Journal of Business & Economic Statistics, October 1990, Vol. 8, No. 4

Predictive Accuracy Gain From Disaggregate Sampling in ARIMA Models

Theo E. Nijman
Department of Econometrics, Tilburg University, 5000 LE Tilburg, the Netherlands

Franz C. Palm
Department of Economics, University of Limburg, 6200 MD Maastricht, the Netherlands

We compare the forecast accuracy of autoregressive integrated moving average (ARIMA) models based on data observed with high and low frequency, respectively. We discuss how, for instance, a quarterly model can be used to predict one quarter ahead even if only annual data are available, and we compare the variance of the prediction error in this case with the variance if quarterly observations were indeed available. Results on the expected information gain are presented for a number of ARIMA models including models that describe the seasonally adjusted gross national product (GNP) series in the Netherlands. Disaggregation from annual to quarterly GNP data has reduced the variance of short-run forecast errors considerably, but further disaggregation from quarterly to monthly data is found to hardly improve the accuracy of monthly forecasts.

KEY WORDS: ARMA models; Disaggregate sampling; Prediction.

1. INTRODUCTION

In recent years, there has been an increased tendency toward collecting and analyzing disaggregate data. In the Netherlands, for instance, the Central Bureau of Statistics publishes quarterly National Accounts, which were until a few years ago only available on an annual basis. In the United States, many series are currently available on a monthly basis. Assuming that the disaggregate data are generated by an autoregressive integrated moving average (ARIMA) model, we show how much additional information is contained in the temporally disaggregate data that can be used to improve the forecast performance at the disaggregate level. Knowledge about the expected gain of information is required to decide on whether it is worthwhile to collect data at the disaggregate level. Moreover, our results can contribute to solving the choice problem of using infrequently sampled data with negligible measurement errors or data at a disaggregate level that generally include larger errors because they are often partly constructed or estimated. Note, however, that we assume that the observations are measured without error.

Palm and Nijman (1984) and Nijman and Palm (in press) considered conditions for identification of a disaggregate model from aggregate data. If the parameters of the forecast function for the disaggregate series are identifiable, the aggregate data can be used to construct disaggregate forecasts, which we compare with forecasts from disaggregate data. We show how analytical expressions for the predictors can be obtained and discuss the use of Kalman filtering techniques for cases that are not analytically tractable. Numerical examples give an indication of the order of magnitude of the improvement of forecast performance resulting from the additional information in the disaggregate data. We restrict ourselves to the comparison of the forecast accuracy of correctly specified univariate ARIMA models based on data observed with a low frequency and a high frequency, respectively.

The implication of temporal aggregation for the model specification and parameter estimation were studied by Brewer (1973) and Weiss (1984) for an autoregressive moving average (ARMA) model and an ARMA model with (lagged) exogenous terms and by Engle and Liu (1972), Geweke (1978), Mundlak (1961), Wei (1978), and Zellner and Montmarquette (1971), among others, for regression models. Palm and Nijman (1984) and Nijman and Palm (1990) considered the identification and estimation of ARIMA models for variables that are sampled with a longer interval than that of the realizations. The estimation of the unobserved realizations has been considered in the literature on interpolation and distribution of time series (e.g., see Chow and Lin 1971; Fernandez 1981; Harvey and Pierse 1984; Litterman 1983; Nijman 1985; Nijman and Palm 1986). The loss of information due to contemporaneous aggregation was analyzed by Kohn (1982), Lütkepohl (1984a,b; 1987), Rose (1977), and Tiao and Guttman (1980). Most closely related to our work are the attempts by, for example, Abraham and Ledolter (1982), Ahsanullah and Wei (1984), Amemiya and Wu (1972), and Lütkepohl (1986, 1987) to quantify the effect of temporal aggregation on the forecast-error variance for the *aggregate* time series. We are concerned, however, with predicting *disaggregate* time series given that the realizations are sampled with a lower frequency. We show how, for instance, a quarterly model can be used to predict one quarter ahead even if only annual data are available.

The plan of the article is as follows. In Section 2, we derive analytical expressions for the minimum mean squared error (MMSE) predictor of the disaggregate series from aggregate data. In Section 3, results on the reduction in the variance of the prediction error due to increasing the frequency of sampling to become identical to that of the realization of the variables are presented. The impact of sampling variation of parameter estimates on the precision of the forecasts is also investigated. In Section 4, we analyze quarterly data on the seasonally adjusted gross national product (GNP) series in the Netherlands that has been recently constructed at The Netherlands Central Bank (see De Nederlandsche Bank 1986). Using results of the previous sections, we show by how much the prediction-error variance of quarterly GNP is reduced through the availability of past quarterly observations on this series, and we examine whether a further disaggregation to monthly data is desirable. Finally, Section 5 contains concluding remarks.

2. MMSE PREDICTION OF HIGH-FREQUENCY SERIES FROM LOW-FREQUENCY DATA

Consider a time series y_t $(t = 1, \ldots, T)$, which is generated by the univariate ARMA(1, 1) model

$$(1 - \rho L)y_t = (1 - \alpha L)\varepsilon_t, \qquad \varepsilon_t \sim NID(0, \sigma_\varepsilon^2), \quad (1)$$

where L is the lag operator defined by $Ly_t = y_{t-1}$. If y_t is a stock variable observed every mth period $(m > 1)$, the sample will consist of the values of y_t for $t \in \{m, 2m, \ldots, [T/m]m\}$, where $[T/m]$ is the largest integer smaller than or equal to T/m. If y_t is a flow variable, $\bar{y}_t = \sum_{i=0}^{m-1} y_{t-i}$ $(t \in T_m)$ will be observed. Throughout, we will assume that T is sufficiently large to neglect the dependence of the MMSE predictors on presample data. Moreover, we assume that the parameters of the forecast function for the disaggregate data are identified.

2.1 Stock Variables

If y_t is a stock variable generated by an AR(1) process $(\alpha = 0)$, the MMSE predictor of y_{T+k} $(k > 0)$ is simply

$$E[y_{T+k} \mid y_t, t \in T_m] = \rho^{k+r} y_{T-r}, \quad (2)$$

where $T_m = \{jm: j = \ldots, -1, 0, 1 \ldots [T/m]m\}$, r is the number of periods between T and the last low-frequency observation, and $r = T - [T/m]m$. If y_t is a flow variable or if y_t is a stock variable generated by a more general ARMA model, however, it is at first sight less straightforward to derive MMSE predictors of the disaggregate series. Later we will show how the structure of the forecasts for the low-frequency data can be used to obtain expressions for the MMSE predictor of the high-frequency series if (1) holds. This approach is no longer attractive for higher-order ARMA models, and we discuss in Appendix A how results from filtering theory (e.g., see Priestley 1981) can be used to obtain analytical expressions for the predictors in these cases. Since the analytical expressions can get cumbersome for high-order models, we also show in this appendix how the required conditional expectations and the associated variances of the prediction errors can be numerically evaluated using the recursive Kalman filter (e.g., see Harvey 1981 or Anderson and Moore 1979).

