad p_{t+12}-p_t = \beta_0 + \Sigma_{j=0}^{q}\beta_j u_{t-j} + \Sigma_{j=0}^{p}\gamma_j \Delta p_{t-j} + \delta' Z_t + \tilde{\epsilon}_t$$

where $p_t = \ln(CPI_t)$, $\pi_t = 1200^*\Delta p_t$ is monthly CPI inflation at an annual rate, u_t is the unemployment rate, and Z_t is a vector of variables that control for supply shocks and/or measurement difficulties. The two variants differ in the supply shock variables Z_t. In one, z_t consists solely of the relative price of food and energy; in the other, this relative price is augmented by Gordon's (1982) variable that controls for the imposition and removal of the Nixon wage and price controls.[8]

The parameters of (5.4) were estimated recursively by ordinary least squares. For each of the two variants, the lag lengths q and p were chosen by recursive BIC, where $0 \leq q \leq 5$ and $0 \leq p \leq 5$.

5.2. Estimation

The estimation and forecasting was carried out in a way to simulate real-time forecasting. This entailed fully recursive parameter estimation, factor extraction, model selection, etc. For example, to construct the first forecast, the parameters and factors were estimated, and the models were selected, using 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). These parameters and models were then used to forecast IP growth and CPI inflation from 1970:1 to 1971:1. All parameters, factors, etc. were then reestimated, and information criteria were recomputed, using data from 1959:1 through 1970:2, and forecasts using these models and parameters were computed for twelve-month growth from 1970:2 to 1971:2. Because all order and model selection is fully recursive, this means that the actual model used to produce the forecasts for a method that uses an information criterion in general changes from one month to the next; what is constant is the rule by which that model is selected.

5.3 Results

Forecasting results. The results of the simulated out of sample forecasting experiments are reported in table 2 for IP and in table 3 for CPI inflation. The entries are the mean squared error (MSE) of the candidate forecasting model, computed relative to the MSE of the autoregressive forecast (so the autoregressive forecast, which is unreported, has a relative MSE of 1.00). Smaller relative MSEs signify more accurate forecasts.

First consider the results for IP. In the table, "DI" denotes the static diffusion index forecasts (p=0 in (5.1)), and "DIAR" denotes the diffusion index forecasts augmented with lagged monthly IP growth (p selected by recursive BIC in (5.1)). The diffusion index forecasts with BIC factor selection represent substantial improvements over the leading indicator multivariate forecasts. The performance of the diffusion index forecasts is similar whether or not lags of industrial production growth are included as predictors. This is rather surprising, because it implies that essentially all the predictable dynamics of industrial production growth are accounted for by the estimated factors. The results for the diffusion index models indicate that, among forecasts with fixed numbers of factors, almost all the improvement is obtained after merely two factors are added; indeed, the MSE actually increases upon the addition of the fourth factor (this is possible because the MSE is for pseudo out of sample forecasts). It is also noteworthy that the BIC-selected forecasts outperform any of the fixed-k forecasts. This suggests that the number of factors useful for forecasting IP evolves over time, and that this time variation is picked up by the recursive BIC procedure. The results for the unbalanced and balanced panels are generally similar. The BIC-selected DI RMSEs are the same for the two panels. However, the fixed-k forecasts are generally better for the larger unbalanced panel.

The diffusion index forecasts of the CPI (table 3) also represent substantial improvements over the benchmark models. Unlike the IP forecasts, the estimated factors do not account for all of the predictable dynamics in CPI inflation, and adding lags of CPI inflation to the diffusion index forecasts improves their performance, both for fixed k and k selected by recursive BIC. The results for fixed numbers of factors and the autoregressive correction indicate that the MSE attains a minimum at five or six factors. In contrast to the case of IP, the best fixed-k forecast is considerably better than the BIC-selected forecast, in both the balanced and unbalanced panel (the relative RMSEs are .62 v. .71 for the balanced panel with the autoregressive terms, respectively).

It is interesting to note that, in constrast to the results for IP, the leading indicator forecast based on a recursively BIC-selected subset of the leading indicators is considerably worse than using all leading indicators and in fact is worse than just using an autoregressive forecast. This is consistent with the view that the correlations between the individual leading indicators and inflation are unstable over time, so that variables selected on the basis of good prior performance become unreliable and thus produce poor out of sample forecasts. It is also noteworthy that the Phillips curve model without the wage and price control variable performs almost as well as the leading indicator forecast, although not nearly as well as the autoregressive-augmented diffusion index forecasts. When the wage-price variable is added, simulated real time performance of the Phillips curve forecast actually deteriorates and barely improves upon the autoregressive forecast.

It should be stressed that the multivariate leading indicator models are sophisticated forecasting models that provide a stiff benchmark against which to judge the diffusion index forecasts. In fact, the performance of the leading indicator models in table 2 arguably overstates their out of sample potential performance, because the lists of leading indicators used to construct the forecasts were chosen by model selection methods using data similar to these, cf. Stock and Watson (1989) and Staiger, Stock and Watson (1997). In this light, we consider the performance of the diffusion index models to be encouraging.

Estimated factors. The previous results suggest that it is of interest to examine the first few estimated factors. The interpretation of the factors is simplest for the case of the balanced panel, for here the factors are the eigenvectors of $N^{-1}\sum_{i=1}^{N}\underline{X}_i\underline{X}_i'$, ordered by the magnitude of the associated eigenvalues. We therefore focus here on the full-sample estimates of the factors using the balanced panel. Only this ordering of the factors is considered. Because the estimates here are for the full sample, these factors in general differ from those used in the recursive out-of-sample forecasting exercise.

Figure 1 displays the R^2s of the regressions of the 170 individual time series in the balanced panel against each of the six factors, 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 inflation and interest rate spreads; the third, on unemployment; the fourth, on housing starts and orders; the fifth, on stock returns and money growth; and the sixth, on new orders. In this sense, the first factor may be thought of as an output factor, the second as an inflation factor, the third as an unemployment factor, etc. Taken together, these six factors account for 47% of the variance of the 170 monthly time series in the balanced panel, as measured by the trace-R^2; the first twelve factors together account for 61% of the variance of these series.[9]

The factors are plotted in figure 2, each along with an individual transformed series suggested by the variance decompositions in figure 1. For example, the first factor is plotted with the monthly growth rate of industrial production, both transformed to have unit standard deviation. Interestingly, the factors generally contain considerable high frequency power, for example, factor #3 has low frequency behavior similar to the unemployment rate, but has much more pronounced monthly fluctuations.

6. Discussion and Conclusions

We find several features of the empirical results surprising and intriguing. Few theoretical macroeconomic models that suggest a linear factor structure for the overall macroeconomy, yet we find that six factors account for almost one-half of the variance of the 170 time series in our balanced panel and twelve factors account for almost two-thirds of this variance. Even if a factor structure describes the joint behavior of these series, there is no reason why a forecast based on static factor estimates should outperform forecasts based on leading indicators or other specialized models that have been fine tuned through years of experience. Yet, forecasts based on just the first six factors perform well for *both* CPI inflation and industrial production growth, series that measure quite different economic concepts (nominal prices, real output) and have quite different univariate time series properties. Thus, these results raise numerous issues for future empirical and theoretical research.

One such issue, only touched on in section 5, is the interpretation of the estimated factors. A feature of traditional diffusion indexes is that they were constructed to have a ready interpretation, such as a measure of how widespread employment growth is across sectors in the economy. Like traditional diffusion indexes, our estimated dynamic factors are averages of many different economic series, but they are identified only up to a nonsingular $k \times k$ transformation. Thus the estimated factors will not in general have the natural interpretation that is a feature of traditional diffusion indexes. This raises the question of how to transform the factors into interpretable diffusion indexes. For work on this topic, see Quah and Sargent (1993) and Forni and Reichlin (1996, 1997, 1998).

Several methodological issues remain. One is to explore estimation methods that might be more efficient in the presence of heteroskedastic and serially correlated uniquenesses. Another is to develop a distribution theory for the estimated factors that goes beyond the consistency results shown here and provides measures of the sampling uncertainty of the estimated factors. A third theoretical extension is to move beyond the I(0) framework of this paper and to introduce persistence into the series, for example by letting some of the factors have a unit autoregressive root which would permit some of the observed series might contain a common stochastic trend.

Another important extension is to real time forecasting with mixed frequency data (weekly, monthly and quarterly). The EM algorithm presented for the unbalanced panel can be extended to panels with mixed periodicities, albeit with some computational complications. Other issues that arise in real time include data revisions and the nonsynchronous timing of data releases. Work on these and related issues is ongoing.

Footnotes

1. It is our understanding (L. Reichlin, personal communication) that a working paper in progress provides proofs of the consistency of the estimated dynamic factors obtained using an alternative method based on dynamic principal components.

2. A body of work applying break tests suggests that the 1/T nesting is empirically plausible for many macroeconomic time series, cf. Stock and Watson (1996, 1998). The 1/T nesting is also the local alternative against which break tests such as the Quandt likelihood ratio test would have nondegenerate asymptotic power were F_t observed.

3. The bounded support condition M(d)(i) is not satisfied if f_t follows a Gaussian vector autoregression. However this assumption is made to simplify the proof of the theorem 1 and arguably is of a technical rather than substantive nature, because $\bar{F}$ can be taken to be quit large and d can be taken to be very close to zero.

4. When $k_T > r_T$, $\hat{F}$ asymptotically has reduced column rank r_T even though $T^{1/2}\tilde{F}$ has orthonormal columns by construction. The source of the difference between $\tilde{F}$ and $\hat{F}$ is that (in a balanced panel) $\tilde{F}$ are the first k_T eigenvectors of $N^{-1}\Sigma_{i=1}^{N}\underline{X}_i\underline{X}_i'$. Asymptotically, the smallest (k_T-r_T) eigenvalues of this matrix are zero, so the columns of $T^{-1}\Sigma_{s=1}^{T}X_s\tilde{F}_s'$ corresponding to these Fs are themselves nearly zero, and in turn the correpsonding columns of $\hat{\Lambda}$, and thus of $\hat{F}_t = \hat{\Lambda}'X_t/N$, are nearly zero.

5. It should be pointed out that theorem 2 simply states sufficient conditions, so the BIC, while not satisfying the conditions of the theorem, nevertheless might provide efficient forecasts.

6. That is, $\lambda_i^* = [R_i^2(R_i^2\text{-}1)^{-1}(\text{var}(\Sigma_{j=0}^{q}\tilde{\lambda}_{ij0}'F_{t-j}))^{-1}]^{1/2}$ (this uses var(e_{it})=1), where R_i^2 is i.i.d. U(0.1,0.8).

7. The list used here consists of the leading indicators used to produce the XRI and the XRI-2, which are released monthly at the web site http://www.nber.org. Additional documentation is available at that site.

8. Most modern specifications of the Phillips curve treat prices as I(2) (cf. Gordon [1982, 1997]), but in (5.4), prices are treated as I(1). An I(2) specification is achieved in (5.4) by imposing $\Sigma_{j=0}^{p}\gamma_j=1$. Forecasts were also computed for this specification. The simulated real

time forecasts for the I(2) specification were slightly worse than those for the I(1) specification, so only results for the better performing I(1) specification are reported below.

9. The contributions to the trace-R^2 by the first six factors are, respectively: 0.156, 0.115, 0.077, 0.046, 0.042, and 0.035, for a total of 0.471.

Appendix A: Proofs of Theorems

The proof of theorem 1 makes use of the following lemma.

Lemma A1. Under conditions R, P and M,

(a) $\delta_{NT}k_T\sup_{s,t}|F_s^{0\prime}\Lambda_0^{\prime}e_t/N| \xrightarrow{p} 0$;

(b) $\delta_{NT}k_T\sup_{s,t}|F_s^{0\prime}(\Lambda_s-\Lambda_0)^{\prime}e_t/N| \xrightarrow{p} 0$;

(c) $\delta_{NT}k_T\sup_{s,t}|e_s^{\prime}e_t/N - \gamma(s-t)| \xrightarrow{p} 0$;

(d) $\delta_{NT}k_T\sup_{s,t}|F_s^{0\prime}(\Lambda_s^{\prime}\Lambda_t/N - \Lambda_0^{\prime}\Lambda_0/N)F_t^0| \xrightarrow{p} 0$;

(e) $\delta_{NT}k_T\sup_{s,t}|X_s^{\prime}X_t/N - F_s^{0\prime}(\Lambda_0^{\prime}\Lambda_0/N)F_t^0 - \gamma(s-t)| \xrightarrow{p} 0$;

(f) $\sup_{i,\ell}\{E\sup_F\Delta_{i\ell}^2(F)\}^{1/2} \leq T^{-1/2}r_T(T\kappa_{2T})\bar{F}^2\{\sup_{i,m}\sum_{u=-\infty}^{u}|\Gamma_{ii,mm}(u)|\}^{1/2}$,

where $\Delta_i(F)=T^{-1}\sum_{t=1}^{T}F_tF_t^{0\prime}(\lambda_{it}-\lambda_{i0})$;

(g) $\sup_F|(NT)^{-1}\sum_{i=1}^{N}\underline{e}_i^{\prime}P_F\underline{e}_i| \xrightarrow{p} 0$;

(h) $\sup_F|(NT)^{-1}\sum_{i=1}^{N}\underline{e}_i^{\prime}P_FF^0\lambda_{i0}| \xrightarrow{p} 0$;

(i) $\sup_F|N^{-1}\sum_{i=1}^{N}(\underline{e}_i^{\prime}F/T)(F^{\prime}F/T)^{-1}\Delta_i| \xrightarrow{p} 0$;

(j) $\sup_F|N^{-1}\sum_{i=1}^{N}\lambda_{i0}^{\prime}(F^{0\prime}F/T)(F^{\prime}F/T)^{-1}\Delta_i| \xrightarrow{p} 0$;

(k) $\sup_F|N^{-1}\sum_{i=1}^{N}\Delta_i^{\prime}(F^{\prime}F/T)^{-1}\Delta_i| \xrightarrow{p} 0$.

Proof of Lemma A1

The proof uses the following results. Let μ_t and $\nu_{t,s}$ be random matrices indexed by $t=1,\ldots,T$, $s=1,\ldots,T$:

(A.1a) If $T\sup_t E(\|\mu_t\|^q) \to 0$ for $q \geq 1$, then $\sup_t\|\mu_t\| \xrightarrow{p} 0$;

(A.1b) If $T^2\sup_{s,t}E(\|\nu_{s,t}\|^q) \to 0$ for $q \geq 1$, then $\sup_{s,t}\|\nu_{s,t}\| \xrightarrow{p} 0$; and

(A.1c) If both $\sup_{s,t}\|E\nu_{s,t}\| \to 0$ and $T^2\sup_{s,t}E(\|\nu_{s,t}-E\nu_{s,t}\|^2)\to 0$, then $\sup_{s,t}\|\nu_{s,t}\| \xrightarrow{p} 0$.

Also note that condition R implies the following limits, which in turn imply limits used in the proof: $k_T^2/\delta_{NT} \to 0$, $\delta_{NT}k_T^6r_T^2/T \to 0$, and $\delta_{NT}^2T^2k_T^2r_T^4/N \to 0$.

