INSTITUT DE HAUTES ÉTUDES INTERNATIONALES ET DU DÉVELOPPEMENT
GRADUATE INSTITUTE OF INTERNATIONAL AND DEVELOPMENT STUDIES

Graduate Institute of International and Development Studies

International Economics Department

Working Paper Series

Working Paper No. HEIDWP0013-2017

Factor Models for Non-Stationary Series: Estimates of Monthly U.S. GDP

Martina Hengge
The Graduate Institute Geneva

Seton Leonard
The Graduate Institute Geneva

Chemin Eugène-Rigot 2
P.O. Box 136
CH - 1211 Geneva 21
Switzerland

Factor Models for Non-Stationary Series: Estimates of Monthly U.S. GDP

Martina Hengge* Seton Leonard**

June 2017

Abstract

This paper presents a novel dynamic factor model for non-stationary data. We begin by constructing a simple dynamic stochastic general equilibrium growth model and show that we can represent and estimate the model using a simple linear-Gaussian (Kalman) filter. Crucially, consistent estimation does not require differencing the data despite it being cointegrated of order 1. We then apply our approach to a mixed frequency model which we use to estimate monthly U.S. GDP from May 1969 to January 2016 using 171 series with an emphasis on housing related data. We suggest our estimates may, at a quarterly rate, in fact be more accurate than measurement error prone observations. Finally, we use our model to construct pseudo real-time GDP nowcasts over the 2007 to 2009 financial crisis. This last exercise shows that a GDP index, as opposed to real time estimates of GDP itself, may be more helpful in highlighting changes in the state of the macroeconomy.

Keywords: Forecasting; Factor model; Large data sets; Mixed-frequency data; Nowcasting; Non-stationarity; Real-time data.

JEL-codes: E27; E52; C53; C33.

*Martina Hengge: Department of Economics, The Graduate Institute of International and Development Studies, Maison de la Paix, Chemin Eugène-Rigot 2, 1202 Genève, Switzerland, E-mail: martina.hengge@graduateinstitute.ch.

**Seton Leonard, Department of Economics, The Graduate Institute of International and Development Studies, Maison de la Paix, Chemin Eugène-Rigot 2, 1202 Genève, Switzerland, E-mail: seth@srlanalytics.com.

1 Introduction

Dynamic factor models (DFMs) have become a standard tool in the analysis of large macroeconomic data sets. Perhaps surprisingly, these models are able to describe fairly accurately large sets of macroeconomic data in a few series thereby overcoming the curse of dimensionality (Giannone et al., 2005; Stock and Watson, 2016). Popular empirical applications of factor models include indexing economic activity (for example, Stock and Watson (1989), Mariano and Murasawa (2003), Arouba et al. (2008), and Altissimo et al. (2010)), nowcasting (for example, Angelini et al. (2008), Giannone et al. (2008), and Bańbura et al. (2011)), and forecasting (for example, D'Agostino and Giannone (2006) and Bańbura et al. (2015)).

Dynamic factor models, including those in this paper, are built on two simple equations. The measurement or observation equation, which we write as

$$Y_t = HZ_t + \epsilon_t \tag{1}$$

relates factors in the current period Z_t to current period observations Y_t via the loadings H_t. The transition equation

$$Z_{t+1} = AZ_t + e_t \tag{2}$$

describes the evolution of factors. Though equation (2) is written for a single lag of the factors, Z_t may contained stacked factors over p lags with A the companion form of the VAR(p) process. Despite its simplicity many popular time series models fit this format including VAR and ARMA processes.

The majority of current factor model applications require the data to be stationary. We begin with the observation that this approach is not consistent with most macro theory models, which typically look at variables in deviations from trend. In particular, we present a simple dynamic stochastic general equilibrium (DSGE) model that features stochastic growth in the style of Aguiar and Gopinath (2007). That is, the growth rate itself changes over time. We solve this model (to a linear approximation) in terms of both log deviations from scaled steady state values as is traditional in macro and in terms of variables in log levels. The latter solution allows us to filter and smooth noisy observations generated by the model to estimate the true underlying states without either differencing or de-trending the data. Unsurprisingly, as we know the true model, our estimates of states are very good with a mean squared error (MSE) that matches the (time-invariant) factor covariance matrix from the Kalman smoother. Having established the principle of operating on log level data directly we then show that when we do not know the true model we can still consistently estimate it by maximum likelihood (ML) based on Watson and Engle (1983) as in Doz et al. (2012).

Having introduced our approach to estimating cointegrated factor models in a DSGE setting[1], we then present two popular applications of factor models using U.S. data: estimates of monthly GDP and nowcasts for quarterly GDP. In both cases, the novelty of our approach is in the fact that we estimate our model for log level data and consequentially our results are also in log levels, not growth rates as is commonly the case. Because we explicitly allow for measurement error, we suggest that, at a quarterly rate, our monthly GDP estimates may in fact be more accurate than observed GDP. Our nowcasting results are similar to those in Giannone et al. (2008), Bańbura and Modugno (2010), Bańbura et al. (2011), and Higgins (2014) in that estimates tend to improve as more data become available throughout the quarter. However, we note that for large changes in the macroeconomy such as the 2007-2009 financial crisis an index based on common factors may be more telling than a nowcast which includes persistent error components as the latter causes estimates to revert towards previous quarter values.

Our choice of estimation routine for the factor model is not trivial: ML estimation following, for example, Bańbura and Modugno (2010) and Doz et al. (2012) and fully Bayesian estimation[2] as in Kim and Nelson (1999) or Durbin and Koopman (2002) are both proven methods for estimating parameters of dynamic factor models. Fully Bayesian estimation is similar to ML but incorporates a prior in estimates of parameter distributions, though this comes at the cost of computational intensity.[3] Until recently, ML estimation has been considered unfeasible for large data sets because of the large number of parameters that need to be estimated. Additionally, ML estimation typically assumes an exact factor structure, that is, the covariance matrix for shocks to observations is assumed to be diagonal, possibly resulting in misspecification. These drawbacks of MLE have led to a shift away from traditional factor estimation by ML to principal components analysis (for example, Bai (2003)). However, Doz et al. (2012) have shown that ML estimates are consistent despite potential misspecification of the correlation in measurement equation disturbances and that ML is indeed viable for large-dimensional factor models. Moreover, ML estimation incorporates the dynamics of factors in transition equation estimates,[4] can easily incorporate parameter restrictions, provides

[1]Series generated by our DSGE model are cointegrated by construction as growth comes from a single state variable, in our case labor productivity, and the long run ratios of variables remain constant.

[2]Kalman filtering is Bayesian in the sense that one step ahead forecasts form our prior for factors in period t given observations through period $t-1$; this prior is then updated as observations in period t become available. By fully Bayesian we mean that parameter estimates also come from a prior distribution which is updated based on observations and estimated factors.

[3]The relative simplicity of our ML estimates comes from the fact that we do not estimate or simulate the distribution of parameters as we assume some unknown "true" parameter values.

[4]Principle components based estimation and ML estimation coincide when there are no lagged factors in the transition equation.

a clean framework for missing data (Jungbacker and Koopman, 2008; Bańbura and Modugno, 2010), and allows us to identify both persistent and idiosyncratic error components in the measurement equation. This latter consideration is an extension of Bańbura and Modugno (2010) which we find useful in identifying measurement error. We therefore opt for ML estimation of the factors, using the Watson and Engle (1983) EM algorithm as in Bańbura and Modugno (2010) with some slight alterations of our own.

Our econometric framework builds on the long tradition of work that incorporates level information from data in a state-space framework. As vector autoregressions (VARs) are a special case of a factor model, this literature dates back at least to the classic papers Engle and Granger (1987) and Johansen (1988). More recently Bańbura et al. (2010) and Giannone et al. (2015) have shown that Bayesian VARs are suitable for the estimation of large dynamic systems and perform comparably to factor models. Similar to our approach, these models can incorporate level information from a cointegrating vector; one can view the approach in this paper as simply using several cointegrating vectors instead of a single cointegrating vector and first differences.

For non-stationary factor models, Barigozzi et al. (2016a) illustrate that factors that are cointegrated of order one can be represented by a vector error correction model. Barigozzi et al. (2016b) propose an estimator of this error correction representation of factors and discuss the conditions under which it is consistent. Though similar in spirit to the exercise in this paper there are several key differences. Barigozzi et al. (2016b) estimate loadings from the differenced data using principle components, obtain factors using estimated loadings and the level data, and then estimate the error correction model for the factors. In contrast, we estimate factors and loadings by maximum likelihood following Watson and Engle (1983) operating directly on the non-stationary level data. Though the data we use throughout the paper is non-stationary in the sense that it grows over time, we deal only with models in which parameters are stationary.[5] That is, we do not address non-stationarity in the sense of time varying parameters (see, for example, Hamilton (2005), Del Negro and Otrok (2008), or Eichler et al. (2011)).

2 Estimating a Stochastic Growth Model

This paper begins with the following question: supposing data is generated by a simple DSGE growth model, how should we estimate underlying states of the model? To make the exercise interesting, we model stochastic growth in the sense

[5]Note that growth, even the stochastic growth we model here, does not imply that shocks to our transition and measurement equations will explode over time. This ensures consistent parameter estimation.

of Aguiar and Gopinath (2007), that is, we allow the actual rate of growth to fluctuate over time. This feature distinguishes our theoretical model from the literature that looks at fluctuations around a steady state (be it deterministic or risky as defined by Coeurdacier et al. (2011)). We insist on stochastic growth as estimating a model in which variables fluctuate around a deterministic growth path is trivial: one simply de-trends the model to enforce stationarity. Though we too can solve our proposed model by scaling variables to enforce stationarity, using this fact to estimate underlying factors assumes we can perfectly observe the state that drives growth (productive labor in this case). As this assumption is unrealistic, we instead write the system of difference equations that constitutes our solution to the model in terms of log level variables. In the case that we know the parameters of the model, this solution allows us to filter and smooth observations to estimate the true underlying states. When we do not know the parameters of the model, we estimate them by maximum likelihood. In the absence of identifying restrictions, this allows to estimate a set of factors which span the underlying states of the model; given credible identifying assumptions we can again estimate the true states.

2.1 A Simple DSGE Model

Our simple DSGE model follows a long line of macro theory literature including classic papers such as Kydland and Prescott (1982) and Long and Plosser (1983) in modeling a utility maximizing representative agent in a closed economy. In particular, our agent maximizes a simple, time separable utility function

$$\mathcal{U} = E_t \sum_{s=0}^{\infty} \beta^s \frac{C_{t+s}^{1-\sigma}}{1-\sigma} \tag{3}$$

subject to a budget constraint

$$a_t K_t^{\alpha} H_t^{\theta} L_t^{1-\alpha-\theta} + (1-\delta)K_t + (1-\gamma)H_t = C_t + K_{t+1} + H_{t+1} \tag{4}$$

where a_t is productivity, K_t is physical capital, H_t is human capital, L_t productive labor, and C_t is consumption.[6] We suppose that total factor productivity follows an exogenous, stationary process

$$\hat{a}_{t+1} = \rho_a \hat{a}_t + \epsilon_{t+1}$$

where hats denote log deviation from trend. As in Solow (1956) growth is labor augmenting. However, following Aguiar and Gopinath (2007) it is also stochastic

[6] We differentiate between physical and human capital only to add more states to the model which makes estimation slightly more interesting.

such that $L_{t+1} = \tau_{t+1} L_t$. We use upper case letters to denote variables that grow over time. Like productivity, we allow the growth rate τ to vary according to

$$\hat{\tau}_{t+1} = \rho_\tau \hat{\tau}_t + \varepsilon_{t+1}$$

Again following Solow (1956) we enforce stationarity on the model by scaling variables by (productive) labor.[7] For example, the Euler equation

$$C_t^{-\sigma} = \beta E_t \Big[C_{t+1}^{-\sigma} \big(\alpha a_{t+1} K_{t+1}^{\alpha-1} H_{t+1}^{\theta} L_{t+1}^{1-\alpha-\theta} + 1 - \delta \big) \Big]$$

becomes

$$c_t^{-\sigma} = \beta E_t \Big[c_{t+1}^{-\sigma} \tau_{t+1}^{-\sigma} \big(\alpha a_{t+1} k_{t+1}^{\alpha-1} h_{t+1}^{\theta} + 1 - \delta \big) \Big]$$

For the parameter values $\beta = 0.99$, $\tau = 1.0033$, $\sigma = 1.5$, $\gamma = 0.075$, $\delta = 0.1$, $\alpha = 0.25$, $\theta = 0.25$, $\rho_a = 0.7$, and $\rho_\tau = 0.9$, the log linear solution for scaled, stationary next period state variables as a function of current period state variables is

$$\begin{bmatrix} \hat{a}_{t+1} \\ \hat{\tau}_{t+1} \\ \hat{k}_{t+1} \\ \hat{h}_{t+1} \end{bmatrix} = \begin{bmatrix} 0.70 & 0 & 0 & 0 \\ 0 & 0.90 & 0 & 0 \\ 0.26 & -1.41 & 0.38 & 0.48 \\ 0.11 & -1.59 & 0.43 & 0.54 \end{bmatrix} \begin{bmatrix} \hat{a}_t \\ \hat{\tau}_{t+1} \\ \hat{k}_t \\ \hat{h}_t \end{bmatrix} + \begin{bmatrix} \epsilon_{t+1}^a \\ \epsilon_{t+1}^\tau \\ \epsilon_{t+1}^k \\ \epsilon_{t+1}^h \end{bmatrix} \tag{5}$$

To put the model in log levels we first augment the vector of states y_t to include labor L_t, so that equation (5) becomes[8]

(6)

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & 1 \end{bmatrix}}_{H} \underbrace{\begin{bmatrix} \hat{a}_{t+1} \\ \hat{\tau}_{t+1} \\ \hat{k}_{t+1} \\ \hat{h}_{t+1} \\ L_{t+1} \end{bmatrix}}_{\hat{y}_{t+1}} = \underbrace{\begin{bmatrix} 0.70 & 0 & 0 & 0 & 0 \\ 0 & 0.90 & 0 & 0 & 0 \\ 0.26 & -1.41 & 0.38 & 0.48 & 0 \\ 0.11 & -1.59 & 0.43 & 0.54 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_{\hat{A}} \underbrace{\begin{bmatrix} \hat{a}_t \\ \hat{\tau}_{t+1} \\ \hat{k}_t \\ \hat{h}_t \\ L_t \end{bmatrix}}_{\hat{y}_t} + \underbrace{\begin{bmatrix} \epsilon_{t+1}^a \\ \epsilon_{t+1}^\tau \\ \epsilon_{t+1}^k \\ \epsilon_{t+1}^h \\ 0 \end{bmatrix}}_{\hat{\epsilon}_{t+1}}$$

where the last row of comes from the law of motion for productive labor, in logs $L_{t+1} = \tau_{t+1} + L_t$. We can write equation (6) concisely as

$$y_{t+1} - \bar{y} = \hat{\hat{A}}(y_t - \bar{y}) + \hat{\hat{\epsilon}}_{t+1} \tag{7}$$

[7]Appendix A offers a complete derivation of the model results.