To derive the MMSE predictors if y_t is a stock variable generated by the ARMA(1, 1) model in (1), first multiply (1) by $(1 + \rho L + \cdots + \rho^{m-1}L^{m-1})$, which yields

$$(1 - \rho^m L^m)y_t = (1 + \rho L + \cdots + \rho^{m-1}L^{m-1}) \times (1 - \alpha L)\varepsilon_t = \sum_{i=0}^{m} \eta_i \varepsilon_{t-i}. \quad (3)$$

The second equality in (3) defines the η_i. As shown by Amemiya and Wu (1972) and Palm and Nijman (1984), among others, it results from (3) that the low-frequency series y_t $(t \in T_m)$ is generated by the ARMA(1, 1) model

$$(1 - \rho^m L^m)y_t = (1 - \lambda L^m)v_t, \qquad v_t \sim NID(0, \sigma_v^2) \quad \text{if } t \in T_m, \quad (4)$$

where λ and σ_v^2 can be derived from the equality of the mth-order autocorrelation and the variance of the right sides of (3) and (4),

$$-\lambda/(1 + \lambda^2) = \eta_m \Big/ \sum_{i=0}^{m} \eta_i^2 \quad (|\lambda| < 1), \quad (5)$$

and

$$\sigma_v^2 = \sigma_\varepsilon^2 \sum_{i=0}^{m} \eta_i^2/(1 + \lambda^2). \quad (6)$$

Equation (5) is quadratic in λ but yields a unique solution for λ within the unit circle. It is important to notice that if $\alpha = 0$ and m is even, the sign of ρ is not identified [ρ is locally identified; e.g., see Palm and Nijman (1984)]. Moreover, if $m = 2$ and $\rho = 0$, α is not identified. If $k + r$ is odd, the forecast function in (2) can then only be computed if ρ is known or if the sign of ρ is known a priori so that its value can be estimated.

The ARMA(1, 1) model (4) for the low-frequency data can be used to obtain the MMSE predictor of the high-frequency data in the following way. Equation (4) implies that

$$E[y_{T+m} \mid y_t, t \in T_m] = (\rho^m - \lambda) \sum_{i=0}^{\infty} \lambda^i y_{T-im} \quad \text{if } r = 0, \quad (7)$$

a result that will be used to derive $E[y_{T+k} \mid y_t, t \in T_m]$ $(\forall k > 0)$. Define the coefficients a_i by $E[y_{T+1} \mid y_t, t \in$

$T_m] = \sum_{i=0}^{\infty} a_i y_{T-im}$ if $r = 0$. From (1), one obtains $E[y_{T+k} \mid y_t, t \in T_m] = \rho^{k-1}E[y_{T+1} \mid y_t, t \in T_m]$ as long as $k > 1$. Substitution of $k = m$ and comparison with (7) yields

$$a_i = \rho(1 - \lambda\rho^{-m})\lambda^i, \quad (8)$$

and we get the MMSE predictor of the high-frequency variable given the low-frequency observations

$$E[y_{T+k} \mid y_t, t \in T_m] = \rho^{k+r}(1 - \lambda\rho^{-m}) \sum_{i=0}^{\infty} \lambda^i y_{T-im-r}. \quad (9)$$

2.2 Flow Variables

Expressions for the MMSE predictor when y_t is a flow variable can be obtained in a similar way. Premultiplying (3) by $(1 + L + \cdots + L^{m-1})$ yields

$$\begin{aligned}(1 - \rho^m L^m)\bar{y}_t &= (1 + L + \cdots + L^{m-1}) \\ &\quad \times (1 + \rho L + \cdots + \rho^{m-1}L^{m-1})(1 - \alpha L)\varepsilon_t \\ &= \sum_{i=0}^{2m-1} \bar{\eta}_i \varepsilon_{t-i}, \end{aligned} \quad (3')$$

which can be used to show that $\bar{y}_t$ $(t \in T_m)$ is generated by the ARMA(1, 1) model

$$(1 - \rho^m L^m)\bar{y}_t = (1 - \bar{\lambda}L^m)v_t, \quad v_t \sim NID(0, \sigma_v^2) \quad \text{if } t \in T_m, \quad (4')$$

where $\bar{\lambda}$ and σ_v^2 can be obtained using the analog of (5) and (6). Notice that when $m = 2$, ρ is identified in (4') even if $\alpha = 0$. Subsequently, define the coefficients $\bar{a}_i$ by $E[y_{T+1} \mid \bar{y}_t, t \in T_m] = \sum_{i=0}^{\infty} \bar{a}_i \bar{y}_{T-im}$ if $r = 0$. Along the lines adopted for stock variables, two expressions for the m-period-ahead forecast can be used to solve for the weights in the high-frequency forecast:

$$\begin{aligned} E[\bar{y}_{T+m} \mid \bar{y}_t, t \in T_m] &= (\rho^m - \bar{\lambda}) \sum_{i=0}^{\infty} \bar{\lambda}^i \bar{y}_{T-im} \\ &= (1 + \rho + \quad + \rho^{m-1}) \sum_{i=0}^{\infty} \bar{a}_i \bar{y}_{T-im} \quad (\text{if } r = 0), \end{aligned} \quad (10)$$

which yields the following MMSE predictor when flow variables are observed:

$$E[y_{T+k} \mid \bar{y}_t, t \in T_m] = \rho^{k+r-1}(1 + \rho + \cdots + \rho^{m-1})^{-1} \times (\rho^m - \bar{\lambda}) \sum_{i=0}^{\infty} \bar{\lambda}^i \bar{y}_{T-im-r}. \quad (11)$$

In the derivations of (7) and (11), we did not exclude the case $\rho = 1$, so results for IMA(1, 1) models are obtained as a special case. For more general ARIMA models, comparison of low-frequency forecasts no longer yields direct expressions for the high-frequency predictors. In Appendix A, we discuss how predictors in these models can be obtained using classical Wiener–Kolmogorov or recursive Kalman filtering procedures.

3. PREDICTIVE ACCURACY GAIN FROM DISAGGREGATE SAMPLING

In Section 2, we have shown that high-frequency series can be predicted using low-frequency data. Evidently the MMSE of these predictors will be larger than that of MMSE predictors based on high-frequency series. An important point, of course, is how much additional information temporally disaggregate observations contain. This topic is addressed in this section.

A first measure of the accuracy gain of predictors based on disaggregate sampling is the reduction in the variance of the prediction error γ defined for stock variables by

$$\gamma = 100(1 - E)\{y_{T+k} - E[y_{T+k} \mid y_t, t \in T_1]\}^2 \div E\{y_{T+k} - E[y_{T+k} \mid y_t, t \in T_m]\}^2. \quad (12)$$

A similar expression in which the conditioning is on $\bar{y}_t$ $(t \in T_m)$ defines the measure of the information gain in case of flow variables. The reduction in the prediction-error variance depends on k, m, r, and the parameters of the model generating y_t.

