Chapter 1

Bayesian Seasonal Adjustment of Long Memory Time Series

Scott H. Holan and Tucker S. McElroy[1]

1.1 Introduction

Existing approaches to the seasonal adjustment of economic time series are typically either nonparametric or model-based. In both cases, the goal is to remove seasonal variation from the time series. In each of the two paradigms, both the seasonally adjusted series and the seasonal component are latent processes. As such, seasonal adjustment can be viewed as an unobserved components (UC) problem and specifically that of UC estimation. Though the nonparametric approach has a rich history going back to the development of X-11 and X-11 ARIMA (Dagum, 1980; Shiskin et al., 1967), our focus centers on model-based methodology.

Within the model-based framework, two directions have emerged. The first direction, and the direction pursued here, directly specifies models for the components and is known as the

[1] Disclaimer: This chapter is released to inform interested parties of research and to encourage discussion. The views expressed on statistical issues are those of the authors and not necessarily those of the U.S. Census Bureau.

structural time series approach (Harvey, 1990). Alternatively, one could start with a model for the observed time series and derive appropriate models for each component (Hillmer and Tiao, 1982). This latter approach is often referred to as "canonical decomposition."

In the seasonal adjustment of economic time series it is common to "preadjust" the series. This preadjustment often includes interpolation of missing values, outlier adjustment and adjustment for trading day and holiday effects. In addition to the customary preadjustments, many model-based approaches require that the observed series be differenced (and/or seasonally differenced) to handle nonstationarity. One question that naturally arises when implementing such an approach is whether or not the correct number of differencing operations have been imposed. In practice, typically, only integer orders of integration are considered. Nevertheless, it is possible that differencing the data once results in a series that still exhibits nonstationary behavior, whereas, imposing a second difference may result in a series that is "over-differenced" and thus noninvertible. In these cases, a natural alternative is to difference the observed series and then model the residual series as a fractionally differenced process.

The models and signal extraction methodology we propose are applied to nonstationary data. However, the approach we develop assumes that, after suitable differencing, the residual series is stationary but allows for long-range dependence (in the seasonal and/or trend component) or anti-persistance (sometimes referred to as intermediate or negative memory). Long-memory time series modeling has experienced tremendous growth during the past three decades. Beginning with the seminal papers on fractional differencing by Granger and Joyeux (1980) and Hosking (1981), many methods have been proposed for modeling long-memory and seasonal long-memory processes with many of these efforts focused on estimation of the fractional differencing parameter, also known as the memory parameter.

Although some research on long-memory has taken a Bayesian viewpoint, the literature is still very heavily frequentist (see Robinson (2003) for a discussion). In general, the literature on long-memory time series is extensive. Excellent references for long-memory time series that include discussion of seasonal long-memory can be found in Palma (2007), Bisognin and Lopes (2009) and the references therein. General discussion regarding long-memory from a

Bayesian perspective can be found in Palma (2007), Holan et al. (2009) and the references therein. Together, these references, along with Holan et al. (2010), provide a detailed survey of the existing literature in long-memory time series.

Even though there exists a substantial body of research on modeling seasonal long-range dependent processes, relatively few efforts have been made toward the seasonal adjustment of economic time series exhibiting such behavior. Further, methods for seasonally adjusting economic time series using Bayesian methodology are also rather limited. The work of Carlin and Dempster (1989) provides one exception.

Carlin and Dempster (1989) develop a Bayesian approach to seasonal adjustment that considers long-memory behavior. However, due to computational limitations, the method they propose is necessarily empirical Bayes and estimates the models using a *plug-in* approach on the fractional differencing parameter from a grid of values. One of the principal differences between our approach and that of Carlin and Dempster (1989) is that our approach is fully Bayesian. Specifically, our method assumes noninformative priors for the fractional differencing parameters and estimates them as part of the model. Additionally, our method uses finite sample minimum mean squared error (MMSE) signal extraction formulas (McElroy, 2008) to facilitate seasonal adjustment. This relies on a matrix representation of the signal extraction filters that is necessary due to the long-memory behavior; these formulas were unavailable to Carlin and Dempster (1989). In addition, the fully Bayesian framework we propose for conducting finite sample MMSE seasonal adjustment extends the current methodology even when the differenced data does not present long-range dependence. Finally, using the matrix approach, requires efficient computation of autocovariances for models with multiple memory parameters (McElroy and Holan, 2011).

Our approach relies on the seasonal fractionally differenced exponential model (SFEXP) (McElroy and Holan, 2011) and more generally the Gegenbauer exponential (GEXP) model (Hsu and Tsai, 2009; McElroy and Holan, 2011). Additionally, our approach allows for versatile models on the seasonal component as well as easy inclusion of extra components that can be modeled rather flexibly. For example, our model allows for straightforward specification of a cycle component, of unknown frequency, (with or without long-range dependence) and/or a sampling error component using GEXP and EXP models respectively.

Finally, we pose our model from a Bayesian hierarchical perspective and, as a result, it is straightforward to include regression effects (e.g., holiday and trading day effects) and to quantify uncertainty. In this context, there are several advantages to proceeding from a Bayesian hierarchical perspective, rather than taking a maximum likelihood approach to estimation. In particular, the likelihood surface is complex and so convergence of numerical optimization algorithms must be carefully monitored. In addition, in the case of maximum likelihood, standard errors for the parameter estimates are typically obtained through asymptotic arguments (using the estimated inverse Hessian matrix) and the final signal extraction estimates are conditioned on the estimated UC model parameters instead of directly accounting for this extra source of uncertainty. In contrast, we design a block Metropolis-Hastings algorithm for efficient model estimation and explicitly propagate uncertainty from the model fitting stage to the signal extraction. Although, our primary focus is on Bayesian methodology, for comparison, we also present results from maximum likelihood estimation (based on the *exact* likelihood rather than the so-called Whittle approximation), which is also novel in this context. Also, maximum likelihood is useful in cases where rapid estimation is desired and/or as starting values for Bayesian estimation.

The remainder of this chapter proceeds as follows. Section 1.2 introduces the SFEXP model and describes an efficient approach to calculating the necessary model autocovariances. Section 1.3 describes long-memory unobserved component models and their application to seasonal adjustment. The methodology is illustrated in Section 1.4 through two real data examples. Finally, Section 1.5 provides a brief conclusion. Details surrounding our Markov chain Monte Carlo (MCMC) algorithm and Bayesian signal extraction estimator are left to the Appendix.

1.2 The SFEXP Model

The structural models of Section 1.3 depend, in large part, on the SFEXP spectral representation and efficient autocovariance computation. The necessary background material is provided here. In particular, we consider a process $\{Y_t\}$ that has long memory at both trend and seasonal frequencies. We focus on the case of monthly data, so that the seasonal frequencies are $\pi j/6$ for $j = 1, 2, \cdots, 6$; the trend frequency ($j = 0$) is treated separately. Now to each of the seasonal frequencies and the trend frequency, we may associate a pole in the pseudo-spectral density with rate of explosion governed by the seasonal and non-seasonal memory parameter δ_S and δ_N respectively. These may be numbers greater than .5, indicating nonstationarity. Letting $\mu_t = \mathrm{E}(Y_t)$, our basic assumption is that

$$(1-B)^{\delta_N} U(B)^{\delta_S} (Y_t - \mu_t) \tag{1.2.1}$$

is a mean zero stationary process modeled by an exponential model (Bloomfield, 1973) of order m, where B is the backshift operator and $U(z) = 1 + z + z^2 + \cdots + z^{11}$. The parameters δ_N and δ_S broadly define the dependence structure.

Specifically, when $\delta_N \in (0, .5)$, the process has (stationary) long memory at frequency zero, whereas if $\delta_N = 0$ the process has short memory. If $\delta_N \in (-.5, 0)$, the process has intermediate memory at frequency zero, which is also stationary. Similar statements apply to the range of δ_S, except we substitute seasonal frequencies for frequency zero. It will be convenient to separate out the integer portion of δ_N and δ_S so that we can focus on the stationary aspects of the model. Since the process Y_t is stationary if and only if $|\delta_N| < .5$ and $|\delta_S| < .5$, it makes sense to define the integer portion of δ_N and δ_S to be given by rounding to the nearest integer (with fractional values of .5 being rounded upwards), which will be denoted by the symbol $[\cdot]$. Hence $\delta_N = [\delta_N] + d$ and $\delta_S = [\delta_S] + D$, where the Latin letters d and D denote the remainders, which are guaranteed to lie in $[-.5, .5)$. If we let Z_t denote the suitably differenced $(Y_t - \mu_t)$, such that the result is stationary, we have

$$Z_t = (1-B)^{[\delta_N]} U(B)^{[\delta_S]} (Y_t - \mu_t) = (1-B)^{[\delta_N]} U(B)^{[\delta_S]} Y_t - \zeta_t.$$

We use the notation $W_t = (1-B)^{[\delta_N]} U(B)^{[\delta_S]} Y_t$ for the differenced observed process and $\zeta_t = (1-B)^{[\delta_N]} U(B)^{[\delta_S]} \mu_t$ for the differenced time-varying mean. Thus $W_t = \zeta_t + Z_t$, with $\{Z_t\}$ stationary and mean zero. It follows that $\{Z_t\}$ has an integrable spectral density function f, which given our basic assumption (1.2.1) means that

$$f(\lambda) = |1 - z|^{-2d} |U(z)|^{-2D} \exp\left\{ \sum_{j=-m}^{m} g_j z^j \right\}, \tag{1.2.2}$$

where we use the convenient abbreviation $z = e^{-i\lambda}$. This defines SFEXP model (McElroy and Holan, 2011), where $g_j = g_{-j}$ by assumption.

Example. If $\delta_N = 1.33$ and $\delta_S = 1.15$, then the nonstationary integer differencing order is $[\delta_N] = 1$ for trend and $[\delta_S] = 1$ for seasonal (fairly typical values), and $d = .33$ and $D = .15$ are the resulting memory parameters. The values of these memory parameters indicate fairly strong memory for the trend, but weaker long-range dependence for the seasonal frequencies.

The fitting of the SFEXP model to seasonal data involves firstly the identification of integer differencing orders, which are the numbers $[\delta_N]$ and $[\delta_S]$. These are applied to the observed time series, resulting in the differenced series $\{W_t\}$. Subsequently, a short-memory model order m is selected and, conditional on m, values for d and D and the $\{g_j\}_{j=0}^{m}$ are determined using Bayesian (or maximum likelihood) estimation. Alternatively, in principal, taking a Bayesian approach similar to Holan et al. (2009), the choice of m (order selection) can be directly incorporated into the modeling procedure using reversible jump Markov chain Monte Carlo.

Whether a Bayesian or a maximum likelihood procedure is carried out, it is convenient to use the Gaussian likelihood function corresponding to (1.2.2), since, conditional on ζ_t, this only depends on autocovariances that can be computed directly and accurately via algorithms described in McElroy and Holan (2011). Note, as demonstrated below, the main determinants of the seasonal adjustment filters are the parameters d and D, whereas the short memory parameters have little effect. The key to estimation of a Gaussian SFEXP model is efficient, accurate computation of the autocovariances. The method of McElroy and Holan (2011) is extremely general; nonetheless, improvements to speed and precision of parameter estimates can be achieved in the special case of (1.2.2) when only one pole is present and is useful in the case of UC models. We provide the details for relevant cases of this situation below.

