A Note on Bias in First-Differenced AR(1) Models

Kazuhiko Hayakawa
Hitotsubashi University

Abstract

In this note, we derive the finite sample bias of the modified ordinary least squares (MOLS) estimator, which was suggested by Wansbeek and Knaap (1999) and reconsidered by Hayakawa (2006a,b). From the formula for the finite sample bias, we find that the bias of the MOLS estimator becomes small as ρ, the autoregressive parameter, approaches unity. Simulation results indicate that the MOLS estimator has very small bias and that its empirical size is close to the nominal one.

Financial support from the JSPS Fellowship is also gratefully acknowledge. All remaining errors are my own.
Citation: Hayakawa, Kazuhiko, (2006) "A Note on Bias in First-Differenced AR(1) Models." *Economics Bulletin,* Vol. 3, No. 27 pp. 1-10
Submitted: June 27, 2006. **Accepted:** October 4, 2006.
URL: http://economicsbulletin.vanderbilt.edu/2006/volume3/EB-06C20015A.pdf

1 Introduction

In the context of a pure time series AR(1) model, it is well known that the bias of the ordinary least squares (OLS) estimator is not negligible when T, the sample size of a time series, is not large. Early contributions to this small sample bias problem include Orcutt (1948), Hurvicz (1950), Marriott and Pope (1954), and Kendall (1954), and more recent contributions include Phillips (1977), Tanaka (1983), and Shaman and Stine (1988). Their major finding is that the bias of the OLS estimator becomes large when T is small and ρ, an autoregressive parameter, is close to one.

In recent years, several papers have appeared that deal with the estimation and inference of a panel AR(1) model with short time series. [1] One of them includes the modified OLS (MOLS) estimator by Wansbeek and Knaap (1999) and Hayakawa (2006a,b). The major finding of these papers is that the MOLS estimator has very small bias even when T is not so large, and ρ is close to one.

Phillips and Han (2006) consider the same estimator as the MOLS estimator in the context of a time series AR(1) model. They showed by simulations that the MOLS estimator has a very small bias. In this note, we derive the finite sample bias of the MOLS estimator in the context of time series AR(1) models, and show the reason why the MOLS estimator has a small bias when T is not large and ρ is close to unity. This result helps to explain theoretically why the MOLS estimator has small bias in the estimation of dynamic panel data models with cross-section dependence (Hayakawa, 2006b), and in the estimation of time series AR(1) models (Phillips and Han, 2006).

The remainder of this note is organized as follows. In Section 2, we provide the setup and main results of this note. In Section 3, we compare the performance of the MOLS estimator and alternative estimators by a Monte Carlo simulation. Finally, Section 4 concludes.

[1] For a recent review, see Arellano (2003).

2 Setup and Main Results

We consider an AR(1) model given by

$$y_t = \mu + \rho y_{t-1} + +u_t \qquad t = 1, ..., T \tag{1}$$

where ρ is the parameter of interest with $|\rho| < 1$. We assume that $u_t \sim iidN(0, \sigma^2)$, and $y_0 = \mu/(1-\rho) + \sum_{j=0}^{\infty} \rho^j u_{-j}$.

By first-differencing model (1), we have

$$\Delta y_t = \rho \Delta y_{t-1} + \Delta u_t \qquad t = 2, ..., T. \tag{2}$$

The OLS estimator of this model is given by

$$\hat{\rho}_{fdols} = \frac{T_0^{-1} \sum_{t=2}^{T} \Delta y_{t-1} \Delta y_t}{T_0^{-1} \sum_{t=2}^{T} \Delta y_{t-1}^2} = \frac{X}{Y} \tag{3}$$

where $T_0 = T - 1$.

The MOLS estimator suggested by Wansbeek and Knaap (1999) and reconsidered by Hayakawa (2006a,b) takes the following form:

$$\hat{\rho}_{mols} = 2\hat{\rho}_{fdols} + 1. \tag{4}$$

It is easy to show that $\text{plim}_{T\to\infty} \hat{\rho}_{fdols} = (\rho - 1)/2$, and $\text{plim}_{T\to\infty} \hat{\rho}_{mols} = \rho$.

The following theorem gives the formulas of the finite sample biases of $\hat{\rho}_{fdols}$ and $\hat{\rho}_{mols}$.

