Bayesian Analysis (2024) **TBA**, Number TBA, pp. 1–25

Modeling Extreme Events in Time Series and Their Impact on Seasonal Adjustment in the Post-Covid-19 Era

Anindya Roy[*,†] and Tucker S. McElroy[†]

Abstract. A conventional approach to the extraction of latent components in a time series is to first model extreme values (including level shifts and seasonal outliers) as fixed effects, followed by their removal. Then the extreme-value adjusted series can be filtered using linear (Gaussian) techniques. A drawback is that identification of the epochs of extreme values is needed, and the uncertainty about this identification – as well as the removal of extremes – goes unmeasured. Alternatively, each outlier effect can be modeled as a particular type of latent stochastic process driven by heavy-tailed innovations; extraction of latent components then follows non-linear techniques and does not require identification of extreme epochs. We model monthly retail data impacted by the Covid-19 epidemic by incorporating additive outliers and level shifts as heavy-tailed latent processes, and estimate the unknown parameters through a Bayesian approach that utilizes Gibbs sampling. As a result, we can extract retail trends that incorporate stochastic level shifts and a full measure of the extraction uncertainty. An added benefit of the proposed approach is an estimate of a counterfactual trend following an extreme event. The posterior estimate of the counterfactual trend can be used to quantify the impact of an extreme event.

MSC2020 subject classifications: Primary 62F15, 62M10; secondary 91B84.

Keywords: Covid-19, impact of extreme events, signal extraction, structural time series.

1 Introduction

Economic time series are typically sensitive to global shocks, such as natural disasters, terrorist events, or epidemics. However, the form that such a shock will take depends on the particular variables and the type of event. Although in most examples of extreme events the timing of the event may be known, the impact of the event could be an instantaneous effect on the variable of interest, or could be a prolonged series of effects over a period of time following the event. The nature of the impact during a period following the event may be quite involved, with aftershocks and echoes of the event producing a complex structure of decay and dissipation of shocks. Possibly the timing of an event (or the event itself), which could be constructed from more complex dynamics in the global market, may elude detection.

[*]University of Maryland Baltimore County, Baltimore, MD, anindya@umbc.edu
[†]U.S. Census Bureau, Suitland, MD, Tucker.S.McElroy@census.gov

 https://doi.org/10.1214/24-BA1424

Regardless of whether the exact event timing is known or not, extreme events can have a substantial impact on the trajectory of a time series, and can adversely impact statistical modeling, estimation, and prediction. In particular, the important task of seasonal adjustment may be adversely affected if extreme events are not accounted for – see McElroy and Penny (2019). Typically, every attempt is made at statistical agencies to detect extreme events, and account for their impact prior to seasonally adjusting the time series. This is often done by attributing unusual movements in the trajectory to event-specific regressors (this is the regression-ARIMA approach described in Findley et al. (1998)). This process, while popular, suffers from the usual problems of multistage procedures. In widespread events – such as the Covid-19 pandemic – that affect several different aspects of global dynamics, attribution of structural changes in the dynamics to specific regressors may be difficult and misleading.

In this paper, we acknowledge that most economic time series are exposed to unusual shocks – not just during the epochs associated with known extreme events, but also during times of market regularity. Hence, each process can be described as the sum of an orthodox part, generally modeled using a Gaussian process, and an aberrant part, which historically has been modeled using deterministic dummy regressors (Tsay (1986)). We propose a framework to accommodate the modeling of the orthodox and aberrant portions together, using a single stochastic model that allows for potentially large changes in the dynamics via heavy-tailed distributions. This synthesis removes some of the criticism of the multistage modeling exercise. We use a Bayesian Structural Time Series (BSTS) model to automatically detect and adjust extreme values.

The idea of using heavy-tailed shocks to accommodate structural changes is not new. For example, Trimbur (2010) used a heavy-tailed model to describe a 'local-level' model with large changes; Maiz et al. (2012), Harvey and Luati (2014), Calvet et al. (2015), and Crevits and Croux (2018), among several others, have used dynamic linear models with heavy-tailed errors to propose robust filtering that can accommodate outliers in structural time series models. Our proposal builds upon these approaches, providing a treatment of seasonality and more general dynamics, while using a fully Bayesian framework that can automate the seasonal adjustment and extreme-value adjustment of time series.

We also provide an approach for evaluating the impact of an extreme event by constructing a counterfactual idealized trend following the event. The comparison of the idealized trend and the estimated trend provides an insight into the changes in the long-term behavior of the series due to the shocks infused by the extreme event. Estimation of the impact of an extreme event is closely related to several procedures that are designed for evaluating a causal estimate of the impact. Often economists and scientists are interested in estimating a counterfactual quantity, viz. the values of the series beyond the time of the extreme event in the absence of such an event. Economists may ask what the trend would look like had the extreme event not occurred; this is in the purview of causal inference and counterfactual estimation. Traditionally, causal inference for time series is done in panel data using the *difference-in-difference* (Did) approach (Angrist and Krueger (1999), Angrist and Pischke (2008), Lechner (2010)) where it is possible get control time series that are not subjected to a treatment or intervention, and hence differences of post- and pre-treatment differences can be evaluated between treatment and control time series. For a single time series, exposed to an extreme event/intervention, pre- and post-treatment control is a counterfactual. Brodersen et al. (2014) delineated a modeling framework based on BSTS that allows prediction of the counterfactual control time series following an extreme event, hence being amenable to causal inference on the effect of the extreme event. The methodology produces a posterior predictive distribution of a series post-event based on the data from prior events. Several articles spanning diverse disciplines have used the Brodersen et al. (2014) framework to evaluate the impact of extreme events – particularly the Covid-19 pandemic – on time series of interest; for example, see Takyia and Bentum-Enninb (2021) and Dey et al. (2021).

Our proposed framework for extreme-value adjustment is also based on BSTS. However, the proposed framework does not set out to estimate a causal estimand, but rather produces an estimate of a latent orthodox component of the time series, i.e., that portion that is free from the influence of any aberrant shocks due to extreme events. Being a fully Bayesian methodology wherein the volatilities associated with extreme events are also estimated, the framework is able to discount the extra volatility in the dynamic components following an event, and thereby estimate the trend as it would appear without the extreme event. The Bayesian framework is able to quantify uncertainty in the estimates measuring the impact of extreme events.

The paper makes several contributions. First, general types of outlier processes are defined, starting with the deterministic setup of Tsay (1986, 1988). Second, we use these outlier processes to model trend, seasonal, and irregular components in economic time series, obtaining the extracted components via Gibbs sampling, and thereby constructing seasonally adjusted series as a sum total of all non-seasonal components. This provides a feasible way of performing seasonal adjustment in the presence of numerous outliers of various magnitudes and types without the manual intervention of choosing time-specific outlier regressors. The fully Bayesian framework allows quantification of the uncertainty in the extraction while avoiding the pitfalls of selecting specific intervention dates. Third, the paper introduces the framework of a counterfactual trend, and develops the corresponding estimation of orthodox components in the time series. The framework provides the quantification of posterior uncertainty of all unknown quantities, including the counterfactual estimates.

2 Background

The impact of extreme events/interventions on different economic time series can be disparate in its scope, depending on the type of event and the type of the series. Accounting for event- and series-specific effects can pose considerable challenges in the statistical modeling of time series with extreme values. Here we study five monthly time series (not seasonally adjusted) at the U.S. aggregate level (units of millions of dollars); the span of each series is January 2001 through August 2022. The series[1] are

- Retail Trade and Food Services 722: Food Services and Drinking Places: Data Extracted on October 18, 2022 (11:23 am)

[1]Downloaded from the U.S. Census Bureau website https://www.census.gov/econ/currentdata/.

- Manufacturers' Shipments, Inventories, and Orders Total Manufacturing: Data Extracted on October 18, 2022 (11:28 am)
- Total for Durable Goods: Data Extracted on October 18, 2022 (11:26 am)
- Grocery Stores, series 4451: Data Extracted on October 18, 2022 (11:25 am)
- Building Materials and Garden Equipment, series 444: Data Extracted on October 18, 2022 (11:30 am)

Retail series for restaurants and bars are known to be "recession-proof," in that such series (shown in the uppermost panel of Figure 1) do not suffer the same dips that may be evident in construction or other retail sectors. But during the Covid-19 crisis, the values of such series have dropped dramatically, suggesting that modeling via a level shift (or sequence of additive outliers) is necessary. The economic reasons for such disparate effects are fairly clear: during a slow-down of the economy due to cyclical factors (or a housing bust) citizens continue to frequent their favorite pubs and eateries, whereas during March and April of 2020 the states of America had various degrees of a lockdown status imposed by the governors, so that only take-out orders for these retailers were legal. However, what is not clear is how to account for the lingering effect of the pandemic and lockdown using the traditional static regression approaches.

On the other hand, the effect of the pandemic and the recession are very different for Grocery Stores (second panel from the top on Figure 1). While the Great Recession effect on this series is similar to that observed in the Food Services and Drinking Places, with a slight flattening of the trend, the Grocery Stores series had a significant uptick at the beginning of the pandemic due to the hoarding behavior of consumers.

Figure 1 displays three other time series over the period January 2001 through August 2022: Total Manufacturing, Durable Goods, and Building Materials and Garden Equipment. All series show some effect of the two documented extreme events, viz. the Great Recession of 2007-2009 and the Covid-19 pandemic that began in March 2020. However, the effect is also very different for the remaining three series. Total Manufacturing and Durable Goods both suffered a significant drop during the recession, and it took several months for the series to return to the pre-recession levels. Subsequent to the recession there seems to be a level shift, in terms of the long-term trend in both series. The effect of Covid-19 is more short-lived for both series, resembling a temporary change, with the series returning to the pre-pandemic trend by the summer of 2020. In contrast, Building Materials and Garden Equipment seem to be less affected by the pandemic, although it also exhibited a shift in the trend during the Great Recession. These plots illustrate that models of macroeconomic time series experiencing extreme events must be flexible yet specific.

