# Chapter 1

# The Error in Business Cycle Estimates Obtained from Seasonally Adjusted Data

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

## 1.1 Introduction

Economists have great interest in measuring the business cycle inherent to economic time series; see Baxter and King (1999); Kaiser and Maravall (2001); Stock and Watson (2004) and the references therein. A popular method for estimating the cycle is the Hodrick-Prescott (HP) filter (Hodrick and Prescott, 1997). Since the HP is designed for nonseasonal series, it can only be sensibly applied to seasonally adjusted data. This serves as the context of this chapter: cycle estimation for univariate time series. Given that it is not uncommon for economists to start with seasonally adjusted data, which is typically produced without taking cyclical effects into account, it is natural to ask the question: how is the estimation of the cycle affected? If

[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.*

there is additional signal extraction error due to the prior activity of seasonal adjustment, can this error be quantified? We set out to provide answers to these important questions.

In general, economic time series undergo several distinct patterns of stochastic variation. Apart from trading-day and holiday effects (Bell and Hillmer, 1983; Roberts et al., 2010; Soukup and Findley, 2000), the major stochastic components are the trend, seasonal, cycle and irregular (see Peña et al. (2001), Harvey and Proietti (2005), Durbin and Koopman (2001) and the references therein). In this paper, we focus on an observed series $Y$ that can be viewed, after suitable transformations and removal of fixed regression effects, as a sum of four unobserved components: cycle $C$, trend $T$, seasonal $S$, and irregular $I$. In particular we assume

$$Y_t = C_t + T_t + S_t + I_t \tag{1.1.1}$$

for times $t = 1, 2, \cdots, n$. This basic model of economic time series plays a prominent role in ARIMA model-based signal extraction methods, such as structural time series models (Harvey, 1989) and SEATS (Maravall and Caporello, 2004).

Define the "two-stage approach" to cycle estimation via first obtaining an estimate of $T + I$ given $Y$ (typically ignoring the putative presence of $C$), followed by estimation of $C$ from the seasonally adjusted data. This is clearly not an ideal approach; ignoring $C$ may corrupt the estimation of $S$ when these components share common spectral power. However, it arises in practice due to two principal considerations: (1) some of the most popular cycle estimation filters (e.g., the HP) do not achieve suppression of seasonal frequencies; (2) it is difficult to estimate cyclical models from unadjusted data (see the discussion in Kaiser and Maravall (2005)). However, were this point resolved then one could proceed with direct estimation of the cycle for the unadjusted series, and the result would be less prone to error. Recognizing this difficulty, Kaiser and Maravall (2005) developed the so-called "recast" approach, which can be viewed as generating a HP-hybrid filter that suppresses seasonality. This technique utilizes the information about the seasonal, trend, and irregular models from the seasonal adjustment stage. More dubious is the approach that models the seasonally adjusted data as $C + T + I$, resulting in a mis-statement of the cycle mean squared error (MSE)[2].

This chapter first provides an in-depth discussion of the above techniques in Section 1.2. One can give a precise theoretical description of the approximate values of maximum likelihood estimates obtained by fitting mis-specified models via the machinery of "pseudo-true" values, and this will help us to see exactly what

[2]Mis-stated MSE refers to a reported MSE quantity that is calculated based upon erroneous assumptions.

is happening in the two-stage approach. Knowing the pseudo-true parameter values, we can then precisely quantify the asymptotic MSE of cycle estimates, using signal extraction formulas derived in Sections 1.3 and 1.4; these are derived in generality for possible application to other scenarios. Having written computer programs to compute both pseudo-true values and MSEs for each method, given an initial specification of Data Generating Process (DGP), we display some of the most interesting results and provide discussion in Section 1.5. A full illustration is provided on an employment series in Section 1.6, and Section 1.7 concludes. All proofs are left to the Appendix.

## 1.2 Methods of Cycle Estimation

As alluded to in the Introduction, there are several methods of model-based cycle estimation that are utilized in practice. The best method from a theoretical standpoint models all the dynamics present in (1.1.1) and constructs the cycle extraction filter in a MSE optimal fashion. This will be called the *direct* method, and will serve as a benchmark for other techniques. In particular, one obtains models for the cycle process $C_t$ and the data $Y_t$, and computes the conditional expectation $\mathrm{E}[C_t|Y]$ as a linear transformation of the data vector $Y$, utilizing estimated parameter values.

The *two-stage* method begins with a seasonal-trend-irregular model that does not explicitly account for cyclical behavior. In practice (as is subsequently demonstrated) the cyclical dynamics, when actually present, become incorporated in the models for seasonal and trend, in the sense that the parameter estimates for these models are somewhat different than they would be if no cyclicality were present. Mathematically, the situation is described through the behavior of the pseudo-true values of the parameters, which are by definition the limit of maximum likelihood estimates (mles) when a mis-specified model is fitted to the data (Taniguchi and Kakizawa, 2000). Although this material has existed in the time series literature for many decades, we review it in the Appendix. Seasonal adjustment filters can then be computed by using the fitted model (either using model-based filters or a fixed forecast-extended filter as in X-11-ARIMA), and the resulting seasonal adjustment is the input for the second stage: one can model this as cycle-trend-irregular. The idea is that any cyclical effects initially present in the data are carried through into the seasonal adjustment. One then computes the MSE optimal cycle estimate from the seasonally adjusted data.

The actual MSE for this *two-stage* procedure can be somewhat different from what would be derived from the cycle signal extraction formulas, which assume that the given data has no seasonality. However, some seasonality remains to distort the error formulas, so that ignoring this contribution results in a mis-statement of the MSE; this is demonstrated theoretically and empirically below.

If instead of applying a cycle extraction filter in the second stage, we use a nonparametric filter, such as the HP, we obtain the *recast* method of Kaiser and Maravall (2005). Using the presumed dynamics of the seasonally adjusted output as given, one can know the exact dynamics of a putative cycle implied by use of such a filter. However, the actual cycle (if one exists) may differ in its dynamics from those suggested by the *recast* approach, which can result in a mis-statement of the MSE.

This defines the three techniques considered here: *direct*, *two-stage*, and *recast*. In Sections 1.3 and 1.4 below we present mathematical formulas to correctly compute, in finite sample, the MSE of the cycle. These formulas do not take parameter uncertainty into account. Instead the MSEs are computed as if the signal extraction filters are given, whereas in practice they actually depend on the data through the parameter estimates[3]. We also present a method for computing the MSE for a bi-infinite sample, when the mles have converged to their pseudo-true values. For large samples the mles for the parameters will be close to the pseudo-true values, so the method for the bi-infinite sample can provide us with an asymptotic picture of the true MSE, including the mis-specified case. Results for both the finite sample and bi-infinite sample MSEs are presented in Section 1.5.

The work of this chapter is predicated on a model-based approach to signal extraction; see McElroy (2008a) for a discussion of model estimation and optimal signal extraction in the context of finite sample econometric time series data. We remind the reader that the two most popular approaches to modeling are the decomposition and structural techniques. The decomposition method (Burman, 1980; Hillmer and Tiao, 1982) starts with a Seasonal AutoRegressive Integrated Moving Average (SARIMA) model for the time series, identified according to standard unit root tests and diagnostics such as Ljung-Box. Then component models for seasonal, trend, cycle, and irregular can be deduced using a partial fraction decomposition. In contrast, the structural method (Gersch and Kitagawa, 1983; Harvey, 1989) begins with putative SARIMA models for each component in (1.1.1) (so-called structural models are parameter-restricted SARIMA models) and derives a model for the data, called the reduced form, and uses the resulting likelihood as the objective

[3]It is quite difficult to determine analytically the impact of parameter uncertainty on MSE, although this could be done numerically by conducting Monte Carlo studies.

function. We will have recourse to both approaches in this paper.

## 1.3 Basic Results on Two-Stage Signal Extraction

This section presents formulas for signal extraction MSE from a fairly generic perspective. We begin by writing the decomposition (1.1.1) in vector form:

$$Y = C + S + T + I.$$

So $Y = \{Y_1, Y_2, \cdots, Y_n\}'$, and similarly for the components. For most applications, the seasonal and trend components are difference stationary, with associated differencing polynomials $\delta^S(z)$ and $\delta^T(z)$ respectively, whereas the cycle and irregular are stationary. However, Theorems 1 and 2 of this section are proved generally, in that any of the components may be stationary or nonstationary (i.e., integrated); the one assumption is that their differencing polynomials are relatively prime, i.e., share no common zeroes. The differencing polynomial for $Y_t$ is $\delta(z) = \delta^S(z)\delta^T(z)$, and $\partial Y_t = \delta(B)Y_t$ is stationary, where $B$ is the backshift operator. These polynomials $\delta$, $\delta^S$, and $\delta^T$ all have their zeroes located on the unit circle of the complex plane, and their leading coefficient is one by convention. We let the differenced components be defined as $\partial S_t = \delta^S(B)S_t$ and $\partial T_t = \delta^T(B)T_t$, which are mean zero weakly stationary time series. This introduces a general notation: if $X_t$ consists of any combination of stationary or integrated components, then $\partial X_t = \delta^X(B)X_t$, where $\delta^X$ is the polynomial containing all necessary differencing factors, with coefficients [4] $\delta_j^X$. Let $d$ be the order of $\delta$, and let $d_S$ and $d_T$ be the orders of $\delta^S$ and $\delta^T$ respectively. Clearly $d = d_S + d_T$. For example, the seasonal operator $\delta^S$ for monthly data would be $U(z) = 1 + z + z^2 + \cdots + z^{11}$, and $\delta^T(z) = (1-z)^2$ is appropriate for a second-order trend.

We assume Assumption A of Bell (1984) holds on the component decomposition, appropriately generalized to four components. Assumption A states that the initial values, i.e., the variables $Y_* = (Y_1, Y_2, \cdots, Y_d)'$, are uncorrelated with the differenced component series $\{\partial S_t\}$, $\{\partial T_t\}$, $C_t$, and $I_t$. Bell (1984), Bell and Hilmer (1988), McElroy and Sutcliffe (2006), and McElroy (2008a) all discuss the implications of this assumption. Note that mean square error optimal signal extraction filters derived under Assumption A agree exactly with the filters implicitly used by a properly initialized state space smoother (Bell and Hillmer, 1991). We will

[4] The sign convention is $\delta^X(z) = 1 + \delta_1^X z + \cdots + \delta_d^X z^d$, which differs from Bell (1984).

also assume that the differenced components $\{\partial S_t\}$, $\{\partial T_t\}$, $C_t$, and $I_t$ are uncorrelated with one another.

