The Out-of-sample Performance of an Exact Median-Unbiased Estimator for the Near-Unity AR(1) Model

Medel, Carlos and Pincheira, Pablo

University of Nottingham, UK, School of Business, Adolfo Ibáñez University, Chile

4 March 2015

The Out-of-sample Performance of an Exact Median-Unbiased Estimator for the Near-Unity AR(1) Model

Carlos A. Medel*
University of Nottingham
United Kingdom

Pablo M. Pincheira†
School of Business
Adolfo Ibáñez University, Chile

March 4, 2015

Abstract

We analyse the multihorizon forecasting performance of several strategies to estimate the stationary AR(1) model in a near-unity context. We focus on the Andrews' (1993) exact median-unbiased estimator (BC), the OLS estimator, and the driftless random walk (RW). In addition, we explore the forecasting performance of pairwise combinations between these individual strategies. We do this to investigate whether the Andrews' (1993) correction of the OLS downward bias helps in reducing mean squared forecast errors. Via simulations, we find that BC forecasts typically outperform OLS forecasts. When BC is compared to the RW we obtain mixed results, favouring the latter as the persistence of the true process increases. Interestingly, we also find that the combination of BC and RW performs well when the persistence of the process is high.

JEL-Codes: *C22; C52; C53; C63.*

Keywords: *Near-unity autoregression*; *median-unbiased estimation*; *unbiasedness*; *unit root model*; *forecasting*; *forecast combinations*.

Highlights:

→We evaluate the out-of-sample behaviour of the Andrews' (1993) exact median-unbiased estimator for the simple stationary AR(1) process.

→We find that forecasts generated using the bias correction typically outperform forecasts generated with the OLS estimator.

→ The combination of the corrected AR(1) model with a driftless RW performs very well for persistent processes.

1 Introduction

The fit of the AR(1) model $y_t = \widetilde{c} + \phi y_{t-1} + \varepsilon_t$, with $-1 \leq \phi < 1$, $\widetilde{c} = c(1-\phi)$, and $\varepsilon_t \sim iid\mathcal{N}(0, \sigma_\varepsilon^2)$, for a finite-sample autocorrelated variable has been fairly used in financial and applied economics for either modelling or forecasting. It has also been used for instrumental or auxiliary estimation and many other applications. The case where the key parameter ϕ is not necessarily estimated, and it is instead assumed equal to unity, delivers the well-known random walk (RW) model. This particular specification has proved to be convenient for a number of economic time series with specific statistical characteristics. It is also common, especially within economics, that when the parameter ϕ is estimated, it is done through the OLS method, which provides a biased estimator in finite samples (Marriott and Pope, 1954; Kendall, 1954).

According to the Frisch-Waugh-Lovell theorem,[1] the AR(1) estimation with OLS is equivalent to estimate it from the equation $(y_t - \overline{y}) = \phi(y_{t-1} - \overline{y}) + \varepsilon_t$, where $\overline{y} = T^{-1}\sum_t^T y_t$. The downward

*E-mail: `lexcm6@nottingham.ac.uk`.

†Diagonal Las Torres 2640, Santiago 7941169, Chile. E-mail: `pablo.pincheira@uai.cl` (*corresponding author*).

[1] See Frisch and Waugh (1933) and Lovell (2008) for details.

bias comes from the correlation between the terms $(y_{t-1} - \overline{y})$ and ε_t. It is important to remark that there is no analytical expression for this bias, but it could be approximated via simulation methods (Hansen, 1999).[2]

Given this shortcoming, several bias-correction methods emerge, emphasizing their in-sample usage for statistical inference.[3] One of these methods, specially made for the AR(1) case, is the median-unbiased estimator developed by Andrews (1993). A number of applications are based on this estimator, covering from similar bias-correction procedures to several macroeconomic estimations (see for instance, Yu, 2012, to name an interesting application).

Surprisingly, relative to the in-sample inference, less attention has been devoted to the behaviour of these bias-correction procedures when evaluating forecast accuracy out-of-sample. Although, Maekawa (1997), Gospodinov (2002), and Kim and Durmaz (2012) may be considered some exceptions.

To fill this gap, the aim of this article is to analyse the multihorizon forecasting performance of several strategies to generate forecasts from the simple stationary AR(1) model in a near-unity context. We focus on the Andrews' (1993) exact median-unbiased estimator (BC), the simple OLS estimator (OLS), and the artificial imposition of a unit-root in the autoregressive coefficient, which is equivalent to generate forecasts following a driftless RW. In addition, we explore the forecasting performance of two pairwise combinations between these individual strategies: the equally-weighted combination of BC and RW (labelled C1), and the equally-weighted combination of OLS and BC (labelled C2). For completeness, we analyse the behaviour of these forecasts with both Gaussian and fat-tails innovations. The latter case is typically associated to financial time series. Notice that our investigation differs from the existing one by focusing on Andrews' (1993) methodology and by considering long-horizon forecasts.

