Bias correction of OLSE in the regression model with lagged dependent variables

Tanizaki, Hisashi

(Citation)
Computational Statistics & Data Analysis,34(4):495-511

(Issue Date)
2000-10

(Resource Type)
journal article

(Version)
Accepted Manuscript

(URL)
https://hdl.handle.net/20.500.14094/90000573

Bias Correction of OLSE in the Regression Model with Lagged Dependent Variables*

Hisashi Tanizaki
Faculty of Economics, Kobe University
Nadaku, Kobe 657-8501, Japan
(tanizaki@kobe-u.ac.jp)

Key Words: AR(p) Model, OLSE, Median-Unbiased Estimator, Mean-Unbiased Estimator, Lagged Dependent Variable, Exogenous Variables, Nonnormal Error.

Abstract

It is well known that the ordinary least squares estimates (OLSE) of autoregressive models are biased in small sample. In this paper, an attempt is made to obtain the unbiased estimates in the sense of median or mean. Using Monte Carlo simulation techniques, we extend the median-unbiased estimator proposed by Andrews (1993) to the higher-order autoregressive processes, the nonnormal error term and inclusion of any exogenous variables. Also, we introduce the mean-unbiased estimator, which is compared with OLSE and the medium-unbiased estimator. Some simulation studies are performed to examine whether the proposed estimation procedure works well or not, where AR(p) for $p = 1, 2, 3$ models are examined. We obtain the results that it is possible to recover the true parameter values from OLSE and that the proposed procedure gives us the less biased estimators than OLSE. Finally, using actually obtained data, an empirical example of the median- and mean-unbiased estimators are shown.

*This paper is motivated by Diebold and Rudebusch (1991), where they investigated unbiased estimation in the case of autoregressive models. Therefore, first the author would like to acknowledge F.X. Diebold and G.D. Rudebusch for valuable suggestions. Moreover, the author is also grateful to two anonymous referees for helpful comments and suggestions. However, responsibility for any errors remains entirely with the author.

1 Introduction

The regression model with lagged dependent variables has been one of the research topics in econometrics. It is known that the OLSE's of autoregressive models are biased. Quenouille (1956) introduced the jackknife estimator of the AR parameter which is median-unbiased to order $1/T$ as T goes to infinity, where the trend term is not taken into account. Hurwicz (1950), Marriott and Pope (1954), Kendall (1954) and White (1961) established the mean-bias of the OLSE. Orcutt and Winokur (1969) constructed approximately mean-unbiased estimates of the AR parameter in stationary models. Sawa (1978), Tanaka (1983), Tsui and Ali (1994) and Ali (1996) also examined the AR(1) models, where the exact moments of OLSE are discussed. Shaman and Stine (1988) established the mean-bias of the OLSE to order $1/T$ in stationary AR(p) (also see Maekawa (1987) for the AR(p) models). Grubb and Symons (1987) gave an expression to order $1/T$ for bias to the estimated coefficient on a lagged dependent variable when all other regressors are exogenous (also see Tse (1982) and Maekawa (1983) for the AR models including the exogenous variables). Peters (1989) studied the finite sample sensitivity of OLSE of the AR(1) term with nonnormal errors. In Abadir (1993), an analytical formula was derived to approximate the finite sample bias of OLSE of the AR(1) term when the underlying process has a unit root. Moreover, Andrews (1993) derived the exactly median-unbiased estimator of the first-order autoregressive model, utilizing the Imhof (1961) algorithm. Andrews and Chen (1994) obtained the approximately median-unbiased estimator of autoregressive models, where Andrews (1993) is applied by transforming AR(p) models into AR(1) and taking the iterative procedure.

Thus, the autoregressive models have been studied with respect to the following four directions:

(i) a stationary model or a unit root model,

(ii) the first-order autoregressive model or the higher-order autoregressive models,

(iii) an autoregressive model with or without exogenous variables,

(iv) a normal error or a nonnormal error.

In this paper, based on a simulation technique, we propose the estimation procedure which can be applied to all the cases of (i) – (iv). That is, in more general formulation including the AR(p) terms and the other exogenous variables, we derive the asymptotically exact estimates of the regression coefficients in the sense of median- or mean-unbiasedness. Furthermore, the proposed estimation procedure can be easily applied to any nonnormal models.

Now we introduce two unbiased estimators (for example, see Andrews (1993) for the two estimators). Let θ be an unknown parameter and $\overline{\theta}$ and $\widetilde{\theta}$

be the estimates of θ. Suppose that the distribution functions of $\overline{\theta}$ and $\widetilde{\theta}$ are given by $f_{\overline{\theta}}(\cdot)$ and $f_{\widetilde{\theta}}(\cdot)$, respectively.

- $\widetilde{\theta}$ is called an *median-unbiased estimator* if we have the following relationship between θ and $\widetilde{\theta}$:

$$\theta = \mathrm{Med}(\widetilde{\theta}), \quad \text{where} \quad 0.5 = \int_{-\infty}^{\mathrm{Med}(\widetilde{\theta})} f_{\widetilde{\theta}}(x)\mathrm{d}x. \tag{1}$$

  That is, $\mathrm{Med}(\widetilde{\theta})$ denotes the median of $\widetilde{\theta}$ when the density function of $\widetilde{\theta}$ is given by $f_{\widetilde{\theta}}(\cdot)$.

- $\overline{\theta}$ is called an *mean-unbiased estimator* if we have the following relationship between θ and $\overline{\theta}$:

$$\theta = \mathrm{E}(\overline{\theta}) \equiv \int_{-\infty}^{+\infty} x f_{\overline{\theta}}(x)\mathrm{d}x. \tag{2}$$

The latter is widely known as an *unbiased estimator*. To distinguish the latter with the former, in this paper the latter is called the mean-unbiased estimator.

The underlying idea in this paper is described as follows. Let θ be an unknown parameter and $\widehat{\theta}$ be the biased estimate of θ. Suppose that the distribution function of $\widehat{\theta}$ is given by $f_{\widehat{\theta}}(\cdot)$. Since $\widehat{\theta}$ is assumed to be biased, we have $\theta \neq \mathrm{Med}(\widehat{\theta})$ and $\theta \neq \mathrm{E}(\widehat{\theta})$. For both the median- and mean-unbiased estimators, the following equations are obtained:

$$0.5 \equiv \int_{-\infty}^{\mathrm{Med}(\widehat{\theta})} f_{\widehat{\theta}}(x)\mathrm{d}x, \tag{3}$$

$$\mathrm{E}(\widehat{\theta}) \equiv \int_{-\infty}^{+\infty} x f_{\widehat{\theta}}(x)\mathrm{d}x. \tag{4}$$

