Auburn University
Department of Economics
Working Paper Series

Bias Correction and Out-of-Sample Forecast Accuracy

Hyeongwoo Kim and Nazif Durmaz

Auburn University

AUWP 2010-02

Bias Correction and Out-of-Sample Forecast Accuracy*

Hyeongwoo Kim† and Nazif Durmaz‡

Auburn University

May 2010

Abstract

We evaluate the usefulness of bias-correction methods for autoregressive (AR) models in terms of out-of-sample forecast accuracy, employing two popular methods proposed by Hansen (1999) and So and Shin (1999). Our Monte Carlo simulations show that these methods do not necessarily achieve better forecasting performances than the bias-uncorrected Least Squares (LS) method, because bias correction tends to increase the variance of the estimator. There is a gain from correcting for bias only when the true data generating process is sufficiently persistent. Though the bias arises in finite samples, the sample size (N) is not a crucial factor of the gains from bias-correction, because both the bias and the variance tend to decrease as N goes up. We also provide a real data application with 7 commodity price indices which confirms our findings.

Keywords: Small-Sample Bias, Grid Bootstrap, Recursive Mean Adjustment, Out-of-Sample Forecast

JEL Classification: C52, C53

---

*We thank Ken West and Henry Thompson for helpful comments. Special thanks go to two anonymous referees who provided extremely useful suggestions.

†Department of Economics, Auburn University, 0339 Haley Center, Auburn, AL 36849. Tel: 334-844-2928. Fax: 334-844-4615. Email: gmmkim@gmail.com

‡Department of Agricultural Economics and Rural Sociology, Auburn University, Auburn, AL 36849. Tel: 334-844-1949. Fax: 334-844-5639. Email: nzd0001@auburn.edu

1 Introduction

It is a well-known statistical fact that the least squares (LS) estimator for autoregressive (AR) processes suffers from serious downward bias in the persistence coefficient when the stochastic process includes a non-zero intercept and/or deterministic time trend. The bias can be substantial especially when the stochastic process is highly persistent (Andrews, 1993).

Since the seminal work of Kendall (1954), an array of bias-correction methods has been put forward. To name a few, Andrews (1993) proposed a method to obtain the exactly median-unbiased estimator for an AR(1) process with Gaussian errors.[1] Andrews and Chen (1994) extends the work of Andrews (1993) to get approximately median-unbiased estimator for higher order AR(p) processes. Hansen (1999) developed the grid bootstrap (GB) method, a nonparametric bias correction method, which is robust to distributional assumptions. These median-bias correction methods have been widely employed by many researchers, among others, Kim and Ogaki (2009), Kim (2009), Steinsson (2008), Karanasos *et al.* (2006), and Murray and Papell (2002).

An alternative approach has been also proposed by So and Shin (1999) who develop the recursive mean adjustment (RMA) method that belongs to a class of (approximately) mean-unbiased estimators.[2] The RMA estimator is computationally convenient to implement yet powerful and used in the work of Choi *et al.* (2009), Kim *et al.* (2009), Sul *et al.* (2005), Taylor (2002), and Cook (2002), for instance.

By construction, the LS estimator provides the best in-sample fit among the class of linear estimators notwithstanding its bias, because the LS estimator is obtained by minimizing the sum of squared residuals. Of course, this does not guarantee good out-of-sample forecasting performances. It is not clear either whether correcting for bias will enhance forecasting performances. Even though bias-correction reduces the mean square error (MSE), it tends to increase the variance of the estimator at the same time, resulting in an increase in MSE. If the bias is not large enough to

[1] Mikusheva (2007) shows that Andrews' (1993) method is asymptotically valid for non-Gaussian errors.

[2] Tanaka (1984) and Shaman and Stine (1988) extend Kendall's (1954) exact mean-bias correction method to AR(p) models. However, their methods are computationally complicated when the lag order is large. Further, one may need to derive the bias-correction formula for higher-order processes than AR(6), because Shaman and Stine provide formulae for up to AR(6).

compensate for increased variance, there will be no gains from bias-correction.

A natural question then arises: when do we need to correct for the bias to achieve better out-of-sample forecasting performances? We attempt to answer to this question via Monte Carlo simulation analysis for an array of AR processes. Our analysis shows that bias-correction indeed increases variances of both the GB and the RMA methods (more substantially for the GB method). We find substantial gains from correcting for bias when the *true* data generating process is very persistent, and/or when the forecast-horizon is fairly short. Though the bias arises in finite samples, the number of observations (N) itself is not a crucial factor, because increases in N tend to reduce both the bias and the variance.

As a real data application, we employ the GB and the RMA approaches along with the LS estimator to out-of-sample forecast quarterly commodity price indices for the period of 1974.Q1 to 2008.Q3, obtained from the Commodity Research Bureau (CRB). We find that both bias correction methods overall outperform the LS estimator with an exception of the Textile CRB index that exhibits fairly low persistence. These findings are consistent with our Monte Carlo simulation results.

Organization of the paper is as follows. In Section 2, we explain the source of bias and how each method corrects for the bias. Section 3 reports our simulation results. In Section 4, we discuss our major empirical findings from our real data experiments. Section 4 concludes.

2 Bias-Correction Methods

Consider the following AR(1) process with an intercept.

$$y_t = c + \rho y_{t-1} + \varepsilon_t, \tag{1}$$

where $|\rho| < 1$ and ε_t is a white noise process. The Frisch-Waugh-Lovell theorem implies that estimating ρ by the LS estimator is equivalent to estimating the following.

$$(y_t - \bar{y}) = \rho\,(y_{t-1} - \bar{y}) + \varepsilon_t, \tag{2}$$

where $\bar{y} = T^{-1}\sum_{j=1}^{T} y_j$. The LS estimator for ρ, however, is biased because ε_t is correlated with the explanatory variable $(y_{t-1} - \bar{y})$.[3] Therefore, the LS estimator for AR processes with an intercept creates the mean-bias. The bias has an analytical representation, and as Kendall (1954) shows, the LS estimator $\hat{\rho}_{LS}$ is biased downward.[4]

There is no analytical representation of the median-bias. Monte Carlo simulations, however, can easily demonstrate that the LS estimator produces significant median-bias for ρ when ρ gets close to unity (see Hansen, 1999).

