Time series seasonal adjustment using regularized singular value decomposition[1]

Wei Lin[1], Jianhua Z. Huang[2], and Tucker McElroy[3]

[1] *International School of Economics and Management, Capital University of Economics and Business, Beijing China 100070*

[2] *Department of Statistics, Texas A&M University, College Station, TX 77840*

[3] *Center for Statistical Research and Methodology, U.S. Census Bureau, 4600 Silver Hill Road, Washington, D.C. 20233*

November 3, 2016

[1] We are grateful to the seminar 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 for helpful comments.

Abstract

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 non-seasonal component. The method also has a variant that can handle that 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 method on both simulated and real data examples.

Key Words: Seasonal adjustment, regularized singular value decomposition, X-13ARIMA-SEATS
JEL Classification: C14, C22.

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 helps 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. A fairly recent overview of seasonal adjustment methodology and software is given in Findley (2005).

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 non-seasonal 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 non-seasonal 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 component, 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; Burridge and Wallis, 1990; Harvey and Scott 1994; Proietti, 2004). Because of the direct specification 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 non-seasonal components in the time series, which subsequently causes misinterpretations of dynamic relationships across different time series. In order 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 pre-treat 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.

Given the availability of many seasonal adjustment methods, many national statistical agencies prefer the empirical-based approach because of its simplicity and nonreliance on model assumptions. In this paper, we adopt the empirical-based approach and propose a flexible and robust seasonal adjustment method based on regularized singular value decomposition (RSVD; Huang, Shen, and Buja, 2008, 2009). 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 non-seasonal 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.

The remaining part of this paper is organized as follows. Section 2 brief 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 non-seasonal component is stationary. Sections 5 and 6 extend 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 7, and three real data examples are provided in Section 8. Section 9 concludes.

2 A brief review of regularized SVD

Regularized singular valued decomposition (RSVD) is a variant of singular value decomposition that takes into account the intrinsic smoothness structure of a data matrix data (Huang, Shen and Buja 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 a $n \times p$ 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,

$$(\widehat{\mathbf{u}}, \widehat{\mathbf{v}}) = \arg\min_{\mathbf{u},\mathbf{v}} \|\mathbf{X} - \mathbf{u}\mathbf{v}^{\intercal}\|_F^2, \tag{2.1}$$

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,

$$(\widehat{\mathbf{u}}, \widehat{\mathbf{v}}) = \arg\min_{\mathbf{u},\mathbf{v}} \|\mathbf{X} - \mathbf{u}\mathbf{v}^\intercal\|_F^2 + \alpha \mathbf{u}^\intercal \boldsymbol{\Omega} \mathbf{u} \tag{2.2}$$

where is $\boldsymbol{\Omega}$ is a $n \times n$ non-negative definite roughness penalty matrix, α is smoothing parameter, and $\mathbf{v}^\intercal\mathbf{v} = 1$ for identification purpose.

A simple variant of the power algorithms in Huang et al. (2008, 2009) gives the following Algorithm 1 for solving the problem (2.2).

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:

(a) $\mathbf{v} \leftarrow \dfrac{\mathbf{X}^\intercal\mathbf{u}}{\|\mathbf{X}^\intercal\mathbf{u}\|}$.

(b) $\mathbf{u} \leftarrow (\mathbf{I}_n + \alpha\boldsymbol{\Omega})^{-1}\mathbf{X}\mathbf{v}$ *with* α *selected by minimizing the following generalized cross-validation criterion,*

$$\mathrm{GCV}(\alpha) = \frac{1}{n}\frac{\|[\mathbf{I}_n - \mathbf{M}(\alpha)]\mathbf{X}\mathbf{v}\|^2}{\left(1 - \frac{1}{n}\operatorname{tr}\{\mathbf{M}(\alpha)\}\right)^2}, \tag{2.3}$$

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 (2.3) is similar to Huang et al. (2008, 2009) and can be found in the supplementary materials. The difference of Algorithm 1 from previous algorithms is that, in Huang et al. (2008) the roughness penalty is imposed only on $\mathbf{v}$ and in Huang et al. (2009) on both $\mathbf{u}$ and $\mathbf{v}$. If there is no penality, i.e., $\alpha = 0$, the algorithm is essentially the power algorithm for standard SVD and solves the problem (2.1).

In general, the regularized SVD attempts to find a rank-r decomposition ($r \leq p$) such that $\mathbf{X} = \mathbf{U}\mathbf{V}^\intercal$, where $\mathbf{U}$ is a $n \times r$ matrix, and $\mathbf{V}$ is a $p \times r$ matrix. The j-th column in matrix $\mathbf{U}$ and $\mathbf{V}$ is called the j-th 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 $\mathbf{X} - \widehat{\mathbf{u}}\widehat{\mathbf{v}}^\intercal$. 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 (c.f., Buys Ballot (1847)) as a matrix with p columns, whose each row represents one period of the seasons, as follows

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,p} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,p} \\ \cdots & \cdots & \cdots & \cdots \\ \vdots & x_{i(t),j(t)} & \vdots & \vdots \\ \cdots & \cdots & \cdots & \cdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,p} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{1\cdot} \\ \mathbf{x}_{2\cdot} \\ \vdots \\ \mathbf{x}_{i\cdot} \\ \vdots \\ \mathbf{x}_{n\cdot} \end{bmatrix},$$

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

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

where the function $\mathrm{Vec}(\cdot)$ converts a matrix into a column vector by stacking the columns of the matrix. The subsecripts 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 j-th **season** in the i-th **period**,

$$\begin{aligned} I : \mathbb{N} &\mapsto \mathbb{N}^2 \\ t &\to (i(t), j(t)) \equiv (\lceil t/p \rceil, t - \lfloor t/p \rfloor p), \end{aligned} \tag{3.1}$$

Let $n \equiv T/p$ denote the total number of time span included in the time series, so that we have that $1 \leq i \leq n$, $1 \leq j \leq p$, and $t = (i(t) - 1)p + j(t)$.[1]

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, i.e., $\mathbf{Q}_n \equiv \mathbf{I}_n - \mathbf{i}_n\mathbf{i}_n^\intercal/n$ so that $\mathbf{Q}_n\mathbf{a} = \mathbf{a} - \bar{\mathbf{a}}$, for a vector $\mathbf{a} = (a_1, \ldots, a_n)^\intercal$, where $\bar{\mathbf{a}} = \bar{a}\,\mathbf{i}_n$, and $\bar{a} = \sum_{1 \leq n} a_i/n$.

Let Δ be the first order difference operator such that,

$$\Delta \equiv \begin{bmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix}_{(n-1)\times n}. \tag{3.2}$$

Then the second order difference operator Δ^2 is,

$$\Delta^2 \equiv \begin{bmatrix} 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \end{bmatrix}_{(n-2)\times n}.$$

Using these difference operators, one widely used choice of the penalty matrix in (2.2) can take the form $\mathbf{\Omega} \equiv (\Delta^2)^\intercal\Delta^2$.

4 Basic seasonal adjustment

This section discusses seasonal adjustment when the non-seasonal component of a time series is stationary.

[1] Here we assume that T/p is an integer for simplicity of exposition.

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 non-seasonal component* e_t in the additive form,

$$x_t = s_t + e_t, \quad t = 1, \ldots, T, \tag{4.1}$$

where the non-seasonal component e_t is a stationary process. Using the mapping I defined in (3.1), we rewrite (4.1) 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,

$$\mathbf{X} = \mathbf{S} + \mathbf{E}. \tag{4.2}$$

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

$$\mathbf{S} = \mathbf{i}_n \cdot \mathbf{f}^\intercal = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{n\times 1} \begin{bmatrix} f_1 & f_2 & \cdots & f_p \end{bmatrix}_{1\times p} = \begin{bmatrix} f_1 & f_2 & \cdots & f_p \\ f_1 & f_2 & \cdots & f_p \\ \vdots & \vdots & & \vdots \\ f_1 & f_2 & \cdots & f_p \end{bmatrix}_{n\times p}. \tag{4.3}$$

In this case, a single seasonal pattern $(f_1,\ f_2,\ \cdots,\ f_p)$ repeats itself in each period. In general, the seasonal effects may change over time, we use a rank-r reduced SVD of $(\mathbf{S} - \mathbf{i}_n \cdot \mathbf{f}^\intercal)$ to

represent the time-varying seasonality:

$$
\begin{aligned}
\mathbf{S} &= \mathbf{i}_n \cdot \mathbf{f}^\intercal + \mathbf{U}\mathbf{V}^\intercal \\
&= \begin{bmatrix} 1 \\ \vdots \\ \vdots \\ 1 \end{bmatrix}_{n\times 1} \begin{bmatrix} f_1 & \cdots & f_p \end{bmatrix}_{1\times p} + \begin{bmatrix} u_{1,1} & \cdots & u_{1,r} \\ \cdots & \cdots & \cdots \\ \cdots & u_{i(t),k} & \cdots \\ \cdots & \cdots & \cdots \\ u_{n,1} & \cdots & u_{n,r} \end{bmatrix}_{n\times r} \begin{bmatrix} v_{1,1} & \cdots & v_{1,p} \\ \vdots & \vdots & \vdots \\ \vdots & v_{k,j(t)} & \vdots \\ \vdots & \vdots & \vdots \\ v_{r,1} & \cdots & v_{r,p} \end{bmatrix}_{r\times p}
\end{aligned}
\tag{4.4}
$$

where $\mathbf{U}$ is a $n \times r$ matrix, and $\mathbf{V}$ is a $p \times r$ matrix with $\mathbf{V}^\intercal\mathbf{V} = \mathbf{I}_r$ and $r \leq p$. For identification, we require the columns of $\mathbf{U}$ to be orthogonal to $\mathbf{i}_n$ or, $\mathbf{U}^\intercal\mathbf{i}_n = \mathbf{0}$, which is equivalent to $\mathbf{Q}_n^\intercal\mathbf{U} = \mathbf{U}$. The second term in the decomposition (4.4) provides an intuitive explanation for the seasonality. The j-th column vector $\mathbf{v}_j$ in $\mathbf{V}$ represents the j-th **seasonal pattern**; and the corresponding j-th column vector $\mathbf{u}_j$ in $\mathbf{U}$ is called **pattern coefficients**, since its elements delineate how the j-th seasonal pattern changes across different periods. Equations (4.2) and (4.4) comprise our basic seasonal adjustment method.

