2001

Seasonality, Nonstationarity and the Structural Forecasting of the Index of Industrial Production

Eugene Kouassi

Walter C. Labys
wlabys@wvu.edu

François B. Aka

Follow this and additional works at: https://researchrepository.wvu.edu/rri_pubs

Part of the [Regional Economics Commons](#)

Digital Commons Citation

Kouassi, Eugene; Labys, Walter C.; and Aka, François B., "Seasonality, Nonstationarity and the Structural Forecasting of the Index of Industrial Production" (2001). *Regional Research Institute Publications and Working Papers*. 153.
https://researchrepository.wvu.edu/rri_pubs/153

This Working Paper is brought to you for free and open access by the Regional Research Institute at The Research Repository @ WVU. It has been accepted for inclusion in Regional Research Institute Publications and Working Papers by an authorized administrator of The Research Repository @ WVU. For more information, please contact ian.harmon@mail.wvu.edu.

Seasonality, Nonstationarity and the Structural Forecasting of the Index of Industrial Production

By

**Eugene Kouassi, Walter C. Labys,
and François B. Aka**

RESEARCH PAPER 2001-2

Visiting Research Fellow and Fulbright Scholar
Natural Resource Economics Program, West Virginia University

Professor of Resource Economics
Faculty Research Associate, Regional Research Institute
Natural Resource Economics Program, West Virginia University
Morgantown, WV 26506-6108
wlabys@wvu.edu

and
C.R.E.F.A. and Dept. of Economics, University of Laval, Qc, Canada

JEL Classification: C22, E23, E27

Key-words: Index of industrial production, forecasting, structural time series models.

Abstract: In this paper we focus on two ‘STS’ models suitable for forecasting the index of industrial production. The first model requires that the index be transformed with a first and seasonal difference filters. The second model considers the index in its second difference filter, while seasonality is modeled with a constant and seasonal dummy variables. Tests designed to discriminate empirically between these two models are also conducted. Our results prefer the performance of the second model, particularly when the conventional ML estimation procedure is replaced by the ALS procedure. This process together with appropriate seasonal adjustment advances the possibility of using the suggested index forecasts to help to predict business cycle turning points.

Acknowledgements: The authors express their gratitude to Professors A. C. Harvey (University of Cambridge, UK), S. J. Koopman (Tilburg University, The Netherlands), J. Goldstein (Bowdoin College, USA), D. Bolduc (University of Laval, Canada), Dr. C. C. A. Winder (De Nederlandsche Bank NV, Amsterdam, The Netherlands), and all the participants to the 39th Canadian Economic Association Meeting in 1999 in Hull (Quebec, Canada) for helpful discussions and suggestions. They also wish to acknowledge the helpful comments and detailed and valuable reports on an earlier version of this paper by two anonymous referees and an editor. The usual disclaimer applies.

SEASONALITY, NONSTATIONARITY AND THE STRUCTURAL FORECASTING OF THE INDEX OF INDUSTRIAL PRODUCTION

E. Kouassi, W. C. Labys and F. B. Aka

1. Introduction and Motivation

The analysis and forecasting of indexes of industrial production, both in aggregate and for individual sectors, continues to capture interest, because of their pertinence as measures of economic performance and international cross-performance comparisons. Despite this notable role, consensus has not yet been reached as to which appropriate modeling and forecasting strategy to adopt. In this paper we consider two structural models as possibilities for forecasting the index of industrial production. The first is the well-known basic structural time series (STS) model advocated by Harvey and Durbin (1986), which requires that the series under consideration be transformed by first and seasonal difference filters [e.g., see Harvey, 1989 and 1994] respectively to achieve stationary. The second is the STS model that makes use of a second difference filter and assumes that the sum of the seasonal components in twelve consecutive periods equals zero and hence the seasonal pattern can be described with a constant and seasonal dummy variables [e.g., see Harvey, 1989 and 1994].

The rationale and primary motive of the present study is the observation that the forecasts for the index of industrial production from use of the first model, following Harvey and Todd (1983), Harvey and Durbin (1986) and Harvey (1989, 1994), fluctuate too widely. This might indicate that the model is mis-specified; and that the second would be more appropriate.

This paper continues as follows. In Section 2, the two competing forecasting models are introduced and their theoretical backgrounds discussed. In Section 3, a brief account is given of a

method to test for (seasonal) unit roots in monthly data, as a procedure to choose between the models. Section 4 introduces and extends the asymptotic least squares (ALS) estimation method for the case of monthly data, as an alternative to the classic and well-known maximum likelihood (ML) estimation method of STS models developed by Harvey and Peters (1990). In Section 5, an ALS method is applied to estimate both STS seasonal models, while in Section 6 forecasting schemes are applied to the industrial production series under investigation. From an extensive predictive performance analysis the appropriate model for forecasting the index of industrial production emerges. Finally, in Section 7, some concluding remarks and suggestions for future research are given.

2. Model Specification

Consider y_t a time series to be modeled and which can be decomposed as:

$$y_t = \mu_t + \gamma_t + \vartheta_t, \quad t = 1, \dots, T \quad (1)$$

where μ_t is the trend component, γ_t is the seasonal component, and ϑ_t is the irregular component.

In the STS methodology described by Harvey (1989, 1994), the process generating the trend is regarded as a local approximation to a linear trend, i.e.,

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \quad t = 1, \dots, T \quad (2)$$

and

$$\beta_t = \beta_{t-1} + \zeta_t, \quad t = 1, \dots, T \quad (3)$$

where, β_t is the slope and η_t and ς_t are distributed independently of each other and over time with means zero and variances σ_η^2 and σ_ς^2 , respectively.

The process generating the seasonal component, γ_t , (in the case of monthly data), is

$$\gamma_t = -\sum_{j=1}^{11} \gamma_{t-j} + \omega_t, \quad t = 1, \dots, T \quad (4)$$

where ω_t is an independently distributed term with mean zero and variance σ_ω^2 . Here, we assume that the seasonal pattern is slowly changing, but by a mechanism which ensures that the sum of the seasonal components over any s consecutive time periods has an expected value of zero and a variance which remains constant over time [e.g., see Harvey, 1989].

Unlike the trend and the seasonal component, the irregular component is assumed to be stationary. In the basic model it is a white-noise disturbance term with mean zero and variance, σ_ϑ^2 .

Model 1

Substitution of the processes for the trend and seasonal component in equation (1) leads to the first competing STS forecasting model which is:

$$\begin{aligned} \Delta\Delta_{12}y_t = & \left(1 + L + L^2 + L^3 + L^4 + L^5 + L^6 + L^7 + L^8 + L^9 + L^{10} + L^{11}\right)\varsigma_{t-1} \\ & + \Delta_{12}\eta_t + \Delta^2\omega_t + \Delta\Delta_{12}\vartheta_t, \quad t = 14, \dots, T \end{aligned} \quad (5)$$

with L the lag operator, $\Delta = 1 - L$ and $\Delta_{12} = 1 - L^{12}$, and thus $\Delta\Delta_{12}y_t \sim$ a MA(13) process¹, but with a number of non-linear restrictions on the parameters [e.g., see Harvey, 1989]. In equation (5), $\Delta\Delta_{12}y_t$ is stationary since it has been obtained by a first and seasonal difference filters,

¹ More generally the autocovariance function of the basic STS model $\Delta\Delta_s y_t$ based on seasonal dummies indicates that the reduced form is such that $\Delta\Delta_s y_t \sim MA(s+1)$. The unrestricted reduced form therefore contains more parameters than the structural form, namely $s+2$ as opposed to s [e.g. Harvey, 1989, pp.69, 75 and 511]

respectively, and hence has thirteen roots [solving for the equations $1 - L = 0$ and $1 - L^{12} = 0$ in $\mathbb{C}$, the vector space of complex numbers], a double unit root, $L = 1$, and eleven seasonal unit roots which are.

$$-1; \pm i; -\frac{1}{2}(1 \pm i\sqrt{3}); \frac{1}{2}(1 \pm i\sqrt{3}); -\frac{1}{2}(\sqrt{3} \pm i); \frac{1}{2}(\sqrt{3} \pm i) \quad (6)$$

where $i^2 = -1$. The seasonal roots correspond to 6, 3, 9, 8, 4, 2, 10, 7, 5, 1, and 11 cycles per year, respectively. The frequencies of these roots are $\pi, \pm\pi/2, \pm2\pi/3, \pm\pi/3, \pm5\pi/6$, and $\pm\pi/6$, respectively.

