No. 8101 February 1981

COEFFICIENT BIAS FROM THE OBSERVATION INTERVAL OF A TIME SERIES*

James M. Holmes,
State University of New York at Buffalo

Gary D. Praetzel,
Federal Reserve Bank of Dallas

Donald H. Dutkowsky,
State University of New York at Buffalo

Research Paper

Federal Reserve Bank of Dallas

No. 8101 February 1981

COEFFICIENT BIAS FROM THE OBSERVATION INTERVAL OF A TIME SERIES*

James M. Holmes,
State University of New York at Buffalo

Gary D. Praetzel,
Federal Reserve Bank of Dallas

Donald H. Dutkowsky,
State University of New York at Buffalo

*Reference to this paper in publications should receive clearance from the authors. The authors wish to thank Norman Cameron, Gary Hawes, David Laidler, James Pearce, Ronald Teigen, William White, and especially Cliff Wymer for most helpful comments. All errors, of course, are the responsibility of the authors. The typing work of Michele Slaten and Janice Deering is gratefully acknowledged.

ABSTRACT

This paper reports on the bias in econometric estimators due to data not being measured in the same observation interval (time unit). The conditions necessary for the existence of this bias, and the factors affecting its magnitude are examined theoretically. Empirical evidence concerning the relevance of observation interval bias is illustrated in Goldfeld's [2] "standard" money demand model for various assumptions concerning the trend and scale of the regressors. Predictions of the theoretical bias expressions are found to be consistent with the empirical estimators. In one data construction, the average bias of all the estimators exceeded 130% with some estimates biased over 230%. Since data series are not constructed uniformily, inconsistent observation intervals (time units) in the data can easily arise, and, thus, induce bias in estimators. Clearly, this appears to be a problem applicable to empirical work in all fields; it can arise in any regression estimating a time series relationship.

I. INTRODUCTION

This study seeks to ascertain the nature of the bias affecting coefficient estimators from varying the observaton interval (time unit) of variables in a model and, thus, the information utilized in estimating a time series equation. Our contribution lies in providing a general theoretical framework in which to analyze time unit bias, and an empirical example which demonstrates the relevancy of this bias.

Section II summarizes the theoretical discussion of observation interval (time unit) bias which is formally presented in the Appendix. The model considered is a discrete one in which the true model is comprised of daily observations on the independent variable, Y, and the matrix of regressors, X. For estimation purposes a quarterly model is assumed. The quarterly model is probably subject to an error of approximation in that it is highly unlikely that quarterly intervals coincide with the underlying economic model generating the data. We assume that this approximation error is the same for all the quarterly models estimated so that we can compare the coefficients for time unit bias.

Time unit bias arises in the model from data observed at different points in the quarter. For illustration, data points are assumed to occur either daily or as one observation point at the end of the quarter. This permits the generation of two extreme time unit models: one using quarterly average data; the second using end of quarter point data. It is shown in the Appendix that estimators are unbiased for uniform time unit models, but are biased for models which mix time units.

Section III provides a theoretical examination of the nature of time unit bias which is empirically addressed in section IV using a money demand model. Gibson [1] and Teigen [5][6] addressed the existence of a lagged money stock term by estimating quarterly money demand models in which the money stock is measured as quarterly average data and as end of quarter point data. Goldfeld [2] estimated a quarterly money demand model using three point measures and quarterly averages of the money stock to ascertain if quarterly average measurement of the money stock leads to a relatively more rapid speed of adjustment coefficient. Our results, which generalize time unit bias, are summarized in section V.

Finally, it should be noted that the results presented here are relevant for all fields of research in which estimation of time series equations occur. Economic models, like money demand, often include both stock variables, measured at a point in time, and flow variables, measured over time. Flow variables are like daily average variables in that they are centered at the middle of the interval. But for some discrete series data limitations may not permit the generation of daily average observations. Wealth data, for example, may only be avaiable at a point in the interval. This leads to bias in estimators of some consumption and money demand functions. Capital stock data is not available on a daily average basis, but the labor input is measured over the interval; hence, production function estimates are subject to time unit bias. Clearly, numerous examples can be given in which data availability does not permit the uniform construction of data in a model or where a careless researcher has used poor judgment in collecting data.

II. THE CAUSE OF TIME UNIT BIAS

This section summarizes the results presented formally in the Appendix.

The first two models considered are ones in which the observation intervals are uniform across each equation. Estimators of both the quarterly average model, $GY = GX\beta + G\varepsilon$ (4A), and the point model, $HY = HX\beta + H\varepsilon$ (9A), are shown to be unbiased given the usual regression assumptions.[1]

Biased estimators occur, though, when observation intervals are not uniform for all variables in the model. Equation (11A), $HY = GX\beta + G\varepsilon$, considers a model in which the dependent variable is observed at only end of quarter points, but the regressors are quarterly averages of daily observations. The estimator of this mixed time unit model is shown to be biased and inefficient relative to the estimator of the quarterly average model. The estimator is consistent, though, if in large samples the expected value of the point data equals the expected value of the daily average data. Consistent estimators are not obtainable if the data is subject to trend.