3.1 Stock Variables

If y_t is a stock variable generated by the AR(1) model that is obtained if $\alpha = 0$ in (1), the reduction in the prediction-error variance can be expressed as

$$\gamma = 100\rho^{2k}(1 - \rho^{2r})(1 - \rho^{2k+2r})^{-1} \quad (13)$$

irrespective of the value of m. This is because, for the AR(1) model, the optimal forecast only depends on the most recent observation. In this model, the potential gain of information is caused purely by the fact that y_t might have been observed after period $[T/m]m$. An upper bound on the information gain is obtained here by putting r at its maximum value, $r = m - 1$. Numerical results are presented in Table 1. We assume that ρ is identifiable; that is, at least the sign of ρ is a priori known. For this model the information gain appears to be substantial only in short-term forecasting when the autoregressive parameter is large in absolute value.

One could argue that (13) underestimates the true efficiency gain because it ignores the fact that in applications parameters have to be estimated and can be estimated more accurately if the sampling frequency is increased. Therefore, we also present results for the case in which the assumption of known parameters has been replaced by approximations up to order T^{-1} for the unconditional prediction-error variances when the

Table 1. Upper Bounds in Percentage Points for the Reduction of the Prediction-Error Variance When a Stock Variable Is Generated by an AR(1) Model

	ρ = +/− .8			ρ = +/− .4		
k	m = 2	m = 3	m = 4	m = 2	m = 3	m = 4
1	39 (43)	51 (56)	57 (62)	14 (19)	16 (32)	16 (69)
2	20 (26)	29 (36)	34 (42)	2 (7)	3 (21)	3 (55)
3	12 (19)	17 (25)	21 (29)	0 (2)	0 (8)	0 (30)
12	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)

NOTE: Results in parentheses refer to the cases in which the parameters have been estimated (T = 100).

parameters have been estimated. Evidently, the result in (13) is valid if T is sufficiently large. The effects of parameter estimation on unconditional prediction-error variances if $m = 1$ have been analyzed, for example, by Yamamoto (1976) and Baillie (1979), who assumed, however, that the parameters are estimated from samples independent of the values to be predicted. Effects on conditional variances, which are outside the scope of this article, were considered by Phillips (1979), Ansley and Kohn (1986), and Hamilton (1986).

If the parameter ρ is unknown, an estimate $\hat{\rho}$ will typically be substituted for ρ in (2), which yields the operational predictor

$$\hat{E}[y_{T+k} \mid y_t, t \in T_m] = \hat{\rho}^{k+r} y_{T-r}. \qquad (14)$$

The corresponding prediction error can be written as the sum of two components of which only the second one depends on the estimation error in the parameter because

$$y_{T+k} - \hat{E}[y_{T+k} \mid y_t, t \in T_m] = \sum_{i=0}^{k+r-1} \rho^i \varepsilon_{T+k-i} - (\hat{\rho}^{k+r} - \rho^{k+r}) y_{T-r}. \qquad (15)$$

The mean squared error (MSE) of the second term can be approximated up to order T^{-1} by

$$E\{(\hat{\rho}^{k+r} - \rho^{k+r}) y_{T-r}\}^2 \sim T^{-1}(k + r)^2 \rho^{2(k+r-1)} E\{\sqrt{T}(\hat{\rho} - \rho) y_{T-r}\}^2. \qquad (16)$$

In the literature, the (unrealistic) assumption is often made that parameters are estimated from samples independent of the sample to be predicted. In that case, we get

$$E\{y_{T+k} - \hat{\rho}^{k+r} y_{T-r}\}^2 \sim E\left\{\sum_{i=0}^{k+r-1} \rho^i \varepsilon_{T+k-i}\right\}^2 + dT^{-1}(k + r)^2 \rho^{2(k+r-1)} E\{\sqrt{T}(\hat{\rho} - \rho)\}^2 E y_{T-r}^2 \qquad (17)$$

when $d = 1$. If the independence assumption is not made, one can bound the covariance between $\sqrt{T}(\hat{\rho} - \rho)$ and y_{T-r} using the Cauchy–Schwarz inequality, which implies that the right side of (17) yields an upper bound for the MSE for $d = 3$ and a lower bound for $d = 0$ if one sample is used both for estimation and prediction.

To evaluate the expectations in (17), it is easily verified that

$$E\left\{\sum_{i=0}^{k+r-1} \rho^i \varepsilon_{T+k-i}\right\}^2 = (1 - \rho^{2k+2r})(1 - \rho^2)^{-1} \sigma_\varepsilon^2 \qquad (18)$$

and that

$$E y_{T-r}^2 = (1 - \rho^2)^{-1} \sigma_\varepsilon^2. \qquad (19)$$

To evaluate $E\{\sqrt{T}(\hat{\rho} - \rho)\}^2$, note that if the high-frequency data are generated by the AR(1) model—that is, if $\alpha = 0$ in (1)—the model for the low-frequency data (4) specializes to become

$$y_t = \psi y_{t-m} + v_t \qquad v_t \sim NID(0, \sigma_v^2) \quad \text{if } t \in T_m, \qquad (20)$$

when $\psi = \rho^m$ and $\sigma_v^2 = (1 - \rho^{2m})(1 - \rho^2)^{-1}\sigma_\varepsilon^2$. The parameter ρ is identified from y_t ($t \in T_m$) if m is odd or if its sign is known a priori. The asymptotic distribution of the maximum likelihood (ML) estimator $\hat{\psi}$ of ψ from data on y_t ($t \in T_m$) is easily shown to be

$$\sqrt{(T/m)}(\hat{\psi} - \psi) \overset{d}{\sim} N(0, 1 - \psi^2), \qquad (21)$$

from which one obtains the asymptotic variance of $\hat{\rho}$

$$E\{\sqrt{T}(\hat{\rho} - \rho)\}^2 \sim \rho^2(1 - \rho^{2m})/\{m\rho^{2m}\}, \qquad (22)$$

where we used $\hat{\rho} = \hat{\psi}^{1/m}$.

Substitution of (18), (19), and (22) into (17) yields upper bounds on the relative efficiency of predictors based on high-frequency data in case of estimated parameters. For given values of m and k, these upper bounds are the maximum over r of the quotient of the right side of (17) evaluated at m, k, r, and $d = 3$ and the same expression evaluated at $m = 1$, k, r, and $d = 0$. The numerical results are presented in Table 1 for the case in which $T = 100$. The information loss caused by temporal aggregation increases if the parameters have to be estimated, but the effect is not very substantial unless ρ is small in absolute value and m is large. Note that in that case ρ is almost unidentified (see Palm and Nijman 1984, table 1).

