Constructing seasonally adjusted data with time-varying confidence intervals[*]

Siem Jan Koopman[†]

Department of Econometrics, Free University Amsterdam

and

Philip Hans Franses[‡]

Econometric Institute, Erasmus University Rotterdam

January 24, 2001

Abstract

Seasonal adjustment methods transform observed time series data into estimated data, where these estimated data are constructed such that they show no or almost no seasonal variation. An advantage of model-based methods is that these can provide confidence intervals around the seasonally adjusted data. One particularly useful time series model for seasonal adjustment is the basic structural time series [BSM] model. The usual premise of the BSM is that the variance of each of the components is constant. In this paper we address the possibility that the variance of the trend component in a macro-economic time series in some way depends on the business cycle. One reason for doing so is that one can expect that there is more uncertainty in recession periods. We extend the BSM by allowing for a business-cycle dependent variance in the level equation. Next we show how this affects the confidence intervals of seasonally adjusted data. We apply our extended BSM to monthly US unemployment and we show that the estimated confidence intervals for seasonally adjusted unemployment change with past changes in the oil price.

---

1 Introduction

A key reason for the application of seasonal adjustment methods is that intra-year seasonality can blur the analysis of the trend and the business cycle in macro-economic time series variables. Seasonal adjustment methods transform observed time series data into estimated data, where these estimated data are constructed such that they show no or almost no seasonal variation. The two main methods for seasonal adjustment are based on (versions of) the Census X-12 program, see Findley et al. (1998), and on various parametric models, see for example Maravall (1995). One of the advantages of the model-based methods is that these methods not only give seasonally adjusted data but also can provide their associated confidence intervals. Indeed, to enable a proper interpretation of seasonally adjusted data, it seems important that the estimated adjusted data come along with a measure of their uncertainty.

One particularly useful time series model for seasonal adjustment is the basic structural time series model [BSM], see Harvey (1989) among others. This model assumes that an economic time series can be decomposed into trend, cycle, seasonal and irregular components. Given certain assumptions on the model structure, these unobserved components can be estimated for the available data. Once the seasonal component is determined, a seasonally adjusted time series can be constructed in a fairly trivial way. A nice feature of the BSM is also that confidence bounds for such a seasonally adjusted series can be easily constructed. These bounds are a function of all the estimated variance parameters in the structural model, see Burridge and Wallis (1985).

The usual premise of the BSM is that the variance of each of the components is constant. Burridge and Wallis (1990) consider the extension to the case where the seasonal component can have some seasonal heteroscedasticity, and, given the recent empirical findings in Jaditz (2000), this can sometimes be a useful extension. In this paper we take a slightly different perspective by addressing the possibility that the variance of the seasonal component in a macroeconomic time series somehow depends on the business cycle. The empirical results in Canova and Ghysels (1994)

1

and Franses (1995), among others, suggest that seasonal variation in several macro-economic variables differs with the business cycle, and hence it seems important to take account of such variation. The main purpose for doing so is that one can expect that there is more uncertainty in recession periods, and somehow this uncertainty should be assigned to each of the components. Ooms and Franses (1997) document that in times of recession the seasonal component in US unemployment series becomes larger in the first quarter and smaller in other quarters, but the origin of such variation is not exactly clear. On the other hand, while calculating confidence bounds for seasonally adjusted data, one would explicitly want to allow for possibly more uncertainty during recession periods. In this paper we therefore extend the BSM by allowing for a business-cycle dependent variance in the level equation. Next we show how this effects the confidence intervals of seasonally adjusted data.

The outline of our paper is as follows. In §2, we discuss some aspects of the basic structural model, with a specific focus on the construction of confidence intervals for seasonally adjusted data. In §3, we extend the BSM by allowing for a business-cycle dependent variance, and we show how this changes the confidence intervals for adjusted data. In §4, we apply our extended BSM to monthly US unemployment. As an indicator for business cycle variation we take changes in the oil price[1]. We find that our model fits the data well, and in fact significantly improves upon the BSM. Additionally, we show that the estimated confidence intervals for seasonally adjusted unemployment change with past changes in the oil price. In §5, we conclude with some remarks.

2   Basic Structural Time Series Model

In this section we discuss various aspects of the BSM, including the computation of confidence bounds for seasonally adjusted data.

---

[1]Needless to say that lots of other useful explanatory variables (indicators) can be considered. Practitioners may want to make a choice between available leading indicators, business survey data, or all kinds of key economic variables. In this paper we take the oil price mainly for illustrative purposes and for convenience.

2.1 Basic components

The structural time series model is based on the basic principle that a time series consists of interpretable unobserved components such as a trend, seasonal, cycle and irregular. The basic structural time series model (BSM) is given by

$$y_t = \mu_t + \gamma_t + \varepsilon_t, \qquad t = 1, \ldots, n, \tag{1}$$

where y_t represents the actual time series with n observations. The time series components, that is, level μ_t, seasonal γ_t and irregular ε_t are unobserved and modeled by stochastic processes. A simple model for the trend is the random walk as given by

$$\mu_{t+1} = \mu_t + \xi_t, \qquad \xi_t \sim \mathrm{N}(0, \sigma_\xi^2), \tag{2}$$

with $\mu_1 \sim \mathrm{N}(0, \kappa)$ where one can take κ as arbitrarily large. This initial condition for μ_1 indicates that no information is available about μ_1. By introducing a slope term ν_t which is also generated by a random walk, we obtain the model