When ε_t is serially correlated, it is convenient to express (1) as follows.

$$y_t = c + \rho y_{t-1} + \sum_{j=1}^{k} \beta_j \Delta y_{t-j} + u_t, \tag{3}$$

where u_t is a white noise process that generates ε_t.[5]

For Hansen's (1999) grid bootstrap (GB) method, we define the following grid-t statistic.

$$t_N(\rho_i) = \frac{\hat{\rho}_{LS} - \rho_i}{se(\hat{\rho}_{LS})},$$

where $\hat{\rho}_{LS}$ is the LS point estimate for ρ, $se(\hat{\rho}_{LS})$ denotes the corresponding LS standard error, and ρ_i is one of M fine grid points in the neighborhood of $\hat{\rho}_{LS}$. Implementing LS estimations for B bootstrap samples at each of M grid points, we obtain the $\alpha\%$ quantile function estimates, $\hat{q}^*_{N,\alpha}(\rho_i) = \hat{q}^*_{N,\alpha}(\rho_i, \varphi(\rho_j))$, where φ denotes nuisance parameters such as βs that are functions of ρ_i. After smoothing quantile function estimates, the (approximately) median-unbiased estimate is obtained by,

$$\hat{\rho}_G = \rho_i \in R, \text{ s.t. } t_N(\rho_i) = \tilde{q}^*_{N,50\%}(\rho_i),$$

where $\tilde{q}^*_{N,50\%}(\rho_i)$ is the smoothed 50% quantile function estimates obtained from $\hat{q}^*_{N,\alpha}$.[6] To correct

[3] Put differently, ε_t is correlated with y_j, for $j = t, t+1, \cdots, T$, thus with $\bar{y}$.

[4] This may not be true for higher order AR(p) processes with *low* persistence. See Stine and Shaman (1989).

[5] When the stochastic process is of higher order than AR(1), exact bias-correction is not possible because the bias becomes random due to the existence of nuisance parameters. For higher order AR(p) models, the RMA and the GB methods yield approximately mean- and median-unbiased estimators, respectively.

[6] We use the Epanechinikov kernel $K(u) = 3(1-u^2)/4I(|u| \leq 1)$, where $I(\cdot)$ is an indicator function.

for median-bias in β_j estimates, we treat other βs as well as ρ as nuisance parameters and follow the procedures described above.

So and Shin's (1999) recursive mean adjustment (RMA) method utilizes demeaning variables using the partial mean instead of the global mean $\bar{y}$. Rather than implementing the LS for (2), the RMA estimator is obtained by the LS estimator for the following regression equation.

$$(y_t - \bar{y}_{t-1}) = \rho\,(y_{t-1} - \bar{y}_{t-1}) + \eta_t,$$

where $\bar{y}_{t-1} = (t-1)^{-1}\sum_{j=1}^{t-1} y_j$ and $\eta_t = \varepsilon_t + c - (1-\rho)\bar{y}_{t-1}$. Note that ε_t is orthogonal to the recursive mean adjusted regressor $(y_{t-1} - \bar{y}_{t-1})$, which results in a substantial bias reduction for the RMA estimator

$$\hat{\rho}_R = \frac{\sum_{t=2}^{T}(y_{t-1} - \bar{y}_{t-1})(y_t - \bar{y}_{t-1})}{\sum_{t=2}^{T}(y_{t-1} - \bar{y}_{t-1})^2} \tag{4}$$

For a higher order AR process such as (3), the RMA estimator can be similarly obtained as in Hansen's (1999) GB method.

3 Monte Carlo Simulations

We first consider AR(1) models. Our simulation analysis uses fine grid points for $\rho \in [0.100, 1.025]$ and an array of sample sizes, $N \in \{50, 75, 100, 125, 150, 200, 300, 400, 500\}$ with $5,000$ iterations for each pair (ρ, N). The data generating process is specified as $y_t = \rho y_{t-1} + \varepsilon_t$ with ε_t being independently distributed Gaussian innovations. We set the constant term to zero, because the distribution of ρ does not depend on the intercept term (Andrews, 1993). Each replication generates $2N$ observations with the initial value $y_0 = 0$, then discards the first N observations to minimize initialization effects. We report only a subset of our full results to save space.[7]

Figure 1 reports the estimated mean and median bias from AR(1) processes.[8,9] As expected,

[7] Full results are available from authors upon request.

[8] Similar results were obtained from AR(2) processes. For AR(2) models, we controled for nuisance parameters (coefficients on lagged differenced variables) as described in Section 2.

[9] We define mean (median) bias as a difference between the true value of ρ and the mean (median) value of the least squares point estimates from $5,000$ simulations.

both types of bias increase as ρ becomes larger, and become greater when the sample size decreases. For instance, when $\rho = 0.95$ and $N = 100$, the mean bias is -0.043, while the median bias is -0.035. For a pair $(\rho, N) = (0.70, 200)$, bias is still non-negligible but much smaller, -0.016 and -0.013 for the mean and the median bias, respectively.

Figure 1 about here

We also investigate the characteristics of the variance of the estimators, which serves as the other important factor of the mean square error. Figure 2 reports two sets of variance ratios, $V(RMA)/V(LS)$ and $V(GB)/V(LS)$, for the forecasting horizon from 1 to 6. Major findings are as follows.

Correcting for bias overall increases the variance no matter which bias-correction method applies with an exception of the RMA method for highly persistent data with fairly small N. It is interesting to see that relative variances tend to be smaller as ρ increases toward one. Given that bias is greater for more persistent data, this implies that one has to correct for bias when forecasting highly persistent time series variables.

It should be also noted that relative variances overall decrease as N increases, which implies that the gains from correcting for bias depend only weakly to the number of observations, because increases in N tend to reduce both the bias and the variance when bias correction methods (especially for the GB) apply. Relative variances tend to increase as the forecast horizon increases, which implies that bias-correction methods become less attractive for longer-horizon out-of-sample forecasting. Overall relative variances are greater when the GB method applies than RMA.

Figure 2 about here

Next, we compare simulated out-of-sample forecast performances of the bias-correction methods with those of the LS. Results for AR(1) processes are reported in Table 1, while Table 2 reports results for AR(2) models.

Table 1 confirms most of our conjectures above. When the true data generating process is highly persistent, say ρ is 0.975 or above, there are evident gains from correcting for bias no matter which method applies. Even for low persistence data, bias-correction may be needed if the forecast horizon is very short. It is also confirmed that more observations (N) do not necessarily lead to a diminishing role of bias-correction methods for achieving better out-of-sample forecasting performances, because both the bias and the variance tend to decrease as N goes up.

Table 1 about here

In what follows, we provide evidence that these findings are valid for higher order AR processes as well. Since bias-correction seems desirable only when ρ is sufficiently large, we consider fairly persistent AR(2) processes to compare forecast performances for higher order AR models, that is, $\rho \in \{0.65, 0.90, 0.95, 0.99\}$ and $\beta \in \{-0.20, 0.30\}$.[10]

As in AR(1) processes, we find gains from bias-correction only when the true data generating process is persistent. The sample size (N) matters only weakly and the length of the forecast horizon seems important again. The nuisance parameter (β) does not matter because bias-correction is made conditional on the realized β estimates (Hansen 1999).

Table 2 about here

4 Empirical Results

This section provides a real data application to examine the validity of the findings from our Monte Carlo simulation experiments in Section 3.

[10]Note that, for example, $y_t = \rho y_{t-1} + 0.30\Delta y_{t-1} + u_t$ and $y_t = \rho y_{t-1} - 0.20\Delta y_{t-1} + u_t$ share the same degree persistence (ρ), and are equivalent to $y_t = (\rho + 0.30)y_{t-1} - 0.30y_{t-2} + u_t$ and $y_t = (\rho - 0.20)y_{t-1} + 0.20y_{t-2} + u_t$, respectively.