3.2 Flow Variables

Empirical results on the prediction-accuracy gain from disaggregate sampling if a flow variable is generated by an AR(1) model can be obtained along the same lines. The derivation of analytical expressions for the accuracy gain as in (13) is intricate for flow variables and results in complicated formulas that do not give much insight. Therefore, the results were obtained numerically along the lines described in Appendix A, where we derive the forecast function and prediction-error variance [see (A.11)]. Rewriting (11) as

$$E[y_{T+k} \mid \bar{y}_t, t \in T_m] = \sum_{i=0}^{\infty} a_{ki}\bar{y}_{T-im-r}, \quad (23)$$

upper and lower bounds up to order T^{-1} can be derived along lines similar to (17) using

$$E\left\{y_{T+k} - \sum_{i=0}^{\infty} \hat{a}_{ki}\bar{y}_{T-im-r}\right\}^2$$
$$\sim E\left\{y_{T+k} - \sum_{i=0}^{\infty} a_{ki}\bar{y}_{T-im-r}\right\}^2$$
$$+ dT^{-1}E\{\sqrt{T}(\hat{\rho} - \rho)\}^2$$
$$\times \sum_{i,j=0}^{\infty} \partial a_{ki}/\partial\rho\partial a_{kj}/\partial\rho E\bar{y}_{T-im}\bar{y}_{T-jm}. \quad (24)$$

As before, upper and lower bounds of the MSE for the case in which estimation and prediction are based on the same sample are obtained for $d = 3$ and $d = 0$, respectively, whereas $d = 1$ and $d = 0$ yield results for the case in which the parameters are estimated from samples independent of the sample to be predicted and for the case of known parameters, respectively. Again the determination of the asymptotic variance of the estimator of the unknown parameter is the most difficult part of the evaluation of (24). An expression for $E\{\sqrt{T}(\hat{\rho} - \rho)\}^2$ is derived in Appendix B. This expression has been used to determine the upper bounds on the predictive-accuracy gain for flow variables generated by an AR(1) model in Table 2. Results similar to Table 1 for $T = \infty$ and for $T = 100$ are provided. In the worst possible case considered in Table 2, the variance of the prediction error approximately doubles if low-frequency data are analyzed, but in many cases the loss of information is much smaller. Note also that the variance of the prediction error can be smaller if the data are aggregated over three periods than if they are aggregated over two periods. The estimation of the parameter ρ affects the conclusions only if ρ is highly negative and m is even, in which case it is very difficult to estimate ρ, as becomes evident from table 1 of Palm and Nijman (1984). The effects of parameter estimation on the predictive accuracy gain from disaggregate sampling are different between stock and flow variables, as shown in Tables 1 and 2. For stock variables, the results indicate that for any given values of ρ and k the gain increases with m. For flow variables, this is no longer true when ρ is negative. This is because with m being odd and $\rho < 0$ the observations contain much information about the (alternating) pattern of the series at the disaggregate level and about ρ, which therefore can be estimated with reasonable accuracy from the aggregate data.

Before we turn to the empirical application in Section 4, we present results on the predictive accuracy gain from disaggregate sampling in IMA(1, 1) models and ARI(1, 1) models in Table 3. The predictors for the IMA(1, 1) model ($\Delta y_t = \varepsilon_t - \alpha\varepsilon_{t-1}$) are special cases of (9) and (11). The predictors for the ARI(1, 1) model ($\Delta y_t = \rho\Delta y_{t-1} + \varepsilon_t$) have been derived using the approaches presented in Appendix A. Only results on the impact of parameter estimation on the predictive accuracy gain have been presented for the ARI(1, 1) model for stock variables. For a special IMA(1, 1) model for a flow variable, results will be given in Section 4.

For the nonstationary models in Table 3, the information gain caused by the use of high-frequency data is usually much larger than that for the stationary models in Tables 1 and 2 except for the IMA(1, 1) models with positive coefficient α. For these models, the variance of the error of predictions based on the incomplete data is no longer a nondecreasing function of the number of periods to be predicted ahead. Therefore, the upper bounds are not simply obtained by putting $r = m - 1$. Notice that from the results in the table it appears that the upper bounds (overall values of r) are decreasing in k, not the variances.

Table 2. Upper Bounds in Percentage Points for the Reduction of the Prediction-Error Variance When a Flow Variable Is Generated by an AR(1) Model

k	$m = 2$	$m = 3$	$m = 4$	$m = 2$	$m = 3$	$m = 4$
		$\rho = .8$			$\rho = .4$	
1	43 (44)	54 (55)	59 (60)	15 (16)	16 (17)	16 (17)
2	22 (25)	32 (33)	36 (40)	2 (3)	3 (3)	3 (3)
3	13 (16)	19 (21)	22 (24)	0 (1)	0 (1)	0 (1)
12	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
		$\rho = -.8$			$\rho = -.4$	
1	61 (99)	57 (61)	63 (98)	15 (69)	16 (27)	16 (50)
2	38 (98)	34 (40)	40 (97)	3 (44)	3 (11)	3 (25)
3	24 (97)	21 (28)	25 (95)	0 (19)	0 (3)	0 (19)
12	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)

NOTE: Results in parentheses refer to the case in which the parameters have been estimated (T = 100).

Table 3. Upper Bounds in Percentage Points for the Reduction of the Prediction-Error Variance When Stock of Flow Variables Are Generated by IMA(1, 1) or ARI(1, 1) Models

k	$m = 2$	$m = 3$	$m = 4$	$m = 2$	$m = 3$	$m = 4$
			ARI(1, 1); stock variables			
		$\rho = .8$			$\rho = .4$	
1	79 (80)	92 (92)	96 (96)	67 (68)	82 (83)	88 (88)
2	62 (64)	81 (81)	88 (89)	47 (48)	64 (65)	73 (74)
3	50 (52)	70 (71)	80 (81)	34 (36)	51 (52)	61 (62)
12	14 (17)	25 (25)	33 (33)	9 (9)	16 (16)	22 (22)
		$\rho = -.8$			$\rho = -.4$	
1	58 (67)	54 (53)	64 (63)	29 (98)	50 (59)	60 (98)
2	62 (54)	52 (47)	67 (59)	32 (98)	45 (58)	55 (97)
3	36 (35)	36 (38)	45 (49)	22 (97)	35 (41)	45 (96)
12	15 (8)	15 (15)	24 (20)	8 (15)	15 (15)	20 (25)
			ARI(1, 1); flow variables			
		$\rho = .8$			$\rho = .4$	
1	84	94	97	73	86	91
2	69	85	92	53	70	78
3	57	76	85	39	57	67
12	17	30	39	11	20	27
		$\rho = -.8$			$\rho = -.4$	
1	26	50	53	31	52	63
2	43	48	59	32	47	58
3	13	32	37	22	38	48
12	7	14	19	1	15	22
			IMA(1, 1); stock variables			
		$\alpha = .8$			$\alpha = .4$	
1	9	15	20	29	44	54
2	8	15	19	23	37	46
3	7	11	15	19	32	40
12	7	11	15	7	14	19
		$\alpha = -.8$			$\alpha = -.4$	
1	79	88	91	67	80	86
2	47	62	71	41	58	67
3	33	48	58	30	45	55
12	9	16	22	8	15	21
			IMA(1, 1); flow variables			
		$\alpha = .8$			$\alpha = .4$	
1	4	7	12	29	45	55
2	4	7	11	22	38	48
3	4	7	11	19	32	42
12	3	6	8	7	14	20
		$\alpha = -.8$			$\alpha = -.4$	
1	82	90	93	72	84	88
2	53	68	76	47	64	72
3	39	55	64	35	51	61
12	12	19	26	11	43	25

NOTE: Results in parentheses refer to the case in which the parameters have been estimated.

