CRED
WORKING
PAPERS

# Identification of Macroeconomic Factors in Large Panels

## CRED WP 2010/10

Lasse Bork, Hans Dewachter, and Romain Houssa

Center for Research in the Economics of Development
University of Namur

# Identification of Macroeconomic Factors in Large Panels*

Lasse Bork†, Hans Dewachter‡ and Romain Houssa§

Current Version: October 6, 2010

First Draft: May 2008

**Abstract**

This paper presents a dynamic factor model where the extracted factors and shocks are given a clear economic interpretation. The economic interpretation of the *factors* is obtained by means of a set of over-identifying loading restrictions, while the structural *shocks* are estimated following standard practices in the SVAR literature. Estimators based on the EM algorithm are developed. We apply this framework to a large panel of US monthly macroeconomic series. In particular, we identify five macroeconomic factors and discuss the economic impact of monetary policy shocks. The results are theoretically more plausible than those implied by standard SVAR models and indicate a significant role for monetary policy shocks in macroeconomic dynamics.

**JEL classifications:** E3, E43, C51, E52, C33

**Keywords:** Monetary policy, Business Cycles, Factor Models, EM Algorithm.

*Please address any comments to romain.houssa@fundp.ac.be. We thank Paul Söderlind, Ken Nyholm, Tom Engsted, Priscilla Toffano and participants at EEA 2010 for helpful comments and suggestions. We also thank Jean Boivin for kindly making the data set available on his website, HEC-Montréal. Hans Dewachter acknowledges financial support from FWO grant N° G.0626.07.

†Finance Research Group, Aarhus School of Business, University of Aarhus and CREATES at the University of Aarhus, a research center funded by Danish National Research Foundation.

‡CES, University of Leuven, RSM Rotterdam and CESIFO.

§CRED and CEREFIM, University of Namur

# 1 Introduction

In recent years, factor models have become a standard tool in applied macroeconomics and finance.[1] The increasing popularity of these models can be explained by two model features. First, factor models distinguish measurement errors and other idiosyncratic disturbances from common structural shocks. As such, factor models provide a direct mapping from observed data to their theoretical and structural counterparts[2]. Second, large data sets are becoming increasingly available and classical multivariate regression models generally perform poorly in fitting them. By contrast, factor models can handle large panels by exploiting the dynamic and cross-sectional structure of the panel. Specifically, various techniques have recently been developed to estimate large-dimensional factor models. For instance, Stock and Watson (2002*a*,*b*) and Forni et al. (2000) propose a non-parametric estimation approach based, respectively, on static and dynamic principal components. In related work, Otrok and Whiteman (1998) and Kim and Nelson (1999) propose a Bayesian estimation technique, whereas Doz et al. (2006, 2007) and Jungbacker and Koopman (2008) use an estimation approach based on the EM algorithm.

In this paper we discuss the economic identification of factors and shocks in the context of the dynamic factor model (DFM) introduced by Bai and Ng (2007) and Forni et al. (2000). In particular, we propose a procedure that *imposes* a specific and well-defined economic interpretation on the dynamic factors and the structural common shocks. The economic interpretation of the factors is

[1] In empirical macroeconomics they have been used for predictions (Bernanke and Boivin (2003), Forni et al. (2005), and Stock and Watson (2002*a*,*b*)); for structural analysis (Forni and Reichlin (1998), Forni and Gambetti (2010), Forni et al. (2008), Giannone et al. (2004, 2002), Houssa (2008*a*), Bernanke et al. (2005), Bork (2008) and Stock and Watson (2005)); and for constructing business cycle indicators (Altissimo et al. (2007)Forni et al. (2001), Kose et al. (2003), Houssa (2008*b*), and Otrok and Whiteman (1998)). Applications of factor models in finance include the arbitrage pricing theory (Chamberlain and Rothschild (1983) and Ingersoll (1984)); the measurement of risks (Campbell et al. (1997)); the estimation of the conditional risk-return relation in Ludvigson and Ng (2007); bond market applications (Mönch (2008), Ludvigson and Ng (2009) and Diebold et al. (2008)); and the prediction of the volatility of asset returns (Alessi et al. (2007)).

[2] Typically, these theoretical counterparts are defined within a DSGE model (see for example Altug (1989), Sargent and Sims (1977), Sargent (1989) and, recently, Boivin and Giannoni (2006)).

based on a set of over-identifying restrictions on factor loadings[3], while a set of standard restrictions on the impulse response functions are used to identify the structural shocks. We integrate these identification restrictions within the iterative maximum likelihood estimation approach proposed by Doz et al. (2006, 2007).

We illustrate our procedure by revisiting the panel data analyzed in Bernanke et al. (2005). We aim at identifying and extracting from the data panel five macroeconomic factors, respectively related to inflation, economic activity, commodity prices, money demand and monetary policy. Given the identification of these factors, we assess and analyze (as in Bernanke et al. (2005)) the impact of monetary policy shocks on a number of key macroeconomic variables through impulse response analysis and variance decompositions. We find that our identification procedure generates a more precise assessment of the impact of the monetary policy shocks, when compared to the standard SVAR or FAVAR.

Our paper is closely related to a number of recent studies. Boivin et al. (2009) and Reis and Watson (2008) impose loadings restrictions to identify a measure of pure inflation for the US economy. In the same way, Forni and Reichlin (2001) and Kose et al. (2003) use loading restrictions to differentiate between world, regional and country factors. Finally, Boivin and Giannoni (2006) employ loading restrictions to estimate the theoretical concepts, defined in a DSGE model. Alternatively, recent studies provide an economic interpretation to structural shocks in DFM, see for example Giannone et al. (2004); Forni and Gambetti (2010), Houssa (2008*a*) and Forni et al. (2008). The contribution of this paper is twofold. First, we combine the identification of both the dynamic factors and the structural shocks in a DFM framework. As such, we obtain a clear macroeconomic interpretation for both the (static and dynamic[4]) factors and shocks (see sections 2 and 3). Second, by directly integrating the linear identification restrictions in the EM algorithm, we obtain closed-form solutions for factor loadings and dynamics.

[3] Alternative types of identification schemes in DFMs, among which exclusion restrictions and loading restrictions, are discussed in the literature; see for instance Stock and Watson (2005), Reis and Watson (2008), Forni and Reichlin (2001) and Kose et al. (2003).

[4] *Static* factors are related to the variance-covariance matrix of the data while *dynamic* factors capture the property of their spectral density matrix. See Bai and Ng (2007) and Forni et al. (2000) for details.

The remainder of the paper is organized as follows. First, the methodological approach is explained in Section 2. We introduce a dynamic factor model and discuss the identification restrictions. In addition, closed-form solutions for the parameter estimates, consistent with the identification schemes and using results from Shumway and Stoffer (1982) and Wu et al. (1996), are presented. An empirical illustration is provided in Section 3. Section 4 concludes.

# 2 Methodology

We first introduce the DFM (see for instance Bai and Ng (2007) and Forni et al. (2000)). Subsequently, we employ the quasi maximum likelihood estimation approach as in Doz et al. (2006, 2007). We take this approach one step further by imposing over-identifying restrictions on the loadings and on the impulse response function (IRF). This allows a clear economic interpretation of the (static and dynamic) factors and a structural identification of the shocks.

## 2.1 Dynamic Factor Model

Consider a panel of observable economic variables $y_{i,t}$, where $i$ denotes the cross-section unit, $i = 1, ..., N$ while $t$ refers to the time index, $t = 1, ..., T$. The panel of observed economic variables is transformed into stationary variables with zero mean and unit variance. These transformed variables are labeled by $x_{i,t}$. Dynamic factor models assume that a variable $x_{i,t}$ can be decomposed into two components, the *common component*, $\chi_{i,t}$, and the *idiosyncratic component* $\xi_{it}$:

$$x_{i,t} = \chi_{i,t} + \xi_{i,t}. \tag{1}$$

Furthermore, in exact dynamic factor models it is assumed that the idiosyncratic and common components are uncorrelated at all leads and lags and across all variables, i.e. $E(\xi_{i,t}\chi_{j,s}) = 0, \forall\ s, t, i, j$. The common component, $\chi_{i,t}$, is assumed to be driven by a small number $q$, $q << N$, of common dynamic factors $f_t = (f_{1,t}$

$f_{2,t}, \cdots, f_{q,t})'$:

$$x_{i,t} = \lambda_i(L) f_t + \xi_{i,t}, \tag{2}$$

with

$$\lambda_i(L) = \lambda_{i,0} I + \lambda_{i,1} L + ... + \lambda_{i,s} L^s$$

where $\lambda_{i,j}$ denotes a $1 \times q$ vector containing the loadings for observable series $i$ on the $j$-th lag of the factors; the typical element of $\lambda_{i,j}$, i.e. $\lambda_{i,j}^k, (k = 1, ..., q)$, denotes the loading on the $k$-th factor at lag $j$ for series $i$. Stacking equation (2) over all cross-section units, $x_{i,t}$, $i = 1, ..., N$, gives:

$$X_t = \lambda_0 f_t + \lambda_1 f_{t-1} + \ldots + \lambda_s f_{t-s} + \xi_t, \tag{3}$$

where $X_t = (x_{1,t}, \ldots, x_{N,t})'$, $\xi_t = (\xi_{1,t}, \ldots, \xi_{N,t})'$, and $\lambda_j$, $j = 0, ..., s$, is a $N \times q$ matrix of series-specific factor loadings, $\lambda_j = [\lambda'_{1,j}, \lambda'_{2,j}, .., \lambda'_{N,j}]'$.

To close the model, we assume that the $q$-dimensional vector of common dynamic factors $f_t$ has a VAR($p$) representation:

$$\phi(L) f_t = \eta_t, \tag{4}$$

where $\phi(L) = I - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p$, with $\phi_j$ denoting a $q \times q$ matrix of autoregressive coefficients ($j = 1, \ldots, p$). Moreover, given the stationarity of the transformed panel, we impose stationarity on the DFM by requiring that the modulus of the roots of $\phi(L)^{-1}$ lie outside the unit circle. The $q$-dimensional vector of dynamic factor innovations is denoted by $\eta_t$. As in Doz et al. (2006), we make additional distributional assumptions: $\eta_t \sim i.i.d\ N(0, Q)$ and $\xi_t \sim i.i.d\ N(0, R)$, with $Q$ and $R$ denoting (semi-) positive definite matrices[5].

[5]Note that, by assuming *i.i.d* idiosyncratic components, (3)-(4) define an *exact* DFM as opposed to an *approximate* factor model where some correlation is allowed among idiosyncratic components. An exact factor structure is certainly a strong assumption, particularly in the case of large panel data sets where cross-sectional and serial correlations are expected to be

Following Bai and Ng (2007), Forni et al. (2000) and Stock and Watson (2002*b*) the model (equations (3) and (4)) can be restated as a static factor model with a $r \times 1$, $r = q(s+1)$, static factors $F_t$, $F_t = (f_t', ..., f_{t-s}')'$ :

$$X_t = \Lambda F_t + \xi_t, \tag{5}$$

$$F_t = \Phi F_{t-1} + VSu_t, \tag{6}$$

where $\Lambda = (\lambda_0, ..., \lambda_s)$ is the $N \times r$ matrix loading, implied by the dynamic factor loadings $\lambda_j$, $j = 0, ..., s$, $\Phi$ is the $r \times r$ companion matrix corresponding to $\phi(L)$, $V = \left(I', 0'_{(r-q)\times q}\right)'$, and $u_t$ represents the structural common shocks that are identified through the matrix $S$ (see sub-section 2.2.2 below). Inverting the VAR in (6) and substituting $F_t$ in (5) gives

$$X_t = B(L)u_t + \xi_t, \tag{7}$$

where $B(L) = \Lambda(I - \Phi L)^{-1}VS$, represents the IRF to $u_t$.

The state-space system, defined by equations (5) and (6), is not uniquely identified. We address the econometric identification as well as the economic interpretation of the factors in section 2.2.1. Finally, the identification of the structural shocks $u_t$ is discussed in section 2.2.2.

## 2.2 Economic interpretation

Economic interpretation of the factors and shocks is achieved by imposing two types of identification restrictions: (*i*) loading restrictions allowing for a clear macroeconomic interpretation of the factors, and (*ii*) restrictions on the IRF identifying the structural common shocks.

found. As such, (3)-(4) represent a missspecified model. However, Doz et al. (2006) show that, for large $N$ and $T$ the exact factor model estimators are consistent quasi-maximum likelihood estimators for the approximate factor model.

### 2.2.1 Economic factors

We impose a set of linear restrictions on the loading matrix $\Lambda$, (equation (5)), and denote the restricted loading matrix by $\Lambda^* = (\lambda_0^*, ..., \lambda_s^*)$. $\lambda_j^*$ denotes an $N \times q$ matrix of restricted factor loadings at lag $j$; with typical entry $\lambda_{i,j}^{k*}$, the possibly restricted loading for series $i$ on factor $k$ at lag $j$. The linear loading restrictions take the following general form:

$$H_\Lambda vec(\Lambda^*) = \kappa_\Lambda, \tag{8}$$

where $\kappa_\Lambda$ refers to a $\ell \times 1$ vector of $\ell$ linear combinations of restrictions of factor loadings defined by $H_\Lambda$, $H_\Lambda \in \mathbb{R}^{\ell \times Nr}$.

We use two types of loading restrictions, depending on the information content of the observables. In particular, economic identification is achieved by means of *(i)* unbiasedness restrictions and/or *(ii)* exclusion restrictions. Both types of restrictions are imposed on the dynamic factor loadings $\lambda_{i,j}^k$.