(a) Let $\nu_{s,t}=\delta_{NT}k_TF_s^{0\prime}\Lambda_0'e_t/N$ and use (A.1b) with q=2. Then

$$T^2\sup_{s,t}E\nu_{s,t}^2 = T^2\delta_{NT}^2k_T^2E[F_s^{0\prime}\Lambda_0'e_t/N]^2$$
$$\leq \delta_{NT}^2k_T^2E(F_s^{0\prime}F_s^0)E(e_t'\Lambda_0\Lambda_0'e_t/N^2).$$

Now $E(F_s^{0\prime}F_s^0) \leq r_T\bar{F}^2$ and $E(e_t'\Lambda_0\Lambda_0'e_t/N^2) \leq (r_T\bar{\lambda}^2/N)N^{-1}\Sigma_{i=1}^N\Sigma_{j=1}^N|\tau_{ij}|$, so

$$T^2\sup_{s,t}E\nu_{t,s}^2 \leq (\delta_{NT}^2T^2k_T^2r_T^2/N)\bar{F}^2\bar{\lambda}^2N^{-1}\Sigma_{i=1}^N\Sigma_{j=1}^N|\tau_{ij}| \to 0$$

by condition M(a)(ii) and lemma A1.

(b) Let $\nu_{s,t} = \delta_{NT}k_Tr_T^{1/2}\bar{F}(\Lambda_s-\Lambda_0)'e_t/N$ and note that, with probability one,

$$|\delta_{NT}k_TF_s^{0\prime}(\Lambda_s-\Lambda_0)'e_t/N| \leq \delta_{NT}k_T\|F_s^0\|\,\|(\Lambda_s-\Lambda_0)'e_t/N\| \leq \|\nu_{s,t}\|.$$

Thus the result follows if $\sup_{s,t}\|\nu_{s,t}\| \xrightarrow{p} 0$, which is shown using (A.1c). Now $Ee_t'(\Lambda_s-\Lambda_0)/N = (h/NT)\Sigma_{i=1}^N\Sigma_{r=1}^sE\zeta_{ir}e_{it} = \kappa_{1T}N^{-1}\Sigma_{i=1}^N\Sigma_{r=1}^s\Psi_{ii}(t-r)$, so

$$\sup_{s,t}\|E\nu_{s,t}\| = \sup_{s,t}\delta_{NT}k_Tr_T^{1/2}\bar{F}\|(h/NT)\Sigma_{i=1}^N\Sigma_{r=1}^s\Psi_{ii}(t-r)\|$$
$$\leq (\delta_{NT}k_Tr_T/T)\bar{F}(T\kappa_{1T})\sup_i\Sigma_{u=-\infty}^{\infty}r_T^{-1/2}\|\Psi_{ii}(u)\|$$

which tends to zero by conditions R and M(c)(i). In addition,

$$T^2 \sup_{s,t} E\|\nu_{s,t} - E\nu_{s,t}\|^2 = T^2 \sup_{s,t} \delta_{NT}^2 k_T^2 r_T \bar{F}^2 \Sigma_{m=1}^{r_T} \mathrm{var}[N^{-1}\Sigma_{i=1}^{N} e_{it}(\lambda_{is,m} - \lambda_{i0,m})]$$
$$\leq (\delta_{NT}^2 k_T^2 r_T^2 T^2/N)\bar{F}^2 (T\kappa_{2T})^2 \{(\sup_{i,t} Ee_{it}^4)^{1/2}(\sup_{i,m,s,} E\zeta_{is,m}^4)^{1/2}$$
$$+ N^{-1}\Sigma_{i=1}^{N}\Sigma_{j=1}^{N} \sup_{t,s,s'} r_T^{-1} |\mathrm{tr}[\mathrm{cov}(e_{it}\zeta_{is}, e_{jt}\zeta_{js'})]|\}$$

which tends to zero by conditions R, M(a)(iii), M(b)(iii), and M(c)(ii).

(c) Let $\nu_{s,t} = \delta_{NT} k_T (e_s' e_t/N - \gamma(t-s))$ and use (A.1b) with q=2. Now

$$T^2 \sup_{s,t} E\|\nu_{s,t}\|^2 = (\delta_{NT}^2 k_T^2 T^2/N)\sup_{s,t}|N^{-1}\Sigma_{i=1}^{N}\Sigma_{j=1}^{N} \mathrm{cov}(e_{is}e_{it}, e_{js}e_{jt})|$$

which tends to zero by conditions R and M(a)(iii).

(d) $\sup_{s,t}\|\delta_{NT} k_T F_s^{0\prime}[(\Lambda_s'\Lambda_t - \Lambda_0'\Lambda_0)/N]F_t^0\|$
$$\leq \sup_{s,t} \delta_{NT} k_T \|F_s^0\|^2 \|(\Lambda_s'\Lambda_t - \Lambda_0'\Lambda_0)/N]\|$$
$$\leq \bar{F}^2 \sup_{s,t}\|\nu_{s,t}\| + 2\bar{F}^2 \sup_t \|\mu_t\|$$

where $\nu_{s,t} = \delta_{NT} k_T r_T (\Lambda_s - \Lambda_0)'(\Lambda_t - \Lambda_0)/N$ and $\mu_t = \delta_{NT} k_T r_T (\Lambda_t - \Lambda_0)'\Lambda_0/N$.

First, show that $\sup_{s,t}\|\nu_{s,t}\| \xrightarrow{p} 0$ using (A.1c):

$$\sup_{s,t}\|E\nu_{s,t}\| = \sup_{s,t} \delta_{NT} k_T r_T \|N^{-1}\Sigma_{i=1}^{N} \kappa_{2T}^2 \Sigma_{r=1}^{s}\Sigma_{r'=1}^{t} E\zeta_{ir}\zeta_{ir'}'\|$$
$$\leq (\delta_{NT} k_T r_T^2/T)(T\kappa_{2T})^2 \sup_i \Sigma_{u=-\infty}^{\infty} \|\Gamma_{ii}(u)\|/r_T$$

which tends to zero by conditions R and M(b)(i). Also, by conditions R and M(b)(iii),

$$
\begin{aligned}
&T^2 \sup_{t,s} E\|\nu_{t,s} - E\nu_{t,s}\|^2 \\
&= \sup_{t,s} \delta_{NT}^2 k_T^2 r_T^2 T^2 E\|N^{-1}\Sigma_{i=1}^N \Sigma_{r=1}^s \Sigma_{r'=1}^t [h_{iT}^2 \zeta_{ir}\zeta_{ir'}' - E(h_{iT}^2 \zeta_{ir}\zeta_{ir'}')]\|^2 \\
&= \sup_{t,s} \delta_{NT}^2 k_T^2 r_T^2 T^2 \\
&\quad \times \Sigma_{\ell=1}^{rT} \Sigma_{m=1}^{rT} E\{N^{-1}\Sigma_{i=1}^N \Sigma_{r=1}^s \Sigma_{r'=1}^t [h_{iT}^2 \zeta_{ir,\ell}\zeta_{ir',m} - E(h_{iT}^2 \zeta_{ir,\ell}\zeta_{ir',m})]\}^2 \\
&\leq (\delta_{NT}^2 k_T^2 r_T^4 T^2/N)(T\kappa_{4T})^4 \{\sup_{i,s,m} E\zeta_{is,m}^4 \\
&\qquad + \sup_{\ell,m} N^{-1}\Sigma_{i=1}^N \Sigma_{j=1}^N \sup_{t,u_1,u_2,u_3} |\mathrm{cov}(\zeta_{it,\ell}\zeta_{it+u_1,m}, \zeta_{jt+u_2,\ell}\zeta_{jt+u_3,m})| < \infty \\
&\to 0,
\end{aligned}
$$

so that $\sup_{t,s}\|\nu_{t,s}\| \xrightarrow{p} 0$. Next, show that $\sup_t \|\mu_t\| \xrightarrow{p} 0$ using (A.1a) with q=2:

$$
\begin{aligned}
T\sup_t E\|\mu_t\|^2 &= T\sup_t \delta_{NT}^2 k_T^2 r_T^2 E\|(\Lambda_t - \Lambda_0)'\Lambda_0/N\|^2 \\
&\leq (\delta_{NT}^2 k_T^2 r_T^4/N)\bar{\lambda}^2 (T\kappa_{2T})^2 \sup_m N^{-1}\Sigma_{i=1}^N \Sigma_{j=1}^N \Sigma_{u=-\infty}^{\infty} |\Gamma_{ij,mm}(u)|
\end{aligned}
$$

which tends to zero by conditions R and M(b)(ii), so $\sup_t \|\mu_t\| \xrightarrow{p} 0$.

(e) Write $\delta_{NT}k_T[X_s'X_t/N - F_s^{0\prime}(\Lambda_0'\Lambda_0/N)F_t^0 - \gamma(s-t)] = \Sigma_{i=1}^4 \nu_{i,st}$, where

$$
\begin{aligned}
\nu_{1,st} &= \delta_{NT}k_T F_s^{0\prime}[(\Lambda_s'\Lambda_t - \Lambda_0'\Lambda_0)/N)F_t^0 \\
\nu_{2,st} &= \delta_{NT}k_T[e_s'e_t/N - \gamma(s-t)] \\
\nu_{3,st} &= 2\delta_{NT}k_T F_s^{0\prime}(\Lambda_s - \Lambda_0)'e_t/N \\
\nu_{4,st} &= 2\delta_{NT}k_T F_s^{0\prime}\Lambda_0'e_t/N.
\end{aligned}
$$

It was shown in parts (a)-(d) that $\sup_{s,t}\|\nu_{i,st}\| \xrightarrow{p} 0$, i=1,...,4, and the desired result follows.

(f) $E\sup_F \Delta_{i\ell}^2(F) = E\sup_F[T^{-1}\Sigma_{t=1}^T F_{\ell t}F_t^{0\prime}(\lambda_{it}-\lambda_{i0})]^2$

$$\le E\bar{F}^4 \Sigma_{\ell=1}^{rT}\Sigma_{m=1}^{rT} T^{-2}\Sigma_{t=1}^T \Sigma_{s=1}^T |(\lambda_{it}-\lambda_{i0})_\ell(\lambda_{is}-\lambda_{i0})_m|$$
$$\le \bar{F}^4 \Sigma_{\ell=1}^{rT}\Sigma_{m=1}^{rT}(\kappa_{2T}^2/T)\Sigma_{t=1}^T\Sigma_{s=1}^T \sup_i\sup_m \Sigma_{u=-\infty}^{\infty}|\Gamma_{ii,mm}(u)|$$
$$\le (r_T^2/T)(T\kappa_{2T})^2\bar{F}^4\sup_i\sup_m\Sigma_{u=-\infty}^{\infty}|\Gamma_{ii,mm}(u)|$$

so $\sup_i\{E\sup_F\Delta_{i\ell}^2(F)\}^{1/2} \le (r_T^2/T)^{1/2}(T\kappa_{2T})\bar{F}^2\{\sup_i\sup_m\Sigma_{u=-\infty}^{\infty}|\Gamma_{ii,mm}(u)|\}^{1/2} \to 0$ by condition M(b)(i).

(g) Let $(F'F/T)^{\ell m}$ denote the (ℓ,m) element of $(F'F/T)^{-1}$, and note that condition M(d) implies that $\sup_{\ell,m}(F'F/T)^{\ell m} \le d^{-1}$. Now

$$(NT)^{-1}\Sigma_{i=1}^N \underline{e}_i' P_F \underline{e}_i$$
$$= \delta_{NT}^{-1}\Sigma_{\ell=1}^{kT}\Sigma_{m=1}^{kT}(F'F/T)^{\ell m}T^{-2}\Sigma_{s=1}^T\Sigma_{t=1}^T F_{\ell t}F_{ms}\delta_{NT}(e_t'e_s/N - \gamma(s-t))$$
$$+ \Sigma_{\ell=1}^{kT}\Sigma_{m=1}^{kT}(F'F/T)^{\ell m}T^{-2}\Sigma_{s=1}^T\Sigma_{t=1}^T F_{\ell t}F_{ms}\gamma(s-t)$$

so, by result (c), condition R, and condition M(a)(i),

$$\sup_F (NT)^{-1}\Sigma_{i=1}^N \underline{e}_i' P_F \underline{e}_i$$
$$\le (k_T^2/\delta_{NT})d^{-1}\bar{F}^2\sup_{t,s}\|\delta_{NT}[e_t'e_s/N-\gamma(s-t)]\| + (k_T^2/T)d^{-1}\bar{F}^2\Sigma_{u=-\infty}^{\infty}|\gamma(u)| \xrightarrow{p} 0.$$

(h) Let $B=(F'F/T)^{-1}(F'F^0/T)$, and note that condition M(d) implies that $|B_{\ell m}| \le k_T\bar{F}^2 d^{-1}$. Now

$$\sup_F|(NT)^{-1}\Sigma_{i=1}^N \underline{e}_i'P_F F^0\lambda_{i0}| = \sup_F|\Sigma_{\ell=1}^{kT}\Sigma_{m=1}^{rT}B_{\ell m}[N^{-1}\Sigma_{i=1}^N\lambda_{i0,m}(\underline{e}_i'F/T)_\ell]|$$
$$\le r_T(k_T^3/\delta_{NT})^{1/2}\bar{F}^2 d^{-1}\bar{\lambda}\{\delta_{NT}k_T\sup_{\ell,F}N^{-1}\Sigma_{i=1}^N(\underline{e}_i'F/T)_\ell^2\}^{1/2}$$

Now

$$\delta_{NT}k_T\sup_{\ell,F}N^{-1}\Sigma_{i=1}^{N}(\underline{e}_i'F/T)_\ell^2$$
$$\leq \bar{F}^2\{\sup_{t,s}|\delta_{NT}k_T[e_t'e_s/N - \gamma(t-s)]| + (\delta_{NT}k_T/T)\Sigma_{u=-\infty}^{\infty}|\gamma(u)|\} \xrightarrow{p} 0$$

by result (c), condition R, and condition M(a)(i), and the result follows because $r_T(k_T^3/\delta_{NT})^{1/2}\to 0$.

(i) Now $\sup_F|N^{-1}\Sigma_{i=1}^{N}(\underline{e}_i'F/T)(F'F/T)^{-1}\Delta_i|$
$$\leq d^{-1}(k_T/\delta_{NT})^{1/2}\{k_T\delta_{NT}\sup_{\ell,F}N^{-1}\Sigma_{i=1}^{N}(\underline{e}_i'F/T)_\ell^2\}^{1/2}\sup_F\Sigma_{m=1}^{k_T}[N^{-1}\Sigma_{i=1}^{N}\Delta_{im}^2]^{1/2}.$$

By result (f),

$$E\sup_F\Sigma_{m=1}^{k_T}[N^{-1}\Sigma_{i=1}^{N}\Delta_{im}^2]^{1/2} \leq k_T\sup_{i,\ell}\{E\sup_F\Delta_{i,\ell}^2\}^{1/2}$$
$$\leq (k_Tr_T/T^{1/2})T\kappa_{2T}\bar{F}^2\{\sup_{i,m}\Sigma_{u=-\infty}^{\infty}|\Gamma_{ii,mm}(u)|\}^{1/2} \to 0$$

so $\sup_F\Sigma_{m=1}^{k_T}[N^{-1}\Sigma_{i=1}^{N}\Delta_{im}^2]^{1/2} \xrightarrow{p} 0$. Also, $\delta_{NT}k_T\sup_{\ell,F}N^{-1}\Sigma_{i=1}^{N}(\underline{e}_i'F/T)_\ell^2 \xrightarrow{p} 0$ by the result in the proof of (h). It follows from condition R that $\sup_F|N^{-1}\Sigma_{i=1}^{N}(\underline{e}_i'F/T)(F'F/T)^{-1}\Delta_i| \xrightarrow{p} 0$.