Note that the biased estimate $\widehat{\theta}$ should be a function of the true parameter θ, i.e., $\widehat{\theta} = \widehat{\theta}(\theta)$. To obtain the numerical relationship between $\widehat{\theta}$ and θ, let $\{\widehat{\theta}_1^*, \widehat{\theta}_2^*, \cdots, \widehat{\theta}_n^*\}$ be a sequence of the biased estimates of θ, which are taken as the random numbers generated from $f_{\widehat{\theta}}(\cdot)$. Since $\widehat{\theta}_i^*$ is the i-th estimate of the parameter θ, it depends on the true parameter value θ, i.e., $\widehat{\theta}_i^* = \widehat{\theta}_i^*(\theta)$ for all $i = 1, 2, \cdots, n$. Using the n random draws, we can interpret equations (3) and (4) numerically as follows:

$$\widehat{\theta} = \text{Median of } \{\widehat{\theta}_1^*(\theta), \widehat{\theta}_2^*(\theta), \cdots, \widehat{\theta}_n^*(\theta)\}, \tag{5}$$

$$\widehat{\theta} = \frac{1}{n}\sum_{i=1}^{n} \widehat{\theta}_i^*(\theta), \tag{6}$$

where equation (5) implies that the median of n random draws ($\widehat{\theta}_i^*$, $i = 1, 2, \cdots, n$) should be equal to the biased estimate $\widehat{\theta}$ and equation (6) indicates that the arithmetic average of n random draws ($\widehat{\theta}_i^*$, $i = 1, 2, \cdots, n$) should be equal to $\widehat{\theta}$. Note that the θ which satisfies equation (5) is defined as the median-unbiased estimate, which is denoted by $\widetilde{\theta}$, while the θ which satisfies equation (6) is called the mean-unbiased estimate, which is denoted by $\overline{\theta}$. When n is sufficiently large, the obtained $\widetilde{\theta}$ and $\overline{\theta}$ should be the unbiased estimates of θ in the sense of median and mean. The two equations shown above are practically solved by an iterative procedure or a simple grid search. The problem in the procedure above is to compute n biased estimates of θ, i.e., $\{\widehat{\theta}_1^*, \widehat{\theta}_2^*, \cdots, \widehat{\theta}_n^*\}$. In a framework of the regression models, generating a series of the dependent variable given the explanatory variables and the unknown paramater, we obtain the OLS estimate of the parameter. Repeating the procedure, the n biased estimates, i.e., $\{\widehat{\theta}_1^*, \widehat{\theta}_2^*, \cdots, \widehat{\theta}_n^*\}$, can be generated. Thus, in this paper we consider generating the data series and obtain the θ which satisfies equations (5) and (6).

2 Coefficient Estimates of Lagged Dependent Variables

In this section, we discuss how much bias OLSE has in the case of AR models. To start with general formulation, we take the autoregressive model which may include the exogenous variables, say x_t. That is, consider the following simple regression model:

$$y_t = x_t\beta + \sum_{j=1}^{p} \alpha_j y_{t-j} + u_t, \tag{7}$$

for $t = p+1, p+2, \cdots, T$, where x_t and β are a $1 \times k$ vector and a $k \times 1$ vector, respectively. u_t is assumed to be distributed with mean zero and variance σ^2, which is usually normal. In this paper, the initial values y_p, y_{p-1}, $\cdots$, y_1 are assumed to be constant.

Rewriting equation (7) in a matrix form, we have:

$$y = X\beta + Y_{-1}\alpha + u,$$

where y, X, β and Y_{-1} are denoted by:

$$y = \begin{pmatrix} y_{p+1} \\ y_{p+2} \\ \vdots \\ y_T \end{pmatrix}, \quad X = \begin{pmatrix} x_{p+1} \\ x_{p+2} \\ \vdots \\ x_T \end{pmatrix}, \quad \alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{pmatrix},$$

$$Y_{-1} = \begin{pmatrix} y_p & y_{p-1} & \cdots & y_1 \\ y_{p+1} & y_p & \cdots & y_2 \\ \vdots & \vdots & \vdots & \vdots \\ y_{T-1} & y_{T-2} & \cdots & y_{T-p} \end{pmatrix}$$

Then, OLSE of $(\beta', \alpha')'$ is given by:

$$\begin{pmatrix} \widehat{\beta} \\ \widehat{\alpha} \end{pmatrix} = \begin{pmatrix} X'X & X'Y_1 \\ Y'_{-1}X & Y'_{-1}Y_{-1} \end{pmatrix}^{-1} \begin{pmatrix} X'y \\ Y'_{-1}y \end{pmatrix}$$

$\widehat{\beta}$ is computed as:

$$\widehat{\beta} = (X'X - X'Y_{-1}(Y'_{-1}Y_{-1})^{-1}Y'_{-1}X)^{-1}X'(y - Y_{-1}\widehat{\alpha}^*),$$

where $\widehat{\alpha}^* = (Y'_{-1}Y_{-1})^{-1}Y'_{-1}y$, which represents OLSE of α^* in the AR(p) model $y = Y_{-1}\alpha^*$+error, i.e., OLSE of the autoregressive coefficient vector in the AR(p) model. Also, the autoregressive coefficient vector $\widehat{\alpha}$ from equation (7) is represented as:

$$\widehat{\alpha} = \widehat{\alpha}^* - (Y'_{-1}Y_{-1})^{-1}Y'_{-1}X\widehat{\beta}.$$

Thus, both $\widehat{\beta}$ and $\widehat{\alpha}$ depend on $\widehat{\alpha}^*$. It is well known that OLSE of the autoregressive coefficient vector in the AR(p) model (i.e.,$\widehat{\alpha}^*$) is biased in small sample (see, for example, Andrews (1993), Andrews and Chen (1994), Diebold and Rudebusch (1991), Hurwicz (1959), Kendall (1954), Marriott and Pope (1954), Quenouille (1956) and so on). Therefore, both $\widehat{\beta}$ and $\widehat{\alpha}$ are biased because they depend on $\widehat{\alpha}^*$. From the above equations, when $\widehat{\alpha}^*$ is downward-biased, $\widehat{\alpha}$ is downward-biased while $\widehat{\beta}$ is upward-biased.

Let Model 1 be the case of $k = 0$, Model 2 be the case of $k = 1$ and $x_t = 1$ and Model 3 be the case of $k = 2$ and $x_t = (1,\ x_{1t})$, i.e,

$$\text{Model 1: } y_t = \sum_{j=1}^{p} \alpha_j y_{t-j} + u_t,$$

$$\text{Model 2: } y_t = \beta_1 + \sum_{j=1}^{p} \alpha_j y_{t-j} + u_t,$$

$$\text{Model 3: } y_t = \beta_1 + \beta_2 x_{1t} + \sum_{j=1}^{p} \alpha_j y_{t-j} + u_t,$$

for $t = p+1, p+2, \ldots, T$, given the initial condition $y_1 = y_2 = \cdots = y_p = 0$.

Now we examine by Monte Carlo simulations how large the OLSE bias is, where we focus only on the case of $p = 1$, i.e., the AR(1) model. Suppose that the true model is represented by Model 1 with $p = 1$. When $x_{1t} = t$ (time trend) is taken in Model 3, it is known that OLSE of α_1 from Model 3 gives us the largest bias of the OLSE's from Models 1 – 3, while OLSE of α_1 from

Model 1 yields the smallest one (see, for example, Andrews (1993)). Figures 1 and 2 show the relationship between the true autoregressive coefficient (i.e., α_1), the median and mean of OLSE's from 10,000 simulation runs. To draw the two figures, we take the following simulation procedure.

(i) Generate y_2, y_3, $\cdots$, y_T by Model 1 given α_1, $u_t \sim N(0,1)$ and $y_1 = 0$, where $T = 20$.

(ii) Compute OLSE of α_1 by estimating Model 1, that of (β_1, α_1) by Model 2, and that of $(\beta_1, \beta_2, \alpha_1)$ by Model 3. Note in Model 3 that $x_{1t} = t$ (time trend) is taken in this simulation study.

(iii) Repeat (i) and (ii) 10,000 times.

(iv) Obtain medium and mean from the 10,000 OLSE's of α_1.

