Journal of Business & Economic Statistics

ISSN: 0735-0015 (Print) 1537-2707 (Online) Journal homepage: https://www.tandfonline.com/loi/ubes20

Time Series Seasonal Adjustment Using Regularized Singular Value Decomposition

Wei Lin, Jianhua Z. Huang & Tucker McElroy

To cite this article: Wei Lin, Jianhua Z. Huang & Tucker McElroy (2018): Time Series Seasonal Adjustment Using Regularized Singular Value Decomposition, Journal of Business & Economic Statistics, DOI: 10.1080/07350015.2018.1515081

To link to this article: https://doi.org/10.1080/07350015.2018.1515081

Time Series Seasonal Adjustment Using Regularized Singular Value Decomposition

Wei LIN
School of International Trade and Economics, University of International Business and Economics, Beijing, 100029, China; International School of Economics and Management, Capital University of Economics and Business, Beijing 100070, China

Jianhua Z. HUANG
Department of Statistics, Texas A&M University, College Station, TX 77840

Tucker MCELROY
Center for Statistical Research and Methodology, U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233

We propose a new seasonal adjustment method based on the Regularized Singular Value Decomposition (RSVD) of the matrix obtained by reshaping the seasonal time series data. The method is flexible enough to capture two kinds of seasonality: the fixed seasonality that does not change over time and the time-varying seasonality that varies from one season to another. RSVD represents the time-varying seasonality by a linear combination of several seasonal patterns. The right singular vectors capture multiple seasonal patterns, and the corresponding left singular vectors capture the magnitudes of those seasonal patterns and how they change over time. By assuming the time-varying seasonal patterns change smoothly over time, the RSVD uses penalized least squares with a roughness penalty to effectively extract the left singular vectors. The proposed method applies to seasonal time-series data with a stationary or nonstationary nonseasonal component. The method also has a variant that can handle the case that an abrupt change (i.e., break) may occur in the magnitudes of seasonal patterns. Our proposed method compares favorably with the state-of-art `X-13ARIMA-SEATS` program on both simulated and real data examples.

KEY WORDS: Regularized singular value decomposition; Seasonal adjustment; X-13ARIMA-SEATS

1. INTRODUCTION

Seasonal adjustment of economic and business time-series data is of great importance in economic analysis and business decisions. Proper use of seasonal adjustment methodology removes the calendrical fluctuations from the seasonal time series, while minimizing distortions to other dynamics in the data, such as trend. Seasonally adjusted time-series data can be used to evaluate and study the present economic situation (e.g., by examining the business cycle), and therefore help policy-makers and economic agents make correct and timely decisions. Moreover, seasonally adjusted time-series data can be entered into time-series econometric models that analyze the non-seasonal dynamic relationships among economic and business variables. Findley (2005) is the most recent review article on the subject; Bell, Holan, and McElroy (2012) contains a volume of articles on recent developments in seasonality and seasonal adjustment; the monograph of Dagum and Bianconcini (2016) provides a comprehensive presentation of various seasonal adjustment methods.

Generally speaking, there are two approaches for seasonal adjustment, the model-based approach and the empirical-based approach. The model-based approach directly incorporates seasonality in the econometric model and *jointly* studies the seasonal and nonseasonal characteristics in time-series data. It can be argued that the seasonality in one economic variable can be related to other economic variables, or to the nonseasonal components within the same variable, and therefore seasonality should not be regarded as a single and isolated factor; see Lovell (1963), Sims (1974), and Bunzel and Hylleberg (1982), among others. There are many different modeling strategies of seasonal components, which can be generally categorized into several types. One modeling strategy treats seasonality as deterministic linear (nonlinear) additive (multiplicative) seasonal components; see, for example, Barsky and Miron (1989), Franses (1998), and Cai and Chen (2006). Another popular modeling strategy considers seasonality as stochastic, where seasonality can be defined as the sum of a stationary stochastic process and a deterministic process (Canova 1992), a nonstationary process with seasonal unit roots (Hylleberg et al. 1990; Osborn 1993), a periodic process in which the coefficients vary periodically with seasonal changes (Gersovitz and MacKinnon 1978; Osborn 1991; Hansen and Sargent 1993), or an unobservable component in a structural time series model (Harrison and Stevens 1976; Harvey 1990; Eiurridge and Wallis 1990; Harvey and Scott 1994; Proietti 2004). Because of the direct specification

Journal of Business & Economic Statistics
XXXX, 2019, Vol. 00, No. 0
DOI: 10.1080/07350015.2018.1515081
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/r/jbes.

and estimation of the seasonal component in an econometric model, the model-based approach is statistically more efficient than the empirical-based approach. The disadvantage of the model-based approach is that the extracted seasonal component can be sensitive to the dynamic and distributional specifications that are imposed on the econometric model.

The empirical-based approach uses ad hoc methods to extract or remove seasonality and delivers plausible empirical results with real data. One example is the X-11 method proposed by the U.S. Census Bureau (Shiskin, Young, and Musgrave 1965), which uses weighted moving averages to remove seasonality. This simple empirical method can be criticized for its inflexibility, lack of support from statistical theory, and possible distortions of those nonseasonal components in the time series, which subsequently causes misinterpretations of dynamic relationships across different time series. To correct the drawbacks of X-11, researchers have proposed various improved empirical-based methods, such as X-11-ARIMA (proposed by Dagum 1980) and X-12-ARIMA (proposed by the U.S. Census Bureau, and described in Findley et al. 1998). These improved methods pretreat the time-series data with ARIMA model to eliminate outliers and (ir)regular calendar effects, and perform forecasting and backcasting techniques to complete the data points at both ends of the time series before the weighted moving averages of X-11 are applied to remove seasonal fluctuations.

Nowadays, there may not exist a very clear-cut distinction between model-based and empirical-based approaches on seasonal adjustment, as researchers prefer to synthesize the two approaches by incorporating statistical models into empirical methods: The two state-of-the-art seasonal adjustment methods, X-12-ARIMA and TRAMO-SEATS, widely used by many national statistical agencies are of this kind. In this article, we also take this synthesized approach. Under the innocuous assumption that the nonseasonal component is stationary or nonstationary with a stochastic trend, we propose a flexible and robust seasonal adjustment method based on Regularized Singular Value Decomposition (RSVD) proposed by Huang, Shen, and Buja (2008) and Huang, Shen, and Buja (2009). Hereafter, these two papers are referred by HSB (2008) and HSB (2009) respectively. We first transform the vector of seasonal time-series data into a matrix whose rows represent periods and columns represent seasons. Then, we perform the RSVD on this matrix, the obtained right singular vectors represent seasonal patterns and left singular vectors represent the magnitudes of the seasonal patterns for different periods. RSVD applies regularization to ensure that the extracted seasonal patterns changes over time slowly. Such regularization improves stability of the extracted seasonal patterns and their magnitudes. Our new method has merits in the following aspects. First, it is flexible enough to handle both fixed and time-varying seasonality, with or without abrupt changes in seasonality. Second, it can accommodate both stationary and nonstationary stochastic nonseasonal components. Third, because the regularization parameter is fully data-driven by generalized cross-validation, it is robust and applicable to some irregular seasonal data for which popular seasonal adjustment methods may fail to deliver reasonable results.

There are similarities and differences between the Seasonal-Trend decomposition procedure based on Regression (STR) approach proposed by Dokumentov and Hyndman (2015) and our RSVD method in modeling seasonality. Essentially, both STR and RSVD methods introduce Tikhonov regularization terms that are motivated by the smoothness feature of the seasonal component. The RSVD method considers the singular value decomposition of the seasonal matrix and only imposes roughness penalties on the left singular vectors that capture the variation of seasonal patterns (i.e., corresponding right singular vectors) across consecutive seasonal cycles. In contrast, the STR method directly imposes roughness penalties on seasonal terms (or the coefficients in the linear combination of spline basis functions that approximate the seasonal terms) across consecutive seasonal cycles and/or within each seasonal cycle. However, compared to the STR method, the main advantage of our RSVD method is that, with the merit of dimension reduction due to a low rank approximation, the parameterization of our method is much more parsimonious than that of the STR method, which directly estimates each seasonal term. Moreover, the RSVD method decomposes the seasonal component into fixed and time-varying seasonal patterns, which can provide much rich information about the complexity and composition of seasonality.

In this article, using both simulated and real economic data, we also compare our proposed seasonal adjustment methods with two state-of-the-art and widely used seasonal adjustment methods (X-12-ARIMA and SEATS) provided in the latest `X-13ARIMA-SEATS` program developed by U.S. Census Bureau. We find that (i) when seasonality is moderate or weak, traditional X-12-ARIMA and SEATS methods tend to outperform our proposed seasonal adjustment method, which is especially the case if the seasonality is weak; (ii) however, in comparison to X-12-ARIMA and SEATS methods, our proposed seasonal adjustment method is good at capturing strong seasonal variations in the series; (iii) our proposed method is robust to some irregular seasonal data for which X-12-ARIMA and SEATS may need additional delicate performance tuning. Moreover, compared to X-12-ARIMA and SEATS, our proposed method provides a more transparent and meaningful explanation for seasonality. Our proposed method decomposes the seasonal component into different seasonal patterns, traces the dynamics of seasonality by time-varying pattern coefficients, and identifies important seasonality break times automatically, which provides rich insights into seasonality.

The remaining part of this article is organized as follows. Section 2 briefly reviews the RSVD. Section 3 introduces some notations for the matrix representation of seasonal time series. Section 4 gives our basic seasonal adjustment method when nonseasonal component is stationary or difference is stationary. Section 5 extends our basic seasonal adjustment method to accommodate stochastic trend and abrupt changes in seasonality. Simulation results under different data generating processes (DGPs) are reported in Section 6, and three real data examples are provided in Section 7. Section 8 concludes. Due to the space limitation, some technical details, additional simulation results, and further discussions on some important issues concerning our proposed RSVD seasonal adjustment method are provided in the Online Appendices.

2. A BRIEF REVIEW OF REGULARIZED SVD

As a well-known matrix factorization technique, the singular value decomposition (SVD) has been widely used in tackling many real practical problems. In the context of latent semantic analysis, Deerwester et al. (1990) propose a new approach to automatic indexing and retrieval using the SVD method. Sarwar et al. (2000) make use of SVD for recommender systems that make product recommendations during a live customer interaction. More recently and publicly known, Bell et al. (2009), the \$1M Grand Prize winner of the Netflix Prize contest, employ SVD in the challenge of predicting user preferences.

Regularized singular valued decomposition (RSVD) is a variant of singular value decomposition that takes into account the intrinsic smoothness structure of a data matrix (HSB, 2008, 2009). The basic idea of RSVD is quite intuitive. The data matrix is considered as discretized values of a bivariate function with certain smoothness structure evaluated at a grid of design points. To impose smoothness in singular value decomposition, RSVD imposes roughness penalties on the left and/or right singular vectors when singular value decomposition is implemented on the data matrix.