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

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

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 non-seasonal component $\mathbf{E}$ in (4.2). This smoothness also suggests the use of regularized SVD for finding the decomposition (4.2) 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.

There are two reasons that prevent direct application of Algorithm 1 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, i.e., 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, i.e., $\sum_{j=1}^{p} s_{i,j} = 0$ for each $i = 1, \cdots, n$. (Otherwise, the seasonal component would incorporate part of the overall level of the series.) Next, taking all these into account, we develop a procedure that is based on a modification of Algorithm 1.

4.2 Seasonal adjustment procedure

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 non-seasonal components. The three steps are elaborated below.

Step One: estimating seasonal pattern coefficients in U

To apply Algorithm 1, we first eliminate the fixed seasonal effects in (4.4) by pre-multiplying the column-wise de-meaning 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 $\mathbf{Q}_n\mathbf{U} = \mathbf{U}$,

$$\begin{aligned}\widetilde{\mathbf{X}} &= \mathbf{Q}_n\mathbf{S} + \mathbf{Q}_n\mathbf{E} = \mathbf{Q}_n(\mathbf{i}_n \cdot \mathbf{f}^\intercal + \mathbf{U}\mathbf{V}^\intercal) + \mathbf{Q}_n\mathbf{E} \\ &= \mathbf{U}\mathbf{V}^\intercal + \widetilde{\mathbf{E}}.\end{aligned} \tag{4.5}$$

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*, $\mathbf{f}^\intercal \mathbf{i}_p = 0$ and $\mathbf{V}^\intercal \mathbf{i}_p = \mathbf{0}_r$. 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(a) now becomes,

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

In Step 2(a) of this algorithm, pre-multiplication with $\mathbf{Q}_p$ ensures $\mathbf{v}^\intercal \mathbf{i}_p = 0$ for the zero-sum of seasonal pattern requirement, and the normalization is to ensure $\mathbf{v}^\intercal \mathbf{v} = 1$ 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 $\widetilde{\mathbf{X}} - \widehat{\mathbf{u}}\widetilde{\mathbf{v}}^\intercal$, 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 f and V

Recall that $X_T = \mathrm{Vec}(\mathbf{X}^\intercal)$. 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,

$$(\widehat{\mathbf{f}}, \widehat{\mathbf{V}}) = \arg\min_{\mathbf{f}, \mathbf{V}} \left[X_T - \mathrm{Vec}(\mathbf{f} \cdot \mathbf{i}_n^\intercal + \mathbf{V}\widehat{\mathbf{U}}^\intercal)\right]^\intercal \left[X_T - \mathrm{Vec}(\mathbf{f} \cdot \mathbf{i}_n^\intercal + \mathbf{V}\widehat{\mathbf{U}}^\intercal)\right], \tag{4.6}$$
$$\text{such that } \mathbf{f}^\intercal \cdot \mathbf{i}_p = 0, \quad \text{and} \quad \mathbf{V}^\intercal \cdot \mathbf{i}_p = \mathbf{0}_r.$$

Note that the minimization problem in (4.6) can be rewritten as,

$$\widehat{\beta} = \arg\min_{\beta} (X_T - \mathbf{Z}\beta)^\intercal (X_T - \mathbf{Z}\beta) \quad \text{with} \quad \mathbf{R}\beta = \mathbf{0}_{r+1}, \tag{4.7}$$

where

$$\begin{aligned} \mathbf{Z} &\equiv \begin{bmatrix} \mathbf{i}_n \otimes \mathbf{I}_p & \widehat{\mathbf{u}}_1 \otimes \mathbf{I}_p & \cdots & \widehat{\mathbf{u}}_r \otimes \mathbf{I}_p \end{bmatrix}, \\ \beta &\equiv (\mathbf{f}^\intercal, \mathbf{v}_1^\intercal, \cdots, \mathbf{v}_r^\intercal)^\intercal, \\ \mathbf{R} &\equiv \mathbf{I}_{r+1} \otimes \mathbf{i}_p^\intercal. \end{aligned}$$

Then the estimate $\widehat{\beta}$ can be written explicitly as,

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

where $\mathbf{b}$ is the unconstrained least squares estimate for the problem (4.7), i.e.,

$$\mathbf{b} = (\mathbf{Z}^\intercal\mathbf{Z})^{-1}\mathbf{Z}^\intercal X_T.$$

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,

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

Step Three (optional): estimating ARMA parameters in non-seasonal component E

In the second step, we obtain the estimated seasonal matrix $\widehat{\mathbf{S}}$, which can be re-written in vector form as $\{\widehat{s}_t\}_{t=1}^T$. Correspondingly, the estimated non-seasonal component can be extracted by subtracting $\widehat{s}_t$ from the original time series x_t, i.e., $\widehat{e}_t \equiv x_t - \widehat{s}_t$. If the stochastic component of x_t is assumed to follow a stationary ARMA(p, q) process

$$e_t \sim \text{ARMA}(p, q), \quad t = 1, \ldots, T, \tag{4.8}$$

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 (4.6) with the inverse of estimated variance-covariance matrix of the stochastic non-seasonal component $\widehat{e}_t$. Although such an iterated procedure could potentially improve estimation accuracy, the gain is quite limited. Thus we recommend to use the unweighted ordinary least squares estimation in (4.6). 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, this step can be omitted from the procedure.

5 Seasonal adjustment when there is stochastic trend

The basic seasonal adjustment assumes that the non-seasonal component of a seasonal time series is stationary. This section discusses the situation more commonly encountered in economic data, wherein the non-seasonal component is nonstationary and has a stochastic trend. More specifically, we assume the non-seasonal component e_t in the decomposition (4.1) $x_t = s_t + e_t$ is an integrated process, i.e., the first difference process of e_t is stationary.

To see why the basic adjustment procedure can not be applied in this situation, suppose, for example, $e_t = \sum_{l=1}^{t} \varepsilon_l$ with white noise ε_l. Re-index using the mapping I introduced in Section 3 and rewrite e_t's into an $n \times p$ matrix $\mathbf{E}$, we can easily see that the matrix $\mathbf{E}$ also has a stochastic (nonstationary) trend in each column. Considering the first column in matrix $\mathbf{E}$, $(e_{1,1}, e_{2,1}, \ldots, e_{n,1})^\intercal$, we have

$$\begin{aligned}
e_{1,1} &= \varepsilon_{1,1}, \\
e_{2,1} &= e_{1,1} + \sum_{j=2}^{p} \varepsilon_{1,j} + \varepsilon_{2,1}, \\
e_{3,1} &= e_{2,1} + \sum_{j=2}^{p} \varepsilon_{2,j} + \varepsilon_{3,1}, \\
&\vdots \\
e_{n,1} &= e_{n-1,1} + \sum_{j=2}^{p} \varepsilon_{n-1,j} + \varepsilon_{n,1}.
\end{aligned}$$

Defining $v_{1,1} = \varepsilon_{1,1}$ and new white noise $v_{i,1} \equiv \sum_{j=2}^{p} \varepsilon_{i-1,j} + \varepsilon_{i,1}$ for $i \geq 2$ and $e_{0,1} \equiv 0$, we

have

$$e_{i,1} = e_{i-1,1} + v_{i,1},$$

which means that the first column of matrix $\mathbf{E}$ also follows an I(1) process. Similarly, it can be shown that all the p columns of the matrix have stochastic trends.

Existence of a stochastic trend in each column of $\mathbf{E}$ invalidates the use of regularized SVD in the basic adjustment procedure, as examination of (4.2) indicates. When the non-seasonal component is stationary, 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 $\mathbf{E}$, the trajectory of each column of $\mathbf{E}$ also looks smooth and thus the regularized SVD will fail to separate $\mathbf{S}$ from $\mathbf{E}$. Moreover, the nonstationarity of the non-seasonal component also invalidates the use of the least squares in Step Two of the basic adjustment procedure. 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 U

We first remove the stochastic trend in the non-seasonal 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 $\mathbf{S}$.

Let $p \times (p-1)$ dimensional matrix Δ_c be the transpose of Δ given in (3.2), i.e.,

$$\Delta_c \equiv \begin{bmatrix} -1 & & & & \\ 1 & -1 & & & \\ & 1 & \ddots & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \\ & & & & 1 \end{bmatrix}_{p\times(p-1)}.$$

It works as a right-hand-side first order difference operator: multiplying a matrix with Δ_c

from the right hand side gives the first order difference of the columns.

