Volume Title: Seasonal Analysis of Economic Time Series

Volume Author/Editor: Arnold Zellner, ed.

Volume Publisher: NBER

Volume URL: http://www.nber.org/books/zell79-1

Publication Date: 1979

Chapter Title: Some Consequences of Temporal Aggregation in Seasonal Time Series Models

Chapter Author: William W. S. Wei

Chapter URL: http://www.nber.org/chapters/c3908

Chapter pages in book: (p. 433 - 448)

# SOME CONSEQUENCES OF TEMPORAL AGGREGATION IN SEASONAL TIME SERIES MODELS

William W. S. Wei
Temple University

## ABSTRACT

Given a basic stochastic seasonal time series model, developed by Box and Jenkins, the corresponding model for temporal aggregates is derived. Insofar as forecasting future aggregates is concerned, the loss in information due to aggregation is substantial if the nonseasonal component of the model is nonstationary. This is not so serious for long-term forecasting, especially when the nonseasonal component is stationary. In forecasting future aggregates, there is no loss in information if the basic series is a purely seasonal model. In terms of parameter estimation, aggregation causes a tremendous loss in efficiency, regardless of the given model. The results are shown by both theory and actual data.

## INTRODUCTION

Temporal aggregation poses an important problem in time series analysis. It is so because, in working with data, one must decide on the time unit he is going to use for his basic observations. If the model for a phenomenon under investigation is regarded appropriate in terms of a small basic time unit (e.g., a month), then proper inferences about the underlying basic model should be drawn from the analysis of data in terms of this basic time unit. Improper use of data in some larger time scale (e.g., a quarter or 1 year) to make inferences could be very misleading and, hence, seriously bias the views of policy-makers unless the effects of aggregation are accurately examined and, hence, properly taken into account.

The temporal aggregation problem was first studied in the field of econometrics in the context of some simple distributed lag or regression models, e.g., Theil [8], Mundlak [4], Zellner and Montmarquette [11], Sims [6], Tiao and Wei [10], and others. In the analysis of a univariate time series, the problem was investigated by Quenouille [5], Amemiya and Wu [1], Brewer [3], Telser [7], Tiao [9], and others. However, all previous work has been restricted to the case where the underlying models are nonseasonal. The important problem of aggregation's effect on seasonal models is still relatively unexplored.

In this paper, we study some consequences of temporal aggregation in discrete stochastic seasonal time series models developed by Box and Jenkins [2, ch. 9]. In the next section, we solve a fundamental problem in temporal aggregation. That is, for a given seasonal time series model in terms of a basic time unit, we derive the corresponding model for temporal aggregates and also discuss the relationship and properties of the models. Since one of the principal purposes of time series analysis is to forecast, the third section will study the effect of aggregation on forecasting. In the fourth section, the loss of information due to aggregation in terms of parameter estimation is examined. In the fifth section, the results are illustrated with an actual example of the U.S. employment data. Finally, in the last section, a summary of findings and some concluding remarks are given.

## MODEL STRUCTURE OF TEMPORAL AGGREGATES

### The Basic Model

Assume that the basic series $z_t$ follows a general multiplicative seasonal model with period $s$, e.g.,

$$\alpha_P(B^s)\phi_p(B)(1-B^s)^D(1-B)^d z_t=\theta_q(B)\beta_Q(B^s)a_t \quad (1)$$

where the $a_t$'s are independently and identically distributed as $N(O, \sigma_a^2)$, $B$ is the backshift operator, such that $Bz_t=z_{t-1}$, $\alpha_P(B^s)$ and $\beta_Q(B^s)$ are polynominals in $B^s$ of degrees $P$ and $Q$ and $\phi_p(B)$ and $\theta_q(B)$ are polynomials in $B$ of degrees $p$ and $q$, respectively. We also assume that all of these polynomials satisfy stationarity and invertibility conditions. That is, all the polynomials have their roots lying outside the unit circle. This model, which has been developed by Box and Jenkins [2, ch. 9] and called a model of order $(p, d, q) \times (P, D, Q)_s$, provides a useful representation for a variety of seasonal time series.

In many economic and business time series, $t$=month and $s=12$. The model then implies that the time series $z_t$ is the product of two correlated random components. The seasonal component

$$\alpha_P(B^s)\,(1-B^s)^D z_t=\beta_Q(B^s)b_t \qquad (2)$$

describes the relationship from year to year for a certain month, and the nonseasonal component

$$\phi_p(B)(1-B)^d b_t=\theta_q(B)a_t \qquad (3)$$

relates the remaining stochastic factors from month to month for all years.

Under the normality assumption, any stationary and invertible stochastic process $\{u_t\}$ is uniquely characterized by its autocovariance generating function $\gamma_u(B)$, defined by

$$\gamma_u(B)=\sum_{j=-\infty}^{\infty}\gamma_u(j)B^j \qquad (4)$$

where

$$\gamma_u(j)=E\,(u_t-E\,(u_t))\,(u_{t-j}-E\,(u_t)).$$

Let

$$w_t=(1-B^s)^D(1-B)^d z_t=[\beta(B^s)\theta(B)]/[\alpha(B^s)\phi(B)]a_t$$

It can be easily seen that the autocovariance generating function of $\{w_t\}$ is given by

$$\gamma_w(B)=\sigma_a^2G\,(B^s)g\,(B) \qquad (5)$$

where

$$G\,(B^s)=\sum_{j=-\infty}^{\infty}G_{js}B^{js}=\beta(B^s)\beta(F^s)/\alpha(B^s)\alpha(F^s)$$

$$g\,(B)=\sum_{i=-\infty}^{\infty}g_iB^i=\theta\,(B)\theta\,(F)/\phi\,(B)\phi\,(F),\text{ and }F=B^{-1}$$

More explicitly,

$$\gamma_w(B)=\sigma_a^2\sum_{\ell=-\infty}^{\infty}\gamma_\ell B^\ell$$

where

$$\gamma_\ell=\sum_{j=-\infty}^{\infty}G_{js}\,g_{\ell-js} \qquad (6)$$

## Temporal Aggregate and Its Model Structure

Let the $m$-component nonoverlapping sums

$$Z_T=\left(\sum_{j=0}^{m-1}B^j\right)z_{mT} \qquad (7)$$

be the desired temporal aggregates, where $T$ is the aggregate time unit. For example, if $t$ is a month, and $m$ equals 3, then $T$ is a quarter. Define $X_t=(\sum_{j=0}^{m-1}B^j)z_t$ and note that $Z_T=X_{mT}$. For practical purposes, we assume that the number of aggregation components $m$ is chosen such that $s=mS$ for some integer $S$, which is usually the case in economic and business applications.

Let

$$W_t=\left(\sum_{j=0}^{m-1}B^j\right)^{d+1}w_t=(1-B^{mS})^D(1-B^m)^dX_t$$

$$=\left(\frac{1-B^m}{1-B}\right)^{d+1}\frac{\beta(B^s)\theta(B)}{\alpha(B^s)\phi(B)}a_t \qquad (8)$$

The autocovariance generating function $\gamma_W(B)$ is given by

$$\gamma_W(B)=\sigma_a^2G\,(B^s)\left(\frac{1-B^m}{1-B}\right)^{d+1}\left(\frac{1-F^m}{1-F}\right)^{d+1}g\,(B) \qquad (9)$$

Letting $t=mT$, $\mathscr{B}=B^m$, and $V_T=W_{mT}$, we can write $V_T=(1-\mathscr{B}^S)^D(1-\mathscr{B})^dZ_T$ with the autocovariance generating function being given by

$$\gamma_V(\mathscr{B})=\sum_{K=-\infty}^{\infty}\gamma_V(K)\mathscr{B}^K \qquad (10)$$

where $\gamma_V(K)$ is the coefficient of $B^{-Km}$ in $\gamma_W(B)$ in (9) and equals

$$\gamma_V(K)=E\,(V_TV_{T-K})$$

$$=\sigma_a^2\sum_{j=-\infty}^{\infty}G_{js}\sum_{\ell=-(d+1)(m-1)}^{(d+1)(m-1)}g_{Km+\ell-js}\sum_{i=0}^{2d+2}\binom{2d+2}{i}(-1)^i\binom{\ell+(d+1-i)m+d}{2d+1} \qquad (11)$$

## Remarks on Some Special Cases

1. If $P=D=Q=0$, the model (1) becomes a nonseasonal model, (5) becomes $\gamma_w(B)=\sigma_a^2g\,(B)$ and (11) reduces to

$$\gamma_V(K)=\sigma_a^2\sum_{\ell=-(d+1)(m-1)}^{(d+1)(m-1)}g_{Km+\ell}\sum_{i=0}^{2d+2}\binom{2d+2}{i}(-1)^i\binom{\ell+(d+1-i)m+d}{2d+1} \qquad (12)$$

On the other hand, letting $\phi_p(B)=\prod_{j=1}^{p}(1-\delta_jB)$ and multiplying $\prod_{j=1}^{p}[(1-\delta_j^mB^m)(1-B^m)^{d+1}/(1-\delta_jB)(1-B)^{d+1}]$ on both sides of (1), we have

$$Y_t\equiv\prod_{j=1}^{p}(1-\delta_j^mB^m)(1-B^m)^dX_t$$

$$=\prod_{j=1}^{p}\left(\frac{1-\delta_j^mB^m}{1-\delta_jB}\right)\left(\frac{1-B^m}{1-B}\right)^{d+1}\theta_q(B)a_t \qquad (13)$$

and $E\,(Y_{mT}Y_{mT-mK})=0$ for $K>r$ where $r=\left[p+d+1+\dfrac{q-p-d-1}{m}\right]$ and $[x]$ denotes the integer part of $x$. The aggregate $Z_T=X_{mT}$, thus, follows an ARIMA $(p, d, r)$ process with the autocovariance structure given by (12). This is a generalization of the result given in Amemiya and Wu [1], Telser [7], Brewer [3], and Tiao [9], who studied the consequences of aggregation on a stationary AR $(p)$, ARMA $(p, q)$, and a nonstationary MA $(q)$ process, respectively.