4. PREDICTION ACCURACY GAIN FROM DISAGGREGATING THE GNP SERIES FOR THE NETHERLANDS

In this section, we illustrate how in practice one can determine whether it is worthwhile to increase the frequency of collecting observations on a variable. We consider the quarterly GNP series for the Netherlands that has recently been provided by the Dutch Central Bank (see De Nederlandsche Bank 1986). Using the results of the previous sections, we show how much the availability of quarterly observations reduces the prediction-error variance of quarterly GNP and whether further disaggregation into monthly data is desirable.

4.1 From Annual to Quarterly Observations

First, we consider seasonally adjusted GNP in millions of guilders in prices of 1980. For the period 1957:1–1984:4, a Box–Jenkins analysis leads us to select two

Table 4. Reduction in Percentage Points of the Prediction-Error Variance of Quarterly Seasonally Adjusted GNP in the Netherlands due to the Use of Quarterly Instead of Annual Observations (r is the number of periods since the last observation)

True model	r	Number of quarters to be predicted ahead (k) 1	2	3	4	8	12
ARI(1, 1) in (25)	0	21	19	13	11	7	5
	1	44	39	30	25	15	11
	2	58	50	41	36	23	17
	3	66	59	49	44	29	22
IMA(1, 1) in (26)	0	17	13	10	8	5	4
	1	39	31	25	22	14	10
	2	51	43	36	32	21	16
	3	60	51	44	39	27	21

quarterly models that are both fairly well in agreement with the information in the data. If a month is chosen as the appropriate time unit, the two models are

$$\Delta_3 \bar{y}_t = \underset{(113)}{543} - \underset{(.09)}{.33}\ \Delta_3 \bar{y}_{t-3} + \hat{v}_t \tag{25}$$

and

$$\Delta_3 \bar{y}_t = \underset{(70)}{408} + \hat{v}_t - \underset{(.09)}{.35}\ \hat{v}_{t-3}, \tag{26}$$

respectively. The parameters have been estimated by ML. Standard errors are given in parentheses. Since $\hat{\rho}$ is approximately $-.4$ and $\hat{\alpha}$ is approximately .4, it is now obvious from Tables 3–5 that the increase in forecast accuracy due to the availability of quarterly instead of annual data can be more than 50%. The details are given in Table 4. Note that here the efficiency strictly increases with r, the number of periods since the last observation. Moreover, even if one is interested in annual forecasts only, quarterly data can be substantially more informative than annual observations.

4.2 From Quarterly to Monthly Observations

The quarterly data can also be used to forecast monthly GNP and to estimate the reduction in the variance of monthly prediction errors if monthly data were indeed collected. The first step is to estimate a monthly model from the quarterly data. As discussed in Section 2, the monthly IMA(1, 1) model

$$\Delta y_t = c + \varepsilon_t - \alpha \varepsilon_{t-1}, \qquad \varepsilon_t \sim NID(0, \sigma_\varepsilon^2), \quad t \in T_1, \tag{27}$$

implies that $\Delta_3 \bar{y}_t = (1 + L + L^2)^2 \Delta y_t$ is generated by the quarterly IMA(1, 1) model

$$\Delta_3 \bar{y}_t = 9c + v_t - \bar{\lambda} v_{t-3}, \qquad v_t \sim NID(0, \sigma_v^2), \quad t \in T_3, \tag{28}$$

where $\bar{\lambda}$ can be determined using

$$\Delta_3 \bar{y}_t = 9c + \sum_{i=0}^{5} \bar{\eta}_i \varepsilon_{t-i} \tag{29}$$

Table 5. The Reduction in Percentage Points of the Prediction-Error Variance of Quarterly Seasonally Adjusted GNP in the Netherlands due to the Use of Monthly Instead of Quarterly Series

Model	r	Number of periods to be predicted ahead (k) 1	2	3	6	9	12
c, α known	0	.8	.8	.7	.6	.5	.4
	1	8.0	7.4	6.9	5.8	5.0	4.5
	2	14.2	13.3	12.5	10.1	8.8	8.2
c, α estimated	0	5.6	6.6	7.8	11.5	15.4	19.2
	1	13.4	14.1	14.8	17.4	20.4	23.5
	2	20.3	20.6	20.9	22.8	25.1	27.6

when $\bar{\eta}_0 = 1$, $\bar{\eta}_1 = 2 - \alpha$, $\bar{\eta}_2 = 3 - 2\alpha$, $\bar{\eta}_3 = 2 - 3\alpha$, $\bar{\eta}_4 = 1 - 2\alpha$, $\bar{\eta}_5 = -\alpha$, and the equality of the first-order autocorrelation of (28) and (29) is

$$-\bar{\lambda}/(1 + \bar{\lambda}^2) = \sum_{i=3}^{5} \bar{\eta}_i \bar{\eta}_{i-3} \Big/ \sum_{i=0}^{5} \bar{\eta}_i^2. \tag{30}$$

Substituting the ML estimate $\hat{\bar{\lambda}} = .35$ for $\bar{\lambda}$ in (30), one gets the ML estimate $\hat{\alpha} = .72$. No other plausible time series model appears to explain the empirical findings in (26), which are almost equivalent to (25).

In Section 2, we have derived the MMSE predictor (11) of a monthly series from quarterly data assuming that the data are generated by (27) when $c = 0$. Along the same lines (the derivations will be made available by the authors on request), one can show that if $c \neq 0$ the MMSE predictor can be written as

$$E[y_{T+k} \mid \bar{y}_t, t \in T_3] = [k + r + 1 + 3\bar{\lambda}/(1 - \bar{\lambda})]c + (1 - \bar{\lambda}) \sum_{i=0}^{\infty} \bar{\lambda}^i \bar{y}_{T-r-3i}/3. \tag{31}$$

The variance of the prediction error of (31) can be compared with that of the optimal predictor from monthly data

$$E[y_{T+k} \mid y_t, t \in T_1] = [k + \alpha/(1 - \alpha)]c + (1 - \alpha) \sum_{i=0}^{\infty} \alpha^i y_{T-i}. \tag{32}$$

The empirical results on the predictive accuracy gain are presented in the first part of Table 5, in which it is assumed that α and c are known a priori to coincide with the ML estimates. This table suggests that monthly data on GNP in the Netherlands would hardly contain more information than the existing quarterly series.

