STATISTICAL RESEARCH DIVISION REPORT SERIES

SRD Research Report Number: CENSUS/SRD/RR-92/12

BAYESIAN ASSESSMENT OF UNCERTAINTY IN SEASONAL ADJUSTMENT WITH SAMPLING ERROR PRESENT

William R. Bell and Mark C. Otto

Statistical Research Division
U.S. Census Bureau
Washington D.C. 20233

Report First Issued: October 15, 1992; *Revised*: March 22, 1993

Abstract

Seasonal adjustment of economic time series is subject to the following sources of error: nonsampling survey errors, sampling error, model uncertainty, parameter uncertainty, seasonal plus nonseasonal decomposition uncertainty, and inherent signal extraction error. While it is difficult to allow for nonsampling errors and model uncertainty, we develop Bayesian methods to allow for the contributions of the other four error sources in assessing uncertainty in seasonal adjustment. There are four elements to our approach. (1) We develop ARIMA time series models for the signal (true, unobserved series) and sampling error components of a time series. (2) We use the model-based framework of Hillmer and Tiao, and Burman, for decomposition of the signal model into seasonal and nonseasonal component models. (3) We treat the model in a Bayesian framework that facilitates the combining of information about the sampling error model available from estimated sampling error variances and covariances with the information contained in the observed time series data. (4) We use Monte Carlo integration to obtain posterior means and variances of the seasonal and nonseasonal components. Uncertainty about the seasonal plus nonseasonal decomposition is reflected by using a flat prior for a key parameter in the approach of Hillmer and Tiao, and Burman. Different values of this parameter lead to different seasonal and nonseasonal component models that are still consistent with the signal model, and hence with the model for the observed time series, so the data contain no information about this parameter. Comparing the posterior means and variances with those obtained conditional on the model parameters, or on the decomposition parameter, or both, provides an assessment of the contributions of parameter uncertainty and decomposition uncertainty to seasonal adjustment uncertainty.

We apply our methods to a Census Bureau time series of five or more unit housing starts in the South region of the U.S. For this series we find that posterior means of the seasonal and nonseasonal components are insensitive to parameter uncertainty, but are very sensitive to decomposition uncertainty. We also find that parameter uncertainty has a generally small effect on posterior variances, though the effects vary erratically over time. We find that decomposition uncertainty has a large effect on posterior variances. The general conclusion is that in seasonal adjustment, careful thought should be devoted to the decomposition assumptions, the assumptions made to yield models for the seasonal and nonseasonal components consistent with the signal model. The decomposition assumptions have a profound effect on the results, and the data carry no information to discriminate between alternative model decompositions that yield the same model for the observed data.

Key Words: Seasonal decomposition, signal extraction, ARIMA model, Bayes theorem

1. Introduction

Let y_t be a time series of estimates from a repeated sample survey taken more frequently than annually, e.g. monthly or quarterly. We let y_t represent both the survey estimators (random variables) and the realized estimates (observed values). We consider the decomposition

$$y_t = S_t + N_t + e_t \qquad t = 1, \ldots, n \tag{1.1}$$

where S_t and N_t are the seasonal and nonseasonal components, and e_t represents sampling error. We write $Y_t = S_t + N_t$ for the series that would be observed if a census rather than a sample survey were taken at each time t. Thus, $y_t = Y_t$ if the estimates contain no sampling error. Notice that the seasonal + nonseasonal decomposition applies to Y_t rather than to y_t. With some exceptions (e.g. Wolter and Monsour 1981, Hausman and Watson 1985, Carlin and Dempster 1989, Pfeffermann 1990), the seasonal adjustment literature has focused on seasonal + nonseasonal decomposition of the observed series y_t without recognizing the presence of the sampling error e_t. The resulting estimators implicitly assign part of e_t to S_t and part to N_t.

In the context of (1.1) seasonal adjustment can be viewed either as (i) estimation of N_t, or (ii) estimation and removal of S_t, i.e. estimation of $y_t - S_t = N_t + e_t$. The first of these concepts is appropriate if one wishes to extract N_t from both the seasonal S_t and sampling error e_t. This is a natural successor to use of time series signal extraction methods to estimate Y_t from data on y_t, something suggested by Scott and Smith (1974), and Scott, Smith, and Jones (1977), and investigated more recently by Binder and Dick (1989), Bell and Hillmer (1990, 1992), Pfeffermann (1991), and others. Since statistical agencies conducting repeated surveys have not yet adopted such techniques, one may instead be led to the second concept as more consistent with current practice. The reluctance to use time series signal extraction techniques in repeated survey estimation

may derive from the need to balance the potential benefits of doing so against both the additional effort involved and the risks associated with possible failure of the time series models used. These concerns seem less important in seasonal adjustment than in repeated survey estimation, however, since seasonal adjustment inherently requires some sort of time series signal extraction. We consider both of the seasonal adjustment concepts here by examining estimation of both S_t and N_t.

In many cases the decomposition (1.1) is more appropriate for the time series of the logarithms of the original survey estimates. This is discussed for signal extraction estimation of Y_t in Bell and Hillmer (1990). To keep notation simple, we let y_t denote either the original estimates or their logarithms. In the latter case a decomposition for the original series is obtained simply by exponentiating (1.1). The resulting point estimates of S_t, and standard errors for estimates of S_t and N_t, may be conveniently interpreted in percentage terms.

Typically, N_t in (1.1) is further decomposed as $N_t = T_t + I_t$ (trend + irregular). In this paper we use the three component decomposition, $Y_t = S_t + T_t + I_t$, only as a tool in studying the two component decomposition, $Y_t = S_t + N_t$. Our methods readily extend to explicit study of the three component decomposition, however.

Our objective in this paper is to develop a model–based Bayesian methodology to assess uncertainty about S_t and N_t in (1.1). Bell (1989) identified the following sources of uncertainty in seasonal adjustment: (1) sampling error; (2) nonsampling error; (3) model uncertainty; (4) parameter uncertainty; (5) decomposition uncertainty; and (6) inherent uncertainty (signal extraction error). The effects of (2) and (3) are extremely difficult to assess, and we shall not attempt to do so here. We deal with the other sources of error by developing Gaussian time series models for the sampling error e_t and "true" (apart from nonsampling error) series Y_t, and then using Bayesian methods with noninformative priors on model parameters to develop posterior means and variances of S_t and N_t. Integral to our methodology is the model–based approach to seasonal adjustment of Burman (1980)

and Hillmer and Tiao (1982). This approach starts with a seasonal ARIMA (autoregressive–integrated–moving average) model for Y_t, and then makes assumptions about seasonality and nonseasonality that uniquely determine models for S_t and N_t that are consistent with the model for $Y_t = S_t + N_t$. Putting a noninformative prior on a key parameter involved in these assumptions permits explicit treatment of decomposition uncertainty.

The Bayesian approach we adopt has several advantages for this application. First, it is particularly convenient for handling the nonlinear sources of seasonal adjustment error, including parameter uncertainty and decomposition uncertainty. Second, the Bayesian approach provides a natural framework for recognizing uncertainty about the model parameters for both the sampling error (e_t) and signal (Y_t) processes. This important problem has not yet been dealt with in the literature on using time series signal extraction methods for repeated survey estimation. This literature typically assumes that the model for e_t, and often the model for Y_t, are known with certainty. Third, the Bayesian approach is convenient for assessing the relative contributions of the various sources of error to overall seasonal adjustment uncertainty. This can be done by comparing the usual posterior means and variances with those that result when the variances of particular error sources are set to zero.

The Bayesian approach requires explicit distributional assumptions about the innovations in the time series models for the components of y_t. Throughout the paper we assume normal distributions. The basic approach would work under other distributional assumptions, but use of nonnormal models in time series is difficult, and we shall not pursue this here.

Section 2 gives a general overview of our approach. The details are best illustrated in the context of an example, which is done in section 3 using a monthly time series of 5 or more unit housing starts in the south region of the United States. Sections 4 and 5 discuss extensions of the approach as presented in sections 2 and 3 to handle autocorrelated

sampling errors and to handle signal models with regression terms.

2. General Approach

Bell and Hillmer (1990, 1992) discuss development of time series models for the sampling error (e_t) and signal process (Y_t). They use ARIMA models for both, with the model for Y_t augmented with regression terms as appropriate. Other enhancements are used with the e_t model to account for particular features of survey designs and estimators. We take this work as a point of departure, and will not discuss model construction here. We focus instead on Bayesian treatment of the models, culminating in the development of posterior means and variances for the S_t and N_t components. Our primary interest is in how these are affected by the various error sources, particularly parameter uncertainty and decomposition uncertainty. The presentation of our approach in this section is very general; details of the approach are illustrated by the example of section 3.

2.1 The Sampling Error (e_t) Model – Bayesian Development

We assume that sampling error covariances, $\mathrm{Cov}(e_t, e_j)$, have been estimated for some set of time points t, j = 1, ... , m. We let c_{tj} denote the estimates, and denote the m×m matrix of these covariance estimates by $C = [c_{tj}]$. (For convenience of notation, and without loss of generality, we are writing this as if the c_{tj} correspond to the first m time points at which the series y_t is observed. The c_{tj} can actually be for any set of m successive time points with no change in our results.) We generally assume stationarity of e_t, so $\mathrm{Cov}(e_t, e_{t+k})$ depends only on k, not t. If y_t represents the logarithms of the original survey estimates, then we are actually assuming relative covariance stationarity of the original sampling errors (call these ξ_t), and e_t is approximately the relative sampling error ($e_t = \xi_t/\exp(Y_t)$). In either case, we assume that, based on the appearance of the c_{tj}, and on knowledge of the survey design and estimation procedures, a stationary ARMA(pe, qe) model,

$$(1 - \phi_1^e B - \cdots - \phi_{pe}^e B^{pe})e_t = (1 - \theta_1^e B - \cdots - \theta_{qe}^e B^{qe})b_t \qquad (2.1)$$

where the b_t are iid $N(0, \sigma_b^2)$, has been selected for e_t. Modifications can be made to (2.1) to deal with certain known nonstationarities of the sampling errors, such as variances fluctuating over time in known ways or independent redrawings of the sample. All the above aspects of sampling error modeling are discussed by Bell and Hillmer (1990, 1992).

Let $\underset{\sim}{\alpha}_e = (\phi_1^e, \ldots, \phi_{pe}^e, \theta_1^e, \ldots, \theta_{qe}^e, \sigma_b^2)'$ denote the parameters of (2.1). We seek to develop a posterior distribution for $\underset{\sim}{\alpha}_e$ given the available data, C, denoted $p(\underset{\sim}{\alpha}_e | C)$. This "posterior" then serves as a "prior" in the next stage of the analysis where we include information from data on the time series y_t.

If the e_t are uncorrelated over time, then $\underset{\sim}{\alpha}_e$ is just $\sigma_b^2 = \mathrm{Var}(e_t) \equiv \sigma_e^2$, and we can reduce the data from C to $s_1^2, \ldots, s_m^2$, where the s_t^2 are the diagonal elements of C, which are the estimates of the $\mathrm{Var}(e_t)$'s. The posterior then simplifies to $p(\sigma_e^2 | s_1^2, \ldots, s_m^2)$. We suggest two approaches for this case. One approach assumes the s_t^2 follow scaled χ_ν^2 distributions for some ν. Though estimates of variances (or relative variances) in repeated surveys are rarely a simple mean sum of squares that would suggest the χ_ν^2 distribution, nevertheless, this may be a useful approximation. The second approach assumes the s_t^2 follow lognormal distributions. Both these approaches are illustrated and compared for the example in section 3.

In the general case where e_t is autocorrelated the development of $p(\underset{\sim}{\alpha}_e | C)$ is more involved. To avoid obscuring the discussion at this point, we defer remarks on this topic to section 4. For the remainder of this section we assume that $p(\underset{\sim}{\alpha}_e | C)$ is available, in the sense that it can be computed for any given value of $\underset{\sim}{\alpha}_e$.

2.2 The Signal (Y_t) Model and the Posterior Distribution of Model Parameters

We use a seasonal ARIMA model for Y_t of the following general form:

$$\phi(B)(1-B)^d(1-B^{12})\, Y_t = \theta(B)(1-\theta_{12}B^{12})a_t \tag{2.2}$$

where $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ and $\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$ are nonseasonal AR and MA polynomials of degrees $p \geq 0$ and $q \geq 0$, $d \geq 0$ is the order of nonseasonal differencing, and a_t is iid $N(0, \sigma_a^2)$. Notice the presence of the seasonal difference and seasonal MA term in (2.2); this is important for the seasonal + nonseasonal decomposition discussed in section 2.3. We can also augment (2.2) with linear regression terms as

$$\phi(B)(1-B)^d(1-B^{12})[Y_t - \sum_{i=1}^{r} x_{it}\beta_i] = \theta(B)(1 - \theta_{12}B^{12})a_t \tag{2.3}$$

where the x_{it} might be used to explain the effects of, e.g., calendar variation or outliers. The usefulness of models of form (2.2) and (2.3) in seasonal adjustment is discussed in Hillmer, Bell, and Tiao (1983). For simplicity of presentation here we focus on the simpler model (2.2), and defer discussion of the model (2.3) to section 5.

Let $\underset{\sim}{\alpha} = (\phi_1,\ldots,\phi_p, \theta_1,\ldots,\theta_q, \theta_{12}, \sigma_a^2)$ be the parameters of (2.2). Given $p(\underset{\sim}{\alpha}_e|C)$ as discussed in section 2.1, and a prior density $p(\underset{\sim}{\alpha})$, the joint posterior density of the parameters is (let $\underset{\sim}{y} = (y_1,\ldots,y_n)'$ be the observed data on the time series y_t)

$$p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e|\underset{\sim}{y}, C) \propto p(\underset{\sim}{y}|\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, C)\, p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e|\, C) \tag{2.4}$$

$$= p(\underset{\sim}{y}|\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e)\, p(\underset{\sim}{\alpha}_e|C)\, p(\underset{\sim}{\alpha})\,. \tag{2.5}$$

We make two assumptions in going from (2.4) to (2.5). First, we assume C tells us nothing about $\underset{\sim}{y}$ beyond the information in $\underset{\sim}{\alpha}_e$, so that $p(\underset{\sim}{y}|\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, C) = p(\underset{\sim}{y}|\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e)$, i.e., $\underset{\sim}{y}$ and C are assumed conditionally independent given $\underset{\sim}{\alpha}$ and $\underset{\sim}{\alpha}_e$. (See sections 2.4 and 4 for related comments.) Modeling conditional dependence of $\underset{\sim}{y}$ and C, if it were present, seems a difficult task. Second, we assume that $\underset{\sim}{\alpha}$ is a priori independent of $\underset{\sim}{\alpha}_e$ and C, so that $p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e|C) = p(\underset{\sim}{\alpha}_e|C)\, p(\underset{\sim}{\alpha})$.

In the special situation when e_t is white noise we can replace $\underset{\sim}{\alpha}_e$ by σ_e^2 and C by $(s_1^2, \cdots, s_m^2)$ in (2.4) and (2.5). We shall use the noninformative prior, $p(\underset{\sim}{\alpha}) \propto 1/\sigma_a^2$, over the parameter space with $\sigma_a^2 > 0$ and with the ARMA parameters restricted to lie within the stationarity and invertibility regions. This last condition means that $|\theta_{12}| < 1$, and that the zeroes of $\phi(B)$ and $\theta(B)$ lie outside the unit circle.

Computing (2.5) requires computation of the likelihood function $p(\underset{\sim}{y}|\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e)$. For nonstationary time series models the likelihood is defined as the joint density of the differenced data, $w_t = (1-B)^d(1-B^{12})y_t$. Let $\underset{\sim}{w} = (w_{12+d+1}, \cdots, w_n)'$ and $\Sigma_w \equiv \Sigma_w(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e) = \mathrm{Var}(\underset{\sim}{w})$. Then the likelihood is defined as

$$p(\underset{\sim}{y}|\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e) = p(\underset{\sim}{w}|\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e) \propto |\Sigma_w|^{-1/2} \exp\{(-1/2)\, \underset{\sim}{w}'\Sigma_w^{-1}\underset{\sim}{w}\}. \tag{2.6}$$

Ansley and Kohn (1985) justify this choice of the likelihood function through "diffuse priors" on starting values, and propose a modified Kalman filter for its computation. As discussed in Bell and Hillmer (1992), (2.6) may be evaluated by the ordinary Kalman filter using an initialization given in Bell and Hillmer (1991).

With a sufficiently long time series, $p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e|\underset{\sim}{y}, C)$ might be approximated by a (truncated) normal density with mean given by the values $(\hat{\underset{\sim}{\alpha}}, \hat{\underset{\sim}{\alpha}}_e)$ that maximize the posterior density (2.5), and with variance matrix given by the negative inverse Hessian

(second derivative) matrix of $\log[p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e|\underset{\sim}{y}, C)]$ evaluated at $(\hat{\underset{\sim}{\alpha}}, \hat{\underset{\sim}{\alpha}}_e)$. (Box and Jenkins (1970, section 7.4) discuss asymptotic Bayesian inference for nonseasonal ARMA model parameters when the time series Y_t is observed (no sampling error).) Reparameterization of σ_a^2 and σ_e^2 to $\log(\sigma_a^2)$ and $\log(\sigma_e^2)$ may improve the normal approximation, and will remove part of the need for truncating the asymptotic normal distribution. We still need to truncate the normal distribution to the stationarity and invertibility regions for the ARMA parameters. An important reason for considering this asymptotic normal approximation to the posterior distribution is that simulation of parameter values $(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e)$ from such a distribution is easy. This greatly facilitates the computation of posterior means and variances of S_t and N_t by Monte Carlo integration, as discussed in section 2.4.