Suppose first that $D = 0$ in (1.2.2), producing the FEXP model. The best technique is to compute autocovariances using the splitting method (Bertelli and Caporin, 2002) (which involves the convolution of short and long memory autocovariances), since the autocovariance sequence for $|1-z|^{-2d}$ is known analytically and the autocovariance sequence for the short memory $g(\lambda) = \exp\{\sum_j g_j z^j\}$ decays at geometric rate:

$$\gamma_h = \frac{1}{2\pi}\int_{-\pi}^{\pi} |1-z|^{-2d} g(\lambda) z^{-h} d\lambda = \sum_k \gamma_k(g)\xi_{h-k}(d), \tag{1.2.3}$$
$$\gamma_k(g) = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(\lambda) z^{-k} d\lambda,$$
$$\xi_j(d) = \frac{\Gamma(j+d)\Gamma(1-2d)}{\Gamma(j-d+1)\Gamma(d)\Gamma(1-d)}.$$

Here $\xi_j(d)$ is the autocovariance sequence for an ARFIMA(0,d,0) as in Brockwell and Davis (1991), utilizing the Gamma function. The expression for γ_h involves an infinite sum that can be truncated safely, since $\gamma_k(g)$ will tend to decay rapidly (since g has no poles).

Secondly, suppose that $d = 0$ and $D > 0$ in (1.2.2). Then rewrite the spectral density as $f(\lambda) = |1-z^{12}|^{-2D}|1-z|^{2D}g(\lambda)$. Observe that $k(\lambda) = |1-z|^{2D}g(\lambda)$ is a combination of negative memory (hyperbolic zeroes in spectrum) and short memory behavior, and also has rapidly decaying autocovariance sequence. This autocovariance sequence need not decay geometrically, but will tend to decay rapidly nevertheless. The autocovariance function (acvf) of the spectrum $|1-z^s|^{-2D}$ is given by the following result.[2]

Proposition 1. *The autocovariance sequence of* $|1-z^s|^{-2D}$ *is*

$$\xi_j(D) = \frac{\Gamma(j/s+D)\Gamma(1-2D)}{\Gamma(j/s-D+1)\Gamma(D)\Gamma(1-D)} 1_{\{j\equiv 0\, mod s\}}.$$

[2]Propoistion 1 is fairly common in the literature on long-memory time series and can be readily deduced from results found in Brockwell and Davis (1991). Thus, it is presented here without proof.

As a result, the desired autocovariance sequence is

$$\gamma_h = \sum_j \gamma_j(k)\xi_{h-j}(D) = \sum_j \gamma_{h-12j}(k) \frac{\Gamma(j+D)\Gamma(1-2D)}{\Gamma(j-D+1)\Gamma(D)\Gamma(1-D)}. \tag{1.2.4}$$

The calculation of these quantities is therefore fast and accurate.

If $D < 0$, the process exhibits negative memory and has a rapidly decaying autocovariance sequence. Therefore, $f(\lambda) = |1 - z^{12}|^{-2D}|1 - z|^{2D} g(\lambda) = |U(z)|^{-2D} g(\lambda)$ can be accurately and efficiently computed. In this case, several methods could be employed including McElroy and Holan (2011), a multiple splitting approach or direct Fourier inversion of the spectral density using numerical integration.

1.3 Long-Memory UC Models and Seasonal Adjustment

In order to apply formulas for finite sample MMSE[3] we need autocovariance generating functions for signal and noise. As previously alluded to, we take a structural approach, which posits models for signal and noise and, thus, an inferred autocovariance structure is obtained for the data process via summing the spectra of the component models. Specifically, we illustrate this for the SFEXP case.

Let Y_t equal the sum of two components S_t and N_t, the seasonal and nonseasonal respectively. In terms of (1.2.1), these have $\delta_N = 0$ for S_t and $\delta_S = 0$ for N_t, and potentially different orders m_S and m_N for the exponential model portions. These are working assumptions that are sensible, since we seek a seasonal component that has no trend dynamics, and also a nonseasonal component with no seasonal dynamics. Including regression effects (i.e., trading day and/or holiday effects), the two component model can be written as $Y_t = \mu_t + S_t + N_t$, where $\mu_t = X_t'\theta$ denotes a regression component. Thus, it follows that

$$W_t = \zeta_t + (1-B)^{[\delta_N]} U_t + U(B)^{[\delta_S]} V_t,$$

[3]The described estimators are MMSE for Gaussian time series, or alternatively are MMSE among linear estimators in the data; this depends on certain signal extraction conditions described in McElroy (2008).

where $\{U_t\}$ and $\{V_t\}$ are suitably differenced versions of $\{S_t\}$ and $\{N_t\}$ that are independent of one another, i.e., $U_t = U(B)^{[\delta_S]} S_t$ and $V_t = (1-B)^{[\delta_N]} N_t$. Note that if we suppose that S_t and N_t follow (1.2.1) then, in general, $Z_t = W_t - \zeta_t$ does not follow an exponential model. In particular, we have

$$f_S(\lambda) = |U(z)|^{-2[\delta_S]} f_U(\lambda) \qquad f_U(\lambda) = |U(z)|^{-2D} g_S(\lambda)$$
$$f_N(\lambda) = |1-z|^{-2[\delta_N]} f_V(\lambda) \qquad f_V(\lambda) = |1-z|^{-2d} g_N(\lambda)$$

for exponential models g_S and g_N given by $g_S(\lambda) = \exp\{\sum_{|j| \leq m_S} u_j z^j\}$ and $g_N(\lambda) = \exp\{\sum_{|j| \leq m_N} v_j z^j\}$. Thus it follows that the spectrum for the differenced, mean-centered process Z_t is

$$|1-z|^{2[\delta_N]} |U(z)|^{-2D} g_S(\lambda) + |U(z)|^{2[\delta_S]} |1-z|^{-2d} g_N(\lambda), \tag{1.3.1}$$

which, in general, will not have the form of f given in (1.2.1), as the exponential portion would have infinitely many nonzero coefficients. Nevertheless, (1.3.1) can easily be used to construct a Gaussian likelihood function: given d, D, and the parameters of g_S and g_N (with $[\delta_N]$ and $[\delta_S]$ determined before-hand as in Section 2), we compute the autocovariances for each of the two terms and sum. Let $\Gamma(g)$ denote the (Toeplitz) covariance matrix associated with a stationary process having spectral density g. Then for a sample from $\{Y_t\}$ of size n, the autocovariance matrix associated with Z_t can be expressed as

$$\Sigma_Z = \underline{\Delta}_N \Gamma(f_U) \underline{\Delta}_N' + \underline{\Delta}_S \Gamma(f_V) \underline{\Delta}_S'$$

where $\underline{\Delta}_N$ and $\underline{\Delta}_S$ are $(n - 12[\delta_S] - [\delta_N]) \times (n - 12[\delta_S])$ and $(n - 12[\delta_S] - [\delta_N]) \times (n - [\delta_N])$-dimensional differencing matrices with entries given by the coefficients of $(1-B)^{[\delta_N]}$ and $U(B)^{[\delta_S]}$ respectively, appropriately shifted. Explicit examples of these matrices can be found in Bell (2004) and McElroy and Gagnon (2008).

Let ψ denote the full parameter vector for the model (excluding regression parameters); then $\psi = (d, u_0, u_1, \cdots, u_{m_S}, D, v_0, v_1, \cdots, v_{m_N})'$. Note that this vector partitions into the first $m_S + 2$ components corresponding to the seasonal component S, and the latter $m_N + 2$ components for the nonseasonal component N. Using maximum likelihood, and the concepts of Section 1.2, the actual estimation of the structural model for the data is fast and accurate to compute. Conditional on the regression parameters, we only need to compute the autocovariance sequence associated with (1.3.1) to obtain likelihood evaluations. The first term is a simple linear combination of autocovariances computed using (1.2.4). The second term is another finite linear combination of autocovariances computed using (1.2.3). Then one sums the autocovariance sequences to get the autocovariances for the differenced data.

If one utilizes this procedure in the heart of a maximum likelihood routine, the end result at convergence is parameter estimates for ψ (and θ - the vector of regression coefficients). This gives complete models for both the seasonal and nonseasonal as well as the autocovariance sequence for Z_t. Additionally, it is clear from the previous discussion that the data spectrum f will have stationary long memory poles at the trend and seasonal frequencies, of order d and D respectively, since it follows from (1.3.1) that

$$f_\psi(\lambda) = \frac{|1-z|^{2\delta_N} g_S(\lambda) + |U(z)|^{2\delta_S} g_N(\lambda)}{|1-z|^{2d} |U(z)|^{2D}},$$

where $f_\psi(\lambda)$ denotes the spectral density associated with the parameters ψ. For the MMSE signal extraction formulas, we need the autocovariances associated with both components. Under the typical assumptions (given in Bell (1984b) and McElroy (2008)), we apply the matrix formula (1.3.2) given below to the finite sample of data $Y = (Y_1, Y_2, \cdots, Y_n)'$ (the presence of mean effects changes things slightly; see the discussion in the Appendix). Let Δ_S and Δ_N denote differencing matrices for seasonal and nonseasonal, such that when applied to the signal and noise vectors, yield U and V respectively. Specifically, Δ_S and Δ_N are $(n - 12[\delta_S]) \times n$ and $(n - [\delta_N]) \times n$ differencing matrices respectively and are constructed similar to $\underline{\Delta}_N$ and $\underline{\Delta}_S$ (see McElroy (2008) for more details). Then the signal extraction matrix associated with ψ is given by

$$F(\psi) = \left(\Delta_S' \Gamma^{-1}(f_U) \Delta_S + \Delta_N' \Gamma^{-1}(f_V) \Delta_N \right)^{-1} \Delta_S' \Gamma^{-1}(f_U) \Delta_S. \tag{1.3.2}$$

The dependence on the parameter vector ψ enters through f_U and f_V, whose reliance on ψ is suppressed in the notation.