Consider an $n \times p$ dimensional data matrix $\mathbf{X} = (x_{ij})$ whose column mean is zero. The first pair of singular vectors, $\mathbf{u}$ and $\mathbf{v}$, respectively, solves the following minimization problem:

$$(\widehat{\mathbf{u}}, \widehat{\mathbf{v}}) = \underset{\mathbf{u},\mathbf{v}}{\arg\min} \|\mathbf{X} - \mathbf{u}\mathbf{v}^\top\|_F^2, \quad (1)$$

which does not assume any smoothness structure of the data matrix. In contrast, RSVD explores such smoothness structure by imposing roughness penalties on singular vectors $\mathbf{u}$ and $\mathbf{v}$. In the context of seasonal adjustment, the seasonal time series can be represented as a matrix whose each row represents one period of all seasons. We later argue that the data matrix should have smooth changes across rows, and thus the changes in left singular vector $\mathbf{u}$ are expected to be smooth. Therefore, a relevant RSVD solves the following minimization problem:

$$(\widehat{\mathbf{u}}, \widehat{\mathbf{v}}) = \underset{\mathbf{u},\mathbf{v}}{\arg\min} \|\mathbf{X} - \mathbf{u}\mathbf{v}^\top\|_F^2 + \alpha \mathbf{u}^\top \boldsymbol{\Omega} \mathbf{u}, \quad (2)$$

where $\boldsymbol{\Omega}$ is an $n \times n$ nonnegative definite roughness penalty matrix, α is smoothing parameter, and $\mathbf{v}^\top\mathbf{v} = 1$ for identification purpose.

A simple variant of the power algorithms in HSB (2008, 2009) gives the following Algorithm 1 for solving the problem (2):

Algorithm 1 (Regularized singular value decomposition of $\mathbf{X}$*).*
Step 1. Initialize $\mathbf{u}$ using the standard SVD for $\mathbf{X}$.
Step 2. Repeat until convergence:

1. $\mathbf{v} \leftarrow \dfrac{\mathbf{X}^\top\mathbf{u}}{\|\mathbf{X}^\top\mathbf{u}\|}$.
2. $\mathbf{u} \leftarrow (\mathbf{I}_n + \alpha\boldsymbol{\Omega})^{-1}\mathbf{X}\mathbf{v}$ with α selected by minimizing the following generalized cross-validation criterion:

$$\mathrm{GCV}(\alpha) = \frac{1}{n} \frac{\|[\mathbf{I}_n - \mathbf{M}(\alpha)]\mathbf{X}\mathbf{v}\|^2}{\left(1 - \frac{1}{n}\mathrm{tr}\{\mathbf{M}(\alpha)\}\right)^2}, \quad (3)$$

where $\mathbf{I}_n$ is the $n \times n$ identity matrix, and $\mathbf{M}(\alpha) = (\mathbf{I}_n + \alpha\boldsymbol{\Omega})^{-1}$ is the smoothing matrix.

The derivation of the generalized cross-validation criterion used in (3) is similar to HSB (2008, 2009) and can be found in the Online Appendices. The difference of Algorithm 1 from previous algorithms is that, in HSB (2008) the roughness penalty is imposed only on $\mathbf{v}$ and in HSB (2009) on both $\mathbf{u}$ and $\mathbf{v}$. If there is no penalty, that is, $\alpha = 0$, the algorithm is essentially the power algorithm for standard SVD and solves the problem (1).

In general, the regularized SVD attempts to find a rank-r decomposition ($r \le p$) such that $\mathbf{X} = \mathbf{U}\mathbf{V}^\top$, where $\mathbf{U}$ is a $n \times r$ matrix, and $\mathbf{V}$ is a $p \times r$ matrix. The jth column in matrix $\mathbf{U}$ and $\mathbf{V}$ is called the jth left and right regularized singular vector of matrix $\mathbf{X}$, respectively. Algorithm 1 finds the first regularized singular vector pair. The subsequent regularized singular vector pairs can be obtained by repeatedly applying Algorithm 1 to the residual matrix $\mathbf{X} - \widehat{\mathbf{u}\mathbf{v}}^\top$. Below, we propose some variants of Algorithm 1 for different scenarios of seasonal adjustment.

3. MATRIX REPRESENTATION OF SEASONAL TIME SERIES

For a seasonal time series $\{x_t : t = 1 \cdots, T\}$ with p seasons, we can represent it (see Buys and Ballot 1847) as a matrix with p columns, whose each row represents one period of the seasons, as follows:

$$\mathbf{X} = \left[\mathbf{x}_{1\cdot}^\top\ \mathbf{x}_{2\cdot}^\top\ \ldots\ \mathbf{x}_{i\cdot}^\top\ \ldots\ \mathbf{x}_{n\cdot}^\top\right]^\top,$$

where the $1 \times p$ row vector $\mathbf{x}_{i\cdot}$ denotes the ith row of matrix $\mathbf{X}$. Hence, the $T \times 1$ column vector form of time-series x_t can be written as

$$\begin{aligned} X_T \equiv \mathrm{Vec}(\mathbf{X}^\top) &= (x_1, \cdots, x_t, \cdots, x_T)^\top \\ &= (\mathbf{x}_{1\cdot}, \cdots, \mathbf{x}_{i\cdot}, \cdots, \mathbf{x}_{n\cdot})^\top, \end{aligned}$$

where the function $\mathrm{Vec}(\cdot)$ converts a matrix into a column vector by stacking the columns of the matrix. The subscripts of the elements in the matrix representation can be obtained using a mapping of the one-dimensional time subscript $t \in \mathbb{N}$ to the two-dimensional time subscripts, $(i, j) \in \mathbb{N}^2$, denoting the jth *season* in the ith *period*,

$$\begin{aligned} I : \mathbb{N} &\mapsto \mathbb{N}^2 \\ t &\to (i(t), j(t)) \equiv (\lceil t/p \rceil, t - \lfloor t/p \rfloor p). \end{aligned} \quad (4)$$

Let $n \equiv T/p$ denote the total number of time span included in the time series, so that we have that $1 \le i \le n$, $1 \le j \le p$, and $t = (i(t) - 1)p + j(t)$. Here, we assume that T/p is an integer for simplicity of exposition.

For later use of notations, let $\mathbf{i}_p$ and $\mathbf{0}_p$ denote the $p \times 1$ column vector of ones and zeros, respectively. Moreover, let

$\mathbf{Q}_n$ denote the n-dimensional column-wise de-meaning matrix, that is, $\mathbf{Q}_n \equiv \mathbf{I}_n - \mathbf{i}_n\mathbf{i}_n^\top/n$ so that $\mathbf{Q}_n\mathbf{a} = \mathbf{a} - \bar{\mathbf{a}}$, for a vector $\mathbf{a} = (a_1, \ldots, a_n)^\top$, where $\bar{\mathbf{a}} = \bar{a}\,\mathbf{i}_n$, and $\bar{a} = \sum_{1\le i\le n} a_i/n$. Let the $(d-1)\times d$ matrix Δ_d be the first-order difference operator, that is,

$$\Delta_d \equiv \left[\mathbf{0}_{d-1}\ \mathbf{I}_{d-1}\right] - \left[\mathbf{I}_{d-1}\ \mathbf{0}_{d-1}\right].$$

Then the second-order difference operator is $\Delta_d^2 \equiv \Delta_{d-1}\Delta_d$. Using these difference operators, one widely used choice of the penalty matrix in (2) can take the form $\mathbf{\Omega} \equiv (\Delta_n^2)^\top\Delta_n^2$.

4. BASIC SEASONAL ADJUSTMENT

This section discusses seasonal adjustment based on regularized SVD. We motivate the use of regularized SVD on seasonal adjustment in Section 4.1, and then propose the seasonal adjustment procedures when the nonseasonal component of a time series is stationary or difference stationary in Section 4.2 and 4.3 respectively. Section 4.4 introduces how to select the number of seasonal patterns in the RSVD method.

4.1. Motivation of Using Regularized SVD for Seasonal Adjustment

We decompose the seasonal time series $\{x_t\}_{t=1}^T$ into the *deterministic seasonal component* s_t and *stochastic nonseasonal component* e_t in the additive form,

$$x_t = s_t + e_t, \quad t = 1, \ldots, T, \tag{5}$$

where the nonseasonal component e_t is a stationary process. Using the mapping I defined in (4), we rewrite (5) as $x_{i,j} = s_{ij} + e_{i,j}$, where the seasonal component satisfies $\sum_{j=1}^p s_{i,j} = 0$ for identification. The decomposition can also be written in matrix from,

$$\mathbf{X} = \mathbf{S} + \mathbf{E}. \tag{6}$$

When the seasonal effects are fixed, that is, the seasonal pattern does not change from period to period, $s_t = f_{j(t)}$, the seasonal component $\mathbf{S}$ can be represented as $\mathbf{S} = \mathbf{i}_n \cdot \mathbf{f}^\top$. In this case, a single seasonal pattern $\mathbf{f}^\top = (f_1, f_2, \cdots, f_p)$ repeats itself in each period. In general, the seasonal effects may change over time, we use a rank-r reduced SVD of $(\mathbf{S} - \mathbf{i}_n \cdot \mathbf{f}^\top)$ to represent the time-varying seasonality:

$$\mathbf{S} = \mathbf{i}_n \cdot \mathbf{f}^\top + \mathbf{U}\mathbf{V}^\top, \tag{7}$$

where $\mathbf{U}$ is an $n \times r$ matrix, and $\mathbf{V}$ is a $p \times r$ matrix with $\mathbf{V}^\top\mathbf{V} = \mathbf{I}_r$ and $r \le p$. For identification, we require the columns of $\mathbf{U}$ to be orthogonal to $\mathbf{i}_n$ or, $\mathbf{U}^\top\mathbf{i}_n = \mathbf{0}$, which is equivalent to $\mathbf{Q}_n^\top\mathbf{U} = \mathbf{U}$. The second term in the decomposition (7) provides an intuitive explanation for the seasonality. The jth column vector $\mathbf{v}_j$ in $\mathbf{V}$ represents the jth *seasonal pattern*; and the corresponding jth column vector $\mathbf{u}_j$ in $\mathbf{U}$ is called *pattern coefficients*, since its elements delineate how the jth seasonal pattern changes across different periods. Equations (6) and (7) comprise our basic seasonal adjustment method. For example, in a special case of $r = 1$ with $\mathbf{f} = \mathbf{0}$, the seasonal matrix is $\mathbf{S} = \mathbf{u}\mathbf{v}^\top$ with $\mathbf{v}$ consisting of the seasonal pattern and $\mathbf{u}$ giving its time evolution across different periods.

Now, we argue that there is intrinsic smoothness in the seasonal signal that warrants using the regularized SVD. For notational simplicity, assume the fixed seasonality term is void. The ith row of $\mathbf{S}$, denoted by $\mathbf{s}_i$, represents the seasonal behavior of series x_t during the ith period, which is a linear combination of all the seasonal patterns in $\mathbf{V}$ with the ith row of $\mathbf{U}$ as the coefficients, that is,

$$\mathbf{s}_i = \mathbf{u}_i\mathbf{V}^\top = \sum_{j=1}^p u_{i,j}\mathbf{v}_j.$$

A necessary condition for seasonality is persistence of a seasonal pattern from one year to the next; for a stochastic approach, persistence is assessed through correlation, whereas in a deterministic context the concept of smoothness is used instead. Essentially, seasonality imposes that the $\mathbf{u}_i$'s, or, $u_{i,j}$'s for fixed j (i.e., the elements in each column of matrix $\mathbf{U}$) *change smoothly* with i. Based on this smoothness on the decomposition of seasonal matrix $\mathbf{S}$, we deem that the roughness of each column in the observed data matrix $\mathbf{X}$ is due to the "contamination" of the stochastic nonseasonal component $\mathbf{E}$ in (6). This smoothness also suggests the use of regularized SVD for finding the decomposition (6) with a roughness penalty applied on the columns of $\mathbf{U}$. On the other hand, it is usually not appropriate to apply a roughness penalty on the columns of $\mathbf{V}$, since seasonal behaviors usually have sharp increases and falls *within* a period.