We use 7 quarterly commodity price indices, the CRB Spot Index and its six sub-indices, obtained from the Commodity Research Bureau (CRB) for the period of 1974.Q1 to 2008.Q3.[11,12] The number of lags (k) was chosen by the general-to-specific rule with a maximum lag length of 6. The 95% confidence interval was obtained by $10,000$ bootstrap simulations from the empirical distributions.

Table 3 reports the estimates for the persistence parameter (ρ) in (3). We find that both the RMA and the GB methods yield substantial bias-corrections. For example, the ρ estimate for the Spot Index increases from 0.950 (LS) to 0.969 (RMA) and 0.975 (GB) after correcting for bias. This is far from being negligible because corresponding half-life estimates are 3.378, 5.503, and 6.844 years, respectively.[13] Note that all median-bias corrected (GB) 95% confidence bands for the half-life extend to positive infinity, which occurs often in other studies (see Rossi, 2005, and Murray and Papell, 2002, for similar results), while all 95% confidence intervals by RMA remain finite (see Kim and Moh, 2010, for similar results). Note also that median-unbiased estimates by the GB are not restricted to be less than one, because the GB is based on the local-to-unity framework and allows even mildly explosive processes.[14]

Table 3 about here

We now evaluate and compare the out-of-sample forecasting ability of the LS, the RMA, and the GB methods with two alternative forecasting strategies. First, we utilize first 69 out of 139

[11] In order to reduce noise in the data, we converted monthly frequency raw data to quarterly data by taking end-of-period values. Alternatively, one may use quarterly averages. Averaging time series data, however, creates time aggregation bias as pointed out by Taylor (2001).

[12] There is a known structural break in 1973, the year of the demise of the Bretton Woods system. Because of the devaluation of the US dollar, all CRB indices, which are based on commodity prices in the US dollar term, rapidly adjusted to higher levels in 1973. Since our main objective is to evaluate relative forecast performances of competing estimators, we use observations starting from 1974.Q1 instead of using a dummy variable for the Bretton Woods era.

[13] The half-life was calculated by $\ln(0.5)/\ln(\rho)$, adjusted for the data frequency, assuming that deviations diminish monotonically. This is exactly true for AR(1) processes. For higher order AR(p) models, this method is only approximately valid, and the impulse-response function can be used for precise estimates (see Cheung and Lai, 1995). We do not pursue this because it is not our major concern.

[14] When the true data generating process is I(1), one may use AR models with differenced variables, then correct for biases. Median/Mean bias for such models, however, tends to be small, because differenced variables often exhibit much weaker persistence. Since we are interested in evaluating the usefulness of bias-corrected estimators, we do not consider such models.

observations to obtain h-step ahead forecasts. Then, we keep forecasting recursively by adding one observation in each iteration until we forecast the last observation. Second, we obtain h-step ahead forecasts using first 69 observations, then keep forecasting with a rolling window by adding and dropping one observation in each iteration, maintaining 69 observations, until we reach the end of full sample. We report the relative root mean squared prediction errors (RRMSPE; LS/RMA and LS/GB) for 1 through 6 forecasting horizons in Tables 4 and 5.

Overall, we find that both bias-correction methods outperform the LS estimator (RRMSPE greater than one) with an exception of the CRB Textile sub-index, which has relatively low persistence. Similarly weaker evidence is found for the Fats and Oil sub-index. These results are consistent with the findings in Section 3. That is, our real data application confirms that there are gains from correcting for bias only when the data is highly persistent. The forecast horizon may seem less important in this application, because forecasting performances are uniformly better with bias correction over different forecast horizons (except the Textile index). That is not true because there should be gains from bias correction when the data is highly persistent, which is the case here for majority indices we use. Neither of forecasting strategies strictly dominates the other one in terms of the RRMSPE.

Tables 4 and 5 about here

5 Concluding Remarks

This paper evaluates relative forecast performances of two bias-correction methods, Recurisve Mean Adjustment (RMA) and the grid bootstrap (GB) method, to the least squares (LS) estimator without bias-correction. When an intercept or an intercept and linear time trend are included in AR models, the LS estimator for the persistent parameter is downward-biased. However, it is not clear whether bias-correction will achieve better out-of-sample forecasting performances, because correcting for bias tends to increase the variance of the estimator. The present paper attempts to provide some guidance on when to correct for bias to enhance out-of-sample forecast accuracy.

Our Monte Carlo simulation experiments reveal that there is a gain from bias-correction only when the data is substantially persistent. Interestingly, the sample size (N) itself does not matter much. This is because both the bias and the variance tend to decrease as N increases. The forecast horizon may matter only when one forecasts low persistence variables. Our real data application with 7 commodity price indices overall confirms these findings.

References

1. Andrews, D.W.K. (1993), "Exactly Median-Unbiased Estimation of First Order Autoregressive Unit Root Models," *Econometrica* 61, 139-165.

2. Andrews, D.W.K., and H.-Y. Chen (1994), "Approximately Median-Unbiased Estimation of Autoregressive Models," *Journal of Business and Economic Statistics* 12, 187-204.

3. Andrews, D.W.K., and J.C. Monahan (1992), "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator," *Econometrica* 60, 953-966.

4. Cheung, Y.-W. and K.S. Lai (2000), "On the Purchasing Power Parity Puzzle," *Journal of International Economics* 52, 321-330.

5. Choi, C.-Y., N.C. Mark, and D. Sul (2009), "Bias Reduction in Dynamic Panel Data Models by Common Recursive Mean Adjustment," *Oxford Bulletin of Economics & Statistics*, forthcoming.

6. Cook, S. (2002), "Correcting Size Distortion of the Dickey-Fuller Test via Recursive Mean Adjustment," *Statistics and Probability Letters* 60, 75-79.

7. Hansen, B.E. (1999), "The Grid Bootstrap and the Autoregressive Model," *Review of Economics and Statistics* 81, 594-607.

8. Karanasos, M., S.H. Sekioua, and N. Zeng (2006), "On the Order of Integration of Monthly US Ex-Ante and Ex-Post Real Interest Rates: New Evidence from over a Century of Data," *Economics Letters* 90, 163-169.

9. Kendall, M.G. (1954), "Note on Bias in the Estimation of Autocorrelation," *Biometrika* 41, 403-404.

10. Kim, H. and Y. Moh (2010), "Examining the Evidence of Purchasing Power Parity by Recursive Mean Adjustment," *MPRA Working Paper* No.22712.

11. _____ (2009), "On the Usefulness of the Contrarian Strategy across National Stock Markets: A Grid Bootstrap Analysis," *Journal of Empirical Finance* 16, 734-744.

12. _____ and M. Ogaki (2009), "Purchasing Power Parity and the Taylor Rule," *Ohio State University Department of Economics Working Paper* No. 09-03.

13. _____, L. Stern, and M. Stern (2009), "Half-Life Bias Correction and the G7 Stock Markets," *manuscript*, Auburn University.