Conditional on m_S and m_N being known, it is straightforward to write down the Gaussian likelihood associated with a sample of size n. Letting $W = (W_1, W_2, \ldots, W_{n'})'$, n' equal the length of the differenced series and $1_{n'}$ the vector of ones with length n' we have

$$p(W|\psi, \zeta) = (2\pi)^{-n/2}|\Sigma(f_\psi)|^{-1/2} \exp\left\{-\frac{1}{2}\left(W - \zeta 1_{n'}\right)' \Sigma(f_\psi)^{-1}\left(W - \zeta 1_{n'}\right)\right\}.$$

It is readily seen that our UC model naturally possesses a hierarchical structure. Ultimately, for signal extraction, we are interested in estimating the model parameters so that we can compute the necessary autocovariances. Specifically, the finite sample MMSE formulas will depend on ψ; the Appendix contains a comprehensive treatment of Bayesian finite sample MMSE. To this end, we use Bayes rule to obtain the posterior distribution of ψ and ζ (and ultimately the autocovariances, unobserved components and seasonal adjustments) given the data

$$p(\psi,\, \zeta|W) = p(D, d, u, v, \sigma_u^2, \sigma_v^2, \zeta|W) \propto p(W|\psi, \zeta)p(u|\sigma_u^2)p(v|\sigma_v^2)p(\sigma_u^2)p(\sigma_v^2)p(d)p(D)p(\zeta),$$

where $u = (u_0, u_1, \ldots, u_{m_S})'$, $v = (v_0, v_1, \ldots, v_{m_N})'$, $\sigma_u^2 = (\sigma_{u_0}^2, \ldots, \sigma_{u_{m_S}}^2)'$ and $\sigma_v^2 = (\sigma_{v_0}^2, \ldots, \sigma_{v_{m_N}}^2)'$ and we have assumed conditional independence between components. To completely specify a Bayesian model requires us to choose priors for $(d, u, \sigma_u^2, D, v, \sigma_v^2, \zeta)$. For ease of exposition, and to be consistent with our empirical case studies (cf. Section 1.4), we assume the differenced data to be mean centered via ζ, noting that it is straightforward to include regression effects such as trading day and holiday effects. Similar to Holan et al. (2009), for $j = 1, \ldots, m_S$ and $k = 1, \ldots, m_N$, we assign hierarchical priors on the unknown parameters as follows: $D \sim \text{Uniform}(-1/2, 1/2)$; $u_j|\sigma_{u_j}^2 \sim N(0, \sigma_{u_j}^2)$; $\sigma_{u_j}^2 \sim IG(A_u, B_u)$; $d \sim \text{Uniform}(-1/2, 1/2)$; $v_k|\sigma_{v_k}^2 \sim N(0, \sigma_{v_k}^2)$; $\sigma_{v_k}^2 \sim IG(A_v, B_v)$; $\zeta \sim N(\zeta_0, \sigma_\zeta^2)$. As usual IG denotes the inverse gamma distribution, so that σ_ℓ^2 has pdf $p(\sigma_\ell^2) \propto (\sigma_\ell^2)^{-(A+1)} \exp\left(-B/\sigma_\ell^2\right)$. Typically, as is the case in our illustrations, the hyperparameters A_u, B_u, A_v, B_v, ζ_0 and σ_ζ^2 are specified so that the prior distributions are vague or noninformative. Comprehensive details regarding the full conditional distributions and exact MCMC algorithm can be found in the Appendix.

1.4 Empirical Case Studies

In order to demonstrate the effectiveness and flexibility of our approach we consider the seasonal adjustment of two time series. Firstly, we considered a dataset from the Current Population Survey (CPS) (Source: U.S. Bureau of Labor Statistics, http://www.bls.gov/data/) that consists of Employed Males, aged 16-19, from 1/1976-6/2010; henceforth referred to as the "EM1619" series. Figure 1.1 (top left panel) provides a plot of the time series. Next, we consider data from the Current Employment Statistics Survey (CES) (Source: U.S. Bureau of Labor Statistics, http://www.bls.gov/data/) that consists of U.S. Total Non-farm Employment from 1/1939-7/2009. Figure 1.2 (top left panel) provides a plot of the observed time series.

Figure 1.1 approximately here

Figure 1.2 approximately here

For both datasets maximum likelihood is conducted using the *optim* command in R (R Development Core Team, 2010) to numerically determine the maximum of the likelihood surface. Conversely, the Bayesian procedure uses the prior specification detailed in Section 1.3. Specifically, for both datasets we take $A_u = B_u = A_v = B_v = .1$, $\zeta_0 = 0$, $\sigma^2_\zeta = 10^3$ and implement the MCMC algorithm described in the Appendix. For model fitting we run a single MCMC chain for 10,000 iterations discarding the first 1000 iterations for burn-in and keeping every third iteration for inference, leaving 3000 iterations total. Convergence of the MCMC is verified through trace plots of the sample chains. All estimated parameters are taken as the posterior means.

In order to arrive at a model for illustration we informally consider AIC and BIC for models estimated using maximum likelihood (Beran et al., 1998). In particular, Beran et al. (1998) provided formal justification for these criterion in the context of fractional autoregressive models. Nevertheless, we shall use these criteria here to narrow down the candidate models.

In the case of the Bayesian hierarchical approach we use deviance information criterion

(DIC) (Spiegelhalter et al., 2002) to evaluate candidate models. DIC has become commonplace in Bayesian model selection and, similar to AIC and BIC, models with smaller DIC are preferable. Ultimately, for each dataset, we choose a Bayesian model for illustration based on DIC. Again, since the goal of our illustration is to demonstrate effective seasonal adjustment, we have not conducted an exhaustive search to find an "optimal" model. Instead, we choose a competitive model and evaluate its effectiveness in seasonal adjustment.

In order to determine if a transformation and/or differencing is required for either dataset, we begin with an exploratory analysis. Figures 1.1 and 1.2 provide plots of the observed series, the first-differenced series and the series obtained from taking the first difference and then applying the differencing operator $U(B)$. In addition, both figures display the associated autocovariance functions, acvfs, and AR(30) spectral densities. From an initial assessment of the plots in Figures 1.1 and 1.2, it seems reasonable to work with the data obtained from applying the $(1-B)U(B) = 1-B^{12}$ differencing operator to both observed series.

1.4.1 Current Population Survey - Employed Males

In seasonal adjustment applications it is common to test for outliers, trading day and holiday effects in monthly time series such as the EM1619 series. Running different specifications through X-12 ARIMA, neither trading day (TD) or Easter holiday effects were found significant and thus no TD or holiday adjustments were made. The current model used at the Bureau of Labor Statistics involves removing three level shifts (LS). We fit models with both the three LS removed and retained and found no appreciable difference. Thus, in what follows, we detail the analysis with the three LS retained.

Several models were fit using maximum likelihood and although many models appeared to provide a good fit to the observed data, not all of these models provided adequate seasonal adjustment. One model identified for further consideration was a long-memory UC model with one seasonal and three nonseasonal short-memory coefficients (i.e., LM-UC(1,3)). The estimated long-memory parameters were $(\widehat{D}, \widehat{d}) = (.460, .159)$ with standard errors of .032

and .084 respectively. The estimated seasonal short-memory coefficient, $\widehat{u}_0$, equals -4.719 with standard error .312. The estimated nonseasonal short-memory coefficients $(\widehat{v}_0, \widehat{v}_1, \widehat{v}_2) = (-.751, -.749, -.420)$ with standard errors of .085, .204 and .157 respectively. Finally, $\widehat{\zeta}$ the estimated mean of the differenced data, equals -.555 with standard error .579. Note that $\overline{x}_d$, the sample mean of the differenced data, equals -.524 and closely agrees with our estimate.

Figure 1.3 displays the estimated nonseasonal component and associated pointwise 95% confidence band. To assess whether the seasonality has been removed we plot the AR(30) spectral density for the first-differenced estimated trend (Figure 1.4). Figure 1.4 clearly illustrates that the seasonality has been satisfactorily removed. Finally, one can examine the seasonal adjustment filter through a plot of the gain function for the seasonal and nonseasonal components. As seen in Figure 1.5, the induced filters sensibly suppress the dynamics at the correct frequencies.

Figure 1.3 approximately here

Figure 1.4 approximately here

Figure 1.5 approximately here

In contrast, we conducted seasonal adjustment using Bayesian methodology. Specifically, based on DIC, we estimated a LM-UC(1,5) model. It is important to emphasize that we have not conducted an exhaustive model selection, but rather chose a candidate model to illustrate Bayesian long-memory seasonal adjustment. The estimated model parameters were $(\widehat{D}, \widehat{d}) = (.381, .097)$ with 95% credible intervals (CI) given by (.224,.471) and (-.143,.321) respectively. The estimated seasonal short-memory coefficient is -4.251 with 95% CI of (-5.020,-3.246) whereas the estimated nonseasonal short-memory coefficients are given by (-.794,-.491,-.377,.233,-.231) with 95% CIs given by (-1.047,-.570), (-1.029,.009), (-.786,.007), (-.066,.525) and (-.546,.006) respectively. Finally, $\widehat{\zeta}$, the estimated mean of the differenced data equals -.542 with 95% CI equal to (-.587, -.495). Figure 1.6 displays the observed series along with the estimated nonseasonal component and pointwise 95% CIs. An AR(30) spectral density and acf plot identical to Figure 1.4 indicate that the induced filters sensibly suppress the dynamics at the correct frequencies (not displayed).

Figure 1.6 approximately here

In both the maximum likelihood and Bayesian approach, the estimated nonseasonal long-memory parameter is not statistically significant (at the .05 level). For comparison, we estimated a model (using classical and Bayesian methodology) with the nonseasonal long-memory parameter set identically equal to zero. We found that this produced qualitatively similar seasonal adjustments and, thus, is not presented here. In the Bayesian context this parameter can be viewed as a nuisance parameter, since our target is an estimated trend component. The method we propose averages over the distribution of this parameter to produce a pointwise distribution of trend components. Therefore, it is advantageous to include this parameter in the model, since a significant portion of the distribution, for the model parameter, is located away from zero (see Figure 1.7). Importantly, the pointwise 95% CI for the estimated nonseasonal component takes into account parameter uncertainty and is narrower than the corresponding interval from maximum likelihood that appeals to large sample approximations (see Holan et al., 2009, for further discussion). Finally, we also estimated models with three components – trend, seasonal and irregular. The results for the seasonal adjustments were qualitatively similar in this case and thus are not presented here.

Figure 1.7 approximately here

1.4.2 U.S. Total Non-farm Employment

As previously discussed, preadjustment of economic time series in the context of seasonal adjustment is commonplace in practice. To assess the need for preadjustment, we ran several specifications in X-12 ARIMA; it was determined that no transformation was necessary but that stock-trading day effects were significant. Although stock - trading day effects could be directly incorporated into our model as a regression effect (cf., Section 1.3), and estimated using Bayesian methods, our focus is on illustrating the long memory aspects of the model and thus we removed this effect prior to model estimation and seasonal adjustment. While this may not be preferred from a modeling perspective, this is also consistent with current seasonal adjustment practices in federal statistical agencies and advantageous from a computational viewpoint.

