Forecasting with Mixed Frequency Samples: The Case of Common Trends

BY

Peter Fuleky and Carl S. Bonham

Working Paper No. 2013-5

April 2013

Forecasting with Mixed Frequency Samples: The Case of Common Trends

Peter Fuleky and Carl S. Bonham

April 11, 2013

Abstract

We analyze the forecasting performance of small mixed frequency factor models when the observed variables share stochastic trends. The indicators are observed at various frequencies and are tied together by cointegration so that valuable high frequency information is passed to low frequency series through the common factors. Differencing the data breaks the cointegrating link among the series and some of the signal leaks out to the idiosyncratic components, which do not contribute to the transfer of information among indicators. We find that allowing for common trends improves forecasting performance over a stationary factor model based on differenced data. The "common-trends factor model" outperforms the stationary factor model at all analyzed forecast horizons. Our results demonstrate that when mixed frequency variables are cointegrated, modeling common stochastic trends improves forecasts.

JEL classifications: E37, C32, C53, L830

Keywords: Dynamic Factor Model, Mixed Frequency Samples, Common Trends, Forecasting.

1 Introduction

Empirical research generally avoids the direct use of mixed frequency data by first aggregating higher frequency series and then performing estimation and testing at the lowest frequency. As a result, information available in the high frequency dataset is not fully exploited. For example, at the end of the sample, when low frequency data has not yet been released, the most recent observations of the high frequency series are discarded. This end-of-sample information loss may be crucial when the task is to estimate current economic conditions or forecast current-quarter indicators (nowcasts).[1]

One potential solution to this problem is to use both high and low frequency data in the estimated model. In recent years there has been a growing interest in estimating macroeconomic coincident indices based on samples of mixed frequency indicators. Several studies rely on the probability model described by Stock and Watson (1991) to extract an unobserved common factor from a vector of stationary macroeconomic variables (see for example Mariano and Murasawa (2003) and Aruoba et al. (2009)). While the common factor may be a useful measure of the unobserved business cycle, the model can also be used for estimating the unobserved value of low frequency indicators at the end of the sample and beyond.

Several studies have shown improved forecasting performance using mixed frequency factor models. For example, Camacho and Perez-Quiros (2010) show that forecasts from a one-factor model of GDP growth dominate a group of institutional forecasts on a mean squared error basis, and Nunes (2005) reports an improvement in nowcasting performance of a mixed frequency model over a low frequency AR(1) model of GDP growth. Hyung and Granger (2008) find that GDP growth rate forecasts from their Linked-ARMA mixed-frequency model are more accurate than quarterly forecasts from a low-frequency model. Similarly, the GDP growth nowcasts of Evans (2005) show an improvement over advanced

---

[1]Friedman (1962) and Chow and Lin (1971) proposed an alternative approach where lower frequency series are disaggregated to higher frequency ones. Additional treatment of low frequency series with missing observations is provided by Dempster et al. (1977), Palm and Nijman (1984) and Little and Rubin (1987). A benefit of the mixed frequency factor models referenced throughout this paper is that they implicitly generate interpolated values of the low frequency series.

or preliminary GDP releases or the median Money Market Services forecast. All of the studies mentioned above estimate one-factor models using stationary monthly or quarterly macroeconomic indicators; any non-stationary levels are converted to growth rates. A potential drawback of this approach is that trends are eliminated from all input series. Yet, if the trends are shared among the indicators, modeling them as common components could improve forecasts.

Some studies extract unobserved components from mixed frequency data in levels. For example, Proietti and Moauro (2006) decompose a vector of time series in levels into a single common trend and non-stationary idiosyncratic components. Koopman and Lucas (2005) and Azevedo et al. (2006) decompose a vector of time series into common cycles and individual trend components. Rather than first differencing to remove trends, they model idiosyncratic trends and focus on extracting and analyzing business cycle indicators. Non-stationarity in the idiosyncratic components implies that individual indicator levels are allowed to diverge from each other. However, if the indicators are cointegrated, then capturing non-stationarity in the form of common trends will facilitate the transfer of information from high frequency indicators to low frequency ones, which in turn will benefit nowcasting and forecasting performance. Although not using a common factor structure, Seong et al. (2012) obtain improved forecasting performance from a mixed frequency error correction model compared to a single-frequency model.

A related literature makes use of the mixed data sampling (MIDAS) regression models first developed by Ghysels et al. (2004) and Ghysels et al. (2007). Clements and Galvão (2008, 2009) study the forecasting performance of MIDAS regression models, and Bai et al. (2013) examine the relationship between MIDAS and the Kalman filter used in mixed-frequency dynamic factor models. The latter study finds that the two methods, when applied to stationary series, produce similar forecasts. Götz et al. (2012) allow for unit roots in the data, and analyze the forecasting performance of a MIDAS based error correction model. In line with research on single-frequency models, they find that ignoring cointegration and

estimating a misspecified differenced model significantly deteriorates forecast accuracy.

In contrast to studies that extract business cycle components from a set of indicators, the focus of this paper is on the forecasting performance of small mixed-frequency dynamic factor models. We extend the existing literature by modeling the indicators in levels and allowing for multiple common factors to capture any cointegrating relationship among the indicators. In general, we expect the observed time series to follow common stochastic trends, the number of which determines the number of factors in the model. We compare the forecasting performance of the common-trends factor model in levels with the stationary factor model in differences, typically used in the literature, and illustrate the misspecification of the latter when applied to a data set containing common stochastic trends. Using both, simulated and observed data, we find that the common-trends factor model (CTFM) outperforms the stationary factor model (SFM) at all analyzed horizons. Our results demonstrate that when the indicators are integrated and cointegrated, modeling common stochastic trends, as opposed to eliminating them, will improve forecasts. In addition, we show that including high frequency data in mixed frequency models improves their forecasting performance compared to models with only low frequency aggregates.

The remainder of the paper is organized as follows. In Section 2 we give the general formulation and describe the estimation of strict dynamic multi-factor models with mixed frequency samples containing stochastic trends. In Section 3 we discuss the misspecification of the stationary factor model when the indicators contain common trends. In Section 4 we illustrate the improvement in forecasting performance of the CTFM over the SFM in a Monte Carlo setting. In the empirical application of Section 5, we contrast the forecasting performance of the common-trends factor model with the stationary factor model and a naive random walk model using macroeconomic data. Section 6 concludes.

2 Methodology

Below we give the general formulation and describe the estimation of a strict dynamic multi-factor model with mixed frequency samples. We pay special attention to non-stationary data and the handling of common stochastic trends.

2.1 Mixed Frequency Dynamic Multi Factor Model

Our analysis is based on the assumption that economic indicators can be modeled as linear combinations of two types of unobserved orthogonal processes. The first one is an $s \times 1$ vector of common factors, $\boldsymbol{f}_t$, that captures the co-movements of indicators. The second is an $n \times 1$ vector of idiosyncratic components, $\boldsymbol{\epsilon}_t$, that is driven by indicator specific shocks. In contrast to Stock and Watson (1991) and the mixed frequency implementations of their framework, we do not restrict our analysis to a single common factor. Instead, as in Macho et al. (1987), we choose the number of factors to match the number of common stochastic trends, s, in our dataset.

The underlying data generating process is assumed to evolve at a high frequency. In empirical applications the base frequency is typically set to the highest available sampling frequency. Correspondingly, the $n \times 1$ vector of observed indicators, $\boldsymbol{y}_t$, is subject to missing values at the base frequency. Indicators sampled at lower frequencies are assumed to be period aggregates of their latent base frequency counterparts. Let $\mathring{\boldsymbol{y}}_t$ denote an $n \times 1$ vector capturing the evolution of the indicators at the base frequency. For example, if $\boldsymbol{y}_t$ contains one monthly indicator and one quarterly indicator, then $\mathring{\boldsymbol{y}}_t$ contains the observed values of the monthly indicator and latent monthly values of the quarterly indicator. The $\mathring{\boldsymbol{y}}_t$ vector can be modelled as a linear combination of the two mutually uncorrelated unobserved stochastic components $\boldsymbol{f}_t$ and $\boldsymbol{\epsilon}_t$

$$\mathring{\boldsymbol{y}}_t = \mathring{\boldsymbol{\Lambda}}_f \boldsymbol{f}_t + \mathring{\boldsymbol{\Lambda}}_\epsilon \boldsymbol{\epsilon}_t \ , \tag{1}$$

where $\mathring{\boldsymbol{\Lambda}}_f$ is an $n \times s$ matrix of factor loadings and $\mathring{\boldsymbol{\Lambda}}_\epsilon$ is a diagonal $n \times n$ matrix containing

the loading parameters of idiosyncratic shocks.