2. If $p=d=q=0$, model (1) becomes a purely seasonal model of period $s$, (5) becomes $\gamma_w(B)=\sigma_a^2 G(B^s)$, and (11) reduces to

$$\gamma_V(K)=\begin{cases} m\sigma_a^2 G_{js} & \text{if } K=jS \text{ for some } j \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

Thus, the aggregates follow a purely seasonal model of period $S$. Moreover, (14) implies that aggregation does not change the model form. In fact, the order and parameters of the model remain exactly the same as the basic model, except that the variance of the noise term in the aggregate model inflates $m$ times as expected. Note that if $S=1$, i.e., $m=s$, (14) becomes $\gamma_V(K)=m\sigma_a^2 G_{Ks}$ and, thus, aggregation reduces a purely seasonal model to a nonseasonal model. In other words, the seasonal effect in this case has been smoothed out.

To further characterize our aggregate model, we quote the following lemma, which was proved by Tiao and Wei [10]:

Let $\{z_t\}$ be a stationary and invertible process and $\{Z_T\}$ be the series of temporal aggregates defined in (7). Then, $\{Z_T\}$ is also a stationary and invertible process.

Summarizing these results, we obtain theorem 1:

Suppose the basic series $z_t$ follows a process of order $(p, d, q) \times (P, D, Q)_s$ given in (1) and $Z_T$ is the aggregate series defined in (7). Then $Z_T$ is a process of order $(p, d, r) \times (P, D, Q)_S$ given by

$$\alpha_P(\mathscr{B}^S)\lambda_p(\mathscr{B})(1-\mathscr{B}^S)^D(1-\mathscr{B})^d Z_T=\nu_r(\mathscr{B})\beta_Q(\mathscr{B}^S)C_T \quad (15)$$

where

$$r=\left[p+d+1+\frac{q-p-d-1}{m}\right]$$

$C_T$'s are independently and identically distributed as $N(0,\sigma_C^2)$, $\alpha_P(\mathscr{B}^S)$, and $\beta_Q(\mathscr{B}^S)$ are polynomials in $\mathscr{B}^S$ of degrees $P$ and $Q$, respectively, and $\lambda_p(\mathscr{B})$ and $\nu_r(\mathscr{B})$ are polynomials in $\mathscr{B}$ of degrees $p$ and $r$, respectively. The $\sigma_C^2$, $\alpha_P(\mathscr{B}^S)$, $\lambda_p(\mathscr{B})$, $\nu_r(\mathscr{B})$, and $\beta_Q(\mathscr{B}^S)$ are obtained by solving the equations induced from equating the coefficient of $\mathscr{B}$ in the following relationship, such that stationarity and invertibility conditions are satisfied

$$\gamma_V(\mathscr{B})=\sigma_C^2\frac{\beta_Q(\mathscr{B}^S)\nu_r(\mathscr{B})}{\alpha_p(\mathscr{B}^S)\lambda_p(\mathscr{B})}\frac{\beta_Q(F^S)\nu_r(F)}{\alpha_P(F^S)\lambda_p(F)}$$

$$=\sum_{K=-\infty}^{\infty}\gamma_V(K)\mathscr{B}^K \quad (16)$$

where $\gamma_V(K)$ was given in (11) and $F=\mathscr{B}^{-1}$.

It should be noted that the aggregate model in (15) is derived under the practical consideration that the number of aggregation components $m\le s$. It is readily seen from comparing (11) and (12) that in this case aggregation contaminates the model structure through its nonseasonal component in (3). If the number of aggregation components $m\ge s$, it was pointed out in remark 2 that the seasonal model reduces to a regular nonseasonal autoregressive integrated moving average (ARIMA) process. Furthermore, it can be easily seen from (13) that, for a given nonseasonal ARIMA $(p, d, q)$ process, as $m$ becomes large, autoregressive parameters tend to zero, and, hence, the aggregate model reduces to an integrated moving average (IMA) $(d, d)$ process, which was also shown in Tiao [9]. Thus, given a general model in (1), if $m$ becomes a large multiple of $s$, which is usually the case in economic and business applications, then the limiting aggregate model (15) becomes an IMA $(D+d, D+d)$ process. In particular, if $D=d=0$, the limiting aggregate model further reduces to a process of white noise. Thus, temporal aggregation will, in general, complicate the model structure. However, as the number of aggregation components $m$ becomes large, it tends to simplify the model form. This may give an explanation why the modeling of time aggregates is sometimes much more involved and sometimes relatively simpler than the modeling of basic disaggregated series.

## EFFECT OF AGGREGATION ON FORECASTING

One serious information loss due to aggregation in forecasting is obvious. If the basic data are available, we can use the basic model to forecast any future aggregates. However, if only aggregates are available, we cannot use the aggregate model to predict desired future disaggregates.

Now, suppose we are only interested in forecasting a future aggregate $Z_{T+\ell}$ at time $T$. We may construct it either from basic data $z_t$ or from aggregates $Z_T$. Employing a general result in Box and Jenkins [2, 128], the optimal forecast of $Z_{T+\ell}$, given its past history, is the conditional expectation $E(Z_{T+\ell}|\text{past history})$.

Now, the model (1) for the basic series $\{z_t\}$ is invertible. It can be written as

$$z_t=\sum_{j=1}^{\infty}\pi_j z_{t-j}+a_t \quad (17)$$

where the $\pi_j$'s are a convergent series obtained by equating coefficients in

$$\theta_q(B)\beta_Q(B^s)(1-\pi_1 B-\pi_2 B^2 \ldots)$$
$$=\alpha_P(B^s)\phi_p(B)(1-B^s)^D(1-B)^d \quad (18)$$

Thus, it is really shown that

$$E(z_{t+\ell}|z_t, z_{t-1}, \ldots)=\sum_{j=1}^{\infty}\pi_j^{(\ell)}z_{t-j+1} \quad (19)$$

where

$$\pi_j^{(\ell)}=\pi_{j+\ell-1}+\sum_{h=1}^{\ell-1}\pi_h\pi_j^{(\ell-h)}$$

and

$$\pi_j^{(1)}=\pi_j \text{ for } j=1,\ 2,\ \ldots$$

Given that the basic data $z_{mT}$, $z_{mT-1}$, ... are available, the optimal forecast $\hat{Z}_T(\ell)$ of $Z_{T+\ell}$ at time $T$ is, hence, given by the following convergent series of all available observations

$$\hat{Z}_T(\ell)=\sum_{j=1}^{m\ell}E(z_{mT}(j)|z_{mT},\ z_{mT-1},\ \ldots)=\sum_{j=1}^{\infty}\omega_j z_{mT-j+1} \quad (20)$$

where

$$\omega_j=\sum_{i=1}^{\ell m}\pi_j^{(i)}$$

Proceeding in the same way, the optimal forecast $\hat{\hat{Z}}_T(\ell)$ of $Z_{T+\ell}$ at time $T$, given the aggregates $Z_T$, $Z_{T-1}$, ..., can be written as a convergent series of all available aggregates

$$\hat{\hat{Z}}_T(\ell)=\sum_{j=1}^{\infty}\Omega_j Z_{T-j+1} \quad (21)$$

Given $z_{mT}$, $z_{mT-1}$, ..., since $\hat{Z}_T(\ell)$ is the optimal forecast of $Z_{T+\ell}$, we have

$$\begin{aligned} E(Z_{T+\ell}-\hat{Z}_T(\ell))^2 &= E(Z_{T+\ell}-\sum_{j=1}^{\infty}\omega_j z_{mT-j+1})^2 \\ &\le E(Z_{T+\ell}-\sum_{j=1}^{\infty}R_j z_{mT-j+1})^2 \\ &= E(Z_{T+\ell}-\hat{\hat{Z}}_T(\ell))^2 \end{aligned} \quad (22)$$

where

$$R_j=\begin{cases}\Omega_1 & \text{if } j=1,\ 2,\ \ldots\ m \\ \Omega_2 & \text{if } j=m+1,\ \ldots,\ 2m \\ \text{etc.} & \end{cases}$$

Thus, the basic model also gives a better precision in forecasting the future aggregates than the aggregate model does.