In sum, any seasonal matrix $\mathbf{S}$ can be decomposed uniquely into the SVD form in (7) with some $r \le p$ according to matrix theory, which implies that any seasonal component is driven by at most a fixed seasonal patterns $\mathbf{f}$ and the r time-varying seasonal patterns in the columns of $\mathbf{V}$. Given the smoothness feature of seasonality, the pattern coefficients in the columns of $\mathbf{U}$ should be smooth over time. The regularization with roughness penalty effectively separates the seasonal variations in $\mathbf{S}$ from the irregular component in $\mathbf{E}$. Based on a selection criterion, we select those significant seasonal patterns from the data matrix $\mathbf{X}$ that drives the seasonal behavior and discard those indiscernible seasonal patterns that are submerged in noise. Therefore, our RSVD method should be good at capturing seasonality that has strong variations compared to the irregular component.

There are three reasons that prevent direct application of Algorithm 1 in HSB (2008, 2009) to the data matrix $\mathbf{X}$ for seasonal adjustment. First, because of the existence of fixed seasonality, it is unrealistic to restrict the sample mean of each column of $\mathbf{X}$ to zero, that is, to simply subtract the mean from each column. Instead, the fixed seasonality $\mathbf{f}$ should be explicitly estimated in the seasonal adjustment procedure. Second, for identification, the sum of seasonal terms within a period should be zero, that is, $\sum_{j=1}^p s_{i,j} = 0$ for each $i = 1, \cdots, n$. Otherwise, the seasonal component would incorporate part of the overall level of the series. Third, if the nonseasonal component $\{e_t\}$ is nonstationary and has a stochastic trend which is more commonly encountered in economic time series data, Algorithm 1, assuming stationarity in $\{e_t\}$, is invalid. Next, taking all these into account, we develop a procedure that is based on a modification of Algorithm 1.

4.2. Seasonal Adjustment With Stationary e_t

Our basic seasonal adjustment procedure has three steps: 1. Estimate the seasonal pattern coefficients in $\mathbf{U}$ using a modified version of Algorithm 1 to satisfy the zero-sum restriction on seasonal effects; 2. Estimate the fixed seasonal pattern $\mathbf{f}$ and the time-varying seasonal patterns in $\mathbf{V}$; 3 (optional). Estimate the parameters of the stationary nonseasonal components. The three steps are elaborated below.

Step One: Estimating Seasonal Pattern Coefficients in $\mathbf{U}$.

To apply Algorithm 1, we first eliminate the fixed seasonal effects in (7) by pre-multiplying the column-wise demeaning matrix $\mathbf{Q}_n$ to data matrix $\mathbf{X}$ to obtain $\widetilde{\mathbf{X}} = \mathbf{Q}_n\mathbf{X}$. Since $\mathbf{Q}_n\mathbf{i}_n = \mathbf{0}_n$ and $\mathbf{Q}_n\mathbf{U} = \mathbf{U}$,

$$\widetilde{\mathbf{X}} = \mathbf{Q}_n\mathbf{S} + \mathbf{Q}_n\mathbf{E} = \mathbf{Q}_n(\mathbf{i}_n \cdot \mathbf{f}^\top + \mathbf{U}\mathbf{V}^\top) + \mathbf{Q}_n\mathbf{E} = \mathbf{U}\mathbf{V}^\top + \widetilde{\mathbf{E}}.$$

The resulting column-centered data matrix $\widetilde{\mathbf{X}}$ does not have a fixed seasonality. To guarantee the *zero-sum seasonal effects* requirement $\mathbf{S} \cdot \mathbf{i}_p = \mathbf{0}_n$, we enforce the sufficient conditions of *zero-sum seasonal patterns*, $\mathbf{f}^\top\mathbf{i}_p = 0$ and $\mathbf{V}^\top\mathbf{i}_p = \mathbf{0}_r$. Combining above gives the following modified version of Algorithm 1.

Algorithm 2.
It is the same as Algorithm 1 except that
1. the data matrix $\mathbf{X}$ is replaced by $\widetilde{\mathbf{X}} = \mathbf{Q}_n\mathbf{X}$, and
2. the updating equation in Step 2.1 now becomes

$$\mathbf{v} \leftarrow \frac{\mathbf{Q}_p\widetilde{\mathbf{X}}^\top\mathbf{u}}{\|\mathbf{Q}_p\widetilde{\mathbf{X}}^\top\mathbf{u}\|}.$$

In Step 2.1 of this algorithm, premultiplication with $\mathbf{Q}_p$ ensures $\mathbf{v}^\top\mathbf{i}_p = 0$ for the zero-sum of seasonal pattern requirement, and the normalization is to ensure $\mathbf{v}^\top\mathbf{v} = 1$ for identification.

Applying Algorithm 2 we obtain the first pair of estimated right singular vector, denoted as $\widetilde{\mathbf{v}}$, and the left singular vector, denoted as $\widehat{\mathbf{u}}$. The subsequent pair of singular vectors can be extracted by applying Algorithm 2 to the residual matrix $\widetilde{\mathbf{X}} - \widehat{\mathbf{u}}\widetilde{\mathbf{v}}^\top$, in which the preceding effect of the first pair of singular vectors is subtracted from data matrix $\widetilde{\mathbf{X}}$. Applying this procedure r times sequentially, we obtain r pairs of regularized singular vectors, concatenating them into the $n \times r$ matrix $\widehat{\mathbf{U}} = (\widehat{\mathbf{u}}_1, \cdots, \widehat{\mathbf{u}}_r)$ and the $p \times r$ matrix $\widetilde{\mathbf{V}} = (\widetilde{\mathbf{v}}_1, \cdots, \widetilde{\mathbf{v}}_r)$. We keep $\widehat{\mathbf{U}}$ for use in the next step.

Step Two: Estimating Fixed/Time-Varying Seasonal Patterns in $\mathbf{f}$ *and* $\mathbf{V}$.

Recall that $X_T = \text{Vec}(\mathbf{X}^\top)$. Given the estimates of seasonal pattern coefficients $\widehat{\mathbf{U}}$ in Step One, and that the pattern coefficients of fixed seasonal pattern $\mathbf{f}$ all take value 1, the estimates of the time varying seasonal patterns in $\mathbf{V}$ and fixed seasonal pattern $\mathbf{f}$ can be obtained jointly by solving a constrained least squares problem,

$$(\widehat{\mathbf{f}}, \widehat{\mathbf{V}}) = \arg\min_{\mathbf{f},\mathbf{V}} \left[X_T - \text{Vec}(\mathbf{f} \cdot \mathbf{i}_n^\top + \mathbf{V}\widehat{\mathbf{U}}^\top)\right]^\top \times \left[X_T - \text{Vec}(\mathbf{f} \cdot \mathbf{i}_n^\top + \mathbf{V}\widehat{\mathbf{U}}^\top)\right], \tag{8}$$

$$\text{such that } \mathbf{f}^\top \cdot \mathbf{i}_p = 0, \quad \text{and} \quad \mathbf{V}^\top \cdot \mathbf{i}_p = \mathbf{0}_r.$$

Note that the minimization problem in (8) can be rewritten as,

$$\widehat{\beta} = \arg\min_{\beta}(X_T - \mathbf{Z}\beta)^\top(X_T - \mathbf{Z}\beta) \quad \text{with} \quad \mathbf{R}\beta = \mathbf{0}_{r+1}, \tag{9}$$

where $\mathbf{Z} \equiv [\mathbf{i}_n \otimes \mathbf{I}_p, \widehat{\mathbf{u}}_1 \otimes \mathbf{I}_p, \cdots, \widehat{\mathbf{u}}_r \otimes \mathbf{I}_p]$, $\beta \equiv (\mathbf{f}^\top, \mathbf{v}_1^\top, \cdots, \mathbf{v}_r^\top)^\top$, and $\mathbf{R} \equiv \mathbf{I}_{r+1} \otimes \mathbf{i}_p^\top$. Then, the estimate $\widehat{\beta}$ can be written explicitly as,

$$\widehat{\beta} \equiv (\widehat{\mathbf{f}}^\top, \widehat{\mathbf{v}}_1^\top, \cdots, \widehat{\mathbf{v}}_r^\top)^\top = \mathbf{b} - (\mathbf{Z}^\top\mathbf{Z})^{-1}\mathbf{R}^\top[\mathbf{R}(\mathbf{Z}^\top\mathbf{Z})^{-1}\mathbf{R}^\top]^{-1}\mathbf{R}\mathbf{b},$$

where $\mathbf{b} = (\mathbf{Z}^\top\mathbf{Z})^{-1}\mathbf{Z}^\top X_T$ is the unconstrained least-squares estimate for the problem (9).

Given the estimates of fixed and time-varying seasonal patterns in $\widehat{\mathbf{f}}$ and $\widehat{\mathbf{V}}$ obtained from the constrained least squares regression, we obtain the estimated seasonal component as,

$$\widehat{\mathbf{S}} = \mathbf{i}_n\widehat{\mathbf{f}}^\top + \widehat{\mathbf{U}}\widehat{\mathbf{V}}^\top.$$

Step Three (optional): Estimating ARMA Parameters in Nonseasonal Component $\mathbf{E}$.

In the second step, we obtain the estimated seasonal matrix $\widehat{\mathbf{S}}$, which can be rewritten in vector form as $\{\widehat{s}_t\}_{t=1}^T$. Correspondingly, the estimated nonseasonal component can be extracted by subtracting $\widehat{s}_t$ from the original time series x_t, that is, $\widehat{e}_t \equiv x_t - \widehat{s}_t$. If the stochastic component of x_t is assumed to follow a stationary ARMA(p, q) process, that is, $e_t \sim$ ARMA(p, q) for $t = 1, \ldots, T$, the ARMA parameters can then be obtained by fitting the ARMA model to $\widehat{e}_t$. (This is a “nuisance” model; other stationary models could be used without affecting the methodology.)

Based on the fitted ARMA model, a feasible GLS estimation can be obtained by weighting the least squares in (8) with the inverse of estimated variance–covariance matrix of the stochastic nonseasonal component $\widehat{e}_t$. Although such an iterated procedure could potentially improve estimation accuracy, we find (using a simulation study) that the efficiency gain of the feasible GLS in terms of reductions in AMSE and AMPE for estimating the seasonal component is only marginal (around 2%) even when the first-order autocorrelation of e_t reaches 0.8. Thus, in general, we recommend using the unweighted ordinary least squares estimation in (8) instead of GLS unless the nonseasonal component e_t exhibits very strong persistence. This also has the benefit of avoiding the additional computation burden. Moreover, as the ARMA model is usually not of particular interest for seasonal adjustment, Step Three can be omitted from the procedure.

4.3. Seasonal Adjustment With Difference Stationary e_t

The basic seasonal adjustment assumes that the nonseasonal component of a seasonal time series is stationary. This section discusses the situation more commonly encountered in economic data, wherein the nonseasonal component is nonstationary and has a stochastic trend. More specifically, we assume the nonseasonal component e_t in the decomposition (5) $x_t = s_t + e_t$ is an integrated process, that is, the first difference process of e_t is stationary.