(j) Now $\sup_F|N^{-1}\Sigma_{i=1}^{N}\lambda_{i0}'(F^{0\prime}F/T)(F'F/T)^{-1}\Delta_i| \leq k_Tr_T\bar{F}^2d^{-1}\bar{\lambda}\Sigma_{m=1}^{k_T}(N^{-1}\Sigma_{i=1}^{N}\Delta_{im}^2)^{1/2}$, so, by result (f),

$$E\sup_F|N^{-1}\Sigma_{i=1}^{N}\lambda_{i0}'(F^{0\prime}F/T)(F'F/T)^{-1}\Delta_i| \leq k_T^2r_T\bar{F}^2d^{-1}\bar{\lambda}\{\sup_{i,m}E\sup_F\Delta_{im}^2\}^{1/2}$$
$$\leq (k_T^4r_T^4/T)^{1/2}\bar{F}^4d^{-1}\bar{\lambda}(T\kappa_{2T})\{\sup_i\sup_m\Sigma_{u=-\infty}^{\infty}|\Gamma_{ii,mm}(u)|\}^{1/2},$$

which converges to zero by by condition R; the desired result follows from the Markov inequality.

(k) Now
$$N^{-1}\Sigma_{i=1}^{N}\Delta_i'(F'F/T)^{-1}\Delta_i = tr[(F'F/T)^{-1}N^{-1}\Sigma_{i=1}^{N}\Delta_i\Delta_i']$$
$$\le maxeval[(F'F/T)^{-1}]tr[N^{-1}\Sigma_{i=1}^{N}\Delta_i\Delta_i']$$
$$\le c\Sigma_{\ell=1}^{kT}N^{-1}\Sigma_{i=1}^{N}\Delta_{i\ell}^2.$$

Thus $Esup_F|N^{-1}\Sigma_{i=1}^{N}\Delta_i'(F'F/T)^{-1}\Delta_i| \le ck_T sup_{i,\ell}Esup_F\Delta_{i\ell}^2(F)$, which converges to zero by result (f) and conditions R and TV; the desired result follows from the Markov inequality.

Proof of theorem 1

(a) The proof consists of two parts. First, it is shown that (minus) the objective function is asymptotically equivalent to $tr[T^{-1}F^{0\prime}P_F F^0 D]$, in the sense that the difference between these two objective functions converges to zero in probability uniformly in F. This permits an asymptotic characterization of $\tilde{F}$. Second, this is used to show the uniform consistency of $\hat{F}_t$.

First turn to the asymptotic objective function. Write $\hat{V}_{NT}(F) = (NT)^{-1}\Sigma_{i=1}^{N}\underline{X}_i'\underline{X}_i - \hat{Q}_{NT}(F)$, where $\hat{Q}_{NT}(F) = (NT)^{-1}\Sigma_{i=1}^{N}\underline{X}_i'P_F\underline{X}_i$. Thus minimization of $\hat{V}_{NT}(F)$ is identical to maximization of $\hat{Q}_{NT}(F)$. Let $\tilde{Q}_{NT}(F) = (NT)^{-1}\Sigma_{i=1}^{N}\lambda_{i0}'F^{0\prime}P_F F^0\lambda_{i0}$. Algebra reveals that $\hat{Q}_{NT}(F)-\tilde{Q}_{NT}(F) = \Sigma_{i=1}^{5}A_{iT}(F)$, where

$$A_{1T}(F) = N^{-1}\Sigma_{i=1}^{N}\Delta_i'(F'F/T)^{-1}\Delta_i$$
$$A_{2T}(F) = 2N^{-1}\Sigma_{i=1}^{N}\Delta_i'(F'F/T)^{-1}(F'F^0/T)\lambda_{i0}$$
$$A_{3T}(F) = 2N^{-1}\Sigma_{i=1}^{N}(\underline{e}_i'F/T)(F'F/T)^{-1}\Delta_i$$
$$A_{4T}(F) = 2N^{-1}\Sigma_{i=1}^{N}(\underline{e}_i'F/T)(F'F/T)^{-1}(F'F^0/T)\lambda_{i0}$$
$$A_{5T}(F) = N^{-1}\Sigma_{i=1}^{N}(\underline{e}_i'F/T)(F'F/T)^{-1}(F'\underline{e}_i/T)$$

By lemma A1(g)-(k), $sup_F|A_{iT}(F)| \xrightarrow{p} 0$, i=1,...,5. Thus $sup_F|\hat{Q}_{NT}(F)-\tilde{Q}_{NT}(F)| \xrightarrow{p} 0$.

If $\tilde{Q}_{NT}(F)$ is bounded away from zero for at least some F that satisfies $F'F/T = I_{k_T}$ and $\sup_{it}|F_{it}| \le \bar{F}$, it follows that $\tilde{F}$ is asymptotically the maximizer of $\tilde{Q}_{NT}(F)$. This condition is now shown to hold; more precisely, it is shown that there exists a constant c and some sequence $\{F_{it}\}$, say $\{F^*_{it}\}$, such that $\Pr[\tilde{Q}_{NT}(F^*) \ge c] \to 1$.

First consider the case $k_T \le r_T$, and choose $F^* = F^0(F^{0\prime}F^0/T)^{-\frac{1}{2}\prime}R'$, where $R = [I_{k_T}\ 0]$. Note that for positive semidefinite $r \times r$ matrices A and B, $tr(AB) \ge mineval(A)tr(B)$. Now

$$\begin{aligned}\tilde{Q}_{NT}(F^*) &= tr[(T^{-1}F^{0\prime}P_{F^*}F^0)(\Lambda_0'\Lambda_0/N)] \\ &\ge mineval(\Lambda_0'\Lambda_0/N)tr[RR'(F^{0\prime}F^0/T)] \\ &\ge mineval(\Lambda_0'\Lambda_0/N)mineval(F^{0\prime}F^0/T)tr(RR')\end{aligned}$$

where $tr(RR') = k_T$. Next, consider the case $k_T > r_T$ and let $F^* = [F^0(F^{0\prime}F^0/T)^{-\frac{1}{2}\prime}\ \ F^{\perp}]$, where $F^{\perp}$ is a $T \times (k_T - r_T)$ matrix such that $F^{\perp\prime}F^0 = 0$, $|F^{\perp}_{it}| \le \bar{F}$, and $0 < d \le eig(F^{\perp\prime}F^{\perp}/T) \le c < \infty$. Then $F^{0\prime}P_{F^*}F^0/T = F^{0\prime}F^0/T$, so $\tilde{Q}_{NT}(F^*) = tr[(F^{0\prime}F^0/T)(\Lambda_0'\Lambda_0/N)] \ge mineval(\Lambda_0'\Lambda_0/N)mineval(F^{0\prime}F^0/T)r_T$. Thus, for general k_T,

$$\tilde{Q}_{NT}(F^*) \ge [r_T mineval(\Lambda_0'\Lambda_0/N)]mineval(F^{0\prime}F^0/T)\min(k_T/r_T, 1).$$

Now $r_T mineval(\Lambda_0'\Lambda_0/N) \ge d$ by condition FL. It is shown below that $k_T r_T^4 \| (F^{0\prime}F^0/T) - \Sigma_{F,T} \|^2 \xrightarrow{p} 0$, so $mineval(F^{0\prime}F^0/T) - mineval(\Sigma_{F,T}) \xrightarrow{p} 0$; by condition M(d), $mineval(\Sigma_{F,T}) \ge d$. Also $k_T/r_T \to \mu > 0$ by condition R. Thus $\Pr[\tilde{Q}_{NT}(F^*) \ge \frac{1}{2}d^2\min(\mu, 1) > 0] \to 1$.

Next turn to the uniform consistency result. Let

$$\begin{aligned}H_{NT} &= (\tilde{F}'\tilde{F}/T)^{-1}(\tilde{F}'F^0/T)(\Lambda_0'\Lambda_0/N) \\ \xi_{st} &= X_s'X_t/N - F_s^{0\prime}(\Lambda_0'\Lambda_0/N)F_t^0 - \gamma(s-t)\end{aligned}$$

so $\hat{F}_t - H_{NT}F_t^0 = (\tilde{F}'\tilde{F}/T)^{-1}\{T^{-1}\Sigma_{s=1}^{T}\tilde{F}_s\gamma(s-t) + T^{-1}\Sigma_{s=1}^{T}\tilde{F}_s\xi_{st}\}$. Thus

$$\delta_{NT}\sup_t\|\hat{F}_t - H_{NT}F_t^0\|$$
$$\leq \delta_{NT}\|(\tilde{F}'\tilde{F}/T)^{-1}\|\{\sup_t\|T^{-1}\Sigma_{s=1}^{T}\tilde{F}_s\gamma(s-t)\| + \sup_t\|T^{-1}\Sigma_{s=1}^{T}\tilde{F}_s\xi_{st}\|\}.$$

Now $\|(\tilde{F}'\tilde{F}/T)^{-1}\| \leq ck_T^{1/2}$, $\delta_{NT}\sup_t\|T^{-1}\Sigma_{s=1}^{T}\tilde{F}_s\gamma(s-t)\| \leq (\delta_{NT}k_T/T)\bar{F}\Sigma_{u=-\infty}^{\infty}|\gamma(u)|$, and $\delta_{NT}\sup_t\|T^{-1}\Sigma_{s=1}^{T}\tilde{F}_s\xi_{st}\| \leq \delta_{NT}k_T^{1/2}\bar{F}\sup_{s,t}|\xi_{st}|$. Thus, $\delta_{NT}\sup_t\|\hat{F}_t - H_{NT}F_t^0\| \leq c(k_T^{3/2}\delta_{NT}/T)\bar{F}\Sigma_{u=-\infty}^{\infty}|\gamma(u)| + ck_T\delta_{NT}\bar{F}\sup_{s,t}|\xi_{st}| \xrightarrow{p} 0$ by lemma A1(e) and condition R.

(b) We now provide a limiting characterization of H_{NT} when $\tilde{F}$ is normalized so $\tilde{F}'\tilde{F}/T = I_{k_T}$. Let

$$\tilde{H}_{NT} = (\tilde{F}'F^0/T)(\Lambda_0'\Lambda_0/N)^{1/2}$$
$$H_{NT}^* = J'G_1^{*\prime}(F^{0\prime}F^0/T)^{1/2}(\Lambda_0'\Lambda_0/N)^{1/2}$$
$$H = J'G_1^{*\prime}\Sigma_{F,T}^{1/2}(\Lambda_0'\Lambda_0/N)$$

where J ($k_T \times k_T$) and G_1^* ($r_T \times k_T$) are defined below. Then $H_{NT} - H = (\tilde{H}_{NT} - H_{NT}^*)(\Lambda_0'\Lambda_0/N)^{1/2\prime} + J'G_1^{*\prime}[(F^{0\prime}F^0/T)^{1/2\prime} - \Sigma_{F,T}^{1/2\prime}](\Lambda_0'\Lambda_0/N)$, so

$$\|H_{NT} - H\|^2 \leq \|\tilde{H}_{NT} - H_{NT}^*\|^2\,\|(\Lambda_0'\Lambda_0/N)^{1/2\prime}\|^2$$
$$+ \|J'G_1^{*\prime}\|^2\|(F^{0\prime}F^0/T)^{1/2} - \Sigma_{F,T}^{1/2}\|^2\|\Lambda_0'\Lambda_0/N\|^2.$$

Now $\|\Lambda_0'\Lambda_0/N\| \leq \|(\Lambda_0'\Lambda_0/N)^{1/2\prime}\|^2 = tr(\Lambda_0'\Lambda_0/N) \leq c < \infty$ by condition FL. Also, it follows from the discussion below that J and G_1^* are orthonormal so that $\|J'G_1^{*\prime}\|^2 = k_T$. Thus

$\|H_{NT}-H\|^2 \xrightarrow{p} 0$ if $k_T\|(F^{0\prime}F^0/T)^{1/2}-\Sigma_{F,T}^{1/2}\|^2 \xrightarrow{p} 0$ and $\|\tilde{H}_{NT}-H_{NT}^*\|^2 \xrightarrow{p} 0$.

First consider $k_T\|(F^{0\prime}F^0/T)^{1/2}-\Sigma_{F,T}^{1/2}\|^2$. For specificity, let $(F^{0\prime}F^0/T)^{1/2}$ and $\Sigma_{F,T}^{1/2}$ be the Cholesky factorizations of $F^{0\prime}F^0/T$ and $\Sigma_{F,T}$. Because these Cholesky factors are Lipschitz functions of $F^{0\prime}F^0/T$ and $\Sigma_{F,T}$, $k_T\|(F^{0\prime}F^0/_T)^{1/2}-\Sigma_{F,T}^{1/2}\|^2 \xrightarrow{p} 0$ if $k_T r_T^4\|(F^{0\prime}F^0/T)-\Sigma_{F,T}\|^2 \xrightarrow{p} 0$. This is readily shown to hold using Markov's inequality and,

$$\begin{aligned} Ek_Tr_T^4\|(F^{0\prime}F^0/T)-\Sigma_{F,T}\|^2 &= Ek_Tr_T^4\Sigma_{\ell=1}^{r_T}\Sigma_{m=1}^{r_T}[T^{-1}\Sigma_{t=1}^{T}(F_{\ell t}^0F_{mt}^0-\Sigma_{F,T,\ell m})]^2 \\ &\le (k_Tr_T^6/T)\sup_{\ell,m,t}\Sigma_{u=-\infty}^{\infty}|\text{cov}(F_{\ell t}^0F_{mt}^0,F_{\ell t+u}^0F_{mt+u}^0)| \end{aligned}$$

which converges to zero by conditions R and M(d)(ii).

Finally, consider the term $\|\tilde{H}_{NT}-H_{NT}^*\|^2$. Without loss of generality, any F can be written,

$$F = F^0(F^{0\prime}F^0/T)^{-1/2\prime}G_1 + F^{\perp}(F^{\perp\prime}F^{\perp}/T)^{-1/2\prime}G_2$$

where, under the normalization $F'F/T=I_{k_T}$, $G_1'G_1+G_2'G_2=I_{k_T}$, where $F^{\perp}$ is the orthogonal complement of F^0 in $\Re^T$ and G_1 and G_2 are respectively $r_T\times k_T$ and $(T-r_T)\times k_T$. Let F* denote a maximizer of $\tilde{Q}_{NT}(F)$ (recall that F* is unique only up to a k×k rotation). Associate G_1^* and G_2^* with F* and $\tilde{G}_1$ and $\tilde{G}_2$ with $\tilde{F}$. Now, for general F

$$\begin{aligned} \tilde{Q}_{NT}(F) &= \text{tr}[T^{-1}F^{0\prime}P_FF^0(\Lambda_0'\Lambda_0/N)] \\ &= \text{tr}[G_1'A_TG_1] \end{aligned}$$

where $A_T = (F^{0\prime}F^0/T)^{1/2}(\Lambda_0'\Lambda_0/N)(F^{0\prime}F^0/T)^{1/2\prime}$, where the final equality uses the normalization $F'F/T=I_{kT}$. We therefore have,

$$\tilde{Q}_{NT}(F^*) - \tilde{Q}_{NT}(\tilde{F}) = tr[G_1^{\dagger\prime}A_TG_1^\dagger] - tr[\tilde{G}_1'A_T\tilde{G}_1] \xrightarrow{p} 0.$$

Thus there is a $k_T \times k_T$ full rank matrix J such that $\tilde{G}_1$ asymptotically equals $G_1^\dagger J$, where $J'G_1^{\dagger\prime}G_1^\dagger J/T = I_{k_T}$. For J thus defined, $\|\tilde{H}_{NT}-H_{NT}^*\|^2 = tr[(\tilde{G}_1-G_1^\dagger J)'A_T(\tilde{G}_1-G_1^\dagger J)] \xrightarrow{p} 0$. Finally recall that $\|(\Lambda_0'\Lambda_0/N)-D\| \xrightarrow{p} 0$ and $\|(F^{0\prime}F^0/T)-\Sigma_{F,T}\| \xrightarrow{p} 0$. Thus $\|H_{NT}-H\| \xrightarrow{p} 0$, where $H = R'\Sigma_{F,T}^{\frac{1}{2}\prime}D$, where $R = J'G_1^{\dagger\prime}$, and also $\|A_T-\Sigma_{F,T}^{\frac{1}{2}\prime}D\Sigma_{F,T}^{\frac{1}{2}}\| \xrightarrow{p} 0$. □

Next turn to theorem 2. It is useful first to set out some additional notation and preliminary results. When $q > r$, it is convenient to consider forecasts based on $\tilde{F}$ rather than $\hat{F}$ (this is done without loss of generality because they have identical column spaces). Partition $\tilde{F}$ as $\tilde{F} = [\tilde{F}^a \;\; \tilde{F}^b]$, where $\tilde{F}^a$ is $T\times r$ and $\tilde{F}^b$ is $T\times(q\text{-}r)$, where the column space of $\tilde{F}^a$ equals the column space of the first r eigenvectors of $N^{-1}\sum_{i=1}^N \underline{X}_i'\underline{X}_i$, and the columns of $\tilde{F}^b$ are (in order) the next q-r eigenvectors of this matrix. The uniqueness of the eigenvectors implies that, given X, the column space of $\tilde{F}^a$ does not depend on q as long as $q \geq r$. The specific rotation of these first r eigenvectors adopted here is $\tilde{F}^a = \hat{F}^a(H_{NT}^a)^{-1}$, where $\hat{F}^a$ denotes $\hat{F}$ estimated with $k=r$ and H_{NT}^a is the corresponding H_{NT} matrix from theorem 1. This rotation is also made without loss of generality and does not depend on q for $q \geq r$. In this notation, theorem 1 implies that $\delta_{NT}\sup_t\|\tilde{F}_t^a-F_t^0\| \xrightarrow{p} 0$.