[8]Note here all variables are in logs.

where $\hat{\hat{A}} = H^{-1}\hat{A}$ and $\hat{\hat{\epsilon}}_t = H^{-1}\hat{\epsilon}_t$. Defining $\hat{B} = \bar{y} - \hat{\hat{A}}\bar{y}$ our final step is to introduce the helper matrix

$$\theta = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

so that

$$\theta y_t = \begin{bmatrix} a_t \\ \tau_t \\ K_t \\ H_t \\ L_t \end{bmatrix}$$

Then our solution for the model in log levels is

$$\underbrace{\theta y_{t+1}}_{Y_{t+1}} = \underbrace{\theta \hat{\hat{A}} \theta^{-1}}_{A} \underbrace{\theta y_t}_{Y_t} + \underbrace{\theta \hat{B}}_{B} + \underbrace{\theta \hat{\hat{\epsilon}}_{t+1}}_{\epsilon_{t+1}} \tag{8}$$

or, using the parameters specified above,

$$\begin{bmatrix} a_{t+1} \\ \tau_{t+1} \\ K_{t+1} \\ H_{t+1} \\ L_{t+1} \end{bmatrix} = \begin{bmatrix} 0.70 & 0 & 0 & 0 & 0 \\ 0 & 0.90 & 0 & 0 & 0 \\ 0.26 & -0.51 & 0.38 & 0.48 & 0.14 \\ 0.11 & -0.69 & 0.43 & 0.54 & 0.03 \\ 0 & 0.90 & 0 & 0 & 1.00 \end{bmatrix} \begin{bmatrix} a_t \\ \tau_t \\ K_t \\ H_t \\ L_t \end{bmatrix} + \begin{bmatrix} 0 \\ 0.0003 \\ 0.1229 \\ 0.1665 \\ 0.0003 \end{bmatrix} + \epsilon_{t+1} \tag{9}$$

Before filtering, smoothing, and estimating this model we briefly look at some of its properties; Aguiar and Gopinath (2007) provide a much more in depth look at models of stochastic growth. Figure 1 illustrates impulse response functions for de-trended log deviations from the deterministic steady state (left hand panel) and for log levels where initial values have been subtracted so that all variables begin at zero. Capital initially falls in response to a positive growth rate shock for both de-trended and level variables corresponding to the dis-saving Aguiar and Gopinath (2007) find for a growth shock; because a growth shock permanently affects income, on impact agents consume more and invest (save) less in anticipation of higher lifetime consumption. Note also that the responses to a fairly small level shock (0.1 in this example) are relatively large — much more so than in the case of a stationary TFP shock.

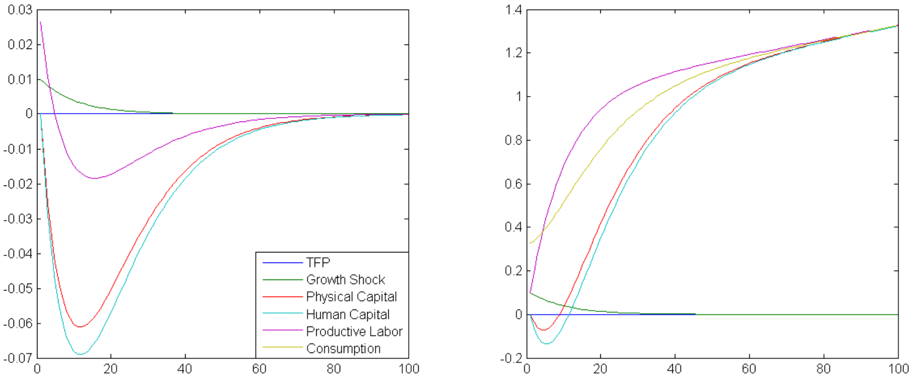

Figure 1: Impulse response functions for de-trended stationary variables (left hand panel) log levels (right hand panel)

2.2 Estimating States when Parameters are Known

Despite the fact that the system of equations in (9) is non-stationary (non-stationary in the sense that the transition matrix has a unit root due to the process for productive labor), the covariance of shocks is stationary and thus the system poses no problems for standard Kalman filtering and smoothing. As a simple first exercise we simulate and then estimate the states of the model assuming we in fact know the parameters in (9), including the jump variable consumption so that there are six observables, five of which are states. Explicitly, in state space our model is described by the measurement equation

$$\begin{bmatrix} a_t \\ \tau_t \\ K_t \\ H_t \\ L_t \\ C_t \end{bmatrix} = \begin{bmatrix} 1.0000 & 0 & 0 & 0 & 0 \\ 0 & 1.0000 & 0 & 0 & 0 \\ 0 & 0 & 1.0000 & 0 & 0 \\ 0 & 0 & 0 & 1.0000 & 0 \\ 0 & 0 & 0 & 0 & 1.0000 \\ 0.0968 & 2.6377 & 0.1700 & 0.2171 & 0.6129 \end{bmatrix} \begin{bmatrix} a_t \\ \tau_t \\ K_t \\ H_t \\ L_t \end{bmatrix} + \epsilon_t$$

and the transition equation (9). Over 1000 repetitions of 600 observations the mean squared error of our estimated states is: for TFP, 0.0076; for the growth shock, 0.0001; for physical capital, 0.0009; for human capital, 0.0008, and for labor 0.0007. These values exactly match the factor variance given all observations arising from the Kalman smoother.

2.3 Estimating Unobserved Parameters

Before we move on to estimating unobserved parameters, several observations about the solutions in equations (5) and (9) are important to our analysis. First, the model cannot be written in first differences. For example, subtracting y_t from both sides of (5) yields $\Delta y_{t+1} = (A - I)y_t + (I - A)\bar{y} + \epsilon_{t+1}$ where $\bar{y}$ denotes the points of approximation. Thus estimation with variables in first differences to enforce stationarity will be misspecified. This autoregressive process can, however, be written as a system of independent AR(1) processes (though shocks will be correlated). For the solution in log levels, using an eigendecomposition of the transition matrix $A = V\Lambda V^{-1}$ we can define a new set of states $\tilde{Y}_t = V^{-1}Y_t$ so that

$$\tilde{Y}_{t+1} = \Lambda\tilde{Y}_t + V^{-1}\bar{Y} + V^{-1}\epsilon_{t+1} \tag{10}$$

Though correlated, shocks to these otherwise independent series will still be normal. For the parameters we use, the eigenvalues of A are

$$diag(\Lambda) = \begin{bmatrix}1.00 & 0.92 & 0.90 & 0.70 & 0.00\end{bmatrix}$$

so that $\tilde{Y}_t$ consists of one random walk, three stationary AR(1) processes, and one white noise process. Put differently, we can write the states in (9) as a random walk and four cointegrating vectors.[9] When we estimate the parameters of the model we will be estimating this form — a single non-stationary factor, with the remainder of factors representing cointegrating relationships.

The novelty of our approach lies in how we specify the non-stationary factor — other elements of the model are simply taken from the existing literature. Several interesting possibilities exist to capture growth in the log variables. The simplest approach that is consistent with the model is to simply specify a random walk with drift. Letting x_t^L denote the non-stationary factor in levels, we could write

$$x_{t+1}^L = \mu + x_t^L + v_{t+1} \tag{11}$$

where the shock v_t denotes velocity, the first difference of x_{t+1}^L. Alternatively, if we wanted to allow the non-stationary factor to be integrated of order 2 we could write

$$\begin{bmatrix}x_{t+1}^L \\ v_{t+1}\end{bmatrix} = \begin{bmatrix}1 & 1 \\ 0 & 1\end{bmatrix}\begin{bmatrix}x_t^L \\ v_t\end{bmatrix} + \begin{bmatrix}a_{t+1} \\ a_{t+1}\end{bmatrix}$$

where the shock a_{t+1} now corresponds, in a physical model, to acceleration.

[9]Two of these, those corresponding to the eigenvalues 0.9 and 0.7, are trivial as they are simply the stationary processes for τ_t and a_t. Though the state corresponding to the eigenvalue 0 is not useful for predicting future states, it still matters in our factor model framework as a determinant of contemporaneous observables.

We opt for something between these two possibilities in which v_t is stationary but may depend on its own lags and the lags of other non-stationary factors. For the illustrative case of three lags and two cointegrating vectors (that is, two additional stationary factors) we write the transition equation for stationary factors — the only part of the transition equation we will estimate — as

$$\begin{bmatrix} v_t \\ x_t^1 \\ x_t^2 \end{bmatrix} = \mu + B_1 \begin{bmatrix} v_{t-1} \\ x_{t-1}^1 \\ x_{t-1}^2 \end{bmatrix} + B_2 \begin{bmatrix} v_{t-2} \\ x_{t-2}^1 \\ x_{t-2}^2 \end{bmatrix} + B_3 \begin{bmatrix} v_{t-3} \\ x_{t-3}^1 \\ x_{t-3}^2 \end{bmatrix} + e_t \tag{12}$$

where the evolution of the $I(1)$ factor is given by (11).

	$T = 60$	$T = 120$	$T = 400$	$T = 600$
$k = 20$	2.77	1.62	1.12	1.07
$k = 50$	2.47	1.53	1.09	1.06
$k = 100$	2.28	1.43	1.09	1.05

Table 1: Average MSE for T observations of k series for 1000 simulations

As we do not use a normalization which would allow us to identify the true states[10], we evaluate the performance of our model on its ability to estimate missing high frequency observations in a mixed frequency framework. This metric feels appropriate as our interest in section 4 will be in estimating a quarterly-monthly mixed frequency model for U.S. GDP. Table 1 provides results for 20 observable series under several different scenarios; sample size refers to the number of observations T. We estimate the model by maximum likelihood based on Watson and Engle (1983) but deffer a more thorough discussion of methodology until section 3. For each simulation we consider five series to be quarterly so that we observe the mean of the current and previous two monthly observations for these series every third period. The mean squared error is calculated from the error in our high frequency (monthly) estimates of series we observe as quarterly. As the number of periods we observe increases, the MSE for high frequency estimates of low frequency variables approaches its estimated variance arising from the Kalman filter,[11] which is close to the variance of shocks to the measurement equation.

The results in Table 1 are for the model estimated in log levels. However, we can difference our level estimates to provide a metric against which we can compare our model with the more standard approach of differencing variables first

[10] Recall that in this framework $Y_t = HX_t + \epsilon_t$ is observationally equivalent to $Y_t = H\theta^{-1}\theta X_t + \epsilon_t$ where θX_t is some alternative linear combination of the factors X_t.

[11] The actual estimated variance — the variance for estimated series taking the parameters of our model as true — is series specific and depends on the draw for the factor loadings H at each iteration.

Log Levels Model

	$T = 60$	$T = 120$	$T = 400$	$T = 600$
$k = 20$	6.54	3.57	2.29	2.18
$k = 50$	5.70	3.34	2.22	2.14
$k = 100$	5.15	3.03	2.20	2.13

Log First Difference Model

	$T = 60$	$T = 120$	$T = 400$	$T = 600$
$k = 20$	12.15	4.42	4.01	3.24
$k = 50$	5.35	3.12	2.26	2.27
$k = 100$	3.87	2.82	2.16	2.13

Table 2: Average MSE for T observations of k series for 1000 simulations

and then estimating a model for stationary variables. Table 2 presents the results of this exercise. Our levels model is able to reduce the MSE for monthly estimates of variables observed quarterly when the number of observable series k is low or the number of periods observed T is high. However, for large k and small T the model in first differences in fact performs better; we attribute this result to the fact that the levels model estimates an additional factor, the $I(1)$ factor, so that in this case the model for differenced data is more parsimonious.

3 Maximum Likelihood Estimation

Popular approaches to estimating the parameters of dynamic factor models include two step principle components as in Doz et al. (2011), maximum likelihood estimation (MLE), or fully Bayesian estimation as in Koop and Korobilis (2010). From a forecasting perspective, Bayesian estimation offers most of the advantages of maximum likelihood estimation with the additional possibility of reducing parameter uncertainty via prior distributions. The big disadvantage of Bayesian estimation versus MLE following Watson and Engle (1983) is the computational burden arising from the need to simulate posterior distributions; this is particularly important for the simulations in sections 2 and 4.3. Two step principal components is faster than MLE still — it is in fact where we begin our iterations of Watson and Engle (1983)'s EM algorithm. However, we find that MLE has a number of advantages. Perhaps most importantly, principal components is a static problem resulting in noisy initial factor estimates. These noisy initial estimates tend to bias parameters of the transition equation towards zero. Additionally, MLE allows for cleaner handling of missing data, the ability to distinguish between persistent and idiosyncratic error components as detailed in section 3.1, allows us to re-estimate initial factor values via the Kalman smoother at each iteration, and does not require the data is standardized as covariances and intercepts are estimated.

To fix the notation we will use throughout the rest of the paper, we will write the measurement equation as

$$Y_t = \mu_1 + H \begin{bmatrix} x_t^L \\ x_t^2 \\ \vdots \\ x_t^m \end{bmatrix} + \epsilon_t \tag{13}$$

where x_t^L is the non-stationary factor in levels, x_t^j, where $j = 2, ..., m$, are the stationary factors, and ϵ_t are error terms potentially correlated across time. If a subset of current observations Y_t^Q depends on both current and lagged factors then, denoting the vector of factors in period t as X_t, we can write equation (13) as

$$Y_t^Q = \mu^Q + H^Q \begin{bmatrix} X_t \\ X_{t-1} \\ \vdots \\ X_{t-(s-1)} \end{bmatrix} + u_t^Q$$

where factors from s periods determine the observations in Y_t^Q. Defining $x_t^D = \Delta x_t^L$, we model the evolution of factors as

$$\begin{bmatrix} x_t^D \\ x_t^2 \\ \vdots \\ x_t^m \end{bmatrix} = \mu_1 + B_1 \begin{bmatrix} x_{t-1}^D \\ x_{t-1}^2 \\ \vdots \\ x_{t-1}^m \end{bmatrix} + B_2 \begin{bmatrix} x_{t-2}^D \\ x_{t-2}^2 \\ \vdots \\ x_{t-2}^m \end{bmatrix} + \ldots + B_p \begin{bmatrix} x_{t-p}^D \\ x_{t-p}^2 \\ \vdots \\ x_{t-p}^m \end{bmatrix} + v_t \tag{14}$$

where p is the number of lags in the vector autoregression equation (14).

