Online Appendices for "Time series seasonal adjustment using regularized singular value decomposition"

By Wei Lin, Jianhua Z. Huang, and Tucker McElroy

S.1 Derivation of the GCV criterion in Algorithm 1

The Step 2(b) of Algorithm 1 is the key part of this algorithm, which imposes roughness penalty on the left singular vector $\mathbf{u}$, and uses the criterion (2.3) to select the smoothing parameter. In this document, we show that (2.3) can be derived from leave-out-one-period cross-validation.

We observe that, given $\mathbf{v}$, the Step 2(b) in Algorithm 1 is equivalent to a ridge regression

$$\widehat{\mathbf{u}} = \arg\min_{\mathbf{u}} \|\mathbf{y} - \mathbf{Z}\mathbf{u}\|^2 + \alpha \mathbf{u}^\intercal \boldsymbol{\Omega} \mathbf{u} \tag{S.1}$$

with $\mathbf{y} \equiv \mathrm{Vec}(\mathbf{X}^\intercal)$ and $\mathbf{Z} \equiv \mathbf{I}_n \otimes \mathbf{v}$, and α is the smoothing parameter. Suppose the data matrix $\mathbf{X}$ has n rows and m columns. If α is fixed, the $\widehat{\mathbf{u}}$ defined in (S.1) is

$$\widehat{\mathbf{u}} = (\mathbf{Z}^\intercal\mathbf{Z} + \boldsymbol{\Omega}_\alpha)^{-1}\mathbf{Z}^\intercal\mathbf{y} = \mathbf{M}(\alpha)\mathbf{Z}^\intercal\mathbf{y},$$

and

$$\widehat{\mathbf{y}} = \mathbf{Z}\mathbf{M}(\alpha)\mathbf{Z}^\intercal\mathbf{y} = \mathbf{H}\mathbf{y}.$$

Consider the cross-validation that removes one row (one period) of $\mathbf{X}$ at a time. It corresponds to remove a block of size m from $\mathbf{y}$ at a time. Note that $\mathbf{y}$ and $\mathbf{H}$ can be partitioned into n blocks of equal size $m \times 1$ and $m \times m$ respectively. Let $\widehat{\mathbf{u}}^{(-i)}$ minimize (S.1) with the i-th block of $\mathbf{y}$ and the corresponding rows of $\mathbf{Z}$ are removed.

Let right singular vector $\mathbf{v}$ be fixed, and $\mathbf{x}_i$ be the i-th column of matrix $\mathbf{X}^\intercal$. Note that $\widehat{\mathbf{u}}^{(-i)}$ also solves the ridge regression when the i-th block of $\mathbf{y}$ is replaced by $\mathbf{v}\widehat{u}_i^{(-i)}$. Therefore, we have

$$\mathbf{v}\widehat{u}_i^{(-i)} = \sum_{k\neq i} \mathbf{H}_{ik}\mathbf{x}_k + \mathbf{H}_{ii}\mathbf{v}\widehat{u}_i^{(-i)}.$$

Subtracting $\mathbf{x}_i$. from both sides, we have,

$$\begin{aligned}
\widehat{\boldsymbol{\epsilon}}_i^{(-i)} \equiv \mathbf{v}\widehat{u}_i^{(-i)} - \mathbf{x}_i &= \sum_k \mathbf{H}_{ik}\mathbf{x}_k - \mathbf{x}_i + \mathbf{H}_{ii}(\mathbf{v}\widehat{u}_i^{(-i)} - \mathbf{x}_i) \\
&= \mathbf{v}\widehat{u}_i - \mathbf{x}_i + \mathbf{H}_{ii}(\mathbf{v}\widehat{u}_i^{(-i)} - \mathbf{x}_i) \\
&\equiv \widehat{\boldsymbol{\epsilon}}_i + \mathbf{H}_{ii}\widehat{\boldsymbol{\epsilon}}_i^{(-i)}.
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\widehat{\boldsymbol{\epsilon}}_i^{(-i)} &= (\mathbf{I}_m - \mathbf{H}_{ii})^{-1}\widehat{\boldsymbol{\epsilon}}_i = (\mathbf{I}_m - \gamma_i\mathbf{v}\mathbf{v}^\intercal)^{-1}\widehat{\boldsymbol{\epsilon}}_i = \left(\mathbf{I}_m + \frac{\gamma_i}{1-\gamma_i}\mathbf{v}\mathbf{v}^\intercal\right)\widehat{\boldsymbol{\epsilon}}_i \\
&= \widehat{\boldsymbol{\epsilon}}_i + \frac{\gamma_i}{1-\gamma_i}(\mathbf{v}^\intercal\widehat{\boldsymbol{\epsilon}}_i)\mathbf{v}
\end{aligned}$$

where γ_i is the i-th diagonal element of matrix $\mathbf{M}(\alpha)$. Note that $\widehat{\boldsymbol{\epsilon}}_i \equiv \mathbf{v}\widehat{u}_i - \mathbf{x}_i$ and

$$\begin{aligned}
\|\widehat{\boldsymbol{\epsilon}}_i\|^2 &= (\mathbf{v}\widehat{u}_i - \mathbf{x}_i)^\intercal(\mathbf{v}\widehat{u}_i - \mathbf{x}_i) = \mathbf{x}_i^\intercal\mathbf{x}_i - 2\mathbf{x}_i^\intercal\mathbf{v}\widehat{u}_i + \|\mathbf{v}\|^2\widehat{u}_i^2 \\
&= \mathbf{x}_i^\intercal\mathbf{x}_i - (\mathbf{x}_i^\intercal\mathbf{v})^2 + (\widehat{u}_i - \mathbf{x}_i^\intercal\mathbf{v})^2.
\end{aligned}$$

Therefore, we have that

$$\begin{aligned}
\|\widehat{\boldsymbol{\epsilon}}_i^{(-i)}\|^2 &= \|\widehat{\boldsymbol{\epsilon}}_i\|^2 + \frac{2\gamma_i}{1-\gamma_i}(\mathbf{v}^\intercal\widehat{\boldsymbol{\epsilon}}_{\boldsymbol{i}})^2 + \left(\frac{\gamma_i}{1-\gamma_i}\right)^2(\mathbf{v}^\intercal\widehat{\boldsymbol{\epsilon}}_{\boldsymbol{i}})^2 \\
&= \|\widehat{\boldsymbol{\epsilon}}_i\|^2 + (\mathbf{v}^\intercal\widehat{\boldsymbol{\epsilon}}_{\boldsymbol{i}})^2\left[\frac{1}{(1-\gamma_i)^2} - 1\right] \\
&= \mathbf{x}_i^\intercal\mathbf{x}_i - (\mathbf{x}_i^\intercal\mathbf{v})^2 + (\widehat{u}_i - \mathbf{x}_i^\intercal\mathbf{v})^2 + (\widehat{u}_i - \mathbf{v}^\intercal\mathbf{x}_i)^2\left[\frac{1}{(1-\gamma_i)^2} - 1\right] \\
&= \mathbf{x}_i^\intercal\mathbf{x}_i - (\mathbf{x}_i^\intercal\mathbf{v})^2 + \frac{(\widehat{u}_i - \mathbf{x}_i^\intercal\mathbf{v})^2}{(1-\gamma_i)^2}.
\end{aligned}$$