Model 2

One may also assume that seasonal pattern is not stochastic but constant over time. In terms of the basic STS model, we have $\sigma_\omega^2 = 0$. For the observed y_t it follows that:

$$\Delta^2 y_t = \varsigma_{t-1} + \Delta \eta_t + \Delta^2 \gamma_t + \Delta^2 \vartheta_t, \quad t = 14, \dots, T \quad (7)$$

The assumption that the sum of the seasonal components in twelve consecutive periods equals zero implies $\gamma_t = \gamma_{t-12}$ - (and therefore $\Delta^2 \gamma_t = \Delta^2 \gamma_{t-12}$) - and hence the seasonal pattern can be described with seasonal dummy variables [e.g., see Winder, 1997, pp.99]. According to equation (7), $\Delta^2 \gamma_t \sim$ a MA(10) process² with deterministic seasonal components [e.g., see Harvey, 1989, pp.69], and γ_t has, therefore, only 10 unit roots [e.g., see Harvey, 1989, pp.69, 75 and 511].

Problems of Mis-specification

These two structural forecasting competing models and, more generally, the class of STS models are often used in modeling exercises [e.g., see Harvey, 1989, 1994]. A practical problem which often occurs in empirical studies is that this class of models may produce forecasts that are

² In fact $S(L)\gamma_t \sim \text{MA}(s-2)$, [e.g., see Harvey, 1989, pp.69, 75 and 511]

either too low or too high. This may be caused by the fact that the appropriate model for y_t is equation (5) while using equation (7), or vice versa. This results in over-differencing and mis-specification. Transforming a series with the $\Delta\Delta_{12}$ filter assumes the presence of 13 roots on the unit circle, two of which are at the zero frequency. Hence, in the case which only the Δ filter [or more generally the Δ^k filter applied k times ($k \in \mathbb{Z}^+ \setminus \{0\}$)] is sufficient to remove non-stationarity, the incorrect assumption of the presence of the other roots implies over-differencing. The mis-specification also can originate from treating deterministic seasonality incorrectly, as being stochastic or vice versa. Osborn (1990) and Franses (1991) empirically demonstrate that this type of mis-specification often occurs.

Since over-differencing and mis-specification are serious dangers in modeling, we will examine the time series structure of the index of industrial production for the countries under investigation without any pre-judgement regarding the nature of seasonality and/or the non-stationarity they may have.

3. Testing for (Seasonal) Unit Roots

To determine what might be the most appropriate STS specification for a particular data series, we first must understand the underlying data generating process. Indexes of industrial production for (i) Canada, (ii) Germany, (iii) Japan, and (iv) the United States are considered in the present study. The data are from the OECD (1998) historical series “Main Economic Indicators” covering the period January 1960 until April 1998³. In the empirical analysis that follows, we consider the log-transformed series.

³ The data for Canada start in January 1961 while that of Germany are available only on the period March 1990 to April 1998 because of the reunification of the western and eastern parts.

A preliminary step in industrial production modeling strategy is to examine the seasonality and/ or non-stationarity issues related to these series. Therefore, we tested all indexes of industrial production for the presence of (seasonal) unit roots. Since the differencing operator L^{12} assumes the presence of 12 roots on the unit circle; this becomes clear by observing that:

$$\begin{aligned}
(1-B^{12}) = & (1-B)(1+B)(1-iB)(1+iB) \left[1 + (\sqrt{3}+i)B/2\right] \left[1 + (\sqrt{3}-i)B/2\right] \\
& \bullet \left[1 - (\sqrt{3}-i)B/2\right] \left[1 + (i\sqrt{3}+1)B/2\right] \left[1 - (i\sqrt{3}-1)B/2\right] \\
& \bullet \left[1 - (i\sqrt{3}+1)B/2\right] \left[1 + (i\sqrt{3}-1)B/2\right]
\end{aligned} \tag{8}$$

where all terms other than $(1-B)$ correspond to seasonal unit roots.

Following the methodology developed by Franses (1991), the presence of (seasonal) unit roots can be tested using the auxiliary regression which is given by:

$$\begin{aligned}
\Psi(B)y_{8,t} = & \varphi_1 y_{1,t-1} + \varphi_2 y_{2,t-1} + \varphi_3 y_{3,t-1} + \varphi_4 y_{3,t-1} + \varphi_5 y_{4,t-1} + \varphi_6 y_{4,t-2} + \varphi_7 y_{5,t-1} + \varphi_8 y_{5,t-2} + \\
& \varphi_9 y_{6,t-1} + \varphi_{10} y_{6,t-2} + \varphi_{11} y_{7,t-1} + \varphi_{12} y_{7,t-2} + \zeta + \xi_t
\end{aligned} \tag{9}$$

The $y_{i,t}$ variables are constructed from the original series, y_t , as follows:

$$\begin{aligned}
y_{1,t} &= (1+B) (1+B^2) (1+B^4+B^8)y_t \\
y_{2,t} &= -(1-B) (1+B^2) (1+B^4+B^8)y_t \\
y_{3,t} &= -(1-B^2) (1+B^4+B^8)y_t \\
y_{4,t} &= -(1-B^4) (1-\sqrt{3}B+B^2) (1+B^2+B^4)y_t \\
y_{5,t} &= -(1-B^4) (1+\sqrt{3}B+B^2) (1+B^2+B^4)y_t \\
y_{6,t} &= -(1-B^4) (1-B^2+B^4) (1-B+B^2)y_t \\
y_{7,t} &= -(1-B^4) (1-B^2+B^4) (1+B+B^2)y_t \\
y_{8,t} &= (1-B^{12})y_t
\end{aligned}$$

where B is the backshift operator given by $(1-B^i)y_t = y_t - y_{t-i}$ and all the industrial production index series are in natural logarithms. Lagged values of the dependent variables $y_{8,t}$ are included as right-hand-side variables in the regression to account for a possible ARMA error process where the lagged dependent variables enter the regression through the $\Psi(B)$ lagged polynomial term [e.g., see Joutz et al. 1995]. The number of lagged dependent variables is chosen by minimizing the Akaike FPE and the Bayesian Schwartz Criterion. Following the guidelines above, the ζ term includes a constant, eleven seasonal dummies, and a trend. For comparison purposes, we also test for unit roots in an equation where the ζ includes a constant with eleven seasonal dummies, but no trend.

Applying ordinary least squares (OLS) to equation (9) gives estimates of the φ_i . In case there are (seasonal) unit roots, the corresponding φ_i are zero. Due to the fact pairs of complex unit roots are conjugates, it should be noted that these roots are only present when pairs of φ_i 's are equal to zero simultaneously. For example, the roots i and $-i$ are only present when φ_3 and φ_4 are equal to zero. There will be no seasonal unit roots if φ_2 through φ_{12} are significantly different from zero. If $\varphi_1 = 0$ then the presence of root 1 can not be rejected. When $\varphi_1 = 0$, and when

φ_2 through φ_{12} are significantly different from zero, this implies the presence of a unit root but no seasonal unit root. Therefore, a first difference seasonal dummy variable model is appropriate. However, $\varphi_i = 0$ for $i = 1, 2, \dots, 12$ implies the presence of both a unit root and eleven seasonal roots. In this case a multiplicative seasonal difference model of the form $(1 - B)(1 - B^{12})y_t$ would be the model of choice.

The general result of testing the null hypothesis that $\varphi_i = 0$ for $i = 1, 2, \dots, 12$ is that seasonality and non-stationarity in the series of industrial production can be appropriately modeled with a STS model with a constant and seasonal dummy variables [e.g., see Tables 1a and 1 b]. Therefore the regularly and commonly applied Δ_{12} not to mention the $\Delta\Delta_{12}$ filter, is inappropriate for the data under investigation. This latter conclusion is consistent with Beaulieu and Miron (1993), Osborn (1990), Franses (1991) and Joutz et al. (1995) who argue in favor of seasonal dummies rather than seasonal differencing.

Table 1a: Testing for seasonal unit roots (Constant, Trend and Seasonal Dummies)

t-statistics	Country			
	Canada	Germany	Japan	USA
ϕ_1	0.0210	-0.0602	0.0008	0.0609 ^a
ϕ_2	-0.1010 ^a	0.0373	0.0070	-0.1242 ^a
ϕ_3	-0.0298 ^c	0.1633 ^b	0.0085	-0.0498 ^b
ϕ_4	-0.0121	-0.1899 ^a	-0.0490 ^b	-0.0572 ^b
ϕ_5	-0.0529 ^c	-0.1348	-0.0846 ^a	-0.1027 ^b
ϕ_6	0.0035	-0.2261 ^a	-0.0889 ^a	0.0065
ϕ_7	0.0011	0.2281 ^a	0.0332	0.0136
ϕ_8	-0.0403	-0.3588 ^a	-0.1046 ^a	-0.0723 ^b
ϕ_9	-0.0605	0.0254	-0.0428	-0.1884 ^a
ϕ_{10}	0.0098	-0.1745	-0.0578	0.0472
ϕ_{11}	0.0041	0.0703	0.0459	-0.0044
ϕ_{12}	-0.0072 ^b	-0.2418 ^b	-0.1739 ^a	-0.1180 ^a
F-statistics				
ϕ_3, ϕ_4	1.8104	4.2160 ^a	2.6494 ^c	8.6111 ^a
ϕ_5, ϕ_6	4.5489 ^a	3.2171 ^b	4.2730 ^a	8.4041 ^a
ϕ_7, ϕ_8	11.3400 ^a	12.3412 ^a	27.1688 ^a	16.3382 ^a
ϕ_9, ϕ_{10}	1.4224	1.8846	1.0239	6.1451 ^a
ϕ_{11}, ϕ_{12}	5.6636 ^a	3.7671 ^b	13.8730 ^a	13.0988 ^a
$\phi_3, \dots, \phi_{12}$	5.7823 ^a	4.3591 ^a	8.9659 ^a	11.0993 ^a

Notes: (i) The auxiliary regression contains constant, trend and seasonal dummies, while $\phi(B)$ is $(1-\phi B^{12})$ and the number of observations for each series equals 447.
(ii) a, b and c denote statistical significance at 1, 5, and 10% levels respectively.