A model that has been adequate for a time series during an epoch of normalcy will, at time of crisis, fail unless it is updated to account for new features. One approach is to use an indicator – this could be another time series, or an appropriately chosen sequence of dummies – in the model of the mean function, essentially introducing the indicator as a regressor to the model. Intervention analysis (Box et al. (1975)) takes a single impulse

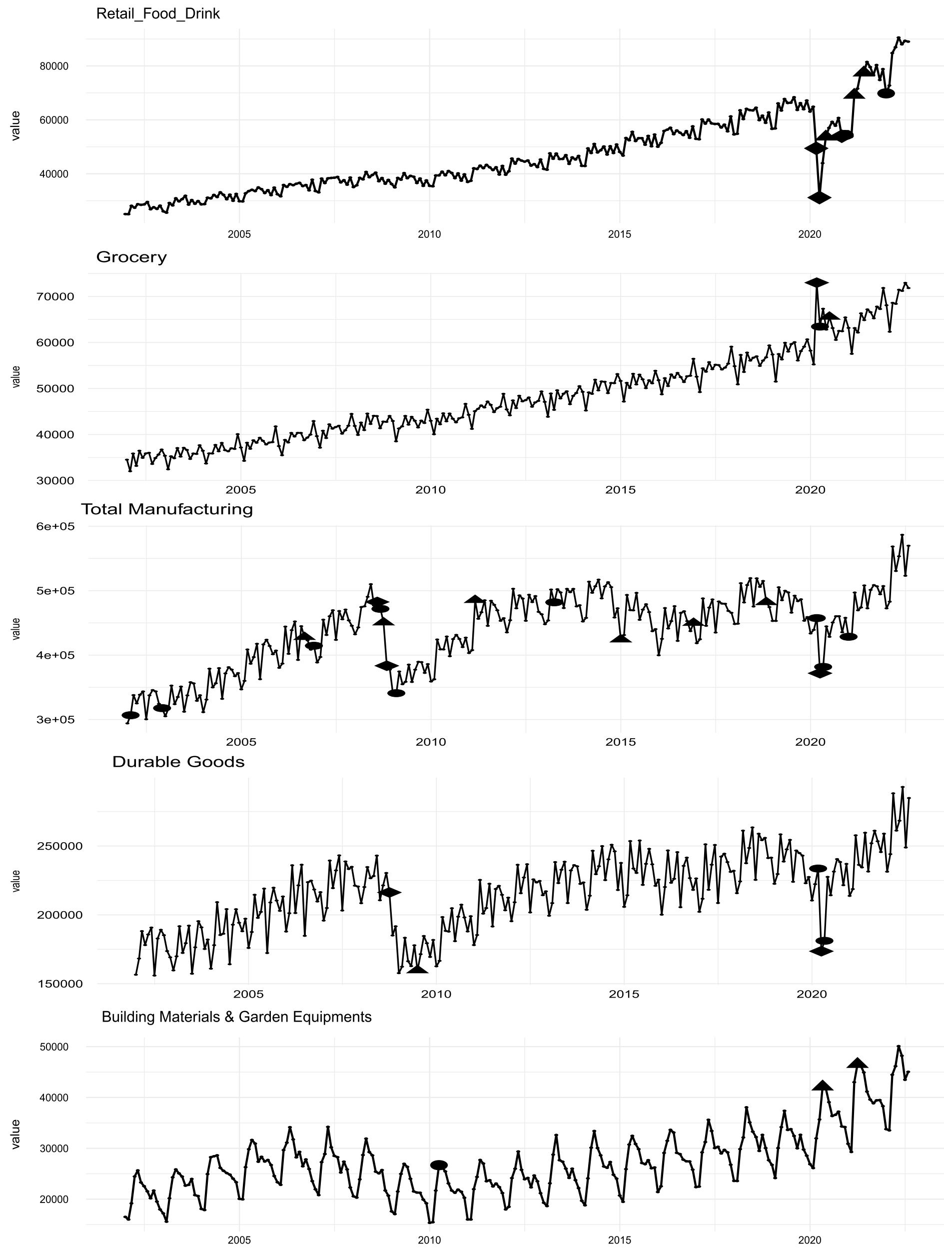

Figure 1: Five aggregate monthly series (in millions of dollars): Food Services and Drinking places; Grocery; Total Manufacturing; Durable goods; Building Materials and Garden Equipment (top to bottom); the outlier points detected by the seasonal R package, as given in Table 1, are marked by circle (AO), triangle (LS) and diamond (TC).

dummy as the indicator of an additive extreme, and this has been generalized in many ways.

Such an approach is predicated on knowing the timing of the crisis and treating the offset as a fixed additive effect. Alternatively, a heavy-tailed distribution could be used to model the magnitude of distortion caused by a shock, while remaining agnostic about the exact timing. This could be advantageous in situations where we are unsure whether there is a lag between the onset of the shock and its impact. For example, we know that on September 11, 2001, a deadly terrorist attack was carried out in the USA, but the timing of this event's impact on airline passenger sales data is harder to pinpoint. Likewise, with the Covid-19 crisis, the lockdown orders were issued with different timing in the various states of America, so that national-level data is a mingling of staggered, local, self-inflicted level shifts.

Conventionally at the US Census Bureau (USCB), sustained crisis effects have been modeled with various types of intervention effects, such as an additive outlier (AO), a level shift (LS), a temporary change (TC), or a quadratic ramp, etc. This methodology is discussed in Maravall and Pérez (2012). First the timing and type of intervention effect must be determined (often by expensive sweeps through the series), followed by extreme-value adjustment; finally, a linear extraction of components of interest, such as the non-seasonal component, is obtained – and extremes may be re-introduced to the final result if they pertain to the target component. In particular, for seasonal adjustment, both additive outliers and level shifts must be identified and removed before filtering, but are finally added back to the resulting extraction (because additive outliers pertain to the irregular movements, and level shifts pertain to the trend). In essence, the trend (as well as the irregular) is implicitly being divided into two portions: the normal part given by a Gaussian process, and the aberrant part given by a level shift dummy. Instead of this "Jekyll and Hyde" decomposition of the trend, we can envision a synthesis where the orthodox and aberrant portions are modeled as a single non-Gaussian process. The drawback, however, is that linear Gaussian techniques of signal extraction are no longer applicable; see McElroy (2016) for an overview of extreme processes and the signal extraction problem in the heavy-tailed context, and Tsay et al. (2000) for a framework for outlier detection and identification in time series.

Non-Gaussian signal extraction is naturally approached through a Bayesian framework. McCulloch and Tsay (1994) handle the modeling of an aberrant trend process in much the way described above, though with the random walk innovations being given by a Bernoulli variable times a heavy-tailed variable. Trimbur (2010) used Student's t innovations for the random walk trend. The Bayesian approach treats all latent components as parameters of interest, which in turn are governed by hyper-parameters specified by the latent models; then posterior means for the components correspond to point estimates of signal extractions in the frequentist context, and the posterior distribution allows us to quantify all the uncertainty – see Holan and McElroy (2012) for Bayesian seasonal adjustment. The Bayesian framework is naturally suited for extracting latent component models with ARMA components. Ravishanker and Ray (1997) have discussed Gibbs sampling algorithms for ARMA time series. More recently, Scott and Varian (2014) have developed efficient Bayesian computation for structural time series.

3 Extreme values and seasonal adjustment

The procedure of seasonally adjusting a time series goes through several important pre-processing steps before an appropriate linear filter can be applied to remove seasonality. One of the key pre-processing steps is the detection and removal of outliers, which here are defined as unusually large (in either direction from the mean level) observations. Linear filters are optimal for Gaussian processes, which do not exhibit many outliers; when numerous substantial outliers are present, linear filtering will be adversely affected. Whereas unusual observations can sometimes be attributed to known events (such as a change in policy or a significant geopolitical event) that could produce large shocks or significantly alter the trajectory of the series, it is also common that observed outliers can not be attributed to any specific cause.

Additive outliers – or single impulses added to the observed series – are the simplest and most common form of outliers. Such impulses could occur in the latent innovation process (in which case they are called innovation outliers). Additive outliers can be modeled using deterministic additive outlier regressors of the form $AO_t^{(t_0)} = I(t = t_0)$, where $I(A)$ is the indicator function of an event A. For idiosyncratic shocks, the time (t_0) may be known, e.g. if there is an extreme value that can be attributed to a particular event whose timing is known. However, if large values are occurring frequently and cannot be attributed to particular events, then the modeler must sweep through the series to identify the times t_0 where additive outliers occur − this is expensive and requires multiple tests.

For example, analyzing the five series Food Services and Drinking Places, Grocery Stores, Durable Goods, Total Manufacturing and Building Materials and Garden Equipment with the deterministic outlier regressor framework of X-13ARIMA-SEATS,[2] the software flags several observations as potential outliers. The results are displayed in Figure 1; the identified extreme regressors are shown at the top of the panel of Table 1. The bottom rows show the results of a similar outlier analysis based on outlier regressors obtained when the outlier critical threshold has been reduced from the default value of 4 to 3.5. Even when care is taken to minimize subjective choices on part of an analyst, the results still depend on the type of regressors used, i.e., the change points in each of the regressors.

In order to minimize the disruptions caused by such subjective choices − as well as to mitigate the effects of multiple testing − we propose a stochastic modeling of the extremes. Such a model will be parsimonious, and will be able to adapt to the unknown locations of the outliers. Thus, unless the outliers can be attributed to specific events, and hence can be modeled using regressors of the form $AO_t^{(t_0)}$, we advocate using a stochastic outlier generating process that allows for large values, e.g. heavy-tailed processes. This approach unifies identification and estimation, thereby reducing uncertainty and biases due to sequential testing.

[2]X-13ARIMA-SEATS Reference Manual, Washington, DC: U.S. Census Bureau, U.S. Department of Commerce, available at https://www.census.gov/ts/x13as/docX13ASHTML.pdf.