When $\mathbf{v}$ is fixed, the first two terms in the above equation are irrelevant when we obtain optimal smoothing parameter α by minimizing the cross-validation criteria. Averaging the last term in the above equation over i, we have cross validation function,

$$\mathrm{CV}(\alpha) = \frac{1}{n}\sum_{i=1}^n \frac{(\widehat{u}_i - \mathbf{x}_i^\intercal\mathbf{v})^2}{(1-\gamma_i)^2} = \frac{1}{n}\sum_{i=1}^n \frac{(\{\mathbf{M}(\alpha)\mathbf{X}\mathbf{v}\}_i - \{\mathbf{X}\mathbf{v}\}_i)^2}{(1-\{\mathbf{M}(\alpha)\}_{ii})^2}$$

$$= \frac{1}{n}\sum_{i=1}^{n}\frac{[\{[\mathbf{I}_n - \mathbf{M}(\alpha)]\mathbf{X}\mathbf{v}\}_i]^2}{(1-\{\mathbf{M}(\alpha)\}_{ii})^2}. \tag{S.2}$$

If the term $\{\mathbf{M}(\alpha)\}_{ii}$ in the cross-validation criterion (S.2) is replaced by the average value $(1/n)\,\mathrm{tr}\{\mathbf{M}(\alpha)\}$, we obtain the generalized cross-validation criterion,

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

which is the objective function in (2.3).

S.2 Simulation with artificial 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 $\mathbf{b} = (b_1, \ldots, b_n)^\top$ and $\mathbf{a} \equiv (a_1, \ldots, a_{12})^\top$ take following values,

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

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,

$$\mathbf{S}^o = \mathbf{b}\mathbf{a}^\top \equiv \mathbf{i}_n \cdot \mathbf{f}^\top + \mathbf{u}\mathbf{v}^\top \tag{S.3}$$

where $\mathbf{v} = \mathbf{a}$, $\mathbf{f} = \overline{\mathbf{b}} \cdot \mathbf{a}^\top$, $\mathbf{u} = \mathbf{b} - \overline{\mathbf{b}}$, with $\overline{\mathbf{b}} = n^{-1}\sum_{i=1}^{n} b_i$. In Figure S1, 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 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,

$$s_t = w s_t^o \equiv \kappa \frac{\text{SD}(e_t)}{\text{SD}(s_t^o)} s_t^o,$$

where $\text{SD}(\cdot)$ is the unconditional standard deviation function, so as to control the sample Unconditional Standard Deviation Ratio (USDR), defined by $\text{SD}(s_t)/\text{SD}(e_t)$, of simulated time series data $x_t = s_t + e_t$ to be exactly κ in each replication of DGPs. For the three DGPs, we choose κ = 0.2, 0.4, 0.6, 0.8, 1, 1.2, 1.4, 1.6, 1.8, 2 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 X-12-ARIMA and SEATS methods on the accuracy of estimation of the seasonal component. They are the mean square errors (MSE) and mean percentage errors (MPE):

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

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

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

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

[Table S1]

Table S1 reports the comparison results among X-12-ARIMA, SEATS and our RSVD methods. Several findings are in order. First, for all three DGPs, the absolute loss (AMSE) of the X-12-ARIMA and SEATS methods keep 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 all three methods keeps decreasing as κ increases, and AMPE of our RSVD method decreases faster. Third, for stationary DGP1 and DGP2 cases the SEATS method has smallest losses of AMSE and AMPE. However, as κ increases, our RSVD method outperforms SEATS in both AMSE and AMPE criteria; for nonstationary DGP3 cases, our RSVD method uniformly dominates the X-12-ARIMA and SEATS methods by delivering smaller losses of AMSE and AMPE regardless of κ values. Moreover, the average selected number of seasonal patterns r is generally the same and is close to one across different values of κ capturing the strength of seasonal variation. It signifies that the selected number of seasonal patterns based on BIC is robust to the irregular component. In general, no additional numbers of seasonal patterns are added due to the irregular variation.

S.3 Simulation with seasonality from real economic time series

In addition to the artificial seasonalities used in Section 6.1 and 6.2 of the main paper, we also use three real seasonal components that prevail in three seasonal economic time series. They are Industrial Production Index, Total Non-Farm Payrolls, and the Inflation Rate calculated from Consumer Price Index for All Urban Consumers, which are abbreviated as IPI, NFP, and INFL respectively and available from Federal Reserve Economic Data (FRED) website. All of the three monthly series are not seasonally adjusted, and the sample period is from January 1967 to December 2016, covering last five decades with 600 observations. In this section, we exclude the RSVDB method from the simulation exercises due to the intensive computational burdens involved in the exhaustive search for abrupt seasonality breaks. However, RSVDB encompasses RSVD as a special case, and hence the RSVDB method performs at least as good as the RSVD method does.

S.3.1 Data generating processes

We adopt two different schemes of simulation with real seasonal economic time series. In the first simulation scheme, we first apply one of the TRAMO-SEATS, X-12-ARIMA, and RSVD seasonal adjustment methods to each of the three seasonal economic time series (IPI, NFP, and INFL) alternately, obtaining the real seasonal component s_t^{method} and seasonally adjusted series e_t^{method}. These two series exclude the calendar effects and outliers in the original series, so that only the seasonal decomposition is evaluated in the simulation. Then, we obtain the simulated seasonal time series $x_t^{\text{sim, method}}$ by

$$x_t^{\text{sim, method}} = a \cdot s_t^{\text{method}} + e_t^{\text{method}} + \sigma u_t^{\text{sim}}$$

where method = {SEATS, X-12-ARIMA, RSVD}, $a = \{1, 2, 4, 8, 16, 32\}$ is the relative amplitude of seasonality, u_t^{sim} is the i.i.d. normally distributed perturbation with $var(u_t^{\text{sim}}) = var(s_t^{\text{method}})$, and $\sigma = \{0.05, 0.10\}$ captures the magnitude of the perturbation. After that, we apply the three seasonal adjustment methods to the simulated series $\{x_t^{\text{sim, method}}\}$ For each of the three series, a and σ 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.

The second simulation scheme follows the suggestion of a reviewer: consider the SEATS-type seasonality as the true underlying one, and evaluate X-12-ARIMA and RSVD methods on their ability to recover the SEATS-type seasonality in the simulated series. In this spirit, we apply the TRAMO-SEATS seasonal adjustment procedure to the seasonal economic time series, and obtain the SEATS seasonal component $\{s_t^{\text{seats}}\}$ and seasonally adjusted series $\{e_t^{\text{seats}}\}$. These two series exclude the calendar effects and outliers in the original series, so that only the seasonal decomposition is evaluated in the simulation. Second, we find the optimal ARIMA$(p, d, q)_{\text{opt}}$ model for the seasonally adjusted series $\{e_t^{\text{seats}}\}$ suggested by the TRAMO-SEATS. Table S5 reports the three selected ARIMA models. For the seasonally adjusted NFP series, the ARIMA model suggested by TRAMO-SEATS is ARIMA(0, 3, 3), and we find that the associated simulated series simply explodes and has much larger magnitude than the adjusted NFP series. Therefore, we only use IPI and INFL series for these simulation exercises. Third, we generate the simulated non-seasonal component e_t^{sim} from ARIMA$(p, d, q)_{\text{opt}}$, generate the simulated seasonal component s_t^{sim} by

$$s_t^{\text{sim}} = \kappa \frac{\text{SD}(e_t^{\text{sim}})}{\text{SD}(e_t^{\text{seats}})} s_t^{\text{seats}} \tag{S.4}$$