In the case of non-trending data, the kth factor $f_{k,t}$, is assumed to follow an $AR(p)$ process at the base frequency

$$\phi_k(L)f_{k,t} = \eta_{k,t} , \quad \phi_k(L) = 1 - \sum_{i=1}^{p} \phi_{k,i}L^i , \quad \eta_{k,t} \sim N(0,\sigma_k) , \quad k = 1\ldots s . \quad (2)$$

If the data contains stochastic trends, the kth factor can be modeled as difference stationary process

$$\Delta f_{k,t} = \mu_k + \zeta_{k,t} , \quad \rho_k(L)\zeta_{k,t} = \eta_{k,t} , \quad \rho_k(L) = 1 - \sum_{i=1}^{p} \rho_{k,i}L^i , \quad \eta_{k,t} \sim N(0,\sigma_k) . \quad (3)$$

The s factors are assumed to be mutually uncorrelated. The idiosyncratic component associated with each variable, $\epsilon_{j,t}$, is assumed to follow an $AR(m)$ process at the base frequency

$$\gamma_j(L)\epsilon_{j,t} = \eta_{j,t} , \quad \gamma_j(L) = 1 - \sum_{i=1}^{m} \gamma_{j,i}L^i , \quad \eta_{j,t} \sim N(0,\sigma_j) , \quad j = 1\ldots n. \quad (4)$$

By definition, the idiosyncratic shocks, $\eta_{j,t}$, are mutually uncorrelated across all n indicators. To simplify notation, in the remainder of this section we will assume that the factors and the idiosyncratic components follow first order $AR(1)$ dynamics.

The problem is to estimate the parameters and the unobserved processes from the fluctuations of the observed indicators, and then use the estimated model to produce forecasts. Estimation requires that the latent high frequency variables in $\mathring{\boldsymbol{y}}_t$, or their components in $\boldsymbol{f}_t$ and $\boldsymbol{\epsilon}_t$, be aggregated to match the observed indicators in $\boldsymbol{y}_t$, which may include flows and stocks. A flow type indicator can be modeled as the accumulated sum of the common factor and idiosyncratic component during the observation period. A stock type indicator can either be modeled as a snapshot in time or as a period average of the latent high frequency variables. To illustrate the accumulation process, let's assume that indicators are sampled at monthly and quarterly frequencies. To deal with both, stock and flow variables, one can

aggregate the kth monthly common factor, $f_{k,t}$, and the jth monthly idiosyncratic component $\epsilon_{j,t}$, into $\tilde{f}_{k,t}$ and $\tilde{\epsilon}_{j,t}$, respectively, according to

$$\tilde{f}_{k,t} = \psi_t \tilde{f}_{k,t-1} + \theta f_{k,t} , \quad r = 1 \ldots s , \tag{5}$$

$$\tilde{\epsilon}_{j,t} = \psi_t \tilde{\epsilon}_{j,t-1} + \theta \epsilon_{j,t} , \quad j = 1 \ldots n , \tag{6}$$

where tilde $(\tilde{\cdot})$ denotes the aggregated value of the unobserved component (see also Harvey, 1989, Section 6.3). The cumulator variable ψ_t, is defined as

$$\psi_t = \begin{cases} 0 & \text{if } t \text{ is the first month of the quarter} \\ 1 & \text{otherwise} \end{cases} \tag{7}$$

for flows and time averaged stocks, and $\psi_t = 0$ for snapshots of stocks. The scaling variable, θ, takes on the values

$$\theta = \begin{cases} 1 & \text{for flow type variables and for snapshots of stocks} \\ 1/3 & \text{for time averaged stock variables.} \end{cases} \tag{8}$$

The model can be cast in state-space form, and we can take advantage of the Kalman filter to estimate the unobserved components. In the state space representation of the model, all unobserved components are collected in the state vector $\boldsymbol{\alpha}_t$. For the case of one factor, one monthly, and one quarterly indicator, the state vector takes the form

$$\boldsymbol{\alpha}_t = (f_t, \tilde{f}_t, \epsilon_{M,t}, \epsilon_{Q,t}, \tilde{\epsilon}_{Q,t})' , \tag{9}$$

where $f_t, \epsilon_{M,t}, \epsilon_{Q,t}$ are the base frequency (monthly) values of the factor, and the idiosyncratic components corresponding to the monthly and the quarterly variable, respectively, and tilde $(\tilde{\cdot})$ denotes the aggregated (quarterly) value of the unobserved components. The transition

6

equation

$$\boldsymbol{\alpha}_t = \boldsymbol{T}_t \boldsymbol{\alpha}_{t-1} + \boldsymbol{\eta}_t \ , \quad \boldsymbol{\eta}_t \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}) \ , \quad \boldsymbol{\alpha}_t \sim N(\boldsymbol{a}, \boldsymbol{P}) \ , \quad t = 1 \ldots T \ , \tag{10}$$

describes the evolution of the state vector. The block diagonal and time varying transition matrix $\boldsymbol{T}_t$ contains the coefficients specifying the dynamics and the temporal aggregation of the state

$$\boldsymbol{T}_t = diag(\boldsymbol{\Phi}_t, \gamma_M, \boldsymbol{\Gamma}_{Q,t}) \ , \tag{11}$$

where

$$\boldsymbol{\Phi}_t = \begin{bmatrix} \phi & 0 \\ \theta\phi & \psi_t \end{bmatrix} \ , \quad k = 1 \ldots s \ , \quad \boldsymbol{\Gamma}_{Q,t} = \begin{bmatrix} \gamma_Q & 0 \\ \theta\gamma_Q & \psi_t \end{bmatrix} . \tag{12}$$

In $\boldsymbol{\Phi}_t$ and $\boldsymbol{\Gamma}_{Q,t}$ the first row specifies the dynamics and the second row the accumulation of the unobserved components. The block diagonal covariance matrix of transition shocks, $\boldsymbol{\Sigma}$ takes the form

$$\boldsymbol{\Sigma} = diag(\boldsymbol{\Sigma}_f, \sigma_M^2, \boldsymbol{\Sigma}_Q) \ , \tag{13}$$

so that the factors and the idiosyncratic components are mutually uncorrelated. The aggregation scheme implies that

$$\boldsymbol{\Sigma}_f = \sigma_f^2 \begin{bmatrix} 1 & \theta \\ \theta & \theta^2 \end{bmatrix} \ , \quad \boldsymbol{\Sigma}_Q = \sigma_Q^2 \begin{bmatrix} 1 & \theta \\ \theta & \theta^2 \end{bmatrix} . \tag{14}$$

The measurement equation relates the observed indicators, $\boldsymbol{y}_t$, to the unobserved state vector

$$\boldsymbol{y}_t = \boldsymbol{Z}\boldsymbol{\alpha}_t \ , \tag{15}$$

where $\boldsymbol{Z}$ is a sparse matrix containing the loading coefficients of the common and idiosyn-

cratic components

$$\boldsymbol{Z} = \left[\begin{array}{ccc} \boldsymbol{\Lambda}_{f,2\times2} & \boldsymbol{\Lambda}_{M,2\times1} & \boldsymbol{\Lambda}_{Q,2\times2} \end{array} \right] = \left[\begin{array}{ccccc} \lambda_{f,M} & 0 & \lambda_M & 0 & 0 \\ 0 & \lambda_{f,Q} & 0 & 0 & \lambda_Q \end{array} \right].$$
(16)

In $\boldsymbol{\Lambda}_f$ and $\boldsymbol{\Lambda}_Q$ the first column corresponds to the monthly frequency and the second to the quarterly. Because each observed indicator is related to the unobserved components accumulated to the indicator's own frequency, each row of these sub-matrices contains at most one parameter.

The model parameters are estimated by maximum likelihood using the Kalman filter's prediction error decomposition (see Harvey, 1989, p. 125). At the end of the sample, the Kalman filter can be used to produce out-of-sample predictions of the state variables. Forecasts of the indicators are then obtained by plugging the predicted state into the measurement equation. By iterating the Kalman filter from the end of the sample h periods forward, we obtain the h-step-ahead forecasts of the variables entering the model.

2.2 Identification, Stocks and Flows

A multi-factor model is only identified up to a rotation of the factors. In a single-frequency model, column k of the factor-loading matrix corresponds to the loading of factor k on the observed indicators, and identification can be achieved by zeroing out the elements above the diagonal of the loading matrix. In a multi-frequency model, a sub-matrix $\boldsymbol{\Lambda}_{f,k}$ corresponds to the loading of the kth factor on the observed indicators, and identification requires restrictions on the rows of $\boldsymbol{\Lambda}_{f,k}$. Specifically, identification can be achieved by setting the $i < k$ rows of $\boldsymbol{\Lambda}_{f,k}$ to zero for $k = 1 \ldots s$.

Equations (5)-(8) imply that at the end of the quarter, a given unobserved component accumulated as a flow is 3 times larger than the same component accumulated as a time averaged stock. In multi-frequency factor models, the scale of the aggregated unobserved components can be controlled by multiplying the corresponding columns of the $\boldsymbol{Z}_t$ matrix by

an arbitrary δ, and multiplying the corresponding row of the state equation by $1/\delta$. Because the loading parameters in the $\boldsymbol{Z}_t$ matrix implicitly cancel out the effects of $\theta = 1/3 = 1/\delta$, the distinction between summation and averaging of the unobserved components is unidentified. Thus, it is the researcher's choice to accumulate the unobserved components as flows or time averaged stocks.

