Can we use seasonally adjusted variables in dynamic factor models?[*]

Maximo Camacho[+]
Universidad de Murcia
mcamacho@um.es

Yuliya Lovcha
Universidad Rovira-i-Virgili
yuliya.lovcha@gmail.com

Gabriel Perez Quiros
Banco de España and CEPR
gabriel.perez@bde.es

Abstract

We examine the short-term performance of two alternative approaches of forecasting from dynamic factor models. The first approach extracts the seasonal component of the individual variables before estimating the model, while the alternative uses the non seasonally adjusted data in a model that endogenously accounts for seasonal adjustment. Our Monte Carlo analysis reveals that the performance of the former is always comparable to or even better than that of the latter in all the simulated scenarios. Our results have important implications for the factor models literature because they show the that the common practice of using seasonally adjusted data in this type of models is very accurate in terms of forecasting ability. Using five coincident indicators, we illustrate this result for US data.

Keywords: Dynamic factor models, seasonal adjustment, short-term forecasting

JEL Classification: E32, C22, E27.

[*] Maximo Camacho thanks CICYT for its support through grants ECO2010-19830 and ECO2013-45698. The views in this paper are those of the authors and do not represent the views of the Bank of Spain or the Eurosystem.

[+] Corresponding Author: Universidad de Murcia, Facultad de Economia y Empresa, Departamento de Metodos Cuantitativos para la Economia y la Empresa, 30100, Murcia, Spain. E-mail: mcamacho@um.es.

1. Introduction

The late-2000s recession, sometimes referred to as the Great Recession, magnified the interest of economic agents in having efficient short-term forecasting models that help monitor ongoing economic developments. This could explain the recent resurgence of dynamic factor models, first developed by Stock and Watson (1991), which have proven to be useful in growth and inflation forecasting. Among others, recent examples are Aruoba, Diebold and Scotti (2009), Aruoba and Diebold (2010) and Camacho and Perez-Quiros (2010).

To our knowledge, all of the forecasting analyses developed in this related literature use seasonally adjusted data, where the seasonal components are extracted individually from each variable either by the official statistical offices that publish the data or by the analyst (when seasonally adjusted data are not available) before estimating the models.[1] Therefore, only one common factor and several idiosyncratic components are estimated in these dynamic factor models. We will call this approach *traditional*, because it is the standard procedure in the literature.

This traditional approach has some limitations. First, behind the individual seasonal adjustments there exists the implicit assumption that the seasonal component for each variable is necessarily idiosyncratic (not common). Second, removing the seasonal component from the individual variables before estimating the models may lead to losses of information about the seasonal components that could potentially be useful for forecasting.

As an alternative to this traditional approach, the *structural* dynamic factor models have the advantage of being formulated in terms of common components, such as trends, seasonal components and cycles that have a direct interpretation. Modelling these features inside the model could be of great benefit since they could be easily projected into the future, leading to potential forecasting improvements.

This paper aims to evaluate the performance of *traditional* versus *structural* factor models. We use a Monte Carlo exercise to show that when the data generating process exhibits idiosyncratic seasonal components the traditional dynamic factor model that uses seasonally adjusted data (the outcomes of TRAMO-SEATS) outperforms the structural dynamic factor model, especially when the idiosyncratic seasonal components are erroneously

[1] Note that the literature on large-scale dynamic factor models, which include a vast number of indicators, also uses seasonally adjusted data. Although our results can be extended to large-scale models, we focus on small-scale models for the sake of simplicity. In addition, Boivin and Ng (2006), Poncela and Ruiz (2012) and Banbura and Modugno (2014) recently show that large specifications could perform worse than small specifications due to difficulties in extracting a relevant signal in the presence of indicators of different quality.

modelled as if they were common across series.[2] Interestingly, when the data are generated with common seasonal components, the performance of traditional factor model is still comparable to or even better than that of *structural* factor models, even in the case that the seasonal components are correctly modelled as common across the series.

One potential explanation is that if seasonality is idiosyncratic, the common seasonality model would be clearly misspecified; if there is common seasonality, extracting idiosyncratic seasonal terms is inefficient and suffers from the curse of dimensionality. The *traditional* approach apparently provides the best of both worlds: not making incorrect assumptions about common seasonality, while keeping a limited number of parameters to estimate. These results have important implications for the literature on factor models since they show the good forecasting performance of the standard models that use seasonally adjusted data with respect to alternative models that handle seasonally adjustments endogenously.

The results obtained in the Monte Carlo analysis are confirmed by using a set of five coincident US economic indicators. Our empirical results also suggest that the standard strategy of forecasting from dynamic factor models that use seasonally adjusted data is the most advisable way to compute the forecasts.

The paper is structured as follows. Section 2 describes the main features of structural and traditional dynamic factor models. Section 3 outlines the Monte Carlo simulation and discusses the results. Section 4 addresses the empirical analysis. Section 5 concludes.

2. Methodological framework

2.1. Structural factor decomposition

The stationary economic variables are assumed to admit a *structural* factor decomposition.[3] Therefore, each of the N stationary variables, y_{it}, can be written as the sum of three stochastic components: a common component, f_t, which represents the overall business cycle conditions; an idiosyncratic component, u_{it}, which refers to the particular dynamics of the series; and a seasonal component, s_{it}, which refers to the periodic patterns and are allowed to be either

[2] We use the TRAMO-SEATS version dated March 11, 2011, as downloaded from the Bank of Spain database. Alternative filters as X-11, X-12, and ARIMA models would lead to qualitatively similar results.

[3] We focus the analysis on stationary variables. The updates of Aruoba Diebold and Scotti (2009) showed that modeling the stochastic trends were very disappointing since the growth rate transformation facilitates better handling of benchmark revisions, which typically affect levels more than growth rates.

idiosyncratic or common. According to this decomposition, the *structural* dynamic factor model can be stated as

$$y_{it} = \alpha_i f_t + u_{it} + s_{it}, \tag{1}$$

where i=1,…,N, and the α_i are the loading factors.

We assume the following dynamic specifications for the three components. The common component and the idiosyncratic components follow autoregressive processes of orders $p1$ and $p2$, respectively:

$$f_t = a_1 f_{t-1} + ... + a_{p1} f_{t-p1} + \varepsilon_{ft}, \tag{2}$$

where $\varepsilon_{ft} \sim i.i.d.N\left(0, \sigma_f^2\right)$, and

$$u_{it} = b_{i1} u_{it-1} + ... + b_{ip2} u_{it-p2} + \varepsilon_{it}, \tag{3}$$

where $\varepsilon_{it} \sim i.i.d.N\left(0, \sigma_i^2\right)$, with i=1,…,N.

For the purposes of the paper, the treatment of the seasonal components deserves special comments. In standard applications that use factor decomposition analyses, which we called *traditional* models in this paper, the seasonal component of the series is extracted before estimating the model and, therefore, model selection, estimation and forecasting is carried on from seasonally adjusted series. The seasonal adjustment techniques are developed either by the researcher, usually with the help of automatic procedures, such as TRAMO-SEATS or X11, or by the statistical agencies, which in some cases publish only the seasonally adjusted versions of the time series. In expression (1), this implies that $s_{it} = 0$, i=1,…,N.