The *unbiasedness restriction* implies that the contemporaneous value of the observable $x_{i,t}$ is an unbiased and direct information variable for the $k$-th factor $f_{k,t}, k = 1, 2, \ldots, q,$:

$$\begin{aligned} &\lambda_{i,0}^{k*} = 1 \\ &\lambda_{i,0}^{l*} = 0 \text{ for } l \neq k \\ &\lambda_{i,j}^{l*} = 0 \text{ for } j = 1, ..., s, \text{ for } l = 1, ..., q. \end{aligned} \tag{9}$$

This type of restrictions is used on observables that are assumed to be a direct measure (up to an idiosyncratic component) of the underlying factor. For instance, our empirical application assumes that the observable "*CPI-u all items*" inflation is a direct measure for the inflation factor. As such, the unbiasedness restrictions imply a unit loading of "*CPI-u all items*" inflation on the contemporaneous inflation factor and zero loadings on all other factors and all lagged factors. Note that by imposing at least one unbiasedness restriction on each of the $q$ dynamic factors allows for the *econometric identification* of the DFM (see Geweke and Singleton (1981)) and the *economic interpretation* of the factors.

*Exclusion restrictions*, i.e. the case where $x_{i,t}$ is unrelated to $f_{k,t}$ or lags of $f_{k,t}$. These restrictions take the form of:

$$\lambda_{i,j}^{l*} = 0 \text{ for } j = 0, ..., s, \text{ for } l = k. \tag{10}$$

In the empirical application we use this type of restriction to exclude variables that do not have a direct information content on the factors.

Note that this identification scheme formalizes and extends the standard informal identification procedures used in the literature. The standard approach is to identify the factors from the dominant factor loadings of the economic variables, disregarding the smaller loadings. Our identification procedure formalizes this approach by *(i)* imposing exclusion restrictions on the non-informative variables, which ensures that only information of relevant variables is incorporated in the factor and *(ii)* facilitating interpretation of the factors by means of the unbiasedness restrictions imposing a direct mapping between the observables and the (static and dynamic) factors.

Finally, we allow for feedback effects across factors. Specifically, through the VAR specification of the transition equation (equation (6)), we allow for dynamic interactions among factors. As such, factors can be correlated and structural shocks are eventually transmitted across all observables.

### 2.2.2 Structural shocks

In equation (7), structural shocks are identified. We follow the standard identification procedure in the SVAR literature by choosing an appropriate matrix $S$ such that the implied restricted IRF, $B(L)^*$, has an economic justification. For instance, the Blanchard and Quah (1989) long-run restrictions can be obtained by choosing $S$ such that appropriate elements of $B(1)^*$ are equal to zero. Sign restrictions, recently introduced by Uhlig (2005), can also be fulfilled by choosing $S$ such that the time path of some elements of $B(L)^*$ have an appropriate sign. Finally, structural identification can be obtained by imposing the Sims (1980)'s triangular representation on the matrix $S$. This is the approach followed in our empirical application in section 3. We use the exclusion restrictions implied by the Cholesky decomposition of $Q = SS'$, with $S$ lower triangular. The structural interpretation of the shocks is then implied by the ordering of the factors and discussed in more detailed in section 3.

### 2.3 Estimation: the EM algorithm

Given the latent nature of the static factors, a standard EM algorithm is used to estimate the parameters and to extract the implied factors. Denote by $\Theta = \{\Lambda^*, R, \Phi, Q\}$ the set of parameters to be estimated with $\Lambda^*$ satisfying the set of identification restrictions listed in equation (8). Conditional on the estimates of the factors, $\hat{F}$ (and matrices measuring uncertainty $\hat{P}$), the elements of $\Theta$ can be estimated by (<u>M</u>aximization step):

$$\begin{aligned}\mathrm{vec}\left(\Lambda^*\right) &= \mathrm{vec}\left(DC^{-1}\right) \\ &\quad + \left(C^{-1} \otimes R\right) H_\Lambda' \left[H_\Lambda \left(C^{-1} \otimes R\right) H_\Lambda'\right]^{-1} \left\{\kappa_\Lambda - H_\Lambda \,\mathrm{vec}\left(DC^{-1}\right)\right\},\end{aligned} \tag{11}$$

where

$$\begin{aligned}&R = \tfrac{1}{T} G; \\ &\mathrm{vec}\left(\Phi\right) = \mathrm{vec}\left(BA^{-1}\right), \\ &\Omega = VQV' = \tfrac{1}{T}\left[C - BA^{-1}B'\right],\end{aligned} \tag{12}$$

and the estimator for $\Lambda^*$ follows from extending results in Wu et al. (1996).[6]

Conditional on the estimated parameters, $\Theta$, the latent factors can be extracted by means of the Kalman smoother and the required moments can be computed

[6] A derivation of the estimator is available on request.

(Expectation step). In particular, the following expectations are generated:

$$
\begin{aligned}
A &= \sum_{t=1}^{T} \left( \hat{P}_{t-1|T} + \hat{F}_{t-1|T}\hat{F}'_{t-1|T} \right), \\
B &= \sum_{t=1}^{T} \left( \hat{F}_{t|T}\hat{F}'_{t-1|T} + \hat{P}_{\{t,t-1\}|T} \right), \\
C &= \sum_{t=1}^{T} \left( \hat{F}_{t|T}\hat{F}'_{t|T} + \hat{P}_{t|T} \right), \\
D &= \sum_{t=1}^{T} X_t \hat{F}'_{t|T}, \\
G &= \sum_{t=1}^{T} (X_t - \Lambda^* \hat{F}_{t|T})(X_t - \Lambda^* \hat{F}_{t|T})' + \Lambda^* \hat{P}_{t|T} \Lambda^{*\prime},
\end{aligned}
\tag{13}
$$

with:

$$
\begin{aligned}
\hat{F}_{t|T} &= E(F_t \mid \mathcal{X}_T), \\
\hat{P}_{t|T} &= E((F_t - \hat{F}_{t|T})(F_t - \hat{F}_{t|T})' \mid \mathcal{X}_T), \\
\hat{P}_{\{t,t-1\}|T} &= E((F_t - \hat{F}_{t|T})(F_{t-1} - \hat{F}_{t-1|T})' \mid \mathcal{X}_T),
\end{aligned}
\tag{14}
$$

where $E(\cdot \mid \cdot)$ denotes the conditional expectations operator implied by the Kalman smoother (as a function of $\Theta$), see for instance de Jong and Mackinnon (1988) and de Jong (1989). $\mathcal{X}_T = \{X_1, \ldots, X_T\}$ denotes the information set. We iterate sequentially over the M-step in equation (11) and the E-step in equation (13) until convergence of the likelihood starting from different sets of initial values.[7]

[7] We define convergence using a relative tolerance of $10^{-4}$ for the log-likelihood. In our empirical application discussed in section 3 the unrestricted model involves $N(q(s+1)+1)+q^2(s+1)+\frac{q(q+1)}{2}$ parameters. Although the numbers of parameters to be estimated is considerable, it is computationally feasible with the EM algorithm. Doz et al. (2006) suggest to initialize the Kalman filter by the parameters implied by principal components and then filter the factors. We follow a somewhat different approach by filtering the factors implied by a loading structure imposing a one-to-one contemporaneous relation between the respective factors and their closest observable variable. All other loadings are initialized at zero. Given these loadings we filter the initial factors, using a dynamic representation with relatively small eigenvalues (i.e. 0.4.).

# 3 Empirical Application

We illustrate our procedure by revisiting the large data panel analyzed in Bernanke et al. (2005). This data set captures the dynamics of a wide range of macroeconomic developments in the US economy over the last decades. In particular, the sample consists of 120 time series (monthly frequency) over the period 1959 : 1 to 2001 : 8.[8] The main focus of our empirical analysis is to *i)* extract a number of factors with an unambiguous (macro) economic interpretation and *ii)* analyze the impact of monetary policy shocks on the US economy. We first discuss the identification restrictions of the factors and shocks. Subsequently, we analyze the extracted factors and we use impulse response functions (IRFs) and variance decompositions to study the impact of monetary policy shocks on the US economy.

## 3.1 Identification

The identification proceeds in two steps. First, we select the number of dynamic factors, $q$, and the number of lags, $p = s + 1$, in the VAR of the dynamic factors (see equation 4). Second, restrictions are imposed to identify and interpret in macroeconomic terms the factors and structural shocks.

### 3.1.1 Number of factors

Our preferred specification contains five dynamic factors $(\widehat{q} = 5)$. This choice is primarily based on information criteria of Hallin and Liska (2007), suggesting values of $q$ within a range of 3 to 6 (see Figure 1). The choice of five dynamic factors does not conflict with the statistical tests, and is in line with the range proposed in the literature. For example, Giannone et al. (2004) argue that the number of shocks (dynamic factors) driving the US economy is equal to two

[8] The data are already transformed by Bernanke et al. (2005) to reach stationarity; see Bernanke et al. (2005) for details on the data set and on the transformations. Prior to the estimation, we de-mean the series and divide them by their standard deviation such that the resulting series have zeros mean and unit variance.

(i.e. $\widehat{q} = 2$). Stock and Watson (2005) analyzing the same data set argue that seven dynamic factors and nine static factors are required ( $\widehat{q} = 7$). Bai and Ng (2007) and Hallin and Liska (2007) opt for $\widehat{q} = 4$. Bernanke et al. (2005) prefer a model specification with four factors ($\widehat{q} = 4$). Finally, Forni and Gambetti (2010), combining various information criteria, estimate the number of dynamic factors for the US economy in between 4 and 7.

Insert Figure 1 and Table 1

We select the number of lags in equation (4) based on an evaluation of the BIC and AIC information criteria. Specifically, for each lag specification, $p = 1, \ldots, 13$, of the dynamic factors we estimate the implied DFM, extract the identified economic factors and report the BIC and AIC of the respective transition equations. Table 1 shows that the optimal number of lags suggested by this procedure is $\widehat{p} = 2$. In the empirical section we will both discuss the optimal model $\widehat{p} = 2$ (implying 10 static factors) as well as, following Bernanke et al. (2005) or Banbura et al. (2010), a $\widehat{p} = 13$ model (implying 65 static factors).

#### 3.1.2 Economic interpretation of factors and shocks

We identify five dynamic factors, capturing a relatively wide array of economic concepts or interpretations, relevant for empirical monetary policy analysis. Given the economic interpretation of the five dynamic factors all remaining static factors (being the lagged values of identified dynamic factors) inherit the respective economic interpretations.

The identification of the first three dynamic factors is motivated by the small-scale theoretical macroeconomic models. In particular, we retain three main macroeconomic factors: an *inflation factor* ($\pi$); an *economic activity factor* ($y$); and a *monetary policy factor* ($i$).[9] Given that the focus of the empirical application is on the impact of monetary policy shocks, we additionally introduce two information factors, facilitating the identification of monetary policy shocks. Specifically,

[9] These are also some of the prime factors discussed by Stock and Watson (2005).

we identify a *commodity price factor* ($pcom$) capturing information on expected inflation pressures, and a *money market factor* ($m$) allowing to distinguish between general money market shocks and monetary policy shocks (see for instance Sims (1986)) and Christiano et al. (1999)).

The identification of the respective factors is obtained in two steps. In the first step, we fix the interpretation of the factors by imposing a set of unbiasedness restrictions on the five observables closest to the economic interpretation of each of the factors. This results in an exactly identified system (along the lines of *Proposition 2* in Geweke and Singleton (1981)). This exactly identified latent factor model is labelled as the "*unrestricted model*". The target observables of the factors are: the CPI-all items index (series 108) for the inflation factor ($\pi$); the Industrial Production-total index (series 16) for the economic activity factor ($y$); the Index of Sensitive Materials prices (series 107) for the commodity price factor ($pcom$); the effective federal funds rate (series 77) for the monetary policy rate factor ($i$).[10]

In the next step, over-identifying restrictions are imposed in the form of exclusion restrictions and additional unbiasedness restrictions. First, the specific set of additional over-identifying restrictions can be summarized as follows; the *inflation factor* ($\pi$) is identified by unbiasedness restrictions on all CPI and PPI variables (excluding commodity price indices). With the inflation factor being a nominal factor, we exclude from the information set all real variables, e.g. industrial production. The *economic activity factor* ($y$), identified by the unbiasedness restriction on all the Industrial Production (IP) indices, uses next to IP variables other real variables as information sources[11], including (un)employment, income, capacity utilization, and consumption series. The *monetary policy factor* ($i$) is uniquely identified by the federal funds rate while the *money market factor* ($m$) is filtered using unbiasedness restrictions on all monetary aggregates ($M_0, .., M_3$) and using deposit and credit variables as additional information variables. Finally, the *commodity price factor* ($pcom$) is identified through the unbiasedness restriction on the Index of Sensitive Material Prices while both the Crude Material PPI and

[10] See appendix A for the definition and numbers assigned to each observable in the data panel.

[11] Information variables are allowed to load on contemporaneous and on lagged factors. Hence, their loadings are estimated as free parameters.

NAPM commodity price indices are used as additional information variables.

Second, exclusion restrictions are primarily imposed on slow-moving variables.[12] This modeling choice is motivated by the idea that fast moving variables, containing a speculative component, can be considered as general and timely information variables for macroeconomic developments. We differentiate between nominal (inflation), real (economic activity), information[13] (money market and commodity price factors), and policy (Fed rate) factors. In our identification strategy, nominal factors exclude all types of real variables as (contemporaneous an lagged) information sources. In the same way, real factors exclude nominal variables. Information factors exclude all slow-moving real and nominal variables.

A final set of exclusion restrictions identifies the structural shocks through a standard Cholesky decomposition of the variance-covariance matrix of disturbances in the state equation (equation(6)). The ordering used in the analysis is as follows: $pcom$, $\pi$, $y$, $i$, $m$. This ordering is in line with the identification of monetary policy shocks in the literature (see for example Sims (1986) and Christiano et al. (1999)). Note that in the empirical application we focus on the impact of monetary policy shocks on the state of the economy. The exact ordering of shocks before the monetary policy shock hence does not matter and results will be robust against any reordering of $\pi, pcom$ or $y$.