Next, we formulate these notations in matrix form. Let $\Delta$ be a $n - d \times n$ matrix with entries given by $\Delta_{ij} = \delta_{i-j+d}$ (the convention being that $\delta_k = 0$ if $k < 0$ or $k > d$). The matrices $\Delta_S$ and $\Delta_T$ have entries given by the coefficients of $\delta^S(z)$ and $\delta^T(z)$, but are $n - d_S \times n$ and $n - d_T \times n$ dimensional respectively. This means that each row of these matrices consists of the coefficients of the corresponding differencing polynomial, horizontally shifted in an appropriate fashion. Hence we can write $\partial Y = \Delta Y$, $\partial S = \Delta_S S$, $\partial T = \Delta_T T$. Now to relate these quantities we need to define further differencing matrices $\underline{\Delta}_T$ and $\underline{\Delta}_S$, which have row entries given by the coefficients of $\delta^T(z)$ and $\delta^S(z)$ respectively, but are $n - d \times n - d_S$ and $n - d \times n - d_T$ dimensional. Then the equation that relates the differenced components is

$$\partial Y = \underline{\Delta}_T \partial S + \underline{\Delta}_S \partial T + \Delta C + \Delta I,$$

which uses the relationship

$$\Delta = \underline{\Delta}_T \Delta_S = \underline{\Delta}_S \Delta_T \tag{1.3.1}$$

proved in McElroy and Sutcliffe (2006). We will be principally interested in making estimates of $C$. For each $1 \leq t \leq n$, the minimum mean squared error signal extraction estimate is $\widehat{C}_t = \mathrm{E}[C_t|Y]$. This can be expressed as a certain linear function of the data vector $Y$ when the data are Gaussian. This estimate is also the minimum mean squared error *linear* estimate when the data are non-Gaussian. For the remainder of the paper, we do not assume Gaussianity, and by optimality we always refer to the minimum mean squared error *linear* estimate. Writing $\widehat{C} = (\widehat{C}_1, \widehat{C}_2, \cdots, \widehat{C}_n)'$, the coefficients of these linear functions form the rows of a matrix $F$, namely $\widehat{C} = FY$. Letting $\Sigma_X$ denote the covariance matrix of any random vector $X$, the formula for $F$ is

$$F = \Sigma_C \Delta' \Sigma_{\partial Y}^{-1} \Delta, \tag{1.3.2}$$

see, e.g. McElroy (2008a). From this formula, it is clear that we only need a model for $C$ and a model for the differenced data $\partial Y$; it is not necessary to compute models for $S$, $T$, and $I$, but only for their aggregate.

Now since we will be analyzing two-stage filtering procedures, we require a more flexible notation, which generalizes ideas in McElroy and Sutcliffe (2006). Since we wish to state our results more generally, we will consider a scenario with three components $\{\alpha_t\}$, $\{\beta_t\}$, and $\{\gamma_t\}$ that are possibly integrated, with relatively prime differencing polynomials $\delta^\alpha$, $\delta^\beta$, and $\delta^\gamma$; let $\{\partial\alpha_t\}$, $\{\partial\beta_t\}$, and $\{\partial\gamma_t\}$ be the differenced components.

Denote by $F^X_{XZ}$ the signal extraction filter matrix for signal $X$ and noise $Z$ derived under the appropriate analogue of Assumption A. In this fashion we define $F^{\alpha}_{\alpha\beta}$, $F^{\alpha}_{\alpha\beta\gamma}$, $F^{\alpha\beta}_{\alpha\beta\gamma}$, etc., the formulas for these being given explicitly in Theorem 1 of McElroy (2008a) – although the initial value assumptions under which they are derived are inconsistent[5]. The following theorem shows how the filter matrices are related.

**Theorem 1.** *The filter matrices are algebraically related via* $F^{\alpha}_{\alpha\beta\gamma} = F^{\alpha}_{\alpha\beta} F^{\alpha\beta}_{\alpha\beta\gamma}$.

Letting $\alpha$, $\beta$, and $\gamma$ denote trend, cycle, seasonal, irregular, or combinations thereof as appropriate, we have the following corollary.

**Corollary 1.** *The following relations hold:* $F^C_{CSTI} = F^C_{CT} F^{CT}_{CSTI} = F^C_{CTI} F^{CTI}_{CSTI} = F^C_{CSI} F^{CSI}_{CSTI}$.

Now $F^C_{CSTI}$ is the optimal cycle filter in the *direct* approach, while the other three expressions in Corollary 1 all represent different two-step approaches, where the parameter values used in the first and second stage are assumed to be those of the true DGP. $F^C_{CT} F^{CT}_{CSTI}$ corresponds to the *recast* approach, since the trend and cycle are initially lumped together in the trend-cycle component $CT$, and this is followed by a "high-pass" filter for cycle estimation. $F^C_{CTI} F^{CTI}_{CSTI}$ corresponds to a two-step approach to cycle estimation; the two-stage approach used in practice – and is studied further in this chapter – is $F^C_{CTI} F^{TI}_{STI}$, but this does not appear in Corollary 1 since it does not yield the optimal filter. The fourth expression is not pursued further in this paper, but illustrates yet another way of viewing cycle estimation. The Appendix contains some related results that do not directly enter into the main trajectory of the paper, but are of independent interest.

## 1.4 Errors in Cycle Estimation

In this section, we consider the various cycle estimation methods discussed in the Introduction, and provide formulas for the cycle estimation filter $F$ and error covariance matrix $M$. In general, the MSEs are the diagonal entries of the error covariance matrix $M$. We employ the following notation. The true covariance matrix of a random vector $X$ is denoted by $\Sigma_X$, while a proxy for this quantity, perhaps based upon some fitted model for $\{X_t\}$, is denoted by $\dot{\Sigma}_X$. If there is a second proxy, it is denoted by $\ddot{\Sigma}_X$. Below, we will consider proxies $\dot{\Sigma}_X$ for $\Sigma_X$ used in the first stage of the filtering, obtained perhaps by fitting a model to the data; in the second stage we have different proxies $\ddot{\Sigma}_X$ that arise from different models. Note that even

[5] That is, we assume Assumption A holds for the decomposition $\alpha + \beta + \gamma$, while also assuming Assumption A holds for the decomposition $\alpha + \beta$, etc.

though the mles that enter into these proxies are data-dependent, and hence random, we treat them as deterministic in the variance formulas developed below. Finally, $\dot{F}$ and $\ddot{F}$ will denote filter matrices utilized in the first and second stages, respectively. Derivations are provided in the Appendix.

### 1.4.1 Direct Approach

The *direct* method was discussed in Section 1.3, and from (1.3.2) it follows that

$$\begin{aligned} \dot{F}^{C}_{CSTI} &= \dot{\Sigma}_C \Delta' \dot{\Sigma}^{-1}_{\partial Y} \Delta \\ M &= \left( \dot{\Sigma}_C^{-1} + \Delta' \dot{\Sigma}^{-1}_{\partial STI} \Delta \right)^{-1} \left( \dot{\Sigma}_C^{-1} \Sigma_C \dot{\Sigma}_C^{-1} + \Delta' \dot{\Sigma}^{-1}_{\partial STI} \Sigma_{\partial STI} \dot{\Sigma}^{-1}_{\partial STI} \Delta \right) \left( \dot{\Sigma}_C^{-1} + \Delta' \dot{\Sigma}^{-1}_{\partial STI} \Delta \right)^{-1}, \end{aligned} \quad (1.4.1)$$

where $\partial STI = \Delta(S + T + I)$. Typically, the proxies $\dot{\Sigma}_C$ and $\dot{\Sigma}_{\partial STI}$ are substituted for $\Sigma_C$ and $\Sigma_{\partial STI}$ in (1.4.1), resulting in the stated error covariance of

$$\left( \dot{\Sigma}_C^{-1} + \Delta' \dot{\Sigma}^{-1}_{\partial STI} \Delta \right)^{-1} = \dot{\Sigma}_C - \dot{\Sigma}_C \Delta' \dot{\Sigma}^{-1}_{\partial Y} \Delta \dot{\Sigma}_C.$$

### 1.4.2 Two-stage Approach

As mentioned in Section 1.2 the *two-stage* method corresponds to using the filter $\dot{F}^{TI}_{STI}$ followed by $\ddot{F}^{C}_{CTI}$. Note that typically no accounting of the cycle is made in the first stage, so cyclical behavior gets incorporated into the other three components. In the second stage it is now assumed that a cycle is present in the seasonally-adjusted data, so that a $C + T + I$ type model is fitted. The filter and error covariance matrices are

$$\begin{aligned} \ddot{F}^{C}_{CTI} \dot{F}^{TI}_{STI} &= \ddot{\Sigma}_C \Delta'_T \ddot{\Sigma}^{-1}_{\partial CTI} \dot{\Sigma}_{\partial TI} \underline{\Delta}'_S \dot{\Sigma}^{-1}_{\partial STI} \Delta, \\ M &= \Sigma_C + \ddot{\Sigma}_C \Delta'_T \ddot{\Sigma}^{-1}_{\partial CTI} \dot{\Sigma}_{\partial TI} \underline{\Delta}'_S \dot{\Sigma}^{-1}_{\partial STI} \Sigma_{\partial Y} \dot{\Sigma}^{-1}_{\partial STI} \underline{\Delta}_S \dot{\Sigma}_{\partial TI} \ddot{\Sigma}^{-1}_{\partial CTI} \Delta_T \ddot{\Sigma}_C \\ &\quad - \ddot{\Sigma}_C \Delta'_T \ddot{\Sigma}^{-1}_{\partial CTI} \dot{\Sigma}_{\partial TI} \underline{\Delta}'_S \dot{\Sigma}^{-1}_{\partial STI} \Delta \Sigma_C - \Sigma_C \Delta' \dot{\Sigma}^{-1}_{\partial STI} \underline{\Delta}_S \dot{\Sigma}_{\partial TI} \ddot{\Sigma}^{-1}_{\partial CTI} \Delta_T \ddot{\Sigma}_C. \end{aligned}$$

This presumes that the order of trend differencing used in the first and second stage models is the same. This $M$ is problematic to approximate via proxies, since it requires a knowledge of both the first and second stage models. In practice it is mis-stated via $\ddot{\Sigma}_C - \ddot{\Sigma}_C \Delta' \ddot{\Sigma}^{-1}_{\partial CTI} \Delta \ddot{\Sigma}_C$, which is the error covariance for extracting

$C$ from a $C+T+I$ process, with proxies substituted.