Table 1b: Testing for seasonal unit roots (Constant and Seasonal dummies)

t-statistics	Country			
	Canada	Germany	Japan	USA
ϕ_1	0.0206	-0.0027	0.0004	0.0599 ^a
ϕ_2	-.1028 ^a	0.0454	0.0067	-0.1337 ^a
ϕ_3	-0.0288	0.1200	0.0088	-0.0489 ^b
ϕ_4	-0.0088	-0.2350 ^a	-0.0483 ^b	-0.0437 ^c
ϕ_5	-0.0502	-0.1766 ^b	-0.0840 ^a	-0.0961 ^b
ϕ_6	0.0061	-0.2534 ^b	-0.0883 ^a	0.0162
ϕ_7	-0.0026	0.1977 ^b	0.0329	0.0034
ϕ_8	-0.0376	-0.3348	-0.1042 ^a	-0.0637 ^b
ϕ_9	-0.057	-0.0361	-0.0420	-0.1758 ^a
ϕ_{10}	0.0138	-0.2085 ^b	-0.0568	0.0677
ϕ_{11}	0.0020	0.0597	0.0453	-0.0153
ϕ_{12}	-0.0701	-0.2600 ^b	-0.1733 ^a	-0.1058 ^a
F-statistics				
ϕ_3, ϕ_4	1.5637	4.2054 ^a	2.5955 ^c	6.4298 ^a
ϕ_5, ϕ_6	4.3684 ^a	3.3550 ^b	4.2333 ^a	8.1403 ^a
ϕ_7, ϕ_8	11.9587 ^a	10.0507 ^a	27.2199 ^a	16.9767 ^a
ϕ_9, ϕ_{10}	1.3569	1.6944	0.9941	5.9401 ^a
ϕ_{11}, ϕ_{12}	5.7489 ^a	3.9862 ^a	13.8455 ^a	12.2980 ^a
$\phi_3, \dots, \phi_{12}$	5.9927 ^a	3.6329 ^b	8.9839 ^a	10.9079 ^a

Notes: (i) The auxiliary regression contains constant and seasonal dummies, while $\phi(B)$ is $(1-\phi B^{12})$ and the number of observations for each series equals 447.
(ii) a, b and c denote statistical significance at 1, 5, and 10% levels respectively.

4. Asymptotic Least Squares Estimation Procedure

There are a number of estimation techniques in the econometric literature that can be used to estimate STS models including the traditional ML [e.g., see Harvey, 1989; Harvey and Peters, 1990] estimation procedure. ML estimation of the parameters of unobserved components models is generally carried out with the Kalman filter. The review of the Harvey and Peters (1990) indicates, however, that this can be quite complex, even if relatively simple STS models are considered; not to mention non-convergence and/ or non-optimality issues often encountered in applied econometric studies with the Kalman filter. Various studies have therefore examined frequency domain estimation as an alternative. Young et al. (1991), for instance, apply the sequential spectral decomposition approach which enabled them to carry out the estimation process in completely recursive terms, albeit at the cost of strict optimality in ML sense.

In this paper, however, we consider the ALS procedure⁴ suggested by Gouriéroux, Monfort and Trognon (1985), Kodde et al. (1990), and Monfort and Rabemananjara (1990). ALS is a two-stage estimation procedure and explores the relationships between univariate ARIMA models and STS models. In the first stage an ARIMA model is estimated and in the second a non-linear optimization problem subject to inequality constraints is solved. This approach yields estimators which are asymptotically efficient and ALS is therefore an alternative for ML estimation via the Kalman filter.

In its most simple form, the ALS procedure first estimates an ARIMA model and then performs a GLS regression.

Description of the ALS Procedure

The basic idea of ALS estimation is to use a general specification which is easy to estimate containing the model of interest as a special case. The correspondence between the general and specific model is described by a set of relationships between the parameters of the general model and those of the specific one. ALS estimation is carried out in two stages. In the first stage the parameters of the general model are estimated. Given these estimates and the set of relationships between the parameters of the general and specific model, the ALS estimators of interest are determined in the second stage. This approach yields estimators which are asymptotically equivalent to ML, if the parameters of the general model have been estimated with ML.

Equation (5) implies a restricted MA (13) process for $\Delta\Delta_{12}y_t$. For the general model, which is easy to estimate, an unrestricted MA (13) process is chosen:

$$\Delta\Delta_{12}y_t = (1 + \theta_1 L + \theta_2 L^2 + \theta_3 L^3 + \theta_4 L^4 + \theta_5 L^5 + \theta_6 L^6 + \theta_7 L^7 + \theta_8 L^8 + \theta_9 L^9 + \theta_{10} L^{10} + \theta_{11} L^{11} + \theta_{12} L^{12} + \theta_{13} L^{13})\zeta_t, \quad \zeta_t \sim NID(0, \sigma_\zeta^2) \quad (10)$$

⁴ ALS estimation technique is very attractive and easy to implement, although it has some limitations [e.g., see Dagenais and Dufour, 1991 ; Hansen et al., 1996]

The relationships between the parameters in equations (5) and (7) can be established with the autocovariance of $\Delta\Delta_{12}y_t$, yielding:

$$\begin{aligned}
(1 + \theta_1^2 + \theta_2^2 + \theta_3^2 + \theta_4^2 + \theta_5^2 + \theta_6^2 + \theta_7^2 + \theta_8^2 + \theta_9^2 + \theta_{10}^2 + \theta_{11}^2 + \theta_{12}^2 + \theta_{13}^2)\sigma_\varsigma^2 &= 12\sigma_\varsigma^2 + 2\sigma_\eta^2 + 6\sigma_\omega^2 + 4\sigma_\vartheta^2 \\
(\theta_1 + \theta_2\theta_1 + \theta_3\theta_2 + \theta_4\theta_3 + \theta_5\theta_4 + \theta_6\theta_5 + \theta_7\theta_6 + \theta_8\theta_7 + \theta_9\theta_8 + \theta_{10}\theta_9 + \theta_{11}\theta_{10} + \theta_{12}\theta_{11} + \theta_{13}\theta_{12})\sigma_\varsigma^2 &= \\
11\sigma_\varsigma^2 - 4\sigma_\omega^2 - 2\sigma_\vartheta^2 \\
(\theta_2 + \theta_3\theta_1 + \theta_4\theta_2 + \theta_5\theta_3 + \theta_6\theta_4 + \theta_7\theta_5 + \theta_8\theta_6 + \theta_9\theta_7 + \theta_{10}\theta_8 + \theta_{11}\theta_9 + \theta_{12}\theta_{10} + \theta_{13}\theta_{11})\sigma_\varsigma^2 &= 10\sigma_\varsigma^2 + \sigma_\omega^2 \\
(\theta_3 + \theta_4\theta_1 + \theta_5\theta_2 + \theta_6\theta_3 + \theta_7\theta_4 + \theta_8\theta_5 + \theta_9\theta_6 + \theta_{10}\theta_7 + \theta_{11}\theta_8 + \theta_{12}\theta_9 + \theta_{13}\theta_{10})\sigma_\varsigma^2 &= 9\sigma_\varsigma^2 \\
(\theta_4 + \theta_5\theta_1 + \theta_6\theta_2 + \theta_7\theta_3 + \theta_8\theta_4 + \theta_9\theta_5 + \theta_{10}\theta_6 + \theta_{11}\theta_7 + \theta_{12}\theta_8 + \theta_{13}\theta_9)\sigma_\varsigma^2 &= 8\sigma_\varsigma^2 \\
(\theta_5 + \theta_6\theta_1 + \theta_7\theta_2 + \theta_8\theta_3 + \theta_9\theta_4 + \theta_{10}\theta_5 + \theta_{11}\theta_6 + \theta_{12}\theta_7 + \theta_{13}\theta_8)\sigma_\varsigma^2 &= 7\sigma_\varsigma^2 \\
(\theta_6 + \theta_7\theta_1 + \theta_8\theta_2 + \theta_9\theta_3 + \theta_{10}\theta_4 + \theta_{11}\theta_5 + \theta_{12}\theta_6 + \theta_{13}\theta_7)\sigma_\varsigma^2 &= 6\sigma_\varsigma^2 \\
(\theta_7 + \theta_8\theta_1 + \theta_9\theta_2 + \theta_{10}\theta_3 + \theta_{11}\theta_4 + \theta_{12}\theta_5 + \theta_{13}\theta_6)\sigma_\varsigma^2 &= 5\sigma_\varsigma^2 \\
(\theta_8 + \theta_9\theta_1 + \theta_{10}\theta_2 + \theta_{11}\theta_3 + \theta_{12}\theta_4 + \theta_{13}\theta_5)\sigma_\varsigma^2 &= 4\sigma_\varsigma^2 \\
(\theta_9 + \theta_{10}\theta_1 + \theta_{11}\theta_2 + \theta_{12}\theta_3 + \theta_{13}\theta_4)\sigma_\varsigma^2 &= 3\sigma_\varsigma^2 \\
(\theta_{10} + \theta_{11}\theta_1 + \theta_{12}\theta_2 + \theta_{13}\theta_3)\sigma_\varsigma^2 &= 2\sigma_\varsigma^2 \\
(\theta_{11} + \theta_{12}\theta_1 + \theta_{13}\theta_2)\sigma_\varsigma^2 &= \sigma_\varsigma^2 + \sigma_\vartheta^2 \\
(\theta_{12} + \theta_{13}\theta_1)\sigma_\varsigma^2 &= -\sigma_\eta^2 - 2\sigma_\vartheta^2 \\
(\theta_{13})\sigma_\varsigma^2 &= \sigma_\vartheta^2
\end{aligned}$$

