# Does climate change affect economic data?

In Choi[1] 

## Abstract

This paper derives the seasonal factors from the US temperature, gasoline price, and fresh food price data sets using the Kalman state smoother and the principal component analysis. Seasonality in this paper is modeled by the autoregressive process and added to the random component of the time series. The derived seasonal factors show a common feature: their volatilities have increased over the last four decades. Climate change is undoubtedly reflected in the temperature data. The three data sets' similar patterns from the 1990s suggest that climate change may have affected the prices' volatility behavior.

**Keywords** Kalman state smoother · Seasonal factor · Principal component analysis · Climate change · Temperature · Gasoline price · Fresh food price

## 1 Introduction

Climate change is currently one of the most critical global issues. It affects the physical world in various ways: sea-level rise, flooding, harsh and extended droughts, extreme weather events like heat waves and severe tropical storms, changes in the global fauna and flora, etc. It also has an impact on the economy, as documented by Deke et al. (2001), Deschênes and Greenstone (2007), Igbal (2022), Nordhaus (1991, 1994, 2019), Stern (2008), Tol (2009, 2011), among others.

Evidence for climate change has originated from scientific research. Masson-Delmotte et al. (2021) report to IPCC (Intergovernmental Panel on Climate Change) about the differences in the temperature in the last 170 years. According to them, global surface temperature has increased by 0.99 °C from 1850–1900 to 2001–2020 and by 1.09 °C from 1850–1900 to 2011–2020. Moreover, the temperature rise has been most pronounced since 1980, as shown in Figure 3.1 of Trenberth et al. (2007).

✉ In Choi
   inchoi@gmail.com; inchoi@sogang.ac.kr
   http://www.inchoi.site

[1] Sogang University, Seoul, Korea

Using climate proxies such as tree rings, ice cores, fossil leaves, researchers have reconstructed global temperature for the last 2000 years. Most notably, (Mann et al. 1998, 2009) found using global temperature reconstructed from tree rings, coral, and other temperature proxies that the twentieth century was warmer than any other time during the last 400 years. PAGES et al. (2019) present 2000-year-long global mean temperature reconstructions and conclude that the most significant warming trends occur during the second half of the twentieth century, highlighting recent decades' unusual global warming. One can find updates on the most recent changes in global surface temperature from the GISS (Goddard Institute for Space Studies) Surface Temperature Analysis (https://data.giss.nasa.gov/gistemp/). GISS provides evidence for global warming using various kinds of data. The reader is referred to Chapter 6 of Mathez (2018) for an accessible account of the history of climate change.

The evidence for climate change reported thus far is abundant, as discussed above but confined to data from Mother Nature. No one has attempted to find evidence for climate change using data generated by human activities, as far as we know. If the climate change of the last few decades is truly unusual and vital for humanity's living environment, it must also be reflected in economic data.

The purpose of this paper is to try to find evidence for climate change in economic data. For this purpose, we use the last four decades' panel data of US gasoline and fresh food prices and temperature data of several regions in the USA. We focus on their seasonal variability rather than their trend. There are some intuitive reasons why climate change makes gasoline and fresh food prices more variable seasonally. It has been known that climate change has affected not only mean temperatures but also their variances (see, for example, Olonscheck et al. (2021)). Higher temperature variability is often believed to bring climate extremes (extended droughts, hot spells, severe tropical storms, floods, etc.), which can have enormous impacts on agriculture, forestry and aquafarming (see Katz and Brown 1992). These climate extremes can cause the supply of agricultural products to fluctuate much, making their prices more variable seasonally. On the other hand, the variability of mean temperatures affects the demand side of economic activities too. Travels and leisure activities, for example, are often affected by temperatures, and their variability can cause the prices of the products related to those activities to change much seasonally. Gasoline is one of those products.

The empirical strategy of this paper is to extract the seasonal component from each time series by using the Kalman state smoother and construct the seasonal factor from each panel data set by applying the principal component analysis to the collection of the estimated seasonal components. Variability of the seasonal factor of the temperature panel data has increased over the last four decades. Furthermore, we find that the seasonal factors of the gasoline and fresh food price data follow a similar pattern. The three data sets' similar patterns from the 1990s suggest that climate change may have affected the prices' volatility behavior. However, this conclusion is undoubtedly not definitive, and further research is needed to confirm or deny it.

The method we use to extract seasonal components differs from the conventional seasonal adjustment methods adopted by statistical agencies and researchers world-

wide. The primary purpose of the traditional methods like X12-ARIMA,[1] TRAMO (time series regression with ARIMA noise, missing observations, and outliers), and SEATS [2] (signal extraction in ARIMA time series) is to eliminate seasonality from time series, while we are more interested in the seasonal component of the time series in this paper. In conventional deseasonalization methods, one usually adopts the restriction that the sum of the seasonal elements is zero over one year. It makes the deseasonalized time series cohere with the original annual one when they are summed up over the seasons. On the other hand, this practice implies that seasons and their characteristics should not impact economic data. This seems unreasonable in the light of ample evidence that climate change affects the economy (see, e.g., Deschênes and Greenstone 2007; Igbal 2022), in particular its agricultural sector. We model the seasonal component in this paper to be free of such restrictions because we aim to find the impact of seasons (or climate change more precisely) on economic data.

The seasonal component of this paper is stochastic and assumed to follow the seasonal autoregressive model. Stochastic seasonality is not new in the time series literature and was introduced in Chen et al. (2022), Mirantes et al. (2012, 2013), and Proietti (2000). They use either stochastic processes based on trigonometric functions or seasonal random walk processes. So and Chung (2014) employ the seasonal autoregressive model but assumes that the sum of the seasonal elements is zero.

To extract the seasonal component from each time series, we employ the Kalman state smoother. The seasonal component can be distinguished from the random part using the Kalman state smoother. We focus on using the Kalman state smoother because they provide a time series showing smoother behavior than the Kalman filter. One may consider estimating the seasonal factor by applying the Kalman filtering method to panel data (cf. Poncela et al. 2021). This procedure, however, requires inverting large variance-covariance matrices and is not proper for the large data sets we are dealing within this paper.

This paper is planned as follows. Section 2 introduces the model and estimation methods. Section 3 reports empirical results. Section 4 provides summary and further remarks. Appendix I contains mathematical derivations and Appendix II lists the data used in Sect. 3.

A few words on our notation. $\mathbf{0}_a$, $\mathbf{0}_{a \times b}$ and $I_b$ denote the null column vector of dimension $a$, the $a \times b$ null matrix and the identity matrix of dimension $b$, respectively.