Now we examine whether as a result of the assumption of known parameters the true predictive accuracy gain from increasing the frequency of data collection is substantially underestimated in the first part of Table 5. The relative efficiency of the ML estimator of α in (27) from monthly data compared with that from quarterly data can be found in Palm and Nijman (1984) for various values of α. When the true value of α is .6, the relative efficiency is only 2.7, which suggests that the results in Table 5, in which $\alpha = .72$, should not be too sensitive to the assumption of known parameters. Numerical results on upper bounds for the reduction of the prediction-error variance are given in the second part of Table 5. For the derivation, we refer to Appendix C. From Table 5, it becomes clear that the impact of parameter estimation is not sufficiently large to alter the main conclusion that the prediction-accuracy gain expected from monthly sampling instead of quarterly sampling is low compared to the gain from quarterly observations relative to annual ones. Nevertheless, the change in the predictive accuracy for longer forecast horizons is probably larger than suggested by the results for known parameters.

5. CONCLUDING REMARKS

In this article, we analyzed the predictive accuracy gain of k-step-ahead forecasts from univariate ARIMA models resulting from increasing the frequency of sampling. Results on the expected information gain have been presented for a number of ARIMA models. Subsequently, we evaluated the additional information content of recently collected quarterly GNP data for the Netherlands and considered whether it is worthwhile to construct monthly data.

The main conclusions are as follows:

1. For variables generated by a first-order autoregressive model with known parameters, the information gain is substantial only in short-run forecasting when subsequent realizations are strongly correlated. We conjecture that this result can be extended to more general stationary processes. Note, however, that the information gain can substantially increase if the parameters in the model have to be estimated as suggested in Table 1. Once again, we would like to emphasize that the forecast function of the components that are aggregated over time has to be identified from the aggregate data only. We limited ourselves to the case in which these components are generated by a homogeneous ARMA structure.
2. For variables generated by nonstationary models, the efficiency gain of more frequent sampling can be very large in short-run forecasting but will often be negligible when the forecast horizon becomes large.
3. The results for the GNP series in the Netherlands suggest that the construction of quarterly GNP data has reduced the variance of prediction errors considerably but that further disaggregation into monthly data would hardly yield extra information to forecast that specific series.

Although we limited ourselves to univariate time series models, the results are likely to contain relevant indications for multivariate models, since the variances of the prediction errors for univariate and multivariate models often have similar properties. Finally, since many macroeconomic variables can be adequately described by an IMA(1, 1) process with positive parameter α, the results in this article are expected to be useful when deciding whether more frequent sampling will be worthwhile.

ACKNOWLEDGMENTS

We would like to thank two referees and an associate editor for helpful comments.

APPENDIX A: MMSE PREDICTORS OF HIGH-FREQUENCY SERIES FROM LOW-FREQUENCY DATA FOR GENERAL ARIMA MODELS

In this appendix, we show how classical filtering theory and recursive Kalman filtering can be used to derive the predictors for higher-order ARMA models. For simplicity, we restrict the discussion to ARMA models and to flow variables. Extensions to ARIMA models and/or stock variables are straightforward.

To derive an analytical expression for the MMSE predictor of the high-frequency data if only low-frequency data are available, we start with the high-frequency MA(∞) representation of the data-generating process, which we express in terms of polynomials in L^m operating on ε_{t-i} as

$$y_t = \sum_{i=0}^{\infty} \alpha_i \varepsilon_{t-i} = \sum_{i=0}^{m-1} \omega_i(L^m)\varepsilon_{t-i}, \qquad \text{(A.1)}$$

where $\omega_i(L^m)$ is a polynomial in the lag operator for the low-frequency data; $\omega_i(L^m) = \sum_{j=0}^{\infty} \omega_{ij} L^{jm}$. The corresponding MA($\infty$) expression for $\bar{y}_{t-r}$ is

$$\begin{aligned}\bar{y}_{t-r} &= (1 + L + \cdots + L^{m-1}) \sum_{i=0}^{\infty} \alpha_i \varepsilon_{t-r-i} \\ &= \sum_{i=0}^{m-1} \bar{\omega}_i(L^m)\varepsilon_{t-i}. \end{aligned} \qquad \text{(A.2)}$$

For instance, for the ARMA(1, 1) model (1) when $m = 2$ and $r = 1$, one has $\omega_0(L^2) = (1 + \alpha\rho L^2)/(1 - \rho^2 L^2)$ and $\omega_1(L^2) = (\alpha + \rho)/(1 - \rho^2 L^2)$, which imply $\bar{\omega}_0(L^2) = \{(1 + \alpha + \rho)L^2 + \alpha\rho L^4\}/(1 - \rho^2 L^2)$ and $\bar{\omega}_1(L^2) = \{1 + (\alpha + \rho + \alpha\rho)L^2\}/(1 - \rho^2 L^2)$.

From (A.1) and (A.2), the covariance generating function of $(y_t, \bar{y}_{t-r})$ can easily be shown to be

$$\begin{aligned}&\sigma_\varepsilon^2 \begin{pmatrix} \sum_{i=0}^{m-1} \omega_i(z)\omega_i(z^{-1}) & \sum_{i=0}^{m-1} \omega_i(z)\bar{\omega}_i(z^{-1}) \\ \sum_{i=0}^{m-1} \bar{\omega}_i(z)\omega_i(z^{-1}) & \sum_{i=0}^{m-1} \bar{\omega}_i(z)\bar{\omega}_i(z^{-1}) \end{pmatrix} \\ &= \begin{pmatrix} g_{11}(z) & g_{12}(z) \\ g_{21}(z) & g_{22}(z) \end{pmatrix}. \end{aligned} \qquad \text{(A.3)}$$

Subsequently, factorize $g_{22}(z) = d(z)d(z^{-1})$, where $d(z) = \sum_{i=0}^{\infty} d_i z^i$. A well-known result in classical filtering theory (e.g., see Priestley 1981, chap. 10) now states that

$$E[y_{T+k} \mid \bar{y}_t, t \in T_m] = h_k(L^m)\bar{y}_{T-r} \quad (A.4)$$

with

$$h_k(z) = [z^k d^{-1}(z^{-1})g_{12}(z)]_+ d^{-1}(z), \quad (A.5)$$

where $[\ ]_+$ indicates that only positive powers of z are to be taken into account.

For the case of a variable y_t generated by the AR(1) model obtained if $\alpha = 0$ in (1), observations on $\bar{y}_t$ ($t \in T_2$) and $r = 0$, one can easily check that

$$g_{12}(z) = \sigma_\varepsilon^2\{1 + \rho + \rho^2 + \rho z^{-1}\}/(1 - \rho^2 z)(1 - \rho^2 z^{-1}) \quad (A.6)$$

and

$$\begin{aligned} g_{22}(z) &= \sigma_\varepsilon^2\{2(1 + \rho + \rho^2) + \rho(z + z^{-1})\} \\ &\quad \div (1 - \rho^2 z)(1 - \rho^2 z^{-1}) \\ &= \sigma_\varepsilon \mu(1 - \bar{\lambda} z^{-1})(1 - \bar{\lambda} z) \\ &\quad \div (1 - \rho^2 z^{-1})(1 - \rho^2 z) \end{aligned} \quad (A.7)$$

when $\bar{\lambda} = -\rho^{-1}\{1 + \rho + \rho^2 - (1 + \rho)\sqrt{(1 + \rho)}\}$ and $\mu = -\rho/\bar{\lambda}$ so that

$$\begin{aligned} h_1(z) &= \mu^{-1}\left(\frac{(1 + \rho + \rho^2)z + \rho}{(1 - \rho^2 z)(1 - \bar{\lambda} z^{-1})}\right)_+ \frac{1 - \rho^2 z}{1 - \bar{\lambda} z} \\ &= \frac{\eta}{1 - \bar{\lambda} z} \end{aligned} \quad (A.8)$$

when $\eta = -(1 + \rho + \rho^2 + \rho^3)(1 - \rho^2\bar{\lambda})^{-1}\bar{\lambda} = (1 + \rho)^{-1}(\rho^2 - \bar{\lambda})$. Of course, the predictor obtained in this way coincides with (11) if $k = 1$, $r = 0$, and $m = 2$, as can be easily verified.