Certain features of the posterior density, $p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e|\underset{\sim}{y}, C)$, may be of independent interest, apart from the role of $p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e|\underset{\sim}{y}, C)$ in the computation of posterior means and variances of S_t and N_t. Two particular features come to mind. The first is the marginal posterior density of θ_{12}, $p(\theta_{12}|\underset{\sim}{y}, C)$. As discussed in Hillmer, Bell, and Tiao (1983), the value of θ_{12} reflects the stability of the underlying seasonal component. Values of θ_{12} near to 1 reflect a seasonal component that follows a nearly fixed annual pattern, while values of θ_{12} far from 1 (e.g., .5 or less) reflect a seasonal component whose annual pattern can change much more rapidly over time. (This statement is most accurate in the absence of "decomposition uncertainty" as discussed in the next section, e.g., if we are using the canonical decomposition defined there.) Examination of $p(\theta_{12}|\underset{\sim}{y}, C)$ tells us what information the data provide regarding the stability of the seasonal component. Secondly, it is of interest to compare the marginal posterior density, $p(\underset{\sim}{\alpha}_e|\underset{\sim}{y}, C)$, with the "prior" $p(\underset{\sim}{\alpha}_e|C)$. This tells us how information from the time series data $\underset{\sim}{y}$ adds to our knowledge about the sampling error model parameters $\underset{\sim}{\alpha}_e$, beyond what we know about $\underset{\sim}{\alpha}_e$ from just the estimates of sampling error variances and covariances contained in C. These and other

analyses of the posterior density $p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e | \underset{\sim}{y}, C)$ might be done using its asymptotic normal approximation. If this approximation is inadequate, such analyses will require specialized Bayesian simulation techniques, discussed briefly in section 2.4.

2.3 Seasonal + Nonseasonal Decomposition

Burman (1980) and Hillmer and Tiao (1982) address the following problem. Suppose we have a given seasonal ARIMA model of form (2.2) for Y_t. What ARIMA models for the independent components in $Y_t = S_t + T_t + I_t$ are consistent with the model for Y_t, while satisfying certain assumptions about the nature of seasonality, trend, and the irregular? "Consistent" means that the corresponding spectral densities satisfy the relation $f_Y(\lambda) = f_S(\lambda) + f_T(\lambda) + f_I(\lambda)$. We refer to the determination of component models consistent with a given model for Y_t as a seasonal + trend + irregular decomposition. Combining $T_t + I_t$ as N_t yields a seasonal + nonseasonal decomposition. (The spectral density $f_Y(\lambda)$ of Y_t following (2.2) is the nonnegative function of $\lambda \in [-\pi,\pi]$ defined by $f_Y(\lambda) = |\theta(e^{i\lambda})(1-\theta_{12}e^{12i\lambda})|^2\sigma_a^2/\{2\pi|\phi(e^{i\lambda})(1-e^{i\lambda})^d(1-e^{12i\lambda})|^2\}$. The spectral densities of S_t, T_t, and I_t ($f_S(\lambda)$, $f_T(\lambda)$, and $f_I(\lambda)$) are defined analogously in terms of their models. Because stationarity is formally required for a spectral density to exist, for models that involve differencing these are often called pseudo–spectral densities.)

Burman (1980) and Hillmer and Tiao (1982) make assumptions leading to a unique decomposition called the "canonical decomposition," thereby defining the canonical seasonal, trend, and irregular component models. We write the canonical decomposition as $Y_t = \bar{S}_t + \bar{T}_t + \bar{I}_t$. For the model (2.2), the canonical decomposition models can be written as follows:

$$
\begin{aligned}
&\text{(a)} \quad (1+B+\cdots+B^{11})\bar{S}_t = (1-\theta_{S,1}B-\cdots-\theta_{S,11}B^{11})\eta_{1t} \\
&\text{(b)} \quad \phi(B)(1-B)^{d+1}\,\bar{T}_t = \theta_T(B)\eta_{2t} \\
&\text{(c)} \quad I_t \text{ is white noise} \\
&\text{(d)} \quad \mathrm{Var}(\bar{I}_t) \equiv \bar{\sigma}_I^2 \text{ , is the maximum value consistent with (a) – (c) and (2.2).}
\end{aligned}
\tag{2.7}
$$

Here $\theta_T(B)$ is an MA polynomial whose degree depends on the particular ARIMA model (2.2), and η_{1t} and η_{2t} are white noise. The restriction of I_T to be white noise is without loss of generality for our purposes, since our focus will be on $N_t = T_t + I_t$. Burman (1980) discusses canonical decomposition for models somewhat more general than (2.2).

Bell and Hillmer (1984) discuss justifications for the assumptions leading to (2.7). Relaxing one or more of these assumptions permits other decompositions, $Y_t = S_t + T_t + I_t$, that are still consistent with the model (2.2). All such decompositions are called "admissible decompositions". The assumption that is easiest to relax is (2.7.d). In fact, any $\sigma_I^2 \in [0, \bar{\sigma}_I^2]$ corresponds to an admissible decomposition in the following way. Let $\gamma \in [0,1]$, and let ε_t^{γ} and $I_t^{1-\gamma}$ be independent white noise series with variances $\gamma\bar{\sigma}_I^2$ and $(1-\gamma)\bar{\sigma}_I^2$, respectively. Then

$$
Y_t = S_t + N_t = (\bar{S}_t + \varepsilon_t^{\gamma}) + (\bar{T}_t + I_t^{1-\gamma}) \tag{2.8}
$$

defines an admissible seasonal + nonseasonal decomposition for any $\gamma \in [0, 1]$. Notice that, in general, S_t and N_t depend on γ, and hence posterior means and variances of S_t and N_t will depend on what we assume about γ. Sometimes we will explicitly write S_t^{γ} $(= \bar{S}_t + \varepsilon_t^{\gamma})$ and N_t^{γ} $(= \bar{T}_t + I_t^{1-\gamma})$ instead of just S_t and N_t to emphasize the dependence on γ. Notice γ is the proportion of white noise variance $\bar{\sigma}_I^2$ added to the canonical $f_S(\lambda)$, and $1-\gamma$ is the proportion of $\bar{\sigma}_I^2$ added to the canonical $f_T(\lambda)$, to define the spectral densities of S_t^{γ} and N_t^{γ}.

We define (seasonal + nonseasonal) decomposition uncertainty as uncertainty about

γ in (2.8). From a Bayesian perspective such uncertainty should be expressed in a probability distribution about γ on [0, 1]. Since (2.8) will be consistent with the model (2.2) for any $\gamma \in [0, 1]$, the data contain no information about γ. Therefore, knowledge about γ comes entirely from its prior distribution. We focus on two possibilities: (1) believing the canonical decomposition is appropriate, so that γ is degenerate at 0, i.e., $\Pr(\gamma = 0) = 1$; (2) ignorance about γ, expressed as $\gamma \sim U(0,1)$. We let $p(\gamma)$ be the prior density of γ. If the prior distribution for γ is degenerate at 0 (or any other given value, e.g. $\gamma = .5$), $p(\gamma)$ can be thought of as a "delta function" (Jenkins and Watts, 1968, pp. 31–32).

It should be noted that decomposition uncertainty is an inherent feature of the seasonal adjustment problem that must be faced by all approaches to its solution, both model–based and empirical. The distinction of the work of Burman (1980) and Hillmer and Tiao (1982) is that they explicitly recognize decomposition uncertainty and deal with it through stated assumptions (leading to the canonical decomposition), whereas most other approaches to seasonal adjustment simply ignore the problem of decomposition uncertainty, effectively dealing with it by default — see Bell and Hillmer (1984) for a discussion of this point. Watson (1987) is one exception. Taking a frequentist approach, he treats γ in (2.8) as an unknown, nonestimable parameter, and develops mini–max estimates of S_t and N_t in this framework.

2.4 Posterior Means and Variances of S_t and N_t (Signal Extraction)

We assume $y_t = Y_t + e_t$ with e_t and Y_t following the models (2.1) and (2.2). The parameters are $\underset{\sim}{\alpha}$, $\underset{\sim}{\alpha}_e$, and γ. We assume we have developed a posterior density, $p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e | \underset{\sim}{y}, C)$, as discussed in section 2.2. We also assume $p(\gamma)$, a prior on γ, has been specified. The joint posterior density of $(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma)$ is then simply $p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e | \underset{\sim}{y}, C)p(\gamma)$, since γ is independent of the data. For the moment, we assume we can simulate values of $(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma)$ from the posterior distribution corresponding to this density.

For given values of $\underset{\sim}{\alpha}$, $\underset{\sim}{\alpha}_e$, and γ, the decomposition (2.8) is determined, and then, under suitable assumptions, the conditional distributions given $\underset{\sim}{y}$ of S_t^γ and N_t^γ in $y_t = S_t^\gamma + N_t^\gamma + e_t$ are normal with means and variances given by nonstationary signal extraction results discussed by Bell (1984a), Ansley and Kohn (1985), Kohn and Ansley (1987), and Bell and Hillmer (1988, 1991). The "suitable assumptions" that we make have to do with starting values for the ARIMA difference equations; either assumption A of Bell (1984a) or the diffuse prior assumptions of Ansley and Kohn (1985) may be used. Application of signal extraction results to repeated survey estimation and to seasonal adjustment is discussed in many of the references to these topics cited previously. In particular, Bell and Hillmer (1992) show how these calculations may be carried out by coupling the Kalman filter (initialized as discussed in Bell and Hillmer (1991)) with the fixed point smoother of reduced dimension described in Anderson and Moore (1979). This approach has been implemented in the REGCMPNT program developed by the time series staff of the Census Bureau. This software was used for the application in the next section.

The signal extraction techniques thus allow us to compute $E(S_t \mid \underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y})$ and $Var(S_t \mid \underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y})$ for $t = 1, \ldots, n$, and similarly for N_t. We assume it is unnecessary to also condition on C, i.e., $E(S_t \mid \underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y}) = E(S_t \mid \underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y}, C)$ and $Var(S_t \mid \underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y}) = Var(S_t \mid \underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y}, C)$, and similarly for N_t. (Analogous to what was done in section 2.2, we could go slightly further and assume $\underset{\sim}{\alpha}_e$ provides all the relevant information about the sampling errors, so that S_t and C are conditionally independent, and so are N_t and C.) These quantities determine the corresponding <u>marginal</u> conditional normal distributions of S_t and N_t. <u>Joint</u> conditional distributions of S_t and N_t, or of S_t and S_j for $j \neq t$, etc., would also require calculation of the appropriate conditional covariance.

Our objective is to obtain posterior means and variances of S_t and N_t. These are

$E(S_t|\underset{\sim}{y}, C)$ and $Var(S_t|\underset{\sim}{y}, C)$, and similarly for N_t. In light of the above remarks, these depend on C only through the posterior distribution of the model parameters, i.e., through $p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e|\underset{\sim}{y}, C)$. To simplify notation, we let $\underset{\sim}{y}^C = (\underset{\sim}{y}, C)$, with this notation used to emphasize that conditioning on both $\underset{\sim}{y}$ and C really uses C only through $p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e|\underset{\sim}{y}, C)$. To obtain $E(S_t|\underset{\sim}{y}, C) = E(S_t|\underset{\sim}{y}^C)$ and $Var(S_t|\underset{\sim}{y}, C) = Var(S_t|\underset{\sim}{y}^C)$ we use the following relations:

$$E(S_t|\underset{\sim}{y}^C) = E[E(S_t|\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y})|\underset{\sim}{y}^C]$$

$$= \int E(S_t|\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y})\, p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e|\underset{\sim}{y}, C)\, p(\gamma)\, d\underset{\sim}{\alpha}\, d\underset{\sim}{\alpha}_e\, d\gamma \qquad (2.9)$$

$$Var(S_t|\underset{\sim}{y}^C) = Var[E(S_t|\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y})|\underset{\sim}{y}^C] + E[Var(S_t|\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y})|\underset{\sim}{y}^C] \qquad (2.10)$$

We can approximate (2.9) via Monte Carlo integration by: (1) simulating $\underset{\sim}{\alpha}$, $\underset{\sim}{\alpha}_e$, and γ from the distributions corresponding to $p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e|\underset{\sim}{y}, C)$ and $p(\gamma)$; (2) computing $E(S_t|\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y})$ for each simulated $\underset{\sim}{\alpha}$, $\underset{\sim}{\alpha}_e$, γ; and (3) taking the sample mean of the $E(S_t|\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y})$'s over the simulations. To approximate (2.10) we proceed similarly, evaluating both $E(S_t|\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y})$ and $Var(S_t|\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y})$ at step (2), then taking the sample variance and the sample mean over the simulations of these respective results, and then adding these together to get $Var(S_t|\underset{\sim}{y}^C)$. Analogous procedures obviously yield $E(N_t|\underset{\sim}{y}^C)$ and $Var(N_t|\underset{\sim}{y}^C)$.

The above assumes that we can simulate from the posterior distribution with density $p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e|\underset{\sim}{y}, C)p(\gamma)$. As is discussed shortly, simulation of γ is actually unnecessary for the

cases we are interested in. Simulation of the model parameters from the posterior distribution with density $p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e | \underset{\sim}{y}, C)$, is a problem of some concern, however. If the (truncated) asymptotic normal approximation to $p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e | \underset{\sim}{y}, C)$ discussed in section 2.2 is deemed adequate, then this distribution can be easily used for simulation. If asymptotic normality does not provide an adequate approximation, then specialized simulation techniques are needed. How to simulate from a nonstandard distribution whose density can be calculated is currently a problem of great interest in Bayesian statistics; see Gelfand and Smith (1990) for a general discussion. Some of the approaches discussed there (e.g. Gibbs sampling) have been applied to time series models, although most such applications have been to autoregressive models. Marriott, et. al. (1992) and Chib and Greenberg (1992) discuss applications to nonseasonal ARMA models (the latter allow the model to be augmented with regression terms). We are unaware of any applications as yet of such simulation techniques to time series models as general as those we consider here. For our example we used a variant of acceptance–rejection sampling (Ripley 1987), as discussed in section 3.3. This approach was reasonably convenient for our particular example, but cannot be recommended as a general technique.

Importance sampling (see Geweke (1989)) seems a promising general approach for calculating the posterior means and variances of S_t and N_t. This approach replaces (2.9), for given γ, by

$$E(S_t | \underset{\sim}{y}^c, \gamma) = \int E(S_t | \underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y}) \, [p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e | \underset{\sim}{y}, C)/f(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e)] \, f(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e) \, d\underset{\sim}{\alpha} \, d\underset{\sim}{\alpha}_e \quad (2.11)$$

where $f(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e)$ is a probability density corresponding to a standard distribution from which simulation is easy, e.g., the truncated asymptotic normal approximation to the posterior. Computation of $E(S_t | \underset{\sim}{y}^c, \gamma)$ from (2.11) via Monte Carlo is done as for (2.9), but with a different integrand, and with $(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e)$ simulated from the distribution corresponding to

$f(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e)$. The integrand requires computation of $E(S_t | \underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y})$ as before, and of the importance function, $p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e | \underset{\sim}{y}, C)/f(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e)$. The numerator of the latter is computed as discussed in section 2.2, and computation of the denominator is generally easy (e.g., evaluation of a low dimensional multivariate normal density). The terms in the posterior variance expression (2.10) can be computed in the same way, but using different integrands. One drawback to importance sampling is that it does not produce simulations of $(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e)$ from the exact posterior distribution with density $p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e | \underset{\sim}{y}, C)$, and so does not provide a means for examining the exact posterior distribution of the model parameters.

Regarding γ, the two cases of primary interest are: (1) $\gamma = 0$, i.e., the canonical decomposition; and (2) $\gamma \sim U(0,1)$. If $p(\gamma)$ is degenerate at any value we simply fix γ in (2.9) and (2.10), and we obviously need not simulate γ. If $\gamma \sim U(0,1)$, we can analytically integrate out γ to obtain the following results for the posterior means and variances of $S_t = S_t^{\gamma}$:

$$E(S_t^{\gamma} | \underset{\sim}{y}^c) = E(S_t^{1/2} | \underset{\sim}{y}^c) \tag{2.12}$$

$$Var(S_t^{\gamma} | \underset{\sim}{y}^c) = Var(S_t^{1/2} | \underset{\sim}{y}^c) + (1/3)\{Var(I_t^{1/2} | \underset{\sim}{y}^c) + [E(I_t^{1/2} | \underset{\sim}{y}^c)]^2\} - E(\bar{\sigma}_I^2 | \underset{\sim}{y}^c)/6 \tag{2.13}$$

with analogous results for the posterior means and variances of $N_t = N_t^{\gamma}$. (The notation $S_t^{1/2}$ means S_t^{γ} for $\gamma = 1/2$, which is $\bar{S}_t + \varepsilon_t^{1/2}$ — see (2.8). It does not mean $\sqrt{S_t}$.) The results (2.12) and (2.13) are derived in the appendix. The quantities needed to evaluate (2.12) and (2.13), along with the analogous results for N_t^{γ}, can be obtained by performing the Monte Carlo signal extraction computations for both the three component decomposition $Y_t = S_t^{1/2} + \bar{T}_t + I_t^{1/2}$, and the two component decomposition $Y_t = S_t^{1/2} + N_t^{1/2}$. The posterior mean of $\bar{\sigma}_I^2$, $E(\bar{\sigma}_I^2 | \underset{\sim}{y}^c) = E(\bar{\sigma}_I^2 | \underset{\sim}{y}, C)$, can be approximated via Monte Carlo by taking the average of the $\bar{\sigma}_I^2$ values over the simulations from the exact posterior

distribution defined by $p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e | \underset{\sim}{y}, C)$. Alternatively, importance sampling could be used to approximate $E(\bar{\sigma}_I^2 | \underset{\sim}{y}, C)$.