## 2 The model and estimation methods

We are concerned with the unobserved components model for the panel data $\{y_{it}\}$

$$y_{it} = s_{it} + x_{it}, \ (i = 1, \ldots, N; \ t = 1, \ldots, T), \tag{1}$$

where $i$ and $t$ are indices for cross-sectional units (e.g., individuals, firms, countries, etc.) and time, respectively; and $s_{it}$ and $x_{it}$ are seasonal and random components,

---

[1] See Findley et al. (1998) and Dagum and Bianconcini (2016) for X-12 ARIMA. Pollock (2021) and Lin20 provide methods that improve on the conventional deseasonalizing methods.

[2] See Gomez and Maravall (1996, 1998) for TRAMO and SEATS.

respectively. Most time series have a trend component, but we assume that the trend component was already eliminated. We can consider the linear detrending, the Hodrick–Prescott filter and Hamilton (2018) regression method to eliminate the trend in the data. We will employ the linear detrending method in Sect. 3. The reason for this is explained there.

The seasonal component is modeled as

$$s_{it} = \beta_i s_{i,t-\tau} + e_{it}, \ e_{it} \sim iid\ (0, \sigma_{e_i}^2),$$

where $\tau$ denotes the number of seasons (e.g., $\tau = 4$ for quarterly data). We assume $|\beta_i| < 1$ for all $i$.[3] The random component is modeled as either

$$x_{it} = \sum_{k=1}^{p} \phi_{ik} x_{i,t-k} + u_{it} + \sum_{l=1}^{q} \theta_{il} u_{i,t-l}, \ u_{it} \sim iid\ (0, \sigma_{u_i}^2) \qquad (2)$$

or

$$\Delta x_{it} = \sum_{k=1}^{p} \phi_{ik} \Delta x_{i,t-k} + u_{it} + \sum_{l=1}^{q} \theta_{il} u_{i,t-l}, \ u_{it} \sim iid\ (0, \sigma_{u_i}^2), \qquad (3)$$

where $\Delta x_{it} = x_{it} - x_{i,t-1}$. We assume that $\{x_{it}\}$ is stationary in the former case, and that $\{\Delta x_{it}\}$ is so in the latter case. Additionally, we assume that $\{e_{it}\}$ and $\{u_{it}\}$ are independent temporally and cross-sectionally.

Our primary interest lies in estimating the seasonal component, while by contrast most other studies focus on eliminating it. For our purpose, the Kalman filtering and smoothing methods are appropriate because they can provide estimates of each component separately. In Sect. 3, we will apply the Kalman smoothing method to each individual, univariate time series and obtain the estimates of $s_{it}$ (more precisely, those of their conditional expectations). Then, we will use the principal component method to extract a common factor from all of them, which we call the seasonal factor.

## 2.1 The case of a stationary random component

This subsection deals with the stationary random component (2) for the Kalman filter and smoother.

### 2.1.1 State space form

The transition equation for the seasonal component is written as

$$\xi_{i,t+1} = V_i \xi_{it} + E e_{i,t+1}, \qquad (4)$$

---

[3] According to some experienced researchers whom the author knows personally, most economic time series do not seem to have seasonal unit roots, although they show quite persistent behavior. See also Chapter 6 of Choi (2015).

where

$$\xi_{it} = \begin{pmatrix} \begin{bmatrix} s_{it} \\ s_{i,t-\tau+1} \\ \vdots \\ s_{i,t-1} \end{bmatrix} \\ \beta_i \end{pmatrix}, \quad V_i = \begin{bmatrix} \mathbf{0}_{\tau-1} & I_{\tau-1} \\ \beta_i & \mathbf{0}'_{\tau-1} \end{bmatrix}, \quad E = \begin{pmatrix} 1 \\ \mathbf{0}_{\tau-1} \end{pmatrix},$$

and that for the random component is

$$\varsigma_{i,t+1} = W_i \varsigma_{i,t} + U_i u_{i,t+1}, \tag{5}$$

where

$$\varsigma_{it} = \begin{bmatrix} x_{it} \\ \phi_{i2} x_{i,t-1} + \ldots + \phi_{ir} x_{i,t-r+1} + \theta_{i1} u_{it} + \ldots + \theta_{i,r-1} u_{i,t-r+2} \\ \phi_{i3} x_{i,t-1} + \ldots + \phi_{ir} x_{i,t-r+2} + \theta_{i2} u_{it} + \ldots + \theta_{i,r-1} u_{i,t-r+2} \\ \vdots \\ \phi_{ir} x_{i,t-1} + \theta_{i,r-1} u_{i,t} \end{bmatrix},$$

$$W_i = \begin{bmatrix} \phi_{i1} \\ \vdots & I_{r-1} \\ \phi_{i,r-1} \\ \phi_{ir} & \mathbf{0}'_{r-1} \end{bmatrix}, \quad U_i = \begin{pmatrix} 1 \\ \theta_{i1} \\ \vdots \\ \theta_{i,r-1} \end{pmatrix},$$

and $r = \max(p, q+1)$. In the special case $r = 1$, $W_i = \phi_{i1}$ and $U_i = 1$.

Let

$$\alpha_{it} = \begin{bmatrix} \xi'_{it} & \varsigma'_{i,t} \end{bmatrix}',$$

$$T_i = \begin{bmatrix} V_i & \mathbf{0}_{\tau \times r} \\ \mathbf{0}_{r \times \tau} & W_i \end{bmatrix},$$

$$\eta_{it} = \begin{pmatrix} e_{i,t+1} \\ u_{i,t+1} \end{pmatrix}, \quad R_i = \begin{bmatrix} 1 & 0 \\ \mathbf{0}_{\tau-1} & \mathbf{0}_{\tau-1} \\ \mathbf{0}_r & U_i \end{bmatrix}.$$

Putting (4) and (5) together, we may write model (1) in state space form as

$$y_{it} = Z\alpha_{it}, \quad Z = [1 \ \mathbf{0}'_{\tau-1} \ 1 \ \mathbf{0}'_{r-1}],$$

$$\alpha_{i,t+1} = T_i \alpha_{it} + R_i \eta_{it}, \quad \eta_{it} \sim iid \ (\mathbf{0}_2, Q_i), \quad Q_i = \begin{bmatrix} \sigma^2_{e_i} & 0 \\ 0 & \sigma^2_{u_i} \end{bmatrix}. \tag{6}$$

### 2.1.2 Kalman filter and state smoother

This Subsubsection briefly introduces the Kalman filter and state smoother for model (6) with a stationary random component, referring the reader to Durbin and Koopman (2012) for the detailed explanations on these procedures. The Kalman filter is

recursions for calculating $a_{i,t+1} = E(\alpha_{i,t+1} \mid y_{it}, ..., y_{i1})$ and $P_{i,t+1} = Var(\alpha_{i,t+1} \mid y_{it}, ..., y_{i1})$, and the state smoother is those for $\hat{a}_{i,t+1} = E(\alpha_{i,t+1} \mid y_{iT}, ..., y_{i1})$ and $V_{i,t+1} = Var(\alpha_{i,t+1} \mid y_{iT}, ..., y_{i1})$.