Alternatively, the dynamic properties of the seasonal components could be accounted for within the structural dynamic factor model. In line with trigonometric seasonality models (see Harvey, 1989), we assume that the seasonal component can be viewed as the sum of its s/2 cyclical components

$$s_{it} = \sum_{j=1}^{s/2} s_{ijt}, \tag{4}$$

where s is the number of observations per year. In this expression, the cyclical components are modelled as trigonometric terms at the seasonal frequencies, $\lambda_j = 2\pi j/s$, through the model

$$\begin{pmatrix} s_{ijt} \\ s_{ijt}^* \end{pmatrix} = \begin{pmatrix} \cos\lambda_j & \sin\lambda_j \\ -\sin\lambda_j & \cos\lambda_j \end{pmatrix} \begin{pmatrix} s_{ijt-1} \\ s_{ijt-1}^* \end{pmatrix} + \begin{pmatrix} \xi_{ijt} \\ \xi_{ijt}^* \end{pmatrix}, \tag{5}$$

where j=1,…,s/2, i=1,…,N, ξ_{ijt} and ξ_{ijt}^* are mutually uncorrelated noises with common variance $\sigma_{\xi_{ij}}^2$, and the term s_{ijt}^* appears by construction to form s_{ijt}. In addition, we use the standard assumption that the error terms exhibit the same variance across frequencies, i.e.,

$\sigma^2_{\xi_{ij}} = \sigma^2_{\xi_i}$ for all j=1,…,s/2. To complete the statistical specification of the model, we assume that all the disturbances driving the three stochastic components are mutually and serially uncorrelated.

To facilitate simulations and estimations, we prove in the Appendix that this seasonal component can be alternatively expressed by using a seasonal autoregressive integrated moving average specification. For quarterly data,[4] the seasonal components are

$$\left(1 + L + L^2 + L^3\right)s_{it} = \left(1 + 0.3187L + 0.1869L^2\right)\zeta_{it}, \tag{6}$$

where L is the backshift operator, $\zeta_{it} \sim i.i.d.N\left(0, \sigma^2_{\zeta_i}\right)$ reflects that the seasonal effect is allowed to change over time, and i=1,…,N.[5] To derive this expression, we used s=4, since the seasonal behaviour of our quarterly variables is often related to the time of a year.[6]

Let us consider some identification issues regarding the structural dynamic factor model described in (1). Let k be the number of common components in this model. In the case of the *traditional* model and the *structural* model with idiosyncratic seasonality, the factor is the only common component, which implies that k=1. Therefore, the minimum number of time series required to identify these models is N=2k+1=3.

To achieve identification in the *traditional* model, we also assume that the factor loading of the first variable is one. To achieve identification in the *structural* model with idiosyncratic seasonality, we also assume that the matrix of factor loadings is lower-triangular with units on the main diagonal. It means that the factor loading of the first variable is one, that first variable does not contain seasonal components ($s_{1t} = 0$), and that the seasonal patterns are proportional across series $s_{it} = \beta_i s_t$, i=2,3,4,5, with $\beta_2 = 1$.

In the case of the *structural* model with common seasonality, the common components are the factor and the season, which implies that k=2 and the minimum number of time series required to identify the model is N=2k+1=5. Therefore, we work with five variables since this is the minimum number of variables to ensure that the common factor is identified in all models.

It is worth pointing out that the structural model that assumes idiosyncratic seasonality could assign part of the seasonal variability to the idiosyncratic component or part of the

---

[4] For the sake of simplicity, we derive all the expressions for quarterly data. Although the expressions would be larger, all the results obtained in the paper could easily be obtained for monthly data.

[5] In this case, the yearly sum of the seasonal effects is expected to be zero, since the disturbance term has zero expectation. A model of deterministic seasonality is easily obtained by imposing $\zeta_{it} = 0$.

[6] Although we focus on trigonometric seasonality as in Harvey (1989), there are alternative ways of allowing seasonal variables to change over time, as in Hannan et al. (1970) or Harrison and Stevens (1976). However, the Hannan et al. (1970) seasonal model and the Harvey (1989) model with non-equal variances are the same models in the Gaussian case, or when innovations follow a mixture of normal distribution as in Bruce and Jurke (1992). The Harrison and Stevens (1976) seasonality with correlated disturbances model and the Hannan's model are the same models, which are also identical to the model that we use in the Gaussian case.

idiosyncratic variability could be modelled as seasonal. Therefore, the noise can be transmitted from the seasonal component, s_{it}, to the idiosyncratic component, u_{it}, and vice versa, which may influence the in-sample fitting performance of the model adversely.[7]

2.2. State-space representation

To estimate model's parameters and to infer unobserved components by using the Kalman filter, it is convenient to rewrite the equations that describe the model's dynamics in a state-space representation. In the case of N economic variables, which are collected in the vector Y_t, the appropriate state-space form of the model requires the specification of both the measurement equation, $Y_t = Hh_t + e_t$, with $e_t \sim i.i.d.N(0,R)$, and the prediction equation $h_t = \Phi h_{t-1} + \eta_t$, with $\eta_t \sim i.i.d.N(0,\Omega)$.

For this purpose, it is worth pointing out that the seasonal components

$$s_{it} = \frac{\left(1+0.3187L+0.1869L^2\right)}{\left(1+L+L^2+L^3\right)}\zeta_{it} \tag{7}$$

can be written as

$$s_{it} = \left(1+0.3187L+0.1869L^2\right)\chi_{it} \tag{8}$$

where $\left(1+L+L^2+L^3\right)\chi_{it} = \zeta_{it}$, with $\mathrm{var}(\zeta_{it}) = \sigma^2_{\zeta_i}$, i=2,…,5.

The specific forms of these two equations depend on the assumption about the seasonal component. Using the assumptions that N=5, p1=p2=1, when seasonal components are common across the different variables, the state space representation of the model becomes:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \\ y_{4t} \\ y_{5t} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \alpha_2 & 0 & 1 & 0 & 0 & 0 & 1 & 0.3187 & 0.1869 \\ \alpha_3 & 0 & 0 & 1 & 0 & 0 & \beta_3 & 0.3187\beta_3 & 0.1869\beta_3 \\ \alpha_4 & 0 & 0 & 0 & 1 & 0 & \beta_4 & 0.3187\beta_4 & 0.1869\beta_4 \\ \alpha_5 & 0 & 0 & 0 & 0 & 1 & \beta_5 & 0.3187\beta_5 & 0.1869\beta_5 \end{pmatrix} \begin{pmatrix} f_t \\ u_{1t} \\ u_{2t} \\ u_{3t} \\ u_{4t} \\ u_{5t} \\ \chi_t \\ \chi_{t-1} \\ \chi_{t-2} \end{pmatrix}, \tag{9}$$

[7] See Geweke (1977) and Geweke and Singleton (1981) for a general discussion of identification in dynamic factor models.

and

$$\begin{pmatrix} f_t \\ u_{1t} \\ u_{2t} \\ u_{3t} \\ u_{4t} \\ u_{5t} \\ \chi_t \\ \chi_{t-1} \\ \chi_{t-2} \end{pmatrix} = \begin{pmatrix} a & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & b_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & b_3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & b_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & b_5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} f_{t-1} \\ u_{1t-1} \\ u_{2t-1} \\ u_{3t-1} \\ u_{4t-1} \\ u_{5t-1} \\ \chi_{t-1} \\ \chi_{t-2} \\ \chi_{t-3} \end{pmatrix} + \begin{pmatrix} \varepsilon_{ft} \\ \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \\ \varepsilon_{5t} \\ \varsigma_t \\ 0 \\ 0 \end{pmatrix} \tag{10}$$

where R=0, and Ω is a diagonal matrix with main diagonal

$$diag(\Omega) = \left(\sigma_f^2, \sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2, \sigma_5^2, \sigma_\varsigma^2, 0, 0\right)'. \tag{11}$$

The state-space form of *traditional* dynamic factor models that use either the official seasonally adjusted data sets or the seasonally adjusted outcomes from TRAMO-SEATS can easily be derived from these expressions. In particular, it is obtained by imposing $\chi_t = \chi_{t-1} = \chi_{t-2} = \chi_{t-3} = 0$, and $\varsigma_t = 0$.