The difficulty with these equations lies in the fact that the variables entering into equation (13) are not the same as those entering into equation (14). To circumvent this problem we define an augmented vector of factors; with one lag in the transition equation this vector is

$$Z_t = \begin{bmatrix} x_t^L \\ x_t^D \\ x_t^2 \\ \vdots \\ x_t^m \end{bmatrix}$$

Simply stacking these augmented vectors when the transition equation contains more than one lag would cause the companion matrix A to be singular (since the current level of the non-stationary variable is just the past level plus the current difference). Thus for p lags in the transition equation, the state vector is

$$Z_t = \begin{bmatrix} x_t^L \\ x_t^D \\ x_t^2 \\ \vdots \\ x_t^m \\ x_{t-1}^D \\ x_{t-1}^2 \\ \vdots \\ x_{t-1}^m \\ \vdots \\ x_{t-p}^D \\ x_{t-p}^2 \\ \vdots \\ x_{t-p}^m \end{bmatrix}$$

Our measurement equation is then $Y_t = \tilde{H}Z_t + \epsilon_t$, where $\tilde{H} = HJ$ and J is a helper matrix to extract the relevant elements of Z. For example, where $m = 3$ for all variables, we have

$$J = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 1 & 0 & \dots \end{bmatrix}$$

Our transition equation is the companion matrix for the coefficients $B = \begin{bmatrix} B_1 & B_2 & \dots & B_p \end{bmatrix}$ in equation (14) modified to incorporate the non-stationary factor. Specifically, we write the matrix A in equation (2), $Z_t = AZ_{t-1} + e_t$, as

$$A = \begin{bmatrix} 1 & b_{1,1}^1 & b_{1,2}^1 & \dots & b_{1,m}^1 & b_{1,1}^2 & \dots \\ 0 & b_{1,1}^1 & b_{1,2}^1 & \dots & b_{1,m}^1 & b_{1,1}^2 & \dots \\ 0 & b_{2,1}^1 & b_{2,2}^1 & \dots & b_{2,m}^1 & b_{2,1}^2 & \dots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \\ 0 & b_{m,1}^1 & b_{m,1}^1 & \dots & b_{m,m}^1 & b_{m,1}^2 & \dots \\ 0 & 1 & 0 & \dots & 0 & 0 & \dots \\ 0 & 0 & 1 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots \end{bmatrix}$$

where $b_{i,j}^1$ is element (i, j) of the matrix B_1 in equation (14), $b_{i,j}^2$ is element (i, j) of the matrix B_2 in equation (14), and so on. While the transition equation is non-stationary, the equation we estimate, equation (14), remains stationary as it does not include the series x_t^L thereby ensuring consistency.

3.1 Including AR(1) Error Terms

An important feature of the real data not present in our simulations is the fact that series depart from their long run relationships to the factors in a persistent way. That is, in our measurement equation

$$Y_t = \mu_1 + \tilde{H} Z_t + \epsilon_t$$

the individual error terms ϵ_t^i are autocorrelated. Ignoring this feature leads to estimates which tend towards the conditional mean of the data in every period. To correct for the persistence of ϵ_t^i in section 4 we follow Bańbura et al. (2011) in modeling each series of error terms as AR(1) so that $\epsilon_t^i = u_t^i + \varepsilon_t^i$ and $u_t^i = \rho u_{t-1}^i + e_t$ and then include the vector of error terms u_t as a state. That is, our modified state-space model becomes

$$Y_t = \begin{bmatrix} \mu_1 \\ 0 \end{bmatrix} + \underbrace{\begin{bmatrix} \tilde{H} & I_k \end{bmatrix}}_{\mathbf{H}} \begin{bmatrix} Z_t \\ u_t \end{bmatrix} + \varepsilon_t \tag{15}$$

and

$$\begin{bmatrix} Z_t \\ u_t \end{bmatrix} = \begin{bmatrix} \mu_2 \\ 0 \end{bmatrix} + \underbrace{\begin{bmatrix} A & 0 \\ 0 & B_\rho \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} Z_{t-1} \\ u_{t-1} \end{bmatrix} + \begin{bmatrix} V_t \\ e_t \end{bmatrix} \tag{16}$$

where I_k is a $k \times k$ identity matrix and B_ρ is a diagonal matrix of the AR(1) coefficients on the error terms. Note that unlike Bańbura et al. (2011) we distinguish between a persistent component of errors terms u_t and an idiosyncratic component ε_t. We have found this distinction to be essential in constructing maximum likelihood estimates of model parameters following Watson and Engle (1983); omitting the idiosyncratic component implies that the covariance matrix of the augmented vector of factors $F_t = \begin{bmatrix} Z_t^a & u_t^a \end{bmatrix}'$ (Z_t^a and u_t^a are defined in detail below) will be singular, preventing the use of the Kalman smoother.

3.2 Estimation of the State-Space Model

As our transition equation in (16) consists of two independent systems, we estimate the parameters for A and B_ρ separately. Define the vector of observations in equation (14), the parameters necessary to construct A in (16), as

$$X_t^s = \begin{bmatrix} x_t^D \\ x_t^2 \\ \vdots \\ x_t^m \end{bmatrix}, \qquad Z_t^s = \begin{bmatrix} 1 \\ X_t^s \\ X_{t-1}^s \\ \vdots \\ X_{t-p+1}^s \end{bmatrix}$$

where the superscript s denotes the fact that X_t^s and Z_t^s contain only stationary elements. Then we can write the transition equation (14) concisely as

$$X_t^s = BZ_{t-1}^s + v_t \tag{17}$$

The parameters of (17) we need to estimate in the M step of the EM algorithm are B and the covariance matrix for v_t which we denote as q. The maximum likelihood estimates of these matrices are given by[12]

$$B = E\big(X_t^s(Z_{t-1}^s)'\big)\big[E\big(Z_{t-1}^s(Z_{t-1}^s)'\big)\big]^{-1} \tag{18}$$

$$q = E(v_t v_t') \tag{19}$$

As detailed by Watson and Engle (1983), we cannot calculate the expectations in (18) and (19) directly as factors are not observed but estimated. Instead, we estimate $E(Z_{t-1}^s(Z_{t-1}^s)')$ as

$$\frac{1}{T}\Big[\sum_t Z_{t-1|T}^s(Z_{t-1|T}^s)' + \sum_t P_{t-1|T}\Big]$$

and estimate $E(X_t^s(Z_{t-1}^s)')$ as

$$\frac{1}{T}\Big[\sum_t X_{t|T}^s(Z_{t-1|T}^s)' + \sum_t C_{t|T}\Big]$$

where $Z_{t-1|T}^s$ denotes values of Z_{t-1}^s estimated by the Kalman smoother, $P_{t-1|T}$ is the variance of $Z_{t-1|T}^s$ from the smoother, $X_{t|T}^s$ are estimates of X_t^s from the smoother, and $C_{t|T}$ estimates $E\Big(\big(X_{t|T}^s - X_t^s\big)\big(Z_{t-1|T}^s - Z_t^s\big)\Big)$, also obtained from the smoother by including an extra lag of X_t^s in the state vector. Thus the augmented vector of states in the Kalman filter and smoother is[13]

$$F_t = \begin{bmatrix} Z_t^a \\ u_t^a \end{bmatrix}$$

[12] We omit hats for parameter estimates to keep the notation clean.

[13] See Watson and Engle (1983) for more detail.

where

$$Z_t^a = \begin{bmatrix} X_t^s \\ X_{t-1}^s \\ \vdots \\ X_{t-p+1}^s \\ X_{t-p}^s \end{bmatrix}, \qquad u_t^a = \begin{bmatrix} u_t \\ u_{t-1} \end{bmatrix}$$

Lastly, for q we estimate $E(v_t v_t')$ as

$$\frac{1}{T}\Big[\sum_t v_{t|T} v_{t|T} + \sum_t P_{t|T}^x + \sum_t B P_{t-1|T} B' - \sum_t B C_{t|T} - \sum_t C_{t|T}' B'\Big]$$

where $P_{t|T}^x$ is the covariance matrix for $X_{t|T}^s$ obtained from the smoother. We estimate B_ρ and the covariance of e_t analogously.

The parameters of the measurement equation we need to estimate are H in equation (13), used to construct $\tilde{H}$ in (15), μ_1, and the covariance of ε_t in equation (15). Define $\tilde{Z}_t = J Z_t$. Then our maximum likelihood estimates of these parameter matrices are given by

$$[\mu \quad H] = E\left(Y_t \begin{bmatrix} 1 & \tilde{Z}_t' \end{bmatrix}\right) \left[E\left(\begin{bmatrix} 1 \\ \tilde{Z}_t \end{bmatrix} \begin{bmatrix} 1 & \tilde{Z}_t' \end{bmatrix}\right)\right]^{-1} \tag{20}$$

and

$$R = E\left(\varepsilon_t \varepsilon_t'\right) \tag{21}$$

Here the difficulty lies in the fact that again $\tilde{Z}_t$ is estimated, not observed, and additionally Y_t may contain missing values. We follow Bańbura and Modugno (2010) in addressing the latter issue;[14] however, our application of Watson and Engle (1983) is slightly different from that in Bańbura and Modugno (2010) as we allow for an idiosyncratic error component in the measurement equation so that our left hand side variable is in fact $\tilde{y}_t^i = y_t^i - u_t^i$. For $[\mu \quad H]$ the term $E\left(\tilde{Y}_t \begin{bmatrix} 1 & \tilde{Z}_t' \end{bmatrix}\right)$ does not require any adjustment due to the fact that factors are estimated. For the second term

$$E\left(\begin{bmatrix} 1 \\ \tilde{Z}_t \end{bmatrix} \begin{bmatrix} 1 & \tilde{Z}_t' \end{bmatrix}\right) = E\left(\begin{bmatrix} 1 \\ \tilde{Z}_{t|T} \end{bmatrix} \begin{bmatrix} 1 & \tilde{Z}_{t|T}' \end{bmatrix}\right) + \begin{bmatrix} 0 & 0 \\ 0 & E(P_{t|T}^h) \end{bmatrix}$$

where $P_{t|T}^h$ is the covariance matrix for $\tilde{Z}_{t|T}$ (as opposed to $Z_{t|T}^s$ above) obtained from the Kalman smoother. Finally, for R we have

$$E(\varepsilon_t \varepsilon_t') = E(\varepsilon_{t|T} \varepsilon_{t|T}') + E\big(\mathbf{H} P_{t|T}^\varepsilon \mathbf{H}'\big)$$

[14] Effectively this means only using periods in which $y_{i,t}$ is observed to calculate parameters for each series i.

where $\mathbf{H}$ is defined in equation (15) and $P^{\varepsilon}_{t|T}$ is the covariance matrix of $\begin{bmatrix} Z_{t|T} \\ u_{t|T} \end{bmatrix}$ obtained from the Kalman smoother.

4 Empirical Application

In this section we use our econometric model to construct both estimates of monthly U.S. GDP and pseudo real-time nowcasts of current-quarter GDP. These applications require modifying our econometric model to account for the fact that GDP is a quarterly series.

4.1 Data set

Our data set consists of 171 monthly and quarterly time series from 1968:10 to 2016:01. It comprises both national series and regional series for the four U.S. Census Bureau regions: Northeast, Midwest, South, and West. We obtained the data from Thomson Reuters Datastream and grouped them into 14 categories: housing prices (10); housing starts and sales (14); housing other (8); production (11); inventories, orders and sales (6); employment and unemployment (45); income and earnings (10); prices (17); interest rates and yields (22); money and credit (8); mortgage debt and delinquencies (8); stock prices (4); exchange rates (6); and other (2). We then transformed the series as needed to ensure cointegration of order one.[15] Our series for GDP is deflated by consumer prices less energy and when available we have used seasonally adjusted data. The data series and their transformation are described in more detail in appendix C.1.

4.2 Mixed-frequency estimation

An important challenge when using actual data is to identify monthly GDP when in fact we have no observations of that series. Denoting log monthly GDP at an annual rate as y_t^M and log quarterly GDP at an annual rate as y_t^Q what we in fact observe is

$$e^{y_t^Q} = \frac{1}{3}e^{y_t^M} + \frac{1}{3}e^{y_{t-1}^M} + \frac{1}{3}e^{y_{t-2}^M} \tag{22}$$

In addition to equation (22), we assume that monthly log GDP is a function of our estimated monthly factors, that is,

$$y_t^M = \alpha X_t + u_t^M \tag{23}$$

[15] This included taking logs of certain series and standardizing the variance of each series around its linear trend.

The difficulty in estimating monthly GDP is that equation (22) is in levels while equation (23) is in logs. To circumvent this problem we take a linear approximation of equation (22) so that

$$e^{y^Q}(y_t^Q - y^Q) = \frac{1}{3}e^{y^M}\Big((y_t^M - y^M) + (y_{t-1}^M - y^M) + (y_{t-2}^M - y^M)\Big)$$

where variables without time subscripts indicate points of approximation. Because both monthly and quarterly GDP are at annualized rates we have that $y^Q = y^M$ so that, as in Mariano and Murasawa (2003), our approximation simplifies to

$$y_t^Q = \frac{1}{3}y_t^M + \frac{1}{3}y_{t-1}^M + \frac{1}{3}y_{t-2}^M$$

Plugging equation (23) into the above yields the equation we in fact estimate in our model

$$y_t^Q = \frac{1}{3}\alpha X_t + \frac{1}{3}\alpha X_{t-1} + \frac{1}{3}\alpha X_{t-2} + \frac{1}{3}u_t^M + \frac{1}{3}u_{t-1}^M + \frac{1}{3}u_{t-2}^M \tag{24}$$

Thus the states that determine quarterly GDP are the factors X_t, X_{t-1}, and X_{t-2} as well as the error terms u_t^M, u_{t-1}^M, and u_{t-2}^M since we maintain the assumption from section 3.1 that $u_t^i = \rho_i u_{t-1}^i$. Including these additional lags of the error term with the properly specified transition matrix and measurement equation allows us to construct estimates for monthly GDP as described by equation (23).

4.3 Estimates of Monthly GDP

Though the observable series we choose to include have a greater impact on our estimation than the number of factors m and lags p, the latter choice still merits careful thought. We opt for an external validity approach to this choice by dropping some observations, estimating the model, and calculating mean squared errors for dropped observations. This does not provide a definitive answer as results for nowcasting, when we only observe lagged values, are different than results for reconstructing missing data, in which we have both leading and lagged observations. We thus bias our selection towards parsimony setting $m = 3$ and $p = 3$. Once we determine m and p we estimate monthly GDP as outlined in the previous two subsections. Figure 2 illustrates our results over the 2007-2009 financial crisis using the full data set. Shaded bars indicate recessions as dated by the NBER's Business Cycle Dating Committee. We find that our model tracks the evolution of GDP well. GDP peaked in November 2007, declined for 4 months, went up slightly in April and May 2008, and then fell over the next 12 months, reaching a trough in May 2009. Not only do our estimates provide a high frequency series for GDP; as we suggest below, it may even be the case that our estimates are more accurate than the observed figures for expenditure side GDP. Estimates over the full sample from May 1969 to January 2016 are available in appendix D.