Series	AO	LS	TC
Types and times of flagged outliers with default threshold 4.0			
Food & Drink	12/20, 01/22	06/20, 03/21, 06/21	03/20, 04/20 06/20, 11/20
Grocery Stores	04/20		03/20
Total Manufacturing	09/09, 03/20, 05/20	10/08, 11/08	11/08, 04/20
Durable Goods	03/20, 05/20	10/08	10/08, 04/20
Building Material	04/10	05/20, 03/21	
Types and times of flagged outliers with threshold 3.5			
Food & Drink	12/20, 01/22	06/20, 03/21, 06/21	03/20, 04/20 06/20, 11/20
Grocery Stores	04/20	07/20	03/20
Total Manufacturing	02/01, 12/01, 12/06, 09/08, 02/09, 04/13, 03/20, 05/20, 01/21	09/06, 10/08, 11/08, 03/11, 01/15, 12/16, 11/18	08/08, 11/08 04/20
Durable Goods	03/20, 05/20	10/08, 07/09,	10/08, 04/20
Building Material	04/10	05/20, 03/21	

Table 1: Detected outliers using default specification in the seas function of the 'seasonal' R package (Sax, 2018; Sax, C. and Eddelbuettel, D., 2018) for the X-13ARIMA-SEATS software; default outlier detection threshold of 4.0 (top panel) and custom outlier detection threshold of 3.5 (bottom panel).

4 Model, estimation and signal extraction

4.1 Specification of the components

The primary goal of the paper is to propose stochastic models for additive outliers (AO), level shifts (LS), and temporary change (TC) outliers. For the application of seasonal adjustment, the outlier regressors that are of interest are

$$AO_t^{t_0} = \begin{cases} 1 & \text{if } t = t_0, \\ 0 & \text{otherwise}, \end{cases} \quad LS_t^{t_0} = \begin{cases} 0 & \text{if } t < t_0, \\ 1 & \text{otherwise}, \end{cases} \quad TC_t^{t_0} = \begin{cases} 0 & \text{if } t < t_0, \\ \phi^{t-t_0} & \text{otherwise}. \end{cases} \tag{1}$$

Here $\phi \in (0,1)$ is a user-defined decay rate for the temporary change outlier. Note that the first difference of the level shift is the additive outlier, $(1-B)LS_t^{(t_0)} = AO_t^{(t_0)}$ (where B is the backshift operator) and $(1-\phi B)TC_t^{t_0} = AO_t^{(t_0)}$. Thus the level shift regressor can be obtained as an integrated additive outlier regressor and the temporary change can be obtained by applying the $(1-\phi B)^{-1}$ operator to an AO regressor. In analogy to using a heavy-tailed white noise process model for additive outliers, stochastic modeling of level shifts can be achieved by using integrated heavy-tailed white noise processes. This paradigm was adopted by Trimbur (2010), who modified the additive outliers in a local level model using a Student's t process.

Seasonal adjustment is often done under a structural time series model:

$$X_t = \mu_t + \gamma_t + S_t + \epsilon_t,$$

where $\{\mu_t\}$, $\{\gamma_t\}$, $\{S_t\}$, and $\{\epsilon_t\}$ are the trend, cycle, seasonal (with s seasons per year), and irregular components, respectively. The different components satisfy individual dynamic equations, given by $(1-B)\mu_t = \xi_t$, $(1-\phi B)\gamma_t = \zeta_t$, and $U(B)S_t = \eta_t$, where $\{\xi_t\}$, $\{\zeta_t\}$, $\{\eta_t\}$, and $\{\epsilon_t\}$ are independent Gaussian white noise processes, and $U(B) = (1 + B + \ldots + B^{s-1})$ is the seasonal aggregation operator.

To account for extreme observations in a particular dynamical component, the component is usually augmented with a dummy outlier regressor, e.g. LS for the trend, TC for the cycle, and AO for the irregular. Since each outlier regressor can be reduced to the additive outlier dummy through the differencing operator for the associated stochastic component, the latent (unobserved) dynamical components with extreme changes are thus, after differencing, modeled as the sum of a Gaussian white noise and additive outlier regressors. Instead of using a combination of white noise and AO regressors as the driving process for the extreme dynamics, we use a single heavy-tailed white noise process that would adapt to the shocks via large disturbances. In particular we use a scaled Student's t distribution with low degrees of freedom to model a component that is traditionally modeled as the sum of a Gaussian white noise with several AO effects.

Applying this formulation to the trend and cycle components, the full model becomes

$$X_t = \mu_t + \gamma_t + S_t + \epsilon_t; \quad \mu_t = (1-B)^{-1}\zeta_t; \quad \gamma_t = (1-\phi B)^{-1}\xi_t, \tag{2}$$

where $\{\xi_t\}$, $\{\zeta_t\}$, and $\{\epsilon_t\}$ are independent scaled Student's t white noise processes, but $\{\eta_t\}$ (the seasonal innovations) is still a Gaussian white noise process. This specific model structure can be expressed as a special case of a more general unobserved component model. Let

$$X_t = \boldsymbol{H}_t \boldsymbol{\eta} + \sum_{j=1}^{J} C_t^j + \epsilon_t, \tag{3}$$

where $\boldsymbol{H}_t = (h_{t,1}, \ldots, h_{t,m})$ are fixed regressors associated with known effects, such as trading day and moving holidays, $\eta = (\eta_1, \ldots, \eta_M)'$ are the regression coefficients for the fixed effects, ϵ_t is the irregular component, and $C_t^j, j = 1, \ldots, J$ are J latent components. It will be convenient to identify ϵ_t as another unobserved component, C_t^0. In the specific application of seasonal adjustment with the structural time series model (2), the model can be recognized as a special case of (3) with $J = 3$ and the unobserved components being the trend $C_t^1 = \mu_t$, the cycle $C_t^2 = \gamma_t$, the seasonal $C_t^3 = S_t$, and the irregular component $C_t^0 = \epsilon_t$. The $(J+1)$ unobserved components are assumed to satisfy ordinary difference equations

$$\Phi_j(B) C_t^j = \Theta_j(B) W_t^j, \; j = 0, \ldots, J.$$

The operators $\Phi_j(B)$ and $\Theta_j(B)$ are given by

$$\begin{aligned} \Phi_j(B) &= 1 - \phi_1^j B - \ldots - \phi_{d_j}^j B^{d_j}, \\ \Theta_j(B) &= 1 - \theta_1^j B - \ldots - \theta_{q_j}^j B^{q_j}, \end{aligned}$$

with polynomial order d_j and q_j, respectively. The degrees of differencing corresponding to the irregular component (i.e., for $j = 0$) are assumed to zero, so that $d_0 = q_0 = 0$

and $\Phi_0(B) = \Theta_0(B) = 1$. All the polynomials are assumed to have roots on or outside the unit circle. We will assume $d_j \geq q_j$ for identification; see Hotta (1989).

Given a length n sample of the time series, in order to describe the joint distribution of the components we need to know the initial observations $\tilde{\boldsymbol{C}}^j = [C_0^j, C_{-1}^j, \ldots, C_{-d_j+1}^j]'$, as well as $\tilde{\boldsymbol{W}}^j = [W_0^j, W_{-1}^j, \ldots, W_{-q_j+1}^j]'$. Conditional on the initial observations, the joint distributions for the latent components are described by the equations

$$\boldsymbol{\Phi}^j \boldsymbol{C}^j + \tilde{\boldsymbol{\Phi}}^j \tilde{\boldsymbol{C}}^j = \boldsymbol{\Theta}^j \boldsymbol{W}^j + \tilde{\boldsymbol{\Theta}}^j \tilde{\boldsymbol{W}}^j, \tag{4}$$

where the (u,v)th element of the $n \times n$ matrix $\boldsymbol{\Phi}^j$ is defined as $\Phi_{u,v}^j = -\phi_{u-v}^j$ and that of the $n \times d_j$ matrix $\tilde{\boldsymbol{\Phi}}^j$ is defined as $\tilde{\Phi}_{u,v}^j = -\phi_{d_j-v+u}^j$. Here it is understood that $\phi_r^j = 0$ if either $r > d_j$ or $r < 0$, and $-\phi_0^j = 1$. The entries of the $n \times n$ matrix $\boldsymbol{\Theta}^j$ and the $n \times q_j$ matrix $\tilde{\boldsymbol{\Theta}}^j$ are defined analogously as $\Theta_{u,v}^j = \theta_{u-v}^j$ and $\tilde{\Theta}_{u,v}^j = \theta_{q_j-v+u}^j$, where $\theta_r^j = 0$ if either $r > q_j$ or $r < 0$, and $\theta_0^j = 1$. Note that for $j = 0$, $\boldsymbol{\Phi}^0$ and $\boldsymbol{\Theta}^0$ are identity matrices and $\tilde{\boldsymbol{\Phi}}^0$ and $\tilde{\boldsymbol{\Theta}}^0$ are zero matrices.

The first J_1 (an integer selected by the modeler) latent processes, as well as the irregular process, will be assumed to be heavy-tailed, whereas processes of higher index are assumed to be Gaussian. Thus, we assume that each $\boldsymbol{W}^j$ is a vector of n consecutive values from a white noise sequence, such that for $j = 0, \ldots, J_1$ the entries in $\boldsymbol{W}^j$ are independently and identically distributed as a scaled Student's t distribution (with scale σ_j^2 and degrees of freedom ν_j), and such that for $j = J_1 + 1, \ldots, J$ the entries in $\boldsymbol{W}^j$ are distributed as Gaussian variables with variance σ_j^2. Specifically, $W_t^j \overset{iid}{\sim} \sigma_j t_{\nu_j}$ for $t = 1, \ldots, n$ and $j = 1, \ldots, J_1$, and $W_t^j \overset{iid}{\sim} N(0, \sigma_j^2)$ for $t = 1, \ldots, n$ and $j = J_1+1, \ldots, J$. Furthermore, the latent component vectors are all assumed to be independent of each other.

4.2 Gibbs sampling

A Gaussian unobserved component model can be cast into a state-space framework, and the computation of likelihood and estimators are done using standard Kalman filtering/smoothing techniques (Durbin and Koopman (2001)). While that is a general framework, specific features may be more directly incorporated in the model using alternative methods, and direct Bayesian computation steps can be more explicitly formulated for specific models.