## 3.2 Empirical Results

### 3.2.1 Model performance

In this section we provide a statistical test on the over-identifying restrictions. In particular, we perform an $LR$-test of our restricted model against the unrestricted

[12] We use the definition of fast- and slow-moving variables of Bernanke et al. (2005) except for monetary aggregates determining the money market factor or the NAPM commodity price index determining the commodity price factor.

[13] Information variables (or information factors) are assumed to be monitored by central banks because they may display relevant information that is not available in typical macroeconomic variables. See Leeper et al. (1996), Christiano et al. (1999) and very recently Bjørnland and Leitemo (2009) for a discussion. Examples of fast moving variables include auction market commodity prices, stock prices, and options on financial instruments.

(exactly identified) model. We complement this test by a number of measures of fit including $R^2$, $AIC$, $BIC$ and the log-likelihood value. Table 2 reports the results. As expected, the over-identifying restrictions are rejected at the usual significance level. Moreover, there is a significant drop in the overall explanatory power of the model. While the unrestricted model explains about 50 percent our specification has an average R-squared of about 40 percent. While this drop in overall explanatory power is significant, we gain an unambiguous interpretation of the underlying factors. Interestingly both the AIC ad BIC information criteria marginally prefer the restricted version of the model, which is due to the large number of restrictions imposed (see Table 2).

Insert Table 2

#### 3.2.2 Implied factors

Figure 2 presents the times series of the filtered factors (with imposed economic interpretation) together with the main observed target variables for the VAR(2) model with all restrictions imposed. For instance, the top left panel shows the extracted inflation factor (obtained from the imposed restrictions) and contrasts it with the realized CPI (all items) inflation (series 108). For economic activity (top right panel), we display the retrieved factor and contrast it with the growth rate of total industrial production (series 16). Visual inspection shows that the retrieved factors capture well the low and medium frequency dynamics of the target series. In particular, the inflation factor captures about 40 percent of the overall variation in the CPI all-items series, and over 61 percent of the low and intermediate variation.[14] Analogously, the economic activity factor captures about 32 percent of the growth variation in industrial production total index (all items). Removing short term variation in the growth rate, this explained variation increases to over 64 percent. For the commodity price index we only explain about 28 percent of total variation. The money market factor explains

[14]We use an HP filter to filter out the higher frequency components. To that end we use a HP filter with $\lambda = 100$. This choice of lambda allows sufficient amount of variability in the trend time series.

more than 32 percent of total variation for the M1 money growth rate, and about 33 percent of the low and intermediate frequencies. The monetary policy factor fits the Fed rate by construction.

Insert Figures 2 till 6

Additionally, Figures 3 till 6 contrast the respective factors to each of the series that were considered informative for the respective factors i.e. were subjected to unbiasedness restrictions. The main conclusion emerging from these figures is that, overall, the respective factors capture well the underlying low frequency dynamics in each of the observed series. These observations corroborate the economic interpretation of the factors as imposed through the cross-equation restrictions, allowing to interpret the retrieved factors according to the targeted economic concepts.

Figures 7 till 8 give a graphical representation of the estimated factor loadings for each of the 120 observable series for each of the factors.[15] Note that in the figure we sum the loadings per factor over the lags.

Insert Figures 7 and 8

The figures are illustrative for two reasons. First, they clearly express the set of restrictions imposed in the identification procedure and second, they display the unrestricted loadings. As can be observed from panel (a) of Figure (7), loading restrictions for the inflation factor impose a unit loading on all observed inflation series (i.e. series 102 till 117) and zero restrictions on slow-moving real variables. Equivalently, for the economic activity factor unit loading restrictions were imposed on all production indices (series 1 till 16) where we allow for additional free loading estimates on other real variables (e.g. employment, consumption, income...). In the same line of reasoning, the commodity price factor loads on index of sensitive material while the money market factor loads on respective money supply series. Finally, we use a one to one relation between the monetary policy factor and the effective federal funds rate. In general no loading restric-

[15] The list of variables and their exact definition, along with the qualification of fast- or slow moving variables can be found in Appendix 1.

tions were imposed on the fast moving variables. Instead, these loadings were estimated freely and as can be observed these variables often load significantly on the respective factors. For example, we find significant and positive loadings of stock market returns, housing starts and inventories on the economic activity factor.

#### 3.2.3 Measuring the impact of monetary policy

We use our model to analyze the overall impact of monetary policy shocks on the US economy. To facilitate comparison with the literature we perform two types of analyses. First, following e.g. Bernanke et al. (2005), Forni and Gambetti (2010) and Marcellino et al. (2005) we compare the IRF implied by our DFM to the standard SVAR specification. Second, as in Bernanke et al. (2005), Banbura et al. (2010) and Forni and Gambetti (2010), we focus on the impact of monetary policy shocks on twenty key indicators for the US economy.

Insert Figure 9

Figure 9 compares the IRFs of a 25 basis points contractionary monetary policy shocks between the SVAR (left panel) and our DFM (right panel). All reported IRFs are measured in terms of unconditional standard deviations. We focus on the VAR (2) specification, modeling the dynamics of the five target series; CPI-all items; commodity price index, money (M1), and the Fed rate. In line with the literature, the IRFs implied by the SVAR display both the price puzzle and overall smaller and more delayed responses to monetary policy shocks. In contrast, the IRFs implied by our DFM are theoretically more plausible. First, we no longer observe the price puzzle. Specifically, we identify a significant decrease in prices following a contractionary monetary policy shock. This result corroborates recent findings in the DFM literature attributing the price puzzle to the restrictive nature of information set used in the SVAR estimation (see for instance, Bernanke et al. (2005), Forni and Gambetti (2010) and Marcellino et al. (2005)). Note however, that using our identification scheme, we recover a statistically significant price effect of monetary policy shocks. Second, in accordance with Forni and

Gambetti (2010) we find larger impact of monetary policy shocks. For instance, our results show that a 25 basis points contractionary monetary policy shock leads to a 25 basis points decrease in the price level at the the two year horizon. Also, the maximum effect, a 50 basis points drop, on industrial production is reached after 1.5 years. This typical hump-shape IRF is more in line with the theory than the strongly delayed response of industrial production as implied by the SVAR.

We now turn to the second part, analyzing in more detail the impact of monetary policy shocks on the US economy. More specifically, we analyze the IRFs of the following economic indicators: federal funds rate, the yen per US dollar exchange rate, the level of industrial production, the consumer price level (CPI), monetary aggregates, the capacity utilization, the (un)employment level, the average hourly earnings, the level of consumption and consumer confidence expectations as key indicators for the macroeconomy. Additionally, we cover housing starts and two financial market indicators: the return on the NYSE composite and the five year treasury yield. We present two versions of the impulse response functions, a first one based on the parsimonious two lag model and an additional one where we allow for thirteen lags. We include the latter model as the thirteen lag model to compare our results with Bernanke et al. (2005) and Banbura et al. (2010).

Insert Figures 10 and 11

Figures 10 and 11 display the IRFs of each of these variables to a 25 basis points monetary policy shock for respectively the two and the thirteen lag models. The sign and magnitude of the IRFs are in line with the literature (see Christiano et al. (1999)) and suggest that the model is able to identify accurately the key macroeconomic transmission mechanisms for the monetary policy shock.

Several observations can be made in this respect. First, unlike standard small-scale VAR models, we do no longer observe a persistent price puzzle. In the two lag model no price puzzle appears, while in the 13 lag model we observe a short-lived and small initial price increase, followed by strong price drops. Second, in line with the findings in the literature, a contractionary monetary policy shock has a negative impact on production where the maximal effect is reached within one to three years, depending on the model specification. Note that similar hump-shape (shorter lived) patterns are observed for alternative measures of economic activity e.g. capacity utilization, employment. Third, long-run neutrality of monetary policy cannot be rejected. In particular, monetary policy shocks only have a temporary effect on production, consumption, capacity utilization, and (un)employment levels. Our IRFs (not shown) indicate a return to the pre-shock situation within a ten year period. Fourth, the impact of temporary policy shocks is initially negative on the consumption expectations but then reverses before the impact becomes neutral in the long-run. Finally, the results show a significant impact of monetary policy shocks on financial markets. Monetary policy tightening increases the bond yields with the short-term yields responding more than the long-term yields, as illustrated by the IRF of the 3 month and 5 year yield. However, given the moderate persistence of the policy shocks (see the IRF of the federal funds rate), the impact on bond yields of monetary policy shocks remains relatively small and temporary. Real estate markets, as illustrated by the IRF of the housing starts, initially contract strongly to the monetary policy shock although there is no long-run effect. Following a monetary tightening, we observe an initial drop in the stock prices while the yen tends to depreciate against the US dollar. These IRFs match both the responses reported in Banbura et al. (2010), using a BVAR and Bernanke et al. (2005) and Forni and Gambetti (2010) using factor models. Moreover, the main features of these IRFs are robust to the specification of the lag length (for lags between 2 and 13) of the model.

Insert Table 3

Table 3 presents the variance decomposition of the selected variables at alternative forecasting horizons. This tool allows us to assess the relative importance of monetary policy shocks in the overall variation of the series. Our results are broadly in line with those reported both in Forni and Gambetti (2010). In line with this study, we observe that monetary shocks have an important long-run (60 month) impact on the forecast error variance of a broad selection of key macroeconomic and financial variables. Specifically, we find that a monetary policy shock explains in between 10% and 30% of the variation in industrial production, consumer prices, commodity prices, (un)employment, new orders, and housing. Moreover, in line with results reported in Banbura et al. (2010), using a BVAR and Bernanke et al. (2005) monetary policy shocks explain only marginal amounts of the variation in consumption indicators. Finally, the estimates reported in Table 3 indicate that monetary policy shocks are mildly persistent and account for approximately 27 percent of total long-run variation in the federal funds rate (at the 60 month horizon) and of about 20 percent for the bond yields. Our estimates hence suggest a smaller long-run impact of monetary policy shocks than in for instance Bernanke et al. (2005), reporting up to 45 percent. More in line with our results are Banbura et al. (2010) and Forni and Gambetti (2010), reporting similarly small numbers.

# 4 Conclusion

This paper has proposed a methodology to identify factors within the framework of dynamic factor models. We impose an economic interpretation on the (static and dynamic) factors through a set of over-identifying restrictions on the factor loadings. We modify the standard estimation methodology to incorporate these over-identifying loading restrictions. In particular, following Shumway and Stoffer (1982) and Wu et al. (1996), the appropriate parameter estimators and filters based on the EM algorithm are discussed.

In the empirical application the paper focuses on identifying a set of five factors with economic interpretation. These factors represent key measures of the US economy such as inflation, economic activity, commodity prices, money market pressure and monetary policy. The obtained factors are empirically plausible measures for each of the targeted key concepts, listed above. Subsequently, we use the model to assess the overall impact of monetary policy on the US economy. Our results are in line with those obtained using alternative methods on large panels, e.g. FAVARs or large BVARs, and suggest identify an important role for monetary policy shocks in macroeconomic dynamics.

# References


Alessi, L., Barigozzi, M. & Capasso, M. (2007), 'Dynamic factor GARCH: Multivariate volatility forecast for a large number of series', *Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies* .

Altissimo, F., Cristadoro, R., Forni, M., Lippi, M. & Veronese, G. (2007), 'New eurocoin: Tracking economic growth in real time', *The Review of Economics and Statistics.(forthcoming)* (631).

Altug, S. (1989), 'Time to build and aggregate fluctuations: some new evidence', *International Economic Review* **30**(4), 889–920.

Bai, J. & Ng, S. (2007), 'Determining the number of primitive shocks in factor models', *Journal of Business & Economic Statistics* **25**, 52–60.

Banbura, M., Giannone, D. & Reichlin, L. (2010), 'Large bayesian VARs', *Journal of Applied Econometrics* **25(1)**, 71–92.

Bernanke, B. S. & Boivin, J. (2003), 'Monetary policy in a data-rich environment', *Journal of Monetary Economics* **50**(3), 525–546.

Bernanke, B. S., Boivin, J. & Eliasz, P. (2005), 'Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach', *The Quarterly Journal of Economics* pp. 387–422.

Bjørnland, H. & Leitemo, K. (2009), 'Identifying the interdependence between US monetary policy and the stock market', *Journal of Monetary Economics (forthcoming)* .

Blanchard, O. J. & Quah, D. (1989), 'The dynamic effects of aggregate demand and supply disturbances', *American Economic Review* **79**(4), 655–73.

Boivin, J. & Giannoni, M. (2006), DSGE models in a data-rich environment, Working Paper 12772, National Bureau of Economic Research.

Boivin, J., Giannoni, M. P. & Mihov, I. (2009), 'Sticky prices and monetary policy: Evidence from disaggregated us data', *American Economic Review* **99**(1), 350–84.

Bork, L. (2008), Estimating US monetary policy shocks using a factor-augmented vector autoregression: An EM algorithm approach. CREATES Research Paper 2009-11, University of Aarhus.

Campbell, J. Y., Lo, A. & MacKinlay, C. (1997), *The Econometrics of Financial Markets*, Princeton University Press, Princeton.

Chamberlain, G. & Rothschild, M. (1983), 'Arbitrage, factor structure, and mean-variance analysis on large asset markets', *Econometrica* **51**(5), 1281–1304.

Christiano, L. J., Eichenbaum, M. & Evans, C. L. (1999), Monetary policy shocks: What have we learned and to what end?, *in* J. B. Taylor & M. Woodford, eds, 'Handbook of Macroeconomics', Vol. 1A, Elsevier Science, North Holland, New York, pp. 65–148.

de Jong, P. (1989), 'Smoothing and interpolation with the state-space model', *Journal of the American Statistical Association* **84**, 1085–1088.

de Jong, P. & Mackinnon, M. J. (1988), 'Covariances for smoothed estimates in state space models', *Biometrika* (75), 601–602.

Diebold, F. X., Li, C. & Yue, V. Z. (2008), 'Global yield curve dynamics and interactions: A dynamic Nelson-Siegel approach', *Journal of Econometrics* **146**(2), 351–363.