We note in passing that if one can simulate from the posterior distribution of the parameters, then one can also produce simulations from the marginal posterior distributions of S_t and N_t, by Monte Carlo. This is done by simulating values of S_t from the normal distribution with mean $E(S_t | \underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y})$ and variance $Var(S_t | \underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y})$ for each set of simulated values of $(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma)$. This will generally require more simulations than are needed for the posterior means and variances (to produce reliable results), it will produce a very large file of the simulated S_t (and N_t) values for each time point, and it will leave one with the problem of how to summarize these simulations from a large number of marginal posterior distributions. We will not pursue this topic in this paper.

3. Application: 5 or More Unit Housing Starts, South Region of the U. S.

To illustrate our approach, we analyze data on 5 or more unit housing starts in the south region of the U. S. Figure 1 shows a plot of the time series of original monthly survey estimates (z_t) from January 1975 through November 1988 (167 observations). (The data were obtained from the Construction Statistics Division of the Census Bureau. The scale is omitted from the plot to prevent identification of the actual numbers of this unpublished series.) This data was previously analyzed by Bell and Hillmer (1992), who estimated the following model for $y_t = \log(z_t)$:

$$y_t = Y_t + e_t \qquad t = 1, \ldots, 167$$

$$(1 - B)(1 - B^{12})Y_t = (1 - \theta_1 B)(1 - \theta_{12} B^{12})a_t \qquad a_t \text{ i.i.d. } N(0, \sigma_a^2) \qquad (3.1)$$

$$e_t \text{ i.i.d. } N(0, \sigma_e^2)$$

The parameter estimates were $\hat{\theta}_1 = .53$, $\hat{\theta}_{12} = .80$, $\hat{\sigma}_a^2 = .0510$, and $\hat{\sigma}_e^2 = .0162$. To develop the e_t model, estimates of relative sampling variances and autocorrelations at lags 1, ... , 23 were available for each of the 60 months of the years 1982 through 1986. Averaging autocorrelations for a given lag over time yielded –.06, –.11, –.02, –.02, and .04 at lags 1 through 4 and 12, suggesting little autocorrelation is present in the sampling error, thus leading to the white noise model for e_t. σ_e^2 was then estimated by the average of the 60 estimates of relative sampling error variance, $\hat{\sigma}_e^2 = .0162$. This corresponds to a sampling coefficient of variation of $\hat{\sigma}_e = .13$, indicating a substantial amount of sampling error is present. In fact, the Census Bureau does not publish monthly estimates of regional housing starts by type of structure (single family, 5 or more unit, etc.) because of the generally large amount of sampling error in these estimates. The remaining parameters of (3.1) were estimated by maximum likelihood using the time series data $\underset{\sim}{y} = (y_1, \cdots, y_{167})'$, but holding σ_e^2 fixed at its estimated value of .0162. Bell and Hillmer (1992) then used this estimated model for signal extraction estimation of Y_t. Here, we shall instead be concerned with obtaining posterior means and variances of S_t and N_t, allowing for uncertainty in the model parameters and in the seasonal + nonseasonal decomposition.

3.1 Sampling Error Model

For this example the only parameter of (2.1) is $\sigma_b^2 = \text{Var}(e_t) \equiv \sigma_e^2$. We assume this variance is constant over time, and that the relative sampling variance estimates, $s_1^2, \cdots, s_m^2$ are m independent, unbiased estimates of σ_e^2. As in section 2.1, it is convenient to label these $s_1^2, \cdots, s_m^2$ despite their actual time points. We shall consider two possible assumptions regarding the distribution of the s_i^2.

First, we assume that $\nu s_i^2/\sigma_e^2 \sim \chi_\nu^2$ where ν is the degrees of freedom in each s_i^2. This

assumption is suggested only as a potentially useful approximation. In fact, the survey of construction (from which the estimates are produced) uses a stratified three–stage cluster sample, and sampling variance estimates S_i^2 for the original estimates z_i are obtained by collapsing pairs of strata and using the Keyfitz two–per–stratum method. The estimate we use of the sampling error variance of $y_i = \log(z_i)$ is the relative variance, $s_i^2 = S_i^2/z_i^2$. Thus, s_i^2 is not of a simple form that directly leads to a χ_ν^2 distribution. Furthermore, the appropriate value for ν, the degrees of freedom assigned to each s_i^2, is unknown.

Proceeding with the assumption that $\nu s_i^2/\sigma_e^2 \sim \chi_\nu^2$, and using the noninformative prior $p(\sigma_e^2) \propto 1/\sigma_e^2$, from Bayes theorem the posterior density of σ_e^2 is

$$p(\sigma_e^2|\nu, s_1^2, \ldots, s_m^2) \propto (1/\sigma_e^2) \prod_{i=1}^{m} p(s_i^2|\sigma_e^2, \nu)$$

$$= [\Gamma(\nu/2)]^{-m}(\nu/2)^{m\nu/2}(\prod_i s_i^2)^{\nu/2-1}(\sigma_e^2)^{-(m\nu/2)-1}\exp\{(-\nu/2\sigma_e^2)\sum_i s_i^2\}$$

$$= g(\nu)(\nu\sum_i s_i^2/\sigma_e^2)^{(m\nu/2)+1}\exp\{-(\nu/2\sigma_e^2)\sum_i s_i^2\} \qquad (3.2)$$

where $g(\nu) = [\Gamma(\nu/2)]^{-m}\, 2^{-m\nu/2}\, \nu^{-1}\, (\sum_i s_i^2)^{-m\nu/2}\, (\prod_i s_i^2)^{\nu/2-1}$. Conditional on ν, (3.2) defines a scaled inverse $\chi_{m\nu}^2$ distribution for σ_e^2 (a $\chi_{m\nu}^2$ distribution for $\nu\sum_i s_i^2/\sigma_e^2$). This is a generalization of a standard Bayesian result for the case m = 1 (Box and Tiao 1973, Theorem 2.3.1).

We suggest two ways to deal with the problem that ν is unknown. First, we can use the result that $\mathrm{Var}(s_i^2)/[E(s_i^2)]^2 = 2/\nu$, to estimate ν by $\hat{\nu} = 2(\bar{s}^2)^2/\hat{\mathrm{Var}}(s_i^2)$, where $\bar{s}^2$ and $\hat{\mathrm{Var}}(s_i^2)$ are the sample mean and variance of $s_1^2, \ldots, s_m^2$. For this example, $\bar{s}^2 = .0162$ and m = 60, as given above, and $\hat{\mathrm{Var}}(s_i^2) = 1.04 \times 10^{-4}$, giving $\hat{\nu} = 5.1$. Using the approximation $\nu = \hat{\nu}$, $p(\sigma_e^2|s_1^2, \ldots, s_{60}^2)$ is then defined by $m\hat{\nu}\bar{s}^2/\sigma_e^2 = 4.96/\sigma_e^2 \sim \chi_{m\hat{\nu}}^2$,

where $m\hat{\nu} = 306$. Alternatively, to recognize uncertainty about ν via a fully Bayesian approach, we can multiply (3.2) by $p(\nu)$, a prior distribution for ν, and numerically integrate the product over ν to get the marginal posterior of σ_e^2. Two natural choices of noninformative priors are a flat prior ($p(\nu) \propto$ constant) or the Jeffreys' prior. (See Box and Tiao (1973, pp. 51–54) for a discussion of Jeffreys' rule.) For our problem, the Jeffreys' prior can be shown to be $p(\sigma_e^2, \nu) \propto (1/\sigma_e^2)[1 + (\nu/2)\psi'(\nu/2)]^{1/2}$, where $\psi'(x) = d^2[\log(\Gamma(x)]/dx^2$, is the trigamma function (Abramowitz and Stegun 1970, p. 260). Except for $\nu < 2$, however, $[1 + (\nu/2)\psi'(\nu/2)]^{1/2}$ is a nearly constant function of ν, and so, for this example, the Jeffreys' prior yields virtually identical results to those from the flat prior for ν.

The second general approach we consider is to assume that the $\log(s_i^2)$ are approximately normally distributed, i.e., $\log(s_i^2) \sim N(\mu, v)$. From properties of the lognormal distribution, and the assumption that $E(s_i^2) = \sigma_e^2$, we have $\sigma_e^2 = \exp\{\mu + v/2\}$, and so $\log(\sigma_e^2) = \mu + v/2$. We want $p(\log(\sigma_e^2)|\ s_1^2, \ldots, s_m^2)$, from which $p(\sigma_e^2|\ s_1^2, \ldots, s_m^2)$ is easily obtained. Using the noninformative prior, $p(\mu, v) \propto 1/v$, from standard Bayesian results on inference for the mean and variance of a normal distribution (Box and Tiao 1973, section 2.4), the joint posterior density of (μ, v) is

$$p(\mu, v|s_1^2, \ldots, s_m^2) \propto p(\mu|v, s_1^2, \ldots, s_m^2)\, p(v|s_1^2, \ldots, s_m^2) \tag{3.3}$$

where $p(\mu|v, s_1^2, \ldots, s_m^2)$ is a $N(\hat{\mu}, v/m)$ density, and $p(v|\ s_1^2, \ldots, s_m^2)$ is a scaled inverse χ^2_{m-1} density ($(m-1)\hat{v}/v \sim \chi^2_{m-1}$). Here $\hat{\mu}$ and $\hat{v}$ are the sample mean and sample variance of the $\log(s_i^2)$. An approximate posterior is obtained by setting $v = \hat{v}$ and using $\mu \sim N(\hat{\mu}, \hat{v}/m)$, which implies $\log(\sigma_e^2) \sim N(\hat{\mu} + \hat{v}/2, \hat{v}/m)$. For this problem $\hat{\mu} = -4.285$ and $\hat{v} = .33$, yielding an approximate posterior for $\log(\sigma_e^2)$ which is a $N(-4.45, .0055)$ distribution. The approximate posterior of σ_e^2 is then the corresponding lognormal distribution. Alternatively, the exact posterior can be obtained by transforming from

(μ, v) to $(\mu + v/2, v) = (\log(\sigma_e^2), v)$, obtaining the joint posterior density of $(\log(\sigma_e^2), v)$ using (3.3) and the Jacobian of the logarithmic transformation, and numerically integrating out v to get $p(\log(\sigma_e^2)|\ s_1^2, \ldots, s_m^2)$, which is then transformed to $p(\sigma_e^2|\ s_1^2, \ldots, s_m^2)$.

Figure 2.a shows $p(\sigma_e^2|\ s_1^2, \ldots, s_m^2)$ resulting from the exact and approximate treatments of both the chi–squared and lognormal models. The numerical integrations required for the exact treatments were performed with the Mathematica computer package. We see that all four posterior densities peak near .0162, the estimate of σ_e^2 obtained by Bell and Hillmer (1992). The posteriors are all concentrated between about .013 and .020, and they convey some uncertainty about σ_e^2. (The corresponding range on σ_e, the coefficient of variation of the original survey estimates, is from 11.4% to 14.1%.) The difference between the exact and approximate posteriors under the chi–squared assumption is interesting. Notice that using the noninformative (Jeffreys') prior for ν produces a posterior for σ_e^2 that is more concentrated than the conditional posterior for σ_e^2 obtained by setting $\nu = \hat{\nu} = 5.1$, a result that may be contrary to expectations. Figure 2.b shows $p(\nu|s_1^2, \ldots, s_m^2)$ for the chi–squared model, with virtually identical results obtained under the flat or Jeffreys' (trigamma) priors for ν.

We can use a model selection criterion such as AIC or BIC to decide on which version of $p(\sigma_e^2|s_1^2, \ldots, s_m^2)$ to use. (More formal Bayesian procedures could also be used.) Since both the chi–squared and lognormal models involve two parameters, such criteria reduce to 2×(log–likelihood) comparisons at the maximum likelihood parameter estimates. For our example, the log–likelihood difference is about 1.5 in favor of the lognormal model. Notice from Figure 2.a that the exact posterior under the lognormal model is very close to the approximate posterior with $\nu = 5.1$ under the chi–squared model. We shall therefore proceed with the second of these (which is easier to evaluate) so that $p(\sigma_e^2|s_1^2, \ldots, s_m^2)$ is defined by (3.2) with $\nu = 5.1$. This now becomes the "prior" for σ_e^2 for the analysis of the next section.

3.2 Signal Model

We now use the time series data $\underset{\sim}{y}$ to develop a posterior distribution for the parameters $(\theta_1, \theta_{12}, \sigma_a^2, \sigma_e^2)$. From (2.5) the posterior density is

$$p(\theta_1, \theta_{12}, \sigma_a^2, \sigma_e^2 | \underset{\sim}{y}, s_1^2, \ldots, s_{60}^2) \propto$$
$$p(\underset{\sim}{y} | \theta_1, \theta_{12}, \sigma_a^2, \sigma_e^2)\, p(\sigma_e^2 | s_1^2, \ldots, s_{60}^2)\, p(\theta_1, \theta_{12}, \sigma_a^2) \qquad (3.4)$$

where $p(\sigma_e^2 | s_1^2, \ldots, s_{60}^2)$ is defined by (3.2) with $\nu = 5.1$, and $p(\underset{\sim}{y} | \theta_1, \theta_{12}, \sigma_a^2, \sigma_e^2)$ is defined as in (2.6) and can be computed by the suitably initialized Kalman filter. We use the noninformative prior $p(\theta_1, \theta_{12}, \sigma_a^2) \propto 1/\sigma_a^2$ truncated to $|\theta_1| \leq 1$, $|\theta_{12}| \leq 1$, $\sigma_a^2 > 0$. A multivariate normal approximation to the posterior for $(\theta_1, \theta_{12}, \log(\sigma_a^2), \log(\sigma_e^2))'$ (restricted to $|\theta_1| \leq 1$, $|\theta_{12}| \leq 1$) has mean vector given by the maximum posterior density estimates, $(\hat{\theta}_1, \hat{\theta}_{12}, \log(\hat{\sigma}_a^2), \log(\hat{\sigma}_e^2))'$, and covariance matrix given by the inverse Hessian (second derivative) matrix of $-\log[p(\theta_1, \theta_{12}, \log(\sigma_a^2), \log(\sigma_e^2) | \underset{\sim}{y}, s_1^2, \ldots, s_{60}^2)]$, evaluated at $(\hat{\theta}_1, \hat{\theta}_{12}, \log(\hat{\sigma}_a^2), \log(\hat{\sigma}_e^2))'$. A modification of the REGCMPNT program was used to maximize (3.4) over $(\theta_1, \theta_{12}, \log(\sigma_a^2), \log(\sigma_e^2))'$ to get the following maximum posterior density estimates, with standard errors from the inverse Hessian in parentheses:

$$\hat{\theta}_1 = .52 \quad (.083) \qquad\qquad \hat{\theta}_{12} = .79 \quad (.11) \qquad (3.5)$$
$$\log(\hat{\sigma}_a^2) = -2.98 \quad (.17) \quad \Rightarrow \quad E(\sigma_a^2 | \underset{\sim}{y}, s_1^2, \ldots, s_{60}^2) = .0516.$$
$$\log(\hat{\sigma}_e^2) = -4.12 \quad (.081) \quad \Rightarrow \quad E(\sigma_e^2 | \underset{\sim}{y}, s_1^2, \ldots, s_{60}^2) = .0162.$$

The Hessian matrix was evaluated by taking numerical second derivatives of the negative log of (3.4), using formulas (25.3.23) and (25.3.27) given in Abramowitz and Stegun (1970). The resulting posterior correlation matrix of the parameters is

$$\begin{array}{c} \\ \theta_1 \\ \theta_{12} \\ \log(\sigma_a^2) \\ \log(\sigma_e^2) \end{array} \begin{bmatrix} 1.00 & & & \\ .01 & 1.00 & & \\ .29 & -.12 & 1.00 & \\ -.11 & -.03 & -.22 & 1.00 \end{bmatrix} \tag{3.6}$$

Several things are worth noting about the results in (3.5) and (3.6). First, the posterior mode estimates are very close to the parameter estimates obtained by Bell and Hillmer (1992). Second, the (approximate) posterior mean of σ_e^2, $E(\sigma_e^2 | \underset{\sim}{y}, s_1^2, \ldots, s_{60}^2) = .0162$, is unchanged from $E(\sigma_e^2 | s_1^2, \ldots, s_{60}^2)$, the "prior" mean of σ_e^2 without using the time series data $\underset{\sim}{y}$. Third, (3.6) shows little posterior correlation between the parameters. There is a small negative correlation between $\log(\sigma_a^2)$ and $\log(\sigma_e^2)$, which would be expected, and a small positive correlation between θ_1 and $\log(\sigma_a^2)$, the reason for which is unclear.

Figure 3 shows marginal densities for θ_1, θ_{12}, $\log(\sigma_a^2)$, and $\log(\sigma_e^2)$ from the normal approximation to the posterior, as well as corresponding histograms of parameter values simulated from the exact joint posterior. (1178 sets of parameters were simulated as discussed in the next subsection.) In comparison with the normal approximation, the marginal posterior of θ_1, and possibly those of θ_{12} and $\log(\sigma_e^2)$, appear to exhibit some skewness. The restriction to the invertibility region has an important effect on the posterior of θ_{12}, and in fact the exact posterior density of θ_{12} appears to "flatten out" into the invertibility boundary $\theta_{12} = 1$ at a somewhat higher level than does the normal approximation. These deficiencies in the normal approximation may not be very important for our purpose of computing posterior means and variances of S_t and N_t by Monte Carlo integration.