Existence of a stochastic trend in each column of $\mathbf{E}$ invalidates the use of regularized SVD in the basic adjustment procedure, the direct use of which may produce an inconsistent estimate of the seasonal component S. As examination of (6) indicates, when the nonseasonal component is stationary, that there is no clear smooth pattern in each column of $\mathbf{E}$, while the seasonal component changes smoothly in each column of $\mathbf{S}$; therefore the regularized SVD can separate $\mathbf{S}$ from $\mathbf{E}$. However, if there is a stochastic trend in $\mathbf{E}$, each column of $\mathbf{E}$ has a stochastic trend and thus is quite smooth. Intuitively, the trend smoothness in $\mathbf{E}$ "contaminates" the seasonality smoothness in $\mathbf{S}$. Thus, the basic regularized SVD in Section 4 will fail to separate $\mathbf{S}$ from $\mathbf{E}$, resulting in inconsistent estimates of smooth pattern coefficients $\widehat{\mathbf{u}}$'s. Moreover, the nonstationarity of the nonseasonal component also invalidates the use of least squares for estimating seasonal patterns $\widehat{\mathbf{v}}$'s in Step Two of the basic adjustment procedure in Section 4.2.

Our new procedure is a modification of the basic procedure to address these issues. It also has three steps, as elaborated below:

Step One: Estimating Seasonal Pattern Coefficients in $\mathbf{U}$.

We first remove the stochastic trend in the nonseasonal component and then apply the regularized SVD. To this end, we take the first-order column difference of matrix $\mathbf{X}$. This differencing removes the stochastic trend in $\mathbf{E}$ but will not change the column-wise smoothness of the seasonal component matrix $\mathbf{S}$. In matrix form, we postmultiply Equation (6) by $\Delta_p^\top$ and obtain

$$\mathbf{X} \equiv \mathbf{X}\Delta_p^\top = \mathbf{S}\Delta_p^\top + \mathbf{E}\Delta_p^\top \equiv \mathbf{S} + \mathbf{E}.$$

As in the basic adjustment procedure, we represent the seasonal component matrix using a reduced SVD as in (7), that is, $\mathbf{S} = \mathbf{i}_n\mathbf{f}^\top + \mathbf{U}\mathbf{V}^\top$. Then,

$$\mathbf{S} = \mathbf{S}\Delta_p^\top = (\mathbf{i}_n\mathbf{f}^\top + \mathbf{U}\mathbf{V}^\top)\Delta_p^\top \equiv \mathbf{i}_n\mathbf{f}^\top + \mathbf{U}\mathbf{V}^\top. \tag{10}$$

Equations (7) and (10) show that the seasonal matrix $\mathbf{S}$ and its first-order column-differenced matrix $\mathbf{S}$ share the same left singular matrix $\mathbf{U}$, as the first-order differencing operates from the right side of the matrix. The first-order column-difference on $\mathbf{E}$ removes the nonstationary trend in ARIMA$(p, 1, q)$, so that $\mathbf{E}$ is weakly stationary.

We eliminate the fixed seasonal effects in (10) by premultiplying the column-wise demeaning matrix $\mathbf{Q}_n$ by the first-order column-differenced data matrix $\mathbf{X}$ to obtain $\widetilde{\mathbf{X}} = \mathbf{Q}_n\mathbf{X}$. We repeatedly apply Algorithm 1 r times to the matrix $\widetilde{\mathbf{X}}$ (or the residual matrices) to sequentially extract the regularized left singular vectors. Here, unlike in the basic procedure of the previous section, there is no need to enforce the zero-sum seasonal effects requirement on the right singular vectors, since we are working on the column-differenced data matrix. Denote the so-extracted $\mathbf{U}$ matrix as $\widehat{\mathbf{U}}$, for use in Step Two.

Step Two: Estimating Fixed/Time-Varying Seasonal Patterns in $\mathbf{f}$ *and* $\mathbf{V}$.

Given the estimated left singular vectors in $\widehat{\mathbf{U}}$, we estimate the fixed and time-varying seasonal patterns in $\mathbf{f}$ and $\mathbf{V}$ jointly by solving a constrained least-squares problem. In contrast to the basic adjustment procedure, we need to work with the differenced series to remove the effect of nonstationarity.

Let ΔX_T denote the first difference of $X_T = \text{Vec}(\mathbf{X}^\top)$, where Δ is the differencing operator. The constrained least-squares problem is similar to that in (8) and can also be written as

$$\widehat{\beta} = \arg\min_{\beta}(\Delta X_T - \Delta\mathbf{Z}\beta)^\top(\Delta X_T - \Delta\mathbf{Z}\beta) \tag{11}$$
$$\text{with} \quad \mathbf{R}\beta = \mathbf{0}_{r+1},$$

where $\mathbf{Z}$, β, and $\mathbf{R}$ are defined in the same manner as (9). After solving the constrained least-squares problem above, we obtain the estimated seasonal component as $\widehat{\mathbf{S}} = \mathbf{i}_n\widehat{\mathbf{f}}^\top + \widehat{\mathbf{U}}\widehat{\mathbf{V}}^\top$.

Step Three (optional): Estimating Parameters in Nonseasonal Component $\mathbf{E}$.

If we assume that the nonseasonal component x_t follows the dynamics of an ARIMA$(p, 1, q)$ process, then the ARIMA parameters can be obtained by fitting an ARIMA model to the residual series $\widehat{e}_t = x_t - \widehat{s}_t$. Then, a feasible GLS estimation of $\mathbf{f}$ and $\mathbf{V}$ can be obtained by weighting the least squares in (11) with the inverse of the estimated variance-covariance matrix of the differenced non-seasonal component $\Delta\widehat{e}_t$. As we discussed in the description of the basic seasonal adjustment procedure, this step is usually not necessary.

4.4. Selecting the Number of Seasonal Patterns

We propose to select the number of seasonal patterns r by the following information criteria. For each seasonal time series, the number of period within each season p is fixed, and the total number of seasons n increases as the total number of observations $T = np$ increases. We use standard Bayesian Information Criterion (BIC) in time-series applications, in which the penalty for overfitting $(\log n)/n$ only involves n. If the nonseasonal component of the seasonal time series is stationary, the information criterion is,

$$\text{BIC}(r) = \ln\left[\frac{1}{T}\sum_{t=1}^{T}(x_t - \widehat{s}_t)^2\right] + r\frac{\log n}{n}; \tag{12}$$

if the nonseasonal component of the seasonal time series is nonstationary, the information criterion is,

$$\text{BIC}(r) = \ln\left[\frac{1}{T-1}\sum_{t=2}^{T}(\Delta x_t - \Delta\widehat{s}_t)^2\right] + r\frac{\log n}{n}; \tag{13}$$

where n is the total number of seasons, $\{x_t\}$ is the original seasonal time series, $\{\widehat{s}_t\}$ is the estimated seasonal component, and Δ is first-order difference operator.

5. SEASONAL ADJUSTMENT WHEN THERE IS ABRUPT CHANGE TO SEASONALITY

In previous discussions, it is assumed that the elements in each column of matrix $\mathbf{U}$ (representing the magnitude of a seasonal pattern in one period) changes smoothly across periods. This has two implications. First, the magnitude of each seasonal pattern only changes in a smooth fashion. Second, all seasonal patterns appear in all periods. In reality, these assumptions may be violated due to sudden changes of statistic criteria (such as sampling method and scope) or social economic environment (such as economic policies and enforcement of laws affecting behavior). Hence, seasonal patterns do not necessarily prevail all the time in a time series: some seasonal patterns may transiently exist with nonzero magnitudes, and abruptly vanish. Moreover, the change in magnitude of seasonal pattern does not necessarily have the same "smoothness" across all time spans: the magnitudes of seasonality may present mild changes for early periods, and then have sharp changes in other periods. This section discusses how to perform seasonal adjustment to handle these complicated scenarios.

To address abrupt changes (also referred to as breaks or change points) in seasonality, our method is a modification of procedures presented in the previous two sections. We take the procedure from Section 4.3 as an example to show how to modify it. The basic seasonal adjustment procedure in Section 4.2 can be modified in a similar manner.

In Section 4.3, the seasonal adjustment procedure has three steps. To handle the abrupt seasonality change, we only need to modify Step One. It is sufficient to allow for at most one abrupt change for each seasonal pattern but the timing of break may be different for each seasonal pattern. Step One of the previous procedure is based on $\widetilde{\mathbf{X}} = \mathbf{U}\mathbf{V}^\top + \widetilde{\mathbf{E}}$. By assuming each column of $\mathbf{U}$ is smooth, the previous procedure extracts the columns of $\mathbf{U}$ using the regularized SVD. Since the columns of $\mathbf{U}$ are sequentially extracted, we only need to discuss how to modify the procedure for one column of $\mathbf{U}$, denoted as $\mathbf{u}$, corresponding to a seasonal pattern $\mathbf{v}$.

Now, suppose a nonsmooth change of seasonality happens after ℓ seasonal periods and $\ell = 0$ if there is no break. The period index ℓ separates the entire time span into two portions: one part starts from the beginning and ends at period ℓ, and the second part contains the rest. If we know ℓ, the timing of the break, we can apply the following modification of Algorithm 1 to extract $\mathbf{u}$. Since the change point naturally separates $\mathbf{u}$ into two parts $\mathbf{u}_1$ and $\mathbf{u}_2$, the modified algorithm updates these two parts separately using different smoothing parameters.

Algorithm 3.
It is the same as Algorithm 1 except that
1. the data matrix $\mathbf{X}$ is replaced by $\widetilde{\mathbf{X}} = \mathbf{Q}_n\mathbf{X}\Delta_p^\top$, and
2. the updating equation in Step 2.2 now becomes two equations that update $\mathbf{u}_1$ and $\mathbf{u}_2$ separately by applying Step 2.2 of the original algorithm to the first ℓ rows and the last $n-\ell$ rows of $\widetilde{\mathbf{X}}$ respectively.

In Algorithm 3, using different smoothing parameters for the first ℓ elements and last $n-\ell$ elements of $\mathbf{u}$ enhances the flexibility of the procedure to handle abrupt changes in seasonal behaviors across time spans. After applying this algorithm r times to sequentially extract the columns of $\widehat{\mathbf{U}}$, Step Two of the procedure in Section 4.3 can be used to obtain $\widehat{\mathbf{f}}$, $\widehat{\mathbf{V}}$. Including the dependence on ℓ in our notation, we can obtain the estimated seasonal component matrix $\widehat{\mathbf{S}}(\ell) = \mathbf{i}_n\widehat{\mathbf{f}}(\ell)^\top + \widehat{\mathbf{U}}(\ell)\widehat{\mathbf{V}}(\ell)^\top$.

In practice, we don't know ℓ and so we need to specify it using data. Since the roughness penalty involves second-order differencing, we have $3 \le \ell \le n-3$. Including the no-break case of $\ell = 0$, there are $(n-5+1)$ possible values of ℓ for each seasonal pattern. For r seasonal patterns, the set of all configurations of breaks is $\mathcal{L} = \{\boldsymbol{\ell} = (\ell_1, \ldots, \ell_r)\}$, and the total number of all possible configurations is $\#(\mathcal{L}) = (n-5+1)^r$. When n is large, $\#(\mathcal{L})$ can be so large that exhaustive search for the optimal breaks impractical due to intensive computational burden. When applying RSVDB method to real data, we set a maximal number of seasonal patterns $r_{\max}$ to alleviate this problem.