2.3   Levels vs. Differences

Economic time series are often characterized as unit root processes, and this gives rise to different approaches to specification and estimation of mixed-frequency factor models. The choice is to either explicitly model any long-run equilibrium relationships that exist among the indicators, or remove any non-stationarity from each indicator by differencing them before modeling.[2] In the first approach, the number of common factors is objectively determined by the number of common stochastic trends, s, in the n observed series, and can be deduced from the $n - s$ cointegrating relationships in the system. The cointegrating vectors are the rows of a matrix $\boldsymbol{A}_{(n-s)\times n}$ which has the property $\boldsymbol{A}\mathring{\boldsymbol{\Lambda}}_f = \boldsymbol{0}$, so that premultiplying (1) with $\boldsymbol{A}$ gives

$$\boldsymbol{A}\mathring{\boldsymbol{y}}_t = \boldsymbol{A}\mathring{\boldsymbol{\Lambda}}_\epsilon \boldsymbol{\epsilon}_t \ , \tag{17}$$

an $(n - s) \times 1$ stationary process. In contrast, when the stochastic trends are removed by differencing the series, the existing literature on mixed-frequency strict factor models generally restricts the number of common factors to a single measure of the latent business cycle.

Conversion to differences eliminates the need to estimate multiple factors and focuses attention on a single index of current business conditions. However, mixed-frequency stationary one-factor models are also frequently used for forecasting[3], and modeling differences when the indicators in levels contain common stochastic trends may leave the model mis-

---

[2]A third approach, followed by Koopman and Lucas (2005) and Azevedo et al. (2006), allows indicator specific trends to filter out the low frequency components of each indicator.

[3]Predicted levels of the indicators are derived by reversing the differencing transformation.

specified (see Section 3) and lead to poor forecasts. The conversion to differences discards the relationship between the level variables, and may amplify the noise relative to the signal in the series: high frequency indicators usually contain a large amount of noise, and differencing them further weakens their signal to noise ratio. If stationary linear combinations of the non-stationary indicators exist, that is, they are cointegrated, it may be optimal even for short horizon forecasting to keep the indicators in levels and let the factors capture the common stochastic trends (Christoffersen and Diebold, 1998).

3 Misspecification of SFM for Data with Common Trends

Below we illustrate the misspecification of stationary factor models if the indicators contain common stochastic trends. To keep the analysis as simple as possible, our data generating process contains only first order dynamics and we begin by restricting our attention to the base frequency. Specifically, assume that the true data generating process (DGP) is given by the common-trends factor model (CTFM) at the base frequency,

$$
\begin{aligned}
\mathring{\boldsymbol{y}}_t \ &= \mathring{\boldsymbol{\Lambda}}_f \boldsymbol{f}_t + \mathring{\boldsymbol{\Lambda}}_\epsilon \boldsymbol{\epsilon}_t, \\
\text{where} \quad \boldsymbol{f}_t \ &= \boldsymbol{f}_{t-1} + \boldsymbol{\zeta}_t, \\
\text{with} \quad \boldsymbol{\zeta}_t \ &= \boldsymbol{T}_\zeta \boldsymbol{\zeta}_{t-1} + \boldsymbol{\eta}_{\zeta,t}, \quad \boldsymbol{\eta}_{\zeta,t} \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_\zeta) \\
\text{and} \quad \boldsymbol{\epsilon}_t \ &= \boldsymbol{T}_\epsilon \boldsymbol{\epsilon}_{t-1} + \boldsymbol{\eta}_{\epsilon,t}, \quad \boldsymbol{\eta}_{\epsilon,t} \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}_\epsilon) \ .
\end{aligned}
\tag{18}
$$

Transforming the DGP by first differencing gives us

$$
\begin{aligned}
\Delta \mathring{\boldsymbol{y}}_t \ &= \mathring{\boldsymbol{\Lambda}}_f \Delta \boldsymbol{f}_t + \mathring{\boldsymbol{\Lambda}}_\epsilon \Delta \boldsymbol{\epsilon}_t \\
\text{with} \quad \Delta \boldsymbol{f}_t \ &= \boldsymbol{\zeta}_t = \boldsymbol{T}_\zeta \boldsymbol{\zeta}_{t-1} + \boldsymbol{\eta}_{\zeta,t}, \\
\text{and} \quad \Delta \boldsymbol{\epsilon}_t \ &= \boldsymbol{T}_\epsilon \Delta \boldsymbol{\epsilon}_{t-1} + \Delta \boldsymbol{\eta}_{\epsilon,t} \ .
\end{aligned}
\tag{19}
$$

Note, the implied model in equation (19) is different from the stationary one-factor model commonly used in the literature,

$$\Delta \mathring{\boldsymbol{y}}_t = \boldsymbol{\Lambda}_f^* f_t^* + \boldsymbol{\Lambda}_\epsilon^* \boldsymbol{\epsilon}_t^*$$

$$\text{with} \quad f_t^* = T_f^* f_{t-1}^* + \eta_{f,t}^* \tag{20}$$

$$\text{and} \quad \boldsymbol{\epsilon}_t^* = \boldsymbol{T}_\epsilon^* \boldsymbol{\epsilon}_{t-1}^* + \boldsymbol{\eta}_{\epsilon,t}^* \ .$$

The stationary model in (19) requires the same number of factors as present in the DGP, and an idiosyncratic component that follows an MA(1) process with a unit root. In contrast, the SFM in equation (20) contains a single common factor regardless of the number of factors in the DGP, and the idiosyncratic components are assumed to follow stationary iid processes. The models in (19) and (20) would only be equivalent if both contained the same number of common factors, and $\boldsymbol{\eta}_{\epsilon,t}^* = \boldsymbol{\eta}_{\epsilon,t} - \boldsymbol{\eta}_{\epsilon,t-1}$ were imposed.

To explore the relationship between the SFM and the CTFM further, we integrate the differences, $\Delta \mathring{\boldsymbol{y}}_t$, in equation (20),

$$\begin{aligned} \mathring{\boldsymbol{y}}_t &= \mathring{\boldsymbol{y}}_1 + \sum_{\tau=2}^{t} \Delta \mathring{\boldsymbol{y}}_\tau \\ &= \mathring{\boldsymbol{y}}_1 + \boldsymbol{\Lambda}_f^* \sum_{\tau=2}^{t} f_\tau^* + \boldsymbol{\Lambda}_\epsilon^* \sum_{\tau=2}^{t} \boldsymbol{\epsilon}_\tau^* \\ &= \mathring{\boldsymbol{y}}_1 + \boldsymbol{\Lambda}_f^* \breve{f}_t^* + \boldsymbol{\Lambda}_\epsilon^* \breve{\boldsymbol{\epsilon}}_t^* \ . \end{aligned} \tag{21}$$

where $\breve{f}_t^*$ is a *common* stochastic trend and $\breve{\boldsymbol{\epsilon}}_t^*$ is a vector of *idiosyncratic* stochastic trends. This is in contrast with the CTFM (DGP), where the idiosyncratic components are assumed to be stationary. Note, the idiosyncratic components in (21) will be non-stationary even if the number of factors in the SFM matches the number of common trends in the DGP unless $\boldsymbol{\eta}_{\epsilon,t}^* = \Delta \boldsymbol{\eta}_{\epsilon,t}$ is imposed in (20) so that the autoregressive and the moving-average unit roots cancel out in $\breve{\boldsymbol{\epsilon}}_t^*$. The presence of idiosyncratic stochastic trends, $\breve{\boldsymbol{\epsilon}}_t^*$, in (21) implies that the level of the extracted factor, $\breve{f}_t^*$, will arbitrarily diverge from the level of the indicators. In

contrast, cointegration ties the series together so that valuable high frequency information is incorporated into low frequency predictions through the common factors. Differencing the data breaks the cointegrating link among the indicators, and some of the signal leaks out to the idiosyncratic components, which do not contribute to the transfer of information from high frequency to low frequency series. Consequently, the forecasting performance of the SFM will be affected by inefficient transfer of information.

An additional form of misspecification may occur with differenced indicators at the aggregation stage.[4] Let's consider the multi-frequency model in which $y_{Q,t}$, the low frequency (quarterly) variable in $\boldsymbol{y}_t$ is defined as the accumulated value of its latent high frequency (monthly) counterpart in $\mathring{\boldsymbol{y}}_t$. Specifically, by aggregating the corresponding element of $\mathring{\boldsymbol{y}}_t$ in (18) to the quarterly observation frequency, we obtain

$$y_{Q,t} = \frac{1}{3}\sum_{\delta=1}^{3}\mathring{y}_{Q,t-3+\delta} \ , \tag{22}$$

where t falls on the last month of a quarter. Differencing at the quarterly frequency gives

$$\Delta_3 y_{Q,t} \;=\; y_{Q,t} - y_{Q,t-3} = \frac{1}{3}\sum_{\delta=1}^{3}\Delta_3\mathring{y}_{Q,t-3+\delta} \ , \tag{23}$$

where Δ_3 denotes the change during a quarter. Note, the change of $\mathring{y}_{Q,t}$ during a quarter can be written as the sum of monthly differences $\Delta_3\mathring{y}_{Q,t} = \sum_{\iota=1}^{3}\Delta\mathring{y}_{Q,t-3+\iota}$, so that

$$\Delta_3 y_{Q,t} \;=\; y_{Q,t} - y_{Q,t-3} = \frac{1}{3}\sum_{\delta=1}^{3}\sum_{\iota=1}^{3}\Delta\mathring{y}_{Q,t-6+\delta+\iota} \ . \tag{24}$$