However, when the seasonal component is assumed to be idiosyncratic for each economic variable, the state-space representation of the model is

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \\ y_{4t} \\ y_{5t} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} \\ \alpha_2 & 0 & 1 & 0 & 0 & 0 & A & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} \\ \alpha_3 & 0 & 0 & 1 & 0 & 0 & 0_{1\times3} & A & 0_{1\times3} & 0_{1\times3} \\ \alpha_4 & 0 & 0 & 0 & 1 & 0 & 0_{1\times3} & 0_{1\times3} & A & 0_{1\times3} \\ \alpha_5 & 0 & 0 & 0 & 0 & 1 & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & A \end{pmatrix} \begin{pmatrix} f_t \\ u_{1t} \\ u_{2t} \\ u_{3t} \\ u_{4t} \\ u_{5t} \\ X_{1t} \\ X_{2t} \\ X_{3t} \\ X_{4t} \end{pmatrix}, \tag{12}$$

and

$$
\begin{pmatrix} f_t \\ u_{1t} \\ u_{2t} \\ u_{3t} \\ u_{4t} \\ u_{5t} \\ X_{2t} \\ X_{3t} \\ X_{4t} \\ X_{5t} \end{pmatrix} = \begin{pmatrix} a & 0 & 0 & 0 & 0 & 0 & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} \\ 0 & b_1 & 0 & 0 & 0 & 0 & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} \\ 0 & 0 & b_2 & 0 & 0 & 0 & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} \\ 0 & 0 & 0 & b_3 & 0 & 0 & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} \\ 0 & 0 & 0 & 0 & b_4 & 0 & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} \\ 0 & 0 & 0 & 0 & 0 & b_5 & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} & 0_{1\times3} \\ 0_{3\times1} & 0_{3\times1} & 0_{3\times1} & 0_{3\times1} & 0_{3\times1} & 0_{3\times1} & B & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times1} & 0_{3\times1} & 0_{3\times1} & 0_{3\times1} & 0_{3\times1} & 0_{3\times1} & 0_{3\times3} & B & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times1} & 0_{3\times1} & 0_{3\times1} & 0_{3\times1} & 0_{3\times1} & 0_{3\times1} & 0_{3\times3} & 0_{3\times3} & B & 0_{3\times3} \\ 0_{3\times1} & 0_{3\times1} & 0_{3\times1} & 0_{3\times1} & 0_{3\times1} & 0_{3\times1} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & B \end{pmatrix} \begin{pmatrix} f_{t-1} \\ u_{1t-1} \\ u_{2t-1} \\ u_{3t-1} \\ u_{4t-1} \\ u_{5t-1} \\ X_{2t-1} \\ X_{3t-1} \\ X_{4t-1} \\ X_{5t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{ft} \\ \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \\ \varepsilon_{5t} \\ Z_{1t} \\ Z_{2t} \\ Z_{3t} \\ Z_{4t} \end{pmatrix}, \quad (13)
$$

where $A = (1, 0.3187, 0.1869)'$, $X_{it} = \left(\chi_{it}, \chi_{it-1}, \chi_{it-2}\right)'$, $Z_{it} = \left(\zeta_{it}, 0, 0\right)'$, and i=2,..,5.

$$
B = \begin{pmatrix} -1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad (14)
$$

$R = 0$, and Ω is a diagonal matrix with main diagonal

$$
diag(\Omega) = \left(\sigma_f^2, \sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2, \sigma_5^2, \sigma_{\varsigma_2}^2, 0, 0, \sigma_{\varsigma_3}^2, 0, 0, \sigma_{\varsigma_4}^2, 0, 0, \sigma_{\varsigma_5}^2, 0, 0\right)'. \quad (15)
$$

3. Monte Carlo simulations

In this section, we design a Monte Carlo experiment to study some of the finite-sample properties of *structural* dynamic factor models that account for common or idiosyncratic seasonal components against *traditional* dynamic factor models that manage seasonally adjusted data. As shown in Table 1, the experiment is conducted with a comprehensive set of coefficients in order to capture a wide range of specifications, allowing for different degrees of common factor correlation, different persistence of idiosyncratic components, and idiosyncratic components that are heterogeneous.

To cover a large variety of combinations, Table 1 reports that the loading factor of the first variable is set to unity in order to achieve identification, while other factor loadings are either positive for variables 2 and 3 (0.7 and 1.1, respectively) or negative for variables 4 and 5 (-0.8 and -0.5, respectively). We generate two alternative scenarios for the seasonal components.[8] The first scenario, called M1, tries to mimic the empirical forecasting exercise where seasonal components are idiosyncratic.[9] The second scenario, called M2, tries to mimic

[8] As mentioned before, for identification purposes, the seasonal component does not affect the first variable.

[9] In line with our empirical results, we set $\sigma_{\zeta_i}^2 < \sigma_j^2 \ \forall \ i, j$. In particular, we set $\sigma_{\zeta_i}^2$ around 0.1.

the case of common seasonal components, where the seasonal factor loadings are either positive (0.9 for variable 3) or negative (-0.8 and -0.7 for variables 4 and 5, respectively).

The common non-seasonal factor, f_t, and the individual components, u_{it}, are generated as first order autoregressive processes. According to Table 1, in simulations S1, S2, and S3 we replicate situations where the economic variables share a strong persistent non-seasonal common component (autoregressive parameter of 0.9). However, in simulations S4, S5 and S6 the persistence of the factor is weak (autoregressive parameter of 0.2) while it is moderate in simulations S7, S8 and S9 (autoregressive parameter of 0.5).

In addition, these potential empirical cases are combined with several degrees of autoregressive parameters of the idiosyncratic components. The persistence is strong (between 0.1 and 0.4) in simulations labelled S1, S4 and S7, it is weak (between 0.6 and 0.9) in simulations labelled as S2, S5 and S8 and mixed (between 0.2 and 0.9) in simulations S3, S6 and S9. In line with the empirical results of Section 4, the simulations labelled as S10 uses moderate persistence of the non-seasonal common component mixed (positive and negative, weak and strong) autoregressive parameters for the idiosyncratic components.

For all of these data-generating processes, the signal-to-noise ratio or proportion of the variance attributed to the common factor to the variance of the idiosyncratic component at business cycle frequencies (3 to 8 years) is also included in the table.