where SD$(\cdot)$ is the unconditional standard deviation function, and the scalar κ controls the relative strength of seasonality such that, in each replication of DGPs, the simulated sample Unconditional Standard Deviation Ratio (abbreviated to USDR), $\text{SD}(s_t^{\text{sim}})/\text{SD}(e_t^{\text{sim}})$, is exactly κ times the original sample USDR value, $\text{SD}(s_t^{\text{seats}})/\text{SD}(e_t^{\text{seats}})$. The smaller this ratio is, the weaker the seasonality is relative to the nonseasonal dynamics. This ratio is 0.040 for IPI series, 0.035 for NFP series, and 0.43 INFL series, which implies that the seasonality in all three series is quite weak, especially for IPI and NFP series. Hence, to examine the capability of the seasonal decomposition in a wide range of magnitude in seasonality, we set $\kappa = \{4, 8, 16, 32, 64, 128\}$. As the value κ increases, the seasonality becomes more prominent, and has stronger variation. Finally, we obtain the simulated seasonal time series x_t^{sim} by

$$x_t^{\text{sim}} = s_t^{\text{sim}} + e_t^{\text{sim}} + \sigma u_t^{\text{sim}}$$

where u_t^{sim} is the i.i.d. normally distributed perturbation with $var(u_t^{\text{sim}}) = var(s_t^{\text{sim}})$ and $\sigma = \{0.05, 0.10\}$ captures the magnitude of the perturbation. After that, we only apply the X-12-ARIMA and RSVD methods to the simulated series $\{x_t^{\text{sim}}\}$, since the generation of a simulated seasonal component, generated from the TRAMO-SEATS procedure, is biased towards the SEATS method. For each of the three series 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.

S.3.2 Simulation results

Through these simulation exercises, we find that the results from the first scheme convey much richer information than that from the second one. In addition, because the seasonality of the three series are relatively weak and their seasonally adjusted series are non-stationary, simulated ARIMA series $\{e_t^{\text{sim}}\}$ usually overwhelm variations in the seasonal component, which renders the seasonal behavior in the simulated series very weak. In contrast, the simulated seasonal time series in the first scheme maintains the stochastic trend and dynamics of the originally adjusted series and therefore is more relevant compared those in the second scheme. We report the simulation results for the first scheme in Table S2 – S4 and those for the second scheme in Table S6.

Several findings in these simulation results are in order.

1. Just as no exact definition of seasonality exists, three different seasonal adjustment methods recognize seasonality differently in the data. If the simulated seasonality is extracted from one seasonal adjustment method, then this same seasonal adjustment method tends to outperform the other two methods in simulation. This is especially the case for the SEATS and RSVD methods. However, X-12-ARIMA is an exception. When the simulated seasonality is generated by X-12-ARIMA in the simulation, SEATS outperforms X-12-ARIMA when the magnitude of seasonality a is large. Moreover, SEATS also outperforms RSVD for all combinations of σ and a. This could be due to the fact that the filters employed by the X-12-ARIMA method could be consistent with an airline model, which tends to be in favor of the SEATS method.

These results imply that both SEATS and our RSVD methods are robust enough to their own type of seasonality in large magnitude, while the X-12-ARIMA method is not. Moreover, when the true seasonality in the data generating processes comes from the X-12-ARIMA or SEATS, the RSVD method increases the number of patterns needed to capture the seasonal component. The X-12-ARIMA or SEATS type seasonality does not exactly conform to our RSVD seasonality defined in (4.3), and therefore some small seasonal variations are ignored by our RSVD method because of the regularization. As the USDR or a increases, those ignored small seasonal variations become more salient and are recognized as additional seasonal patterns by our RSVD method.

2. Among the three seasonal adjustment methods, the RSVD method is the most robust to irregular variations. In Table S2 – S4, when the simulated seasonality s_t^{rsvd} is extracted by the RSVD method and the simulated series includes an irregular component with $\sigma = 10\%$ standard deviation of s_t^{rsvd}, i.e. $\mathrm{SD}(\sigma u_t^{\mathrm{sim}}) = 0.1\mathrm{SD}(s_t^{\mathrm{rsvd}})$, the AMPE losses of the RSVD method in simulation, decreasing as the magnitude of seasonality a increases, range from 32.96% to 2.13%, 22.17% to 4.55%, and 26.79% to 0.94% for the IPI, NFP, and INFL series respectively. Moreover, the average selected number of seasonal patterns r is generally the same when σ increases from 0.05 to 0.1, i.e. the irregular variations becomes stronger. This signifies that the selected number of seasonal patterns based on BIC is robust to the irregular component. In general, no additional numbers of seasonal patterns are added due to the irregular variation.

In comparison, when the SEATS method is used to adjust the simulated series in which the seasonality is also extracted by SEATS and additional irregular noises with $\sigma = 10\%$ standard deviation of s_t^{seats} are added, the AMPE losses of SEATS also decrease as a increases, and range from 42.54% to 8.94%, 21.80% to 1.97%, and 96% to 12.03% for the three series respectively. In the same scenario for X-12-ARIMA, the AMPE losses range from 26.08% to 14.61%, 209.04% to 18.97%, and 95.83% to 45.93% for the three series.

3. In particular, if we consider the model-based SEATS type seasonality as the true one and compare the performance of the two empirical-based adjustment methods, X-12-ARIMA and RSVD, we find that RSVD outperforms X-12-ARIMA with smaller AMSE and AMPE losses for all three series when the magnitude of seasonality a is large enough, while X-12-ARIMA delivers smaller losses when a is small. In other words, RSVD is better at capturing strong SEATS-type seasonality, and X-12-ARIMA is better at capturing a weak structure.

This conclusion also holds true under the second simulation scheme, in which SEATS-type seasonality is also deemed as the true one. However, the difference is that SEATS seasonally adjusted series are simulated directly by the optimal ARIMA process selected by AIC. Table S5 shows the optimal ARIMA models for the three series. Table S6 reports the comparison results between X-12-ARIMA and RSVD. The last column shows the average value of selected number of seasonal patterns r. For all three cases, when the simulated sample Unconditional Standard Deviation Ratio (USDR) is relatively small, the performance of X-12-ARIMA is marginally better than the RSVD method with slightly smaller AMSE and AMPE losses. But when the simulated sample USDR keeps increasing, while the RSVD method gains in superiority and has an overwhelming advantage against X-12-ARIMA with much smaller AMSE and AMPE losses. In addition, as USDR increases, the seasonality becomes more salient. As a result, the number of seasonal patterns r also increases on average accordingly. These simulation results imply that the main advantage of RSVD begins to manifest itself when the seasonality is strong. Moreover, the RSVD method is also robust to moderate SEATS type seasonality with almost the same or better finite sample performance than X-12-ARIMA.

S.3.3 Data revision

When new observations become available as time evolves, a seasonally adjusted series sometimes may need to be changed accordingly after incorporating the new information in the updated time series. If a seasonal adjustment method requires practically no or only a little revision, its estimated seasonal component should be quite robust to new observations.