Define

$$\mathbf{X}^{\dagger} = \mathbf{X}\Delta_c = \begin{bmatrix} x_{1,2} - x_{1,1} & \cdots & x_{1,p} - x_{1,p-1} \\ x_{2,2} - x_{2,1} & \cdots & x_{2,p} - x_{2,p-1} \\ \cdots & \cdots & \cdots \\ x_{n,2} - x_{n,1} & \cdots & x_{n,p} - x_{n,p-1} \end{bmatrix},$$

$$\mathbf{S}^{\dagger} = \mathbf{S}\Delta_c = \begin{bmatrix} s_{1,2} - s_{1,1} & \cdots & s_{1,p} - s_{1,p-1} \\ s_{2,2} - s_{2,1} & \cdots & s_{2,p} - s_{2,p-1} \\ \cdots & \cdots & \cdots \\ s_{n,2} - s_{n,1} & \cdots & s_{n,p} - y_{n,p-1} \end{bmatrix},$$

and

$$\mathbf{E}^{\dagger} = \mathbf{E}\Delta_c = \begin{bmatrix} e_{1,2} - e_{1,1} & \cdots & e_{1,p} - e_{1,p-1} \\ e_{2,2} - e_{2,1} & \cdots & e_{2,p} - e_{2,p-1} \\ \cdots & \cdots & \cdots \\ e_{n,2} - e_{n,1} & \cdots & e_{n,p} - e_{n,p-1} \end{bmatrix},$$

so that right multiplication of (4.2) by Δ_c yields

$$\mathbf{X}^{\dagger} = \mathbf{S}^{\dagger} + \mathbf{E}^{\dagger}. \tag{5.1}$$

As in the basic adjustment procedure, we represent the seasonal component matrix using a reduced SVD as in (4.4), that is, $\mathbf{S} = \mathbf{i}_n\mathbf{f}^{\intercal} + \mathbf{U}\mathbf{V}^{\intercal}$. Then

$$\mathbf{S}^{\dagger} = \mathbf{S}\Delta_c = (\mathbf{i}_n\mathbf{f}^{\intercal} + \mathbf{U}\mathbf{V}^{\intercal})\Delta_c \equiv \mathbf{i}_n\mathbf{f}^{\intercal\dagger} + \mathbf{U}\mathbf{V}^{\intercal\dagger}, \tag{5.2}$$

where

$$\mathbf{f}^{\intercal\dagger} = \begin{bmatrix} f_2 - f_1 & \cdots & f_p - f_{p-1} \end{bmatrix}, \quad \mathbf{V}^{\intercal\dagger} \equiv \mathbf{V}^{\intercal}\Delta_c = \begin{bmatrix} v_{1,2} - v_{1,1} & \dots & v_{1,p} - v_{1,p-1} \\ v_{2,2} - v_{2,1} & \dots & v_{2,p} - v_{2,p-1} \\ \vdots & & \vdots \\ v_{r,2} - v_{r,1} & \dots & v_{r,p} - v_{r,p-1} \end{bmatrix}.$$

Equations (4.4) and (5.2) show that the seasonal matrix $\mathbf{S}$ and its first order column-differenced matrix $\mathbf{S}^\dagger$ 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 $\mathbf{E}^\dagger$ is weakly stationary.

We eliminate the fixed seasonal effects in (5.2) by pre-multiplying the column-wise demeaning matrix $\mathbf{Q}_n$ by the first order column-differenced data matrix $\mathbf{X}^\dagger$ to obtain $\widetilde{\mathbf{X}}^\dagger = \mathbf{Q}_n\mathbf{X}^\dagger$. Since $\mathbf{Q}_n\mathbf{i}_n = \mathbf{0}_n$ and $\mathbf{Q}_n\mathbf{U} = \mathbf{U}$,

$$\begin{aligned} \widetilde{\mathbf{X}}^\dagger &\equiv \mathbf{Q}_n\mathbf{S}^\dagger + \mathbf{Q}_n\mathbf{E}^\dagger = \mathbf{Q}_n(\mathbf{i}_n \cdot \mathbf{f}^{\intercal\dagger} + \mathbf{U}\mathbf{V}^{\intercal\dagger}) + \mathbf{Q}_n\mathbf{E}^\dagger \\ &= \mathbf{U}\mathbf{V}^{\intercal\dagger} + \widetilde{\mathbf{E}}^\dagger. \end{aligned} \tag{5.3}$$

We repeatedly apply Algorithm 1 r times to the matrix $\mathbf{X}^\dagger$ (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 f and 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 $X_T = \text{Vec}(\mathbf{X}^\intercal)$, where Δ is the differencing operator.

The constrained least squares problem is

$$(\widehat{\mathbf{V}}, \widehat{\mathbf{f}}) = \arg\min_{\mathbf{V},\mathbf{f}} \left[\Delta X_T - \Delta\,\mathrm{Vec}(\mathbf{f}\cdot\mathbf{i}_n^\intercal + \mathbf{V}\widehat{\mathbf{U}}^\intercal)\right]^\intercal \left[\Delta X_T - \Delta\,\mathrm{Vec}(\mathbf{f}\cdot\mathbf{i}_n^\intercal + \mathbf{V}\widehat{\mathbf{U}}^\intercal)\right],$$
$$\text{such that } \mathbf{V}^\intercal\mathbf{i}_p = \mathbf{0}_r \quad \text{and} \quad \mathbf{f}^\intercal\cdot\mathbf{i}_p = 0. \tag{5.4}$$

This minimization problem in (5.4) can also be written as,

$$\widehat{\beta} = \arg\min_{\beta}(\Delta X_T - \Delta\mathbf{Z}\beta)^\intercal(\Delta X_T - \Delta\mathbf{Z}\beta) \quad \text{with} \quad \mathbf{R}\beta = \mathbf{0}_{r+1}, \tag{5.5}$$

where

$$\begin{aligned}
\mathbf{Z} &\equiv \left[\mathbf{i}_n\otimes\mathbf{I}_p \quad \widehat{\mathbf{u}}_1\otimes\mathbf{I}_p \quad \cdots \quad \widehat{\mathbf{u}}_r\otimes\mathbf{I}_p\right],\\
\beta &\equiv (\mathbf{f}^\intercal, \mathbf{v}_1^\intercal, \cdots, \mathbf{v}_r^\intercal)^\intercal,\\
\mathbf{R} &\equiv \mathbf{I}_{r+1}\otimes\mathbf{i}_p^\intercal.
\end{aligned}$$

After solving the constrained least squares problem above, we obtain the estimated seasonal component as $\widehat{\mathbf{S}} = \mathbf{i}_n\widehat{\mathbf{f}}^\intercal + \widehat{\mathbf{U}}\widehat{\mathbf{V}}^\intercal$.

Step Three (optional): estimating parameters in non-seasonal component E

If we assume that the non-seasonal 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 $\widehat{e}_t = x_t - \widehat{s}_t$. Based on the fitted ARIMA model, a feasible GLS estimation of $\mathbf{f}$ and $\mathbf{V}$ can be obtained by weighting the least squares in (5.4) with the inverse of the estimated variance-covariance matrix of the differenced non-seasonal component $\Delta\widehat{e}_t$. Although such an update (or its iterative version) could potentially improve estimation accuracy, the gain is quite limited and thus in this paper we only consider the unweighted ordinary least squares estimation in (5.4). In general, the ARIMA model is not of interest for seasonal adjustment, and this step can be omitted from the procedure.

6 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 seasaonal 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 5 as an example to show how to modify it. The basic seasonal adjustment procedure in Section 4 can be modified in a similar manner.

In Section 5, 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 (5.3), which is

$$\widetilde{\mathbf{X}}^{\dagger} = \mathbf{U}\mathbf{V}^{\intercal\dagger} + \widetilde{\mathbf{E}}^{\dagger}.$$

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 non-smooth 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* $\widetilde{\mathbf{X}}^{\dagger} = \mathbf{Q}_n \mathbf{X} \Delta_c$*, and*

(2) the updating equation in Step 2(b) now becomes two equations that update $\mathbf{u}_1$ *and* $\mathbf{u}_2$ *separately by applying the Step 2(b) of the original algorithm to the first* ℓ *rows and the last* $n - \ell$ *rows of* $\widetilde{\mathbf{X}}^{\dagger}$ *respectively.*

In Algorithm 3, using different smoothing parameters for the first ℓ elements and last $n-\ell$ elements of $\mathbf{u}$ enhances the flexibilty 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 5 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 $\widehat{\mathbf{S}}(\ell) = \mathbf{i}_n \widehat{\mathbf{f}}(\ell)^{\intercal} + \widehat{\mathbf{U}}(\ell)\widehat{\mathbf{V}}(\ell)^{\intercal}$.

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 $3 \leq \ell \leq n-3$. 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 $\#(\mathcal{L}) = (n - 5 + 1)^r$.

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:

$$\widehat{\boldsymbol{\ell}} \;=\; \arg\min_{\boldsymbol{\ell} \in \mathcal{L}} \frac{1}{T-1} \sum_{t=2}^{T} [\Delta x_t - \Delta \widehat{s}_t(\boldsymbol{\ell})]^2, \tag{6.1}$$

Note the criterion equals

$$\frac{1}{T-1}\sum_{t=2}^{T}[\Delta s_t - \Delta\widehat{s}_t(\boldsymbol{\ell})]^2 + \frac{1}{T-1}\sum_{t=2}^{T}\Delta e_t^2 + \frac{2}{T-1}\sum_{t=2}^{T}[\Delta s_t - \Delta\widehat{s}_t(\boldsymbol{\ell})]\Delta e_t. \tag{6.2}$$

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 (6.2), 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).

7 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, 2016), 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) and X13, which refers to the use of the TRAMO/SEATS methodology within the X-13ARIMA-SEATS program. The data generating processes follow (4.1), and include seasonality with and without abrupt breaks and stationary/nonstationary ARIMA error terms.

7.1 Seasonality without abrupt breaks

We consider a deterministic monthly seasonal component

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

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

$$\begin{aligned} \mathbf{b} &= (1 + 1/10, \cdots, 1 + i/10, \cdots, 1 + n/10)^\intercal, \\ \mathbf{a} &= (-1.25, -2.25, -1.25, 0.75, -1.25, -0.25, 2.75, -0.25, 0.75, -0.25, 0.75, 1.75)^\intercal. \end{aligned}$$

The vector $\mathbf{a}$ represents the reoccurring variation within each seasonal period. The magnitude of the seasonal component, captured by the multipliers in $\mathbf{b}$, increases slowly every year in linear fashion. The seasonality can be also expressed in matrix form as follows,

$$\mathbf{S}^o = \mathbf{b}\mathbf{a}^\intercal \equiv \mathbf{i}_n \cdot \mathbf{f}^\intercal + \mathbf{u}\mathbf{v}^\intercal \tag{7.1}$$

where $\mathbf{v} = \mathbf{a}$, $\mathbf{f} = \overline{\mathbf{b}} \cdot \mathbf{a}^\intercal$, $\mathbf{u} = \mathbf{b} - \overline{\mathbf{b}}$, with $\overline{\mathbf{b}} = n^{-1} \sum_{i=1}^{n} b_i$. In Figure 1, 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 $\mathbf{i}_n\mathbf{f}^\intercal$ and $\mathbf{u}\mathbf{v}^\intercal$ in lower-left panel, and total seasonality $\mathbf{S}^o$ in lower-right panel.