Theorem 1. *The expectations of $\hat{\rho}_{fdols}$, and $\hat{\rho}_{mols}$ up to $O(T^{-1})$ are given by*

$$E(\hat{\rho}_{fdols}) = \frac{\rho - 1}{2} \left[1 - \frac{1}{T-1} \right] + o(T^{-1}) \tag{5}$$

$$E(\hat{\rho}_{mols}) = \rho + \frac{1-\rho}{T-1} + o(T^{-1}). \tag{6}$$

The proof is given in the appendix.

Remark 1 We find that as ρ approaches unity, the finite sample bias of $\hat{\rho}_{mols}$ becomes small. This is in contrast to the usual OLS estimator whose bias increases as ρ approaches unity.[2] Figure 1 depicts the bias of $\hat{\rho}_{ols}$ and $\hat{\rho}_{mols}$ for the case of $T = 25$. From this figure, we find that the bias of $\hat{\rho}_{mols}$ is much smaller than that of $\hat{\rho}_{ols}$, and tends to be small as ρ approaches one. This supports the simulation results of Phillips and Han (2006).

[2] The form of the finite sample bias of the OLS estimator in (1), $\hat{\rho}_{ols}$, is given by $E(\hat{\rho}_{ols}) - \rho = -(1+3\rho)/T$. See Marriott and Pope (1954) and Tanaka (1983).

Remark 2 This result explains the reason why the bias of the MOLS estimator in the estimation of dynamic panel data models with cross-section dependence is small when T is not so large and ρ is close to one. For a detailed discussion, see Hayakawa (2006b).

3 Monte Carlo Simulation

In this section, we compare the performance of the MOLS estimator with those of the OLS estimator and the recursive mean adjusting (RMA) estimator by So and Shin (1999), in terms of the bias and inference.[3] Observations are simulated from $y_t = 1 + \rho y_{t-1} + u_t$ with $u_t \sim iidN(0,1)$ and $y_1 \sim iidN(1/(1-\rho), 1/(1-\rho^2))$. We set $T = 15, 20, 25, 50, 100, 200$, and $\rho = 0.5, 0.9, 0.95$. The number of replications is 100,000. We computed $\hat{\rho}_{ols}$, $\hat{\rho}_{mols}$, the recursive mean adjusting estimator, $\hat{\rho}_{rma}$, and bias-corrected versions of $\hat{\rho}_{ols}$, $\hat{\rho}_{mols}$,

$$\tilde{\rho}_{ols} = \frac{T\hat{\rho}_{ols} + 1}{T-3} \tag{7}$$

$$\tilde{\rho}_{mols} = \frac{(T-1)\hat{\rho}_{mols} - 1}{T-2}. \tag{8}$$

It is easy to verify that $E(\tilde{\rho}_{ols}) = E(\tilde{\rho}_{mols}) = \rho + o(T^{-1})$. $\tilde{\rho}_{ols}$, and $\tilde{\rho}_{mols}$ are useful in bias reduction when we are just interested in the estimation of AR(1) models. However, in some cases, they are not. For example, $\hat{\rho}_{ols}$ and $\hat{\rho}_{mols}$ appear implicitly in the estimators of dynamic panel data models with cross-section dependence. From Proposition 3 in Phillips and Sul (2006), we find that $\hat{\rho}_{ols}$ appears implicitly in the probability limit of the least squares dummy variables estimator. In this case, we cannot use $\tilde{\rho}_{ols}$. Similarly, as shown in Hayakawa (2006b), $\hat{\rho}_{mols}$ appears implicitly in the MOLS estimator of dynamic panel data models with cross-section dependence. Thus, in these cases, we cannot use $\tilde{\rho}_{ols}$ or $\tilde{\rho}_{mols}$.

The simulation results are summarized in Table 1. We computed mean (mean), the root mean squared error (RMSE) and the empirical size of t-test for $H_0 : \rho = \rho_0$ with 5% significance.