Finally, though not shown here, the normal approximation to

$p(\log(\sigma_e^2)|\underset{\sim}{y}, s_1^2, \ldots, s_{60}^2)$ was transformed to a lognormal approximation to $p(\sigma_e^2|\underset{\sim}{y}, s_1^2, \ldots, s_{60}^2)$, which was then plotted on the same graph as $p(\sigma_e^2|s_1^2, \ldots, s_{60}^2)$, the "prior" for σ_e^2 developed in the previous section (from (3.2) with $\nu = 5.1$). The two curves were virtually identical. Thus, apart from possible deviations of the true posterior of $\log(\sigma_e^2)$ from the normal approximation, we can say that $p(\sigma_e^2|\underset{\sim}{y}, s_1^2, \ldots, s_{60}^2) \approx p(\sigma_e^2|s_1^2, \ldots, s_{60}^2)$. In other words, the time series data $\underset{\sim}{y}$ add nothing to our knowledge about σ_e^2 beyond the information contained in the variance estimates $s_1^2, \ldots, s_{60}^2$.

3.3 Simulating from the Exact Posterior Distribution of the Model Parameters

We produced simulated values of $(\theta_1, \theta_{12}, \sigma_a^2, \sigma_e^2)$ from their exact posterior distribution (with density defined by (3.4)) for use in the Monte Carlo integration results of section 3.4, as well as for the comparisons with the normal approximation to the posterior made in the last section. The simulations were done using a variation on acceptance–rejection sampling, as follows. First, a set Ω was defined in the four dimensional parameter space of $\underset{\sim}{\omega} \equiv (\theta_1, \theta_{12}, \log(\sigma_a^2), \log(\sigma_e^2))$, such that Ω was thought to contain virtually all the posterior probability. Then, values of $\underset{\sim}{\omega}$ were simulated from a uniform distribution on this set. (The definition of Ω and simulation on Ω are discussed below.) These $\underset{\sim}{\omega}$ values were accepted or rejected as simulations from the true posterior according to the following scheme:

For each $\underset{\sim}{\omega}$ simulated from the uniform distribution on Ω:

(1) Simulate U from the U(0,1) distribution.

(2) If $U \leq p(\underset{\sim}{\omega}|\underset{\sim}{y}, C) \,/\, \max_{\underset{\sim}{\omega}}[p(\underset{\sim}{\omega}|\underset{\sim}{y}, C)]$, accept $\underset{\sim}{\omega}$; otherwise, reject $\underset{\sim}{\omega}$.

At step (2) $\max_{\underset{\sim}{\omega}}[p(\underset{\sim}{\omega}|\underset{\sim}{y}, C)] = p(\underset{\sim}{\hat{\omega}}|\underset{\sim}{y}, C)$ where $\underset{\sim}{\hat{\omega}} = (\hat{\theta}_1, \hat{\theta}_{12}, \log(\hat{\sigma}_a^2), \log(\hat{\sigma}_e^2))$ is the set of

maximum posterior density estimates defined in (3.5). It can be seen from Ripley (1987, pp. 60–61) that the accepted $\underset{\sim}{\omega}$ values constitute a random sample from the true posterior of $\underset{\sim}{\omega}$, with density $p(\underset{\sim}{\omega}|\underset{\sim}{y}, C)$. Finally, simulated values of $\log(\sigma_a^2)$ and $\log(\sigma_e^2)$ were exponentiated to give simulated values of σ_a^2 and σ_e^2.

The set Ω was defined as the intersection of (i) the standard four dimensional 99.95 per cent confidence ellipsoid for $\underset{\sim}{\omega}$ from the asymptotic normal approximation to the posterior given in section 3.2, and (ii) the invertibility region ($|\theta_1| < 1$, $|\theta_{12}| < 1$). Values of $p(\underset{\sim}{\omega}|\underset{\sim}{y}, C)/p(\hat{\underset{\sim}{\omega}}|y, C)$ were computed at the points $\underset{\sim}{\omega}$ where the ellipsoid (i) intersected the $\underset{\sim}{\omega}$–axes, and also where it intersected the principal component axes, to verify that the exact posterior density $p(\underset{\sim}{\omega}|\underset{\sim}{y}, C)$ was very small at these boundary points of Ω ("very small" was defined as $p(\underset{\sim}{\omega}|\underset{\sim}{y}, C)/p(\hat{\underset{\sim}{\omega}}|y, C) < .001$) . To simulate from a uniform distribution on Ω we first defined a four dimensional "box" just containing Ω whose sides were parallel to the principal component axes. We then simulated uniformly on this box by independently simulating each ζ_i in $\underset{\sim}{\zeta} \equiv (\zeta_1, \cdots, \zeta_4)' = P\underset{\sim}{\omega}$ on the appropriate interval defined by the box, where P is the matrix of eigenvectors of the variance–covariance matrix for the asymptotic normal approximation to the posterior of $\underset{\sim}{\omega}$. After translating each $\underset{\sim}{\zeta}$ to $\underset{\sim}{\omega} = P'\underset{\sim}{\zeta}$, we rejected any values $\underset{\sim}{\omega}$ lying outside of Ω (values $\underset{\sim}{\omega}$ lying outside either region defined in (i) and (ii) above).

The general drawback to this use of acceptance–rejection sampling in multiple dimensions is that the acceptance rate in step (2) above is very low. We generated 50,000 simulations from the uniform distribution on Ω to produce 1178 simulations from the exact posterior, an acceptance rate of about 2.4 per cent. For most problems this acceptance rate would be unacceptably low. This was not of such great concern to us because the computational cost of doing the 50,000 simulations from the uniform distribution on Ω was trivial, and the computational cost of applying the accept–reject rule (2) above to the simulated $\underset{\sim}{\omega}$ values was low, relative to the cost of computing the conditional means and

variances of S_t and N_t for the simulated $\underset{\sim}{\omega}$ values (signal extraction). We needed to perform the latter only for the 1178 accepted parameter values. Although it was convenient for this particular example, we cannot generally recommend the acceptance–rejection scheme for problems of this type, particularly for models with more than four parameters.

3.4 Analysis of Posterior Means and Variances of S_t and N_t

Having produced sets of simulated values of the parameters θ_1, θ_{12}, σ_a^2, and σ_e^2 from their exact posterior distributions, we can approximate the posterior means and variances of S_t and N_t (for t = 1, ... , 167) by Monte Carlo integration as discussed following (2.9) and (2.10). This directly gives us the results we want for any specified γ, such as $\gamma = 0$ (the canonical decomposition). The corresponding results when $\gamma \sim U(0,1)$ are obtained using (2.12) and (2.13). Comparing results when $\gamma \sim U(0,1)$ with those obtained when $\gamma = 0$ (or any other fixed value) provides information about the effects of decomposition uncertainty. Also, comparing the true posterior means and variances with signal extraction results obtained with parameters held at given values (e.g. the maximum posterior density estimates $\hat{\underset{\sim}{\omega}}$ given in (3.5)) provides information about the effects of parameter uncertainty on posterior means and variances. We now investigate these issues for our example.

Figure 4 shows posterior seasonal factors, $\exp\{E(S_t \mid \underset{\sim}{y}^c)\}$, for the canonical decomposition ($\gamma = 0$) and for the case of $\gamma \sim U(0,1)$. From (2.12), the latter are the same as the seasonal factors for $\gamma = 1/2$. The canonical seasonal factors show sharp drops in the winter (particularly January and February), spring and fall peaks, and a slow evolution over time. The general pattern of peaks and drops can also be found in the graph of the seasonal factors when $\gamma \sim U(0,1)$, but there it is obscured by pronounced year–to–year fluctuations. These fluctuation arise from the additional white noise assigned to S_t since,

when $\gamma \sim U(0,1)$, $E(S_t^{\gamma} | \underset{\sim}{y}^c) = E(S_t^{1/2} | \underset{\sim}{y}^c)$, and $S_t^{1/2} = \bar{S}_t + \varepsilon_t^{1/2}$ includes 1/2 the maximum possible white noise that can be assigned to the seasonal component. Corresponding plots of posterior means for N_t (not shown) exhibit the reverse behavior: $E(N_t^{\gamma} | \underset{\sim}{y}^c)$ is smoother when $\gamma \sim U(0,1)$ (equivalently, $\gamma = 1/2$) than when $\gamma = 0$. These results show that decomposition uncertainty has a marked effect on posterior means.

Figure 5 shows the ratios of the posterior seasonal and nonseasonal factors for the canonical decomposition to the corresponding canonical factors obtained from signal extraction calculations with the model parameters set at the estimates $\hat{\underset{\sim}{\omega}}$. That is, Figure 5 shows $\exp\{E(S_t | \gamma = 0, \underset{\sim}{y}^c)\} \; / \; \exp\{E(S_t | \underset{\sim}{\omega} = \hat{\underset{\sim}{\omega}}, \gamma = 0, \underset{\sim}{y}^c)\}$, and similarly for N_t. While these ratios vary erratically over time, few deviate from 1 by more than plus or minus one per cent. The corresponding plots (not shown) for the case of $\gamma \sim U(0,1)$ (equivalently, $\gamma = 1/2$) also fluctuate over time, but are even closer to 1: few deviate from 1 by as much as $\pm .5\%$. These results show that parameter uncertainty has little affect on the posterior means for this example. Hence, if point estimates were the only quantities of interest, simply doing signal extraction with the estimated parameters would be sufficient.

We turn now to consideration of posterior variances. The effects of parameter uncertainty on posterior variances can be examined in two ways that turn out to be roughly equivalent. First, we can compare posterior variances with variances conditional on parameters set at their estimated values. Alternatively, we can examine both terms in the posterior variance expression (2.10): the variance (over the simulations) of the conditional expectations (the contribution of parameter uncertainty), plus the average (over the simulations) of the conditional variances. The two approaches are roughly equivalent for this example because the second term in (2.10) turns out to be very close to the variance conditional on $\underset{\sim}{\omega} = \hat{\underset{\sim}{\omega}}$. (For the canonical decomposition, percentage differences between these two were generally less than 6% for S_t and 1% for N_t.)

Figure 6 shows, for the canonical decomposition, posterior variances (solid curves) of

S_t and N_t along with variances conditional on estimated parameters (dotted curves). That is, Figure 6 shows $\mathrm{Var}(S_t \mid \gamma = 0, \underset{\sim}{y}^c)$ and $\mathrm{Var}(S_t \mid \underset{\sim}{\omega} = \underset{\sim}{\hat{\omega}}, \gamma = 0, \underset{\sim}{y}^c)$, and similarly for N_t. Several points are worth noting from the graphs. First, the variances conditional on estimated parameters are smooth functions of time that are actually symmetric about the midpoint of the series ($t = 84$), while the posterior variances are erratic functions of time. Thus, the effect of parameter uncertainty on the posterior variances is itself an erratic function of time. (Note: The differences between $E[\mathrm{Var}(N_t \mid \underset{\sim}{\omega}, \gamma=0, \underset{\sim}{y}) \mid \underset{\sim}{y}^c]$ and $\mathrm{Var}(N_t \mid \underset{\sim}{\omega} = \underset{\sim}{\hat{\omega}}, \gamma=0, \underset{\sim}{y})$, though small, are responsible for the solid curve dipping slightly below the dotted curve at a few time points in the lower graph of Figure 6.) Second, the posterior variances of N_t tend to be about twice as large as those for S_t (for the canonical decomposition). Finally, both the variances conditional on estimated parameters and the posterior variances increase at either end of the series. This behavior is standard in seasonal adjustment, and arises because signal extraction point estimates can be thought of as deriving from application of symmetric linear filters to the time series y_t extended with forecasts and backcasts. This "forecast extension" implies additional uncertainty near the ends of the series where the symmetric signal extraction filters make heavy use of the forecasts and backcasts in place of the unobserved y_ts.

For a direct measure of the magnitude of the contribution of parameter uncertainty to posterior variances in the canonical decomposition we referred to (2.10), and computed and examined graphs (not shown here) of $\mathrm{Var}[E(S_t \mid \underset{\sim}{\omega}, \gamma=0, \underset{\sim}{y}) \mid \underset{\sim}{y}^c] / \mathrm{Var}(S_t \mid \gamma=0, \underset{\sim}{y}^c)$, and similarly for N_t. These ratios are very erratic over time, as would be expected from Figure 6. The ratios for S_t rarely exceed .2, and are usually less than .1, while those for N_t rarely exceed .1, and are usually less than .05. From these results we conclude that the contribution of parameter uncertainty to the posterior variances in the canonical decomposition is generally small, though erratic over time, and not negligible at all time points. We examined graphs (also not shown here) of the analogous ratios measuring

contribution of parameter uncertainty to posterior variance when $\gamma \sim U(0,1)$. In this case the ratios are also erratic over time but are smaller than those for $\gamma = 0$. For S_t these ratios rarely exceed .04 and for N_t these ratios rarely exceed .03. We conclude for this example that, when $\gamma \sim U(0,1)$, the contribution of parameter uncertainty to the posterior variance is negligible. This was confirmed by an examination of graphs analogous to Figure 6 but for the case $\gamma \sim U(0,1)$, which showed very small differences between the solid and dotted curves. A final point worth mentioning is that these graphs also showed that when $\gamma \sim U(0,1)$ even the variances conditional on estimated parameters are not smooth functions of time. This latter result is due to the contribution of a term, $\mathrm{Var}(I_t^{1/2} | \hat{\omega}, \underset{\sim}{y}^c) + [\mathrm{E}(I_t^{1/2} | \hat{\omega}, \underset{\sim}{y}^c)]^2$, as in (2.13), which is not a smooth function of time, even though it is conditional on the estimated parameters.

Figure 7 compares posterior variances when $\gamma \sim U(0,1)$ with those from the canonical decomposition. The former are always substantially larger than the latter for S_t. For N_t the posterior variances for $\gamma \sim U(0,1)$ are more often than not smaller than those for $\gamma = 0$, but occasionally the posterior variances for $\gamma \sim U(0,1)$ are much larger. This behavior can be explained by reference to (2.13), which shows that in this case $\mathrm{Var}(N_t | \underset{\sim}{y}^c)$ equals $\mathrm{Var}(N_t^{1/2} | \underset{\sim}{y}^c)$ plus additional terms. It turns out that $\mathrm{Var}(N_t^{1/2} | \underset{\sim}{y}^c)$ and $\mathrm{Var}(N_t^0 | \underset{\sim}{y}^c)$ are not very different, so the difference between the two curves in the bottom graph of Figure 7 is due mostly to the "additional terms" from (2.13). The additional terms vary erratically over time, being negative more often than they are positive, but occasionally being positive and large. In any case, Figure 7 shows that posterior variances are very sensitive to decomposition uncertainty.

3.5 Conclusions (for the example)

Based on the preceeding analysis, we can offer the following conclusions for this example.

1. The time series data $\underset{\sim}{y}$ tells us nothing about σ_e^2 beyond the information in the sampling error variance estimates $s_1^2, \ldots, s_{60}^2$.

2. Parameter uncertainty has a very minor effect on point estimates (posterior means) of S_t and N_t. Parameter uncertainty has a generally small effect on posterior variances for the canonical decomposition, but its contribution varies erratically over time and is not always negligible. The contribution of parameter uncertainty to the posterior variances when $\gamma \sim U(0,1)$ is negligible.

3. Decomposition uncertainty has an important effect on posterior means of S_t and N_t, and can have a <u>huge</u> effect on posterior variances of S_t and N_t.

The obvious general conclusion is that careful thought should be devoted to assumptions about the decomposition parameter γ, since (i) what is assumed about γ has a profound effect on the results, and (ii) the data carry no information about γ.

4 Extension: Autocorrelated Sampling Error

Suppose estimates $c_{tj} = \text{Cov}(e_t, e_j)$ t,j = 1,...,m are available, and let $C = [c_{tj}]$ be the m×m matrix of these covariances. The multivariate analog to the scaled χ_ν^2 distribution used in section 3.1 is to assume C has a scaled Wishart distribution with ν degrees of freedom — see DeGroot (1970, section 5.5) for a discussion. The mean of this distribution is the covariance matrix of m successive observations of e_t following the model (2.1) — call this matrix $\Sigma(\underset{\sim}{\alpha}_e)$. Coupling a prior $p(\underset{\sim}{\alpha}_e)$ with the Wishart likelihood, the posterior for $\underset{\sim}{\alpha}_e$ is (for known ν)

$$p(\underset{\sim}{\alpha}_e | \nu, C) \propto p(\underset{\sim}{\alpha}_e)\, p(C | \underset{\sim}{\alpha}_e, \nu)$$

$$= g_2(\nu)\, p(\underset{\sim}{\alpha}_e)\, |\Sigma(\underset{\sim}{\alpha}_e)|^{-\nu/2} \exp\{-(1/2)\mathrm{tr}[\Sigma(\underset{\sim}{\alpha}_e)^{-1}C]\} \qquad (4.1)$$

where $g_2(\nu)$ includes those terms in the Wishart distribution not explicitly present in (4.1) that involve ν. A possible noninformative prior for $\underset{\sim}{\alpha}_e$ is $p(\underset{\sim}{\alpha}_e) \propto 1/\sigma_b^2$. If ν is unknown, the variances $c_{11}, \ldots, c_{mm}$ may be used to estimate ν as in section 3.1, since the $c_{tt} = s_t^2$ have scaled χ^2_ν distributions. Alternatively, we could consider multiplying (4.1) by a prior density $p(\nu)$ and numerically integrating over ν to determine $p(\underset{\sim}{\alpha}_e | C)$. Depending on the dimension of $\underset{\sim}{\alpha}_e$, use of this fully Bayesian approach may be difficult in practice, since simply evaluating $p(\underset{\sim}{\alpha}_e | C)$ for any given $\underset{\sim}{\alpha}_e$ then requires a numerical integration.