[INSERT FIGURE 1]

For non-seasonal component, we consider three different data generating processes for stochastic non-seasonal component e_t:

- DGP1: $e_t \sim i.i.d.\, N(0, 1)$, for all t,
- DGP2: $e_t \sim \text{ARMA}(1, 1)$, with $\phi = 0.8$ and $\psi = 0.1$ with $N(0, 1)$ innovations.
- DGP3: $e_t \sim \text{ARIMA}(1, 1, 1)$, with $\phi = 0.8$ and $\psi = 0.1$ with $N(0, \sigma^2)$ innovations and $\sigma^2 = 0.04$.

In DGP1 and DGP2, the non-seasonal component is stationary with standard normally distributed innovations. In DGP2, there exists ARMA(1,1) linear time dependence in e_t, while there is no time dependence in DGP1. Given the nonstationarity of DGP3, we set the variance of innovation in DGP3 to be small so that the sample unconditional variance of simulated series e_t is not too large (practically infinite) in finite sample.

After the original seasonal component s_t^o and non-seasonal component e_t are generated, we

scale the original seasonal component s_t^o by w, to obtain the working seasonal component s_t,

$$s_t = ws_t^o \equiv \kappa\sqrt{\frac{\mathrm{var}(e_t)}{\mathrm{var}(s_t^o)}}s_t^o,$$

so as to control the sample unconditional standard deviation ratio $\sqrt{\mathrm{var}(s_t)/\mathrm{var}(e_t)}$ of simulated time series data $x_t = s_t + e_t$ to be exactly κ in each replication of DGPs. For stationary DGP1 and DGP2, we choose $\kappa = 0.2, 0.4, 0.6, 0.8, 1, 1.2, 1.4, 1.6, 1.8, 2$ in our simulation setups; for nonstationary DGP3, we choose $\kappa = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1$ in our simulation setups. For each combination of DGPs and κ values, we simulate monthly time series data with sample size $T = 600$ (i.e., $n = 50$ and $p = 12$). We repeat the simulation $B = 500$ times for each setup.

We use two evaluation criteria to compare our proposed seasonal adjustment methods with the X13 method on the accuracy of estimation of the seasonal component. They are the mean square errors (MSE) and mean percentage errors (MPE):

$$\mathrm{MSE} = E[(\hat{s}_t - s_t)^2], \quad \mathrm{MPE} = E\left|\frac{\hat{s}_t - s_t}{s_t}\right| \times 100\%,$$

where $\hat{s}_t$ is the estimated seasonal component, and they capture absolute and relative losses respectively. From the simulation, the average values of these two criteria are calculated by

$$\mathrm{AMSE} = \frac{1}{B}\sum_{b=1}^{B}\left(\frac{1}{T}\sum_{t=1}^{T}(\hat{s}_t^{(b)} - s_t)^2\right), \quad \mathrm{AMPE} = \frac{1}{B}\sum_{b=1}^{B}\left(\frac{1}{T}\sum_{t=1}^{T}\left|\frac{\hat{s}_t^{(b)} - s_t}{s_t}\right|\right) \times 100\%,$$

where b is the b-th replication and B is the total number of replications.

[INSERT TABLE 1]

Table 1 reports the comparison results between X13 and our RSVD methods. Several findings are in order. First, for all three DGPs, the absolute loss (AMSE) of the X13 method keeps increasing as the ratio κ increases, while that of our RSVD method keeps decreases to a stable value. Second, for all three DGPs, the relative loss (AMPE) of both the X13 and our RSVD methods keeps decreasing as κ increases, and AMPE of our RSVD method decreases faster. Third, for stationary DGP1 and DGP2 cases the X13 method has smaller losses of AMSE and AMPE. However, as κ increases, our RSVD method outperforms X13 in both

AMSE and AMPE criteria; for nonstationary DGP3 cases, our RSVD method uniformly dominates the X13 method by delivering smaller losses of AMSE and AMPE regardless of κ values.

7.2 Seasonality with abrupt breaks

Now we consider a deterministic monthly seasonal component with a non-smooth 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 $\mathbf{b} = (b_1, \ldots, b_n)^\intercal$ and $\mathbf{a} = (a_1, \ldots, a_{12})^\intercal$ take the following values,

$$\begin{aligned} b_i &= \begin{cases} 1 + i/10, & \text{if } 1 \leq i \leq n/2, \\ 1 + (n+1-i)/5, & \text{if } n/2+1 \leq i \leq n. \end{cases} \\ \mathbf{a} &= (-1.25, -2.25, -1.25, 0.75, -1.25, -0.25, 2.75, -0.25, 0.75, -0.25, 0.75, 1.75)^\intercal. \end{aligned}$$

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 $\mathbf{s}^o = \mathbf{b}\mathbf{a}^\intercal = \mathbf{i}_n\mathbf{f}^\intercal + \mathbf{u}\mathbf{v}^\intercal$ where the terms are defined the same way in (7.1). In Figure 2, 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 $\mathbf{i}_n\mathbf{f}^\intercal$ and $\mathbf{u}\mathbf{v}^\intercal$ in lower-left panel, and total seasonality $\mathbf{s}^o$ in lower-right panel.

[INSERT FIGURE 2]

For the non-seasonal 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

$$x_t = s_t + e_t \equiv \kappa \sqrt{\frac{\mathrm{var}(e_t)}{\mathrm{var}(s_t^b)}} s_t^b + e_t,$$

and the sample unconditional standard deviation ratio $\sqrt{\mathrm{var}(s_t)/\,\mathrm{var}(e_t)}$ is fixed to be exactly κ in each replication of DGPs. For nonstationary DGP3, we choose $\kappa = 0.1$, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1 in our setups. For each combination of DGP3 and κ values, we simulate monthly time series data with sample size $T = 600$ (i.e., $n = 50$ and $p = 12$). We repeat the simulation $B = 500$ times for each setup.

[INSERT TABLE 2]

Table 2 reports the comparison results between X13, RSVD without break, and RSVD allowing for break. Both RSVD methods outperforms X13 by delivering smaller absolute and relative losses, and the RSVD allowing for break has the smallest errors among the three methods. The absolute loss (AMSE) of RSVD and X13 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.

8 Real data

In the section, we use some real time series data with seasonal behaviors to compare our proposed RSVD seasonal adjustment method with the X13 method. 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*). The three empirical examples showcase that our proposed seasonal adjustment method could produce (i) similar seasonal components as the X13 method does; (ii) better seasonal components when X13 fails; and (iii) seasonal components for other than quarterly and monthly frequencies, such as daily and weekly.

We apply X13 to the first two time series data sets, and also employ RSVD method with 3

seasonal patterns ($r = 3$), allowing for a non-smooth break in each of the three corresponding left singular vectors with roughness penalties for all the three time series data sets.

8.1 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.

[INSERT FIGURE 3]

[INSERT FIGURE 4]

Figure 3 and 4 show and compare the seasonal adjustment results of the retail time series data using both the X13 and RSVD methods. In Figure 3(a) and (b), 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 coefficients in $\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$. The red, green, and blue vertical lines in Figure 3(b) represents the abrupt break detected in $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ respectively. Figure 3(c) presents the fixed seasonality $\mathbf{i}_n \cdot \mathbf{f}^\intercal$ and the time-varying seasonality $\sum_{r=1}^{3} \mathbf{u}_r \mathbf{v}_r^\intercal$. In Figure 3(d), we plot the two seasonal components extracted by X13 and the RSVD method. Finally, Figure 4 shows the original time series and seasonal adjustments by the X13 and RSVD methods.

First, in Figure 3(c) we find that the fixed seasonal component is larger than the time-varying seasonal component. Second, the seasonal component s_t extracted via X13 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 3(d) and Figure 4. In the period between red and blue vertical lines, the RSVD seasonal component is slightly more volatile than the X13 seasonal component, and the RSVD seasonal adjusted series is slightly smoother than the X13 seasonal adjusted series. In the other periods, the RSVD seasonal components and adjusted time series are almost the same as their X13 counterparts.

8.2 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.

[INSERT FIGURE 5]

[INSERT FIGURE 6]

Figure 5 and 6 show and compare the seasonal adjustment results of the berry production time series data using both X13 and RSVD. In Figure 5(a) and (b), 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 coefficients in $\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$. The fixed seasonal pattern has a much larger scale than the time-varying seasonal patterns, so it is associated with the right vertical axis. The red, green, and blue vertical lines in Figure 5(b) represents the abrupt break detected in $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ respectively. Figure 5(c) presents the fixed seasonality $\mathbf{i}_n \cdot \mathbf{f}^\intercal$ and the time-varying seasonality $\sum_{r=1}^{3} \mathbf{u}_r \mathbf{v}_r^\intercal$. In Figure 5(d), we plot the two seasonal components extracted by X13 and the RSVD method. Finally, Figure 6 shows the original time series and seasonal adjusted ones by the X13 and RSVD methods.

Because of the strong seasonal behavior of the New Zealand agricultural industry, the actual berry production in the fall quarter is very close to zero. It turns out that the X13 method does not deliver a stable and plausible seasonal component: the X13 seasonal component in Figure 5(d) is excessively negative at certain periods, and therefore the X13 seasonal adjusted series in Figure 6 is excessively high at those periods. In contrast, the RSVD method produces much more reasonable seasonal component, and the seasonal adjusted series is much more smoother than the original time series. Moreover, just like the retail volume data, the fixed seasonal component is much more salient than the time-varying component, and dominates the seasonality.

8.3 Online submission count data

Lastly, 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.