Similar to the EM1619 series, several candidate models were investigated. One model identified for further consideration (under both maximum likelihood and Bayesian estimation) was a LM-UC(2,3) model. Thus, using the Bayesian framework, we conducted seasonal adjustment under this model. In this case, the estimated long-memory parameters were $(\widehat{D}, \widehat{d}) = (.053, .325)$ with 95% CIs equal to (-.061,.179) and (.208,.429) respectively. The estimated seasonal short-memory coefficients, $(\widehat{u}_0, \widehat{u}_1)$, equal (-6.431, 2.562) with 95% CIs equal to (-7.168, -5.832) and (2.041, 3.025) respectively. The estimated nonseasonal short-memory coefficients $(\widehat{v}_0, \widehat{v}_1, \widehat{v}_2) = (-3.931, 0.278, 0.569)$ with 95% CIs equal to (-4.104, -3.771) and (.004, .595) and (.337, .828) respectively. Finally, $\widehat{\zeta}$ the estimated mean of the differenced data, equals 1.621 with 95% CI (1.609, 1.634). Figure 1.8 shows a plot of the observed series along with the estimated nonseasonal component. The estimated pointwise 95% CI is intentionally suppressed since the width of this interval is uniformly less than .25. Finally, to assess whether the seasonality has been adequately removed we plot the AR(30) spectral density for the first-differenced estimated trend (Figure 1.9). This figure clearly demonstrates that the seasonality has been removed. Again, similar to the EM1619 series, one could study properties of the seasonal adjustment filter through an investigation of the gain function. However, we defer such investigation here, noting that the investigation yields similar conclusions to the EM1619 series.

Figure 1.8 approximately here

Figure 1.9 approximately here

Alternatively, using maximum likelihood, the estimated long-memory parameters were $(\widehat{D}, \widehat{d}) = (.064, .349)$ with standard errors of .062 and .069 respectively. The estimated seasonal short-memory coefficients, $(\widehat{u}_0, \widehat{u}_1)$, equal (-6.475, 2.583) with standard errors of (.337, .260) respectively. The estimated nonseasonal short-memory coefficients $(\widehat{v}_0, \widehat{v}_1, \widehat{v}_2) = (-3.923, 0.231, 0.533)$ with standard errors of .082, .168 and .127 respectively. Finally, $\widehat{\zeta}$, the estimated mean of the differenced data, equals 1.613 with standard error 1.054 and agrees closely with $\overline{x}_d = 1.597$ (the sample mean of the differenced data). Plots of the estimated nonseasonal component, from maximum likelihood, and the AR(30) spectral density for the first differenced estimated trend demonstrate that the seasonality has been convincingly removed. These plots are similar to the Bayesian case (Figures 1.8 and 1.9) and, thus, are not displayed for the sake of brevity.

Similar to the EM1619 series, we could set $D \equiv 0$. However, our preference is to include this parameter in the model since its distribution has considerable mass away from zero (see Figure 1.10). Additionally, our Bayesian approach views this parameter as a nuisance parameter and averages over the distribution of D to obtain the distribution of the unobserved components. To verify that this approach was reasonable we fit a model with $D \equiv 0$ and found that the results were consistent with those reported here.

Figure 1.10 approximately here

1.5 Discussion

Research into long-memory processes has recently spread to the modeling of seasonality through the use of generalized exponential time series models. This chapter considers the application of seasonal long memory modeling to the problem of seasonal adjustment of economic time series. In particular, we introduced the new SFEXP model, and explored its fit to economic time series data. Subsequently, we discussed a structural approach to obtaining component models for seasonal and trend in the context of long memory, and use these models to obtain finite sample MMSE. The approach we propose is fully Bayesian, producing distributions for the unobserved components, and thus naturally quantifies the uncertainty in the signal extraction estimates.

One interesting direction for future research is to model the regressors (i.e., Trading Day and Holiday effects) dynamically. In particular, in order to observe and effectively model long-range dependence typically requires a long time series, as was the case in our examples. Thus, it is conceivable that the regression coefficients might change over time. In principal, modeling these parameters dynamically at another level in the hierarchical model would constitute a relatively straightforward extension to the models we propose.

More specifically, the general modeling approach we propose extends seasonal adjustment methodology in several ways. First, our methodology provides the first attempt at exact finite sample MMSE signal extraction for long-memory time series. In addition, we propose a fully Bayesian framework for conducting finite sample MMSE seasonal adjustment, which extends the current methodology even when the differenced data does not present long-range dependence.

In order to facilitate Bayesian estimation we develop an efficient block Metropolis-Hastings (M-H) sampler. The sampling algorithm provides efficient computation by minimizing the number of expensive evaluations of the likelihood. In addition, we propose an effective method for computing the necessary model autocovariances. These computational tools allow us to effectively estimate the SFEXP model and LM-UC models that were introduced in Sections 1.2 and 1.3.

The methodology is illustrated using two real economic time series, the CPS – Employed Males, aged 16-19, from 1/1976-6/2010 and the CES - U.S. Total Non-farm Employment from 1/1939-7/2009. These empirical case studies demonstrate the flexibility and utility of our approach. In short, we have shown that our proposed methodology provides a necessary and timely extension to the current practice of seasonal adjustment.

1.6 Appendix

1.6.1 Full Conditionals and MCMC

Estimation of the UC model presented in Section 1.3 is computationally demanding due to expensive likelihood evaluations. Thus, it is essential to minimize the number of likelihood evaluations in the MCMC algorithm. As in Section 1.3, we assume that the components U_t and V_t are uncorrelated. Therefore, it is natural to sample the parameters in blocks according to their respective components. Further, since some of the full conditionals are not of standard form we have used a Metropolis within Gibbs algorithm (Gelman et al., 2003). Below we list the necessary full conditional distributions.

First, since the prior for $\sigma^2_{u_j}$ $(j = 1, \ldots, m_S)$ does not depend on D, d, v, σ^2_v or ζ and since u_j $(j = 1, \ldots, m_S)$ are independent of each other it follows that

$$p(\sigma^2_{u_j}|d, D, u, v, \sigma^2_{u_{-j}}, \sigma^2_v, \zeta, W) \sim IG\left\{(A_u + 1/2),\, (u_j^2/2 + B_u)\right\}, \tag{1.6.1}$$

where $\sigma^2_{u_{-j}}$ is the vector of variances for $u_0, u_1, \ldots, u_{j-1}, u_{j+1}, \ldots, u_{m_S}$. Similarly, for $k = 1, \ldots, m_N$, the full conditional of $\sigma^2_{v_k}$ is given by

$$p(\sigma^2_{v_k}|d, D, u, v, \sigma^2_u, \sigma^2_{v_{-k}}, \zeta, W) \sim IG\left\{(A_v + 1/2),\, (v_k^2/2 + B_v)\right\}. \tag{1.6.2}$$

Furthermore, the joint full conditional of D and u is not conjugate under this model but is straightforward to derive

$$\begin{aligned} p(D, u|d, v, \sigma^2_u, \sigma^2_v, \zeta, W) \quad &\propto \quad |\Sigma(f_\psi)|^{-1/2} \exp\left\{Z'\Sigma(f_\psi)^{-1}Z\right\} \times |\Sigma_u|^{-1/2} \exp\left\{u'\Sigma_u^{-1}u\right\} \\ &\times \quad I_{(-1/2,\,1/2)}(D), \end{aligned} \tag{1.6.3}$$

where $Z = (W - \zeta 1_{n'})$, $\Sigma_u = \text{diag}(\sigma^2_{u_1}, \ldots, \sigma^2_{u_{m_S}})$ and $I_{(-1/2,\,1/2)}(D)$ is the indicator function (i.e., equal to 1 if $D \in (-1/2,\, 1/2)$ and 0 otherwise). Similarly, the joint full conditional of d and v can be expressed as

$$\begin{aligned} p(d, v|D, u, \sigma^2_u, \sigma^2_v, \zeta, W) \quad &\propto \quad |\Sigma(f_\psi)|^{-1/2} \exp\left\{Z'\Sigma(f_\psi)^{-1}Z\right\} \times |\Sigma_v|^{-1/2} \exp\left\{v'\Sigma_v^{-1}v\right\} \\ &\times \quad I_{(-1/2,\,1/2)}(d). \end{aligned} \tag{1.6.4}$$

Finally, the full conditional of ζ is given by

$$p(\zeta|d, D, u, v, \sigma^2_u, \sigma^2_v, W) \sim N\left(\overline{\zeta}, \overline{\sigma}^2_\zeta\right), \tag{1.6.5}$$

where $\overline{\zeta} = c_2/c_1$, and $\overline{\sigma}^2_\zeta = 1/c_1$ with $c_1 = \left\{1'_{n'}\Sigma^{-1}(f_\psi)1_{n'} + \sigma_\zeta^{-2}\right\}$ and $c_2 = \left\{W'\Sigma^{-1}(f_\psi)1_{n'} + \zeta_0\sigma_\zeta^{-2}\right\}$.

Given the calculated likelihood, implementation of the MCMC requires M-H updates in order to sample from the joint full conditional distributions of (D, u) and (d, v). To summarize, our MCMC algorithm proceeds as follows:

Step 1: Set initial values for all parameter values.

Step 2: For $j = 1, \ldots, m_S$, generate samples from (1.6.1).

Step 3: For $k = 1, \ldots, m_N$, generate samples from (1.6.2).

Step 4: Using a Random-Walk M-H step, jointly sample (D, u) from (1.6.3).

Step 5: Using a Random-Walk M-H step, jointly sample (d, v) from (1.6.4).

Step 6: Generate samples from (1.6.5).

Step 7: Repeat until convergence.

In many cases it is possible to estimate the parameters of the UC model using maximum likelihood. Under these circumstances it is advantageous to use the maximum likelihood values for the initial values described in Step 1. In doing so, the MCMC algorithm essentially starts in the stationary distribution or at least close to it.

For implementation of the Random-Walk M-H (RW M-H) algorithm, one needs a candidate generating density. Chib and Greenberg (1995) have several proposals in this regard. However, since d and D both have bounded support it is beneficial to take a transformation and work with a proposal distribution on the transformed space. Specifically, let $D_\infty = \text{logit}(D + 1/2)$ where $\text{logit}(r) = \log\{r/(1 - r)\}$, and let d_∞ be defined analogously. For specificity, let $\mu_S^* = (D_\infty^*, u^*)$ denote the current state for the parameters $\mu_S = (D, u)$; we then draw a candidate value of μ_S^* using a $N(\mu_S, \Sigma_S)$ proposal distribution where Σ_S is chosen as $\{2.4^2/(m_S + 1)\}C_S$ with C_S equal to the empirical covariance matrix of μ_S determined from a pilot simulation (Gelman et al., 2003). Choosing Σ_S in this manner produces a jumping rule shaped like an estimate of the target distribution and thus produces efficient simulation (i.e., acceptance rates around 25% with adequate mixing).

The algorithm accepts μ_S^* as a new value of μ_S with acceptance probability

$$\alpha_{\mu_S} = \min\left\{1, \frac{p(W|D^*, u^*, d, v, \sigma_u^2, \sigma_v^2, \zeta)p(u^*|\Sigma_u)|J^*|}{p(W|D, u, d, v, \sigma_u^2, \sigma_v^2, \zeta)p(u|\Sigma_u)|J|}\right\},$$

where $|J^*| = \{1 + \exp(D_\infty^*)\}^{-2}$ and $|J| = \{1 + \exp(D_\infty)\}^{-2}$ denote the necessary Jacobians for the transformation described above. Finally, a completely analogous RW M-H step is used to sample from the joint full conditional distribution of (d, v).