Next, we discuss how to specify the timing of the breaks. We select the optimal specification of the change points $\widehat{\boldsymbol{\ell}}$ by minimizing the following criterion:

$$\widehat{\boldsymbol{\ell}} = \arg\min_{\boldsymbol{\ell}\in\mathcal{L}} \frac{1}{T-1}\sum_{t=2}^{T}[\Delta x_t - \Delta\widehat{s}_t(\boldsymbol{\ell})]^2.$$

Note the criterion equals

$$\frac{1}{T-1}\sum_{t=2}^{T}[\Delta s_t - \Delta\widehat{s}_t(\boldsymbol{\ell})]^2 + \frac{1}{T-1}\sum_{t=2}^{T}\Delta e_t^2 + \frac{2}{T-1}\sum_{t=2}^{T}[\Delta s_t - \Delta\widehat{s}_t(\boldsymbol{\ell})]\Delta e_t. \tag{14}$$

Here, by taking a first-order difference of the times series (i.e., Δx_t and $\Delta\widehat{s}_t(\boldsymbol{\ell})$), we avoid working with a nonstationary series and the associated difficulties. By the ergodic theorem, on the right-hand side of (14), the second term converges to a constant and the third term converges to zero. Thus, minimizing this criterion essentially finds the best configuration by matching the extracted seasonal component with the true seasonal component (i.e., focusing on the first term).

The TRAMO-SEATS method has a feature to handle breaks in seasonality as the seasonal outlier that allows for an abrupt increase or decrease in the level of the seasonal pattern that is compensated for in the other months or quarters. To use this feature, one needs to specify in advance the types of breaks and the break times in the official `X-13ARIMA-SEATS` program from the U.S. Census Bureau. Such seasonal outliers can be automatically detected in the Eurostat software JDemetra+. In contrast, our RSVD method not only can automatically detect seasonal breaks but also allows for more diverse abrupt breaks in seasonality.

6. SIMULATION

In this section, we use simulated monthly time-series data to evaluate the finite sample performance of our proposed seasonal adjustment methods and compare them with one state-of-art method used by U.S. Census Bureau. The benchmark for our comparison is the `X-13ARIMA-SEATS` (U.S. Census Bureau 2017), which is a hybrid program that integrates the model-based TRAMO/SEATS software developed at the Bank of Spain, described in Gómez and Maravall (1992, 1997) and the X-12-ARIMA program developed at the U.S. Census Bureau. In this section and the next, we abbreviate our seasonal adjustment methods as RSVD (since RSVD plays a critical role in our procedure), the TRAMO/SEATS and X-11 style methodologies in `X-13ARIMA-SEATS` program as SEATS and X-12-ARIMA respectively. The data generating processes follow (5). Section S.2 (in the online appendices) and 6.1 consider artificial seasonality with and without abrupt breaks and stationary/nonstationary ARIMA error terms, and Section 6.2 considers seasonality from three real economic time series.

6.1. Seasonality With Abrupt Breaks

Now, we consider a deterministic monthly seasonal component with a nonsmooth break

$$s_t^b \equiv s_{i,j}^b = b_i a_j$$

where $i = 1, \ldots, n$ and $j = 1, \ldots, 12$ indicate year and month respectively, and the elements in vector $\mathbf{b} = (b_1, \ldots, b_n)^\top$ and $\mathbf{a} = (a_1, \ldots, a_{12})^\top$ take the following values,

$$b_i = \begin{cases} 1 + i/10, & \text{if } 1 \le i \le n/2, \\ 1 + (n+1-i)/5, & \text{if } n/2 + 1 \le i \le n. \end{cases}$$

$$\mathbf{a} = (-1.25, -2.25, -1.25, 0.75, -1.25, -0.25, 2.75, \\ -0.25, 0.75, -0.25, 0.75, 1.75)^\top.$$

The vector $\mathbf{a}$ represents the reoccurring variation within each seasonal period, which is the same as that in Section 6.1. The magnitude of the seasonal component, captured by the multipliers in vector $\mathbf{b}$, increases slowly in the first $n/2$ years linearly, doubles at $n/2 + 1$ year, and then decreases slowly in the last $n/2$ years linearly. The seasonality can be also expressed in matrix form $\mathbf{s}^o = \mathbf{b}\mathbf{a}^\top = \mathbf{i}_n\mathbf{f}^\top + \mathbf{u}\mathbf{v}^\top$ where the terms are defined the same way in (S.3). In Figure S2, we plot fixed/time-varying seasonal patterns $\mathbf{f}$ and $\mathbf{v}$ in upper-left panel, fixed/time-varying pattern coefficients $\mathbf{i}_n$ and $\mathbf{u}$ in upper-right panel, fixed/time-varying seasonality $\mathbf{i}_n\mathbf{f}^\top$ and $\mathbf{u}\mathbf{v}^\top$ in lower-left panel, and total seasonality $\mathbf{s}^o$ in lower-right panel.

For the nonseasonal component, we only consider the nonstationary ARIMA(1,1,1) process in DGP3: $e_t \sim \text{ARIMA}(1, 1, 1)$, with $\phi = 0.8$ and $\psi = 0.1$ with $N(0, \sigma^2)$ innovations and $\sigma^2 = 0.04$. The results of stationary cases for DGP1 and DGP2, which are similar to the nonstationary DGP3, are omitted here.

After the seasonal component s_t^b and non-seasonal component e_t are generated, we use the following formula to obtain simulated time series data

$$x_t = s_t + e_t \equiv \kappa \frac{\text{SD}(e_t)}{\text{SD}(s_t^b)} s_t^b + e_t,$$

and the sample unconditional standard deviation ratio $\text{SD}(s_t)/\text{SD}(e_t)$ is fixed to be exactly κ in each replication of DGPs. For nonstationary DGP3, we choose κ = 0.2, 0.4, ..., 2 in our setups. For each combination of DGP3 and κ values, we simulate monthly time-series data with sample size $T = 240$ (i.e., $n = 20$ and $p = 12$). We repeat the simulation $B = 500$ times for each setup.

Table 1 reports the results of the two benchmark methods, RSVD without break, and RSVD allowing for break. Both RSVD methods outperform the benchmarks by delivering smaller absolute and relative losses, and the RSVD allowing for break has the smallest error among the three methods. The absolute loss (AMSE) of RSVD and the benchmarks increases as the ratio κ increases, while that of RSVD allowing for break decreases and stabilizes. The relative loss, AMPE, of the three methods decreases as κ increases, and that of the RSVD allowing for break decreases most quickly among the three methods. Moreover, similar to the cases in Table S1, the average selected numbers of seasonal patterns r for both RSVD methods are generally the same and are close to one across different values of κ, and no additional numbers of seasonal patterns are added due to the irregular variation.

6.2. Seasonality From Real Economic Time Series

The simulation in Section 6.1 favors our proposed methods since the artificial seasonality takes exactly the form in (7) that the X-12-ARIMA and SEATS may disagree. We also use the real seasonalities extracted from three seasonal economic time series to conduct simulation. They are *Industrial Production Index*, *Total Nonfarm Payrolls*, and the Inflation Rate calculated from *Consumer Price Index for all Urban Consumers*, which are available on Federal Reserve Economic Data website.

Table 1. Evaluation of estimates of seasonal component with break (DGP3)*

	AMSE ($\times 10^{-2}$)				AMPE (%)				Avg. r	
κ	X-12-ARIMA	SEATS	RSVD	RSVDB	X-12-ARIMA	SEATS	RSVD	RSVDB	RSVD	RSVDB
0.2	4.3774	5.7199	1.9851	1.7623	33.18	40.89	24.32	23.87	1.030	1.026
0.4	11.2136	11.3763	2.5562	1.6082	19.51	28.93	11.82	11.49	1.012	1.016
0.6	21.4088	16.3613	3.8361	1.5588	15.23	23.38	8.06	7.54	1.006	1.014
0.8	34.7795	20.2008	5.6820	1.5439	13.16	19.62	6.27	5.63	1.008	1.014
1.0	50.3344	23.9295	8.0542	1.5366	11.92	16.90	5.19	4.49	1.004	1.014
1.2	68.6043	27.3233	10.9465	1.5318	11.16	14.49	4.49	3.74	1.004	1.014
1.4	86.4549	31.5762	14.4462	1.5296	10.56	12.82	4.04	3.20	1.012	1.014
1.6	108.0307	36.1908	18.4037	1.5276	10.12	11.46	3.67	2.80	1.012	1.014
1.8	130.6624	41.2645	22.9642	1.5264	9.77	10.37	3.39	2.49	1.004	1.014
2.0	155.3913	46.3343	28.0056	1.5254	9.49	9.49	3.18	2.24	1.008	1.014

*The nonseasonal component $\{e_t\}$ follows Gaussian ARIMA(1,1,1) with AR(1) coefficient 0.8 and MA(1) coefficient 0.1. The official X-13ARIMA-SEATS program can only manually specify seasonal outliers as breaks in seasonality. To make the evaluation more fair, we use the X-12-ARIMA and SEATS provided in JDemetra+ that is capable of detecting seasonal outliers automatically.

We adopt two different schemes of simulation and discuss the simulation results thoroughly. The simulation results and detailed discussion are reported in the online appendices. In general, the main messages conveyed by these simulation exercises are in line with those in the subsections above: When seasonality is strong, our RSVD seasonal adjustment method is superior to X-12-ARIMA and SEATS methods by delivering much smaller AMSE and AMPE losses, and X-12-ARIMA and SEATS methods tend to outperform our RSVD method when the seasonality is weak.

7. REAL DATA

In the section, we use some real time-series data with seasonal behaviors to compare our proposed RSVDB seasonal adjustment method with the X-12-ARIMA and SEATS methods. They are (i) monthly retail volume data (henceforth *retail*), (ii) quarterly berry production data of New Zealand (henceforth *berry*), and (iii) daily online submission counts (henceforth *counts*). These three empirical examples are specifically selected to showcase that our proposed RSVDB seasonal adjustment method could produce (i) similar seasonal components as the X-12-ARIMA and SEATS methods do; (ii) better seasonal components when X-12-ARIMA and SEATS fail; and (iii) seasonal components for other than quarterly and monthly frequencies, such as daily and weekly. Furthermore, the seasonality of the first series (monthly retail volume data) is steady and mild, which is similar to that of the simulated series in Section 6.2 where the seasonality is from real economic time series. In contrast, the seasonality of the second series (quarterly berry production data of New Zealand) is strong and has some possible breaks, which is similar to that of simulated series in Section 6.1 where the artificial seasonality has an abrupt break in the middle of the time period.

Since no exact definition of seasonality exists and the true underlying seasonality is always unknown in real data, different seasonal adjustment methods recognize seasonality differently given the same data. It is hard to formally compare the results from different seasonal adjustment methods, especially when they are very close. For empirical applications, these seasonal adjustment methods can only be compared and evaluated in a qualitative fashion with visual inspection. More importantly, our proposed RSVDB method is able to decompose the seasonal component into different seasonal patterns, trace the dynamics of seasonality by time-varying pattern coefficients, and identify important seasonality break times automatically. We also use these three applications to illustrate that the RSVDB method can provide a very transparent and meaningful explanation to economic seasonality in real data.

