The logo of the University of Hawaii is a large, light green circular emblem in the background. It features a central torch with a flame, set against a stylized globe. Below the torch is the Hawaiian word 'MĀLAMALAMA'. The year '1907' is at the bottom center. The outer ring of the emblem contains the text 'UNIVERSITY OF HAWAII' at the top and 'U A MAU KE EA O KA 'ĀINA I KA PONO' at the bottom.

University of Hawai‘i at Mānoa
Department of Economics
Working Paper Series

Saunders Hall 542, 2424 Maile Way,
Honolulu, HI 96822 · Phone: (808) 956-8496
economics.hawaii.edu

Working Paper No. 11-10

Forecasting Based on Common Trends
in Mixed

by
Peter Fuleky and Carl S. Bonham

June 13, 2011

Forecasting Based on Common Trends in Mixed Frequency Samples

PETER FULEKY* and CARL S. BONHAM†

June 13, 2011

Abstract

We extend the existing literature on small mixed frequency single factor models by allowing for multiple factors, considering indicators in levels, and allowing for cointegration among the indicators. We capture the cointegrating relationships among the indicators by common factors modeled as stochastic trends. We show that the stationary single-factor model frequently used in the literature is misspecified if the data set contains common stochastic trends. We find that taking advantage of common stochastic trends improves forecasting performance over a stationary single-factor model. The common-trends factor model outperforms the stationary single-factor model at all analyzed forecast horizons on a root mean squared error basis. Our results suggest that when the constituent indicators are integrated and cointegrated, modeling common stochastic trends, as opposed to eliminating them, will improve forecasts.

JEL classifications: E37, C32, C53, L830

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

*Corresponding author—Post Doctoral Fellow, University of Hawaii Economic Research Organizaion, 540 Saunders Hall, 2424 Maile Way, Honolulu, HI 96822; e-mail: fuleky@hawaii.edu

†UHERO, and Department of Economics, University of Hawaii at Manoa; e-mail: bonham@hawaii.edu

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).

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 a *single* 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 (coincident index) may be a useful measure of the unobserved business cycle, it can also be used when estimating the unobserved value of low frequency indicators at the end of the sample. Furthermore, forecasts of the coincident index can be used when forecasting other indicators in the model.

Unfortunately, the forecasting performance of mixed frequency factor models is decidedly mixed. On the one hand, Camacho and Perez-Quiros (2010) show that forecasts from a single-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. On the other hand, Hyung and Granger (2008) find that GDP growth rate forecasts from their Linked-ARMA mixed-frequency model are less accurate than quarterly forecasts from a low-frequency model. Similarly, the GDP growth nowcasts of Evans (2005) show no significant improvement over advanced or preliminary GDP releases or the median Money Market Services forecast. All

of the studies mentioned above estimate single-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 some of the trends are common, the estimated model is misspecified, and forecasting performance may suffer.

We extend the existing literature on small mixed-frequency single-factor models by allowing for multiple factors, considering indicators in levels, and allowing for cointegration among the indicators. In general, we expect indicators to be driven by one or more common stochastic trends, the number of which determines the number of factors in the model. Consequently, rather than restrict our attention to the estimation of a single coincident index, we focus on the model's performance in nowcasting and forecasting the constituent indicators.

In the empirical section of the paper we compare forecasts from our common-trends factor model in levels with forecasts from a single-factor model in differences. The short horizon forecasts from both models are more precise than those from random walk models of individual indicators at their observation frequencies. However, the reduction in root mean squared forecast error for the common-trends factor model (CTFM) can be significantly greater than that for the stationary single-factor model (SSFM). Our results suggest that when the constituent indicators are integrated and cointegrated, modeling common stochastic trends, as opposed to eliminating them, will improve forecasts. Despite the potential benefits of modeling and forecasting cointegrated series using a common-trends mixed-frequency factor model, to our knowledge this paper is the first attempt at such an exercise.

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. We pay special attention to the misspecification of the stationary single-factor model when the indicators contain common trends. In the empirical application of Section 3, we contrast the forecasting performance of the common trends factor model with the stationary single-factor model and a naive random

walk model. Section 4 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 a $s \times 1$ vector of common factors, \mathbf{f}_t , that captures the co-movements of indicators. The second is a $n \times 1$ 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; for tractability we set the base frequency to daily. The $n \times 1$ vector of observed indicators, \mathbf{y}_t , is subject to missing values at the base frequency. The observed indicators are period aggregates of an $n \times 1$ vector of latent variables, $\mathring{\mathbf{y}}_t$, that can be expressed as a linear combination of the two mutually uncorrelated unobserved stochastic components \mathbf{f}_t and $\boldsymbol{\epsilon}_t$

$$\mathring{\mathbf{y}}_t = \mathring{\mathbf{L}}\mathbf{f}_t + \mathring{\mathbf{G}}\boldsymbol{\epsilon}_t, \quad (1)$$

where $\mathring{\mathbf{L}}$ is an $n \times s$ matrix of factor loadings and $\mathring{\mathbf{G}}$ is a diagonal $n \times n$ matrix containing the loading parameters of idiosyncratic shocks.¹ The problem is to estimate the parameters

¹In a multi-frequency model the dimensions of the loading matrices will depend on the number of different

and the unobserved processes from the fluctuations of the observed indicators, and then use the estimated model to produce forecasts.

Because the model is linear, it can be cast in state-space form, and we can take advantage of the Kalman filter to estimate the unobserved components. In general, the r^{th} factor f_t^r , is assumed to follow an $AR(p)$ process at the daily base frequency²

$$f_t^r = \phi_1^r f_{t-1}^r + \phi_2^r f_{t-2}^r + \dots + \phi_p^r f_{t-p}^r + e_t^r, \quad e_t^r \sim N(0, \sigma_f^r), \quad r = 1 \dots s, \quad (2)$$

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

$$\epsilon_{j,t} = \gamma_{j,1} \epsilon_{j,t-1} + \gamma_{j,2} \epsilon_{j,t-2} + \dots + \gamma_{j,m} \epsilon_{j,t-m} + e_{j,t}, \quad e_{j,t} \sim N(0, \sigma_{\epsilon,j}), \quad j = 1 \dots n. \quad (3)$$

By definition, the idiosyncratic shocks, $e_{j,t}$, are mutually uncorrelated across all n indicators.

Indicators of economic performance are generally recorded at a lower frequency implying missing data points at the base frequency. There are several ways to relate the latent high frequency variables, \mathbf{y}_t° , to the observed indicators, \mathbf{y}_t , which may include stock and flow variables. If the indicators are recorded as levels, flow types can be modeled as period sums, and stock types can be modeled either as snapshots in time or more commonly as period averages. Specifically, when the observed indicator is a flow type variable, it can be related to the accumulated value of the common factor and idiosyncratic component during the observation period. When the observed indicator is a stock type variable, it can be related to the average value of the common factor during the period. To deal with stock and flow variables as well as their temporal aggregation, one can aggregate an unobserved daily factor,

frequencies of the indicators in \mathbf{y}_t . These dimensions will be made precise in (20), the measurement equation of the state space model.

²Note that the $AR(p)$ process allows for the possibility that factors may follow a random walk.

f_t^r , and idiosyncratic component, $e_{j,t}$, into $\tilde{f}_{i,t}^r$ and $\tilde{\epsilon}_{j,t}$, respectively, according to

$$\tilde{f}_{i,t}^r = \psi_{i,t} \tilde{f}_{i,t-1}^r + \theta_i f_t^r, \quad i = 1 \dots k, \quad r = 1 \dots s, \quad (4)$$

$$\tilde{\epsilon}_{j,t} = \psi_{j,t} \tilde{\epsilon}_{j,t-1} + \theta_j \epsilon_{j,t}, \quad j = 1 \dots n, \quad (5)$$

where k is the number of frequencies in the model, ψ_t is a cumulator variable defined as

$$\psi_{i,t}, \psi_{j,t} = \begin{cases} 0 & \text{if } t \text{ is the first day of the period} \\ 1 & \text{otherwise,} \end{cases} \quad (6)$$

and

$$\theta_i = \begin{cases} 1 \\ 1/d_i \end{cases}, \quad \theta_j = \begin{cases} 1 & \text{for flow type variables} \\ 1/d_j & \text{for stock type variables,} \end{cases} \quad (7)$$

where d_i and d_j are the number of days per period for frequencies i and j respectively. We follow Nunes (2005) and Aruoba et al. (2009) and apply a similar aggregation method for indicators converted to growth rates. However, as we show in Section 2.2, in an unobserved component model with multiple frequencies the distinction between stocks and flows is not identified, and therefore we follow the stock type convention for all variables.

