BUREAU OF THE CENSUS
STATISTICAL RESEARCH DIVISION REPORT SERIES

SRD Research Report Number: CENSUS/SRD/RR–90/01

ALTERNATIVE APPROACHES TO THE ANALYSIS OF TIME SERIES COMPONENTS

by

W. R. Bell
U.S. Bureau of the Census
Statistical Research Division
Room 3000, FOB #4
Washington, DC 20233 U.S.A.

and

M. G. Pugh
Department of Biostatistics
Harvard School of Public Health
677 Huntington Avenue
Boston, MA 92115 U.S.A.

Recommended by: Nash J. Monsour

Report completed: June 11, 1990

Report issued: June 11, 1990

ALTERNATIVE APPROACHES TO THE ANALYSIS OF TIME SERIES COMPONENTS

W. R. Bell and M. G. Pugh

ABSTRACT

In the time series literature of recent years one finds different approaches to the analysis of time series postulated to follow some type of component structure. There are alternatives to the now familiar ARIMA (autoregressive–integrated–moving average) modeling approach, perhaps the most popular being the "structural modeling" approach of Harvey and others, which uses an explicit components structure. Despite the considerable research on these models, remarkably little work has appeared comparing results from the alternative approaches. Questions arise regarding the comparative fit of alternative models, and the effect of model choice on applications such as model–based seasonal adjustment and use of time series methods in repeated survey estimation. As these are empirical questions, we attempt to address them here through comparing results from applying such alternative models to some Census Bureau time series.

KEY WORDS: ARIMA Model; Components Model; AIC; Seasonal Adjustment; Repeated Survey Estimation.

The authors would like to thank Larry Bobbitt, Brian Monsell, and Mark Otto of the Time Series Staff of the Statistical Research Division for assistance with the computations, programming, and graphs. This paper was presented at the Statistics Canada Symposium on Analysis of Data in Time in October 1989, held in Ottawa, Ontario. A condensed version of this paper is to appear in the proceedings of the symposium.

1. INTRODUCTION

The analysis of the components of time series has a long history (discussed in Nerlove, Grether, and Carvalho 1979), going back to work in astronomy, meteorology, and economics in the 17th through 19th centuries, and to early seasonal analysis by Buys–Ballot (1847). Empirical methods of seasonal adjustment were developed in the early part of this century leading utlimately to the development of the well–known X–11 method in 1967. As discussed in Bell and Hillmer (1984), these methods were developed in advance of adequate seasonal time series models, which have only become widely available and computationally feasible in the last 20 years or so.

This well–established interest in time series components has had important influences on time series modeling; in particular, it has led to two rather different approaches to modeling and model–based seasonal adjustment. For the autoregressive–integrated–moving average (ARIMA) models (Box and Jenkins 1976), several approaches to seasonal adjustment have been developed. The most successful of these, in our view, is the "canonical" approach of Burman (1980) and Hillmer and Tiao (1982). In contrast, a "component modeling" approach has developed that uses simple ARIMA models for seasonal, trend, irregular, etc. components. This approach is exemplified in the work of Akaike (1980), Gersch and Kitagawa (1983) and Kitagawa and Gersch (1984), and Harvey and Todd (1983) and Harvey (1985). Nerlove, Grether, and Carvalho (1979) suggested a somewhat different approach that appears not to have caught on, possibly because their ARIMA component models are too flexible to even assure that the model structure is identified (Hotta 1989), and because their treatment of nonstationarity (by polynomial detrending) is now viewed as inadequate.

While there has been considerable developmental work on both modeling appproaches, there is surprisingly little literature comparing results for the two different approaches. Harvey and Todd (1983) compared the forecast performance of their "basic structural model" (BSM) with that of ARIMA models fitted by Prothero and Wallis (1976) to six quarterly macroeconomic time series. Their results were rather inconclusive, also some of the ARIMA models used were of unusual form, featuring long lags in the seasonal operators. (In fairness, Prothero and Wallis' (1976) work was in the early stages of development of seasonal ARIMA modeling, before such refinements as exact maximum likelihood and outlier treatment were readily available.) Expanding the BSM, Harvey (1985) developed components models to explain cyclical behavior (with nonseasonal series) and gave some discussion of their relation to ARIMA models. Maravall (1985) observed that the BSM could yield an overall model close to Box and Jenkins (1976) popular ARIMA $(0,1,1)\times(0,1,1)_{12}$ "airline model," by showing that autocorrelations for the differenced series could be similar for the two models (depending on parameter values). This raised the important possibility that the BSM and certain ARIMA models could be about the same for some series. Carlin and Dempster (1989), in a detailed analysis of two series, found only small differences between canonical ARIMA seasonal adjustments and those from a fractionally–integrated–moving average (FRIMA) components model, and more major differences when comparing the FRIMA adjustment with the X–11 adjustment used in practice for another series.

The literature seems to leave two important questions unanswered, namely: (1) do ARIMA or components models provide a better fit to actual data or can available data even discriminate between them, and (2) how different are the results from ARIMA and components models in practical applications? The former question is one of statistical significance, the latter one of practical significance. Both questions are largely empirical, and an empirical investigation into them shall be the focus of this paper. In section 2 we describe the specific models we shall consider in detail, and use the AIC criterion of Akaike (1973) to compare the fit of ARIMA models and the BSM for a set of 45 seasonal time series. In general, AIC expresses a strong preference for ARIMA models.

Section 3 considers seasonal adjustment. Bell and Hillmer (1984) noted that component modelers have ignored the inherent uncertainty about seasonal–nonseasonal decompositions consistent with any given fitted model. To address this we consider the range of admissible decompositions consistent with a given components model, and obtain a "canonical decomposition" for component models in the same way this was done for ARIMA models by Burman (1980) and Hillmer and Tiao (1982). The canonical decomposition turns out to be trivially simple to obtain and very easy to use in signal extraction for seasonal adjustment. However, it also turns out to be very close to the original fitted components model for the series considered here, suggesting that seasonal adjustments for the original and canonical components models may typically be virtually identical. We then compare ARIMA model and BSM seasonal adjustments for two series and find negligible differences in signal extraction point estimates and proportionally large differences in signal extraction variances, though the signal extraction variances all seem small in an absolute sense.

In section 4 we investigate the effects of using ARIMA versus component models in applying time series signal extraction techniques to estimation for repeated surveys. This idea was originally suggested by Scott and Smith (1974) and Scott, Smith, and Jones (1977), but has seen intensive investigation more recently following theoretical and computational developments in estimation and signal extraction for nonstationary time series models. For the two series we consider the signal extraction point estimates using ARIMA models and the BSM are quite close, but for one series the signal extraction variances are quite different. Finally, in section 5 we draw some tentative conclusions.

2. ARIMA AND COMPONENTS MODELS

Let Y_t for $t=1,\ldots,n$ be observations on a time series, which will often be the logarithm of some original time series. We write

$$Y_t = \underset{\sim}{X}_t'\underset{\sim}{\beta} + Z_t \tag{2.1}$$

where $\underset{\sim}{X}_t'\underset{\sim}{\beta}$ is a linear regression mean function with $\underset{\sim}{X}_t$ the vector of regression variables at time t and $\underset{\sim}{\beta}$ the vector of regression parameters, and Z_t is the (zero mean) stochastic part of Y_t. The regression variables used here will be to account for trend constants, calendar variation, fixed seasonal effects, and outlier effects (Findley, et. al. 1988). We will be interested in decompositions of Z_t such as

$$Z_t = S_t + N_t = S_t + T_t + I_t \tag{2.2}$$

where S_t is a (stochastic) seasonal component, and N_t a (stochastic) nonseasonal component that can be further decomposed into a trend component T_t and an irregular component I_t. If Y_t is the logarithm of the time series of interest, note (2.1) and (2.2) imply multiplicative decompositions for the original time series.

One approach to analyzing time series components involves modeling Z_t directly, then making assumptions that lead from this model to definitions of and models for the components. The other approach is to directly specify models for the components, which then implies a model for Z_t that can be fitted to data. We shall consider ARIMA models as a basis for both approaches. While other models have certainly received attention in recent years (long memory, ARCH, and nonlinear models come to mind), ARIMA and ARIMA component models seem to have been the most popular, and so focusing attention on these two seems an appropriate starting point.

The ARIMA models we shall use for Z_t can be written in the form (c.f. Box and Jenkins 1976):

$$\phi(B)(1-B)^d(1-B^{12})Z_t = \theta(B)(1-\theta_{12}B^{12})a_t \tag{2.3}$$

where B is the backshift operator ($BZ_t = Z_{t-1}$), $d \geq 0$ (if d=0, $(1-B)^d = 1$), $\phi(B) = 1-\phi_1 B-\ldots-\phi_p B^p$ and $\theta(B) = 1-\theta_1 B-\ldots-\theta_q B^q$ are AR and MA operators of low order (usually p, q $\leq$ 3), and a_t is white noise (iid $N(0,\sigma_a^2)$.) This model is for monthly seasonal data; the modifications for data with other seasonal periods (e.g. quarterly) are obvious, and the $1-B^{12}$ and $1-\theta_{12}B^{12}$ are removed for nonseasonal data. We could have included a seasonal autoregressive operator in (2.3), though we rarely use these. If $\theta_{12} = 1$ we can "cancel" the $1-B^{12}$ factor on both sides of (2.3) and add seasonal mean variables to $\underset{\sim}{X}_t$ (Abraham and Box 1978, Bell 1987). Identification, estimation, and diagnostic checking of these models proceeds with by now well-established procedures — see Box and Jenkins (1976) for pure ARIMA models, Bell and Hillmer (1983) and Findley et al. (1988) for models with regression terms. Estimation is by maximum likelihood where the likelihood function is defined as the joint density of the differenced data $(1-B)^d(1-B^{12})Y_t$ for $t=d+13,\ldots,n$.

Component models specify simple ARIMA models for the components in (2.2). Harvey and Todd's (1983) basic structural model (BSM) can be written

$$
\begin{aligned}
Z_t &= S_t + T_t + I_t \\
U(B)\,S_t &= \epsilon_{1t} \qquad \epsilon_{1t} \sim \text{iid } N(0,\sigma_1^2) \\
(1-B)^2 T_t &= (1-\eta B)\epsilon_{2t} \qquad \epsilon_{2t} \sim \text{iid } N(0,\sigma_2^2) \\
I_t &\sim \text{iid } N(0,\sigma_3^2)
\end{aligned}
\tag{2.4}
$$

where $U(B) = 1 + B \ldots + B^{11}$ sums a series over 12 consecutive months. They actually begin with T_t following a random walk with stochastic drift, where the drift also follows a random walk; this leads to the (0,2,1) model for T_t in (2.4) with the constraint $\eta \geq 0$. While we shall not enforce this constraint, it turns out to be easily satisfied for all our example series here. If the "stochastic" drift has zero innovation variance (i.e. it is actually a constant) then $\eta = 1$ and the model for T_t reduces to $(1-B)T_t = \beta_0 + \epsilon_{1t}$, and we can account for β_0 by adding the time trend variable t to $\underset{\sim}{X}_t$. If $\sigma_1^2 = 0$ then S_t becomes fixed and can be handled with appropriate variables in $\underset{\sim}{X}_t$ analogous to what was noted when $\theta_{12} = 1$ in the ARIMA model (2.3).

Gersch and Kitagawa (1983) (see also Kitagawa and Gersch 1984) consider models similar to (2.4), but with T_t following the model

$$(1-B)^{\delta}T_t = \epsilon_{2t} \qquad \delta = 1,2, \text{ or } 3. \tag{2.5}$$

We whall refer to (2.4) but with T_t following (2.5) as the GK model. Notice that the GK model with $\delta = 2$ becomes the BSM with $\eta = 0$, while the BSM with η=1 is the GK with δ=1 and a trend constant. Akaike (1980) suggested similar models, but with S_t following a model that now seems unattractive.