In latent component models, Bayesian data augmentation methods are typically used for ease of computation. In latent component time series models, there is a long literature of such augmentation algorithms, using versions of the Kalman filter/smoother; see Carter and Kohn (1994), de Jong and Shepard (1995), McCausland et al. (2011), and Scott and Varian (2014), and the references therein. In the present context, if the model is written in a state-space format, the error variances for the aberrant components will be time-dependent since the mixing parameters will change at every time point. The simple conjugate update is not possible any longer in such cases, but each conditional variance parameter could be updated using univariate methods such as rejection sampling.

The structural models for the latent components in (4), along with the data model in (3), provide a natural framework for formulating a joint distribution for all the unknown variables. Once a prior distribution for the parameters in the structural models of the latent components is specified, the full joint distribution of the data and the unknown components can be obtained. To facilitate posterior computation we further use a hierarchical formulation, writing the Student's t distributions as a scale mixture of Gaussian distributions. This will allow one to use conjugacy in the full conditional specification.

Specifically, if $W_t \sim \sigma^2 t_\nu$ then the hierarchical formulation gives $W_t|\sigma^2, \alpha \sim N(0, \sigma^2\alpha)$ and $\alpha|\sigma^2 \sim IG(\nu/2, \nu/2)$, where IG is the *inverse gamma* density and the degrees of freedom ν is assumed to be known. Following this idea, we specify the full set of priors for the latent components and parameters for their distributions as follows: for $j = 0, \ldots, J_1$,

$$\begin{aligned} \boldsymbol{C}^j|\sigma_j^2, \boldsymbol{A}_j &\sim N(\boldsymbol{\theta}^j, \sigma_j^2 \boldsymbol{B}_j \boldsymbol{A}_j \boldsymbol{B}_j'), \\ \alpha_t^j|\sigma_j^2 &\stackrel{iid}{\sim} IG(\nu_j/2, \nu_j/2),\ t = 1, \ldots, n, \\ \sigma_j^2 &\sim IG(\nu_{j0}, \sigma_{j0}^2), \end{aligned} \tag{5}$$

and for $j = J_1 + 1, \ldots, J$,

$$\begin{aligned} \boldsymbol{C}^j|\sigma_j^2 &\sim N(\boldsymbol{\theta}^j, \sigma_j^2 \boldsymbol{B}_j \boldsymbol{B}_j'), \\ \sigma_j^2 &\sim IG(\nu_{j0}, \sigma_{j0}^2), \end{aligned} \tag{6}$$

where $\boldsymbol{A}_j = diag(\alpha_1^j, \ldots, \alpha_n^j)$, $\boldsymbol{\theta}^j = \tilde{\boldsymbol{\Theta}}^j \tilde{\boldsymbol{W}}^j - \tilde{\boldsymbol{\Phi}}^j \tilde{\boldsymbol{C}}^j$, $\boldsymbol{B}_j = (\boldsymbol{\Phi}^j)^{-1}\boldsymbol{\Theta}^j$, and the hyperparameters $\nu_1, \ldots, \nu_{J_1}, \nu_{10}, \ldots, \nu_{J0}, \sigma_{10}^2, \ldots, \sigma_{J0}^2$ are assumed to be pre-specified. The different components are assumed to be independently distributed a priori. In addition, the fixed effects are given standard conjugate normal priors, viz. $\boldsymbol{\eta}|\sigma_0^2 \sim N(\boldsymbol{\eta}_0, \sigma_0^2 \boldsymbol{A}_{\boldsymbol{\eta}})$, where the prior mean vector $\boldsymbol{\eta}_0$ and the prior variance matrix $\boldsymbol{A}_{\boldsymbol{\eta}}$ are pre-specified.

With the full conditional specification, it is natural to use Gibbs sampling for posterior sampling. For updating the latent component, normal-normal conjugacy provides the conditional posterior distribution for the component vector given the rest of the parameters. Specifically,

$$\begin{aligned} \boldsymbol{\eta} &\sim N(\hat{\boldsymbol{\eta}}, \sigma_0^2 \hat{\boldsymbol{A}}_\eta) \\ \sigma_j^2 &\sim IG((\nu_{j,0} + n)/2, (\nu_{j,0}\sigma_{j,0}^2 + \boldsymbol{\ell}_j'\boldsymbol{\ell}_j)/2) \\ \alpha_{j,t} &\sim IG((\nu_j + 1)/2, (\nu_j + \boldsymbol{\ell}_{j,t}^2/\sigma_j^2)/2) \\ \boldsymbol{C}^j &\sim N(\tilde{\boldsymbol{\theta}}^j, \boldsymbol{\Omega}_j^{-1}),\ j > 0, \end{aligned} \tag{7}$$

where $\hat{\boldsymbol{A}}_\eta = [\boldsymbol{H}'\boldsymbol{A}_0^{-1}\boldsymbol{H} + \boldsymbol{A}_\eta^{-1}]^{-1}$, $\hat{\boldsymbol{\eta}} = \hat{\boldsymbol{A}}_\eta[\boldsymbol{H}'\boldsymbol{A}_0^{-1}(\boldsymbol{X} - \sum_{k=1}^J \boldsymbol{C}^k) + \boldsymbol{A}_\eta^{-1}\boldsymbol{\eta}_0]$, $\boldsymbol{\ell}_j = \boldsymbol{A}_j^{-1/2}\boldsymbol{w}^j$, and $\boldsymbol{w}^j = (\boldsymbol{\Theta}^j)^{-1}[\boldsymbol{\Phi}^j\boldsymbol{C}^j + \tilde{\boldsymbol{\Phi}}^j\tilde{\boldsymbol{C}}^j - \tilde{\boldsymbol{\Theta}}^j\tilde{\boldsymbol{W}}^j]$ for $j = 1, \ldots, J$. Also, $\boldsymbol{w}^0 =$

$\boldsymbol{X}-\boldsymbol{H}\boldsymbol{\eta}-\sum_{k\geq 1}\boldsymbol{C}^k$, $\boldsymbol{\Omega}^j=(\sigma_0^2\boldsymbol{B}_0\boldsymbol{A}_0\boldsymbol{B}_0')^{-1}+(\sigma_j^2\boldsymbol{B}_j\boldsymbol{A}_j\boldsymbol{B}_j')^{-1}$, and the conditional mean of the jth component given the rest of the unknowns is $\tilde{\boldsymbol{\theta}}_j=\boldsymbol{\Omega}_j^{-1}[(\sigma_0^2\boldsymbol{B}_0\boldsymbol{A}_0\boldsymbol{B}_0')^{-1}(\boldsymbol{X}-\boldsymbol{H}\boldsymbol{\eta}-\sum_{k\neq j}\boldsymbol{C}^k+(\sigma_j^2\boldsymbol{B}_j\boldsymbol{A}_j\boldsymbol{B}_j')^{-1}\boldsymbol{\theta}_j]$. Here $\boldsymbol{A}_j=\boldsymbol{I}$ for $j>J_1$ and $\boldsymbol{H}=(\boldsymbol{H}_1',\ldots,\boldsymbol{H}_n')'$. The expressions in (7) follow directly from the model and the priors described in (4), (5), and (6), using the standard derivation of conjugate posterior under Normal-Inverse Gamma conjugacy; see Gelman et al. (2013).

The latent component block can be sampled together in a block-Gibbs step. However, for the application at hand, it is more convenient to successively sample the components of each latent component. It is understood that, at each Gibbs step, only the most recent value of each conditioning variable is used. For sampling the coordinates of a latent component $\boldsymbol{C}^j$ having conditional distribution $N(\tilde{\boldsymbol{\theta}}^j,\boldsymbol{\Omega}_j^{-1})$, the tth component $\boldsymbol{C}_t^j$ is sampled from $N(\tilde{\boldsymbol{\theta}}^j_{t\setminus t},\boldsymbol{\Omega}^{-1}_{j,t\setminus t})$, where the mean and the variance, $\tilde{\boldsymbol{\theta}}^j_{t\setminus t}$ and $\boldsymbol{\Omega}^{-1}_{j,t\setminus t}$, of $\boldsymbol{C}_t^j$ given the rest of the coordinates (the set denoted by $\setminus t$) are obtained using standard formulas for the conditional mean and variance of coordinates of a multivariate normal vector.

5 Numerical results and illustrations

In this section, we present results from analyzing real data time series using the proposed method. In addition, results from a limited simulation experiment are given in the supplement (Roy and McElroy, 2024). Since this is purely for illustration, we do not use any fixed regressors. Thus, throughout this section, $\boldsymbol{H}$ in (2) is assumed to be zero, and hence $\boldsymbol{\eta}$ is not estimated.

We fit model (2) to several macroeconomic time series, and study the flexibility of the proposed model with regard to fitting data with multiple extreme events. We fixed the degrees of freedom for the Student's t processes to be $\nu_0=\nu_1=\nu_2=2.1$. Thus, the variances (apart from the scales) for the Student's t processes – corresponding to the trend, the temporary change, and the irregular innovations – were 3, 3, and 3, respectively. The parameters of the Inverse Gamma priors for all the scale parameters were given defaults values of $\nu_{0,0}$, $\nu_{1,0}=\nu_{2,0}=\nu_{3,0}=3$, and the scales were $\sigma_{0,0}^2=\sigma_{1,0}^2=\sigma_{2,0}^2=\sigma_{3,0}^2=400$. When fitting in the log scale the scale parameters were chosen to be $\sigma_{0,0}^2=\sigma_{1,0}^2=\sigma_{2,0}^2=\sigma_{3,0}^2=0.5$. The decay parameter in the temporary change processes was fixed at $\phi=0.7$ following the recommendation in Tsay et al. (2000). The seasonal period is taken to be $s=12$, mimicking a monthly series. The automatic default values of the hyperparameters were obtained by tuning. However, we did a sensitivity analysis based on the inverse gamma parameters, and the results (presented in the supplement) show the models to be robust with respect to specification of the scale parameters. The degrees of freedom parameters for the aberrant components were set to 2.1 to have finite variance, while retaining the capacity to generate very large shocks.