In the simulations, we try to cover various scenarios, according to the contribution of the common factor to the variation of the series. In simulations labeled as S2 and S4, the common factor and the idiosyncratic components account for about the same portion of the variance (signal-to-noise ratio close to 1). In simulations S1, S3 and S7, the variance of the common factor is larger than that of the idiosyncratic components (signal-to-noise ratios higher than 1), while it is much smaller (signal-to-noise ratios lower than 1) in simulations S5, S6 and S8. Finally, simulations S9 and S10 account for mixed scenarios.

For each of these cases, we generate a total of M=1000 sets of time series of length T=120 observations.[10] We use them to mimic three different empirical forecasting scenarios. The first scenario, called EId, mimics the case in which an analyst fits a *structural* dynamic factor model to the non seasonally adjusted data, whose seasonal components are treated as idiosyncratic. The second scenario, called ECo, refers to a similar case but where the seasonal component is common to the last four time series. The third scenario, called EsaTS, mimics the case in which the analyst uses seasonally adjusted data before estimating the standard dynamic factor model, i.e., the *traditional* approach. In our analysis, the seasonal components are extracted from the generated time series using TRAMO-SEATS.

[10] The length of the generated time series is 120 since it would refer to 30 years of quarterly observations.

In each replication, m, we estimate the two *structural* factor models and the *traditional* factor model that uses seasonally adjusted data. We examine the performance of these models in Tables 2 and 3. In each of these tables, the figures in brackets analyze the ability of the models to infer the factor while the rest of the figures refer to the accuracy of the models to infer the time series.

In Table 2, we examine the in-sample fit of the models by computing the averaged squared difference across the T observations between the generated and the estimated time series (Mean Squared Errors, MSE), which are also averaged across the M replications. In Table 3, we compare the out-of-sample forecasting accuracy by computing the errors in forecasting (one-step-ahead) the generated target series. For each m-th replication, the one-step-ahead forecasts are obtained by estimating the models with data from t=1 to t=T-1, and by computing the forecasts for T.

To facilitate interpretation, the tables report fractions of MSEs, where the denominator is the MSE from forecasting using the specification that agrees with the data-generating parameters. That is, the figures compare the MSEs of all models against the situation where an oracle has given the researcher the correct model to forecast from. Therefore, figure below one indicate that the forecasting model does the job better than the oracle.

The main results of the Monte Carlo experiment are the following. First, there are two main potential sources of seasonal misspecification in *structural* dynamic factor models: when the data are generated with idiosyncratic seasons but the model incorrectly assumes common seasons (columns labelled as M1) and when the data are generated with common seasons but the model uses the erroneous assumption that the seasons are idiosyncratic (columns labelled as M2). According to the magnitude of the figures reported in the tables, the second source of misspecification seems to be much less damaging than the first.

Second, when the seasonal component is generated idiosyncratically across the time series, the *traditional* approach of dynamic factor models that use seasonally adjusted data unequivocally achieves the best performance. The figures reported in the third column of the panels labelled as M1 show that this strategy outperforms the *structural* factor model that assumes idiosyncratic seasons. The potential explanation is that the *structural* factor model may suffer from an identification problem since it is hard to identify separately the variances of the individual components from those of the seasonal components when they are idiosyncratic. Another explanation would be that the greater number of parameters to be estimated within the *structural* approach generates larger uncertainty and noise in the estimation.

Third, the *structural* model that correctly treats the seasonal components as common when they are actually generated as common (fifth columns of the tables) usually exhibits the best performance. However, it is worth emphasizing that the accuracy of this model is comparable to that of the *traditional* factor model that uses seasonally adjusted data, which in many cases exhibits the lowest MSEs.

Fourth, the persistence of the idiosyncratic and the common components increases the size of the differences across specifications but does not alter the nature of the results. Regardless of whether the seasonal components are common or idiosyncratic, the *traditional* factor model achieves relatively better forecasting performance in the case of high persistence, which motivates the use of this approach in case of doubts about the nature of the seasonal components.

Fifth, the conclusions obtained by analysing the MSEs achieved by the models on inferring the factor and those achieved by the models on fitting the variables are of the same nature, i.e., good factor estimation implies good fitting of the data. In addition, the results of the out-of-sample analysis (Table 3) are qualitatively similar to those of the in-sample performance, although a little weaker. The intuition is that there is more noise in the out-of-sample analysis, which generates higher uncertainty across the models and makes it difficult to extract conclusions from the analysis.

Summing up, these results agree with the general strategy followed by analysts that routinely apply *traditional* dynamic factor models to time series that exhibit common, idiosyncratic, and seasonal components. This approach consists, prior to fitting the factor model, of removing the seasonal components. When seasonality is idiosyncratic, this strategy leads to the best results. When the seasonality is common across series, it leads to very good results, which are comparable to the results of estimating the *structural* factor model associated with the data generating process. Notably, the *traditional* approach exhibits the advantage of eliminating the potential damage of using *structural* factor models that assume common seasonality when it is actually idiosyncratic.

4. Empirical analysis

4.1. In-sample analysis

The five quarterly indicators used in the empirical analysis are the University of Michigan consumer sentiment index, new passenger car and truck sales, median usual weekly earnings in constant dollars, total housing starts (new privately owned housing units started), and employees on nonagricultural payrolls from 1978.1 to 2007.4.[11] According to our preliminary analysis of unit roots, we find that all of them contain unit roots; therefore all variables are used in growth rates.

The University of Michigan consumer sentiment index is a consumer confidence index published by the University of Michigan and Thomson Reuters. The index is normalized to have a value of 100 in December 1964 and it is based on at least 500 telephone interviews, which are conducted each month in a United States sample to assess near-time consumer attitudes on the business climate, personal finance, and spending. The index does not contain seasonality. New passenger car and truck sales and housing starts were obtained from the Department of Commerce's Bureau of Economic Analysis (BEA), the median usual weekly earnings and nonagricultural payrolls were obtained from the Bureau of Labor Statistics.

These economic indicators exhibit a key advantage for our study: they are available as both non-seasonally adjusted and seasonally adjusted. In addition, the selection of these indicators follows the line suggested by the influential paper of Stock and Watson (1991). We start the analysis with a set of indicators that includes an indicator from the supply side of the economy (housing starts), an indicator from the demand side (car and truck sales), an indicator from the income side (weekly earnings), and an indicator of the labor market (employees on nonagricultural payrolls).[12] Then, we enlarge the initial set of indicators with the University of Michigan consumer sentiment in order to incorporate a non-seasonal series which agrees with the evolution of the business cycle.

Table 4 displays the maximum likelihood estimates of the *structural* dynamic factor models that account for seasonal adjustments and the *traditional* factor models that use seasonally adjusted data where the seasonal components are extracted before estimation. The figures that appear in brackets refer to their standard deviations.[13] The choice of model specifications is always based on the Schwarz criterion. The table also shows the log-Likelihood achieved by these models and their signal-to-noise ratios.

There are several noteworthy features from the estimates reported in Table 4. First, the estimated common factor is moderate (or even weak) since the estimates for its first order autocorrelation range from 0.06 to 0.52, which refer to cases S4 to S10 in the Monte Carlo experiment. Second, the persistence of the individual components of the first four variables is mixed since some autoregressive parameters are small while some others are large, which

[11] The Great Recession was not included to overcome the problems associated to the large breaks of the time series in these years.
[12] To manage both seasonally and non-seasonally adjusted series, we substitute manufacturing and trade sales, originally used in Stock and Watson (1991) for car and truck sales. The same applies to the substitution of real personal income less transfers by weekly earnings.
[13] Although not included to save space, we obtained similar results when the *traditional* model consists of using the official seasonally adjusted version of the variables.

agrees with simulations S4, S9 and S10. Finally, some of these autoregressive parameters are negative, which agrees with simulation S10. Third, employment exhibits large positive autoregressive coefficients for the idiosyncratic components, which leads to the low signal-to-noise ratios in EId and EsaTS models.