When we compare $\hat{\rho}_{ols}$, $\hat{\rho}_{mols}$, and $\hat{\rho}_{rma}$, we find that $\hat{\rho}_{mols}$ and $\hat{\rho}_{rma}$ have almost the same bias in the case of $\rho = 0.5$. However, when $\rho = 0.9, 0.95$, the bias of $\hat{\rho}_{mols}$ is smallest

[3] Sul, Phillips, and Choi (2005, p.540–542) gives an intuitive reason why the RMA estimator has smaller bias than the OLS estimator.

among the three estimators, especially when T is not so large. With regard to inference, the empirical size of $\hat{\rho}_{mols}$ is very close to the nominal one in all cases, although those of $\hat{\rho}_{ols}$ and $\hat{\rho}_{rma}$ are not so close in the case of small T. One drawback of $\hat{\rho}_{mols}$ is its variability. Since the variance of $\hat{\rho}_{mols}$ is four times as large as that of $\hat{\rho}_{fdols}$, in terms of the RMSE, $\hat{\rho}_{mols}$ does not perform best. In other words, at a cost of efficiency, $\hat{\rho}_{mols}$ gains precision with regard to the bias.

When we compare all five estimators, we find that the bias of $\tilde{\rho}_{mols}$ is smallest in almost all the cases, although both $\tilde{\rho}_{mols}$ and $\tilde{\rho}_{ols}$ are unbiased up to $O(T^{-1})$. It might be conjectured that this is because the third-order bias of $\hat{\rho}_{ols}$ is not negligible when ρ is close to one. In terms of inference, the empirical sizes of $\hat{\rho}_{mols}$ and $\tilde{\rho}_{mols}$ are close to the nominal one in almost all cases, although those of other estimators are not, especially when T is small. With regard to the RMSE, the RMSE of $\tilde{\rho}_{mols}$ is largest among the five estimators, although its bias is smallest. Hence, as in the above case, $\tilde{\rho}_{mols}$ has very a small bias at a cost of efficiency.

4 Conclusion

In this paper, we derived the finite sample bias of the modified OLS estimator in the context of a time series AR(1) model. From the formula of the finite sample bias, we found that as ρ approaches unity, the bias becomes small, unlike the usual OLS estimator. This supported the simulation results of Phillips and Han (2006) theoretically. This result was also useful in explaining why the MOLS estimator performs well in the estimation of dynamic panel data models with cross-section dependence even if T is not so large. Simulation results showed that although the RMSE of the MOLS estimator is larger than those of other alternative estimators, in terms of the bias and inference, the MOLS estimator performs best.

References

[1] Arellano M. (2003) *Panel Data Econometrics*, Oxford University Press, New York.

[2] Brockwell, P.J. and R.A. Davis (1991) *Time Series: Theory and Methods*, 2nd ed. Springer-Verlag, New York.

[3] Hayakawa, K. (2006a) "Consistent OLS Estimation of AR(1) Dynamic Panel Data Models with Short Time Series", forthcoming in *Applied Economics Letters*.

[4] Hayakawa, K. (2006b) "A Modified OLS Estimator in AR(1) Dynamic Panel Data Models with Cross Section Dependence and Heteroskedasticity," mimeo.

[5] Hurvicz, L. (1950). "Least Squares Bias in Time Series". In T. Koopmans (ed.), *Statistical Inference in Dynamic Economic Models*. 365–83. New York: Wiley.

[6] Kendall, M.G. (1954) "Note on the Bias in the Estimation of Autocorrelation", *Biometrika*, 41, 403–404.

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

[8] Mood, A.M., F.A. Graybill, and D.C. Boes (1974) *Introduction to the Theory of Statistics*, 3rd ed. McGraw-Hill, Singapore.

[9] Orcutt, G.H. (1948) "A Study of the Autoregressive Nature of the Times Series Used for Tinbergenfs Model of the Economic System of the United States". *Journal of the Royal Statistical Society*, Series B 1-45.

[10] Phillips, P.C.B. (1977) "Approximations to Some Finite Sample Distributions Associated with a First Order Stochastic Difference Equation", *Econometrica*, 45, 463–485.

[11] Phillips, P.C.B. and D. Sul (2006) "Bias in Dynamic Panel Estimation with Fixed Effects, Incidental Trends and Cross Section Dependence," forthcoming in *Journal of Econometrics*.

[12] Phillips, P.C.B. and C. Han (2006) "Gaussian Inference in AR(1) Time Series with or without a Unit Root", Cowles Foundation Discussion Paper, No.1546, Yale University.