Since the process is nonstationary and starts at an arbitrary time point, to get the initial values for the trend and the seasonal components we used the output from a

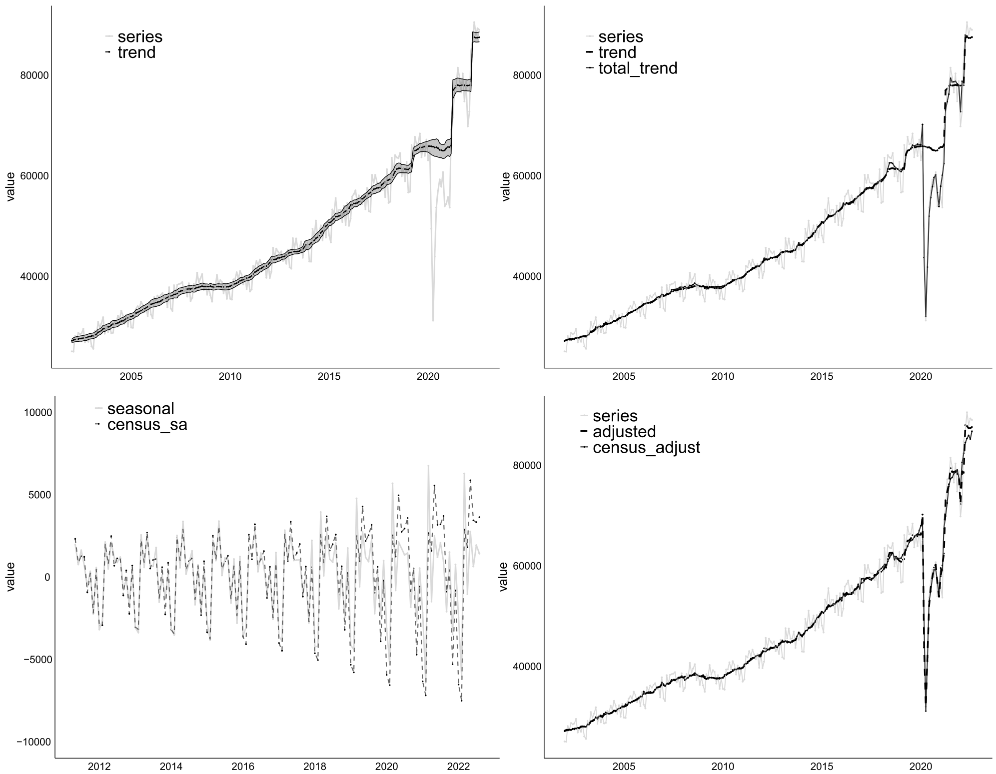

Figure 2: Food Services and Drinking Places: trend component with 90% confidence band (upper left panel), long-term component consisting of trend together with temporary changes (upper right panel), extracted seasonal component along with extraction using *seasonal* (bottom left panel), seasonally adjusted series (bottom right panel).

fitting of the data (with start date $s-1$, i.e. 11 months prior to the start date of the fitting period) using the *seasonal* package in R with the default specification. The initial value for the temporary change process was set to zero. These initial values were fixed throughout the modeling exercise. To initialize the chain, parameters were drawn randomly from their prior distribution and then given the initial parameter values and the initial values obtained from the *seasonal* package; the components were drawn randomly from their corresponding structural models.

The aggregate Food Services and Drinking Places series' dynamics went through a period of change during the Great Recession, where the overall trend flattened before returning to normalcy. However, the effect of the Covid-19 crisis is much starker in the early months of the pandemic (in early 2020). Figure 2 shows that much of the departure from normalcy at the end of the series is being attributed to a level shift in the trend, plus a large disturbance captured by the temporary change component. The seasonal part in the right middle panel of Figure 2 seems to be similar in pre- and post-Covid times. The behavior of the model fit during the Great Recession also shows

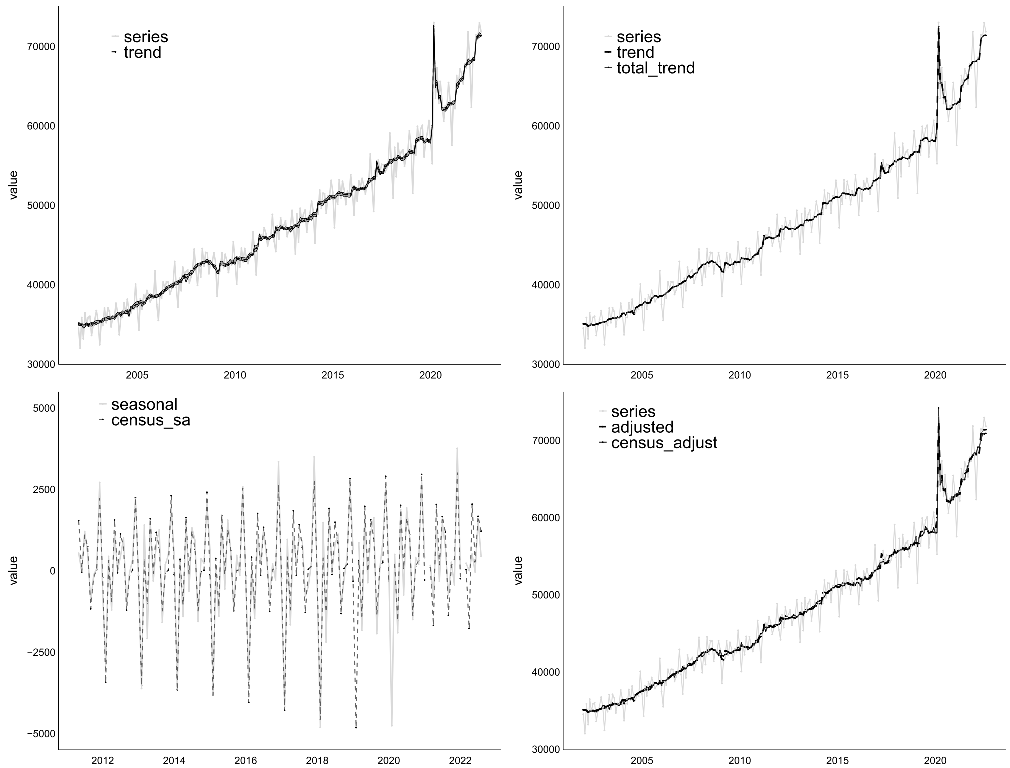

Figure 3: Grocery Stores: trend component with 90% confidence band (upper left panel), long-term component consisting of trend together with temporary changes (upper right panel), extracted seasonal component along with extraction using *seasonal* (bottom left panel), seasonally adjusted series (bottom right panel).

a change in the overall trend. However, the attribution to the temporary change is of much greater magnitude during the Covid-19 era. While there is a considerable shift in the trend during the pandemic, it is still difficult to discern whether (and when) the dynamics will stabilize. The uncertainty due to challenges that emerged during the pandemic continues to exert an influence upon the data; hence, a model that can rapidly adapt to these changes is preferable.

Grocery Stores, despite a partial lockdown, have fared reasonably well throughout the crisis period. Indeed, there was a substantial upward movement in February 2020, possibly due to hoarding; the series returned to previous patterns afterward. Figure 3 shows that the uptick in the single month of February 2020 is mostly attributed to the aberrant irregular component. The extracted seasonality in Figure 3 does not show any aberrant behavior, and can be used in the usual way to seasonally adjust the series.

We also examined the U.S. monthly total series for Total Manufacturing (Figure 4), Durable Goods (Figure 5), and Building Materials and Garden Equipment (Figure 6).

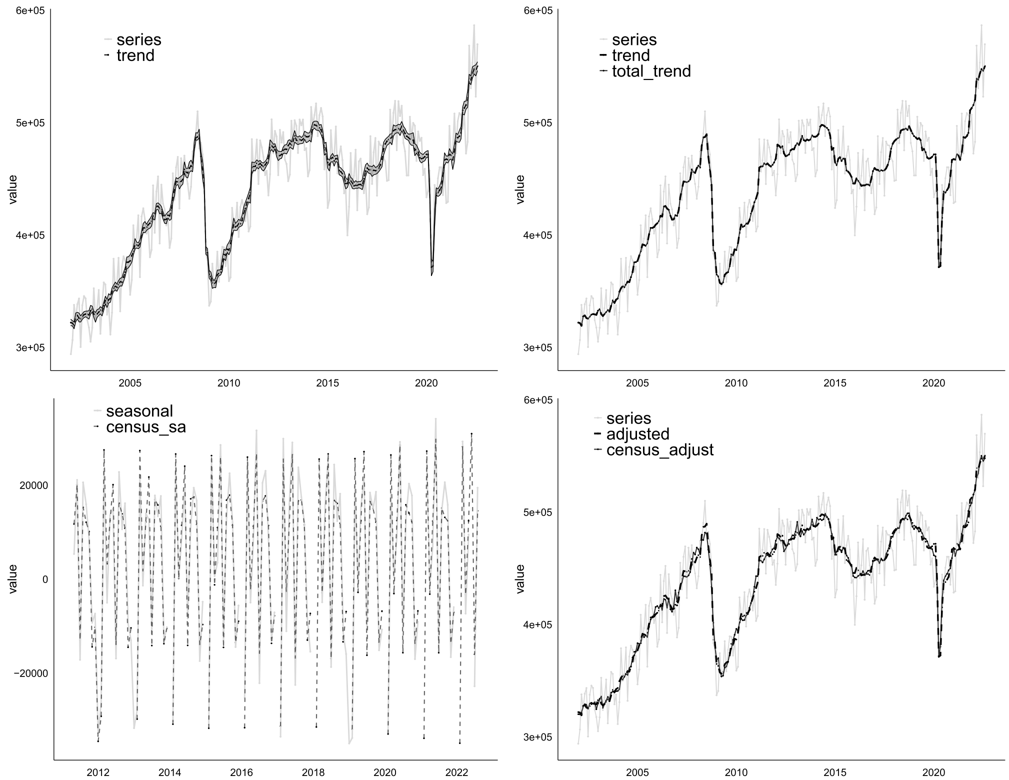

Figure 4: Total Manufacturing: trend component with 90% confidence band (upper left panel), long-term component consisting of trend together with temporary changes (upper right panel), extracted seasonal component along with extraction using *seasonal* (bottom left panel), seasonally adjusted series (bottom right panel).