Doz, C., Giannone, D. & Reichlin, L. (2006), A quasi maximum likelihood approach for large approximate dynamic factor models. European Central Bank Working Paper no. 674.

Doz, C., Giannone, D. & Reichlin, L. (2007), A two-step estimator for large approximate dynamic factor models based on Kalman filtering, CEPR Discussion Paper 6043, Centre for Economic Policy Research.

Forni, M. & Gambetti, L. (2010), 'The dynamic effects of monetary policy: A structural factor model approach', *Journal of Monetary Economics* **57**(2), 203–216.

Forni, M., Giannone, D., Lippi, M. & Reichlin, L. (2008), 'Opening the black box: Structural factor models with large cross-sections', *Econometric Theory, Vol. 25, No. 05, Pages 1319-1347* .

Forni, M., Hallin, M., Lippi, M. & Reichlin, L. (2000), 'The generalized dynamic-factor model: Identification and estimation', *The Review of Economics and Statistics* **82**(4), 540–554.

Forni, M., Hallin, M., Lippi, M. & Reichlin, L. (2001), 'Coincident and leading indicators for the euro area', *Economic Journal* **111**(471), 62–85.

Forni, M., Hallin, M., Lippi, M. & Reichlin, L. (2005), 'The generalized dynamic factor model: one sided estimation and forecasting', *Journal of the American Statistical Association* **100**(471), 830–840.

Forni, M. & Reichlin, L. (1998), 'Let's get real: A factor analytical approach to disaggregated business cycle dynamics', *Review of Economic Studies* **65**(3), 453–73.

Forni, M. & Reichlin, L. (2001), 'Federal policies and local economies: Europe and the us', *European Economic Review* **45**(1), 109–134.

Geweke, J. F. & Singleton, K. J. (1981), 'Maximum likelihood "confirmatory" factor analysis of economic time series', *International Economic Review* **22**(1), 37–54.

Giannone, D., Reichlin, L. & Sala, L. (2002), Tracking Greenspan: Systematic and unsystematic monetary policy revisited. CEPR Discussion Papers no. 3550.

Giannone, D., Reichlin, L. & Sala, L. (2004), Monetary policy in real time, *in* M. Gertler & K. Rogoff, eds, 'NBER Macroeconomic Annual 2004', Vol. 19, MIT Press.

Hallin, M. & Liska, R. (2007), 'Determining the number of factors in the generalized factor model', *Journal of the American Statistical Association* **102**, 603–617.

Houssa, R. (2008*a*), 'Monetary union in west africa and asymmetric shocks: A dynamic structural factor model approach', *Journal of Development Economics* **Vol. 85**, 319–347.

Houssa, R. (2008*b*), Sources of Fluctuations: World, Regional, and National Factors, In Macroeconomic Fluctuations in Developing Countries, PhD thesis, KULeuven.

Ingersoll, Jonathan E, J. (1984), 'Some results in the theory of arbitrage pricing', *Journal of Finance* **39**(4), 1021–39.

Jungbacker, B. & Koopman, S. J. (2008), Likelihood-based analysis for dynamic factor models, Tinbergen Institute Discussion Papers 08-007/4, Tinbergen Institute.

Kim, C.-J. & Nelson, C. R. (1999), *State-Space Models with Regime Switching*, The MIT Press.

Kose, A. M., Otrok, C. & Whiteman, C. (2003), 'International business cycles: World, region, and country-specific factors', *American Economic Review* **93**(4), 1216–1239.

Leeper, E. M., Sims, C. A. & Zha, T. (1996), 'What does monetary policy do?', *Brookings Papers on Economic Activity* **27**(1996-2), 1–78.

Ludvigson, S. C. & Ng, S. (2007), 'The empirical risk-return relation: A factor analysis approach', *Journal of Financial Economics* **83**(1), 171–222.

Ludvigson, S. C. & Ng, S. (2009), 'Macro factors in bond risk premia', *The Review of Financial Studies* **22(12)**, 5027–506.

Marcellino, M., Favero, C. A. & Neglia, F. (2005), 'Principal components at work: the empirical analysis of monetary policy with large data sets', *Journal of Applied Econometrics* **20**(5), 603–620.

Mönch, E. (2008), 'Forecasting the yield curve in a data-rich environment: A no-arbitrage factor-augmented VAR approach', *Journal of Econometrics* **146**(1), 26–43.

Otrok, C. & Whiteman, C. H. (1998), 'Bayesian leading indicators: Measuring and predicting economic conditions in Iowa', *International Economic Review* **39**(4), 997–1014.

Reis, R. & Watson, M. W. (2008), Relative goods' prices, pure inflation, and the Philips correlation. American Economic Journal Macroeconomics (forthcoming).

Sargent, T. J. (1989), 'Two models of measurements and the investment accelerator', *Journal of Political Economy* **97**(2), 251–287.

Sargent, T. J. & Sims, C. A. (1977), Business cycle modeling without pretending to have too much a priori economic theory, Working Papers 55, Federal Reserve Bank of Minneapolis.

Shumway, R. H. & Stoffer, D. S. (1982), 'An approach to time series smoothing and forecasting using the EM algorithm', *Journal of Time Series Analysis* **3**, 253–226.

Sims, C. (1980), 'Macroeconomics and reality', *Econometrica* **48**, 1–48.

Sims, C. A. (1986), 'Are forecasting models usable for policy analysis?', *Quarterly Review* (win), 2–16.

Stock, J. H. & Watson, M. W. (2002*a*), 'Forecasting using principal components from a large number of predictors', *Journal of the American Statistical Association* **97**, 1167–1179.

Stock, J. H. & Watson, M. W. (2002*b*), 'Macroeconomic forecasting using diffusion indexes', *Journal of Business Economics and Statistics* **XX:II**, 147–162.

Stock, J. H. & Watson, M. W. (2005), Implications of dynamic factor models for VAR analysis, Working Paper 11467, National Bureau of Economic Research.

Uhlig, H. (2005), 'What are the effects of monetary policy on output? Results from an agnostic identification procedure', *Journal of Monetary Economics* **52**(2), 381–419.

Wu, L. S.-Y., Pai, J. S. & Hosking, J. (1996), 'An algorithm for estimating parameters of state-space models', *Statistics and Probability Letters* **28**, 99–106.

# A Data description

Data are from Bernanke et al. (2005).

The second column is a mnemonic and a * indicates a "slow-moving" variable. Fourth column contains transformation codes. "level" indicates an un-transformed variable, say $x_t$. "ln" means $\ln x_t$ and "$\Delta \ln$" means $\ln x_t - \ln x_{t-1}$.

Real output and income

| | | | | |
|---|---|---|---|---|
| 1 | IPP* | 1959:01–2001:08 | $\Delta \ln$ | Industrial production: products, total (1992 = 100,SA) |
| 2 | IPF* | 1959:01–2001:08 | $\Delta \ln$ | Industrial production: final products (1992 = 100,SA) |
| 3 | IPC* | 1959:01–2001:08 | $\Delta \ln$ | Industrial production: consumer goods (1992 = 100,SA) |
| 4 | IPCD* | 1959:01–2001:08 | $\Delta \ln$ | Industrial production: durable cons. goods (1992 = 100,SA) |
| 5 | IPCN* | 1959:01–2001:08 | $\Delta \ln$ | Industrial production: nondurable cons. goods (1992 = 100,SA) |
| 6 | IPE* | 1959:01–2001:08 | $\Delta \ln$ | Industrial production: business equipment (1992 = 100,SA) |
| 7 | IPI* | 1959:01–2001:08 | $\Delta \ln$ | Industrial production: intermediate products (1992 = 100,SA) |
| 8 | IPM* | 1959:01–2001:08 | $\Delta \ln$ | Industrial production: materials (1992 = 100,SA) |
| 9 | IPMD* | 1959:01–2001:08 | $\Delta \ln$ | Industrial production: durable goods materials (1992 = 100,SA) |
| 10 | IPMND* | 1959:01–2001:08 | $\Delta \ln$ | Industrial production: nondur. goods materials (1992 = 100,SA) |
| 11 | IPMFG* | 1959:01–2001:08 | $\Delta \ln$ | Industrial production: manufacturing (1992 = 100,SA) |
| 12 | IPD* | 1959:01–2001:08 | $\Delta \ln$ | Industrial production: durable manufacturing (1992 = 100,SA) |
| 13 | IPN* | 1959:01–2001:08 | $\Delta \ln$ | Industrial production: nondur. manufacturing (1992 = 100,SA) |
| 14 | IPMIN* | 1959:01–2001:08 | $\Delta \ln$ | Industrial production: mining (1992 = 100,SA) |
| 15 | IPUT* | 1959:01–2001:08 | $\Delta \ln$ | Industrial production: utilities (1992 = 100,SA) |
| 16 | IP* | 1959:01–2001:08 | $\Delta \ln$ | Industrial production: total index (1992 = 100,SA) |
| 17 | IPXMCA* | 1959:01–2001:08 | level | Capacity util rate: manufac., total (% of capacity,SA) (frb) |
| 18 | PMI* | 1959:01–2001:08 | level | Purchasing managers' index (SA) |
| 19 | PMP* | 1959:01–2001:08 | level | NAPM production index (percent) |
| 20 | GMPYQ* | 1959:01–2001:08 | $\Delta \ln$ | Personal income (chained) (series #52) (bil 92$,SAAR) |
| 21 | GMYXPQ* | 1959:01–2001:08 | $\Delta \ln$ | Personal inc. less trans. payments (chained) (#51) (bil 92$,SAAR) |

(Un)employment and hours

| | | | | |
|---|---|---|---|---|
| 22 | LHEL* | 1959:01–2001:08 | $\Delta$ ln | Index of help-wanted advertising in newspapers (1967 = 100;SA) |
| 23 | LHELX* | 1959:01–2001:08 | ln | Employment: ratio; help-wanted ads: no. unemployed clf |
| 24 | LHEM* | 1959:01–2001:08 | $\Delta$ ln | Civilian labor force: employed, total (thous.,SA) |
| 25 | LHNAG* | 1959:01–2001:08 | $\Delta$ ln | Civilian labor force: employed, nonag. industries (thous.,SA) |
| 26 | LHUR* | 1959:01–2001:08 | level | Unemployment rate: all workers, 16 years and over (%,SA) |
| 27 | LHU680* | 1959:01–2001:08 | level | Unemploy. by duration: average (mean) duration in weeks (SA) |
| 28 | LHU5* | 1959:01–2001:08 | level | Unemploy. by duration: pers unempl. less than 5 wks (thous.,SA) |
| 29 | LHU14* | 1959:01–2001:08 | level | Unemploy. by duration: pers unempl. 5 to 14 wks (thous.,SA) |
| 30 | LHU15* | 1959:01–2001:08 | level | Unemploy. by duration: pers unempl. 15 wks = (thous.,SA) |
| 31 | LHU26* | 1959:01–2001:08 | level | Unemploy. by duration: pers unempl. 15 to 26 wks (thous.,SA) |
| 32 | LPNAG* | 1959:01–2001:08 | $\Delta$ ln | Employees on nonag. payrolls: total (thous.,SA) |
| 33 | LP* | 1959:01–2001:08 | $\Delta$ ln | Employees on nonag. payrolls: total, private (thous.,SA) |
| 34 | LPGD* | 1959:01–2001:08 | $\Delta$ ln | Employees on nonag. payrolls: goods-producing (thous.,SA) |
| 35 | LPMI* | 1959:01–2001:08 | $\Delta$ ln | Employees on nonag. payrolls: mining (thous.,SA) |
| 36 | LPCC* | 1959:01–2001:08 | $\Delta$ ln | Employees on nonag. payrolls: contract construc. (thous.,SA) |
| 37 | LPEM* | 1959:01–2001:08 | $\Delta$ ln | Employees on nonag. payrolls: manufacturing (thous.,SA) |
| 38 | LPED* | 1959:01–2001:08 | $\Delta$ ln | Employees on nonag. payrolls: durable goods (thous.,SA) |
| 39 | LPEN* | 1959:01–2001:08 | $\Delta$ ln | Employees on nonag. payrolls: nondurable goods (thous.,SA) |
| 40 | LPSP* | 1959:01–2001:08 | $\Delta$ ln | Employees on nonag. payrolls: service-producing (thous.,SA) |
| 41 | LPTU* | 1959:01–2001:08 | $\Delta$ ln | Employees on nonag. payrolls: trans. and public util. (thous.,SA) |
| 42 | LPT* | 1959:01–2001:08 | $\Delta$ ln | Employees on nonag. payrolls: wholesale and retail (thous.,SA) |
| 43 | LPFR* | 1959:01–2001:08 | $\Delta$ ln | Employees on nonag. payrolls: finance, ins. and real est (thous.,SA) |
| 44 | LPS* | 1959:01–2001:08 | $\Delta$ ln | Employees on nonag. payrolls: services (thous.,SA) |
| 45 | LPGOV* | 1959:01–2001:08 | $\Delta$ ln | Employees on nonag. payrolls: government (thous.,SA) |
| 46 | LPHRM* | 1959:01–2001:08 | level | Avg. weekly hrs. of production wkrs.: manufacturing (sa) |
| 47[5] | LPMOSA* | 1959:01–2001:08 | level | Avg. weekly hrs. of prod. wkrs.: mfg., overtime hrs. (sa) |
| 48 | PMEMP* | 1959:01–2001:08 | level | NAPM employment index (percent) |

Consumption

| | | | | |
|---|---|---|---|---|
| 49[4] | GMCQ* | 1959:01–2001:08 | $\Delta$ ln | Pers cons exp (chained)—total (bil 92$,SAAR) |
| 50 | GMCDQ* | 1959:01–2001:08 | $\Delta$ ln | Pers cons exp (chained)—tot. dur. (bil 96$,SAAR) |
| 51 | GMCNQ* | 1959:01–2001:08 | $\Delta$ ln | Pers cons exp (chained)—nondur. (bil 92$,SAAR) |
| 52 | GMCSQ* | 1959:01–2001:08 | $\Delta$ ln | Pers cons exp (chained)—services (bil 92$,SAAR) |
| 53 | GMCANQ* | 1959:01–2001:08 | $\Delta$ ln | Personal cons expend (chained)—new cars (bil 96$,SAAR) |