[INSERT FIGURE 7]

[INSERT FIGURE 8]

Figure 7 and 8 show the seasonal adjustment results of the submission counts time series data using the RSVD method. In Figure 7(a) and (b), 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 coefficients in $\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$. The red, green, and blue vertical lines in Figure 7(b) represents the abrupt break detected in $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ respectively. Figure 7(c) presents the fixed seasonality $\mathbf{i}_n \cdot \mathbf{f}^\intercal$ and the time-varying seasonality $\sum_{r=1}^{3} \mathbf{u}_r \mathbf{v}_r^\intercal$. In Figure 7(d), we plot the seasonal component extracted by the RSVD method.

Figure 8 plots the seasonal adjustment results of the daily online submission count data in logarithms. Because the data occurs at a daily frequency, the X13 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 RSVD method is still well applicable to the daily data with weekly seasonality. In Figure 7(c), the fixed and time-varying seasonal components have similar magnitude. In Figure 7(d), the RSVD seasonal component shows that the seasonal behavior is quite different at the beginning, middle, and end of the time series. In Figure 8, the seasonal adjusted series is much smoother than the original series: the submission counts series increases and reaches its peak in the first week, decrease in the second week, and first increase 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.

9 Conclusion

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 paper 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, 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.

References

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

[2] Bunzel, H., & 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(2-3), 345-366.

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

[4] Cai, Z., & Chen, R. (2006). Flexible seasonal time series models. *Econometric Analysis of Financial and Economic Time Series,* 20, 63-87.

[5] Canova, F. (1992). An alternative approach to modeling and forecasting seasonal time series. *Journal of Business & Economic Statistics,* 10(1), 97-108.

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

[7] Burridge, P., & Wallis, K. F. (1990). Seasonal adjustment and Kalman filtering: extension to periodic variances. *Journal of Forecasting,* 9(2), 109-118.

[8] Findley, D.F. (2005). Some recent developments and directions in seasonal adjustment. *Journal of Official Statistics* **21**, 343–365.

[9] Findley, D. F., Monsell, B. C., Bell, W. R., Otto, M. C., & Chen, B. C. (1998). New capabilities and methods of the X-12-ARIMA seasonal-adjustment program. *Journal of Business & Economic Statistics,* 16(2), 127-152.

[10] Franses, P. H. (1998). Time series models for business and economic forecasting. Cambridge university press.

[11] Gersovitz, M., & MacKinnon, J. G. (1978). Seasonality in regression: An application of smoothness priors. *Journal of the American Statistical Association,* 73(362), 264-273.

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

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

[14] Hansen, L. P., & Sargent, T. J. (1993). Seasonality and approximation errors in rational expectations models. *Journal of Econometrics,* 55(1), 21-55.

[15] Harrison, P. J., & Stevens, C. F. (1976), Bayesian forecasting. *Journal of Royal Statistical Society, Series B,* 38, 205-247.

[16] Harvey, A. C. (1990). Forecasting, structural time series models and the Kalman filter. Cambridge university press.

[17] Harvey, A., & Scott, A. (1994). Seasonality in dynamic regression models. *The Economic Journal,* 1324-1345.

[18] Huang, J. Z., Shen, H., & Buja, A. (2008). Functional principal components analysis via penalized rank one approximation. *Electronic Journal of Statistics,* Vol 2: 678-695.

[19] Huang, J. Z., Shen, H., & 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.

[20] Hylleberg, S., Engle, R. F., Granger, C. W., & Yoo, B. S. (1990). Seasonal integration and cointegration. *Journal of econometrics,* 44(1), 215-238.

[21] Kohn, R., Ansley, C. F., & Wong, C. (1992). Nonparametric spline regression with autoregressive moving average errors. *Biometrika,* 79, 335-346.

[22] Lovell, M. C. (1963). Seasonal adjustment of economic time series and multiple regression analysis. *Journal of the American Statistical Association,* 58(304), 993-1010.

[23] McElroy, T., & Monsell, B. (2015). Model estimation, prediction, and signal extraction for nonstationary stock and flow time series observed at mixed frequencies. *Journal of the American Statistical Association,* 110(511), 1284-1303.

[24] Monsell, B. C., Lytras, D. P., & Findley, D. F. (2013). Getting Started with X-13ARIMA-SEATS Input Files (Accessible Version), U.S. Census Bureau.

[25] Osborn, D. R. (1991). The implications of periodically varying coefficients for seasonal time-series processes. *Journal of Econometrics,* 48(3), 373-384.

[26] Osborn, D. R. (1993). Discussion: Seasonal cointegration. *Journal of Econometrics,* 55(1), 299-303.

[27] Proietti, T. (2004). Seasonal specific structural time series. *Studies in Nonlinear Dynamics & Econometrics,* 8(2).

[28] Shiskin, J., Young, A.H., and Musgrave, J.C. (1965). The X-11 variant of the Census method II seasonal adjustment program. No. 15. US Department of Commerce, Bureau of the Census.

[29] Sims, C. A. (1974). Seasonality in regression. *Journal of the American Statistical Association,* 69(347), 618-626.

[30] U.S. Census Bureau (2016). "X-13ARIMA-SEATS Reference Manual". http://www.census.gov/ts/x13as/docX13AS.pdf

Table 1: Evaluation of estimates of seasonal component with no break

		AMSE$\times 10^{-2}$		AMPE(%)	
	κ	X13	RSVD	X13	RSVD
S_t^o and DGP1: $e_t \sim$ i.i.d. $N(0, \sigma^2)$.	0.2	3.8027	4.6657	170.5288	254.4890
	0.4	3.6791	4.1408	102.6348	121.0693
	0.6	4.1825	4.0750	76.0336	81.0652
	0.8	4.5937	3.9338	60.7023	59.8170
	1.0	4.9999	3.8731	51.2793	47.4248
	1.2	5.2961	3.7851	44.5730	38.9227
	1.4	5.9294	3.8602	40.0335	33.5560
	1.6	6.1705	3.7273	36.1888	29.4866
	1.8	6.3099	3.6876	33.1291	25.8607
	2.0	6.8946	3.7938	31.1219	23.5752
S_t^o and DGP2: $e_t \sim$ ARMA(1,1), $\phi = 0.8$, $\psi = 0.1$, and $N(0, 1)$ innovations.	0.2	3.3401	5.0410	133.2620	174.1002
	0.4	4.2161	4.4404	79.2533	83.2970
	0.6	5.3341	4.2963	59.7956	54.9942
	0.8	6.0918	4.1395	48.3261	41.0960
	1.0	6.8841	4.1380	41.5572	32.8152
	1.2	7.5151	4.0523	35.7913	26.6444
	1.4	8.2039	4.0463	32.4293	23.1987
	1.6	9.4137	4.2533	30.3949	20.8738
	1.8	9.9428	4.1257	27.6120	17.9895
	2.0	10.5706	4.1151	25.5900	16.3031
S_t^o and DGP3: $e_t \sim$ ARIMA(1,1,1), $\phi = 0.8$, $\psi = 0.1$, and $N(0, 0.04)$ innovations.	0.1	1.5272	0.3819	39.5679	21.5201
	0.2	2.3386	0.3826	25.3175	11.0715
	0.3	3.2513	0.3863	19.4587	7.1949
	0.4	4.0294	0.3957	16.1756	5.5927
	0.5	4.7071	0.3983	14.0200	4.3952
	0.6	5.8034	0.4003	12.5008	3.6737
	0.7	6.3456	0.3735	11.2921	2.9698
	0.8	6.9465	0.3679	10.3224	2.5888
	0.9	7.2107	0.3870	9.6452	2.4161
	1.0	7.8357	0.3777	9.1788	2.1820

Table 2: Evaluation of estimates of seasonal component with break

		AMSE$\times 10^{-2}$			AMPE(%)		
	κ	X13	RSVD	RSVD-b	X13	RSVD	RSVD-b
S_t^b and DGP3: $e_t \sim$ ARIMA(1,1,1), $\phi = 0.8$, $\psi = 0.1$, and $N(0, 0.04)$ innovations.	0.1	2.5198	0.6291	0.5677	52.0431	22.6619	22.8086
	0.2	4.2578	0.9121	0.5423	31.5513	10.6038	10.5901
	0.3	5.7358	1.4183	0.5526	24.0914	7.3358	7.1500
	0.4	6.9938	2.0564	0.5681	19.9925	5.7384	5.5052
	0.5	8.3978	3.0317	0.5648	16.5410	4.5683	4.2532
	0.6	10.3549	3.8793	0.5442	14.9055	4.0221	3.6791
	0.7	11.3345	5.0405	0.5470	13.0074	3.5411	3.1538
	0.8	13.0095	6.7009	0.5380	11.3756	3.1185	2.6947
	0.9	14.9500	8.6697	0.5431	10.1673	2.7676	2.3334
	1.0	16.9045	9.5812	0.5511	9.7780	2.6702	2.2164