More explicitly, the variance of the forecast error, based on the aggregate model is

$$\text{Var}(Z_{T+\ell}-\hat{\hat{Z}}_T(\ell))=\sigma_{\mathcal{E}}^2\sum_{j=0}^{\ell-1}\Psi_j^2 \quad (23)$$

where $\Psi_0=1$ and $\Psi$'s are obtained from the relationship

$$\alpha(\mathcal{B}^S)\lambda(\mathcal{B})(1-\mathcal{B}^S)^D(1-\mathcal{B})^d\left(\sum_{j=0}^{\infty}\Psi_j\mathcal{B}^j\right)=\nu(\mathcal{B})\beta(\mathcal{B}^S)$$

On the other hand, model (1) can be written as $z_t=\left(\sum_{j=0}^{\infty}\psi_j B^j\right)a_t$, where $\psi_0=1$ and the $\psi$'s are obtained from the relationship

$$\alpha(B^s)\phi(B)(1-B^s)^D(1-B)^d\left(\sum_{j=0}^{\infty}\psi_j B^j\right)=\theta(B)\beta(B^s)$$

Hence,

$$X_t=\left(\sum_{j=0}^{m-1}B^j\right)\left(\sum_{i=0}^{\infty}\psi_i B^i\right)a_t=\left(\sum_{j=0}^{\infty}\Phi_j B^j\right)a_t$$

where

$$\Phi_j=\sum_{i=0}^{m-1}\psi_{j-i}$$

The variance of the forecast error, based on the basic model, is, hence,

$$\begin{aligned}&\text{Var}(Z_{T+\ell}-\hat{Z}_T(\ell)) \\ &=\text{Var}(X_{mT+m\ell}-\hat{X}_{mT}(m\ell))^2=\sigma_a^2\sum_{j=0}^{\ell m-1}\Phi_j^2\end{aligned} \quad (24)$$

The efficiency of forecasting future aggregates using the aggregate model can be measured by the variance ratio

$$\zeta(m,\ \ell)=\frac{\text{Var}(Z_{T+\ell}-\hat{Z}_T(\ell))}{\text{Var}(Z_{T+\ell}-\hat{\hat{Z}}_T(\ell))} \quad (25)$$

Since $\hat{Z}_T(\ell)$ and $\hat{\hat{Z}}_T(\ell)$ are obviously unbiased forecasts for $Z_{T+\ell}$, (22) implies that $0\le\zeta(m,\ \ell)\le 1$.

## Some Remarks

1. For a general model (1), it has been pointed out in the section on model structure of temporal aggregates that when the number of aggregation components $m$ becomes large, the limiting aggregate model tends to a process of white noise or an IMA process, depending on whether the basic series is stationary or nonstationary. In the first case, when the basic series is stationary, Amemiya and Wu [1] and Tiao [9] showed that the limiting efficiency $\zeta(\ell)\equiv\lim_{m\to\infty}\zeta(m,\ \ell)=1$ for all $\ell$. In the second case, when the basic series is nonstationary, Tiao [9] showed that the limiting efficiency $\zeta(\ell)$ is a small number, much less than 1. It approaches 1 only when $\ell\to\infty$.
2. When $p=d=q=0$, the basic model (1) becomes a purely seasonal model

$$\alpha_P(B^s)(1-B^s)^D z_t=\beta_Q(B^s)a_t \quad (26)$$

As shown in the section on model structure of temporal aggregates, for the number of aggregation components $m$ be such that $s=mS$ for some integer $S$, which is usually so for economic and business applications, the aggregate model becomes

$$\alpha_P(\mathscr{B}^S)(1-\mathscr{B}^S)^D Z_T=\beta_Q(\mathscr{B}^S)C_T \qquad (27)$$

where $C_T$'s are i.i.d. $N(0,\ m\sigma_a^2)$. By applying (23) and (24), it is readily shown that in this case $\zeta(m,\ \ell)=1$ for all $m$ and $\ell$.

Summarizing the above results, we have theorem 2:

Suppose the basic series $z_t$ follows a model in (1) and let $Z_T$ be the temporal aggregate defined in (7). Let $\hat{Z}_T(\ell)$ and $\hat{\hat{Z}}_T(\ell)$ be the optimal forecasts of $Z_{T+\ell}$ at time $T$, based on the past history of the basic series $z_t$ and aggregate series $Z_T$, respectively. Define $\zeta(m,\ \ell)=\mathrm{Var}\,(Z_{T+\ell}-\hat{Z}_T(\ell))/\mathrm{Var}(Z_{T+\ell}-\hat{\hat{Z}}_T(\ell))$. Then—

1. $0\le\zeta(m,\ \ell)\le 1$ for all $m$ and $\ell$.
2. $\zeta(\ell)\equiv\lim_{m\to\infty}\zeta(m,\ \ell)=1$ for all $\ell$, if $d=0$.
3. $\zeta(\ell)<<1$ and $\zeta(\ell)\to 1$ only when $\ell\to\infty$ if $d>0$.
4. $\zeta(m,\ \ell)=1$ for all $m$ and $\ell$ if $p=d=q=0$ and $s=mS$ for some integer $S$.

In other words, what the theorem says is that, insofar as forecasting the future aggregates is concerned, the loss in efficiency through aggregation can still be substantial if the nonseasonal component of the model is nonstationary. It is not so serious for long-term forecasting particularly when the nonseasonal component is stationary. There is no loss in efficiency due to aggregation if the basic model is a purely seasonal process.

## INFORMATION LOSS DUE TO AGGREGATION IN PARAMETER ESTIMATION

### Parameter Estimation of a Seasonal Model

Assume that the set of $N=(n+d+Ds)$ observations $z_1$, $z_2$, $\ldots$, $z_N$ are generated by a general multiplicative seasonal model of order $(p,\ d,\ q)\times(P,\ D,\ Q)_s$, given in (1). Note that the model can be parameterized in terms of the zeroes of $\phi_p(B)$, $\alpha_P(B^s)$, $\theta_q(B)$, and $\beta_Q(B^s)$, so that the process can be rewritten as

$$\prod_{j=1}^{p}(1-\delta_j B)\prod_{j=1}^{P}(1-\Gamma_j B^s)w_t=\prod_{j=1}^{q}(1-h_j B)\prod_{j=1}^{Q}(1-H_j B^s)a_t \qquad (28)$$

where

$$w_t=(1-B^s)^D(1-B)^d z_t,$$

$$\phi_p(B)=\Pi_{j=1}^{p}(1-\delta_j B),$$

$$\alpha_P(B^s)=\Pi_{j=1}^{P}(1-\Gamma_j B^s),$$

$$\theta_q(B)=\Pi_{j=1}^{q}(1-h_j B)\text{ and}$$

$$\beta_Q(B^s)=\Pi_{j=1}^{Q}(1-H_j B^s)$$

Let $\underset{\sim}{\delta}=[\delta_1,\ \ldots,\ \delta_p]'$, $\underset{\sim}{\Gamma}=[\Gamma_1,\ \ldots,\ \Gamma_P]'$, $\underset{\sim}{h}=[h_1,\ \ldots,\ h_q]'$, and $H=[H_1,\ \ldots,\ H_Q]'$. Under the normality assumption of $a_t$'s, the likelihood function for the parameters $\delta$'s, $\Gamma$'s, $h$'s, and $H$'s is given by

$$L(\underset{\sim}{\delta},\ \underset{\sim}{\Gamma},\ \underset{\sim}{h},\ \underset{\sim}{H},\ \sigma_a^2|\underset{\sim}{w})=(2\Pi)^{-n/2}(\sigma_a^2)^{-n/2}\exp\left[-\frac{1}{2\sigma_a^2}\sum_{t=1}^{n}\hat{a}_t^2(\underset{\sim}{\delta},\underset{\sim}{\Gamma},\underset{\sim}{h},\underset{\sim}{H})\right] \qquad (29)$$

where

$$\underset{\sim}{w}=[w_1,\ \ldots,\ w_n]' \text{ and } \hat{a}_t(\underset{\sim}{\delta},\underset{\sim}{\Gamma},\underset{\sim}{h},\underset{\sim}{H})=E[a_t|\underset{\sim}{\delta},\underset{\sim}{\Gamma},\underset{\sim}{h},\underset{\sim}{H},\underset{\sim}{w}]$$

denotes the expectation of $a_t$ conditional on $\underset{\sim}{\delta}$, $\underset{\sim}{\Gamma}$, $\underset{\sim}{h}$, $\underset{\sim}{H}$ and $\underset{\sim}{w}$.