14. Mikusheva, A. (2007), "Uniform Inference in Autoregressive Models," *Econometrica* 75, 1411-1452.

15. Murray, C.J. and D.H. Papell (2002), "The Purchasing Power Parity Persistence Paradigm," *Journal of International Economics* 56, 1-19.

16. Ng, S. and P. Perron (2001), "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power," *Econometrica* 69, 1519-1554.

17. Rossi, B. (2005), "Confidence Intervals for Half-Life Deviations from Purchasing Power Parity," *Journal of Business and Economic Statistics* 23, 432-442.

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

19. Stine, R.A., P. Shaman (1989), "A Fixed Point Characterization for Bias of Autoregressive Estimators," *Annals of Statistics* 17, 1275–1284.

20. So, B.S. and D.W. Shin (1999), "Recursive Mean Adjustment in Time-Series Inferences," *Statistics & Probability Letters* 43, 65-73.

21. Steinsson, J. (2008), "The Dynamic Behavior of the Real Exchange Rate in Sticky-Price Models," *American Economic Review* 98, 519-533.

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

23. Tanaka, K. (1984), An Asymptotic Expansion Associated with the Maximum Likelihood Estimators in ARMA Models, *Journal of the Royal Statistical Society*, Ser. B 46, 58-67.

24. Taylor, A.M. (2001), "Potential Pitfalls for the Purchasing-Power-Parity Puzzle? Sampling and Specification Biases in Mean-Reversion Tests of the Law of One Price," *Econometrica* 69, 473-498.

25. Taylor, R. (2002), "Regression-Based Unit Root Tests with Recursive Mean Adjustment for Seasonal and Nonseasonal Time Series," *Journal of Business and Economic Statistics* 20, 269-281.

Table 1. Simulated Forecast Performance Comparison: AR(1) Models

$y_t = c + \rho y_{t-1} + u_t$

One-Period Ahead Forecast

	LS/RMA				LS/GB			
ρ	$T=50$	100	150	200	$T=50$	100	150	200
0.500	0.99840	1.00004	1.00041	0.99944	0.99805	0.99983	1.00024	0.99979
0.600	0.99956	0.99988	1.00001	1.00034	0.99842	0.99958	1.00010	0.99997
0.700	1.00110	0.99890	1.00044	1.00031	0.99850	0.99903	1.00008	1.00013
0.750	1.00332	1.00051	0.99949	1.00018	0.99980	1.00056	0.99947	1.00026
0.800	1.00290	1.00067	1.00053	1.00007	0.99830	1.00000	1.00022	1.00014
0.850	1.00467	0.99926	1.00035	0.99971	0.99922	0.99820	1.00055	0.99937
0.900	1.00961	1.00053	1.00082	0.99991	1.00450	0.99747	0.99993	0.99898
0.925	1.01383	1.00254	1.00337	1.00014	1.00863	0.99929	1.00167	0.99946
0.950	1.03537	1.00649	1.00087	1.00215	1.03052	1.00265	0.99838	1.00172
0.975	1.08439	1.01947	1.00713	1.00309	1.07599	1.01854	1.00490	1.00249
0.990	1.16831	1.06879	1.03002	1.01381	1.15020	1.05775	1.02899	1.01170

Two-Period Ahead Forecast

	LS/RMA				LS/GB			
ρ	$T=50$	100	150	200	$T=50$	100	150	200
0.500	0.99613	0.99938	1.00006	0.99961	0.99587	0.99929	1.00005	0.99982
0.600	0.99708	0.99882	0.99982	0.99998	0.99566	0.99887	0.99989	0.99999
0.700	0.99719	0.99795	1.00013	0.99963	0.99412	0.99819	1.00023	0.99986
0.750	1.00008	0.99953	0.99911	0.99993	0.99465	0.99926	0.99904	0.99968
0.800	0.99831	0.99991	1.00047	0.99909	0.98977	0.99884	0.99972	0.99928
0.850	1.00206	0.99931	0.99927	1.00016	0.99400	0.99760	0.99941	0.99981
0.900	1.00928	0.99987	1.00098	0.99949	1.00048	0.99494	0.99919	0.99873
0.925	1.01920	1.00098	1.00420	1.00023	1.01219	0.99567	1.00126	0.99945
0.950	1.03956	1.00714	1.00376	1.00113	1.03462	0.99958	1.00049	1.00021
0.975	1.10045	1.03406	1.01282	1.00366	1.10113	1.03122	1.00945	1.00239
0.990	1.19067	1.10375	1.05162	1.02950	1.18415	1.09281	1.04821	1.02405

Note: LS/RMA and LS/GB are the relative root mean squared prediction errors (RMSPE) for the Least Squares (LS) and to those of the Recursive Mean Adjustment (RMA), and to those of the grid bootstrap (GB) estimators, respectively.

Table 1 (Continued). Simulated Forecast Performance Comparison: AR(1) Models

$$y_t = c + \rho y_{t-1} + u_t$$

Three-Period Ahead Forecast

	LS/RMA				LS/GB			
ρ	$T = 50$	100	150	200	$T = 50$	100	150	200
0.500	0.99606	0.99978	0.99939	0.99958	0.99595	0.99972	0.99957	0.99960
0.600	0.99676	0.99805	0.99972	0.99979	0.99530	0.99852	0.99964	0.99985
0.700	0.99369	0.99780	0.99918	0.99951	0.98896	0.99816	0.99932	0.99956
0.750	0.99627	0.99781	0.99848	1.00030	0.98946	0.99769	0.99836	0.99991
0.800	0.99347	0.99755	1.00073	0.99860	0.98204	0.99677	0.99965	0.99886
0.850	0.99835	0.99732	0.99919	0.99994	0.98684	0.99519	0.99907	0.99951
0.900	1.00297	0.99731	1.00031	0.99945	0.99086	0.99098	0.99814	0.99795
0.925	1.01666	0.99951	1.00466	0.99918	1.00747	0.99209	1.00043	0.99823
0.950	1.04037	1.00811	1.00349	1.00088	1.03528	0.99934	0.99870	0.99930
0.975	1.10400	1.03839	1.01489	1.00684	1.11335	1.03220	1.01011	1.00448
0.990	1.19202	1.12465	1.06915	1.04102	1.19803	1.11383	1.06245	1.03288

Four-Period Ahead Forecast

	LS/RMA				LS/GB			
ρ	$T = 50$	100	150	200	$T = 50$	100	150	200
0.500	0.99685	1.00018	0.99975	0.99982	0.99673	0.99997	0.99980	0.99997
0.600	0.99508	0.99814	0.99991	0.99936	0.99392	0.99864	0.99983	0.99946
0.700	0.99115	0.99693	0.99857	0.99954	0.98592	0.99756	0.99893	0.99965
0.750	0.98907	0.99600	0.99784	0.99998	0.98075	0.99649	0.99785	0.99976
0.800	0.98913	0.99716	0.99924	0.99836	0.97568	0.99626	0.99836	0.99870
0.850	0.99228	0.99499	0.99921	0.99965	0.97738	0.99352	0.99870	0.99913
0.900	0.99628	0.99524	0.99897	1.00001	0.98158	0.98765	0.99605	0.99795
0.925	1.00602	0.99626	1.00244	0.99877	0.99626	0.98554	0.99788	0.99738
0.950	1.03222	1.00865	1.00299	0.99950	1.02937	0.99904	0.99718	0.99741
0.975	1.10167	1.04261	1.01716	1.00713	1.11807	1.03781	1.01170	1.00441
0.990	1.18337	1.13849	1.08121	1.04893	1.20919	1.13134	1.07112	1.03940

Note: LS/RMA and LS/GB are the relative root mean squared prediction errors (RMSPE) for the Least Squares (LS) and to those of the Recursive Mean Adjustment (RMA), and to those of the grid bootstrap (GB) estimators, respectively.