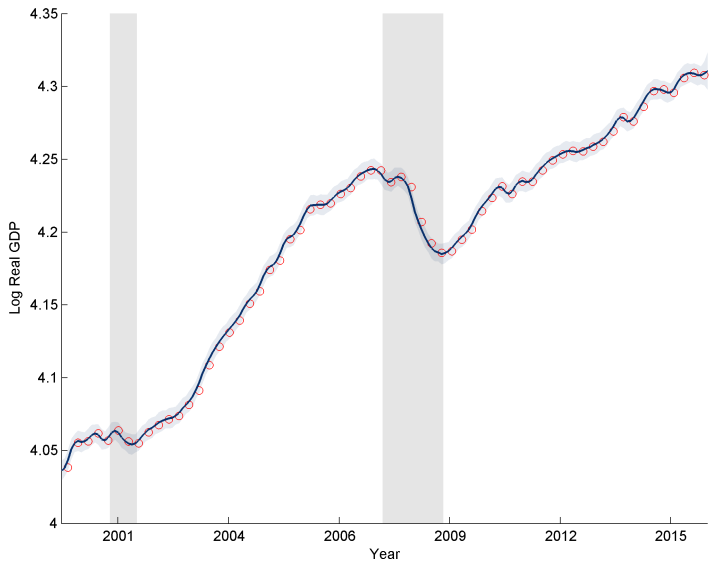

Figure 2: Estimates for monthly log real GDP at an annual rate with two standard deviation confidence intervals (estimated from the filter and smoother as opposed to simulation) over the financial crisis. Realizations for quarterly GDP are marked by circles.

As monthly GDP is not observed we cannot directly assess the accuracy of our GDP estimates presented above. We therefore simulate data using the estimated parameters of the state-space model. Following the approach outlined in section 4.2, we extract monthly simulated GDP and compute its MSE. For 1000 repetitions of our simulation, the MSE of estimated monthly GDP growth relative to the variance of simulated GDP growth is 0.49. This indicates that our estimates perform reasonably well.[16]

The MSE estimates of GDP growth from the simulation assume that our model is correctly specified — as mentioned above the parameters we use to simulate data are those we estimate. The real world data generating process will of course not fit this framework exactly. In particular, while assuming persistent errors u_t^i are AR(1) as opposed to AR(p) reduces the number of parameters we need to estimate, any true data generating process is unlikely to be so simple. For this reason we also simulate data in which errors are highly persistent and AR(3) so

[16]We calculate GDP growth by differencing the log level series.

that our estimated model will be misspecified. For 1000 repetitions, the ratio of the MSE of monthly GDP growth relative to the unconditional variance of GDP growth is 0.68. Unsurprisingly, misspecification of the error term in the measurement equation reduces the accuracy of our estimates. However, if we assume that i.i.d. shocks to GDP represent measurement error while true shocks to GDP are persistent then estimated quarterly GDP — the three month aggregate of our estimated monthly GDP — is still more accurate than observed quarterly GDP; the ratio of the MSE of estimated quarterly GDP to the MSE of observed quarterly GDP (again, assuming that i.i.d. errors to GDP represent measurement error) is 0.49.

Our ability to identify transient shocks, which we call measurement error, is due to the fact that we distinguish between three types of shocks. Shocks to common factors, v_t, have a strong cross-sectional component; a shock to a factor will contemporaneously impact many if not all of the series we observe. Shocks to persistent errors, e_t, have a strong intertemporal component; these shocks impact a single variable over many periods and identification comes from filtering and smoothing. Finally, transient shocks, ϵ_t, have neither a strong cross-sectional or intertemporal component and do not enter into our estimated series.

4.4 Pseudo Real-Time Nowcasts of Quarterly GDP

Economic decision making is complicated by uncertainty about the present state of the economy. A large number of high frequency data are available in real time. Yet, many key macroeconomic indicators, including GDP, are released at low frequency and with a publication lag. Consequently, forecasting the present — that is, nowcasting — is an important task for economic policy makers. To exploit all the information available in real time, nowcasting typically takes into consideration that variables in a multivariate data set are available at mixed frequencies. In addition, differing publication lags and release dates across variables result in an intricate pattern of missing observations towards the end of the sample — the so-called ragged edge. Our econometric model takes into account both of these features of the data.

In this section we nowcast current-quarter GDP based on data that would have been available for a specified nowcast date. Within the quarter, we consider 12 potential nowcast dates: weeks 1 through 4 of each of the three months that comprise the quarter.[17] Each week, we let the data set expand based on a stylized release calendar.[18] The impact of each release on the updated current-quarter

[17]Of course there may not be four weeks, or at least four Fridays, in every month. We break down every month into four sub-periods and consider each a week, regardless of the actual date.

[18]Appendix C.2 provides detailed information on the stylized release calendar.

nowcast will depend on how much information the newly released series contain for our target variable, GDP.[19] We apply this framework to a specific question: how well could this model have nowcasted U.S. GDP over the 2007-2009 financial crisis?[20]

Recall from equation 24 that our nowcasts for GDP are made up of two components: the contribution of common factors X_t^M through X_{t-2}^M and the persistent error terms u_t^M through u_{t-2}^M. These two components play very different roles in our nowcasts. The factors' common movements in the data depend less on intertemporal smoothing; as new developments in the economy evolve this is where we would expect to see the action. As noted in the previous section, identification of persistent error terms comes primarily from smoothing the data over time. As real-time analysis necessarily precludes forward values, these estimates will be less reliable. Figure 3 illustrates the contribution of common factors and the contribution of persistent errors to our GDP nowcasts.

[19] As in Giannone et al. (2008) the highest frequency we consider is monthly. We thus assume that data which are weekly or daily (such as financial data) become available at the end of the month. This implies that our January nowcasts for Q1 GDP are in fact an average of an one-step ahead forecast for January, a two-step ahead forecast for February, and a three-step ahead forecast for March. The February nowcasts then combine data from January with one-step ahead and two-step ahead forecasts for February and March, respectively. Finally, in March, current-quarter GDP nowcasts combine data available up to February with a one-step ahead forecast for March.

[20] Note that this procedure does not take into account revisions to the data. Schumacher and Breitung (2008) have shown that data revisions tend not have a substantial impact on forecast accuracy.

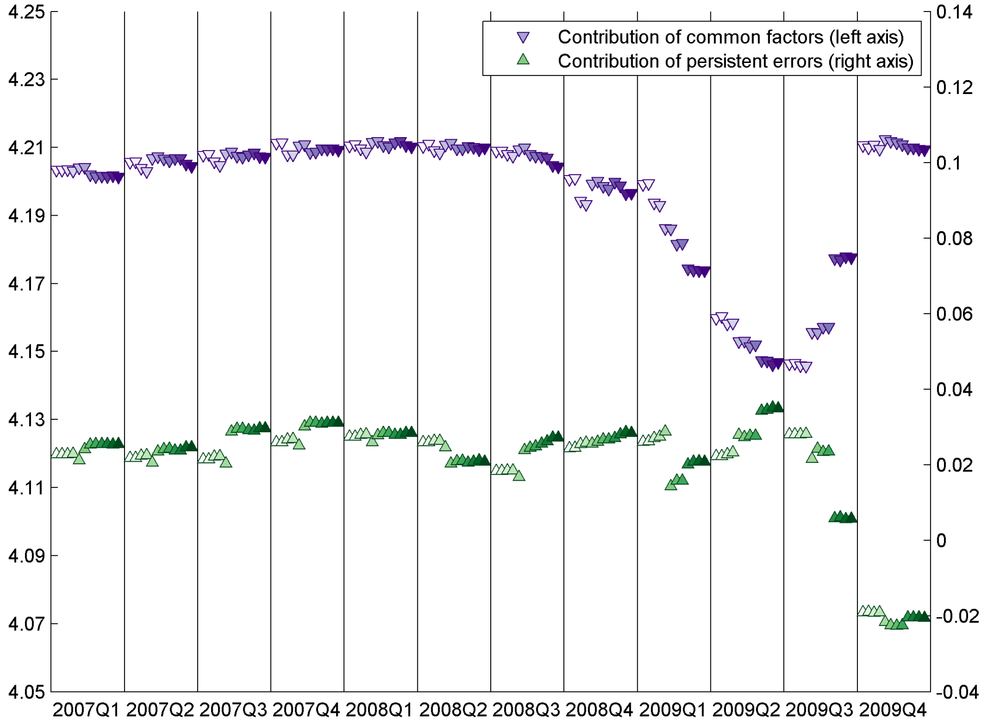

Figure 3: Contribution of common factors X_t^M through X_{t-2}^M and contribution of persistent error terms u_t^M through u_{t-2}^M to GDP nowcasts.

Note that as new data become available, our GDP estimate for 2008 Q4 based on common factors is overall revised downwards. The opposite is true for the persistent errors. This part of our estimate is continuously revised upwards. Estimated persistent errors were positive as realized GDP was well above its level implied by the common factors. Put differently, persistent errors have a tendency to automatically smooth nowcasts when factor estimates change. This in fact covered up the deteriorating economic conditions; as illustrated by figure 4, our 2008 Q4 GDP nowcast underestimates the scale of the contraction. At the through in 2009 Q2, however, the persistence in the errors pushed the current-quarter GDP nowcasts closer towards realized GDP. The nowcasts thus provide a better estimate than the common factors — which we might consider a GDP index — on their own.

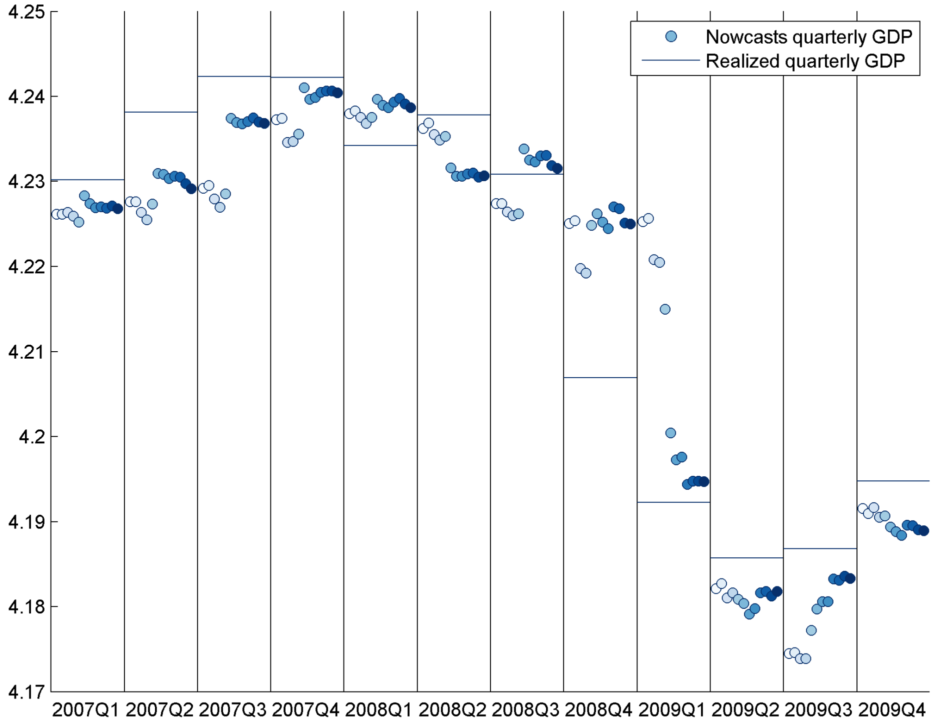

Figure 4: Nowcasts of current-quarter GDP over 2007 Q1- 2009 Q4

5 Conclusion

This paper makes several contributions to the macroeconomics literature. First, we present a novel method for estimating state-space models when data is cointegrated of order 1. This econometric model bridges a gap that still remains between macroeconomic theory and macroeconometrics. Theoretical models typically describe results in deviations from trend and only consider shocks to productivity. Econometric models, on the other hand, typically difference data to enforce stationarity and often allow what would in theory be considered jump variables to enter VARs. By considering observables as a function of unobserved states of the economy we are able to make our econometric estimates consistent with theory, so long as our theoretical framework allows for shocks to all state variables. Thus we are able to show that for our chosen data generating process our econometric approach offers an improvement over the standard practice of differencing even if the econometrician is not directly interested in the levels of observations.

Second, we use our model to construct estimates of monthly log real GDP in levels from October 1969 to January 2016. These estimates allow us to describe the monthly evolution of GDP over the recent financial crisis with a peak in November 2007 and a trough in June 2009. Additionally, simulation shows that our estimated GDP may be a more accurate measure of GDP than observed expenditure side GDP due to measurement error in the observed series.

Finally, we nowcast quarterly GDP in log levels differentiating between a component based on common factors and a component based on persistent error terms. While the common factor component does better at capturing changes in GDP, that is, nowcasts in first differences, level nowcasts require the inclusion of persistent error terms.

In this paper we have proposed a novel application of state-space models to cointegrated data generated by a simple dynamic stochastic general equilibrium model, yet much work remains to be done. None of the elements that go into our estimation are in fact new — we simply use principle components as an initial guess, the Kalman filter, and maximum likelihood estimation. However, we do not know a-priori the number of factors which will in fact be non-stationary; this is only something we learn after our estimation. What, then, are the conditions under which there will be a single non-stationary factor and all other factors will be stationary? In a broader context, how important is the differentiation between state and jump variables in empirical estimation? What are the implications of the misspecification entailed in estimating jump variables in a transition equation (that is, a VAR)? While we believe the framework we present in this paper offers a promising direction for improving both nowcasts and forecasts more generally, we also hope the questions it raises will stimulate further debate on their construction and estimation.

References

Aguiar, M. and Gopinath, G., 2007. Emerging Market Business Cycles: The Cycle Is the Trend. *Journal of Political Economy*, 115:69–102.

Altissimo, F., Cristadora, R., Forni, M., Lippi, M., and Veronese, G., 2010. New Eurocoin: Tracking Economic Growth in Real Time. *The Review of Economics and Statistics*, 92(4):1024–1034.

Angelini, E., Bańbura, M., and Rünstler, G., 2008. Estimating and Forecasting the Euro Area Monthly National Accounts from a Dynamic Factor Model. Working Paper No. 953, European Central Bank.

Arouba, S. B., Diebold, F. X., and Scotti, C., 2008. Real-Time Measurement of Business Conditions. Working Paper No. 14349, NBER.

Bai, J., 2003. Inferential Theory for Factor Models of Large Dimensions. *Econometrica*, 71(1):135–171.

Bańbura, M. and Modugno, M., 2010. Maximum Likelihood Estimation of Factor Models on Data Sets with Arbitrary Pattern of Missing Data. Working Paper No. 1189, European Central Bank.

Bańbura, M., Giannone, D., and Reichlin, L., 2010. Large Bayesian Vector Auto Regressions. *Journal of Applied Econometrics*, 25(1):71–92.

Bańbura, M., Giannone, D., and Reichlin, L., 2011. Nowcasting. In Clements, M. P. and Hendry, D. F., editors, *The Oxford Handbook of Economic Forecasting.* Oxford, Oxford Handbooks.