$$\begin{array}{rcll} \mu_{t+1} & = & \mu_t + \nu_t + \xi_t, & \xi_t \sim \mathrm{N}(0, \sigma_\xi^2), \\ \nu_{t+1} & = & \nu_t + \zeta_t, & \zeta_t \sim \mathrm{N}(0, \sigma_\zeta^2), \end{array} \tag{3}$$

with $\nu_1 \sim \mathrm{N}(0, \kappa)$. If $\xi_t = \zeta_t = 0$ then $\nu_{t+1} = \nu_t = \nu$, say, and $\mu_{t+1} = \mu_t + \nu$. In that case, the trend is exactly linear and (3) reduces to the deterministic linear trend model. The specification (3) with $\sigma_\xi^2 > 0$ and $\sigma_\zeta^2 > 0$ allows the trend level and slope to vary over time.

Various specifications for the seasonal component γ_t exist. In this paper we adopt the trigonometric form which for the case of a monthly time series is given by

$$\gamma_t = \gamma_{1t} + \ldots + \gamma_{6t}, \tag{4}$$

where

$$\begin{pmatrix} \gamma_{j,t+1} \\ \gamma_{j,t+1}^* \end{pmatrix} = \begin{bmatrix} \cos \lambda_j & \sin \lambda_j \\ -\sin \lambda_j & \cos \lambda_j \end{bmatrix} \begin{pmatrix} \gamma_{jt} \\ \gamma_{jt}^* \end{pmatrix} + \begin{pmatrix} \omega_{jt} \\ \omega_{jt}^* \end{pmatrix},$$

with frequency $\lambda_j = \pi j / 6$, for $j = 1, \ldots, 6$. The disturbances are serially and mutually uncorrelated and they are normally distributed with mean zero and variance matrix

$$\mathrm{Var} \begin{pmatrix} \omega_{jt} \\ \omega_{jt}^* \end{pmatrix} = \sigma_{\omega j}^2 I_2,$$

3

where I_k is the $k \times k$ identity matrix. Each initial seasonal value γ_{j1} and γ_{j1}^* for $j = 1, \ldots, 6$ is modeled by the non-informative prior distribution $\mathrm{N}(0, \kappa)$. Finally, for γ_{6t} we have $\lambda_6 = \pi$ and $\sin \lambda_6 = 0$ so that γ_{6t}^* does not have an influence on γ_{6t} and it can therefore be excluded from the model. The ω_{jt}'s and ω_{jt}^*'s can be restricted to have a common variance σ_ω^2 but we allow them to have different variances. Finally, we take the irregular as a normally random variable with mean zero and variance σ_ε^2. The irregular and all other disturbances are mutually uncorrelated, both contemporaneously and between different time periods. More details about the basic model and its dynamic properties are given by, for example, Harvey (1989) and Proietti (2000).

The unknown parameters in this basic model are the variances σ_ε^2, σ_ξ^2, σ_ζ^2 and $\sigma_{\omega1}^2, \ldots, \sigma_{\omega6}^2$. These can be estimated by maximum likelihood and since these are restricted to be non-negative, we estimate the logs of the variances rather than the variances themselves. These transformed parameters will be collected in the parameter vector ψ.

2.2 Parameter estimation

The Kalman filter equations produce the one-step-ahead prediction errors $v_t = y_t - \mathrm{E}(y_t | y_1, \ldots, y_{t-1}) = y_t - \mathrm{E}(\mu_t + \gamma_t | y_1, \ldots, y_{t-1})$ together with its variance $f_t = \mathrm{Var}(y_t | y_1, \ldots, y_{t-1}) = \mathrm{Var}(v_t)$ for a given time series y_t ($t = 1, \ldots, n$) and for a given model represented in state space. An introduction to these matters is given by Harvey (1993) and Durbin and Koopman (2001), among others. The classical result of Schweppe (1965) implies that the likelihood can be presented in terms of prediction errors using the so-called "prediction error decomposition". The log-likelihood function is then given by

$$\log L(y; \psi) = -\{ n \log 2\pi + \sum_{t=d+1}^{n} (\log f_t + \frac{v_t^2}{f_t}) \}/2,$$

where d is the number of initial components with the arbitrarily large variance κ in the model. For a given model, the errors v_t depend on the given time series y_t and the parameter vector ψ and the variances f_t depend only on ψ. Numerically

maximizing $\log L(y; \psi)$ with respect to ψ leads to maximum likelihood estimates of the unknown variances of the model.

2.3 Model specification diagnostics

Once the parameter vector has been estimated, it is common practice to check whether the model under investigation is correctly specified. For this purpose, we can compute some basic diagnostic statistics for testing the underlying assumptions of normally distributed, serially uncorrelated and homoskedastic one-step-ahead prediction residuals, which we standardize to obtain $u_t = v_t / \sqrt{f_t}$ ($t = d + 1, \dots, n$).