If the covariance estimates were of the form $c_{tj} = K^{-1} \sum_{k=1}^{K} (y_{tk} - y_t)(y_{jk} - y_j)$, where the y_{tk} are observations on individual units $k = 1, \ldots, K$ at time t, the survey estimators y_t are the sample means ($y_t = K^{-1} \sum_{k=1}^{K} y_{tk}$), and the y_{tk} for $k = 1, \ldots, K$ are assumed iid normal, then C would indeed have a Wishart distribution. Furthermore, C would be independent of the time series y_t, an assumption used in obtaining (2.5) in section 2. Though estimates of repeated survey variances and covariances (or relative variances and relative covariances) are rarely of this simple form, the Wishart likelihood may still provide a useful approximation. In any case, this approach provides an objective means of combining the information from all the available c_{tj} in developing the posterior for $\underset{\sim}{\alpha}_e$.

As an alternative, let $\tilde{\underset{\sim}{\alpha}}_e$ be a transformation of $\underset{\sim}{\alpha}_e$ with $\log(\sigma_b^2)$ replacing σ_b^2. (We might also consider some transformation applied to a partial correlation type representation of the AR or MA polynomials, if any of the roots of these polynomials is likely to be near the unit circle.) If m is large, and using a flat prior for $\tilde{\underset{\sim}{\alpha}}_e$, we might approximate $p(\tilde{\underset{\sim}{\alpha}}_e | \nu, C)$ by an asymptotic multivariate normal distribution $N(\hat{\tilde{\underset{\sim}{\alpha}}}_e, \hat{V}(\tilde{\underset{\sim}{\alpha}}_e))$,

where $\hat{\underset{\sim}{\alpha}}_e$ maximizes the posterior density, $p(\underset{\sim}{\alpha}_e | \nu, C) \propto p(C | \underset{\sim}{\alpha}_e, \nu)$, and $\hat{V}(\underset{\sim}{\alpha}_e)$ is the negative inverse Hessian (second derivative) matrix of the log–posterior. This normal approximation to the posterior density is then easily evaluated for any given $\underset{\sim}{\alpha}_e$.

One final alternative, that does not rely on the Wishart assumption, would be multivariate time series modeling of $\log(c_{tt})$ and estimated autocorrelations, $r_t(k) = c_{t,t+k}/(c_{tt}\, c_{t+k,t+k})^{1/2}$ (possibly transformed). The estimate of the mean vector from the multivariate time series model, and its associated covariance matrix, can be used to define an asymptotic normal posterior distribution for the log variance and autocorrelations. (The estimate of $\log[\mathrm{Var}(e_t)]$ should be bias adjusted as in section 3.1.) Functionally relating the log variance and autocorrelations to $\underset{\sim}{\alpha}_e$ defines a posterior for $\underset{\sim}{\alpha}_e$. This approach is most feasible when (2.1) is a low order AR model (pe is small), since then only a few (pe) lag correlations need be modeled, and they can be functionally related to the AR parameters and $\mathrm{Var}(e_t)$ through the well–known Yule–Walker equations (Box and Jenkins 1970, pp. 55–56).

5. Extension: Regression Terms in the Signal Model

The generalization of the results of section 2.2 to the model (2.3) with regression terms is straightforward. In this case (2.5) becomes

$$p(\underset{\sim}{\alpha}, \underset{\sim}{\beta}, \underset{\sim}{\alpha}_e | \underset{\sim}{y}, C) \propto p(\underset{\sim}{y} | \underset{\sim}{\alpha}, \underset{\sim}{\beta}, \underset{\sim}{\alpha}_e)\, p(\underset{\sim}{\alpha}_e | C)\, p(\underset{\sim}{\alpha}) .$$

Let $X_D = [x_{Dit}]$ be the (n–12–d) × r matrix of differenced regression variables, $x_{Dit} = (1-B)^d(1-B^{12})x_{it}$. Then

$$p(\underset{\sim}{y}|\underset{\sim}{\alpha}, \underset{\sim}{\beta}, \underset{\sim}{\alpha}_e) = p(\underset{\sim}{w}|\underset{\sim}{\alpha}, \underset{\sim}{\beta}, \underset{\sim}{\alpha}_e)$$

$$\propto |\Sigma_w|^{-1/2} \exp\{(-1/2)(\underset{\sim}{w}-X_D\underset{\sim}{\beta})'\Sigma_w^{-1}(\underset{\sim}{w} - X_D\underset{\sim}{\beta})\}$$

$$= |\Sigma_w|^{-1/2} \exp\{(-1/2)[\underset{\sim}{w} - X_D\hat{\underset{\sim}{\beta}} + X_D(\underset{\sim}{\beta}-\hat{\underset{\sim}{\beta}})]'\Sigma_w^{-1}[\underset{\sim}{w} - X_D\hat{\underset{\sim}{\beta}} + X_D(\underset{\sim}{\beta}-\hat{\underset{\sim}{\beta}})]\}$$

$$= \{|X_D'\Sigma_w^{-1}X_D|^{1/2} \exp\{(-1/2)[\underset{\sim}{\beta} - \hat{\underset{\sim}{\beta}}]'X_D'\Sigma_w^{-1}X_D[\underset{\sim}{\beta} - \hat{\underset{\sim}{\beta}}]\} \times$$
$$\{|\Sigma_w|^{-1/2} |X_D'\Sigma_w^{-1}X_D|^{-1/2} \exp\{(-1/2)[\underset{\sim}{w}-X_D\hat{\underset{\sim}{\beta}}]'\Sigma_w^{-1}[\underset{\sim}{w}-X_D\hat{\underset{\sim}{\beta}}]\} \quad (5.1)$$

where, given $\underset{\sim}{\alpha}$ and $\underset{\sim}{\alpha}_e$, $\hat{\underset{\sim}{\beta}} = E(\underset{\sim}{\beta}|\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \underset{\sim}{y}) = (X_D'\Sigma_w^{-1}X_D)^{-1}X_D'\Sigma_w^{-1}\underset{\sim}{w}$ is the generalized least squares estimate of $\underset{\sim}{\beta}$, with variance $V(\hat{\underset{\sim}{\beta}}) = Var(\underset{\sim}{\beta}|\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \underset{\sim}{y}) = (X_D'\Sigma_w^{-1}X_D)^{-1}$. As noted in Bell and Hillmer (1992), for any given $\underset{\sim}{\alpha}$ and $\underset{\sim}{\alpha}_e$, $\hat{\underset{\sim}{\beta}}$ and $V(\hat{\underset{\sim}{\beta}})$ may be computed by passing $\underset{\sim}{y}$ and the columns of X through the Kalman filter (which includes differencing) to get "innovations" $\tilde{\underset{\sim}{y}}$ and $\tilde{X}$, and then taking $\hat{\underset{\sim}{\beta}} = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{\underset{\sim}{y}}$, $V(\hat{\underset{\sim}{\beta}}) = (\tilde{X}'\tilde{X})^{-1}$.

Examining (5.1), we see its first term is proportional to a $N(\hat{\underset{\sim}{\beta}}, V(\hat{\underset{\sim}{\beta}}))$ density, which is therefore $p(\underset{\sim}{\beta}|\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \underset{\sim}{y}, C)$. The second term in (5.1) must then be proportional to $p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e|\underset{\sim}{y}, C)$. This second term is similar to (2.6), and can be calculated by applying a (suitably initialized) Kalman filter to $\underset{\sim}{w} - X_D\hat{\underset{\sim}{\beta}}$ to evaluate a $N(X_D\hat{\underset{\sim}{\beta}}, \Sigma_w^{-1})$ density for $\underset{\sim}{w}$, and then multiplying this by the factor $|X_D'\Sigma_w^{-1}X_D|^{-1/2} = |\tilde{X}'\tilde{X}|^{-1/2}$, which is readily calculated given $\tilde{X}$.

To perform signal extraction for the model (2.3) we must first decide how to assign the regression effects, $\underset{\sim}{x}_t'\underset{\sim}{\beta} = \sum_i^r x_{it}\beta_i$, to the seasonal and nonseasonal components. This is discussed in Bell (1984b). Assume the regression effects are broken down so that $y_t = (S_t + \underset{\sim}{x}_{1t}'\underset{\sim}{\beta}_1) + (N_t + \underset{\sim}{x}_{2t}'\underset{\sim}{\beta}_2) + e_t \equiv S_{xt} + N_{xt} + e_t$, where $\underset{\sim}{\beta} = (\underset{\sim}{\beta}_1', \underset{\sim}{\beta}_2')'$ and

$x_t = (\underset{\sim}{x}'_{1t}, \underset{\sim}{x}'_{2t})'$. Then, we wish to obtain posterior means and variances of S_{xt} and N_{xt}. From results in Bell and Hillmer (1992), it turns out that $E(S_{xt} | \underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y}) = E(S_t | \underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y}) + \underset{\sim}{x}'_{1t}\hat{\underset{\sim}{\beta}}_1$, where $\hat{\underset{\sim}{\beta}} = (\hat{\underset{\sim}{\beta}}'_1, \hat{\underset{\sim}{\beta}}'_2)'$ is the GLS estimate of $\underset{\sim}{\beta}$ for given $\underset{\sim}{\alpha}$, $\underset{\sim}{\alpha}_e$ discussed above. Also, $\mathrm{Var}(S_{xt} | \underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y})$ can be computed as discussed in Bell and Hillmer (1992).

Given that $E(S_{xt} | \underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y})$ and $\mathrm{Var}(S_{xt} | \underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \gamma, \underset{\sim}{y})$ can be computed (the REGCMPNT program will do this), we can compute $E(S_{xt} | \underset{\sim}{y}^c)$ and $\mathrm{Var}(S_{xt} | \underset{\sim}{y}^c)$ (and similarly for N_{xt}) by Monte Carlo integration as discussed in section 2.4. We again encounter the problem of how to simulate from a posterior distribution of nonstandard form. This problem now arises in regard to $p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e | \underset{\sim}{y}, C)$ given by the second term in (5.1). As discussed in section (2.4), this problem might be addressed by using an asymptotic normal approximation to $p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e | \underset{\sim}{y}, C)$ (now obtained using the expression in (5.1)), or by specialized Bayesian simulation techniques such as importance sampling. Since the conditional posterior distribution of β given $\underset{\sim}{\alpha}$ and $\underset{\sim}{\alpha}_e$ is exactly $N(\hat{\underset{\sim}{\beta}}, V(\hat{\underset{\sim}{\beta}}))$, once we have simulated $(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e)$ we can easily simulate $\underset{\sim}{\beta}$.

Appendix: Derivation of Results (2.12) and (2.13)

We use the decompostion (2.8): $Y_t = (\bar{S}_t + \varepsilon_t^{\gamma}) + \bar{T}_t + I_t^{1-\gamma} \equiv S_t^{\gamma} + N_t^{\gamma}$, where ε_t^{γ} and $I_t^{1-\gamma}$ are i.i.d. $N(0, \gamma\bar{\sigma}_I^2)$ and $N(0, (1-\gamma)\bar{\sigma}_I^2)$, respectively, and all four components of Y_t are independent of one another. We assume $\gamma \sim U(0,1)$ independently of all other random variables in the problem.

Let $\tilde{E}$ and $\tilde{V}$ denote $E(\cdot | \underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \underset{\sim}{y})$ and $\mathrm{Var}(\cdot | \underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e, \underset{\sim}{y})$. (Also, $\tilde{E}$ and $\tilde{V}$ condition on regression parameters $\underset{\sim}{\beta}$, if these are present; see section 5.) Let $\underset{\sim}{\varepsilon}^{\gamma} = (\varepsilon_1^{\gamma}, \ldots, \varepsilon_n^{\gamma})'$. Bell and Hillmer (1988, equations (4.3) and (4.5)) give results for signal extraction of a stationary component in a nonstationary time series. For signal extraction of the white

noise series ε_t^{γ}, these results specialize to

$$\tilde{E}(\underset{\sim}{\varepsilon}^{\gamma}|\,\gamma) = \gamma\bar{\sigma}_I^2\Delta'\Sigma_w^{-1}\underset{\sim}{w} \qquad \tilde{V}(\underset{\sim}{\varepsilon}^{\gamma}|\,\gamma) = \gamma\bar{\sigma}_I^2 I_{n\times n} - \gamma^2\bar{\sigma}_I^4\Delta'\Sigma_w^{-1}\Delta \tag{A.1}$$

where Δ is an (n–12–d)×n differencing matrix defined so that $\Delta\underset{\sim}{y} = \underset{\sim}{w}$, and $I_{n\times n}$ is the n×n identity matrix. As discussed in section 2.4, the results (A.1) hold under either assumption A of Bell (1984a) or the diffuse prior assumptions of Ansley and Kohn (1985).

Upon setting $\gamma = 1$, ε_t^{γ} has the same stochastic structure as $\bar{I}_t$ (i.i.d. $N(0,\bar{\sigma}_I^2)$ and independent of the other components). Thus, $\tilde{E}(\varepsilon_t^{\gamma}|\,\gamma{=}1) = \tilde{E}(\bar{I}_t)$ and $\tilde{V}(\varepsilon_t^{\gamma}|\,\gamma{=}1) = \tilde{V}(\bar{I}_t)$. Therefore, after minor manipulation, (A.1) implies that

$$\tilde{E}(\varepsilon_t^{\gamma}|\,\gamma) = \gamma\tilde{E}(\bar{I}_t) \qquad \tilde{V}(\varepsilon_t^{\gamma}|\,\gamma) = \gamma^2\tilde{V}(\bar{I}_t) + \gamma(1{-}\gamma)\bar{\sigma}_I^2. \tag{A.2}$$

We can now obtain $\tilde{E}(S_t^{\gamma})$ and then $E(S_t|\underset{\sim}{y}^c)$. Since $\gamma \sim U(0,1)$ has mean 1/2, and $\tilde{E}(\bar{I}_t)$ does not depend on γ, we have that $\tilde{E}(\varepsilon_t^{\gamma}) = E_{\gamma}[\tilde{E}(\varepsilon_t^{\gamma}|\,\gamma)] = E_{\gamma}[\gamma\tilde{E}(\bar{I}_t)] = E_{\gamma}(\gamma)\tilde{E}(\bar{I}_t) = (1/2)\tilde{E}(\bar{I}_t) = \tilde{E}(\varepsilon_t^{1/2})$, the last equality following from (A.2) on setting $\gamma = 1/2$. (E_{γ} denotes expectation taken over the distribution of γ.) Therefore, $\tilde{E}(S_t^{\gamma}) = \tilde{E}(\bar{S}_t + \varepsilon_t^{\gamma}) = \tilde{E}(\bar{S}_t) + \tilde{E}(\varepsilon_t^{1/2}) = \tilde{E}(S_t^{1/2})$. Then, letting $\bar{E}(\cdot)$ denote expectation over the posterior distribution for $(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e)$, $E(S_t^{\gamma}|\underset{\sim}{y}^c) = \bar{E}[\tilde{E}(S_t^{\gamma})] = \bar{E}[\tilde{E}(S_t^{1/2})] = E(S_t^{1/2}|\underset{\sim}{y}^c)$, which establishes (2.12).