[13] Shaman, P. and R.A. Stine (1988) "The Bias of Autoregressive Coefficient Estimators", *Journal of the American Statistical Association*, 83, 842–848.

[14] So, B.S. and Shin, D.W. (1999a) "Recursive Mean Adjustment in Time Series Inferences", *Statistics and Probability Letters*, 43, 65–73.

[15] Sul, D., P.C.B. Phillips, and C.Y. Choi (2005) "Prewhitening Bias in HAC Estimation" *Oxford Bulletin of Economics and Statistics*, 67, 517–546.

[16] Tanaka, K. (1983) "Asymptotic Expansion Associated with the AR(1) Model with Unknown Mean", *Econometrica*, 51, 1221–1232.

[17] Wansbeek, T.L. and T. Knaap (1999) "Estimating a Dynamic Panel Data Model with Heterogeneous Trend", *Annales D'Économie et de Statistique*, 55–56, 331–349.

A Appendix

Proof of Theorem 1 We follow Marriott and Pope (1954). Note that $E(X/Y)$ can be expanded up to $O(T^{-1})$ as follows:[4]

$$E\left(\frac{X}{Y}\right) = \frac{E(X)}{E(Y)}\left[1 - \frac{cov(X,Y)}{E(X)E(Y)} + \frac{var(Y)}{[E(Y)]^2}\right] + o(T^{-1}). \quad (9)$$

To derive the finite sample bias, we need to obtain $E(Y^2)$ and $E(XY)$:

$$\begin{aligned} E(Y^2) &= \frac{1}{T_0^2}E\left[\left(\sum_{t=2}^{T} y_{t-1}^2\right)^2 + \left(\sum_{t=2}^{T} y_{t-2}^2\right)^2 + 4\left(\sum_{t=2}^{T} y_{t-2}y_{t-1}\right)^2\right. \\ &\quad +2\left(\sum_{t=2}^{T} y_{t-2}^2\right)\left(\sum_{t=2}^{T} y_{t-1}^2\right) + -4\left(\sum_{t=2}^{T} y_{t-2}^2\right)\left(\sum_{t=2}^{T} y_{t-2}y_{t-1}\right) \\ &\quad \left. + - 4\left(\sum_{t=2}^{T} y_{t-1}^2\right)\left(\sum_{t=2}^{T} y_{t-2}y_{t-1}\right)\right] \\ &= A_1 + A_2 + 4A_3 + 2A_4 - 4A_5 - 4A_6 \end{aligned}$$

$$\begin{aligned} E(XY) &= \frac{1}{T_0^2}E\left[\left(\sum_{t=2}^{T} y_{t-1}^2\right)\left(\sum_{t=2}^{T} y_{t-1}y_t\right) - \left(\sum_{t=2}^{T} y_{t-1}^2\right)^2 - \left(\sum_{t=2}^{T} y_{t-1}^2\right)\left(\sum_{t=2}^{T} y_{t-2}y_t\right)\right. \\ &\quad + \left(\sum_{t=2}^{T} y_{t-1}^2\right)\left(\sum_{t=2}^{T} y_{t-2}y_{t-1}\right) - \left(\sum_{t=2}^{T} y_{t-2}^2\right)\left(\sum_{t=2}^{T} y_{t-1}y_t\right) \end{aligned}$$

[4]See, for example, Mood, Graybill and Boes (1974, p.181).

$$
\begin{aligned}
& -\left(\sum_{t=2}^{T} y_{t-2}^2\right)\left(\sum_{t=2}^{T} y_{t-1}^2\right) + \left(\sum_{t=2}^{T} y_{t-2}^2\right)\left(\sum_{t=2}^{T} y_{t-2}y_t\right) \\
& -2\left(\sum_{t=2}^{T} y_{t-2}^2\right)\left(\sum_{t=2}^{T} y_{t-2}y_{t-1}\right) + 2\left(\sum_{t=2}^{T} y_{t-2}y_{t-1}\right)\left(\sum_{t=2}^{T} y_{t-1}y_t\right) \\
& +2\left(\sum_{t=2}^{T} y_{t-2}y_{t-1}\right)\left(\sum_{t=2}^{T} y_{t-1}^2\right) + 2\left(\sum_{t=2}^{T} y_{t-2}y_{t-1}\right)\left(\sum_{t=2}^{T} y_{t-2}y_t\right) \\
& \left. -2\left(\sum_{t=2}^{T} y_{t-2}y_{t-1}\right)^2\right] \\
= \; & B_1 - B_2 - B_3 + B_4 - B_5 - B_6 + B_7 - 2B_8 + 2B_9 + 2B_{10} + 2B_{11} - 2B_{12}
\end{aligned}
$$