Table 1 (Continued). Simulated Forecast Performance Comparison: AR(1) Models

$$y_t = c + \rho y_{t-1} + u_t$$

Five-Period Ahead Forecast

	LS/RMA				LS/GB			
ρ	$T = 50$	100	150	200	$T = 50$	100	150	200
0.500	0.99815	0.99981	0.99968	0.99988	0.99769	0.99984	0.99978	0.99990
0.600	0.99586	0.99886	0.99936	0.99973	0.99425	0.99901	0.99954	0.99965
0.700	0.99002	0.99703	0.99912	0.99974	0.98393	0.99717	0.99957	0.99978
0.750	0.98800	0.99541	0.99752	0.99993	0.97795	0.99581	0.99791	0.99955
0.800	0.98692	0.99492	0.99966	0.99811	0.97192	0.99401	0.99888	0.99849
0.850	0.98392	0.99280	0.99731	0.99878	0.96619	0.99134	0.99733	0.99810
0.900	0.98590	0.99262	0.99810	0.99898	0.96676	0.98384	0.99408	0.99662
0.925	0.99861	0.99588	1.00053	0.99793	0.98736	0.98399	0.99552	0.99580
0.950	1.02373	1.00388	1.00248	0.99961	1.02217	0.99185	0.99504	0.99713
0.975	1.08599	1.04117	1.01561	1.00811	1.11412	1.03553	1.01007	1.00410
0.990	1.16565	1.13669	1.08768	1.05338	1.20979	1.13403	1.07822	1.04403

Six-Period Ahead Forecast

	LS/RMA				LS/GB			
ρ	$T = 50$	100	150	200	$T = 50$	100	150	200
0.500	0.99844	0.99979	0.99991	0.99999	0.99823	0.99972	0.99993	1.00000
0.600	0.99628	0.99888	0.99963	0.99977	0.99497	0.99905	0.99974	0.99984
0.700	0.98970	0.99741	0.99910	0.99969	0.98217	0.99756	0.99942	0.99981
0.750	0.98778	0.99538	0.99859	0.99956	0.97675	0.99563	0.99872	0.99923
0.800	0.98392	0.99497	0.99941	0.99764	0.96698	0.99431	0.99886	0.99809
0.850	0.97782	0.99338	0.99655	0.99760	0.95720	0.99109	0.99663	0.99735
0.900	0.98295	0.99010	0.99626	0.99831	0.96116	0.97825	0.99179	0.99568
0.925	0.98936	0.99459	0.99832	0.99613	0.97657	0.98076	0.99244	0.99357
0.950	1.00840	1.00032	1.00058	0.99975	1.00981	0.98810	0.99206	0.99619
0.975	1.06642	1.03615	1.01587	1.00788	1.10566	1.03284	1.00941	1.00240
0.990	1.14025	1.13640	1.09457	1.05454	1.20787	1.13675	1.08467	1.04832

Note: LS/RMA and LS/GB are the relative root mean squared prediction errors (RMSPE) for the Least Squares (LS) and to those of the Recursive Mean Adjustment (RMA), and to those of the grid bootstrap (GB) estimators, respectively.

Table 2. Simulated Forecast Performance Comparison: AR(2) Models

$y_t = c + \rho y_{t-1} + \beta \Delta y_{t-1} + u_t$

One-Period Ahead Forecast

		LS/RMA				LS/GB			
ρ	β	$T = 50$	100	150	200	$T = 50$	100	150	200
0.65	0.30	0.99931	0.99987	1.00091	1.00060	0.99204	0.99987	1.00052	0.99717
	−0.20	0.99847	0.99966	1.00083	1.00089	0.99023	1.00057	1.00253	0.99663
0.90	0.30	1.00147	1.00195	1.00187	1.00173	0.98590	0.99499	1.00053	1.00015
	−0.20	1.00800	1.00310	1.00335	1.00206	0.99318	0.99952	1.00084	0.99982
0.95	0.30	1.01728	1.00496	1.00400	1.00250	0.99761	0.99776	1.00206	1.00082
	−0.20	1.03743	1.01015	1.00696	1.00278	1.02575	1.00147	1.00547	0.99866
0.99	0.30	1.13617	1.05408	1.02546	1.01309	1.11007	1.03543	1.02187	1.01250
	−0.20	1.14977	1.09092	1.04462	1.02604	1.14345	1.06239	1.04044	1.02342

Two-Period Ahead Forecast

		LS/RMA				LS/GB			
ρ	β	$T = 50$	100	150	200	$T = 50$	100	150	200
0.65	0.30	0.99349	0.99917	1.00006	1.00084	0.98973	0.99907	1.00140	0.99963
	−0.20	0.99565	0.99919	1.00005	1.00059	0.98597	0.99955	0.99995	1.00072
0.90	0.30	1.00086	1.00138	1.00236	1.00132	0.98063	0.99173	1.00199	1.00067
	−0.20	1.01054	1.00285	1.00413	1.00081	0.99419	0.98857	1.00421	1.00047
0.95	0.30	1.03049	1.00610	1.00610	1.00167	1.00558	0.99092	1.00339	0.99987
	−0.20	1.05513	1.01507	1.00957	1.00051	1.03905	1.00102	1.00732	0.99719
0.99	0.30	1.18774	1.08876	1.04209	1.01261	1.16307	1.06188	1.03743	1.01293
	−0.2	1.18876	1.14043	1.07372	1.02763	1.19374	1.10578	1.06532	1.02728

Three-Period Ahead Forecast

		LS/RMA				LS/GB			
ρ	β	$T = 50$	100	150	200	$T = 50$	100	150	200
0.65	0.30	0.98842	0.99916	1.00031	1.00091	0.98221	0.99999	1.00198	1.00183
	−0.20	0.99128	0.99890	0.99998	1.00129	0.97313	0.99691	1.00165	0.99881
0.90	0.30	0.99390	0.99912	1.00153	1.00363	0.96279	0.98807	1.00151	1.00193
	−0.20	1.00367	0.99853	1.00260	1.00484	0.97560	0.97953	0.99952	1.00334
0.95	0.30	1.02841	1.00306	1.00537	1.00487	0.99635	0.98127	1.00028	1.00128
	−0.20	1.05528	1.01044	1.00796	1.00654	1.03213	0.99044	1.00011	1.00160
0.99	0.30	1.19806	1.10806	1.04978	1.02011	1.17780	1.07750	1.03976	1.01951
	−0.20	1.19690	1.16083	1.08359	1.04177	1.20726	1.12422	1.06700	1.04364

Note: LS/RMA and LS/GB are the relative root mean squared prediction errors (RMSPE) for the Least Squares (LS) and to those of the Recursive Mean Adjustment (RMA), and to those of the grid bootstrap (GB) estimators, respectively.