Bańbura, M., Giannone, D., and Lenza, M., 2015. Conditional forecasts and scenario analysis with vector autoregressions for large cross-sections. *International Journal of Forecasting*, 31(3):739–756.

Barigozzi, M., Lippi, M., and Luciani, M., 2016a. Dynamic Factor Models, Cointegration, and Error Correction Mechansims. Technical report, http://arxiv.org/abs71510.02399.

Barigozzi, M., Lippi, M., and Luciani, M., 2016b. Non-Stationary Dynamic Factor Models for Large Datasets. Finance and economics discussion series 2016-024, Federal Reserve Board.

Coeurdacier, N., Rey, H., and Winant, P., 2011. The risky steady state. *The American Economic Review*, 101(3):398–401.

D'Agostino, A. and Giannone, D., 2006. Comparing Alternative Predictors Based on Large-Panel Factor Models. Working Paper No. 680, European Central Bank.

Del Negro, M. and Otrok, C., 2008. Dynamic factor models with time-varying parameters: measuring changes in international business cycles. Staff Reports No. 326, Federal Reserve Bank of New York.

Doz, C., Giannone, D., and Reichlin, L., 2011. A two-step estimator for large approximate dynamic factor models based on Kalman filtering. *Journal of Econometrics*, 164(1):188–205.

Doz, C., Giannone, D., and Reichlin, L., 2012. A Quasi-Maximum Likelihood Approach for Large, Approximate Dynamic Factor Models. *Review of Economics and Statistics*, 94(4):1014–1024.

Durbin, J. and Koopman, S., 2002. A Simple and Efficient Simulation Smoother for State Space Time Series Analysis. *Biometrika*, 89(3):603–615.

Eichler, M., Motta, G., and von Sachs, R., 2011. Fitting dynamic factor models to non-stationary time series. *Journal of Econometrics*, 163(1):51–70.

Engle, R. F. and Granger, C. W. J., 1987. Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55(2):251–276.

Giannone, D., Reichlin, L., and Sala, L., 2005. Monetary Policy in Real Time. In Gertler, M. and Rogoff, K., editors, *NBER Macroeconomics Annual 2004*, volume 19. MIT Press.

Giannone, D., Reichlin, L., and Small, D., 2008. Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55 (4):665–676.

Giannone, D., Lenza, M., and Primiceri, G. E., 2015. Prior Selection for Vector Autoregressions. *The Review of Economics and Statistics*, 97(2):436–451.

Hamilton, J., 2005. Regime-Switching Models. Prepared for: Palgrave dictionary of economics, Palrave Dictionary of Economics. URL http://econweb.ucsd.edu/~jhamilto/palgrav1.pdf.

Higgins, P., 2014. GDPNow: A Model for GDP "Nowcasting". Working Paper No. 2014-7, Federal Reserve Bank of Atlanta.

Johansen, S., 1988. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2-3):231–254.

Jungbacker, B. and Koopman, J. S., 2008. Likelihood-based Analysis for Dynamic Factor Models. Discussion Paper No. 007/4, Tinbergen Institute.

Kim, C.-J. and Nelson, C. R., 1999. *State-Space Models with Regime Switching*. MIT Press.

Koop, G. and Korobilis, D., 2010. *Bayesian Multivariate Time Series Methods for Empirical Macroeconomics*. now Publishers.

Kydland, F. E. and Prescott, E. C., 1982. Time to build and aggregate fluctuations. *Econometrica*, 50(6):1345–1370.

Long, J. B. and Plosser, C., 1983. Real business cycles. *Journal of Political Economy*, 91(1):39–69.

Mariano, R. S. and Murasawa, Y., 2003. A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics*, 18(4): 427–443.

Schumacher, C. and Breitung, J., 2008. Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data. *International Journal of Forecasting*, 24(3):386–398.

Solow, R., 1956. A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, 70(1):65–94.

Stock, J. H. and Watson, M. W., 1989. New Indexes of Coincident and Leading Economic Indicators. In Blanchard, O. J. and Fischer, S., editors, *NBER Macroeconomic Annual 1989*, volume 4. MIT Press.

Stock, J. H. and Watson, M. W., 2016. Factor Models and Structural Vector Autoregressions in Macroeconomics. In Taylor, J. B. and Uhlig, H., editors, *Handbook of Macroeconomics*, volume 2A, pages 415–525. North-Holland.

Watson, M. W. and Engle, R. F., 1983. Alternative Algorithms for the Estimation of Dynamic Factor, Mimic and Varying Coefficient Regression Models. *Journal of Econometrics*, 23(3):385–400.

A Derivations for a Simple Theoretical Growth Model

This appendix contains the derivations for the simple theoretical growth models used to simulate data in section **??**. The two models are identical except for the fact that in the first case, when we estimate the correctly specified models, the stationary productivity shock follows an AR(1) process, while in the second case, when we estimated the misspecified model, this productivity shock is AR(2).

The model is described by a representative agent who maximizes utility

$$\mathcal{U} = E_t \sum_{s=0}^{\infty} \beta^s \frac{C_{t+s}^{1-\sigma}}{1-\sigma} \tag{25}$$

subject to

$$a_t K_t^{\alpha} H_t^{\theta} L_t^{1-\alpha-\theta} + (1-\delta)K_t + (1-\gamma)H_t = C_t + K_{t+1} + H_{t+1} \tag{26}$$

$$\hat{a}_{t+1} = \rho_a \hat{a}_t + \varepsilon_t^a \tag{27}$$

and

$$\hat{\tau}_{t+1} = \rho_\tau \hat{\tau}_t + \varepsilon_t^\tau \tag{28}$$

where $L_{t+1} = \tau_{t+1} L_t$. Upper case letters represent levels of variables that grow over time while lower case letters represent stationary variables. We take first order conditions in levels. Writing the Lagrangian for this constrained maximization problem as

$$\mathcal{L} = E_t \sum_{s=0}^{\infty} \beta^s \left[\frac{C_{t+s}^{1-\sigma}}{1-\sigma} + \lambda_{t+s} \Big(a_{t+s} K_{t+s}^{\alpha} H_{t+s}^{\theta} L_{t+s}^{1-\alpha-\theta} + (1-\delta)K_{t+s} + (1-\gamma)H_{t+s} - C_{t+s} - K_{t+s+1} - H_{t+s+1} \Big) \right]$$

the combined first order conditions for consumption (C_t) and physical capital (K_t) is

$$C_t^{-\sigma} = \beta E_t \big[C_{t+1}^{-\sigma} (\alpha a_{t+1} K_{t+1}^{\alpha-1} H_{t+1}^{\theta} L_{t+1}^{1-\alpha-\theta} + 1 - \delta) \big] \tag{29}$$

The combined first order condition for consumption and human capital (H_t) is, similarly,

$$C_t^{-\sigma} = \beta E_t \big[C_{t+1}^{-\sigma} (\theta a_{t+1} K_{t+1}^{\alpha} H_{t+1}^{\theta-1} L_{t+1}^{1-\alpha-\theta} + 1 - \gamma) \big] \tag{30}$$

The structure of our model (in particular, the fact that production is homogeneous of degree one) makes it convenient to scale by productive labor (L_t) so that equation (29) becomes

$$\left(\frac{C_t}{L_t}\right)^{-\sigma} = \beta E_t \left[\left(\frac{C_{t+1}}{L_{t+1}}\right)^{-\sigma} \tau_{t+1}^{-\sigma} (\alpha a_{t+1} \left(\frac{K_{t+1}}{L_{t+1}}\right)^{\alpha-1} \left(\frac{H_{t+1}}{L_{t+1}}\right)^{\theta} \left(\frac{L_{t+1}}{L_{t+1}}\right)^{1-\alpha-\theta} + 1 - \delta) \right]$$

where we have used equation (28) on the right hand side. Thus

$$c_t^{-\sigma} = \beta E_t \big[c_{t+1}^{-\sigma} \tau_{t+1}^{-\sigma} (\alpha a_{t+1} k_{t+1}^{\alpha-1} h_{t+1}^{\theta} + 1 - \delta) \big] \tag{31}$$

where $x_t = \frac{X_t}{L_t}$ and equation (29) becomes

$$c_t^{-\sigma} = \beta E_t \big[c_{t+1}^{-\sigma} \tau_{t+1}^{-\sigma} (\theta a_{t+1} k_{t+1}^{\alpha} h_{t+1}^{\theta-1} + 1 - \gamma) \big] \tag{32}$$

The system of difference equations describing this model is completed by the budget constraint, which in terms of scaled variables is

$$a_t k_t^{\alpha} h_t^{\theta} + (1-\delta) k_t + (1-\gamma) h_t = c_t + \tau_{t+1} k_{t+1} + \tau_{t+1} h_{t+1} \tag{33}$$

In matrix form the log linearized system of difference equations described by this model is

$$E_t \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & \tau(k+h) & \tau k & \tau h & 0 \\ \beta\tau^{-\sigma}\alpha k^{\alpha-1}h^{\theta} & -\sigma\beta\tau^{-\sigma}(\alpha k^{\alpha-1}h^{\theta} + (1-\delta)) & (\alpha-1)\beta\tau^{-\sigma}\alpha k^{\alpha-1}h^{\theta} & \theta\beta\tau^{-\sigma}\alpha k^{\alpha-1}h^{\theta} & -\sigma\beta\tau^{-\sigma}(\alpha k^{\alpha-1}h^{\theta} + (1-\delta)) \\ \beta\tau^{-\sigma}\theta k^{\alpha}h^{\theta-1} & -\sigma\beta\tau^{-\sigma}(\theta k^{\alpha}h^{\theta-1} + (1-\gamma)) & \alpha\beta\tau^{-\sigma}\theta k^{\alpha}h^{\theta-1} & (\theta-1)\beta\tau^{-\sigma}\theta k^{\alpha}h^{\theta-1} & -\sigma\beta\tau^{-\sigma}(\theta k^{\alpha}h^{\theta-1} + (1-\gamma)) \end{bmatrix}}_{H} \begin{bmatrix} \hat{a}_{t+1} \\ \hat{\tau}_{t+1} \\ \hat{k}_{t+1} \\ \hat{h}_{t+1} \\ \hat{c}_{t+1} \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} \rho_a & 0 & 0 & 0 & 0 \\ 0 & \rho_\tau & 0 & 0 & 0 \\ k^{\alpha}h^{\theta} & 0 & \alpha k^{\alpha}h^{\theta} + (1-\delta)k & \theta k^{\alpha}h^{\theta} + (1-\gamma)h & -c \\ 0 & 0 & 0 & 0 & -\sigma \\ 0 & 0 & 0 & 0 & -\sigma \end{bmatrix}}_{N} \begin{bmatrix} \hat{a}_t \\ \hat{\tau}_t \\ \hat{k}_t \\ \hat{h}_t \\ \hat{c}_t \end{bmatrix}$$

Hats indicate log deviation from steady state values and variables without time subscripts indicate steady state values. Letting $A = H^{-1}N$, defining Λ the diagonal matrix of stable eigenvalues of A, and C_s the square upper 4×4 submatrix of the eigenvectors associated with Λ, the the solution to this model, in terms of state variables in $t+1$ as a function of state variables in t, is

$$Y_{t+1} = W Y_t + \varepsilon_{t+1} \tag{34}$$

where $Y_t = \begin{bmatrix} \hat{a}_t & \hat{\tau}_t & \hat{k}_t & \hat{h}_t & \hat{c}_t \end{bmatrix}'$, $W = C_s \Lambda C_s^{-1}$, and ε_{t+1} is a vector of shocks to state variables. In the macro theory literature typically only a_{t+1}, total factor productivity, and perhaps τ_{t+1}, labor augmenting productivity, would be subject to shocks. However, as the econometrics literature allows for shocks to all state variables, we do so as well (that is, none of the elements of ε_{t+1} are restricted to zero), though in deference to the macro theory, shocks to a_{t+1} are much larger; in our parameterized model ε_{t+1}^{a} has a standard deviation of .01 and all other shocks have a standard deviation of 10^{-6}. The other parameter values we use in our simulations are $\beta = 0.99$, $\tau = 1.0033$, $\sigma = 1.5$, $\gamma = 0.075$, $\delta = 0.1$, $\alpha = 0.25$,

$\theta = 0.25$, $\rho_a = 0.7$, and $\rho_\tau = 0.9$. Using these values we can write the solution to our model, equation (34), as

$$\begin{bmatrix} \hat{a}_{t+1} \\ \hat{\tau}_{t+1} \\ \hat{k}_{t+1} \\ \hat{h}_{t+1} \end{bmatrix} = \begin{bmatrix} 0.70 & 0 & 0 & 0 \\ 0 & 0.90 & 0 & 0 \\ 0.23 & -0.98 & 0.34 & 0.52 \\ 0.09 & -1.17 & 0.41 & 0.62 \end{bmatrix} \begin{bmatrix} \hat{a}_t \\ \hat{\tau}_{t+1} \\ \hat{k}_t \\ \hat{h}_t \end{bmatrix} + \begin{bmatrix} \epsilon^a_{t+1} \\ \epsilon^\tau_{t+1} \\ \epsilon^k_{t+1} \\ \epsilon^h_{t+1} \end{bmatrix}$$

In the case that productivity follows an AR(2) process we simply augment the vector of state variables to include a_{t-1}. That is, we can write the AR(2) productivity shock in matrix form as

$$\begin{bmatrix} \hat{a}_{t+1} \\ \hat{a}_t \end{bmatrix} = \begin{bmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a_t \\ a_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon^a_{t+1} \\ 0 \end{bmatrix}$$

The Matlab programs used to simulate these models are available upon request.

B Simulated GDP

Our first set of simulations in which our model is correctly specified is analogous to the construction of IRFs for frequentest VARs by simulation. We begin by estimating the 8 factor 12 lag model from our 171 observed series as outlined in the paper. We then use the estimated transition matrix for stationary factors B to simulate factors where shocks are normally distributed with mean zero and covariance matrix q, the estimated covariance of the errors v_t. The non-stationary factor in the simulated augmented vector of factors Z_t^a is simply the sum of the differenced factor over the preceding periods. Because this first factor is non-stationary the influence of initial values will not fade over time; for this reason we start every simulation using estimated initial values from the data. We simulate the persistent component of errors in the measurement equation using the estimated AR(1) terms for each series where shocks are zero mean with variance given by the estimated variance of e_t. For our second set of simulations, those for the misspecified model, the only difference is the evolution of the persistent deviations in variables u_t; in these second simulations the evolution for u_t for every series is

$$u_t = .3u_{t-1} + .3u_{t-2} + .3u_{t-3} + e_t$$

where the variance of e_t is again given by the estimated variance. In this alternative process for errors the largest eigenvalue of the associated companion matrix is 0.95 so that shocks to errors are highly persistent.

Once we have simulated the factors Z_t and u_t we construct observations using the estimated matrix H and the helper matrix J described in section 3.1. The i.i.d. shocks to variables, which we consider to be measurement error, are zero mean with covariance matrix R where again R comes from our original estimations. We also save what we consider to be the unobserved true monthly realizations for all variables including low frequency variables given by

$$Y_t^{true} = \begin{bmatrix} H & I_k \end{bmatrix} \begin{bmatrix} X_t \\ u_t \end{bmatrix}$$

That is, what we consider to be the true realizations of variables do not include i.i.d. errors ϵ_t.