The following lemma is used in the proof of theorem 2.

Lemma A2.

Under the assumptions of theorem 2,

(a) Let x be a $T\times 1$ random vector with $\text{plim}\|T^{-\frac{1}{2}}x\| \leq \infty$ and $\text{plim}\|T^{-1}F^{0\prime}x\| < \infty$.

Then $\delta_{NT}|x'(P_{\tilde{F}^a} - P_{F^0})x/T| \xrightarrow{p} 0$.

(b) $\delta_{NT}\beta'F^{0\prime}P_{\tilde{F}^b}F^0\beta/T \xrightarrow{p} 0$

(c) $\delta_{NT}\epsilon' P_{\tilde{F}^b}\epsilon/T \xrightarrow{p} 0$

(d) $\delta_{NT} y' P_{\tilde{F}^b} y/T \xrightarrow{p} 0$

(e) $\delta_{NT}|y'(P_{\tilde{F}^a} - P_{F^0})y/T| \xrightarrow{p} 0$

(f) $\delta_{NT}|y'(P_{\tilde{F}} - P_{F^0})y/T| \xrightarrow{p} 0.$

Proof

(a) let $V_{00} = F^{0\prime}F^0/T$, $V_{x0} = x'F^0/T$, $V_{xa} = x'\tilde{F}^a/T$, $V_{aa} = \tilde{F}^{a\prime}\tilde{F}^a/T$, etc. Then,

$$\begin{aligned}\delta_{NT}|x'(P_{\tilde{F}^a}-P_{F^0})x/T| \le\ & \delta_{NT}\|V_{xa}-V_{x0}\|\,\|V_{aa}^{-1}\|\,\|V_{xa}\| \\ &+ \delta_{NT}\|V_{x0}\|\,\|V_{aa}^{-1}-V_{00}^{-1}\|\,\|V_{xa}\| \\ &+ \delta_{NT}\|V_{x0}\|\,\|V_{00}^{-1}\|\,\|V_{xa}-V_{x0}\|\end{aligned}$$

By condition M(d), $\|V_{00}\| \le r^{1/2}d$ and $\|V_{00}^{-1}\| \le r^{1/2}c$. Also, $\delta_{NT}\|V_{xa}-V_{x0}\| \le \|T^{-1/2}x\|\delta_{NT}\sup_t\|\tilde{F}_t^a-F_t^0\| \xrightarrow{p} 0$, and $\delta_{NT}\|V_{00}-V_{aa}\| \le \delta_{NT}\|(\tilde{F}^a-F^0)'(\tilde{F}^a-F^0)/T\| + 2\delta_{NT}\|(\tilde{F}^a-F^0)'F^0/T\| \xrightarrow{p} 0$. By the continuity of the inverse, $\delta_{NT}\|V_{00}^{-1}-V_{aa}^{-1}\| \xrightarrow{p} 0$. By assumption, $\|T^{-1/2}x\|$ and $\|V_{0x}\|$ are $O_p(1)$. The result follows.

(b) Use $\tilde{F}^{a\prime}\tilde{F}^b=0$ to write, $M_{\tilde{F}} = I-P_{\tilde{F}^a}-P_{\tilde{F}^b} = M_{\tilde{F}^a}-P_{\tilde{F}^b}$, so $x'M_{\tilde{F}^a}x = x'P_{\tilde{F}^b}x + x'M_{\tilde{F}}x \ge x'P_{\tilde{F}^b}x$. Also, $x'M_{\tilde{F}^a}x = x'M_{F^0}x + x'(P_{F^0}-P_{\tilde{F}^a})x$. Thus, $\delta_{NT}x'P_{\tilde{F}^b}x/T \le \delta_{NT}x'M_{F^0}x/T + \delta_{NT}|x'(P_{\tilde{F}^a}-P_{F^0})x|/T$. Now let $x = F_0\beta$. Then $\beta'F^{0\prime}M_{F^0}F^0\beta = 0$ and $\delta_{NT}|\beta'F^{0\prime}(P_{\tilde{F}^a}-P_{F^0})F^0\beta/T| \xrightarrow{p} 0$ by part a of this lemma, because $\|T^{-1/2}x\| = \|T^{-1/2}F^0\beta\|^2 \xrightarrow{p} \beta'\Sigma_{F,T}\beta < \infty$ and $\|V_{0x}\| = \|T^{-1}F^{0\prime}F^0\beta\| \xrightarrow{p} \|\Sigma_{F,T}\beta\| \le \infty$; thus $\delta_{NT}\beta'F^{0\prime}P_{\tilde{F}^b}F^0\beta/T \xrightarrow{p} 0$.

(c) This follows because ϵ_{t+1} is a martingale difference sequence with respect to $\{X_t, y_t, F_t, X_{t-1}, y_{t-1}, F_{t-1},\ldots\}$ and because $P_{\tilde{F}^b}$ is idempotent with rank q-r.

(d) $\delta_{NT}y'P_{\tilde{F}^b}y/T \leq [(\delta_{NT}\beta'F^{0\prime}P_{\tilde{F}^b}F^0\beta/T)^{½} + (\delta_{NT}\epsilon'P_{\tilde{F}^b}\epsilon/T)^{½}]^2$, which converges in probability to zero by parts (b) and (c).

(e) This follows from (a) with x=y because $\|T^{-½}y\|^2 \xrightarrow{p} Ey^2$ and $\|T^{-1}F^{0\prime}y\| \xrightarrow{p} \|\Sigma_{F,T}\beta\|$.

(f) This follows from (d) and (e).

Proof of theorem 2

(a) Let $\hat{\epsilon}_0 = y_t-\hat{\beta}^{0\prime}F^0_t$, where $\hat{\beta}^0 = (F^{0\prime}F^0)^{-1}(F^{0\prime}y)$, and write

$$\hat{\sigma}^2_\epsilon(q)-\sigma^2_\epsilon = [\hat{\epsilon}'\hat{\epsilon}/T-\hat{\epsilon}^{0\prime}\hat{\epsilon}^0/T] + [\hat{\epsilon}^{0\prime}\hat{\epsilon}^0/T-\epsilon'\epsilon/T] + [\epsilon'\epsilon/T-\sigma^2_\epsilon].$$

Consider the three bracketed terms separately.

(i) $\hat{\epsilon}'\hat{\epsilon}/T-\hat{\epsilon}^{0\prime}\hat{\epsilon}^0/T = y'(P_{\tilde{F}} - P_{F^0})y/T \xrightarrow{p} 0$ by lemma A2(f).

(ii) The moment conditions imply that $\hat{\beta}^0 \xrightarrow{p} \beta$, from which it follows that $\hat{\epsilon}^{0\prime}\hat{\epsilon}^0/T-\epsilon'\epsilon/T \xrightarrow{p} 0$.

(iii) This follows from the moment assumptions on ϵ_t. □

(b)The proof proceeds by showing (i) $Pr[\hat{r}>r]\to 0$ and (ii) $Pr[\hat{r}<r]\to 0$, from which it follows that $Pr[\hat{r}=r]\to 1$. The results $\hat{\sigma}^2_\epsilon(\hat{r}) \xrightarrow{p} \sigma^2_\epsilon$ follows from the consistency of r and from part (a) of this theorem.

(i) This holds trivially if k=r so suppose that k>r. Now

$$Pr[\hat{r}>r] \leq Pr[\min_{q=r+1,\ldots,k} IC_q < IC_r]$$

$$\leq \Sigma_{q=r+1}^{k} \Pr[\delta_{NT}(IC_q - IC_r) < 0]$$
$$= \Sigma_{q=r+1}^{k} \Pr[\delta_{NT}\ln(\hat{\sigma}_\epsilon^2(q)/\hat{\sigma}_\epsilon^2(r)) + (q-r)\delta_{NT}g(T) < 0]$$

By assumption, $\delta_{NT}g(T)\rightarrow\infty$. Thus, because (q-r) is positive, to prove $\Pr[\hat{r} > r]\rightarrow 0$ it suffices to show that $\delta_{NT}\ln[\hat{\sigma}_\epsilon^2(q)/\hat{\sigma}_\epsilon^2(r)] \xrightarrow{p} 0$ for $q=r+1,\ldots,k$. But because $\hat{\sigma}_\epsilon^2(k) \leq \hat{\sigma}_\epsilon^2(k-1) \leq \cdots \leq \hat{\sigma}_\epsilon^2(r)$, it suffices to show this for $q=k$. Now,

$$\delta_{NT}\ln[\hat{\sigma}_\epsilon^2(k)/\hat{\sigma}_\epsilon^2(r)] = \delta_{NT}\ln[1+\{\delta_{NT}(\hat{\sigma}_\epsilon^2(k)-\hat{\sigma}_\epsilon^2(r))/\hat{\sigma}_\epsilon^2(r)\}/\delta_{NT}]$$
$$= \delta_{NT}\ln[1+\{\delta_{NT}(y'P_{\tilde{F}^b}y/T)/\hat{\sigma}_\epsilon^2(r)\}/\delta_{NT}].$$

From part (a), $\hat{\sigma}_\epsilon^2(r)\xrightarrow{p}\sigma_\epsilon^2$, and from Lemma A2(d), $\delta_{NT}(y'P_{\tilde{F}^b}y/T)\xrightarrow{p}0$, so the final expression above converges to zero in probability and the desired result follows. □

(ii) This holds trivially if $r=1$ so suppose that $r>1$. Following the reasoning in (i), it suffices to show that $\Pr[IC_q \leq IC_r]\rightarrow 0$, $1\leq q<r$. Now $IC_q - IC_r = \ln\{[\hat{\sigma}_\epsilon^2(q)/\hat{\sigma}_\epsilon^2(r)] + (q-r)g(T)\}$. Because $g(T)\rightarrow 0$ and $\hat{\sigma}_\epsilon^2(r)\xrightarrow{p}\sigma_\epsilon^2$, $\Pr[IC_q \leq IC_r]\rightarrow 0$ if $\hat{\sigma}_\epsilon^2(r)-\hat{\sigma}_\epsilon^2(q)\xrightarrow{p}d>0$. From theorem 1, $\sup_t \| \hat{F}_T - H_{NT}F_t^0 \| \xrightarrow{p} 0$, where H_{NT} is $q\times r$ and $H_{NT}\xrightarrow{p}H$. Thus

$$\hat{\sigma}_\epsilon^2(r)-\hat{\sigma}_\epsilon^2(q) = T^{-1}\beta'F^{0\prime}F^0\beta - T^{-1}F^{0\prime}P_{\hat{F}}F^0\beta + o_p(1)$$
$$= T^{-1}\beta'F^{0\prime}F^0\beta - T^{-1}F^{0\prime}P_{F^0H'}F^0\beta + o_p(1)$$
$$= \beta'[\Sigma_{F,T} - \Sigma_{F,T}H'(H\Sigma_{F,T}H')^{-1}H\Sigma_{F,T}]\beta + o_p(1).$$

The term in brackets in the final line is positive definite, so for general β, $\beta'[\Sigma_{F,T} - \Sigma_{F,T}H'(H\Sigma_{F,T}H')^{-1}H\Sigma_{F,T}]\beta > 0$, which yields the desired result. □

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. An asterisk after the date denotes a series which is available for less than the full period and thus was included in the unbalanced panel but not the balanced panel. 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;

Real output and income (Out)

1. IP 1959:01-1997:09 5 INDUSTRIAL PRODUCTION: TOTAL INDEX (1992=100,SA)
2. IPP 1959:01-1997:09 5 INDUSTRIAL PRODUCTION: PRODUCTS, TOTAL (1992=100,SA)
3. IPF 1959:01-1997:09 5 INDUSTRIAL PRODUCTION: FINAL PRODUCTS (1992=100,SA)
4. IPC 1959:01-1997:09 5 INDUSTRIAL PRODUCTION: CONSUMER GOODS (1992=100,SA)
5. IPCD 1959:01-1997:09 5 INDUSTRIAL PRODUCTION: DURABLE CONSUMER GOODS (1992=100,SA)
6. IPCN 1959:01-1997:09 5 INDUSTRIAL PRODUCTION: NONDURABLE CONDSUMER GOODS (1992=100,SA)
7. IPE 1959:01-1997:09 5 INDUSTRIAL PRODUCTION: BUSINESS EQUIPMENT (1992=100,SA)
8. IPI 1959:01-1997:09 5 INDUSTRIAL PRODUCTION: INTERMEDIATE PRODUCTS (1992=100,SA)
9. IPM 1959:01-1997:09 5 INDUSTRIAL PRODUCTION: MATERIALS (1992=100,SA)
10. IPMD 1959:01-1997:09 5 INDUSTRIAL PRODUCTION: DURABLE GOODS MATERIALS (1992=100,SA)
11. IPMND 1959:01-1997:09 5 INDUSTRIAL PRODUCTION: NONDURABLE GOODS MATERIALS (1992=100,SA)
12. IPMFG 1959:01-1997:09 5 INDUSTRIAL PRODUCTION: MANUFACTURING (1992=100,SA)
13. IPD 1959:01-1997:09 5 INDUSTRIAL PRODUCTION: DURABLE MANUFACTURING (1992=100,SA)
14. IPN 1959:01-1997:09 5 INDUSTRIAL PRODUCTION: NONDURABLE MANUFACTURING (1992=100,SA)
15. IPMIN 1959:01-1997:09 5 INDUSTRIAL PRODUCTION: MINING (1992=100,SA)
16. IPUT 1959:01-1997:09 5 INDUSTRIAL PRODUCTION: UTILITIES (1992-=100,SA)
17. IPX 1967:01-1997:09* 1 CAPACITY UTIL RATE: TOTAL INDUSTRY (% OF CAPACITY,SA)(FRB)
18. IPXMCA 1959:01-1997:09 1 CAPACITY UTIL RATE: MANUFACTURING,TOTAL(% OF CAPACITY,SA)(FRB)
19. IPXDCA 1967:01-1997:09* 1 CAPACITY UTIL RATE: DURABLE MFG (% OF CAPACITY,SA)(FRB)
20. IPXNCA 1967:01-1997:09* 1 CAPACITY UTIL RATE: NONDURABLE MFG (% OF CAPACITY,SA)(FRB)
21. IPXMIN 1967:01-1997:09* 1 CAPACITY UTIL RATE: MINING (% OF CAPACITY,SA)(FRB)
22. IPXUT 1967:01-1997:09* 1 CAPACITY UTIL RATE: UTILITIES (% OF CAPACITY,SA)(FRB)
23. PMI 1959:01-1997:09 1 PURCHASING MANAGERS' INDEX (SA)
24. PMP 1959:01-1997:09 1 NAPM PRODUCTION INDEX (PERCENT)
25. GMPYQ 1959:01-1997:09 5 PERSONAL INCOME (CHAINED) (SERIES #52) (BIL 92$,SAAR)
26. GMYXPQ 1959:01-1997:09 5 PERSONAL INCOME LESS TRANSFER PAYMENTS (CHAINED) (#51) (BIL 92$,SAAR)