For moderate and large values of $n$, the likelihood function in (29) is dominated by the sum of squares function given by

$$S(\underset{\sim}{\delta},\underset{\sim}{\Gamma},\underset{\sim}{h},\underset{\sim}{H})=\sum_{t=1}^{n}\hat{a}_t^2(\underset{\sim}{\delta},\underset{\sim}{\Gamma},\underset{\sim}{h},\underset{\sim}{H}) \qquad (30)$$

and $\hat{a}_t(\underset{\sim}{\delta},\underset{\sim}{\Gamma},\underset{\sim}{h},\underset{\sim}{H})$ can be explicitly written as

$$\hat{a}_t(\underset{\sim}{\delta},\underset{\sim}{\Gamma},\underset{\sim}{h},\underset{\sim}{H})=E\{w_t-\Lambda_1 w_{t-1}\ldots-\Lambda_{p+Ps}w_{t-(p+Ps)}+\tau_1 a_{t-1}+\ldots+\tau_{q+Qs}a_{t-(q+Qs)}\} \qquad (31)$$

where $E(w_t)=w_t$ for $t=1,\ 2,\ \ldots,\ n$ and is the back forecast for $w_t$ for $t\le 0$, and $\Lambda$'s are functions of $\delta$'s and $\Gamma$'s, and $\tau$'s are functions of $h$'s and $H$'s, obtained by the relation

$$w_t-\Lambda_1 w_{t-1}\ldots-\Lambda_{p+Ps}w_{t-(p+Ps)}=\Pi_{j=1}^{p}(1-\delta_j B)\Pi_{j=1}^{P}(1-\Gamma_j B^s)w_t$$

and

$$a_t-\tau_1 a_{t-1}\ldots-\tau_{q+Qs}a_{t-(q+Qs)}=\Pi_{j=1}^{q}(1-h_j B)\,\Pi_{j=1}^{Q}(1-H_j B^s)a_t$$

The least squares estimates obtained by minimizing the sum of squares in (30) will usually provide a good approximation to the maximum likelihood estimates. However, it can be easily seen from (31) that the calculation of $\hat{a}_t$ depends on $(p+Ps)$ values of the $w$'s and $(q+Qs)$ values of $a$'s prior to the commencement of the $w$ series. One solution is to start the recursive calculation of the $\hat{a}$'s at $t=\max.\ (p+Ps,\ q+Qs)$, setting previous $a$'s equal to zero. An alternative solution to this starting value problem is to note that the model in (28) can also be written as

$$\prod_{j=1}^{p}(1-\delta_j F)\prod_{j=1}^{P}(1-\Gamma_j F^s)w_t=\prod_{j=1}^{q}(1-h_j F)\prod_{j=1}^{Q}(1-H_j F^s)e_t \qquad (32)$$

where $e_t$'s are i.i.d $N(0,\ \sigma_e^2)$. Thus,

$$w_t=\Lambda_1 w_{t+1}\ \ldots\ +\Lambda_{p+Ps}w_{t+(p+Ps)}+e_t-\tau_1 e_{t+1}\ \ldots\ -\tau_{q+Qs}e_{t+(q+Qs)} \qquad (33)$$

By letting $\hat{e}_t = E[e_t | \underset{\sim}{\delta}, \underset{\sim}{\Gamma}, \underset{\sim}{h}, \underset{\sim}{H}] = 0$ for $t \geq n-(p+Ps)$ and given observations $w_n, w_{n-1}, \ldots, w_2, w_1$, we can use (33) to backforecast the conditional expectation $\hat{e}_t$'s and, hence, to calculate the backforecast of $w_t$ for $t \leq 0$. The desired terms $\hat{a}_t$'s can then be obtained from (31). The details of this technique can be found in Box and Jenkins [2, 212-220].

Let $\underset{\sim}{\eta}' = [\underset{\sim}{\delta}', \underset{\sim}{\Gamma}', \underset{\sim}{h}', \underset{\sim}{H}'] = [\eta_1, \ldots, \eta_K]$, where $K = p + P + q + Q$ and let $\hat{\underset{\sim}{\eta}}$ be the estimator of $\underset{\sim}{\eta}$, based on the basic model (28). It is known that, under fairly general conditions, the large sample variance-covariance matrix of this estimator is given by $V(\hat{\underset{\sim}{\eta}}) = I_a^{-1}(\underset{\sim}{\eta})$, where $I_a(\underset{\sim}{\eta}) = E(2\sigma_a^2)^{-1}\left[\dfrac{\partial^2 S(\underset{\sim}{\eta})}{\partial\eta_i \partial\eta_j}\right]$ and $S(\underset{\sim}{\eta}) = S(\underset{\sim}{\delta}, \underset{\sim}{\Gamma}, \underset{\sim}{h}, \underset{\sim}{H})$. It is readily seen that $I_a(\underset{\sim}{\eta})$ is, in fact, equal to $E(U'U)\sigma_a^{-2}$, where $U$ is the $n \times K$ matrix of derivatives given by

$$\frac{\partial a_t}{\partial \delta_i} = -(1-\delta_i B)^{-1} a_{t-1}$$

$$\frac{\partial a_t}{\partial \Gamma_i} = -(1-\Gamma_i B^s)^{-1} a_{t-s}$$

$$\frac{\partial a_t}{\partial h_i} = (1-h_i B)^{-1} a_{t-1}$$

$$\frac{\partial a_t}{\partial H_i} = (1-H_i B^s)^{-1} a_{t-s} \tag{34}$$

### A Measure of Information Loss Due to Aggregation

Given the basic model in (1), the corresponding aggregate model has been derived in (15), which can be rewritten in terms of the zeros of $\lambda_p(\mathscr{B})$, $\alpha_P(\mathscr{B}^S)$, $v_r(\mathscr{B})$, and $\beta_Q(\mathscr{B}^S)$ as

$$\prod_{j=1}^{p} (1-\bar{\delta}_j \mathscr{B}) \prod_{j=1}^{P} (1-\Gamma_j \mathscr{B}^S) V_T = \prod_{j=1}^{r} (1-\bar{h}_j \mathscr{B}) \prod_{j=1}^{Q} (1-H_j \mathscr{B}^S) C_T \tag{35}$$

where

$V_T = (1-\mathscr{B}^S)^D (1-\mathscr{B})^d Z_T$, $\alpha_P(\mathscr{B}^S) = \Pi_{j=1}^P (1-\Gamma_j \mathscr{B}^S)$,

$\beta_Q(\mathscr{B}^S) = \Pi_{j=1}^Q (1-H_j \mathscr{B}^S)$, $\lambda_p(\mathscr{B}) = \Pi_{j=1}^p (1-\bar{\delta}_j \mathscr{B})$,

$v_r(\mathscr{B}) = \Pi_{j=1}^r (1-\bar{h}_j \mathscr{B})$

and $\bar{\delta}_j$'s and $\bar{h}_j$'s are functions of $\delta_j$'s and $h_j$'s.

Hopefully, by using the estimation procedure in the subsection on parameter estimation of a seasonal model, we would also be able to obtain the estimator $\hat{\hat{\underset{\sim}{\eta}}}$ of $\underset{\sim}{\eta}$ and find $V(\hat{\hat{\underset{\sim}{\eta}}}) = I_a^{-1}(\underset{\sim}{\eta})$, where $I_a(\underset{\sim}{\eta})$ is the large sample information matrix, based on the aggregate model in (35). Define

$$\xi(m) = 1 - \frac{\det V(\hat{\underset{\sim}{\eta}})}{\det V(\hat{\hat{\underset{\sim}{\eta}}})} = 1 - \frac{\det I_a(\underset{\sim}{\eta})}{\det I_d(\underset{\sim}{\eta})} \tag{36}$$

We can then use $\xi(m)$ to assess the information loss in estimation due to aggregation. Unfortunately, the relationship between the parameters $\bar{\delta}$'s and $\bar{h}$'s in the aggregate model and $\delta$'s and $h$'s in the basic model is so confounded that it is almost impossible to locate $\delta$'s and $h$'s through $\bar{\delta}$'s and $\bar{h}$'s. However, by considering some common parameters in the basic model (28) and aggregate model (35), such as $\Gamma$'s and $H$'s, $\xi(m)$ can still give us an idea how serious is the information loss in estimation due to aggregation.

### Minimum Information Loss in Estimation Due to Aggregation

We have shown that, insofar as forecasting future aggregates is concerned, there is no loss in information if the basic model is a purely seasonal model and the number of aggregation components $m$ be such that $s = mS$ for some integer $S$. The result is not that surprising, because, in this case, the aggregate model has exactly the same form as the given basic model. This represents the best situation we can have under temporal aggregation. It is of interest to know whether this result also remains true in the case of parameter estimation.

If $p = d = q = 0$, the model (28) becomes

$$\prod_{j=1}^{P} (1-\Gamma_j B^s) w_t = \prod_{j=1}^{Q} (1-H_j B^s) a_t \tag{37}$$

where $a_t$'s are i.i.d. $N(0, \sigma_a^2)$. The large sample information matrix for $\Gamma_j$'s and $H_j$'s, based on the basic model, can be shown through (34) to be

$$I_d(\eta) = n \begin{bmatrix} (1-\Gamma_1^2)^{-1} & (1-\Gamma_1\Gamma_2)^{-1} & \cdots & (1-\Gamma_1\Gamma_P)^{-1} & -(1-\Gamma_1 H_1)^{-1} & \cdots & -(1-\Gamma_1 H_Q)^{-1} \\ \vdots & & & & & & \\ (1-\Gamma_P\Gamma_1)^{-1} & (1-\Gamma_P\Gamma_2)^{-1} & \cdots & (1-\Gamma_P^2)^{-1} & -(1-\Gamma_P H_1)^{-1} & \cdots & -(1-\Gamma_P H_Q)^{-1} \\ -(1-H_1\Gamma_1)^{-1} & -(1-H_1\Gamma_2)^{-1} & \cdots & -(1-H_1\Gamma_P)^{-1} & (1-H_1^2)^{-1} & \cdots & (1-H_1 H_Q)^{-1} \\ \vdots & & & & & & \vdots \\ -(1-H_Q\Gamma_1)^{-1} & -(1-H_Q\Gamma_2)^{-1} & \cdots & -(1-H_Q\Gamma_P)^{-1} & (1-H_Q H_1)^{-1} & \cdots & (1-H_Q^2)^{-1} \end{bmatrix}$$

The corresponding aggregate model (35) in this case becomes

$$\prod_{j=1}^{P} (1-\Gamma_j \mathscr{B}^S) V_T = \prod_{j=1}^{Q} (1-H_j \mathscr{B}^S) C_T \tag{39}$$

where $C_T$'s are i.i.d. $N(0, m\sigma_a^2)$. The large sample information matrix for $\Gamma$'s and $H$'s, based on this aggregate model, is then easily seen to be $I_a(\underset{\sim}{\eta})/m$. Thus,

$$\xi(m) = 1 - m^{-(P+Q)} \tag{40}$$

To see the implication of (40), assume $P=1$, $Q=0$, and the basic series is a monthly series with $s=12$, we have $\xi(2)=1/2$, $\xi(3)=2/3$, $\xi(6)=5/6$, and $\xi(12)=11/12$. In estimating the parameters in the basic model, temporal aggregation, hence, leads to a tremendous loss in information. In fact, $\xi(m)$ in (40) is an increasing function of $m$ and the number of parameters in the model. The larger the $m$ and the more parameters in the model, the more information loss is caused by aggregation.