1.6.2 Derivation of Signal Extraction Estimator

For this section only we reserve S for signal and N for noise, whereas in the remainder of this chapter N is the nonseasonal – the actual signal of interest – and S is the seasonal noise. Our goal is estimation of the signal conditional on the data Y, which has mean vector μ. Suppose that the signal consists of a mean zero stochastic component S plus its mean effect μ^S. This fixed component is typically viewed as a regression component $X\theta$ for deterministic θ, though in the Bayesian framework θ is just a subset of the full parameter vector ψ and hence is random as well. Likewise, the noise component is assumed to consist of stochastic N plus its mean μ^N, and $Y = S + N + \mu^S + \mu^N$, the sum of the signal and noise.

Note that conditional on ψ, the mean effects μ^S and μ^N are the expectations of the signal and noise respectively. Write $S = \{S_1, S_2, \cdots, S_n\}'$. The MMSE solution to the estimation problem is given by conditional expectations.

$$\mathrm{E}(S + \mu^S | Y) = \int s \, p_{S+\mu^S|Y}(s) \, ds.$$

This equality is to be understood component-wise, i.e., that $\mathrm{E}(S_j + \mu_j^S | Y) = \int s p_{S_j+\mu_j^S|Y}(s) \, ds$ for all $j = 1, 2, \cdots, n$. Integrals are over all the real numbers and $p_{S_j+\mu_j^S|Y}$ denotes the probability density function of $S_j + \mu_j^S$ conditional on Y. Further, this density can be expressed as

$$p_{S_j+\mu_j^S|Y}(s) = \int p_{S_j+\mu_j^S|Y,\Psi}(s) \, p_{\Psi|Y}(\psi) d\psi.$$

Here we represent multiple integration through a single integral sign, for economy of notation. This introduces the posterior for the parameter random vector Ψ, denoted $p_{\Psi|Y}$. This function is assumed to be already known, being determined during the model estimation phase via usual sampling methods. If we substitute into the expression for the conditional expectation we obtain

$$\mathrm{E}(S_j + \mu_j^S|Y) = \int\int s\, p_{S_j+\mu_j^S|Y,\Psi}(s)\, ds\, p_{\Psi|Y}(\psi)\, d\psi.$$

The expression in the interior is equal to $\mathrm{E}(S_j + \mu_j^S|Y,\psi)$, which is given by a simple formula. This is because, conditional on ψ, both S_j and Y are Gaussian, and μ_j^S and μ^N are deterministic. Simple extensions of McElroy (2008) to the case of fixed effects reveals that $\mathrm{E}(S + \mu^S|Y,\psi) = F(\psi)\,[Y-\mu] + \mu^S$. That is, we first remove the fixed effects from Y (this is possible since they are known conditional on ψ) then apply the matrix $F(\psi)$, and then add μ^S back in. In some cases $F(\psi)\mu = \mu^S$ so that we can just apply $F(\psi)$ to Y, but this need not always be the case (see below). So the final solution can be expressed as

$$\mathrm{E}(S + \mu^S|Y) = \int \left(\mu^S + F(\psi)\,[Y-\mu]\right)\, p_{\Psi|Y}(\psi)\, d\psi,$$

interpreted component-wise. One way to approach this computation is integrate on the coefficients of the matrix $F(\psi)$ and the fixed effects. For example, component j can be written

$$\mathrm{E}(S_j+\mu_j^S|Y) = \int \mu_j^S\, p_{\Psi|Y}(\psi)\, d\psi - \sum_{k=1}^{n}\int F_{jk}(\psi)\mu_k\, p_{\Psi|Y}(\psi)\, d\psi + \sum_{k=1}^{n}\int F_{jk}(\psi)p_{\Psi|Y}(\psi)\, d\psi\; Y_k.$$

So one could take simulations of ψ from its posterior distribution, plug these values into the formula for $F(\psi)$, μ^S, and μ^N, to get $F_{jk}(\psi)$ and the fixed effects for each $j,k = 1,2,\cdots,n$, and average over the whole chain. This gives the Monte Carlo approximation to the above integrals. When this is finished, we apply the smoothed matrix to Y and add the appropriately smoothed fixed effects.

Application. We now discuss the particular application of the above exposition to the framework of this paper. To avoid confusion, we now let S denote the seasonal and N the nonseasonal once again. Consider the case of seasonal differencing $1 - B^{12}$ for data with no trading day or other fixed effects (or assume they have been previously removed). Then the mean effects μ_t naturally break into two portions: a centered periodic effect μ_t^S and a linear trend effect μ_t^N (see Bell, 1984a, 1995, 2004, for discussion). The former is annihilated by $U(B)$, whereas the latter can be written $\alpha_0 + \alpha_1 t$. Then it follows that $\zeta = 12\alpha_1$; thus when we estimate the mean from the differenced data, it is interpretable as being proportional to the slope in the trend mean effect. With F the signal extraction matrix for the nonseasonal, from (1.3.2) we know that $F\mu^S$ is identically zero (essentially, F contains the $U(B)$ operator). Moreover, letting I_n denote the n-dimensional identity matrix, $I_n - F$ is the seasonal extraction matrix and hence contains a $1-B$ factor; thus we can write $F\mu^N = \mu^N - (I_n - F)\,\mu^N = \mu^N - \alpha_1 \cdot H 1_n$, where $I_n - F = H\Delta_N$ defines the matrix H. This is true because single differencing on the linear trend reduces it to constancy, represented through the vector $\alpha_1 1_n$. Hence it follows that $\mathrm{E}(N + \mu^N | Y, \psi) = FY + \alpha_1 H 1_n$, which we note does not depend on α_0. Finally, $\mathrm{E}(N + \mu^N | Y) = \int (\alpha_1 H(\psi) 1_n + F(\psi) Y) p_{\Psi|Y}(\psi)\, d\psi$.

Acknowledgments

The authors would like to thank the Editor, William Bell, and two anonymous referees for useful comments that helped improve this chapter.

Bibliography

Bell, W. (1984a). Seasonal Decomposition of Deterministic Effects. Technical report, Research Report Number 84/01, Statistical Research Division, U.S. Census Bureau.

Bell, W. (1984b). Signal Extraction for Nonstationary Time Series. *The Annals of Statistics*, 12:646 – 664.

Bell, W. (1995). Correction to Seasonal Decomposition of Deterministic Effects. Technical report, Research Report 95/01, Statistical Research Division, U.S. Census Bureau.

Bell, W. R. (2004). On RegComponent Time Series Models and Their Applications. In Harvey, A., Koopman, S., and Shephard, N., editors, *State Space and Unobserved Component Models: Theory and Applications*. Cambridge University Press.

Beran, J., Bhansali, R. J., and Ocker, D. (1998). On Unified Model Selection for Stationary and Nonstationary Short- and Long-Memory Autoregressive Processes. *Biometrika*, 85:921–934.

Bertelli, S. and Caporin, M. (2002). A Note on Calculating Autocovariances of Long-memory Processes. *Journal of Time Series Analysis*, 23(5):503–508.

Bisognin, C. and Lopes, S. R. C. (2009). Properties of seasonal long memory processes. *Mathematical and Computer Modelling*, 49:1837–1851.

Bloomfield, P. (1973). An Exponential Model for the Spectrum of a Scalar Time Series. *Biometrika*, 60:217–226.

Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods*. Springer-Verlag.

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. (1995). Understanding the Metropolis-Hastings Algorithm. *American Statistician*, 49:327–335.

Dagum, E. B. (1980). *The X-11-ARIMA Seasonal Adjustment Method*. Catalogue No. 12-564E, Ottawa: Statistics Canada.

Gelman, A., Carlin, J., Stern, H., and Rubin, D. (2003). *Bayesian Data Analysis, Second Edition*. Chapman & Hall.

Granger, C. W. J. and Joyeux, R. (1980). An Introduction to Long-memory Time Series Models and Fractional Differencing. *Journal of Time Series Analysis*, 1:15–29.

Harvey, A. C. (1990). *Forecasting, Structural Time Series Models, and the Kalman Filter*. Cambridge University Press.

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.

Holan, S., McElroy, T., and Chakraborty, S. (2009). A Bayesian Approach to Estimating the Long Memory Parameter. *Bayesian Analysis*, 4:159–190.

Holan, S. H., Lund, R., and Davis, G. (2010). The ARMA Alphabet Soup: A Tour of ARMA Model Variants. *Statistics Surveys*, 4:232–274.

Hosking, J. R. M. (1981). Fractional Differencing. *Biometrika*, 68:165–176.

Hsu, N.-J. and Tsai, H. (2009). Semiparametric Estimation for Seasonal Long-memory Time Series Using Generalized Exponential Models. *Journal of Statistical Planning and Inference*, 139(6):1992–2009.

McElroy, T. (2008). Matrix Formulas for Nonstationary ARIMA Signal Extraction. *Econometric Theory*, 24:1–22.

McElroy, T. and Gagnon, R. (2008). Finite Sample Revision Variances for ARIMA Model-Based Signal Extraction. *Journal of Official Statistics*, 24(3):451.

McElroy, T. and Holan, S. (2011). On the Computation of Autocovariances for Generalized Gegenbauer Processes. Technical report, U.S. Census Bureau.

Palma, W. (2007). *Long-Memory Time Series: Theory and Methods*. Wiley-Interscience.

R Development Core Team (2010). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Robinson, P. (2003). *Time Series With Long Memory*. Oxford University Press.

Shiskin, J., Young, A., and Musgrave, J. (1967). The X-11 Variant of the Census II Seasonal Adjustment Program. Technical report, Bureau of the Census, Washington D.C.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linde, A. (2002). Bayesian Measures of Model Complexity and Fit (Pkg: P583-639). *Journal of the Royal Statistical Society, Series B: Statistical Methodology*, 64(4):583–616.