Total Manufacturing suffered less impact than Food Services and Drinking Places, and the dip produced by the pandemic appears to be of a transient nature – especially in comparison to the Great Recession. The story is similar for Durable Goods, but Building Materials and Garden Equipment show two stochastic level shifts in the Covid-19 period, reflecting changes in those particular markets.

6 Estimation of the orthodox trend

6.1 Definition of an idealized trend

Bayesian structural time series and related models have been used to make causal inference on the effect of extreme events. There is a growing literature on estimation of counterfactuals to quantify the impact of extreme events; see Brodersen et al. (2014). The methodology used in Brodersen et al. (2014) and related work essentially relies on

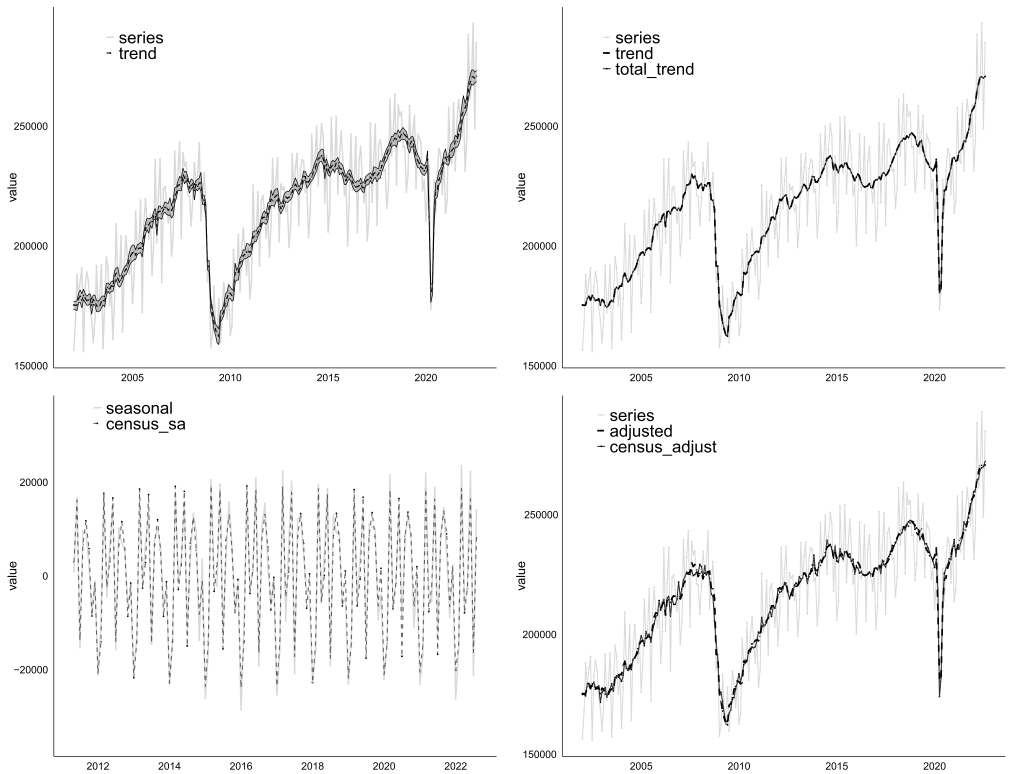

Figure 5: Durable Goods: trend component with 90% confidence band (upper left panel), long-term component consisting of trend together with temporary changes (upper right panel), extracted seasonal component along with extraction using *seasonal* (bottom left panel), seasonally adjusted series (bottom right panel).

forecasting the post-event time series using a posterior predictive distribution obtained from a BSTS fitted to the pre-event time series. The method regards the impact of an event as a singular quantity that changes the trajectory of the time series from that time point onward. Thus, a difference between pointwise forecasts and the observed value – after the onset of the extreme event – provides a counterfactual. A related body of work on causal inference for time series data is based on the DiD analysis using a parallel trend assumption; see Lechner (2010). The parallel trend assumption says that the trajectory of the time series with and without the extreme event should be parallel, beginning from the time of the event. A particular advantage of Brodersen et al. (2014) over the traditional DiD framework is that – being based on stochastic regressors and Bayesian posterior analysis, as opposed to static regressors as in the classical DiD analysis – uncertainty quantification in the estimated counterfactual quantities are more precise.

In the present work we use a fully stochastic framework to describe counterfactuals and quantify uncertainty. We assume a counterfactual series where the latent white

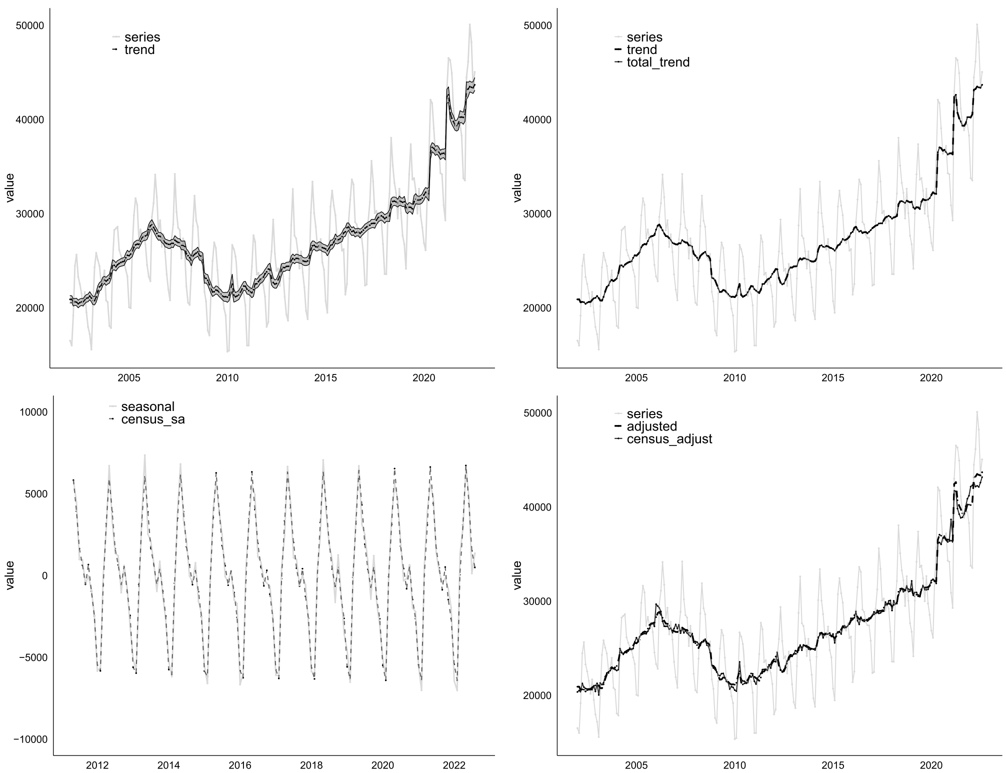

Figure 6: Building Materials and Garden Equipment: trend component with 90% confidence band (upper left panel), long-term component consisting of trend together with temporary changes (upper right panel), extracted seasonal component along with extraction using *seasonal* (bottom left panel), seasonally adjusted series (bottom right panel).

noise shocks are Gaussian instead of t-distributed. The Gaussian case signifies the behavior of the series during 'normal times' as opposed to one that is exposed to extreme shocks, where the Gaussian white noise are multiplied by the latent scales $\sqrt{\alpha_t}$ to have a student's t distribution. Since the Bayesian model estimates all latent quantities, including the latent shocks and the scales, one can also generate posterior samples of the normalized Gaussian white noise by rescaling using the estimated scales.

We focus our treatment on the latent trend process $\{\mu_t\}$, though the discussion could be extended to the latent cycle or seasonal as well if practitioners were interested in counterfactuals for cyclic or seasonal dynamics. Letting $\Delta = 1 - B$, recall from (2) that $\zeta_t = \Delta \mu_t = \sqrt{\alpha_t} W_t$, where $\{W_t\}$ is i.i.d. Gaussian. We shall define a counterfactual trend $\{m_t\}$ (using the Latin form of the Greek μ_t) to be the so-called orthodox portion of the stochastic process obtained by removing the presence of the inverse Gamma variable α_t; this corresponds to enforcing that the extreme-value distribution is concentrated at unity. Thus by definition $\Delta m_t = W_t$, which in turn equals $\Delta\mu_t/\sqrt{\alpha_t}$, and hence

$$m_t = \sum_{s=1}^{t} {\alpha_s}^{-1/2}\, \Delta\mu_s.$$

Here we have defined the orthodox trend for times $t \geq 1$, but the starting point could be taken to be more remotely in the past. The so-called idealized trend is defined to be the actual trend $\{\mu_t\}$ for times up to some onset t^* of aberrant behavior, and thereafter is given by the orthodox trend, viz.

$$m_t^{(t^*)} = \begin{cases} \mu_t, & \text{if } t \leq t^*, \\ \mu_{t^*} + \sum_{s=t^*+1}^{t} {\alpha_s}^{-1/2}\, \Delta\mu_s, & \text{if } t > t^*. \end{cases}$$

This idealized trend $\{m_t^{(t^*)}\}$ differs significantly from the real trend $\{\mu_t\}$ only at times subsequent to the crisis, and allows us via comparison to assess what would have happened if the crisis had not occurred; $m_t^{(t^*)}$ can be interpreted as the counterfactual trend of interest. Of course, $m_t^{(t^*)}$ is a latent stochastic process that is not observed, but for a given posterior sample we have draws for α_t and μ_t from the sampling iteration of the Gibbs loop described in (7), and we can therefore construct a posterior estimate $\widehat{m}_t^{(t^*)}$ by averaging over these draws. Below, we use $\widehat{m}_t^{(t^*)}$ to evaluate the impact of specific events such as temporary changes, level shifts, and additive outliers in simulated series.

6.2 Estimation of the impact of an extreme event

Based on the estimated idealized trends from the post-onset event time t^*, one is able to quantify the direct impact of the event in terms of changes to the stochastic processes. The difference between the estimated total trend, i.e. the sum of the trend and the cycle, provides a noise-free version of the observed series after the seasonal effect has been removed. This quantity can be compared with the estimated idealized total trend, i.e., the sum of the latent orthodox part of the trend and the cycle that are bereft of any effect of the extreme shocks.