Table 2 (Continued). Simulated Forecast Performance Comparison: AR(2) Models

$y_t = c + \rho y_{t-1} + \beta \Delta y_{t-1} + u_t$

Four-Period Ahead Forecast

		LS/RMA				LS/GB			
ρ	β	$T=50$	100	150	200	$T=50$	100	150	200
0.65	0.30	0.98752	0.99815	0.99896	1.00000	0.98354	0.99772	0.99973	1.00052
	−0.20	0.99095	0.99772	0.99836	1.00043	0.96868	0.99710	0.99816	1.00016
0.90	0.30	0.98438	0.99780	0.99867	1.00354	0.94345	0.98729	0.99860	1.00172
	−0.20	0.99467	1.00008	0.99996	1.00321	0.95551	0.98028	0.99485	1.00018
0.95	0.30	1.01547	1.00399	1.00283	1.00521	0.97936	0.97939	0.99515	0.99972
	−0.20	1.04109	1.01686	1.00711	1.00444	1.01601	0.99540	0.99251	0.99669
0.99	0.30	1.18230	1.12465	1.05784	1.02303	1.16863	1.08644	1.04018	1.01776
	−0.20	1.17219	1.17928	1.10411	1.04772	1.18932	1.12734	1.07816	1.03931

Five-Period Ahead Forecast

		LS/RMA				LS/GB			
ρ	β	$T=50$	100	150	200	$T=50$	100	150	200
0.65	0.30	0.99295	0.99909	0.99868	0.99995	0.98812	0.99971	0.99930	0.99980
	−0.20	0.99277	0.99923	0.99793	1.00061	0.96597	0.99892	0.99811	1.00020
0.90	0.30	0.97454	0.99808	0.99563	1.00293	0.92484	0.98438	0.99637	1.00064
	−0.20	0.98751	1.00085	0.99753	1.00235	0.93425	0.97484	0.99225	0.99814
0.95	0.30	1.00105	1.00534	1.00006	1.00429	0.95635	0.97608	0.99087	0.99769
	−0.20	1.03355	1.01938	1.00552	1.00355	0.99031	0.99068	0.98780	0.99394
0.99	0.30	1.16269	1.13533	1.06378	1.02530	1.14667	1.09225	1.04160	1.01749
	−0.20	1.15964	1.19051	1.11638	1.05401	1.16676	1.13769	1.08844	1.04479

Six-Period Ahead Forecast

		LS/RMA				LS/GB			
ρ	β	$T=50$	100	150	200	$T=50$	100	150	200
0.65	0.30	0.99452	0.99969	0.99929	0.99972	0.99015	1.00106	0.99930	0.99956
	−0.20	0.99021	0.99976	0.99828	1.00030	0.96006	0.99852	0.99848	1.00011
0.90	0.30	0.96393	0.99760	0.99357	1.00303	0.90392	0.98083	0.99496	1.00065
	−0.20	0.97502	1.00012	0.99675	1.00399	0.90649	0.96954	0.99052	1.00039
0.95	0.30	0.98524	1.00543	0.99830	1.00522	0.92761	0.97172	0.98778	0.99805
	−0.20	1.01819	1.02051	1.00595	1.00664	0.95959	0.99121	0.98615	0.99550
0.99	0.30	1.14182	1.14123	1.06983	1.03033	1.11420	1.10154	1.04510	1.01949
	−0.20	1.14107	1.19439	1.12748	1.06501	1.13391	1.14489	1.09725	1.05014

Note: LS/RMA and LS/GB are the relative root mean squared prediction errors (RMSPE) for the Least Squares (LS) and to those of the Recursive Mean Adjustment (RMA), and to those of the grid bootstrap (GB) estimators, respectively.

Table 3. Persistence Parameter Estimation Results

Index	ρ_{L}	CI	ρ_{R}	CI	ρ_{G}	CI
Spot	0.950	[0.856,0.972]	0.969	[0.872,0.985]	0.975	[0.910,1.022]
Livestock	0.933	[0.770,0.966]	0.972	[0.795,0.986]	0.990	[0.875,1.044]
Fats and Oil	0.933	[0.776,0.965]	0.951	[0.800,0.985]	0.997	[0.864,1.049]
Foodstuff	0.952	[0.813,0.976]	0.977	[0.836,0.993]	1.008	[0.890,1.049]
Raw Industrials	0.940	[0.847,0.966]	0.969	[0.863,0.979]	0.955	[0.907,1.009]
Textiles	0.917	[0.807,0.951]	0.947	[0.824,0.967]	0.932	[0.874,1.003]
Metals	0.963	[0.870,0.981]	0.974	[0.887,0.993]	0.996	[0.929,1.024]
Index	HL_{L}	CI	HL_{R}	CI	HL_{G}	CI
Spot	3.378	[1.114,6.102]	5.503	[1.265,11.47]	6.844	[1.837, ∞]
Livestock	2.499	[0.663,5.010]	6.102	[0.755,12.29]	17.24	[1.298, ∞]
Fats and Oil	2.499	[0.683,4.864]	3.449	[0.777,11.47]	57.68	[1.185, ∞]
Foodstuff	3.523	[0.837,7.133]	7.447	[0.967,24.70]	∞	[1.487, ∞]
Raw Industrials	2.801	[1.044,5.010]	5.503	[1.176,8.165]	3.764	[1.775, ∞]
Textiles	2.000	[0.808,3.449]	3.182	[0.895,5.164]	2.461	[1.287, ∞]
Metals	4.596	[1.244,9.033]	6.578	[1.445,24.70]	43.24	[2.353, ∞]

Note: i) The number of lags (k) was chosen by the general-to-specific rule as recommended by Ng and Perron (2001). ii) ρ_{L}, ρ_{R}, and ρ_{G} denote the least squares (LS), recursive mean adjustment (RMA, So and Shin 1999), and the grid bootstrap (GB, Hansen 1999) estimates for persistence parameter, respectively. iii) 95% confidence intervals (CI) were constructed by 10,000 nonparametric bootstrap simulations for the LS and RMA estimators, and by 10,000 nonparametric bootstrap simulations on 30 grid points over the neighborhood of the LS estimate for the GB estimator. iv) HL_{L}, HL_{R}, and HL_{G} denote the corresponding half-lives in years, calculated by $(\ln(0.5)/\ln(\rho))/4$.