### 1.4.3 Recast Approach

As alluded to in Section 1.2, the *recast* method is given by $\ddot{F}^{C}_{CT} \dot{F}^{CT}_{CSTI}$, where $\ddot{F}$ here denotes a filter matrix chosen in some *ad hoc* fashion; e.g., using model-based HP filters with pre-determined parameter values. The filter and error covariance matrices are

$$\ddot{F}^{C}_{CT} \dot{F}^{CT}_{CSTI} = \ddot{\Sigma}_C \Delta' \dot{\Sigma}^{-1}_{\partial STI} \Delta,$$
$$M = \Sigma_C + \ddot{\Sigma}_C \Delta' \dot{\Sigma}^{-1}_{\partial STI} \Sigma_{\partial Y} \dot{\Sigma}^{-1}_{\partial STI} \Delta \ddot{\Sigma}_C - \ddot{\Sigma}_C \Delta' \dot{\Sigma}^{-1}_{\partial STI} \Delta \Sigma_C - \Sigma_C \Delta' \dot{\Sigma}^{-1}_{\partial STI} \Delta \ddot{\Sigma}_C.$$

Here $\dot{\Sigma}_{\partial STI}$ refers to the covariance matrix of a non-cyclical model $\Delta(S+T+I)$, which is mis-specified when a cycle is truly present. The recasting method of SEATS assumes that $\Sigma_C = \ddot{\Sigma}_C$, essentially taking this as definition of the cyclical model. Using the proxy $\dot{\Sigma}_{\partial STI}$ for $\Sigma_{\partial Y}$, the error covariance would then be stated as $\ddot{\Sigma}_C - \ddot{\Sigma}_C \Delta' \dot{\Sigma}^{-1}_{\partial STI} \Delta \ddot{\Sigma}_C$; this is optimal when the true cycle process is define as the HP high-pass filter applied to the trend.

## 1.5 Simulation Results

We now make an assessment of the cycle MSE for the methods discussed above. For each of the three methods, there is a true MSE and a mis-stated MSE based on using certain proxies or other dubious assumptions. In this section we simulate time series of various sample sizes, apply each of the three procedures, and generate both cycle estimates and MSE curves. By repeating the simulations, we obtain a picture of the overall effect. We also look at the bi-infinite sample situation, by computing the pseudo-true values and generating the asymptotic MSEs for all three methods via the techniques described in the Appendix. In both approaches (finite and infinite) we ignore the contribution of mle variability to the stated MSEs.

In particular, we begin by simulating a process with both a cycle $\{C_t\}$ and seasonal-trend-irregular $\{X_t\}$, the latter being given by an Airline process, i.e.,

$$(1-B)^2 U(B) X_t = (1-\theta_X B)(1-\Theta_X B^{12}) \epsilon^X_t.$$

We take the parameters $\theta_X = .6$, $\Theta_X = .6$, and $\sigma_X^2 = 1$ for our initial simulation. The cyclical component is given by an AR(2) model:

$$(1 - 2\rho\cos(\omega)B + \rho^2 B^2)C_t = \epsilon_t^C. \tag{1.5.1}$$

Our DGP has $\rho = .9$, $\omega = \pi/12$, and $\sigma_C^2 = \kappa\sigma_X^2$; we set $\kappa = 1$ corresponding to a strong cycle. So our DGP is $Y_t = C_t + X_t$, though it is only required to simulate $\partial Y_t = (1-B)^2 U(B) Y_t$ since all cycle estimates can be written in terms of the differenced data. Note carefully that from inception this exercise favors the *direct* and *two-stage methods*; to fairly evaluate the *recast* technique, we should set $\kappa = 0$ and define $C_t = (1 - H(B))T_t$ in the DGP, where $H(B)$ is the HP filter and $\{T_t\}$ is the trend process (from the canonical decomposition of $\{X_t\}$).

The next step is to estimate the correctly specified model by fitting the $C + X$ model to the DGP; this is done using a likelihood based on structural ARIMA models. We can immediately compute *direct* cycle estimates and the true MSE via (1.4.1), as well as the mis-stated MSE. The diagonal entries of these MSE matrices are of principal interest. For both the *recast* and *two-stage* methods we next fit an Airline model to the DGP, which of course is a mis-specified model when $\kappa > 0$. From these estimates we obtain the canonical decomposition, and hence the signal extraction filters $\dot{F}_{STI}^{T}$ and $\dot{F}_{STI}^{TI}$. In the *recast* method the cycle covariances $\ddot{\Sigma}_C$ are obtained as described in the Appendix, utilizing the HP filter with $q = 1/130,000$ (the value appropriate for monthly series – Hodrick and Prescott (1997)) and the spectral representation described in McElroy (2008b). Then we obtain at once the cycle estimates, the true cycle MSE, and the mis-stated cycle MSE.

Finally, the seasonally adjusted data from the previous step is taken as input to the *two-stage* method, and we fit a cycle-trend-irregular model with components $C$, $T$, and $I$. The cycle follows (1.5.1) as before, the irregular $I$ is white noise, and the trend follows a Smooth Trend structural model, i.e., $(1-B)^2 T_t$ is white noise. These model choices were made to mimic structural models popular with econometricians (though our cycle model (1.5.1) differs from that of Harvey (1989)). Once we have the parameter estimates, again obtained via a structural likelihood, we can compute $\ddot{\Sigma}_C$ and $\ddot{\Sigma}_{\partial CTI}$ and thereby obtain the cycle estimate, true cycle MSE, and mis-stated cycle MSE[6].

Our numerical studies examine several specifications of the basic DGP outlined above for three sample sizes, computing for each of the three methods the cycle estimate, the true MSE curve (i.e., the diagonal

[6]The same procedure can be followed for real data, except for the construction of true MSE (for any of the three techniques), since the DGP is unknown.

of the $M$ matrices described in Section 1.4), and the mis-stated MSE curve. We simulated 200 realizations of each DGP with $n = 120, 180, 240$ and $\kappa = 1$, and the other parameter choices mentioned above. These sample sizes were chosen to reflect lengths typical of series being seasonally adjusted at statistical agencies. The mles tended to be accurate for the *direct* method for $n = 180, 240$: the root MSE over the 200 simulations for $\rho$ and $\omega$ is .0382 and .0497 at sample size 180, and .0297 and .0303 at sample size 240. The results of this simulation study are extensive and, as a result, we only provide a summary of the most salient aspects along with a plot from a representative realization (with $n = 180$), presented in Figures 1.1, 1.2, and 1.3.

***Figure 1 about here***

Figure 1.1 displays the various cycle estimates along with the true generating process for a particular simulation. The *direct* estimate is closest to the true cycle, with the *recast* approach doing an excellent job as well, and the *two-stage* method performing worst. The temporal averages of squared differences – for the *direct*, *two-stage*, and *recast* methods versus the truth – for this particular simulation were 2.19, 15.98, and 3.90 respectively. In the spectral domain (Figure 1.2) it is apparent that the *direct* method captures the underlying second order dynamics more closely than either the *two-stage* or *recast* approach, though it should be noted that even the MSE optimal *direct* estimate will always have different spectral characteristics from the true cycle, as is well-known in the signal extraction literature. In addition, Figure 1.3 illustrates that the *direct* method has the lowest MSE, followed by the *recast* approach and finally the *two-stage* method. Also, mis-stated MSEs tend to be higher than the true MSEs for these latter two methods. One key conclusion from the empirical studies is that the mis-stated MSEs and true MSEs for the *direct* method tend to be very close indeed.

***Figure 2 about here***

***Figure 3 about here***

The situation can be more easily digested by considering the limiting case of a bi-infinite sample size. We next present MSEs for all three methods, using pseudo-true values for parameters and frequency domain calculations described in the Appendix. This is equivalent to examining the center values of the matrix-based finite-sample MSEs when the sample size $n$ is extremely large, so that the mles have converged. Of course this calculation is intractable for large $n$; hence we adopt the frequency domain methodology outlined in the Appendix to thereby obtain the exact asymptotic quantities. Note that the MSE is no longer time-varying. Tables 1.1 and 1.2 display pseudo-true values and asymptotic MSEs for the three methods and for a variety of DGPs.

***Table 1 about here***

***Table 2 about here***

The results indicate that the *two-stage* approach has slightly higher MSE than the optimal *direct* estimate. The *recast* method fares poorly for large values of $\kappa$, which is to be expected given the parametric form of our DGP, but improves somewhat as $\kappa$ shrinks. It is interesting to see which pseudo-true values are produced from the fitting of the generally mis-specified airline model to the DGP, and which cycle parameters arise from the *two-stage* method. In this latter case, fairly accurate cycle parameters can be obtained even when seasonal adjustment is first done (Table 1.1) so long as $\rho$ is large. Most interesting, perhaps, are the results when $\kappa = 0$, i.e., no cycle actually exists in the DGP. In this case the cycle parameter values obtained for the *direct* method are meaningless, but the airline model parameters match those of the DGP.

## 1.6 Illustration

The preceding results highlight the dangers of the *two-stage* method. We next provide a further illustration on the time series of Unemployment Level (Job losers in thousands, 16 years and over) published by the Bureau of Labor Statistics (http://data.bls.gov). This has the series identification LBU03023621, denoted hereafter as LBU for short. We examine the period from January 1967 through December 2007, purposefully excluding the current economic crisis. Figure 1.4 displays a plot of the logged series, along with its seasonal adjustment in the upper panel.

***Figure 4 about here***

No calendar effects were found to be significant. The data was fitted with the cycle-plus-airline model of the previous section, as well as the airline model and the model of the *two-stage* approach. The cycle-plus-airline model had mles of $\widehat{\rho} = .973$, $\widehat{\omega} = .09$ (a period of 69.69 months), $\widehat{\kappa} = .016$, $\widehat{\theta}_X = 0$, $\widehat{\Theta}_X = .817$, and $\widehat{\sigma}^2_X = .0028$. Fitting the airline model to the same data yields $\widehat{\theta}_X = -.080$, $\widehat{\Theta}_X = .612$, and $\widehat{\sigma}^2_X = .2092$. Unfortunately, the *two-stage* approach yields mles of $\widehat{\rho} = .471$, $\widehat{\omega} = .366$ (a period of 17.16 months), $\widehat{\sigma}^2_C = .0516$, $\widehat{\sigma}^2_T = .0218$, and $\widehat{\sigma}^2_I = .0370$. The signal-to-noise ratio is adequate to capture a component, but the values of $\rho$ and $\omega$ cannot be associated with a business cycle.