Housing starts and sales

| | | | | |
|---|---|---|---|---|
| 54 | HSFR | 1959:01–2001:08 | ln | Housing starts: nonfarm (1947–1958); tot. ( |
| 55 | HSNE | 1959:01–2001:08 | ln | Housing starts: northeast (thous.u.)s.a. |
| 56 | HSMW | 1959:01–2001:08 | ln | Housing starts: midwest (thous.u.)s.a. |
| 57 | HSSOU | 1959:01–2001:08 | ln | Housing starts: south (thous.u.)s.a. |
| 58 | HSWST | 1959:01–2001:08 | ln | Housing starts: west (thous.u.)s.a. |
| 59 | HSBR | 1959:01–2001:08 | ln | Housing authorized: total new priv housing (thous.,SAAR) |
| 60 | HMOB | 1959:01–2001:08 | ln | Mobile homes: manufacturers' shipments (thous. of units,SAAR) |

Real inventories, ordes and unfilled orders

| | | | | |
|---|---|---|---|---|
| 61 | MNV | 1959:01–2001:08 | level | NAPM inventories index (percent) |
| 62 | PMNO | 1959:01–2001:08 | level | NAPM new orders index (percent) |
| 63 | PMDEL | 1959:01–2001:08 | level | NAPM vendor deliveries index (percent) |
| 64 | MOCMQ | 1959:01–2001:08 | $\Delta \ln$ | New orders (net)—consumer goods and materials, 1992 $ (bci) |
| 65 | MSONDQ | 1959:01–2001:08 | $\Delta \ln$ | New orders, nondefense capital goods, in 1992 $s (bci) |

Stock prices

| | | | | |
|---|---|---|---|---|
| 66 | FSNCOM | 1959:01–2001:08 | $\Delta \ln$ | NYSE composite (12/31/65 = 50) |
| 67 | FSPCOM | 1959:01–2001:08 | $\Delta \ln$ | S&P's composite (1941–1943 = 10) |
| 68 | FSPIN | 1959:01–2001:08 | $\Delta \ln$ | S&P's industrials (1941–1943 = 10) |
| 69 | FSPCAP | 1959:01–2001:08 | $\Delta \ln$ | S&P's capital goods (1941–1943 = 10) |
| 70 | FSPUT | 1959:01–2001:08 | $\Delta \ln$ | S&P's utilities (1941–1943 = 10) |
| 71 | FSDXP | 1959:01–2001:08 | level | S&P's composite common stock: dividend yield (% per annum) |
| 72 | FSPXE | 1959:01–2001:08 | level | S&P's composite common stock: price-earnings ratio (%,NSA) |

Foreign exchange rates

| | | | | |
|---|---|---|---|---|
| 73 | EXRSW | 1959:01–2001:08 | $\Delta \ln$ | Foreign exchange rate: Switzerland (swiss franc per US$) |
| 74 | EXRJAN | 1959:01–2001:08 | $\Delta \ln$ | Foreign exchange rate: Japan (yen per US$) |
| 75 | EXRUK | 1959:01–2001:08 | $\Delta \ln$ | Foreign exchange rate: United Kingdom (cents per pound) |
| 76 | EXRCAN | 1959:01–2001:08 | $\Delta \ln$ | Foreign exchange rate: Canada (canadian $ per US$) |

Interest rates and spreads

| | | | | |
|---|---|---|---|---|
| 77 | FYFF | 1959:01–2001:08 | level | Interest rate: federal funds (effective) (% per annum,nsa) |
| 78 | FYGM3 | 1959:01–2001:08 | level | Interest rate: us tbill,sec mkt,3-mo. (% per ann,nsa) |
| 79 | FYGM6 | 1959:01–2001:08 | level | Interest rate: us tbill,sec mkt,6-mo. (% per ann,nsa) |
| 80 | FYGT1 | 1959:01–2001:08 | level | Interest rate: ust const matur., 1-yr. (% per ann,nsa) |
| 81 | FYGT5 | 1959:01–2001:08 | level | Interest rate: ust const matur., 5-yr. (% per ann,nsa) |
| 82 | FYGT10 | 1959:01–2001:08 | level | Interest rate: ust const matur., 10-yr. (% per ann,nsa) |
| 83 | FYAAAC | 1959:01–2001:08 | level | Bond yield: moody's aaa corporate (% per annum) |
| 84 | FYBAAC | 1959:01–2001:08 | level | Bond yield: moody's baa corporate (% per annum) |
| 85 | SFYGM3 | 1959:01–2001:08 | level | Spread fygM3—fyff |
| 86 | SFYGM6 | 1959:01–2001:08 | level | Spread fygm6—fyff |
| 87 | SFYGT1 | 1959:01–2001:08 | level | Spread fygt1—fyff |
| 88 | SFYGT5 | 1959:01–2001:08 | level | Spread fygt5—fyff |
| 89 | SFYGT10 | 1959:01–2001:08 | level | Spread fygt10—fyff |
| 90 | SFYAAAC | 1959:01–2001:08 | level | Spread fyaaac—fyff |
| 91 | SFYBAAC | 1959:01–2001:08 | level | Spread fybaac—fyff |

## Money and credit quantity aggregates

| | | | | |
|---|---|---|---|---|
| 92 | FM1 | 1959:01–2001:08 | $\Delta$ ln | Money stock: M1 (bil$,SA) |
| 93 | FM2 | 1959:01–2001:08 | $\Delta$ ln | Money stock: M2 (bil$,SA) |
| 94 | FM3 | 1959:01–2001:08 | $\Delta$ ln | Money stock: M3 (bil$,SA) |
| 95 | FM2DQ | 1959:01–2001:08 | $\Delta$ ln | Money supply—M2 in 1992 $s (bci) |
| 96 | FMFBA | 1959:01–2001:08 | $\Delta$ ln | Monetary base, adj for reserve requirement changes (mil$,SA) |
| 97 | FMRRA | 1959:01–2001:08 | $\Delta$ ln | Depository inst reserves: total, adj for res. req chgs (mil$,SA) |
| 98 | FMRNBA | 1959:01–2001:08 | $\Delta$ ln | Depository inst reserves: nonbor., adj res req chgs (mil$,SA) |
| 99 | FCLNQ | 1959:01–2001:08 | $\Delta$ ln | Commercial and indust. loans outstanding in 1992 $s (bci) |
| 100 | FCLBMC | 1959:01–2001:08 | level | Wkly rp lg com. banks: net change com and ind. loans (bil$,SAAR) |
| 101 | CCINRV | 1959:01–2001:08 | $\Delta$ ln | Consumer credit outstanding nonrevolving g19 |

## Price indexes

| | | | | |
|---|---|---|---|---|
| 102 | PMCP | 1959:01–2001:08 | level | NAPM commodity prices index (%) |
| 103 | PWFSA* | 1959:01–2001:08 | $\Delta$ ln | PPI: finished goods (82 = 100,SA) |
| 104 | PWFCSA* | 1959:01–2001:08 | $\Delta$ ln | PPI: finished consumer goods (82 = 100,SA) |
| 105 | PWIMSA* | 1959:01–2001:08 | $\Delta$ ln | PPI: intermed mat. sup and components (82 = 100,SA) |
| 106 | PWCMSA* | 1959:01–2001:08 | $\Delta$ ln | PPI: crude materials (82 = 100,SA) |
| 107 | PSM99Q* | 1959:01–2001:08 | $\Delta$ ln | Index of sensitive materials prices (1990 = 100) (bci-99a) |
| 108 | PUNEW* | 1959:01–2001:08 | $\Delta$ ln | CPI-u: all items (82–84 = 100,SA) |
| 109 | PU83* | 1959:01–2001:08 | $\Delta$ ln | CPI-u: apparel and upkeep (82–84 = 100,SA) |
| 110 | PU84* | 1959:01–2001:08 | $\Delta$ ln | CPI-u: transportation (82–84 = 100,SA) |
| 111 | PU85* | 1959:01–2001:08 | $\Delta$ ln | CPI-u: medical care (82–84 = 100,SA) |
| 112 | PUC* | 1959:01–2001:08 | $\Delta$ ln | CPI-u: commodities (82–84 = 100,SA) |
| 113 | PUCD* | 1959:01–2001:08 | $\Delta$ ln | CPI-u: durables (82–84 = 100,SA) |
| 114 | PUS* | 1959:01–2001:08 | $\Delta$ ln | CPI-u: services (82–84 = 100,SA) |
| 115 | PUXF* | 1959:01–2001:08 | $\Delta$ ln | CPI-u: all items less food (82–84 = 100,SA) |
| 116 | PUXHS* | 1959:01–2001:08 | $\Delta$ ln | CPI-u: all items less shelter (82–84 = 100,SA) |
| 117 | PUXM* | 1959:01–2001:08 | $\Delta$ ln | CPI-u: all items less medical care (82– 84 = 100,SA) |

## Average hourly earnings

| | | | | |
|---|---|---|---|---|
| 118 | LEHCC* | 1959:01–2001:08 | $\Delta$ ln | Avg hr earnings of constr wkrs: construction ($,SA) |
| 119 | LEHM* | 1959:01–2001:08 | $\Delta$ ln | Avg hr earnings of prod wkrs: manufacturing ($,SA) |

## Miscellaneous

| | | | | |
|---|---|---|---|---|
| 120 | HHSNTN | 1959:01–2001:08 | level | U. of mich. index of consumer |

**Table 1: Lag Selection transition equation.**

| Lag: $p = s + 1$ | No. static Factor: $r$ | AIC | BIC | $R^2$ (total model) |
|---|---|---|---|---|
| Lag 1 | 5 | -15.984 | -15.777 | 36.86 |
| Lag 2 | 10 | **-19.814** | **-19.399** | 40.07 |
| Lag 3 | 15 | -16.441 | -15.82 | 42.07 |
| Lag 4 | 20 | -19.492 | -18.663 | 43.70 |
| | ... | ... | ... | ... |
| Lag13 | 65 | -14.717 | -12.023 | 47.95 |

Notes: $R^2$ is a simple average of the R-squared of the 120 series; *AIC* denotes Akaike Information Criterion; *BIC* is Bayesian Information Criterion

**Table 2: Model Performance.**

| | $R^2$ | *AIC* | *BIC* | *Log Lik* | $p$-value for *LR test* |
|---|---|---|---|---|---|
| Exactly Identified Model | 49.8 | 1.797 | 1.907 | $-54371$ | – |
| Our (Restricted) Model | 40.1 | 1.789 | 1.799 | $-54787$ | 0.0000 |

$R^2$ is a simple average of the R-squared of the 120 series; *AIC* denotes Akaike Information Criterion; *BIC* is Bayesian Information Criterion; and *Log Lik* is the Log-Likelihood value.

**Table 3: Forecast error variance due to monetary policy shocks.**

| Variables | *0 month* | *1 months* | *12 months* | *24 months* | *60 months* |
|---|---|---|---|---|---|
| **77) Federal funds** | 0.72 | 0.66 | 0.27 | 0.24 | 0.27 |
| **16) IP: total index** | 0.00 | 0.01 | 0.17 | 0.18 | 0.18 |
| **108) CPI-U: all items** | 0.00 | 0.00 | 0.09 | 0.21 | 0.31 |
| **78) US Tbill. 3m.** | 0.30 | 0.36 | 0.21 | 0.19 | 0.23 |
| **81) Tbond const 5yr.** | 0.22 | 0.25 | 0.20 | 0.14 | 0.17 |
| **96) Monetary base** | 0.07 | 0.13 | 0.13 | 0.14 | 0.13 |
| **92) Money stock: M1** | 0.07 | 0.14 | 0.13 | 0.14 | 0.14 |
| **74) FX : Japan** | 0.04 | 0.06 | 0.09 | 0.09 | 0.09 |
| **107) Index of sensitive mat.** | 0.00 | 0.01 | 0.19 | 0.17 | 0.16 |
| **17) Capacity util rate** | 0.00 | 0.00 | 0.20 | 0.22 | 0.22 |
| **49) Pers cons exp: total** | 0.00 | 0.00 | 0.03 | 0.03 | 0.03 |
| **50) Pers cons exp: tot.** | 0.00 | 0.00 | 0.02 | 0.02 | 0.02 |
| **51) Pers cons exp: nondur.** | 0.00 | 0.00 | 0.02 | 0.02 | 0.02 |
| **26) Unempl. Rate: all wrks** | 0.00 | 0.00 | 0.11 | 0.12 | 0.12 |
| **48) NAPM Empl. Index** | 0.00 | 0.00 | 0.26 | 0.27 | 0.27 |
| **119) Avg hr earnings manuf.** | 0.00 | 0.00 | 0.02 | 0.02 | 0.02 |
| **54) Housing starts: nonfarm** | 0.06 | 0.16 | 0.53 | 0.49 | 0.48 |
| **62) NAPM new orders** | 0.00 | 0.04 | 0.33 | 0.31 | 0.32 |
| **68) NYSE composite** | 0.06 | 0.07 | 0.08 | 0.08 | 0.08 |
| **120) Consumer expec. (Mich.)** | 0.00 | 0.00 | 0.03 | 0.05 | 0.18 |

.