The covariance-generating function of the prediction error of the preceding example for $m = 2$ and $r = 0$ is

$$g^*(z) = [1 - \eta(1 - \bar{\lambda} z)^{-1}]g(z)(1 - \eta(1 - \bar{\lambda} z^{-1})^{-1}). \quad (A.9)$$

The variance of the prediction error of y_T in this case, denoted by v, is the constant term in $g^*(z)$, which after some manipulation can be expressed as

$$v = [(1 + \rho^2)^{1/2} - 1]\sigma_\varepsilon^2/\rho^2(1 + \rho). \quad (A.10)$$

For the variance of the prediction error of y_{T+k} if $r \neq 0$, we get

$$\bar{v}_{k,r}^m = [1 - \rho^{2(k+r)}][1 - \rho^2]^{-1} + \rho^{2(k+r)}v, \quad (A.11)$$

from which $\bar{\gamma}_{k,r}^m$ can be determined. The magnitudes $\bar{v}_{k,r}^m$ and $\bar{\gamma}_{k,r}^m$ are defined as $\bar{v}_{k,r}^m$ and $\bar{\gamma}_{k,r}^m$ except that the observations are on flow variables. Equation (A.11) is also valid for $m \geq 2$, although no simple expression for v has been obtained.

Although (A.4) yields an explicit expression for the required MMSE predictor, this expression requires factorization of $g_{22}(z)$ and will not be analytically tractable for higher-order models. In such cases, the recursive Kalman filter is an attractive instrument to compute forecasts of the high-frequency series from low-frequency data. To use the filter in the present context, the ARIMA model has to be written in state-space form with transition equation

$$\xi_t = T\xi_{t-1} + Re_t, \qquad e_t \sim NID(0, \sigma_e^2), \quad (A.12)$$

and measurement equation

$$\begin{aligned} \bar{y}_t &= \bar{Z}\xi_t \quad \text{if } t \in T_m \\ &= 0 \quad \text{if } t \notin T_m. \end{aligned} \quad (A.13)$$

Harvey (1981, p. 103) discussed suitable choices of the vector ξ_t and the matrices T, R, and $\bar{Z}$. We assume that one of the elements of ξ_t coincides with y_t. Harvey (1981) showed that, if

$$a_{t|s} = E[\xi_t \mid \bar{y}_r, r \in T_m, r \leq s] \quad (A.14)$$

and

$$P_{t|s} = E[(\xi_t - a_{t|s})(\xi_t - a_{t|s})' \mid \bar{y}_r, r \in T_m, r \leq s], \quad (A.15)$$

the following recursive prediction equations of the Kalman filter hold true:

$$\begin{aligned} a_{t|t-1} &= Ta_{t-1|t-1} \\ P_{t|t-1} &= TP_{t-1|t-1}T' + \sigma_e^2 RR'. \end{aligned} \quad (A.16)$$

Moreover, if $\bar{y}_t$ is observed, the updating equations of the Kalman filter imply that

$$\begin{aligned} a_{t|t} &= a_{t|t-1} + P_{t|t-1}\bar{Z}'\{\bar{Z}P_{t|t-1}\bar{Z}'\}^{-1}(\bar{y}_t - \bar{Z}a_{t|t-1}) \\ P_{t|t} &= P_{t|t-1} - P_{t|t-1}\bar{Z}'\{\bar{Z}P_{t|t-1}\bar{Z}'\}^{-1}\bar{Z}P_{t|t-1}. \end{aligned} \quad (A.17)$$

If $\bar{y}_t$ is not observed, $a_{t|t} = a_{t|t-1}$ and $P_{t|t} = P_{t|t-1}$. The recursions can be started up by setting $a_{1|0}$ and $P_{1|0}$ equal to the unconditional mean and variance, respectively.

APPENDIX B: DERIVATION OF THE LARGE-SAMPLE VARIANCE OF THE ML ESTIMATOR OF THE PARAMETERS IN AN AR(1) MODEL FROM LOW-FREQUENCY FLOW VARIABLES

In this appendix, we will consider the derivation of the large-sample variance of the ML estimator of the parameter ρ in (1) if $\alpha = 0$ a priori from data on $\bar{y}_t$ ($t \in T_m$). Similar to the derivation given in (20)–(22) for the case of stock variables, the derivation starts with the derivation of an expression for the large-sample variance of the unrestricted ML estimator of the parameters in the low-frequency model. If the ML estimator of $\vartheta = (\psi, \bar{\lambda}, \sigma_v^2)'$ in (4′) is denoted to $\hat{\vartheta}$, it is well known (e.g., see Harvey 1981, p. 132) that $\sqrt{(T/m)}(\hat{\vartheta} - \vartheta) \sim N(0, V)$, where

$$V = \begin{pmatrix} (1 - \psi^2)^{-1} & (1 - \psi\bar{\lambda})^{-1} & 0 \\ (1 - \psi\bar{\lambda})^{-1} & (1 - \bar{\lambda}^2)^{-1} & 0 \\ 0 & 0 & (2\sigma_v^4)^{-1} \end{pmatrix}^{-1}. \quad (B.1)$$

The parameters in ϑ depend on the parameters in the AR(1) model in (1). Denoting $\eta = (\rho, \sigma_\varepsilon^2)$ by ML estimator $\hat{\eta}$, it is easily shown that

$$\sqrt{T}(\hat{\eta} - \eta) \overset{d}{\sim} N(0, m\{\partial\vartheta'/\partial\eta V^{-1}\partial\vartheta/\partial\eta'\}^{-1}). \quad (B.2)$$

Some straightforward algebra finally yields the required result—

$$E\{\sqrt{T}(\hat{\rho} - \rho)\}^2 = m\{m^2\rho^{2m-2}/(1 - \rho^{2m}) - [2m\rho^{m-1}\partial\bar{\lambda}/\partial\rho]/(1 - \rho^m\bar{\lambda}) + (\partial\bar{\lambda}/\partial\rho)^2/(1 - \bar{\lambda}^2)\}^{-1}. \quad (B.3)$$