Furthermore, we can compute estimators of disturbances using all observations and we refer to them as smoothed estimators. For example, the smoothed level residual is defined as $\hat{\xi}_t = \mathrm{E}(\xi_t | y_1, \dots, y_n)$ with variance $\mathrm{Var}(\xi_t | y_1, \dots, y_n)$. Such full-sample estimators can be computed using so-called smoothing algorithms. Time series plots and basic diagnostics for testing against normality are useful in identifying outlying residuals. In the case of ε_t, an outlying residual may indicate an outlying observation while in the case of ξ_t it may indicate a structural break in the time series. Thus smoothing disturbances are used to identify and to distinguish between outliers, structural breaks, slope breaks and other irregularities in the time series process. Such diagnostic procedures have been discussed by Harvey and Koopman (1992) and de Jong and Penzer (1998).

2.4 Seasonal adjustment with confidence intervals

A satisfactorily estimated model can be used as a tool to further investigate the properties of a time series. For example, we may estimate the level and seasonal components using all observations. The estimated trend reflects the underlying long-term movement of the time series. The seasonal component informs us about the seasonal variation within a year (or some other interval) which may change slowly over time.

Economic theories usually apply to stable relationships in the long-term and hence seasonal variation can be viewed as part of the noise in time series. This explains why empirical work in economics is often based on seasonally adjusted data.

5

In our view, the optimal strategy would be to analyze seasonally unadjusted data and to model the seasonal variation simultaneously with other parts (components) of the model. In any case, when seasonal adjustment is required, it should be based on a model. When the structural time series model is used, seasonal adjustment is carried out by simply subtracting the estimated seasonal component $\hat{\gamma}_t = \mathrm{E}(\gamma_t | y_1, \ldots, y_n)$ from the time series y_t ($t = 1, \ldots, n$). Since an analysis based on the Kalman filter and the associated smoothing algorithm also provides variances of the estimated components such as $\mathrm{Var}(\gamma_t | y_1, \ldots, y_n)$, we can construct confidence intervals for the estimated seasonal component but also for the seasonally adjusted time series. The latter will be the main focus of the paper.

3   The BSM with heteroscedasticity

In this section we discuss various extensions of the BSM, which allow for time-varying confidence intervals around the estimated components in the BSM.

3.1   Seasonal heteroscedasticity

Various forms of heteroscedasticity may exist in time series data and they should be modeled explicitly to obtain efficient estimates. A particular form of seasonal heteroscedasticity appears when different variances of disturbances apply to different seasonal periods. For example, observations of December can perhaps be more difficult to predict than observations of other months due to the Christmas celebration. Therefore the variance associated with the measurement errors of December should be larger compared to the errors of other months. Examples of this type of heteroscedasticity are given by Harvey, Koopman and Riani (1997) and Proietti (1998). Similar forms of heteroscedasticity can be applied to other disturbances in a structural time series model.

3.2   Trigonometric seasonal heteroscedasticity

Another form of seasonal heteroscedasticity is associated with the trigonometric specification of the seasonal component in the structural time series model. For the

trigonometric seasonal component (4) we usually assume that the disturbances ω_{jt} and ω_{jt}^* are homogeneous for $j = 1, \ldots, 6$ such that we restrict the variances of these disturbances to be equal to each other, see Harvey (1989, §2.3.4). However, it is regularly found for monthly time series that the seasonal cycle at a higher frequency, say γ_{1t}, varies more over time compared to the seasonal cycle at a lower frequency, say γ_{6t}. This requires that we should allow for different variances for different frequencies. The implied generality is already introduced for the seasonal component (4) of the structural time series model. The variances for ω_{jt} and ω_{jt}^* remain restricted to be equal since they both contribute equally to the variation of γ_{jt}, for $j = 1, \ldots, 6$.

3.3 Trend heteroscedasticity

Economic conditions vary over time and many economic variables display a cyclical behaviour which is usually referred to as the business cycle. When the economy is in a transition from a recession to an expansion, for example, the local dynamic properties of the trend may be different compared to the trend when stable economic conditions apply. Therefore, time series properties of economic data are not necessarily homogeneous between different time periods. For example, the local properties of a time series of house construction figures will be different depending on the state of the economy.

An alternative strategy for allowing for this type of heteroscedasticity is to identify, say, two different conditions of the economy such as decline (recession) and growth. Parameters affected by the different conditions can switch between two possible values. When the state of the economy in the business cycle cannot be observed, Hamilton (1989) proposed a Markov-switching model in which the switching depends on a latent unobserved variable. The Markov-switching technique has been later introduced to unobserved components models by Kim and Nelson (1999) and Luginbuhl and de Vos (1999). In this paper we explicitly include explanatory variables to indicate a recession.

Trend heteroscedasticity can be incorporated in the structural time series model (1) by introducing time-varying variances for the level disturbances. In particular,

we can specify the variance of the disturbance associated with the level component as

$$\sigma_{\xi,t}^2 = \exp(\lambda_0 + \lambda_1 z_t), \tag{5}$$

where z_t is an exogenous variable which may provide some information concerning the business cycle. This specification is easily generalized by including other exogenous variables and lagged values of such variables.