The initial state vector is written as

$$\alpha_{i1} = a_i + \lambda_{i0},$$

where $a_i = \mathbf{0}_{\tau+r}$ and $\lambda_{i0} \sim N(\mathbf{0}_{\tau+r}, \Lambda_{i0})$. Due to the assumed independence of $\{e_{it}\}$ and $\{u_{it}\}$, we may write

$$\Lambda_{i0} = \begin{bmatrix} S_{i0} & \mathbf{0}_{\tau \times r} \\ \mathbf{0}_{r \times \tau} & X_{i0} \end{bmatrix}. \tag{7}$$

Since $\{s_{it}\}$ and $\{x_{it}\}$ are weakly stationary, $S_{i0}$ and $X_{i0}$ should satisfy the equations

$$S_{i0} = V_i S_{i0} V_i' + \sigma_{e_i}^2 \begin{bmatrix} 1 & \mathbf{0}_{\tau-1}' \\ \mathbf{0}_{\tau-1} & \mathbf{0}_{(\tau-1)\times(\tau-1)} \end{bmatrix},$$
$$X_{i0} = W_i X_{i0} W_i' + \sigma_{u_i}^2 U_i U_i'.$$

Solving these equations for $S_{i0}$ and $X_{i0}$ is not easy in most cases. Alternatively, we may use the linear process representations of $\xi_{it}$ and $\varsigma_{it}$ to obtain $S_{i0}$ and $X_{i0}$ in practice.

The variance of $\alpha_{i1}$ is given as $P_{i1} = \Lambda_{i0}$, and the Kalman filter is

$$v_{it} = y_{it} - Z a_{it}, \quad F_{it} = Z P_{it} Z',$$
$$a_{i,t+1} = T_i a_{i,t} + K_{it} v_{it}, \quad P_{i,t+1} = T_i P_{it} L_{it}' + R_i Q_i R_i', \tag{8}$$

with $a_{i1} = \mathbf{0}_{\tau+r}$, $K_{it} = T_i P_{it} Z' F_{it}^{-1}$ and $L_{it} = T_i - K_{it} Z$. The state smoother is written as

$$\hat{\alpha}_{it} = a_{it} + P_{it} r_{i,t-1}, \quad r_{i,t-1} = Z' F_{it}^{-1} v_{it} + L_{it}' r_{it},$$
$$V_{it} = P_{it} - P_{it} N_{i,t-1} P_{it}, \quad N_{i,t-1} = Z' F_{it}^{-1} Z + L_{it}' N_{it} L_{it}, \tag{9}$$

where $r_{iT}$ and $N_{iT}$ are the null vector and matrix, respectively.

The log-likelihood function for the estimation of unknown parameters is

$$\log L(y_{i1}, ..., y_{in}) = c - \frac{1}{2} \sum_{t=1}^{n} (\ln |F_{it}| + v_{it}' F_{it}^{-1} v_{it}),$$

where $c$ is a constant. The maximum likelihood estimators of the unknown parameters $\beta_i, \sigma_{e_i}^2, \phi_{i1}, ..., \phi_{ip}, \theta_{i1}, .., \theta_{1q}$ and $\sigma_{u_i}^2$ are obtained by maximizing this function.

## 2.2 The case of a random component with a unit root

This subsubsection considers the Kalman filter and state smoother for model (6) with the random component (3) having a unit root. We will use the diffuse initialization for the random component. Some US time series seem to have two unit roots (cf. Choi and Jeong 2020), but extending the case of a unit root to that of double unit roots is straightforward and will be mentioned very briefly in due course.

### 2.2.1 State space form

Only the parts related to the random component need to be changed in the state space form. Letting $x_{it}^* = \Delta x_{it}$, the random component in the state vector is written as

$$
\tilde{\varsigma}_{it} = \begin{bmatrix}
x_{i,t-1} \\
x_{it}^* \\
\phi_{i2}x_{i,t-1}^* + \cdots + \phi_{ir}x_{i,t-r+1}^* + \theta_{i1}u_{it} + \cdots + \theta_{i,r-1}u_{i,t-r+2} \\
\phi_{i3}x_{i,t-1}^* + \cdots + \phi_{ir}x_{i,t-r+2}^* + \theta_{i2}u_{it} + \cdots + \theta_{i,r-1}u_{i,t-r+2} \\
\vdots \\
\phi_{ir}x_{i,t-1}^* + \theta_{i,r-1}u_{i,t}
\end{bmatrix},
$$

which follows the dynamic relation

$$
\tilde{\varsigma}_{i,t+1} = \tilde{W}_i \varsigma_{i,t} + \tilde{U}_i u_{i,t+1},
$$

with

$$
\tilde{W}_i = \begin{bmatrix}
1 & 1 & \mathbf{0}'_{r-1} \\
& \phi_{i1} & \\
\mathbf{0}_r & \vdots & I_{r-1} \\
& \phi_{i,r-1} & \\
& \phi_{ir} & \mathbf{0}'_{r-1}
\end{bmatrix} \text{ and } \tilde{U}_i = \begin{pmatrix}
0 \\
1 \\
\theta_{i1} \\
\vdots \\
\theta_{i,r-1}
\end{pmatrix}.
$$

If there are two unit roots in the random component, $x_{i,t-1}$ and $\Delta x_{i,t-1}$ should be the first two elements of $\varsigma_{it}$ and $x_{it}^*$ should be redefined as $x_{it}^* = \Delta^2 x_{it}$. Some obvious changes should follow for $W_i$ and $U_i$ too.