4.2. Out-of-sample analysis

In this section, we develop an out-of-sample forecasting analysis. For this purpose, we assume that the time series of interest to be forecasted are the seasonally adjusted outcomes of TRAMO-SEATS of the five economic indicators described in the previous section.[14] The h-period-ahead forecasts were computed recursively and the analysis was conducted to simulate real-time forecasting. The first forecast is obtained by estimating the models with data from t=1 to $t=\tau$, and by computing the forecast $\tau+h$. Then, the models are re-estimated with data from t=1 to $t=\tau+1$, and the forecast is obtained for $\tau+h+1$. This process is repeated until $\tau=T-h$. The first simulated out-of-sample forecast was made in 1998.1 and we consider forecast horizons of h=1 and h=4 periods.[15]

For each time series, the averaged squared differences between the forecasts and the targeted variables are computed. The results, which are reported in Table 5, suggest some conclusions that are in line with the findings obtained in the Monte Carlo analysis. Regardless of the targeted series and the forecast horizon, the forecasts computed from the *traditional* dynamic factor model that uses seasonally adjusted series exhibit the lowest MSE.[16] In addition, the table shows the p-values of Clark and West (2007), which compare the accuracy of the traditional approach with the two versions of the structural approach and the seasonal ARIMA forecasts. Overall, the table shows that the better performance of the traditional approach is statistically significant at 5% level at 1-period forecast horizon.

Therefore, cleaning up the economic indicators from seasonality either by using the seasonally adjusted series or by using automatic univariate procedures before using the variables in the dynamic factor model seems to be a reasonably strategy to follow.

5. Conclusion

[14] The results obtained from the official seasonally adjusted time series are quite similar to the seasonally adjusted outcomes of TRAMO-SEATS. Accordingly, if the former were the series of interest, the results would be qualitatively similar to those presented in the paper.

[15] Each quarter, we updated the database as if all the variables had been observed in that quarter. Therefore, we did not develop a pseudo real-time analysis since data revisions or publication delays are not treated. For a careful analysis of these forecasting problems, see for example Camacho, Perez-Quiros and Poncela (2012).

[16] There is only one exception: the one-step-ahead forecasts of cars. Notably, the gain with respect to the *traditional* factor model is not large.

Despite the efforts of recent studies to evaluate the empirical short-term forecasting performance of dynamic factor models, it still remains an open question whether it is better to use seasonally adjusted indicators before estimating the model or to account for the seasonal components of the raw data within a factor model. The first strategy implicitly assumes that the seasonal components are idiosyncratic and the latter strategy could lead to unnecessary complexity, especially for practitioners that are not familiar with seasonal analysis.

We use Monte Carlo experiments to analyze the extent to which these two alternatives exhibit relative forecasting performance gains. Our simulation results suggest that when the data are generated under the assumption that the seasonal components are idiosyncratic the dynamic factor model that uses seasonally adjusted indicators exhibits the best forecasting performance. Interestingly, when the seasonal components are common to all the time series, its forecasting deterioration with respect to a dynamic factor model that account for the common seasonality is usually negligible in our experiment. Notably, the former improves on the latter in several cases.

In empirical applications, it is difficult to decide a priori if the seasonality is common or idiosyncratic across series. Given that the deterioration of the in-sample fitting and out-of-sample forecasting performance of the dynamic factor model that uses seasonally adjusted indicators is very small, while the performance of the common seasonal component model is very poor in the case of idiosyncratic seasonal factors, we strongly recommend the use of seasonally adjusted series in factor models.

We illustrate these results by using US data from 1978.1 to 2011.1 of the University of Michigan consumer sentiment index, new passenger car and truck sales, median usual weekly earnings, housing starts, and employees on nonagricultural. In line with the simulations results, the forecasting performance of a dynamic factor model that uses the seasonally adjusted versions of these series is better than those of dynamic factor models that assume common or idiosyncratic seasonal components.

Appendix

Since the empirical data are quarterly, the seasonal component of each time series, s_i, is the sum of two cyclical components, $s_{it} = s_{i1t} + s_{i2t}$, which are evaluated at the seasonal frequencies, $\lambda_1 = \pi/2$, and $\lambda_2 = \pi$. According to (4), the dynamics of the first cyclical component is

$$\begin{pmatrix} s_{i1t} \\ s_{i1t}^{*} \end{pmatrix} = \begin{pmatrix} \cos\pi/2 & \sin\pi/2 \\ -\sin\pi/2 & \cos\pi/2 \end{pmatrix} \begin{pmatrix} s_{i2t-1} \\ s_{i2t-1}^{*} \end{pmatrix} + \begin{pmatrix} \xi_{i1t} \\ \xi_{i1t}^{*} \end{pmatrix}. \tag{A1}$$

Using $\cos\pi/2 = 0$ and $\sin\pi/2 = 1$ and rearranging terms, one can obtains

$$\begin{pmatrix} s_{i1t} \\ s_{i1t}^{*} \end{pmatrix} = \frac{1}{1+L^2} \begin{pmatrix} 1 & L \\ -L & 1 \end{pmatrix} \begin{pmatrix} \xi_{i1t} \\ \xi_{i1t}^{*} \end{pmatrix}, \tag{A2}$$

which implies that $\left(1+L^2\right)s_{i1t} = \zeta_{i1t}$, where $\zeta_{i1t} = \xi_{i1t} + \xi_{i1t-1}^{*}$. If $\mathrm{var}(\xi_{i1t}) = \mathrm{var}(\xi_{i1t}^{*}) = \sigma_{\xi_{i1}}^2$, $\forall t$, then $\mathrm{var}(\zeta_{i1t}) = 2\sigma_{\xi_{i1}}^2$.

Similarly, the dynamics of the second cyclical component can be obtained from

$$\begin{pmatrix} s_{i2t} \\ s_{i2t}^{*} \end{pmatrix} = \begin{pmatrix} \cos\pi & \sin\pi \\ -\sin\pi & \cos\pi \end{pmatrix} \begin{pmatrix} s_{i2t-1} \\ s_{i2t-1}^{*} \end{pmatrix} + \begin{pmatrix} \xi_{i2t} \\ \xi_{i2t}^{*} \end{pmatrix}, \tag{A3}$$

which, using $\cos\pi = -1$ and $\sin\pi = 0$, leads to

$$\begin{pmatrix} s_{i2t} \\ s_{i2t}^{*} \end{pmatrix} = \frac{1}{(1+L)^2} \begin{pmatrix} 1+L & 0 \\ 0 & 1+L \end{pmatrix} \begin{pmatrix} \xi_{i2t} \\ \xi_{i2t}^{*} \end{pmatrix}. \tag{A4}$$

This expression implies that $(1+L)s_{i2t} = \zeta_{i2t}$, where $\zeta_{i2t} = \xi_{i2t}$ and $\mathrm{var}(\zeta_{i2t}) = \sigma_{\xi_{i2}}^2$. Let us additionally assume that $\sigma_{\xi_{i1}}^2 = \sigma_{\xi_{i2}}^2 = \sigma_{\xi_i}^2$.