In the state space representation of the model, all unobserved components are collected in the state vector $\boldsymbol{\alpha}_t$. Its first lag

$$\boldsymbol{\alpha}_{t-1} = (f_{t-1}^1 \dots f_{t-p}^1, \tilde{f}_{1,t-1}^1 \dots \tilde{f}_{k,t-1}^1, \dots, f_{t-1}^s \dots f_{t-p}^s, \tilde{f}_{1,t-1}^s \dots \tilde{f}_{k,t-1}^s, \epsilon_{1,t-1} \dots \epsilon_{1,t-m}, \tilde{\epsilon}_{1,t-1}, \dots, \epsilon_{n,t-1} \dots \epsilon_{n,t-m}, \tilde{\epsilon}_{n,t-1})'. \quad (8)$$

contains p lags of the s factors, the factors aggregated to k frequencies, m lags of the n idiosyncratic shocks, and the value of each shock aggregated to the observation frequency of the corresponding indicator. The transition equation

$$\boldsymbol{\alpha}_t = \mathbf{T}_t \boldsymbol{\alpha}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim N(\mathbf{0}, \mathbf{Q}), \quad \boldsymbol{\alpha}_t \sim N(\mathbf{a}, \mathbf{P}), \quad t = 1 \dots T, \quad (9)$$

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

$$\mathbf{T}_t = \begin{bmatrix} \mathbf{K}_t^1 & \cdots & \mathbf{0}_{(p+k) \times (p+k)} & \mathbf{0}_{(p+k) \times (m+1)} & \cdots & \mathbf{0}_{(p+k) \times (m+1)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{(p+k) \times (p+k)} & \cdots & \mathbf{K}_t^s & \mathbf{0}_{(p+k) \times (m+1)} & \cdots & \mathbf{0}_{(p+k) \times (m+1)} \\ \mathbf{0}_{(m+1) \times (p+k)} & \cdots & \mathbf{0}_{(m+1) \times (p+k)} & \mathbf{M}_{1,t} & \cdots & \mathbf{0}_{(m+1) \times (m+1)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{(m+1) \times (p+k)} & \cdots & \mathbf{0}_{(m+1) \times (p+k)} & \mathbf{0}_{(m+1) \times (m+1)} & \cdots & \mathbf{M}_{n,t} \end{bmatrix}, \quad (10)$$

where

$$\mathbf{M}_{j,t} = \begin{bmatrix} \gamma_{j,1} \cdots \gamma_{j,m-1} & \gamma_{j,m} & 0 \\ \mathbf{I}_{(m-1)} & \mathbf{0}_{(m-1) \times 1} & \mathbf{0}_{(m-1) \times 1} \\ \theta_j \gamma_{j,1} \cdots \theta_j \gamma_{j,m-1} & \theta_j \gamma_{j,m} & \psi_{j,t} \end{bmatrix}, \quad j = 1 \dots n, \quad (11)$$

specifies the dynamics (rows 1 through m) and accumulation (last row) of the idiosyncratic component, and

$$\mathbf{K}_t^r = \begin{bmatrix} \phi_1^r \cdots \phi_{p-1}^r & \phi_p^r & \mathbf{0}_{1 \times k} \\ \mathbf{I}_{(p-1)} & \mathbf{0}_{(p-1) \times 1} & \mathbf{0}_{(p-1) \times k} \\ \Theta \Phi_1^r \cdots \Theta \Phi_{p-1}^r & \Theta \Phi_p^r & \Psi_t \end{bmatrix}, \quad r = 1 \dots s, \quad (12)$$

specifies the dynamics (rows 1 through p), and accumulation (last k rows) of the factors with

$$\Theta = \text{diag}(\theta_1 \dots \theta_k) \quad (13)$$

$$\Phi_h = \phi_h \mathbf{1}_{k \times 1}, \quad h = 1 \dots p \quad (14)$$

$$\Psi_t = \text{diag}(\psi_{1,t} \dots \psi_{k,t}). \quad (15)$$

The block diagonal covariance matrix of transition shocks, \mathbf{Q} takes the form

$$\mathbf{Q} = \begin{bmatrix} \mathbf{H}_f^1 & \cdots & \mathbf{0}_{(p+k) \times (p+k)} & \mathbf{0}_{(p+k) \times (m+1)} & \cdots & \mathbf{0}_{(p+k) \times (m+1)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{(p+k) \times (p+k)} & \cdots & \mathbf{H}_f^s & \mathbf{0}_{(p+k) \times (m+1)} & \cdots & \mathbf{0}_{(p+k) \times (m+1)} \\ \mathbf{0}_{(m+1) \times (p+k)} & \cdots & \mathbf{0}_{(m+1) \times (p+k)} & \mathbf{H}_{\epsilon,1} & \cdots & \mathbf{0}_{(m+1) \times (m+1)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{(m+1) \times (p+k)} & \cdots & \mathbf{0}_{(m+1) \times (p+k)} & \mathbf{0}_{(m+1) \times (m+1)} & \cdots & \mathbf{H}_{\epsilon,n} \end{bmatrix}, \quad (16)$$

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

$$\mathbf{H}_{\epsilon,j} = \sigma_{\epsilon,j}^2 \begin{bmatrix} 1 & \mathbf{0}_{1 \times (m-1)} & \theta_j \\ \mathbf{0}_{(m-1) \times 1} & \mathbf{0}_{(m-1) \times (m-1)} & \mathbf{0}_{(m-1) \times 1} \\ \theta_j & \mathbf{0}_{1 \times (m-1)} & \theta_j^2 \end{bmatrix}, \quad j = 1 \dots n. \quad (17)$$

and

$$\mathbf{H}_f^r = (\sigma_f^r)^2 \begin{bmatrix} 1 & \mathbf{0}_{1 \times (p-1)} & (\theta_1 \dots \theta_k) \\ \mathbf{0}_{(p-1) \times 1} & \mathbf{0}_{(p-1) \times (p-1)} & \mathbf{0}_{(p-1) \times k} \\ (\theta_1 \dots \theta_k)' & \mathbf{0}_{k \times (p-1)} & (\theta_1 \dots \theta_k)'(\theta_1 \dots \theta_k) \end{bmatrix}, \quad r = 1 \dots s, \quad (18)$$

The measurement equation relates the observed variables, \mathbf{y}_t , to the unobserved state vector

$$\mathbf{y}_t = \mathbf{Z}\boldsymbol{\alpha}_t, \quad (19)$$

where \mathbf{Z} is a sparse matrix containing the loading coefficients of the common and idiosyncratic components

$$\mathbf{Z} = \begin{bmatrix} \mathbf{L}_{n \times (p+k)}^1 & \cdots & \mathbf{L}_{n \times (p+k)}^s & \mathbf{G}_{n \times n(m+1)} \end{bmatrix}. \quad (20)$$

Because each variable is related to the factors accumulated to the variable's own frequency, each row of \mathbf{L}^r ; $r = 1 \dots s$ contains only one parameter: the $(j, p+i)$ element of \mathbf{L}^r represents the loading of the r^{th} factor accumulated to the i^{th} frequency onto the j^{th} observed indicator. The $(j, j(m+1))$ element of the matrix \mathbf{G} contains the coefficient of the idiosyncratic component corresponding to indicator $y_{j,t}$, and the rest of \mathbf{G} contains zeros.

2.2 Identification, Stocks and Flows

A multi-factor model is only identified up to a rotation of the factors. In a single-frequency multi-factor model, column r of the loading matrix corresponds to the loading of factor r on the observed variables, and identification is usually achieved by zeroing out the elements above the diagonal of the loading matrix and fixing the covariance matrix of the state shocks (see for example Bai and Ng (2010)). In a multi-frequency multi-factor model, a submatrix \mathbf{L}^r corresponds to the loading of factor r on the observed variables, and identification requires restrictions on the rows of the \mathbf{L}^r matrices and on the variance of the unobserved components. Specifically, identification can be achieved by setting the rows $j < r$ of \mathbf{L}^r to zero, and setting $\sigma_f^r = 1$ and $\sigma_{\epsilon,j} = 1$ for $r = 1 \dots s, j = 1 \dots n$.