or in matrix notation

$$g(\delta) = A\rho \tag{11}$$

with g a vector of functions

$$in \quad \delta = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7, \theta_8, \theta_9, \theta_{10}, \theta_{11}, \theta_{12}, \theta_{13}, \sigma_\varsigma^2)', \rho = (\sigma_\varsigma^2, \sigma_\eta^2, \sigma_\omega^2, \sigma_\vartheta^2)' \text{ and therefore:}$$

$$\mathbf{A}_{Stochastic\ Seasonality} = \begin{pmatrix} 12 & 2 & 6 & 4 \\ 11 & 0 & -4 & -2 \\ 10 & 0 & 1 & 0 \\ 9 & 0 & 0 & 0 \\ 8 & 0 & 0 & 0 \\ 7 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & -1 & 0 & -2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Further details of this solution are provided in Appendix 1.

Accounting for the restriction imposed in Section 2 (fixed seasonality), we have a special case of matrix $\mathbf{A}_{Stochastic\ Seasonality}$:

$$\mathbf{A}_{Fixed\ Seasonality} = \begin{pmatrix} 1 & 2 & 6 & 6 \\ 0 & -1 & -4 & -6 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

see Appendix 2 for more details

Turning to the general case, given the first-stage estimate of δ , $\hat{\delta}$, ALS minimizes the distances of $g(\hat{\delta}) - A\rho$ in the metric of the matrix

$$S^{-1} = \left[\frac{\partial g}{\partial \delta'} \Omega \frac{\partial g'}{\partial \delta} \right] \quad (12)$$

where Ω is the covariance matrix of $\hat{\delta}$, i.e.

$$\left[g(\hat{\delta}) - A\rho \right]' S \left[g(\hat{\delta}) - A\rho \right] \quad (13)$$

yielding the ALS estimate

$$\hat{\rho} = [A'SA]^{-1} A'S[g(\hat{\delta})] \quad (14)$$

Asymptotic Properties

If δ consistently, $\hat{\rho}$ will also be consistent. Moreover, the large sample distribution of $\hat{\rho}$ is

$$\sqrt{T}(\hat{\rho} - \rho) \xrightarrow{L} N(0, (A'SA)^{-1}) \quad (15)$$

if $\hat{\rho}$ has a large sample normal distribution.

Since the variances are non-negative, an iterative algorithm minimizing the quadratic form (11) subject to inequality restrictions for the parameters is necessary. If, however, all the variance parameters are strictly positive the ALS estimator in equation (12) is equivalent to the GLS estimator in the model:

$$g(\hat{\delta}) = A\rho + u, E(u) = 0, \quad \text{and} \quad E(uu') = S. \quad (16)$$

The unrestricted MA (13) model (equation 5) contains fourteen parameters. Since the basic STS model has only twelve unknown parameters, two restrictions have been imposed. These restrictions are based on the Wald test. The Wald statistic has a value equal to the minimum of the objective function (11) multiplied by the number of observations.

5. Asymptotic Least Squares Estimation Results

The ALS estimation procedure is now used to estimate the above STS models. For the sake of clarity and conciseness, the econometric analysis is based upon the following steps:

Step 1: STS Model Specification

Since the basic STS model has often been applied successfully [e.g., see Harvey 1989, Harvey and Todd 1983, Harvey and Durbin 1986, and Winder, 1997], we adopt its specification.

Also, since unit root tests strongly reject the presence of seasonal unit roots, we examine the alternative model using a deterministic seasonal specification. This model is described by equations (1)-(3) and $\sum_{j=0}^{11} \gamma_{t-j} = 0$. It has been argued above that this model implies a MA (10) process with deterministic seasonals for $\Delta^2 y_t$. This ARIMA process can be used as the general model. In this case, both the STS model and the corresponding ARIMA model have twenty-two parameters. Equation (7) is, therefore, only a reparameterization of the general model. The relationships between the parameters of interest and those of the MA (10) with seasonal dummies can be derived in an analogous way, as performed in the previous section for the basic STS model. Summarizing the two STS models, we have:

Model 1:

$$\Delta\Delta_{12}y_t = (1 + L^2 + L^3 + L^4 + L^5 + L^6 + L^7 + L^8 + L^9 + L^{10} + L^{11})\zeta_{t-1} + \Delta_{12}\eta_t + \Delta^2\omega_t + \Delta\Delta_{12}\vartheta_t$$

Model 2:

$$\Delta^2 y_t = \zeta_{t-1} + \Delta\eta_t + \Delta^2\gamma_t + \Delta^2\vartheta_t$$

Preliminary Tests

Before estimating these two STS models, we first investigate the underlying time series properties of the level and first differences of the data. Our results indicate an important increase in volatility occurred in the post-1970 and -1980 periods. These periods correspond to:

- (i) 1973 – first oil price shock,
- (ii) 1979 – second oil price shock, and
- (iii) 1987 – stock market crash.

These findings have important implications for the specification of STS models [e.g., see Goldstein, 1997]. The heteroscedasticity and trend shifts can be theorized in one of four ways as a

result of [e.g., see Goldstein, 1997]: (i) the normal evolution of the stochastic components of the STS model; (ii) a structural change in the data generation process; (iii) an increase in volatilities, a consequence of (significant) random shocks which increase variances and levels in an autocorrelated (damped, but persistent) manner; or (iv) in the case of increased volatility only, a simple heteroscedasticity pattern (as a continuous function of time). Each of these four perspectives supports, respectively, a different statistical specification: (i) a STS model for the entire sample; (ii) two (or more) separate STS models, one estimated prior to the structural break and the other after; (iii) an autoregressive conditional heteroscedasticity (ARCH) model integrated with a single period STS model; and (iv) a STS model for $\log(Y_t)$ rather than Y_t .

The heteroscedasticity pattern in the data and the possible trend shifts cast serious doubt on the appropriateness of a single uncorrected STS model with assumed homoscedasticity error variances for the entire sample range [e.g., see Goldstein, 1997]. The existence of a heteroscedasticity pattern which is a positive step-function, rather than a continuous function of time, also suggests that a log transformation may (not necessarily) improve the properties of the statistical estimates.

The ARCH and structural change models are competing perspectives on the evolution of a heteroscedasticity pattern. The former is a stochastic explanation, whereas the latter is structural/deterministic. An ARCH interpretation implies that the damping process is slow enough such that significant supply shocks in the 1970s and 1980s have resulted in increased volatility into the 1990s without a return of $\sigma_{\vartheta_t}^2$ to its steady state. In contrast, a structural break argument focuses on major regime shifts, circa 1970 and 1980, associated with a significantly lower rate of productivity growth, increased levels of indebtedness, the intensification of international competition, a decline in profitability, and weakness in the international monetary

system. These fundamental changes in both the economic and institutional structures created a permanently more uncertain and volatile environment for macroeconomic time series and particularly for the index of industrial production; accounting for these changes is thus, crucial in the modeling process.

Based upon a model selection strategy and the structural shift arguments which are the most likely, we thus report estimation results for three distinct periods:

[1960:01–1973:03], [1973:05–1987:09] and [1987:11–1998:04]. These break points have been identified in two recent and separate papers by Goldstein (1997) and Badillo et al. (1999). The results reported below are robust for alternative break points between these sub-periods and the log-transformation appears to improve diagnostics and post-sample testing.