Similar results are obtained in equation (16A), $GY = GX\beta + G\varepsilon$, in which the dependent variable is quarterly averages and the regressors are end of quarter point observations. The estimator is biased and inefficient relative to the point data model estimator, but, likewise, is consistent if the expected value of the point data equals the expected value of the daily

---

[1]The "A" refers to the Appendix.

average data. Interestingly, an unbiased estimator is obtained for this model by actually disregarding some information on the dependent variable by measuring all variables at a point in time.

Equations (21A), $HY = GX\beta + HZ\alpha + H\varepsilon$, and (26A), $HY = GX\beta + GZ\alpha + H\varepsilon$, generalize further by considering partitioned models. Both models measure the dependent variable at end of quarter points, but two groups of regressors are considered for each model. In equation (21A) one group is measured as quarterly averages, the second at end of quarter points; in equation (26A) both groups are quarterly averages. The estimators are biased but consistent under the condition previously discussed. It is shown that the bias worsens for an estimator measured in a time unit different from the dependent variable when more variables of the model are measured in time units inconsistent with the dependent variable.

Secton III assumes time unit bias exists and analyzes those conditions which exacerbate it.

III. THE NATURE OF TIME UNIT BIAS

The behavior of the bias is examined in the Y point-X average model. This model is used to predict time unit bias in coefficients of the money demand model estimated in section IV.

From equation (14A) the bias in the Y point-X average model is given by:

(1) $$-[I_n - (X'G'GX)^{-1} (X'G'HX)]\beta.$$

For expository purposes, consider the special case of this model in which the X matrix contains only one independent variable. In this case time unit bias is given by:

$$(2)\ \beta \left[\frac{[x_1, 91 \sum_{i=1}^{91} x_{1,i} + x_2, 91 \sum_{i=1}^{91} x_{2,i} + \ldots + x_k, 91 \sum_{i=1}^{91} x_{k,i}]}{[(\frac{1}{91} \sum_{j=1}^{91} x_{1,j}) \sum_{i=1}^{91} x_{1,i} + (\frac{1}{91} \sum_{j=1}^{91} x_{2,j}) \sum_{i=1}^{91} x_{2,i} + \ldots + (\frac{1}{91} \sum_{j=1}^{91} x_{k,j}) \sum_{i=1}^{91} x_{k,i}]} - 1 \right].$$

The first subscript represents the quarter and the second subscript represents the day in the quarter. For example, $x_{1,91}$ is the point observation of x at the last day in the first quarter (91st day); hence, the average of daily observations of x for the first quarter is $\frac{1}{91} \sum_{j=1}^{91} x_{1,j}$.

The coefficient estimator of this model is, of course, biased since the time unit of the regressor is inconsistent with the time unit of the dependent variable; but, interestingly, the magnitude of the bias depends solely on the nature of the independent variable. In particular, the greater the trend in the regressor, the greater the bias in the estimator. It is apparent that if Y are end of quarter point data, and if X is a positive (negative) trend variable, then end of quarter point observations consistently exceed (are less than) daily average observations over the quarter, and, consequently, bias the estimator upward (downward) in absolute terms. If the Y data are selected at the beginning of the quarter, and if X is a positive (negative) trend variable, then beginning of the quarter point observations are consistently less than (exceed) daily average observations over the quarter, and, consequently, bias the estimator downward (upward) in absolute terms. If the Y data is chosen at the middle of the period, and if the X data is of either positive or negative uniform trend over all the quarters, there exists relatively little bias in the estimator.

The general expression for the bias in the Y point-X average model is:

$$\text{(3)} \qquad \begin{aligned} \text{Bias} &= \{[(GX)'(GX)]^{-1}(GX)'(HX) - I_n\}\beta \\ &= \{[(GX)'(GX)]^{-1}(GX)'(HX - GX)\}\beta \\ &= \frac{1}{\sigma^2} \text{Var}\,(\tilde{\beta})V\beta \end{aligned}$$

where

$$\text{(4)} \qquad V \equiv (GX)'(HX - GX).$$

The time unit bias of the k^{th} coefficient in an equation is given by:

$$\text{(5)} \qquad \text{Bias } \beta^*_k = \frac{1}{\sigma^2} \sum_{i=1}^{n} \beta_i \sum_{j=1}^{n} \text{Cov}(\tilde{\beta}_k, \tilde{\beta}_j)V_{ji}.$$

The bias of the k^{th} coefficient depends on all the β_i's (true coefficients), the covariances of $\tilde{\beta}_k$ (k^{th} estimator of the quarterly average model) with all the estimators of the daily average model, and all the elements of matrix V. Clearly, there are offsetting effects on the bias due to differences in signs of the β_i's, and in the covariances of regression coefficients in the daily average model. This leads one to focus on large V_{ji}'s in predicting the bias.[2]

The matrix V, by construction, is a square nxn matrix of the average daily observations (GX) of the n independent variables times the measurement error caused by measuring X at one point in the interval (HX) rather than using daily averages. Attention in predicting time unit bias logically focuses on those independent variables which display a strong trend, and are measured in large units.