Via simulations, we find that BC forecasts typically outperform OLS forecasts. When compared to the RW we obtain mixed results, favouring the latter as the persistence of the process increases. Interestingly, we also find that the combination of BC and RW performs well, which resembles results reported by Hansen (2010).[4]

2 Econometric details

We focus on the following simple specification, labelled as "Model 2" in Andrews' (1993) notation: $y_t = \widetilde{c} + \phi y_{t-1} + \varepsilon_t$; with $-1 \leq \phi < 1$; $\widetilde{c} = c(1-\phi)$; and $\varepsilon_t \sim iid\mathcal{N}(0, \sigma_\varepsilon^2)$. Given that many financial variables seem to display fat tails, we also consider the case in which $\varepsilon_t \sim t_8$ according to the findings of Rachev, Racheva-Iotova, and Stoyanov (2010). Interestingly, Mikusheva (2007) points out that Andrews' (1993) methodology could be also a valid procedure with fat-tails innovations.

While OLS estimation is widely known, it is important to provide some details of Andrews' (1993) bias correction. Suppose that $\widehat{\phi}$ is an estimator with a median function given by $m(\phi)$, that is

[2] As Patterson (2000) points out, some analytical expressions may be derived, but relying on the *true* parameters of the model.

[3] Some examples, also applied to the higher order AR(p) model, are the methods proposed in Orcutt and Winokur (1969), Shaman and Stine (1988), Stine and Shaman (1989), Andrews and Chen (1994), Hansen (1999), So and Shin (1999), Roy and Fuller (2001), Kim (2003), and Withers and Nadarajah (2011) among others. Moreover, Patterson (2000) shows that this bias could lead to a very misleading statistic when is accumulated across time, as the impulse response function does.

[4] Note that as Hansen (2010) suggests, model selection procedures coexist with the literature pioneered by Bates and Granger (1969) related to *forecast combinations* and the so-called "combination puzzle" (Stock and Watson, 2004). It basically states that fuzzy estimations of the weights that compute the average forecast do not guarantee a superior out-of-sample performance than the equally-weighted forecast. The results obtained in Hansen (2010) basically state that a combination based on Mallows weights performs fairly well.

uniquely defined and strictly increasing on the finite parameter space $-1 \leq \phi < 1$. Thus, the function $m(\phi)$ is the unique median of $\widehat{\phi}$ when ϕ is the true parameter. Then, the following estimator $\widehat{\phi}_{BC}$:

$$\widehat{\phi}_{BC} = \begin{cases} 1 & if\ \widehat{\phi} > m(1) \\ \\ m^{-1}(\widehat{\phi}) & if\ m(-1) < \widehat{\phi} \leq m(1) \\ \\ -1 & if\ \widehat{\phi} \leq m(-1), \end{cases} \tag{1}$$

corresponds to the median-unbiased estimator of ϕ (see Andrews, 1993). Notice that $m(-1) = \lim_{\phi \to -1} m(\phi)$, and the function $m^{-1} : (m(-1), m(1)] \to (-1, 1]$ acts as the inverse function of $m(\cdot)$.

Aside from these two methods (OLS and BC) to estimate and generate forecasts for the AR(1) process, another forecasting method originates when one assumes that ϕ is exactly equal to one, and, consequently, the drift is exactly equal to zero. Forecasts from this misspecified model have been analyzed by Clements and Hendry (2001) and Pincheira and Medel (2012) among others. When the true model is very close to having a unit root, RW-based forecasts perform relatively well; a result also reported in a number of articles. In particular, Pincheira and Medel (2015) notice that forecasts from the incorrectly specified RW model have the appealing property of being unbiased. To see this, we can iterate forward the AR(1) model to obtain:

$$y_{t+h} = \widetilde{c}\left[\frac{1-\phi^h}{1-\phi}\right] + \phi^h y_t + \sum_{i=0}^{h-1} \phi^i \varepsilon_{t+h-i}.$$