The state space form is now written as

$$
y_{it} = Z\alpha_{it}, \ Z = [(1 \ \mathbf{0}'_{\tau-1}) \ (1 \ 1 \ \mathbf{0}'_{r-1})]
$$

$$
\alpha_{i,t+1} = T_i\alpha_{it} + R_i\eta_{it}, \ \eta_{it} \sim iid \ (\mathbf{0}_2, Q_i), \ Q_i = \begin{bmatrix} \sigma_{e_i}^2 & 0 \\ 0 & \sigma_{u_i}^2 \end{bmatrix}
$$

with

$$\alpha_{it} = \begin{bmatrix} \xi'_{it} & \tilde{\varsigma}'_{i,t} \end{bmatrix}',$$

$$T_i = \begin{bmatrix} V_i & \mathbf{0}_{\tau \times (r+1)} \\ \mathbf{0}_{(r+1) \times \tau} & \tilde{W}_i \end{bmatrix},$$

$$\eta_{it} = \begin{pmatrix} e_{i,t+1} \\ u_{i,t+1} \end{pmatrix}, \quad R_i = \begin{bmatrix} 1 & 0 \\ \mathbf{0}_{\tau-1} & \mathbf{0}_{\tau-1} \\ \mathbf{0}_{r+1} & \tilde{U}_i \end{bmatrix}.$$

### 2.2.2 Exact initial Kalman filter and state smoother

The initial state vector is written as

$$\alpha_{i1} = a_i + A\delta_i + B\lambda_{i0},$$

where $a_i = \mathbf{0}_{\tau+r+1}$, $A = \begin{bmatrix} \mathbf{0}_\tau \\ 1 \\ \mathbf{0}_r \end{bmatrix}$, $\delta_i = x_{i1}$, $B = \begin{bmatrix} I_\tau & \mathbf{0}_{\tau \times r} \\ \mathbf{0}'_\tau & \mathbf{0}'_r \\ \mathbf{0}_{r \times \tau} & I_r \end{bmatrix}$ and $\lambda_{i0} \sim$ $N(\mathbf{0}_{\tau+r}, \Lambda_{i0})$ with $\Lambda_{i0}$ defined by Eq. (7). We assume $\delta_i \sim N(0, \kappa_i)$ with $\kappa_i \to \infty$. This assumption, together with the initial state vector given above, constitutes the diffuse initialization of the filter. See Durbin and Koopman (2012)'s Chapter 5 for the details of the diffuse initialization.

The variance of $\alpha_{i1}$ is given as

$$P_{i1} = \kappa_i P_{i,\infty,1} + P_{i,*,1}, \tag{10}$$

where $P_{i,\infty,1} = AA'$ and $P_{i,*,1} = B\Lambda_{i0}B'$. Analogously to (10), we have

$$P_{it} = \kappa_i P_{i,\infty,t} + P_{i,*,t},$$

where $P_{\infty,t}$ and $P_{*,t}$ do not depend on $\kappa$. For $t > d$ (a positive integer), we should have $P_{i,\infty,t} = 0$ under the assumption $\kappa_i \to \infty$. Otherwise, it is a contradiction to the finiteness of $P_{it}$. For $t > d$, the usual Kalman filter (8) applies.

As $\kappa_i \to \infty$, we have the following recursive relations for $t = d, \ldots, 1$

$$P_{i,\infty,t+1} = T_i P_{i,\infty,t} L_{it}^{(0)'},$$
$$P_{i,*,t+1} = T_i P_{\infty,t} L_{it}^{(1)'} + T_i P_{i,*,t} L_{it}^{(0)'} + R_i Q_i R_i', \tag{11}$$

where

$$L_{it}^{(0)} = T_i - K_{it}^{(0)} Z, \quad L_{it}^{(1)} = -K_{it}^{(1)} Z,$$
$$K_{it}^{(0)} = T_i P_{i,\infty,t} Z' F_{it}^{(1)}, \quad K_{it}^{(1)} = T_i P_{i,*,t} Z' F_{it}^{(1)} + T_i P_{i,\infty,t} Z'_t F_{it}^{(2)},$$
$$F_{it}^{(1)} = F_{i,\infty,t}^{-1}, \quad F_{it}^{(2)} = -F_{i,\infty,t}^{-1} F_{i,*,t} F_{it}^{(1)},$$
$$F_{i,\infty,t} = Z P_{i,\infty,t} Z', \quad F_{i,*,t} = Z P_{i,*,t} Z'.$$

In addition, as $\kappa_i \to \infty$,

$$a_{i,t+1}^{(0)} = T a_{it}^{(0)} + K_{it}^{(0)} v_{it}^{(0)} \tag{12}$$

with $a_1^{(0)} = a$ and $v_t^{(0)} = y_t - Z_t a_t^{(0)}$. The detailed derivations of (11) and (12) can be found in Koopman (1997) and Durbin and Koopman (2012).

For model (1), we have $P_{i,\infty,2} = \mathbf{0}_{(\tau+r+1)\times(\tau+r+1)}$ as reported in Appendix I. Thus, for $t = 2, 3, \ldots$, the usual Kalman filter (8) can be applied using $a_{i2}$ and $P_{i,*,2}$ (given in Appendix I) as initial variables.

Next, we consider the exact initial state smoother. As $\kappa_i \to \infty$, we have for $t = d, \ldots, 1$

$$\hat{\alpha}_{it} = a_{it}^{(0)} + P_{i,*,t} r_{i,t-1}^{(0)} + P_{i,\infty,t} r_{i,t-1}^{(1)},$$

where

$$r_{i,t-1}^{(0)} = L_{it}^{(0)'} r_{it}^{(0)}, \; r_{id}^{(0)} = r_{id},$$
$$r_{i,t-1}^{(1)} = Z' F_{it}^{(1)} v_{it}^{(0)} + L_{it}^{(0)'} r_{it}^{(1)} + L_{it}^{(1)'} r_{it}^{(0)}, \; r_{id}^{(1)} = 0,$$

and $r_{id}$ is obtained from (9) below.