## APPLICATION OF ANALYSIS

In this section, we illustrate the results discussed in previous sections through a simple example, using the monthly data of the U.S. employed civilian workers from January 1949 through December 1974 as our basic series $z_t$. The aggregate series $Z_T$ is the quarterly observations of the employment data during the same period. The data are given in the appendix. However, in order to compare the actual forecasting performance between the basic monthly model and the aggregate quarterly model, we use the data from 1949 to 1973 as our basis to identify the underlying process and estimate its parameters.

### Identification of Basic Monthly Model

Table 1 shows the sample autocorrelations of 300 monthly observations from 1949 through 1973. By applying the three-stage iterative procedure, proposed by Box and Jenkins [2, 18], the monthly model would be $(0, 1, 1) \times (0, 1, 1)_{12}$

$$(1-B^{12})(1-B)z_t=(1-\theta B)(1-HB^{12})a_t \qquad (41)$$

where $a_t$'s are independently and identically distributed as $N(0, \sigma_a^2)$.

Let $w_t=(1-B^{12})(1-B)z_t$. Since $G(B^{12})=(-HF^{12}+(1+H^2)-HB^{12})$ and $g(B)=(-\theta F+(1+\theta^2)-\theta B)$, (5) and (6) imply that the autocovariance generating function of $\{w_t\}$ is given by

$$\gamma_w(B)=\sigma_a^2 \sum_{\ell=-\infty}^{\infty} \gamma_\ell B^\ell \qquad (42)$$

where $\gamma_0=(1+H^2)(1+\theta^2)$, $\gamma_1=-\theta(1+H^2)$, $\gamma_{11}=\gamma_{13}=\theta H$, $\gamma_{12}=-H(1+\theta^2)$ and $\gamma_\ell=0$ otherwise.

### Aggregate Quarterly Model

For a basic model of order $(0, 1, 1) \times (0, 1, 1)_{12}$ and $m=3$, theorem 1 implies that the quarterly model should be of order $(0, 1, 1) \times (0, 1, 1)_4$

$$(1-\mathscr{B}^4)(1-\mathscr{B})Z_T=(1-v\mathscr{B})(1-H\mathscr{B}^4)C_T \qquad (43)$$

where $C_T$'s are i.i.d. $N(0, \sigma_C^2)$. The parameters $H$ in both (41) and (43) are the same, while the parameters $v$ and $\sigma_C^2$ are related to the basic parameters $\theta$ and $\sigma_a^2$ through (16) as follows:

$$\sigma_C^2(1+v^2)=\sigma_a^2(19\theta^2-32\theta+19)$$

$$-\sigma_C^2 v=\sigma_a^2(4\theta^2-11\theta+4) \qquad (44)$$

Table 2 shows the sample autocorrelations of 100 quarterly observations from 1949 through 1973. Applying, again, the Box-Jenkins three-stage iterative procedure, we would come up with a model of order $(0, 1, 1) \times (0, 1, 1)_4$, which confirms the theoretical model implied by theorem 1.

### Estimation of Parameters in Monthly and Quarterly Models

If the parameters $\theta$, $H$, and $\sigma_a^2$ in the monthly model (41) are known, the parameters $v$, $H$, and $\sigma_C^2$ in the quarterly model can be easily obtained from (44). Since these parameters in both models are actually unknown, they are estimated by the nonlinear least squares procedure, subject to some starting values discussed in the subsection on the parameter estimation of a seasonal model. More specifically, $\theta$ and $H$ in the monthly model (41) are estimated by minimizing $\sum_{t=14}^{300} \hat{a}_t^2(\theta, H)$, where $\hat{a}_t=w_t+\theta a_{t-1}+Ha_{t-12}-\theta Ha_{t-13}$ and $a_t \doteq 0$ for $t\leq 13$. The estimates and confidence intervals for $\theta$ and $H$ in this case are

| Parameter | Estimate (standard error) | 95-percent confidence interval |
|---|---|---|
| $\theta$ | 0.21 (.05918) | [0.09, 0.32] |
| $H$ | 0.664 (.04647) | [0.57, 0.75] |

Similarly, by letting $V_T=(1-\mathscr{B}^4)(1-\mathscr{B})Z_T$, $v$ and $H$ in the quarterly model (43) are estimated by minimizing $\sum_{T=6}^{100} \hat{C}_T^2$, where $\hat{C}_T=V_T+vC_{T-1}+HC_{T-4}-vHC_{T-5}$ and $C_T \doteq 0$ for $T\leq 5$. The estimates and confidence intervals for $v$ and $H$ in this case are

| Parameter | Estimate (standard error) | 95-percent confidence interval |
|---|---|---|
| $v$ | −0.32 (0.10106) | [−0.52, −0.12] |
| $H$ | 0.659 (.08137) | [ 0.50, 0.82] |

Also, we have $\hat{\sigma}_a=374.71$ and $\hat{\sigma}_C=1292.4$.

It should be noted that, other than direct estimation from the quarterly model the estimates of $v$, $H$, and $\sigma_C$ can also be obtained through $\hat{\theta}$, $\hat{H}_a$, and $\hat{\sigma}_a$ from (44) and the fact that $H$ in both models are the same. Thus, given $\hat{\sigma}_a=374.71$, $\hat{\theta}=0.21$, and $\hat{H}_a=0.66$, we have $\hat{H}_a=0.66$, $\hat{v}=-0.15$, and $\hat{\sigma}_C=1343.05$. As expected, they are very close to the direct estimation result through the quarterly model.

As pointed out in the subsection on a measure of information loss due to aggregation, to obtain a rough idea of information loss through aggregation in terms of parameter estimation, we can compare the efficiencies of estimates of the common parameter $H$ in monthly and

Table 1. **AUTOCORRELATIONS OF THE MONTHLY EMPLOYMENT DATA**

| Series | Lags | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $z_t$ | 1-12 | 0.98 | 0.96 | 0.94 | 0.93 | 0.91 | 0.89 | 0.88 | 0.87 | 0.87 | 0.86 | 0.86 | 0.86 |
| | ST. E. | .06 | .10 | .13 | .15 | .17 | .18 | .20 | .21 | .22 | .23 | .24 | .25 |
| | 13-24 | .82 | .82 | .80 | .79 | .77 | .76 | .75 | .74 | .74 | .74 | .74 | .73 |
| | ST. E. | .26 | .27 | .28 | .29 | .29 | .30 | .31 | .31 | .32 | .32 | .33 | .33 |
| | 25-26 | .72 | .70 | .69 | .67 | .66 | .65 | .64 | .64 | .64 | .64 | .63 | .63 |
| | ST. E. | .34 | .35 | .35 | .35 | .36 | .36 | .37 | .37 | .37 | .38 | .38 | .38 |
| $(1-B)z_t$ | 1-12 | .22 | .07 | -.19 | .02 | -.35 | -.39 | -.31 | -.01 | -.17 | .04 | .28 | .77 |
| | ST. E. | .06 | .06 | .06 | .06 | .07 | .08 | .08 | .08 | .08 | .08 | .08 | .08 |
| | 13-24 | .26 | .06 | -.20 | -.01 | -.36 | -.35 | -.33 | .04 | -.20 | .06 | .24 | .74 |
| | ST. E. | .11 | .11 | .11 | .11 | .11 | .11 | .12 | .12 | .12 | .12 | .12 | .12 |
| | 25-36 | .25 | .05 | -.18 | -.03 | -.33 | -.36 | -.28 | .02 | -.16 | .06 | .24 | .71 |
| | ST. E. | .14 | .14 | .14 | .14 | .14 | .14 | .14 | .15 | .15 | .15 | .15 | .15 |
| $(1-B^{12})$ $(1-B)z_t$ | 1-12 | -.30 | .17 | -.06 | .16 | .04 | -.08 | .17 | -.24 | .21 | -.19 | .21 | -.49 |
| | ST. E. | .06 | .06 | .06 | .06 | .07 | .07 | .07 | .07 | .07 | .07 | .07 | .08 |
| | 13-24 | .14 | -.01 | -.09 | -.03 | -.18 | .17 | -.23 | .21 | -.17 | .03 | -.08 | .04 |
| | ST. E. | .09 | .09 | .09 | .09 | .09 | .09 | .09 | .09 | .09 | .09 | .09 | .09 |
| | 25-36 | .03 | -.12 | .08 | -.04 | .12 | -.14 | .17 | -.12 | .03 | .07 | .02 | -.04 |
| | ST. E. | .09 | .09 | .09 | .10 | .10 | .10 | .10 | .10 | .10 | .10 | .10 | .10 |