This expression can be simplified by expanding the double summation into

$$\Delta_3 y_{Q,t} = \frac{1}{3}(\Delta\mathring{y}_{Q,t-4} + 2\Delta\mathring{y}_{Q,t-3} + 3\Delta\mathring{y}_{Q,t-2} + 2\Delta\mathring{y}_{Q,t-1} + \Delta\mathring{y}_{Q,t}) \ , \tag{25}$$

---

[4]See also Mariano and Murasawa (2003). For simplicity we assume that the indicators have not undergone a non-linear transformation. Issues related to the temporal aggregation of log-transformed data were discussed by Proietti and Moauro (2006).

but this aggregation becomes cumbersome to implement if the base frequency is daily and the dataset contains quarterly observations. As an alternative, Evans (2005) and Aruoba et al. (2009), who used a daily base frequency in their studies, applied the aggregation scheme (5) to differenced data. This is equivalent to an approximation of the weighted average in (25) by the simple average of a latent variable $\Delta \dot{y}_{Q,t}$

$$\Delta_3 y_{Q,t} = \frac{1}{3}(\Delta \dot{y}_{Q,t-2} + \Delta \dot{y}_{Q,t-1} + \Delta \dot{y}_{Q,t}) = \frac{1}{3}\sum_{\delta=1}^{3} \Delta \dot{y}_{Q,t-3+\delta} \ , \tag{26}$$

where $\Delta \dot{y}_{Q,t}$ is an element of $\Delta \dot{\boldsymbol{y}}_t$, which can be defined similarly to (20)

$$\Delta \dot{\boldsymbol{y}}_t = \dot{\boldsymbol{L}}^* \dot{f}_t^* + \dot{\boldsymbol{G}}^* \dot{\boldsymbol{\epsilon}}_t^* \quad \text{with} \quad \dot{f}_t^* = \dot{T}_f^* \dot{f}_{t-1}^* + \ddot{\eta}_{f,t}^* \quad \text{and} \quad \ddot{\boldsymbol{\epsilon}}_t^* = \dot{\boldsymbol{T}}_\epsilon^* \ddot{\boldsymbol{\epsilon}}_{t-1}^* + \dot{\boldsymbol{\eta}}_{\epsilon,t}^* \ . \tag{27}$$

Note that while (25) spans two quarterly observation periods, (26) only spans one, and therefore the approximation directly affects the dynamics of the model.

To summarize, when the DGP contains common stochastic trends, the SFM has misspecified dynamics both at the base frequency and the observation frequency, and introduces idiosyncratic stochastic trends. These forms of misspecification will affect the decomposition of the observed data into common and idiosyncratic components, lead to inefficient transfer of information among the variables, and deteriorate the forecasting performance of the model.

4   Simulations

In this section we investigate the relative forecasting performance of the CTFM and SFM when the data generating process coincides with the CTFM. In addition, to illustrate the benefit of high frequency data in these models, we include in our comparison the single-frequency counterpart of the CTFM that uses temporally aggregated indicators. To keep the exercise simple, we restrict our mixed frequency model to three indicators (two monthly

and one quarterly), one or two stochastic common trends, and first order dynamics in the idiosyncratic components. Each stochastic trend contains drift and autocorrelated errors. The monthly variables are assumed to be released with a one-month lag, and the quarterly variable with a three-month lag. Figure 1 illustrates the timing of data releases and forecast horizons relative to the forecast date. The passage of time is indicated by a shift of the forecast date, T, forward in time. The information set increases when there is a new release of the monthly and quarterly indicators. In the notation for the quarterly forecasts, $Q_{i,T+j}$, i indicates the quarter for which the forecast is being made, and j indicates the forecast horizon in months. Note, at any given forecast date, only a subset of forecast horizons falls on an end-of-quarter month.

Having specified the data generating process, we simulate a 21-year long sample of monthly and quarterly observations from a given random seed. We reserve the first 20 years of the sample for estimation of the CTFM and SFM models, and then produce forecasts for horizons ranging from -2 months to $+10$ months, from which only every third one coincides with end-of-quarter months. To obtain forecasts for the horizons that did not fall on an end-of-quarter month, we repeat the forecasting exercise recursively (extend the size of the sample by one month, repeat estimation and forecasting) two more times. Once we obtained a quarterly forecast for each horizon between -2 months to $+10$ months, we compare the predictions to the corresponding values in the remainder of the simulated sample.

We repeat the forecasting exercise using the CTFM and SFM described above for 1000 different seeds and 6 different parameterizations listed in Table 1. In these scenarios the loading parameters across the common stochastic trends always add up to 1, and the idiosyncratic components follow an autoregressive process with either low or high persistence, denoted by subscript LP or HP, respectively. Scenario A represents a situation when two common stochastic trends have similar loadings on all three variables. Scenario B illustrates the case when the two monthly variables do not share common stochastic trends. Finally, Scenario C illustrates the case when all variables are affected by only a single com-
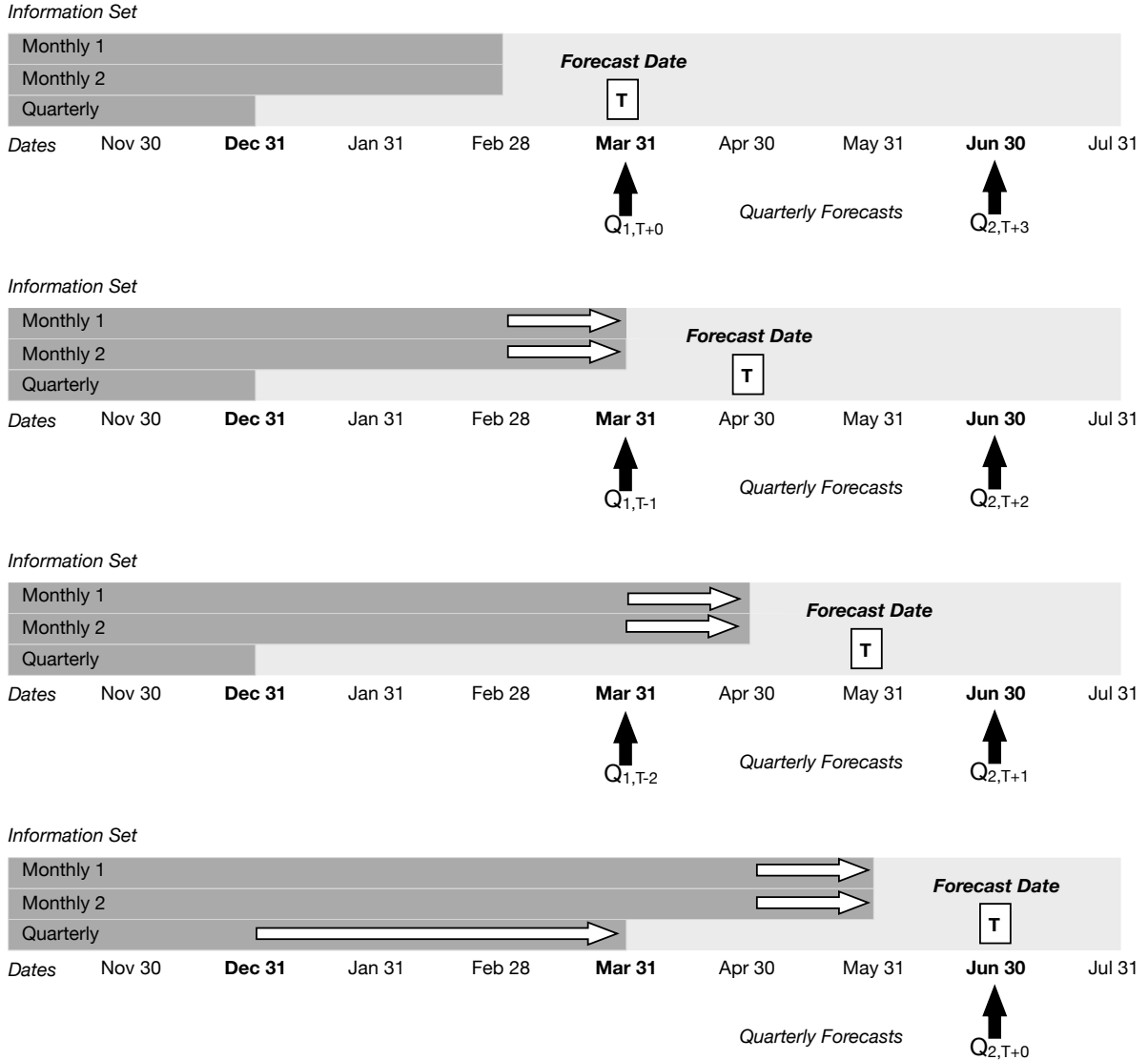
Figure 1: Illustration of the expansion of the information set and the monthly forecast horizons for the quarterly variable relative to the forecast date T. In $Q_{i,T+j}$, i indicates the quarter for which the forecast is being made, and j indicates the forecast horizon in months. The figure captures the passage of time from the end of March (top panel) through the end of June (bottom panel) in monthly increments.

mon stochastic trend. We compare the CTFM to SFMs with one and two factors under scenarios A and B, but only to a one-factor SFM under scenario C. We also compare the mixed frequency CTFM to its single-frequency counterpart: the quarterly common trends factor model (QCTFM) uses indicators aggregated to the quarterly frequency. While the

15

comparison of the CTFM with the SFM illustrates the effect of eliminating common trends in mixed frequency models, the comparison of the CTFM with its quarterly counterpart illustrates the benefit of harnessing high frequency data in the correctly specified common trends factor model. For each parameterization and each forecast horizon, we evaluate the improvement in forecasting performance of the CTFM relative to the SFM and QCTFM by calculating the reduction in root mean squared forecast error (RMSE) across the 1000 different random seeds. The results, categorized by the various forms of the data generating process and estimated models, are listed in Table 2.