Gersch and Kitagawa extend their model with the addition of a stationary autoregressive component. This can be written as

$$\begin{gathered} Z_t = S_t + T_t + I_t + V_t \\ (1-\alpha_1 B - \dots - \alpha_p B^p)V_t = \epsilon_{4t} \qquad \epsilon_{4t} \sim \text{iid } N(0,\sigma_4^2) \end{gathered} \tag{2.6}$$

with S_t and I_t as in (2.4), and T_t as in (2.5). Harvey (1985) also considers such an extension to his models, with the autoregressive parameters constrained so that V_t tends to exhibit cyclical behavior. He also considers an ARMA(2,1) formulation for V_t.

Modeling procedures for these component models are more automatic than for ARIMA models and are discussed in the references cited. Estimation is again by maximum likelihood, with the likelihood evaluated using the Kalman filter. Since the models are nonstationary this presents problems for initialization of the Kalman filter that have been recently addressed by Kohn and Ansley (1986) and Bell and Hillmer (1990). These approaches produce a likelihood function that is again the joint density of the differenced data, which is now determined by the components models.

The ARIMA models for the components imply an ARIMA model for the aggregate Z_t, as has been observed by G. C. Tiao (reported in Findley 1983) and Maravall (1985). Taking (2.4) for illustration, applying $(1-B)^2U(B) = (1-B)(1-B^{12})$ to Z_t gives $(1-B)^2\epsilon_{1t} + U(B)(1-\eta B)\epsilon_{2t} + (1-B)(1-B^{12})\epsilon_{3t}$, which follows a moving average model of order 13 whose parameters are determined by σ_1^2, σ_2^2, σ_3^2, and η. While (2.4) is thus equivalent to an ARIMA$(0,1,13)\times(0,1,0)_{12}$ model for Z_t, the high regular MA order and the constraints on the parameters make it unlikely that direct ARIMA modeling of Z_t would yield such a model exactly. Thus, there is potential for difference between the ARIMA and component model approaches, though Maravall (1985) notes that certain parameter values for (2.4) can yield a model close to the popular ARIMA$(0,1,1)\times(0,1,1)_{12}$ "airline model" of Box and Jenkins (1976). For nonseasonal series or series whose seasonality is modeled as fixed through the regression function $\underline{X}_t'\underline{\beta}$, the ARIMA model implied by (2.4) for $Z_t = T_t + I_t$ depends on $(1-\eta B)\epsilon_{2t} + (1-B)^2 I_t$, which follows an MA(2) model whose 3 parameters are determined by σ_2^2, σ_3^2, and η. We could easily get exactly the same model by direct modeling of Z_t as ARIMA (0,2,2). Similar results obtain for other nonseasonal components models. While the potential for difference between nonseasonal ARIMA and components models is difficult to judge, the potential for ARIMA and components models to be effectively the same seems greater in the nonseasonal than in the seasonal case.

This discussion raises questions about how much ARIMA and components models will differ in practice, and which will fit better when they do differ? We will make a preliminary investigation into this by comparing the fit of ARIMA and components models on a set of time series. As the models we wish to compare are generally nonnested (one is not obtained by simple constraints on the parameters of the other) traditional hypothesis tests or confidence intervals would be difficult to apply. We shall use the AIC criterion of Akaike (1973), which is defined as

$$\mathrm{AIC} = -2\hat{L} + 2m$$

where $\hat{L}$ is the maximized log–likelihood and m is the number of parameters estimated. The model with the smaller AIC is to be preferred. To compare two models, 1 and 2 say, we present the difference in their AIC's, DAIC = $\mathrm{AIC}_1 - \mathrm{AIC}_2$. A positive value of DAIC favors model 2, a negative value model 1. Judging when there is a "significant" difference between models as measured by DAIC is not necessarily straightforward (see Findley 1988), but users of AIC often view differences of 1 or 2 as significant. We shall use 2 as a rough significance boundary. As a crude justification, notice that if we add a parameter to a model $\hat{L}$ cannot decrease, so if the parameter yields no improvement in fit, $\hat{L}$ remains the same and AIC increases by 2.

We shall use AIC to compare the fit of ARIMA and components models on a set of Census Bureau seasonal time series analyzed by Burman and Otto (1988) using ARIMA models. (Many were analyzed previously in Hillmer, Bell, and Tiao (1983), though with fewer years of data available. We also include one series, labelled ENM20, from the U.S. Bureau of Labor Statistics, analyzed in Bell and Hillmer 1984.) These series have the advantage of having readily available ARIMA models with careful treatment of regression terms for calendar variation, fixed seasonal effects (occasionally), and outliers. We exclude a few series Burman and Otto (1988) analyzed that are not published, as well as the foreign trade series they analyzed since these have undergone significant revisions in recent years to correct some major data problems. This leaves 45 series for analysis which are listed in Table 1. The series are broadly representative of the series seasonally adjusted by the Census Bureau, but are not a random sample, so the analysis here might be best viewed as a pilot study.

For a given series we shall use the same regression terms with both ARIMA and components models, and also will restrict comparisons to models with the same order of differencing. Comparing models with different orders of differencing poses some problems since the likelihood functions for the two models are then based on different (differenced) data. This restriction means that we will compare ARIMA models (2.3) with d=1 to the BSM as in (2.4). ARIMA models with d=0 will be compared to a model as in (2.4), but with T_t following (2.5) with δ=1. Models with a fixed seasonal and d=1 in the ARIMA structure will be compared to a components model with a fixed seasonal (no stochastic S_t), and with T_t again following (2.5) with δ=1. The latter two cases correspond to particular cases of both the BSM and GK models. When the ARIMA model has d=1 and a stochastic seasonal, we shall not make comparisons with the GK model that would use (2.5) with δ=2. As a special case of (2.4) with η=0, at best this GK model would avoid one extraneous parameter and have an AIC 2 less than that of (2.4). At worst, it can have a substantially higher AIC than (2.4) if the maximum likelihood estimator $\hat{\eta}$ is not near 0 (though if $\hat{\eta} \approx 1$ we can think of (2.4) as overdifferencing the GK model with δ=1.)

The ARIMA models used and their AICs, the fitted BSMs and their AICs, and the AIC differences are given in Table 2. The table below provides a summary. The results are obvious: AIC exhibits a strong preference for ARIMA models overall, with large AIC differences (> 8) for about one half of the series. DAIC's for the two series for which the BSM was preferred were only –2.1 and –2.7.

BSM versus ARIMA: Number of Series by Intervals of DAIC

DAIC range	Order of Differencing (1,1)	(0,1)	(1,0)
< -2	2	0	0
−2 to 2	6	1	0
2 to 8	9	2	3
8 to 20	10	3	2
20 to 40	5	0	1
> 40	4	0	0
	36	6	6

Three series appear twice in the table since they were refit with fixed seasonals after first getting $\hat{\theta}_{12} \approx 1$.

In looking for possible explanations for the poor fit of the BSM we examined DAICs and corresponding $\hat{\theta}_{12}$'s, $\hat{\eta}$'s, etc., but found no obvious patterns. Selection bias was considered as a possible explanation, even though the ARIMA models were selected with the usual identification approach based on autocorrelations and partial autocorrelations, and not by searching a set of models for the model with minimum AIC. To check for selection bias, the BSM AICs were compared with those for the ARIMA$(0,1,1)\times(0,1,1)_{12}$ "airline model", which seems a reasonable choice if one were to use a single ARIMA model. The results are given in Table 3. Although the BSM fit much better than the airline model for two series (DAICs of −11.7 and −25.6), aside from this the results changed little from those in Table 1. This is perhaps not surprising since 15 of the selected ARIMA models were airline models, and others were not very different from the airline model. The airline model performed much better in comparison to the selected ARIMA models than the BSM, though seven series favored the selected ARIMA model over the airline model by an AIC greater than 8, suggesting that use of any single model for all series will occasionally lead to poor fits.

This report would not be complete without some comments on our experience fitting components models. The results presented here were obtained using a computer program for fitting time series models with ARIMA components and regression terms recently developed by ourselves, other members of the Time Series Staff of the Statistical Research Division at Census, and Steven Hillmer of the University of Kansas. We found the components models much more difficult to fit than regular ARIMA models. For example, getting good starting values for nonlinear iteration over the component model parameters (something not addressed in the literature, to our knowledge) seems important, whereas we find getting good starting values for ARIMA model parameters not at all important. We have not presented results for models with a fourth component as in (2.6) because we were unable to successfully fit such models. For every series adding a fourth component caused the nonlinear search to go outside the stationarity region for V_t, causing the program to crash. While there are means of programming around this problem, and while inclusion of a fourth component might improve the fits, we found these difficulties discouraging. Though we did not make a formal study of the numerical problems we experienced with components models, they seemed due to the likelihood being rather flat in certain directions in the parameter space. Given this, we find the oft–claimed advantages of "simplicity" and "interpretability" for components models difficult to accept.

The computational difficulties we experienced suggest a final possible explanation for our results – that there is something wrong with our software and it is not actually maximizing the likelihood. While we have checked our program thoroughly, and do not believe this to be the case, we cannot rule this out with certainty. We will gladly provide our data to anyone interested in checking our results. We would be even more interested in seeing a study done with other series to see if similar results are obtained .

3. SEASONAL ADJUSTMENT

While section 2 suggests that ARIMA models may fit a time series substantially better than components models, there is still the question of what difference choice of a model makes in practice? Here we consider the effect of model choice on seasonal adjustment. For a given components model, seasonal adjustment can be done by applying a Kalman smoother to the series (see, e.g., Gersch and Kitagawa 1983). With ARIMA models one must first make sufficient assumptions leading from simple ARIMA models for observed series to unique component models. This is addressed by Burman (1980) and Hillmer and Tiao (1982), who consider a range of possible decompositions and suggest a choice leading to a unique decomposition into component models. (The two approaches differ some for certain models that do not seem to occur often.) The underlying assumptions are set out and discussed further by Bell and Hillmer (1984). As will be seen shortly, we can also consider a range of decompositions for any given components model.

For Y_t following (2.1) and (2.3), Burman (1980) and Hillmer and Tiao (1982) achieve a decomposition of form (2.2) by making a partial fractions decomposition of the covariance generating function (CGF), $\gamma_Z(B)$, of Z_t, yielding CGF's $\gamma_S(B)$, $\gamma_T(B)$, and $\gamma_I(B)$, and corresponding ARIMA models for the components. This yields a range of admissible decompositions corresponding to $\gamma_Z(B) = [\gamma_S(B) - \gamma_1] + [\gamma_T(B) - \gamma_2] + [\gamma_I(B) + \gamma_1 + \gamma_2]$, for any γ_1 and γ_2 such that each bracketed term is ≥ 0 for all $B = e^{i\lambda}$. The range reflects inherent uncertainty about the decomposition; specifying γ_1 and γ_2 yields a particular decomposition that can be used for seasonal adjustment. Burman (1980) and Hillmer and Tiao (1982) suggest picking the maximum possible γ_1 and γ_2 ($\bar{\gamma}_1 = \min_\lambda \gamma_S(e^{i\lambda})$ and $\bar{\gamma}_2 = \min_\lambda \gamma_T(e^{i\lambda})$), leading to what is called the canonical decomposition, which has several attractive properties. Focusing in particular on the seasonal–nonseasonal decomposition now, the components corresponding to any admissible γ_1 can be written as $S_t = \bar{S}_t + \nu_t$ and $N_t = \bar{N}_t - \nu_t$, where $\bar{S}_t$ and $\bar{N}_t$ are the canonical seasonal and nonseasonal, and ν_t is white noise with variance $\bar{\gamma}_1 - \gamma_1$. Thus, the canonical decomposition can be viewed as removing as much white noise as possible from the seasonal component and putting it in the nonseasonal through the irregular. Since there is no apparent reason to include additional white noise in the seasonal, this is a good argument for using the canonical decomposition. (Watson (1987) gives an approach that avoids assuming a particular decomposition.)