3.4 Representation and estimation

The heteroscedastic model can be presented in a time-varying state space model. The stochastic components are placed in the state vector α_t to obtain the 13×1 vector

$$\alpha_t = \begin{pmatrix} \mu_t \\ \nu_t \\ \gamma_{1t} \\ \gamma_{1t}^* \\ \vdots \\ \gamma_{6t} \end{pmatrix}.$$

The basic model in state space is given by

$$y_t = Z\alpha_t + \varepsilon_t, \qquad \alpha_{t+1} = T\alpha_t + \eta_t, \tag{6}$$

where row vector Z selects the appropriate elements for y_t and is given by

$$Z = (1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1).$$

The transition matrix T is the 13×13 identity matrix except for the elements

$$\begin{array}{rclcl} T(1,2) &=& 1, \\ T(1+2j, 1+2j) &=& T(2+2j, 2+2j) &=& \cos \lambda_j, \\ -T(2+2j, 1+2j) &=& T(1+2j, 2+2j) &=& \sin \lambda_j, \\ T(13, 13) &=& -1, \end{array}$$

for $j = 1, \ldots, 5$. The disturbance ε_t is normally distributed with mean zero and variance σ_ε^2. The vector η_t consists of disturbances associated with the elements of the state vector α_t and is normally distributed with mean vector zero and diagonal variance matrix Ω with elements

$$(\sigma_\xi^2, \sigma_\zeta^2, \sigma_{\omega 1}^2, \sigma_{\omega 1}^2, \sigma_{\omega 2}^2, \sigma_{\omega 2}^2, \sigma_{\omega 3}^2, \sigma_{\omega 3}^2, \sigma_{\omega 4}^2, \sigma_{\omega 4}^2, \sigma_{\omega 5}^2, \sigma_{\omega 5}^2, \sigma_{\omega 6}^2).$$

All disturbances are mutually and serially uncorrelated. The initial state vector is assumed to be a draw from a diffuse distribution, that is,

$$\alpha_1 \sim \mathrm{N}(0, \kappa I_{13})$$

with κ arbitrarily large.

The state space form of the basic model with trend heteroscedasticity is the same except that the first diagonal element of Ω, corresponding to the level residual ξ_t, is time-varying and given by (5). The variance matrix for the state disturbance vector η_t is therefore time-varying and denoted by Ω_t for $t = 1, \ldots, n$. The parameter vector ψ consists of logged variances and of the coefficients λ_0, λ_1 of the logged variance $\sigma_{\xi,t}^2$.

When the model in state space form is time-varying, the Kalman filter can still be used to compute one-step-ahead prediction errors together with their variances. The likelihood function of the model can therefore be computed via the Kalman filter. Seasonally adjusted data with confidence intervals can be obtained by applying a smoothing algorithm which allows for time-varying state space matrices.

3.5   Explanatory variables and interventions

In our empirical analysis below we will introduce explanatory and intervention variables in the models. Explanatory variables and intervention effects are easily allowed for in the basic structural model. Suppose that we have k regressors $x_{1t}, \ldots, x_{kt}$ with regression coefficients $\beta_1, \ldots, \beta_k$ which are constant over time and that we also wish to measure the change in level due to an intervention at time τ. We define this intervention variable w_t as follows:

$$
\begin{aligned}
w_t &= 0, & t < \tau, \\
&= 1, & t \geq \tau.
\end{aligned}
$$

Adding the above variables to the model (1) gives

$$y_t = \mu_t + \gamma_t + \sum_{j=1}^{k} \beta_j x_{jt} + \beta^* w_t + \varepsilon_t, \qquad t = 1, \ldots, n. \tag{7}$$

We see that β^* measures the change in the level of the series at a known time τ due to an intervention at time τ. The resulting model can readily be put into state space

9

form. For example, if $\gamma_t = \beta^* = 0$, $k = 1$ and if μ_t is determined by a random walk model, we can take

$$
\begin{aligned}
\alpha_t &= \begin{pmatrix} \mu_t & \beta_{1t} \end{pmatrix}', & Z_t &= \begin{pmatrix} 1 & x_{1t} \end{pmatrix}, \\
T &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, & \eta_t &= \begin{pmatrix} \xi_t \\ 0 \end{pmatrix},
\end{aligned}
$$

in (6). Here, although we have attached a suffix t to β_1 it is made to satisfy $\beta_{1,t+1} = \beta_{1t}$ so it is constant. We note here that row vector Z has become time-varying and therefore requires a subscript t. Another example of an intervention variable is the outlier intervention variable defined by

$$
\begin{aligned}
w_t &= 0, & t < \tau, \quad t > \tau, \\
&= 1, & t = \tau.
\end{aligned}
$$

In the next section we consider the application of (7) with heteroscedasticity.

4 Seasonal adjustment of unemployment figures

In this section we illustrate that taking into account explanatory variables and heteroscedasticity matters substantially for the width of confidence intervals around seasonally adjusted unemployment data.

4.1 Data and software

We investigate the seasonal adjustment of monthly US total unemployment (UN-EMPL) for the period between January 1960 and July 1997, that is a total of 451 observations. The data had been obtained from the Bureau of Labor Statistics (BLS) from which we also have obtained monthly oil prices (in logs) p_t.

The time-invariant BSM of §2, including its extensions with explanatory and intervention variables, can be analyzed using the program STAMP 6.0 of Koopman, Harvey, Doornik and Shephard (2000) . Models with time-varying heteroscedastic features such as the ones discussed in §3 cannot be analyzed using STAMP. For this purpose we have used the object-oriented matrix language Ox 2.2 of Doornik (1998) together with the state space functions in SsfPack 2.2 of Koopman, Shephard and Doornik (1999) [2].