(v) Repeat (i) – (iv) given the exactly same random draws for u_t (i.e., $10,000 \times (T-p)$ random draws for $T = 20$ and $p = 1$) and the different parameter value for α_1 (i.e., $\alpha_1 = -1.20, -1.49. - 1.48, \cdots, 1.20$).

Thus, we have median and mean from the 10,000 OLSE's corresponding to the true parameter value for each model. In both Figures 1 and 2, the true model is given by Model 1 and it is estimated by Models 1 – 3. The horizontal line implies the true parameter value of the AR(1) coefficient and the vertical line indicates the OLSE corresponding to the true parameter value. Unless the OLSE is biased, the 45° degree line represents the relationship between the true parameter value and the OLSE.

Each line indicates the arithmetic mean of the 10,000 OLSE's in Figure 1 and the median of the 10,000 OLSE's in Figure 2. Taking either median or mean, there is the largest bias around $\alpha_1 = 1$ for all the Models 1 – 3. From Figures 1 and 2, bias drastically increases as the number of the exogenous variables increases. That is, in the case where α_1 is positive, OLSE of Model 3 has the largest downward-bias and OLSE of Model 1 represents the smallest downward-bias, which implies that inclusion of more extra variables results in larger bias of OLSE.

Thus, from Figures 1 and 2 we can see how large the OLSE bias is. That is, discrepancy between the 45° degree line and the other lines increases as the number of the extra variables increases. Now we consider correcting the OLSE bias. In Figures 1 and 2, we see median and mean from the 10,000 OLSE's given the true coefficient, respectively. It is also possible to read the figures reversely. For example, in Figure 1, when OLSE is obtained as $\widehat{\alpha}_1 = 0.5$ from actually observed data, the true parameter value α_1 can be estimated as 0.526 for Model 1, 0.642 for Model 2 and 0.806 for Model 3. Similarly, in Figure 2, when we have $\widehat{\alpha}_1 = 0.50$, the mean-unbiased estimate is given by 0.552 for Model 1, 0.673 for Model 2 and 0.849 for Model 3. For each estimator, it is possible to consider shifting the distribution of OLSE toward the distribution around the true value in the sense of median or mean.

Figure 1: Median from 10,000 OLSE's of AR(1) Coefficient ($T = 20$)

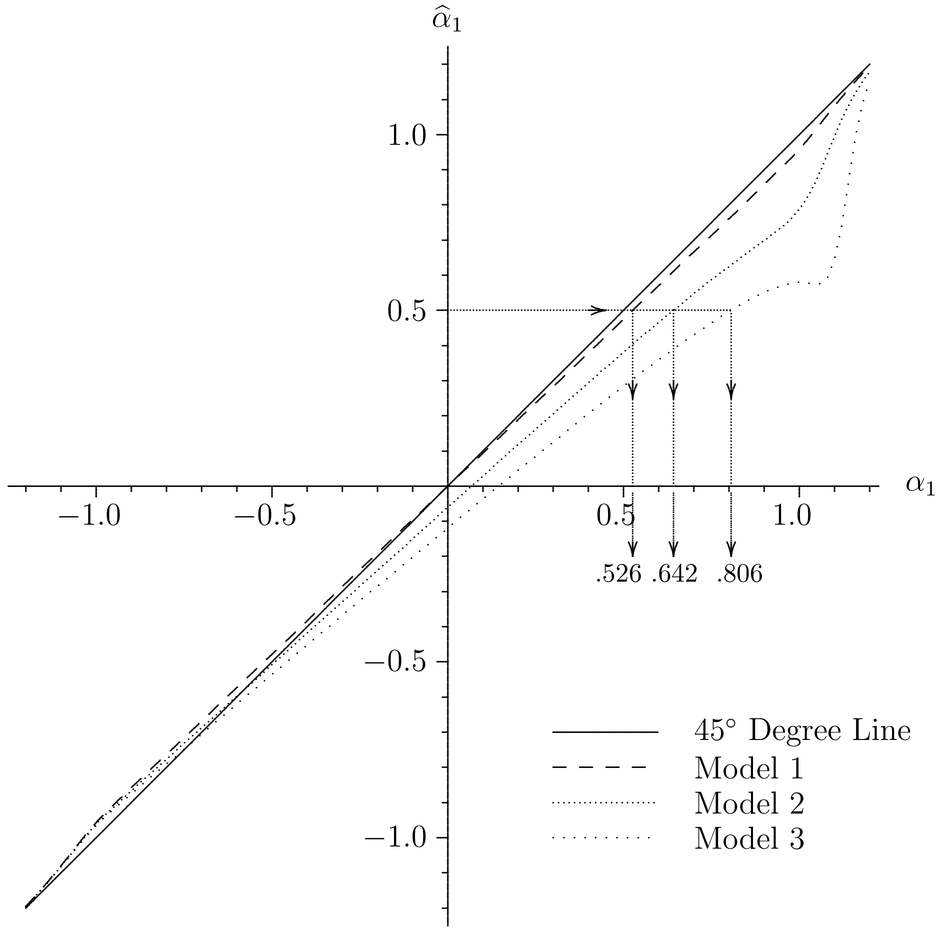
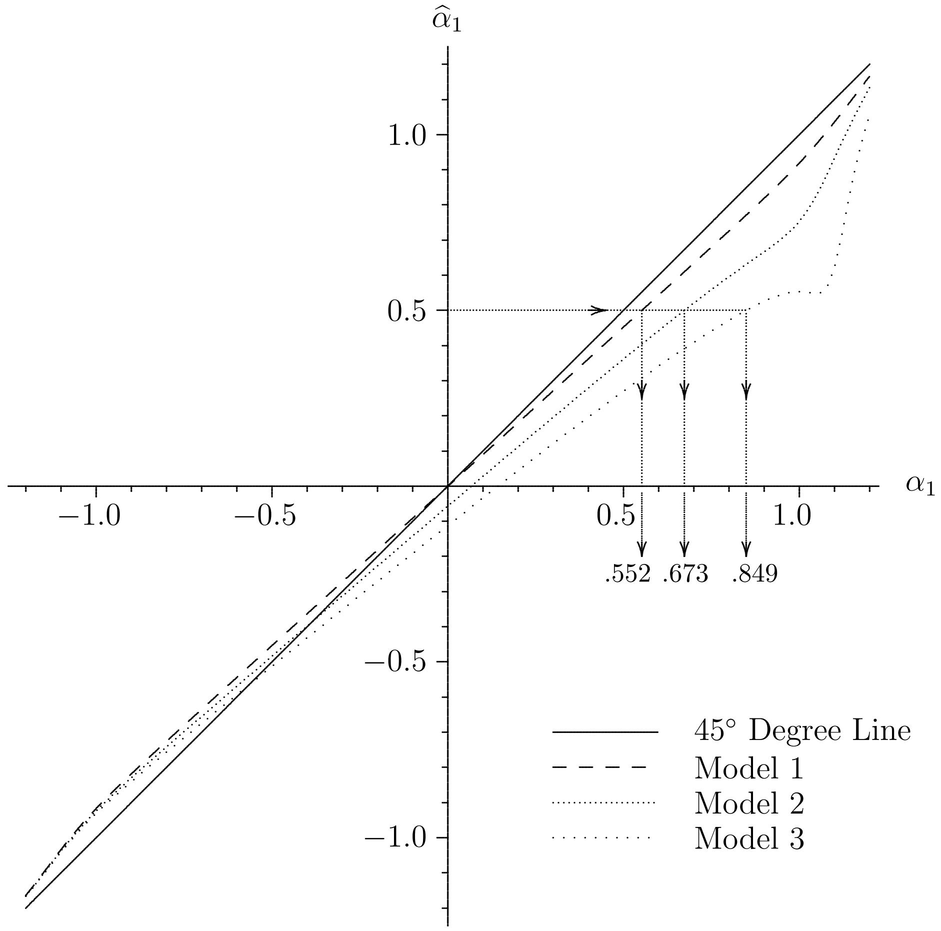


Figure 2: Mean from 10,000 OLSE's of AR(1) Coefficient ($T = 20$)



In practice, however, no one knows the true model. What we can do is to estimate the model assumed by a researcher. Figures 1 and 2 indicate that inclusion of more extra variables possibly yields serious biased OLSE and furthermore that the true parameter values can be recovered from the estimated model even if we do not know the true model. In the next section, based on this idea, we obtain the median- and mean-unbiased estimators, which can be applied to any case of the higher-order autoregressive models, the nonnormal error term and inclusion of the exogenous variables other than the constant and trend terms. Here, we take the constant term and the time trend as x_t, although any exogenous variables can be included in the model.