(As an aside, we note that it is also necessary to decompose the deterministic regression effects, $X_t'\beta$, into seasonal and nonseasonal parts. This is discussed in Bell (1984), but since there is no reason to do this differently for ARIMA and components models we need not go into it here.)

Bell and Hillmer (1984) criticize component modelers for simply taking the component models for adjustment as those obtained in modeling the observed series, and thus ignoring the uncertainty inherent in the basic decomposition into components. We can address this decomposition uncertainty for component models by defining a "canonical decomposition" in an analogous way to that defined for ARIMA models — subtracting as much white noise as possible from S_t and adding it to N_t through I_t. (A canonical trend for components models is also discussed in Appendix A.2.) We show in Appendix A.1 that the resulting canonical components model decomposition, $Z_t = \bar{S}_t + \bar{N}_t = \bar{S}_t + [T_t + \bar{I}_t]$, has a canonical irregular $\bar{I}_t$ with variance $\bar{\sigma}_3^2 = \sigma_3^2 + \sigma_1^2/144$, and a canonical seasonal $\bar{S}_t$ that follows the model

$$U(B)\bar{S}_t = \psi(B)\bar{\epsilon}_{1t} \qquad \bar{\epsilon}_{1t} \sim \text{iid } N(0,\bar{\sigma}_1^2) \tag{3.1}$$

where $\psi(B)$, of order 11, is given in Table 5., and $\bar{\sigma}_1^2 = .8081\ \sigma_1^2$. This is in fact the same form as the canonical seasonal model of Burman (1980) and Hillmer and Tiao (1982), though their seasonal model will generally have a different $\psi(B)$ and $\bar{\sigma}_1^2$ (that depend on the ARIMA model). As with ARIMA models, using any other admissible decomposition (corresponding to any valid decomposition of the covariance generating function), including that defined by the original fitted components model, can be viewed as adding white noise to the canonical seasonal $\bar{S}_t$. Notice that, given a components model, the model for $\bar{S}_t$ in (3.1) is trivial to obtain. Also, signal extraction for canonical seasonal adjustment may be performed in the usual way with a Kalman smoother using the model (3.1) for $\bar{S}_t$ and increasing the irregular variance to $\bar{\sigma}_3^2$.

Notice that the amount of variance removed from the components model seasonal, $\sigma_1^2/144$, will be small unless σ_1^2 is large relative to σ_2^2 and σ_3^2. A quick glance at Table 2 reveals the opposite to be true for the series considered here: σ_1^2 is generally quite small relative to $\sigma_2^2 + \sigma_3^2$. This has two implications: (1) the estimated component model typically implies a very nearly fixed seasonal, and (2) the original component model decomposition will often be very close to the canonical component model decomposition. In fact, for the examples we have tried, seasonal adjustments from the original and canonical component model decompositions have been virtually identical. Since this aspect of decomposition choice appears to make little difference we shall not consider it further here. This is not to say choosing some other decomposition than the canonical cannot have important effects, though we shall not consider that here either.

To examine potential differences in seasonal adjustments arising from model choice we examine seasonal adjustments for two series: IHAPVS (value of U.S. household appliances shipped from 1/62–12/81), and ENM20 (thousands of employed males 20 and older in nonagricultural industries from 1/65 – 8/79), a series analyzed by Bell and Hillmer (1984). IHAPVS was one of the series which the BSM fit best (DAIC = –.7), while the BSM fit for ENM20 was rather poor (DAIC = 13.7), though far from the worst. ENM20 was the one series for which logarithms were not taken so an additive decomposition is used here.

Figure 1.a. shows the estimated ARIMA and BSM seasonal components for IHAPVS. Close inspection is required to detect any difference. As this is also true of the seasonal adjustments we do not present these. Figure 1.b. shows the signal extraction standard deviations for IHAPVS expressed as coefficients of variation. Here substantial differences appear with the ARIMA CV's being 20 percent or more higher near the end of the series. (This does not mean that the results for the ARIMA model are necessarily bad.) However, the CV's might all be considered small: none exceed about 1.6 percent.

Figure 2.a shows the ARIMA and BSM seasonals for ENM20. Here we can see a difference: the ARIMA seasonal evolves steadily over time while the BSM seasonal remains relatively fixed. (Notice from Table 2 for the BSM that for ENM20 $\hat{\sigma}_1^2 = 27$ while $\hat{\sigma}_2^2 = 16{,}500$.) Figure 2.b portrays seasonal adjustment results for the last 5 years of the data. While differences can be seen they may not be important since the month–to–month changes themselves are not large, seldom exceeding .5 percent. Figure 2.c. shows even larger differences for signal extraction standard deviations than we saw for IHAPVS. The BSM standard deviations rise very little at the end of the series because an essentially fixed seasonal is being estimated. Still, the most noteworthy aspect of Figure 2.c. may be how small the standard deviations are relative to series values of 40,000 to 50,000.

We conjecture that $\text{Var}(S_t - \hat{S}_t) \longrightarrow 0$ as $\theta_{12} \longrightarrow 1$ in the ARIMA model and as $\sigma_1^2 \longrightarrow 0$ in the BSM, which probably explains the small signal extraction standard deviations observed in the two examples. However, if we decide $\theta_{12} = 1$ or $\sigma_1^2 = 0$ and use a model with fixed seasonal regression effects instead, the signal extraction variances will not be 0 since we will have error in estimating the seasonal regression parameters. A curious aspect of these results is the apparent discontinuity between results for $\theta_1 < 1$ (or $\sigma_1^2 > 0$) and $\theta_{12} = 1$ (or $\sigma_1^2 = 0$).

4. REPEATED SURVEY ESTIMATION

Scott and Smith (1974) and Scott, Smith and Jones (1977) suggested using time series signal extraction techniques for estimation in periodic surveys. If s_t denotes the true population quantity (the signal) and e_t the sampling error at time t, then we use signal extraction to estimate s_t in

$$Y_t = s_t + e_t , \qquad (4.1)$$

If Y_t is the logarithm of the original series, then $\exp(s_t)$ and $\exp(e_t)$ are the true population quantity and multiplicative sampling error in the original series. Any of the models discussed in section 2 can be used for s_t; Binder and Dick (1989) and Bell and Hillmer (1989) use ARIMA models, while Pfefferman (1989) uses a BSM. Generally, any regression terms in the model are also part of s_t.

Model building for the survey estimation problem is discussed in the references cited above. A primary distinction between this application and what we have considered before, is that the model for e_t can be estimated, in some fashion, using survey microdata. The sampling error model is then held fixed when estimating the parameters of the s_t model using the time series data on Y_t. Questions arise about the sensitivity of the survey estimation results to any of the aspects of the modeling. Here we shall examine the sensitivity of results to the choice between an ARIMA model and a BSM for s_t.

We consider two time series. For the first, U.S. teenage unemployment (in 1000's) from 1/72 to 12/83, Bell and Hillmer (1987) develop the following model for $Y_t = s_t + e_t$:

$$(1-B)(1-B^{12})s_t = (1-.27B)(1-.68B^{12})a_t \qquad \sigma_a^2 = 4294$$

$$e_t = h_t\tilde{e}_t \quad (1-.6B)\tilde{e}_t = (1-.3B)c_t \quad \sigma_c^2 = .8767 \quad h_t^2 = -.0000153\, Y_t^2 + 1.971\, Y_t$$

The model for s_t has been reestimated, yielding slightly different parameter values than those reported in Bell and Hillmer (1987). With $\sigma_c^2 = .8767$, $\mathrm{Var}(\tilde{e}_t) = 1$, so h_t is the (estimated) sampling error standard deviation, which is time–varying. The modeling of the second series we consider here, U.S. 5 or more unit housing starts, is very similar to that for U.S. single family housing starts, also considered in Bell and Hillmer (1987). The sampling errors for 5 or more unit housing starts appear approximately uncorrelated over time with relative variance .00729, which is the approximate variance of the logged multiplicative sampling errors. The estimated ARIMA model for the signal in the logged time series is

$$(1-B)(1-B^{12})s_t = (1 - .47B)(1 - .89B^{12})a_t \qquad \sigma_a^2 = .0215.$$

We used the above models in signal extraction estimation of s_t, and then did the same with a BSM fitted for s_t with the same e_t models given above. The BSM model fitted relatively well for both these series, with DAIC = AIC(BSM) – AIC(ARIMA) = – 3.1 for teenage unemployment and DAIC = 1.8 for housing starts. (The appropriateness of these AIC comparisons is in some question since the e_t models are not fitted with the time series data.) Figure 3.a. shows the signal extraction point estimates for teenage unemployment using both models; $(1-B^{12})\hat{s}_t$ is shown to avoid the obscuring effects of seasonality. The BSM estimates less variance in the signal than the ARIMA model, and thus yields slightly smoother estimates. Figure 3.b. shows substantial differences in the signal extraction variances for the two models. The two signal extraction estimates for the housing starts series were virtually identical, and so are not shown. Figure 4 shows 10^6 times the signal extraction variances for the logged series, for the last half of the housing starts series — those for the first half would be a mirror image.

While there are some interesting differences in pattern, the magnitude of the differences is small.

5. CONCLUSIONS

Even the conclusions drawn in section 2 must be somewhat tentative; it would be interesting to see similar studies with other sets of time series. Because of the limited examples considered in sections 3 and 4, the conclusions there can only be suggestive. To summarize:

1. Data can frequently discriminate between ARIMA and components models. For the 45 series analyzed, AIC showed a strong general preference for ARIMA models over the BSM. To the extent that model fit is important, merely assuming the BSM provides an adequate fit could be dangerous.
2. We found fitting components models more difficult than fitting ARIMA models. While we would have liked to see if the addition of a stationary AR component or other cycle term could improve the component model fits, we were unable to fit such models due to numerical problems.
3. Signal extraction point estimates for seasonal adjustment and survey estimation using ARIMA models and using the BSM differed little for the examples considered. Signal extraction variances showed much larger differences, though for the seasonal adjustment examples the variances using both models might be regarded as quite small. This last point is worth more investigation, to see if model–based seasonal adjustment variances with canonical, or approximately canonical, decompositions are typically very small.

APPENDIX

A.1 Canonical Seasonal for Components Models

We obtain a "canonical" decomposition $Z_t = \bar{S}_t + T_t + \bar{I}_t$, starting from $Z_t = S_t + T_t + I_t$ with $U(B)S_t = \epsilon_{1t}$ ~ iid $N(0,\sigma_1^2)$ and I_t ~ iid $N(0,\sigma_3^2)$ as in (2.4). Here we only work with the model for S_t, so there are actually no restrictions on the model for T_t or on the white noise I_t (the possibility $I_t = 0$ ($\sigma_3^2 = 0$) is allowed). Following Burman (1980) and Hillmer and Tiao (1982), the canonical decomposition is obtained by subtracting as much white noise as possible from S_t (to get $\bar{S}_t$) and adding this to I_t (to get $\bar{I}_t$). In terms of covariance generating functions (CGF's), we consider

$$\gamma_Z(B) = \left[\frac{\sigma_1^2}{U(B)U(F)} - \gamma_1 \right] + \gamma_T(B) + (\sigma_3^2 + \gamma_1)$$

for any γ_1 such that this is a valid decomposition (all 3 terms ≥ 0 for all $B = e^{i\lambda}$). The minimum γ_1 is $-\sigma_3^2$ (though if $\min_\lambda \gamma_T(e^{i\lambda}) > 0$ we can combine $T_t + I_t$ and pick a smaller γ_1) and the maximum (canonical) γ_1 is

$$\bar{\gamma}_1 = \min_\lambda (\sigma_1^2 / |U(e^{i\lambda})|^2) = \sigma_1^2 / (\max_\lambda |U(e^{i\lambda})|^2).$$