In section IV, empirical evidence is given to demonstrate the potential importance of time unit bias by considering differing degrees of time unit mixing, trend, and scale in a money demand model.

IV. TIME UNIT BIAS: AN APPLICATION

The potential relevance of time unit bias is considered using Goldfeld's [2] money demand model in which he explicity addressed this question. Goldfeld's sole concern, though, was to test Gibson's [1] contention that quarterly average money stock data relative to end of quarter point data leads to a faster speed of adjustment of actual money balances to their desired level. Goldfeld estimated a partial adjustment quarterly money demand model in which real money demand, M/P, is a function of itself lagged, $(M/P)_{-1}$, real GNP, GNP/P, the interest rate on commercial paper, i^{cp}, and the interest rate on time deposits, i^{td}. Both interest

---

[2]It is interesting to note that even in the case of orthogonal independent variables (i.e., $Cov(\tilde{\beta}_k, \tilde{\beta}_j) = 0$ for $k \neq j$), bias exists.

rates are measured as daily averages. Income, being a flow variable, is also centered at the middle of the quarter. For the M-1 money variable, we employ the four observation intervals used by Goldfeld.

The four observation intervals for the quarterly money stock are: daily average data in equation A; daily averages for only the last month of the quarter in equation B; the average of daily observations for the last month in the current quarter averaged with the average of daily observations for the first month of the subsequent quarter in equation C; and end of quarter call figures in equation D. Equation A represents an unbiased time unit of measurement model since all variables are measured uniformly, but inconsistent time measurement units are represented in models B, C, and D. Time unit bias is clearly worse in models C and D relative to B if there are pronounced trends in GNP and interest rates. Logically, equation A represents the preferred model since one desires to explain money demand over the quarter, and it can best be explained by movements in its determinants over the quarter.

Equations B, C, and D are variants of the partitioned model $HY = GX\beta + HZ\alpha + H\varepsilon$ (21A). It should be noted that the bias in the estimator of β in the partitioned model equals the bias in the estimator of β in the Y point-X average model, $HY = GX\beta + H\varepsilon$ (11A), if N_2, $I_k - HZ[(HZ)'(HZ)]^{-1}Z'H'$ (see 21A, 22A, 12A), equals the identity matrix. This result is easy to see if Z is a one column matrix. As the number of observations increase, the second term in N_2 becomes smaller and approaches zero as the number of observations become very large. Since our sample contains 83 observations, the general bias expression derived from the Y point-X average model approximates well the bias existing in the quarterly average estimators of the partitioned model.

In general, the bias in the k^{th} daily average coefficient is approximately:

(6) $$\text{Bias } \beta_k^* = \frac{1}{\sigma^2} \sum_{i=1}^{3} \beta_i \sum_{j=1}^{3} \text{Cov}(\tilde{\beta}_k, \tilde{\beta}_j) V_{ji}, \quad k = 1,2,3$$

where

x^1 is the time deposit rate, i^{td};

x^2 is the commercial paper rate, i^{cp};

x^3 is real GNP, GNP/P;

and ~ and * refer to the daily average model and the Y point-X average model, respectively.

The dominant V_{ji} term, which is important in predicting the bias, comprises the variable, X_j, measured in the largest absolute scale, and the variable, X_i, having the greatest trend. The choice is GNP on both counts since, numerically, it greatly exceeds interest rates, and it trends strongly upward over the period.

Given that V_{33} is the dominant term in affecting coefficient bias, the direction of the bias for the interest rate and GNP coefficients is:

(7) $$\text{Bias } \beta_1^* = \underset{(+)}{\frac{1}{\sigma^2}} \underset{(+)}{\beta_3} \underset{(-)}{\text{Cov}(\tilde{\beta}_1, \tilde{\beta}_3)} \underset{(+)}{V_{33}} < 0;$$

(8) $$\text{Bias } \beta_2^* = \underset{(+)}{\frac{1}{\sigma^2}} \underset{(+)}{\beta_3} \underset{(-)}{\text{Cov}(\tilde{\beta}_2, \tilde{\beta}_3)} \underset{(+)}{V_{33}} < 0;$$

$$\text{(9)} \qquad \text{Bias } \beta_3^* = \underset{(+)}{\frac{1}{\sigma^2}} \underset{(+)}{\beta_3} \underset{(+)}{\text{Cov}(\tilde{\beta}_3, \tilde{\beta}_3)} \underset{(+)}{V_{33}} > 0;^3$$

Therefore, if the money variable is measured at end of quarter points the two interest rate coefficients are biased downward, and the GNP/P coefficient biased upward, with the bias increasing as trend increases in the regressors.