In multi-frequency factor models, the scale of the aggregated unobserved components can be controlled by multiplying the corresponding row of the state equation by δ and multiplying the corresponding columns of the \mathbf{Z}_t matrix by $1/\delta$. Equation (7) implies that at the end of the period, a given variable accumulated as a flow is d times larger than the same variable accumulated as a stock. Because the loading parameters in the \mathbf{Z}_t matrix implicitly cancel out the effects of $\delta = \theta = 1/d$, the difference between summation and averaging of the unobserved components is unidentified. As a result, the distinction between stock and flow type variables is not identified for multi-frequency unobserved component models. Thus, it is the researcher's choice to let the accumulated unobserved components be flow or stock type variables. An advantage of dealing with the accumulated unobserved components as stocks, that is period averages, is that their end of period value will have the same variance

across all frequencies, and the loadings in the measurement equation will capture the relative importance of the unobserved components for each variable.

2.3 Estimation

We use the Kalman Filter's prediction error decomposition to evaluate the log-likelihood function and estimate the model parameters by maximum likelihood. The Kalman filter is an algorithm for sequentially updating a linear projection for the system. The initial state vector entering the recursive process is assumed to be random with mean $\mathbf{a}_{1|0} = \mathbf{0}$ and diffuse variance matrix $\mathbf{P}_{1|0}$. Given the starting values $\mathbf{a}_{t|t-1}$ and $\mathbf{P}_{t|t-1}$, the prediction error, \mathbf{v}_t , and the mean squared error (MSE), \mathbf{F}_t , are calculated by

$$\mathbf{v}_t = \mathbf{y}_t - \mathbf{Z}_t \mathbf{a}_{t|t-1} \quad \text{and} \quad \mathbf{F}_t = \mathbf{Z}_t \mathbf{P}_{t|t-1} \mathbf{Z}'_t, \quad (21)$$

where $\mathbf{a}_{t|t-1} = E(\boldsymbol{\alpha}_t | \mathbf{Y}_{t-1})$ with $\mathbf{Y}_{t-1} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{t-1})$ and $\mathbf{P}_t = E[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1})']$. Next, the inference about the current value of $\boldsymbol{\alpha}_t$ is updated based on the observation of \mathbf{y}_t to produce

$$\mathbf{a}_{t|t} = \mathbf{a}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{Z}'_t \mathbf{F}_t^{-1} \mathbf{v}_t, \quad \text{with MSE} \quad \mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} \mathbf{Z}'_t \mathbf{F}_t^{-1} \mathbf{Z}_t \mathbf{P}_{t|t-1}, \quad (22)$$

where $\mathbf{a}_{t|t} = E(\boldsymbol{\alpha}_t | \mathbf{Y}_t)$ and $\mathbf{P}_{t|t} = E[(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t})(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t})']$. Finally, the state equation is used to forecast the state at time $t + 1$

$$\mathbf{a}_{t+1|t} = \mathbf{T} \mathbf{a}_{t|t}, \quad \text{with MSE} \quad \mathbf{P}_{t+1|t} = \mathbf{T} \mathbf{P}_{t|t} \mathbf{T}' + \mathbf{Q}, \quad (23)$$

which in turn are used to find \mathbf{v}_{t+1} and \mathbf{F}_{t+1} . Once the algorithm iterates through the complete time series, the value of the log-likelihood can be calculated from

$$\log L_t = -\frac{\tau N}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^{\tau} (\log |\mathbf{F}_t| + \mathbf{v}'_t \mathbf{F}_t^{-1} \mathbf{v}_t), \quad (24)$$

for given parameter values. Then, the log-likelihood function can be maximized with respect to the unknown parameters of the system.

If some of the indicators are not observed on a given day t , the vector \mathbf{y}_t collapses to \mathbf{y}_t^* with $n^* < n$, and the rows corresponding to the missing \mathbf{y}_t values are eliminated from the measurement equation

$$\mathbf{y}_t^* = \mathbf{Z}_t^* \boldsymbol{\alpha}_t . \quad (25)$$

When $n^* = 0$, that is when the full \mathbf{y}_t vector is missing, the algorithm sets $\mathbf{v}_t = \mathbf{0}$ and $\mathbf{F}_t^{-1} = \mathbf{0}$, and the Kalman filter reduces to the prediction step

$$\mathbf{a}_{t+1|t} = \mathbf{T}\mathbf{a}_{t|t} , \quad \text{with MSE} \quad \mathbf{P}_{t+1|t} = \mathbf{T}\mathbf{P}_{t|t}\mathbf{T}' + \mathbf{Q} . \quad (26)$$

This feature of the Kalman filter is convenient at the end of the sample where it can be used to produce out-of-sample predictions of the state variables. Then, one can obtain forecasts of the indicators by plugging the predicted state into the measurement equation. The Kalman smoother (Harvey, 1989, p. 149) can be used to find in-sample estimates of the state variables, which can then be used to backcast any of the as yet unobserved indicators. By iterating the Kalman filter from the end of the sample (time T) forward, we obtain the h -step-ahead forecasts \mathbf{y}_{T+h} of the variables entering the model.

2.4 Levels vs. Differences

Economic time series are often characterized as unit root processes, and this gives rise to two 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. 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 $n - s$ rows of an $(n - s) \times n$ matrix \mathbf{A} which has the property $\mathbf{A}\mathring{\mathbf{L}} = \mathbf{0}$, so that premultiplying (1) with \mathbf{A} gives

$$\mathbf{A}\mathring{\mathbf{y}}_t = \mathbf{A}\mathring{\mathbf{G}}\epsilon_t, \quad (27)$$

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

Conversion to differences eliminates the need to estimate multiple factors and focuses attention on the single factor as a measure of current business conditions. However, besides extracting a current conditions index, most of the literature also attempts to illustrate an improvement in forecasting performance through the use of mixed-frequency data in stationary single-factor models (SSF_M).³ But modeling differences when the indicators in levels contain common stochastic trends may leave the model misspecified (see section 2.5) and lead to poor forecasts. In addition, the conversion to differences discards the relationship between variable levels, 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 already poor 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 let the factors capture the common stochastic trends (Christoffersen and Diebold, 1998).

³Predicted values of the indicators are derived by reversing the differencing transformation: the predicted differences are integrated from the last observation forward to obtain forecasts of the level of indicators.

2.5 Misspecification of SSFM for Data with Common Stochastic Trends

Below we illustrate the misspecification of stationary single factor models if the data contains multiple 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 observed low frequency variables, \mathbf{y}_t , are aggregates of their latent high frequency counterparts, $\mathring{\mathbf{y}}_t$, and the true data generating process (DGP) is given by the common trends factor model (CTFM) at the base frequency,

$$\begin{aligned} \mathring{\mathbf{y}}_t &= \mathring{\mathbf{L}}\mathbf{f}_t + \mathring{\mathbf{G}}\boldsymbol{\epsilon}_t, \\ \text{where } \mathbf{f}_t &= \mathbf{f}_{t-1} + \boldsymbol{\eta}_{f,t}, \quad \boldsymbol{\eta}_{f,t} \sim N(\mathbf{0}, \mathbf{Q}_f) \\ \text{and } \boldsymbol{\epsilon}_t &= \boldsymbol{\eta}_{\epsilon,t}, \quad \boldsymbol{\eta}_{\epsilon,t} \sim N(\mathbf{0}, \mathbf{Q}_\epsilon). \end{aligned} \tag{28}$$

Transforming the DGP by first differencing gives us

$$\begin{aligned} \Delta\mathring{\mathbf{y}}_t &= \mathring{\mathbf{L}}\Delta\mathbf{f}_t + \mathring{\mathbf{G}}\Delta\boldsymbol{\epsilon}_t \\ \text{with } \Delta\mathbf{f}_t &= \boldsymbol{\eta}_{f,t}, \\ \text{and } \Delta\boldsymbol{\epsilon}_t &= \boldsymbol{\eta}_{\epsilon,t} - \boldsymbol{\eta}_{\epsilon,t-1}. \end{aligned} \tag{29}$$