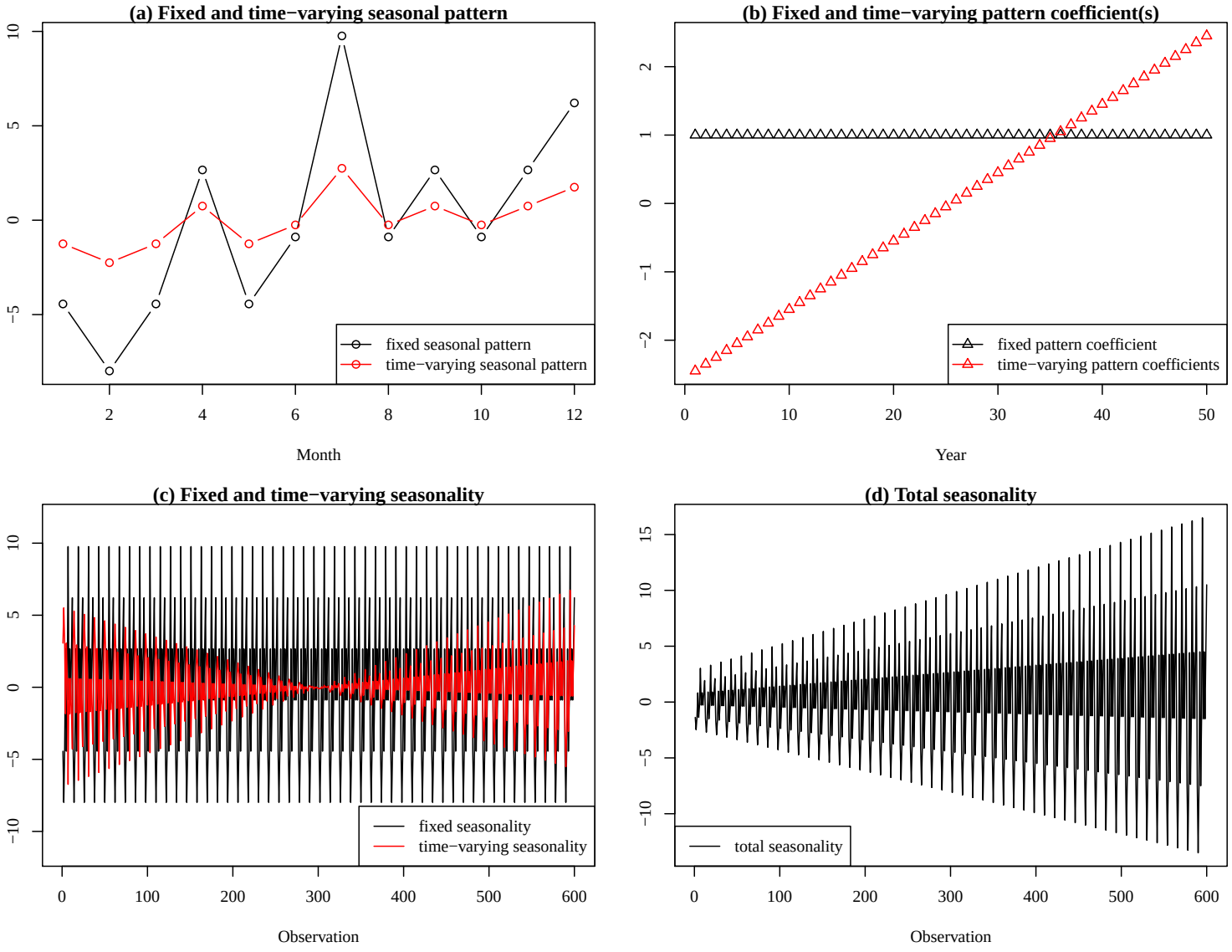

Figure 1: Simulated seasonal component without break

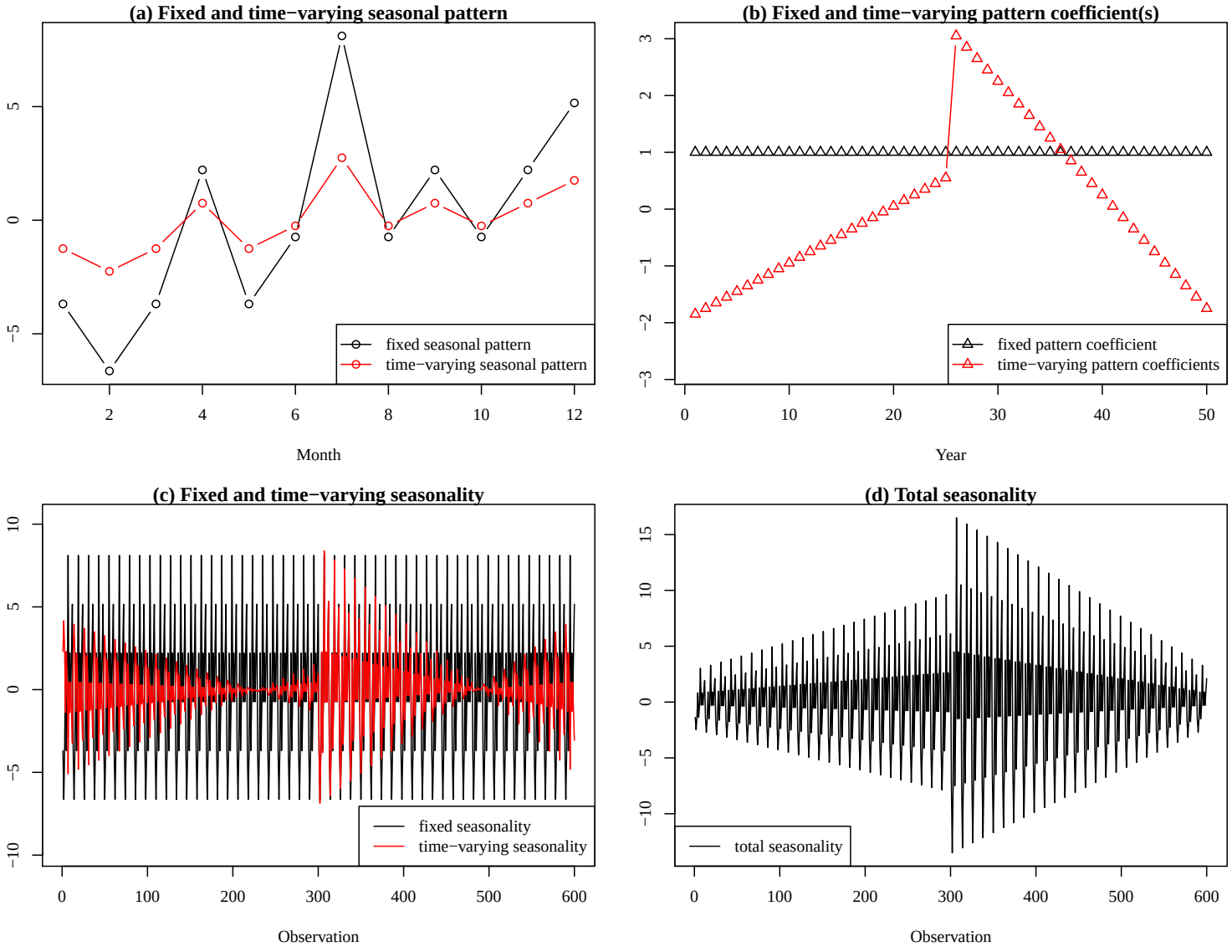

Figure 2: Simulated seasonal component with break

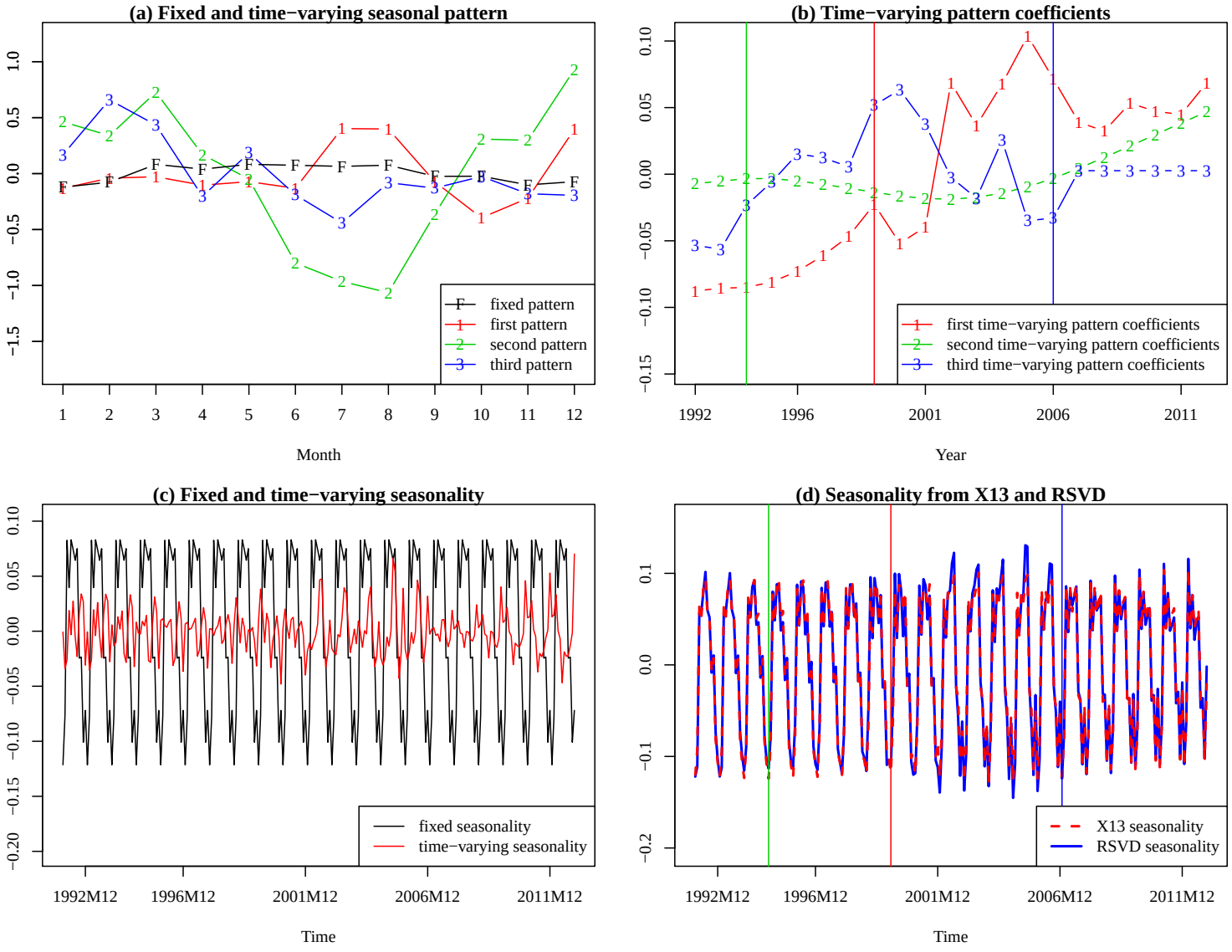

Figure 3: Time series plots of (i) original and seasonal adjusted data, (ii) estimated seasonal components, (iii) left singular vectors **u**'s, and (iv) right singular vectors **v**'s.