Since the scale of the variables may be important in affecting time unit bias, Goldfeld's real money demand model is estimated in Table I in both log-linear and linear functional forms. In the log-linear equation, the coefficients of both interest rates and GNP are consistent with the prediction of the general bias expression since the quarterly average model yields the largest interest rate coefficients, and the smallest GNP/P estimate. In the linear model, the estimates of i^{cp} and GNP/P are biased in the expected manner; however, the quarterly average estimate of i^{td} is not the largest time deposit estimate.

In general, it appears that trend may cause substantial bias in estimators when the data is not time unit consistent. This is gleaned by examining the percentage difference of the biased model coefficient from the quarterly average coefficient. The estimate of the percentage bias reaches a particularly large magnitude in the point estimate model.

---

[3] The sign of the GNP coefficient is expected to be positive, reflecting an increase in money demand occurring when real transactions rise, while the interest rate coefficients are expected to be negative, reflecting an increase in money balances occurring when the opportunity cost (foregone earnings) of holding money decreases. The signs of the covariance matrix of estimated coefficients are from the daily average model.

TABLE I

Comparison of Log-Linear and Linear Functional Forms for Goldfeld's Real Money Demand Model*

LOG-LINEAR FUNCTIONAL FORM

Money Measure	log i^{td}	log i^{cp}	log GNP/P	log $(M/P)_{-1}$	$\hat{\rho}$
(A) Quarterly Average	-.043	-.018	.178	.679	.39
(B) Last Month of Quarter	-.052	-.021	.214	.624	.17
(C) Two Month Average Centered on End of Quarter	-.048	-.022	.207	.642	.29
(D) Point Estimate	-.098	-.025	.328	.347	0
Expected Bias of (B,C,D) Relative to (A)	(-)	(-)	(+)		
Estimated Percentage Bias (B)	20.93%	16.67%	20.22%		
(C)	11.63%	22.22%	16.29%		
(D)	127.91%	38.89%	84.27%		

LINEAR FUNCTIONAL FORM

Money Measure	log i^{td}	log i^{cp}	log GNP/P	log $(M/P)_{-1}$	$\hat{\rho}$
(A) Quarterly Average	-2.08	-.864	.030	.816	.38
(B) Last Month of Quarter	-2.00	-.977	.031	.827	.14
(C) Two Month Average Centered on End of Quarter	-2.13	-1.04	.033	.814	.27
(D) Point Estimate	-7.00	-1.37	.074	.471	0
Expected Bias of (B,C,D) Relative to (A)	(-)	(-)	(+)		
Estimated Percentage Bias (B)	-3.85%	13.08%	3.33%		
(C)	2.40%	20.44%	10.00%		
(D)	236.54%	58.56%	146.67%		

*The sample period is 1952:2-1972:4. The data sources are: i^{td}, from the Federal Reserve-MIT-Pennsylvania Model supplied by Stephen Goldfeld; i^{cp}, from Banking and Monetary Statistics, 1941-1970 and various Federal Reserve Bulletins; GNP and its implicit price deflator, P, from Business Statistics, 1977; the first three money stock measures, from Business Statistics; and the last money measure, from various Federal Reserve Bulletins. The estimated 1st order autocorrelation coefficeint is $\hat{\rho}$.

Concerning the role of scale, the evidence is less clear cut. Estimates of the percentage bias in the linear functional form are not generally in excess of the estimated percentage bias of coefficients in the more compactly measure log-linear functional form. However, in the point model which exhibits substantial time unit bias, the estimated percentage bias is significantly greater in the linear model.

To further examine the role of trend and scale in affecting time unit bias, consider the nominal version of the money demand function in Table II. Since deflation of GNP by the price level reduces both the scale and degree of trend in this regressor, one expects greater time unit bias in a nominal money demand model relative to a real money demand model. Furthermore, time unit bias should be greater in the linear version of this model relative to the log-linear form.

In the estimation of the nominal model the price level term is explicitly included even though, theoretically, its coefficient should not be significantly different from zero. Since our sole purpose is to ascertain the potential importance of time unit bias, this seems justifiable. From equation (9) it follows that the price coefficient, β_4^*, is biased upward since:

$$\text{(10)} \qquad \text{Bias } \beta_4^* = \underset{(+)(+)}{\frac{1}{\sigma^2}\beta_3} \; \underset{(+)\;(+)}{\text{Cov}(\tilde{\beta}_4, \tilde{\beta}_3)V_{33}} > 0.$$

The estimates of the nominal money demand model, for the four time unit measures, are reported in Table II for each functional form. All quarterly average coefficients of the point money stock equations are

TABLE II

Comparison of Log-Linear and Linear Functional Forms for Goldfeld's Nominal Money Demand Model