In this section, only seasonal decompositions of the X-12-ARIMA and SEATS methods in `X-13ARIMA-SEATS` program and the RSVDB method are compared. The first monthly series is already pretreated, and the second quarterly series has very strong seasonal fluctuations which may overwhelm possible calendar effects and outliers. Given the automatic feature of `X-13ARIMA-SEATS` program in seasonal adjustment, we only shut down the options for calendar effects and automatic outlier detection, and apply the X-12-ARIMA, SEATS, and RSVDB methods to these two series directly so that the comparison among the three methods only considers their capabilities of seasonal adjustment. In addition, the number of seasonal patterns r in the RSVDB method is selected by the Bayesian Information Criterion (13), allowing for a non-smooth break in each of the corresponding left singular vectors with roughness penalties for all the three time series datasets. To control the computational burden in the exhaustive search for abrupt seasonality breaks, we limit the maximal number of seasonal patterns to be 3, that is $1 \le r \le r_{\max} = 3$.

7.1. Retail Volume Data (Retail)

We first examine the monthly series of Motor Vehicle and Parts Dealers published by the U.S. Census Bureau's Advance Monthly Sales for Retail and Food Services, covering the period from January 1992 through December 2012.

Figure 1 and 4 show and compare the seasonal adjustment results of the retail time series data using both the X-12-ARMA, SEATS, and RSVD methods. In Figure 4(a) – (c), we plot the fixed and time-varying seasonal patterns in $\mathbf{f}$ and $\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ and their corresponding time-varying pattern

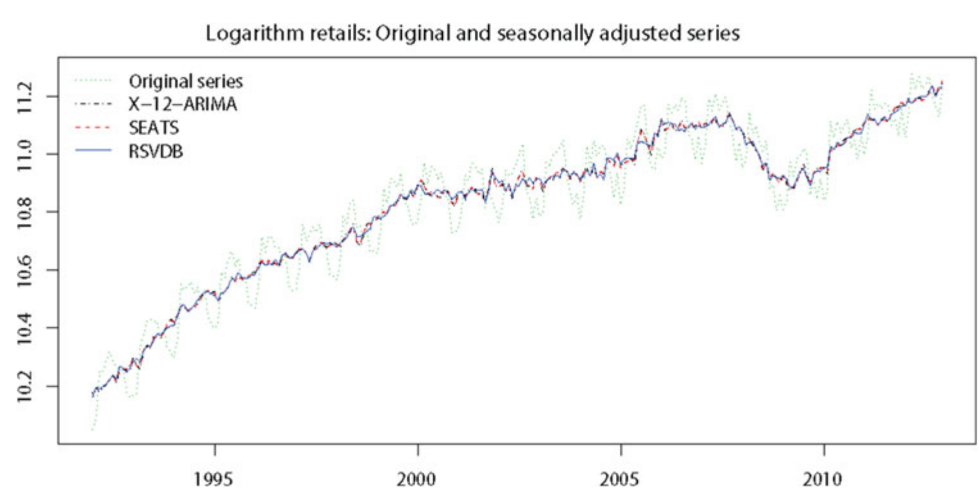

Figure 1. Logarithm retail volume: Original and seasonally adjusted series with the X-12-ARIMA, SEATS, and RSVDB methods

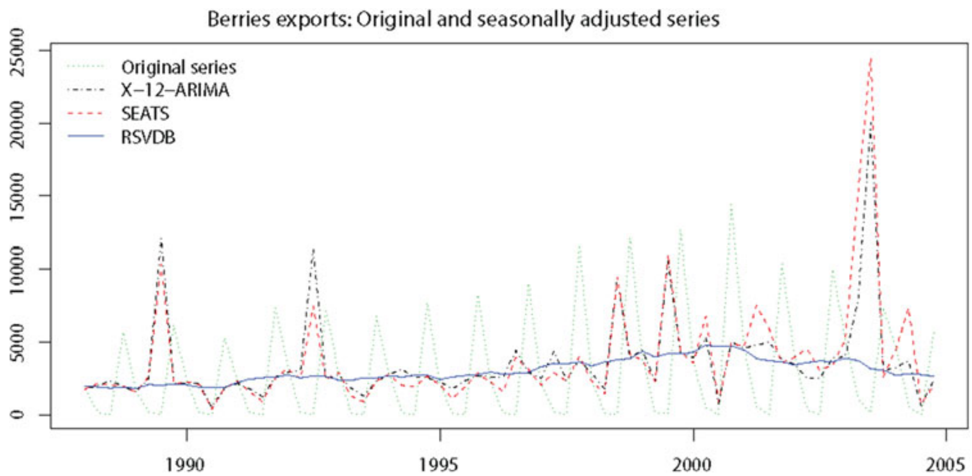

Figure 2. New Zealand berries exports: Original and seasonally adjusted series with the X-12-ARIMA, SEATS, and RSVDB methods

coefficients in $\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$. The black solid, red dashed, and green dotted vertical lines in Figure 4(c) represents the abrupt break detected in $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ respectively. Figure 4(d) presents the fixed seasonality $\mathbf{i}_n \cdot \mathbf{f}^\top$ and the time-varying seasonality $\sum_{r=1}^{3} \mathbf{u}_r \mathbf{v}_r^\top$. In Figure 4(e), we plot the three seasonal components extracted by X-12-ARIMA, SEATS and RSVD method. Figure 1 shows the original time series and seasonal adjustments by the three adjustment methods. Finally, to check whether our RSVD method adjusts seasonality adequately, we plot the periodogram for the adjusted series in Figure S3. Clearly, the sample spectrum does not show any spike at any of the seasonal frequencies, indicating no residual seasonality exists.

First, in Figure 4(d), we find that the fixed seasonal component is larger than the time-varying seasonal component. Second, the seasonal component s_t extracted via X-12-ARIMA, SEATS, and RSVD are very similar for this retail volume time series. The three breaks detected in the three time-varying pattern coefficients $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ segment the time series into four periods, see Figure 4(e) and Figure 1. In the period around red dashed and green dotted vertical lines, the RSVD seasonal component is slightly more volatile than the other two estimated seasonal components, and the RSVD seasonal adjusted series is slightly smoother than the other two seasonal adjusted series. In the other periods, the RSVD seasonal components and adjusted time series are almost the same as their counterparts.

7.2. Berry Production Data of New Zealand

We next examine the quarterly series of New Zealand constant price exports of berries, covering the period from 1988Q1 to 2005Q2.

Figures 2 and 5 show and compare the seasonal adjustment results of the berry production time series data using X-12-ARIMA, SEATS, and RSVD. In Figure 5(a)–(c), we plot the fixed and time-varying seasonal patterns in $\mathbf{f}$ and $\mathbf{V} = (\mathbf{v}_1, \mathbf{v}_2)$ and their corresponding time-varying pattern coefficients in $\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2)$. The fixed seasonal pattern has a much larger scale than the time-varying seasonal patterns. The black solid and red dashed vertical lines in Figure 5(c) represents the abrupt break detected in $\mathbf{u}_1$ and $\mathbf{u}_2$ respectively. Figure 5(d) presents the fixed seasonality $\mathbf{i}_n \cdot \mathbf{f}^\top$ and the time-varying seasonality $\sum_{r=1}^{2} \mathbf{u}_r \mathbf{v}_r^\top$. In Figure 5(e), we plot the three seasonal components extracted by X-12-ARIMA, SEATS, and RSVDB method. Figure 2 shows the original time series and seasonal adjusted ones by the three adjustment methods. Finally, we plot the periodogram for the RSVDB adjusted series in Figure S4. Clearly, the sample spectrum does not show any spike at any of the seasonal frequencies, indicating no residual seasonality exists.

Because of the laws of nature in agricultural production, the actual berry production in the fall quarter is close to zero. Despite this, both X-12-ARIMA and SEATS methods automatically apply the logarithmic transformation and do not deliver reasonable results: Their seasonal components are excessively negative at certain periods in Figure 5(e), and their adjusted series in Figure 2 is excessively high at those periods. It turns out that one need further manually modify the default options of the two methods to produce reasonable outcomes. In contrast, the RSVD method is robust to this irregularity and produce reasonable results. Just like the retail volume data, the fixed seasonal component is much more salient than the time-varying component, and dominates the seasonality.

Moreover, RSVD identifies the year 2000 as a major break time in seasonality, as the first seasonal pattern coefficients drop dramatically after year 2000. This phenomenon is also manifested in Figure 2 and 5(e). The magnitude of seasonality generally increases gradually before year 2000, then has a sudden decrease in year 2001 and decreases gradually thereafter.

7.3. Online Submission Count Data

Finally, we study a daily time series of submission counts for 2015 Census Test, covering March 23 through June 1. The Census Test is described in *www.census.gov/2015censustests*. Submissions cover both self-responses and responses taken over the telephone at one of the Census Bureau telephone centers. In this case, the seasonal component to the data corresponds to day-of-week dynamics, and it is of interest to know whether certain days have systematically higher activity.

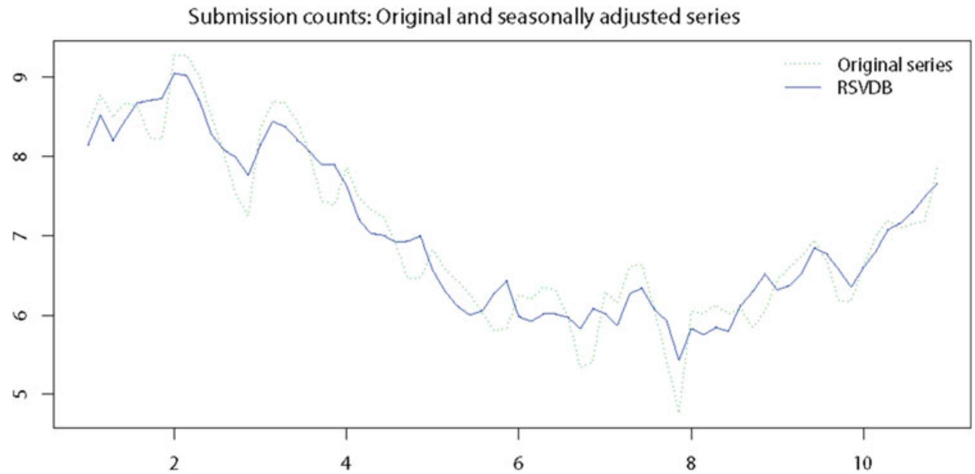

Figure 3. Logarithm submission counts: Original and seasonally adjusted series with the RSVDB method

Figures 3 and 6 show the seasonal adjustment results of the submission counts time series data using the RSVD method. In Figure 6(a)–(c), we plot the fixed and time-varying seasonal patterns in $\mathbf{f}$ and $\mathbf{v}_1$ and their corresponding time-varying pattern coefficients in $\mathbf{u}_1$. The black solid vertical line in Figure 6(c) represents the abrupt break detected in $\mathbf{u}_1$. Figure 6(d) presents the fixed seasonality $\mathbf{i}_n \cdot \mathbf{f}^\top$ and the time-varying seasonality $\mathbf{u}_1\mathbf{v}_1^\top$. In Figure 6(e), we plot the seasonal component extracted by the RSVD method. Finally, we plot the periodogram for the RSVDB adjusted series in Figure S5. The sample spectrum does not show any spike at any of the seasonal frequencies, indicating no residual seasonality exists.