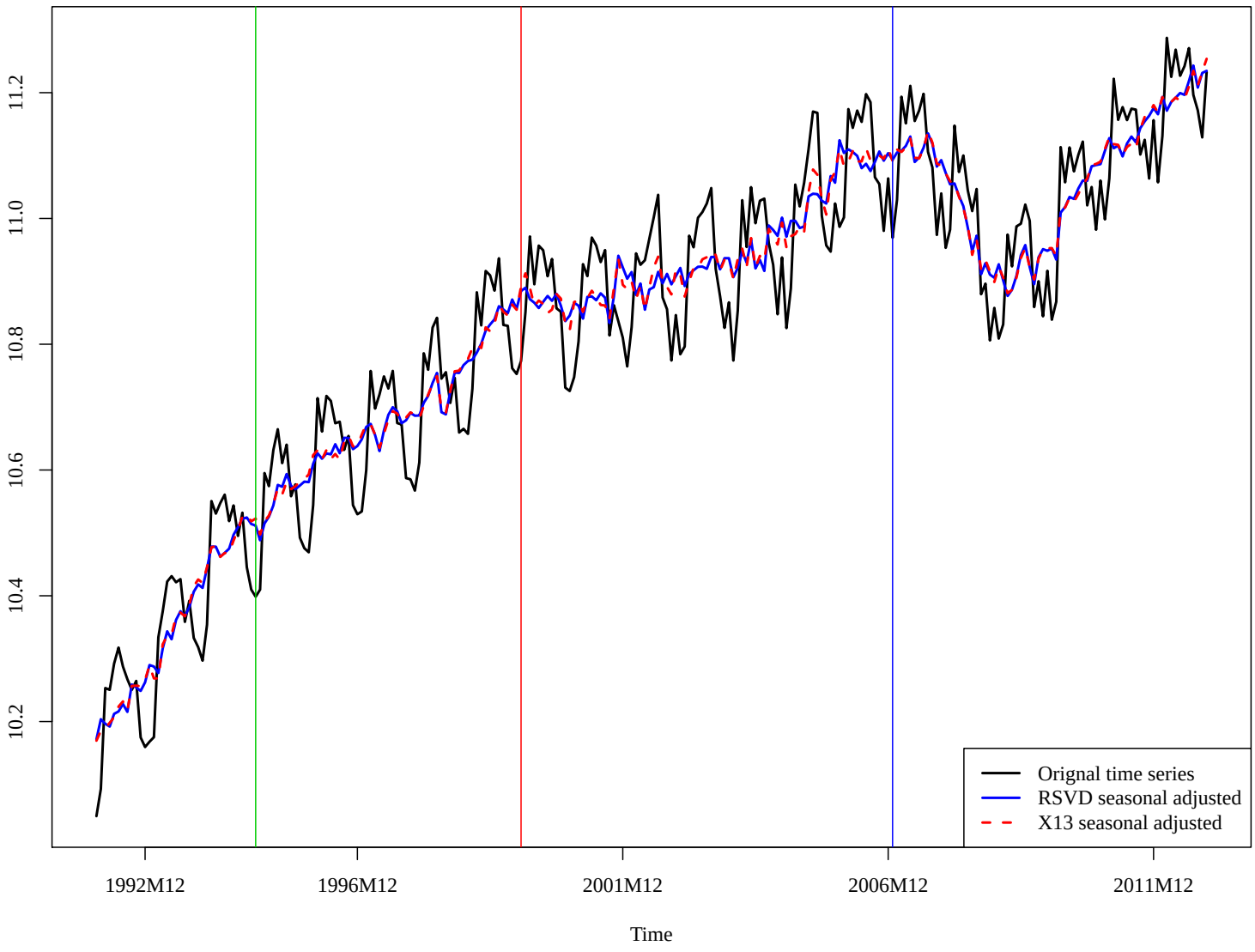

Figure 4: Time series plots of (i) original and seasonal adjusted data

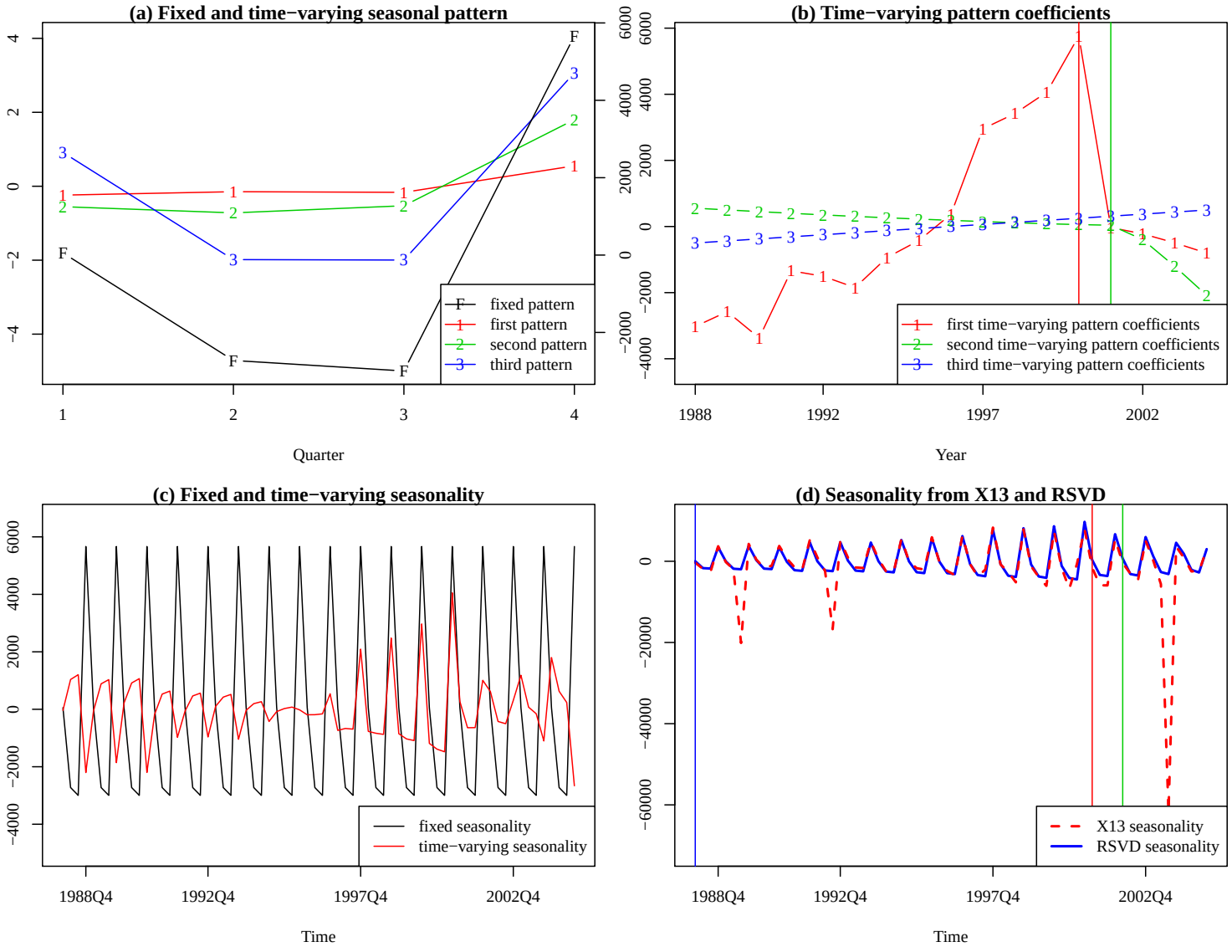

Figure 5: Time series plots of (i) original and seasonal adjusted data, (ii) estimated seasonal components, (iii) left singular vectors **u**'s, and (iv) right singular vectors **v**'s.