Table 1: Data generating process parameters

Parameter	Scenarios					
	A_{LP}	A_{HP}	B_{LP}	B_{HP}	C_{LP}	C_{HP}
$\lambda^1_{M_1}$	0.6	0.6	1.0	1.0	1.0	1.0
$\lambda^1_{M_2}$	0.4	0.4	0.0	0.0	1.0	1.0
λ^1_Q	0.5	0.5	0.5	0.5	1.0	1.0
$\lambda^2_{M_1}$	0.4	0.4	0.0	0.0	0.0	0.0
$\lambda^2_{M_2}$	0.6	0.6	1.0	1.0	0.0	0.0
λ^2_Q	0.5	0.5	0.5	0.5	0.0	0.0
$\gamma_1 = \gamma_2 = \gamma_3$	0.1	0.7	0.1	0.7	0.1	0.7

Note: In the analyzed scenarios only the displayed parameters vary. The remaining parameters are held fixed at the following values: For each idiosyncratic component the loading parameter is set to $g_1 = g_2 = g_3 = 1$, and the standard deviation of each idiosyncratic shock is set to $\sigma_1 = \sigma_2 = \sigma_3 = 0.02$. The two stochastic trends are drifting with speed $\mu^1 = 0.015$ and $\mu^2 = 0.005$, respectively, and have autocorrelated errors with persistence $\rho^1 = 0.85$ and $\rho^2 = 0.75$, and variance $\sigma^1 = 0.01$ and $\sigma^2 = 0.005$, respectively.

As shown in Section 3, when the DGP contains common stochastic trends the SFM is misspecified whether it uses the correct number of factors or not: it introduces spurious idiosyncratic stochastic trends and suffers from misspecified dynamics. Therefore it is not surprising that the CTFM outperforms the SFM for all considered parameterizations and at all analyzed horizons. The different combinations of factor loadings do not seem to have large effects on the results. However, a reduction in idiosyncratic dynamics (LP) results in greater improvements in relative backcasting and nowcasting performance of the CTFM. For short forecast horizons the CTFM outperforms the SFM by 20% to 35%, depending on

Table 2: % Reduction in RMSE: CTFM Compared to SFM and QCTFM

| | Scenarios | | | | | | | |
Horizon	A^1_{LP}	A^2_{LP}	A^Q_{LP}	B^1_{LP}	B^2_{LP}	B^Q_{LP}	C^1_{LP}	C^{Q1}_{LP}
T-2	-33%	-32%	-40%	-30%	-32%	-39%	-31%	-36%
T-1	-32%	-32%	-37%	-29%	-32%	-37%	-31%	-36%
T	-27%	-28%	-29%	-23%	-27%	-29%	-29%	-32%
T+1	-37%	-35%	-52%	-31%	-36%	-52%	-34%	-58%
T+2	-30%	-31%	-42%	-26%	-31%	-43%	-29%	-49%
T+3	-21%	-26%	-28%	-16%	-25%	-30%	-23%	-36%
T+4	-23%	-25%	-41%	-16%	-26%	-42%	-19%	-48%
T+5	-18%	-23%	-32%	-13%	-23%	-33%	-16%	-38%
T+6	-10%	-20%	-20%	-6%	-20%	-20%	-12%	-25%
T+7	-14%	-22%	-31%	-8%	-22%	-32%	-12%	-37%
T+8	-11%	-21%	-24%	-7%	-21%	-25%	-11%	-29%
T+9	-6%	-21%	-15%	-3%	-20%	-15%	-9%	-20%
T+10	-10%	-23%	-25%	-5%	-23%	-25%	-10%	-29%

Horizon	A^1_{HP}	A^2_{HP}	A^Q_{HP}	B^1_{HP}	B^2_{HP}	B^Q_{HP}	C^1_{HP}	C^Q_{HP}
T-2	-17%	-17%	-20%	-13%	-14%	-19%	-16%	-17%
T-1	-17%	-17%	-18%	-13%	-14%	-18%	-15%	-16%
T	-14%	-16%	-13%	-10%	-12%	-13%	-16%	-15%
T+1	-25%	-28%	-33%	-20%	-24%	-33%	-26%	-38%
T+2	-21%	-27%	-27%	-18%	-23%	-28%	-23%	-33%
T+3	-16%	-25%	-18%	-13%	-20%	-20%	-22%	-26%
T+4	-20%	-29%	-31%	-15%	-24%	-32%	-22%	-40%
T+5	-16%	-28%	-23%	-13%	-23%	-25%	-20%	-32%
T+6	-10%	-25%	-15%	-7%	-20%	-16%	-16%	-23%
T+7	-13%	-30%	-26%	-9%	-24%	-26%	-16%	-34%
T+8	-10%	-29%	-19%	-8%	-24%	-21%	-14%	-27%
T+9	-6%	-27%	-12%	-4%	-22%	-13%	-13%	-19%
T+10	-9%	-32%	-21%	-6%	-27%	-22%	-13%	-28%

Note: The superscripts 1, 2, and Q in the scenario designations indicate comparison of the CTFM to the SFM with 1 and 2 factors, and the QCTFM, respectively. The table shows the percentage difference in RMSE for the CTFM forecasts relative to the forecasts produced by the SFM with 1 and 2 factors, and the QCTFM. The results are based on 1000 repetitions for each horizon. The forecast horizon is measured in months relative to the forecast date, T. The Diebold and Mariano (1995) test indicates that the improvements are significant at the 5% level for each scenario at all horizons.

model specification.

The improvements of the CTFM over the one-factor SFM remain similar for one and two factors in the DGP. And, while adding a second factor to the SFM has little effect on the short

horizon results, it does cause a deterioration in long horizon forecast performance. These findings suggest that additional factors in the SFM split the differenced indicators into more unobserved components, but do not improve model specification. As our empirical results in Section 5.1 illustrate, the first factor in the SFM captures most of the signal in the differenced data set, and the second factor is heavily affected by noise. Consequently, the integrated value of the second stationary factor has a limited effect at short horizons, but leads to a deterioration in forecasting performance at longer horizons.
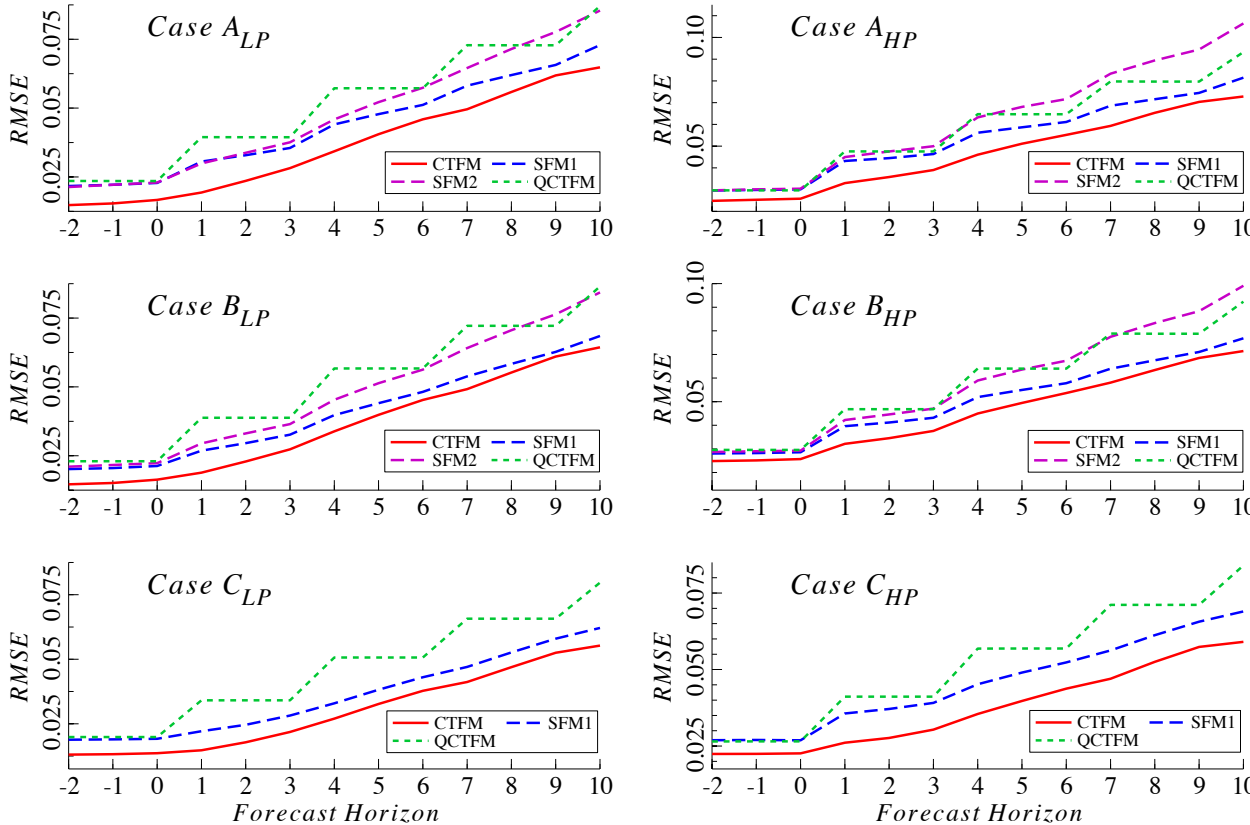

Figure 2: Root mean squared forecast error by scenario, for all considered horizons for the CTFM, SFM with one and two factors, and the QCTFM.