The *direct* estimate successfully captures the cycle of given parametric form, and this is remarkably corroborated by the result of the *recast* method, as seen in the second panel of Figure 1.4. The failure of the *two-stage* method on LBU is disappointing, but highlights the difficulty of fitting cyclical models in practice[7]. One is tempted to conclude that estimation of the cycle has been impacted by the seasonal adjustment procedure. However, it must be acknowledged that our results are contingent upon a certain parametric specification of the cycle. In practice, an economist encountering such a result with the *two-stage* method would have recourse to fixing the value of $\rho$ and/or $\omega$ *a priori* and re-estimating the other parameters – or utilizing the HP filter on the estimated trend. The analysis would not end here.

This example does advocate the *direct* approach, although it must be acknowledged that the cycle-plus-airline model is rarely successful, in our experience, in the sense that many series viewed by economists as having a business cycle (e.g., housing starts and permits series) result in infeasible estimates of $\rho$ and/or $\omega$. However, when it works correctly, a key advantage of the *direct* estimation method is a fairly accurate statement of the MSE, as borne out by our simulation studies. The central values (which also happen to be the minimum values) of the MSE curves are .015 and .759 for the *direct* and *recast* methods respectively. We know that the latter value tends to severely over-estimate the true MSE in practice. Particularly seeing as how the two cyclical estimates are so close (their empirical mean square difference is .0022), it seems reasonable to conclude that they both are providing decent estimates of an underlying cycle, but the *direct* method does a much better job of quantifying the MSE.

## 1.7 Conclusion

This chapter examines the problem of cycle estimation in the presence of seasonality. Under the rubric of a model-based univariate approach to cyclical dynamics, we quantify the real MSE of cycle estimation arising

[7]In our experience with dozens of economic series, rarely does the *direct* method produce viable cycle estimates, even while the *two-stage* approach is sometimes successful.

from three approaches: *direct* estimation, the *recast* approach of SEATS, and a *two-stage* approach involving cyclical analysis of seasonally adjusted data. The mathematical formulas are useful for demonstrating that, in finite samples, the real MSE can be quite different from the quantity that is typically stated. This discrepancy arises for two reasons: cyclical dynamics are ignored in the seasonal adjustment procedure, and the error in seasonal adjustment itself is ignored. The first problem may afflict the *recast* method (depending on the form of the true cycle), while both problems certainly afflict the *two-stage* approach. Another tool we introduce for the study of the problem is the actual limiting MSE based on a bi-infinite sample, where parameter estimates have been assumed to converge to their limiting pseudo-true values. This is informative for showing how model mis-specification and the signal-to-noise ratio of cycle to seasonal-trend dynamics impacts the MSE for the three methods; the pseudo-true values themselves are also helpful for understanding how spurious cycles can arise from seasonal adjustment procedures.

The scope of our numerical study is necessarily limited. We have chosen a particular form of the true cycle that biases the outcomes against the *recast* method; all results are based on certain choices of models (e.g., airline models and basic structural models), which we elected in an effort to mimic economic practice. It would be interesting to consider the impact of seasonal adjustment by X-11 filters, or by a model that encompasses autoregressive effects to allow for cyclical dynamics. To any of these further questions, the formulas and tools of this paper can be readily applied with some additional effort. The issue of spurious cycles certainly merits further study.

Although the *direct* approach can be difficult in practice – due to the challenge of getting reasonable mles – it is quite appealing when it works, because we know empirically that the stated MSE is a decent approximation to the true MSE. Moreover, cycle estimates arising from this approach are likely to be the most accurate when the model is correctly specified. If we wish to remain agnostic about the parametric form of the cycle, the *recast* method can be utilized. The caution here is that the stated MSE can be quite inaccurate if a parametric cycle is indeed present; it can also perform optimally when the cycle takes the form of an HP high-pass of the trend. The performance of the *two-stage* approach can be adequate in some cases, but in general the quantification of the MSE is highly problematic, leading to false inferences about turning points.

## 1.8 Appendix

### 1.8.1 Further Results Related to Theorem 1

The following theorem provides an explicit formula for the filter $F^{\alpha}_{\alpha\beta\gamma}$ in terms of "two-component" filters $F^{\alpha}_{\alpha\beta}$, $F^{\beta}_{\beta\gamma}$, etc. It is a generalization of Theorem 1 of McElroy and Sutcliffe (2006) to the case where all 3 unobserved components are difference stationary. We let 1 denote a shorthand for the $n \times n$ identity matrix.

**Theorem 2.** *The matrices* $1 - F^{\alpha}_{\alpha\gamma} F^{\beta}_{\beta\gamma}$ *and* $1 - F^{\alpha}_{\alpha\beta} F^{\gamma}_{\beta\gamma}$ *are invertible, and*

$$F^{\alpha}_{\alpha\beta\gamma} = \left(1 - F^{\alpha}_{\alpha\gamma} F^{\beta}_{\beta\gamma}\right)^{-1} F^{\alpha}_{\alpha\gamma}(1 - F^{\beta}_{\beta\gamma}) = \left(1 - F^{\alpha}_{\alpha\beta} F^{\gamma}_{\beta\gamma}\right)^{-1} F^{\alpha}_{\alpha\beta}(1 - F^{\gamma}_{\beta\gamma}).$$

The invertibility claimed in Theorem 2 depends upon the assumption that the differencing operators are relatively prime. This result can be applied to cycle estimation via the following corollary:

**Corollary 2.** *We have the following expressions:*

$$F^{C}_{CSTI} = \left(1 - F^{C}_{CTI} F^{S}_{STI}\right)^{-1} F^{C}_{CTI}(1 - F^{S}_{STI}) = \left(1 - F^{C}_{CT} F^{SI}_{STI}\right)^{-1} F^{C}_{CT}(1 - F^{SI}_{STI}).$$

The second expression in this result is interesting, because the filter $F^{C}_{CTI}(1 - F^{S}_{STI})$ corresponds to seasonal adjustment that ignores cyclical dynamics followed by cycle estimation from a nonseasonal specification. The result indicates that this matrix fails to equal the correct filter $F^{C}_{CSTI}$, but requires further application of $\left(1 - F^{C}_{CTI} F^{S}_{STI}\right)^{-1}$. The third expression likewise shows that trend filtering without utilizing cyclical dynamics, followed by $F^{C}_{CT}$ for cycle estimation, does not equal the optimal cycle filter, but the additional application of $\left(1 - F^{C}_{CT} F^{SI}_{STI}\right)^{-1}$ is required.

### 1.8.2 Proofs

**Proof of Theorem 1.** First we note that $F^{\alpha}_{\alpha\beta\gamma} + F^{\beta}_{\alpha\beta\gamma} + F^{\gamma}_{\alpha\beta\gamma} = 1$ implies that $F^{\alpha}_{\alpha\beta\gamma} + F^{\beta}_{\alpha\beta\gamma} = 1 - F^{\gamma}_{\alpha\beta\gamma} = F^{\alpha\beta}_{\alpha\beta\gamma}$. Now the theorem results from a commutativity property of the extraction matrices:

$$F^{\alpha}_{\alpha\beta} F^{\alpha\beta}_{\alpha\beta\gamma} = (1 - F^{\beta}_{\alpha\beta}) F^{\alpha}_{\alpha\beta\gamma} + F^{\alpha}_{\alpha\beta} F^{\beta}_{\alpha\beta\gamma} = F^{\alpha}_{\alpha\beta\gamma} + \left(F^{\alpha}_{\alpha\beta} F^{\beta}_{\alpha\beta\gamma} - F^{\beta}_{\alpha\beta} F^{\alpha}_{\alpha\beta\gamma}\right),$$

which uses $F^{\alpha}_{\alpha\beta} + F^{\beta}_{\alpha\beta} = 1$. The expression in parentheses is similar to a commutator for the matrices; below we show that it is zero, which will prove the theorem. Let the matrices $\underline{\Delta}_{\alpha\beta}$, $\underline{\Delta}_{\alpha\gamma}$, and $\underline{\Delta}_{\beta\gamma}$ be defined similarly to $\underline{\Delta}_T$ and $\underline{\Delta}_S$, but with coefficients of $\delta^\alpha\delta^\beta$, $\delta^\alpha\delta^\gamma$, and $\delta^\beta\delta^\gamma$ respectively, and with appropriate dimensions such that, similar to (1.3.1), $\underline{\Delta}_{\alpha\beta}\Delta_\gamma = \underline{\Delta}_{\alpha\gamma}\Delta_\beta = \underline{\Delta}_{\beta\gamma}\Delta_\alpha = \Delta$ is satisfied. Following McElroy (2008a), the filter formulas are given by

$$F^{\alpha}_{\alpha\beta} = \left(\Delta'_\alpha \Sigma^{-1}_{\partial\alpha}\Delta_\alpha + \Delta'_\beta \Sigma^{-1}_{\partial\beta}\Delta_\beta\right)^{-1} \Delta'_\beta \Sigma^{-1}_{\partial\beta}\Delta_\beta, \qquad \Delta_\beta F^{\beta}_{\alpha\beta\gamma} = \Sigma_{\partial\beta}\underline{\Delta}'_{\alpha\gamma}\Sigma^{-1}_W \Delta.$$

Therefore $F^{\alpha}_{\alpha\beta}F^{\beta}_{\alpha\beta\gamma} = \left(\Delta'_\alpha \Sigma^{-1}_{\partial\alpha}\Delta_\alpha + \Delta'_\beta \Sigma^{-1}_{\partial\beta}\Delta_\beta\right)^{-1} \Delta'\Sigma^{-1}_W\Delta$. Similarly,

$$F^{\beta}_{\alpha\beta} = \left(\Delta'_\alpha \Sigma^{-1}_{\partial\alpha}\Delta_\alpha + \Delta'_\beta \Sigma^{-1}_{\partial\beta}\Delta_\beta\right)^{-1} \Delta'_\alpha \Sigma^{-1}_{\partial\alpha}\Delta_\alpha, \qquad \Delta_\alpha F^{\alpha}_{\alpha\beta\gamma} = \Sigma_{\partial\alpha}\underline{\Delta}'_{\beta\gamma}\Sigma^{-1}_W \Delta,$$

which produces $F^{\beta}_{\alpha\beta}F^{\alpha}_{\alpha\beta\gamma} = \left(\Delta'_\alpha \Sigma^{-1}_{\partial\alpha}\Delta_\alpha + \Delta'_\beta \Sigma^{-1}_{\partial\beta}\Delta_\beta\right)^{-1} \Delta'\Sigma^{-1}_W\Delta$. This concludes the proof. □