To obtain $\mathrm{Var}(S_t^{\gamma}|\underset{\sim}{y}^c)$ requires more preliminary results. First, note that

$$\tilde{V}(S_t^{\gamma}) = \mathrm{Var}_{\gamma}[\tilde{E}(S_t^{\gamma}|\,\gamma)] + E_{\gamma}[\tilde{V}(S_t^{\gamma}|\,\gamma)] \tag{A.3}$$

where $\mathrm{Var}_{\gamma}(\cdot)$ is taken over the distribution of γ. The first term in (A.3) is, using (A.2)

$$\mathrm{Var}_\gamma[\tilde{E}(\bar{S}_t + \varepsilon_t^\gamma | \gamma)] = \mathrm{Var}_\gamma[\tilde{E}(\bar{S}_t) + \gamma\tilde{E}(\bar{I}_t)]. \tag{A.4}$$

Since $\tilde{E}(\bar{S}_t)$ and $\tilde{E}(\bar{I}_t)$ do not depend on γ, $\mathrm{Var}_\gamma[\tilde{E}(\bar{S}_t)] = 0$ and (A.4) becomes (note $\mathrm{Var}_\gamma(\gamma) = 1/12$)

$$\mathrm{Var}_\gamma[\tilde{E}(S_t^\gamma | \gamma)] = \mathrm{Var}_\gamma(\gamma)[\tilde{E}(\bar{I}_t)]^2 = (1/12)[\tilde{E}(\bar{I}_t)]^2 = (1/3)[\tilde{E}(\varepsilon_t^{1/2})]^2. \tag{A.5}$$

Now, $\tilde{V}(S_t^\gamma | \gamma) = \tilde{V}(\bar{S}_t + \varepsilon_t^\gamma | \gamma) = \tilde{V}(\bar{S}_t) + 2\tilde{C}ov(\bar{S}_t, \varepsilon_t^\gamma | \gamma) + \tilde{V}(\varepsilon_t^\gamma | \gamma)$, where $\tilde{C}ov(\cdot)$ indicates a covariance conditional on $\underset{\sim}{\alpha}$, $\underset{\sim}{\alpha}_e$, and $\underset{\sim}{y}$. From results of Bell and Hillmer (1988), it can be shown that $\tilde{C}ov(\bar{S}_t, \varepsilon_t^\gamma | \gamma) = \gamma\tilde{C}ov(\bar{S}_t, \bar{I}_t)$. Therefore, using (A.2), $\tilde{V}(S_t^\gamma | \gamma) = \tilde{V}(\bar{S}_t) + 2\gamma\tilde{C}ov(\bar{S}_t, \bar{I}_t) + \gamma^2\tilde{V}(\bar{I}_t) + \gamma(1-\gamma)\bar{\sigma}_I^2$. For $\gamma \sim U(0,1)$ note that $E(\gamma^2) = 1/3$ and $E[\gamma(1-\gamma)] = 1/6$. Therefore,

$$\begin{aligned} E_\gamma[\tilde{V}(S_t^\gamma | \gamma)] &= \tilde{V}(\bar{S}_t) + 2E_\gamma(\gamma)\tilde{C}ov(\bar{S}_t, \bar{I}_t) + E_\gamma(\gamma^2)\tilde{V}(\bar{I}_t) + E_\gamma[\gamma(1-\gamma)]\bar{\sigma}_I^2 \\ &= \tilde{V}(\bar{S}_t) + 2(1/2)\tilde{C}ov(\bar{S}_t, \bar{I}_t) + (1/3)\tilde{V}(\bar{I}_t) + (1/6)\bar{\sigma}_I^2 \\ &= \tilde{V}(\bar{S}_t) + 2\tilde{C}ov(\bar{S}_t, \varepsilon_t^{1/2}) + (4/3)\tilde{V}(\varepsilon_t^{1/2}) - (1/3)\bar{\sigma}_I^2 + (1/6)\bar{\sigma}_I^2 \end{aligned}$$

using (A.2) with $\gamma = 1/2$. Then

$$\begin{aligned} E_\gamma[\tilde{V}(S_t^\gamma | \gamma)] &= \tilde{V}(\bar{S}_t + \varepsilon_t^{1/2}) + (1/3)\tilde{V}(\varepsilon_t^{1/2}) - (1/6)\bar{\sigma}_I^2 \\ &= \tilde{V}(S_t^{1/2}) + (1/3)\tilde{V}(I_t^{1/2}) - (1/6)\bar{\sigma}_I^2 \end{aligned} \tag{A.6}$$

since $I_t^{1/2}$ and $\varepsilon_t^{1/2}$ have the same stochastic characteristics. From (A.3), (A.5), and (A.6) we have

$$\tilde{V}(S_t^\gamma) = (1/3)[\tilde{E}(I_t^{1/2})]^2 + \tilde{V}(S_t^{1/2}) + (1/3)\tilde{V}(I_t^{1/2}) - (1/6)\bar{\sigma}_I^2. \tag{A.7}$$

Letting $\bar{V}(\cdot)$ denote variance over $p(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}_e|\underset{\sim}{y}, C)$, from (2.12) and (A.7) we have that

$$\begin{aligned}
\mathrm{Var}(S_t^{\gamma}|\underset{\sim}{y}^c) &= \bar{V}[\tilde{E}(S_t^{\gamma})] + \bar{E}[\tilde{V}(S_t^{\gamma})] \\
&= \bar{V}[\tilde{E}(S_t^{1/2})] + \bar{E}\{\tilde{V}(S_t^{1/2}) + (1/3)\tilde{V}(I_t^{1/2}) + (1/3)[\tilde{E}(I_t^{1/2})]^2 - \bar{\sigma}_I^2/6\} \\
&= \mathrm{Var}(S_t^{1/2}|\underset{\sim}{y}^c) + (1/3)\{\mathrm{Var}(I_t^{1/2}|\underset{\sim}{y}^c) - \bar{V}[\tilde{E}(I_t^{1/2})]\} + (1/3)\bar{E}[\tilde{E}(I_t^{1/2})]^2 - \bar{E}(\bar{\sigma}_I^2/6) \\
&= \mathrm{Var}(S_t^{1/2}|\underset{\sim}{y}^c) + (1/3)\mathrm{Var}(I_t^{1/2}|\underset{\sim}{y}^c) + (1/3)\{\bar{E}[\tilde{E}(I_t^{1/2})]\}^2 - E(\bar{\sigma}_I^2|\underset{\sim}{y}^c)/6 \\
&= \mathrm{Var}(S_t^{1/2}|\underset{\sim}{y}^c) + (1/3)\{\mathrm{Var}(I_t^{1/2}|\underset{\sim}{y}^c) + [E(I_t^{1/2}|\underset{\sim}{y}^c)]^2\} - E(\bar{\sigma}_I^2|\underset{\sim}{y}^c)/6
\end{aligned}$$

which is (2.13).

When $\gamma = 1/2$, $1 - \gamma = 1/2$ and $N_t^{1/2} = \bar{T}_t + I_t^{1/2}$. Thus, $N_t^{\gamma} = \bar{T}_t + I_t^{1-\gamma}$ and $S_t^{\gamma} = \bar{S}_t + \varepsilon_t^{\gamma}$ have parallel roles, and the above argument establishes the analogous results for N_t:

$$E(N_t^{\gamma}|\underset{\sim}{y}^c) = E(N_t^{1/2}|\underset{\sim}{y}^c) = E(\bar{T}_t|\underset{\sim}{y}^c) + E(I_t^{1/2}|\underset{\sim}{y}^c)$$

$$\mathrm{Var}(N_t^{\gamma}|\underset{\sim}{y}^c) = \mathrm{Var}(N_t^{1/2}|\underset{\sim}{y}^c) + (1/3)\{\mathrm{Var}(I_t^{1/2}|\underset{\sim}{y}^c) + [E(I_t^{1/2}|\underset{\sim}{y}^c)]^2\} - E(\bar{\sigma}_I^2|\underset{\sim}{y}^c)/6.$$

REFERENCES

Abramowitz, M. and Stegun, I. A. (1970) Handbook of Mathematical Functions (9th ed.), National Bureau of Standards, Applied Mathematics Series No. 55, Washington, D. C. : U. S. Government Printing Office.

Anderson, B.D.O. and Moore, J. B. (1979), Optimal Filtering, Englewood Cliffs: Prentice–Hall.

Ansley, C. F. and Kohn, R. (1985) "Estimation, Filtering, and Smoothing in State Space Models with Incompletely Specified Initial Conditions," Annals of Statistics, 13, 1286–1316.

Bell, W. R. (1984a) "Signal Extraction for Nonstationary Time Series," Annals of Statistics, 12, 646–664.

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

_____ (1989) "Sources of Uncertainty in Seasonal Adjustment", Comment on Carlin, J. B., and Dempster, A. P. 'Sensitivity Analysis of Seasonal Adjustments: Empirical Case Studies,' Journal of the American Statistical Association, 84, 22–24.

Bell, W. R. and Hillmer, S. C. (1984), "Issues Involved with the Seasonal Adjustment of Economic Time Series," (with discussion), Journal of Business and Economic Statistics, 2, 291–320.

_____ (1988), "A Matrix Approach to Signal Extraction and Likelihood Evaluation for ARIMA Component Time Series Models," Research Report Number 88/22, Statistical Research Division, Bureau of the Census.

_____ (1990), "The Time Series Approach to Estimation for Repeated Surveys," Survey Methodology, 16, 195–215.

_____ (1991), "Initializing the Kalman Filter for Nonstationary Time Series Models," Journal of Time Series Analysis, 12, 283–300.

_____ (1992), "Applying Time Series Models in Survey Estimation," submitted for publication.

Binder, D. A. and Dick, J. P. (1989), "Modelling and Estimation for Repeated Surveys," Survey Methodology, 15, 29–45.

Box, G.E.P. and Jenkins, G. M. (1970), Time Series Analysis: Forecasting and Control, San Francisco: Holden Day.

Box, G.E.P. and Tiao, G. C. (1973) Bayesian Inference in Statistical Analysis, Reading, Massachusetts: Addison–Wesley.

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

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.

Chib, S. and Greenberg, E. (1992) "Bayes Inference in Regression Models with ARMA(p,q) Errors," Technical Report, Washington University — St. Louis.

DeGroot, M. H. (1970) Optimal Statistical Decisions, New York: McGraw–Hill.

Gelfand, A. E. and Smith, A. F. M. (1990) "Sampling–Based Approaches to Calculating Marginal Densities," Journal of the American Statistical Association, 85, 398–409.

Geweke, J. (1989) "Bayesian Inference in Econometric Models Using Monte Carlo Integration," Econometrica, 57, 1317–1339.

Hausman, J. A. and Watson, M. W. (1985), "Errors in Variables and Seasonal Adjustment Procedures," Journal of the American Statistical Association, 80, 531–540.

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.

Jenkins, G. M. and Watts, D. G. (1968) Spectral Analysis and Its Applications, San Francisco: Holden–Day.

Kohn, R. and Ansley, C. F. (1987), "Signal Extraction for Finite Nonstationary Time Series," Biometrika, 74, 411–421.

Marriott, J., Ravishanker, N., Gelfand, A., and Pai, J. (1992) "Bayesian Analysis of ARMA Processes: Complete Sampling Based Inference Under Full Likelihoods," Technical Report, University of Connecticut.

Pfeffermann, D. (1990) "A General Method for Estimating the Variances of X–11 Estimators," American Statistical Association, Proceedings of the Survey Research Methods Section, 634–639.

_____ (1991) "Estimation and Seasonal Adjustment of Population Means Using Data from Repeated Surveys," Journal of Business and Economic Statistics, 9, 163–175.

Ripley, B. (1987) Stochastic Simulation, New York: John Wiley.

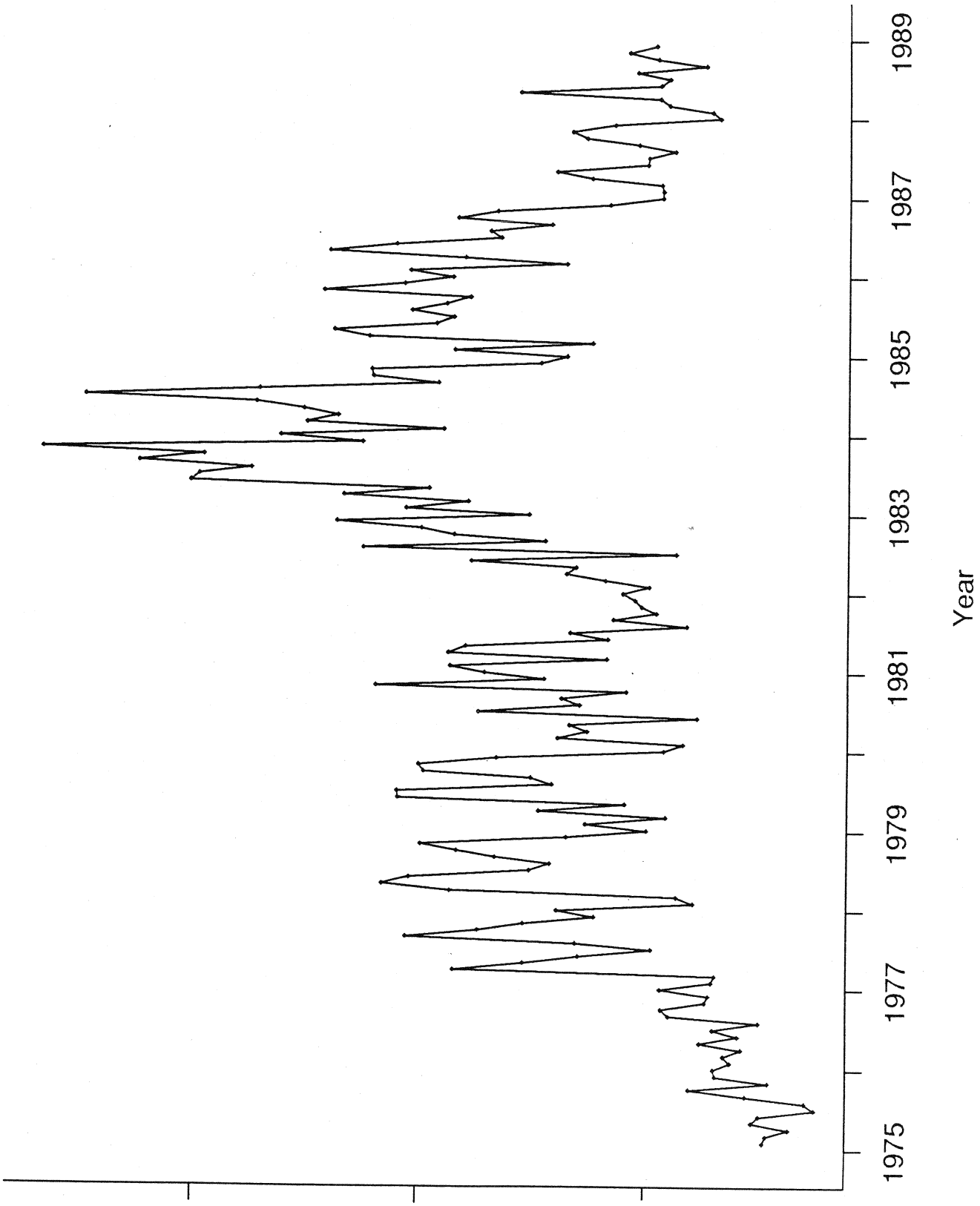
Scott, A. J. and Smith, T.M.F. (1974), "Analysis of Repeated Surveys Using Time Series Methods," Journal of the American Statistical Association, 69, 674–678.

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.

Wolter, K. M. and Monsour, N. J. (1981), "On the Problem of Variance Estimation for a Deseasonalized Series," in Current Topics in Survey Sampling, ed. D. Krewski, R. Platek, and J.N.K. Rao, New York: Academic Press, 199–226.