Table 4. Recursive Out-of-Sample Forecast Results

Index	h	LS/RMA	LS/GB	Index	h	LS/RMA	LS/GB
Spot	1	1.031	1.004	Raw Industrials	1	1.028	1.009
	2	1.059	1.033		2	1.057	1.021
	3	1.065	1.029		3	1.056	1.023
	4	1.050	1.018		4	1.036	1.010
	6	1.026	1.012		6	1.030	1.015
Livestock	1	1.035	1.012	Textiles	1	0.993	0.989
	2	1.066	1.025		2	0.997	0.999
	3	1.035	1.012		3	0.990	0.994
	4	1.039	1.011		4	0.978	0.985
	6	1.034	1.021		6	0.964	0.973
Fats and Oil	1	1.003	0.995	Metals	1	1.020	1.014
	2	1.013	1.018		2	1.031	1.034
	3	1.008	1.011		3	1.033	1.046
	4	1.001	1.003		4	1.016	1.024
	6	0.994	0.992		6	1.019	1.025
Foodstuff	1	1.027	1.029				
	2	1.032	1.040				
	3	1.017	1.022				
	4	1.015	1.020				
	6	1.003	1.004				

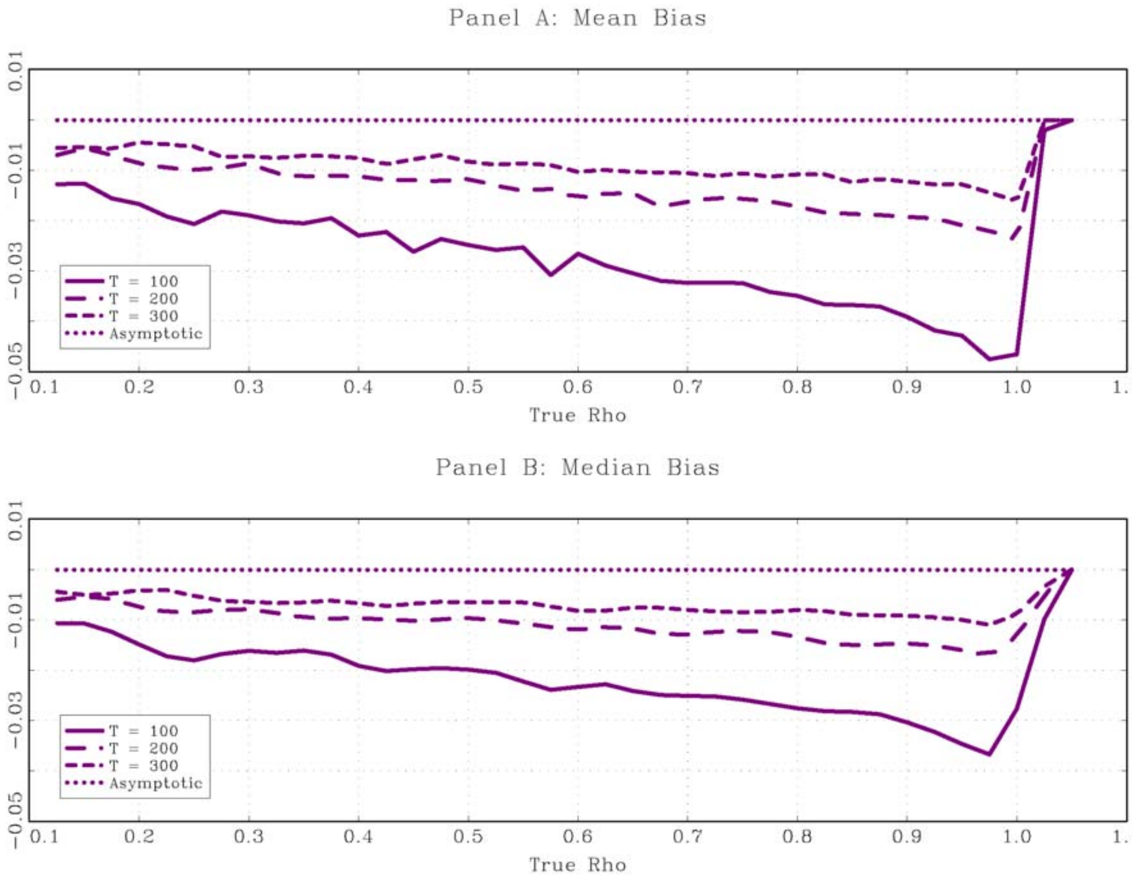
Note: i) Out-of-sample forecasting was recursively implemented by sequentially adding one additional observation from 69 initial observations toward 139 total observations. ii) The number of lags (k) was chosen by the general-to-specific rule recommended by Ng and Perron (2001). iii) h denotes the forecast horizon (quarters). iv) LS/RMA and LS/GB are the relative root mean squared prediction errors (RMSPE) for the Least Squares (LS) and to those of the Recursive Mean Adjustment (RMA), and to those of the grid bootstrap (GB) estimators, respectively.

Table 5. Rolling Window Out-of-Sample Forecast Results

Index	h	LS/RMA	LS/GB	Index	h	LS/RMA	LS/GB
Spot	1	1.006	1.010	Raw Industrials	1	1.007	1.021
	2	1.039	1.054		2	1.014	1.049
	3	1.046	1.066		3	1.004	1.035
	4	1.034	1.046		4	1.007	1.034
	6	1.032	1.043		6	1.000	1.020
Livestock	1	1.014	1.008	Textiles	1	1.017	1.002
	2	1.030	1.058		2	1.029	1.009
	3	1.026	1.039		3	1.010	0.999
	4	1.020	1.036		4	0.990	0.991
	6	1.012	1.036		6	0.985	0.985
Fats and Oil	1	1.001	1.011	Metals	1	0.998	1.006
	2	1.005	1.037		2	0.997	1.014
	3	1.001	1.035		3	1.004	1.035
	4	0.994	1.018		4	1.003	1.028
	6	0.980	0.989		6	1.002	1.019
Foodstuff	1	1.010	1.016				
	2	1.034	1.069				
	3	1.018	1.057				
	4	1.016	1.055				
	6	1.007	1.032				