Figure 1: Hallin and Liska Information Criteria on Dynamic Factors

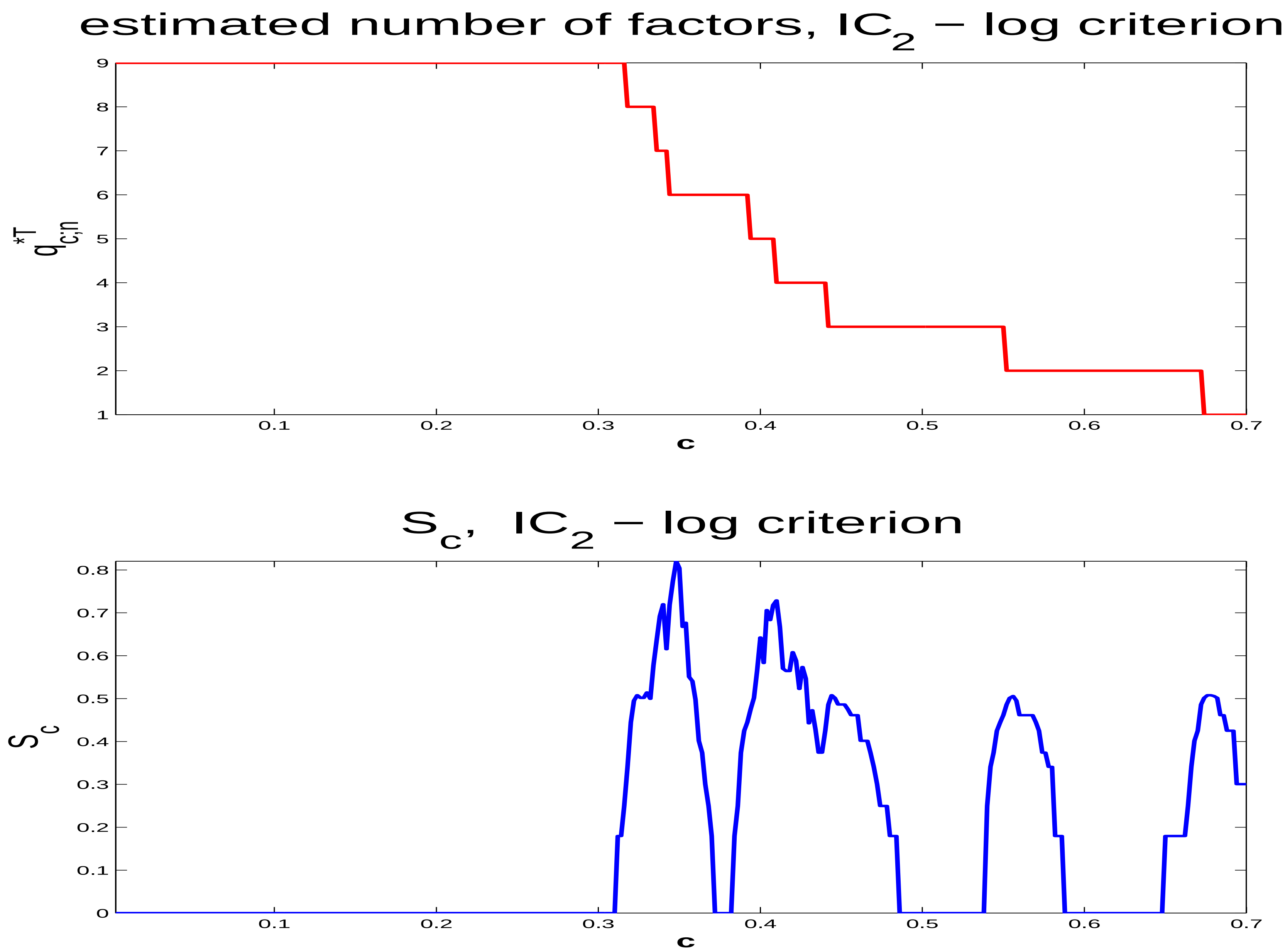
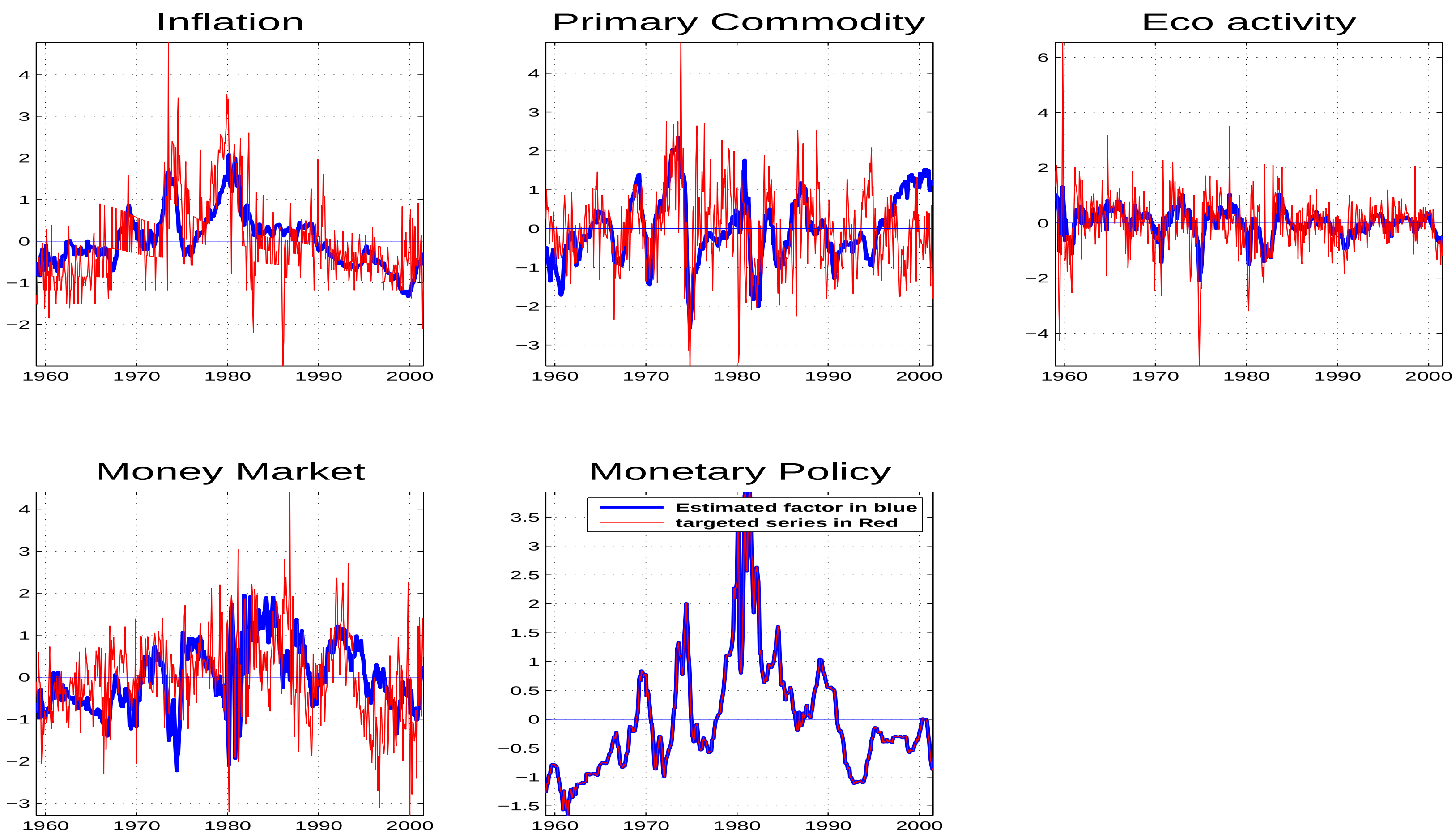


Figure 2: Time series of factors versus related observed variables



## Figure 3: Time series of inflation factor against the inflation series

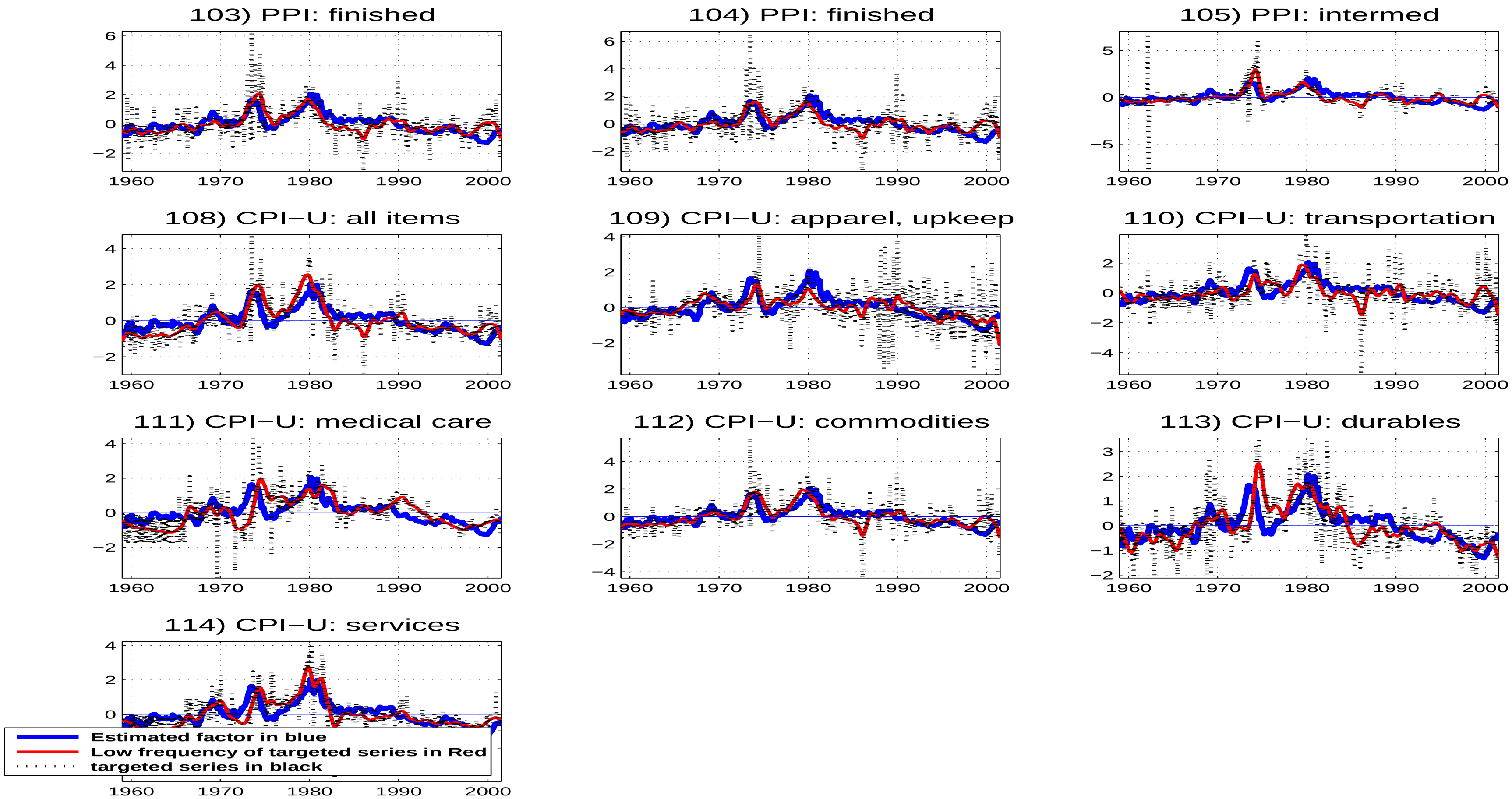


Figure 4: Time series of economic activity factor against the production growth series