Since $|U(e^{i\lambda})|^2 = (e^{11i\lambda} + e^{-11i\lambda}) + 2(e^{10i\lambda} + e^{-10i\lambda}) + \ldots + 11(e^{i\lambda} + e^{-i\lambda}) + 12 = 2\cos(11\lambda) + 4\cos(10\lambda) + \ldots + 22\cos(\lambda) + 12$ is maximized at $\lambda = 0$ with $|U(1)|^2 = 144$, we have $\bar{\gamma}_1 = \sigma_1^2/144$. (This has previously been observed by Maravall 1985, p. 354.) The CGF for $\bar{S}_t$ is then

$$\frac{\sigma_1^2}{U(B)U(F)} - \bar{\gamma}_1 = \frac{\sigma_1^2[1 - (1/144)U(B)U(F)]}{U(B)U(F)}$$

The numerator is the CGF of an MA(11) model, thus the model for $\bar{S}_t$ is

$$U(B)\bar{S}_t = \psi(B)\bar{\epsilon}_{1t} \qquad \bar{\epsilon}_{1t} \sim \text{iid } N(0,\bar{\sigma}_1^2) \tag{A.1}$$

with $\psi(B) = 1 - \psi_1 B - \ldots - \psi_{11}B^{11}$ and $\bar{\sigma}_1^2$ determined by $\bar{\sigma}_1^2\ \psi(B)\psi(F) = \sigma_1^2[1 - (1/144)U(B)U(F)]$. Letting $\bar{\sigma}_1^2 = \sigma_1^2\alpha$, we can determine $\psi(B)$ and α to satisfy

$$(144\alpha)\psi(B)\psi(F) = 144 - U(B)U(F)$$
$$= 132 - 11(B+F) - 10(B^2+F^2) - \ldots - (B^{11}+F^{11}) .$$

Notice that α and $\psi(B)$ will not depend on any of the model parameters. This gives us the important result that the canonical seasonal $\bar{S}_t$ corresponding to any components model (2.4) follows (A.1) with the same $\psi(B)$ and with $\bar{\sigma}_1^2 = \sigma_1^2\alpha$. Thus, given α and $\psi(B)$, which we determine next, the model for the canonical seasonal is trivial to obtain.

To find α and $\psi(B)$ we first compute the zeros of the polynomial $p(x) = x^{22} + 2x^{21} + \ldots + 11x^{12} - 132x^{11} + 11x^{10} + \ldots + 2x + 1 = -(144\alpha)\psi(x)[\psi(x^{-1})x^{11}]$. There is a repeated real root of 1 since $\gamma_S(B) - \bar{\gamma}_1 = 0$ for B=1, and 5 sets of complex roots $\xi_k = a_k \pm b_k i$ and their reciprocals. We want the roots on or outside the unit circle $(|\xi_k|^2 \geq 1)$ for $\psi_1(B)$ — these are given in Table 4. These were computed using the POLYROOT command of the GAUSS programming language (Edlefsen and Jones 1986). We then computed $p_1(x) = (x-1)\prod_{k=1}^{5}(x-\xi_k)(x-\bar{\xi}_k) = (x-1)\prod_{k=1}^{5}(x^2-2a_kx + |\xi_k|^2)$ using the POLYMULT command of GAUSS, getting the coefficients of

$$\psi(B) = (1-B)\prod_{k=1}^{5}(1-\xi_k^{-1}B)(1-\bar{\xi}_k^{-1}B) = -p_1(B)/(\prod_{k=1}^{5}|\xi_k|^2)$$

These are given in Table 5. To find α we equate coefficients of B^0 in $(144\alpha)\psi(B)\psi(F) = -F^{11}p(B) = 132 - 11(B+F) - \ldots - (B^{11}+F^{11})$, giving

$$\alpha = 132/[144(1+\psi_1^2 + \ldots + \psi_{11}^2)] = .808118 .$$

(To check the results we used POLYMULT to compute $-(144\alpha)\psi(x)[\psi(x^{-1})x^{11}] = -(144\alpha)\psi(x)[x^{11} - \psi_1 x^{10} - \ldots - \psi_{10}x - \psi_{11}]$ and got back p(x) as desired, exact to the eight digits printed.) Using $\psi(B)$ from Table 5. and $\bar{\sigma}_1^2 = .808118\sigma_1^2$ completes the specification of model (A.1).

A.2 Canonical Trend for Components Models

A "canonical" trend, $\bar{T}_t$, corresponding to $Z_t = S_t + T_t + I_t$, can also be defined. Maravall (1985, p. 353) has considered this for the case where T_t follows the BSM model (2.4). Here we consider the GK model (2.5) for $\delta = 1$ or 2. In general, $\gamma_T(e^{i\lambda}) = \sigma_2^2/(|1-e^{i\lambda}|^2)^d = \sigma_2^2/[2(1-\cos(\lambda))]^d$ is minimized at $\lambda = \pi$ with value $\bar{\gamma}_2^{(d)} = \sigma_2^2/4^d$; thus, $\bar{\gamma}_2^{(1)} = \sigma_2^2/4$ and $\bar{\gamma}_2^{(2)} = \sigma_2^2/16$. Since $\gamma_T(B) - \bar{\gamma}_2^{(d)} = [\sigma_2^2 - (\sigma_2^2/4^d)(1-B)^d(1-F)^d]/(1-B)^d(1-F)^d$, the model for $\bar{T}_t$ is $(1-B)^d\bar{T}_t = \omega(B)\bar{\epsilon}_{2t}$ with $\omega(B) = 1-\omega_1 B - \ldots -\omega_d B^d$ and $\bar{\epsilon}_{2t} \sim$ iid $N(0,\bar{\sigma}_2^2)$. Since $\gamma_T(e^{i\lambda}) - \bar{\gamma}_2^{(d)} = 0$ for $\lambda = \pi$, $\omega(B)$ has a root of $e^{i\pi} = -1$, implying a factor $(1+B)$. So for d=1, $\omega(B) = 1+B$, and for d=2, $\omega(B) = (1-\omega B)(1+B)$. Also, $\bar{\sigma}_2^2\ \omega(B)\omega(F) = \sigma_2^2[1-4^{-d}(1-B)^d(1-F)^d]$ gives

$$(d{=}1):\ \bar{\sigma}_2^2(1+B)(1+F) = \sigma_2^2[1-(1-B)(1-F)/4] = (\sigma_2^2/4)(1+B)(1+F)$$

$$(d{=}2):\ \bar{\sigma}_2^2(1+B)(1+F)(1-\omega B)(1-\omega F) = \sigma_2^2[1-(1-B)^2(1-F)^2/16]$$

For d=1 we see $\bar{\sigma}_2^2 = \sigma_2^2/4$, and for d=2 one can verify that $\omega = 3-2\sqrt{2} \approx .1716$ and $\bar{\sigma}_2^2 = \sigma_2^2/(16\omega) \approx .36428\ \sigma_2^2$. Thus, the canonical trend models are as follows:

$$(d{=}1)\qquad (1-B)\bar{T}_t = (1+B)\bar{\epsilon}_{2t} \qquad \bar{\sigma}_2^2 = \sigma_2^2/4$$

$$(d{=}2)\qquad (1-B)^2\bar{T}_t = (1-.1716B)(1+B)\bar{\epsilon}_{2t} \qquad \bar{\sigma}_2^2 = .36428\ \sigma_2^2\ .$$

REFERENCES

Abraham, B. and Box, G.E.P. (1978), "Deterministic and Forecast–Adaptive Time–Dependent Models," Applied Statistics, 27, 120–130.

Akaike, H. (1973) "Information Theory and an Extension of the Likelihood Principle," in the 2nd International Symposium on Information Theory, eds. B. N. Petrov and F. Czaki, Budapest: Akademia Kiado, 267–287.

________ (1980), "Seasonal Adjustment by a Bayesian Modeling," Journal of Time Series Analysis, 1, 1–13.

Bell, W. R. (1984) "Seasonal Decomposition of Deterministic Effects," Research Report Number 84/01, Statistical Research Division, Bureau of the Census.

Bell, W. R. (1987) "A Note on Overdifferencing and the Equivalence of Seasonal Time Series Models With Monthly Means and Models With $(0,1,1)_{12}$ Seasonal Parts When $\Theta = 1$," Journal of Business and Economic Statistics, 5, 383–387.

Bell, W. R. and Hillmer, S. C. (1983), "Modeling Time Series with Calendar Variation," Journal of the American Statistical Association, 78, 526–534.

________ (1984), "Issues Involved with the Seasonal Adjustment of Economic Time Series," (with discussion), Journal of Business and Economic Statistics, 2, 291–320.

________ (1987), "Time Series Methods for Survey Estimation," Research Report Number 87/20, Statistical Research Division, Bureau of the Census.

________ (1989), "Modeling Time Series Subject to Sampling Error," Research Report Number 89/01, Statistical Research Division, Bureau of the Census.

________ (1990), "Initializing the Kalman Filter for Nonstationary Time Series Models," Journal of Time Series Analysis, to appear.

Binder, D. A. and Dick, J. P. (1989), "Modelling and Estimation for Repeated Surveys," Survey Methodology, 14, to appear.

Box, G.E.P. and Jenkins, G. M. (1976), Time Series Analysis: ForecastingandControl, San Francisco: Holden Day.

Burman, J. P. (1980), "Seasonal Adjustment by Signal Extraction," Journal of the Royal Statistical Society Series A, 143, 321–337.

Burman, J. P. and Otto, M. (1988), "Outliers in Time Series," Research Report Number 88/14, Statistical Research Division, Bureau of the Census.

Buys Ballot, C. H. D. (1847) Les Changements Periodiques de Temperature, Utrecht: Kemink et Fils.

Carlin, J. B., and Dempster, A. P. (1989) "Sensitivity Analysis of Seasonal Adjustments: Empirical Case Studies," Journal of the American Statistical Association, 84, 6–20.

Edlefsen, L. E. and Jones, S. D. (1986) GAUSS Programming Language Manual, Seattle: Aptech Systems, Inc.

Findley, D. F. (1983), "Comments on 'Comparative Study of the X–11 and BAYSEA Procedures of Seasonal Adjustment' by H. Akaike and M. Ishiguro," in Applied Time Series Analysis of Economic Data, ed. Arnold Zellner, Washington, D.C.: U. S. Department of Commerce, Bureau of the Census.

________ (1988), "Comparing Not Necessarily Nested Models With the Minimum AIC and the Maximum Kullback–Leibler Entropy Criteria: New Properties and Connections," Research Report Number 88/21, Statistical Research Division, Bureau of the Census.

Findley, D. F., Monsell, B. M., Otto, M. C., Bell, W. R., and Pugh, M. G. (1988) "Toward X–12 ARIMA," Proceedings of the Fourth Annual Research Conference, U. S. Department of Commerce, Bureau of the Census.

Gersch, W. and Kitagawa, G. (1983), "The Prediction of Time Series With Trends and Seasonalities," Journal of Business and Economic Statistics, 1, 253–264.

Harvey, A. C. (1985), "Trends and Cycles in Macroeconomic Time Series," Journal of Business and Economic Statistics, 3, 216–227.

Harvey, A. C. and Todd, P. H. J. (1983), "Forecasting Economic Time Series With Structural and Box–Jenkins Models: A Case Study," (with discussion), Journal of Business and Economic Statistics, 1, 299–315.

Hillmer, S. C., Bell, W. R., and Tiao, G. C. (1983), "Modeling Considerations in the Seasonal Adjustment of Economic Time Series," in Applied Time Series Analysis of Economic Data, ed. Arnold Zellner, U.S. Department of Commerce, Bureau of the Census, 74–100.

Hillmer, S. C., and Tiao, G. C. (1982), "An ARIMA–Model–Based Approach to Seasonal Adjustment," Journal of the American Statistical Association, 77, 63–70.

Hotta, L. K. (1989), "Identification of Unobserved Components Models," Journal of Time Series Analysis, 10, 259–270.

Kitagawa, G. and Gersch, W. (1984), "A Smoothness Priors–State Space Modeling of Time Series With Trend and Seasonality," Journal of the American Statistical Association, 79, 378–389.

Kohn, R. and Ansley, C. F. (1986), "Estimation, Prediction, and Interpolation for ARIMA Models With Missing Data," Journal of the American Statistical Association, 81, 751–761.