Accordingly, the seasonal component of each time series can be expressed as

$$s_{it} = \frac{\zeta_{i1t}(1+L) + \zeta_{i2t}\left(1+L^2\right)}{\left(1+L^2\right)(1+L)}, \tag{A5}$$

or

$$\left(1+L+L^2+L^3\right)s_{it} = \zeta_{i1t}(1+L) + \zeta_{i2t}\left(1+L^2\right). \tag{A6}$$

Since the greatest polynomial of the two terms from the right-hand side is of power two, the resulting polynomial (the result of summation) is of power two as well

$$\left(1+L+L^2+L^3\right)s_{it} = \left(1+\alpha L+\beta L^2\right)\zeta_{it}\ . \tag{A7}$$

To find the unknown coefficients α and β, we derive the spectra of right-hand sides of both expressions. On the one hand, the spectrum of right-hand side of (A6) is

$$\begin{aligned}&\left(1+e^{i\omega}\right)\left(1+e^{-i\omega}\right)\sigma^2_{\zeta_{i1}} + \left(1+e^{2i\omega}\right)\left(1+e^{-2i\omega}\right)\sigma^2_{\zeta_{i2}} = \left(2+2\cos\omega\right)\sigma^2_{\zeta_{i1}} + \\ &+\left(2+2\cos 2\omega\right)\sigma^2_{\zeta_{i2}} = \left(6+4\cos\omega+2\cos 2\omega\right)\sigma^2_{\xi_i}.\end{aligned} \tag{A8}$$

The first equality follows from the fact that $e^{i\lambda}+e^{-i\lambda}=2\cos\lambda$, $\forall\lambda$. The last equation uses $\sigma^2_{\zeta_{i1}} = 2\sigma^2_{\xi_i}$ and $\sigma^2_{\zeta_{i2}} = \sigma^2_{\xi_i}$. On the other hand, the spectrum of right-hand side of (A7) is

$$\begin{aligned}&\left(1+\alpha e^{i\omega}+\beta e^{2i\omega}\right)\left(1+\alpha e^{-i\omega}+\beta e^{-2i\omega}\right)\sigma^2_{\zeta_i} = \\ &\left(\left(1+\alpha^2+\beta^2\right)+2\left(\alpha+\alpha\beta\right)\cos\omega+2\beta\cos 2\omega\right)\sigma^2_{\zeta_i}\end{aligned}\ . \tag{A9}$$

Since the two spectra must represent the same dynamics, one can use the system of three equations with three unknowns $6\sigma^2_{\xi_i} = \left(1+\alpha^2+\beta^2\right)\sigma^2_{\zeta_i}$, $4\sigma^2_{\xi_i} = 2\left(\alpha+\alpha\beta\right)\sigma^2_{\zeta_i}$ and $2\sigma^2_{\xi_i} = 2\beta\sigma^2_{\zeta_i}$ to obtain $\alpha^4-4\alpha^3+12\alpha^2-16\alpha+4=0$. The real solutions of this equation are $\alpha_1 = 0.3187$ and $\alpha_2 = 1.6813$, and using again the system of equations, it is easy to obtain that they correspond to values $\beta_1 = 0.1869$ and $\beta_2 = 5.2745$. The first pair of solutions $\left(\alpha = 0.3187, \beta = 0.1869\right)$ produces invertible MA polynomial in (A7), opposite, the second pair of solutions results in non-invertible (A7). Using the first pair of real solutions we find $\sigma^2_{\zeta_i} = 5.3505\sigma^2_{\xi_i}$ from the last equation of the system. In this way, the seasonal component for the series i is given by:

$$\left(1+L+L^2+L^3\right)s_{it} = \left(1+0.3187L+0.1869L^2\right)\zeta_{it}\ .$$

References

Aruoba, B., Diebold, F., and Scotti, C. 2009. Real-time measurement of business conditions. *Journal of Business and Economic Statistics* 7: 417-427.

Aruoba, B., and Diebold, F. 2010. Real-time macroeconomic monitoring: Real activity, inflation, and interactions. *American Economic Review* 100: 20-24.

Banbura, M., and Modugno, M. 2010. Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data. *Journal of Applied Econometrics* 29: 133-160.

Boivin, J., and Ng, S. 2006. Are more data always better for factor analysis? *Journal of Econometrics* 132: 169-194.

Bruce, A.G. and Jurke, S.R. 1996. Non-Gaussian seasonal adjustment: X-12 ARIMA versus robust structural models. *Journal of Forecasting* 15: 305-327.

Camacho, M., and Perez-Quiros, G. 2010. Introducing the Euro-STING: Short Term INdicator of euro area Growth. *Journal of Applied Econometrics* 25: 663-694.

Camacho, M, Perez-Quiros, and Poncela, P. 2012. Markov-switching dynamic factor models in real time. CEPR Discussion Paper no. 8866.

Clark, T., and West, K. 2007. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics* 138: 291-311.

Geweke, J.F. 1977. The dynamic factor analysis of economic time series models, in D. Aigner and A. Goldberger (eds.), Latent Variables in Socioeconomic Models, Latent Variables in Socioeconomic Models.

Geweke, J.F. and Singleton, K.J. 1981. Maximum likelihood "confirmatory" factor analysis of economic time series, *International Economic Review* 22, 37-54.

Hannan, E. J., Terrell, R. D. and Tuckwell, N. E. 1970. The seasonal adjustment of economic time series, *International Economic Review* 11, 24-52.

Harrison, P.J. and Stevens, C.F. 1976. Bayesian Forecasting. *Journal of the Royal Statistical Society, Series B* 38: 205:47.

Harvey, A. 1989. Forecasting. Structural time series models and Kalman filter, Cambridge University Press.

Poncela, P., and Ruiz, E. 2012. More is not always better: back to the Kalman filter in dynamic factor models. UC3M Working papers, Statistics and Econometrics, 12/17.

Stock, J., and Watson, M. 1991. A probability model of the coincident economic indicators. In Kajal Lahiri and Geoffrey Moore editors, *Leading economic indicators, new approaches and forecasting records*. Cambridge University Press, Cambridge.