**Proof of Theorem 2.** We prove the first line of the theorem. To show the invertibility of $1 - F^{\alpha}_{\alpha\gamma}F^{\beta}_{\beta\gamma}$, we compute

$$1 - F^{\alpha}_{\alpha\gamma}F^{\beta}_{\beta\gamma} = 1 - F^{\alpha}_{\alpha\gamma}(1 - F^{\gamma}_{\beta\gamma}) = F^{\gamma}_{\alpha\gamma} + F^{\alpha}_{\alpha\gamma}F^{\gamma}_{\beta\gamma}.$$

We need to define some additional differencing matrices. Let $\underline{\Delta}_\beta$ and $\underline{\Delta}_\gamma$ be defined similarly to the matrices in the proof of Theorem 1 above. Also, the matrix $\Delta_{\beta\gamma}$ does $\delta^\beta\delta^\gamma$ differencing, but is $n - (d_\beta + d_\gamma) \times n$ dimensional. Then we have the relation $\Delta_{\beta\gamma} = \underline{\Delta}_\beta\Delta_\gamma = \underline{\Delta}_\gamma\Delta_\beta$ as in (1.3.1). Let $\partial\beta\gamma$ denote the differenced $\beta+\gamma$ component: $\partial\beta\gamma = \Delta_{\beta\gamma}(\beta+\gamma) = \underline{\Delta}_\gamma\partial\beta + \underline{\Delta}_\beta\partial\gamma$. Then the following formulas hold:

$$\begin{aligned}
F^{\alpha}_{\alpha\gamma} &= \left(\Delta'_\alpha \Sigma^{-1}_{\partial\alpha}\Delta_\alpha + \Delta'_\gamma \Sigma^{-1}_{\partial\gamma}\Delta_\gamma\right)^{-1} \Delta'_\gamma \Sigma^{-1}_{\partial\gamma}\Delta_\gamma, \\
F^{\gamma}_{\alpha\gamma} &= \left(\Delta'_\alpha \Sigma^{-1}_{\partial\alpha}\Delta_\alpha + \Delta'_\gamma \Sigma^{-1}_{\partial\gamma}\Delta_\gamma\right)^{-1} \Delta'_\alpha \Sigma^{-1}_{\partial\alpha}\Delta_\alpha, \\
\Delta_\gamma F^{\gamma}_{\beta\gamma} &= \Sigma_{\partial\gamma}\underline{\Delta}'_\beta \Sigma^{-1}_{\partial\beta\gamma}\Delta_{\beta\gamma}, \\
F^{\alpha}_{\alpha\gamma}F^{\gamma}_{\beta\gamma} &= \left(\Delta'_\alpha \Sigma^{-1}_{\partial\alpha}\Delta_\alpha + \Delta'_\gamma \Sigma^{-1}_{\partial\gamma}\Delta_\gamma\right)^{-1} \Delta'_{\beta\gamma} \Sigma^{-1}_{\partial\beta\gamma}\Delta_{\beta\gamma}, \\
F^{\gamma}_{\alpha\gamma} + F^{\alpha}_{\alpha\gamma}F^{\gamma}_{\beta\gamma} &= \left(\Delta'_\alpha \Sigma^{-1}_{\partial\alpha}\Delta_\alpha + \Delta'_\gamma \Sigma^{-1}_{\partial\gamma}\Delta_\gamma\right)^{-1} \left[\Delta'_\alpha \Sigma^{-1}_{\partial\alpha}\Delta_\alpha + \Delta'_{\beta\gamma} \Sigma^{-1}_{\partial\beta\gamma}\Delta_{\beta\gamma}\right].
\end{aligned}$$

The expression in square brackets is invertible, since it is the sum of two non-negative definite matrices whose null spaces' intersection is the zero vector, due to the fact that $\delta^\alpha$ and $\delta^\beta\delta^\gamma$ are relatively prime; see

Lemma 2 of McElroy and Sutcliffe (2006). This establishes the invertibility of $1 - F^{\alpha}_{\alpha\gamma} F^{\beta}_{\beta\gamma}$, and gives an expression for its inverse. Finally, we compute the result:

$$\begin{aligned}\left(1 - F^{\alpha}_{\alpha\gamma} F^{\beta}_{\beta\gamma}\right)^{-1} F^{\alpha}_{\alpha\gamma}(1 - F^{\beta}_{\beta\gamma}) &= \left[\Delta'_{\alpha}\Sigma^{-1}_{\partial\alpha}\Delta_{\alpha} + \Delta'_{\beta\gamma}\Sigma^{-1}_{\partial\beta\gamma}\Delta_{\beta\gamma}\right]^{-1} \Delta'_{\gamma}\Sigma^{-1}_{\partial\gamma}\Delta_{\gamma} F^{\gamma}_{\beta\gamma}, \\ &= \left[\Delta'_{\alpha}\Sigma^{-1}_{\partial\alpha}\Delta_{\alpha} + \Delta'_{\beta\gamma}\Sigma^{-1}_{\partial\beta\gamma}\Delta_{\beta\gamma}\right]^{-1} \Delta'_{\gamma}\underline{\Delta}'_{\beta}\Sigma^{-1}_{\partial\beta\gamma}\Delta_{\beta\gamma} = F^{\alpha}_{\alpha\beta\gamma}. \quad \square\end{aligned}$$

### 1.8.3 Derivations for Section 1.4

**Direct Approach.** The formula for the filter matrix is (1.3.2); the corresponding error is

$$\dot{F}^{C}_{CSTI} Y - C = -\left(\dot{\Sigma}^{-1}_{C} + \Delta'\dot{\Sigma}^{-1}_{\partial STI}\Delta\right)^{-1} \dot{\Sigma}^{-1}_{C} C + \left(\dot{\Sigma}^{-1}_{C} + \Delta'\dot{\Sigma}^{-1}_{\partial STI}\Delta\right)^{-1} \Delta'\dot{\Sigma}^{-1}_{\partial STI}\partial STI,$$

from which (1.4.1) easily follows.

**Two-Stage Approach.** We use the fact that $\Delta_T \dot{F}^{TI}_{STI} = \dot{\Sigma}_{\partial TI}\underline{\Delta}'_S \dot{\Sigma}^{-1}_{\partial STI}\Delta$ and $\ddot{F}^{C}_{CTI} = \ddot{\Sigma}_C \Delta'_T \ddot{\Sigma}^{-1}_{\partial CTI}\Delta_T$ to derive the filter formula. The error process and error covariance formula then follow by standard calculations.

**Recast Approach.** We begin with the definition of the implied cycle model, utilizing both the fitted model for $\partial Y$ and the HP filter, as described in Kaiser and Maravall (2005); also see McElroy (2008b). We begin with a fitted pseudo-spectral density $\dot{f}_Y$ from which we obtain via canonical decomposition the trend component spectrum $\dot{f}_T$. However, we take the viewpoint that this actually corresponds to a trend-cycle, not just a trend, so we will write $\dot{f}_{CT}$. Utilize the following notation: the covariance matrix associated with a spectral density $g$ has $jk$th entry $\Sigma_{jk}(g) = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(\lambda) e^{i\lambda(j-k)}\, d\lambda$, i.e., the autocovariance at lag $j-k$ corresponding to the spectral density $g$. Let the frequency response function of the HP high-pass filter $H(B)$ be denoted $H(e^{-i\lambda})$ (which is given by the formula $(2-2\cos\lambda)^2/\{(2-2\cos\lambda)^2 + q\}$ for signal-to-noise ratio $q$); then we define implied spectra for the cycle and trend as follows:

$$\ddot{f}_T(\lambda) := H(e^{-i\lambda})\dot{f}_{CT}(\lambda) \qquad \ddot{f}_C(\lambda) := \left(1 - H(e^{-i\lambda})\right)\dot{f}_{CT}(\lambda)$$

Since the HP high-pass accomplishes second differencing, we have $1 - H(e^{-i\lambda})$ equal to $|1 - e^{-i\lambda}|^2$ times a bounded function, see McElroy (2008b). This differencing factor cancels with the poles in $\dot{f}_{CT}$, resulting in a bounded function for $\ddot{f}_C$. Trivially, $\ddot{f}_{CT} := \ddot{f}_C + \ddot{f}_T = \dot{f}_{CT}$. With these definitions, we can define $\ddot{F}^C_{CT}$ as follows: $\ddot{\Sigma}_C := \Sigma(\ddot{f}_C)$ and $\ddot{F}^C_{CT} = \ddot{\Sigma}_C \Delta'_T \ddot{\Sigma}^{-1}_{\partial CT} \Delta_T$. Of course, $\ddot{f}_{\partial CT} = \dot{f}_{\partial CT}$, which shows how to compute the matrix $\ddot{\Sigma}_{\partial CT} = \dot{\Sigma}_{\partial CT}$. Next, note that the estimate of the differenced trend-cycle would be $\dot{\Sigma}_{\partial CT} \underline{\Delta}'_S \dot{\Sigma}^{-1}_{\partial Y} \partial Y$, which should correspond to the action of $\Delta_T \dot{F}^{CT}_{CSTI}$ upon $Y$; therefore

$$\ddot{F}^C_{CT} \dot{F}^{CT}_{CSTI} = \ddot{\Sigma}_C \Delta'_T \ddot{\Sigma}^{-1}_{\partial CT} \Delta_T \dot{F}^{CT}_{CSTI} = \ddot{\Sigma}_C \Delta'_T \underline{\Delta}'_S \dot{\Sigma}^{-1}_{\partial Y} \Delta = \ddot{\Sigma}_C \Delta' \dot{\Sigma}^{-1}_{\partial Y} \Delta,$$

using (1.3.1). This demonstrates the stated filter formula, and the error process is $\ddot{\Sigma}_C \Delta' \dot{\Sigma}^{-1}_{\partial Y} \partial Y - C$, from which the stated error covariance formula follows using the fact that the covariance of $\partial Y$ and $C$ is $\Delta \Sigma_C$.