3 Bias Correction of OLSE

Let us define as $\theta = (\beta', \alpha')'$. Consider the two estimates of θ which are denoted by $\widetilde{\theta}$ and $\overline{\theta}$. We have defined $\widetilde{\theta}$ as θ such that the OLSE given actual data is equal to median of the simulated OLSE's given θ, which is called the median-unbiased estimate (see equation (1) for the definition of median-unbiasedness). Similarly, $\overline{\theta}$ denotes θ such that the OLSE given actual data is equal to arithmetic mean of the simulated OLSE's given θ, which is called the mean-unbiased estimate (see equation (2) for the definition of median-unbiasedness). From equations (5) and (6), we can suppose that a certain relationship between $\widehat{\theta}$ and θ is as follows:

$$\widehat{\theta} = g(\theta), \tag{8}$$

where $\widehat{\theta}$ represents OLSE of θ, given the original data y_t and x_t. $g(\cdot)$ denotes a $(k+p) \times 1$ vector function, which corresponds to the right-hand side in equations (5) and (6).

A functional form of $g(\cdot)$ is different, depending on the median-unbiased estimator (i.e., $\widetilde{\theta}$) or the mean-unbiased estimator (i.e., $\overline{\theta}$), because the definition of the medium-unbiased estimator is different from that of the mean-unbiased estimator as shown in equations (1) and (2). In any case, it is impossible to obtain an explicit functional form of $g(\cdot)$ for both estimators. Therefore, we numerically obtain the two estimators by the simulation technique. For $\widetilde{\theta}$, the vector function $g(\theta)$ is taken as the median of the simulated OLSE's given θ, which is represented by equation (5). In the multi-dimensional case of θ, we take the median for each element of θ. For $\overline{\theta}$, the vector function $g(\theta)$ is defined as the arithmetic mean of the simulated OLSE's given θ, which corresponds to equation (6). In order to obtain $\widetilde{\theta}$ and $\overline{\theta}$, the numerical optimization procedure is applied (see, for example, Tanizaki (1995)). Using equation (8), we update the parameter θ as follows:

$$\theta^{(i+1)} = \theta^{(i)} + \gamma^{(i)} \left(\widehat{\theta} - g(\theta^{(i)}) \right), \tag{9}$$

where i denotes the i-th iteration and $g(\cdot)$ is the vector function defined above. In the first iteration, OLSE of θ is taken for $\theta^{(1)}$, i.e., $\theta^{(1)} = \widehat{\theta}$. $\gamma^{(i)}$ is a scalar, which may depend on the number of iteration i.

For an interpretation of $\gamma^{(i)}$, it might be appropriate to consider that the Newton-Raphson optimization procedure is taken. which is described as follows. Approximating $\widehat{\theta} - g(\theta)$ about $\theta = \theta^*$, we have:

$$\begin{aligned} 0 &= \widehat{\theta} - g(\theta) \\ &\approx \widehat{\theta} - g(\theta^*) - \frac{\partial g(\theta^*)}{\partial \theta'}(\theta - \theta^*). \end{aligned}$$

Then, we can rewrite as:

$$\theta = \theta^* + \left(\frac{\partial g(\theta^*)}{\partial \theta'} \right)^{-1} \left(\widehat{\theta} - g(\theta^*) \right).$$

Regarding θ as $\theta^{(i+1)}$ and θ^* as $\theta^{(i)}$, the following equation is derived:

$$\theta^{(i+1)} = \theta^{(i)} + \left(\frac{\partial g(\theta^{(i)})}{\partial \theta'} \right)^{-1} (\widehat{\theta} - g(\theta^{(i)})),$$

which is equivalent to equation (9) with the following condition:

$$\left(\frac{\partial g(\theta^{(i)})}{\partial \theta'} \right)^{-1} = \gamma^{(i)} I_{k+p},$$

where I_{k+p} denotes a $(k+p) \times (k+p)$ identity matrix. Since $g(\theta)$ cannot be explicitly specified, we take the first derivative of $g(\theta)$ as the diagonal matrix. Moreover, taking into account the convergence speed, $\gamma^{(i)} = c^{i-1}$ is used in this paper, where $c = 0.9$.

Thus, using equation (9), the median- and mean-unbiased estimators can be obtained. When $\theta^{(i+1)}$ is stable, we take it as the estimate of θ, i.e., $\widetilde{\theta}$ or $\overline{\theta}$. As for convergence criterion, in this paper, when each element of $\theta^{(i+1)} - \theta^{(i)}$ is less than 0.001 in absolute value, we consider that $\theta^{(i+1)}$ is stable.

Under the setup above, the computational procedure is shown as follows.

1. Given the actual time series data (i.e., x_t and y_t), estimate θ and σ^2 by OLS, which are denoted by $\widehat{\theta}$ and $\widehat{\sigma}^2$.
2. (i) Given the initial values $y_p, y_{p-1}, \cdots, y_1$, the exogenous variable x_t for $t = p+1, p+2, \cdots, T$ and $\theta^{(i)}$, generate random draws of u_t by $u_t \sim N(0, \widehat{\sigma}^2)$ and obtain random draws of y_t for $t = p+1, p+2, \cdots, T$ using equation (7), where θ in equation (7) is taken as $\theta^{(i)}$. Note that the initial values $y_p, y_{p-1}, \cdots, y_1$ may be taken as the actual data.

(ii) Given the actual data x_t and the simulated data of y_t for $t = p + 1, p+2, \cdots, T$, compute OLSE of θ.

(iii) Performing Steps 2.(i) and 2.(ii) n times, n OLSE's of θ are obtained based on $\theta^{(i)}$. From the n OLSE's of θ, compute median or mean for each element of θ, which corresponds to the function $g(\theta^{(i)})$. In this paper, $n = 10,000$ is taken.

3. Using equation (9), $\theta^{(i)}$ is updated to $\theta^{(i+1)}$.
4. Repeat Steps 2 and 3 until $\theta^{(i+1)}$ is stable. Note that we need to use the same random draws of u_t for each simulation run. In other words, $n \times (T-p)$ random draws of u_t have to be stored.

The limit of $\theta^{(i+1)}$ is taken as $\widetilde{\theta}$ or $\overline{\theta}$. Using the simulation procedure shown in Steps 1 – 4, we can obtain the unbiased estimator of the autoregressive coefficients in the higher-order cases, in the case of inclusion of the exogenous variables, and in stationary or nonstationary cases. Moreover, in Step 2 of the estimation procedure, we possibly assume that the error term u_t is nonnormal. Accordingly, the proposed procedure can be broadly applied to various cases.