Step 2: STS Model Estimation and Results

Tables 2 and 3 present the ALS estimates for the two STS models, respectively. Because of the large amount of information, all results of estimation will not be discussed in detail. Instead, we focus on some general features and mention some salient facts. The first remarkable result is that for all of the industrial production index series, the seasonal dummy variables are significant in the $\Delta^2 \ln y_t$ STS model; this confirms the appropriateness of the inclusion of these variables in this STS model. A second important result is that for the STS models with stochastic seasonality two of the hyper-parameters equal zero, namely the variance of the seasonal disturbance term and the variance of the slope. This might suggest that seasonality is fixed or constant over time rather than stochastic and thus seasonality may be modeled by a constant and seasonal dummy variable. In the case of the STS model with seasonal dummy variables the two hyper-parameters that equal zero are the variance of the level and the slope. In addition, the hyper-parameters from the STS model with seasonal dummy variables are significant and

relatively smaller than that of the basic STS model; this implies that the STS model with seasonal dummy variables ($\Delta^2 \ln y_t$ model) is more accurate than the STS model with stochastic seasonality ($\Delta\Delta_{12} \ln y_t$ model).

Step 3: Diagnostics

The appropriateness of the two STS models ($\Delta\Delta_{12} \ln y_t$ and $\Delta^2 \ln y_t$ models) is now evaluated based on goodness-of-fit statistics⁵ which are reported in Tables 2 (for the $\Delta\Delta_{12} \ln y_t$ model) and 3 (for the $\Delta^2 \ln y_t$ model), respectively. The goodness-of-fit statistics considered are: the prediction error variance for one-step-ahead predictions, the Box-Ljung test for serial correlation of the residuals, Q (lag length), the Jarque-Bera test for normality of the residuals, N , a standard split sample test for heteroscedasticity, H , the percentage improvement in fit over a random walk plus drift model, R_D^2 , and the percentage improvement in fit over a random walk plus drift and fixed seasonals, R_S^2 . The normality statistic has under the null a $\chi^2(2)$ distribution. The statistic for heteroscedasticity [$F(25, 25)$] distributed and the Ljung-Box statistic has under the null of white noise a $\chi^2(12)$ distribution. The prediction error variance for one-step-ahead predictions is globally satisfactory for all of the indexes of industrial production [e.g., see Table 3]. In addition, results in Table 3 show that normality and heteroscedasticity have to be rejected for the majority of the industrial production index series when using the $\Delta\Delta_{12} \ln y_t$ model.

The R_D^2 statistics, which report the percentage improvement in fit over the random walk with drift model, range from 0.03 to 0.15 for the $\Delta\Delta_{12} \ln y_t$ model, while they range from 0.21 to

⁵ The goodness-of-fit statistics used are based on Harvey's definitions and notations [e.g., see Harvey 1989]

0.43 for the $\Delta^2 \ln y_t$ model. This result provides further evidence for the appropriateness of this latter specification ($\Delta^2 \ln y_t$ model) in modeling the index of industrial production.

The R_s^2 statistics, which also report the percentage improvement in fit over the random walk drift model and fixed seasonals, range from 0.05 to 0.11 for the $\Delta\Delta_{12} \ln y_t$ model and 0.15 to 0.35 for the $\Delta^2 \ln y_t$ model. This result is indeed another support for the appropriateness of this latter specification ($\Delta^2 \ln y_t$ model) in modeling the index of industrial production.

Table 2: Estimation Results of Models for $\Delta_1 \Delta_{12} \ln y_t$

	Country						
	Canada	Germany		Japan		USA	
[1960:01 – 1973:03]							
(i) Parameter Estimates							
σ_{ϑ}^2	1.085	---	---	0.980	0.019	0.908	0.0002
σ_{η}^2	0.001	---	---	0.013	0.002	0.024	0
σ_{ξ}^2	0	---	---	0	0	0	0
σ_{ω}^2	0	---	---	0	0	0	0
(ii) Diagnostics							
Prediction Error Variance	1.023	---	---	1.022	0.0003	0.991	0.0003
Heteroskedasticity	3.474 ^a	---	---	0.545	0.544	1.074	1.073
Ljung-Box Q(24) statistic	8.236	---	---	35.508 ^a		22.887 ^b	
Ljung-Box Q(36) statistic	11.552	---	---	49.168 ^a		37.423 ^b	
Jarque-Bera Normality statistic	27.353 ^a	---	---	0.423	2.539	2.344	1.028
R ² _D	0.552	---	---	0.532	0.533	0.436	0.437
R ² _S	0.524	---	---	0.523	0.521	0.428	0.429
(iii) Post-sample Testing							
Chow F-test	3.548 ^a	---	---	0.688		0.330	2.051 ^c
CUSUM-test	0.554	---	---	1.047		-1.265	-1.266

[1973:05 – 1987:09]

(i) Parameter Estimates

σ_{ϑ}^2	1.017	---	---	0.917	0.0004	0.635	0.0001
σ_{η}^2	0.006	---	---	0.023	0	0.117	0
σ_{ξ}^2	0	---	---	0	0	0	0
σ_{ω}^2	0	---	---	0	0	0	0

(ii) Diagnostics

Prediction Error Variance	1.032	---	---	1.008	0.0004	0.913	0.0002
Heteroskedasticity	0.671	---	---	0.661	0.660	0.515	0.517
Ljung-Box Q(24) statistic	24.186 ^b	---	---	29.099 ^a		25.180 ^a	
Ljung-Box Q(36) statistic	38.479 ^b	---	---	40.027 ^b		33.788 ^c	
Jarque-Bera Normality statistic	1.900	---	---	11.056 ^a	1.799	0.055	4.972
R_D^2	0.539	---	---	0.512	0.511	0.307	0.308
R_S^2	0.536	---	---	0.504	0.514	0.300	0.299

(iii) Post-sample Testing

Chow F-test	3.036 ^a	---	---	0.605	1.011	0.531	0.633
CUSUM-test	0.720	---	---	0.596	0.599	0.145	0.147

[1987:11 – 1998:04]

(i) Parameter Estimates

σ_{ϑ}^2	1.057	0.829	0.0002	0.797	0.686
σ_{η}^2	0.0004	0	0	0	0
σ_{ξ}^2	0	0	0	0	0
σ_{ω}^2	0	0.003	0.0002	0.0001	0.0003

(ii) Diagnostics

Prediction Error Variance	0.995	1.165	0.0007	0.772	0.729
Heteroskedasticity	0.936	1.288	1.016	1.102	0.148
Ljung-Box Q(24) statistic	9.313	10.218		14.949	19.888 ^b
Ljung-Box Q(36) statistic	16.620	18.199		30.716	29.841
Jarque-Bera Normality statistic	0.023	0.237	0.019	75.531 ^a	492.999 ^a
R_D^2	0.585	0.555	0.583	0.643	0.628
R_S^2	0.574	0.534	0.557	0.614	0.596

(iii) Post-sample Testing

Chow F-test	2.316 ^a	1.197	2.244 ^a	1.396	3.023 ^a
CUSUM-test	0.091	-0.232	-0.024	-1.527	-0.548

Notes: a, b and c indicate statistical significance at 1%, 5% and 10% levels, respectively.

Table 3: Estimation Results of Models for $\Delta^2 \ln y_t$

	Country				
	Canada	Germany		Japan	USA
[1960:01 – 1973:03]					
(i) Parameter Estimates					
γ_1	1.442 ^a	---	---	-1.928 ^a	0.955 ^a
γ_2	1.425 ^a	---	---	1.367 ^a	0.621 ^a
γ_3	0.351 ^a	---	---	-0.254 ^a	-0.185 ^c
γ_4	1.005 ^a	---	---	-2.152 ^a	-0.040
γ_5	0.839 ^a	---	---	0.234 ^b	0.259 ^b
γ_6	1.072 ^a	---	---	-0.017	0.586 ^a
γ_7	-0.555 ^a	---	---	-0.772 ^a	-1.743 ^a
γ_8	2.395 ^a	---	---	-0.848 ^a	2.511 ^a
γ_9	1.069 ^a	---	---	0.544 ^a	0.102
γ_{10}	0.272 ^b	---	---	-0.935 ^a	-0.492 ^a
γ_{11}	1.025 ^a	---	---	-0.557 ^a	-0.309 ^b
σ^2_{ϑ}	4.17x10 ⁻⁶	---	---	0.0006	0.002
σ^2_{η}	0	---	---	0	0
σ^2_{ζ}	0	---	---	0	0
σ^2_{ω}	0.031	---	---	0.045	0.082
(ii) Diagnostics					
Prediction Error Variance	0.030	---	---	0.046	0.090
Heteroskedasticity	2.640 ^a	---	---	0.594	1.266
Ljung-Box Q(24) statistic	36.316 ^a	---	---	82.996 ^a	20.197
Ljung-Box Q(36) statistic	39.899 ^b	---	---	94.348 ^a	35.288
Jarque-Bera Normality statistic	83.115 ^a	---	---	2.560	1.886
R ² _D	0.985	---	---	0.984	0.968
R ² _S	0.880	---	---	0.856	0.638
(iii) Post-sample Testing					
Chow F-test	1.662	---	---	0.951	2.235 ^a
CUSUM-test	1.006	---	---	1.204	-0.168