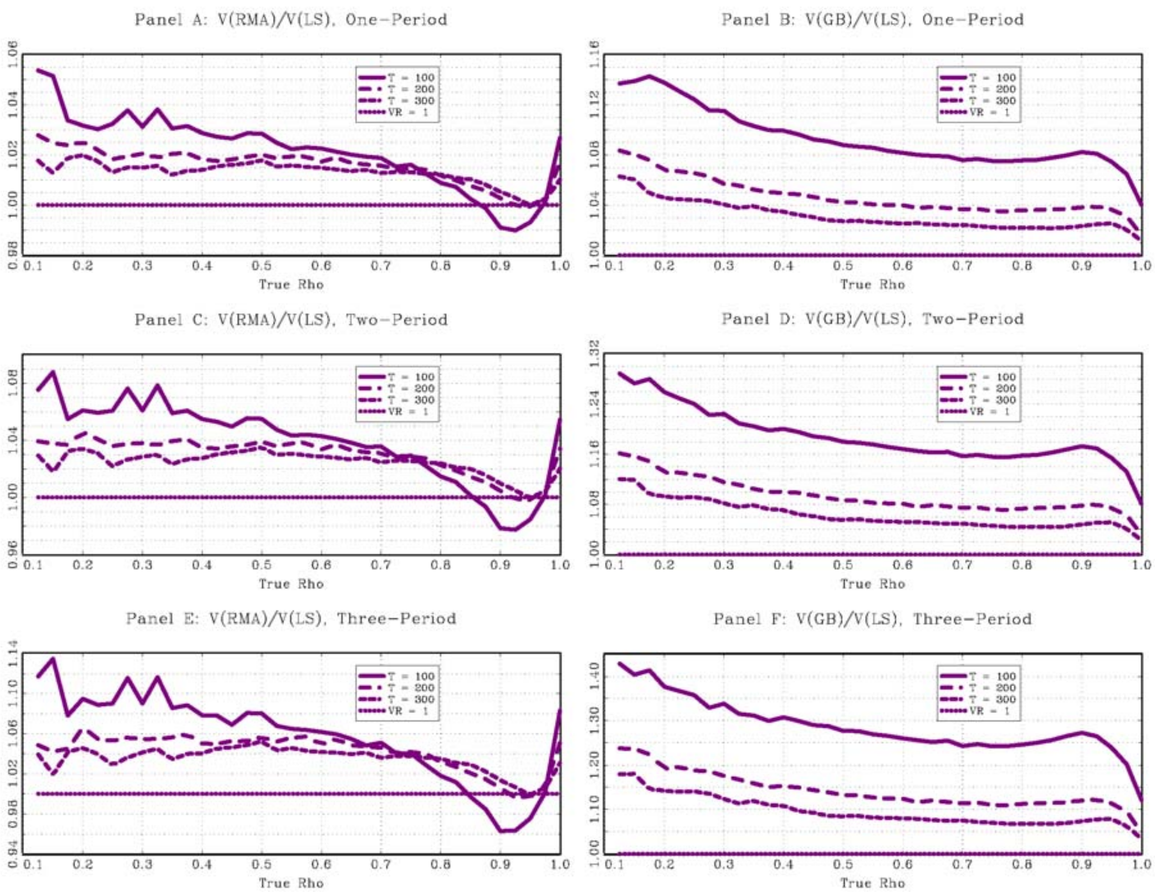
Note: i) Out-of-sample forecasting was implemented by sequentially adding one additional observation and dropping one observation in each iteration, maintaining 69 observations. ii) The number of lags (k) was chosen by the general-to-specific rule recommended by Ng and Perron (2001). iii) h denotes the forecast horizon (quarters). iv) LS/RMA and LS/GB are the relative root mean squared prediction errors (RMSPE) for the Least Squares (LS) and to those of the Recursive Mean Adjustment (RMA), and to those of the grid bootstrap (GB) estimators, respectively.

Figure 1. Small Sample Bias: AR(1) Models



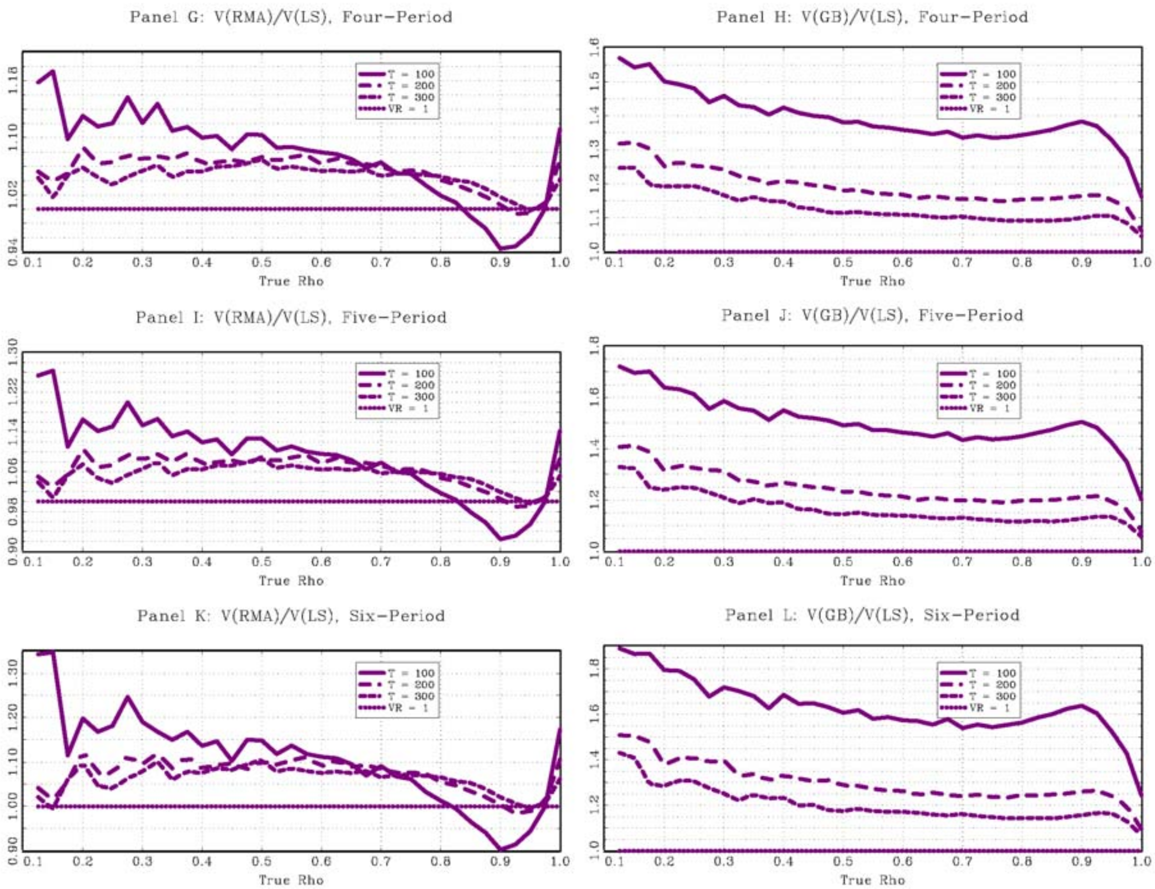
Note: We define the mean (median) bias as a difference between the true value of $\rho$and the mean (median) value of the least squares estimates from 5,000 simulations.

Figure 2. Variance Ratios



Note: V(RMA)/V(LS) and V(GB)/V(LS) denote the estimate variance ratios from 1- to 6-period ahead forecasts by the RMA, the GB, and the LS methods. Each variance was calculated by 5,000 simulations.

Figure 2 (continued). Variance Ratios



Note: V(RMA)/V(LS) and V(GB)/V(LS) denote the estimate variance ratios from 1- to 6-period ahead forecasts by the RMA, the GB, and the LS methods. Each variance was calculated by 5,000 simulations.