Figure 3 plots the seasonal adjustment results of the daily online submission count data in logarithms. Because the data occurs at a daily frequency, the X-13ARIMA-SEATS software cannot be applied, although in principle model-based approaches could be used. However, the seasonal pattern (i.e., the weekly pattern) is very dynamic, and hence presents a challenge for parametric models. In contrast, our proposed RSVDB method is still well applicable to the daily data with weekly seasonality. In Figure 6(d), the fixed and time-varying seasonal components have similar magnitude. In Figure 6(e), the RSVD seasonal component shows that the seasonal behavior is quite different at the beginning, middle, and end of the time series. In Figure 3, the seasonal adjusted series is much smoother than the original series: the submission counts series increases and reaches its peak in the first week, decreases in the second week, and first increases and then decrease in the third week. Then, the adjusted series keeps decreasing and reaches its trough in the sixth week. After that, the adjusted series increases again but with more fluctuations.

More interestingly, RSVD identifies the 4th week as the major break time in seasonality. It is observed that in Figure 6(c), the time-varying pattern coefficients are virtually zero before and on the 4th week, indicating that basically no time-varying seasonal pattern appears during the first 4 weeks. After that, the time-varying pattern coefficients, moving away from zero, first decreases a little and then increases sharply after the 7th week. This means that the time-varying seasonal pattern emerges gradually after the 4th week and finally dominates the seasonal component at the end of this series.

8. DISCUSSION AND CONCLUSION

Other important issues concerning our proposed RSVD method include how to deal with some potential data problems (such as missing values, outliers, and calendar effects), how to obtain confidence intervals for the seasonally adjusted process, and how to use RSVD to deal with multiple types of seasonality for the time series with daily or even high frequency. Due to the space limitation, some further discussions on these issues are provided in Section S.4 of the online appendices.

The bulk of seasonal adjustment methodology and software is divided between the model-based and empirical-based approaches, each with their own proponents among researchers and practitioners. The empirical-based methods all rely upon linear filters, and therefore struggle to successfully adjust highly nonlinear seasonal structures. The model-based methods are more flexible, yielding a wider array of filters, but the methods (whether based on deterministic or stochastic components) still tend to be linear. When seasonality evinces structural changes (perhaps a response of consumers to a change in legislation), systemic extremes (perhaps due to high sensitivity to local weather conditions), or very rapid change (perhaps due to a dynamic marketplace, where new technologies rapidly alter cultural habits) the conventional paradigms tend to be inadequate. While it's possible to specify ever-more complex models, it is arguably more attractive to devise nonparametric (or empirical-based) techniques that automatically adapt to a variety of structures in the data – this approach is especially attractive to a statistical agency involved in adjusting thousands of series each month or quarter, because devising specially crafted models for each problem series requires excessive manpower.

The methodology of this article is empirical in spirit, utilizing a nonparametric method to separate seasonal structure from other time series dynamics. Like X-12-ARIMA, which combines nonparametric filters with model-based forecast extension, our RSVD method combines stochastic models of nuisance structures with the regularized elicitation of seasonal dynamics. The advantages over purely model-based approaches are an ability to avoid model misspecification fallacies, allow for structural change in seasonality, handle seasonal extremes,