27. SIPCD 1959:01-1997:09 1 Spread IPCD - IP
28. SIPMND 1959:01-1997:09 1 Spread IPMND - IP
29. SIPUT 1959:01-1997:09 1 Spread IPUT - IP

Employment and hours (Emp)

30. LHEL 1959:01-1997:09 5 INDEX OF HELP-WANTED ADVERTISING IN NEWSPAPERS (1967=100;SA)
31. LHELX 1959:01-1997:09 4 EMPLOYMENT: RATIO; HELP-WANTED ADS:NO. UNEMPLOYED CLF
32. LHEM 1959:01-1997:09 5 CIVILIAN LABOR FORCE: EMPLOYED, TOTAL (THOUS.,SA)
33. LHNAG 1959:01-1997:09 5 CIVILIAN LABOR FORCE: EMPLOYED, NONAGRIC.INDUSTRIES (THOUS.,SA)
34. LHUR 1959:01-1997:09 1 UNEMPLOYMENT RATE: ALL WORKERS, 16 YEARS & OVER (%,SA)
35. LHU680 1959:01-1997:09 1 UNEMPLOY.BY DURATION: AVERAGE(MEAN)DURATION IN WEEKS (SA)
36. LHU5 1959:01-1997:09 1 UNEMPLOY.BY DURATION: PERSONS UNEMPL.LESS THAN 5 WKS (THOUS.,SA)
37. LHU14 1959:01-1997:09 1 UNEMPLOY.BY DURATION: PERSONS UNEMPL.5 TO 14 WKS (THOUS.,SA)
38. LHU15 1959:01-1997:09 1 UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 WKS + (THOUS.,SA)
39. LHU26 1959:01-1997:09 1 UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 TO 26 WKS (THOUS.,SA)
40. LHU27 1959:01-1997:09 1 UNEMPLOY.BY DURATION: PERSONS UNEMPL.27 WKS + (THOUS,SA)
41. LHCH 1959:01-1995:02* 1 AVERAGE HOURS OF WORK PER WEEK (HOUSEHOLD DATA)(SA)
42. LPNAG 1959:01-1997:09 5 EMPLOYEES ON NONAG. PAYROLLS: TOTAL (THOUS.,SA)
43. LP 1959:01-1997:09 5 EMPLOYEES ON NONAG PAYROLLS: TOTAL, PRIVATE (THOUS,SA)
44. LPGD 1959:01-1997:09 5 EMPLOYEES ON NONAG. PAYROLLS: GOODS-PRODUCING (THOUS.,SA)
45. LPMI 1959:01-1997:09 5 EMPLOYEES ON NONAG. PAYROLLS: MINING (THOUS.,SA)
46. LPCC 1959:01-1997:09 5 EMPLOYEES ON NONAG. PAYROLLS: CONTRACT CONSTRUCTION (THOUS.,SA)
47. LPEM 1959:01-1997:09 5 EMPLOYEES ON NONAG. PAYROLLS: MANUFACTURING (THOUS.,SA)
48. LPED 1959:01-1997:09 5 EMPLOYEES ON NONAG. PAYROLLS: DURABLE GOODS (THOUS.,SA)
49. LPEN 1959:01-1997:09 5 EMPLOYEES ON NONAG. PAYROLLS: NONDURABLE GOODS (THOUS.,SA)
50. LPSP 1959:01-1997:09 5 EMPLOYEES ON NONAG. PAYROLLS: SERVICE-PRODUCING (THOUS.,SA)
51. LPTU 1959:01-1997:09 5 EMPLOYEES ON NONAG. PAYROLLS: TRANS. & PUBLIC UTILITIES (THOUS.,SA)
52. LPT 1959:01-1997:09 5 EMPLOYEES ON NONAG. PAYROLLS: WHOLESALE & RETAIL TRADE (THOUS.,SA)
53. LPFR 1959:01-1997:09 5 EMPLOYEES ON NONAG. PAYROLLS: FINANCE,INSUR.&REAL ESTATE (THOUS.,SA
54. LPS 1959:01-1997:09 5 EMPLOYEES ON NONAG. PAYROLLS: SERVICES (THOUS.,SA)
55. LPGOV 1959:01-1997:09 5 EMPLOYEES ON NONAG. PAYROLLS: GOVERNMENT (THOUS.,SA)
56. LW 1964:01-1997:09* 2 AVG. WEEKLY HRS. OF PROD. WKRS.: TOTAL PRIVATE (SA)
57. LPHRM 1959:01-1997:09 1 AVG. WEEKLY HRS. OF PRODUCTION WKRS.: MANUFACTURING (SA)
58. LPMOSA 1959:01-1997:09 1 AVG. WEEKLY HRS. OF PROD. WKRS.: MFG.,OVERTIME HRS. (SA)
59. PMEMP 1959:01-1997:09 1 NAPM EMPLOYMENT INDEX (PERCENT)

Real retail, manufacturing and trade sales (RTS)

60. MSMTQ 1959:01-1997:09 5 MANUFACTURING & TRADE: TOTAL (MIL OF CHAINED 1992 DOLLARS)(SA)
61. MSMQ 1959:01-1997:09 5 MANUFACTURING & TRADE:MANUFACTURING;TOTAL(MIL OF CHAINED 1992 DOLLARS)(SA)
62. MSDQ 1959:01-1997:09 5 MANUFACTURING & TRADE:MFG; DURABLE GOODS (MIL OF CHAINED 1992 DOLLARS)(SA)
63. MSNQ 1959:01-1997:09 5 MANUFACT. & TRADE:MFG;NONDURABLE GOODS (MIL OF CHAINED 1992 DOLLARS)(SA)
64. WTQ 1959:01-1997:09 5 MERCHANT WHOLESALERS: TOTAL (MIL OF CHAINED 1992 DOLLARS)(SA)
65. WTDQ 1959:01-1997:09 5 MERCHANT WHOLESALERS:DURABLE GOODS TOTAL (MIL OF CHAINED 1992 DOLLARS)(SA)
66. WTNQ 1959:01-1997:09 5 MERCHANT WHOLESALERS:NONDURABLE GOODS (MIL OF CHAINED 1992 DOLLARS)(SA)
67. RTQ 1959:01-1997:09 5 RETAIL TRADE: TOTAL (MIL OF CHAINED 1992 DOLLARS)(SA)
68. RTNQ 1959:01-1997:09 5 RETAIL TRADE:NONDURABLE GOODS (MIL OF 1992 DOLLARS)(SA)

Consumption (PCE)

69. GMCQ 1959:01-1997:09 5 PERSONAL CONSUMPTION EXPEND (CHAINED) - TOTAL (BIL 92$,SAAR)
70. GMCDQ 1959:01-1997:09 5 PERSONAL CONSUMPTION EXPEND (CHAINED) - TOTAL DURABLES (BIL 92$,SAAR)
71. GMCNQ 1959:01-1997:09 5 PERSONAL CONSUMPTION EXPEND (CHAINED) - NONDURABLES (BIL 92$,SAAR)
72. GMCSQ 1959:01-1997:09 5 PERSONAL CONSUMPTION EXPEND (CHAINED) - SERVICES (BIL 92$,SAAR)
73. GMCANQ 1959:01-1997:09 5 PERSONAL CONS EXPEND (CHAINED) - NEW CARS (BIL 92$,SAAR)

Housing starts and sales (HSS)

74. HSFR 1959:01-1997:09 4 HOUSING STARTS:NONFARM(1947-58);TOTAL FARM&NONFARM(1959-)(THOUS.,SA
75. HSNE 1959:01-1997:09 4 HOUSING STARTS:NORTHEAST (THOUS.U.)S.A.
76. HSMW 1959:01-1997:09 4 HOUSING STARTS:MIDWEST(THOUS.U.)S.A.
77. HSSOU 1959:01-1997:09 4 HOUSING STARTS:SOUTH (THOUS.U.)S.A.

78. HSWST 1959:01-1997:09 4 HOUSING STARTS:WEST (THOUS.U.)S.A.
79. HSBP 1959:01-1997:09 4 BUILDING PERMITS FOR NEW PRIVATE HOUSING UNITS (THOUS.)
80. HSBR 1959:01-1997:09 4 HOUSING AUTHORIZED: TOTAL NEW PRIV HOUSING UNITS (THOUS.,SAAR)
81. HSBNE 1960:01-1997:09* 4 HOUSES AUTHORIZED BY BUILD. PERMITS:NORTHEAST(THOU.U.)S.A
82. HSBMW 1960:01-1997:09* 4 HOUSES AUTHORIZED BY BUILD. PERMITS:MIDWEST(THOU.U.)S.A.
83. HSBSOU 1960:01-1997:09* 4 HOUSES AUTHORIZED BY BUILD. PERMITS:SOUTH(THOU.U.)S.A.
84. HSBWST 1960:01-1997:09* 4 HOUSES AUTHORIZED BY BUILD. PERMITS:WEST(THOU.U.)S.A.
85. HNS 1963:01-1997:09* 4 NEW 1-FAMILY HOUSES SOLD DURING MONTH (THOUS,SAAR)
86. HNSNE 1973:01-1997:09* 4 ONE-FAMILY HOUSES SOLD:NORTHEAST(THOU.U.,S.A.)
87. HNSMW 1973:01-1997:09* 4 ONE-FAMILY HOUSES SOLD:MIDWEST(THOU.U.,S.A.)
88. HNSSOU 1973:01-1997:09* 4 ONE-FAMILY HOUSES SOLD:SOUTH(THOU.U.,S.A.)
89. HNSWST 1973:01-1997:09* 4 ONE-FAMILY HOUSES SOLD:WEST(THOU.U.,S.A.)
90. HNR 1963:01-1997:09* 4 NEW 1-FAMILY HOUSES, MONTH'S SUPPLY @ CURRENT SALES RATE(RATIO)
91. HNIV 1963:01-1997:09* 4 NEW 1-FAMILY HOUSES FOR SALE AT END OF MONTH (THOUS,SA)
92. HMOB 1959:01-1997:09 4 MOBILE HOMES: MANUFACTURERS' SHIPMENTS (THOUS.OF UNITS,SAAR)
93. CONTC 1964:01-1997:09* 4 CONSTRUCT.PUT IN PLACE:TOTAL PRIV & PUBLIC 1987$(MIL$,SAAR)
94. CONPC 1964:01-1997:09* 4 CONSTRUCT.PUT IN PLACE:TOTAL PRIVATE 1987$(MIL$,SAAR)
95. CONQC 1964:01-1997:09* 4 CONSTRUCT.PUT IN PLACE:PUBLIC CONSTRUCTION 87$(MIL$,SAAR)
96. CONDO9 1959:01-1997:09 4 CONSTRUCT.CONTRACTS: COMM'L & INDUS.BLDGS(MIL.SQ.FT.FLOOR SP.;SA)

Real inventories and inventory-sales ratios (Inv)

97. IVMTQ 1959:01-1997:09 5 MANUFACTURING & TRADE INVENTORIES:TOTAL (MIL OF CHAINED 1992)(SA)
98. IVMFGQ 1959:01-1997:09 5 INVENTORIES, BUSINESS, MFG (MIL OF CHAINED 1992 DOLLARS, SA)
99. IVMFDQ 1959:01-1997:09 5 INVENTORIES, BUSINESS DURABLES (MIL OF CHAINED 1992 DOLLARS, SA)
100. IVMFNQ 1959:01-1997:09 5 INVENTORIES, BUSINESS, NONDURABLES (MIL OF CHAINED 1992 DOLLARS, SA)
101. IVWRQ 1959:01-1997:09 5 MANUFACTURING & TRADE INV:MERCHANT WHOLESALERS (MIL OF CHAINED 1992 DOLLARS)
102. IVRRQ 1959:01-1997:09 5 MANUFACTURING & TRADE INV:RETAIL TRADE (MIL OF CHAINED 1992 DOLLARS)(SA)
103. IVSRQ 1959:01-1997:09 2 RATIO FOR MFG & TRADE: INVENTORY/SALES (CHAINED 1992 DOLLARS, SA)
104. IVSRMQ 1959:01-1997:09 2 RATIO FOR MFG & TRADE:MFG;INVENTORY/SALES (87$)(S.A.)
105. IVSRWQ 1959:01-1997:09 2 RATIO FOR MFG & TRADE:WHOLESALER;INVENTORY/SALES(87$)(S.A.)
106. IVSRRQ 1959:01-1997:09 2 RATIO FOR MFG & TRADE:RETAIL TRADE;INVENTORY/SALES(87$)(S.A.)
107. PMNV 1959:01-1997:09 1 NAPM INVENTORIES INDEX (PERCENT)

Orders and unfilled orders (Ord)

108. PMNO 1959:01-1997:09 1 NAPM NEW ORDERS INDEX (PERCENT)
109. PMDEL 1959:01-1997:09 1 NAPM VENDOR DELIVERIES INDEX (PERCENT)
110. MOCMQ 1959:01-1997:09 5 NEW ORDERS (NET) - CONSUMER GOODS & MATERIALS, 1992 DOLLARS (BCI)
111. MDOQ 1959:01-1997:09 5 NEW ORDERS, DURABLE GOODS INDUSTRIES, 1992 DOLLARS (BCI)
112. MSONDQ 1959:01-1997:09 5 NEW ORDERS, NONDEFENSE CAPITAL GOODS, IN 1992 DOLLARS (BCI)
113. MO 1959:01-1997:09 5 MFG NEW ORDERS: ALL MANUFACTURING INDUSTRIES, TOTAL (MIL$,SA)
114. MOWU 1959:01-1997:09 5 MFG NEW ORDERS: MFG INDUSTRIES WITH UNFILLED ORDERS(MIL$,SA)
115. MDO 1959:01-1997:09 5 MFG NEW ORDERS: DURABLE GOODS INDUSTRIES, TOTAL (MIL$,SA)
116. MDUWU 1959:01-1997:09 5 MFG NEW ORDERS:DURABLE GOODS INDUST WITH UNFILLED ORDERS(MIL$,SA)
117. MNO 1959:01-1997:09 5 MFG NEW ORDERS: NONDURABLE GOODS INDUSTRIES, TOTAL (MIL$,SA)
118. MNOU 1959:01-1997:09 5 MFG NEW ORDERS: NONDURABLE GDS IND.WITH UNFILLED ORDERS(MIL$,SA)
119. MU 1959:01-1997:09 5 MFG UNFILLED ORDERS: ALL MANUFACTURING INDUSTRIES, TOTAL (MIL$,SA)
120. MDU 1959:01-1997:09 5 MFG UNFILLED ORDERS: DURABLE GOODS INDUSTRIES, TOTAL (MIL$,SA)
121. MNU 1959:01-1997:09 5 MFG UNFILLED ORDERS: NONDURABLE GOODS INDUSTRIES, TOTAL (MIL$,SA)
122. MPCON 1959:01-1997:09 5 CONTRACTS & ORDERS FOR PLANT & EQUIPMENT (BIL$,SA)
123. MPCONQ 1959:01-1997:09 5 CONTRACTS & ORDERS FOR PLANT & EQUIPMENT IN 1992 DOLLARS (BCI)