Maravall, A. (1985), "On Structural Time Series Models and the Characterization of Components," Journal of Business and Economic Statistics, 3, 350–355.

Nerlove, M., Grether, D. M., and Caravallo, J. L. (1979), Analysis of Economic Time Series: A Synthesis, New York: Academic Press.

Pfeffermann, D. (1989) "Estimation and Seasonal Adjustment of Population Means Using Data from Repeated Surveys," paper presented at the annual meeting of the American Statistical Association, Washington, D. C.

Prothero, D. L. and Wallis, K. F. (1976) "Modeling Macroeconomic Time Series," Journal of the Royal Statistical Society Series A, 139, 468–500.

Scott, A. J. and Smith, T.M.F. (1974), "Analysis of Repeated Surveys Using Time Series Methods," Journal of the American Statistical Association, 69, 674–678.

Scott, A. J., Smith, T.M.F., and Jones, R. G. (1977), "The Application of Time Series Methods to the Analysis of Repeated Surveys," International Statistical Review, 45, 13–28.

Watson, M. W. (1987) "Uncertainty in Model–Based Seasonal Adjustment Procedures and Construction of Minimax Filters," Journal of the American Statistical Association, 82, 395–408.

Table 1: Series and ARIMA Models (with Regression Variables) Used in the Study

Series	Years	Selected ARIMA[#] Model	Outliers[*]	Series Description
bapprs	67-83	(010)(011)+TD	6/72	Retail sales of household appliance stores
bautrs	67-82	(110)(011)+TD	3/75	Retail sales of automotive dealers (total)
belgws	67-83	(011)(011)+TD	-	Wholesale sales of electrical goods
bfrnws	67-82	(011)(011)+TD	12/77,12/78,2/79,1/80	Wholesale sales of furniture and home furnishings
bgasrs	67-82	(011)(011)+TD	-	Retail sales of gasoline stations
bgrcrs	67-82	(013)(011)+TD+E	12/70,1/72,4/75,12/77	Retail sales of grocery stores
bgrcws	67-83	(013)(011)+TD	-	Wholesale sales of groceries and related products
bhdwws	67-83	(011)(011)+TD	-	Wholesale sales of hardware, plumbing, heating equipment, and supplies
blqrrs	67-83	(012)(011)+TD	-	Retail sales of liquor stores
bshors	67-83	(011)(011)+TD+E	12/69,1/70	Retail sales of shoe stores
bvarrs	67-83	(013)(011)+TD+E	4/67,4/76 (level shift)	Retail sales of variety stores
bwaprs	67-83	(012)(011)+TD+E	8/73	Retail sales of women's clothing stores
c1ftbp	64-83	(011)(011)+TD	2/66,1/70,12/70,12/78, 3/79	Total 1 family dwelling building permits
c24tbp	64-83	(011)(011)+TD	3/75,8/75,6/78,4/80	Total 2 to 4 unit building permits
c5ptbp	64-83	(013)(011)	12/74	Total 5+ unit building permits
caopvp	64-83	(310)(011)	4/69,8/70,7/77	Value put in place, all other private residences
cnetbp	64-83	(011)(011)+TD	1/67,12/74,3/77,1/82	Total Northeast building permits
cwsths	64-83	(013)(011)	-	Total West housing starts

enm20	65-79	(110)(011)	-	Employed males 20 and older in nonagricultural industries (Bureau of Labor Statistics
iapevs	68-83	(011)(011)	12/71,3/73,1/78	Value shipped of aircraft parts and equipment
ibevti	62-81	(012)(011)	-	Total inventories of beverages
ibevvs	62-81	(014)(011)+TD	-	Value shipped of beverages
icmeti	68-84	(310)(011)	1/69,9/69,11/82	Total inventories of communications equipment
icmevs	68-83	(210)(011)	8/74,12/75,12/76,8/83 12/83	Value shipped of communications equipment
ifatti	62-81	(014)(011)	9/71,10/73,10/74,9/77, 9/78,8/80	Total inventories of fats and oils
ifatvs	62-81	(011)(011)	10/73,8/74	Value shipped of fats and oils
ifmeti	62-81	(210)(011)	10/71,10/75,11/75,10/76, 11/76,2/77,11/77,2/78, 12/78,11/81	Total inventories of farm machinery and equipment
ifrtvs	62-81	(011)(011)	4/78,5/78,1/79,4/79,5/79	Value of fertilizer shipped
iglcvs	62-81	(012)(011)	4/65,2/68,3/68,12/70, 12/74,7/75,3/76,3/77, 8/77	Value of glass containers shipped
ihapti	62-81	(012)(011)	8/66,1/72,4/80	Total inventories of household appliances
ihapvs	62-81	(011)(011)	-	Value of household appliances shipped
inewuo	64-83	(011)(011)	9/68,3/82	Unfilled newspaper, periodical, and magazine orders
irrevs	62-81	(011)(011)	-	Value of railroad equipment shipped
itobvs	64-81	(013)(011)	11/75,10/77,6/79,10/79	Value of tobacco shipped
itvrti	64-83	(011)(011)	1/69,1/76	Total television and radio inventories
itvrvs	62-81	(012)(011)	4/67	Value of televisions and radios shipped

bdptrs	67-83	(101)(011)+TD+E	-	Retail sales of department stores
bfrnrs	67-82	(101)(011)+TD	-	Retail sales of furniture stores
bmncrs	67-83	(101)(011)+TD+E	-	Retail sales of men's and boys' clothing stores
cnctbp	64-83	(100)(011)+TD	12/64,1/65,1/79	Total North Central building permits
cncths	64-83	(101)(011)	2/64,1/73,1/75,1/77, 1/79,2/79	Total North Central housing starts
cneths	64-83	(101)(011)	1/65,2/75,2/78,2/80	Total Northeast housing starts
csoths	64-83	(012)(000)	1/77	Total South housing starts
ifmevs	64-83	(019)(000)	10/69,11/70,10/73,10/76, 1/82	Value shipped of farm machinery and equipment
iglcti	62-81	(013)(000)	3/65,12/70,8/73,7/77	Total inventories of glass containers

Notes: [#]The numbers in parentheses refer to the AR, differencing, and MA orders of the nonseasonal and (monthly) seasonal parts of the ARIMA model. Thus, (011)(011) denotes an airline model. (000) indicates a fixed seasonal modeled with regression terms. "TD" indicates trading-day variables are included in the model, and "E" indicates an Easter holiday variable is included -- see Bell and Hillmer (1983).

[*]Outliers are handled as discussed in Hillmer, Bell, and Tiao (1983). Only additive outliers are used here, resulting in inclusion in the model of an indicator regression variable for the given month. The one exception is the 4/76 level shift outlier indicated for bvarrs, that was discovered in Hillmer, Bell, and Tiao (1983).

Table 2.: AIC Comparisons of BSM and ARIMA Models

	ARIMA Model†					BSM			
Series	Model	$\hat{\theta}_{12}$	AIC	$\hat{\sigma}_1^2$	$\hat{\eta}$	$\hat{\sigma}_2^2$	$\hat{\sigma}_3^2$	AIC	DAIC#
bapprs	(010)(011)	.68	1942.3	23.0	1.00	630	192	1948.6	6.3
bautrs	(110)(011)	.69	1912.9	34.1	1.00	1720	1.5	1927.1	14.2
belgws	(011)(011)	.81	1868.7	9.9	.99#	568	65.9	1871.1	2.4
bfrnws	(011)(011)	.67	1784.1	16.2	1.00	647	90.3	1799.3	15.2
bgasrs	(011)(011)	.79	1540.8	1.2	.90	226	.002	1544.0	3.2
bgrcrs	(013)(011)	.97*	1406.8	.23	.43	2.42	64.0	1406.2	-.6
bgrcws	(013)(011)	.96*	1729.1	.30	.99	141	147	1734.4	5.3
bhdwws	(011)(011)	.72	1867.3	14.0	1.00	525	67.8	1865.2	-2.1
blqrrs	(012)(011)	.75	1722.7	3.52	1.00	140	120	1726.7	4.0
bshors	(011)(011)	.70	1972.6	29.4	1.00#	688	218	1972.8	.2
bvarrs	(013)(011)	.75	1822.8	13.5	.99	120	266	1823.9	1.1
bwaprs	(012)(011)	.65	1858.7	27.4	1.00	326	163	1871.8	13.1
c1ftbp	(011)(011)	.99*	2581.2	.016	.95	4329	3.25	2605.7	24.5
c24tbp	(011)(011)	.89	2823.2	.816	1.00#	8654	1426	2826.4	3.2
c5ptbp	(013)(011)	.80	2903.2	143	.99#	11075	2886	2911.2	8.0
caopvp	(310)(011)	.89	2643.5	.014	1.00	7318	.105	2710.5	67.0
cnetbp	(011)(011)	.82	2897.1	37.3	1.00#	8310	4162	2905.3	8.2
cwsths	(013)(011)	.85	2907.4	68.8	1.00#	10714	3480	2909.9	2.5
enm20	(110)(011)	.77	2073.6	27.8	.89	16488	.024	2087.3	13.7
iapevs	(011)(011)	.80	2081.2	101	.98	1885	1443	2079.9	-1.3
ibevti	(012)(011)	.65	1843.1	1.79	1.00	177.5	0.0	1890.0	46.9
ibevvs	(014)(011)	.78	2300.9	8.46	.99	529	390	2320.0	19.1
icmeti	(310)(011)	.76	1535.8	1.63	.63	95.2	12.4	1542.3	6.5
icmevs	(210)(011)	.74	1856.5	24.8	1.00#	568	410	1865.6	9.1
ifatti	(014)(011)	.64	2553.4	23.1	1.00#	4017	.49	2599.5	46.1
ifatvs	(011)(011)	.82	2409.4	26.0	.99#	1477	99.0	2406.7	-2.7
ifmeti	(210)(011)	.77	1907.6	.009	.73	225	.004	1934.8	27.2
ifrtvs	(011)(011)	.31	2752.4	1024	1.00#	4279	7.81	2774.4	22.0
iglcvs	(012)(011)	.84	2510.1	7.43	1.00	1305	1134	2538.4	28.3
ihapti	(012)(011)	.82	1991.0	.291	1.00	356	.281	2028.3	37.3
ihapvs	(011)(011)	.82	2398.0	17.4	1.00#	1012	379	2397.3	-.7
inewuo	(011)(011)	.60	2402.6	32.0	.98	1960	.115	2447.6	45.0
irrevs	(011)(011)	.83	2671.2	79.8	.95	2287	1918	2671.0	-.2
itobvs	(013)(011)	.71	2117.7	50.9	.96#	90.1	1030	2132.9	15.2
itvrti	(011)(011)	.82	2226.7	.003	.87	958	.101	2244.9	18.2
itvrvs	(012)(011)	.69	2683.7	282	1.00	2807	1496	2699.7	16.0
bdptrs	(101)(011)	.53	1723.8	29.8	--	104	68.3	1729.6	5.8
bfrnrs	(101)(011)	.66	1698.3	15.7	--	316	59.1	1698.4	0.1
bmncrs	(101)(011)	.57	1959.0	40.1	--	644	199	1976.3	17.3
cnctbp	(100)(011)	.79	2872.9	18.0	--	13870	15.9	2889.3	16.4
cncths	(101)(011)	.81	3008.7	24.6	--	13784	6862	3019.8	11.1
cneths	(101)(011)	.88	3088.5	44.3	--	6125	21141	3092.9	4.4

csoths	(012)(000)	--	2940.7	--	--	5126	3953	2947.0	6.3
ifmevs	(019)(000)	--	2718.3	--	--	4049	269	2730.2	11.9
iglcti	(013)(000)	--	2211.1	--	--	542	.087	2217.7	6.6
bgrcrs	(013)(000)	--*	1473.1	--	--	34.2	46.0	1485.1	12.0
bgrcws	(013)(000)	--*	1809.8	--	--	129	145	1813.5	3.7
clftbp	(011)(000)	--*	2688.0	--	--	3999	.777	2710.0	22.0

Notes:

† The numbers in parentheses give the nonseasonal and seasonal parts of the ARIMA model, with (000) used to denote a fixed seasonal modeled with regression terms. Other regression terms in the models are given in Table 1.