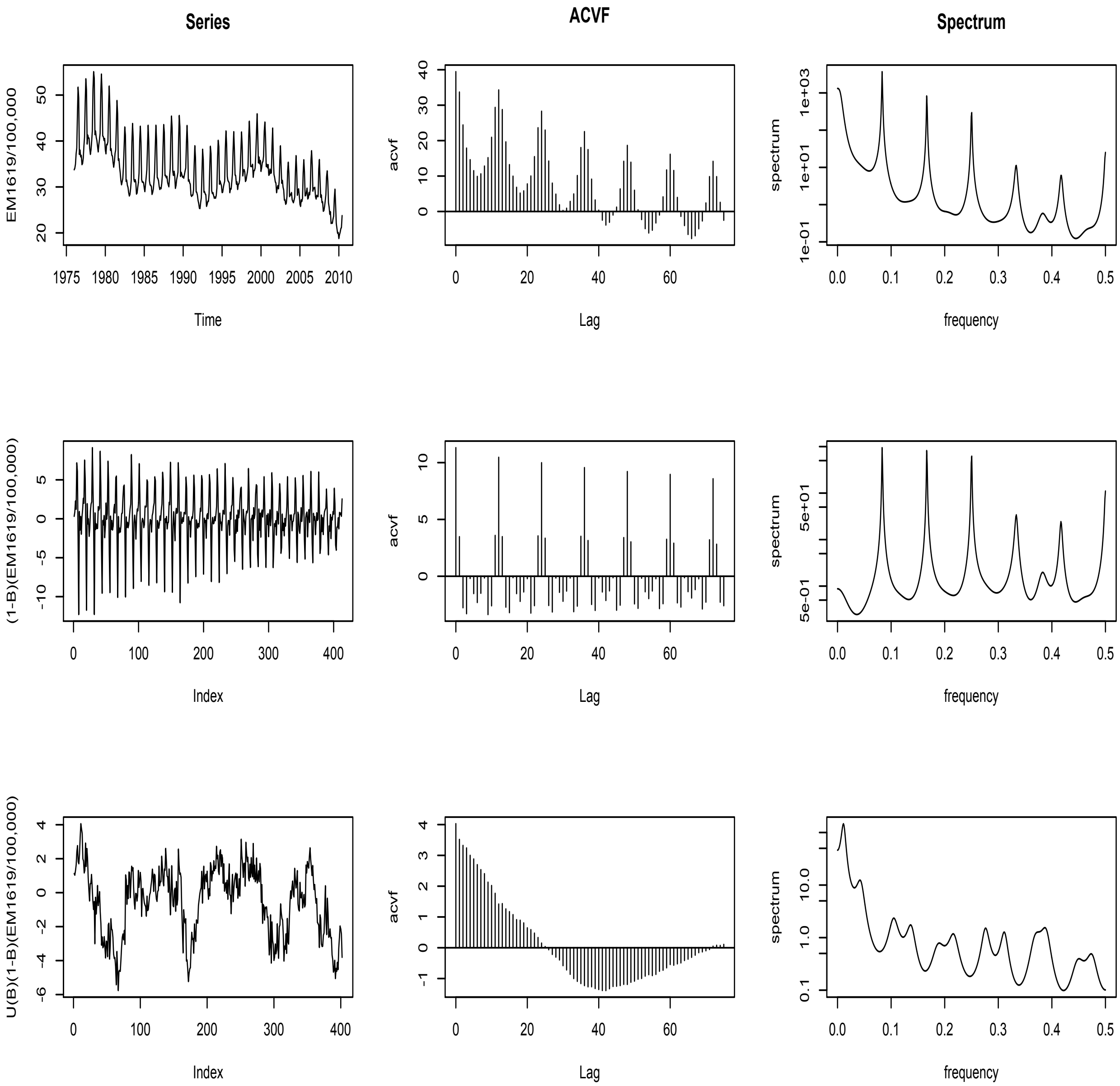

Figure 1.1: Exploratory data analysis plot for the Current Population Survey - Employed Males aged 16 -19 from 1/1976-6/2010 (EM1619). The first row displays the time series plot, the autocovariance (acvf) plot and the AR(30) spectrum. The second row displays the first differenced data (i.e. $(1-B)Y_t$) along with the corresponding acvf plot and AR(30) spectrum. The third row displays $U(B)(1-B)Y_t$ along with the corresponding acvf plot and AR(30) spectrum

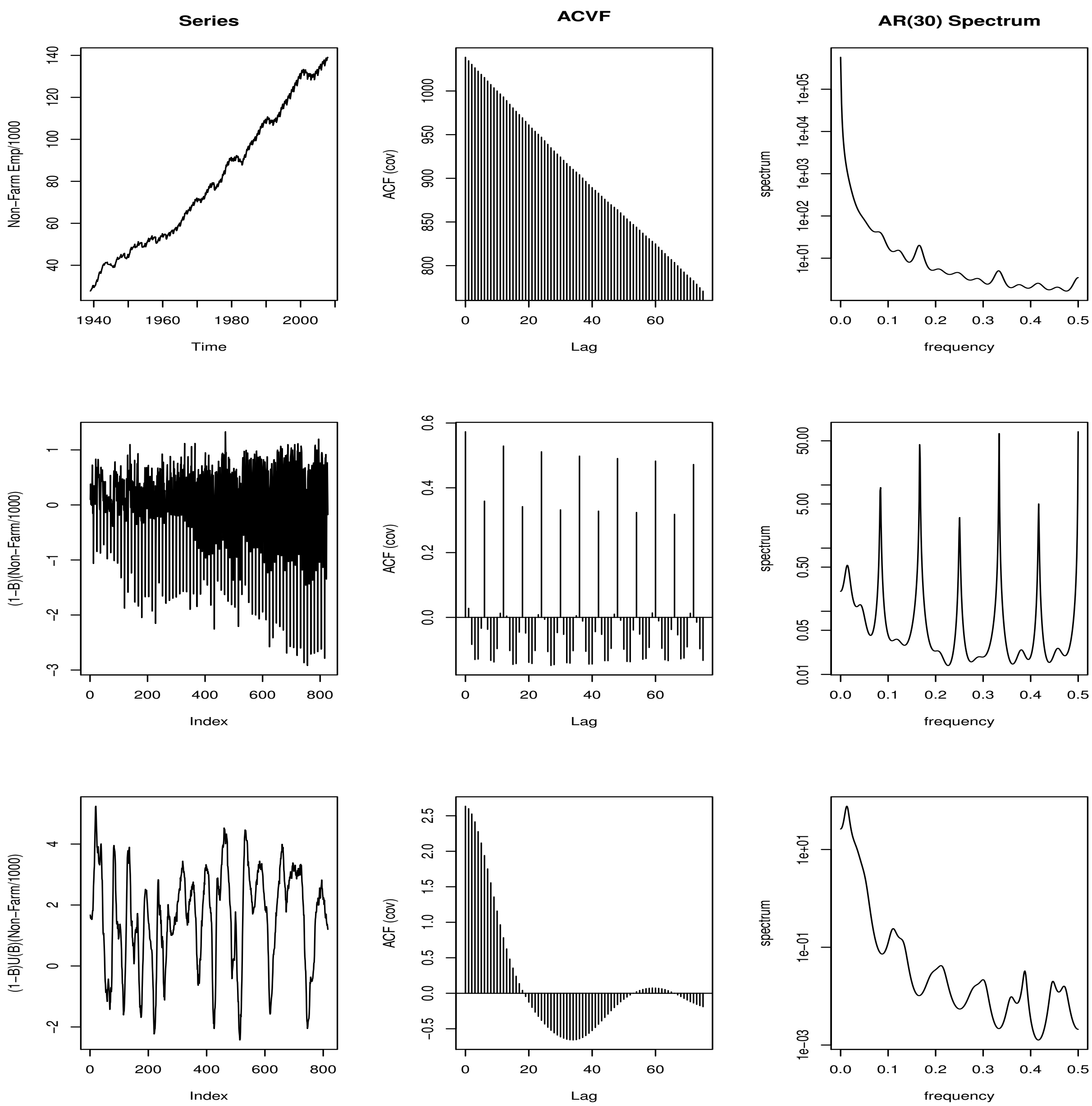

Figure 1.2: Exploratory data analysis plot for the U.S. Total Non-farm Employment from 1/1939-7/2009. The first row displays the time series plot, the autocovariance (acvf) plot and the AR(30) spectrum. The second row displays the first differenced data (i.e. $(1-B)Y_t$) along with the corresponding acvf plot and AR(30) spectrum. The third row displays $U(B)(1-B)Y_t$ along with the corresponding acvf plot and AR(30) spectrum

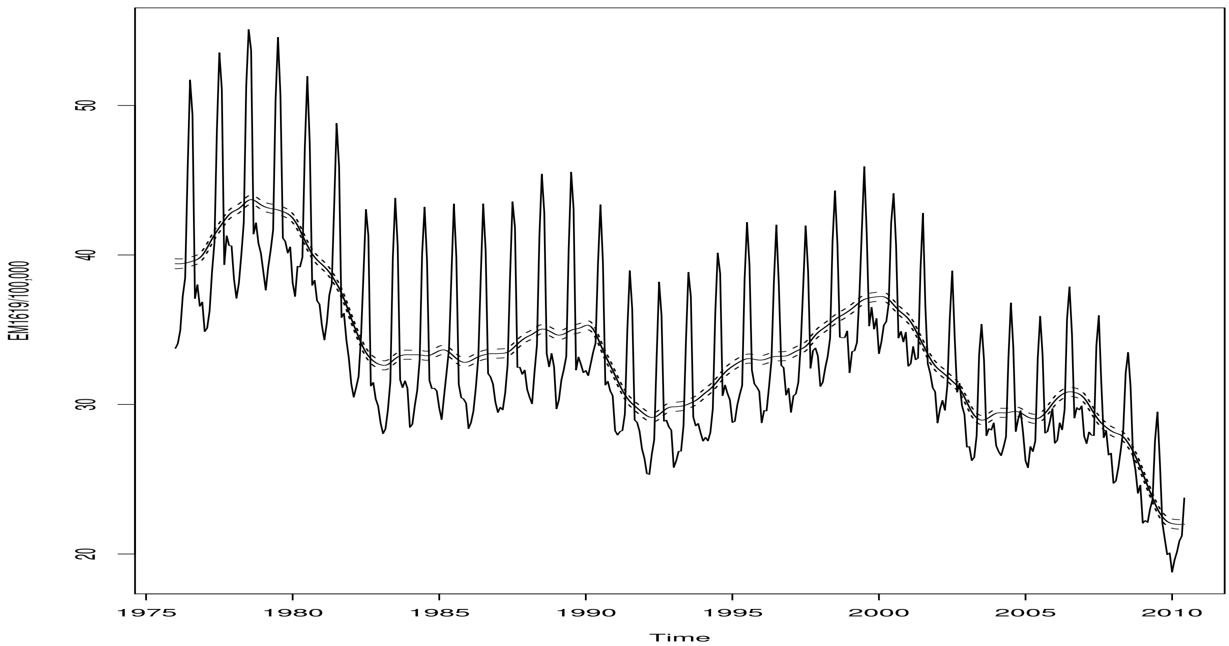

Figure 1.3: Time plot, for the Employed Males series (EM1619), with estimated nonseasonal component derived from a LM-UC(1,3) model using maximum likelihood. The dashed lines represent the estimated pointwise 95% confidence interval

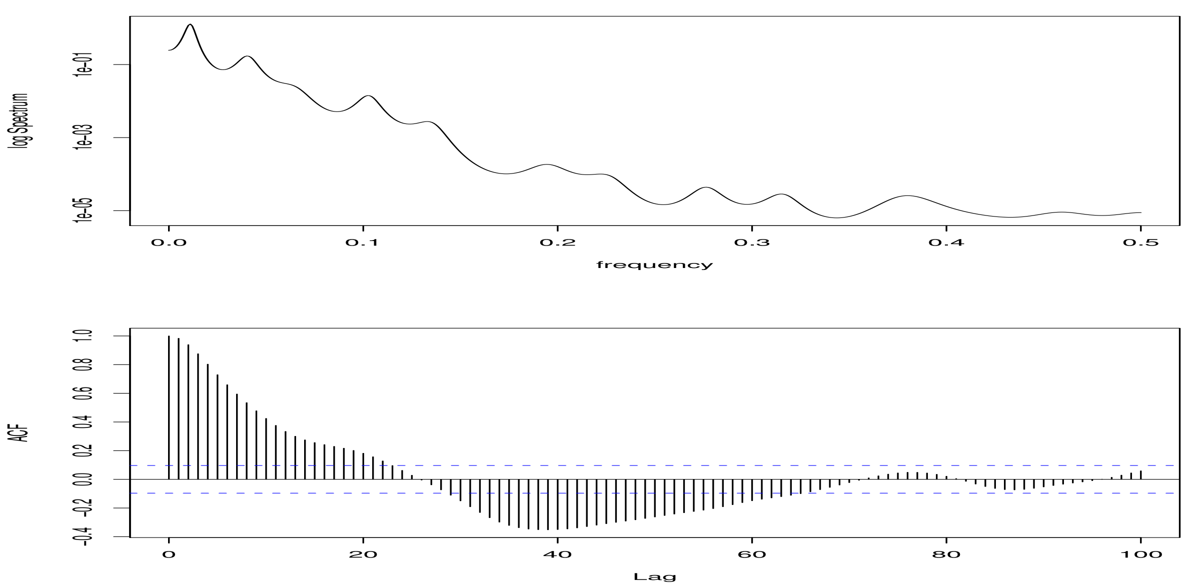

Figure 1.4: Employed Males series (EM1619) - maximum likelihood: The top panel plots the AR(30) spectrum of the differenced estimated trend (nonseasonal) estimate. The bottom plot displays the sample autocorrelation function of the differenced trend (nonseasonal) estimate.