Confidence Interval of the Coefficients: In the above procedure, $\widetilde{\theta}$ or $\overline{\theta}$ is obtained. For statistical inference, we need the distributions of $\widetilde{\theta}$ and $\overline{\theta}$. However, it is impossible to know the distributions explicitly. Therefore, we consider obtaining the distributions numerically.

When we judge that $\theta^{(i+1)}$ is stable in Procedure 4, n OLSE's of θ are available in Procedure 2.(iii). For each of the n OLSE's of θ, we may compute the median-unbiased estimate or the mean-unbiased estimate. Thus, we can obtain the n median-unbiased estimates associated with the n OLSE's and similarly the n mean-unbiased estimates based on the n OLSE's. From the n median- or mean-unbiased estimates, we can compute the percentiles by sorting, the standard errors and etc. In order to obtain one median-unbiased estimate (or one mean-unbiased estimate), we have to compute n OLSE's. Therefore, we need $n \times n$ OLSE's for the n median-unbiased estimates (or the n mean-unbiased estimates). This implies that constructing the confidence interval, the standard error of the coefficients and etc takes an extremely lot of time computationally.

Standard Error of Regression: Once $\widetilde{\theta}$ and $\overline{\theta}$ are computed by Procedures 1 – 4, $\widetilde{\sigma}^2$ and $\overline{\sigma}^2$ are derived by using the following formula:

$$\sigma^2 = \frac{1}{T-p-k} \sum_{t=p+1}^{T} (y_t - x_t\beta - \sum_{j=1}^{p} \alpha_j y_{t-p})^2, \tag{10}$$

where $\theta = (\beta', \alpha')'$ takes $\widetilde{\theta}$ or $\overline{\theta}$. In equation (10), σ^2 is represented by $\widetilde{\sigma}^2$ when θ is taken as $\widetilde{\theta}$, while σ^2 reduces to $\overline{\sigma}^2$ when $\overline{\theta}$ is used for θ. In other words,

(θ, σ^2) is represented by $(\widetilde{\theta}, \widetilde{\sigma}^2)$ for the median-unbiased estimator and $(\overline{\theta}, \overline{\sigma}^2)$ for the mean-unbiased estimator.

4 Monte Carlo Experiments

We have derived the median-unbiased and mean-unbiased estimators in the previous sections. Using the first-order autoregressive model (Section 4.1) and the higher-order autoregressive models (Section 4.2), it is examined whether the proposed procedure works well.

4.1 AR(1) Models

In this section, providing that Model 1 is the true model, we consider estimating the following three models:

Model 1: $y_t = \alpha_1 y_{t-1} + u_t,$

Model 2: $y_t = \beta_1 + \alpha_1 y_{t-1} + u_t,$

Model 3: $y_t = \beta_1 + \beta_2 x_{1t} + \alpha_1 y_{t-1} + u_t,$

where the initial value is fixed as $y_1 = 0$ for Models 1 – 3. The error terms $u_2, u_3, \cdots, u_T$ are mutually independently distributed as normal random variables with mean zero and variance one. We estimate Models 1 – 3 given the artificially generated data from Model 1 and examine $\widehat{\alpha}_1$, $\widetilde{\alpha}_1$ and $\overline{\alpha}_1$ with respect to the arithmetic average (AVE), the root mean square error (RMS) and the median (50%). $\alpha_1 = 0.6, 0.9, 1.0$ and $T = 20, 40, 60$ are taken for each of Models 1 – 3.

In Table 1, under the assumption that the true model is given by Model 1, we estimate Models 1 – 3 by OLS, the median-unbiased estimator and the mean-unbiased estimator. In Model 1 of Table 1, when T is small, the OLSE bias is large. As α_1 goes to one, the OLSE bias increases. When α_1 goes to one, $\widetilde{\alpha}_1$ and $\overline{\alpha}_1$ are better than $\widehat{\alpha}_1$ in the sense of RMS. Furthermore, for all the Models in Table 1, AVE of $\overline{\alpha}_1$ is closer to the true parameter value α_1 than AVE of $\widetilde{\alpha}_1$, while 50% of $\widetilde{\alpha}_1$ is closer to the true parameter value α_1 than 50% of $\overline{\alpha}_1$. Recall that $\widetilde{\alpha}_1$ is the estimator such that the median (50%) is equal to the true value and $\overline{\alpha}_1$ is the estimator such that the arithmetic mean (AVE) is equal to the true parameter. That is, from the computation procedure, AVE of $\overline{\alpha}_1$ and 50% of $\widetilde{\alpha}_1$ should be close to the true parameter value. If the distribution is symmetric, the median (50%) is equal to the mean (AVE). In the case of OLSE with lagged dependent variable (i.e., the AR(1) case), however, it is well known that OLSE is distributed with skewness to the right as the autoregressive coefficient (α_1) goes to one. Accordingly, AVE of $\widetilde{\alpha}_1$ is smaller than AVE of $\overline{\alpha}_1$ and also 50% of $\widetilde{\alpha}_1$ is smaller than 50% of $\overline{\alpha}_1$.

Table 1: AR(1) Models (True Model is Model 1.)

Estimated by:			Model 1			Model 2			Model 3 ($x_{1t} = t$)		
T	α_1		AVE	RMS	50%	AVE	RMS	50%	AVE	RMS	50%
20	0.6	$\widehat{\alpha}_1$	.553	.205	.571	.447	.268	.464	.343	.348	.360
		$\widetilde{\alpha}_1$	.581	.210	.601	.582	.260	.595	.581	.320	.573
		$\overline{\alpha}_1$	.609	.219	.630	.607	.271	.622	.601	.329	.596
	0.9	$\widehat{\alpha}_1$	.827	.176	.866	.678	.295	.713	.534	.432	.558
		$\widetilde{\alpha}_1$	.865	.168	.908	.827	.229	.883	.792	.309	.812
		$\overline{\alpha}_1$	.901	.168	.949	.855	.223	.923	.810	.301	.846
	1.0	$\widehat{\alpha}_1$	.922	.166	.959	.763	.301	.799	.559	.498	.582
		$\widetilde{\alpha}_1$	.957	.151	1.001	.886	.221	.932	.812	.340	.846
		$\overline{\alpha}_1$	.993	.147	1.036	.910	.209	.962	.829	.328	.874
40	0.6	$\widehat{\alpha}_1$	.577	.134	.594	.528	.158	.540	.481	.190	.495
		$\widetilde{\alpha}_1$	.591	.136	.608	.587	.152	.598	.589	.168	.603
		$\overline{\alpha}_1$	.606	.138	.624	.603	.156	.615	.604	.174	.618
	0.9	$\widehat{\alpha}_1$	.862	.102	.884	.787	.163	.808	.716	.227	.734
		$\widetilde{\alpha}_1$	.882	.098	.905	.869	.134	.893	.861	.167	.871
		$\overline{\alpha}_1$	.903	.098	.927	.889	.130	.919	.879	.163	.897
	1.0	$\widehat{\alpha}_1$	.960	.088	.982	.873	.163	.897	.762	.272	.784
		$\widetilde{\alpha}_1$	.978	.079	1.003	.938	.120	.967	.900	.181	.926
		$\overline{\alpha}_1$	.998	.076	1.022	.952	.110	.985	.917	.170	.953
60	0.6	$\widehat{\alpha}_1$	.583	.110	.590	.553	.123	.561	.524	.140	.537
		$\widetilde{\alpha}_1$	.593	.111	.599	.592	.119	.599	.594	.127	.608
		$\overline{\alpha}_1$	.602	.113	.609	.601	.121	.609	.603	.130	.617
	0.9	$\widehat{\alpha}_1$	.873	.076	.889	.829	.112	.845	.785	.151	.801
		$\widetilde{\alpha}_1$	.887	.073	.904	.885	.096	.901	.885	.118	.897
		$\overline{\alpha}_1$	.901	.073	.918	.900	.096	.918	.900	.117	.916
	1.0	$\widehat{\alpha}_1$	.972	.058	.987	.916	.108	.931	.839	.186	.855
		$\widetilde{\alpha}_1$	.984	.052	1.001	.961	.078	.981	.935	.123	.953
		$\overline{\alpha}_1$	.997	.048	1.013	.971	.071	.993	.946	.115	.972