Table 2. **AUTOCORRELATIONS OF QUARTERLY EMPLOYMENT DATA**

| Series | Lags | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $Z_T$ | 1-12 | 0.95 | 0.90 | 0.87 | 0.86 | 0.81 | 0.76 | 0.74 | 0.73 | 0.69 | 0.65 | 0.64 | 0.63 |
| | ST. E. | .10 | .17 | .21 | .24 | .27 | .29 | .31 | .33 | .35 | .36 | .37 | .38 |
| $(1-B)Z_T$ | 1-12 | -.08 | -.65 | -.11 | .89 | -.12 | -.64 | -.11 | .84 | -.13 | -.61 | -.09 | .81 |
| | ST. E. | .10 | .10 | .14 | .14 | .19 | .19 | .21 | .21 | .24 | .24 | .26 | .26 |
| $(1-B^4)$ $(1-B)Z_T$ | 1-12 | .46 | .13 | -.12 | -.42 | -.31 | -.20 | -.15 | -.08 | -.03 | .10 | .08 | -.04 |
| | ST. E. | .10 | .12 | .12 | .12 | .14 | .14 | .15 | .15 | .15 | .15 | .15 | .15 |

Table 3. **FORECASTS OF THE 1974 EMPLOYMENT DATA**

| Lead Time (quarter) | Actual observation | Forecast from quarterly model | Forecast from monthly model | r |
|---|---|---|---|---|
| 1. . . . . . . . . . . . . . . | 253260 | 255040.0 (0.007) | 254189.07 (0.004) | 0.27 |
| 2. . . . . . . . . . . . . . . | 258144 | 260823.4 (.010) | 259967.70 (.007) | .46 |
| 3. . . . . . . . . . . . . . . | 261832 | 264890.4 (.012) | 264026.76 (.008) | .51 |
| 4. . . . . . . . . . . . . . . | 257191 | 264358.2 (.028) | 263484.26 (.025) | .77 |

quarterly models. We first recall from the subsection on minimum information loss in estimation due to aggregation that if the monthly model is of order (0, 0, 0) × (0, 1, 1)$_{12}$, then $\xi(3)=2/3=0.6666$. Now $1-V(\hat{H}_d)/V(\hat{H}_a)=1-0.3263=0.6737$. Thus, in the present case, in terms of parameter estimation, aggregation causes at least a 67-percent loss in efficiency.

### Forecasting Efficiency of Monthly and Quarterly Models

If we are interested in forecasting the quarterly employment figures in 1974, we can either utilize the monthly model to forecast 1974 monthly figures and then aggregate them to obtain the quarterly forecasts or use quarterly model to forecast the desired quarterly values directly.

Table 3 shows the result for these forecasts. The values in parentheses are percentages of actual forecast error, i.e., $\{E(Z_{T+\ell}|\text{past history})-Z_{T+\ell}\}/Z_{T+\ell}$. Also shown in the table is the ratio of forecast error squares between monthly and quarterly models, i.e., $r=(\hat{Z}_T(\ell)-Z_{T+\ell})^2/(\hat{\hat{Z}}_T(\ell)-Z_{T+\ell})^2$. It shows that, even in forecasting quarterly figures, the monthly model gives much more accurate results than the quarterly model. This is especially so if the forecasting lead time $\ell$ is small. In terms of forecasting future aggregates, the loss of information becomes negligible only when the forecasting lead time $\ell$ becomes large, as predicted by our theory in the section on the effect of aggregation on forecasting.

## SUMMARY AND CONCLUDING REMARKS

Since Box and Jenkins developed the so-called Box-Jenkins approach to time series analysis about a decade ago, because of its representation for a wide variety of actual series, the general multiplicative stochastic seasonal time series model, introduced in (1), has become a very popular tool in applied time series analysis, especially in the field of economic and business applications. However, before getting into actual analysis, one must decide first on the time unit he is going to use for his basic observations. The aggregation problem, hence, naturally will come to the mind of any conscientious research worker.

In this paper, we have studied the consequences of temporal aggregation in stochastic seasonal time series model. These results are shown in the following subsections.

### Aggregation Effect on Model Structure

1. Given a stochastic time series model of order $(p, d, q) \times (P, D, Q)_S$, the corresponding model for the aggregates of m-component nonoverlapping sum is of order $(p, d, r) \times (P, D, Q)_S$ where $s=mS$ for some integer $S$ and $r=\left[p+d+1+\frac{q-p-d-1}{m}\right]$.
2. Aggregation contaminates the model structure only through its nonseasonal component. In fact, the order of the process is changed only through the moving average order of the nonseasonal component. Thus, based on the order of the process obtained from modeling time aggregates, the proper limit of the order of the underlying basic series can be obtained.
3. Temporal aggregation will, in general, complicate the model structure. However, as the number of aggregation components m becomes larger, it tends to simplify the model form.

### Aggregation Effect on Forecasting

1. The most serious information loss in forecasting due to aggregation is that, while basic series can be used to forecast any desired future aggregates, temporal aggregates cannot be used to predict desired future disaggregates.
2. As far as forecasting future aggregates is concerned, the loss in efficiency through aggregation depends on the structure of the nonseasonal component of the process. This is expected, because aggregation contaminates model structure only through this component.
3. In forecasting future aggregates, aggregation causes a substantial loss in efficiency when the nonseasonal component of the series is nonstationary; the loss in efficiency is relatively small for long-term forecasting, particularly when the nonseasonal component of the basic model is stationary; there is no loss in efficiency if the basic series is a purely seasonal model.

### Aggregation Effect on Parameter Estimation

1. Given an invertible basic process, there exists a unique set of parameters of the corresponding invertible aggregate model. However, in general, it is almost impossible to locate the parameters of the basic model from the parameters of an aggregate model.
2. In terms of parameter estimation, aggregation causes a tremendous loss in efficiency, regardless of the given model. The larger the number of aggregation components and the more parameters in the model, the more serious information loss is caused by aggregation.

The above results have been supported both by the theory and the numerical results from an empirical application to U.S. employment data.

It is hoped that the results will be useful to time series analysts who are concerned about the implications of temporal aggregation in stochastic time series models. More importantly, it is hoped that the results will bring attention to some research workers who use aggregated data in their statistical analysis and inferences, while being unconscious of the consequences of temporal aggregation.

## REFERENCES


1. Amemiya, T., and Wu, R. Y. "The Effect of Aggregation on Prediction in the Autoregressive Model." Journal of the American Statistical Association 67 (September 1972): 628–632.

2. Box, G. E. P., and Jenkins, G. M. *Time Series Analysis Forecasting and Control*. San Francisco: Holden-Day, Inc., 1970.

3. Brewer, K. R. W. "Some Consequences of Temporal Aggregation and Systematic Sampling for ARMA and ARMAX Models." *Journal of Econometrics* 1 (1973): 133–154.

4. Mundlak, Y. "Aggregation Over Time in Distributed Lag Models." *International Economic Review* 2 (1961): 154–163.

5. Quenouille, M. H. "Discrete Autoregressive Schemes with Varying Time Interval." *Metrika* 1 (1958): 21–27.

6. Sims, C. A. "Discrete Approximations to Continuous Time Distributed Lags in Econometrics." *Econometrica* 39 (1971): 545–563.

7. Telser, L. G. "Discrete Samples and Moving Sums in Stationary Stochastic Process." *Journal of the American Statistical Association* 62 (June 1967): 484–499.

8. Theil, H. *Linear Aggregation of Economic Relationships*. Amsterdam: North-Holland Press, 1954.

9. Tiao, G. C. "Asymptotic Behavior of Time Series Aggregates." *Biometrika* 59 (1972): 525–531.

10. Tiao, G. C., and Wei, W. S. "Effect of Temporal Aggregation on the Dynamic Relationship of Two Time Series Variables." *Biometrika* 63 (1976): 513–523.

11. Zellner, A., and Montmarquette, C. "A Study of Some Aspects of Temporal Aggregation Problems in Econometric Analysis." *Review of Economics and Statistics* 53 (1971): 335–342.