[1973:05 – 1987:09]

(i) Parameter Estimates

γ_1	1.899 ^a	---	---	-1.605 ^a	0.976 ^a
γ_2	1.691 ^a	---	---	1.581 ^a	0.933 ^a
γ_3	0.168 ^c	---	---	0.041	-0.536 ^a
γ_4	0.737 ^a	---	---	-1.501 ^a	0.174
γ_5	1.166 ^a	---	---	-0.119	0.350 ^a
γ_6	1.566 ^a	---	---	0.450 ^a	0.933 ^a
γ_7	-0.799	---	---	-0.515 ^a	-1.592
γ_8	2.928 ^a	---	---	-1.095 ^a	2.192
γ_9	1.317 ^a	---	---	1.316 ^a	-0.132
γ_{10}	0.196 ^b	---	---	-1.105 ^a	-0.451 ^a
γ_{11}	1.215 ^a	---	---	-0.184	-0.061
σ_{ϑ}^2	0.0004	---	---	0.006	0.013
σ_{η}^2	0	---	---	0	0
σ_{ζ}^2	0	---	---	0	0
σ_{ω}^2	0.034	---	---	0.033	0.072

(ii) Diagnostics

Prediction Error Variance	0.035	---	---	0.035	0.113
Heteroskedasticity	0.678	---	---	0.782	0.596
Ljung-Box Q(24) statistic	33.653 ^a	---	---	41.573 ^a	33.497 ^a
Ljung-Box Q(36) statistic	47.031 ^a	---	---	56.586 ^a	45.412 ^b
Jarque-Bera Normality statistic	2.293	---	---	0.273	16.541 ^a
R_D^2	0.987	---	---	0.987	0.961
R_S^2	0.786	---	---	0.812	0.685

(iii) Post-sample Testing

Chow F-test	0.393	---	---	1.136	0.688
CUSUM-test	0.457	---	---	0.489	0.158

[1987:11 – 1998:04]

(i) Parameter Estimates

γ_1	2.223 ^a	0.538 ^b	-0.996 ^a	0.918 ^a
γ_2	1.707 ^a	1.115 ^a	1.223 ^a	-0.060
γ_3	0.229 ^a	1.109 ^a	0.311 ^b	-0.332
γ_4	0.852 ^a	-1.062 ^a	-1.822 ^a	-0.158
γ_5	1.319 ^a	0.739 ^a	0.338 ^b	0.277
γ_6	1.523 ^a	1.116 ^a	0.777 ^a	0.788 ^a
γ_7	-0.570 ^a	-0.489 ^b	-0.472 ^a	-1.657 ^a
γ_8	3.251 ^a	0.292	-1.024 ^a	1.944 ^a
γ_9	0.864 ^a	2.451 ^a	1.362 ^a	-0.946 ^a
γ_{10}	0.568 ^a	-0.759 ^a	-0.937 ^a	-0.312
γ_{11}	1.297 ^a	0.052	-0.027	-0.212
σ_{ϑ}^2	2.01x10 ⁻⁵	0	0	0.226
σ_{η}^2	0	0	1.18x10 ⁻⁵	0
σ_{ζ}^2	0	0.002	0	0
σ_{ω}^2	0.019	0.077	0.030	0.016

(ii) Diagnostics

Prediction Error Variance	0.017	0.073	0.017	0.286
Heteroskedasticity	1.393	1.336	0.073	0.076
Ljung-Box Q(24) statistic	16.688	16.126	16.626	13.793
Ljung-Box Q(36) statistic	28.821	30.853	31.819	22.676
Jarque-Bera Normality statistic	0.221	2.076	67.373 ^a	1834.027 ^a
R_D^2	0.994	0.981	0.994	0.913
R_S^2	0.845	0.139	0.664	0.668

(iii) Post-sample Testing

Chow F-test	3.509 ^a	1.707 ^c	0.306	0.263
CUSUM-test	0.189	-0.974	0.558	-0.429

Notes: a, b and c indicate statistical significance at 1%, 5% and 10% levels, respectively.

Step 4: Post-sample Testing

To further evaluate the adequacy of the estimated models, we also assessed their post-sample predictive performance. In each case, the formal post-sample predictive test was based on the CUSUM test [e.g., see Harvey, 1989 pp. 257 and 272]. An alternative test considered as well is the Chow test. Under the null hypothesis that the model is correctly specified, the Chow statistic is $F(14,119)$ distributed for sub-period 1, $F(17,145)$ distributed for sub-period 2, and $F(12,100)$ distributed for sub-period 3; while on these same sub-periods the CUSUM statistic has a $t(119)$, $t(145)$ and $t(100)$ distribution respectively. Except in a few cases [e.g., see Table 3], the post-sample results indicate rather better prediction performance for the $\Delta^2 \ln y_t$ model than for the $\Delta\Delta_{12} \ln y_t$ model; this provides additional evidence of the appropriateness of the STS with seasonal dummies in modeling the index of industrial production.

6. Forecasting

To evaluate the $\Delta\Delta_{12} \ln y_t$ and $\Delta^2 \ln y_t$ models with respect to their forecasting performance; and following Andrews (1994), forecasts for the 12 months out-of-sample are generated from each of the two models.

To investigate whether there are significant differences between the forecasts, the non-parametric sign (exact binomial), sign (normal approximation), Wilcoxon-ranks and Van der Waerden (normal scores) tests and a number of evaluating criteria are considered in Table 4. Discriminating test results do not establish a clear superiority for either of the two models. The general result with respect to the criteria mean error (ME) through root mean squared error (RMSE) and Theil's U statistic, however, seems to be that the $\Delta^2 \ln y_t$ model outperforms the $\Delta\Delta_{12} \ln y_t$ model, irrespective of country and sub-period consideration. It is

also clear that the number of positive forecast errors M from using the $\Delta^2 \ln y_t$ model are closer to what might have been expected on average, while those when using a $\Delta\Delta_{12} \ln y_t$ model are higher and out of any reasonable range.

From the results of (i) PIMSE, which denotes the percentage improvement of forecasts from the $\Delta^2 \ln y_t$ model with respect to mean squared error, and of (ii) SIGNSE, which reports the number of times the squared error of $\Delta^2 \ln y_t$ model is smaller than that of $\Delta\Delta_{12} \ln y_t$ model in pairwise comparison, it clearly appears that most of the differences between the models are significant and are in favor of the $\Delta^2 \ln y_t$ model.

Table 4: Evaluation of the 12 Months Out-of-Sample Forecasting Performance of Models

Criterion	Canada		Germany		Japan		USA	
	$\Delta\Delta_{12} \ln y_t$	$\Delta^2 \ln y_t$	$\Delta\Delta_{12} \ln y_t$	$\Delta^2 \ln y_t$	$\Delta\Delta_{12} \ln y_t$	$\Delta^2 \ln y_t$	$\Delta\Delta_{12} \ln y_t$	$\Delta^2 \ln y_t$
[1960:01 – 1973:03]								
ME	-0.041	0.010	---	---	-0.007	-0.003	-0.003	0.008
MAE	1.034	0.499	---	---	0.588	0.168	0.742	0.320
MAPE	675.446	47.634	---	---	132.012	72.174	222.099	46.082
RMSE	1.238	0.575	---	---	0.687	0.189	0.868	0.358
M	4	5	---	---	1	0	6	5
Theil's U Statistic	0.634	0.288	---	---	0.456	0.093	0.510	0.161
PIMSE of $\Delta^2 \ln y_t$	44.250 ^b		---		27.825 ^b		41.275 ^b	
SIGNSE	8 ^b		---		12 ^b		12 ^b	
Sign (exact binomial)	74		---		76		81	
Sign (normal approximation)	1.123		---		0.413		1.241	
Wilcoxon Signed-Ranks Test	0.606		---		0.556		1.463	
Van der Waerden Test	-0.426		---		-0.422		1.602 ^c	
[1973:05 – 1987:09]								
ME	0.0005	0.003	---	---	-0.002	-0.004	0.028	-0.032
MAE	0.483	0.143	---	---	0.682	0.194	0.416	0.233
MAPE	243.227	131.848 ---		---	273.899	33.013	201.264	107.741
RMSE	0.563	0.180	---	---	0.743	0.260	0.501	0.297
M	0	4	---	---	2	3	2	2
Theil's U Statistic	0.482	0.101	---	---	0.558	0.116	0.436	0.165
PIMSE of $\Delta^2 \ln y_t$	52.651 ^b				42.528 ^b		24.025 ^b	
SIGNSE	6 ^b				11 ^b		7 ^b	
Sign (exact binomial)	95		---		95		90	
Sign (normal approximation)	1.216		---		1.216		0.455	
Wilcoxon Signed-Ranks Test	1.618 ^c		---		0.985		0.954	
Van der Waerden Test	-1.553		---		-0.768		1.092	