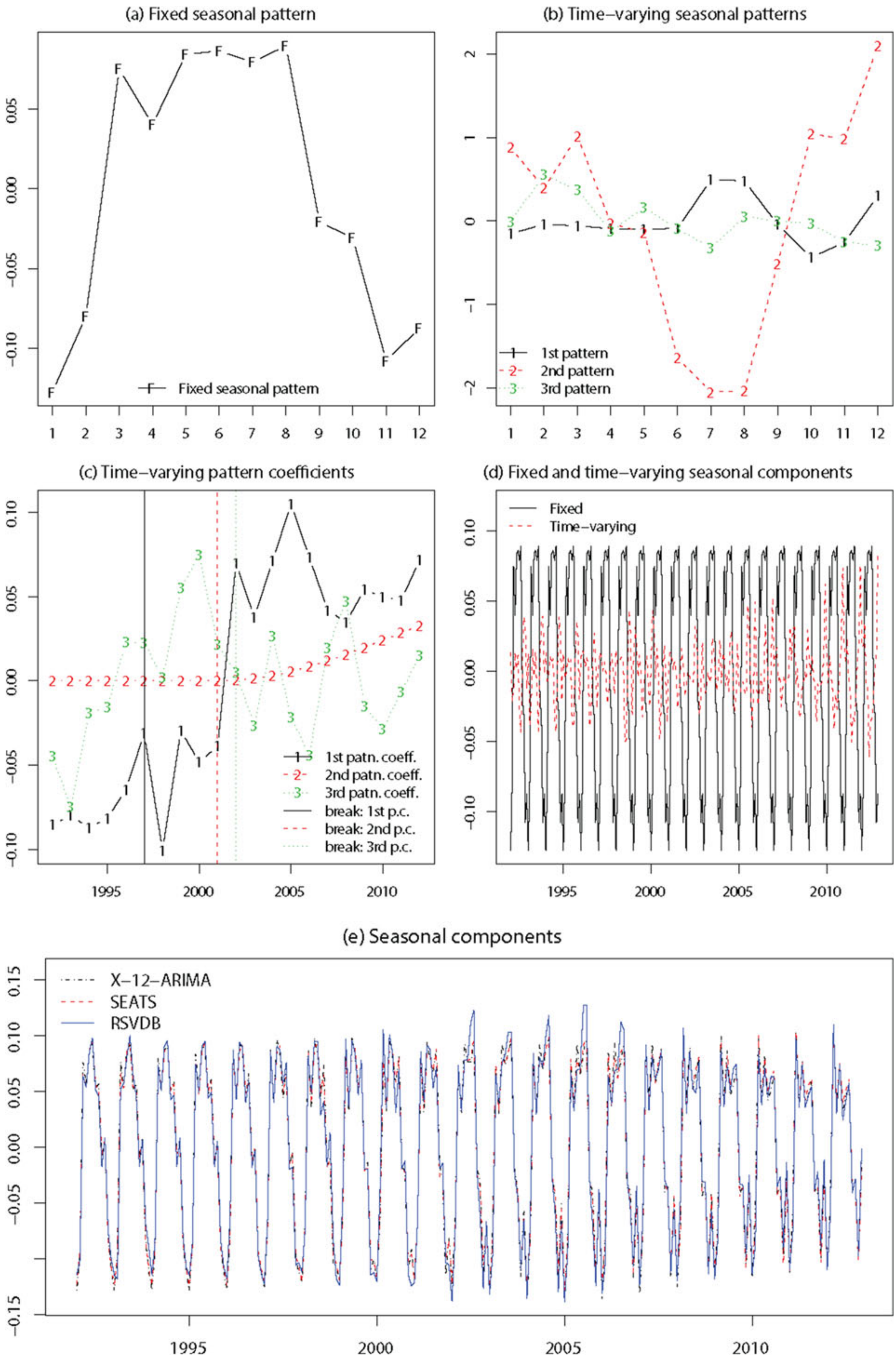

Figure 4. The RSVDB seasonal decomposition of the logarithm retail series

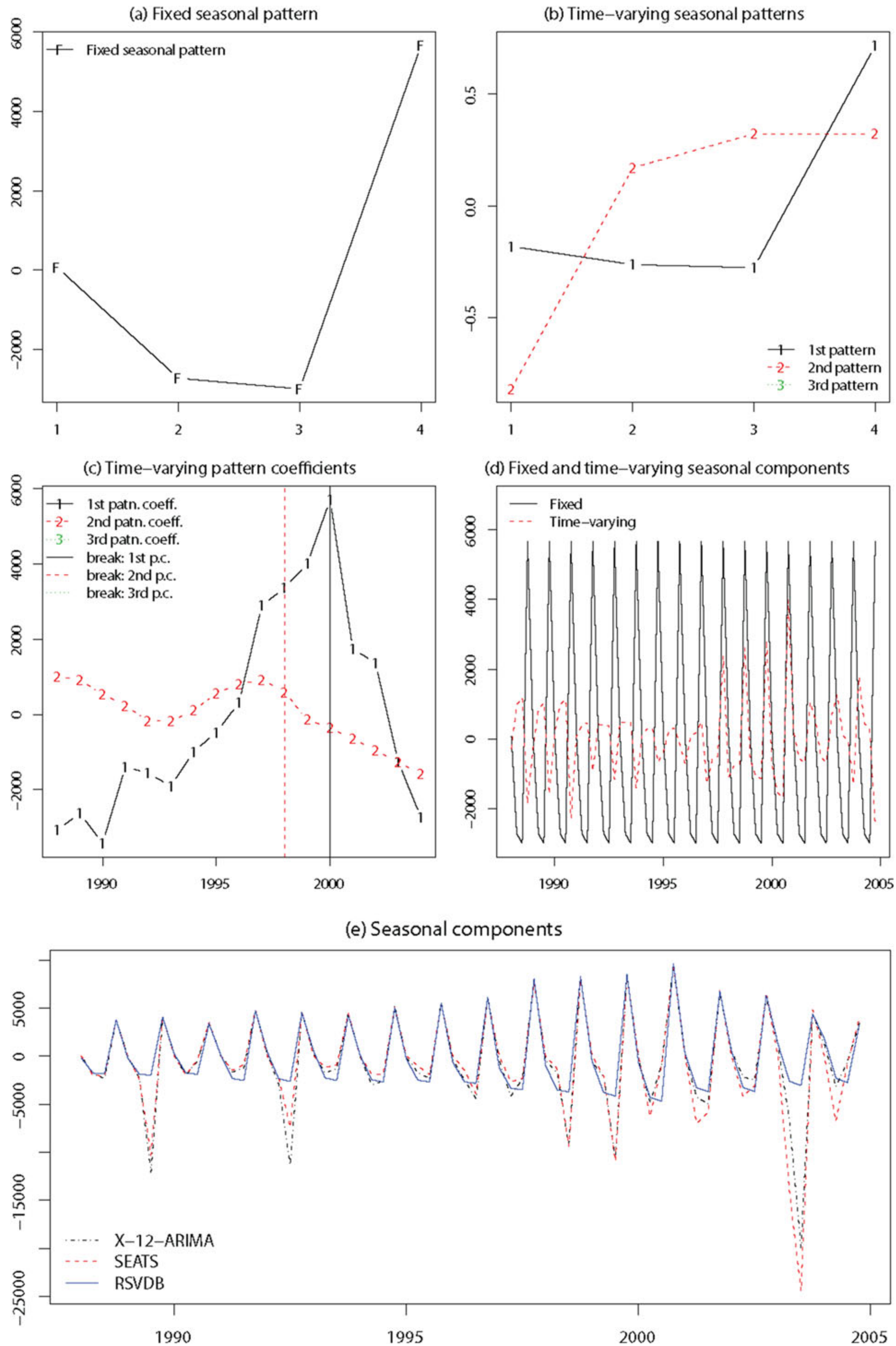

Figure 5. The RSVDB seasonal decomposition of the berries exports series

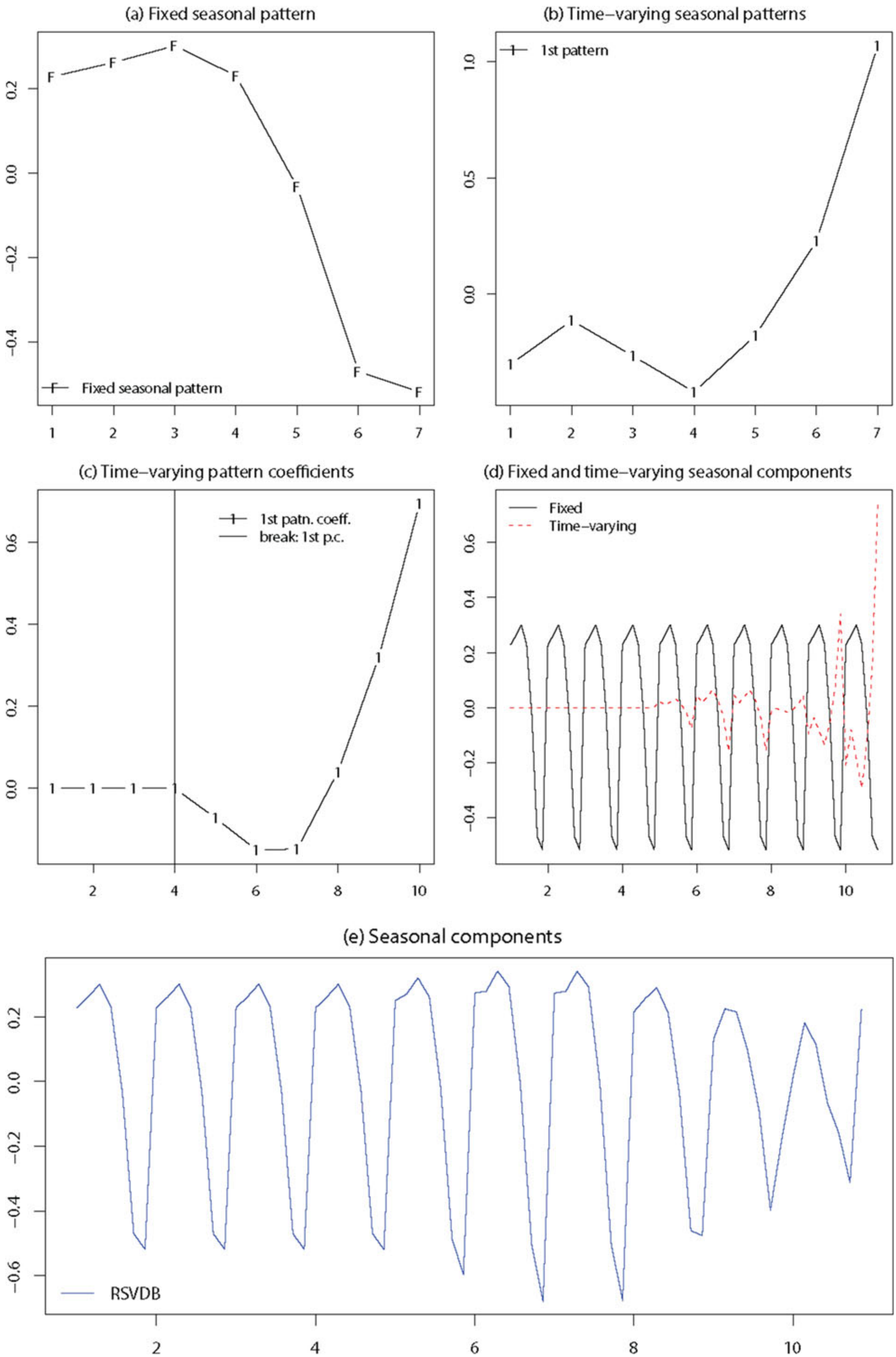

Figure 6. The RSVDB seasonal decomposition of the daily submission counts series

and capture rapidly evolving seasonality. Moreover, the RSVD method is computationally fast and almost automatic (the ARIMA specification does require choices of the user), and hence is attractive in a context where individual attention to thousands of series is a logistical impossibility. An admitted downside is that RSVD does not quantify the estimation error in the seasonal component. With market demands for more data – higher frequency, more granularity – coupled with tightening budgets, the necessity of automation in data processing must drive future research efforts; RSVD takes a substantial step in that direction.

Finally, we mention that there are many fruitful directions for extensions to RSVD: use of the **U** and **V** singular vectors to detect seasonality; multivariate modeling, where **U** vectors may be common to multiple time series; handling multiple frequencies of seasonality (e.g., daily time series with weekly and annual seasonality) through an extension of matrix embedding to an array (tensor) structure. Any of these facets would greatly assist the massive data processing task facing statistical agencies.

SUPPLEMENTARY MATERIALS

The online appendices include the derivation of the GCV criterion in Algorithm 1, simulation results with artificial seasonality without abrupt breaks, simulation results with seasonality from real economic time series, further discussions on some extensions of our RSVD method, and additional tables and figures for simulation, model comparison, and empirical illustration. The data and R codes are also provided.

ACKNOWLEDGMENTS

We are grateful to the seminar and conference participants at Center for Statistical Research & Methodology, U.S. Census Bureau, Washington D.C. 2015, Department of Economics, Texas A&M University, College Station 2015, and School of Economics, Shanghai University of Finance and Economics, Shanghai 2016, School of International Trade and Economics, University of International Business and Economics 2016, Institute of Statistics and Big Data, Renmin University of China 2017, the 2017 International Conference on Econometrics and Statistics in Hong Kong, School of Economics, Liaoning University 2018, and the 2018 Workshop for Economics and Financial Econometrics in the Era of Big Data in Beijing for helpful comments. We are also grateful to the editor, Todd Clark, associate editor, and two anonymous referees whose constructive comments have greatly improved the manuscript. The usual caveat applies. Wei Lin's research is partially supported by National Science Foundation of China grants No. 71501134 (Young Scientists Fund Project) and No. 71671183 (General Program Project).

[Received November 2016. Revised May 2018.]

REFERENCES

Barsky, R. B., and Miron, J. A. (1989), "The Seasonal Cycle and the Business Cycle," *Journal of Political Economy*, 503–534. [1]

Bell, R., Bennett, J., Koren, Y., and Volinsky, C. (2009), "The Million Dollar Programming Prize," *IEEE Spectrum*, 46, 28–33. [3]

Bell, W. R., Holan, S. H., and McElroy, T. S. (Ed.). (2012), Economic Time series: modeling and seasonality, London, UK: Chapman and Hall/CRC. [1]

Bunzel, H., and Hylleberg, S. (1982), "Seasonality in Dynamic Regression Models: A Comparative Study of Finite Sample Properties of Various Regression Estimators Including Band Spectrum Regression," *Journal of econometrics*, 19, 345–366. [1]

Buys Ballot, C. H. D. (1847), *Les Changements Périodiques de Température*, Utrecht: The Netherlands: Kemink et Fils. [3]

Cai, Z., and Chen, R. (2006), "Flexible Seasonal Time Series Models," *Econometric Analysis of Financial and Economic Time Series*, 20, 63–87. [1]

Canova, F. (1992), "An Alternative Approach to Modeling and Forecasting Seasonal Time Series," *Journal of Business and Economic Statistics*, 10, 97–108. [1]

Dagum, E. B. (1980), *The X-11-ARIMA Seasonal Adjustment method*, Ottawa, Canada, Statistics Canada, Seasonal Adjustment and Time Series Staff. [2]

Dagum, E. B., and Bianconcini, S. (2016), Seasonal adjustment methods and real time Trend-Cycle estimation, Berlin, Germany: Springer. [1]

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., and Harshman, R. (1990), "Indexing by Latent Semantic Analysis," *Journal of the American society for information science*, 41, 391–407. [3]

Dokumentov, A., and Hyndman, R. J. (2015), "STR: A Seasonal-Trend Decomposition Procedure based on Regression" (Working paper No. 13/15). Monash University Econometrics and Business Statistics. Available at http://robjhyndman.com/working-papers/str/ [2]

Eiurridge, P., and Wallis, K. F. (1990), "Seasonal Adjustment and Kalman Filtering: Extension to Periodic Variances," *Journal of Forecasting*, 9, 109–118. [1]

Findley, D. F. (2005), "Some Recent Developments and Directions in Seasonal Adjustment," *Journal of Official Statistics*, 21, 343–365. [1]

Findley, D. F., Monsell, B. C., Bell, W. R., Otto, M. C., and Chen, B. C. (1998), "New Capabilities and Methods of the X-12-ARIMA Seasonal-Adjustment Program," *Journal of Business and Economic Statistics*, 16, 127–152. [2]

Franses, P. H. (1998), Time series models for business and economic forecasting, London, UK: Cambridge University Press. [1]

Gersovitz, M., and MacKinnon, J. G. (1978), "Seasonality in Regression: An Application of Smoothness Priors," *Journal of the American Statistical Association*, 73, 264–273. [1]

Gómez, V., and Maravall, A. (1992), "Time Series Regression with ARIMA Noise and Missing Observations—Program TRAMO," EUI Working Paper ECO, No. 92/81. [7]

Gómez, V., and Maravall, A. (1997), "Programs TRAMO and SEATS: Instructions for the User," Working Paper 9628. Servicio de Estudios. Banco de España. [7]

Hansen, L. P., and Sargent, T. J. (1993), "Seasonality and Approximation Errors in Rational Expectations Models," *Journal of Econometrics*, 55, 21–55. [1]

Harrison, P. J., and Stevens, C. F. (1976), "Bayesian forecasting," *Journal of Royal Statistical Society*, Series B, 38, 205–247. [1]

Harvey, A. C. (1990), *Forecasting, Structural Time Series Models and the Kalman Filter*, London, UK: Cambridge University Press. [1]

Harvey, A. and Scott, A. (1994), "Seasonality in Dynamic Regression Models," *The Economic Journal*, 1324–1345. [1]

Huang, J. Z., Shen, H., and Buja, A. (2008), "Functional Principal Components Analysis Via Penalized Rank One Approximation," *Electronic Journal of Statistics*, 2, 678–695. [2]

Huang, J. Z., Shen, H., and Buja, A. (2009), "The Analysis of Two-Way Functional Data using Two-Way Regularized Singular Value Decompositions," *Journal of American Statistical Association*, 104, 1609–1620. [2]

Hylleberg, S., Engle, R. F., Granger, C. W., and Yoo, B. S. (1990), "Seasonal Integration and Cointegration," *Journal of econometrics*, 44, 215–238. [1]

Lovell, M. C. (1963), "Seasonal Adjustment of Economic Time Series and Multiple Regression Analysis," *Journal of the American Statistical Association*, 58, 993–1010. [1]

Osborn, D. R. (1991), "The Implications of Periodically Varying Coefficients for Seasonal Time-Series Processes," *Journal of Econometrics*, 48, 373–384. [1]

——— (1993), "Discussion: Seasonal Cointegration," *Journal of Econometrics*, 55, 299–303. [1]

Proietti, T. (2004), "Seasonal Specific Structural Time Series," *Studies in Nonlinear Dynamics and Econometrics*, 8. [1]

Sarwar, B., Karypis, G., Konstan, J., and Riedl, J. (2000), "Application of Dimensionality Reduction in Recommender System—a Case Study (No. TR-00-043)," Minneapolis, MN: Minnesota Univ Minneapolis Dept of Computer Science. [3]

Shiskin, J., Young, A. H., and Musgrave, J. C. (1965), "The X-11 Variant of the Census Method II Seasonal Adjustment Program," No. 15., Washington, DC: US Department of Commerce, Bureau of the Census. [2]

Sims, C. A. (1974), "Seasonality in Regression," *Journal of the American Statistical Association*, 69, 618–626. [1]

U.S. Census Bureau (2017), "X-13ARIMA-SEATS Reference Manual," http://www.census.gov/ts/x13as/docX13AS.pdf [7]