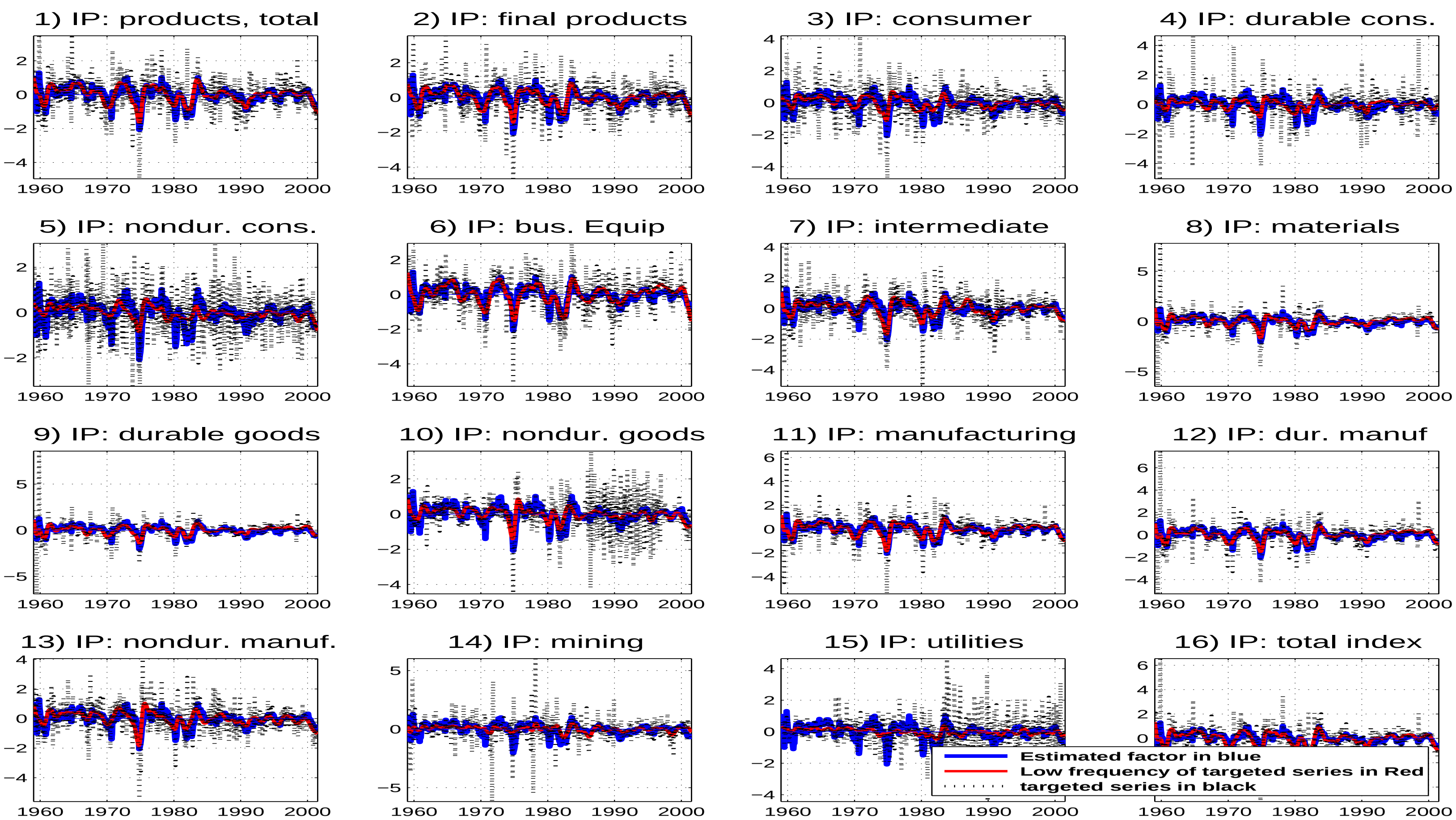


**Figure 5: Time series of commodity price factor against the commodity price series**

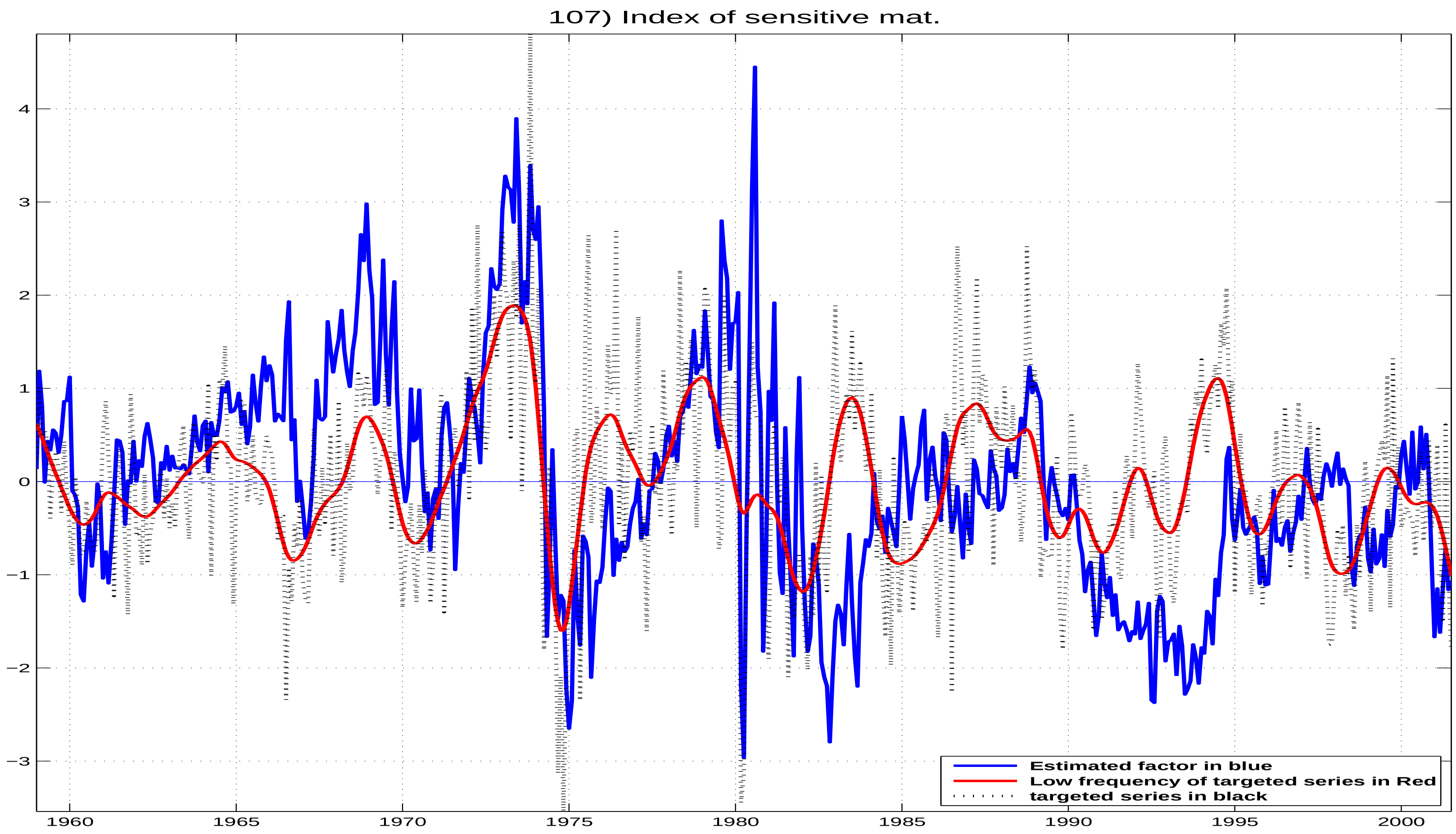


**Figure 6: Time series of money market factor against the money supply series**

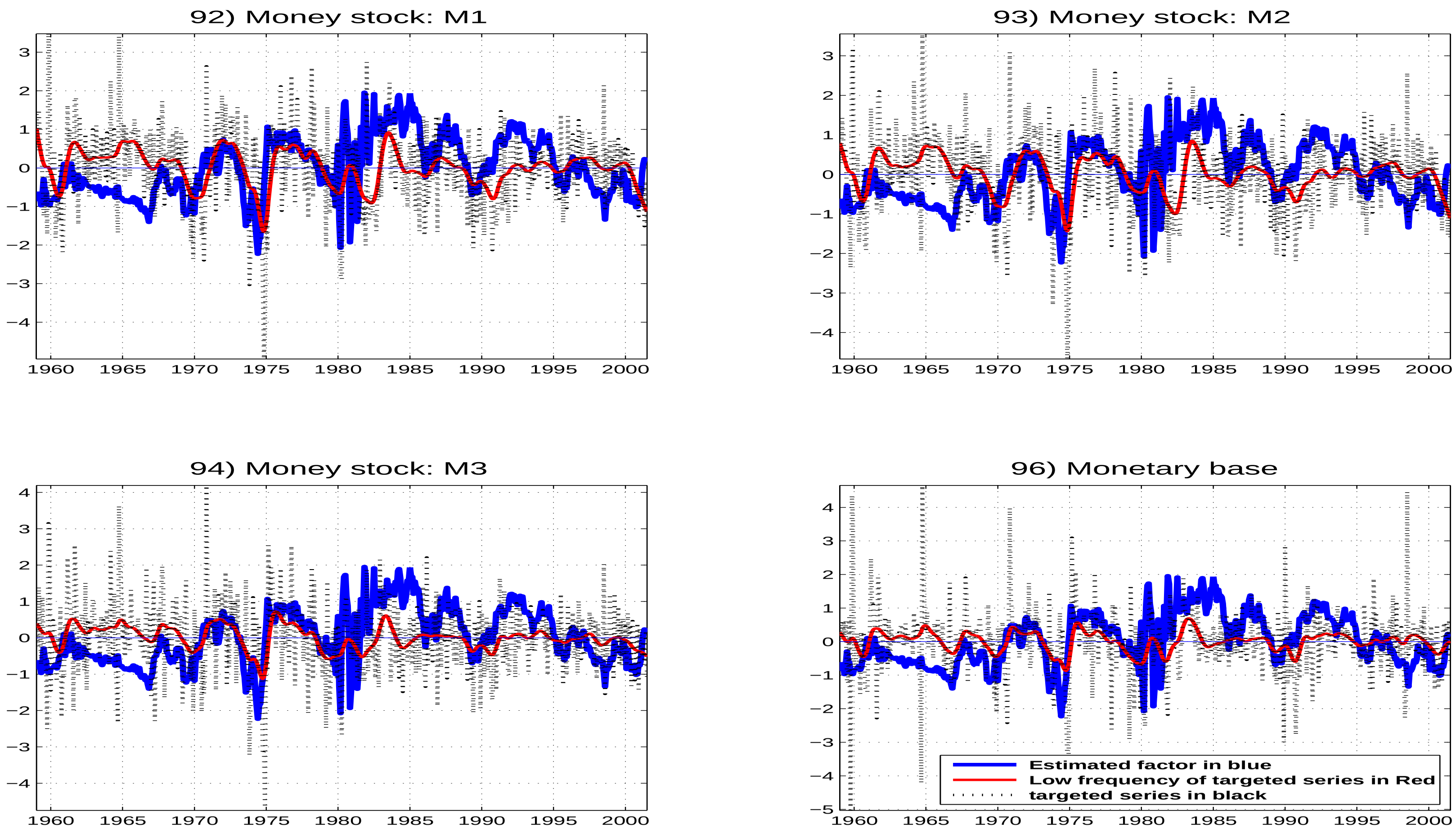


Figure 7: Loadings (sum over lags) for the alternative factors

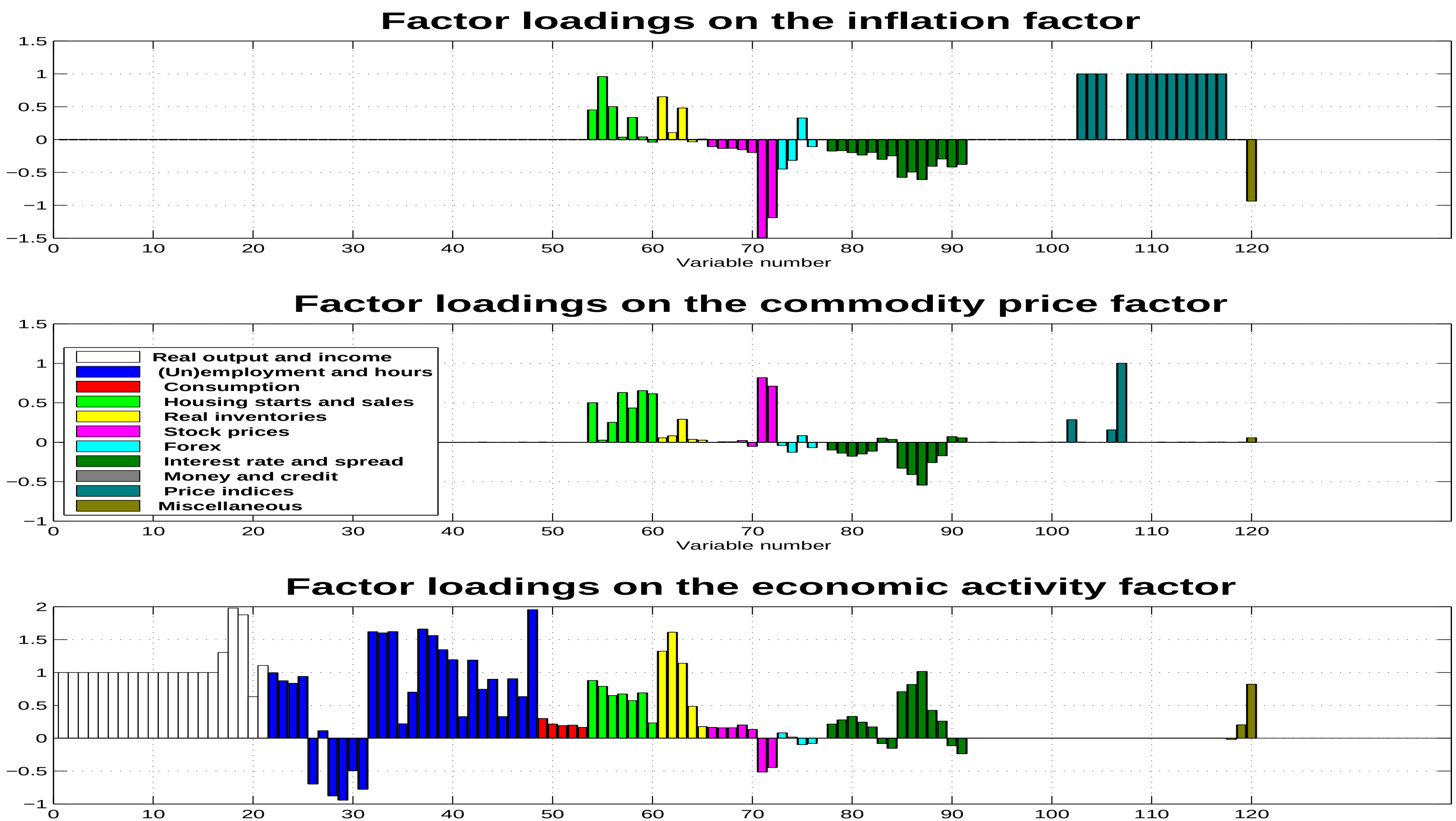


Figure 8: Loadings (sum over lags) for the alternative factors

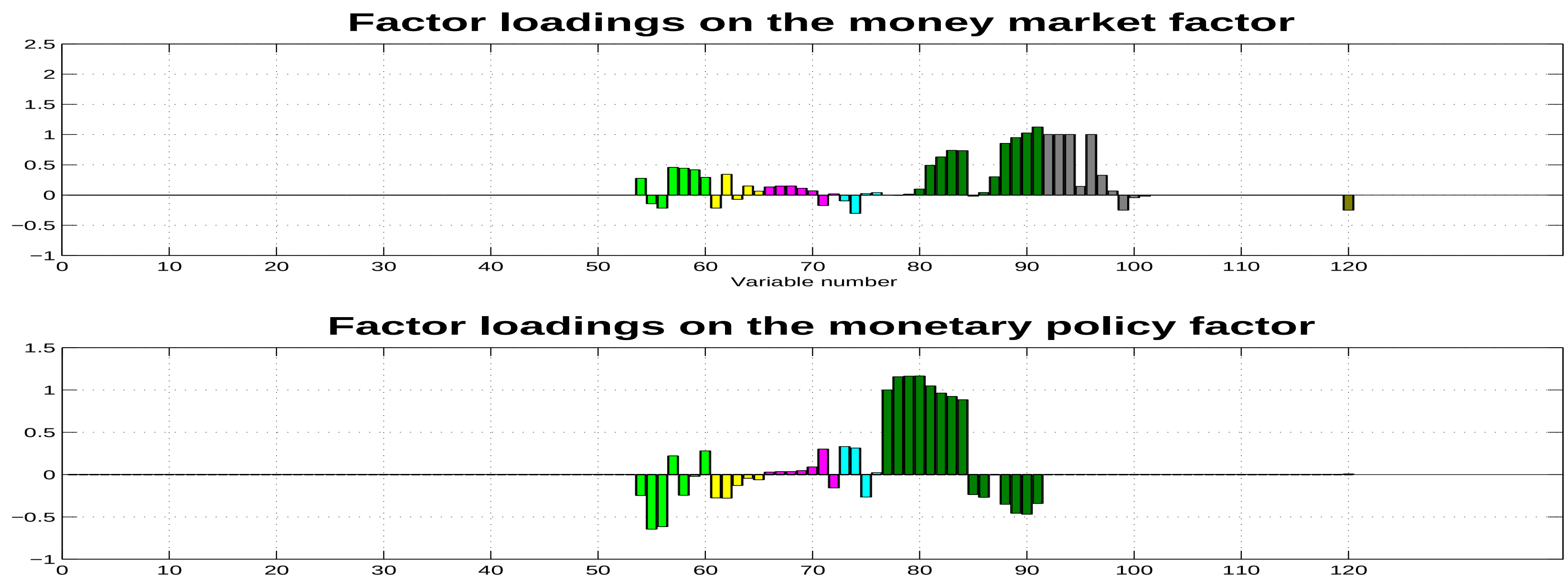
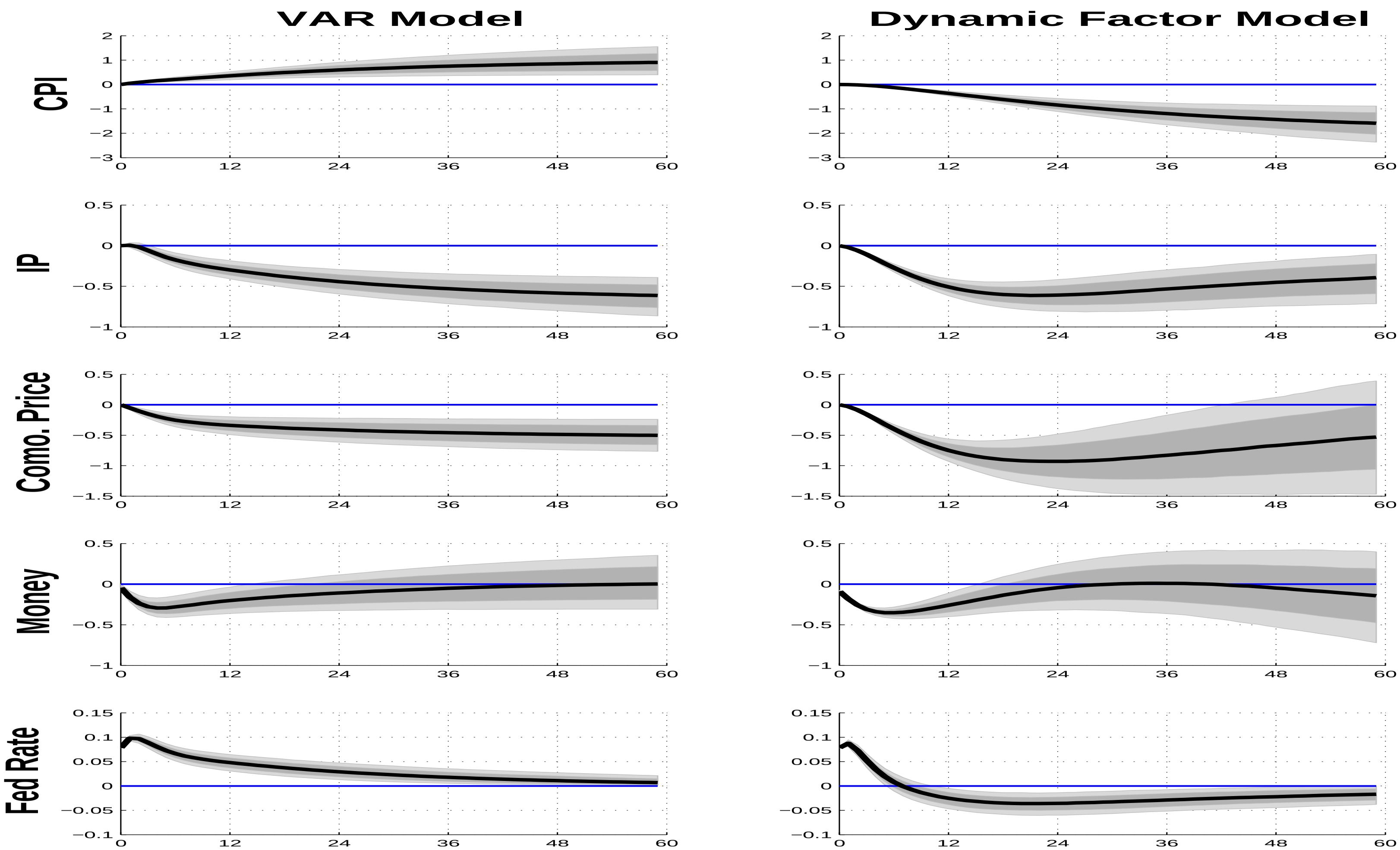
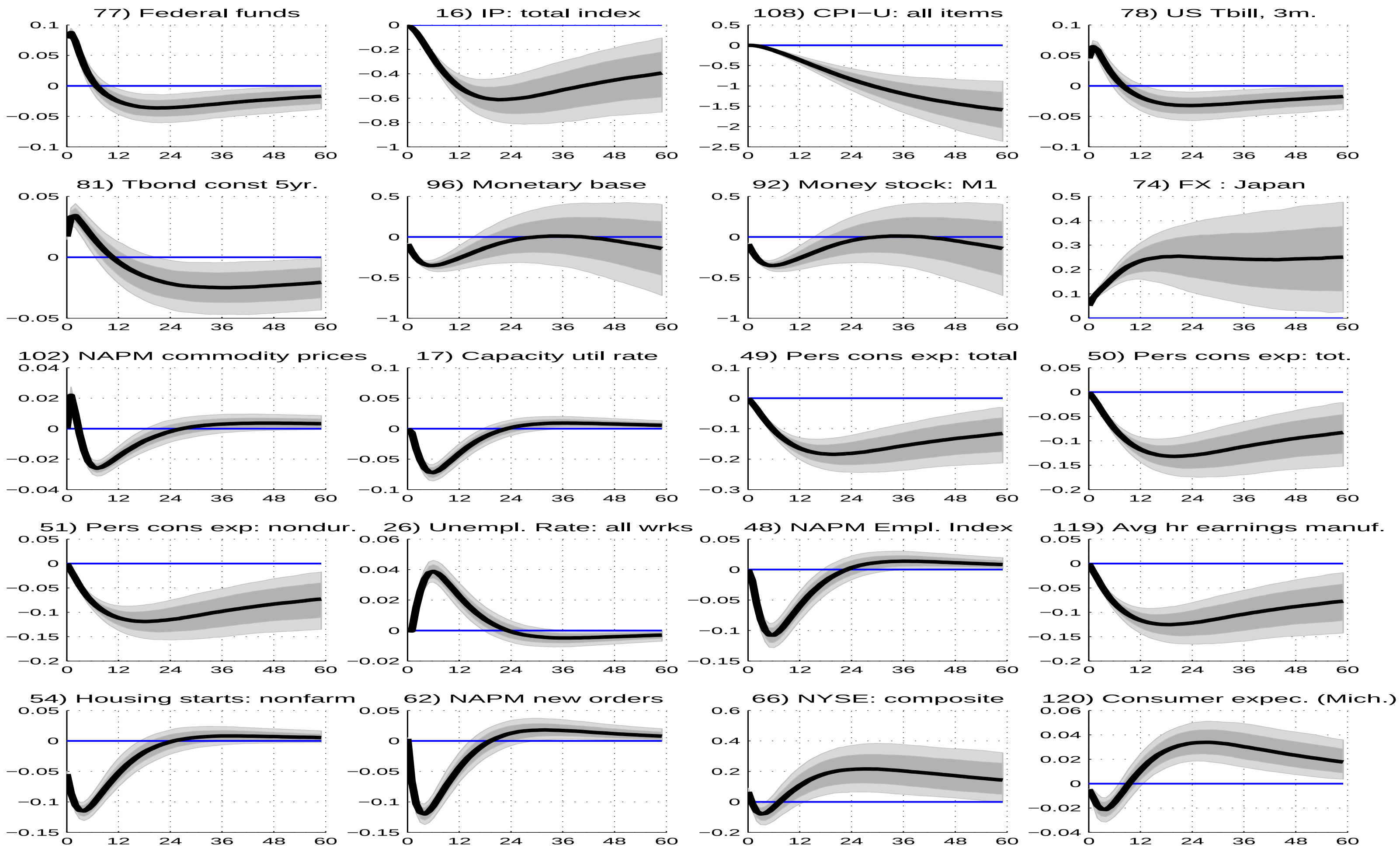