[1987:11 – 1998:04]

ME	-0.008	-0.007	-0.150	-0.001	-0.027	-0.001	-0.029	0.001
MAE	0.802	0.199	1.009	0.249	0.623	0.282	0.312	0.185
MAPE	101.052	219.681	86.757	23.300	179.345	75.234	128.211	64.267
RMSE	0.996	0.251	1.224	0.346	0.715	0.307	0.392	0.229
M	6	3	7	4	3	3	0	0
Theil's U Statistic	0.986	0.133	0.584	0.159	0.630	0.152	0.519	0.137
PIMSE of $\Delta^2 \ln y_t$	16.216 ^b		23.500 ^b		38.600 ^b		58.416 ^b	
SIGNSE	12 ^b		12 ^b		12 ^b		3 ^b	
Sign (exact binomial)	66		42		70		69	
Sign (normal approximation)	0.536		0.222		1.158		0.979	
Wilcoxon Signed-Ranks Test	0.177		0.706		1.124		0.545	
Van der Waerden Test	0.124		0.676		-0.859		0.368	

Notes: (i) The forecast error is defined as the true value y minus the forecasted value f . Forecast evaluation criteria are the mean error, ME, mean absolute error, MAE, mean absolute percentage error, MAPE, and root mean squared error, RMSE. M denotes the number of times y exceeds f . U is Theil's test statistic. PIMSE denotes the percentage improvement of forecasts from the $\Delta^2 \ln y_t$ model with respect to mean squared error. SIGNSE refers to the sign test which reports the number of times the squared error of $\Delta^2 \ln y_t$ model is smaller than that of $\Delta \ln y_t$ model in pairwise comparison. The Binomial sign test statistic reports the value of the sample proportion above and below the true median if the sample were drawn randomly from a binomial distribution. The Normal approximation sign test statistic reports the value of the sample proportion above and below the true median if the sample were drawn randomly from a normal distribution. The Wilcoxon signed-rank test statistic refers to the ranks of positive differences between the forecasts. The Van der Waerden test statistic is similar to the Wilcoxon signed-rank test but it is based on smoothed ranks. Definitions and asymptotic results for these non-parametric test can be found in Conover (1980) and Sheskin (1997).

(ii) a, b and c indicate statistical significance at 1%, 5% and 10% levels, respectively.

7. Final Remarks

In this paper, it has been shown that correctly taking account of the type of seasonality and non-stationarity in monthly data can improve forecasting performance. This is illustrated in the context of the STS models which include a number of variants [STS with trend; STS with trend and cycle; STS with trend and seasonality; STS with trend, cycle and seasonality etc.], and thus feature a high risk of increased mis-specification based on inappropriate data transformations.

The major result of the paper, which is applied to the forecasting of the index of industrial production, is that the recognition of the presence of seasonal unit roots can have important implications for forecasting and model building. Recent additional arguments for not automatically and systematically applying difference or moving average filters can be found in Franses (1991). The understanding and control of the properties of these filters based on their transfer functions is a necessary, though not sufficient, condition for adequately modeling and forecasting of economic time series. Finally, we have shown that employing the ALS estimation procedure together with the STS approach can provide production index forecasts of high accuracy, making it a useful alternative to other production index forecasting approaches based on explanatory variables for forecasting turning points in the business cycle.

APPENDIX 1. AUTOCOVARIANCES OF $\Delta\Delta_{12}y_t$

$$\begin{aligned}\Delta\Delta_{12}y_t &= (1 + L + L^2 + L^3 + L^4 + L^5 + L^6 + L^7 + L^8 + L^9 + L^{10} + L^{11})\varsigma_{t-1} + \Delta_{12}\eta_t + \Delta^2\omega_t + \Delta\Delta_{12}\vartheta_t; \\ &= \varsigma_{t-1} + \varsigma_{t-2} + \varsigma_{t-3} + \varsigma_{t-4} + \varsigma_{t-5} + \varsigma_{t-6} + \varsigma_{t-7} + \varsigma_{t-8} + \varsigma_{t-9} + \varsigma_{t-10} + \varsigma_{t-11} + \varsigma_{t-12} + \eta_t - \eta_{t-12} + \omega_t \\ &\quad - 2\omega_{t-1} + \omega_{t-2} + \vartheta_t - \vartheta_{t-1} - \vartheta_{t-12} + \vartheta_{t-13}\end{aligned}$$

Define $X_t = \Delta\Delta_{12}y_t$, we obtain:

$$\begin{aligned}1 - E[X_t X_t] &= 12\sigma_\varsigma^2 + 2\sigma_\eta^2 + 6\sigma_\omega^2 + 4\sigma_\vartheta^2 \\ 2 - E[X_t X_{t-1}] &= 11\sigma_\varsigma^2 - 4\sigma_\omega^2 - 2\sigma_\vartheta^2 \\ 3 - E[X_t X_{t-2}] &= 10\sigma_\varsigma^2 + \sigma_\omega^2 \\ 4 - E[X_t X_{t-3}] &= 9\sigma_\varsigma^2 \\ 5 - E[X_t X_{t-4}] &= 8\sigma_\varsigma^2 \\ 6 - E[X_t X_{t-5}] &= 7\sigma_\varsigma^2 \\ 7 - E[X_t X_{t-6}] &= 6\sigma_\varsigma^2 \\ 8 - E[X_t X_{t-7}] &= 5\sigma_\varsigma^2 \\ 9 - E[X_t X_{t-8}] &= 4\sigma_\varsigma^2 \\ 10 - E[X_t X_{t-9}] &= 3\sigma_\varsigma^2 \\ 11 - E[X_t X_{t-10}] &= 2\sigma_\varsigma^2 \\ 12 - E[X_t X_{t-11}] &= \sigma_\varsigma^2 + \sigma_\vartheta^2 \\ 13 - E[X_t X_{t-12}] &= -\sigma_\eta^2 - 2\sigma_\vartheta^2 \\ 14 - E[X_t X_{t-13}] &= \sigma_\vartheta^2 \\ 15 - E[X_t X_{t-h}] &= 0, \quad \forall h \geq 14.\end{aligned}$$

APPENDIX 2. AUTOCOVARIANCES OF $\Delta^2 y_t$

$$\begin{aligned}\Delta^2 y_t &= \zeta_{t-1} + \Delta \eta_t + \Delta^2 \gamma_t + \Delta^2 \vartheta_t; \\ &= \zeta_{t-1} + \eta_t - \eta_{t-1} + \gamma_t - 2\gamma_{t-1} + \gamma_{t-2} + \vartheta_t - 2\vartheta_{t-1} + \vartheta_{t-2}\end{aligned}$$

Define $X_t = \Delta^2 y_t$, we obtain

$$\begin{aligned}1 - E[X_t X_t] &= \sigma_\zeta^2 + 2\sigma_\eta^2 + 6\sigma_\gamma^2 + 6\sigma_\vartheta^2 \\ 2 - E[X_t X_{t-1}] &= -\sigma_\eta^2 - 4\sigma_\gamma^2 - 4\sigma_\vartheta^2 \\ 3 - E[X_t X_{t-2}] &= \sigma_\gamma^2 + \sigma_\vartheta^2 \\ 4 - E[X_t X_{t-h}] &= 0, \quad h \geq 3.\end{aligned}$$

APPENDIX 3. DISCUSSION OF ML ESTIMATES VIA THE KALMAN FILTER

In order to compare the ALS estimation results with the ML estimates obtained via the Kalman filter, the two STS models have also been estimated with the computer package STAMP, version 5 (e.g., see Koopman et al., 1995). The results of exact ML estimation in the time domain with numerical optimization carried out by a quasi-Newton algorithm can be obtained upon request from the authors.

In general results from the STS model with seasonal dummy variables based on ML estimation are fairly good and closed to those of ALS estimates discussed in the paper. The estimates of the seasonals are very similar also. The differences for the estimates of the variance parameters are somewhat larger. The t-ratios appear to be larger than those of the ALS estimates, but this result does not hold for all series.

For the results based on the basic STS model, the discrepancies between the Kalman filter estimates and those according to ALS are larger. This holds notably for the series for which two or three variance parameters were found to be zero with ALS.

The appropriateness of the models estimated by ML via the Kalman filter has been examined by the same statistical criteria as in the main text. The results be for diagnostics or post-sample testing are very similar to those for the STS models estimated by ALS. This holds notably for the tests for normality and heteroscedasticity except for Canada and Germany and for the second sub-period. With respect to the Ljung-Box statistic the differences between the ALS – and Kalman filter –estimated models are somewhat larger; Ljung – Box statistic being larger for the Kalman filter – estimated models.