The idealized trend can deviate substantially from the actual series over time, since the accumulation of extreme shocks is likely to shift the actual series from the orthodox part of the series. Hence, to make a meaningful evaluation of the impact of an event the time horizon is limited to a few time points after the event. A large disruptive phenomenon such as the Covid-19 pandemic constitutes a sequence of shocks, mimicking the aftershocks following a large earthquake, and hence accounting for the entire sequence of shocks becomes important while evaluating the impact. For example, the initial lockdown following the beginning of the pandemic had several stages, and the total impact was an aggregate effect of the partial and full closures in different parts of the country. An advantage of the proposed framework is that it naturally allows for a sequence of extreme events, and hence any estimate based on the framework includes the effect of the entire sequence of aberrant shocks during the time period.

For a fixed time horizon H, the impact of an event occurring at time t^* can be measured via the sum of pointwise differences between the estimated total trend and its orthodox part, and can be estimated for each posterior draw. The overall impact can be summarized as the average of the quantity over all posterior samples.

The time horizon H could be chosen to be a fixed number or a quantity estimated from the sample, e.g., the first time after t^* such that the total trend and the idealized total trend intersect (if this happens); the horizon also should have a upper bound, such as the endpoint of the observation period. While most scientific questions require the difference between the two trends to be small in order to declare that the event no longer has an impact, we advance the criterion that the two trends must cross in order to claim that the series has returned to the pre-event level. Note that when H is a sample-dependent quantity (such as the first time point following the event time where the trends intersect), its value will be different for each posterior sample.

6.3 Numerical illustration

Using the orthodox total trend described above, one could evaluate the impact of extreme events such as the Great Recession and the onset of the Covid-19 pandemic. Figure 7 shows the estimated orthodox trends for each series for post-recession and post-covid periods.

For Food Services and Drinking Places and Grocery Stores, the idealized trends after the Great Recession show that the recovery outpaced the orthodox trend. In the former series the Covid-19 effect is mostly a large temporary change with a sharp recovery after the lockdown, and a slight shift in level is estimated in the trend after the recovery. The idealized trend estimate is flat, maintaining the level from the pre-Covid-19 era.

For the Grocery Stores series, the estimated trend shows a level shift at the onset of the pandemic, but the idealized trend is essentially flat at the pre-Covid-19 level.

For the Total Manufacturing series the estimated trend does not show any impact from the pandemic, as the impact of the extreme event is essentially attributed to a temporary change (blue line showing the trend plus the temporary change), and therefore the counterfactual trend is also at the level of the estimated trend. The Great Recession had a significant effect on the trend, about two full years being required for the trend to catch up with the pre-recession trend and make up for the loss.

For the Durable Goods series a majority of Covid-19's impact is attributed to a temporary change at the onset of the pandemic. There is a small level change in the trend that is absent in the idealized trend. During the post-recession there was also a significant drop in the level, and the counterfactual estimate indicates a more permanent change in the level and the trend than it showed for the Total Manufacturing series.

For the Building Materials and Garden Equipment series there is a small level shift at the beginning of the pandemic followed by a temporary change. The recession effect was essentially a level shift, and toward the end of the observation period the idealized trend and the estimated trend seem to be merging together.

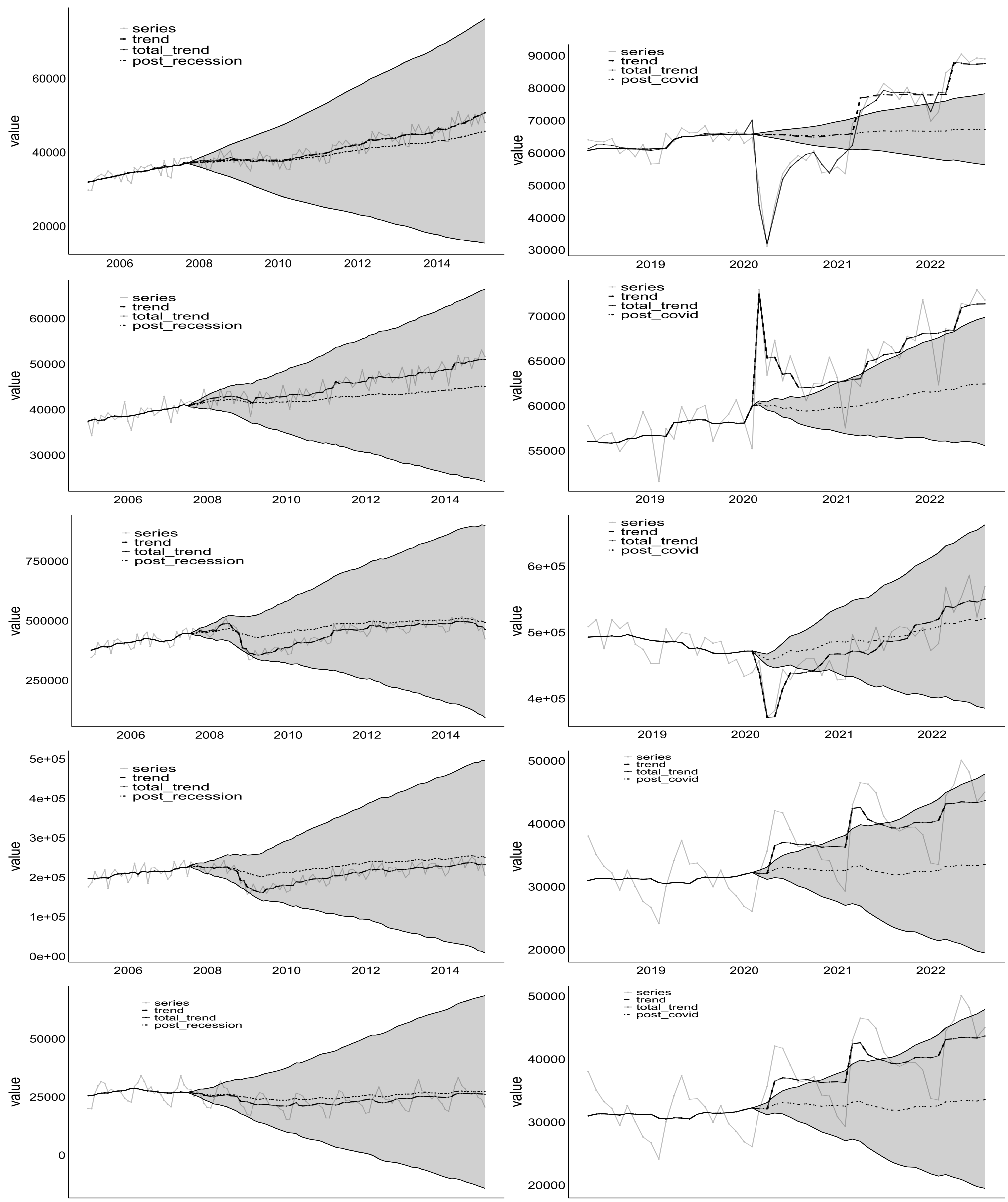

Figure 7: Estimates of the orthodox trend immediately after the Great Recession and the start of the Covid-19 pandemic: Food Services and Drinking Places (first row); Grocery Stores (second row); Total Manufacturing (third row); Durable Goods (fourth row), Building Materials and Garden Equipment (fifth row). The left panel and the right panels are for post-recession trend and post-covid trend, respectively. In each panel the estimated trend and the estimated counterfactual trend are shown against the series. For post-Covid-19 plots, the estimated total trends are also shown to illustrate the attribution of the Covid-19 effect to different dynamics.

6.4 Estimate of impact of the COVID-19 pandemic

We use the Food Services and Drinking Places series to illustrate the effectiveness of the proposed impact measure. Due to the lockdowns the restaurant business (or in-person food service) in the USA suffered great losses in the first year of the pandemic, particularly in the three months following its onset. We used three different values of the time horizon H to measure the impact of the onset of the pandemic. The values were $H_3 = 3$ for evaluating the impact during March, April, and May 2020, $H_{12} = 12$ to evaluate the impact in the first year, and $H = H^*$ equal to the first time point of intersection of the total trend and the idealized total trend following the onset to evaluate the impact prior to returning to normalcy. The posterior mean estimate of H^* was approximately 14 months, indicating about a year-long struggle for the industry to catch up with the pre-pandemic level. The estimated total impacts (in billions of dollars) were 58, 192, and 211 for the three different time horizons H.

In contrast, the USDA Economic Research Service Food Expenditure Survey (FES) reports a decline of about 19.6% in food-away-from-home sales during 2020 compared to 2019 (Zeballos and Dong (2021)). Based on the FES values (in nominal billions of dollars) the total sales in full and limited service restaurants and drinking places are about 724, 622, and 829 in the years 2019, 2020, and 2021, respectively. The expenditures in the Food Service during 2020 based on the pre-pandemic trend (an exponentially weighted moving average forecast based on the annual sales from years 2000-2019) is about 762, indicating a loss of about 140 billion, almost entirely coming from the pandemic months in that year, spanning March through December of 2020. Our estimate of the loss in sales for the retail Food Services and Drinking services from March 2020 through February 2021 (12 months) is about 192 billion. The National Restaurant Association, in its January 15, 2021 article[3] estimates a total of 240 billion dollars of loss to the industry during 2020. However, this figure includes the sales shortfall at eating and drinking places along with a reduction in spending at food service operations in sectors such as lodging, arts/entertainment/recreation, education, healthcare, and retail. Overall, the estimate $\widehat{\Delta}_{t^*}^{H} = 197$ captures the amount of loss in the average level (smooth trend) of the Retail Food and Drinking Services series, over a 12 month period from March 2020.

7 Discussion

The X-13ARIMA-SEATS software provides a choice of different models and filters, and the default procedure performs an automatic model selection to choose the best possible model. Such flexibility is needed to have adequate seasonal adjustment for a wide variety of time series. The structural model (2), while considerably rich in terms of the models it can approximate, can be generalized to obtain a broader spectrum of model options, and hence more flexibility in adjusting seasonal features while accounting for extreme observations.