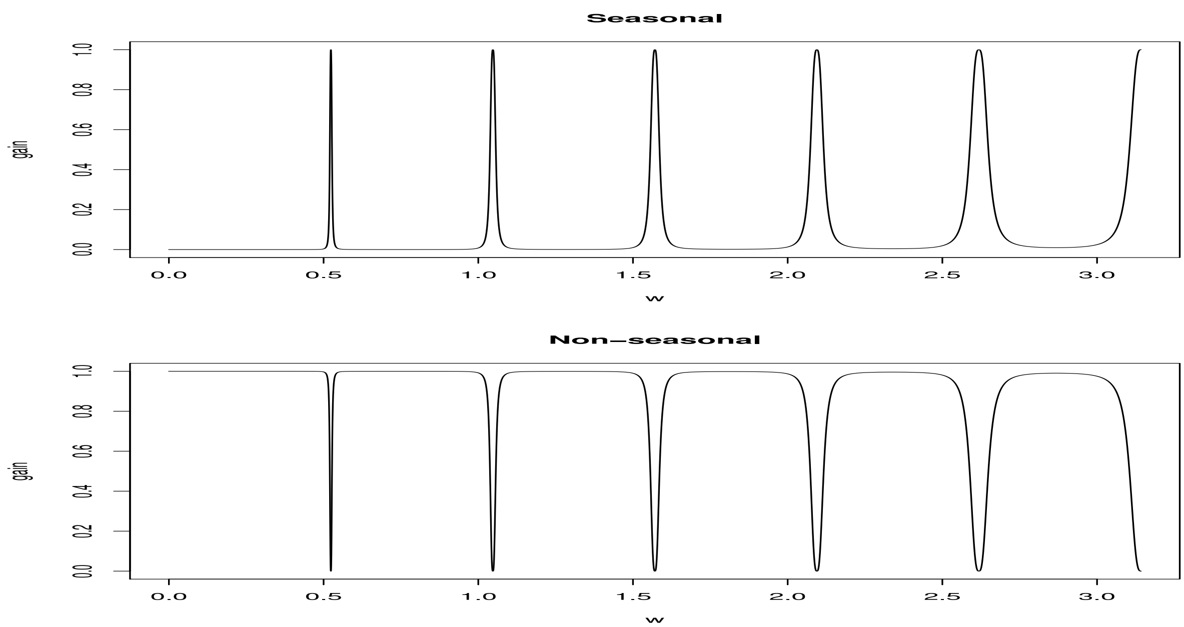

Figure 1.5: Employed Males series (EM1619): The top panel plots the gain function for the seasonal filter. The bottom panel displays the gain function for the nonseasonal filter.

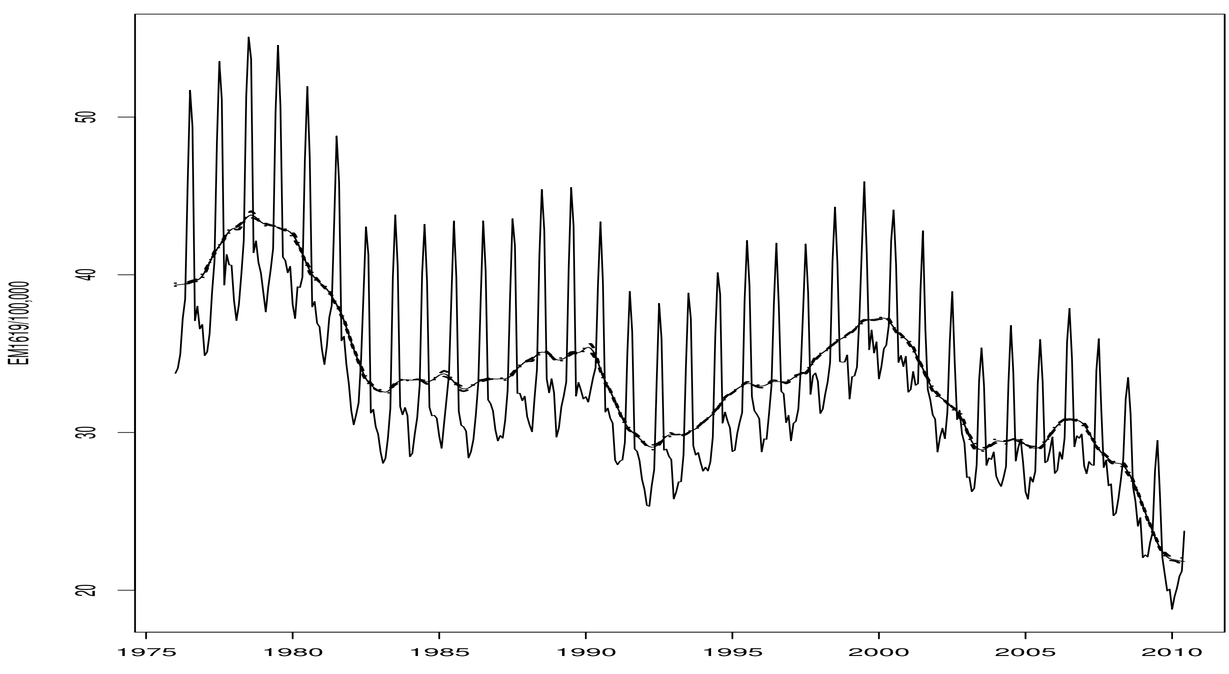

Figure 1.6: Employed Males series (EM1619): Time plot with estimated nonseasonal component derived from a LM-UC(1,5) model using Bayesian estimation. The dashed lines represent the estimated pointwise 95% credible interval

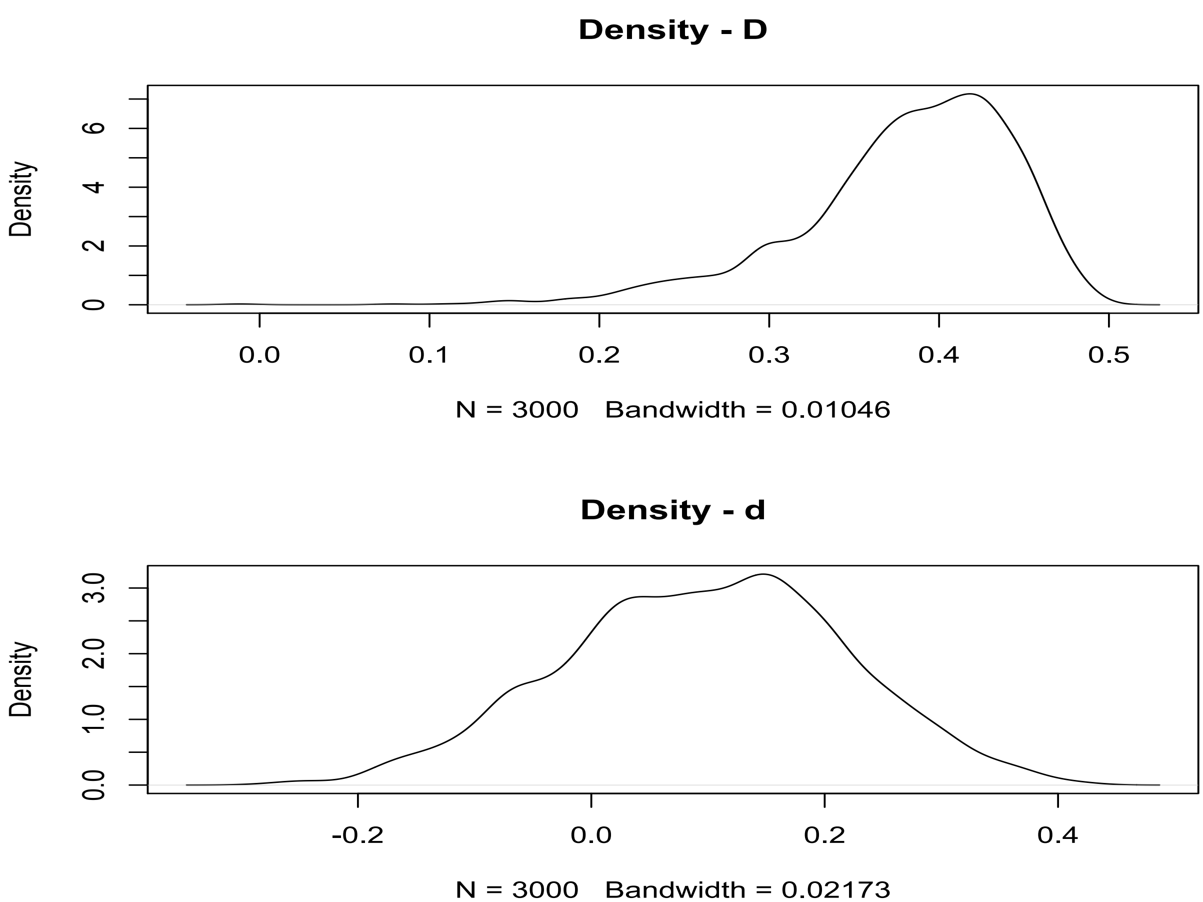

Figure 1.7: Employed Males series (EM1619): Kernel density estimate of the posterior distribution of the memory parameters D and d from a Bayesian LM-UC(1,5) (using the default *density* command in R).