The results on post sample predictive testing (CUSUM and Chow test) also yield a similar picture to those for the ALS–estimated models except in the case of Canada and Japan and for the first sub-period.

APPENDIX 4. FORECAST SUMMARY

The following set of tables provides comparisons of forecast and actual values over time for the index of industrial production.

Table A.1: 12 months-out-of sample forecasts (Model 1)

	CANADA			GERMANY			JAPAN			USA		
	Actual	Predicted	Error	Actual	Predicted	Error	Actual	Predicted	Error	Actual	Predicted	Error
1997:04	120.60	120.57	0.03	100.60	100.60	0.00	100.30	100.30	0.00	123.70	123.70	0.00
1997:05	119.00	119.01	-0.01	94.30	94.34	-0.04	98.70	98.67	0.03	123.00	123.01	-0.01
1997:06	121.70	121.70	0.00	102.90	102.84	0.06	103.40	103.39	0.01	127.10	127.10	0.00
1997:07	110.50	110.48	0.02	101.60	101.57	0.03	107.40	107.42	-0.02	124.10	124.08	0.02
1997:08	118.50	118.51	-0.01	88.50	88.56	-0.06	94.10	94.10	0.00	129.90	129.90	0.00
1997:09	124.00	123.99	0.01	105.60	105.58	0.02	106.60	106.59	0.01	130.20	130.20	0.00
1997:10	125.60	125.59	0.01	109.20	109.18	0.02	103.70	103.72	-0.02	129.60	129.59	0.01
1997:11	127.10	127.12	-0.02	106.70	106.71	-0.01	99.30	99.34	-0.04	128.00	127.99	0.01
1997:12	116.20	116.18	0.02	99.30	99.31	-0.01	100.60	100.58	0.02	126.80	126.79	0.01
1998:01	115.20	115.24	-0.04	96.50	96.46	0.04	93.20	93.21	-0.01	127.30	127.30	0.00
1998:02	127.40	127.37	0.03	99.80	99.79	0.01	96.40	96.41	-0.01	128.60	128.60	0.00
1998:03	128.40	128.40	0.00	113.20	113.15	0.05	108.00	108.01	-0.01	128.90	128.89	0.01
1998-04	--	--	--	105.10	105.16	-0.06	94.50	94.50	0.00	126.80	126.82	-0.02

Table A.2: 12 months-out-of sample forecasts (Model 2)

	CANADA			GERMANY			JAPAN			USA		
	Actual	Predicted	Error	Actual	Predicted	Error	Actual	Predicted	Error	Actual	Predicted	Error
1997:04	120.60	120.55	0.05	100.60	100.60	0.00	100.30	100.30	0.00	123.70	123.70	0.00
1997:05	119.00	119.05	-0.05	94.30	99.33	-5.03	98.70	98.70	0.00	123.00	123.01	-0.01
1997:06	121.70	121.61	0.09	102.90	102.78	0.12	103.40	103.35	0.05	127.10	127.05	0.05
1997:07	110.50	110.62	-0.12	101.60	101.61	-0.01	107.40	107.39	0.01	124.10	124.15	-0.05
1997:08	118.50	118.35	0.15	88.50	88.68	-0.18	94.10	94.20	-0.10	129.90	129.83	0.07
1997:09	124.00	124.00	0.00	105.60	105.28	0.32	106.60	106.46	0.14	130.20	130.24	-0.04
1997:10	125.60	125.67	-0.07	109.20	109.35	-0.15	103.70	103.70	0.00	129.60	129.60	0.00
1997:11	127.10	127.06	0.04	106.70	106.74	-0.04	99.30	99.34	-0.04	128.00	128.00	0.00
1997:12	116.20	116.28	-0.08	99.30	99.34	-0.04	100.60	100.59	0.01	126.80	126.79	0.01
1998:01	115.20	115.15	0.05	96.50	96.42	0.08	93.20	93.26	-0.06	127.30	127.28	0.02
1998:02	127.40	127.27	0.13	99.80	99.77	0.03	96.40	96.33	0.07	128.60	128.59	0.01
1998:03	128.40	128.50	-0.10	113.20	113.12	0.08	108.00	107.89	0.11	128.90	128.90	0.00
1998-04	--	--	--	105.10	105.25	-0.15	94.50	94.63	-0.13	126.80	126.81	-0.01

REFERENCES

- Andrews, R.L., "Forecasting Performance Structural of Time Series Model", *Journal of Business and Economic Statistics*, **12** (1994), 129-133.
- Badillo, D., Labys W.C., and Yangru, Wu., "Identifying Trends and Breaks in Primary Commodity Prices", *European Financial Review*, **5** (1999), 315-330.
- Beaulieu, J.J., and Miron, J.A., "Seasonal Unit Roots in Aggregate U.S. Data [with Discussion]", *Journal of Econometrics*, **55** (1993), 305-331.
- Conover, W.J., *Practical Nonparametric Statistics*, 2nd Edition, New-York, John Wiley and Sons, 1980.
- Engle, R.F., Granger, C.W.J., Hylleberg S. and Lee, H.S., "Seasonal Cointegration: The Japanese Consumption Function [with Discussion]", *Journal of Econometrics*, **55** (1993), 275-303.
- Dagenais, M., and Dufour, J.M., "Invariance, Nonlinear Models and Asymptotic Tests", *Econometrica*, **59** (1991), 1601-1615.
- Flores, B.E., "The Utilization of the Wilcoxon Test to Compare Forecasting Methods: A Note", *International Journal of Forecasting*, **5** (1989), 529-535.
- Franses, P.H., "Seasonality, Non-Stationarity and the Forecasting of Monthly Time Series", *International Journal of Forecasting*, **7** (1991), 199-208.
- Franses, P.H., and Hobijn, B "Critical Values for Unit Root Tests in Seasonal Time Series", *Journal of Applied Statistics*, **24** (1997), 25-47.
- Goldstein, J., "Is the Endogenous Business Cycle Dead?", *Southern Economic Journal*, **63** (1997), 962-977.
- Gourieroux, C., Monfort, A., and Trognon, A., "Moindres Carrés Asymptotiques", *Annales de l'INSEE*, **58** (1985), 91-122.
- Hansen, L.P., Heaton, J., and Yaron, A., "Finite Sample Properties of Some Alternative GMM Estimators", *Journal of Business and Economic Statistics*, **14** (1996), 262-280.
- Harvey, A.C., *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, (1989).

- Harvey, A.C., *Time Series Models*, 2nd. Edition, Cambridge, MA: The MIT Press, (1994).
- Harvey, A.C., and Todd, P.H.J., "Forecasting Economic Time Series with Structural and Box-Jenkins Models [with discussion]", *Journal of Business and Economic Statistics*, **1** (1983), 299-315.
- Harvey, A.C., and Durbin, J., "The Effects of Seat Belt Legislation on British Road Casualties: A Case Study in Structural Time Series Modelling", *Journal of The Royal Statistical Society, Ser.A.*, **149** (1986), 187-227.
- Harvey, A.C., and Peters S., "Estimation Procedures for Structural Time Series Models", *Journal of Forecasting*, **9** (1990), 89-108.
- Hylleberg, S., and Pagan, A.R., "Seasonal Integration and the Evolving Seasonal Model", *International Journal of Forecasting*, **13** (1997), 329-340.
- Joutz, F.L., G.S. Maddala and R.P. Trost "An Integrated Bayesian Vector Autoregression and Error Correction Model for Forecasting Electricity Consumption and Prices", *Journal of Forecasting*, **14** (1995), 287-310.
- Kodde, D., Palm, F.C., and Pfann, G.A., "Asymptotic Least Squares Estimation: Efficiency Considerations and Applications", *Journal of Applied Econometrics*, **3** (1990), 229-243.
- Koopman, S.J., Harvey, A.C., Doornik, J., and Shephard, N., STAMP 5.0 - Structural Time Series Package, Chapman and Hall, (1995).
- Monfort, A., and Rabemananjara, R., "From a VAR Model to a Structural Model, with An Application to the Wage-Price Spiral", *Journal of Applied Econometrics*, **5** (1990), 203-227.

- Osborn, D.R., "A Survey of Seasonality in U.K. Macroeconomic Variables",
International Journal of Forecasting, **6** (1990), 327-336.
- Sheskin, D.J., *Parametric and Nonparametric Statistical Procedures*, CRC Press,
1997.
- Winder, C.C.A., "Structural Time-Series Modelling of Monetary Aggregates: A Case
Study for Eleven European Countries", *Journal of Forecasting*, **16** (1997), 97-
123.
- Young, P.C., Ng, CN., Lane, K., and Parker, D., "Recursive Forecasting, Smoothing
and Seasonal Adjustment of Non-Stationary Environment Data", *Journal of
Forecasting*, **10** (1991), 57-89.