LOG-LINEAR FUNCTIONAL FORM

Money Measure	log i^{td}	log i^{cp}	log GNP	log P	log $(M)_{-1}$	$\hat{\rho}$
(A) Quarterly Average	-.042	-.018	.193	.098	.686	.37
(B) Last Month of Quarter	-.052	-.022	.229	.126	.624	.16
(C) Two Month Average Centered on End of Quarter	-.048	-.023	.222	.116	.643	.28
(D) Point Estimate	-.105	-.027	.384	.256	.310	0
Expected Bias of (B,C,D) Relative to (A)	(-)	(-)	(+)	(+)		
Estimated Percentage Bias						
(B)	23.81%	22.22%	18.65%	28.57%		
(C)	14.29%	27.78%	15.03%	18.37%		
(D)	150.00%	50.00%	98.97%	161.22%		

LOG-LINEAR FUNCTIONAL FORM

Money Measure	log i^{td}	log i^{cp}	log GNP	log P	log $(M)_{-1}$	$\hat{\rho}$
(A) Quarterly Average	-1.71	-.597	.053	9.76	.672	.37
(B) Last Month of Quarter	-2.03	-.715	.061	11.86	.626	.17
(C) Two Month Average Centered on End of Quarter	-1.99	-.777	.062	10.61	.627	.31
(D) Point Estimate	-5.43	-1.07	.132	25.66	.115	0
Expected Bias of (B,C,D) Relative to (A)	(-)	(-)	(+)	(+)		
Estimated Percentage Bias						
(B)	18.61%	19.77%	15.09%	21.49%		
(C)	15.93%	30.15%	16.98%	8.69%		
(D)	216.86%	79.90%	149.06%	162.86%		

biased in the direction predicted by the general bias expression. In particular, the interest rate coefficients of the quarterly average model are negative and closest to zero; the GNP and price coefficients are positive and smallest in the daily average model. Observe, in particular, the large bias incurred by using end of quarter point data. These results demonstrate again the importance of trend in biasing coefficient estimates when data is not measured in uniform time units.

Concerning the relevancy of scale in time unit bias, Table II indicates once again that the larger scale adds to the bias in the point estimate model--the model which is most affected by time unit bias. At the same time the estimated percentage bias is greater for most coefficients in Table II relative to Table I.

Of particular interest is the estimate of the bias when the dependent variable is measured as end of quarter call dates, and the independent variables are measured as averages of daily observations over the quarter. Table III reports the percentage bias of the coefficient estimates from using end of quarter point data for each functional form.

The results indicate that the use of end of period call data for the money stock, rather than daily averages, causes highly biased parameter estimates regardless of functional form. In particular, eight of the fourteen estimates are biased in excess of 100 percent of the unbiased estimate with the average percentage bias in three of the four variables exceeding this figure. This case demonstrates the possibility of bias of a serious magnitude if estimation is done with mismatched data.

Table III

Estimated Bias in Percentage Terms from Measuring the Money Stock at End of Period Call Dates Relative to Using Daily Average Money Stock Data

Functional Form	i^{td}	i^{cp}	GNP	P
Real Log-Linear	127.91%	38.89%	84.27%	--
Real Linear	236.54%	58.56%	146.67%	--
Nominal Log-Linear	150.00%	50.00%	98.97%	161.22%
Nominal Linear	216.86%	79.90%	149.06%	162.86%
Average Percentage Bias	182.83%	56.84%	119.74%	162.04%
Overall Average Percentage Bias	130.36%			

Finally, our bias expression is not applicable for examining the bias of the lagged dependent estimator. In three of the four models estimated, the slowest speed of adjustment occurs in the quarterly average model. Obtaining an unbiased estimate of the adjustment coefficient is important for policy considerations in that it provides an estimate of interest rate variability facing the Federal Reserve when hitting particular money targets.

V. CONCLUSION

This paper has delineated the cause and nature of bias in estimators if data is measured in nonuniform time units. In the general theoretical case it has been shown that for the Y point-X average model time unit bias has been found to depend on all the β_i's, the covariances between the estimated coefficient under examination and all remaining estimators of the daily average model, the magnitude of the regressors, and the degree of measurement error in the regressors.

The money demand function has been used to illustrate time unit bias. Money demand seems appropriate since commonly specified regressors such as interest rates and GNP have exhibited an upward trend over the post war era. Also, the function is commonly estimated in both real and nominal terms using in most instances either a log-linear or linear functional form. Equations were estimated under these alternatives for examing the role of trend and scale in affecting time unit bias.

The money demand estimates reveal that the presence of trend and nonuniform observation intervals may cause substantial bias in estimators. With regard to scale, the results are less clear cut, but did reveal that in estimates of the model most affected by time unit bias the larger scaled specification exhibited a much higher magnitude of time unit bias relative to the more compactly scaled equation.