Let $V_{it} = Var(\alpha_{it} \mid y_{iT}, \ldots, y_{i1})$. Letting $\kappa_i \to \infty$, the smoothed state variance for $t = d, \ldots, 1$ is given by

$$V_{it} = P_{i,*,t} - P_{i,*,t} N_{i,t-1}^{(0)} P_{i,*,t} - (P_{i,\infty,t} N_{i,t-1}^{(1)} P_{i,*,t})'$$
$$- P_{i,\infty,t} N_{i,t-1}^{(1)} P_{i,*,t} - P_{i,\infty,t} N_{i,t-1}^{(2)} P_{i,\infty,t},$$

where

$$N_{i,t-1}^{(0)} = L_{it}^{(0)'} N_{it}^{(0)} L_{it}^{(0)},$$
$$N_{i,t-1}^{(1)} = Z' F_{it}^{(1)} Z + L_{it}^{(0)'} N_{it}^{(1)} L_{it}^{(0)} + L_{it}^{(1)'} N_{it}^{(0)} L_{it}^{(0)},$$
$$N_{i,t-1}^{(2)} = Z' F_{it}^{(2)} Z + L_{it}^{(0)'} N_{it}^{(2)} L_{it}^{(0)} + L_{it}^{(0)'} N_{it}^{(1)} L_{it}^{(1)}$$
$$+ L_{it}^{(1)'} N_{it}^{(1)} L_{it}^{(0)} + L_{it}^{(1)'} N_{it}^{(0)} L_{it}^{(1)}$$

with $N_{id}^{(0)} = N_{id}$ and $N_{id}^{(1)} = N_{id}^{(2)} = 0$. For $t > d$, the usual state smoother (9) can be used with $a_{i,d+1} = a_{i,d+1}^{(0)}$ and $P_{i,d+1} = P_{i,*,d+1}$. Because $d = 1$ for model (1), the usual state smoother can be used for $t = 2, 3, \ldots$.

The diffuse log-likelihood function for the estimation of the unknown parameters $\beta_i, \sigma_{e_i}^2, \phi_{i1}, \ldots, \phi_{ip}, \theta_{i1}, \ldots, \theta_{1q}$ and $\sigma_{u_i}^2$ is

$$\log L(y_{i1}, \ldots, y_{in}) = c - \frac{1}{2} \ln \left| F_{i,\infty,1} \right| - \frac{1}{2} \sum_{t=2}^{n} (\ln |F_{it}| + v_{it} F_{it}^{-1} v_{it})$$

$$= c - \frac{1}{2} \sum_{t=2}^{n} (\ln |F_{it}| + v_{it} F_{it}^{-1} v_{it}),$$

since $F_{i,\infty,1} = 1$ as shown in Appendix. We obtain the maximum likelihood estimators of the unknown parameters by maximizing this function with respect to the unknowns.

### 2.3 Estimation of seasonal factors

Suppose that we have the estimates of $s_{it}$, $\hat{s}_{it}$. We assume that $\hat{s}_{it}$ follows the factor model with the number of factors being equal to one. Then, we may write

$$\hat{s}_{it} = l_i s_t + v_{it},$$

where $l_i$ is the factor loading, $s_t$ is the seasonal factor and $v_{it}$ is the error term (cf. Bai and Ng 2002). Note that $l_i$ and $s_t$ are not identified because $l_i s_t = l_i^* s_t^*$ with $l_i^* = cl_t$ and $s_t^* = s_t/c$ for a nonzero constant $c$. In matrix notation, this model is written as

$$X = SL' + v, \tag{13}$$

where $X = \begin{bmatrix} \hat{s}_{11} & \cdots & \hat{s}_{1N} \\ \vdots & & \vdots \\ \hat{s}_{T1} & \cdots & \hat{s}_{TN} \end{bmatrix}$, $S = \begin{bmatrix} s_1 \\ \vdots \\ s_T \end{bmatrix}$, $L = \begin{bmatrix} l_1 \\ \vdots \\ l_N \end{bmatrix}$ and $v = \begin{bmatrix} v_{11} & \cdots & v_{1N} \\ \vdots & & \vdots \\ v_{T1} & \cdots & v_{TN} \end{bmatrix}$.

The principal component estimator of the factor space, denoted as $\hat{S}$, is $\sqrt{T}$ times the eigenvector corresponding to the largest eigenvalue of the matrix $XX'$ (with the standardization $\hat{S}'\hat{S} = T$), and that of the factor loading is obtained by $\hat{L} = \frac{1}{T} X' \hat{S}$.

## 3 Empirical results

This section reports some empirical results applying the methods described in previous sections. The data we use include temperature, gasoline price, and fresh food price data of US temperatures are most directly connected to climate change. Therefore, their behaviors need to be investigated and compared to those of any other time series data being studied concerning climate change. Gasoline prices are deemed to be affected by climate change because they are used for heating, transportation, and traveling. For example, there tend to be more travelers if spring comes earlier than usual, prompting more gasoline consumption. The supply of fresh food is also expected to be influenced by climate change. For example, higher global temperature will increase the supply

of tropical fruits in the long run, while that of fruits from colder regions will decrease. This section will first study the US temperature data and compare its behavior with those of the US gasoline price and fresh food price data. The sampling period of the three data sets is commonly January 1980–June 2019, and they are seasonally unadjusted, nominal, monthly time series. We deliberately excluded data from 2020 and 2021 because they are likely to be affected by Covid-19 and may contain many unhelpful outliers for data analysis.

The data sets considered in this section are panel data. This section's analysis aims to extract the seasonal factors from the panel data and compare their properties. To this end, we take the following steps for each data set:

(i) Regress each time series in the data set on the time polynomial with order zero or one and calculate residuals.
(ii) Obtain the Kalman state smoother of the seasonal component for each residual series from the previous step.
(iii) Estimate the seasonal factor from the Kalman state smoother using the principal component analysis as in Bai and Ng (2002).

In Step (i), we used the time polynomial with order one when there are noticeable upward trends in the data. Besides this, we can consider two other methods to detrend the data: the Hodrick–Prescott filter and Hamilton (2018) regression method. However, according to Hamilton (2018), the Hodrick–Prescott filter may induce spurious cycles in the seasonal and random components; it is not yet known how Hamilton (2018) methods work for seasonal data. Thus, we used the linear detrending method for the price data sets studied below.

In Step (ii), we focused on the Kalman state smoother because they show much more stable behavior than the Kalman filters. The Kalman filters often generated series with many outliers according to unreported results. In addition, we used BIC for the model selection of the random component.

More precisely, the seasonal factor from Step (iii) estimates the space of the seasonal factor because that is what the principal component analysis delivers for the factor model. Thus, we may use any multiples of the seasonal factors from Step (ii) in practice. We examined whether the seasonal factor and the Kalman state smoothers move close to each other, following a reviewer's suggestion, and found that they indeed do so.