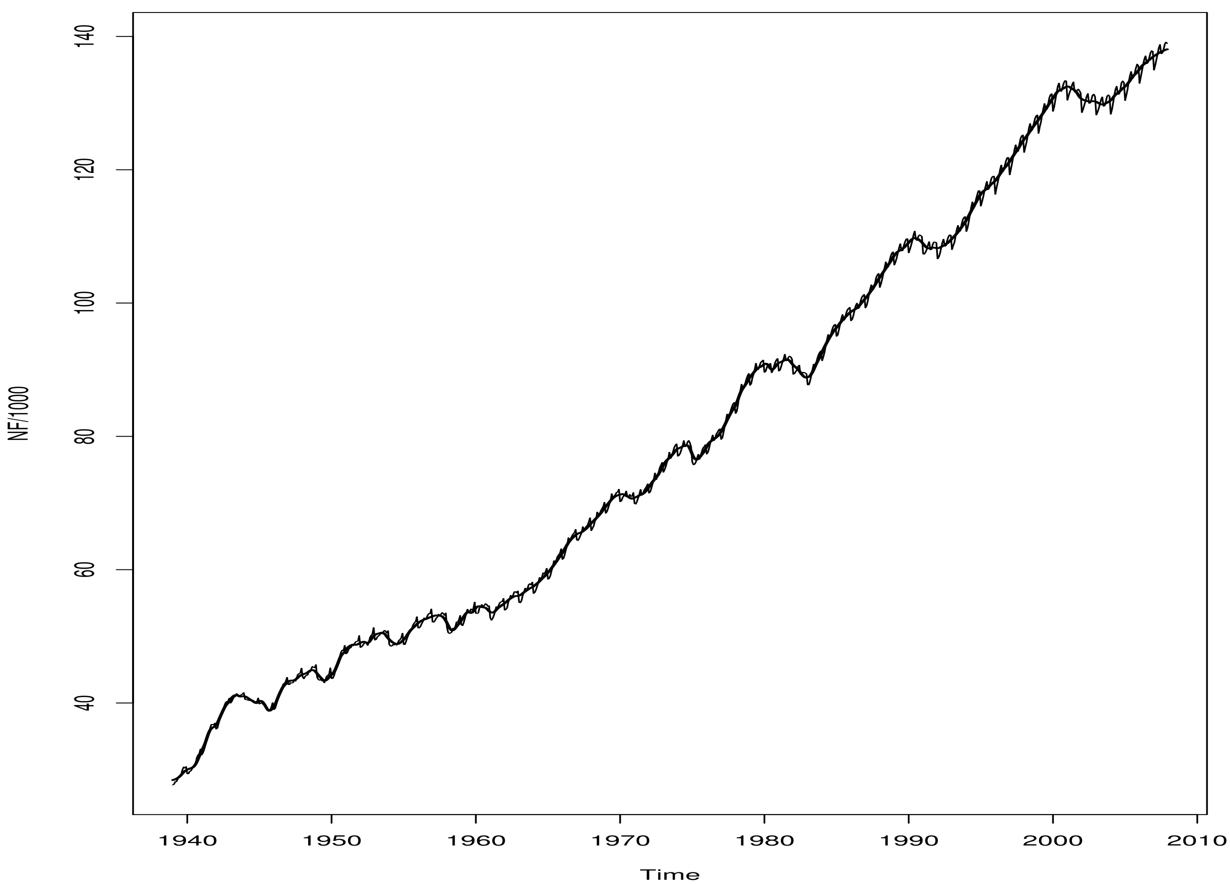

Figure 1.8: Non-farm employment series - Bayesian estimation: Time plot with estimated nonseasonal component derived from a LM-UC(2,3) model using Bayesian estimation. The estimated pointwise 95% CI has been suppressed since the width is uniformly less than .25.

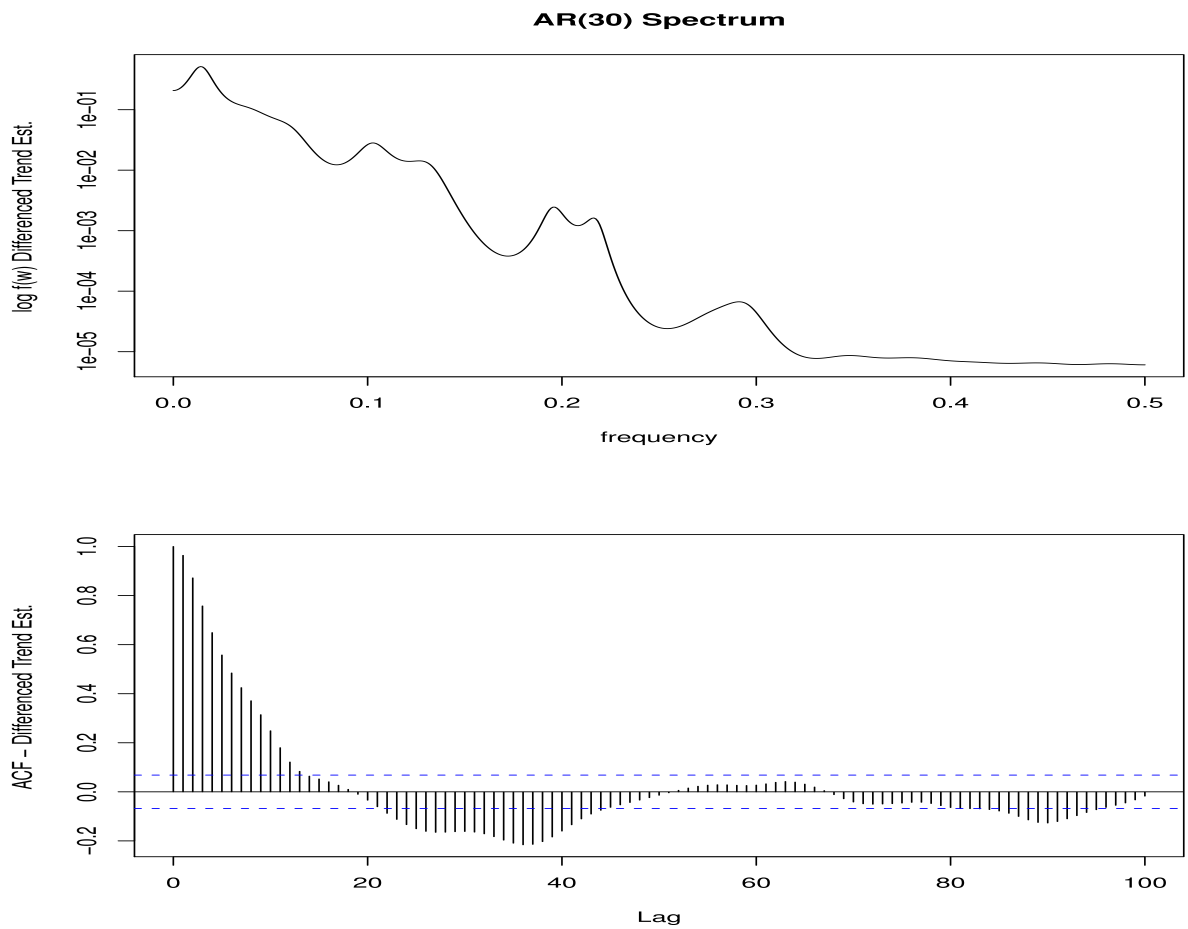

Figure 1.9: Non-farm employment series - Bayesian estimation: The top panel plots the AR(30) spectrum of the differenced estimated trend (nonseasonal) estimate. The bottom plot displays the sample autocorrelation function of the differenced trend (nonseasonal) estimate.

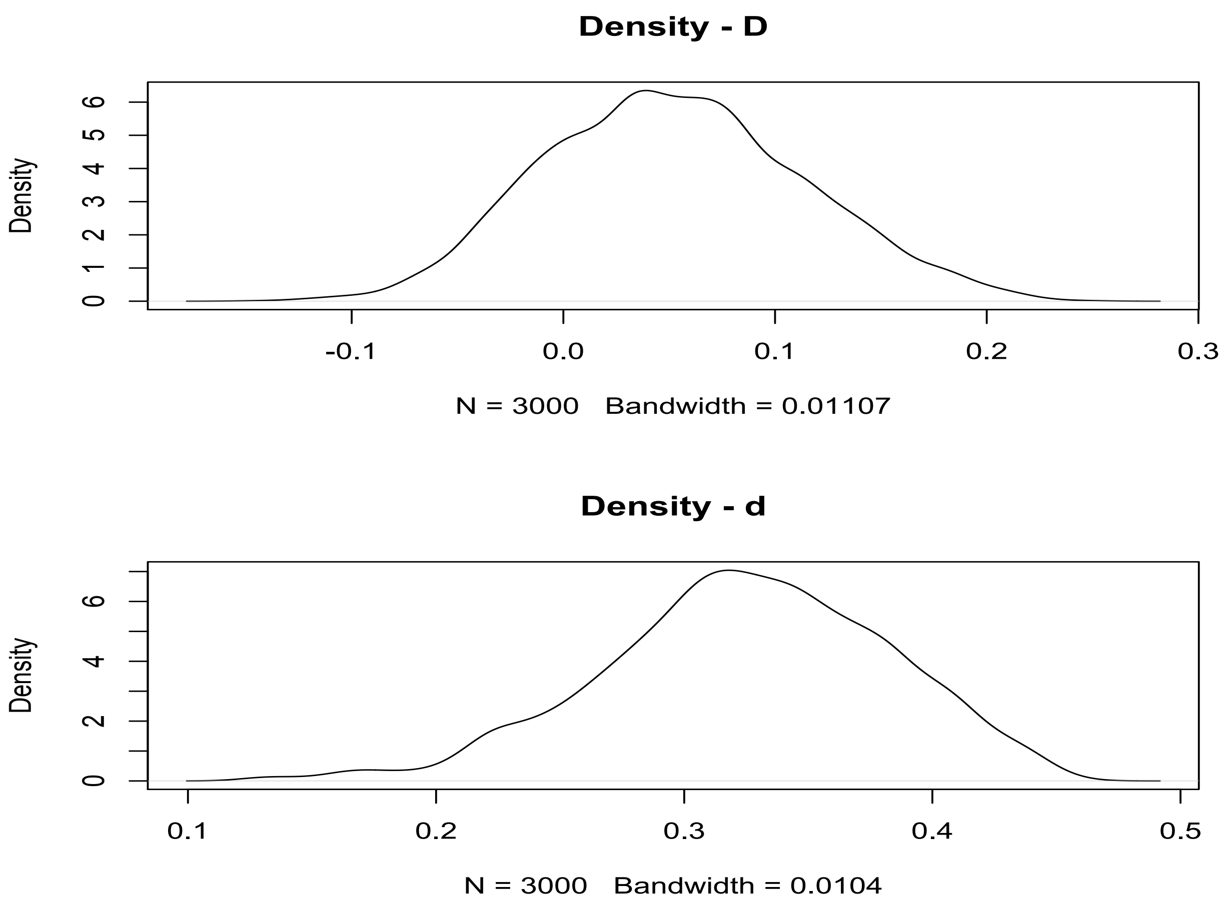

Figure 1.10: Non-farm employment series: Kernel density estimate of the posterior distribution of the memory parameters D and d from a Bayesian LM-UC(2,3) (using the default *density* command in R).