VI. APPENDIX

The analysis begins by specifying a discrete daily model. From this daily model one derives daily average and point data models. Assuming a quarterly model, there are 91 observations per quarter on each variable where Y is the vector of daily observations on the dependent variable and X is the matrix of daily observations on the n independent variables. In matrix form:

(1A)

$$Y \equiv \begin{bmatrix} y_{1,1} \\ \vdots \\ y_{1,91} \\ \vdots \\ y_{k,1} \\ \vdots \\ y_{k,91} \end{bmatrix}_{(t\times 1)} ; \quad X \equiv \begin{bmatrix} x^1_{1,1} & x^2_{1,1} & \cdots & x^n_{1,1} \\ \vdots & \vdots & & \vdots \\ x^1_{1,91} & x^2_{1,91} & \cdots & x^n_{1,91} \\ \vdots & \vdots & & \vdots \\ x^1_{k,1} & x^2_{k,1} & \cdots & x^n_{k,1} \\ \vdots & \vdots & & \vdots \\ x^1_{k,91} & x^2_{k,91} & \cdots & x^n_{k,91} \end{bmatrix}_{(t\times n)}$$

where the first subscript refers to the quarter, the second subscript refers to the day, and the superscripts denote different independent variables.

The average data model is generated from daily observations by a grouping matrix, G, such that

(2A)

$$G \equiv \begin{bmatrix} 1/91,1/91,\ldots,1/91,\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 0 \\ 0\ldots\ldots\ldots\ldots 0,1/91,1/91,\ldots,1/91,0\ldots\ldots\ldots\ldots 0 \\ \quad 0\text{'s} \qquad\qquad \ddots \quad 1/91,1/91,\ldots,1/91 \end{bmatrix}_{(k\times t)}$$

where $k(=t/91)$ is the number of quarters, t is the number of daily observations, and each 1/91 segment runs for a quarter's length. Premultiplying both Y and X by G yields a model of daily observations averaged over the quarter for both the dependent and independent variables such that:

(3A)

$$GX = 1/91 \begin{bmatrix} \sum_{i=1}^{91} x^1_{1,1} & \sum_{i=1}^{91} x^2_{1,i} & \cdots & \sum_{i=1}^{91} x^n_{1,i} \\ \vdots & \vdots & & \vdots \\ \sum_{i=1}^{91} x^1_{k,i} & \sum_{i=1}^{91} x^2_{k,i} & \cdots & \sum_{i=1}^{91} x^n_{k,i} \end{bmatrix}_{(k \times n)} ; \quad GY = 1/91 \begin{bmatrix} \sum_{i=1}^{91} y_{1,i} \\ \vdots \\ \sum_{i=1}^{91} y_{k,i} \end{bmatrix}_{(k \times 1)}.$$

The quarterly average model, in matrix form, is:

(4A) $$GY = GX\beta + G\varepsilon, \ \varepsilon \sim wn(\sigma^2)$$

where ε is a tx1 disturbance vector generated also from daily observations. Estimating the quarterly average model by OLS yields as the estimator of β:

(5A) $$\tilde{\beta} = [(GX)'(GX)]^{-1}(GX)'GY.$$ [4]

The expected value of β is:

(6A) $$\begin{aligned} E(\tilde{\beta}) &= E[(X'G'GX)^{-1}X'G'GY] \\ &= E[(X'G'GX)^{-1}X'G'G(X\beta + \varepsilon)] \\ &= \beta. \end{aligned}$$

Under the usual regression assumptions, the estimator of β in the quarterly average model is unbiased, assuming the daily model is the true model, with variance $\sigma^2(X'G'GX)^{-1}$. Since all the data of this model are generated from daily observations in a uniform manner, no time unit bias exists.

---

[4] Assuming a homoscedastic diagonal disturbance matrix the GLS and OLS estimators are identical.

An extreme opposite model in terms of information utilized, but, also containing uniform time units, is the consistent point data model. In this model, observations of both Y and X are generated from daily observations as end of quarter point figures.[5] The generation of such data from daily observations is accomplished by the grouping matrix, H, which, when postmultiplied by Y and X, yields end of quarter observations as data points. This H matrix is of the form:

(7A)

$$H \equiv \begin{bmatrix} 0.....0,1,0.......................0 \\ 0.......0,0.....0,1,0................0 \\ \text{0's} \qquad \ddots \qquad \text{0's} \\ \qquad\qquad\qquad\qquad 0.....0,1 \end{bmatrix}_{(kxt)}$$

where the 1's are placed in multiples of the 91st column. The data matrices for the uniform point data model are:

(8A)

$$HX = \begin{bmatrix} x^1_{1,91} & x^2_{1,91} & \cdots & x^n_{1,91} \\ \vdots & \vdots & & \vdots \\ x^1_{k,91} & x^2_{k,91} & \cdots & x^n_{k,91} \end{bmatrix}_{(kxn)} ; \quad HY = \begin{bmatrix} y_{1,91} \\ \vdots \\ y_{k,91} \end{bmatrix}_{(kx1)}.$$

Estimating the consistent point data model,

(9A) $$HY = HX\beta + H\varepsilon, \ \varepsilon \sim wn(\sigma^2),$$

by OLS yields:

(10A) $$\tilde{\tilde{\beta}} = [(HX)'(HX)]^{-1}(HX)'HY.$$

---

[5]The discussion is general in the sense that point observations can be selected at any point in the interval. End of period data is common for monetary statistics.