## 3.1 Temperature

The temperature data of four US cities we use—Boston, Los Angeles, Austin TX, and Springfield IL—were taken from the National Centers for Environmental Information (https://www.ncdc.noaa.gov). They are the average temperatures of each month. Although it is widely known that temperatures have gone up globally over the last one hundred years, the data do not show any noticeable trends during the 40 years of the sampling period. Thus, the data were only demeaned for subsequent analysis. In addition, we assumed the presence of a unit root in the random component because
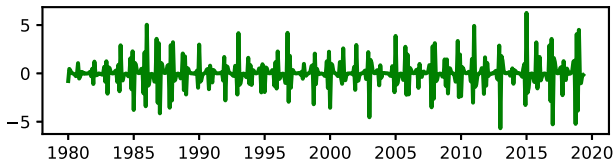
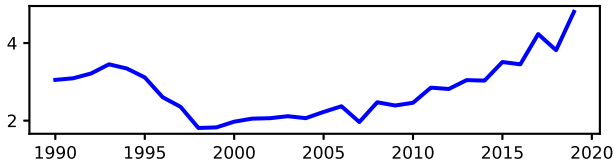**Fig. 1** Seasonal factor from average temperatures



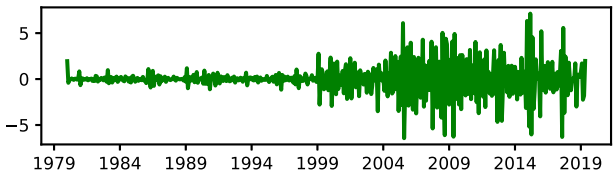**Fig. 2** Variances of seasonal factor from average temperatures



**Fig. 3** Seasonal factor from gasoline prices

we could obtain a better fitting of the data by doing so.[4] The Kalman state smoother sometimes show quite irregular behavior under the stationarity assumption.

We plot the seasonal factor from the 4 Kalman state smoother in Fig. 1. The figure indicates that the volatility of the time series changes over time: the series appears to be more volatile during the last 15 years or so. The figure is typical of seasonal time series because regular peaks and troughs are observed. But the time series in the figure is also atypical seasonal time series because similar patterns are not repeated. This aspect stems from the stochastic seasonal component modeled uniquely in this paper. If the seasonal factors show identical patterns over time, they would not have much to tell about climate change.

Figure 1 indicates that the seasonal factor may have a variance that grows with time. To investigate this, we drew empirical variances of the seasonal factor in Fig. 2 using the rolling widows of the length of 120 months. The time difference between the adjacent windows was set at 12 months. The horizontal axis of Fig. 2 denotes the endpoints of the sampling period of each window. As we can expect from Fig. 1, the rolling variances of the seasonal factor are shown to increase since the 1990s in Fig. 2. Significant aspects of climate change are higher temperatures themselves and greater volatilities of temperatures, which often bring unexpected weather conditions like flooding and hurricanes. Figure 2 squares with these latter historical observations.

---

[4]  When stationarity is assumed for the random component, we often obtained ill-behaved Kalman state smoother having many outliers. In addition, higher values of BIC are observed for some temperature data sets. In the literature on climate change, some authors argued for the presence of a unit root in the temperature; see Chang et al. (2020) and Kaufmann et al. (2013), among others.
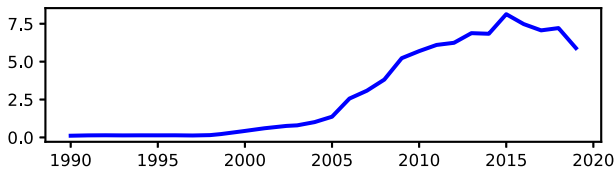
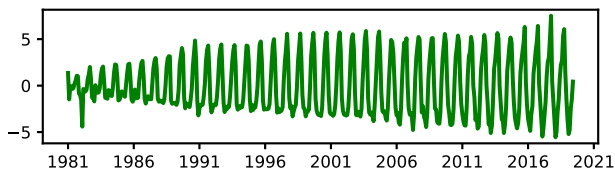Fig. 4 Variances of seasonal factor from gasoline prices



Fig. 5 Seasonal factor from food prices

The patterns in Figs. 1 and 2, however, were obtained by using only the four areas' temperature data in the USA. Whether they can be repeated in other regions, as asked by one of the reviewers, is an important question we need to pursue in the future.

### 3.2 Gasoline prices

We performed the same analysis as the previous subsection using the 13 US gasoline price time series taken from the FRED (https://fred.stlouisfed.org/). The list of the data is reported in Table A.1 in Appendix II. The data contain average, seasonally unadjusted monthly prices of leaded and unleaded types of gasoline in various parts of the USA. Since all the time series appear to have upward trends, we used the time polynomial of order one in Step (i) above. Moreover, we assumed the presence of a unit root in the random component because most time series of nominal prices are known to have a unit root.

We plot the seasonal factor from the 13 Kalman state smoother in Fig. 1. The figure shows that the time series become more volatile beginning from 1999 and that the volatility level has increased since then. The figure also displays regular peaks and troughs, as in Fig. 1. Next, we drew rolling variances of the seasonal factor in Fig. 4 using the same methods as Fig. 2. The seasonal factor's rolling variances have increased since the 1990s in Fig. 4, which is quite similar to Fig. 2.

### 3.3 Fresh food prices

The fresh food price data we use were taken from the FRED (https://fred.stlouisfed.org/) and contain average, seasonally unadjusted monthly prices of 14 food items. The list of the data is reported in Table 2 in Appendix II. The food prices are those of fruits and vegetables in various parts of the USA. The same item has different prices in each region but does not differ much.
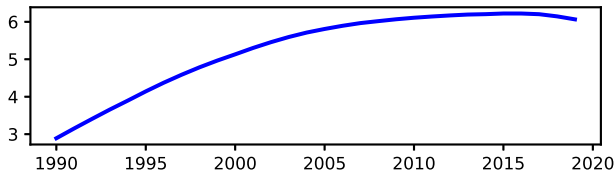
**Fig. 6** Variances of seasonal factor from food prices

All the data show noticeable upward trends. So we used the time polynomial of order one in Step (i) above. Additionally, we assumed the presence of a unit root in the random component because most time series of nominal prices are known to have a unit root.

We plot the seasonal factor from the 14 Kalman state smoothers in Fig. 5. We find from this figure that the seasonal factor shows typical seasonal patterns and displays more wild behavior over time. Next, we drew rolling variances of the seasonal factor in Fig. 6 using the same methods as the previous subsections. The seasonal factor's rolling variances have increased steadily since 1980. Both Figs. 4 and 6 show an upward trend. The value of the correlation coefficient between the two sets of the variances is 0.77, confirming their similar patterns.