Once we have our simulated observations and saved true realizations, we estimate the model using the simulated observations where only every third realization for quarterly series is observed.

We obtain the mean squared error for GDP growth by comparing estimated monthly GDP growth $y_t^* - y_{t-1}^*$ with true monthly GDP growth $y_t^{true} - y_{t-1}^{true}$. Note that estimated GDP growth does not use any measure of actual monthly GDP as monthly GDP is not observed. In our second set of simulations for the misspecified model we also report the mean squared error for quarterly estimated GDP, $y_t^{Q,est} - y_t^{Q,true}$, relative to the mean squared error for quarterly observed GDP, $y^{Q,obs} - y_t^{Q,true}$. The complete set of Matlab programs used for these simulations is available on request.

C Detailed description of the data set

C.1 Data set

Table 3: Data set

Series	Freq.	Start date	End date	Trans- formations	Datastream code
Housing prices					
S&P/Case-Shiller national home price index	M	1975:01	2015:08	ln	USCSHP.ME
S&P/Case-Shiller national home price index - 20 city composite	M	2000:01	2015:08	ln	USCSHP20E
Median price of existing one family homes sold - Midwest	M	1968:11	2015:09	ln	USHPMERMA
Median price of existing one family homes sold - Northeast	M	1968:11	2015:09	ln	USHPMERNA
Median price of existing one family homes sold - South	M	1968:11	2015:09	ln	USHPMERSA
Median price of existing one family homes sold - West	M	1968:11	2015:09	ln	USHPMERWA
Average price of existing one family homes sold - Midwest	M	1989:01	2015:09	ln	USHPAERMA
Average price of existing one family homes sold - Northeast	M	1989:01	2015:09	ln	USHPAERNA
Average price of existing one family homes sold - South	M	1989:01	2015:09	ln	USHPAERSA
Average price of existing one family homes sold - West	M	1989:01	2015:09	ln	USHPAERWA
Housing starts and sales					
Housing started - 5 units or more	M	1968:10	2015:09		USHB5ANDO
Housing started - Midwest	M	1968:10	2015:09		USHBRM..O
Housing started - Northeast	M	1968:10	2015:09		USHBRN..O
Housing started - South	M	1968:10	2015:09		USHBRS..O
Housing started - West	M	1968:10	2015:09		USHBRW..O
Housing authorized - Midwest	M	1968:10	2015:09		USHARM..P
Housing authorized - Northeast	M	1968:10	2015:09		USHARN..P
Housing authorized - South	M	1968:10	2015:09		USHARS..P
Housing authorized - West	M	1968:10	2015:09		USHARW..P
Sales of new one family houses	M	1968:10	2015:09	ln	USHOUSESE
Existing one-family homes sold - Midwest	M	1989:01	2015:09	ln	USHSOERMO
Existing one-family homes sold - Northeast	M	1989:01	2015:09	ln	USHSOERNO
Existing one-family homes sold - South	M	1989:01	2015:09	ln	USHSOERSO
Existing one-family homes sold - West	M	1989:01	2015:09	ln	USHSOERWO
Housing other					
Home ownership rates - Midwest	Q	1968:11	2015:08		USHOWNMWR
Home ownership rates - Northeast	Q	1968:11	2015:08		USHOWNNER
Home ownership rates - South	Q	1968:11	2015:08		USHOWNSOR
Home ownership rates - West	Q	1968:11	2015:08		USHOWNWER
Rental vacancy rate - Midwest	Q	1968:11	2015:08		USHVRRM.%
Rental vacancy rate - Northeast	Q	1968:11	2015:08		USHVRRN.%
Rental vacancy rate - South	Q	1968:11	2015:08		USHVRRS.%
Rental vacancy rate - West	Q	1968:11	2015:08		USHVRRW.%
Production					
Industrial production	M	1968:10	2015:10	ln	USIPTOT.G
Industrial production - automotive products	M	1968:10	2015:10	ln	USIPMAUPG
Industrial production - business equipment	M	1968:10	2015:10	ln	USIPMBUQG
Industrial production - consumer goods	M	1968:10	2015:10	ln	USIPMCOGG
Industrial production - durable consumer goods	M	1968:10	2015:10	ln	USIPMDUCG
Industrial production - energy	M	1968:10	2015:10	ln	USIPMENTG
Industrial production - final products	M	1968:10	2015:10	ln	USIPTOT.G
Industrial production - materials	M	1968:10	2015:10	ln	USIPMMATG
Industrial production - nondurable consumer goods	M	1968:10	2015:10	ln	USIPMNOCG
Capacity utilization rate	M	1968:11	2015:10		USCAPUTLQ

continued on next page

Table 3: Data set

Series	Freq.	Start date	End date	Trans-formations	Datastream code
GDP	Q	1968:10	2015:10	ln	USGDP...B
Inventories, orders and sales					
Inventories/sales ratio - manufacturing	M	1997:01	2015:08		USISSMFGQ
Inventories/sales ratio - retail trade	M	1997:01	2015:08		USISSR..Q
Inventories - manufacturing	M	1992:01	2015:09	ln	USINMFG.B
New orders - manufacturing	M	1992:01	2015:09	ln	USNEWORDB
MFRS new orders	M	1968:10	2015:09		USMNOEACQ
ISM manufacturers survey (supplier delivery index)	M	1968:10	2015:10		USNAPMDL
Employment					
Labor force - Midwest	M	1976:01	2015:09	ln	USLFPI2EO
Labor force - Northeast	M	1976:01	2015:09	ln	USLFR6FOO
Labor force - South	M	1976:01	2015:09	ln	USLFKRPCO
Labor force - West	M	1976:01	2015:09	ln	USLF8Q9EO
Employed - construction	M	1968:10	2015:10	ln	USEM23..O
Employed - durable goods	M	1968:10	2015:10	ln	USEMIMD.O
Employed - education and health services	M	1968:10	2015:10	ln	USEMIE..O
Employed - federal	M	1968:10	2015:10	ln	USEMGF..O
Employed - financial activities	M	1968:10	2015:10	ln	USEMIF..O
Employed - goods producing	M	1968:10	2015:10	ln	USEMPG..O
Employed - government	M	1968:10	2015:10	ln	USEMIG..O
Employed - information	M	1968:10	2015:10	ln	USEM51..O
Employed - leisure and hospitality	M	1968:10	2015:10	ln	USEMIL..O
Employed - local government	M	1968:10	2015:10	ln	USEMGL..O
Employed - manufacturing	M	1968:10	2015:10	ln	USEMPMANO
Employed - mining	M	1968:10	2015:10	ln	USEM21..O
Employed - natural resources and mining	M	1968:10	2015:10	ln	USEMIU..O
Employed - nondurable goods	M	1968:10	2015:10	ln	USEMIMN.O
Employed - nonfarm industries total	M	1968:10	2015:10	ln	USEMPALLO
Employed - other services	M	1968:10	2015:10	ln	USEM81..O
Employed - private service providing	M	1968:10	2015:10	ln	USEMPP..O
Employed - professional and business services	M	1968:10	2015:10	ln	USEMIB..O
Employed - retail trade	M	1968:10	2015:10	ln	USEMIR..O
Employed - state government	M	1968:10	2015:10	ln	USEMGS..O
Employed - utlities	M	1968:10	2015:10	ln	USEM22..O
Employed - wholesale trade	M	1968:10	2015:10	ln	USEM42..O
Employment - Northeast	M	1976:01	2015:09	ln	USLER6FQO
Employment - Midwest	M	1976:01	2015:09	ln	USLEPI2GO
Employment - South	M	1976:01	2015:09	ln	USLEKRPEO
Employment - West	M	1976:01	2015:09	ln	USLE8Q9GO
Unemployed - less than 5 weeks	M	1968:10	2015:10	ln	USUNWK5.O
Unemployed - 5-14 weeks	M	1968:10	2015:10	ln	USUNWK14O
Unemployed - 15-26 weeks	M	1968:10	2015:10	ln	USUNWK26O
Unemployed - 15 weeks and more	M	1968:10	2015:10	ln	USUNPLNGE
Unemployed - 27 weeks and more	M	1968:10	2015:10	ln	USUNWK27O
Average weekly hours - total private nonfarm	M	1968:10	2015:10		USHKIP..O
Average weekly hours - manufacturing	M	1968:10	2015:10		USHKIM..O
Average overtime hours - manufacturing	M	1968:10	2015:10		USHXPMANO
Unemployment rate - 25-54 years	M	1968:10	2015:10		USUR2554Q
Unemployment rate - 55 years and over	M	1968:10	2015:10		USUR55..Q
Unemployment rate - Midwest	M	1976:01	2015:09		USLRPI24Q
Unemployment rate - Northeast	M	1976:01	2015:09		USLRR6FEQ
Unemployment rate - South	M	1976:01	2015:09		USLRKRP2Q

continued on next page

Table 3: Data set

Series	Freq.	Start date	End date	Trans-formations	Datastream code
Unemployment rate - West	M	1976:01	2015:09		USLR8Q94Q
Average weekly initial claims	M	1976:01	2015:09		USUNINSCQ
Income and earnings					
Disposable personal income per capita	M	1968:10	2015:09	ln	USINPERCB
Average hourly earnings - total private nonfarm	M	1968:10	2015:10	ln	USWRIP..B
Average hourly earnings - durable goods	M	1968:10	2015:10	ln	USWRIMD.B
Average hourly earnings - goods producing	M	1968:10	2015:10	ln	USWRPG..B
Average hourly earnings - natural resources and mining	M	1968:10	2015:10	ln	USWRIU..B
Average hourly earnings - nondurable goods	M	1968:10	2015:10	ln	USWRIMN.B
Average hourly earnings - other services	M	1968:10	2015:10	ln	USWR81..B
Average hourly earnings - professional and business services	M	1968:10	2015:10	ln	USWRIB..B
Average hourly earnings - retail trade	M	1972:01	2015:10	ln	USWRIR..B
Average hourly earnings - wholesale trade	M	1972:01	2015:10	ln	USWR42..B
Prices					
CPI - all urban	M	1968:10	2015:10	ln	USCONPRCE
CPI - all items less energy	M	1968:10	2015:10	ln	USCPXENGE
CPI - all items less food	M	1968:10	2015:10	ln	USCPXF..E
CPI - all items less medical care	M	1968:10	2015:10	ln	USCPXMEDE
CPI - all items less shelter	M	1968:10	2015:10	ln	USCPXHS.E
CPI - commodities	M	1968:10	2015:10	ln	USCPCOMME
CPI - durables	M	1968:10	2015:10	ln	USCPD...E
CPI - medical care	M	1968:10	2015:10	ln	USCPMEDCE
CPI - services	M	1968:10	2015:10	ln	USCPSERVE
CPI - transportation services	M	1968:10	2015:10	ln	USCPST..E
PPI - finished consumer goods	M	1968:10	2015:10	ln	USWPCONFE
PPI - intermediate materials, supplies and components	M	1968:10	2015:10	ln	USWPINTME
PPI - petroleum products	M	1968:10	2015:09	ln	USBCIPPEE
PCE	M	1968:10	2015:09	ln	USCP...CE
PCE - durables	M	1968:10	2015:09	ln	USCONDUCE
PCE - nondurables	M	1968:10	2015:09	ln	USCONNDCE
PCE - services	M	1968:10	2015:09	ln	USCONSRCE
Interest rates and yields					
Fed funds effective rate	M	1968:10	2015:11		FRFEDFD
Conventional mortgage points - 15 years	M	1990:01	2015:11		USMFCF1
Conventional mortgage points - 30 years	M	1990:01	2015:11		USMFCF3
FHA mortgage points	M	1990:01	2015:11		USMFGFH
US treasury bonds constant maturity - 1 year	M	1968:10	2015:11		FRTCM1Y
US treasury bonds constant maturity - 5 year	M	1968:10	2015:11		FRTCM5Y
US treasury bonds constant maturity - 10 year	M	1968:10	2015:11		FRTCM10
US treasury bill secondary market - 3 month	M	1968:10	2015:11		FRTBS3M
US treasury bill secondary market - 6 month	M	1968:10	2015:11		FRTBS6M
US commercial paper - 3 month	M	1971:04	2015:09		USI60BC.
US corporate bond yield - Moody's AAA	M	1968:10	2015:10		USCRBYLD
US corporate bond yield - Moody's BAA	M	1968:10	2015:10		USCRBBAA
US rate 3 month Euro-Dollar deposit	M	1968:10	2015:10		USOIR075R
Corporate BAA - Tbond10	M	1968:10	2015:10		
Corporate AAA - Tbond10	M	1968:10	2015:10		
Tbill6 - Tbill3	M	1968:10	2015:11		
Tbond1 - Tbill3	M	1968:10	2015:11		
Tbond5 - Tbill3	M	1968:10	2015:11		
Tbond10 - Tbill3	M	1968:10	2015:11		
CP3 - Tbill3	M	1971:04	2015:09		

continued on next page

Table 3: Data set

Series	Freq.	Start date	End date	Trans-formations	Datastream code
Mortg15 - Tbond10	M	1990:01	2015:11		
Mortg30 - Tbond10	M	1990:01	2015:11		
Money and credit					
Money supply M1	M	1968:10	2015:10	ln	USM1....B
Money supply M2	M	1968:10	2015:10	ln	USM2....B
Monetary base (adjusted for reserve requirements)	M	1968:10	2015:10	ln	USMYBSS.B
Reserve balance of depository institutions with Federal Reserve banks	M	1968:10	2015:10	ln	USRSBALNA
Consumer credit outstanding	M	1968:10	2015:09	ln	USCRDCONB
Consumer credit outstanding as share of GDP	Q	1968:11	2015:08		
Non-revolving consumer credit outstanding	M	1968:10	2015:09	ln	USCRDNRVB
Commercial and industrial loans	M	1968:10	2015:10	ln	USBCACI.B
Mortgage debt and delinquencies					
Credit market debt outstanding - home mortgages	Q	1968:12	2015:06	ln	US15MGHDB
Home mortgages as share of GDP	Q	1968:11	2015:08	ln	
Credit market instruments - total mortgages	Q	1968:12	2015:06	ln	US89MGTAB
Total mortgages -nonfinancial business, nonfarm, noncorporate	Q	1968:12	2015:06	ln	US11MGTLB
Home mortgages - nonfinancial corporate businesses	Q	1968:12	2015:06	ln	US10MGHLB
Delinquent residential mortgage loans - Northeast	Q	1979:03	2015:09		USMGDHN.R
Delinquent residential mortgage loans - South	Q	1979:03	2015:09		USMGDHS.R
Delinquent residential mortgage loans - West	Q	1979:03	2015:09		USMGDHW.R
Stock prices					
Wilshire 5000	M	1971:01	2015:11	ln	WIL5TMK
S&P 500 Composite	M	1968:10	2015:11	ln	S&PCOMP
S&P 500 Industrials	M	1989:09	2015:11	ln	SP5EIND
Dow Jones Industrials	M	1968:10	2015:11	ln	DJINDUS
Exchange rates					
USD nominal effective exchange rate	M	1975:01	2015:10	ln	USE$EF..
Exchange rate - CAD per USD	M	1968:10	2015:09	ln	CNI..AE.
Exchange rate - CHF per USD	M	1968:10	2015:09	ln	SWI..AE.
Exchange rate - EUR per USD	M	1999:01	2015:09	ln	EMI..AE.
Exchange rate - GBP per USD	M	1968:10	2015:09	ln	UKI..AE.
Exchange rate - JPY per USD	M	1968:10	2015:09	ln	JPI..AE.
Other					
ISM purchasing managers index	M	1968:10	2015:10		USCNFBUSQ
Economic policy uncertainy index (news based)	M	1985:01	2015:10		USEPUNEWR