---

[2]Details can be obtained from http://www.ssfpack.com. Data and programs can be obtained
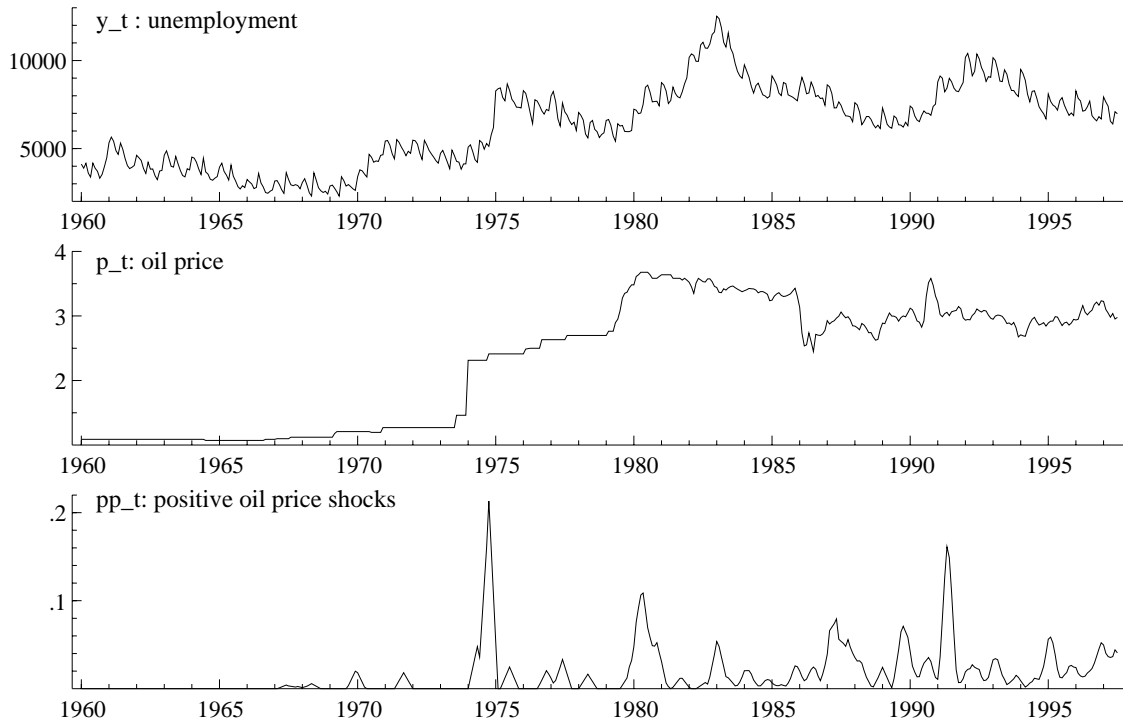
Figure 1: Unemployment y_t (measured by thousands of unemployed individiuals); Oil price p_t (in logs of US dollars); Positive oil price shocks pp_t (positive price difference of oil prices in logs).

4.2 Model-based seasonal adjustment

4.2.1 Mean specification

In the case of a structural time series model without explanatory variables, the mean equation consists of level and seasonal components. During the process of estimation and diagnostic checking, we may detect irregular observations such as outliers and structural breaks; see also the discussion in §2.3. It is standard practice in time series analysis to allow for such observations by introducing intervention variables into the model. The coefficients of the interventions will be part of the state vector and estimation is done by means of the Kalman filter.

4.2.2 Variance specification

The default variance specification is based on homogeneity, that is, constant variances for disturbances associated with the level, slope, seasonal and irregular com-

---

from `http://www.econ.vu.nl/koopman/seasadj/`.

11

ponents. This implies that the default model is (1) with the restriction that all seasonal variances are equal, that is,

$$\sigma_{\omega 1}^2 = \ldots = \sigma_{\omega 6}^2 = \sigma_\omega^2.$$

The inclusion of different variances for different seasonal frequencies in the trigonometric specification of the seasonal component will be referred to as seasonal heteroscedasticity.

The specification of trend heteroscedasticity depends on the choice of the exogenous variable z_t in (5). Since we want to investigate whether the underlying trend uncertainty of unemployment is higher during a recession and since, especially in the 1970s and 1980s, oil price is believed to be an important leading variable for recessions, we use a constructed variable based on the oil price to model the trend heteroscedasticity. Generally speaking, positive oil price shocks lead to price increases and a higher burden for manufacturing and therefore output is bound to decrease after some burning-in period. A simple weighted sum, where weights are distributed by a triangular shape, of positive changes in oil price can be an effective measure. For example, we have found that the indicator

$$pp_t = \sum_{j=-3}^{3} \frac{4 - |j|}{16} \max(0, \Delta p_{t-j}),$$

is appropriate for this purpose. The oil price is a leading indicator and therefore we may want to allow pp_t to enter the variance specification of the trend with some lag P. The constructed pp_t series as defined is given in Figure 1.

4.2.3 Model specifications

We consider the following model specifications in our empirical analysis of the unemployment data:

- (a) basic structural model;

- (b) basic model with seasonal heteroscedasticity;

- (c) basic model with restricted seasonal heteroscedasticity;

12

- (d) basic model with trend heteroscedasticity;

- (e) basic model with both seasonal (c) and trend (d) heteroscedasticity;

4.3 Empirical results

We start with the BSM. We find that it is required to include a level intervention for January 1975 and an outlier intervention for January 1986 in this model. The January 1975 observation corresponds with an unexpected increase of about 1 million unemployed. Furthermore, in January 1986 we have found an irregularity in the data which might have been caused by a redefinition of unemployed workers. An outlier intervention was appropriate to allow for this. These effects for the basic model are estimated as 1122 $(194; 5.785)$ for the January 1975 break and -529 $(136; -3.890)$ for the January 1986 outlier with standard error and t-value given in parentheses. These interventions are also included in all other models considered in this paper and give similar estimation results.