**Figure 9: Impulse Response Functions of SVAR and Structural Factor Models to a monetary policy shock**



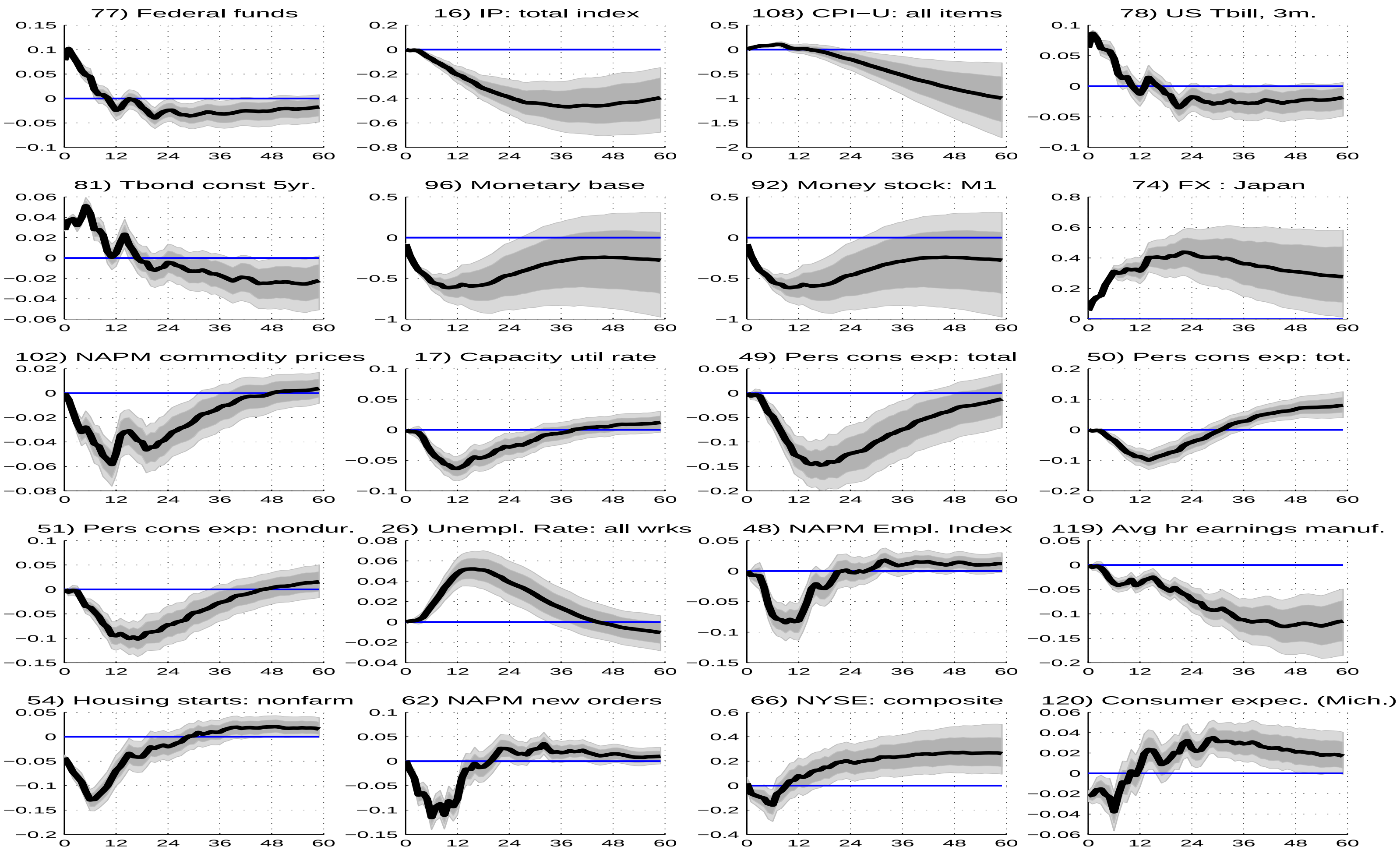

The figure shows median IRFs together with the 68 and 90 percent confidence intervals using the standard bootstrap method.

## Figure 10: Impulse Response Functions of various series to a monetary policy shock lag 2 model




The figure shows median IRFs together with the 68 and 90 percent confidence intervals using the standard bootstrap method.

## Figure 11: Impulse Response Functions of various series to a monetary policy shock lag 13 model




The figure shows median IRFs together with the 68 and 90 percent confidence intervals using the standard bootstrap method.