Substituting the true data model for Y and taking expectations, yields $E(\tilde{\tilde{\beta}}) = \beta$. The estimator of β is unbiased for this consistent time unit model, using the daily model as the true model, with variance $\sigma^2(X'H'HX)^{-1}$. Once again, time unit bias is not present if all data is measured in the same time unit.

As an illustration of time unit bias, consider a mixed model in which the dependent variable is measured at the end of the quarter and the regressors measured as averages over the quarter such that:

(11A) $$HY = GX\beta + H\varepsilon, \ \varepsilon \sim wn(\sigma^2).$$

The OLS estimator of β is:

(12A) $$\beta^* = [(GX)'(GX)]^{-1}(GX)'(HY).$$

The expected value of β^* is:

(13A) $$\begin{aligned} E(\beta^*) &= E[(X'G'GX)^{-1}X'G'HY] \\ &= E[(X'G'GX)^{-1}X'G'H(X\beta + \varepsilon)] \\ &= (X'G'GX)^{-1}X'G'HX\beta. \end{aligned}$$

Since the expected value of β^* does not equal β, the estimator is biased. An expression for the bias is:

(14A) $$[(X'G'GX)^{-1}X'G'HX - I_n]\beta.$$

The mean squared error of the biased estimator, β^*, is:

(15A) $$\begin{aligned} MS(\beta^*) &= E[(\beta^* - \beta)(\beta^* - \beta)'] \\ &= E\{[(X'G'GX)^{-1}X'G'HY-\beta)][(X'G'GX)^{-1}X'G'HY-\beta)]'\} \\ &= [(X'G'GX)^{-1}(X'G'HX)-I_n]\beta\beta'[(X'G'GX)^{-1}(X'G'HX)-I_n]' \\ &\quad +\sigma^2(X'G'GX)^{-1} \end{aligned}$$

which equals the mean squared error (variance) of the unbiased estimator of the quarterly average model plus a positive definite matrix involving the bias of the mixed estimator. Therefore, the estimator of the mixed model is not only biased, but also inefficient compared to the estimator of the quarterly average model.

A second mixed model contains Y measured at the end of each quarter and regressors that are daily averages of observations over the quarter such that:

(16A) $$GY = HX\beta + G\varepsilon, \quad \varepsilon \sim wn(\sigma^2).$$

Applying OLS the estimator of β is:

(17A) $$\beta^{**} = [(HX)'(HX)]^{-1}(HX)'GY.$$

The expected value of β^{**} is:

(18A) $$E(\beta^{**}) = (X'H'HX)^{-1}X'H'GX\beta.$$

The estimator of this mixed model is again biased with the bias equal to:

(19A) $$[(X'H'HX)^{-1}X'H'GX - I_n]\beta.$$

Since β^{**} is a biased estimator of β, the mean squared error is:

(20A) $$\begin{aligned} MS(\beta^{**}) &= E[(\beta^{**} - \beta)(\beta^{**} - \beta)'] \\ &= E\{[((HX)'(HX))^{-1}(HX)'GY-\beta]\,[((HX)'(HX))^{-1}(HX)'GY-\beta]'\} \\ &= [(X'H'HX)^{-1}(X'H'GX)-I_n]\beta\beta'[(X'H'HX)^{-1}(X'H'GX)-I_n] \\ &\quad + \sigma^2(X'H'HX)^{-1} \end{aligned}$$

which, analogous to the first mixed model, equals the variance of the estimator of the consistent point data model plus a positive definite matrix involving the bias of the mixed estimator. Ironically, the estimator from the complete point model is preferred to the estimator of this model even though all information on the dependent variable is not utilized, since the mixed model estimator is biased and less efficient than the complete point data model. Unbiased and efficient estimation is obtained for this model by actually disregarding some information on the dependent variable which, certainly, is a nonintuitive result.

In summary, it has been established that if the estimated model uses data constructed in the same time unit no time unit bias exists. But if the model is mixed such that the variables are not constructed in the same manner from daily observations, then coefficient estimates are biased and of greater mean squared error relative to consistently measured models. These bias results concerning mixed model estimators assume that all the regressors are measured in the same time unit. This assumption is relaxed in the more general partitioned models.