DAIC = BSM(AIC) - ARIMA(AIC). If DAIC > 0 the ARIMA model is preferred.

* $\hat{\theta}_{12} \approx 1$, so model refit with fixed seasonal.

Table 3.: Airline Model versus Selected ARIMA Model and BSM

	Selected ARIMA†		Airline Model				BSM	
Series	Model	AIC	$\hat{\theta}_1$	$\hat{\theta}_{12}$	AIC	DAIC1*	AIC	DAIC2#
bapprs	(010)(011)	1942.3	.18	.68	1941.7	-.6	1948.6	6.9
bautrs	(110)(011)	1912.9	-.19	.69	1911.6	-1.3	1927.1	15.5
belgws	(011)(011)	1868.7	.18	--	--	0	1871.1	2.4
bfrnws	(011)(011)	1784.1	.24	--	--	0	1799.3	15.2
bgasrs	(011)(011)	1540.8	-.13	--	--	0	1544.0	3.2
bgrcrs	(013)(011)	1406.8	.43	.99#	1417.9	11.1	1406.2	-11.7
bgrcws	(013)(011)	1729.1	.40	.96#	1732.4	3.3	1734.4	2.0
bhdwws	(011)(011)	1867.3	.16	--	--	0	1865.2	-2.1
blqrrs	(012)(011)	1722.7	.42	.75	1721.5	-1.2	1726.7	5.2
bshors	(011)(011)	1972.6	.31	--	--	0	1972.8	.2
bvarrs	(013)(011)	1822.8	.57	.75	1823.4	.6	1823.9	.5
bwaprs	(012)(011)	1858.7	.33	.61	1863.2	4.5	1871.8	8.6
c1ftbp	(011)(011)	2581.2	-.27	--	--	0	2605.7	24.5
c24tbp	(011)(011)	2823.2	.12	--	--	0	2826.4	3.2
c5ptbp	(013)(011)	2903.2	.22	.80	2908.6	5.4	2911.2	2.6
caopvp	(310)(011)	2643.5	-.22	.90	2689.8	46.3	2710.5	20.7
cnetbp	(011)(011)	2897.1	.32	--	--	0	2905.3	8.2
cwsths	(013)(011)	2907.4	.24	.86	2907.7	.3	2909.9	2.2
enm20	(110)(011)	2073.6	-.21	.77	2074.3	.7	2087.3	13.0
iapevs	(011)(011)	2081.2	.39	--	--	0	2079.9	-1.3
ibevti	(012)(011)	1843.1	-.11	.67	1843.8	.7	1890.0	56.2
ibevvs	(014)(011)	2300.9	.42	.77	2315.1	14.2	2320.0	4.9
icmeti	(310)(011)	1535.8	-.17	.87	1567.9	32.1	1542.3	-25.6
icmevs	(210)(011)	1856.5	.39	.73	1861.1	4.6	1865.6	4.5
ifatti	(014)(011)	2553.4	-.27	.65	2555.4	2.0	2599.5	44.1
ifatvs	(011)(011)	2409.4	.14	--	--	0	2406.7	-2.7
ifmeti	(210)(011)	1907.6	-.30	.79	1930.8	23.2	1934.8	4.0
ifrtvs	(011)(011)	2752.4	.27	--	--	0	2774.4	22.0
iglcvs	(012)(011)	2510.1	.50	.96#	2535.6	25.5	2538.4	2.8
ihapti	(012)(011)	1991.0	-.28	.80	2001.5	10.5	2028.3	26.8
ihapvs	(011)(011)	2398.0	.26	--	--	0	2397.3	-.7
inewuo	(011)(011)	2402.6	-.22	--	--	0	2447.6	45.0
irrevs	(011)(011)	2671.2	.36	--	--	0	2671.0	-.2
itobvs	(013)(011)	2117.7	.75	.72	2124.6	6.9	2132.9	8.3
itvrti	(011)(011)	2226.7	-.24	--	--	0	2244.9	18.2
itvrvs	(012)(011)	2683.7	.39	.67	2687.0	3.3	2699.7	12.7
bdptrs	(101)(011)	--	.44	.54	1720.6	--	1723.9	3.3
bfrnrs	(101)(011)	--	.23	.67	1689.1	--	1691.1	2.0
bmncrs	(101)(011)	--	.40	.59	1952.7	--	1968.3	15.6
cnctbp	(100)(011)	--	.02	.80	2866.4	--	2878.5	12.1
cncths	(101)(011)	--	.30	.81	3000.6	--	3009.2	8.6
cneths	(101)(011)	--	.59	.87	3080.7	--	3083.1	2.4
csoths	(012)(000)	--	.34	1.00#	2826.1	--	2831.8	5.7
ifmevs	(019)(000)	--	.05	.84	2617.1	--	2619.6	2.5
iglcti	(013)(000)	--	-.16	.95	2127.6	--	2135.7	8.1

† Selected ARIMA Models without a regular or without a seasonal difference cannot be compared to the Airline Model by AIC. For these series only the Airline Model and BSM as in (2.4) are compared.

* DAIC1 = Airline Model AIC - Selected ARIMA Model AIC

DAIC2 = BSM AIC - Airline Model AIC

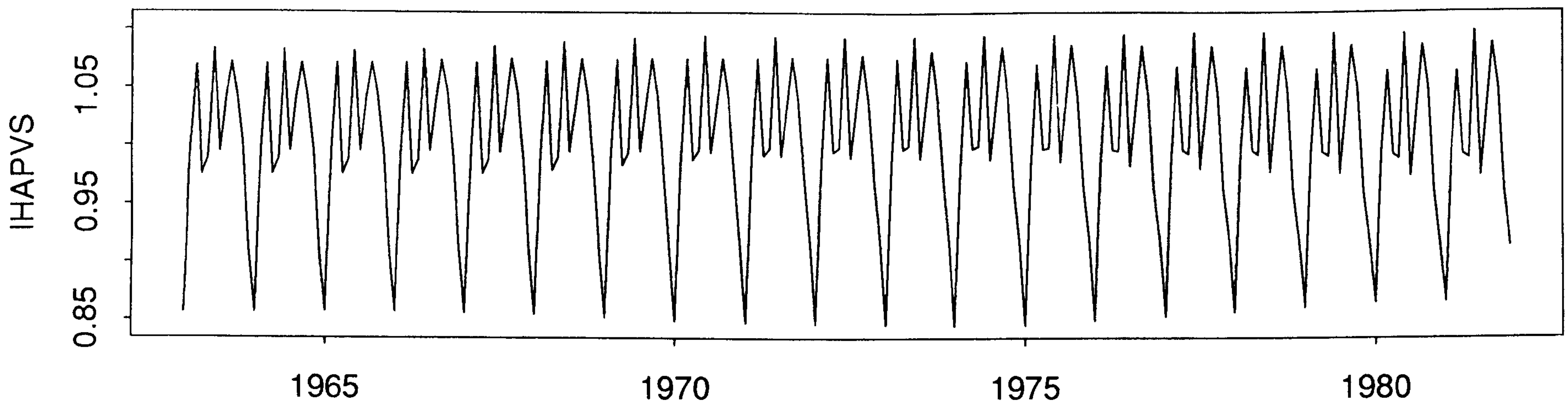
Table 4: $\xi_k = a_k \pm b_k i$ = zeros of $\psi(B)$

k	a_k	b_k	$\lvert\xi_k\rvert = \sqrt{a_k^2 + b_k^2}$
1	−1.6448	.4511	1.7055
2	−1.1702	1.2133	1.6856
3	−.3744	1.6002	1.6434
4	.4820	1.4965	1.5722
5	1.1046	.9429	1.4523
6	1	0	1

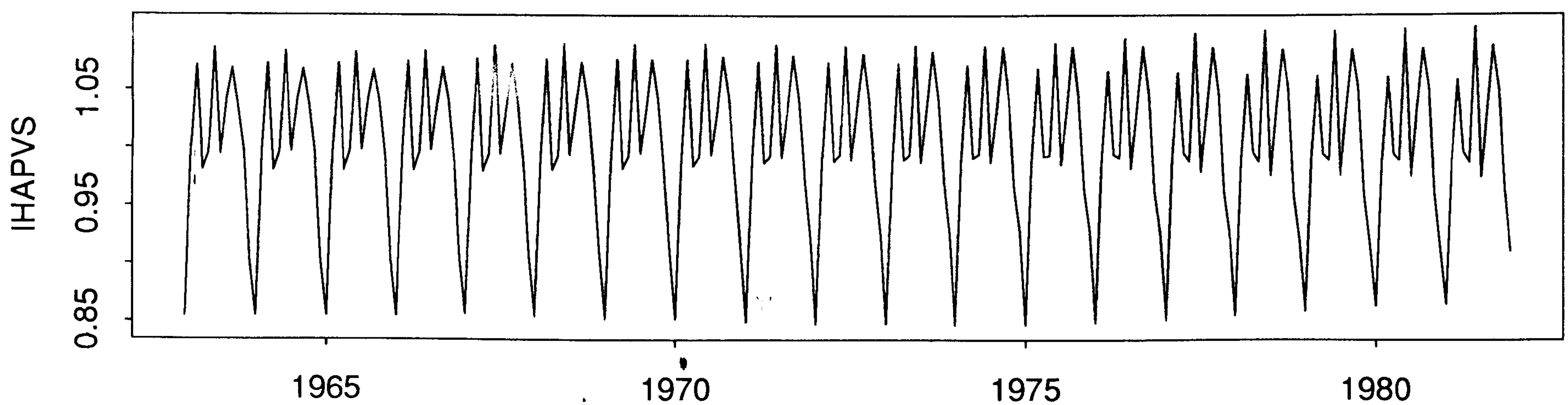
Table 5: Coefficients ψ_k for $\psi(B) = 1-\psi_1 - \dots - \psi_{11}B^{11}$

k	ψ_k	k	ψ_k
1	.205555	7	.061661
2	.175919	8	.045395
3	.148557	9	.031188
4	.123471	10	.018953
5	.100648	11	.008593
6	.080059		