Now, we investigate for each estimator whether bias increases or not as the number of irrelevant variables increases. For $\widehat{\alpha}_1$, the OLSE bias of Model 1 is smaller than that of Model 2 and furthermore that of Model 2 is smaller than that of Model 3. $\widetilde{\alpha}_1$ and $\overline{\alpha}_1$ are closer to the true parameter value than $\widehat{\alpha}_1$ for all the Models 1 – 3, but their bias increases as the number of irrelevant variables increases. Thus, in the classical regression model, inclusion of irrelevant variables results in the unbiased OLSE (see, for example, Greene (1993)). However, in a context of the autoregressive models, inclusion of more extra variables generates more serious OLSE bias. For both Models 2 and 3, AVE of $\overline{\alpha}_1$ and 50% of $\widetilde{\alpha}_1$ should be close to the true parameter value, which result is similar to the case of Model 1. Therefore, judging from the bias criterion, it might be concluded that the OLSE bias is correctly improved using the median- and mean-unbiased estimators even if the true model is different from the estimated model. This result suggests that it might be possible to recover the true parameter values from any estimated model. However, note as follows. For all the three estimators, RMS increases as the number of irrelevant variables increases. That is, RMS's of Model 3 are larger than RMS's of Model 2, which are larger than RMS's of Model 1.

4.2 AR(p) Models

Next, we consider the AR(p) models, where $p = 2, 3$ is taken. Assume that the true model is represented by Model 1, i.e.,

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \cdots + \alpha_p y_{t-p} + u_t,$$

for $t = p+1, p+2, \cdots, T$, where u_t is assumed to be distributed as a standard normal random variable and the initial values are given by $y_1 = y_2 = \cdots = y_p = 0$. The above AR($p$) model is rewritten as:

$$(1 - \lambda_1 L)(1 - \lambda_2 L) \cdots (1 - \lambda_p L) y_t = u_t,$$

where λ_1, λ_2, $\cdots$, λ_p are assumed to be real numbers and L denotes the lag operator. Taking the cases of $p = 2, 3$, we estimate the true model by Model 1. That is, the true model is equivalent to the estimated model. The results are in Table 2 for AR(2) and Table 3 for AR(3). The sample size is taken as $T = 60$ in both Tables 2 and 3. In the second column of Tables 2 and 3, α, $\widehat{\alpha}$, $\widetilde{\alpha}$ and $\overline{\alpha}$ denote the true parameter value, the OLSE, the median-unbiased estimate and the mean-unbiased estimate, respectively. In Tables 2 and 3, note that there is one-to-one correspondence between (λ_1, λ_2, $\cdots$, λ_p) and (α_1, α_2, $\cdots$, α_p).

In Table 2, RMS of the median- and mean-unbiased estimates is smaller than that of OLSE in the case of $\lambda_1 = \lambda_2 = 1$. In the cases of non-zero coefficients of α_1 and α_2, 50% of the median-unbiased estimates and AVE of

Table 2: AR(2) Model ($T = 60$)

(λ_1, λ_2)		α_1 AVE	RMS	50%	α_2 AVE	RMS	50%
(0.0,0.0)	α	0.000			0.000		
	$\widehat{\alpha}$	0.000	0.138	0.003	-0.012	0.136	-0.005
	$\widetilde{\alpha}$	0.001	0.140	0.003	0.007	0.140	0.014
	$\overline{\alpha}$	0.000	0.140	0.002	0.006	0.142	0.013
(0.0,0.5)	α	0.500			0.000		
	$\widehat{\alpha}$	0.492	0.139	0.495	-0.012	0.134	-0.007
	$\widetilde{\alpha}$	0.501	0.140	0.504	0.006	0.138	0.011
	$\overline{\alpha}$	0.500	0.141	0.503	0.006	0.140	0.011
(0.0,1.0)	α	1.000			0.000		
	$\widehat{\alpha}$	0.981	0.141	0.981	-0.012	0.140	-0.017
	$\widetilde{\alpha}$	1.000	0.142	1.001	-0.005	0.144	-0.009
	$\overline{\alpha}$	1.000	0.144	1.001	-0.005	0.146	-0.009
(0.5,0.5)	α	1.000			-0.250		
	$\widehat{\alpha}$	0.984	0.136	0.988	-0.252	0.130	-0.252
	$\widetilde{\alpha}$	0.999	0.135	1.003	-0.244	0.135	-0.245
	$\overline{\alpha}$	1.001	0.137	1.004	-0.248	0.137	-0.247
(0.5,1.0)	α	1.500			-0.500		
	$\widehat{\alpha}$	1.471	0.126	1.480	-0.487	0.123	-0.497
	$\widetilde{\alpha}$	1.497	0.124	1.506	-0.499	0.125	-0.509
	$\overline{\alpha}$	1.504	0.126	1.514	-0.507	0.128	-0.517
(1.0,1.0)	α	2.000			-1.000		
	$\widehat{\alpha}$	1.952	0.087	1.968	-0.953	0.088	-0.968
	$\widetilde{\alpha}$	1.983	0.071	2.000	-0.983	0.073	-1.001
	$\overline{\alpha}$	1.996	0.069	2.013	-0.996	0.071	-1.014

Table 3: AR(3) Model ($T = 60$)