Figure 2 illustrates the root mean squared forecast error for the CTFM, SFM with one and two factors, and the QCTFM for the considered forecast horizons. Because the QCTFM does not benefit from any intra-quarter information, its RMSE remains constant within a quarter but exhibits a decline for horizons that are full quarters (3, 6, 9 and 12 months)

18

away from the release date of the quarterly series. In contrast the CTFM and SFM forecasts do take advantage of the intra-quarter information and have therefore smoother evolution of RMSE as the forecast horizon changes. The low level and smoothness of the CTFM RMSE is an indication of its ability to transmit information in the high frequency indicators to the quarterly forecasts.

5   Empirical Application

To illustrate the forecasting performance of the CTFM relative to the SFM, we make use of monthly real personal income (RPI) and real personal consumption expenditures (PCE) to forecast quarterly real gross domestic product (GDP). Table 3 summarizes the sampling

Table 3: Indicators

	Sampling Freq.	Reporting Lag
Real Personal Income, RPI	monthly	1 month
Real Personal Consumption Expenditures, PCE	monthly	1 month
Real Gross Domestic Product, GDP	quarterly	3 months

Note: All series were obtained from the Federal Reserve Economic Database (FRED). The second revision of real gross domestic product is released three months after a quarter ends. Note that we do not construct a real-time data set as in Giannone et al. (2008); our data set reflects all the revisions prior to January 25, 2013, the day we obtained the data.

frequencies and reporting lags of the indicators in our study. Although monthly real personal consumption expenditures are only available after January 1995, the nominal value of the series and the PCE chain type price index are available since 1959, and we use these two series to calculate an extended history of PCE. Accordingly, we set the start of our dataset to January 1959 (GDP is available from 1947). Because the variables exhibit exponential growth, we apply a logarithmic transformation to the data.

Table 4 reports results from augmented Dickey-Fuller (1979) (ADF) tests for the null hypothesis of a unit root in each of our indicators. We can not reject the null hypothesis of a unit root for any of the series at the 10% significance level. To test for cointegration, we apply Johansen's (1988) rank test to a temporally aggregated quarterly system. While the

Table 4: Augmented Dickey-Fuller tests

$$\Delta y_t = \alpha + \beta y_{t-1} + \sum_{k=1}^{m} \theta_t \Delta y_{t-k} + \epsilon_t \; ; \quad H_0 : \beta = 0$$

	$1 + \hat{\beta}$	ADF t-test	p-value
RPI	0.999	-2.169	0.218
PCE	0.999	-1.836	0.363
GDP	0.997	-2.304	0.171

Note: Column 1 lists the series tested for a unit root; column 2 presents the estimated AR(1) parameter; column 3 the the ADF t-test for the null hypothesis $\beta = 0$; and column 4 presents the marginal significance level for the ADF t-test. The lag-length, m is determined by testing down from the maximum of 14 lags.

number of cointegrating vectors in the system is invariant to temporal aggregation, the finite sample power of tests may fall as the number of observations declines (see Marcellino, 1999). Therefore, we use the rank test to obtain *initial estimates* of the number of cointegrating vectors and *verify* our findings through unit root tests on the residuals from our CTFM. Results in Table 5 indicate that we can reject the null hypothesis of no cointegrating vectors,

Table 5: Cointegration rank tests

Rank	Eigenvalue	Trace test	p-value	λ-max test	p-value
$r = 0$	0.110	37.815	0.004	24.860	0.012
$r = 1$	0.043	12.955	0.117	9.445	0.257
$r = 2$	0.016	3.511	0.061	3.511	0.061

Note: Column 1 lists the null hypothesis of zero, at least one, and at least two cointegrating vectors; column 2 the eigenvalue; column 3 the trace test; column 4 the marginal significance level for the trace tests; column 5 the maximum eigenvalue test; and column 6 the marginal significance level for the maximum eigenvalue test. The test is evaluated with an unrestricted constant, and the lag-length is determined by the Schwarz-Bayesian Information Criterion.

and we are unable to reject the hypothesis of at most one cointegrating vector at the 5% significance level. We tentatively conclude that our three indicators contain two common stochastic trends. We verify this result by estimating our model with one and two factors containing a unit root, and find that the idiosyncratic errors become stationary once two such factors are included in the model. In other words, a linear combination of two common stochastic trends is able to explain the non-stationarity of the individual variables.

We standardize all variables before estimation. Predicted values of the indicators are obtained by reversing the standardization and log-transformation. To keep the model as simple as possible, we follow Nunes (2005), Aruoba et al. (2009), and others by restricting the unobserved components to at most first order dynamics. While these models may suffer from misspecified dynamics, the dynamic structure won't change as new data becomes available, as it would if we allowed for variable lag lengths.

5.1 Estimation Results

Table 6 displays the estimation results for the CTFM with two stochastic trends using the full data set, and Figure 3 displays the decomposition of the standardized log-levels into the two common stochastic trends and the idiosyncratic components.[5] From the two factors, the first one has a slightly larger drift and variance of the errors (μ_1, σ_1). The idiosyncratic components have approximately similar error variances, but their persistence varies considerably. While the idiosyncratic component of GDP is persistent, it is stationary, implying that the two factors have successfully captured the common stochastic trends in the three variables.

Table 6: Estimation Results: Common-Trends Factor Model

$\lambda_{1,RPI}$	$\lambda_{1,PCE}$	$\lambda_{1,GDP}$	$\lambda_{2,RPI}$	$\lambda_{2,PCE}$	$\lambda_{2,GDP}$
1.000	0.721*	0.781*	*0.000*	*1.000*	0.792*
μ_1	ϕ_1	σ_1	μ_2	ϕ_2	σ_2
0.005*	0.703*	0.005*	0.001*	0.737*	0.002*
γ_{RPI}	γ_{PCE}	γ_{GDP}	σ_{RPI}	σ_{PCE}	σ_{GDP}
0.449*	0.258*	0.940*	0.008*	0.007*	0.009*

Note: Maximum likelihood parameter estimates for CTFM containing two common stochastic trends with drift (μ) and autocorrelated errors (ϕ). * denotes significance at the 5% level. To satisfy identification requirements, $\lambda_{2,RPI}$ is fixed at 0, and $\lambda_{1,RPI}$, $\lambda_{2,PCE}$, and the loading parameters of the idiosyncratic components are fixed at 1.

Table 7 displays the estimation results for the SFM with one and two factors using the