### 1.8.4 Frequency Domain Calculations for Asymptotic MSE

We begin by reviewing elementary material on pseudo-true values. Letting $f_{\partial Y}(\cdot; \psi)$ denote the spectral density of a model for the stationary process $\{\partial Y_t\}$, and $\widetilde{f}_{\partial Y}$ its true spectral density, the Kullback-Leibler discrepancy between these spectra is

$$D(\psi) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f_{\partial Y}(\lambda; \psi)\, d\lambda + \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\widetilde{f}_{\partial Y}(\lambda)}{f_{\partial Y}(\lambda; \psi)}\, d\lambda,$$

expressed as a function of the parameter vector $\psi$ (Taniguchi and Kakizawa, 2000). This is essentially the Whittle likelihood. The minimizers of $D$ are called the pseudo-true values of $\psi$, and are denoted by $\dot{\psi}$ when they are unique. When the model is correctly specified, $\widetilde{f}_{\partial Y} = f_{\partial Y}(\cdot; \widetilde{\psi})$ and the pseudo-true values are equal to the true parameter values; but when the model is misspecified, the pseudo-true values are the limit in probability of the mles. As a numerical exercise, given knowledge of a process and a misspecified model, one can compute $D(\psi)$ and calculate its minimizers. Analytical solutions can be found in the case of an AR($p$), in which case $\dot{\psi}$ solves the Yule-Walker equations (with innovation variance given by concentration), but in general the minimization must be computed numerically, similarly to the mle calculations.

When we compute quantities (such as MSEs) using the pseudo-true values, we are accounting for a type of distortion that arises from model misspecification. In viewing a pseudo-true value as an approximation to an mle, we ignore statistical variation in the mle. We next derive the asymptotic MSE quantities. First, consider fitting a correctly specified cycle-plus-airline model to the DGP. We have the following pseudo-spectra, where $z = e^{-i\lambda}$:

$$
\begin{aligned}
f_C(\lambda) &= \frac{\sigma_C^2}{|1 - 2\rho\cos(\omega)z + \rho^2 z^2|^2}, \\
f_X(\lambda) &= \frac{|(1-\theta_X z)(1-\Theta_X z^{12})|^2 \sigma_X^2}{|(1-z)(1-z^{12})|^2}, \\
f_C(\lambda) + f_X(\lambda) &= \frac{|\Theta_{CX}(z)|^2 \sigma_{CX}^2}{|1 - 2\rho\cos(\omega)z + \rho^2 z^2|^2 |(1-z)(1-z^{12})|^2}.
\end{aligned}
$$

The third spectral density is expressed in reduced form, where $\Theta_{CX}(B)$ has degree 15 and is obtained via spectral factorization. Then the WK MSE is the integral of $f_C f_X/(f_C + f_X)$ evaluated at the true parameter values:

$$
\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{|(1-\theta_X z)(1-\Theta_X z^{12})|^2 \, \sigma_C^2 \, \sigma_X^2}{|\Theta_{CX}(z)|^2 \, \sigma_{CX}^2} \, d\lambda.
$$

In the two-stage method, the final cycle filter is given by the composition of the seasonal adjustment filter and the cycle filter, i.e., $\ddot{f}_C \dot{f}_{TI}/(\ddot{f}_{CTI}\dot{f}_{STI})$, and the MSE simplifies to

$$
\begin{aligned}
&\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{|\dot{\Theta}_{TI}(z)|^4 |\tilde{\Theta}_{STI}(z)|^2 |1-z|^4 |U(z)|^2 \ddot{\sigma}_C^4 \dot{\sigma}_{TI}^4 \tilde{\sigma}_{STI}^2}{|\ddot{\Theta}_{CTI}(z)|^4 |\dot{\Theta}_{STI}(z)|^4 \ddot{\sigma}_{CTI}^4 \dot{\sigma}_{STI}^4} \, d\lambda \\
&+ \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{|\dot{\Theta}_{TI}(z)|^4 |1-z|^8 |U(z)|^4 \ddot{\sigma}_C^4 \dot{\sigma}_{TI}^4 \tilde{\sigma}_C^2}{|\ddot{\Theta}_{CTI}(z)|^4 |\dot{\Theta}_{STI}(z)|^4 |1 - 2\tilde{\rho}\cos(\tilde{\omega})z + \tilde{\rho}^2 z^2|^2 \ddot{\sigma}_{CTI}^4 \dot{\sigma}_{STI}^4} \, d\lambda \\
&- \frac{2}{2\pi}\int_{-\pi}^{\pi} \frac{|\dot{\Theta}_{TI}(z)|^2 |1-z|^4 |U(z)|^2 \ddot{\sigma}_C^2 \dot{\sigma}_{TI}^2 \tilde{\sigma}_C^2}{|\ddot{\Theta}_{CTI}(z)|^2 |\dot{\Theta}_{STI}(z)|^2 |1 - 2\tilde{\rho}\cos(\tilde{\omega})z + \tilde{\rho}^2 z^2|^2 \ddot{\sigma}_{CTI}^2 \dot{\sigma}_{STI}^2} \, d\lambda \\
&+ \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{\tilde{\sigma}_C^2}{|1 - 2\tilde{\rho}\cos(\tilde{\omega})z + \tilde{\rho}^2 z^2|^2} \, d\lambda.
\end{aligned}
$$

Here the second dot denotes pseudo-true parameter values for the cycle-trend-irregular model fitted to the limiting estimated seasonal adjustment process with spectrum $\dot{f}_{TI}^2 \tilde{f}_{CSTI}/\dot{f}_{STI}^2$, which is given by

$$
\frac{\dot{f}_{TI}^2 \tilde{f}_{CSTI}}{\dot{f}_{STI}^2} = \frac{|\dot{\Theta}_{TI}(z)|^4 |\tilde{\Theta}_{CSTI}(z)|^2 |U(z)|^2}{|\dot{\Theta}_{STI}(z)|^4 |1 - 2\tilde{\rho}\cos(\tilde{\omega})z + \tilde{\rho}^2 z^2|^2 |1-z|^4} \, \frac{\dot{\sigma}_{TI}^4 \tilde{\sigma}_{CSTI}^2}{\dot{\sigma}_{STI}^4}.
$$

Next, for the *recast* method we fit an Airline model to the DGP and produce the WK trend estimate – the corresponding pseudo-true values are denoted with a dot. In this case, the cycle estimate is $\widehat{C}_t = (1 - H(B))\Psi(B)Y_t$, where $H(B)$ is the HP filter and $\Psi(B)$ is the WK trend extraction filter. The ARMA form of $1 - H(z)$ is useful for computing MSE: $1 - H(z) = \frac{|1-z|^4}{|\phi(z)|^2} \, \frac{c}{q}$ with $q$ the given signal-to-noise ratio of the HP filter, and $c$ and $\phi(B) = 1 + \phi_1 B + \phi_2 B^2$ given by (2.10) through (2.12) of McElroy (2008b); also see (3.2). The error process can then be written as

$$\begin{aligned}\widehat{C}_t - C_t &= \frac{c\dot{\sigma}_T^2}{q\dot{\sigma}_X^2} \, \frac{\dot{\Theta}_T(B)\dot{\Theta}_T(F)(1-F)^2 U(F)}{\phi(B)\phi(F)\dot{\Theta}_X(B)\dot{\Theta}_X(F)} \, \partial X_t \\ &+ \left( \frac{c\dot{\sigma}_T^2}{q\dot{\sigma}_X^2} \, \frac{\dot{\Theta}_T(B)\dot{\Theta}_T(F)(1-F)^2(1-B)^2 U(F)U(B)}{\phi(B)\phi(F)\dot{\Theta}_X(B)\dot{\Theta}_X(F)} - 1 \right) C_t.\end{aligned}$$

Hence (with the tilde on parameters denoting the true values) the Recast MSE can be written as

$$\begin{aligned}&\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{|\dot{\Theta}_T(z)|^4 |1-z|^4 |U(z)|^2 |\tilde{\Theta}_X(z)|^2}{|\phi(z)|^4 |\dot{\Theta}_X(z)|^4} \, \frac{c^2\dot{\sigma}_T^4}{q^2\dot{\sigma}_X^4} \, \tilde{\sigma}_X^2 \, d\lambda \\ &+ \frac{1}{2\pi}\int_{-\pi}^{\pi} \left( \frac{c\dot{\sigma}_T^2}{q\dot{\sigma}_X^2} \, \frac{|\dot{\Theta}_T(z)|^2 |1-z|^4 |U(z)|^2}{|\phi(z)|^2 |\dot{\Theta}_X(z)|^2} - 1 \right)^2 \frac{\tilde{\sigma}_C^2}{|1 - 2\tilde{\rho}\cos(\tilde{\omega})z + \tilde{\rho}^2 z^2|^2} \, d\lambda.\end{aligned}$$

# Bibliography


Baxter, M. and King, R. (1999). Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series. *The Review of Economics and Statistics*, 81:575–593.

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

Bell, W. and Hillmer, S. (1991). Initializing the Kalman Filter for Nonstationary Time Series Models (Corr: V13 P281-282). *Journal of Time Series Analysis*, 12:283–300.

Bell, W. and Hilmer, S. (1988). A Matrix Approach to Likelihood Evaluation and Signal Extraction for ARIMA Component Time Series Models. Technical report, U.S. Census Bureau.

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

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

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

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

Harvey, A. and Proietti, T. (2005). *Readings in Unobserved Component Models.* Oxford University Press.

Harvey, A. C. (1989). *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.

Hodrick, R. and Prescott, E. (1997). Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit, and Banking*, 29:1–16.

Kaiser, R. and Maravall, A. (2001). *Measuring Business Cycles in Economic Time Series*. Springer-Verlag Inc.

Kaiser, R. and Maravall, A. (2005). Combining Filter Design with Model-based Filtering (with an Application to Business-cycle Estimation) (Pkg: P691-715). *International Journal of Forecasting*, 21(4):691–710.

Maravall, A. and Caporello, G. (2004). Program TSW: Revised Reference Manual. Technical report, Research Department, Bank of Spain.

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

McElroy, T. (2008b). NOTES: Exact formulas for the Hodrick-Prescott filter. *The Econometrics Journal Online*, 11(1):209–217.

McElroy, T. and Sutcliffe, A. (2006). An Iterated Parametric Approach to Nonstationary Signal Extraction. *Computational Statistics and Data Analysis*, 50:2206–2231.

Peña, D., Tiao, G. C., and Tsay, R. S. (2001). *A Course in Time Series Analysis*. John Wiley & Sons.

Roberts, C., Holan, S., and Monsell, B. (2010). Comparison of X-12 ARIMA Trading Day Regressors to Country Specific Regressors. *Journal of Official Statistics*, 26:371–394.

Soukup, R. and Findley, D. (2000). Detection and Modeling of Trading Day Effects. ICES Proceedings.

Stock, J. and Watson, M. (2004). Business Cycle Fluctuations in U.S. Macroeconomic Time Series. In Taylor, J. and Woodford, M., editors, *Handbook of Macroeconomics*. Elsevier.