Stock prices (SPr)

124. FSNCOM 1959:01-1997:09 5 NYSE COMMON STOCK PRICE INDEX: COMPOSITE (12/31/65=50)
125. FSNIN 1966:01-1997:09* 5 NYSE COMMON STOCK PRICE INDEX: INDUSTRIAL (12/31/65=50)
126. FSNTR 1966:01-1997:09* 5 NYSE COMMON STOCK PRICE INDEX: TRANSPORTATION (12/31/65=50)
127. FSNUT 1966:01-1997:09* 5 NYSE COMMON STOCK PRICE INDEX: UTILITY (12/31/65=50)
128. FSNFI 1966:01-1997:09* 5 NYSE COMMON STOCK PRICE INDEX: FINANCE (12/31/65=50)
129. FSPCOM 1959:01-1997:09 5 S&P'S COMMON STOCK PRICE INDEX: COMPOSITE (1941-43=10)
130. FSPIN 1959:01-1997:09 5 S&P'S COMMON STOCK PRICE INDEX: INDUSTRIALS (1941-43=10)

131. FSPCAP 1959:01-1997:09 5 S&P'S COMMON STOCK PRICE INDEX: CAPITAL GOODS (1941-43=10)
132. FSPTR 1970:01-1997:09* 5 S&P'S COMMON STOCK PRICE INDEX: TRANSPORTATION (1970=10)
133. FSPUT 1959:01-1997:09 5 S&P'S COMMON STOCK PRICE INDEX: UTILITIES (1941-43=10)
134. FSPFI 1970:01-1997:09* 5 S&P'S COMMON STOCK PRICE INDEX: FINANCIAL (1970=10)
135. FSDXP 1959:01-1997:09 1 S&P'S COMPOSITE COMMON STOCK: DIVIDEND YIELD (% PER ANNUM)
136. FSPXE 1959:01-1997:09 1 S&P'S COMPOSITE COMMON STOCK: PRICE-EARNINGS RATIO (%,NSA)
137. FSNVV3 1974:01-1997:07* 5 NYSE MKT COMPOSITION:REPTD SHARE VOL BY SIZE,5000+ SHRS,%

Exchange rates (EXR)

138. EXRUS 1959:01-1997:09 5 UNITED STATES;EFFECTIVE EXCHANGE RATE(MERM)(INDEX NO.)
139. EXRGER 1959:01-1997:09 5 FOREIGN EXCHANGE RATE: GERMANY (DEUTSCHE MARK PER U.S.$)
140. EXRSW 1959:01-1997:09 5 FOREIGN EXCHANGE RATE: SWITZERLAND (SWISS FRANC PER U.S.$)
141. EXRJAN 1959:01-1997:09 5 FOREIGN EXCHANGE RATE: JAPAN (YEN PER U.S.$)
142. EXRUK 1959:01-1997:09 5 FOREIGN EXCHANGE RATE: UNITED KINGDOM (CENTS PER POUND)
143. EXRCAN 1959:01-1997:09 5 FOREIGN EXCHANGE RATE: CANADA (CANADIAN $ PER U.S.$)

Interest rates (Int)

144. FYFF 1959:01-1997:09 2 INTEREST RATE: FEDERAL FUNDS (EFFECTIVE) (% PER ANNUM,NSA)
145. FYCP 1959:01-1997:09 2 INTEREST RATE: COMMERCIAL PAPER, 6-MONTH (% PER ANNUM,NSA)
146. FYGM3 1959:01-1997:09 2 INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,3-MO.(% PER ANN,NSA)
147. FYGM6 1959:01-1997:09 2 INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,6-MO.(% PER ANN,NSA)
148. FYGT1 1959:01-1997:09 2 INTEREST RATE: U.S.TREASURY CONST MATURITIES,1-YR.(% PER ANN,NSA)
149. FYGT5 1959:01-1997:09 2 INTEREST RATE: U.S.TREASURY CONST MATURITIES,5-YR.(% PER ANN,NSA)
150. FYGT10 1959:01-1997:09 2 INTEREST RATE: U.S.TREASURY CONST MATURITIES,10-YR.(% PER ANN,NSA)
151. FYAAAC 1959:01-1997:09 2 BOND YIELD: MOODY'S AAA CORPORATE (% PER ANNUM)
152. FYBAAC 1959:01-1997:09 2 BOND YIELD: MOODY'S BAA CORPORATE (% PER ANNUM)
153. FWAFIT 1973:01-1994:04* 1 WEIGHTED AVG FOREIGN INTEREST RATE(%,SA)
154. FYFHA 1959:01-1997:09 2 SECONDARY MARKET YIELDS ON FHA MORTGAGES (% PER ANNUM)
155. SFYCP 1959:01-1997:09 1 Spread FYCP - FYFF
156. SFYGM3 1959:01-1997:09 1 Spread FYGM3 - FYFF
157. SFYGM6 1959:01-1997:09 1 Spread FYGM6 - FYFF
158. SFYGT1 1959:01-1997:09 1 Spread FYGT1 - FYFF
159. SFYGT5 1959:01-1997:09 1 Spread FYGT5 - FYFF
160. SFYGT10 1959:01-1997:09 1 Spread FYGT10 - FYFF
161. SFYAAAC 1959:01-1997:09 1 Spread FYAAAC - FYFF
162. SFYBAAC 1959:01-1997:09 1 Spread FYBAAC - FYFF
163. SFYFHA 1959:01-1997:09 1 Spread FYFHA - FYFF

Money and credit quantity aggregates (Mon)

164. FM1 1959:01-1997:09 5 MONEY STOCK: M1(CURR,TRAV.CKS,DEM DEP,OTHER CK'ABLE DEP)(BIL$,SA)
165. FM2 1959:01-1997:09 5 MONEY STOCK:M2(M1+O'NITE RPS,EURO$,G/P&B/D MMMFS&SAV&SM TIME DEP(BIL$,
166. FM3 1959:01-1997:09 5 MONEY STOCK: M3(M2+LG TIME DEP,TERM RP'S&INST ONLY MMMFS)(BIL$,SA)
167. FML 1959:01-1997:09 5 MONEY STOCK:L(M3 + OTHER LIQUID ASSETS) (BIL$,SA)
168. FM2DQ 1959:01-1997:09 5 MONEY SUPPLY - M2 IN 1992 DOLLARS (BCI)
169. FMFBA 1959:01-1997:09 5 MONETARY BASE, ADJ FOR RESERVE REQUIREMENT CHANGES(MIL$,SA)
170. FMBASE 1959:01-1997:09 5 MONETARY BASE, ADJ FOR RESERVE REQ CHGS(FRB OF ST.LOUIS)(BIL$,SA)
171. FMRRA 1959:01-1997:09 5 DEPOSITORY INST RESERVES:TOTAL,ADJ FOR RESERVE REQ CHGS(MIL$,SA)
172. FMRNBA 1959:01-1997:09 5 DEPOSITORY INST RESERVES:NONBORROWED,ADJ RES REQ CHGS(MIL$,SA)
173. FMRNBC 1959:01-1997:09 5 DEPOSITORY INST RESERVES:NONBORROW+EXT CR,ADJ RES REQ CGS(MIL$,SA)
174. FMFBA 1959:01-1997:09 5 MONETARY BASE, ADJ FOR RESERVE REQUIREMENT CHANGES(MIL$,SA)
175. FCLS 1973:01-1997:09* 5 LOANS & SEC @ ALL COML BANKS: TOTAL (BILS,SA)
176. FCSGV 1973:01-1997:09* 5 LOANS & SEC @ ALL COML BANKS: U.S.GOVT SECURITIES (BIL$,SA)
177. FCLRE 1973:01-1997:09* 5 LOANS & SEC @ ALL COML BANKS: REAL ESTATE LOANS (BIL$,SA)
178. FCLIN 1973:01-1997:09* 5 LOANS & SEC @ ALL COML BANKS: LOANS TO INDIVIDUALS (BIL$,SA)
179. FCLNBF 1973:01-1994:01* 5 LOANS & SEC @ ALL COML BANKS: LOANS TO NONBANK FIN INST(BIL$,SA)
180. FCLNQ 1959:01-1997:09 5 COMMERCIAL & INDUSTRIAL LOANS OUSTANDING IN 1992 DOLLARS (BCI)
181. FCLBMC 1959:01-1997:09 1 WKLY RP LG COM'L BANKS:NET CHANGE COM'L & INDUS LOANS(BIL$,SAAR)
182. CCI30M 1959:01-1995:09* 1 CONSUMER INSTAL.LOANS: DELINQUENCY RATE,30 DAYS & OVER, (%,SA)
183. CCINT 1975:01-1995:09* 1 NET CHANGE IN CONSUMER INSTAL CR: TOTAL (MIL$,SA)

184. CCINV 1975:01-1995:09* 1 NET CHANGE IN CONSUMER INSTAL CR: AUTOMOBILE (MIL$,SA)
185. CCINRV 1980:01-1995:09* 1 NET CHANGE IN CONSUMER INSTAL CR: REVOLVING(MIL$,SA)

Price indexes (Pri)
186. PMCP 1959:01-1997:09 1 NAPM COMMODITY PRICES INDEX (PERCENT)
187. PWFSA 1959:01-1997:09 5 PRODUCER PRICE INDEX: FINISHED GOODS (82=100,SA)
188. PWFCSA 1959:01-1997:09 5 PRODUCER PRICE INDEX:FINISHED CONSUMER GOODS (82=100,SA)
189. PWIMSA 1959:01-1997:09 5 PRODUCER PRICE INDEX:INTERMED MAT.SUPPLIES & COMPONENTS(82=100,SA)
190. PWCMSA 1959:01-1997:09 5 PRODUCER PRICE INDEX:CRUDE MATERIALS (82=100,SA)
191. PWFXSA 1967:01-1997:09* 5 PRODUCER PRICE INDEX: FINISHED GOODS,EXCL. FOODS (82=100,SA)
192. PW160A 1974:01-1997:09* 5 PRODUCER PRICE INDEX: CRUDE MATERIALS LESS ENERGY (82=100,SA)
193. PW150A 1974:01-1997:09* 5 PRODUCER PRICE INDEX: CRUDE NONFOOD MAT LESS ENERGY (82=100,SA)
194. PSM99Q 1959:01-1997:09 5 INDEX OF SENSITIVE MATERIALS PRICES (1990=100)(BCI-99A)
195. PUNEW 1959:01-1997:09 5 CPI-U: ALL ITEMS (82-84=100,SA)
196. PU81 1967:01-1997:09* 5 CPI-U: FOOD & BEVERAGES (82-84=100,SA)
197. PUH 1967:01-1997:09* 5 CPI-U: HOUSING (82-84=100,SA)
198. PU83 1959:01-1997:09 5 CPI-U: APPAREL & UPKEEP (82-84=100,SA)
199. PU84 1959:01-1997:09 5 CPI-U: TRANSPORTATION (82-84=100,SA)
200. PU85 1959:01-1997:09 5 CPI-U: MEDICAL CARE (82-84=100,SA)
201. PUC 1959:01-1997:09 5 CPI-U: COMMODITIES (82-84=100,SA)
202. PUCD 1959:01-1997:09 5 CPI-U: DURABLES (82-84=100,SA)
203. PUS 1959:01-1997:09 5 CPI-U: SERVICES (82-84=100,SA)
204. PUXF 1959:01-1997:09 5 CPI-U: ALL ITEMS LESS FOOD (82-84=100,SA)
205. PUXHS 1959:01-1997:09 5 CPI-U: ALL ITEMS LESS SHELTER (82-84=100,SA)
206. PUXM 1959:01-1997:09 5 CPI-U: ALL ITEMS LESS MIDICAL CARE (82-84=100,SA)
207. PCGOLD 1975:01-1997:09* 5 COMMODITIES PRICE:GOLD,LONDON NOON FIX,AVG OF DAILY RATE,$ PER OZ
208. GMDC 1959:01-1997:09 5 PCE,IMPL PR DEFL:PCE (1987=100)
209. GMDCD 1959:01-1997:09 5 PCE,IMPL PR DEFL:PCE; DURABLES (1987=100)
210. GMDCN 1959:01-1997:09 5 PCE,IMPL PR DEFL:PCE; NONDURABLES (1987=100)
211. GMDCS 1959:01-1997:09 5 PCE,IMPL PR DEFL:PCE; SERVICES (1987=100)

Average hourly earnings (AHE)
212. LEH 1964:01-1997:09* 5 AVG HR EARNINGS OF PROD WKRS: TOTAL PRIVATE NONAGRIC ($,SA)
213. LEHCC 1959:01-1997:09 5 AVG HR EARNINGS OF CONSTR WKRS: CONSTRUCTION ($,SA)
214. LEHM 1959:01-1997:09 5 AVG HR EARNINGS OF PROD WKRS: MANUFACTURING ($,SA)
215. LEHTU 1964:01-1997:09* 5 AVG HR EARNINGS OF NONSUPV WKRS: TRANS & PUBLIC UTIL($,SA)
216. LEHTT 1964:01-1997:09* 5 AVG HR EARNINGS OF PROD WKRS:WHOLESALE & RETAIL TRADE(SA)
217. LEHFR 1964:01-1997:09* 5 AVG HR EARNINGS OF NONSUPV WKRS: FINANCE,INSUR,REAL EST($,SA)
218. LEHS 1964:01-1997:09* 5 AVG HR EARNINGS OF NONSUPV WKRS: SERVICES ($,SA)

Miscellaneous (Oth)
219. FSTE 1986:01-1997:09* 5 U.S.MDSE EXPORTS: TOTAL EXPORTS(F.A.S. VALUE)(MIL.$,S.A.)
220. FSTM 1986:01-1997:09* 5 U.S.MDSE IMPORTS: GENERAL IMPORTS(C.I.F. VALUE)(MIL.$,S.A.)
221. FTMD 1986:01-1997:09* 5 U.S.MDSE IMPORTS: GENERAL IMPORTS (CUSTOMS VALUE)(MIL$,S.A.)
222. FSTB 1986:01-1997:09* 2 U.S.MDSE TRADE BALANCE:EXPORTS LESS IMPORTS(FAS/CIF)(MIL$,S.A.)
223. FTB 1986:01-1997:09* 2 U.S.MDSE TRADE BALANCE:EXP.(FAS) LESS IMP.(CUSTOM)(MIL$,S.A.)
224. HHSNTN 1959:01-1997:09 1 U. OF MICH. INDEX OF CONSUMER EXPECTATIONS(BCD-83)

References

Bekker, P., P. Dobbelstein, and T. Wansbeek (1996), "The APT Model as Reduced-Rank Regression," *Journal of Business and Economic Statistics*, 14, 199-202.

Campbell, J.Y., A.W. Lo and A. C. MacKinley (1996), *The Econometrics and Financial Markets*, Princeton University Press, Princeton, New Jersey.

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

Forni, M. and L. Reichlin (1997), "National Policies and Local Economies: Europe and the United States," CEPR Discussion paper no. 1632.

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.

Geweke, J. and G. Zhou (1996), "Measuring the Pricing Error of the Arbitrage Pricing Theory," *Review of Financial Studies*, 9, 557-587.

Mei, J. (1993), "A Semiautoregression Approach to the Arbitrage Pricing Theory," *Journal of Finance*48, 599-620.

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. (1989), "Two Models of Measurements and the Investment Accelerator," *Journal of Political Economy*, 97, 251-287.

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.

Schneeweiss, H. and H. Mathes (1995), "Factor Analysis and Principal Components," *Journal of Multivariate Analysis* 55, 105-124.

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

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), "Median Unbiased Estimation of Coefficient Variance in a Time Varying Parameter Model," *Journal of the American Statistical Association*, 93, 349-358.

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.