$(\lambda_1, \lambda_2, \lambda_3)$		α_1 AVE	RMS	50%	α_2 AVE	RMS	50%	α_3 AVE	RMS	50%
(0.0,0.0,0.0)	α	0.000			0.000			0.000		
	$\widehat{\alpha}$	-0.001	0.137	-0.002	-0.010	0.132	-0.014	0.006	0.135	0.008
	$\widetilde{\alpha}$	-0.001	0.139	-0.002	0.006	0.136	0.002	0.007	0.142	0.009
	$\overline{\alpha}$	-0.002	0.139	-0.002	0.007	0.138	0.002	0.007	0.145	0.009
(0.0,0.0,0.5)	α	0.500			0.000			0.000		
	$\widehat{\alpha}$	0.490	0.137	0.489	-0.010	0.146	-0.009	0.000	0.134	0.002
	$\widetilde{\alpha}$	0.499	0.139	0.497	0.004	0.151	0.005	0.001	0.141	0.001
	$\overline{\alpha}$	0.498	0.139	0.497	0.008	0.154	0.008	0.001	0.143	0.003
(0.0,0.0,1.0)	α	1.000			0.000			0.000		
	$\widehat{\alpha}$	0.979	0.139	0.978	-0.009	0.184	0.000	-0.001	0.132	0.004
	$\widetilde{\alpha}$	0.998	0.140	0.997	-0.001	0.190	0.008	-0.010	0.138	-0.004
	$\overline{\alpha}$	0.999	0.140	0.998	0.008	0.193	0.018	-0.012	0.141	-0.007
(0.0,0.5,0.5)	α	1.000			-0.250			0.000		
	$\widehat{\alpha}$	0.982	0.137	0.978	-0.246	0.180	-0.242	-0.005	0.131	-0.004
	$\widetilde{\alpha}$	0.998	0.139	0.994	-0.245	0.186	-0.241	-0.005	0.138	-0.005
	$\overline{\alpha}$	0.998	0.139	0.995	-0.241	0.190	-0.237	-0.005	0.140	-0.004
(0.0,0.5,1.0)	α	1.500			-0.500			0.000		
	$\widehat{\alpha}$	1.470	0.140	1.467	-0.484	0.227	-0.476	-0.002	0.133	-0.002
	$\widetilde{\alpha}$	1.499	0.140	1.494	-0.495	0.233	-0.485	-0.011	0.139	-0.010
	$\overline{\alpha}$	1.499	0.140	1.494	-0.489	0.237	-0.479	-0.013	0.142	-0.012
(0.0,1.0,1.0)	α	2.000			-1.000			0.000		
	$\widehat{\alpha}$	1.957	0.143	1.954	-0.960	0.271	-0.952	0.002	0.143	-0.003
	$\widetilde{\alpha}$	1.997	0.140	1.994	-1.002	0.277	-0.992	0.002	0.147	-0.002
	$\overline{\alpha}$	1.999	0.141	1.997	-0.998	0.282	-0.989	-0.002	0.150	-0.005
(0.5,0.5,0.5)	α	1.500			-0.750			0.125		
	$\widehat{\alpha}$	1.472	0.137	1.469	-0.721	0.215	-0.716	0.109	0.130	0.114
	$\widetilde{\alpha}$	1.498	0.137	1.494	-0.741	0.220	-0.735	0.115	0.136	0.119
	$\overline{\alpha}$	1.499	0.138	1.495	-0.741	0.224	-0.736	0.117	0.138	0.122
(0.5,0.5,1.0)	α	2.000			-1.250			0.250		
	$\widehat{\alpha}$	1.960	0.140	1.960	-1.202	0.250	-1.195	0.234	0.133	0.235
	$\widetilde{\alpha}$	1.996	0.136	1.997	-1.238	0.253	-1.231	0.239	0.138	0.238
	$\overline{\alpha}$	1.998	0.137	1.999	-1.241	0.257	-1.235	0.240	0.141	0.240
(0.5,1.0,1.0)	α	2.500			-2.000			0.500		
	$\widehat{\alpha}$	2.445	0.136	2.454	-1.914	0.265	-1.928	0.468	0.134	0.475
	$\widetilde{\alpha}$	2.496	0.127	2.506	-1.996	0.257	-2.014	0.499	0.134	0.510
	$\overline{\alpha}$	2.503	0.129	2.512	-2.007	0.262	-2.023	0.504	0.136	0.515
(1.0,1.0,1.0)	α	3.000			-3.000			1.000		
	$\widehat{\alpha}$	2.917	0.126	2.932	-2.831	0.258	-2.864	0.914	0.133	0.932
	$\widetilde{\alpha}$	2.984	0.095	2.999	-2.968	0.197	-3.003	0.984	0.103	1.003
	$\overline{\alpha}$	2.997	0.094	3.013	-2.993	0.196	-3.032	0.997	0.103	1.018

the mean-unbiased estimates are closer to the true parameter values than those of OLSE. In the cases of zero coefficients, the three estimators are close to each other. Therefore, it might be concluded that the OLSE bias is corrected by the median- and mean-unbiased estimators when OLSE is biased. Thus, in the case of the AR(2) models, we obtain the same results as in the case of the AR(1) models.

Now, we examine the AR(3) models and the results are in Table 3. For estimation of zero coefficients, all the three estimators are close to the true parameter value, which is equal to zero. However, for estimation of non-zero coefficients, the median- and mean-unbiased estimators are superior to OLSE, which implies that the median- and mean-unbiased estimators are less biased than OLSE.

Thus, for all the cases of AR(p) for $p = 1, 2, 3$, it is shown from Tables 1 – 3 that OLSE bias is corrected using the median- and mean-unbiased estimators even if the data generating process is not known. Finally, note as follows. In Table 2, the case of $\lambda_1 = 0$ and $\lambda_2 \neq 0$ implies that the data generating process is AR(1). In Table 3, the case of $\lambda_1 = 0$, $\lambda_2 = 0$ and $\lambda_3 \neq 0$ implies that the data generating process is AR(1) and the case of $\lambda_1 = 0$, $\lambda_2 \neq 0$ and $\lambda_3 \neq 0$ implies that the data generating process is AR(2). Thus, in any case, even if the true model is different from the estimated model, we can obtain the bias-corrected coefficient estimate based on the median- and mean-unbiased estimators.

5 Empirical Example

In this section, based on actually observed annual data from 1956 to 1996, Japanese consumption function is estimated as an empirical example of the proposed estimators. We consider the AR(1) model with constant term and exogenous variable x_t, which is specified as follows:

$$y_t = \beta_1 + \beta_2 x_t + \alpha_1 y_{t-1} + u_t,$$

where $u_t \sim N(0, \sigma^2)$ is assumed. y_t and x_t represent consumption and income, respectively.

As for the initial value of y_t, the actual consumption data of 1956 is used. The estimation period is from 1957 to 1996. The following consumption and income data are used for y_t and x_t.

y_t = Domestic Final Consumption Expenditure of Households (Billions of Japanese Yen at Market Prices in 1990)

x_t = National Disposable Income of Households (Billions of Japanese Yen at Market Prices in 1990)

Table 4: Example: Japanese Consumption Function

Estimation Method		β_1	β_2	α_1
OLSE	$\widehat{\theta}$	3301 (1132)	.1323 (.0586)	.8556 (.0684)
	$\widetilde{\theta}$	3492 (1067)	.1101 (.0461)	.8817 (.0544)
Median-	2.5%	1138	.0352	.7557
Unbiased	5.0%	1552	.0459	.7790
Estimator	50.0%	3495	.1103	.8818
	95.0%	5036	.1960	.9565
	97.5%	5354	.2167	.9695
	$\overline{\theta}$	3546 (1035)	.1054 (.0454)	.8871 (.0536)
Mean-	2.5%	1283	.0263	.7687
Unbiased	5.0%	1693	.0376	.7916
Estimator	50.0%	3595	.1012	.8925
	95.0%	5076	.1854	.9665
	97.5%	5376	.2061	.9796

The data are taken from the *Annual Report on National Accounts* (the Economic Planning Agency, Government of Japan, 1998). In the economic interpretation, β_2 is known as the marginal propensity to consume. Using Japanese data, β_1, β_2 and α_1 are estimated using three estimators, i.e., OLSE, the median-unbiased estimator and the mean-unbiased estimator.

Under the setup, the estimation results of Japanese classical consumption function are in Table 4. $\widehat{\theta}$, $\widetilde{\theta}$ and $\overline{\theta}$ denote OLSE, the median-unbiased estimate and the mean-unbiased estimate, respectively. Each value in the parentheses indicates the standard error of the corresponding estimate. 2.5%, 5.0%, 50.0%, 95.0% and 97.5% represent the percentiles obtained from n median- or mean-unbiased estimates, where $n = 10,000$ is taken (see Section 3 for derivation of the percentiles).