To assess the degree of data revision for a seasonal adjustment method, we devise the following evaluation criterion. We consider a monthly seasonal time series $\{x_{i,j}\}$, where the (i, j) pair denotes the j-th month in i-th year with $i = 1, \ldots, n$ and $j = 1, \ldots, 12$. We first set the initial subsample, which covers the first L years of observations. Then, we create a sequence of subsamples from the original seasonal time series. The time span of the first subsample starts from $(1, 1)$ to $(L + 1, 12)$, the time span of the next subsample extends to $(L + 2, 12)$, and so on until the last subsample coincides with the whole sample. In total, we create $n - L$ subsamples from the original time series. After that, we apply a seasonal adjustment method to the initial subsample and those $n - L$ subsamples, and obtain their seasonally adjusted series denoted by $\{e_t^{\mathrm{ini}}\}$ and $\{e_t^s\}$, where $s = 1, 2, \ldots, n - L$ indexes subsamples. After that, we calculate the the sample standard deviation of the difference

$\delta_t^s = e_t^s - e_t^{\text{ini}}$ across $s = \{1, 2, \ldots, n-L\}$ for each time $t \equiv (i-1) \cdot 12 + j$ in the initial subsample,

$$sd_t = \left(\frac{1}{n-L} \sum_{s=1}^{n-L} (\delta_t^s - \overline{\delta_t})^2 \right)^{1/2} \tag{S.5}$$

where $\overline{\delta_t} = \frac{1}{n-L} \sum_{s=1}^{n-L} \delta_t^s$. We use the ratio of $\max_t \{sd_t / e_t^{\text{ini}}\}$ for the initial subsample to capture the relative size of data revision.

Using X-12-ARIMA, SEATS, and our RSVD methods, we calculate this ratio for the IPI and NFP series. The INFL series is not evaluated because the adjusted series can be so close to zero that the ratio explodes. For IPI, the ratios of all three seasonal adjustment methods are smaller than 0.001. For NFP, the ratios of both X-12-ARIMA and SEATS are smaller than 0.001, and the ratio of the RSVD method is about 0.002. Therefore, we believe that our RSVD method is robust enough to new observations and the need for data revision is comparable to X-12-ARIMA and SEATS methods even for the time series with relative weak seasonality.

S.4 Further discussions

S.4.1 Some potential data problems

Regarding the issues of missing values, outliers, and calendar effects that emerge in seasonal adjustment, there exist two popular and sophisticated pre-treatment procedures, TRAMO and RegARIMA, which are specifically designed to deal with these data problems. One could use one of these two pre-treatment steps before applying our RSVD method. Alternatively, our RSVD method can also potentially deal with missing values, outliers, and calendar effects after suitable modifications and extensions.

Missing values. Suppose the missing entry x_{ij} is on the i-th row and j-th column of data matrix $\mathbf{X}$. We propose two different remedies to overcome the missing value problem. The first remedy is to impute the missing values based on some appropriate time series model.

For example, we could fit an ARIMA model for the nonmissing entries in the same i-th row, and replace the missing entry x_{ij} by the fitted value $\widehat{x}_{ij}$. Then, the RSVD method can be applied to the imputed data matrix $\mathbf{X}$.

An alternative approach is to modify the RSVD algorithms slightly by removing the the rows or columns of those matrices that contain the missing entries. For example, in Algorithm 1, we obtain the initial $\mathbf{u}_{-i}$ from the standard SVD of the modified data matrix $\mathbf{X}_{-i,\cdot}$ with the i-th row of $\mathbf{X}$ removed. Then, in Step 2(a), we update the seasonal pattern $\mathbf{v}$ with $\mathbf{X}_{-i,\cdot}$ and $\mathbf{u}_{-i}$. Finally, in Step 2(b), we update the pattern coefficient $\mathbf{u}$ with the modified data matrix $\mathbf{X}_{\cdot,j}$ and vector $\mathbf{v}_{-j}$ which are obtained respectively by removing the j-th column from $\mathbf{X}$ and the j-th entry from $\mathbf{v}$. Similar modifications could also be made for other afore-mentioned RSVD algorithms.

Outliers. The seasonal adjustment results of our current RSVD method could be sensitive to outliers, since the quadratic loss function is used to measure the reconstruction errors of low-rank matrix approximation in the regularized singular value decomposition. In order to enhance the robustness of the RSVD method, we propose to replace the quadratic loss function by some robust loss functions, such as the check function and Huber's function. Zhang, Shen, and Huang (2013) propose a robust version of RSVD using a typical Huber's function and develop an efficient Iterative Reweighted Least Squares (IRLS) algorithm correspondingly.

Calendar effects. Generally speaking, the calendar effect means any repeating market anomaly or economic fluctuation that is attributable to the calendar features, such as trading day effects, fixed and moving holiday effects, etc. The calendar effects can be accommodated by our RSVD method by including regression effects in the RSVD seasonal adjustment procedure. For example, in the Step 2 of the seasonal adjustment procedure in Section 4.2, the constrained least squares problem (4.5) can be modified as follows,

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

where the $np \times L$ matrix $\mathbf{C}$ represents the L calendar regressors and the coefficients in γ capture the calendar effects. Then, the fixed and time-varying seasonal patterns in $\beta \equiv (\mathbf{f}^\top, \mathbf{v}_1^\top, \cdots, \mathbf{v}_r^\top)^\top$ and the pattern coefficients in $\mathbf{U}$ can be further estimated by applying the RSVD method to the matrix $\mathbf{X} - \mathbf{C}\widehat{\gamma}$ after removing the calendar effects from the original data matrix.

S.4.2 Bootstrapped confidence interval

The statistical inference theory for our RSVD seasonal adjustment method is not easy to derive. Although we treat the regularized SVD problem in finite dimensions, it is connected to a smoothing splines estimation whose inference is a long-standing issue in statistics, and the construction of confidence intervals for smoothing splines estimation is difficult.

Instead, we propose a potentially valid bootstrap procedure. Formal investigation of its statistical properties is left for future research. First, if the non-seasonal component e_t is stationary, we can block bootstrap the estimated non-seasonal component $\{\widehat{e}_t\}$ to maintain the time dependence structure in the bootstrapped non-seasonal component $\{\widehat{e}_t^*\}$. If the non-seasonal component e_t has a stochastic trend, the block bootstrap method is invalid for non-stationary time series and we can use the sieve bootstrap procedure similar to Psaradakis (2001) to obtain the bootstrapped non-seasonal component $\{\widehat{e}_t^*\}$. Next, we obtain the bootstrapped seasonal time series $x_t^* = \widehat{s}_t + \widehat{e}_t^*$, and apply the RSVD seasonal adjustment method to $\{x_t^*\}$ to obtain the bootstrapped seasonal component $\{\widehat{s}_t^*\}$. Finally, we take the empirical quantiles from the bootstrap distribution of the seasonal component for each time t, and the bootstrapped confidence interval for the seasonal component is $\{[2\widehat{s}_t - \widehat{s}^*_{t,(1-\alpha/2)}, 2\widehat{s}_t - \widehat{s}^*_{t,(\alpha/2)}]\}$ where $\widehat{s}^*_{t,(1-\alpha/2)}$ denotes the $1-\alpha/2$ percentile of the bootstrapped seasonal component $\widehat{s}_t^*$ at time t. Correspondingly, the bootstrapped confidence interval for the seasonally adjusted series is $\{[x_t - 2\widehat{s}_t + \widehat{s}^*_{t,(\alpha/2)}, x_t - 2\widehat{s}_t + \widehat{s}^*_{t,(1-\alpha/2)}]\}$.

S.4.3 Multiple types of seasonality

Although the proposed RSVD seasonal adjustment method in this paper can only deal with single seasonality, it does shed some light on seasonal adjustment for time series with multiple types of seasonality. For example, an intraday high frequency time series has different seasonal cycles and can possibly exhibit multiple types of seasonality, such as daily, weekly, and annual seasonality. If the time series is reshaped into a multi-dimensional array, seasonal behaviors at those different frequencies incur different smoothness structures on different "facets" of the multi-dimensional array. A straightforward extension of our current RSVD method to tackle multiple types of seasonality is to consider applying regularized SVD appropriately to the different "facets" of the multi-dimensional array and separate out those smooth seasonal variations at different frequencies from the original time series.