Table 1: Estimation output for BSM specification (1)

stand.dev	estimate	q-ratio
σ_ε	15.6	0.0911
σ_ξ	171.	1.0000
σ_ζ	26.4	0.1542
σ_ω	4.07	0.0238
Diagnostics		
N	2.28	
H	1.30	
DW	1.97	
Q(20)	27.76	
R^2	0.20	

Estimates of standard deviations are reported together with q-ratios which are defined as the standard deviation divided by the largest standard deviation. Diagnostics include N, the χ_2^2 normality test statistic, H, the Goldfeld-Quandt $F(150,150)$ test for heteroscedasticity, DW, the Durbin-Watson test for serial correlation, Q(20), the χ_{17}^2 Box-Ljung portmanteau statistic, and R^2, a measure of fit against a random walk plus drift model with fixed seasonal dummies.

Model parameter estimates and some diagnostic test statistics applied to the standardized one-step-ahead prediction errors u_t for the basic model are reported in

Table 1. The estimated components are presented in Figure 2. We may conclude that the basic model is capable of extracting the main features of the model. However, the model does not seem to have captured all dynamics satisfactorily since the Box-Ljung $Q(20)$ statistic is relatively large.
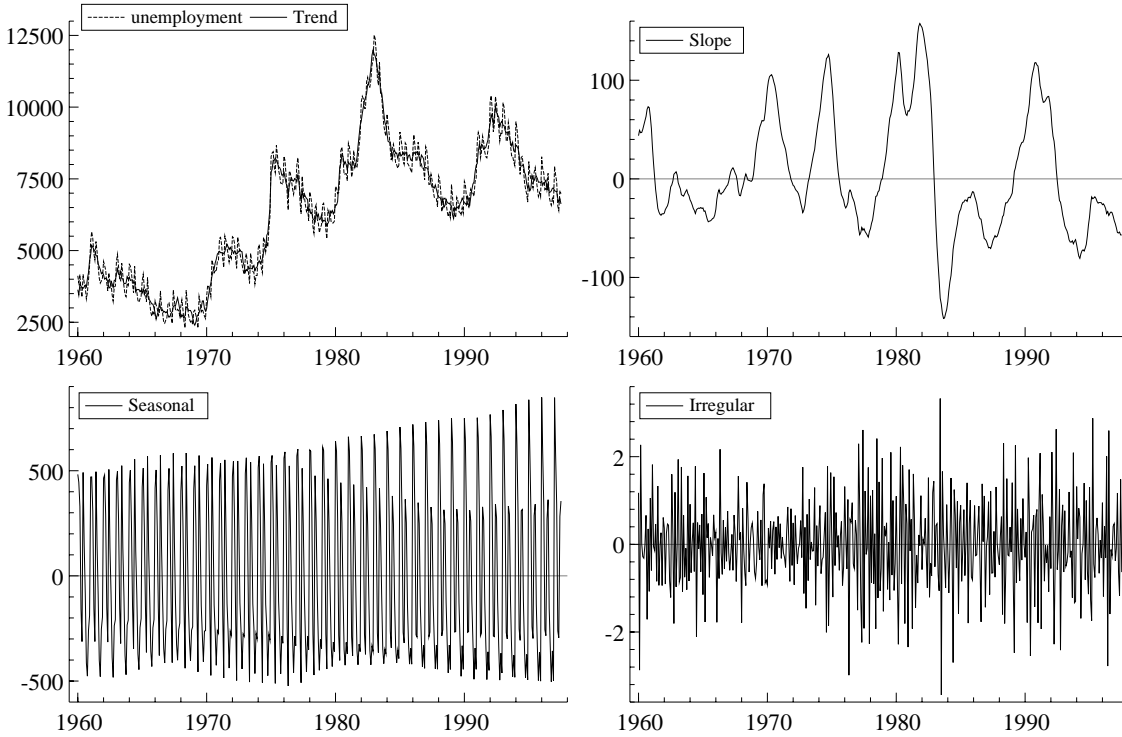

Figure 2: Estimated components: Trend, Slope, Seasonal and Irregular

The first extension of the basic model concerns seasonal heteroscedasticity in terms of the trigonometric specification. Instead of having one seasonal variance, that is model specification (a), we consider six different variances (an increase of five parameters) and this model is labelled as (b). The resulting estimates of the variances are reported in Table 2. The variances associated with the frequencies 2 and 4 are sufficiently close, as well as the ones associated with frequencies 3, 5 and 6, to restrict these variances accordingly. This model is labelled as (c) in Table 2 and it produces a loglikelihood value which was sufficiently larger than its value for the basic model (a) which is reflected by the smallest Akaike information criterion for model (c). It is remarkable that the Box-Ljung statistic $Q(20)$ is much lower for the preferred model (c) compared to (a). This emphasizes that accurate modeling

14

of seasonal dynamics is important in time series modeling.