## 4 Summary and further remarks

We have derived the seasonal factors from the US temperature, gasoline price, and fresh food price data sets using the Kalman state smoother and the principal component analysis. They have a common feature: their volatilities have increased over the last four decades. Climate change is undoubtedly reflected in the temperature data. It is uncertain yet that climate change caused the volatility behaviors of the price data sets. However, the three data sets' similar patterns from the 1990s suggest that climate change may have affected the prices' volatility behavior. We need to pursue further studies to confirm their relationship.

## Declarations

**Conflict of interest** Author In Choi declares that he has no conflict of interest.

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

## Appendix I: Derivation of $a_{i2}$ and $P_{i,\infty,2}$ for the unit root case

Here, $X(a, b)$ and $X(\mathbf{1})$ denote the $(a, b)$th element and the first column of the matrix $X$, respectively. We will use almost the same notation as in Durbin and Koopman (2012).

We have by the given assumptions

$$a_{i1} = \mathbf{0}_{\tau+r+1}, \quad P_{i,\infty,1} = AA' = \begin{bmatrix} \mathbf{0}_{\tau\times\tau} & \mathbf{0}_{\tau} & \mathbf{0}_{\tau\times r} \\ \mathbf{0}'_{\tau} & 1 & \mathbf{0}'_{r} \\ \mathbf{0}_{r\times\tau} & \mathbf{0}_{r} & \mathbf{0}_{r\times r} \end{bmatrix}$$

and

$$P_{i,*,1} = B\Lambda_{i0}B' = \begin{bmatrix} S_{i0} & \mathbf{0}_{\tau} & \mathbf{0}_{\tau\times r} \\ \mathbf{0}'_{\tau} & 0 & \mathbf{0}'_{r} \\ \mathbf{0}_{r\times\tau} & \mathbf{0}_{r} & X_{i0} \end{bmatrix}.$$

Using these, we obtain

$$F_{i,\infty,1} = Z P_{i,\infty,1} Z' = 1,$$
$$F_{i,*,1} = Z P_{i,*,1} Z' = X_{i0}(1,1) + S_{i0}(1,1)$$
$$M_{i,\infty,1} = P_{i,\infty,1} Z' = \begin{bmatrix} \mathbf{0}_{\tau} \\ 1 \\ \mathbf{0}_{r} \end{bmatrix}, \quad M_{i,*,1} = P_{i,*,1} Z' = \begin{bmatrix} S_{i0}(1) \\ 0 \\ X_{i0}(1) \end{bmatrix},$$
$$K_{i1}^{(0)} = T_i M_{i,\infty,1} F_{i,\infty,1}^{-1} = \begin{bmatrix} \mathbf{0}_{\tau} \\ 1 \\ \mathbf{0}_{r} \end{bmatrix},$$
$$L_{i1}^{(0)} = T_i - K_{i1}^{(0)} Z = \begin{bmatrix} V_i & \mathbf{0}_{\tau} & \mathbf{0}_{\tau\times r} \\ (-1\ \mathbf{0}'_{\tau-1}) & 0 & \mathbf{0}'_{r} \\ \mathbf{0}_{r\times\tau} & \mathbf{0}_{r} & W_i \end{bmatrix},$$
$$K_{i1}^{(1)} = T_i M_{i,*,1} F_{i,\infty,1}^{-1} + T_i M_{i,\infty,1}(-F_{i,\infty,1}^{-2} F_{i,*,1}),$$

and

$$L_{i1}^{(1)} = -K_{i1}^{(1)} Z.$$

These results are used to derive

$$a_{i2} = K_{i1}^{(0)} v_{i1}^{(0)} = \begin{bmatrix} \mathbf{0}_{\tau} \\ y_{i1} \\ \mathbf{0}_{r} \end{bmatrix},$$
$$P_{i,\infty,2} = T_i P_{i,\infty,1} L_{i1}^{(0)'} = \mathbf{0}_{(\tau+r+1)\times(\tau+r+1)}$$

and

$$P_{i,*,2} = T_i P_{i,\infty,1} L_{i1}^{(1)'} + T_i P_{i,*,1} L_{i1}^{(0)'} + R_i Q_i R_i'.$$

# Appendix II: Data list

**Table 1** List of gasoline price data

| Series ID | Title |
| --- | --- |
| APU000074714 | Unleaded regular in US City Average |
| APU00007471A | All types in US City Average |
| APU030074714 | Unleaded regular in the South Census Region-Urban |
| APUS11A74714 | Unleaded regular in Boston-Cambridge-Newton, MA-NH (CBSA) |
| APUS12A74714 | Unleaded regular in New York-Newark-Jersey City, NY-NJ-PA (CBSA) |
| APUS12A7471A | All types in New York-Newark-Jersey City, NY-NJ-PA (CBSA) |
| APUS12B74714 | Unleaded regular in Philadelphia-Camden-Wilmington, PA-NJ-DE-MD (CBSA) |
| APUS37A7471A | All types in Dallas-Fort Worth-Arlington, TX(CBSA) |
| APUS37B74714 | Unleaded regular in Houston-The Woodlands-Sugar Land, TX (CBSA) |
| APUS49A74714 | Unleaded regular in Los Angeles-Long Beach-Anaheim, CA (CBSA) |
| APUS49A7471A | All types in Los Angeles-Long Beach-Anaheim, CA (CBSA) |
| APUS49B74714 | Unleaded regular in San Francisco–Oakland–Hayward, CA (CBSA) |
| APUS35B7471A | All types in Miami-Fort Lauderdale-West Palm Beach, FL (CBSA) |

The following table contains the list of gasoline price data studied in Sect. 3.2. The original source of all the series is the US Bureau of Labor Statistics, and their units are US dollars. The prices denote cost per gallon

**Table 2** List of fresh food price data

| Series ID | Title |
| --- | --- |
| APU0000712211 | Potatoes, white (Cost per Pound) in US City Average |
| APU0000712311 | Tomatoes, field grown, Per Lb. in US City Average |
| APU0100711211 | Bananas, PerLb. in Northeast Urban |
| APU0100711411 | Grapefruit, PerLb. in Northeast Urban |
| APU0200711211 | Bananas, PerLb. in Midwest Urban |
| APU0200711411 | Grapefruit, PerLb. in Midwest Urban |
| APU0200711412 | Lemons, PerLb. in Midwest Urban |
| APU0300711211 | Bananas, PerLb. in South Urban |
| APU0300711411 | Grapefruit, PerLb. in South Urban |
| APU0300711412 | Lemons, PerLb. in South Urban |
| APU0300712211 | Lettuce, Iceberg, PerLb. in South Urban |
| APU0400711211 | Bananas, PerLb. in West Urban |
| APU0400711411 | Grapefruit, PerLb. in West Urban |
| APU0400711412 | Lemons, PerLb. in West Urban |