C.2 Release calendar

Table 4: Release calendar

Series	Week of update	Lags (in months)
Housing prices		
S&P/Case-Shiller national home price index	4	2
S&P/Case-Shiller national home price index - 20 city composite	4	2
Median price of existing one family homes sold - Midwest	3	1
Median price of existing one family homes sold - Northeast	3	1
Median price of existing one family homes sold - South	3	1
Median price of existing one family homes sold - West	3	1
Average price of existing one family homes sold - Midwest	3	1
Average price of existing one family homes sold - Northeast	3	1
Average price of existing one family homes sold - South	3	1
Average price of existing one family homes sold - West	3	1
Housing starts and sales		
Housing started - 5 units or more	3	1
Housing started - Midwest	3	1
Housing started - Northeast	3	1
Housing started - South	3	1
Housing started - West	3	1
Housing authorized - Midwest	3	1
Housing authorized - Northeast	3	1
Housing authorized - South	3	1
Housing authorized - West	3	1
Sales of new one family houses	4	1
Existing one-family homes sold - Midwest	3	1
Existing one-family homes sold - Northeast	3	1
Existing one-family homes sold - South	3	1
Existing one-family homes sold - West	3	1
Housing other		
Home ownership rates - Midwest[1)]	4	1
Home ownership rates - Northeast[1)]	4	1
Home ownership rates - South[1)]	4	1
Home ownership rates - West[1)]	4	1
Rental vacancy rate - Midwest[1)]	4	1
Rental vacancy rate - Northeast[1)]	4	1
Rental vacancy rate - South[1)]	4	1
Rental vacancy rate - West[1)]	4	1
Production		
Industrial production	3	1
Industrial production - automotive products	3	1
Industrial production - business equipment	3	1
Industrial production - consumer goods	3	1
Industrial production - durable consumer goods	3	1
Industrial production - energy	3	1
Industrial production - final products	3	1
Industrial production - materials	3	1
Industrial production - nondurable consumer goods	3	1
Capacity utilization rate	3	1
GDP[1)]	2	2

continued on next page

Table 4: Release calendar

Series	Week of update	Lags (in months)
Inventories, orders and sales		
Inventories/sales ratio - manufacturing	4	2
Inventories/sales ratio - retail trade	4	2
Inventories - manufacturing	1	2
New orders - manufacturing	1	2
MFRS new orders	3	1
ISM manufacturers survey (supplier delivery index)	1	1
Employment		
Labor force - Midwest	4	1
Labor force - Northeast	4	1
Labor force - South	4	1
Labor force - West	4	1
Employed - construction	1	1
Employed - durable goods	1	1
Employed - education and health services	1	1
Employed - federal	1	1
Employed - financial activities	1	1
Employed - goods producing	1	1
Employed - government	1	1
Employed - information	1	1
Employed - leisure and hospitality	1	1
Employed - local government	1	1
Employed - manufacturing	1	1
Employed - mining	1	1
Employed - natural resources and mining	1	1
Employed - nondurable goods	1	1
Employed - nonfarm industries total	1	1
Employed - other services	1	1
Employed - private service providing	1	1
Employed - professional and business services	1	1
Employed - retail trade	1	1
Employed - state government	1	1
Employed - utlities	1	1
Employed - wholesale trade	1	1
Employment - Northeast	1	1
Employment - Midwest	1	1
Employment - South	1	1
Employment - West	1	1
Unemployed - less than 5 weeks	1	1
Unemployed - 5-14 weeks	1	1
Unemployed - 15-26 weeks	1	1
Unemployed - 15 weeks and more	1	1
Unemployed - 27 weeks and more	1	1
Average weekly hours - total private nonfarm	1	1
Average weekly hours - manufacturing	1	1
Average overtime hours - manufacturing	1	1
Unemployment rate - 25-54 years	1	1
Unemployment rate - 55 years and over	1	1
Unemployment rate - Midwest	4	1
Unemployment rate - Northeast	4	1
Unemployment rate - South	4	1
Unemployment rate - West	4	1
Average weekly intial claims	3	1

continued on next page

Table 4: Release calendar

Series	Week of update	Lags (in months)
Income and earnings		
Disposable personal income per capita	4	1
Average hourly earnings - total private nonfarm	1	1
Average hourly earnings - durable goods	1	1
Average hourly earnings - goods producing	1	1
Average hourly earnings - natural resources and mining	1	1
Average hourly earnings - nondurable goods	1	1
Average hourly earnings - other services	1	1
Average hourly earnings - professional and business services	1	1
Average hourly earnings - retail trade	1	1
Average hourly earnings - wholesale trade	1	1
Prices		
CPI - all urban	3	1
CPI - all items less energy	3	1
CPI - all items less food	3	1
CPI - all items less medical care	3	1
CPI - all items less shelter	3	1
CPI - commodities	3	1
CPI - durables	3	1
CPI - medical care	3	1
CPI - services	3	1
CPI - transportation services	3	1
PPI - finished consumer goods	3	1
PPI - intermediate materials, supplies and components	3	1
PPI - petroleum products	3	1
PCE	4	1
PCE - durables	4	1
PCE - nondurables	4	1
PCE - services	4	1
Interest rates and yields		
Fed funds effective rate	1	1
Conventional mortgage points - 15 years	1	1
Conventional mortgage points - 30 years	1	1
FHA mortgage points	1	1
US treasury bonds constant maturity - 1 year	1	1
US treasury bonds constant maturity - 5 year	1	1
US treasury bonds constant maturity - 10 year	1	1
US treasury bill secondary market - 3 month	1	1
US treasury bill secondary market - 6 month	1	1
US commercial paper - 3 month	4	1
US coporate bond yield - Moody's AAA	1	1
US coporate bond yield - Moody's BAA	1	1
US rate 3 month Euro-Dollar deposit	1	1
Corporate BAA - Tbond10	1	1
Corporate AAA - Tbond10	1	1
Tbill6 - Tbill3	1	1
Tbond1 - Tbill3	1	1
Tbond5 - Tbill3	1	1
Tbond10 - Tbill3	1	1
CP3 - Tbill3	4	1
Mortg15 - Tbond10	3	1
Mortg30 - Tbond10	3	1

continued on next page

Table 4: Release calendar

Series	Week of update	Lags (in months)
Money and credit		
Money supply M1	2	1
Money supply M2	2	1
Monetary base (adjusted for reserve requirements)	1	1
Reserve balance of depository institutions with Federal Reserve banks	1	1
Consumer credit outstanding	2	2
Consumer credit outstanding as share of GDP[1)]	2	2
Non-revolving consumer credit outstanding	2	2
Commercial and industrial loans	2	1
Mortgage debt and delinquencies		
Credit market debt outstanding - home mortgages[1)]	2	3
Home mortgages as share of GDP[1)]	2	2
Credit market instruments - total mortgages[1)]	2	3
Total mortgages -nonfinancial business, nonfarm, noncorporate[1)]	2	3
Home mortgages - nonfinancial corporate businesses[1)]	2	3
Delinquent residential mortgage loans - Northeast[1,2)]	-	-
Delinquent residential mortgage loans - South[1,2)]	-	-
Delinquent residential mortgage loans - West[1,2)]	-	-
Stock prices		
Wilshire 5000	1	1
S&P 500 Composite	1	1
S&P 500 Industrials	1	1
Dow Jones Industrials	1	1
Exchange rates		
USD nominal effective exchange rate	1	1
Exchange rate - CAD per USD	4	1
Exchange rate - CHF per USD	4	1
Exchange rate - EUR per USD	4	1
Exchange rate - GBP per USD	4	1
Exchange rate - JPY per USD	4	1
Other		
ISM purchasing managers index	1	1
Economic policy uncertainy index (news based)	1	1

[1)] If released in month of update.
[2)] Not used for sequential updates.

D Observed GDP and estimated monthly GDP series

Table 5: Quarterly observed GDP and estimated monthly GDP

Date	Observed GDP	Estimated monthly GDP
1969m5	NaN	3.2895
1969m6	3.2894	3.2923
1969m7	NaN	3.2953
1969m8	NaN	3.2967
1969m9	3.2964	3.2949
1969m10	NaN	3.2923
1969m11	NaN	3.2892
1969m12	3.2892	3.2869
1970m1	NaN	3.2846
1970m2	NaN	3.2837
1970m3	3.2836	3.2834
1970m4	NaN	3.2831
1970m5	NaN	3.2842
1970m6	3.2841	3.2866
1970m7	NaN	3.2907
1970m8	NaN	3.2907
1970m9	3.2913	3.2883
1970m10	NaN	3.2816
1970m11	NaN	3.2815
1970m12	3.2817	3.2898
1971m1	NaN	3.3042
1971m2	NaN	3.3131
1971m3	3.3136	3.3181
1971m4	NaN	3.3188
1971m5	NaN	3.3206
1971m6	3.3204	3.3234
1971m7	NaN	3.3279
1971m8	NaN	3.3304
1971m9	3.3311	3.3328
1971m10	NaN	3.3327
1971m11	NaN	3.3355
1971m12	3.3352	3.3411
1972m1	NaN	3.3506
1972m2	NaN	3.3588
1972m3	3.3590	3.3671
1972m4	NaN	3.3748
1972m5	NaN	3.3806
1972m6	3.3810	3.3854
1972m7	NaN	3.3864
1972m8	NaN	3.3906
1972m9	3.3903	3.3958
1972m10	NaN	3.4044
1972m11	NaN	3.4127
1972m12	3.4125	3.4208
1973m1	NaN	3.4291
1973m2	NaN	3.4351
1973m3	3.4349	3.4378
1973m4	NaN	3.4379
1973m5	NaN	3.4365

continued on next page

Table 5: Quarterly observed GDP and estimated monthly GDP

Date	Observed GDP	Estimated monthly GDP
1973m6	3.4368	3.4336
1973m7	NaN	3.4282
1973m8	NaN	3.4261
1973m9	3.4264	3.4294
1973m10	NaN	3.4361
1973m11	NaN	3.4386
1973m12	3.4385	3.4361
1974m1	NaN	3.4295
1974m2	NaN	3.4263
1974m3	3.4261	3.4260
1974m4	NaN	3.4289
1974m5	NaN	3.4297
1974m6	3.4298	3.4279
1974m7	NaN	3.4232
1974m8	NaN	3.4190
1974m9	3.4189	3.4165
1974m10	NaN	3.4163
1974m11	NaN	3.4133
1974m12	3.4131	3.4084
1975m1	NaN	3.4039
1975m2	NaN	3.4029
1975m3	3.4028	3.4045
1975m4	NaN	3.4089
1975m5	NaN	3.4146
1975m6	3.4142	3.4191
1975m7	NaN	3.4251
1975m8	NaN	3.4306
1975m9	3.4301	3.4334
1975m10	NaN	3.4362
1975m11	NaN	3.4401
1975m12	3.4404	3.4467
1976m1	NaN	3.4558
1976m2	NaN	3.4625
1976m3	3.4625	3.4670
1976m4	NaN	3.4700
1976m5	NaN	3.4714
1976m6	3.4714	3.4717
1976m7	NaN	3.4714
1976m8	NaN	3.4723
1976m9	3.4722	3.4748
1976m10	NaN	3.4790
1976m11	NaN	3.4836
1976m12	3.4836	3.4874
1977m1	NaN	3.4901
1977m2	NaN	3.4942
1977m3	3.4945	3.5003
1977m4	NaN	3.5071
1977m5	NaN	3.5131
1977m6	3.5131	3.5183
1977m7	NaN	3.5226
1977m8	NaN	3.5266
1977m9	3.5267	3.5307
1977m10	NaN	3.5341

continued on next page

Table 5: Quarterly observed GDP and estimated monthly GDP

Date	Observed GDP	Estimated monthly GDP
1977m11	NaN	3.5354
1977m12	3.5356	3.5351
1978m1	NaN	3.5320
1978m2	NaN	3.5352
1978m3	3.5350	3.5441
1978m4	NaN	3.5579
1978m5	NaN	3.5664
1978m6	3.5668	3.5705
1978m7	NaN	3.5694
1978m8	NaN	3.5710
1978m9	3.5706	3.5745
1978m10	NaN	3.5806
1978m11	NaN	3.5842
1978m12	3.5842	3.5843
1979m1	NaN	3.5805
1979m2	NaN	3.5782
1979m3	3.5782	3.5774
1979m4	NaN	3.5767
1979m5	NaN	3.5781
1979m6	3.5781	3.5805
1979m7	NaN	3.5832
1979m8	NaN	3.5845
1979m9	3.5846	3.5841
1979m10	NaN	3.5822
1979m11	NaN	3.5804
1979m12	3.5803	3.5788
1980m1	NaN	3.5772
1980m2	NaN	3.5718
1980m3	3.5718	3.5624
1980m4	NaN	3.5496
1980m5	NaN	3.5416
1980m6	3.5417	3.5390
1980m7	NaN	3.5415
1980m8	NaN	3.5446
1980m9	3.5446	3.5483
1980m10	NaN	3.5531
1980m11	NaN	3.5592
1980m12	3.5592	3.5672
1981m1	NaN	3.5769
1981m2	NaN	3.5818
1981m3	3.5823	3.5826
1981m4	NaN	3.5786
1981m5	NaN	3.5757
1981m6	3.5754	3.5746
1981m7	NaN	3.5756
1981m8	NaN	3.5746
1981m9	3.5746	3.5718
1981m10	NaN	3.5686
1981m11	NaN	3.5640
1981m12	3.5641	3.5581
1982m1	NaN	3.5505
1982m2	NaN	3.5475
1982m3	3.5472	3.5472