# APPENDIX

Table A-1. U.S. EMPLOYED CIVILIAN WORKERS, BY MONTH

| Year | January | February | March | April | May | June | July | August | September | October | November | December |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1949. . . . | 56,486 | 56,320 | 56,809 | 56,929 | 57,669 | 58,231 | 58,171 | 58,504 | 58,324 | 58,050 | 58,616 | 57,712 |
| 1950. . . . | 56,189 | 56,197 | 56,733 | 57,812 | 58,719 | 59,997 | 59,839 | 60,948 | 60,245 | 60,708 | 60,313 | 59,392 |
| 1951. . . . | 58,166 | 58,102 | 59,366 | 59,206 | 60,219 | 60,373 | 60,968 | 61,128 | 60,408 | 60,906 | 60,464 | 60,252 |
| 1952. . . . | 58,884 | 58,834 | 58,912 | 59,232 | 60,250 | 60,988 | 60,775 | 60,872 | 61,162 | 60,992 | 61,394 | 60,748 |
| 1953. . . . | 60,134 | 60,271 | 60,874 | 60,757 | 61,061 | 62,166 | 62,186 | 62,271 | 61,529 | 61,805 | 61,302 | 59,796 |
| 1954. . . . | 58,645 | 59,059 | 59,119 | 59,537 | 60,020 | 60,497 | 60,523 | 60,858 | 60,952 | 61,210 | 60,901 | 59,990 |
| 1955. . . . | 59,354 | 59,336 | 59,850 | 60,861 | 61,780 | 62,568 | 63,497 | 63,876 | 63,676 | 64,138 | 63,840 | 63,268 |
| 1956. . . . | 62,049 | 61,773 | 62,172 | 63,002 | 64,045 | 64,707 | 64,940 | 65,085 | 64,831 | 65,074 | 64,310 | 63,619 |
| 1957. . . . | 61,974 | 62,512 | 63,134 | 63,512 | 64,213 | 65,127 | 65,726 | 65,009 | 64,769 | 65,112 | 64,129 | 63,598 |
| 1958. . . . | 61,508 | 61,265 | 61,567 | 62,116 | 63,098 | 63,652 | 63,810 | 64,018 | 63,766 | 64,480 | 63,890 | 63,266 |
| 1959. . . . | 62,052 | 62,015 | 63,091 | 64,241 | 65,036 | 65,924 | 66,193 | 65,897 | 65,414 | 65,891 | 64,877 | 64,927 |
| 1960. . . . | 63,375 | 63,871 | 63,607 | 65,450 | 66,342 | 67,288 | 67,239 | 67,004 | 66,892 | 66,563 | 66,394 | 65,287 |
| 1961. . . . | 63,797 | 63,869 | 64,700 | 64,957 | 65,831 | 67,151 | 66,911 | 67,028 | 66,036 | 66,786 | 66,348 | 65,531 |
| 1962. . . . | 64,215 | 64,872 | 65,421 | 65,957 | 67,066 | 67,852 | 67,849 | 68,096 | 67,261 | 67,850 | 67,046 | 66,585 |
| 1963. . . . | 65,168 | 65,519 | 66,329 | 67,240 | 67,984 | 68,844 | 69,225 | 69,052 | 68,567 | 68,964 | 68,471 | 67,791 |
| 1964. . . . | 66,468 | 67,197 | 67,695 | 68,947 | 69,952 | 70,448 | 70,839 | 70,676 | 69,849 | 70,147 | 69,892 | 69,543 |
| 1965. . . . | 68,235 | 68,690 | 69,385 | 70,220 | 71,298 | 72,278 | 73,093 | 72,695 | 71,408 | 72,112 | 71,824 | 71,819 |
| 1966. . . . | 70,368 | 70,691 | 71,090 | 72,066 | 72,619 | 74,037 | 74,655 | 74,665 | 73,248 | 73,744 | 73,995 | 73,600 |
| 1967. . . . | 72,161 | 72,505 | 72,560 | 73,445 | 73,638 | 75,393 | 76,220 | 76,170 | 74,632 | 75,180 | 75,218 | 75,337 |
| 1968. . . . | 73,272 | 74,114 | 74,517 | 75,143 | 75,931 | 77,273 | 77,748 | 77,431 | 75,939 | 76,365 | 76,608 | 76,699 |
| 1969. . . . | 75,357 | 76,180 | 76,520 | 77,077 | 77,265 | 78,958 | 79,615 | 79,646 | 78,026 | 78,671 | 78,716 | 78,789 |
| 1970. . . . | 77,313 | 77,489 | 77,957 | 78,408 | 78,357 | 79,382 | 80,291 | 79,895 | 78,254 | 78,916 | 78,740 | 78,515 |
| 1971. . . . | 77,238 | 77,260 | 77,492 | 78,204 | 78,710 | 79,477 | 80,682 | 80,619 | 79,295 | 80,065 | 80,203 | 80,188 |
| 1972. . . . | 79,106 | 79,366 | 80,195 | 80,626 | 81,225 | 82,628 | 83,443 | 83,506 | 82,035 | 82,707 | 82,702 | 82,882 |
| 1973. . . . | 81,043 | 81,837 | 82,814 | 83,299 | 83,759 | 85,566 | 86,367 | 85,920 | 84,842 | 85,994 | 85,858 | 85,644 |
| 1974. . . . | 84,088 | 84,294 | 84,878 | 85,192 | 85,785 | 87,165 | 88,015 | 87,575 | 86,252 | 86,047 | 85,924 | 85,220 |

## Table A-2. U.S. EMPLOYED CIVILIAN WORKERS, BY QUARTER

| Year | 1st quarter | 2d quarter | 3d quarter | 4th quarter |
|---|---|---|---|---|
| 1949 | 169,615 | 172,829 | 174,999 | 174,378 |
| 1950 | 169,119 | 176,528 | 181,032 | 180,373 |
| 1951 | 175,634 | 179,798 | 182,504 | 181,622 |
| 1952 | 176,630 | 180,470 | 182,809 | 183,134 |
| 1953 | 181,279 | 183,984 | 185,986 | 182,903 |
| 1954 | 176,823 | 180,054 | 182,333 | 182,101 |
| 1955 | 178,540 | 185,209 | 191,049 | 191,246 |
| 1956 | 185,994 | 191,754 | 194,856 | 193,003 |
| 1957 | 187,620 | 192,852 | 195,504 | 192,839 |
| 1958 | 184,340 | 188,866 | 191,594 | 191,636 |
| 1959 | 187,158 | 195,201 | 197,504 | 195,695 |
| 1960 | 190,853 | 199,080 | 201,135 | 198,244 |
| 1961 | 192,366 | 197,939 | 199,975 | 198,665 |
| 1962 | 194,508 | 200,075 | 203,566 | 201,481 |
| 1963 | 197,016 | 204,068 | 206,844 | 205,266 |
| 1964 | 201,360 | 209,347 | 211,364 | 209,582 |
| 1965 | 206,310 | 213,796 | 217,196 | 215,755 |
| 1966 | 212,149 | 218,722 | 222,568 | 221,339 |
| 1967 | 217,226 | 222,476 | 227,022 | 225,735 |
| 1968 | 221,903 | 228,347 | 231,118 | 229,672 |
| 1969 | 228,057 | 233,300 | 237,287 | 236,176 |
| 1970 | 232,759 | 236,147 | 238,440 | 236,171 |
| 1971 | 231,990 | 236,391 | 240,596 | 240,456 |
| 1972 | 238,667 | 244,479 | 248,984 | 248,291 |
| 1973 | 245,694 | 252,624 | 257,129 | 257,466 |
| 1974 | 253,260 | 258,144 | 261,832 | 257,191 |

# COMMENTS ON "SOME CONSEQUENCES OF TEMPORAL AGGREGATION IN SEASONAL TIME SERIES MODELS" BY WILLIAM W. S. WEI

P. M. Robinson
Harvard University

The scheme used here for incorporating seasonal and nonstationary affects in a time series model is a convenient one and certainly facilitates study of the affects of temporal aggregation. The model that the author investigates was considered by Box and Jenkins [1, 305]. It is one in which the basic series can be regarded as the output obtained from passing white noise through two filters; both filters have rational frequency response, but one (responsible for the stationary and nonstationary seasonal effects) is restricted to be a rational function of $\lambda^s$ (where $s$ is the seasonal period). The composite model is then the author's autoregressive integrated moving average (ARIMA) model (1);[1] although other approaches to the modeling problem are certainly possible, a great body of knowledge about ARIMA models, and their estimation, is available, and they are relatively easy to use in forecasting. An autoregressive moving average (ARMA) model for a stationary, detrended, series can be justified on theoretical grounds in that its spectral density provides an arbitrarily good uniform approximation to any continuous spectral density. Moreover, the aggregate series $Z_t$ is formed by passing $z_t$ through a simple moving average filter, and, thus, $Z_t$ is also an ARIMA process. Indeed, as Wei notes, some previous authors have obtained results on the relationship between similar models for $z_t$ and $Z_t$. Unfortunately, these authors found also that, except in simple cases, the relation between the parameters in the two models is complicated and tedious to derive, involving the solution of a polynomial, and this greatly hinders Wei's study of information loss in estimation and forecasting. However, Wei obtains some perspicuous results that confirm one's expectations that, in many important cases, the effects of aggregation on parameter estimates and forecasts are likely to be substantial.

In the practical application in the section on application and analysis, Wei finds that the method of Box and Jenkins [1, 18] for identifying the degree of an ARIMA model works well, in that the model suggested for the aggregated series is the same as that which is obtained after using the Box-Jenkins method on the disaggregated series, and then applying the theory of the present paper. The relative efficiency provided by the two models is then considered. In a situation in which the disaggregated values are not available, however, the extent of the efficiency loss will be less easy to estimate, even in simple models. It seems unlikely that support for a particular model for $z_t$ will often be available from economic theory, and, thus, there may be little theoretical basis for modeling $Z_t$. Moreover, there may be doubt about how to choose the basic time unit for $z_t$. The aggregation operation may, in practice, be more complicated than the author's simple sum (7). In these circumstances, a natural interval to be used in specifying the underlying model for $z_t$ may sometimes be suggested by the nature of the economic transactions involved, but not always. If $z_t$ is regarded as defined continuously in time, a continuous time model would seem more appropriate. This might be a stochastic differential equation model of a type considered by a number of authors. It might alternatively be a difference equation model in which the spans are known or unknown real numbers of which the aggregate time interval $m$ is not necessarily an integer multiple; the only statistical treatment of such a model, of which I am aware, is in Robinson [3]. Any of these continuous time models will, however, raise problems that are similar to, but somewhat more difficult than, Wei's discrete time model.