To calculate the expectations, we use the following result:[5]

$$
E(y_t y_{t+k} y_{t+k+l} y_{t+k+l+m}) = r\rho^{k+m}(1 + 2\rho^{2l}) \qquad k, l, m \geq 0 \tag{10}
$$

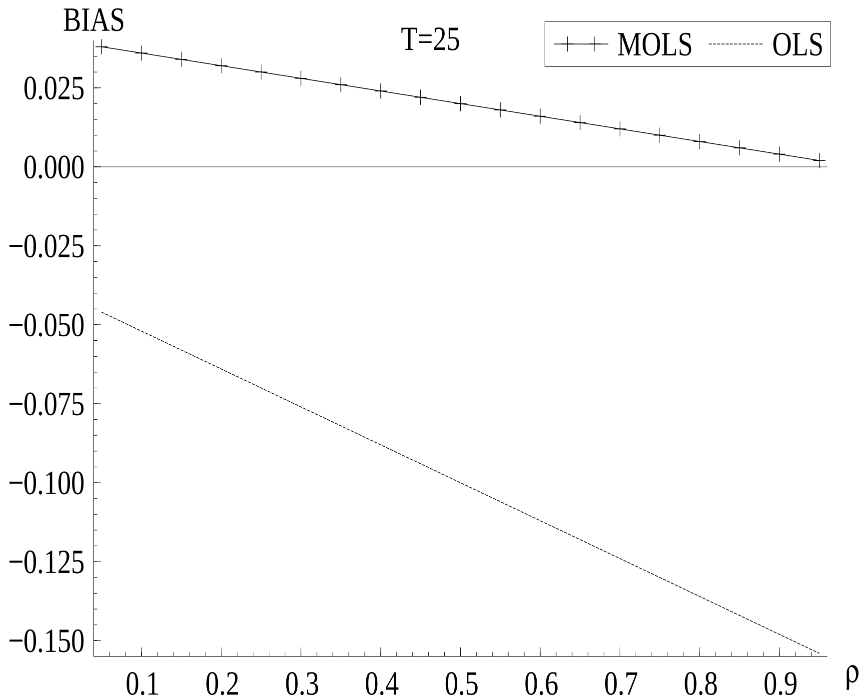
where $r = \sigma^4/(1-\rho^2)^2$. Using (10), we have

$$
\begin{aligned}
A_1 = A_2 = A_4 = B_2 = B_6 &= r + \frac{2r}{T_0}\left(\frac{1+\rho^2}{1-\rho^2}\right) \\
A_3 = B_9 = B_{12} &= \rho^2 r + \frac{1}{T_0}\left(\frac{4\rho^2 r}{1-\rho^2} + (1+\rho^2)r\right) \\
A_5 = A_6 = B_1 = B_4 = B_5 = B_8 = B_{10} &= \rho r + \frac{1}{T_0}\left(\frac{4\rho r}{1-\rho^2}\right) \\
B_3 = B_7 &= \rho^2 r + \frac{2\rho^2 r}{T_0}\left(1 + \frac{2}{1-\rho^2}\right) \\
B_{11} &= \rho^3 r + \frac{2\rho r}{T_0}\left((1+\rho^2) + \frac{2\rho^2}{1-\rho^2}\right).
\end{aligned}
$$

From these results, we have

$$
\begin{aligned}
E(Y^2) &= 4r\left[(1-\rho^2) + \frac{1}{T_0}\frac{(1-\rho)^2(\rho+3)}{1+\rho}\right] && (11) \\
E(XY) &= 2r\left[(\rho-1)^3 + \frac{-2}{T_0}\left(\frac{(\rho-1)^4(\rho+2)}{1-\rho^2}\right)\right] && (12) \\
var(Y) &= \frac{4r}{T_0}\frac{(1-\rho)^2(\rho+3)}{1+\rho} && (13) \\
cov(X,Y) &= \frac{-4r}{T_0}\left(\frac{(\rho-1)^4(\rho+2)}{1-\rho^2}\right). && (14)
\end{aligned}
$$

[5] See Marriott and Pope (1954), and Brockwell and Davis (1991, p.226–227).