continued on next page

Table 5: Quarterly observed GDP and estimated monthly GDP

Date	Observed GDP	Estimated monthly GDP
1982m4	NaN	3.5495
1982m5	NaN	3.5500
1982m6	3.5499	3.5483
1982m7	NaN	3.5456
1982m8	NaN	3.5451
1982m9	3.5450	3.5464
1982m10	NaN	3.5489
1982m11	NaN	3.5528
1982m12	3.5529	3.5577
1983m1	NaN	3.5634
1983m2	NaN	3.5685
1983m3	3.5689	3.5748
1983m4	NaN	3.5813
1983m5	NaN	3.5879
1983m6	3.5880	3.5949
1983m7	NaN	3.6019
1983m8	NaN	3.6073
1983m9	3.6078	3.6140
1983m10	NaN	3.6196
1983m11	NaN	3.6245
1983m12	3.6246	3.6294
1984m1	NaN	3.6341
1984m2	NaN	3.6394
1984m3	3.6392	3.6448
1984m4	NaN	3.6506
1984m5	NaN	3.6553
1984m6	3.6553	3.6583
1984m7	NaN	3.6593
1984m8	NaN	3.6606
1984m9	3.6608	3.6626
1984m10	NaN	3.6648
1984m11	NaN	3.6680
1984m12	3.6678	3.6710
1985m1	NaN	3.6743
1985m2	NaN	3.6768
1985m3	3.6769	3.6792
1985m4	NaN	3.6813
1985m5	NaN	3.6849
1985m6	3.6848	3.6893
1985m7	NaN	3.6946
1985m8	NaN	3.6984
1985m9	3.6986	3.7003
1985m10	NaN	3.6997
1985m11	NaN	3.7001
1985m12	3.7000	3.7014
1986m1	NaN	3.7033
1986m2	NaN	3.7047
1986m3	3.7050	3.7057
1986m4	NaN	3.7058
1986m5	NaN	3.7057
1986m6	3.7055	3.7060
1986m7	NaN	3.7080
1986m8	NaN	3.7089

continued on next page

Table 5: Quarterly observed GDP and estimated monthly GDP

Date	Observed GDP	Estimated monthly GDP
1986m9	3.7089	3.7091
1986m10	NaN	3.7087
1986m11	NaN	3.7091
1986m12	3.7091	3.7106
1987m1	NaN	3.7124
1987m2	NaN	3.7146
1987m3	3.7146	3.7166
1987m4	NaN	3.7184
1987m5	NaN	3.7204
1987m6	3.7204	3.7224
1987m7	NaN	3.7252
1987m8	NaN	3.7282
1987m9	3.7281	3.7321
1987m10	NaN	3.7378
1987m11	NaN	3.7416
1987m12	3.7418	3.7442
1988m1	NaN	3.7452
1988m2	NaN	3.7480
1988m3	3.7477	3.7510
1988m4	NaN	3.7547
1988m5	NaN	3.7578
1988m6	3.7580	3.7603
1988m7	NaN	3.7615
1988m8	NaN	3.7631
1988m9	3.7632	3.7656
1988m10	NaN	3.7689
1988m11	NaN	3.7722
1988m12	3.7722	3.7757
1989m1	NaN	3.7791
1989m2	NaN	3.7821
1989m3	3.7820	3.7844
1989m4	NaN	3.7858
1989m5	NaN	3.7873
1989m6	3.7875	3.7895
1989m7	NaN	3.7916
1989m8	NaN	3.7927
1989m9	3.7928	3.7922
1989m10	NaN	3.7896
1989m11	NaN	3.7897
1989m12	3.7894	3.7912
1990m1	NaN	3.7941
1990m2	NaN	3.7970
1990m3	3.7968	3.7990
1990m4	NaN	3.8002
1990m5	NaN	3.8004
1990m6	3.8004	3.7994
1990m7	NaN	3.7968
1990m8	NaN	3.7933
1990m9	3.7932	3.7895
1990m10	NaN	3.7854
1990m11	NaN	3.7819
1990m12	3.7820	3.7785
1991m1	NaN	3.7749

continued on next page

Table 5: Quarterly observed GDP and estimated monthly GDP

Date	Observed GDP	Estimated monthly GDP
1991m2	NaN	3.7734
1991m3	3.7734	3.7743
1991m4	NaN	3.7770
1991m5	NaN	3.7791
1991m6	3.7794	3.7812
1991m7	NaN	3.7823
1991m8	NaN	3.7828
1991m9	3.7830	3.7834
1991m10	NaN	3.7835
1991m11	NaN	3.7841
1991m12	3.7843	3.7864
1992m1	NaN	3.7900
1992m2	NaN	3.7933
1992m3	3.7933	3.7964
1992m4	NaN	3.7997
1992m5	NaN	3.8024
1992m6	3.8023	3.8045
1992m7	NaN	3.8067
1992m8	NaN	3.8093
1992m9	3.8092	3.8122
1992m10	NaN	3.8155
1992m11	NaN	3.8175
1992m12	3.8177	3.8181
1993m1	NaN	3.8173
1993m2	NaN	3.8174
1993m3	3.8171	3.8177
1993m4	NaN	3.8194
1993m5	NaN	3.8210
1993m6	3.8211	3.8225
1993m7	NaN	3.8239
1993m8	NaN	3.8259
1993m9	3.8259	3.8292
1993m10	NaN	3.8333
1993m11	NaN	3.8370
1993m12	3.8371	3.8404
1994m1	NaN	3.8433
1994m2	NaN	3.8463
1994m3	3.8464	3.8504
1994m4	NaN	3.8549
1994m5	NaN	3.8576
1994m6	3.8578	3.8591
1994m7	NaN	3.8591
1994m8	NaN	3.8606
1994m9	3.8606	3.8639
1994m10	NaN	3.8684
1994m11	NaN	3.8716
1994m12	3.8716	3.8732
1995m1	NaN	3.8729
1995m2	NaN	3.8725
1995m3	3.8724	3.8721
1995m4	NaN	3.8713
1995m5	NaN	3.8719
1995m6	3.8721	3.8742

continued on next page

Table 5: Quarterly observed GDP and estimated monthly GDP

Date	Observed GDP	Estimated monthly GDP
1995m7	NaN	3.8769
1995m8	NaN	3.8796
1995m9	3.8797	3.8820
1995m10	NaN	3.8833
1995m11	NaN	3.8848
1995m12	3.8848	3.8861
1996m1	NaN	3.8866
1996m2	NaN	3.8902
1996m3	3.8899	3.8944
1996m4	NaN	3.9001
1996m5	NaN	3.9047
1996m6	3.9048	3.9076
1996m7	NaN	3.9083
1996m8	NaN	3.9096
1996m9	3.9095	3.9113
1996m10	NaN	3.9137
1996m11	NaN	3.9166
1996m12	3.9165	3.9193
1997m1	NaN	3.9220
1997m2	NaN	3.9254
1997m3	3.9255	3.9293
1997m4	NaN	3.9334
1997m5	NaN	3.9379
1997m6	3.9377	3.9420
1997m7	NaN	3.9466
1997m8	NaN	3.9496
1997m9	3.9497	3.9519
1997m10	NaN	3.9533
1997m11	NaN	3.9549
1997m12	3.9548	3.9567
1998m1	NaN	3.9586
1998m2	NaN	3.9603
1998m3	3.9604	3.9620
1998m4	NaN	3.9638
1998m5	NaN	3.9664
1998m6	3.9662	3.9694
1998m7	NaN	3.9728
1998m8	NaN	3.9772
1998m9	3.9771	3.9819
1998m10	NaN	3.9870
1998m11	NaN	3.9912
1998m12	3.9913	3.9946
1999m1	NaN	3.9966
1999m2	NaN	3.9992
1999m3	3.9990	4.0013
1999m4	NaN	4.0031
1999m5	NaN	4.0061
1999m6	4.0060	4.0096
1999m7	NaN	4.0138
1999m8	NaN	4.0182
1999m9	4.0181	4.0229
1999m10	NaN	4.0291
1999m11	NaN	4.0336

continued on next page

Table 5: Quarterly observed GDP and estimated monthly GDP

Date	Observed GDP	Estimated monthly GDP
1999m12	4.0336	4.0360
2000m1	NaN	4.0362
2000m2	NaN	4.0386
2000m3	4.0384	4.0434
2000m4	NaN	4.0511
2000m5	NaN	4.0552
2000m6	4.0554	4.0568
2000m7	NaN	4.0559
2000m8	NaN	4.0565
2000m9	4.0564	4.0581
2000m10	NaN	4.0607
2000m11	NaN	4.0618
2000m12	4.0619	4.0610
2001m1	NaN	4.0578
2001m2	NaN	4.0572
2001m3	4.0570	4.0590
2001m4	NaN	4.0625
2001m5	NaN	4.0637
2001m6	4.0638	4.0624
2001m7	NaN	4.0589
2001m8	NaN	4.0565
2001m9	4.0563	4.0549
2001m10	NaN	4.0543
2001m11	NaN	4.0550
2001m12	4.0551	4.0570
2002m1	NaN	4.0600
2002m2	NaN	4.0625
2002m3	4.0626	4.0646
2002m4	NaN	4.0658
2002m5	NaN	4.0675
2002m6	4.0675	4.0693
2002m7	NaN	4.0705
2002m8	NaN	4.0714
2002m9	4.0715	4.0722
2002m10	NaN	4.0726
2002m11	NaN	4.0741
2002m12	4.0738	4.0758
2003m1	NaN	4.0790
2003m2	NaN	4.0814
2003m3	4.0814	4.0838
2003m4	NaN	4.0870
2003m5	NaN	4.0915
2003m6	4.0912	4.0968
2003m7	NaN	4.1032
2003m8	NaN	4.1083
2003m9	4.1086	4.1133
2003m10	NaN	4.1176
2003m11	NaN	4.1214
2003m12	4.1213	4.1249
2004m1	NaN	4.1281
2004m2	NaN	4.1310
2004m3	4.1311	4.1338
2004m4	NaN	4.1364

continued on next page

Table 5: Quarterly observed GDP and estimated monthly GDP

Date	Observed GDP	Estimated monthly GDP
2004m5	NaN	4.1394
2004m6	4.1393	4.1427
2004m7	NaN	4.1471
2004m8	NaN	4.1509
2004m9	4.1509	4.1539
2004m10	NaN	4.1562
2004m11	NaN	4.1591
2004m12	4.1593	4.1639
2005m1	NaN	4.1698
2005m2	NaN	4.1740
2005m3	4.1741	4.1764
2005m4	NaN	4.1775
2005m5	NaN	4.1804
2005m6	4.1804	4.1854
2005m7	NaN	4.1914
2005m8	NaN	4.1954
2005m9	4.1951	4.1967
2005m10	NaN	4.1980
2005m11	NaN	4.2016
2005m12	4.2013	4.2059
2006m1	NaN	4.2113
2006m2	NaN	4.2156
2006m3	4.2156	4.2182
2006m4	NaN	4.2185
2006m5	NaN	4.2187
2006m6	4.2187	4.2188
2006m7	NaN	4.2186
2006m8	NaN	4.2194
2006m9	4.2197	4.2216
2006m10	NaN	4.2241
2006m11	NaN	4.2263
2006m12	4.2261	4.2277
2007m1	NaN	4.2287
2007m2	NaN	4.2303
2007m3	4.2302	4.2325
2007m4	NaN	4.2357
2007m5	NaN	4.2381
2007m6	4.2382	4.2399
2007m7	NaN	4.2411
2007m8	NaN	4.2424
2007m9	4.2423	4.2431
2007m10	NaN	4.2434
2007m11	NaN	4.2421
2007m12	4.2422	4.2398
2008m1	NaN	4.2361
2008m2	NaN	4.2345
2008m3	4.2342	4.2346
2008m4	NaN	4.2371
2008m5	NaN	4.2379
2008m6	4.2378	4.2371
2008m7	NaN	4.2351
2008m8	NaN	4.2311
2008m9	4.2308	4.2232

continued on next page

Table 5: Quarterly observed GDP and estimated monthly GDP

Date	Observed GDP	Estimated monthly GDP
2008m10	NaN	4.2140
2008m11	NaN	4.2073
2008m12	4.2069	4.2017
2009m1	NaN	4.1964
2009m2	NaN	4.1922
2009m3	4.1922	4.1889
2009m4	NaN	4.1868
2009m5	NaN	4.1860
2009m6	4.1857	4.1849
2009m7	NaN	4.1856
2009m8	NaN	4.1869
2009m9	4.1868	4.1891
2009m10	NaN	4.1919
2009m11	NaN	4.1946
2009m12	4.1948	4.1968
2010m1	NaN	4.1989
2010m2	NaN	4.2016
2010m3	4.2017	4.2054
2010m4	NaN	4.2102
2010m5	NaN	4.2144
2010m6	4.2143	4.2178
2010m7	NaN	4.2205
2010m8	NaN	4.2233
2010m9	4.2234	4.2264
2010m10	NaN	4.2298
2010m11	NaN	4.2310
2010m12	4.2313	4.2305
2011m1	NaN	4.2272
2011m2	NaN	4.2262
2011m3	4.2261	4.2279
2011m4	NaN	4.2320
2011m5	NaN	4.2343
2011m6	4.2346	4.2352
2011m7	NaN	4.2343
2011m8	NaN	4.2345
2011m9	4.2346	4.2362
2011m10	NaN	4.2395
2011m11	NaN	4.2423
2011m12	4.2423	4.2449
2012m1	NaN	4.2473
2012m2	NaN	4.2492
2012m3	4.2492	4.2506
2012m4	NaN	4.2519
2012m5	NaN	4.2533
2012m6	4.2533	4.2544
2012m7	NaN	4.2555
2012m8	NaN	4.2556
2012m9	4.2557	4.2554
2012m10	NaN	4.2548
2012m11	NaN	4.2554
2012m12	4.2553	4.2564
2013m1	NaN	4.2575
2013m2	NaN	4.2583

continued on next page

Table 5: Quarterly observed GDP and estimated monthly GDP

Date	Observed GDP	Estimated monthly GDP
2013m3	4.2586	4.2597
2013m4	NaN	4.2607
2013m5	NaN	4.2621
2013m6	4.2619	4.2638
2013m7	NaN	4.2661
2013m8	NaN	4.2691
2013m9	4.2690	4.2726
2013m10	NaN	4.2767
2013m11	NaN	4.2788
2013m12	4.2788	4.2783
2014m1	NaN	4.2760
2014m2	NaN	4.2761
2014m3	4.2759	4.2780
2014m4	NaN	4.2820
2014m5	NaN	4.2859
2014m6	4.2859	4.2900
2014m7	NaN	4.2939
2014m8	NaN	4.2966
2014m9	4.2967	4.2980
2014m10	NaN	4.2980
2014m11	NaN	4.2977
2014m12	4.2978	4.2968
2015m1	NaN	4.2957
2015m2	NaN	4.2958
2015m3	4.2955	4.2978
2015m4	NaN	4.3024
2015m5	NaN	4.3057
2015m6	4.3057	4.3077
2015m7	NaN	4.3089
2015m8	NaN	4.3091
2015m9	4.3092	4.3086
2015m10	NaN	4.3077
2015m11	NaN	4.3076
2015m12	4.3076	4.3088
2016m1	NaN	4.3104