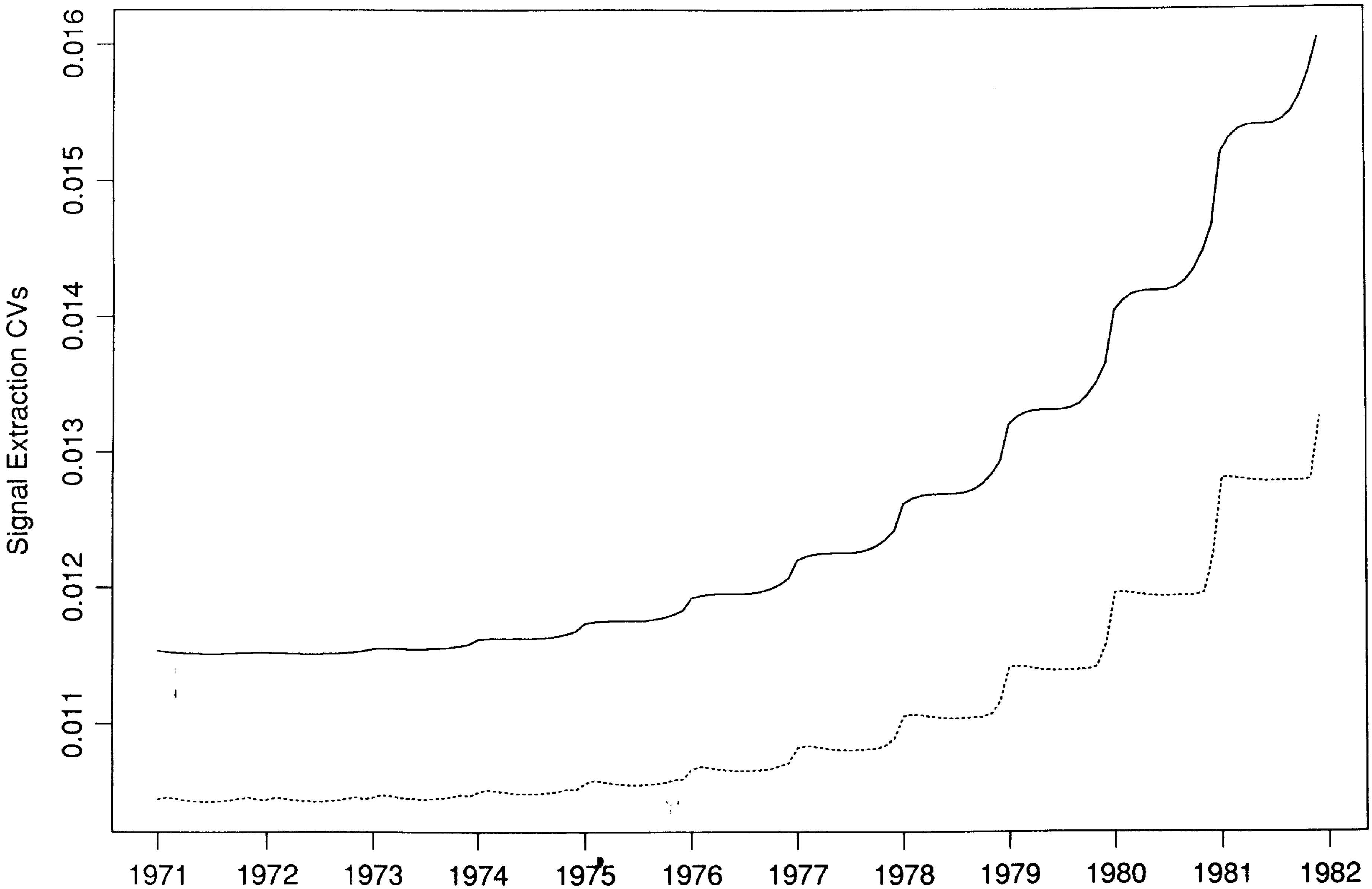
ARIMA Canonical Seasonal



BSM Seasonal

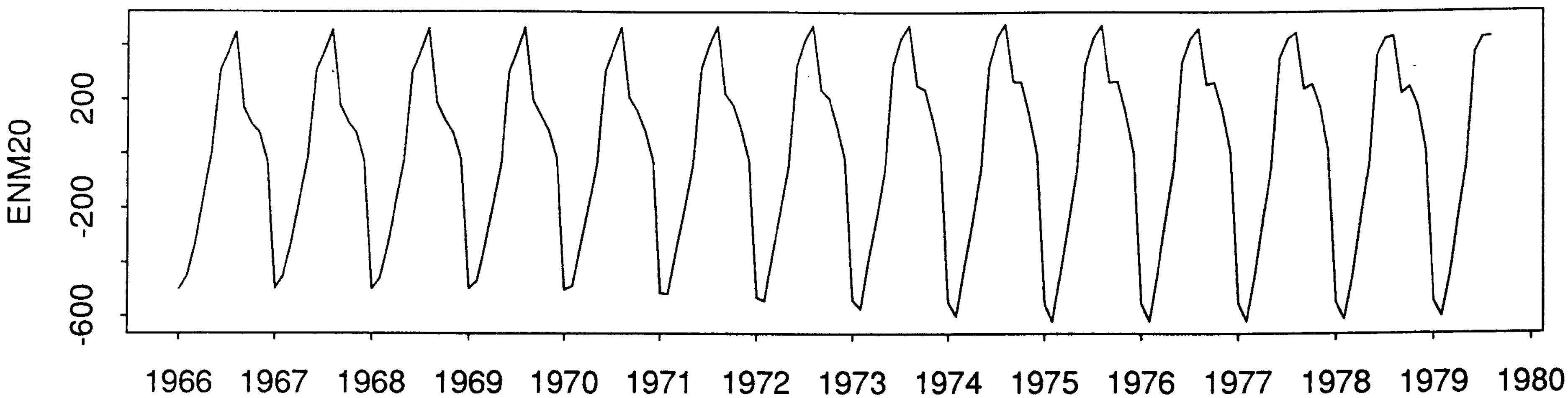


IHAPVS, Signal Extraction CV of Seasonally Adjusted Data (After 1971)

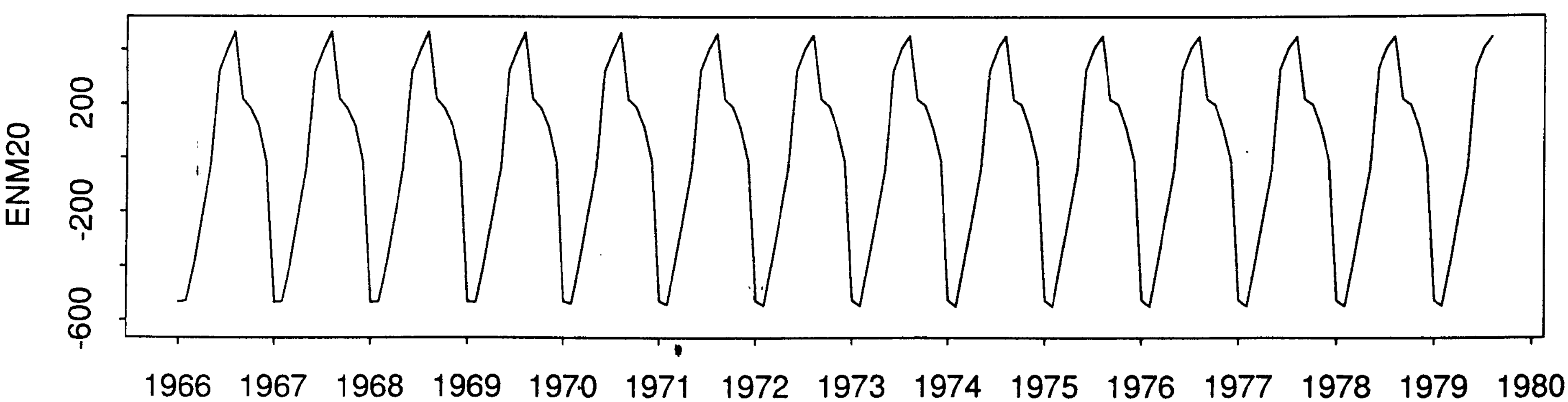


Solid line = Canonical ARIMA Seasonal Adjustment
Dash line = BSM Seasonal Adjustment