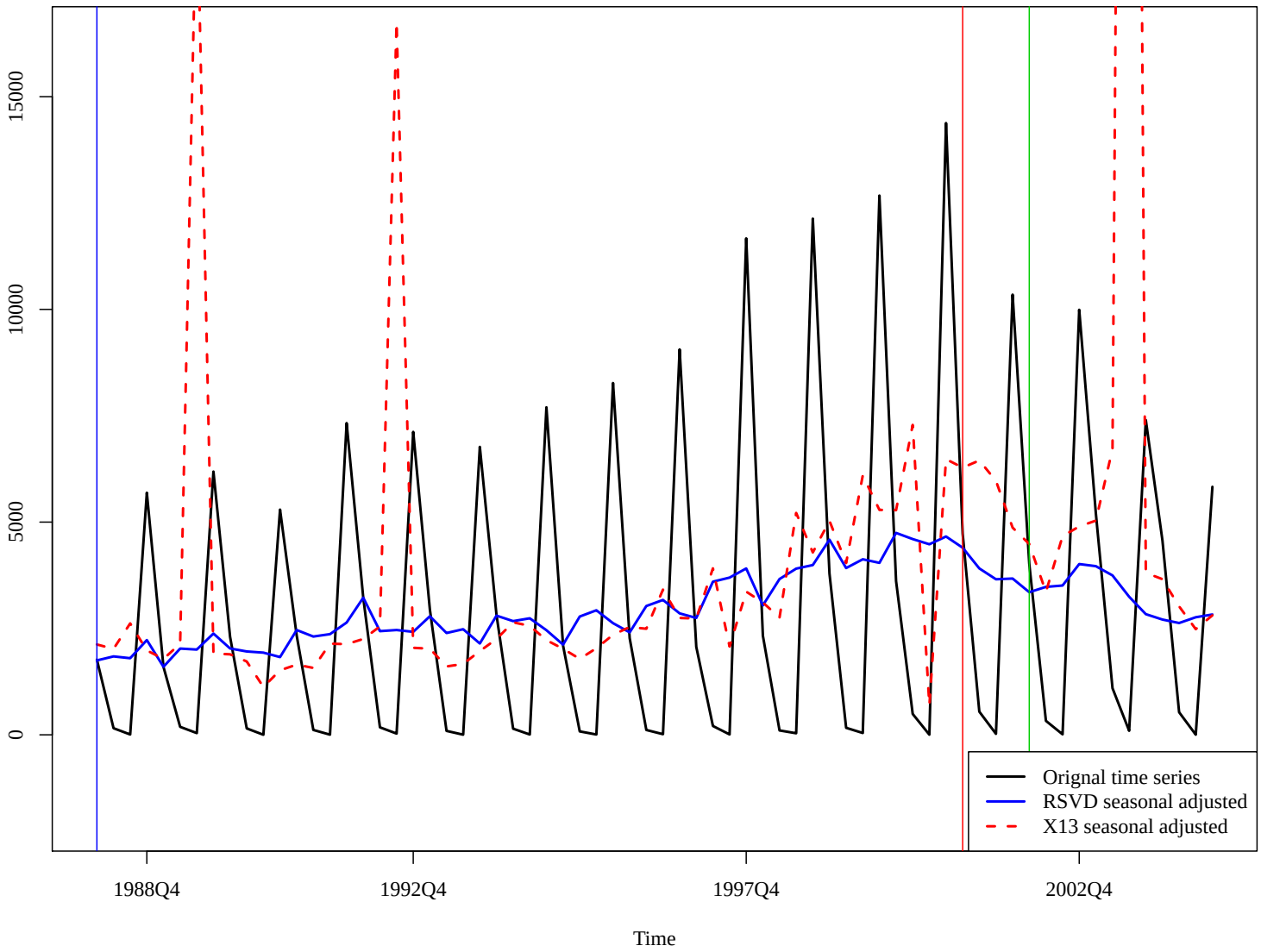

Figure 6: Time series plots of (i) original and seasonal adjusted data

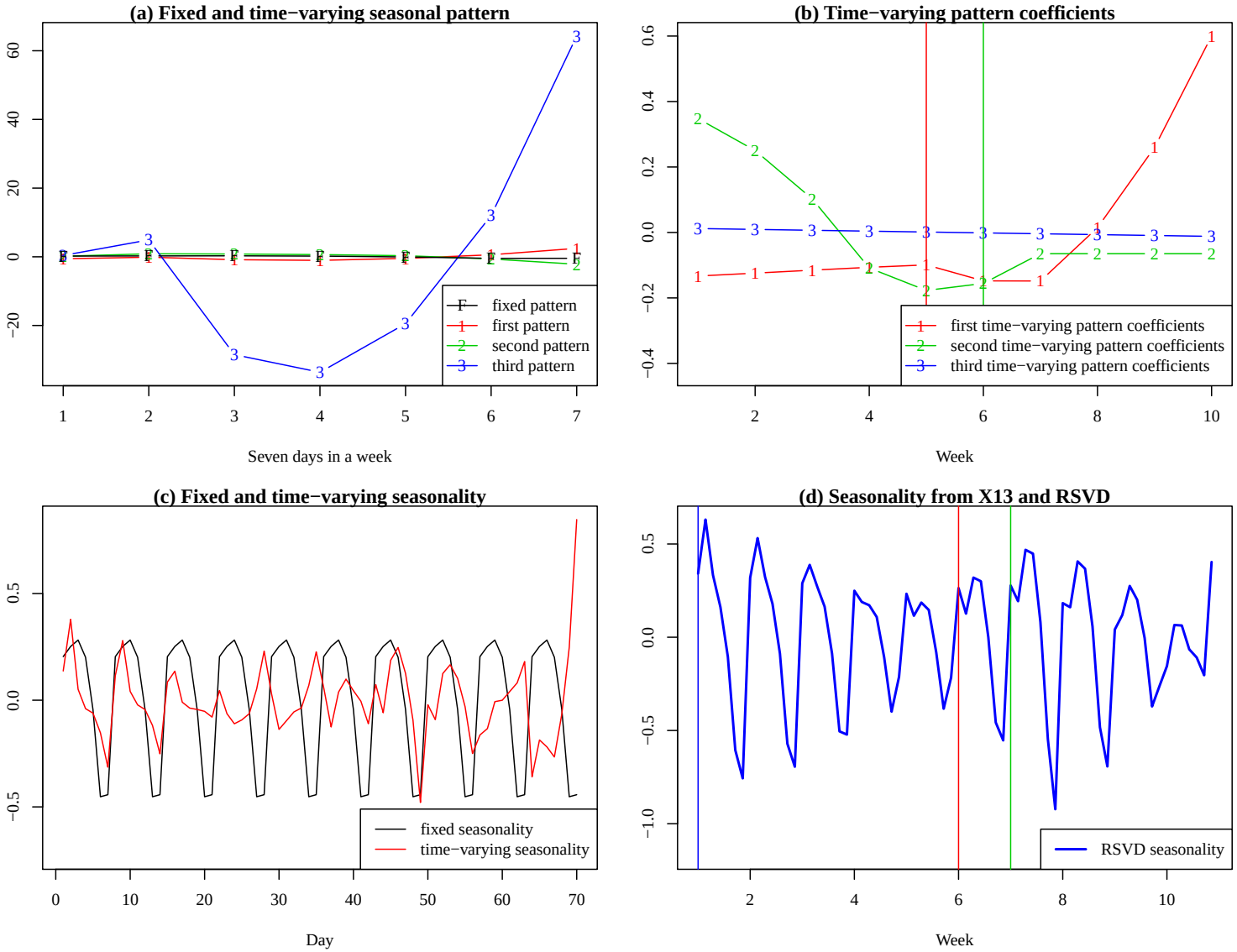

Figure 7: Time series plots of (i) original and seasonal adjusted data, (ii) estimated seasonal components, (iii) left singular vectors **u**'s, and (iv) right singular vectors **v**'s.

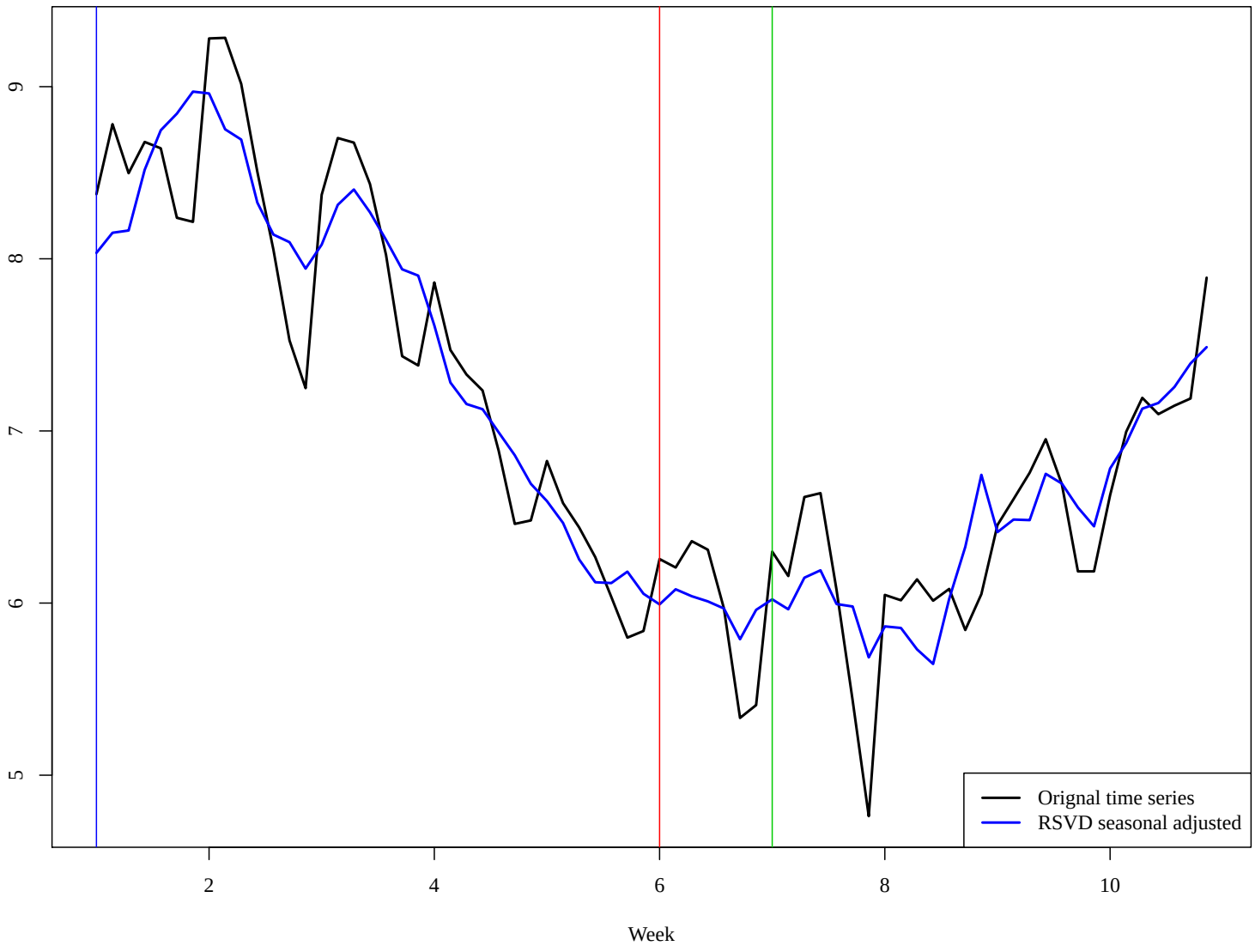

Figure 8: Time series plots of (i) original and seasonal adjusted data