Figure 1. South 5+ Housing Starts



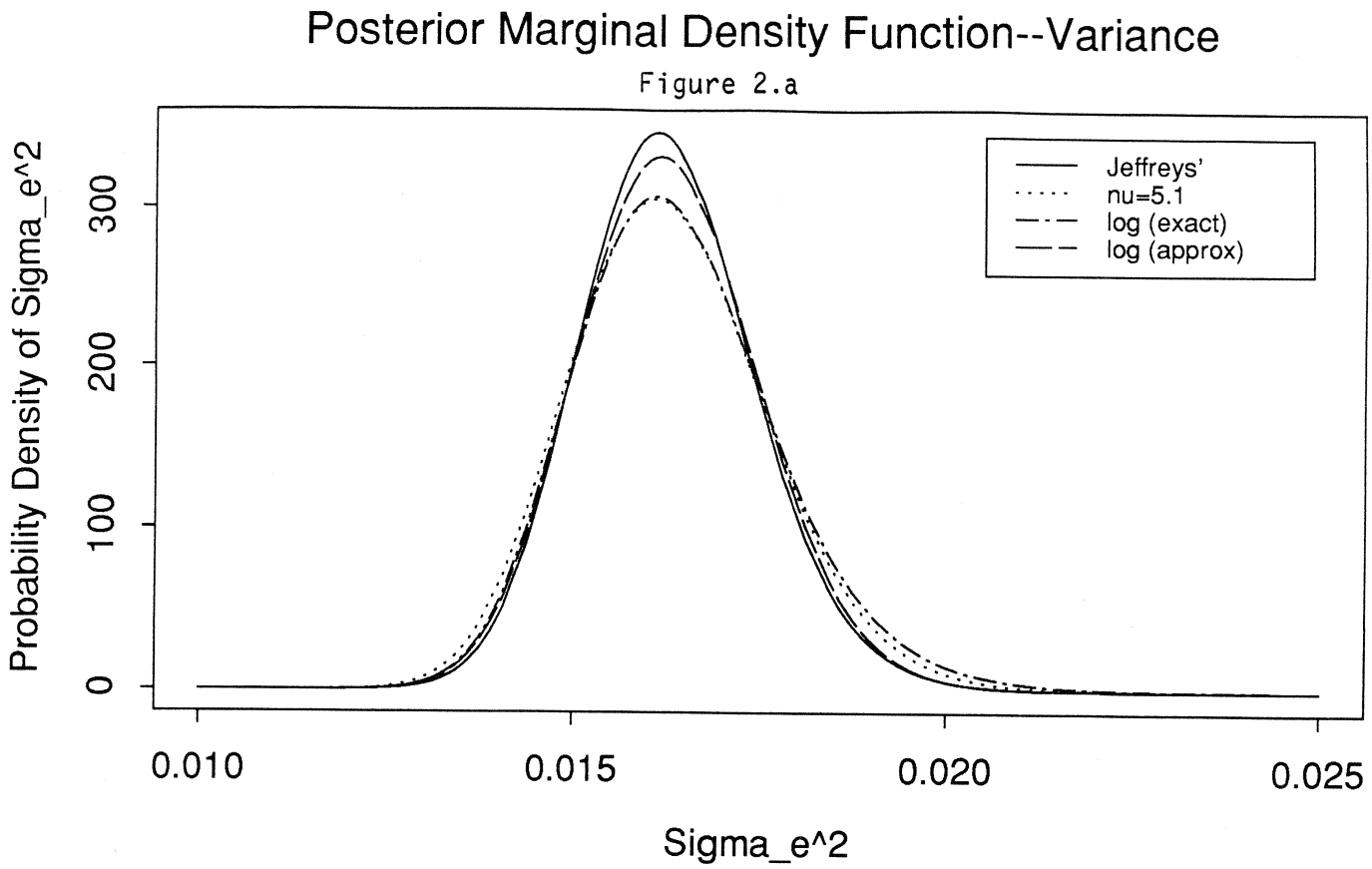


Figure 2.a

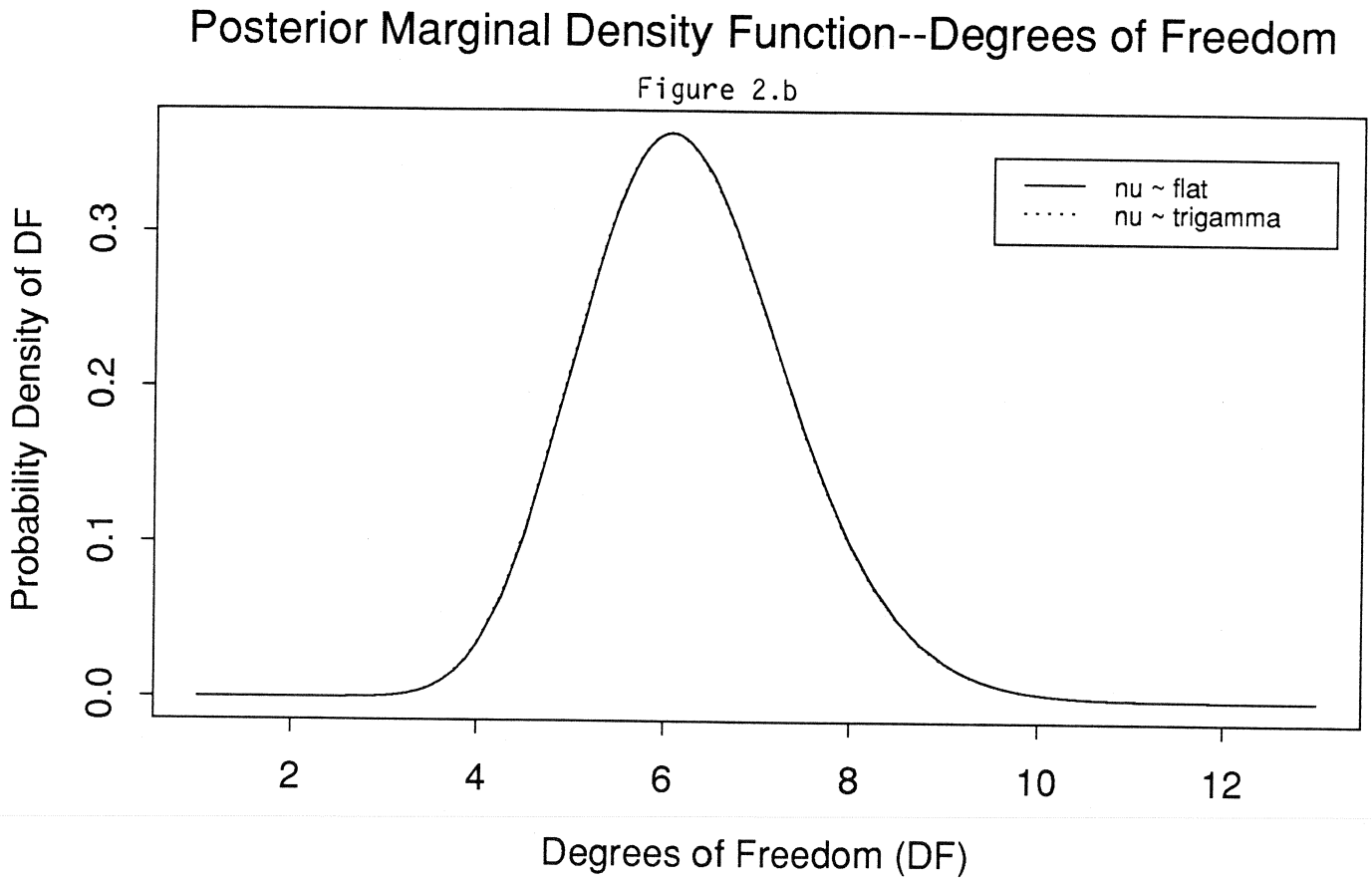
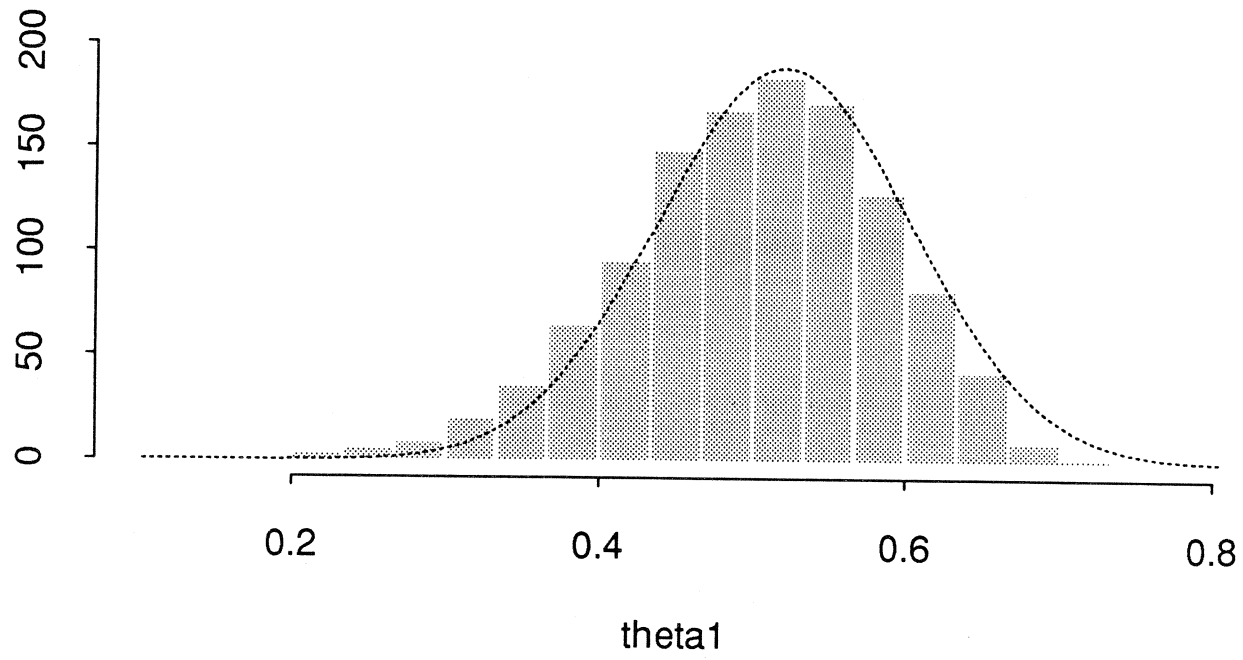


Figure 2.b

Figure 3.

Marginal Posteriors: Histogram = Simulations from Exact Posterior

Dotted Line = Normal Approximation to Posterior Density



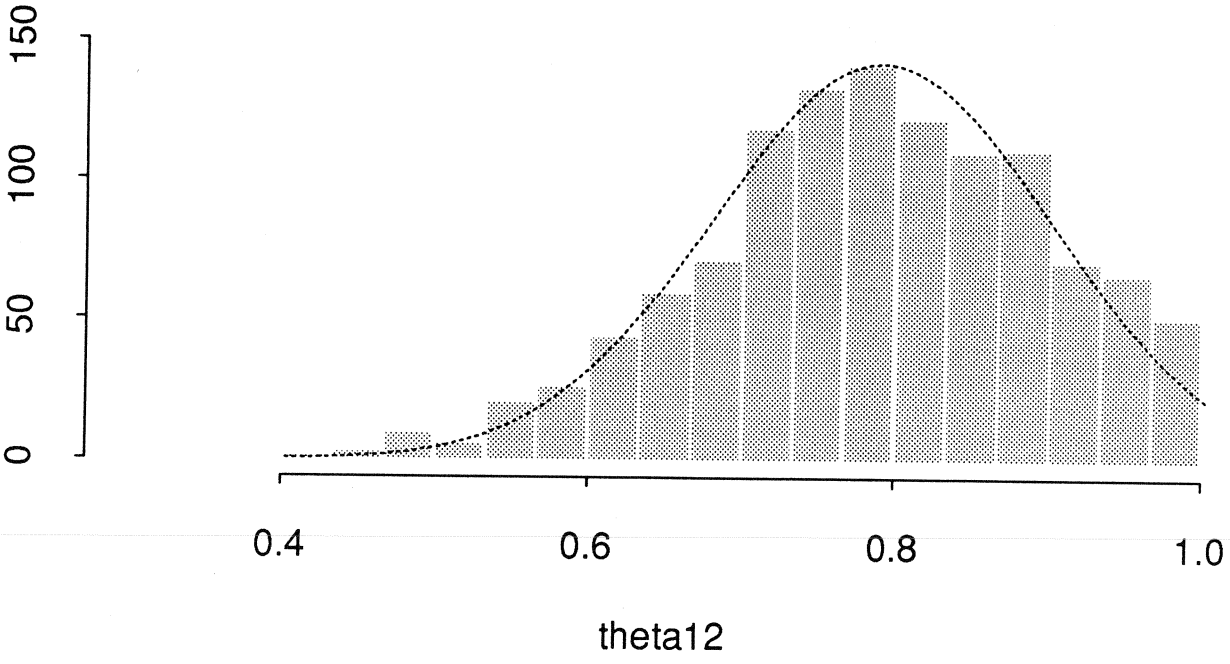
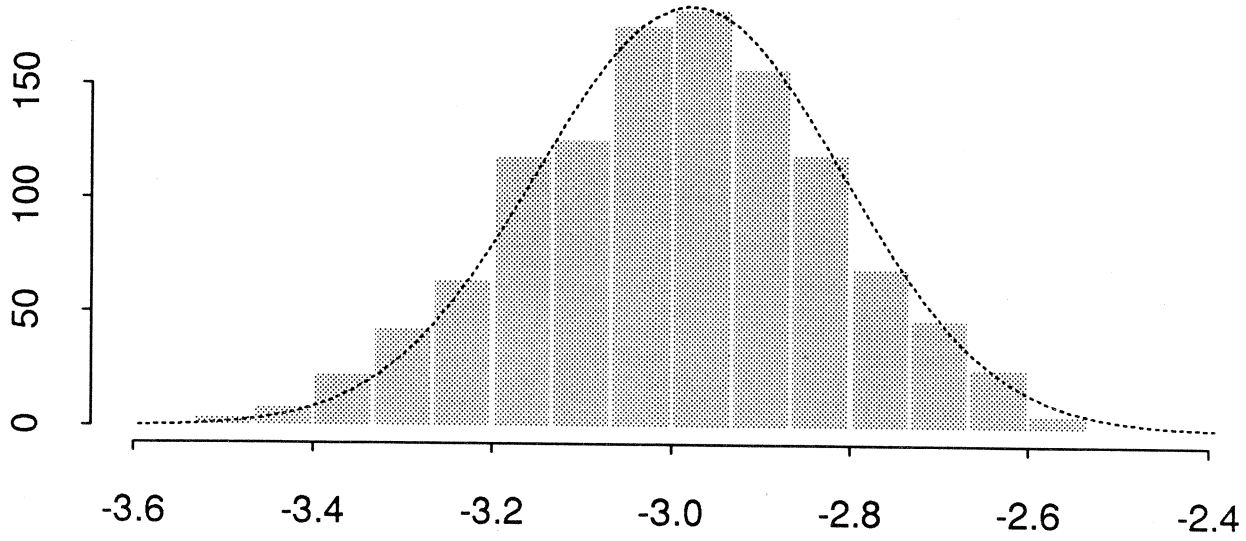


Figure 3. (continued)

Marginal Posteriors: Histogram = Simulations from Exact Posterior

Dotted Line = Normal Approximation to Posterior Density



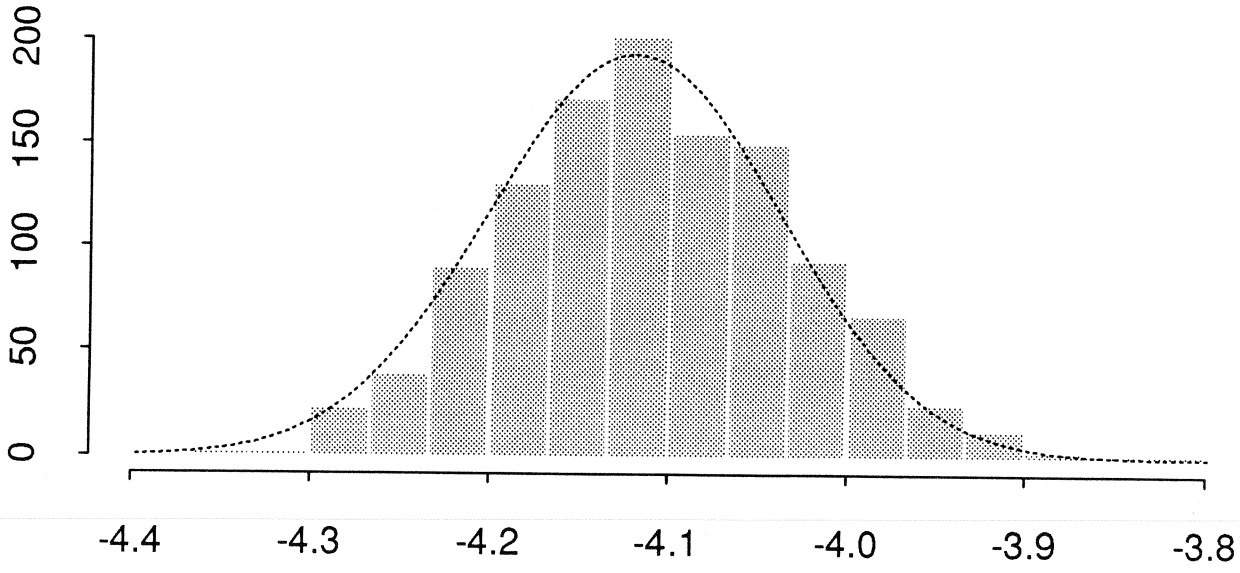
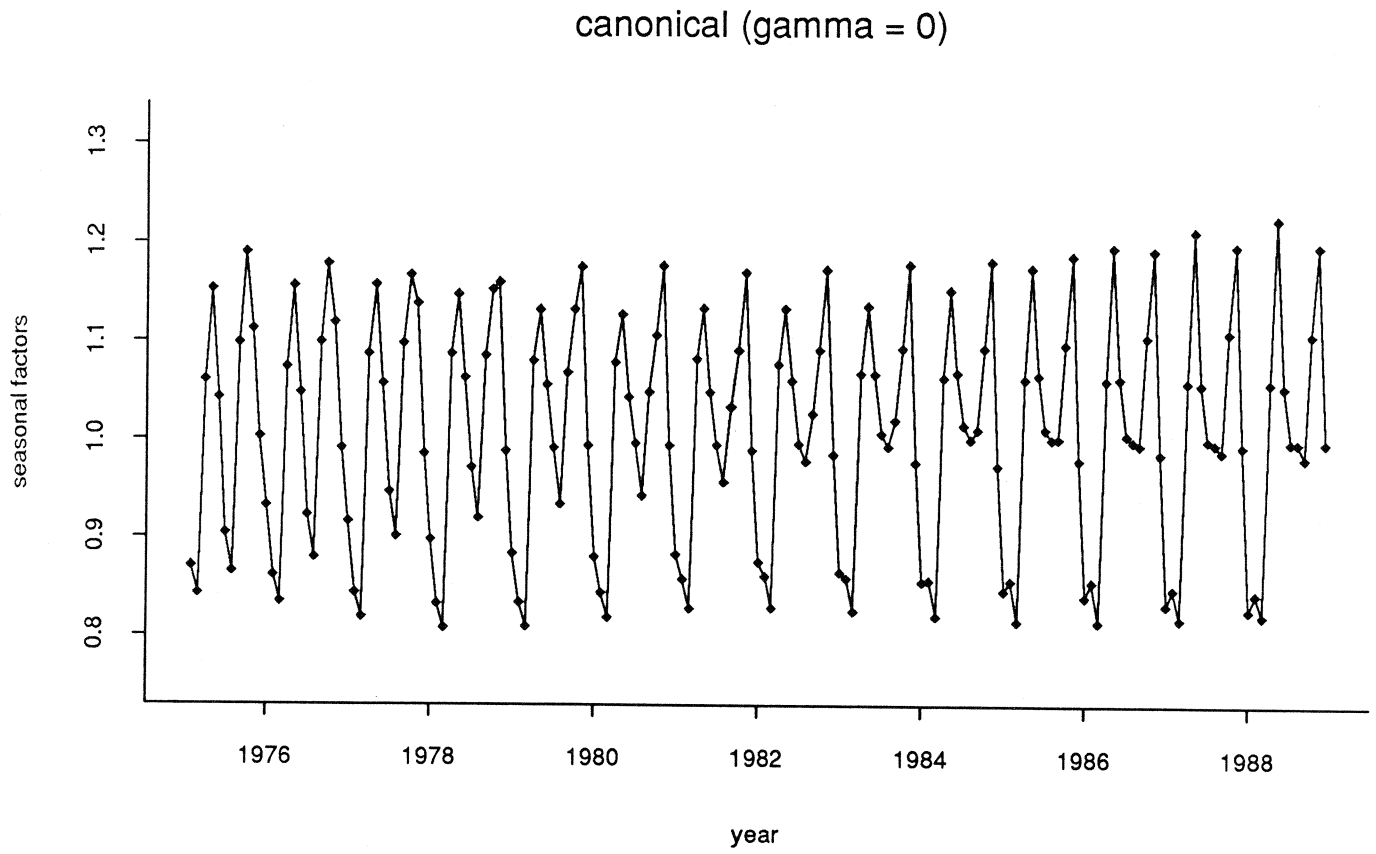