Note, the implied model in equation (29) is different from the SSFM commonly used in the literature,

$$\begin{aligned} \Delta\mathring{\mathbf{y}}_t &= \mathbf{L}^*f_t^* + \mathbf{G}^*\boldsymbol{\epsilon}_t^* \\ \text{with } f_t^* &= T_f^*f_{t-1}^* + \eta_{f,t}^* \\ \text{and } \boldsymbol{\epsilon}_t^* &= \boldsymbol{\eta}_{\epsilon,t}^*. \end{aligned} \tag{30}$$

The stationary model in (29) requires the same number of factors as present in the DGP, and precludes modeling the factor as a dynamic process. In addition, in the first differenced DGP (29), the idiosyncratic component follows an MA(1) process with a unit root. In contrast, the SSFM in equation (30) 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 (29) and (30) are equivalent when $T_f^* = 0$, $\boldsymbol{\eta}_{\epsilon,t}^* = \boldsymbol{\eta}_{\epsilon,t} - \boldsymbol{\eta}_{\epsilon,t-1}$, and the DGP contains a single common trend, but if the DGP contains multiple trends the two models cannot be reconciled.

To explore the relationship between the SSFM and the CTFM further, integrate the differences, $\Delta \mathring{\boldsymbol{y}}_t$, in equation (30),

$$\begin{aligned}
\mathring{\boldsymbol{y}}_t &= \mathring{\boldsymbol{y}}_1 + \sum_{\tau=2}^t \Delta \mathring{\boldsymbol{y}}_\tau \\
&= \mathring{\boldsymbol{y}}_1 + \boldsymbol{L}^* \sum_{\tau=2}^t \boldsymbol{f}_\tau^* + \boldsymbol{G}^* \sum_{\tau=2}^t \boldsymbol{\epsilon}_\tau^* \\
&= \mathring{\boldsymbol{y}}_1 + \boldsymbol{L}^* \check{\boldsymbol{f}}_t^* + \boldsymbol{G}^* \check{\boldsymbol{\epsilon}}_t^* .
\end{aligned} \tag{31}$$

where $\check{\boldsymbol{f}}_t^*$ is a *common* stochastic trend and $\check{\boldsymbol{\epsilon}}_t^*$ is a vector of *idiosyncratic* stochastic trends.⁴ The integrated SSFM, equation (31), differs from the CTFM (DGP) in several respects. In the former, the stochastic trends are not random walks because their increments follow an AR(1) process. In the latter the number of common trends is determined by the cointegrating rank of the system, and the idiosyncratic components are assumed to be stationary.

Now consider a multi-frequency model, in which the low frequency variables \boldsymbol{y}_t are defined as period aggregates of their latent high frequency counterparts, $\mathring{\boldsymbol{y}}_t$. Specifically, write the

⁴The presence of idiosyncratic stochastic trends in (31) implies that the level of the extracted coincident index will arbitrarily diverge from the level of the indicators, and therefore the former will not be a good estimator of the latter.

value of the j^{th} element of $\mathring{\mathbf{y}}_t$ in (28) aggregated to the observation frequency as

$$y_{j,t} = \theta_j \sum_{\delta=1}^{d_j} \mathring{y}_{j,t-d_j+\delta} = \mathring{\mathbf{L}}_j \left(\theta_j \sum_{\delta=1}^{d_j} \mathbf{f}_{t-d_j+\delta} \right) + \mathring{G}_{jj} \left(\theta_j \sum_{\delta=1}^{d_j} \epsilon_{j,t-d_j+\delta} \right) = \mathring{\mathbf{L}}_j \tilde{\mathbf{f}}_t + \mathring{G}_{jj} \tilde{\epsilon}_{j,t} , \quad (32)$$

where $\mathring{\mathbf{L}}_j$ is the j^{th} row of $\mathring{\mathbf{L}}$, and \mathring{G}_{jj} is the j^{th} element on the diagonal of $\mathring{\mathbf{G}}$. Differencing at the observation frequency gives us

$$\begin{aligned} \Delta_{d_j} y_{j,t} &= y_{j,t} - y_{j,t-d_j} = \theta_j \sum_{\delta=1}^{d_j} \Delta_{d_j} \mathring{y}_{j,t-d_j+\delta} \\ &= \mathring{\mathbf{L}}_j \left(\theta_j \sum_{\delta=1}^{d_j} \Delta_{d_j} \mathbf{f}_{t-d_j+\delta} \right) + \mathring{G}_{jj} \left(\theta_j \sum_{\delta=1}^{d_j} \Delta_{d_j} \epsilon_{j,t-d_j+\delta} \right) \\ &= \mathring{\mathbf{L}}_j \Delta_{d_j} \tilde{\mathbf{f}}_t + \mathring{G}_{jj} \Delta_{d_j} \tilde{\epsilon}_{j,t} , \end{aligned} \quad (33)$$

where Δ_{d_j} denotes the change during the observation period. Note, the change during the observation period can be written as the sum of base period differences $\Delta_{d_j} \mathring{y}_{j,t} = \sum_{\iota=1}^{d_j} \Delta \mathring{y}_{j,t-d_j+\iota}$, so that

$$\begin{aligned} \Delta_{d_j} y_{j,t} &= y_{j,t} - y_{j,t-d_j} = \theta_j \sum_{\delta=1}^{d_j} \sum_{\iota=1}^{d_j} \Delta \mathring{y}_{j,t-2d_j+\delta+\iota} \\ &= \mathring{\mathbf{L}}_j \left(\theta_j \sum_{\delta=1}^{d_j} \sum_{\iota=1}^{d_j} \Delta \mathbf{f}_{t-2d_j+\delta+\iota} \right) + \mathring{G}_{jj} \left(\theta_j \sum_{\delta=1}^{d_j} \sum_{\iota=1}^{d_j} \Delta \epsilon_{j,t-2d_j+\delta+\iota} \right) \\ &= \mathring{\mathbf{L}}_j \sum_{\iota=1}^{d_j} \Delta \tilde{\mathbf{f}}_{t-d_j+\iota} + \mathring{G}_{jj} \sum_{\iota=1}^{d_j} \Delta \tilde{\epsilon}_{j,t-d_j+\iota} . \end{aligned} \quad (34)$$

This expression can be simplified by expanding the double summation into

$$\Delta_{d_j} y_{j,t} = \theta_j (\Delta \mathring{y}_{j,t-2d_j+2} + 2\Delta \mathring{y}_{j,t-2d_j+3} + \cdots + d_j \Delta \mathring{y}_{j,t-d_j+1} + \cdots + 2\Delta \mathring{y}_{j,t-1} + \Delta \mathring{y}_{j,t}) , \quad (35)$$

but (35) is still cumbersome to implement if the base frequency is daily and the dataset contains quarterly observations, that is, when $d_j = 90$.⁵ In fact, Evans (2005) and Aruoba et al.

⁵Mariano and Murasawa (2003), Hyung and Granger (2008) and others used (35) in mixed frequency

(2009), who used a daily base frequency in their studies, approximated (35) by considering the daily contribution of a latent variable $\Delta\dot{y}_{j,t}$ to the observed one $\Delta_{d_j}y_{j,t}$

$$\Delta_{d_j}y_{j,t} = \theta_j(\Delta\dot{y}_{j,t-d_j+1} + \cdots + \Delta\dot{y}_{j,t-1} + \Delta\dot{y}_{j,t}) = \theta_j \sum_{\delta=1}^{d_j} \Delta\dot{y}_{j,t-d_j+\delta} , \quad (36)$$

where $\Delta\dot{y}_{j,t}$ is the j^{th} row of $\Delta\dot{\mathbf{y}}_t$ which can be defined similarly to (30)

$$\Delta\dot{\mathbf{y}}_t = \dot{\mathbf{L}}^* \dot{f}_t^* + \dot{\mathbf{G}}^* \dot{\epsilon}_t^* \quad \text{with} \quad \dot{f}_t^* = \dot{T}_f^* \dot{f}_{t-1}^* + \dot{\eta}_{f,t}^* \quad \text{and} \quad \dot{\epsilon}_t^* = \dot{\eta}_{\epsilon,t}^* . \quad (37)$$

Note that while (35) spans two observation periods, (36) only spans one, and therefore the approximation directly affects the dynamics of the model.