---

[5]We estimated the model by Ox 5.10 (Doornik, 2007) and SsfPack 2.2 (Koopman et al., 1999).
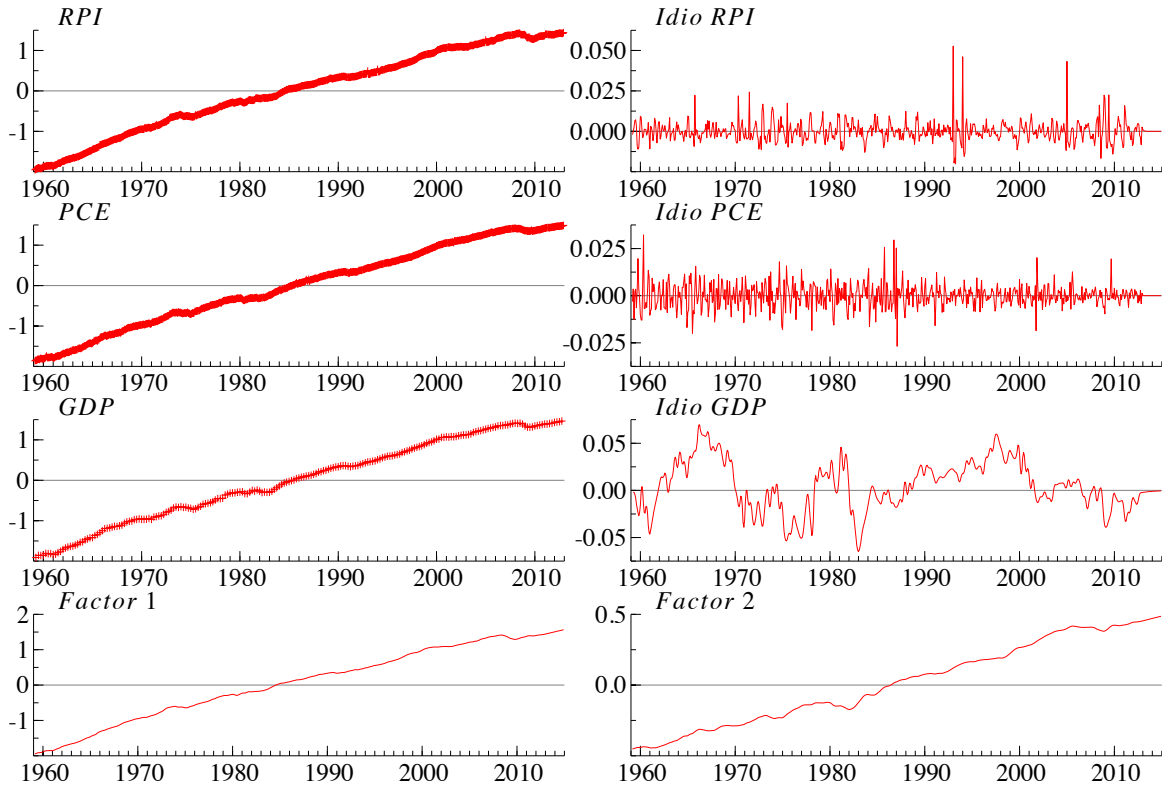
Figure 3: Standardized log-levels of real personal income, real personal consumption expenditures, and real gross domestic product, and their decomposition into common stochastic trends and stationary idiosyncratic components.

full data set. When a second factor is considered, its autocorrelation, variance and loading are all statistically insignificant. This implies that the contribution of the second factor to the explanatory power of the model is limited, and the first factor captures most of the signal in the data. The remaining parameter estimates are significant and fairly similar across the one-factor and two-factor models. Figure 4 displays the decomposition of the standardized differences into the idiosyncratic components and two common factors (except for the second factor, there is no qualitative difference in the subplots when only one factor is considered). The first factor seems to capture most of the cyclical information, whereas the second factor and the idiosyncratic components absorb quickly decaying shocks.

Table 7: Estimation Results: Stationary Factor Models

One-Factor Model					
$\lambda_{1,\Delta RPI}$	$\lambda_{1,\Delta PCE}$	$\lambda_{1,\Delta GDP}$			
1.000	0.876*	2.046*			
	ϕ_1	σ_1			
	0.795*	0.246*			
$\gamma_{\Delta RPI}$	$\gamma_{\Delta PCE}$	$\gamma_{\Delta GDP}$	$\sigma_{\Delta RPI}$	$\sigma_{\Delta PCE}$	$\sigma_{\Delta GDP}$
-0.245*	-0.401*	-0.130*	0.895*	0.891*	1.204*

Two-Factor Model					
$\lambda_{1,\Delta RPI}$	$\lambda_{1,\Delta PCE}$	$\lambda_{1,\Delta GDP}$	$\lambda_{2,\Delta RPI}$	$\lambda_{2,\Delta PCE}$	$\lambda_{2,\Delta GDP}$
1.000	0.883*	2.053*	*0.000*	*1.000*	0.519
	ϕ_1	σ_1		ϕ_2	σ_2
	0.766*	0.263*		0.000	0.050
$\gamma_{\Delta RPI}$	$\gamma_{\Delta PCE}$	$\gamma_{\Delta GDP}$	$\sigma_{\Delta RPI}$	$\sigma_{\Delta PCE}$	$\sigma_{\Delta GDP}$
-0.199*	-0.290*	-0.114*	0.891*	0.886*	1.192*

Note: Maximum likelihood parameter estimates for SFM. * denotes significance at the 5% level. To satisfy identification requirements, $\lambda_{2,\Delta RPI}$ is fixed at 0, and $\lambda_{1,\Delta RPI}$, $\lambda_{2,\Delta PCE}$, and the loading parameters of the idiosyncratic components are fixed at 1.
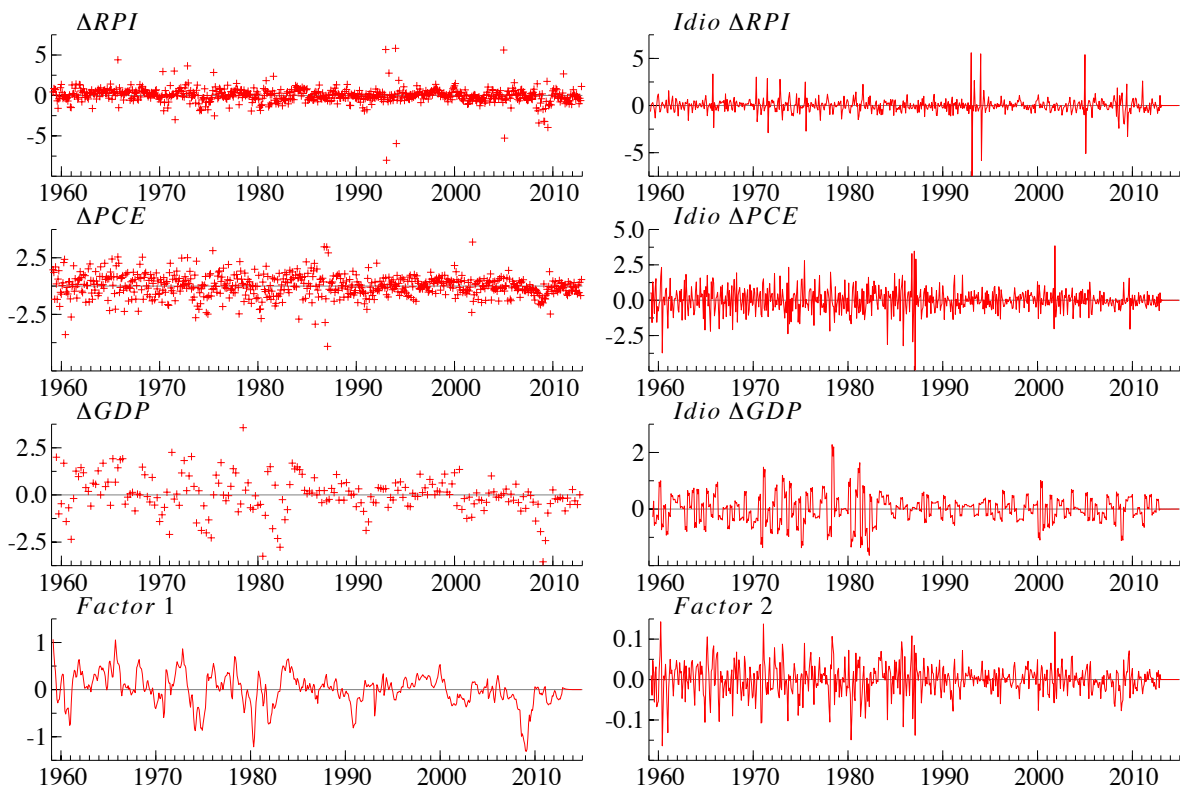
Figure 4: Standardized log-differences of real personal income, real personal consumption expenditures, and real gross domestic product, and their decomposition into two stationary common factors and idiosyncratic components.

24

5.2 Forecasting Results

The quality of forecasts depends in part on what information is available at the time the forecast is made. Although we do not construct a real-time data set as in Giannone et al. (2008), we do replicate the sequence of data releases for the three variables. The model is estimated, and a forecast is made at the end of each month between January 1979 and December 2012. At each forecast date, T, the latest available monthly observation is for the previous month (with a time stamp of $T-1$). The quarterly indicator is only available three months after the end of the quarter. Figure 1 in Section 4 gives an illustration of the timing of data releases and forecast horizons relative to the forecast date T. As the forecast date moves forward, the amount of available information increases, and for a given target date the forecast horizon shrinks. We expect the flow of high frequency information before the release of GDP to improve estimates of the GDP series.

Table 8 compares the accuracy of predictions from the CTFM to those from the SFMs with one and two factors, and a quarterly factor model (QCTFM) with the monthly series aggregated to the quarterly frequency. We report the root mean squared error (RMSE) of forecasts, the percentage difference in RMSE across models, and the 5% marginal significance of the Diebold and Mariano (1995) tests for forecast accuracy. The GDP predictions produced by the mixed-frequency models show an improvement over the single-frequency QCTFM forecasts: the RMSE reduction for the CTFM over the QCTFM model ranges from 17% to 43%. The quarterly factor model of GDP does not benefit from the high frequency information that becomes available within a quarter. Therefore the RMSE from this model is constant *between GDP* release dates but exhibits a drop *on GDP* release dates.