Therefore, it follows that

$$E(\hat{\rho}_{fdols}) = E\left(\frac{X}{Y}\right) = \frac{\rho - 1}{2}\left[1 - \frac{1}{T-1}\right] + o(T^{-1}). \tag{15}$$

$E(\hat{\rho}_{mols})$ is also straightforwardly derived from (15).

Table 1: Simulation Results

Case		Mean					RMSE					Size				
ρ	T	$\hat{\rho}_{ols}$	$\tilde{\rho}_{ols}$	$\hat{\rho}_{mols}$	$\tilde{\rho}_{mols}$	$\hat{\rho}_{rma}$	$\hat{\rho}_{ols}$	$\tilde{\rho}_{ols}$	$\hat{\rho}_{mols}$	$\tilde{\rho}_{mols}$	$\hat{\rho}_{rma}$	$\hat{\rho}_{ols}$	$\tilde{\rho}_{ols}$	$\hat{\rho}_{mols}$	$\tilde{\rho}_{mols}$	$\hat{\rho}_{rma}$
0.5	15	0.338	0.505	0.530	0.494	0.469	0.292	0.303	0.447	0.480	0.279	0.100	0.098	0.053	0.072	0.100
0.5	20	0.377	0.502	0.525	0.499	0.481	0.243	0.246	0.387	0.408	0.232	0.090	0.091	0.052	0.065	0.092
0.5	25	0.401	0.501	0.519	0.498	0.488	0.209	0.210	0.344	0.359	0.201	0.081	0.084	0.050	0.061	0.085
0.5	50	0.451	0.501	0.510	0.500	0.499	0.136	0.135	0.244	0.249	0.133	0.064	0.066	0.050	0.055	0.070
0.5	100	0.475	0.500	0.506	0.501	0.502	0.092	0.091	0.173	0.174	0.091	0.057	0.058	0.049	0.051	0.062
0.5	200	0.488	0.500	0.503	0.501	0.503	0.063	0.063	0.122	0.123	0.063	0.054	0.055	0.049	0.051	0.058
0.9	15	0.640	0.883	0.908	0.901	0.812	0.347	0.288	0.491	0.529	0.258	0.225	0.129	0.053	0.065	0.095
0.9	20	0.699	0.881	0.905	0.900	0.834	0.274	0.220	0.427	0.451	0.200	0.197	0.124	0.053	0.060	0.083
0.9	25	0.738	0.884	0.904	0.900	0.849	0.226	0.179	0.383	0.399	0.165	0.176	0.117	0.052	0.059	0.075
0.9	50	0.819	0.893	0.903	0.901	0.880	0.123	0.098	0.271	0.277	0.091	0.120	0.100	0.050	0.053	0.064
0.9	100	0.860	0.897	0.902	0.901	0.893	0.069	0.058	0.195	0.197	0.055	0.089	0.084	0.051	0.053	0.061
0.9	200	0.881	0.899	0.900	0.900	0.899	0.041	0.036	0.137	0.138	0.035	0.071	0.072	0.050	0.051	0.062
0.95	15	0.671	0.923	0.955	0.952	0.849	0.360	0.287	0.497	0.536	0.259	0.279	0.132	0.054	0.064	0.097
0.95	20	0.734	0.922	0.950	0.947	0.872	0.284	0.218	0.432	0.456	0.200	0.251	0.132	0.054	0.061	0.086
0.95	25	0.775	0.927	0.953	0.951	0.890	0.233	0.176	0.386	0.403	0.163	0.226	0.128	0.051	0.057	0.077
0.95	50	0.861	0.937	0.951	0.950	0.922	0.124	0.093	0.276	0.282	0.087	0.169	0.115	0.049	0.052	0.063
0.95	100	0.907	0.945	0.950	0.950	0.939	0.066	0.051	0.196	0.198	0.048	0.122	0.101	0.051	0.052	0.059
0.95	200	0.929	0.948	0.951	0.951	0.947	0.036	0.030	0.139	0.140	0.028	0.090	0.082	0.050	0.050	0.057