APPENDIX C: APPROXIMATION OF THE ESTIMATION ERROR WHEN SEVERAL PARAMETERS ARE ESTIMATED

To correct the comparison of the forecast accuracy from aggregate and disaggregate data for the effects of parameter estimation, we have to generalize (24) because, for the IMA(1, 1) model in (27), the MMSE predictor depends on a vector of parameters $\vartheta = (c, \lambda)'$. The variance of the prediction error can be approximated up to order T^{-1} by

$$E[y_{T+k} - \hat{y}_{T+k}(\vartheta)]^2 + T^{-1}\sum_{i,j=1}^{2} \times E(\partial\hat{y}_{T+k}/\partial\vartheta_i)(\partial\hat{y}_{T+k}/\partial\vartheta_j)\sqrt{T}(\hat{\vartheta}_i - \vartheta_i)\sqrt{T}(\hat{\vartheta}_j - \vartheta_j). \quad (C.1)$$

When $\sqrt{T}(\hat{\vartheta} - \vartheta) \overset{d}{\sim} N(0, V)$, the Cauchy–Schwarz inequality can be used to obtain an upper bound for (C.1):

$$E(y_{T+k} - \hat{y}_{T+k}(\vartheta)]^2 + T^{-1}\sum_{i,j=1}^{2} E(\partial\hat{y}_{T+k}/\partial\vartheta_i)(\partial\hat{y}_{T+k}/\partial\vartheta_j)V_{ij} + 2T^{-1}\sum_{i,j=1}^{2} (E(\partial\hat{y}_{T+k}/\partial\vartheta_i)^2 E(\partial\hat{y}_{T+k}/\partial\vartheta_j)^2 V_{ii}V_{jj})^{1/2}, \quad (C.2)$$

where V_{ij} denotes the (i, j)th element of V. In the present case, $V_{11} = \sigma_v^2/27$, $V_{22} = 3(1 - \bar{\lambda}^2)$, and $V_{12} = 0$. Using

$$\partial\hat{y}_{T+k}/\partial\bar{\lambda} = \sum_{i=0}^{\infty} \bar{\lambda}^i v_{T-3i}/3, \quad (C.3)$$

the results in Table 5 can be computed.

[*Received May 1988. Revised March 1990.*]

REFERENCES

Abraham, B., and Ledolter, J. (1982), "Forecast Efficiency of Systematically Sampled Time Series," *Communications in Statistics, Part A—Theory and Methods*, 11, 2857–2868.

Ahsanullah, M., and Wei, W. W. S. (1984), "Effects of Temporal Aggregation on Forecasts in an ARMA(1, 1) Process," in *Proceedings of the Business and Economic Statistics Section, American Statistical Association*, pp. 297–302.

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

Anderson, B. D. O., and Moore, J. B. (1979), *Optimal Filtering*, New York: Prentice-Hall.

Ansley, C. F., and Kohn, R. (1986), "Prediction Mean Squared Error for State Space Models With Estimated Parameters," *Biometrika*, 73, 467–473.

Baillie, R. T. (1979), "Asymptotic Prediction Mean Squared Error for Vector Autoregressive Models," *Biometrika*, 66, 675–678.

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

Chow, G. C., and Lin, A. L. (1971), "Best Linear Unbiased Interpolation, Distribution and Extrapolation of Time Series by Related Series," *Review of Economics and Statistics*, 53, 372–375.

De Nederlandsche Bank (1986), *Kwartaalconfrontatie van Middelen en Bestedingen*, Kluwer, the Netherlands: Deventer.

Engle, R. F., and Liu, T.-C. (1972), "Effects of Aggregation over Time on Dynamic Characteristics of an Econometric Model," in *Econometric Models of Cyclical Behaviors* (Vol. 2), ed. B. G. Hickman, New York: Columbia University Press, pp. 673–737.

Fernandez, R. B. (1981), "A Methodological Note on the Estimation of Time Series," *Review of Economics and Statistics*, 63, 471–476.

Geweke, J. (1978), "Temporal Aggregation in the Multiple Regression Model," *Econometrica*, 46, 643–661.

Hamilton, J. D. (1986), "A Standard Error for the Estimated State Vector of a State Space Model," *Journal of Econometrics*, 33, 387–397.

Harvey, A. C. (1981), *Time Series Models*, Oxford, U.K.: Philip Allan.

Harvey, A. C., and Pierse, R. G. (1984), "Estimation of Missing Observations in Economic Time Series," *Journal of the American Statistical Association*, 79, 125–131.

Kohn, R. (1982), "When Is an Aggregate of a Time Series Efficiently Forecast by Its Past?" *Journal of Econometrics*, 18, 337–350.

Litterman, R. B. (1983), "A Random Walk, Markov Model for the Distribution of Time Series," *Journal of Business and Economic Statistics*, 1, 169–173.

Lütkepohl, H. (1984a), "Linear Transformations of Vector ARMA Processes," *Journal of Econometrics*, 19, 283–293.

——— (1984b), "Forecasting Contemporaneously Aggregated Vector ARMA Processes," *Journal of Business and Economic Statistics*, 2, 201–214.

——— (1986), "Forecasting Vector ARMA With Systematically Missing Observations," *Journal of Business and Economic Statistics*, 4, 375–390.

——— (1987), *Forecasting Aggregated Vector ARMA-Processes* (Lecture Notes in Economics and Mathematical Systems No. 284), Berlin: Springer-Verlag.

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

Nijman, T. E. (1985), *Missing Observations in Dynamic Macroeconomic Modeling*, Amsterdam: Free University Press.

Nijman, T. E., and Palm, F. C. (1986), "The Construction and Use of Approximations for Missing Quarterly Observations: A Model-Based Approach," *Journal of Business and Economic Statistics*, 4, 47–58.

——— (in press), "Parameter Identification in ARMA Processes in the Presence of Regular but Incomplete Sampling," *Journal of Time Series Analysis*.

Palm, F. C., and Nijman, T. E. (1984), "Missing Observations in the Dynamic Regression Model," *Econometrica*, 52, 1415–1435.

Phillips, P. C. B. (1979), "The Sampling Distribution of Forecasts From a First Order Autoregression," *Journal of Econometrics*, 9, 241–261.

Priestley, M. B. (1981), *Spectral Analysis and Time Series* (Vols. 1 and 2), London: Academic Press.

Rose, D. E. (1977), "Forecasting Aggregates of Independent ARIMA-Processes," *Journal of Econometrics*, 5, 325–345.

Tiao, G. C., and Guttman, I. (1980), "Forecasting Contemporal Aggregates of Multiple Time Series," *Journal of Econometrics*, 12, 219–230.

Wei, W. W. S. (1978), "The Effect of Temporal Aggregation on Parameter Estimation in Distributed Lag Models," *Journal of Econometrics*, 8, 237–246.

Weiss, A. A. (1984), "Systematic Sampling and Temporal Aggregation in Time Series Models," *Journal of Econometrics*, 26, 271–281.

Yamamoto, T. (1976), "Asymptotic Mean Squared Error Prediction for an Autoregressive Model With Estimated Coefficients," *Applied Statistics*, 25, 123–127.

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