References

Board of Governors of the Federal Reserve System (US), Industrial Production: Total index [IPB50001N], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/IPB50001N, July 8, 2017.

Psaradakis, Z. (2001). Bootstrap tests for an autoregressive unit root in the presence of weakly dependent errors. *Journal of Time Series Analysis,* 22(5), 577–594.

U.S. Bureau of Labor Statistics, All NFPoyees: Total Nonfarm Payrolls [PAYNSA], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/PAYNSA, July 8, 2017.

U.S. Bureau of Labor Statistics, Consumer Price Index for All Urban Consumers: All Items [CPIAUCNS], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CPIAUCNS, July 8, 2017.

Zhang, L., Shen, H., & Huang, J. Z. (2013). Robust regularized singular value decomposition with application to mortality data. *The Annals of Applied Statistics,* 7(3), 1540–1561.

Table S1: Evaluation of estimates of seasonal component with no break

	κ	AMSE ($\times 10^{-2}$) X12-ARIMA	SEATS	RSVD	AMPE (%) X12-ARIMA	SEATS	RSVD	Avg. r RSVD
DGP1 *	0.2	13.6819	**3.5118**	6.7704	449.43	**161.03**	287.39	1.092
	0.4	13.8016	**3.5307**	6.5452	225.45	**101.12**	140.09	1.094
	0.6	13.9597	**3.9594**	6.4137	151.03	**73.62**	91.89	1.128
	0.8	14.0816	**4.6717**	6.0776	113.68	**60.07**	67.38	1.130
	1.0	14.2103	**5.0835**	5.8606	91.30	**50.95**	53.33	1.134
	1.2	14.3884	**5.3678**	5.7873	76.47	**44.32**	44.42	1.136
	1.4	14.6425	5.7749	**5.7494**	66.08	39.65	**37.88**	1.136
	1.6	14.8681	6.1388	**5.7080**	58.26	36.00	**33.05**	1.134
	1.8	15.0391	6.5726	**5.7343**	52.06	33.21	**29.41**	1.138
	2.0	15.3330	6.9360	**5.7271**	47.29	30.81	**26.46**	1.138
DGP2 †	0.2	18.8260	**4.2091**	7.8188	299.08	**133.81**	172.05	1.010
	0.4	18.9328	**5.6554**	10.4065	149.75	**79.74**	91.90	1.006
	0.6	19.0276	**7.0764**	13.7174	99.98	**60.03**	65.11	1.022
	0.8	19.0987	**8.3877**	11.5922	75.06	49.17	**47.52**	1.262
	1.0	19.1916	9.5771	**8.5133**	60.17	42.08	**35.96**	1.296
	1.2	19.2761	10.7163	**7.3184**	50.21	37.14	**28.89**	1.222
	1.4	19.3463	11.8055	**7.1906**	43.09	33.42	**24.33**	1.140
	1.6	19.4087	12.7889	**6.6489**	37.75	30.48	**20.81**	1.082
	1.8	19.4716	13.8061	**6.4423**	33.59	28.16	**18.17**	1.038
	2.0	19.5353	14.8512	**5.8001**	30.26	26.28	**16.00**	1.012
DGP3 ‡	0.2	2.1745	4.0332	**0.5596**	20.61	26.61	**10.60**	1.000
	0.4	2.5360	6.3576	**0.5543**	10.72	16.79	**5.29**	1.000
	0.6	3.1728	8.2954	**0.5538**	7.48	12.70	**3.53**	1.000
	0.8	4.1132	9.2859	**0.5541**	5.91	10.13	**2.65**	1.000
	1.0	5.3071	9.9594	**0.5536**	4.99	8.47	**2.12**	1.000
	1.2	6.7664	9.6011	**0.5533**	4.40	7.04	**1.76**	1.000
	1.4	8.4086	9.5583	**0.5538**	3.99	6.10	**1.51**	1.000
	1.6	10.2984	9.4299	**0.5534**	3.69	5.31	**1.33**	1.000
	1.8	11.8878	9.1184	**0.5500**	3.49	4.71	**1.19**	1.000
	2.0	14.4208	9.1037	**0.5511**	3.30	4.19	**1.07**	1.000

* The non-seasonal component $\{e_t\}$ follows i.i.d. $N(0, \sigma^2)$. † The non-seasonal component $\{e_t\}$ follows Gaussian ARMA(1,1) with AR(1) coefficient 0.8 and MA(1) coefficient 0.1. ‡ The non-seasonal component $\{e_t\}$ follows Gaussian ARIMA(1,1,1) with AR(1) coefficient 0.8 and MA(1) coefficient 0.1.

Table S2: Simulation results with Industrial Production Index (IPI) series

		AMSE ($\times 10^{-4}$)			AMPE (%)			Avg. r
σ	a	X-12-ARIMA	SEATS	RSVD	X-12-ARIMA	SEATS	RSVD	RSVD
Simulated seasonality is from the X-12-ARIMA method								
0.05	1	0.0054	0.0164	0.1072	18.05	30.56	85.54	2.000
0.05	2	0.0131	0.0509	0.1882	15.31	32.10	59.89	3.780
0.05	4	0.0575	0.2821	0.2911	13.29	38.42	30.89	4.000
0.05	8	0.3495	0.4684	0.3805	15.64	23.05	20.26	9.000
0.05	16	1.7365	0.2196	0.5139	17.69	8.31	17.11	9.000
0.05	32	8.5090	0.3322	1.0693	19.99	4.90	11.45	10.000
0.10	1	0.0094	0.0179	0.1091	26.08	33.18	85.47	1.974
0.10	2	0.0170	0.0497	0.1956	18.17	32.01	59.56	3.676
0.10	4	0.0604	0.2479	0.2932	14.61	36.68	31.82	4.004
0.10	8	0.3444	0.5637	0.3847	15.94	25.68	20.96	8.978
0.10	16	1.7346	0.2442	0.5239	18.02	8.80	17.18	9.008
0.10	32	8.4849	0.3354	1.0812	20.14	4.92	11.47	10.000
Simulated seasonality is from the SEATS method								
0.05	1	0.0171	0.0054	0.1082	47.59	33.88	130.61	2.000
0.05	2	0.0367	0.0120	0.2193	33.78	28.49	102.26	4.450
0.05	4	0.1405	0.1105	0.3392	36.43	39.82	59.18	4.000
0.05	8	0.6089	0.3558	0.4090	39.05	33.41	34.11	9.244
0.05	16	2.4498	0.0904	0.8419	39.56	8.81	22.14	10.038
0.05	32	10.7650	0.3553	1.7495	42.04	8.72	17.20	11.000
0.10	1	0.0197	0.0074	0.1102	54.69	42.54	130.00	1.954
0.10	2	0.0389	0.0131	0.2332	37.87	30.27	92.92	3.738
0.10	4	0.1419	0.0952	0.3459	37.78	37.50	61.46	4.004
0.10	8	0.6052	0.3686	0.4200	39.58	33.42	34.20	9.188
0.10	16	2.4654	0.1049	0.8564	40.14	9.51	22.55	10.046
0.10	32	10.7903	0.3657	1.7524	42.62	8.94	17.24	10.998
Simulated seasonality is from the RSVD method								
0.05	1	0.1078	0.0813	0.0162	123.63	113.93	27.13	1.548
0.05	2	0.1149	0.1390	0.0273	66.95	80.25	19.57	2.000
0.05	4	0.1460	0.2253	0.0334	41.80	54.46	12.95	2.000
0.05	8	0.1910	0.4005	0.0354	25.46	37.13	7.15	2.000
0.05	16	0.7374	0.3956	0.0372	19.25	18.74	3.83	2.000
0.05	32	5.2690	0.4809	0.0422	17.36	10.34	2.17	2.000
0.10	1	0.1104	0.0800	0.0183	124.70	112.45	32.96	1.566
0.10	2	0.1181	0.1357	0.0269	67.79	78.75	19.46	2.000
0.10	4	0.1491	0.2220	0.0307	41.84	53.83	11.62	2.000
0.10	8	0.1942	0.3968	0.0344	25.58	36.90	6.50	2.000
0.10	16	0.6514	0.4176	0.0390	19.37	19.16	3.78	2.000
0.10	32	5.3271	0.4700	0.0441	17.55	10.24	2.13	2.000