Zhou, G. (1997), "Security Return Factors as Linear Combinations of Economic Variables," manuscript, Washington University in St. Louis

Table 1

Monte Carlo Results

T	N	r	k	q	α	a	b	h	$R^2_{\hat{F},\hat{F}_0}$	$S^2_{\hat{Y},\hat{Y}_0}$ q=r	BIC	AIC	(3.2), ω = 0.001	0.005	0.010
									A. Static factor models						
25	50	5	5	0	0.0	0.0	0.0	0	0.88	0.70	0.57	0.67	0.71	0.70	0.61
25	100	5	5	0	0.0	0.0	0.0	0	0.92	0.81	0.68	0.76	0.81	0.80	0.76
25	250	5	5	0	0.0	0.0	0.0	0	0.95	0.88	0.74	0.83	0.88	0.88	0.86
25	500	5	5	0	0.0	0.0	0.0	0	0.96	0.90	0.74	0.85	0.90	0.90	0.89
50	50	5	5	0	0.0	0.0	0.0	0	0.89	0.80	0.76	0.80	0.80	0.55	0.15
50	100	5	5	0	0.0	0.0	0.0	0	0.94	0.89	0.84	0.89	0.89	0.78	0.40
50	250	5	5	0	0.0	0.0	0.0	0	0.96	0.94	0.90	0.93	0.94	0.92	0.76
50	500	5	5	0	0.0	0.0	0.0	0	0.97	0.96	0.91	0.95	0.96	0.95	0.89
50	50	10	10	0	0.0	0.0	0.0	0	0.82	0.54	0.39	0.53	0.55	0.17	0.04
50	100	10	10	0	0.0	0.0	0.0	0	0.89	0.72	0.52	0.70	0.72	0.36	0.12
50	250	10	10	0	0.0	0.0	0.0	0	0.95	0.87	0.66	0.84	0.87	0.75	0.34
50	500	10	10	0	0.0	0.0	0.0	0	0.97	0.92	0.71	0.88	0.92	0.86	0.62
100	250	5	5	0	0.0	0.0	0.0	0	0.97	0.96	0.95	0.96	0.96	0.63	0.17
100	500	5	5	0	0.0	0.0	0.0	0	0.98	0.97	0.96	0.97	0.97	0.87	0.37
100	250	5	10	0	0.0	0.0	0.0	0	0.95	0.96	0.95	0.94	0.95	0.56	0.17
100	500	5	10	0	0.0	0.0	0.0	0	0.97	0.98	0.96	0.95	0.96	0.86	0.37
100	100	10	10	0	0.0	0.0	0.0	0	0.90	0.83	0.75	0.83	0.73	0.06	0.01
100	250	10	10	0	0.0	0.0	0.0	0	0.95	0.93	0.87	0.92	0.91	0.16	0.05
100	500	10	10	0	0.0	0.0	0.0	0	0.97	0.95	0.89	0.95	0.94	0.34	0.09
100	100	10	15	0	0.0	0.0	0.0	0	0.86	0.83	0.76	0.81	0.75	0.05	0.00
100	250	10	15	0	0.0	0.0	0.0	0	0.93	0.92	0.85	0.90	0.90	0.15	0.03
100	500	10	15	0	0.0	0.0	0.0	0	0.95	0.96	0.89	0.93	0.93	0.37	0.09

Table 1, continued

T	N	r	k	q	α	a	b	h	$R^2_{F,\hat{F}_0}$	$S^2_{Y,\hat{Y}_0}$ q=r	BIC	AIC	(3.2), ω = 0.001	0.005	0.010
									B. Correlated errors						
100	250	5	5	0	0.0	0.5	0.0	0	0.97	0.96	0.94	0.96	0.95	0.59	0.17
100	500	5	5	0	0.0	0.5	0.0	0	0.98	0.97	0.96	0.97	0.97	0.84	0.32
100	250	5	10	0	0.0	0.5	0.0	0	0.93	0.96	0.95	0.95	0.96	0.61	0.16
100	500	5	10	0	0.0	0.5	0.0	0	0.95	0.97	0.96	0.96	0.96	0.87	0.35
100	250	5	5	0	0.0	0.9	0.0	0	0.86	0.84	0.81	0.83	0.83	0.48	0.14
200	250	5	5	0	0.0	0.9	0.0	0	0.95	0.96	0.96	0.96	0.93	0.08	0.00
100	250	5	5	0	0.0	0.0	1.0	0	0.97	0.96	0.95	0.96	0.96	0.64	0.18
100	500	5	5	0	0.0	0.0	1.0	0	0.98	0.98	0.96	0.97	0.97	0.87	0.36
100	250	5	10	0	0.0	0.0	1.0	0	0.94	0.96	0.94	0.94	0.95	0.60	0.17
100	500	5	10	0	0.0	0.0	1.0	0	0.96	0.98	0.96	0.95	0.96	0.86	0.38
									C. Dynamic factor models						
100	250	5	10	1	0.0	0.0	0.0	0	0.95	0.92	0.85	0.91	0.90	0.18	0.04
100	500	5	10	1	0.0	0.0	0.0	0	0.97	0.95	0.88	0.94	0.94	0.36	0.09
100	250	5	15	1	0.0	0.0	0.0	0	0.93	0.91	0.84	0.89	0.89	0.17	0.05
100	500	5	15	1	0.0	0.0	0.0	0	0.95	0.95	0.88	0.92	0.92	0.35	0.09
100	250	5	5	0	0.9	0.0	0.0	0	0.92	0.78	0.75	0.78	0.77	0.44	0.22
100	500	5	5	0	0.9	0.0	0.0	0	0.93	0.79	0.76	0.79	0.79	0.60	0.39
100	250	5	10	0	0.9	0.0	0.0	0	0.90	0.77	0.73	0.75	0.75	0.45	0.24
100	500	5	10	0	0.9	0.0	0.0	0	0.91	0.79	0.76	0.78	0.78	0.60	0.36
100	250	5	5	0	0.9	0.5	1.0	0	0.91	0.76	0.72	0.76	0.75	0.43	0.21
100	500	5	5	0	0.9	0.5	1.0	0	0.92	0.78	0.76	0.78	0.78	0.58	0.34
100	250	5	10	0	0.9	0.5	1.0	0	0.87	0.79	0.74	0.77	0.77	0.48	0.24
100	500	5	10	0	0.9	0.5	1.0	0	0.89	0.80	0.76	0.79	0.79	0.61	0.35

Table 1, continued

T	N	r	k	q	α	a	b	h	$R^2_{\tilde{F},\hat{F}_0}$	$S^2_{\tilde{Y},\hat{Y}_0}$ q=r	BIC	AIC	(3.2), ω = 0.001	0.005	0.010
D. Time-varying factor loadings															
100	250	5	5	0	0.0	0.0	0.0	5	0.93	0.93	0.92	0.93	0.92	0.55	0.17
100	500	5	5	0	0.0	0.0	0.0	5	0.94	0.94	0.92	0.93	0.93	0.81	0.34
100	250	5	5	0	0.0	0.0	0.0	10	0.87	0.89	0.88	0.89	0.89	0.47	0.11
100	500	5	5	0	0.0	0.0	0.0	10	0.87	0.89	0.88	0.89	0.89	0.73	0.27
100	250	5	10	0	0.0	0.0	0.0	5	0.91	0.93	0.91	0.87	0.90	0.57	0.16
100	500	5	10	0	0.0	0.0	0.0	5	0.92	0.94	0.91	0.86	0.87	0.81	0.33
100	250	5	10	0	0.0	0.0	0.0	10	0.85	0.89	0.85	0.80	0.82	0.44	0.12
100	500	5	10	0	0.0	0.0	0.0	10	0.85	0.90	0.85	0.81	0.81	0.70	0.25
100	250	5	5	0	0.9	0.5	1.0	5	0.89	0.76	0.73	0.76	0.75	0.43	0.21
100	500	5	5	0	0.9	0.5	1.0	5	0.89	0.76	0.72	0.76	0.76	0.58	0.35
100	250	5	5	0	0.9	0.5	1.0	10	0.85	0.74	0.70	0.74	0.72	0.43	0.22
100	500	5	5	0	0.9	0.5	1.0	10	0.85	0.73	0.70	0.73	0.73	0.54	0.32
100	250	5	10	0	0.9	0.5	1.0	5	0.86	0.76	0.73	0.74	0.75	0.45	0.24
100	500	5	10	0	0.9	0.5	1.0	5	0.87	0.79	0.76	0.76	0.77	0.59	0.36
100	250	5	10	0	0.9	0.5	1.0	10	0.83	0.75	0.72	0.69	0.72	0.43	0.25
100	500	5	10	0	0.9	0.5	1.0	10	0.84	0.73	0.70	0.66	0.67	0.55	0.33

Notes: Based on 2000 Monte Carlo draws using the design described in section 4 of the text. The statistic $R^2_{\tilde{F},F_0}$ is the trace R^2 of a regression of the estimated factors on the true factor, and $S^2_{\tilde{Y},\hat{Y}_0}$ is measure of the closeness of the forecast of y_{T+1} based on the estimated factors to factor to the forecast based on the true factor; see the text for definitions. The $S^2_{\tilde{Y},\hat{Y}_0}$ values are reported for k=r and for $\hat{k}$ estimated using an information criterion, where on, $1\le\hat{k}\le k$.

Table 2

Simulated Out-Of-Sample Forecasting Results: Industrial Production Growth

Variables forecasted: $\ln(z_{t+12}/z_t)$, where z_t is the Index of Industrial Production. Data are monthly, 1959:1 - 1997:9

Entries are the ratio of the MSE of the indicated forecast, to the MSE of a univariate autoregressive forecast with lag length selected by BIC.

Forecasting method				
Leading-indicator forecasts				
using full set of indicators	.77			
with BIC model selection	.78			
Diffusion index forecasts				
	balanced panel (N=170)		unbalanced panel (N=224)	
Factor selection	DI	DIAR	DI	DIAR
k = BIC	0.54	0.56	0.54	0.55
k = 1	0.76	0.79	0.77	0.76
k = 2	0.58	0.61	0.62	0.63
k = 3	0.58	0.59	0.55	0.56
k = 4	0.67	0.67	0.54	0.55
k = 5	0.63	0.64	0.58	0.58
k = 6	0.63	0.63	0.56	0.57
k = 7	0.59	0.58	0.58	0.58
k = 8	0.58	0.58	0.56	0.56
k = 9	0.61	0.60	0.55	0.55
k = 10	0.60	0.59	0.59	0.59
k = 11	0.59	0.59	0.58	0.58
k = 12	0.59	0.59	0.61	0.62

Notes: The forecasting models are described in detail in section 5. DI refers to the diffusion index forecasts constructed using only the factors; DIAR augments this by including lagged values of $\Delta\ln(z_t)$ as additional predictors, where the lag length is selected by recursive BIC. The model estimation period begins 1960:1; the forecast period begins 1970:1 and ends in 1997:9.

Table 3

Simulated Out-Of-Sample Forecasting Results: CPI Inflation

Variables forecasted: $\ln(z_{t+12}/z_t)$, where z_t is the Consumer Price Index. Data are monthly, 1959:1 - 1997:9

Entries are the ratio of the MSE of the indicated forecast, to the MSE of a univariate autoregressive forecast with lag length selected by BIC.

Forecasting method				
Leading-indicator forecasts				
using full set of indicators	.89			
with BIC model selection	1.11			
Phillips curve forecasts				
without wage-price control variable	.94			
with wage-price control variable	.97			
Diffusion index forecasts				
	balanced panel (N=170)		unbalanced panel (N=224)	
Factor selection	DI	DIAR	DI	DIAR
k = BIC	0.82	0.71	0.84	0.71
k = 1	2.02	0.88	2.24	1.01
k = 2	1.21	0.69	1.72	0.78
k = 3	0.74	0.62	1.44	0.68
k = 4	0.78	0.63	0.91	0.65
k = 5	0.74	0.62	0.99	0.66
k = 6	0.78	0.66	0.86	0.65
k = 7	0.84	0.73	0.88	0.73
k = 8	0.85	0.72	0.87	0.74
k = 9	0.85	0.73	0.87	0.73
k = 10	0.85	0.71	0.87	0.73
k = 11	0.85	0.71	0.87	0.73
k = 12	0.86	0.72	0.86	0.72

Notes: See the notes to table 2.

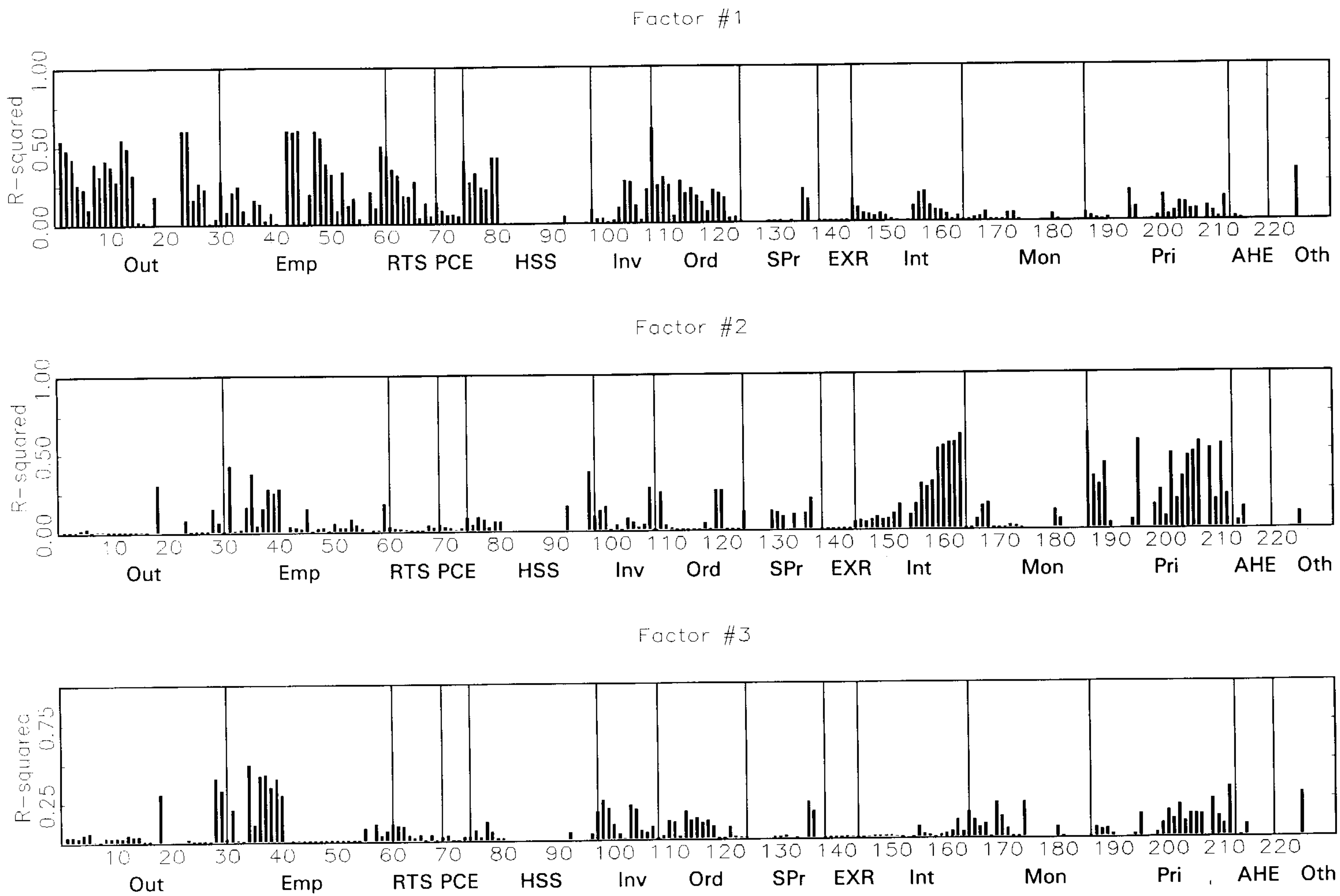

Figure 1. R^2s between factors and individual time series, grouped by category (see Appendix B)