[3] Available at https://restaurant.org/education-and-resources/resource-library/state-of-the-restaurant-industry-report-measures-virus-impact-on-business/.

For example, the seasonal component can be given a moving average structure in order to make the seasonal model more versatile. This is analogous to the popular airline model, where an MA(1) structure is given to the seasonally differenced series. Triantafyllopoulosa and Nason (2007) provides a conjugate sampling procedure for the MA(1) parameter, and this can be easily incorporated in the overall Gibbs sampling scheme (7). In (2), for the specific model used in the examples with $J = 3$ and $C_t^3 = S_t$, instead of assuming that $U(B)S_t = W_t$ (a white noise), we would assume that the seasonal component satisfies $U(B)C_t^3 = Y_t$, where $Y_t = W_t^3 + \theta_1^3 W_{t-1}^3$ is an MA(1) process. Thus, the entries of moving average coefficient matrix $\mathbf{\Theta}^3$ will be a function of the scalar parameter θ_1^3. For updating θ_1^3 with the rest of the parameters given, the conditional posterior will be that of an MA(1) parameter computed from data of length n, from an invertible MA(1) process with the initial value W_0^3 pre-specified. In the Gibbs sampling the main change will only be in the updating steps for the seasonal component S_t, the variance of the seasonal white noise σ_3^2, with the additional updating step for the seasonal moving average parameter θ.

In summary, we have proposed a fully Bayesian framework for accommodating aberrant behavior in the components of a structural time series, particularly in the process of doing seasonal adjustment. The BSTS paradigm has several distinct advantages over the outlier dummy regressor framework currently used to do seasonal adjustment at many statistical agencies. The Bayesian method provides a full measure of uncertainty for all the unknown quantities, and also produces counterfactual trend estimates that allows quantification of the effect of extreme events. One challenge in the heavy-tailed framework is that signal extraction is done based on non-linear filters, which are not explicitly defined. Thus, exercises such as designing a modified filter for addressing features of the seasonal process, or such as forecasting that involves filter coefficients, become extremely challenging, if not impossible.

The current practice of seasonal adjustment has been perfected over many decades, with expert experiences helping to fine-tune the different parameters. In the current proposal we have not done any extensive investigation on how sensitive the results are to the choice of hyper-parameters, or any other pre-specified quantity. If the proposal is to provide a viable supplement to the seasonal adjustment methodology, such investigations need to be carried out.

The world having emerged from the Covid-19 pandemic, the seasonal adjustment community will be well-served to develop a flexible framework that quickly adapts to the changing dynamics of the thousands of time series requiring seasonal adjustment. The present work adds tools to the seasonal adjustment community's arsenal that could be used in the post-Covid-19 era, particularly when the occurrence of change in dynamics can not be attributed to a specific month or quarter, but rather needs to be modeled as a stochastic event.

Acknowledgments

This report is released to inform interested parties of research and to encourage discussion. The views expressed on statistical issues are those of the authors and not those of the U.S. Census Bureau. All time series analyzed in this article are from public data sources.

The authors would like to thank the editors and the anonymous referees for their valuable comments and suggestions that improved the quality of the paper

Supplementary Material

Supplement to 'Modeling Extreme Events in Time Series and Their Impact on Seasonal Adjustment in the Post-Covid-19 Era' (DOI: 10.1214/24-BA1424SUPP; .pdf). Additional simulation results are presented in the supplement along with a discussion about how to utilize seasonal outlier dummies.

References

Angrist, J. D. and Krueger, A. B. (1999). "Empirical strategies in labor economics." *Handbook of Labor Economics*, 3: 1277–1366. 2

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton Univ. Press. 2

Box, G. E. P., and Tiao, G. C. (1975). "Intervention analysis with applications to environmental and economic problems." *JASA*, 70: 70–79. MR0365957. 4

Brodersen, K. H., G., F., Koehler, J., Remy, N., and Scott, S. L. (2014). "Inferring causal impact using Bayesian structural time-series models." *AoAS*, 9: 247–274. MR3341115. doi: https://doi.org/10.1214/14-AOAS788. 3, 15, 16

Calvet, L. E., Czellar, V., and Ronchetti, E. (2015). "Robust Filtering." *JASA*, 110: 1591–1606. MR3449057. doi: https://doi.org/10.1080/01621459.2014.983520. 2

Carter, C. K. and Kohn, R. (1994). "On Gibbs sampling for state-space models." *Biometrika*, 81: 541–553. MR1311096. doi: https://doi.org/10.1093/biomet/81.3.541. 10

Crevits, R. and Croux, C. (2018). "Robust Estimation of Linear State Space Models." *Communications in Statistics: Simulation and Computation*, 48: 1694–1705. MR3945948. doi: https://doi.org/10.1080/03610918.2017.1422752. 2

de Jong, P. and Shepard, N. (1995). "The simulation is smoother for time series models." *Biometrika*, 82: 339–350. MR1354233. doi: https://doi.org/10.1093/biomet/82.2.339. 10

Dey, T., Tyagi, P., and Sabath, M. e. a. (2021). "Counterfactual time series analysis of short-term change in air pollution following the COVID-19 state of emergency in the U.S." *Scientific Reports*, 11: 23517. 3

Durbin, J. and Koopman, S. J. (2001). *Time Series Analysis by State Space Methods.*. Oxford University Press. MR1856951. 10

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

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., and Rubin, D. (2013). *Bayesian Data Analysis*. Chapman and Hall/CRC, 3 edition. MR3235677. 12

Harvey, A. and Luati, A. (2014). "Filtering With Heavy Tails." *JASA*, 109: 1112–1122. MR3265684. doi: https://doi.org/10.1080/01621459.2014.887011. 2

Holan, S. and McElroy, T. (2012). "Bayesian seasonal adjustment of long memory time series." *Econometric Time Series: Modeling and Seasonality*, 403–429. MR3076025. doi: https://doi.org/10.1201/b11823-24. 6

Hotta, L. K. (1989). "Identification of unobserved components models." *Journal of Time Series Analysis*, 10: 259–270. MR1028941. doi: https://doi.org/10.1111/j.1467-9892.1989.tb00027.x. 10

Lechner, M. (2010). "The estimation of causal effects by difference-in-difference methods." *Foundations and Trends in Econometrics*, 4: 165–224. 2, 16

Maiz, C. S., Molanes-Lopez, E. M., Miguez, J., and Djuric, P. M. (2012). "A Particle Filtering Scheme for Processing Time Series Corrupted by Outliers." *IEEE Transactions on Signal Processing,*, 60: 4611–4627. MR2960549. doi: https://doi.org/10.1109/TSP.2012.2200480. 2

Maravall, A. and Pérez, D. (2012). "Applying and interpreting model-based seasonal adjustment - the Euro-area industrial production series." In Bell, W. R., Holan, S. H., and McElroy, T. S. (eds.), *Econ. Time Series: Modeling and Seasonality*, 281–313. New York: Chapman Hall. MR3076020. doi: https://doi.org/10.1201/b11823-17. 6

McCausland, W. J., Miller, S., and Pelletier, D. (2011). "Simulation smoothing for state-space models: a computational efficiency analysis." *Computational Statistics and Data Analysis*, 55: 199–212. MR2736547. doi: https://doi.org/10.1016/j.csda.2010.07.009. 10

McCulloch, R. E. and Tsay, R. S. (1994). "Bayesian Analysis of Autoregressive Time Series via the Gibbs Sampler." *Journal of Time Series Analysis*, 15: 235–250. MR1263893. doi: https://doi.org/10.1111/j.1467-9892.1994.tb00188.x. 6

McElroy, T. (2016). "On the measurement and treatment of extremes in time series." *Extremes*, 19: 467–490. MR3535963. doi: https://doi.org/10.1007/s10687-016-0254-4. 6

McElroy, T. and Penny, R. (2019). "Maximum entropy extreme-value seasonal adjustment." *Australian & New Zealand Journal of Statistics*, 61: 152–174. MR3979793. doi: https://doi.org/10.1111/anzs.12262. 2

Ravishanker, N. and Ray, B. (1997). "Bayesian analysis of vector ARMA models using Gibbs sampling." *Journal of Forecasting*, 16: 177–194. 6

Roy, A. and McElroy T. S. (2024). "Supplementary Material for "Modeling Extreme

Events in Time Series and Their Impact on Seasonal Adjustment in the Post-Covid-19 Era”.” doi: https://doi.org/10.1214/24-BA1424SUPP. 12

Scott, S. L. and Varian, H. R. (2014). “Predicting the present with Bayesian structural time series.” *International Journal of Mathematical Modelling and Numerical Optimisation*, 5: 4–23. 6, 10

Takyia, P. O. and Bentum-Enninb, I. (2021). “The impact of COVID-19 on stock market performance in Africa: a Bayesian structural time series approach.” *Journal of Economics and Business*, 115: 1059–1068. 3

Triantafyllopoulosa, K. and Nason, G. (2007). “A Bayesian analysis of moving average processes with time-varying parameters.” *Computational Statistics and Data Analysis*, 52: 1025 – 1046. MR2418543. doi: https://doi.org/10.1016/j.csda.2007.04.001. 22

Trimbur, T. (2010). “Stochastic level shifts and outliers and the dynamics of oil price movements.” *International Journal of Forecasting*, 26: 162–179. 2, 6, 8

Tsay, R., Peña, D., and Pankratz, A. E. (2000). “Outliers in multivariate time series.” *Biometrika*, 87: 789–804. MR1813975. doi: https://doi.org/10.1093/biomet/87.4.789. 6, 12

Tsay, R. S. (1986). “Time series model specification in the presence of outliers.” *JASA*, 81: 132–141. 2, 3

Tsay, R. S. (1988). “Outliers, level shifts, and variance change in time series.” *J. Forecasting*, 7: 1–20. 3

Zeballos, E. and Dong, X. (2021). “The effect of COVID-19 on food sales.” *Appl Econ Perspect Policy*. Oct 13: 10.1002/aepp.13201. 21