Table S3: Simulation results with total Non-Farm Payroll (NFP) series

		AMSE ($\times 10^{-2}$)			AMPE (%)			Avg. r
σ	a	X-12-ARIMA	SEATS	RSVD	X-12-ARIMA	SEATS	RSVD	RSVD
Simulated seasonality is from the X-12-ARIMA method								
0.05	1	0.0539	0.0677	0.6829	122.92	194.48	93.94	2.000
0.05	2	0.0773	0.1609	1.4655	83.25	214.49	149.89	1.000
0.05	4	0.2470	0.6185	2.2247	48.03	204.11	51.73	3.712
0.05	8	1.6058	1.8047	3.0317	39.77	92.81	29.27	6.704
0.05	16	8.1157	0.8444	5.6101	28.01	33.23	47.05	8.026
0.05	32	38.3131	1.3865	10.5476	17.22	10.54	62.10	9.096
0.10	1	0.1578	0.1191	0.7653	209.04	173.48	152.09	2.018
0.10	2	0.2122	0.2040	1.4457	150.26	170.68	182.45	1.000
0.10	4	0.3556	0.5246	2.3775	70.08	163.88	76.20	3.570
0.10	8	1.5940	1.4518	3.3735	40.16	95.77	46.89	6.578
0.10	16	8.3436	1.2132	5.3953	26.22	35.53	54.91	8.466
0.10	32	38.1979	1.6197	10.4992	18.97	12.30	57.48	9.598
Simulated seasonality is from the SEATS method								
0.05	1	0.0819	0.0329	0.6547	24.70	12.83	29.38	2.000
0.05	2	0.1374	0.0545	1.5050	16.64	7.82	30.00	1.000
0.05	4	0.4502	0.2953	1.8419	17.49	10.26	17.76	5.174
0.05	8	2.1259	0.9523	3.4790	18.21	8.46	10.35	6.528
0.05	16	9.6780	0.6537	6.3728	18.69	3.10	10.59	9.134
0.05	32	40.7692	0.6382	17.2679	19.77	1.74	9.95	9.496
0.10	1	0.1724	0.0918	0.7377	30.67	21.80	34.67	2.022
0.10	2	0.2525	0.1267	1.4880	18.79	12.70	32.10	1.000
0.10	4	0.5466	0.2831	2.1857	15.06	9.79	19.77	4.730
0.10	8	2.2093	0.8431	4.1221	16.85	8.02	11.93	6.188
0.10	16	9.8038	0.8233	6.6204	18.12	3.81	10.51	8.822
0.10	32	40.4162	0.7678	16.5835	19.44	1.97	9.10	9.814
Simulated seasonality is from the RSVD method								
0.05	1	0.6817	0.6141	0.0333	37.38	37.44	9.64	2.000
0.05	2	1.4952	1.4414	1.1970	23.06	22.99	23.49	1.226
0.05	4	4.6168	3.3320	2.6989	15.85	14.63	9.62	2.056
0.05	8	17.7801	7.3443	0.9252	13.36	9.79	4.05	3.000
0.05	16	66.4177	22.3084	3.0364	13.01	10.15	4.27	3.000
0.05	32	258.1326	78.7335	12.3832	14.61	11.02	4.50	3.000
0.10	1	0.7750	0.6341	0.1523	44.47	36.93	19.60	2.036
0.10	2	1.6271	1.5311	1.1246	26.77	23.38	22.17	1.388
0.10	4	4.7254	3.7294	2.4721	18.18	16.31	10.70	2.166
0.10	8	17.7779	8.3874	1.1216	14.21	11.38	5.13	3.008
0.10	16	66.3603	23.7757	3.2494	13.28	10.78	4.55	3.002
0.10	32	260.0433	72.1982	12.5217	14.56	10.51	4.55	3.000

Table S4: Simulation results with Inflation (INFL) series

		AMSE ($\times 10^{-3}$)			AMPE (%)			Avg. r
σ	a	X-12-ARIMA	SEATS	RSVD	X-12-ARIMA	SEATS	RSVD	RSVD
Simulated seasonality is from the X-12-ARIMA method								
0.05	1	1.8280	3.9318	6.4199	92.17	126.49	224.80	1.000
0.05	2	2.3328	5.9490	11.5515	60.90	75.53	110.25	3.000
0.05	4	4.6465	8.0056	15.5058	56.51	49.04	85.75	6.000
0.05	8	15.9655	18.2086	16.7062	58.02	42.60	55.33	9.000
0.05	16	67.0332	32.9560	26.7756	47.91	29.65	22.09	10.000
0.05	32	296.6000	34.9535	71.9444	45.57	15.79	17.83	10.000
0.10	1	1.8765	3.9452	6.4252	95.83	127.24	225.06	1.000
0.10	2	2.3942	5.9798	11.5935	63.00	75.60	111.61	3.000
0.10	4	4.7454	8.0405	15.6287	57.02	49.23	86.94	6.000
0.10	8	15.5110	18.2361	16.9141	57.32	42.57	55.24	9.004
0.10	16	67.2046	33.0768	26.8196	48.15	29.79	22.44	10.000
0.10	32	297.2183	35.0585	71.9575	45.93	15.79	17.84	10.000
Simulated seasonality is from the SEATS method								
0.05	1	3.2342	1.0074	2.6165	155.42	96.43	224.07	1.000
0.05	2	4.5699	1.4792	3.6774	99.41	35.35	106.32	2.000
0.05	4	4.4043	2.4491	4.9616	45.95	27.84	54.39	3.000
0.05	8	5.6330	4.4972	10.1684	26.18	21.66	41.16	4.000
0.05	16	7.3661	8.9068	18.3406	17.74	16.35	25.40	5.000
0.05	32	18.4353	22.1587	31.8693	16.99	12.01	15.83	6.000
0.10	1	3.2775	1.0119	2.6209	157.06	96.68	224.10	1.000
0.10	2	4.6070	1.4891	3.7548	100.05	36.18	105.75	2.000
0.10	4	4.4518	2.4609	4.9836	46.32	28.03	54.54	3.000
0.10	8	5.6474	4.5050	10.2205	26.40	21.72	41.15	4.000
0.10	16	7.3963	8.9543	18.3759	17.73	16.39	25.42	5.000
0.10	32	18.3650	22.1508	31.9860	16.95	12.03	16.02	6.044
Simulated seasonality is from the RSVD method								
0.05	1	5.4888	1.2688	2.31e-03	449.33	249.20	12.94	1.000
0.05	2	7.3696	2.6139	2.36e-03	146.86	144.41	6.49	1.000
0.05	4	7.5841	4.7368	2.40e-03	70.24	81.53	3.40	1.000
0.05	8	8.4923	7.5907	2.66e-03	38.94	55.25	1.66	1.000
0.05	16	12.2855	12.5930	3.93e-03	35.88	52.58	0.96	1.000
0.05	32	22.5479	23.4975	9.11e-03	24.30	47.02	0.62	1.000
0.10	1	5.5698	1.2740	9.23e-03	453.12	250.52	26.79	1.000
0.10	2	7.4306	2.6314	9.42e-03	151.39	145.31	13.28	1.000
0.10	4	7.6622	4.7497	9.11e-03	72.82	81.58	6.39	1.000
0.10	8	8.6801	7.6115	9.76e-03	42.53	55.52	3.38	1.000
0.10	16	12.3124	12.6254	1.09e-02	35.95	52.56	1.69	1.000
0.10	32	22.4030	23.5389	1.61e-02	24.54	47.00	0.94	1.000