To summarize, under the DGP given by the CTFM, the SSFM described by (36) and (37) contains an incorrect number of common stochastic trends, introduces idiosyncratic stochastic trends, and has mis-specified dynamics both at the base frequency and the observation frequency. These forms of misspecification will affect the decomposition of the observed data into common and idiosyncratic components and the forecasting performance of the model.

3 Empirical Application

To illustrate the forecast performance of our CTFM relative to the more common SSFM, we make use of mixed frequency data on Hawaii's tourism industry. Table 1 summarizes the sampling frequencies, reporting lags and sample periods of indicators in our study. ⁶

The indicator series are lined up against time so that observations fall on the last day of the representative period: weekly values are matched with the Sunday of the week for which

models of monthly and quarterly data, that is, with $d_j = 3$.

⁶Weekly passenger count is not simply a high-frequency version of visitor arrivals. The daily passenger count series is based on crew reports on domestic flights and VIP reports for international flights. No daily counts are available for flights from Canada or from charter flights. And, daily passengers include residents returning to Hawai'i. On a monthly basis, all airlines report exact passenger counts, and survey results and customs reports are used to subtract the number of returning residents to obtain monthly visitor arrivals numbers.

Table 1: Tourism Indicators

	Sampling Freq.	Reporting Lag	Start	End
Airline passenger counts, <i>PC</i>	weekly	1 day	10/7/00	8/28/10
Total Visitor arrivals, <i>VIS</i>	monthly	1 month	1/31/00	7/31/10
Real Visitor Expenditures, <i>VEXPR</i>	quarterly	4 months	3/31/01	6/30/10

the value is measured, monthly and quarterly values are matched with the last day of the month and quarter, respectively, for which the value is measured. Figure 1 displays the levels and differences of the series. The plots of differences in the right hand panel of Figure 1 indicate that removing the trend from the indicators results in noisy series. This outcome

Table 2: AR(1) estimates

$\Delta \ln y_t = \beta \Delta \ln y_{t-1} + \epsilon_t$	<i>PC</i>	<i>VIS</i>	<i>VEXPR</i>
estimate	-0.291	-0.095	0.212
p-value	0.000	0.132	0.183

Note: Autoregressive coefficient estimates and their marginal significance level in a first order autoregressive model of differences at the observation frequency.

is confirmed by the negative and insignificant first order autocorrelation of the differences presented in Table 2. Table 3 reports results from augmented Dickey-Fuller (1979) (ADF)

Table 3: Augmented Dickey-Fuller tests

$\Delta y_t = \alpha + \beta y_{t-1} + \sum_{k=1}^m \theta_k \Delta y_{t-k} + \epsilon_t$			
$H_0 : \beta = 0$			
Series	$1 + \hat{\beta}$	ADF t-test	p-value
<i>PC</i>	0.906	-2.437	0.132
<i>VIS</i>	0.890	-2.662	0.084
<i>VEXPR</i>	0.909	-2.086	0.251

Note: Column 1 lists the series tested for a unit root, column 2 presents the estimated AR1 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.

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 5% significance level. To test for

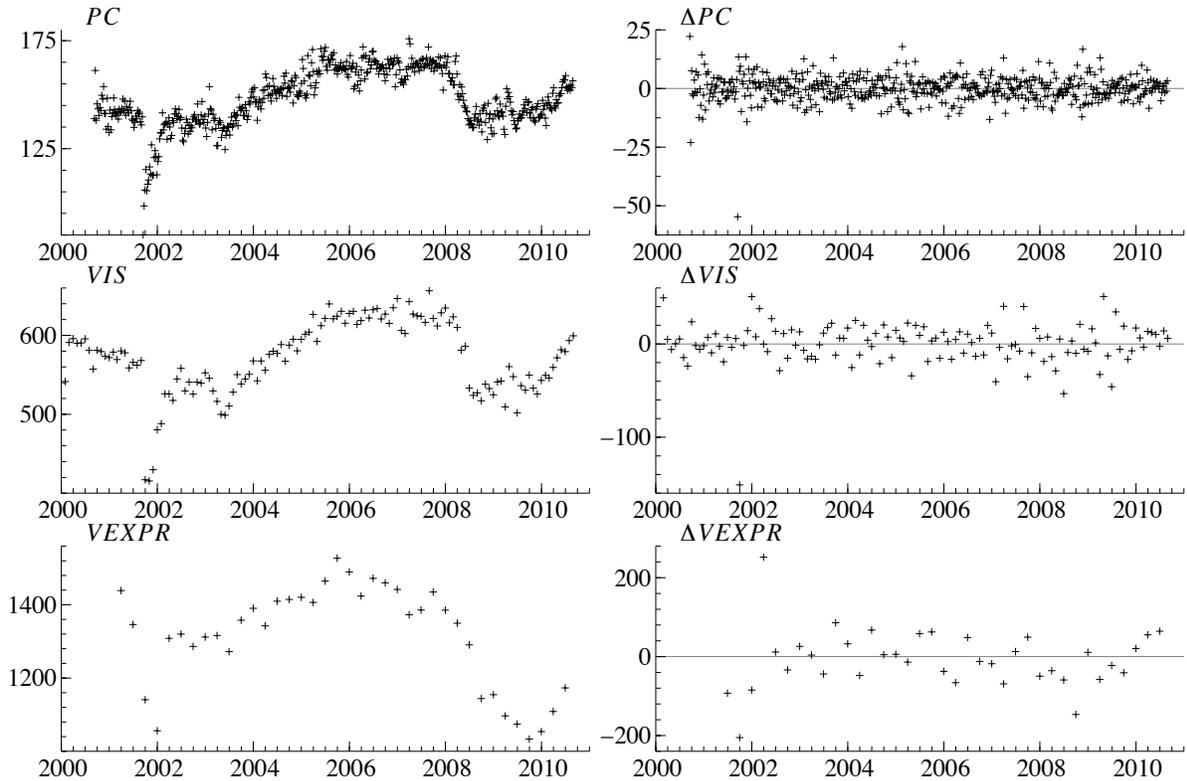


Figure 1: Time series plots of levels and differences of indicators.

cointegration among the indicators, we apply Johansen’s (1988) rank test to a temporally aggregated quarterly system. Because of the very short sample (37 quarterly values), our test likely suffers from low power.⁷ 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 4 indicate that we can reject the null hypothesis of no cointegrating vectors at the 7% significance level based on the Trace test. We are unable to reject the hypothesis of one or fewer cointegrating vectors at the 10% significance level on the basis of either test. 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

⁷While the 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 decline (see Marcellino (1999)).

Table 4: Cointegration rank tests

Estimation period: 2001:1 - 2010:2 (T = 37).					
Rank	Eigenvalue	Trace test	p-value	Lmax test	p-value
$r = 0$	0.410	33.792	0.069	19.513	0.119
$r \leq 1$	0.267	14.279	0.277	11.467	0.227
$r \leq 2$	0.073	2.812	0.623	2.812	0.622

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 Lmax test.

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.

To reduce the number of estimated parameters, we standardize all variables. In addition, we benchmark the shorter non-stationary series to the longest one: we set the mean and variance of the shorter series equal to the mean and variance of a subsample of the longest series, where the subsample is the period over which the long and the short series overlap. This benchmarking is equivalent to setting the mean distance between cointegrated variables to zero. Predicted values of the indicators are obtained by reversing the standardization and the benchmarking of the forecasts produced by the model.

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. Specifically, we contrast the CTFM (32), with the SSFM (37), both described in Section 2.5. 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.