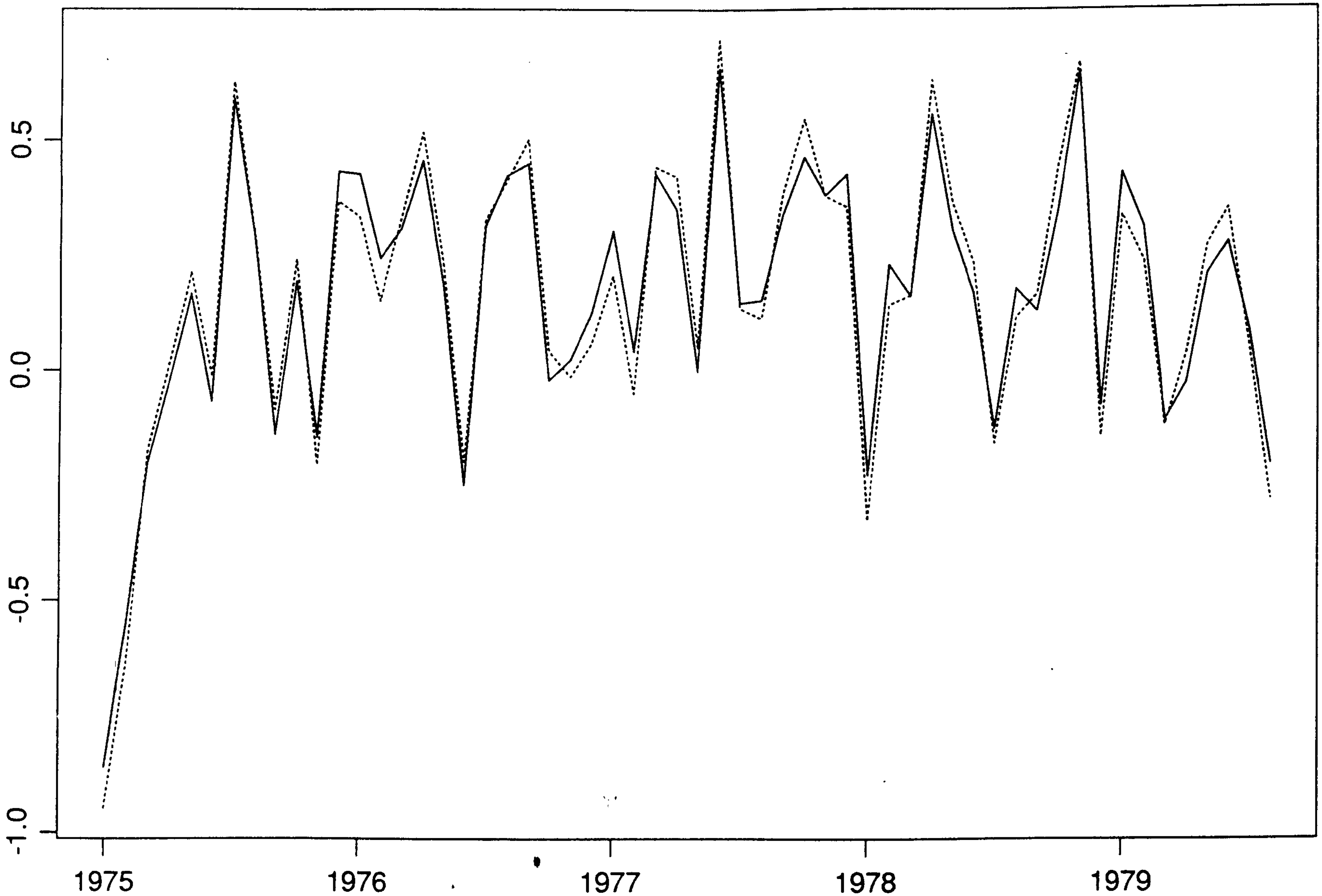
ARIMA Canonical Seasonal



BSM Seasonal

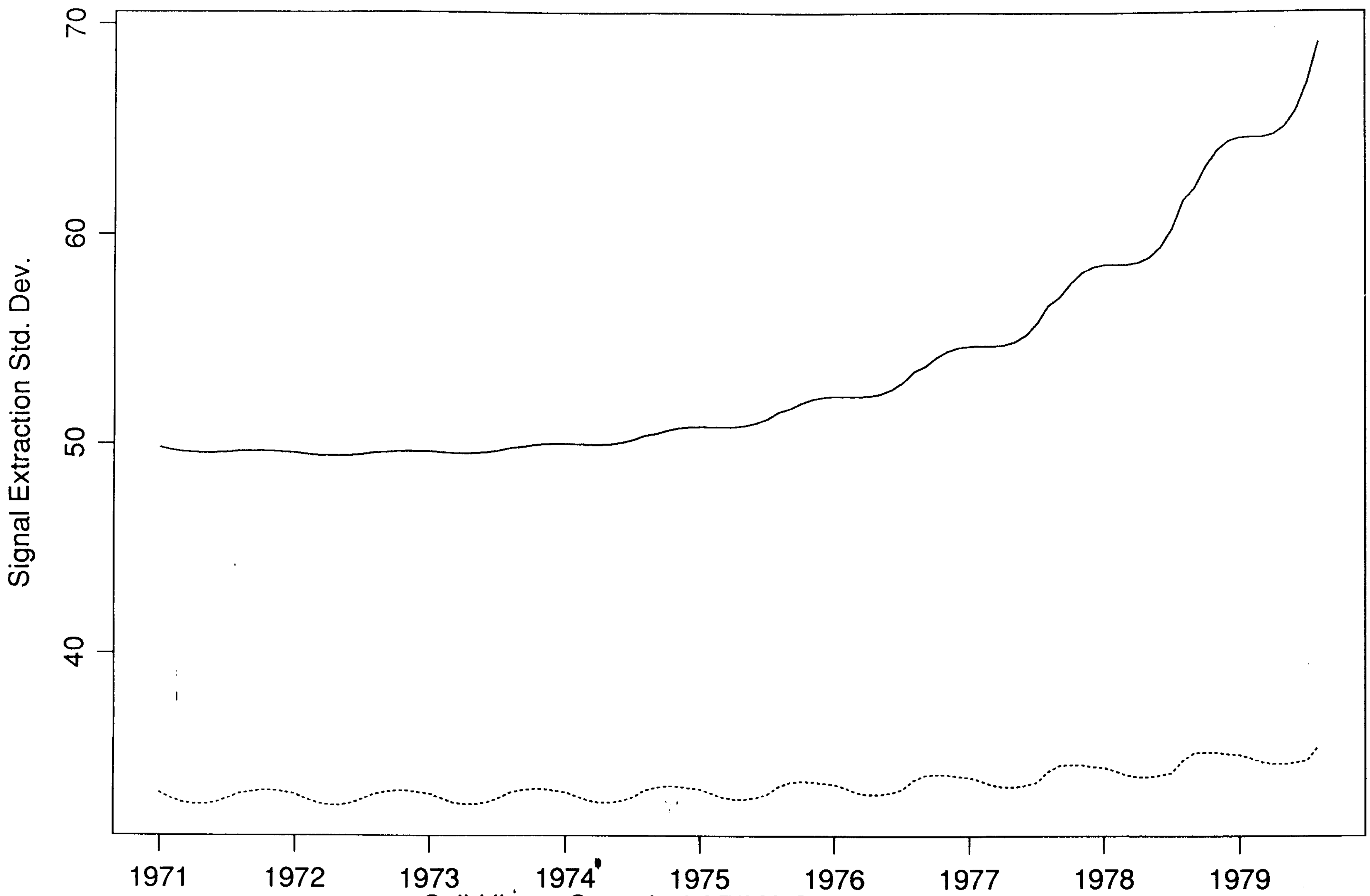


Month-to-Month Percent Changes, Seasonally Adjusted Data (After 1975)



Solid line = Canonical Seasonal Adjustment
Dash line = BSM Seasonal Adjustment

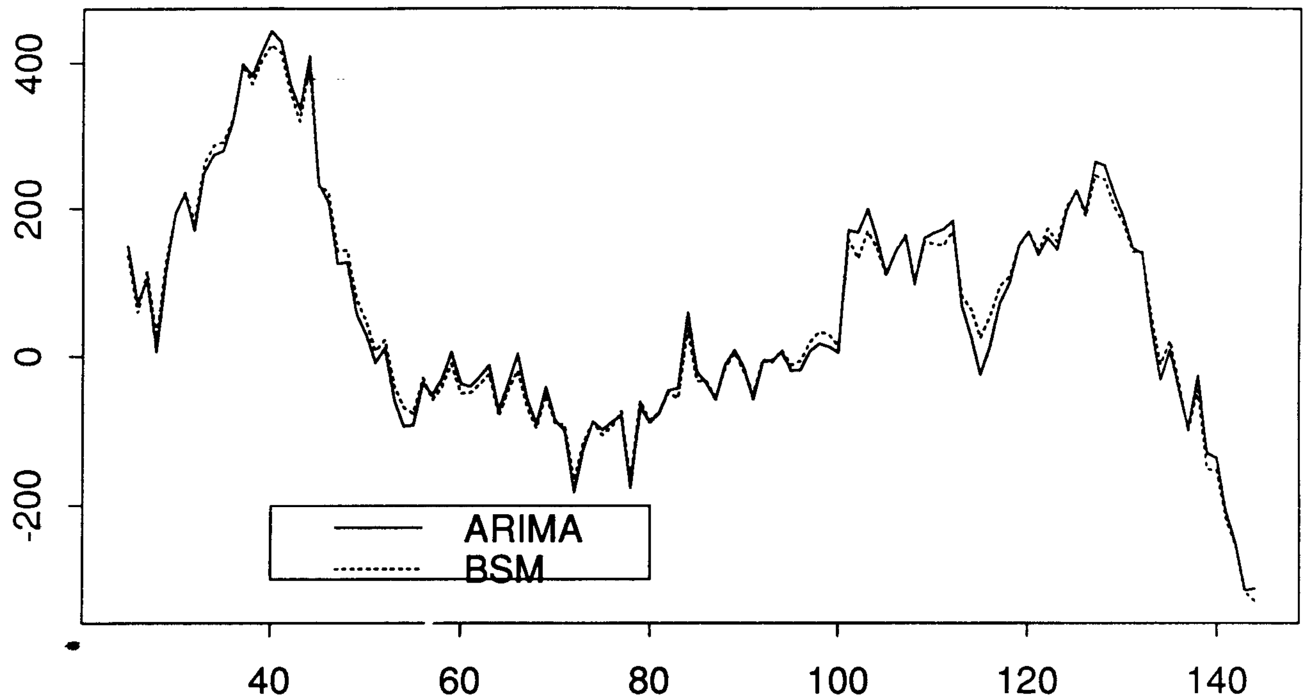
ENM20, Signal Extraction Std. Dev. of Seasonally Adjusted Data (After 1971)



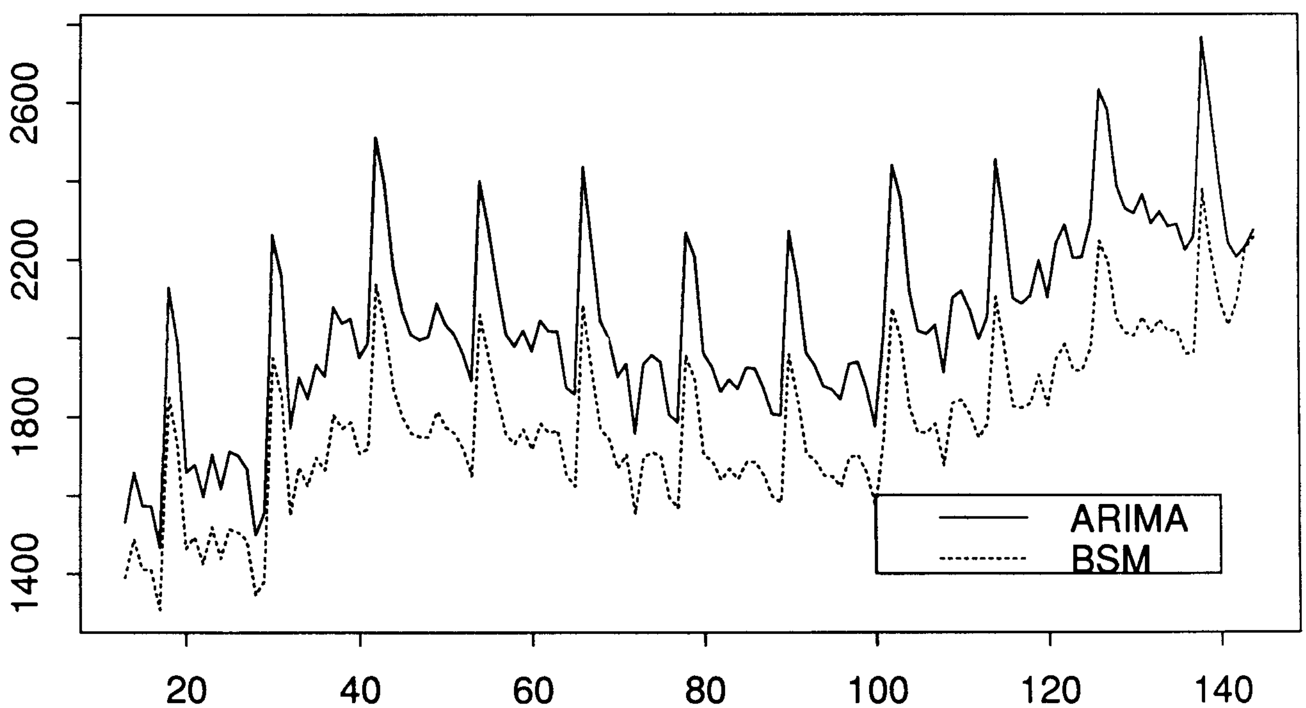
Solid line = Canonical ARIMA Seasonal Adjustment

Dash line = BSM Seasonal Adjustment

Teenage Unemployment

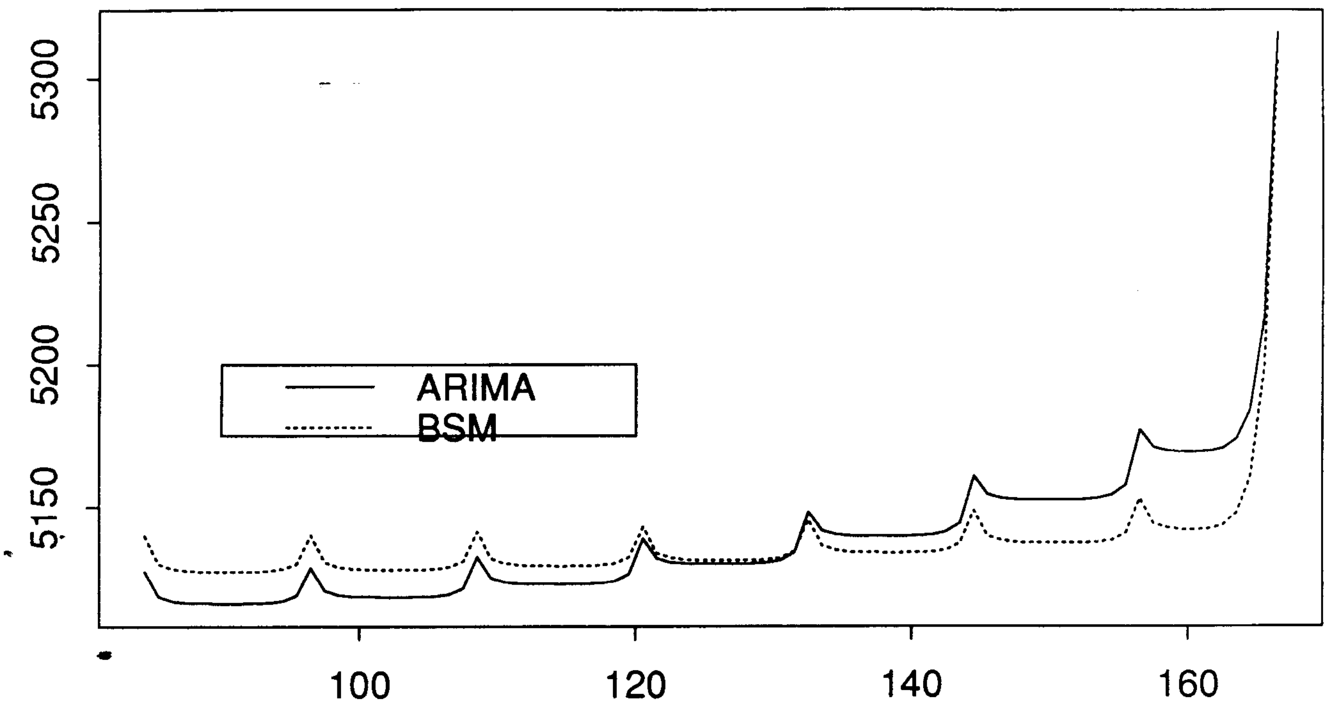


Signal Extraction Estimates



Signal Extraction Variances

Total US 5+ Housing Starts



Signal Extraction Variances