Table 2: Parameter estimates for models with seasonal heteroscedasticity

| | Model specifications | | |
	(a)	(b)	(c)
σ_ε	15.6	50.0	45.4
σ_ξ	171.	149.	152.
σ_ζ	26.4	30.7	30.0
$\sigma_{\omega 1}$	4.07	11.6	11.3
$\sigma_{\omega 2}$	4.07	6.31	5.67
$\sigma_{\omega 3}$	4.07	3.65	3.27
$\sigma_{\omega 4}$	4.07	5.23	5.67
$\sigma_{\omega 5}$	4.07	2.78	3.27
$\sigma_{\omega 6}$	4.07	3.38	3.27
log.lik	-2989.33	-2986.79	-2986.99
nr.pars	4	9	6
nr.state	15	15	15
AIC	13.34	13.36	13.33
Q(20)	27.76	16.97	17.42

Estimates of standard deviations are reported for model specifications (a), (b) and (c). Further the loglikelihood value (log.lik) is reported of the estimated model together with the number of estimated parameters (nr.pars) and the dimension of the state vector (nr.state). Finally, AIC is the Akaike information criterion and Q(20) is the χ^2_{17} Box-Ljung portmanteau statistic.

We now consider the trend heteroscedastic extension of the basic model. The time-varying variance for the trend component is specified in (5) and we take z_t as the lagged positive oil price (in logs), that is

$$\sigma^2_{\xi t} = \exp(\lambda_0 + \lambda_1 pp_{t-P}),$$

for some positive value P. Some experimentation suggested that the optimal lag was equal to $P = 9$. The resulting model is labelled as (d) and the corresponding estimation results are reported in the first column of Table 3. Note that the estimated value for λ_0 is $\log \hat{\sigma}^2_\xi = \log 134. = 4.86$. The standard error of the estimated λ_1 is 1.56 so that we conclude that trend heteroscedasticity significantly improves the fit of our model. This conclusion is confirmed by the increase of the loglikelihood value, that is, when comparing the loglikelihood values of models (a) and (d).

15

The model that allows for trend and seasonal heteroscedasticity is labelled as
(e) and the estimation results for this model are also reported in Table 3. This
model (with $P = 9$) leads to the highest loglikelihood value and the smallest AIC
and it is therefore our preferred model. Since in empirical time series analysis it is
not often that we find significant heteroscedastic effects in variances of unobserved
components, we strongly believe that it may well be essential to include time-varying
variances in the specification of trend components for time series which are subject
to business-cycle effects such as unemployment figures.

Table 3: Parameter estimates for models trend heteroscedasticity

| | Model specifications | |
	(d)	(e)
σ_ε	12.5	40.2
σ_ξ	134.	121.
σ_ζ	28.1	29.0
$\sigma_{\omega 1}$	4.47	13.8
$\sigma_{\omega 2,4}$	4.47	5.64
$\sigma_{\omega 3,5,6}$	4.47	3.52
λ_1	4.55	4.55
log.lik	-2984.75	-2980.31
nr.pars	5	7
nr.state	15	15
AIC	13.32	13.31
Q(20)	13.22	14.04

Estimates of standard deviations are reported for model specifications (d) and (e). Further the
loglikelihood value (log.lik) is reported of the estimated model together with the number of esti-
mated parameters (nr.pars) and the dimension of the state vector (nr.state). Finally, AIC is the
Akaike information criterion and Q(20) is the χ^2_{17} Box-Ljung portmanteau statistic.

4.4   Seasonal adjustment of unemployment figures

The seasonally adjusted figures for unemployment based on our estimated are com-
puted by means of Kalman filtering and smoothing and they are constructed as

$$\hat{y}_t^{SA} = \hat{\mu}_t + \hat{\varepsilon}_t, \qquad t = 1, \ldots, n;$$

16

see also §2.4. These figures are estimates and subject to error for which standard deviations can be computed. For the basic model, the standard error of seasonally adjusted data is constant for all time periods in the middle of the sample and it starts to get larger when we get closer to the end of the sample (or the beginning) since the number of observations surrounding a particular time point decreases. The introduction of seasonal heteroscedasticity will result in different standard errors for different seasonal periods but the differences are usually small. The inclusion of trend heteroscedasticity leads to different standard errors for different time periods. For our model (e) we find that in times of recession the standard errors are higher than in times of no recession. This reflects well on the fact that in times of recession, more uncertainty exists within the economy and it is therefore harder to identify the estimated trend during these periods.

In our model the seasonally adjusted data is the sum of the estimated trend and irregular. The differences in width of the confidence intervals are indicated by twice the standard errors and they are reported in Table 4 for models (a), (c) and (e). We present the results for the typical years of 1980 (recession) and 1996 (no recession). It is concluded that the differences in standard errors among the three models are pronounced in 1980 while in 1996 the confidence intervals are more or less equal. The percentage increase of the width of the confidence interval due to the recession period is, on average, about 34% for model (e) compared to model (a).

5   Concluding remarks

In this paper we have constructed a structural time series model, which allows the confidence intervals around the seasonally adjusted data to depend on business cycle conditions. We discussed representation, estimation, and computational issues. We illustrated the model for a US unemployment series, and we found substantial evidence in favor of our model. Next, we showed that the estimated confidence interval around seasonally adjusted unemployment can be much wider in a recession period than in an expansion period. As such, we believe that our model can contribute to a better understanding of seasonally adjusted data.