3.1 Estimation Results

The estimation results for the CTFM with two stochastic trends are displayed in Table 5.⁸

Both trends have significant loading coefficients, $\lambda_{j,i}^r$, for each indicator. The loading of the

Table 5: Estimation Results: Common-Trends Factor Model

	$\lambda_{PC,W}^1$	$\lambda_{VIS,M}^1$	$\lambda_{VEXP,Q}^1$	$\lambda_{VIS,M}^2$	$\lambda_{VEXP,Q}^2$
estimate	0.076	0.079	0.057	0.004	0.026
p-value	0.000	0.000	0.000	0.002	0.000

	g_{PC}	g_{VIS}	g_{VEXP}
estimate	0.777	1.283	0.341
p-value	0.000	0.000	0.571

Note: Estimate is the maximum likelihood parameter estimate for CTFM with two common trends; p-value is the marginal significance level for each parameter. The value of $\lambda_{PC,W}^2$ is fixed at zero to satisfy identification requirements (see Section 2.2).

quarterly *VEXP* idiosyncratic component is insignificant, implying that the *VEXP* series is mostly determined by the combination of the two stochastic trends.

Estimation results for a SSFM are displayed in Table 6. The noise in the differenced

Table 6: Estimation Results: Stationary Single-Factor Model

	ϕ	$\lambda_{PC,W}$	$\lambda_{VIS,M}$	$\lambda_{VEXP,Q}$
estimate	-0.809	3.352	11.425	13.905
p-value	0.000	0.000	0.000	0.000

	g_{PC}	g_{VIS}	g_{VEXP}
estimate	0.830	0.984	7.248
p-value	0.004	0.755	0.000

Note: Estimate is the maximum likelihood parameter estimate for the SSFM; p-value is the marginal significance level for each parameter.

series causes the common component to be negatively autocorrelated at the daily frequency. The value of ϕ implies that the half-life of a shock is about three days and its impact decays to 23% of its original value within a week.

Figure 2 displays the decomposition of the standardized levels into the two common stochastic trends and the idiosyncratic components, and Figure 3 displays the decomposition

⁸We estimated the model by Ox 5.10 (Doornik, 2007) and SsfPack 2.2 (Koopman et al., 1999).

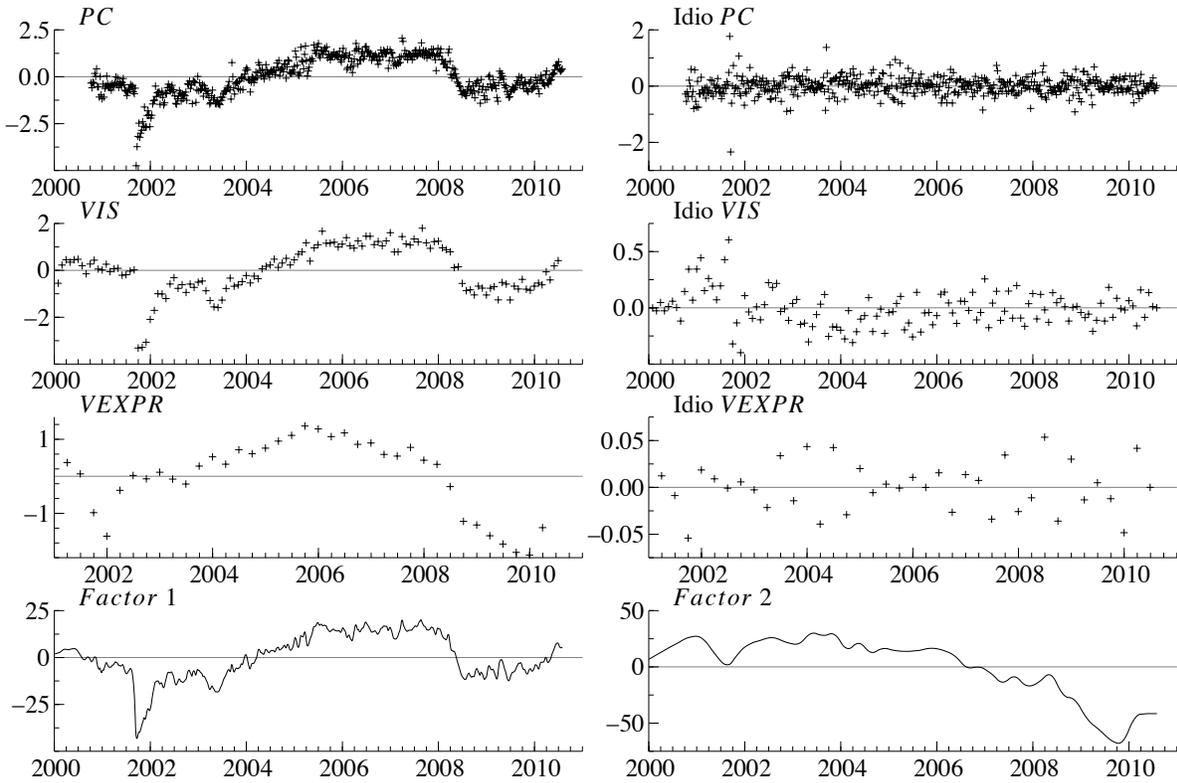


Figure 2: Standardized levels, and their decomposition into common stochastic trends and stationary idiosyncratic components.

of the standardized differences into a common factor and idiosyncratic components. While forecasts from the CTFM will be determined predominantly by the values of two factors, forecasts of the indicators from the SSFM will be determined by the integrated value of a single factor (“Coincident Index”) and the integrated values of the idiosyncratic components.

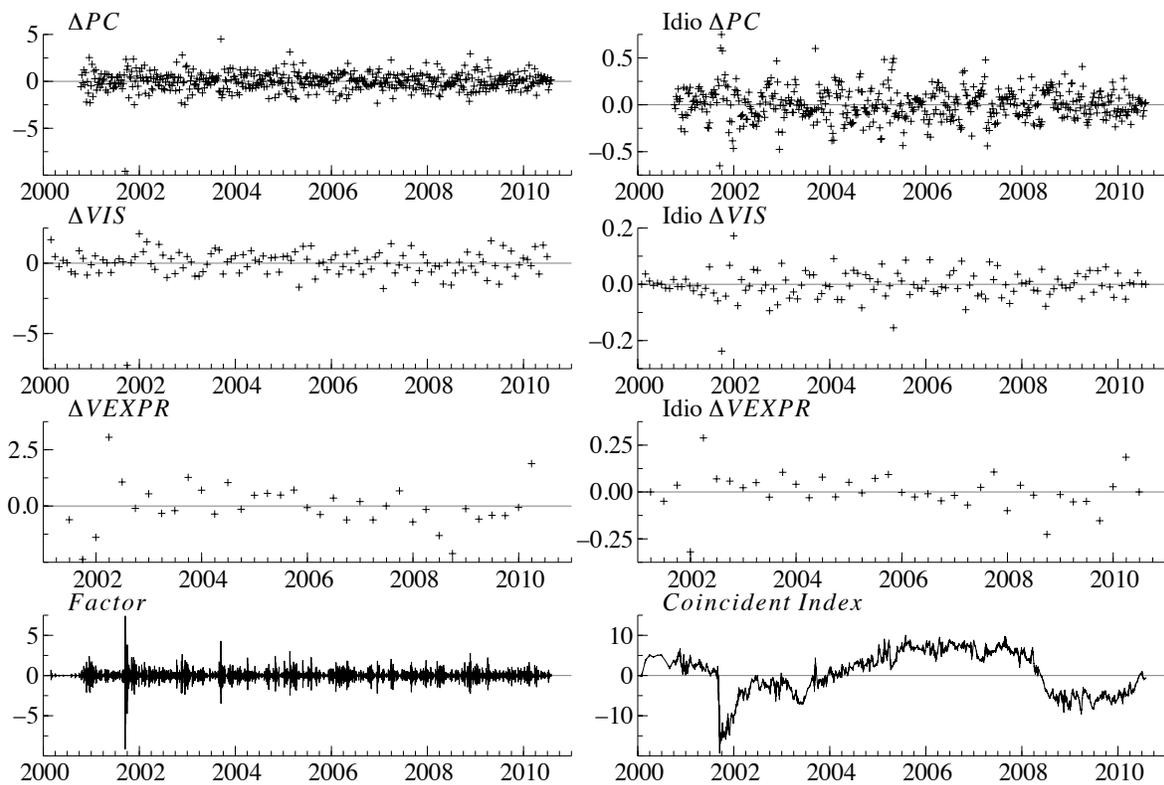


Figure 3: Standardized differences, and their decomposition into a stationary common factor and idiosyncratic components. The integrated common factor can be interpreted as a business conditions index.