Taniguchi, M. and Kakizawa, Y. (2000). *Asymptotic Theory of Statistical Inference for Time Series*. Springer-Verlag.

| Estimation: $\theta_X = .6$; $\Theta_X = .6$; $\omega = \pi/12$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| DGP | Airline | | | Two-Stage | | | | |
| $\rho = .7$ | $\theta_X$ | $\Theta_X$ | $\sigma^2_X$ | $\rho$ | $\omega$ | $\sigma^2_C$ | $\sigma^2_T$ | $\sigma^2_I$ |
| $\kappa = 1$ | -0.017 | 0.731 | 3.009 | 0.674 | 0.084 | 0.900 | 0.0002 | 0.373 |
| $\kappa = .25$ | 0.253 | 0.674 | 1.680 | 0.686 | 0.000 | 0.230 | 0.0003 | 0.361 |
| $\kappa = .1$ | 0.408 | 0.646 | 1.321 | 0.247 | 0.000 | 0.636 | 0.0022 | 0.000 |
| $\kappa = 0$ | 0.600 | 0.600 | 1.000 | 0.832 | 0.000 | 0.011 | 0.0003 | 0.343 |
| $\rho = .8$ | $\theta_X$ | $\Theta_X$ | $\sigma^2_X$ | $\rho$ | $\omega$ | $\sigma^2_C$ | $\sigma^2_T$ | $\sigma^2_I$ |
| $\kappa = 1$ | -0.134 | 0.736 | 3.379 | 0.799 | 0.221 | 0.850 | 0.0002 | 0.394 |
| $\kappa = .25$ | 0.132 | 0.683 | 1.868 | 0.791 | 0.190 | 0.225 | 0.0002 | 0.366 |
| $\kappa = .1$ | 0.302 | 0.657 | 1.440 | 0.781 | 0.147 | 0.104 | 0.0002 | 0.356 |
| $\kappa = 0$ | 0.600 | 0.600 | 1.000 | 0.832 | 0.000 | 0.011 | 0.0003 | 0.343 |
| $\rho = .9$ | $\Theta_X$ | $\theta_X$ | $\sigma^2_X$ | $\rho$ | $\omega$ | $\sigma^2_C$ | $\sigma^2_T$ | $\sigma^2_I$ |
| $\kappa = 1$ | -0.287 | 0.763 | 4.402 | 0.907 | 0.251 | 0.851 | 0.0002 | 0.417 |
| $\kappa = .25$ | -0.035 | 0.710 | 2.316 | 0.906 | 0.246 | 0.214 | 0.0002 | 0.386 |
| $\kappa = .1$ | 0.140 | 0.681 | 1.713 | 0.903 | 0.241 | 0.091 | 0.0003 | 0.374 |
| $\kappa = 0$ | 0.600 | 0.600 | 1.000 | 0.832 | 0.000 | 0.011 | 0.0003 | 0.343 |

Table 1.1: Pseudo-true values for the bi-infinite sample under the three approaches discussed in Section 1.5. The left column gives the DGP parameter values of $\rho$ and $\kappa$. Other columns give pseudo-true values for various parameters, for the fitted airline model and the fitted cycle-trend-irregular model of the two-stage approach. For the direct approach, pseudo-true values equal the DGP values.

| MSEs: $\theta_X = .6$; $\Theta_X = .6$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\rho = .7$ | | | $\rho = .8$ | | | $\rho = .9$ | | |
| $\omega = \pi/12$ | Direct | Two-Stage | Recast | Direct | Two-Stage | Recast | Direct | Two-Stage | Recast |
| $\kappa = 1$ | 1.660 | 1.846 | 1.775 | 2.426 | 2.639 | 2.672 | 3.306 | 3.492 | 3.940 |
| $\kappa = .25$ | 0.743 | 0.879 | 0.942 | 1.079 | 1.220 | 1.183 | 1.453 | 1.540 | 1.526 |
| $\kappa = .1$ | 0.418 | 0.705 | 0.725 | 0.627 | 0.785 | 0.838 | 0.861 | 0.940 | 0.996 |
| $\kappa = 0$ | 0.000 | 0.297 | 0.502 | 0.000 | 0.297 | 0.502 | 0.000 | 0.297 | 0.502 |
| $\omega = \pi/60$ | Direct | Two-Stage | Recast | Direct | Two-Stage | Recast | Direct | Two-Stage | Recast |
| $\kappa = 1$ | 2.449 | 2.587 | 2.642 | 6.055 | 6.211 | 8.641 | 31.495 | 31.765 | 84.301 |
| $\kappa = .25$ | 1.061 | 1.177 | 1.156 | 2.471 | 2.598 | 2.670 | 11.801 | 11.838 | 21.608 |
| $\kappa = .1$ | 0.593 | 0.737 | 0.812 | 1.378 | 1.526 | 1.431 | 6.261 | 6.323 | 9.026 |
| $\kappa = 0$ | 0.000 | 0.297 | 0.502 | 0.000 | 0.297 | 0.502 | 0.000 | 0.297 | 0.502 |

Table 1.2: MSE values associated with the psuedo-true values for a bi-infinite sample under the three approaches discussed in Section 1.5.

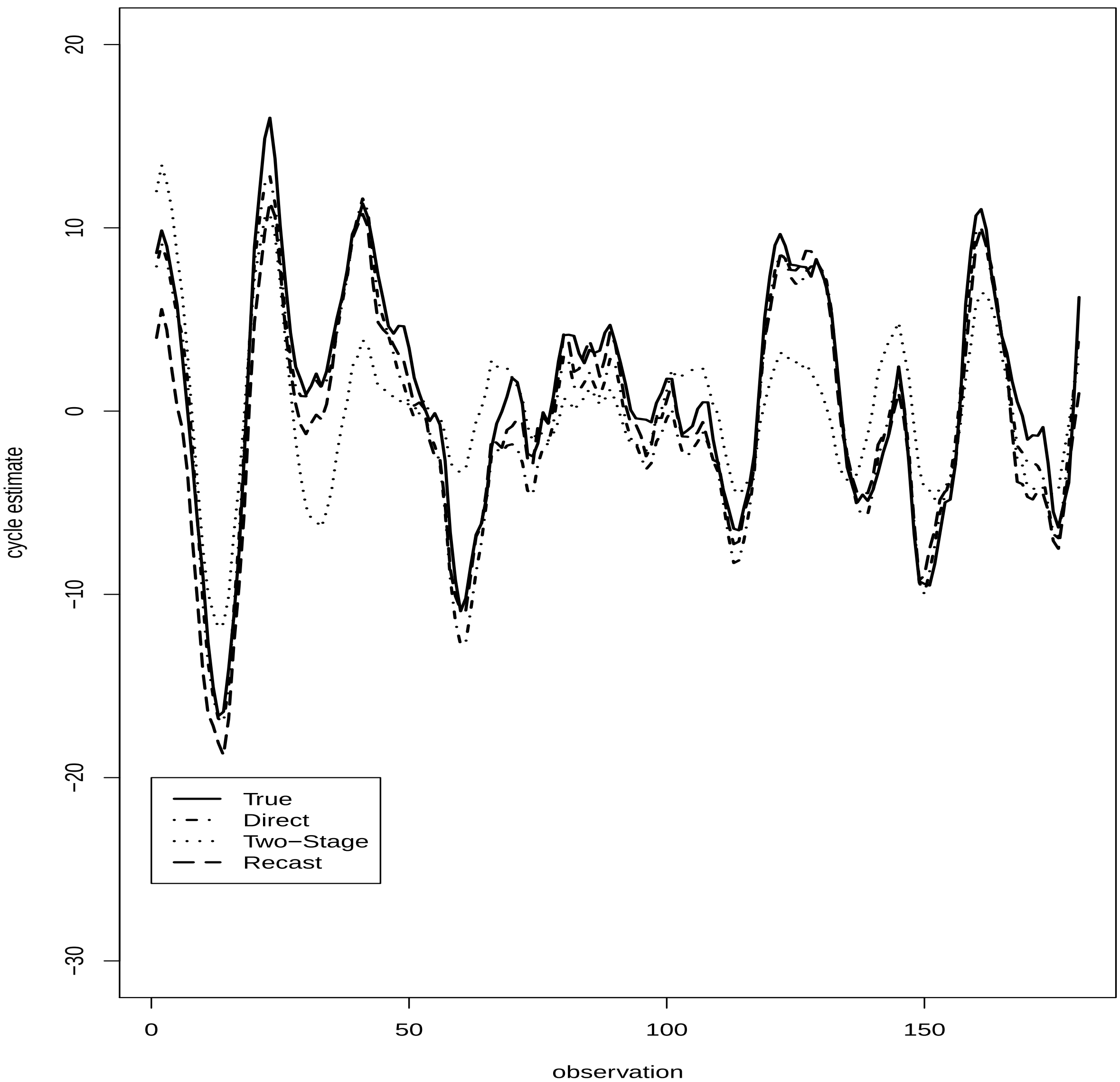


Figure 1.1: Representative cycle estimate in the finite sample simulation discussed in Section 1.5. Note that all three approaches are considered.