Figure 4. Seasonal Posterior Means



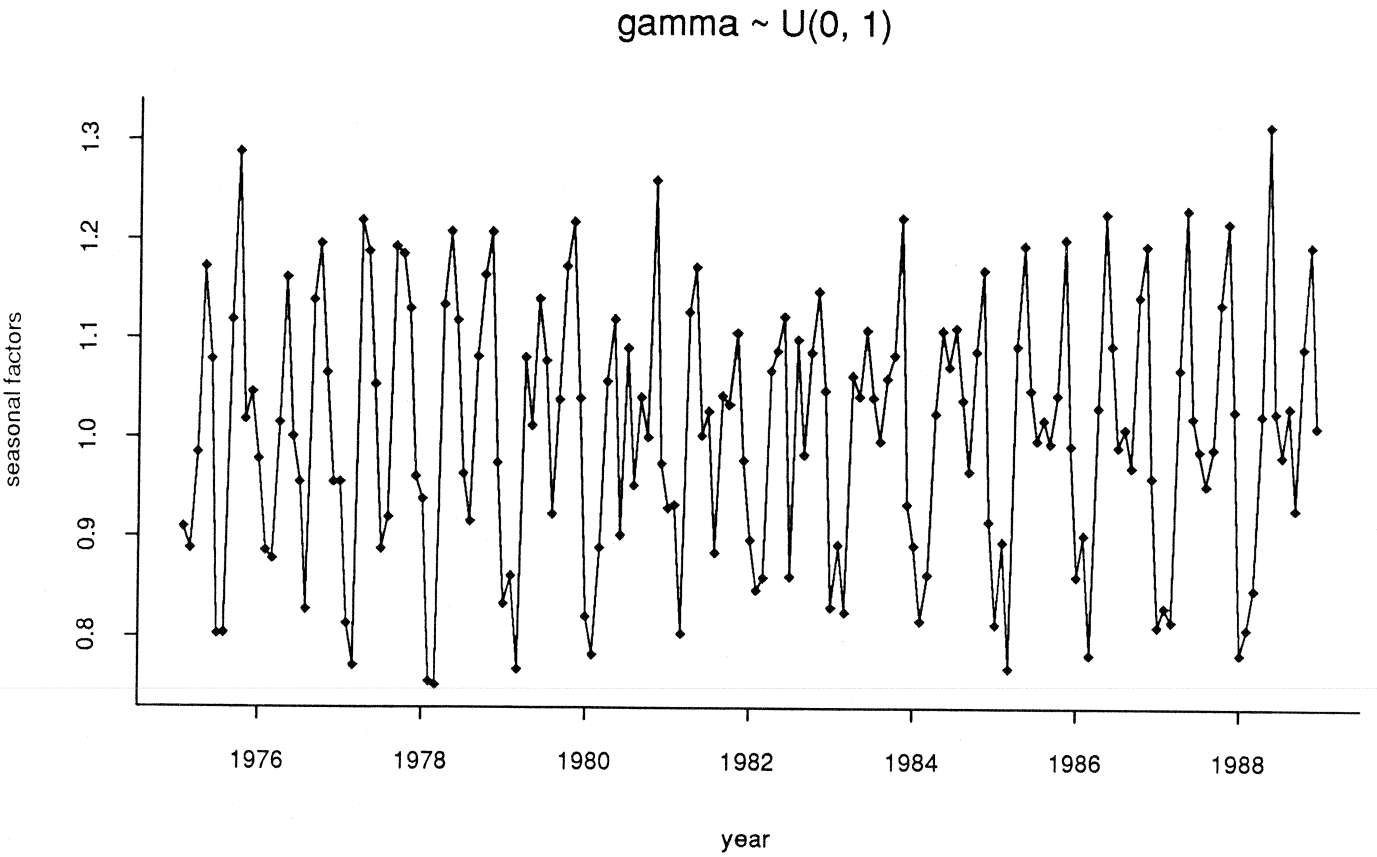
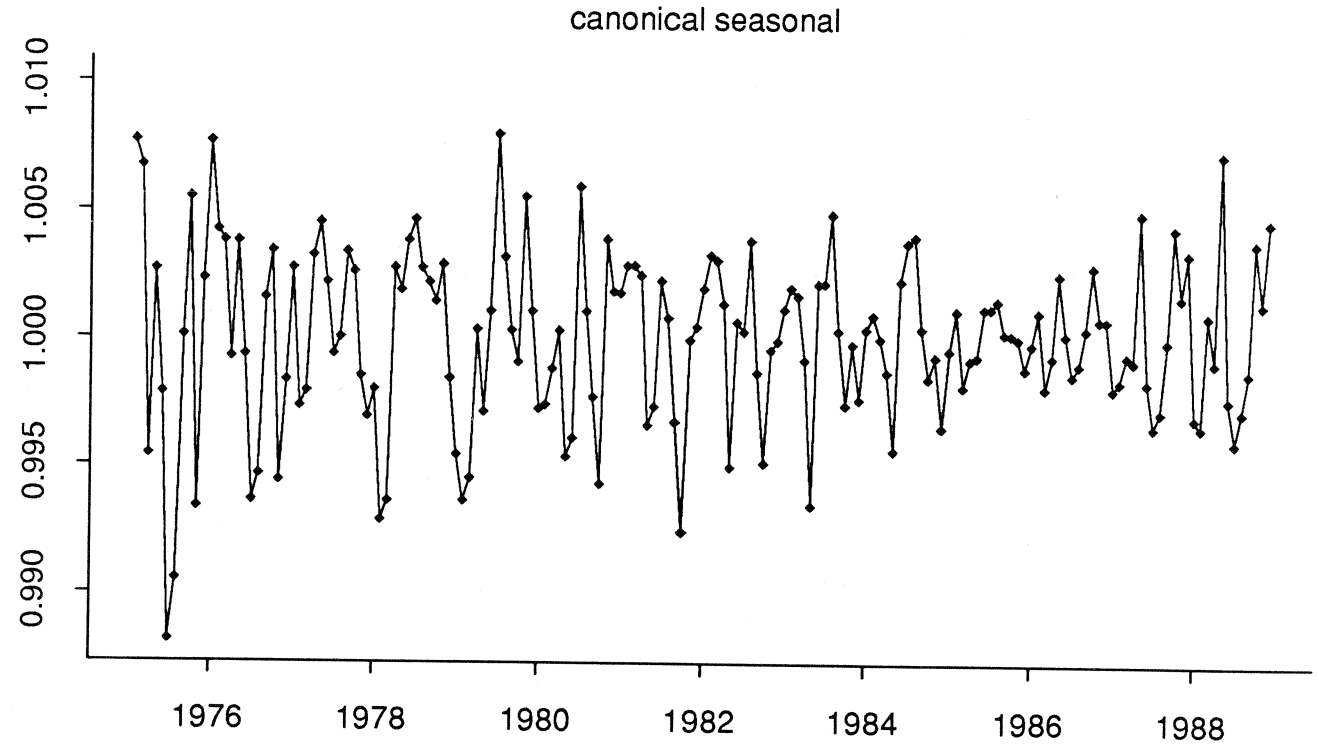


Figure 5. Ratio of Posterior Means to Signal Extraction Estimates Conditional on Parameters



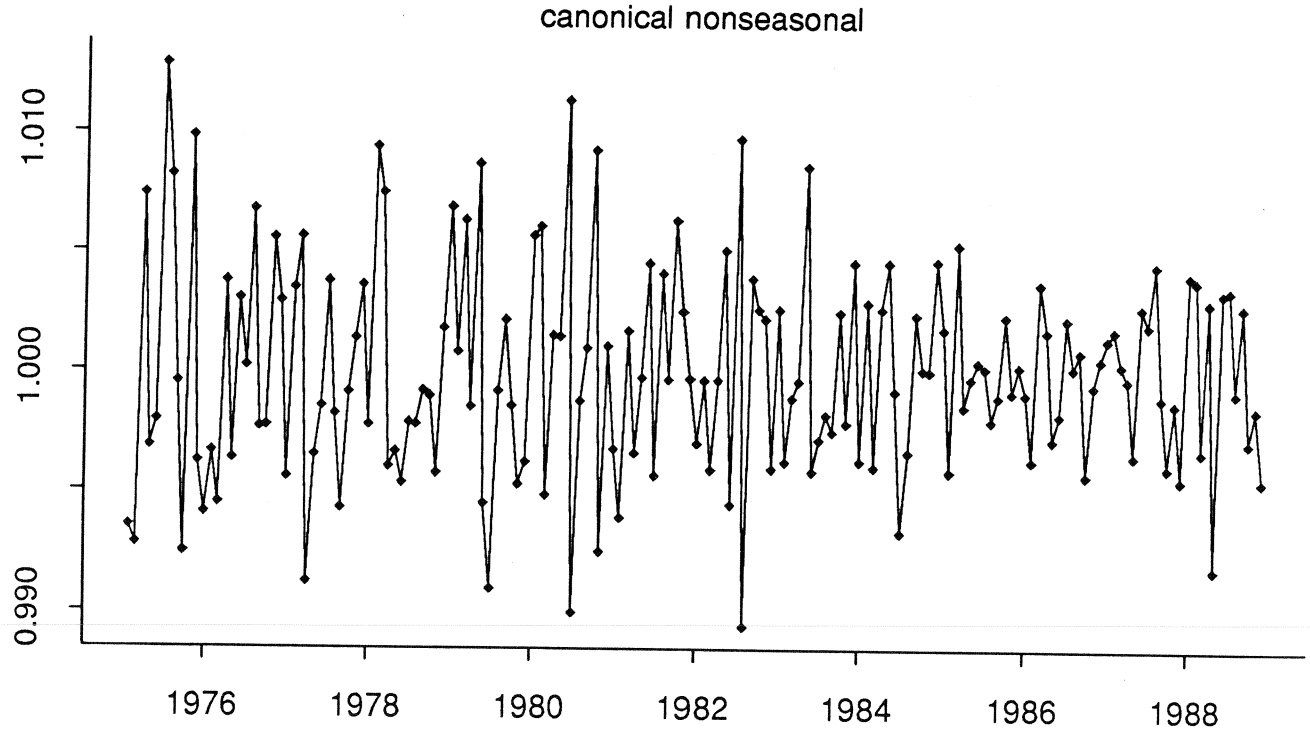
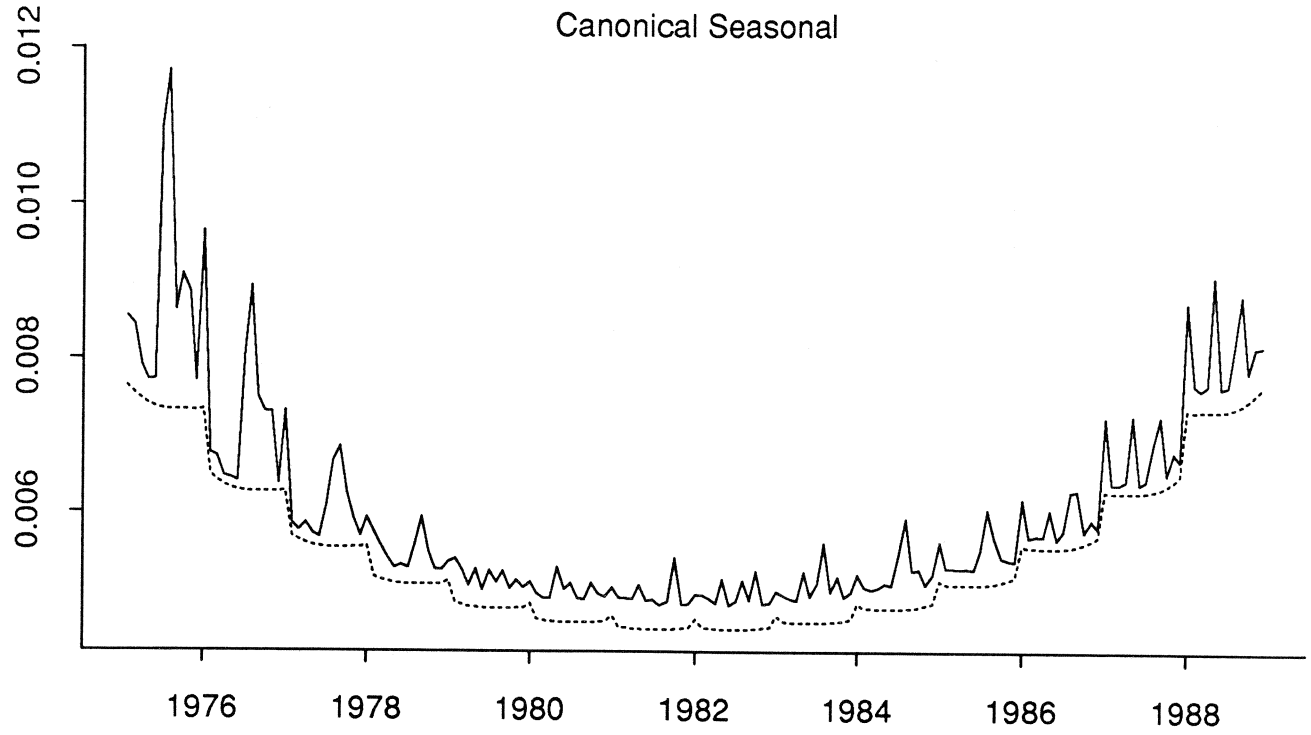


Figure 6. Posterior Variances (solid line) and Variances Conditional on Estimated Parameters (dotted line)



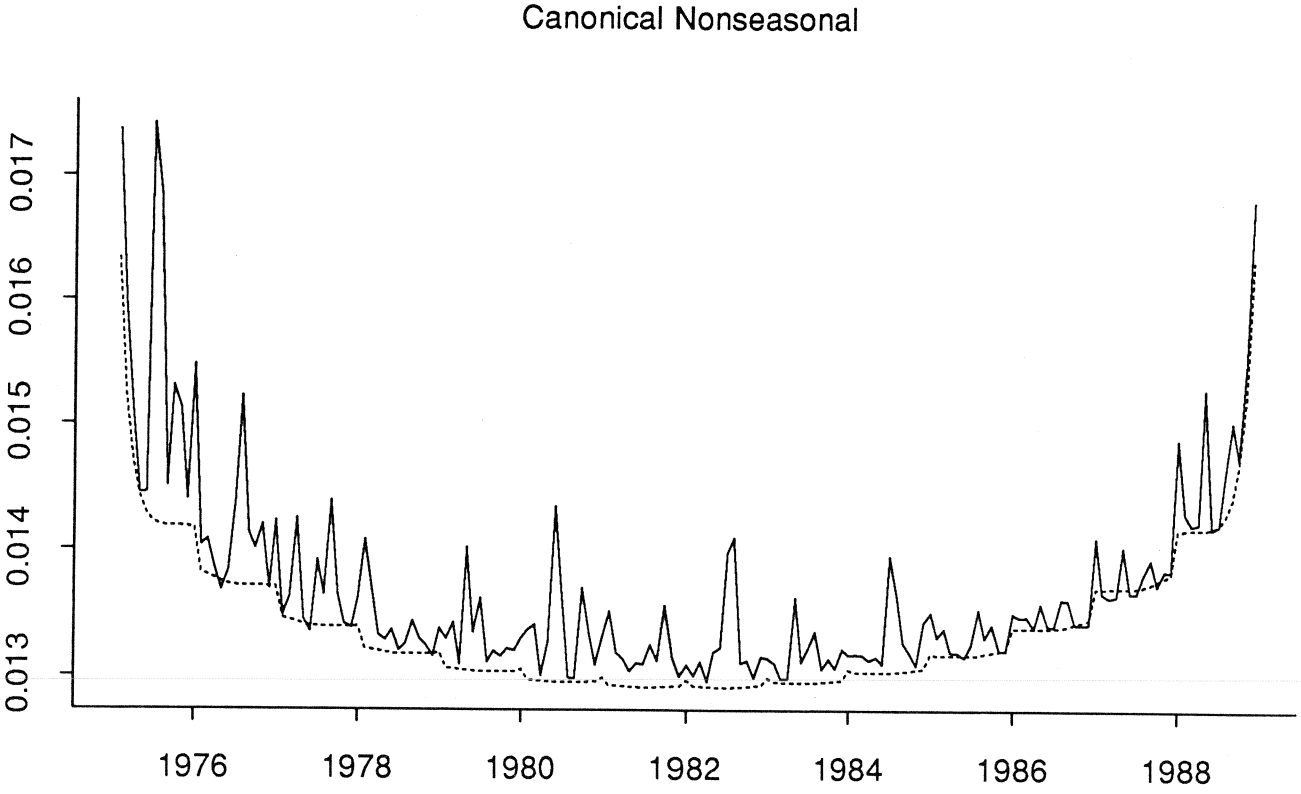


Figure 7. Posterior Variances when gamma ~ U(0, 1) (solid line) and for the Canonical Decompostion (dotted line)