Table 4: Standard errors (two times) for model-based seasonally adjusted figures

period	2SE model (a)	2SE model (c)	2SE model (e)
1980.1	160.12	159.14	214.14
1980.2	160.12	158.32	212.63
1980.3	160.12	161.75	216.83
1980.4	160.12	160.93	216.02
1980.5	160.12	157.38	211.92
1980.6	160.12	158.62	212.48
1980.7	160.12	162.29	217.01
1980.8	160.12	162.08	216.85
1980.9	160.12	157.57	212.14
1980.10	160.12	159.59	213.70
1980.11	160.12	160.93	216.44
1980.12	160.12	161.21	216.29
1996.1	160.12	159.14	158.63
1996.2	160.12	158.32	157.97
1996.3	160.12	161.75	161.29
1996.4	160.12	160.93	160.10
1996.5	160.12	157.38	157.14
1996.6	160.12	158.62	157.98
1996.7	160.12	162.29	161.49
1996.8	160.12	162.08	161.28
1996.9	160.12	157.57	156.34
1996.10	160.12	159.59	158.91
1996.11	160.12	160.93	160.17
1996.12	160.12	161.21	160.94

Admittedly, we illustrated our model for only one unemployment series, while using only one variable that can indicate business cycle fluctuations. Naturally, we are aware of the fact that other macroeconomic variables could have been used for both purposes. In order to see whether our model is useful in general, it should be considered in other situations, and we plan to do so in the future. Another further topic would be to see whether our model can lead to policy decisions, other than those based on the standard model. Finally, it is of interest to see if business cycle conditions have common effects on the confidence intervals of several seasonally adjusted series

18

References

Burridge, P. and K.F. Wallis (1985), Calculating the variance of seasonally adjusted time series, *Journal of the American Statistical Association*, 80, 541–552.

Burridge, P. and K.F. Wallis (1990), Seasonal adjustment and Kalman filtering: Extension to periodic variances, *Journal of Forecasting*, 9, 109–118.

Canova, F. and E. Ghysels (1994), Changes in seasonal patterns: Are they cyclical?, *Journal of Economic Dynamics and Control*, 18, 1143–1171.

de Jong, P. and J. Penzer (1998), Diagnosing shocks in time series, *Journal of the American Statistical Association*, 93, 796–806.

Doornik, J.A. (1998), *Ox: Object oriented matrix programming 2.0*, London: Timberlake Consultants Press.

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

Fiorentini, G. and A. Maravall (1996), Unobserved components in ARCH models: An application to seasonal adjustment, *Journal of Forecasting*, 15, 175–201.

Findley, D.F., B.C. Monsell, W.R. Bell, M.C. Otto and B.-C. Chen (1998), New capabilities and methods of the X-12 ARIMA seasonal-adjustment program (with discussion), *Journal of Business and Economic Statistics*, 16, 127-177.

Franses, P.H. (1995), Quarterly US unemployment: Cycles, seasons and asymmetries, *Empirical Economics*, 20, 717-725.

Hamilton, J.D. (1989), A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica, 57,* 357–384.

Harvey, A.C. (1989), *Forecasting, structural time series models and the Kalman filter,* Cambridge: Cambridge University Press.

Harvey, A.C. (1993), *Time Series Models, 2nd edition,* Hemel Hempstead: Harvester Wheatsheaf.

Harvey, A.C. and S.J. Koopman (1992), Diagnostic checking of unobserved components time series models, *Journal of Business and Economic Statistics,* 10, 377–389.

Harvey, A.C., S.J. Koopman and M. Riani (1997), The modelling and seasonal adjustment of weekly observations, *Journal of Business and Economic Statistics,* 15, 354–368.

Jaditz, T. (2000), Seasonality in variance is common in macro time series, *Journal of Business,* 73, 245–254.

Kim, C.J. and C.R. Nelson (1999), *State space models with regime switching,* Cambridge MA: MIT Press.

Koopman, S.J., A.C. Harvey, J.A. Doornik and N. Shephard (2000), *STAMP 6.0 Structural Time Series Analyser, Modeller and Predictor,* London: Timberlake Consultants Press. `http://www.stamp-software.com`

Koopman, S.J., N. Shephard and J. Doornik (1999), Statistical algorithms for models in state space using SsfPack 2.2, *Econometrics Journal,* 2, 113–166.

Luginbuhl R.E. and A.F. de Vos (1999), *Bayesian analysis of an unobserved component time series model of GDP with Markov switching and time varying growths*, Journal of Business and Economic Statistics, 17, 456–465 .

Maravall, A. (1995), Unobserved components in economic time series, in H. Pesaran, P. Schmidt and M. Wickens (eds.), *Handbook of Applied Econometrics (Volume 1)* Oxford: Basil Blackwell.

Ooms, M. and P.H. Franses (1997), On periodic correlation between estimated seasonal and nonseasonal components in German and US unemployment, *Journal of Business and Economic Statistics*, 15, 470–481.

Proietti, T. (1998), Seasonal heteroscedasticity and trends. *Journal of Forecasting*, 17, 1–17.

Proietti, T. (2000), Forecasting with structural time series models, Working paper, University of Padua, Italy.

Schweppe F. (1965), Evaluation of likelihood functions for Gaussian signals, *IEEE Transactions on Information Theory*, 11, 61–70.