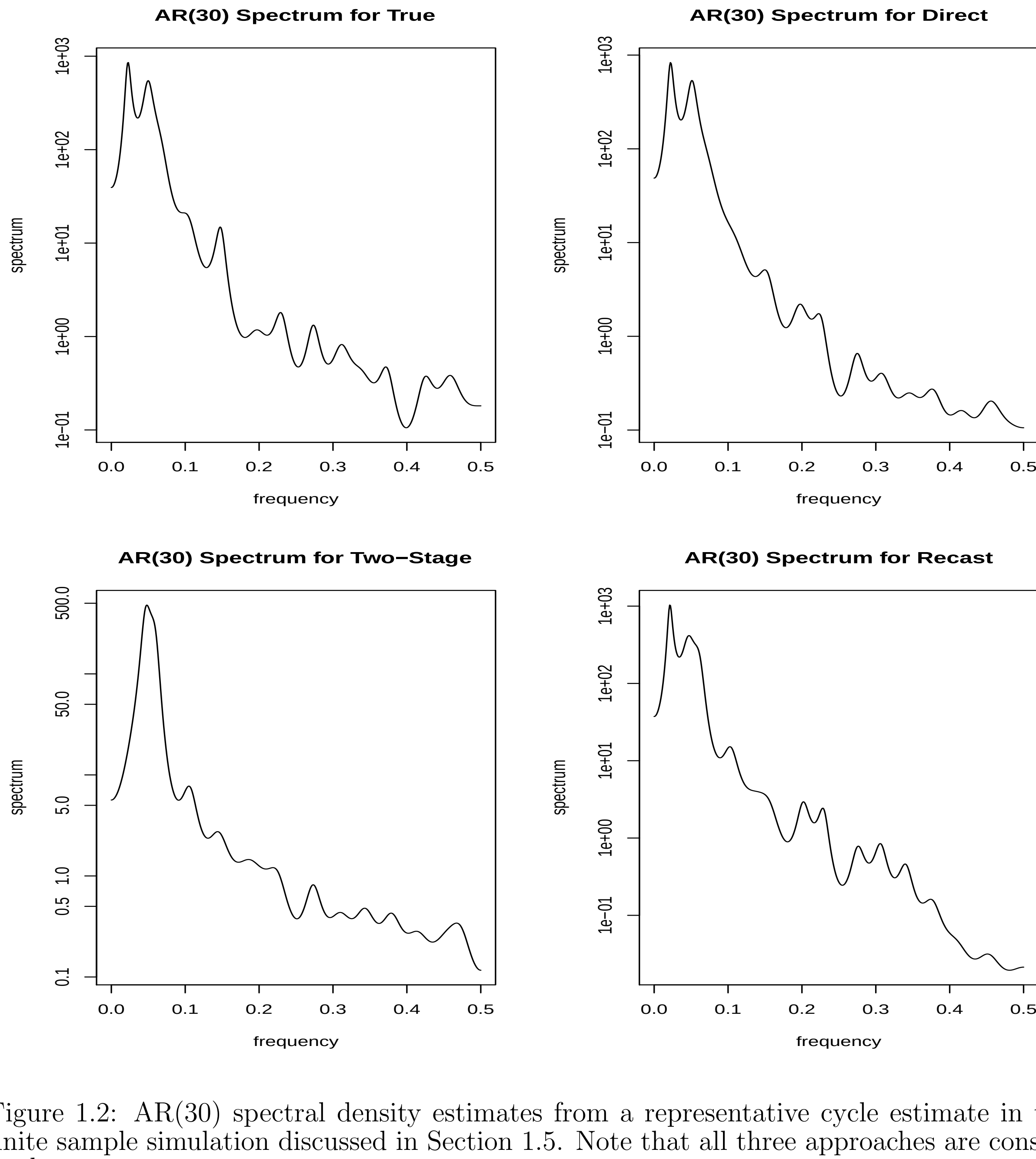


Figure 1.2: AR(30) spectral density estimates from a representative cycle estimate in the finite sample simulation discussed in Section 1.5. Note that all three approaches are considered.

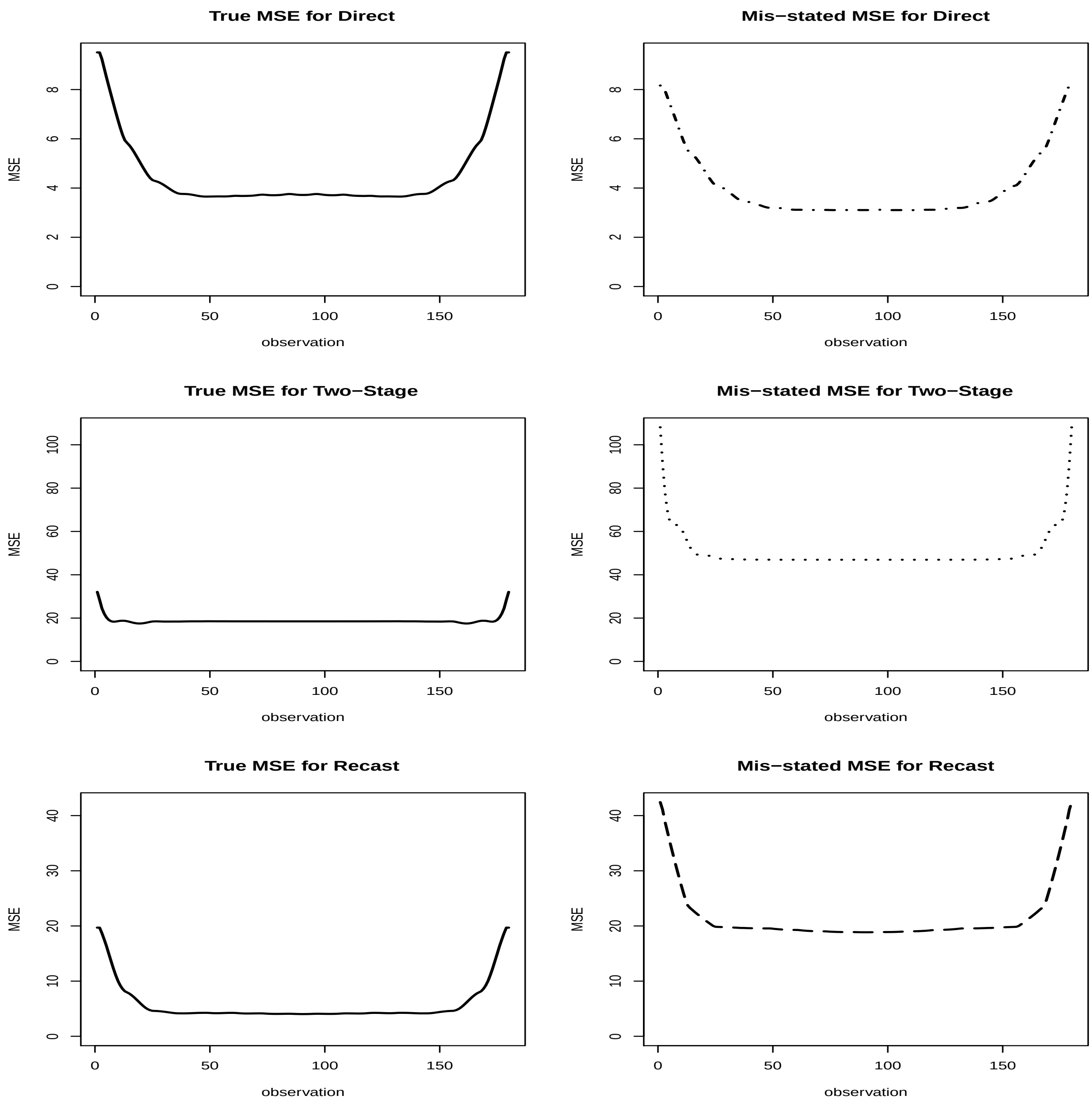


Figure 1.3: True and mis-stated MSEs for a representative cycle estimate in the finite sample simulation discussed in Section 1.5. Note that all three approaches are considered. Also the scales on the y-axes differ according to the method considered.

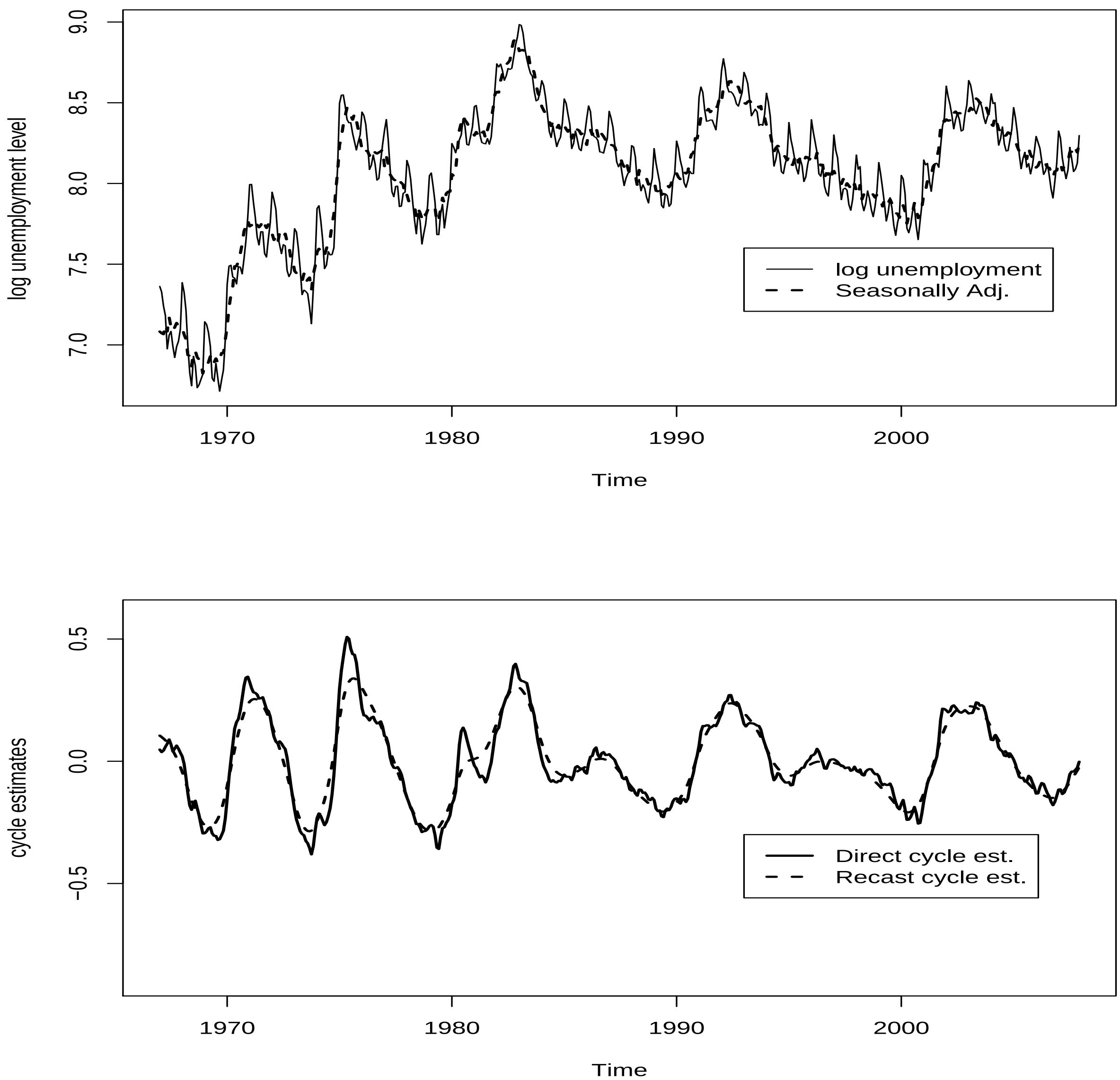


Figure 1.4: The top panel plots the LBU series in logs with its seasonal adjustment arising from a fitted airline model. The bottom panel provides cycle estimates from the *direct* and *recast* methods.