The following table contains the list of fresh food price data studied in Sect. 3.3. The original source of all the series is the US Bureau of Labor Statistics, and their units are US dollars

# References

Bai J, Ng S (2002) Determining the number of factors in approximate factor models. Econometrica 70(1):191–221

Chang Y, Kaufmann RK, Kim CS, Miller JI, Park JY, Park S (2020) Evaluating trends in time series of distributions: A spatial fingerprint of human effects on climate. J Econom 214(1):274–294

Chen SH, Chiou-Wei SZ, Zhu Z (2022) Stochastic seasonality in commodity prices: the case of US natural gas. Empirical Econom 62(5):2263–2284

Choi I (2015) Almost all about unit roots: Foundations, developments, and applications. Cambridge University Press, Cambridge

Choi I, Jeong H (2020) Differencing versus nondifferencing in factor-based forecasting. J Appl Economet 35(6):728–750

Dagum EB, Bianconcini S (2016) Seasonal adjustment methods and real time trend-cycle estimation. Springer International Publishing, Berlin/Heidelberg, Germany

Deke O, Hooss KG, Kasten C, Klepper G, Springer K (2001) Economic impact of climate change: simulations with a regionalized climate-economy model (No. 1065). Kiel working paper

Deschênes O, Greenstone M (2007) The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather. Am. Econ Rev 97(1):354–385

Durbin J, Koopman SJ (2012) Time series analysis by state space methods, vol 38. Oxford University Press, Oxford

Findley DF, Monsell BC, Bell WR, Otto MC, Chen BC (1998) New capabilities and methods of the X-12-ARIMA seasonal-adjustment program. J Business Econ Stat 16(2):127–152

Gomez V, Maravall A (1996) Programs TRAMO (Time series regression with arima noise, missing observations, and outliers) and SEATS (Signal extraction in arima time series). Instructions for the User. Documento de Trabajo, 9628, 56

Gomez V, Maravall A (1998) Seasonal adjustment and signal extraction in economic time series. Banco de Espana, Servicio de Estudios

Hamilton JD (2018) Why you should never use the Hodrick-Prescott filter. Rev Econ Stat 100(5):831–843

Igbal MR (2022) The economic impact of climate change on the agricultural system in Fiji. J Agric Sci 14(2):144

Katz RW, Brown BG (1992) Extreme events in a changing climate: variability is more important than averages. Clim Change 21(3):289–302

Kaufmann RK, Kauppi H, Mann ML, Stock JH (2013) Does temperature contain a stochastic trend: linking statistical results to physical mechanisms? Clim Change 118(3–4):729–743

Koopman SJ (1997) Exact initial Kalman filtering and smoothing for nonstationary time series models. J Am Stat Assoc 92(440):1630–1638

Lin W, Huang JZ, McElroy T (2020) Time Series seasonal adjustment using regularized singular value decomposition. J Business Econ Stat 38(3):487–501

Masson-Delmotte V, Zhai P, Pirani A, Connors SL, Pé an C, Berger S, Zhou B (2021) Climate change 2021: the physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change, 2

Mann ME, Bradley RS, Hughes MK (1998) Global-scale temperature patterns and climate forcing over the past six centuries. Nature 392(6678):779–787

Mann ME, Zhang Z, Rutherford S, Bradley RS, Hughes MK, Shindell D, Ni F (2009) Global signatures and dynamical origins of the little ice age and medieval climate anomaly. Science 326(5957):1256–1260

Mathez EA (2018) Climate change: the science of global warming and our energy future. Columbia University Press, New York

Mirantes AG, Población J, Serna G (2012) The stochastic seasonal behaviour of natural gas prices. Eur Financ Manag 18(3):410–443

Mirantes AG, Población J, Serna G (2013) The stochastic seasonal behavior of energy commodity convenience yields. Energy Econ 40:155–166

Nordhaus WD (1991) To slow or not to slow: the economics of the greenhouse effect. Econ J 101(407):920–937

Nordhaus WD (1994) Managing the global commons: the economics of climate change, vol 31. MIT Press, Cambridge, MA

Nordhaus W (2019) Climate change: The ultimate challenge for economics. Am Econ Rev 109(6):1991–2014

Olonscheck D, Schurer AP, Lücke L, Hegerl GC (2021) Large-scale emergence of regional changes in year-to-year temperature variability by the end of the 21st century. Nat Commun 12(1):1–10

PAGES 2k Consortium, Neukom, R, Barboza LA, Erb MP, Shi F, Emile-Geay J, Evans MN, Franke J, Kaufman DS, Lücke L, Rehfeld K, Schurer A, Zhu F, Brönimann S, Hakim GJ, Henley BJ, Ljungqvist FC, McKay N, Valler V, von Gunten, L (2019) Consistent multi-decadal variability in global temperature reconstructions and simulations over the Common Era. Nature Geosci 12(8):643–649

Pollock DSG (2021) Enhanced methods of seasonal adjustment. Econometrics 9(1):3

Poncela P, Ruiz E, Miranda K (2021) Factor extraction using Kalman filter and smoothing: This is not just another survey. Int J Forecast 37(4):1399–1425

Proietti T (2000) Comparing seasonal components for structural time series models. Int J Forecast 16(2):247–260

So MK, Chung RS (2014) Dynamic seasonality in time series. Comput Stat Data Anal 70:212–226

Stern N (2008) The economics of climate change. Am Econ Rev 98(2):1–37

Tol RS (2009) The economic effects of climate change. J Econ Perspect 23(2):29–51

Tol RJ (2011) The economic impact of climate change in the 20th Century. Papers WP376, Economic and Social Research Institute (ESRI)

Trenberth KE, Jones PD, Ambenje P, Bojariu R, Easterling D, Tank AK, Zhai P (2007) Observations: surface and atmospheric climate change. In: Climate change 2007: the physical science basis. contribution of working group 1 to the 4th assessment report of the intergovernmental panel on climate change. Cambridge University Press