Table 1. Parameters used in Monte Carlo simulations

Fixed parameters for all simulations										
$\alpha_1 = 1, \alpha_2 = 0.7, \alpha_3 = 1.1, \alpha_4 = -0.8, \alpha_5 = -0.5$										
$\sigma_f^2 = 1, \sigma_1^2 = 0.7, \sigma_2^2 = 0.8, \sigma_3^2 = 0.9, \sigma_4^2 = 1, \sigma_5^2 = 0.9$										
Parameters that control idiosyncratic versus common seasons										
	M1 (idiosyncratic seasons)					M2 (common seasons)				
	$\sigma_{\zeta_2}^2 = 0.1, \sigma_{\zeta_3}^2 = 0.09, \sigma_{\zeta_4}^2 = 0.08, \sigma_{\zeta_5}^2 = 0.1$					$\sigma_{\zeta_2}^2 = 0.1, \beta_3 = 0.9, \beta_4 = -0.8, \beta_5 = -0.7$				
Parameters that control non-seasonal factor and individual components										
Not seasonal component	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
Common	strong	strong	strong	weak	weak	weak	mod	mod.	mod	mod
Idiosynchr.	weak	strong	mixed	weak	strong	mixed	weak	strong	mixed	p-n
	a=0.9	a=0.9	a=0.9	a=0.2	a=0.2	a=0.2	a=0.5	a=0.5	a=0.5	a=0.5
	b_1=0.3	b_1=0.7	b_1=0.9	b_1=0.3	b_1=0.7	b_1=0.9	b_1=0.3	b_1=0.7	b_1=0.9	b_1=0.9
	b_2=0.2	b_2=0.8	b_2=0.2	b_2=0.2	b_2=0.8	b_2=0.2	b_2=0.2	b_2=0.8	b_2=0.2	b_2=-0.2
	b_3=0.4	b_3=0.9	b_3=0.5	b_3=0.4	b_3=0.9	b_3=0.5	b_3=0.4	b_3=0.9	b_3=0.5	b_3=-0.4
	b_4=0.1	b_4=0.7	b_4=0.9	b_4=0.1	b_4=0.7	b_4=0.9	b_4=0.1	b_4=0.7	b_4=0.9	b_4=-0.1
	b_5=0.3	b_5=0.6	b_5=0.3	b_5=0.3	b_5=0.6	b_5=0.3	b_5=0.3	b_5=0.6	b_5=0.3	b_5=0.2
	12.46	3.83	2.04	1.61	0.50	0.26	3.50	1.08	0.57	0.57
	4.67	0.81	4.67	0.60	0.10	0.60	1.31	0.23	1.31	2.80
SNR	8.92	1.89	6.72	1.16	0.25	0.87	2.51	0.53	1.89	11.83
	6.12	1.20	0.64	0.79	0.16	0.08	1.72	0.34	0.18	2.51
	1.88	0.81	1.88	0.24	0.10	0.24	0.53	0.23	0.53	0.67

Notes. Parameters α_i refer to the loading factors. Parameters σ_f^2 and σ_i^2 refer to the variance of noises of the common non-seasonal factor and the idiosyncratic components, respectively. Parameters $\sigma_{\zeta_i}^2$ refer to the variances of the noises of the cyclical components. Parameters a and b_i refer to the autoregressive parameters of the common factor and the idiosyncratic components, respectively. SNR is the signal-to-noise of the variance of the common factor at business cycle frequencies (3 to 8 years) to the variance of the idiosyncratic component in the series.

Table 2. In-sample Monte Carlo results

Specification	M1 (idiosyncratic seasons)			M2 (common seasons)		
	EId	ECo	EsaTS	EId	ECo	EsaTS
S1	(1.00)	(1.57)	(0.91)	(1.11)	(1.00)	(0.92)
	1.00	7.62	0.46	1.83	1.00	0.73
	1.00	6.13	0.37	3.13	1.00	0.85
	1.00	9.19	0.49	2.24	1.00	1.14
	1.00	13.62	0.33	3.38	1.00	1.09
S2	(1.00)	(2.52)	(1.00)	(1.07)	(1.00)	(0.98)
	1.00	6.39	0.44	1.87	1.00	0.65
	1.00	7.56	0.37	3.18	1.00	0.92
	1.00	6.84	0.39	2.10	1.00	1.19
	1.00	14.05	0.32	3.36	1.00	0.92
S3	(1.00)	(2.34)	(0.94)	(1.14)	(1.00)	(0.94)
	1.00	12.32	0.48	1.95	1.00	0.77
	1.00	6.05	0.47	3.31	1.00	0.88
	1.00	6.01	0.39	1.84	1.00	0.84
	1.00	13.95	0.41	3.60	1.00	1.15
S4	(1.00)	(1.69)	(0.94)	(1.17)	(1.00)	(0.98)
	1.00	8.21	0.60	1.97	1.00	0.69
	1.00	5.32	0.52	3.17	1.00	1.06
	1.00	7.44	0.45	2.33	1.00	0.99
	1.00	10.06	0.45	3.60	1.00	1.02
S5	(1.00)	(1.72)	(1.01)	(1.09)	(1.00)	(1.02)
	1.00	6.94	0.55	1.88	1.00	0.69
	1.00	6.73	0.58	3.27	1.00	1.58
	1.00	9.45	0.50	2.21	1.00	1.17
	1.00	12.82	0.44	3.32	1.00	1.04
S6	(1.00)	(1.73)	(0.98)	(1.15)	(1.00)	(1.02)
	1.00	16.36	0.51	2.01	1.00	0.82
	1.00	5.50	0.51	4.27	1.00	1.29
	1.00	7.02	0.47	1.95	1.00	1.14
	1.00	12.40	0.44	3.80	1.00	1.06
S7	(1.00)	(1.73)	(0.93)	(1.15)	(1.00)	(0.95)
	1.00	7.68	0.54	1.89	1.00	0.70
	1.00	5.39	0.48	3.35	1.00	0.97
	1.00	6.72	0.46	2.29	1.00	1.14
	1.00	13.31	0.39	3.45	1.00	0.99
S8	(1.00)	(1.81)	(1.00)	(1.08)	(1.00)	(0.99)
	1.00	7.91	0.51	1.87	1.00	0.63
	1.00	5.39	0.54	3.53	1.00	1.06
	1.00	6.98	0.39	2.14	1.00	1.00
	1.00	15.14	0.42	3.59	1.00	1.09
S9	(1.00)	(1.74)	(0.98)	(1.14)	(1.00)	(0.98)
	1.00	9.74	0.54	2.03	1.00	0.85
	1.00	6.62	0.51	3.24	1.00	1.28
	1.00	5.10	0.42	2.18	1.00	0.97
	1.00	14.57	0.39	3.55	1.00	0.94
S10	(1.00)	(1.02)	(1.00)	(1.01)	(1.00)	(0.98)
	1.00	1.05	1.02	1.01	1.00	1.04
	1.00	1.22	1.00	1.00	1.00	0.97
	1.00	1.30	0.81	1.74	1.00	0.99
	1.00	1.33	0.99	1.00	1.00	0.97
	1.00	2.04	0.99	1.00	1.00	0.98

Notes. Expressions S1 to S10 are described in Table 1. Figures in parentheses refer to the MSE of the common factor while other figures refer to the MSE of series 2 to 5. Columns labelled as M1 and M2 refer to data generated processes with idiosyncratic and common seasonal components, respectively. EId, Eco, and EsaTS refer to models with idiosyncratic seasons, common season, and models whose indicators are seasonally adjusted (TRAMO-SEATS) before estimation, respectively. Lowest MSE are highlighted in bold.