It should also be mentioned that loss in efficiency may, in practice, turn out to be even greater than Wei suggests. In the subsection on a measure of information loss due to aggregation, Wei states the model for $Z_t$, in terms of the roots $\bar{\delta}_j$, $\Gamma_j$, $\bar{h}_j$, and $H_j$. The $\bar{\delta}_j$, $\bar{h}_j$ are functions of the $\delta_j$, $h_j$, which are the roots in the model for $z_t$. As Wei acknowledges in this subsection, these functions are very complicated. Therefore, it seems quite possible that one will simply estimate the coefficients of the model for $Z_t$, i.e., the coefficients $\alpha_j$, $\beta_j$ in

$$1+\prod_{j=1}^{p+PS} \alpha_j B^j=\prod_{j=1}^{p} (1-\bar{\delta}_j B) \prod_{j=1}^{P} (1-\Gamma_j B^S)$$

$$1+\prod_{j=1}^{r+QS} \beta_j B^j=\prod_{j=1}^{r} (1-\bar{h}_j B)\prod_{j=1}^{Q}(1-H_j B^S)$$

*Research supported, in part, by NSF Grant Soc. 75-13436.*

[1] To avoid identifiability problems, one must assume that the polynomial operators in (1) are of minimal order; in a practical sense this problem seems to be taken care of by the method in [1, 18] that the author uses to identify the most economical model. Also, in order for the decomposition into $\alpha_P(B^S)$ and $\phi_p(B)$ on the one hand, and $\beta_Q(B^S)$ and $\theta_q(B)$ on the other, to be unique, $\alpha_P$ and $\beta_Q$ are presumably assumed to be of maximal order.

(see (35)), without attempting to deduce the underlying model. Although knowledge of $S$ and the degrees $P$ and $Q$ of the seasonal factors may allow us to take some of the $\alpha_j$, $\beta_j$ to be zero a priori, we may still be estimating rather more parameters than the $p+P+q+Q$ parameters $\delta_j$, $\Gamma_j$, $h_j$, and $H_j$ of the underlying model. Even if the nonseasonal moving average order, $q$, in the underlying model is zero, when $m$ is an integral divisor of $S$, the nonseasonal moving average order, $r$, in the aggregate model is $\left[(p+d+1)\left(\frac{m-1}{m}\right)\right]$, which is nonzero whenever $(p+d)(m-1)\geq 1$. Therefore, the inefficiency stems not only from the fact that relatively few pieces of data are being used, but also from the fact that relatively many parameters are being estimated, although one would expect that the former source will usually predominate.

It should also be said that the measure of efficiency employed is somewhat arbitrary. Wei defines a measure of the information loss in estimation due to aggregation to be

$$\xi(m)=1-\frac{\det I_a(\eta)}{\det I_d(\eta)}$$

(See (36).) There are $P+p+Q+q$ elements in $\eta$, and, thus, when the parameter space is large an alternative measure, such as

$$\bar{\xi}(m)=1-\left[\frac{\det I_a(\eta)}{\det I_d(\eta)}\right]^{1/(P+p+Q+q)}$$

will produce numbers that are somewhat less horrifying.

Whether one tries to estimate the model in terms of the $\delta_j$, $\mu_j$ or in terms of the $\bar{\delta}_j$, $\bar{\mu}_j$, it seems that something can be added to Wei's treatment of the estimation problem in the section on information loss due to aggregation in parameter estimation. He sets up the exact normal[2] likelihood in the subsection on parameter estimation of a seasonal model and approximates it by a sum of squares function (30) with no loss in asymptotic efficiency. He recommends that (30) be minimized. This can probably be done by utilizing a computer package for iterative optimization or, if the model is small like the one in his numerical example, by brute force scanning of (30). However, it may be possible to obtain estimates that are asymptotically as efficient without recourse to such numerical methods. For example, Hannan [2, 377] describes a three-step procedure that would estimate the structural coefficients $\alpha_j$, $\beta_j$ in the model for $Z_t$. But, as earlier noted, there may be more $\alpha_j$, $\beta_j$ than $\delta_j$, $\Gamma_j$, $h_j$, and $H_j$. An efficient two-step procedure for estimating quantities corresponding to the $\delta_j$ in a stochastic differential equation driven by pure noise was described by Robinson [4].[3] It involves finding an initial consistent, but inefficient estimate, in a relatively simple fashion and then making a suitable correction to achieve efficiency. This approach seems capable of adaptation to Wei's discrete time model. It should be added that the methods in both [2, 377] and [4] employ Fourier transformed data, and they do not involve the somewhat messy starting value problem dealt with in the subsection on parameter estimation of a seasonal model of Wei's paper. They are, only approximately nonlinear least squares (NLLS) or maximum likelihood (ML) methods, therefore. In small samples, the various methods may give substantially different results. In the absence of information on finite sample properties, however, it is not clear to me that there are significant grounds for preferring NLLS or ML over other methods that are just as efficient asymptotically.

It is of some interest to analyze Wei's model (1) in the frequency domain. Consider the case of a simple purely seasonal process

$$z_t=\delta z_{t-12}+a_t,\ t=1,2,\cdots,|\delta|<1$$

where the $a_t$ are white noise, with $E(a_t^2)=\sigma^2$. This process has spectral density

$$f(\lambda)=\frac{1}{2\pi}\sum_{j=-\infty}^{\infty}E(z_t z_{t+j})e^{-ij\lambda}=\frac{\sigma^2}{2\pi|1-\delta e^{i12\lambda}|^2},\ -\pi<\lambda\leq\pi \quad (1)$$

The function $f$ is periodic of period $\pi/6$. When $\delta\neq 0$, it has stationary points only at the frequencies $\pi j/12$, $j=-11,\ldots,12$. When $\delta>0$, those at the seasonal frequencies $\lambda_j=\pi j/6$, $j=-5,\ldots,6$, are maximum points, while those at the intermediate $\pi j/12$ are minimums. The amplitudes of the peaks vary directly with $\delta$, and (1) can be thought of as the Abel sum of the Fourier series (see Zygmund [5, 96]) of the limiting generalized function $g(\lambda)$ that gives delta function weight to the $\lambda_j$, while giving zero weight to all other frequencies. The function $g(\lambda)$ corresponds to a purely periodic process, and (1) can be thought of as a smooth approximation to it. This suggests that one might use alternative methods of approximate summation of the Fourier series of $g(\lambda)$ to represent seasonal peaks in the spectrum. However, most of these, unlike the author's, could not be very conveniently incorporated in a time domain model.

The effect of passing a purely seasonal process through the nonseasonal filter $\theta_q(B)/\phi_p(B)$ is to modify the location and amplitude of the seasonal peaks to a greater or lesser extent. There is an alternative, additive model, considered by Hannan [2, 174], that can also employ the ARMA idea. It might have been interesting if Wei had considered the

[2]It may be noted that his normality assumption in the section on model structure of temporal aggregates is not of great importance. Certainly, it motivates the likelihood criterion, and, certainly, it justifies using only autocovariances in the estimation procedure. But, the parameter estimates will have the same asymptotic distribution without the normality assumption as they will if normality is imposed.

[3]It may be noted that, whereas the coefficients in $\alpha_p(B^S)$, $\phi_P(B)$, $\theta_q(B)$, and $\beta_Q(B^S)$ are real, in general, some of the $\delta_j$, $\Gamma_j$, $h_j$, and $H_j$ will be in complex conjugate pairs in which case the distribution in the central limit theorem for the estimates will be complex multivariate normal.

effects of temporal aggregation on this model also and investigated its performance on the data. One writes

$$z_t = q_t + r_t + s_t$$

where $q_t$, $r_t$, and $s_t$ are unobservables, such that $q_t$ is a trend, $r_t$ is a stationary (possibly ARMA) process, and $s_t$ is an evolving seasonal process with representation

$$s_t = \sum_{j=1}^{6} (\alpha_{jt} cos\ \lambda_j t + \beta_{jt} sin\ \lambda_j t)$$

where the $\lambda_j$ are as previously mentioned, and $\alpha_{jt}$ and $\beta_{kt}$ are stationary (possibly ARMA) processes that are incoherent for all $j$, $k$, and $\alpha_{jt}$, $\beta_{jt}$ have identical autocovariance properties for each $j$.

# REFERENCES


1. Box, G. E. P. and Jenkins, G. M. *Time Series Analysis: Forecasting and Control.* San Francisco: Holden-Day, Inc., 1970.
2. Hannan, E. J. *Multiple Time Series.* New York: Wiley and Sons, Inc., 1970.
3. Robinson, P. M. "Stochastic Difference Equations with Non-Integral Differences." *Advances in Applied Probability* 6 (September 1974): 524–545.
4. ———. "The Efficient Estimation of a Rational Spectral Density," 1975. (Manuscript.)
5. Zygmund, A. *Trigonometric Series. Vol. I.* Cambridge: Cambridge University Press, 1959.