From Table 4, 95% confidence intervals of β_1 (i.e., the constant term) are obtained as 1138 – 5354 for the median-unbiased estimator and 1283 – 5376 for the mean-unbiased estimator. 95% confidence intervals of β_2 (i.e., the marginal propensity to consume) are .0352 – .2167 for the median-unbiased estimator and .0263 – .2061 for the mean-unbiased estimator. 95% confidence intervals of α_1 (i.e., the AR(1) coefficient) are .7557 – .9695 for the median-unbiased estimator and .7687 – .9796 for the mean-unbiased estimator. We can see skewness of the distribution by computing the distance between 2.5% and 97.5% values (or 5.0% and 95% values). Judging from both median- and mean-unbiased estimators, the distribution is skewed to the right for the estimates of β_1 and α_1, and to the left for the estimates of β_2. That is, OLSE's of β_1 and α_1 are underestimated while OLSE of β_2 is overestimated. Moreover, we can observe that all the standard errors obtained from OLSE are overestimated compared with those from the median- and mean-unbiased estimators. The standard error obtained from OLSE is computed by the conventional formula, but the standard error from the median- or mean-unbiased estimator is based on the simulation technique. Accordingly, it might be appropriate to consider that the standard error from OLSE is also biased.

6 Summary

It is well known that OLSE yields the biased estimator when it is applied to the autoregressive models, which is displayed in Figures 1 and 2. In the classical regression theory, we do not have the biased OLSE in the case of inclusion of irrelevant variables. However, in the case of the autoregressive models, the OLSE bias is serious as the number of unnecessary exogenous variables increases, at least when the true model is given by the AR(1) model. In order to improve the biased estimator, in this paper, we have proposed the estimation procedure of the median-unbiased and the mean-unbiased estimators using the simulation technique. In practice, the data generating process is not known. Using the simulation technique, an attempt is made to estimate the unknown parameters correctly even if we do not know the true model. The proposed estimation procedure can be applied to the higher-order autoregressive models with the other exogenous variables and furthermore it is also applied to the nonnormal models although we have not taken any example of the nonnormal models.

We have examined several Monte Carlo simulation studies, where OLSE (i.e., $\widehat{\alpha}_1$, $\widehat{\alpha}_2$ and $\widehat{\alpha}_3$), the median-unbiased estimator (i.e., $\widetilde{\alpha}_1$, $\widetilde{\alpha}_2$ and $\widetilde{\alpha}_3$) and the mean-unbiased estimator (i.e., $\overline{\alpha}_1$, $\overline{\alpha}_2$ and $\overline{\alpha}_3$) are compared with the true parameter value (i.e., α_1, α_2 and α_3). In Table 1. the first-order autoregressive models are estimated by Models 1 – 3 when the true model is given by Model 1. In Tables 2 and 3, we estimate the higher-order autoregressive models by Model 1 when the true model is given by Model 1. Judging from Tables 1 – 3, for all the cases the OLSE bias is correctly improved using the median- and mean-unbiased estimators.

Finally, we have shown estimation of Japanese consumption function as an empirical example. The AR(1) model which includes both a constant term and an exogenous variable x_t has been estimated, where the standard error of each coefficient estimate, and the 90% and 95% confidence intervals are shown.

References

Abadir, K.M. (1993) " OLS Bias in A Nonstationary Autoregression, " *Econometric Theory*, Vol.9, No.1, pp.81 – 93.

Ali, M.M. (1996) " Distribution of the Least Squares Estimator in a First-Order Autoregressive Model, " unpublished manuscript (Economics Working Paper Archive, ftp:// econwpa.wustl.edu/ econ-wp/ em/ papers/ 9610/ 9610004.ps).

Andrews, D.W.K. (1993) " Exactly Median-Unbiased Estimation of First Order Autoregressive / Unit Root Models, " *Econometrica*, Vol.61, No.1, pp.139 – 165.

Andrews, D.W.K. and H.Y. Chen (1994) " Approximately Median-Unbiased Estimation of Autoregressive Models, " *Journal of Business and Economic Statistics*, Vol.12, No.2, pp.187 – 204.

Diebold, F.X. and G.D. Rudebusch (1991) " Unbiased Estimation of Autoregressive Models, " unpublished manuscript, University of Pennsylvania.

Greene, W.H. (1993) *Econometric Analysis* (Second Edition), Prentice Hall.

Grubb, D. and J. Symons (1987) " Bias in Regressions with a Lagged Dependent Variable, " *Econometric Theory*, Vol.3, pp.371 – 386.

Hurwicz, L. (1950) " Least-Squares Bias in Time Series, " in *Statistical Inference in Dynamic Economic Models*, ed. T.C. Koopmans, New York. John Wiley, pp.365 – 383.

Imhof, J.P. (1961) " Computing the Distribution of Quadratic Forms in Normal Variates, " *Biometrika*, Vol.48, pp.419 – 426.

Kendall, M.G. (1954) " Note on Bias in the Estimation of Autocorrelations, " *Biometrika*, Vol.41, pp.403 – 404.

Maekawa, K. (1983) " An Approximation to the Distribution of the Least Squares Estimator in an Autoregressive Model with Exogenous Variables, " *Econometrica*, Vol.51, No.1, pp.229 – 238.

Maekawa, K. (1987) " Finite Sample Properties of Several Predictors from an Autoregressive model, " *Econometric Theory*, Vol.3, pp.359 – 370.

Marriott, F.H.C. and J.A. Pope (1954) " Bias in the Estimation of Autocorrelations, " *Biometrika*, Vol.41, pp.390 – 402.

Orcutt, G.H. and H.S. Winokur (1969) " First Order Autoregression: Inference, Estimation, and Prediction, " *Econometrica*, Vol.37, No.1, pp.1 – 14.

Peters, T.A. (1989) “ The Exact Moments of OLS in Dynamic Regression Models with Non-Normal Errors, ” *Journal of Econometrics*, Vol.40, pp.279 – 305.

Quenouille, M.H. (1956) “ Notes on Bias in Estimation, ” *Biometrika*, Vol.43, pp.353 – 360.

Sawa, T. (1978) “ The Exact Moments of the Least Squares Estimator for the Autoregressive Model, ” *Journal of Econometrics*, Vol.8, pp.159 – 172.

Shaman, P. and R.A. Stine (1988) “ The Bias of Autoregressive Coefficient Estimators, ” *Journal of the American Statistical Association*, Vol.83, No.403, pp.842 – 848.

Tanaka, K. (1993) “ Asymptotic Expansions Associated with AR(1) Model with Unknown Mean, ” *Econometrica*, Vol.51, No.4, pp.1221 – 1231.

Tanizaki, H. (1995) “ Asymptotically Exact Confidence Intervals of CUSUM and CUSUMSQ Tests: A Numerical Derivation Using Simulation Technique, ” *Communications in Statistics, Simulation and Computation*, Vol.24, No.4, pp.1019 – 1036.

Tse, Y.K. (1982) “ Edgeworth Approximations in First-Order Stochastic Difference Equations with Exogenous Variables, ” *Journal of Econometrics*, Vol.20, pp.175 – 195.

Tsui, A.K. and M.M. Ali (1994) “ Exact Distributions, Density Functions and Moments of the Least Squares Estimator in a First-Order Autoregressive Model, ” *Computational Statistics & Data Analysis*, Vol.17, No.4, pp.433 – 454.

White, J.S. (1961) “ Asymptotic Expansions for the Mean and Variance of the Serial Correlation Coefficient, ” *Biometrika*, Vol.48, pp.85 – 94.