3.2 Forecasting Results

The quality of forecasts depends in part on what information is available at the time the forecast is made. To evaluate the forecasting performance of our mixed-frequency factor model, we carefully replicate the historical flow of information.⁹ The model is estimated, and a forecast is made on the last day of each month between January 2005 and August 2010. At each forecast date, T , the latest available weekly observation is for the week ending on the last Sunday of the given month (the Sunday falling between T and $T - 6$). The last recorded monthly observation is for the previous month (with a time stamp of about $T - 30$).¹⁰ The quarterly indicator is only available four months after the end of the quarter. That is, on the last day of the first month of every quarter $VEXPR$ becomes available for the quarter ended at $T - 120$. Figure 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 between $T - 30$ and T to improve estimates of VIS , and the flow of information between $T - 120$ and T to improve estimates of $VEXPR$ series.

Tables 7 and 8 compare the accuracy of predictions from the two mixed frequency factor models with those from a benchmark random walk model for the two low frequency indicators at their observation frequency. We report the root mean squared error (RMSE) of forecasts, the percentage difference in RMSE across models, and the marginal significance level of Diebold and Mariano (1995) tests for forecast accuracy.

The nowcasts of visitor arrivals (VIS) produced by the mixed-frequency models benefit from the availability of weekly passenger counts (PC): the CTFM and SSFM show significant reductions in RMSE of 38% and 21%, respectively, over the RW model. While the CTFM

⁹Note that we do not construct a real-time data set as in Giannone et al. (2008). Data on passenger counts is not revised, while data on total visitor arrivals and spending are revised annually. Unfortunately, the preliminary data on arrivals and spending is not available for the purpose of constructing a real-time data set.

¹⁰For simplicity, our notation for indexing time horizons assumes 30 days per month.

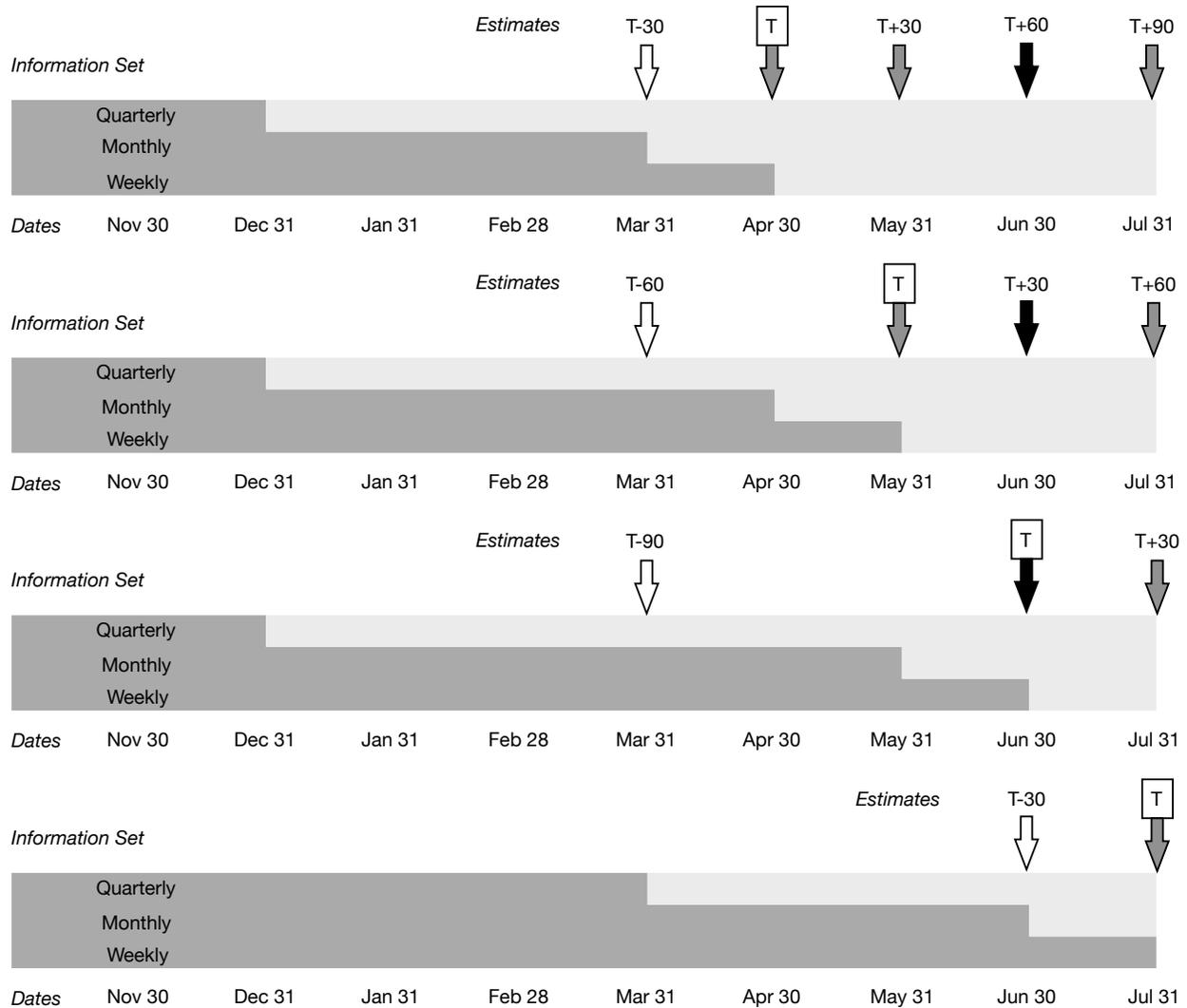


Figure 4: Illustration of the horizon for indicator estimates (backcasts, nowcasts, forecasts) relative to the forecast date T and the expansion of the information set. Negative horizons represent backcasts. VIS estimates with horizon up to 30 days are nowcasts, and with horizon 30 days and higher are forecasts. $VEXP$ estimates with horizon up to 90 days are nowcasts, and with horizon 90 days and higher are forecasts. The white arrows indicate target dates of quarterly $VEXPR$ backcasts. The grey arrows indicate target dates of monthly VIS nowcasts and forecasts. The black arrows indicate target dates of both, monthly VIS and quarterly $VEXPR$ nowcasts and forecasts.

retains its advantage over the RW model for all considered forecast horizons where intra-period information is not available, the forecasting precision of the SSFM declines faster, and becomes worse than that of the RW model at horizon $(T) + 210$.

The forecasts of real visitor expenditures ($VEXPR$) produced by both mixed frequency

Table 7: *VIS* Forecasting Results: Comparison of CTFM, SSFM and RW Model

Horizon	RMSE			% Difference			DM p-value		
	CT	SS	RW	CT/SS	CT/RW	SS/RW	CT/SS	CT/RW	SS/RW
(<i>T</i>) + 300	41.7	44.9	42.6	-7.3	-2.1	5.5	0.01	0.56	0.09
(<i>T</i>) + 270	40.0	42.9	41.0	-6.7	-2.5	4.5	0.05	0.56	0.19
(<i>T</i>) + 240	36.3	39.2	38.3	-7.3	-5.3	2.2	0.01	0.20	0.54
(<i>T</i>) + 210	35.5	37.3	36.3	-4.7	-2.2	2.6	0.10	0.60	0.44
(<i>T</i>) + 180	32.8	34.5	34.8	-5.0	-5.8	-0.9	0.11	0.14	0.81
(<i>T</i>) + 150	28.2	28.9	30.6	-2.5	-8.1	-5.7	0.48	0.16	0.22
(<i>T</i>) + 120	25.2	26.2	27.4	-4.1	-8.4	-4.4	0.34	0.15	0.40
(<i>T</i>) + 90	22.4	23.0	25.1	-2.8	-10.6	-8.0	0.55	0.14	0.15
(<i>T</i>) + 60	20.3	21.2	21.3	-4.1	-4.7	-0.7	0.50	0.67	0.93
(<i>T</i>) + 30	17.1	17.2	19.6	-0.9	-12.9	-12.1	0.88	0.17	0.12
(<i>T</i>) + 0	10.5	13.5	17.0	-22.3	-38.6	-21.0	0.01	0.00	0.01

Note: RMSE of *VIS* forecasts. The forecast horizon is measured relative to the forecast date, *T*. The CTFM and SSFM are compared to each other and a quarterly RW model of *VIS* via the percentage difference between the RMSE of forecasts from each model and the marginal significance of Diebold and Mariano (1995) test.