Table S5: The optimal ARIMA models for the three seasonally adjusted series[†]

ARIMA process	Industrial Production Index (Natural Log, 2012 = $\ln 100$) ARIMA(1, 2, 3)	Total Non-Farm Payroll (Millions of Persons) ARIMA(0, 3, 3)	Inflation (%) ARIMA(1, 1, 3)
ar(1)	0.7713		-0.3082
ma(1)	-1.493	-1.5299	-0.2785
ma(2)	0.5128	0.5466	-0.4951
ma(3)	-0.0009	0.0008	-0.2029
σ^2	4.80e-05	0.0199	0.0369

[†] The optimal models are selected by the TRAMO-SEAT method.

Table S6: Simulation evaluation of seasonal decomposition with real economic seasonality†

σ	USDR	AMSE ($\times 10^{-4}$) X-12-ARIMA	AMSE ($\times 10^{-4}$) RSVD	AMPE (%) X-12-ARIMA	AMPE (%) RSVD	Avg. r RSVD
			DGP based on Industrial Production Index (IPI)			
0.05	0.16	**0.2665**	0.6829	**65.07**	91.83	3.860
0.05	0.32	**0.8257**	1.0173	**53.57**	60.20	5.906
0.05	0.64	3.1946	**1.2802**	49.31	**33.66**	9.222
0.05	1.28	12.6396	**2.0950**	46.83	**21.12**	10.606
0.05	2.56	48.8251	**3.9784**	45.20	**13.96**	10.998
0.05	5.12	191.4524	**10.1929**	44.45	**9.64**	11.000
0.10	0.16	**0.2720**	0.6369	**65.69**	90.23	3.970
0.10	0.32	**0.8189**	0.9648	**53.61**	58.91	6.224
0.10	0.64	3.2024	**1.3095**	49.40	**34.03**	9.094
0.10	1.28	12.6458	**2.0815**	47.20	**21.07**	10.626
0.10	2.56	48.8033	**3.9440**	45.38	**13.89**	10.996
0.10	5.12	191.7176	**10.1590**	44.45	**9.70**	11.000

σ	USDR	AMSE ($\times 10^{-2}$) X-12-ARIMA	AMSE ($\times 10^{-2}$) RSVD	AMPE (%) X-12-ARIMA	AMPE (%) RSVD	Avg. r RSVD
			DGP based on Inflation (INFL)			
0.05	1.72	**0.6756**	0.7087	75.46	**74.36**	3.038
0.05	3.44	**0.8072**	0.8363	41.94	**41.03**	4.400
0.05	6.88	**1.1625**	1.2472	25.79	**24.96**	5.720
0.05	13.76	2.3906	**2.2049**	20.14	**16.61**	7.408
0.05	27.52	8.1580	**3.4680**	19.71	**9.97**	9.562
0.05	55.04	35.7276	**5.3203**	20.25	**6.30**	10.910
0.10	1.72	**0.6768**	0.7152	75.69	**75.04**	3.050
0.10	3.44	**0.8139**	0.8448	41.67	**40.72**	4.376
0.10	6.88	**1.1531**	1.2377	25.88	**25.02**	5.640
0.10	13.76	2.3886	**2.1952**	20.04	**16.62**	7.358
0.10	27.52	8.1359	**3.4183**	19.60	**9.84**	9.584
0.10	55.04	35.9361	**5.2596**	20.27	**6.30**	10.930

† For the DGP based on IPI and INFL series, the nonseasonal component e_t^{sim} follows the corresponding ARIMA process specified in Table S5, and the seasonal component s_t^{sim} is generated by equation (S.4) accordingly.