The mixed frequency models are re-estimated for each expansion of the information set, and the precision of these models for the forecast horizon $T-2$ is only influenced by updated parameter values. The main benefit of mixed frequency models comes from the incorporation of intra-quarter high frequency information into the end-of-quarter GDP estimates. As Figure 1 in Section 4 illustrates, the GDP predictions for horizons $T-1$ and higher are

Table 8: *GDP* Forecasting Results: Comparison of CTFM, SSFM and RW Model

Horizon	RMSE				% Difference		
	CTFM	2SFM	1SFM	QM	CT/2S	CT/1S	CT/QM
T-2	48	54	54	68	-11%*	-11%*	-29%*
T-1	48	56	57	68	-15%*	-14%*	-28%*
T	51	57	58	68	-12%*	-11%*	-25%*
T+1	72	83	83	126	-14%*	-14%*	-43%*
T+2	79	94	94	126	-15%*	-16%*	-37%*
T+3	89	98	98	126	-9%*	-9%*	-29%*
T+4	111	125	124	182	-10%*	-11%*	-38%*
T+5	126	142	141	182	-10%*	-11%*	-30%*
T+6	141	149	148	182	-5%*	-5%*	-22%*
T+7	158	172	170	233	-7%*	-8%*	-32%*
T+8	174	190	189	233	-8%*	-8%*	-25%*
T+9	192	199	197	233	-3%	-3%	-17%*
T+10	206	220	218	280	-6%*	-6%*	-26%*

Note: RMSE of *GDP* forecasts (units: US $ Billion), and percentage difference between the RMSE of *GDP* forecasts. The forecast horizon is measured relative to the forecast date, T. The CTFM is compared to SFMs with one and two factors, and to a quarterly two-factor model. Marginal significance of the Diebold and Mariano (1995) test at the 5% level is indicated by *.

directly affected by the releases of the monthly series. The impact of this high frequency information is then propagated to longer horizon forecasts.

As foreshadowed by the simulation results, the CTFM is more precise than the SFMs at all horizons. Cointegration ties the series together so that valuable high frequency information is passed to low frequency forecasts through the common factors. Therefore two common stochastic trends predict the evolution of the indicators more effectively than one or two stationary factors. The transformation in the SFM removes all long-run trend information leaving the model with a set of relatively noisy variables to analyze. The first stationary factor quite successfully captures common cyclical information, but that is not sufficient to reconstruct all the comovement of the variables in levels. The insignificant loading of the second stationary factor implies that it does not capture any meaningful amount of information, and has a minimal effect on the forecast. Figure 5 illustrates that the integrated idiosyncratic components in the SFM with two factors are much more persistent than the idiosyncratic components in the CTFM in Figure 4. This occurs because differencing breaks

the cointegrating link among the variables, and as a consequence the SFM is unable to fully extract the common stochastic trends in the data. In the SFM some of the signal leaks out to the idiosyncratic components, which do not contribute to the transfer of information from high frequency indicators to low frequency ones. By making use of a linear combination of the two common stochastic trends, the CTFM is better able to capture and transfer the relevant high frequency information from the monthly variables to the quarterly one. Because the CTFM incorporates the intra-period information more accurately than is possible in the misspecified SFM, the forecasts based on the former model have a lower RMSE than forecasts that are based on the latter.
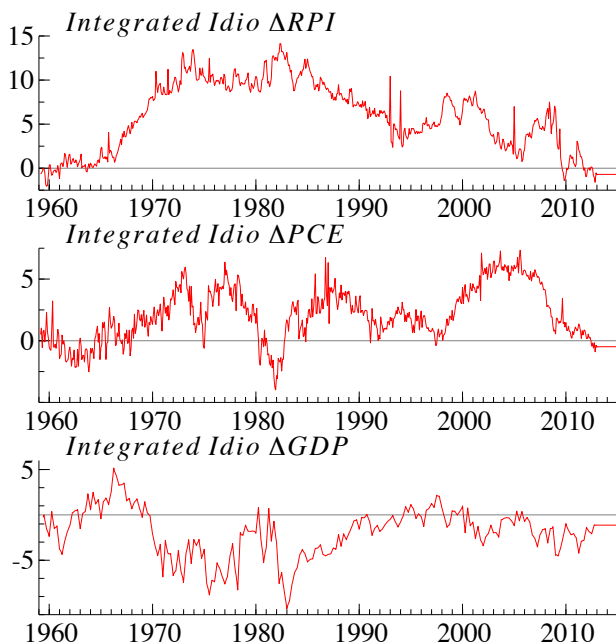

Figure 5: Integrated idiosyncratic components from the SFM with two factors.

6 Conclusion

We analyze the forecasting performance of small mixed frequency factor models when the observed variables share stochastic trends. We allow for multiple common factors to capture potential cointegrating relationships among the levels of the observed variables. Our

comparison of mixed and single-frequency models demonstrates that forecasts benefit from incorporating high frequency information into the model. However, in the presence of common stochastic trends the mixed frequency stationary one-factor models that are frequently used to extract coincident indicators are misspecified. The elimination of common stochastic trends leads to inefficient information transfer from high frequency indicators to low frequency ones. The forecasting performance of the stationary factor model suffers even if the model contains the same number of common factors as the data generating process. The common-trends factor model outperforms the stationary factor model at all forecast horizons, and the gain in precision as measured by the root mean squared forecast error tends to be the strongest for nowcasts and short horizon forecasts. Our results illustrate that when the constituent indicators are cointegrated, modeling common stochastic trends improves short horizon forecasts.

References

Aruoba, S., Diebold, F., and Scotti, C. (2009). Real-time measurement of business conditions. *Journal of Business and Economic Statistics*.

Azevedo, J. V. e., Koopman, S. J., and Rua, A. (2006). Tracking the business cycle of the euro area: A multivariate model-based bandpass filter. *Journal of Business & Economic Statistics*, 24:278–290.

Bai, J., Ghysels, E., and Wright, J. (2013). State space models and MIDAS regressions. *Econometric Reviews*, 32(7):779–813.

Camacho, M. and Perez-Quiros, G. (2010). Introducing the euro-sting: Short-term indicator of euro area growth. *Journal of Applied Econometrics*, 25(4):663–694.

Chow, G. and Lin, A. (1971). Best linear unbiased interpolation, distribution, and extrapolation of time series by related series. *The review of Economics and Statistics*, 53(4):372–375.

Christoffersen, P. and Diebold, F. (1998). Cointegration and long-horizon forecasting. *Journal of Business & Economic Statistics*, 16(4):450–458.

Clements, M. and Galvão, A. (2008). Macroeconomic forecasting with mixed-frequency data. *Journal of Business and Economic Statistics*, 26(4):546–554.

Clements, M. and Galvão, A. (2009). Forecasting us output growth using leading indicators: An appraisal using midas models. *Journal of Applied Econometrics*, 24(7):1187–1206.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 1–38.

Dickey, D. and Fuller, W. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74:427–431.

Diebold, F. and Mariano, R. (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, 13(3).

Doornik, J. (2007). *Object-oriented matrix programming using Ox*. Timberlake Consultants Press and Oxford: www.doornik.com, 3rd ed. edition.

Evans, M. (2005). Where Are We Now? Real-Time Estimates of the Macroeconomy. *International Journal of Central Banking*.

Friedman, M. (1962). The interpolation of time series by related series. *Journal of the American Statistical Association*, 57(300):729–757.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2004). The MIDAS Touch: Mixed Data Sampling Regression Models. *Working paper*.

Ghysels, E., Sinko, A., and Valkanov, R. (2007). Midas regressions: Further results and new directions. *Econometric Reviews*, 26(1):53–90.

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.

Götz, T., Hecq, A., and Urbain, J. (2012). Forecasting mixed frequency time series with ecm-midas models. *DP Maastricht University*.

Harvey, A. (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge Univ Pr.

Hyung, N. and Granger, C. (2008). Linking series generated at different frequencies. *Journal of Forecasting*, 27(2):95–108.

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

Koopman, S. and Lucas, A. (2005). Business and default cycles for credit risk. *Journal of Applied Econometrics*, 20(2):311–323.

Koopman, S., Shephard, N., and Doornik, J. (1999). Statistical algorithms for models in state space using SsfPack 2.2. *Econometrics Journal*, 2(1):107–160.

Little, R. J. and Rubin, D. B. (1987). *Statistical analysis with missing data*, volume 539. Wiley New York.

Macho, F. J. F., Harvey, A. C., and Stock, J. H. (1987). Forecasting and interpolation using vector autoregressions with common trends. *Annales d'Economie et de Statistique*, (6/7):279–287.

Marcellino, M. (1999). Some consequences of temporal aggregation in empirical analysis. *Journal of Business & Economic Statistics*, 17(1):129–136.

Mariano, R. 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.

Nunes, L. (2005). Nowcasting quarterly GDP growth in a monthly coincident indicator model. *Journal of Forecasting*, 24(8):575–592.

Palm, F. C. and Nijman, T. E. (1984). Missing observations in the dynamic regression model. *Econometrica: journal of the Econometric Society*, pages 1415–1435.

Proietti, T. and Moauro, F. (2006). Dynamic factor analysis with non-linear temporal aggregation constraints. *Journal Of The Royal Statistical Society Series C*, 55(2):281–300.

Seong, B., Ahn, S., and Zadrozny, P. (2012). Estimation of vector error correction models with mixed-frequency data. *Journal of Time Series Analysis*.

Stock, J. and Watson, M. (1991). A probability model of the coincident economic indicators. *Leading Economic indicators: new approaches and forecasting records*, 66.