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)

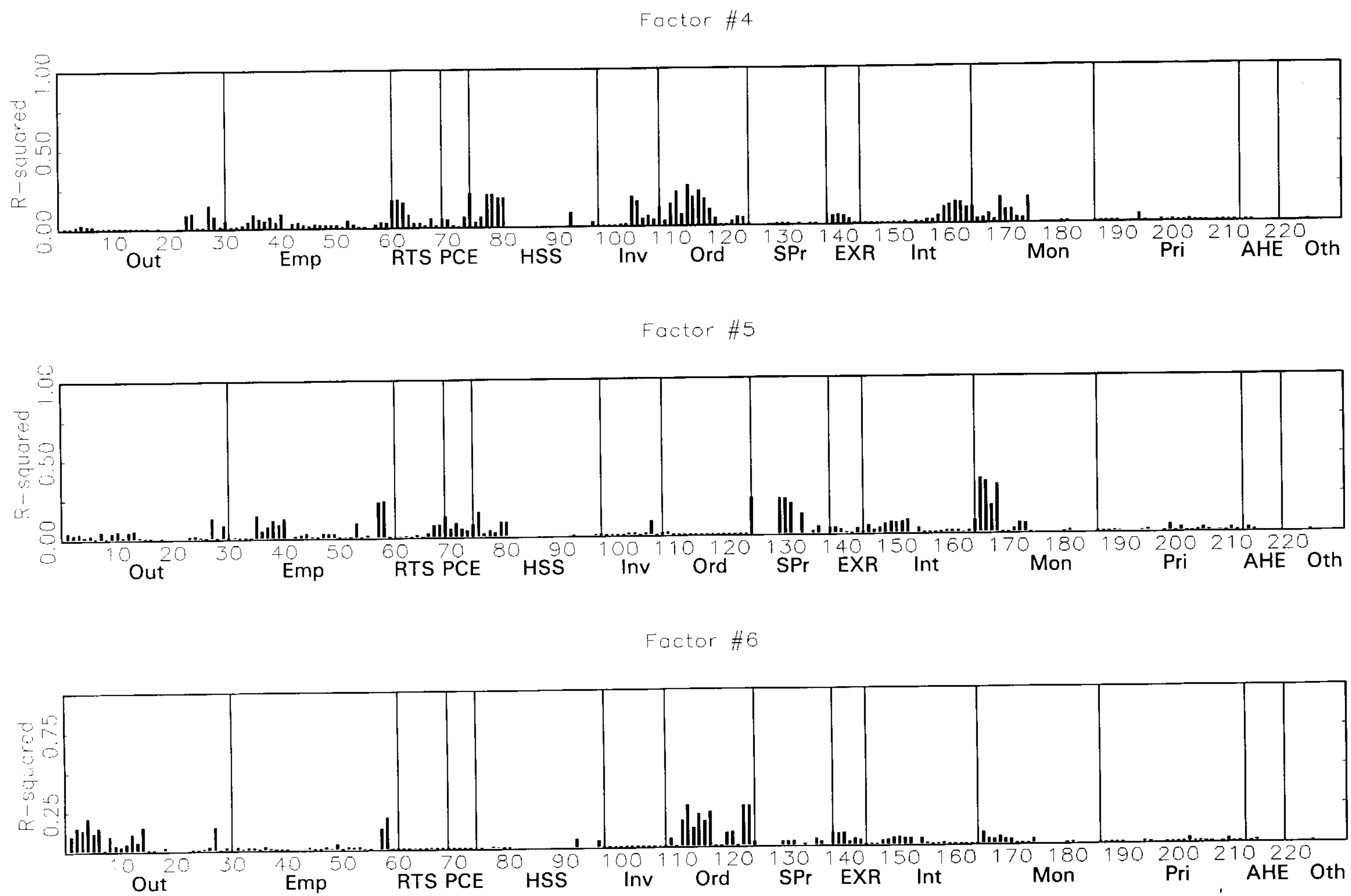

Figure 1 (continued)

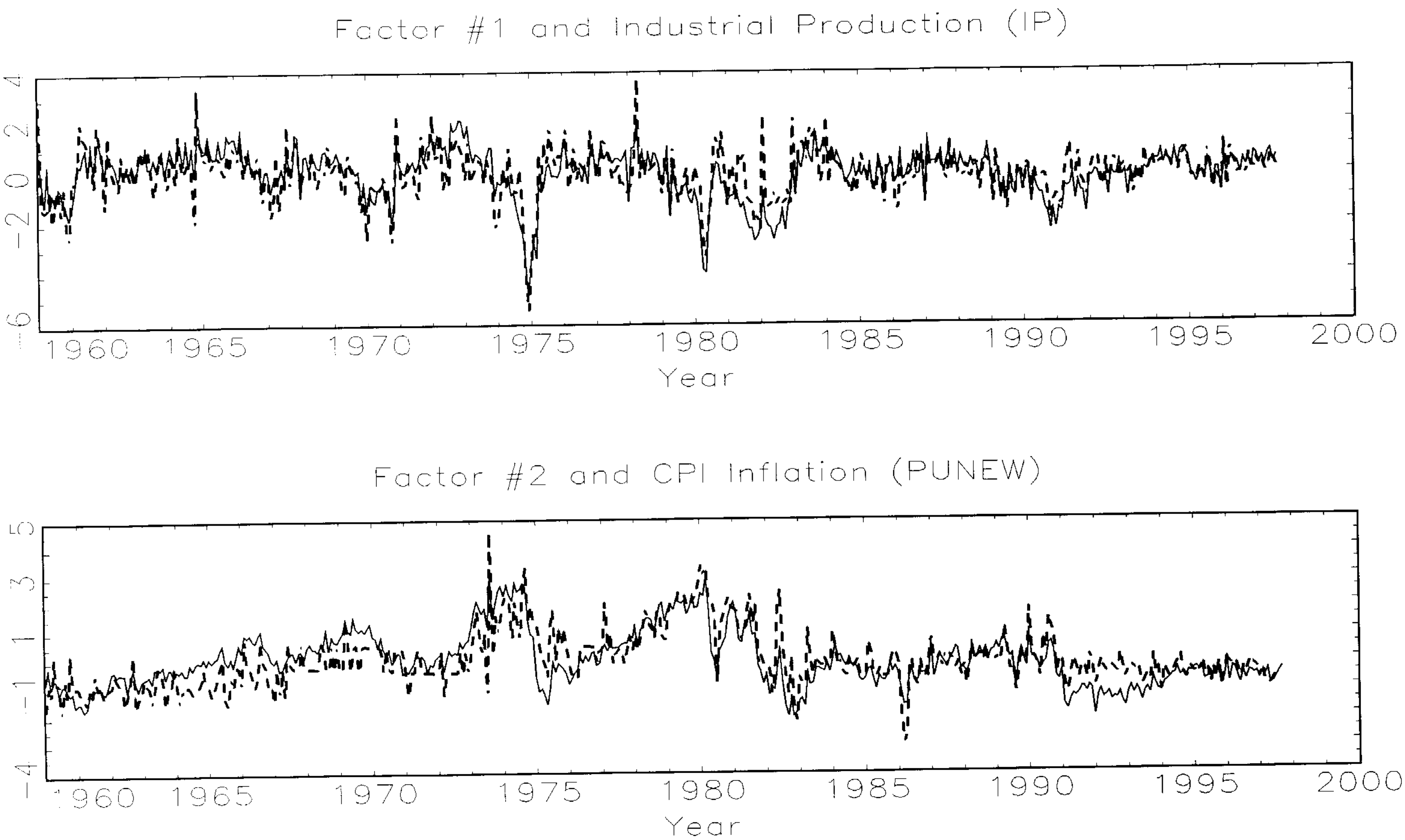

Figure 2. Factors #1 - 6 (solid lines) and selected individual monthly time series (dashed lines), 1960:1 - 1997:9
Series are scaled to have mean zero and unit standard deviation

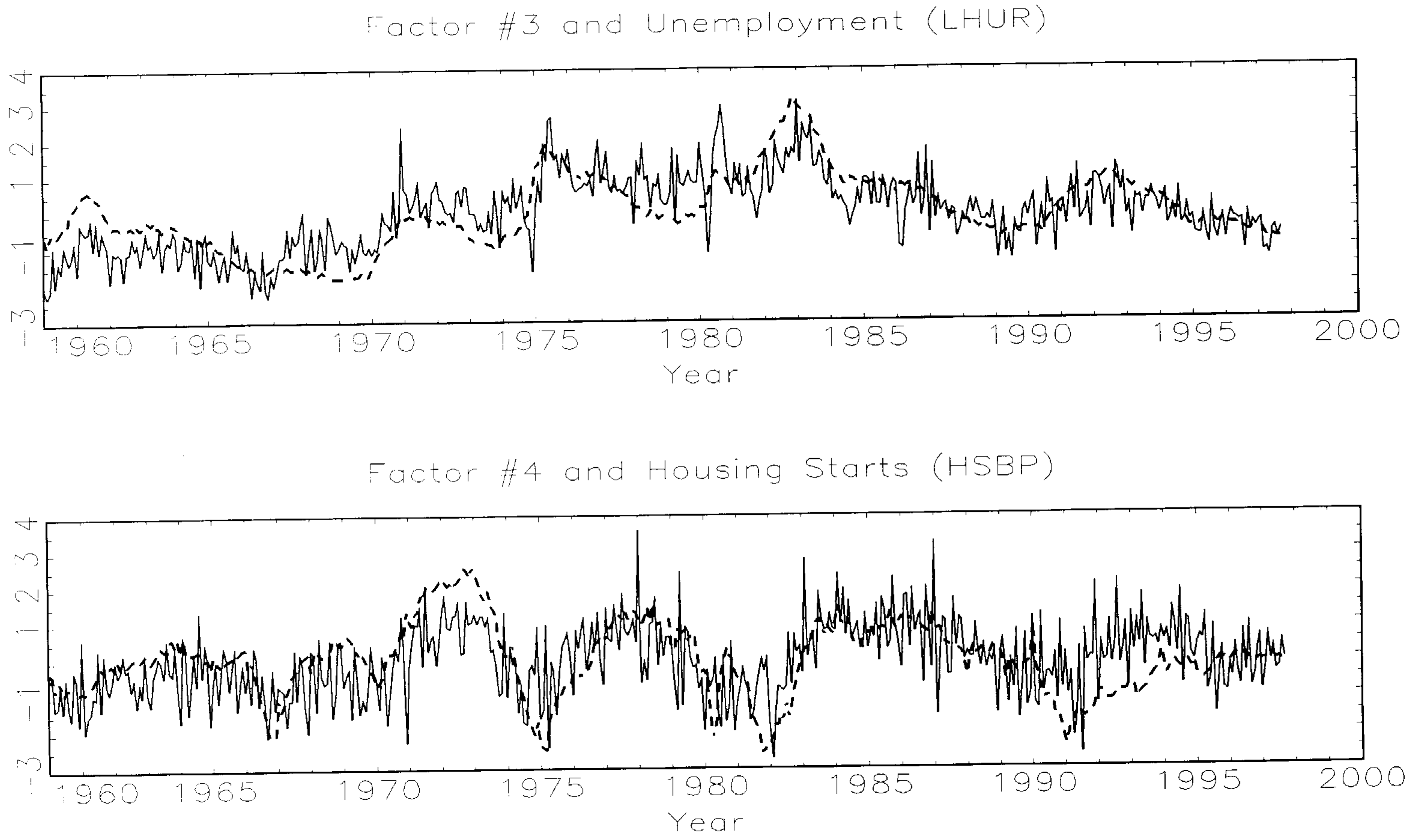

Figure 2 (continued)

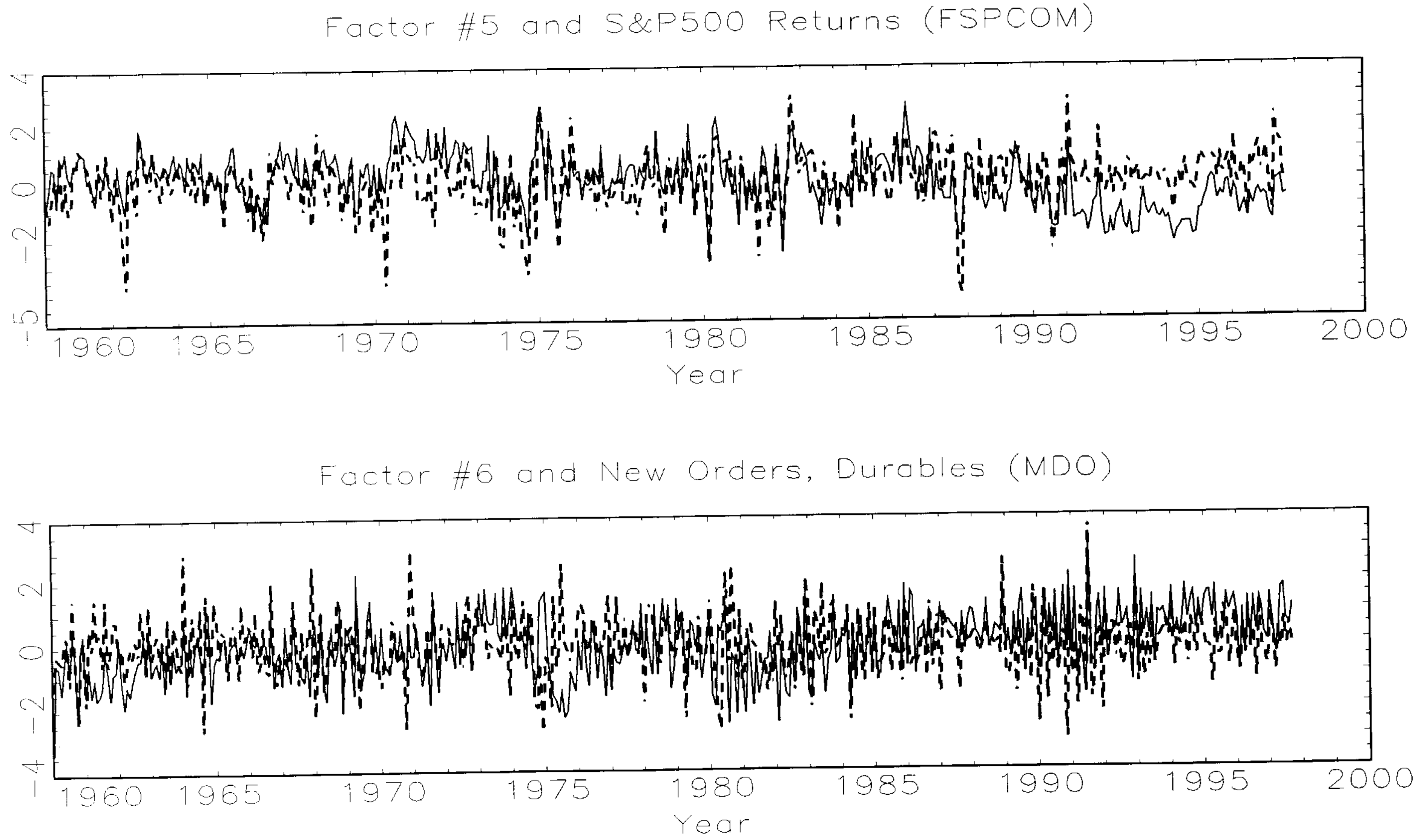

Figure 2 (continued)