Consider a partitioned model such that the independent variable, Y, and a group of regressors, Z, are measured as of the end of the quarter while the remaining regressors, X, are measured as daily averages over the quarter. In matrix notation such a model is:

(21A) $$HY = GX\beta + HZ\alpha + H\varepsilon, \ \varepsilon \sim wn(_{\sigma}2)$$

$$= [GX \vdots HZ] \left[\begin{matrix}\beta \\ \vdots \\ \alpha\end{matrix}\right] + H\varepsilon.$$

Define matrix N_1 as $I_k - GX[(GX)'(GX)]^{-1}X'G'$, and matrix N_2 as $I_k - HZ[(HZ)'(HZ)]^{-1}Z'H'$ such that $N_1GX = 0$ and $N_2HZ = 0$. The estimator of in this first partitioned model is

(22A) $$\hat{\beta} = [(GX)'N_2(GX)]^{-1}[(GX)'N_2HY].$$

Replacing Y by the true model, $Y = X\beta + Z\alpha + \varepsilon$, and taking expectations yields:

(23A) $$E(\hat{\beta}) = [(GX)'N_2(GX)]^{-1}[(GX)'N_2(HX)]\beta.$$

The estimator of β is biased, but is consistent if in large samples the expected value of the point data is equal to the expected value of the average data. The estimator of α, $\hat{\alpha}$, is also biased, but, similarly, it is consistent if in large samples the expected value of the point data equals the expected value of the average data such that HX = GX. The estimator of α and its expected value are given by:

(24A) $$\alpha = [(HZ)'N_1(HZ)]^{-1}[(HZ)'N_1HY];$$

$$E(\hat{\alpha}) = [(HZ)'N_1(HZ)]^{-1}[(HZ)'N_1H(X\beta + Z\alpha + \varepsilon)]$$

(25A) $$= \alpha + [(HZ)'N_1(HZ)]^{-1} \cdot [(HZ)'N_1(HX)]\beta.$$

For comparison, consider a second partitioned model in which both groups of regressors are measured as averages of daily observations generating the model:

(26A) $$HY = GX\beta + GZ\alpha + \varepsilon, \varepsilon \sim wn(\sigma^2).$$

Define N_1 as before and N_2^1 as I_k - $GZ[(GZ)'(GZ)]^{-1}(GZ)'$ such that $N_1GX = 0$ and $N_2^1GZ = 0$. The estimator of β in this second partitioned model is:

(27A) $$\hat{\hat{\beta}} = [(GX)'N_2^1(GX)]^{-1}[(GX)'N_2^1HY].$$

Substituting the true daily model for Y, and taking expected values yields:

(28A) $$E(\hat{\hat{\beta}}) = [(GX)'N_2^1(GX)]^{-1}[(GX)'N_2^1(HX)]\beta+[(GX)'N_2^1(GX)]^{-1}[(GX)'N_2^1HZ]\alpha \neq \beta.$$

As in the first partitioned model considered, $\hat{\hat{\beta}}$ is consistent if H = G in large samples. Observe that the bias in the estimator of β in the second partitioned model exceeds the bias in the first partitioned model which means that the bias increases with an increase in the number of variables measured in a time unit different from the dependent variable.

Turning to α, the estimator of α and its expected value are given by:

(29A) $$\hat{\hat{\alpha}} = [(GZ)'N_1(GZ)]^{-1}[(GZ)'N_1HY];$$

(30A) $$E(\hat{\hat{\alpha}}) = [(GZ)'N_1(GZ]^{-1}[(GZ)'N_1H][X\beta + Z\alpha].$$

Once again, if H = G in large samples then α is consistent. Algebraically, one cannot say if the bias in the estimator of the second partitioned model exceeds the bias in the first partitioned model, but intuitively, it is expected since a larger group of variables in the second case are now measured in a time unit inconsistent with the dependent variable.

REFERENCES

[1] Gibson, William: "Demand and Supply Functions for Money in the United States: Theory and Measurement," Econometrica, 40 (1972), 361-370.

[2] Goldfeld, Stephen: "The Demand for Money Revisited," in Brookings Papers on Economic Activity 3:1973, 577-638.

[3] Griliches, Zvi: "A Note on the Serial Correlation Bias in Estimates of Distributed Lags," Econometrica, 29 (1961), 65-73.

[4] Teigen, Ronald: "Demand and Supply Functions for Money in the United States: Some Structural Estimates," Econometrica, 32 (1964), 476-509.

[5] ________________: "Demand and Supply Functions for Money: Another Look at Theory and Measurement," Econometrica, 44 (1976), 377-385.

[6] Theil, Henri: Principles of Econometrica. New York: John Wiley and Sons, 1971.

[7] White, William: "Improving the Demand for Money Function in Moderate Inflation," International Monetary Fund Staff Papers, 25 (1978), 564-607.

[8] Wymer, Cliff: "Continuous Time Models in Macro-Economics: Specification and Estimation," a paper prepared for the SSRC-Ford Foundation Conference on "Macroeconomic Policy and Adjustment in Open Economies" at Fanhams Hall, Ware, England, April 1976.