Table 3. One-period-ahead Monte Carlo results

Specification	M1 (idiosyncratic seasons)			M2 (common seasons)		
	EId	ECo	EsaTS	EId	Eco	EsaTS
S1	(1.00)	(1.02)	(0.92)	(1.01)	(1.00)	(0.94)
	1.00	0.97	1.00	1.01	1.00	0.97
	1.00	1.43	0.67	1.01	1.00	0.96
	1.00	1.19	0.77	1.03	1.00	0.97
	1.00	1.84	0.51	1.01	1.00	1.01
	1.00	3.48	0.27	1.01	1.00	0.98
S2	(1.00)	(1.28)	(0.79)	(0.97)	(1.00)	(0.96)
	1.00	1.02	0.99	0.99	1.00	0.99
	1.00	1.29	0.70	0.98	1.00	0.93
	1.00	1.47	0.63	1.00	1.00	0.93
	1.00	1.39	0.67	1.00	1.00	0.94
	1.00	2.83	0.31	0.97	1.00	0.92
S3	(1.00)	(1.10)	(0.93)	(1.01)	(1.00)	(0.91)
	1.00	1.00	1.00	1.00	1.00	0.99
	1.00	1.73	0.58	1.01	1.00	0.96
	1.00	1.40	0.70	1.02	1.00	0.97
	1.00	1.23	0.76	1.05	1.00	0.99
	1.00	2.38	0.41	1.01	1.00	0.98
S4	(1.00)	(1.00)	(0.99)	(1.02)	(1.00)	(1.01)
	1.00	0.99	1.02	1.01	1.00	1.01
	1.00	1.31	0.75	1.02	1.00	1.00
	1.00	1.23	0.79	1.02	1.00	1.01
	1.00	1.31	0.75	1.02	1.00	1.01
	1.00	2.01	0.48	1.02	1.00	1.04
S5	(1.00)	(1.01)	(0.98)	(1.01)	(1.00)	(0.99)
	1.00	1.02	1.01	1.00	1.00	1.01
	1.00	1.25	0.75	1.00	1.00	0.96
	1.00	1.27	0.72	0.96	1.00	0.95
	1.00	1.12	0.84	1.02	1.00	0.97
	1.00	2.13	0.43	1.01	1.00	0.99
S6	(1.00)	(0.98)	(1.01)	(1.00)	(1.00)	(1.00)
	1.00	1.09	0.94	1.00	1.00	1.00
	1.00	2.65	0.37	1.01	1.00	0.98
	1.00	1.04	0.94	1.04	1.00	0.98
	1.00	1.22	0.78	1.01	1.00	0.93
	1.00	2.37	0.41	1.00	1.00	0.97
S7	(1.00)	(0.97)	(1.01)	(0.99)	(1.00)	(0.97)
	1.00	0.98	1.00	1.00	1.00	1.00
	1.00	1.18	0.84	1.00	1.00	0.97
	1.00	1.17	0.84	1.03	1.00	0.97
	1.00	1.25	0.80	1.01	1.00	0.98
	1.00	2.89	0.34	1.02	1.00	0.97
S8	(1.00)	(1.10)	(0.88)	(0.98)	(1.00)	(0.99)
	1.00	1.05	0.96	1.00	1.00	1.01
	1.00	1.28	0.76	1.02	1.00	0.93
	1.00	1.25	0.74	1.04	1.00	0.99
	1.00	1.45	0.66	1.02	1.00	1.00
	1.00	3.13	0.30	1.03	1.00	0.99
S9	(1.00)	(1.13)	(0.88)	(0.97)	(1.00)	(0.99)
	1.00	1.02	0.99	1.00	1.00	1.01
	1.00	1.29	0.77	0.99	1.00	0.97
	1.00	1.31	0.74	0.98	1.00	0.93
	1.00	1.27	0.72	1.01	1.00	0.96
	1.00	2.93	0.33	1.02	1.00	1.00
S10	(1.00)	(1.02)	(0.98)	(1.01)	(1.00)	(0.98)
	1.00	1.05	0.97	1.01	1.00	1.04
	1.00	1.22	0.82	1.00	1.00	0.97
	1.00	1.30	0.63	1.74	1.00	0.99
	1.00	1.33	0.74	1.00	1.00	0.97
	1.00	2.04	0.49	1.00	1.00	0.98

Notes. See notes of Table 2.

Table 4. Maximum likelihood estimates

	Estimated models		
	EId	Eco	EsaTS
α_2	0.59 (0.40)	0.84 (0.33)	1.12 (0.40)
α_3	0.37 (0.17)	1.40 (0.50)	0.72 (0.37)
α_4	-0.08 (0.05)	-0.39 (0.21)	-0.39 (0.39)
α_5	1.20 (0.32)	0.38 (0.15)	2.29 (0.97)
a	0.25 (0.09)	0.06 (0.13)	0.51 (0.12)
b_1	-0.14 (0.09)	-0.05 (0.09)	-0.13 (0.09)
b_2	-0.51 (0.09)	-0.32 (0.09)	-0.45 (0.09)
b_3	-0.26 (0.10)	-0.95 (0.04)	-0.26 (0.09)
b_4	0.81 (0.05)	-0.58 (0.08)	0.84 (0.05)
b_5	0.52 (0.22)	0.06 (0.05)	-0.07 (0.31)
σ_f	0.31 (0.08)	0.33 (0.10)	0.26 (0.09)
σ_1	0.94 (0.06)	0.94 (0.06)	0.94 (0.06)
σ_2	0.48 (0.04)	0.68 (0.07)	0.83 (0.06)
σ_3	0.51 (0.03)	0.21 (0.07)	0.91 (0.06)
σ_4	0.18 (0.01)	0.67 (0.04)	0.53 (0.03)
σ_5	0.01 (0.18)	0.26 (0.03)	0.52 (0.16)
σ_{ζ_2}	0.16 (0.04)	0.09 (0.07)	-
σ_{ζ_3}	0.08 (0.03)	-	-
σ_{ζ_4}	0.01 (0.01)	-	-
σ_{ζ_5}	0.05 (0.02)	-	-
β_3	-	-1.94 (0.80)	-
β_4	-	0.70 (0.32)	-
β_5	-	1.88 (0.67)	-
logL	-29.71	153.97	261.62
SNR	0.24	0.15	0.33
	0.56	0.33	0.89
	0.13	20.36	0.23
	0.01	0.10	0.01
	99	0.23	5.36

Notes. See notes of Table 2. SNR is the signal-to-noise ratio of the variance of the common factor at business cycle frequencies (3 to 8 years) to the variance of the idiosyncratic component in the adjusted series.

Table 5. Empirical forecasting analysis

	Sentim	Cars	Wages	Employ	Housing
1-period-ahead forecasts: Relative MSE					
EId	0.82	1.39	1.39	0.03	0.23
Eco	0.78	1.53	6.26	1.26	0.57
EsaTS	**0.71**	1.32	**1.32**	**0.01**	**0.21**
TSW	1.74	**1.07**	14.37	1.06	0.25
1-period-ahead forecasts: Equal accuracy tests					
EsaTS/Eid	0.023	0.056	0.032	0.001	0.209
EsaTS/Eco	0.039	0.013	0.001	0.001	0.028
EsaTS/TSW	0.001	0.026	0.001	0.003	0.140
4-period-ahead forecasts: Relative MSE					
EId	0.71	0.72	0.80	0.40	0.33
Eco	0.64	0.74	6.43	0.60	0.33
EsaTS	**0.63**	**0.71**	**0.80**	**0.38**	**0.33**
TSW	0.65	0.72	0.88	0.66	0.37
4-period-ahead forecasts: Equal accuracy tests					
EsaTS/Eid	0.077	0.338	0.428	0.082	0.354
EsaTS/Eco	0.238	0.109	0.001	0.008	0.318
EsaTS/TSW	0.234	0.184	0.196	0.007	0.473

Notes. See notes of Table 2. To compare the results across time series easily, the figures show the MSE divided by the in-sample standard deviations of each time series.