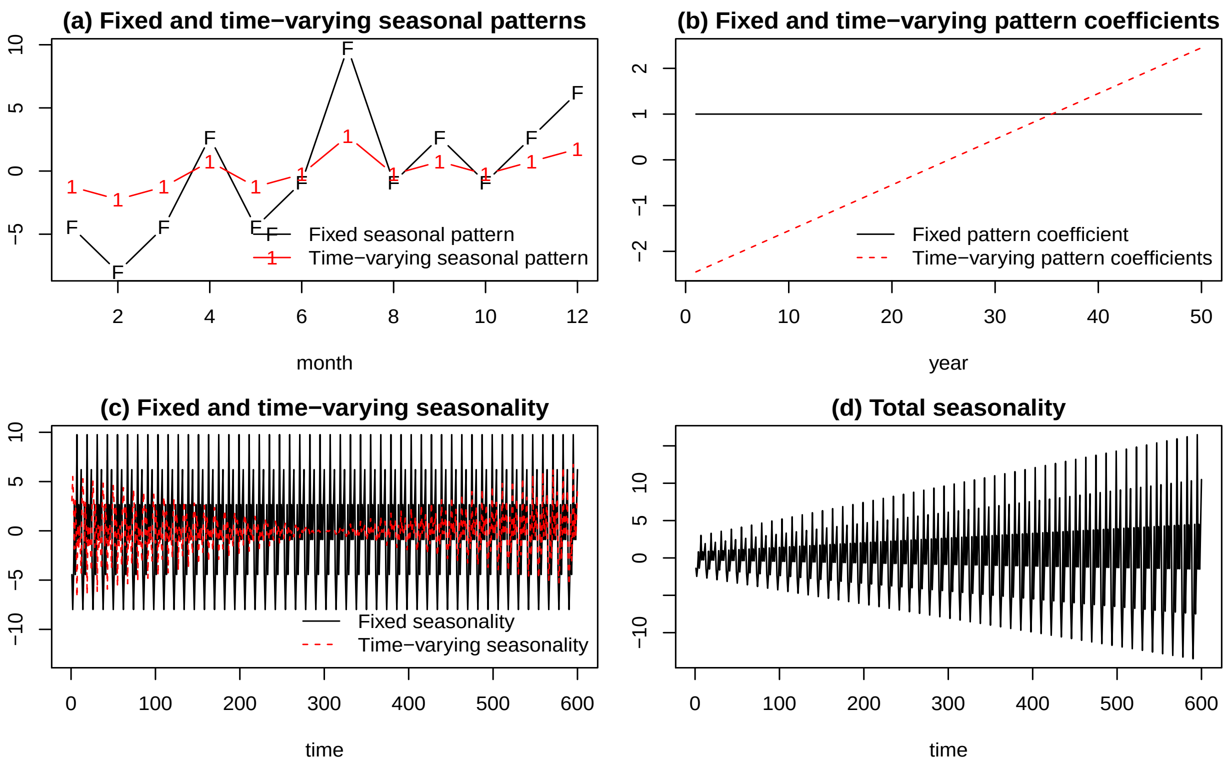

Figure S1: Simulated seasonal component without break

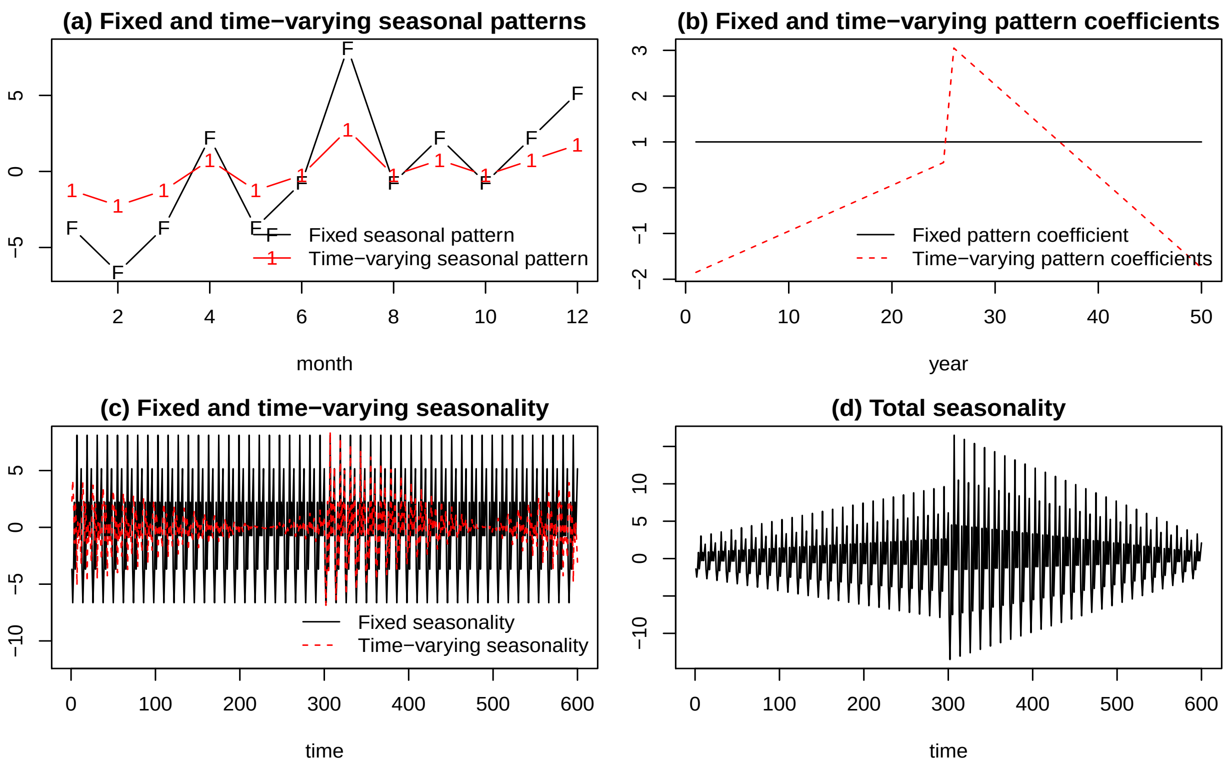

Figure S2: Simulated seasonal component with one break

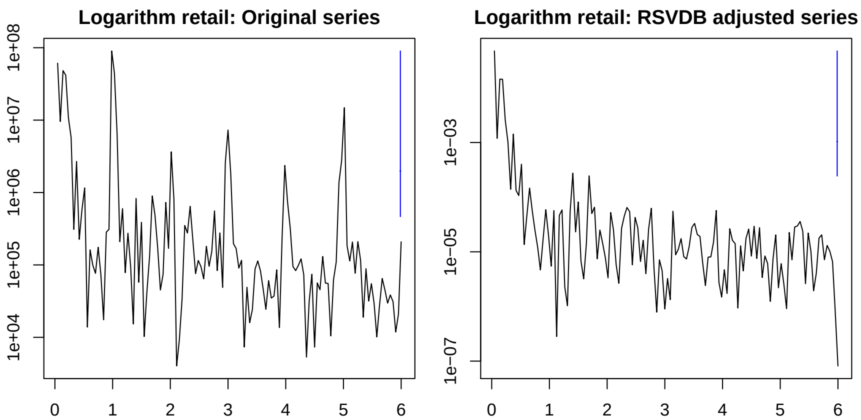

Figure S3: Periodograms of logarithm retail volume series and RSVDB adjusted series

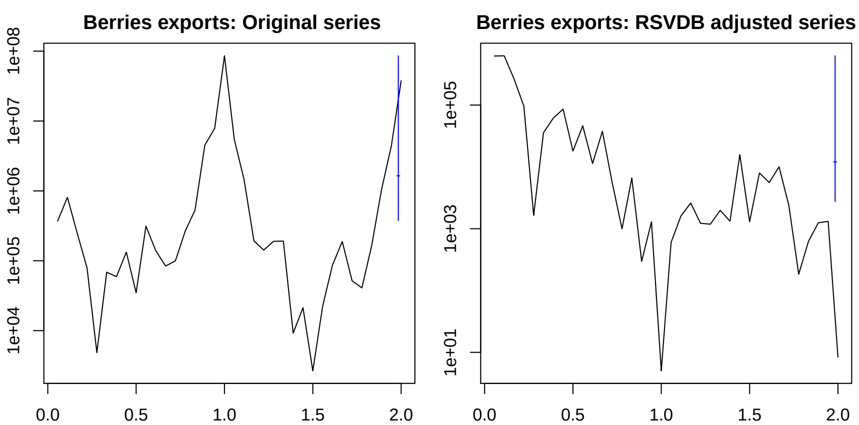

Figure S4: Periodograms of berries exports series and RSVDB adjusted series

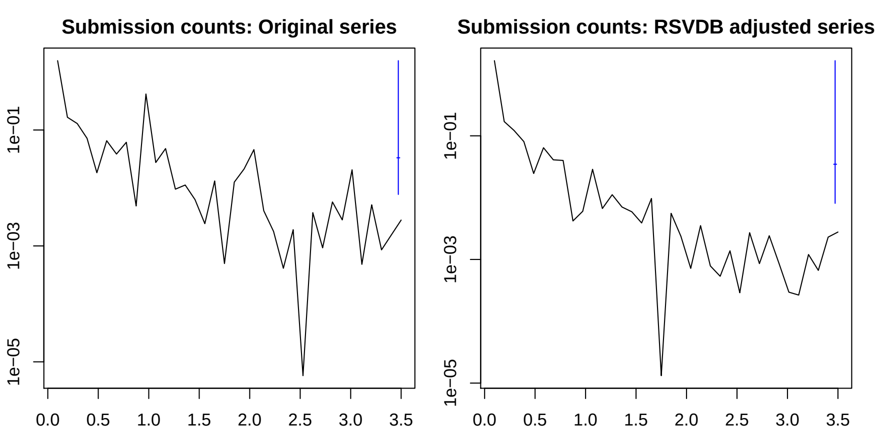

Figure S5: Periodograms of online submission counts series and RSVDB adjusted series