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

Horizon	RMSE			% Difference			DM p-value		
	CT	SS	RW	CT/SS	CT/RW	SS/RW	CT/SS	CT/RW	SS/RW
(<i>T</i>) + 300	117.8	134.1	147.2	-12.2	-20.0	-8.9	0.16	0.10	0.10
(<i>T</i>) + 270	111.1	132.0	144.7	-15.8	-23.2	-8.8	0.05	0.07	0.17
(<i>T</i>) + 240	109.2	116.1	124.9	-5.9	-12.6	-7.1	0.51	0.33	0.21
(<i>T</i>) + 210	94.2	110.0	124.9	-14.4	-24.6	-11.9	0.16	0.07	0.05
(<i>T</i>) + 180	84.3	107.0	122.4	-21.2	-31.1	-12.6	0.02	0.03	0.07
(<i>T</i>) + 150	82.5	91.6	99.5	-9.9	-17.1	-8.0	0.38	0.28	0.31
(<i>T</i>) + 120	69.1	85.7	99.5	-19.4	-30.6	-13.8	0.12	0.07	0.10
(<i>T</i>) + 90	63.7	84.6	97.5	-24.7	-34.6	-13.2	0.04	0.07	0.18
(<i>T</i>) + 60	59.1	70.3	74.9	-16.0	-21.1	-6.2	0.27	0.32	0.55
(<i>T</i>) + 30	51.0	67.1	74.9	-24.1	-32.0	-10.4	0.12	0.12	0.27
(<i>T</i>) + 0	43.3	64.3	73.3	-32.7	-40.9	-12.3	0.03	0.06	0.26
(<i>T</i>) - 30	32.4	41.8	44.6	-22.5	-27.3	-6.3	0.14	0.12	0.54
(<i>T</i>) - 60	31.8	42.9	45.4	-25.8	-29.9	-5.5	0.09	0.10	0.60
(<i>T</i>) - 90	29.8	42.4	44.5	-29.8	-33.1	-4.7	0.07	0.07	0.65

Note: RMSE of *VEXP* forecasts. The forecast horizon is measured relative to the forecast date, *T*. The CTFM and SSFM are compared to each other and a quarterly RW model of *VEXP* via the percentage difference between the RMSE of forecasts from each model and the marginal significance of Diebold and Mariano (1995) test.

factor models are more precise than those produced by the random walk model for all horizons. But the improvement achieved by the CTFM is significantly greater than that of the SSFM at virtually all horizons. The RMSE reduction over the RW model ranges from 12% to 40% for the CTFM, and from 4% to 13% for the SSFM. Apparently, the use of two common stochastic trends captures the evolution of the indicators more effectively than a single stationary factor. The transformation in the SSFM removes all long-run trend information leaving the model with a set of relatively noisy variables to analyze. When the differences are re-integrated to obtain levels forecasts, there is only a single common factor underlying the forecasts. Yet our rank tests and the analysis of the cointegrated system tell us that there are two common stochastic trends present in the three variables, implying that the SSFM is misspecified. Forecasts from the CTFM that make use of a linear combination of the two common stochastic trends have a lower RMSE than forecasts that discard this additional information. Specifically, the CTFM incorporates the intra-period information differently and propagates it to the forecasts more accurately than is possible in the misspecified SSFM.

Figure 5 displays the RMSE for *VEXPR* across forecast horizons. The quarterly RW model of *VEXPR* does not benefit from the high frequency information that becomes available within a quarter. Therefore the RMSE from this model is constant *between* *VEXPR* release dates but exhibits a jump *on* *VEXPR* release dates. The mixed frequency models are re-estimated for each expansion of the information set. While the updated parameter values may affect the forecasts, we expect the main benefit of mixed frequency models to come from the incorporation of intra-quarter high frequency observations into the end-of-quarter *VEXPR* estimates. However, there is no intra-quarter information released during the three months preceding a *VEXPR* target between $T - 30$ and $T - 90$.¹¹ Therefore, for these horizons the precision of the mixed frequency models is only influenced by updated parameter values and remains relatively stable. In contrast, the mixed frequency *VEXPR*

¹¹For example, notice that the information set prior to the March 31 target date is constant in Figure 4.

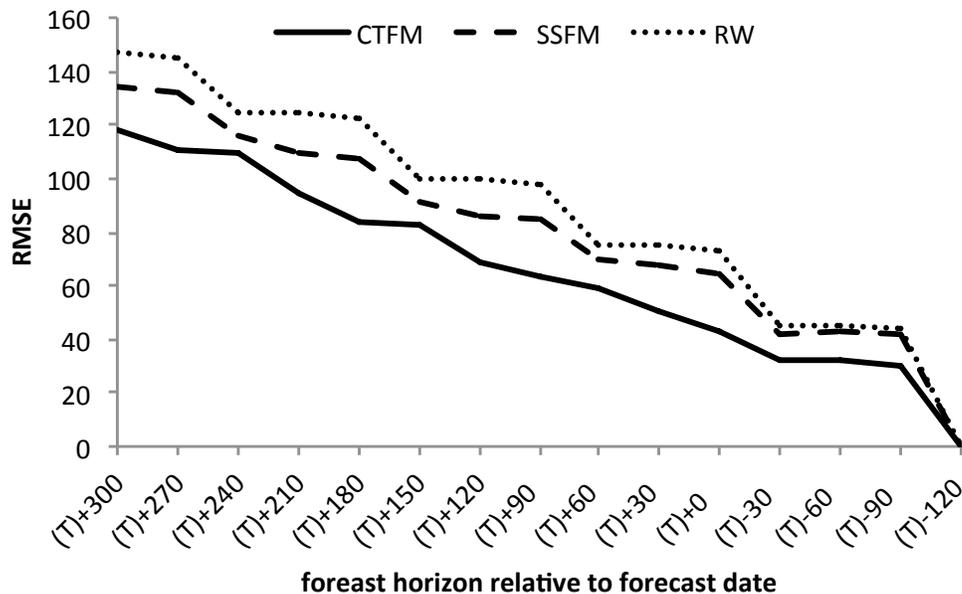


Figure 5: RMSE of *VEXPR* forecasts for an increasing information set and declining forecast horizon. Horizon $(T) + 300$ corresponds to a ten month ahead forecast from the forecast date, T . Horizon $(T) - 120$ corresponds to the release date of *VEXPR*.

estimates for horizons $(T) + 60$ through $(T) - 30$ are directly affected by the releases of weekly and monthly series. The impact of this high frequency information is then propagated to longer horizon forecasts. The almost linear evolution of RMSE and the lack of jumps at *VEXPR* release dates imply that the CTFM incorporates the intra-quarter information more effectively than the SSFM.

4 Conclusion

We extend the existing literature on small mixed frequency single-factor models by allowing for multiple factors, considering indicators in levels, and allowing for cointegration among the indicators. We show that the stationary single-factor models that are frequently used to extract coincident indicators are miss-specified if the data contain common stochastic trends. Our empirical results suggest that when the constituent indicators are cointegrated, modeling common stochastic trends, as opposed to eliminating them, will improve forecasts.

Our common-trends factor model outperforms the stationary mixed frequency single-factor model, at all analyzed forecast horizons with up to 32% gain in precision as measured by the root mean squared forecast error.

References

- Aruoba, S., Diebold, F., and Scotti, C. (2009). Real-time measurement of business conditions. *Journal of Business and Economic Statistics*.
- Bai, J. and Ng, S. (2010). Principal Components Estimation and Identification of the Factors. *Working paper*.
- 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.
- Christoffersen, P. and Diebold, F. (1998). Cointegration and long-horizon forecasting. *Journal of Business & Economic Statistics*, 16(4):450–458.
- 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*.
- 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.
- 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., Shephard, N., and Doornik, J. (1999). Statistical algorithms for models in state space using SsfPack 2.2. *Econometrics Journal*, 2(1):107–160.
- 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.
- Stock, J. and Watson, M. (1991). A probability model of the coincident economic indicators. *Leading Economic indicators: new approaches and forecasting records*, 66.