If the process were a driftless RW, then the optimal forecast would simply be y_t at any forecasting horizon. Accordingly, the expected value of the forecast error associated to the driftless RW h-step-ahead forecast, $\mathbb{E}\left[\varepsilon_t^{RW}(h)\right] = \mathbb{E}\left[y_{t+h} - y_t^{RW}(h)\right]$, would satisfy:

$$\begin{aligned} \mathsf{Bias}(h) \quad &\equiv \quad \mathbb{E}\left[\varepsilon_t^{RW}(h)\right] = \mathbb{E}\left[c\left[\frac{1-\phi^h}{1-\phi}\right] - (1-\phi^h)y_t + \sum_{i=0}^{h-1}\phi^i \varepsilon_{t+h-i}\right], \qquad (2) \\ &= \quad c\left[\frac{1-\phi^h}{1-\phi}\right] - (1-\phi^h)\mathbb{E}\left[y_t\right] = 0, \end{aligned}$$

as $\mathbb{E}\left[y_t\right] = c/(1-\phi)$.

The following proposition, extracted from Pincheira and Medel (2015), shows that the previous result could be straightforwardly generalised to the case in which the target variable is any stationary process.

Proposition 1 | ***Driftless RW-based Forecast Unbiasedness.***[5] *Let y_t be a covariance stationary process: $\mathbb{E}[y_t] = \mu$, $\mathbb{V}\left[y_t\right] = \gamma_0$, $\mathbb{C}\left[y_t, y_{t-s}\right] = \gamma_s$, $\forall t, s$ (not dependent on t). Then, driftless RW-based forecasts are unbiased and display a bounded Mean Squared Forecast Error (MSFE) as the forecasting horizon goes to infinity.*

Proof. Suppose that we forecast y_{t+h} assuming that the true model is a driftless RW that delivers the forecast $y_t^{RW}(h) = y_t$, with forecasting errors, $\varepsilon_t^{RW}(h) = y_{t+h} - y_t$. Because of the stationarity of y_{t+h} we have that:

$$\mathbb{E}\left[y_{t+h} - y_t^{RW}(h)\right] = \mathbb{E}\left[y_{t+h} - y_t\right] = \mathbb{E}\left[y_{t+h}\right] - \mathbb{E}\left[y_t\right] = 0, \tag{P1}$$

[5] We denote the expected value, variance, covariance, and autocovariance with $\mathbb{E}[\cdot]$, $\mathbb{V}[\cdot]$, $\mathbb{C}[\cdot,\cdot]$, and γ_t, respectively.

and therefore, *driftless RW-based forecasts* are unbiased. The MSFE is given by:

$$\begin{aligned}\mathsf{MSFE}(h) &= \mathbb{E}\left[y_{t+h} - y_t^{RW}(h)\right]^2 \\ &= \mathbb{E}\left[y_{t+h} - y_t\right]^2 \\ &= 2\mathbb{V}\left[y_t\right] - 2\gamma_h.\end{aligned} \tag{P2}$$

So,

$$\begin{aligned}\mathsf{MSFE}(h) &= \left|2\mathbb{V}\left[y_t\right] - 2\gamma_h\right| \leq 2\mathbb{V}\left[y_t\right] + 2|\gamma_h| \\ \mathsf{MSFE}(h) &\leq 2\mathbb{V}\left[y_t\right] + 2\sqrt{\mathbb{V}\left[y_{t+h}\right]\mathbb{V}\left[y_t\right]} \\ \mathsf{MSFE}(h) &\leq 4\mathbb{V}\left[y_t\right],\end{aligned} \tag{P2'}$$

and:

$$\lim_{h\to\infty} \mathsf{MSFE}(h) \leq 4\mathbb{V}\left[y_t\right] < \infty, \tag{P3}$$

and then MSFE(h) is a bounded sequence. ■

In the following section we explore the out-of-sample performance of the five different forecasting strategies considered: BC, OLS, RW, C1, and C2.

3 Simulations

3.1 Simulation set-up

We analyse the out-of-sample performance via Monte Carlo simulations. First, we generate observations from a stationary AR(1) model. Second, we pick a rolling estimation window to estimate the model (when needed). Third, we generate forecasts at several horizons. Finally, we compute and compare the sample Root MSFE (RMSFE) between the procedures.

We use the following approach to compute the median unbiased estimator of ϕ. Given a persistence value of ϕ, we generate a grid of AR(1) models with persistence parameter $\phi - \delta_i$, where $\delta_i = 0.01 \times i$, $i \in [1, 20]$. For each of these models we generate 1000 simulations. Then, for each of the simulated series we estimate an AR(1) model, with sample size T, via OLS including the series coming from the true model. This last estimation–that of the true series–delivers the value $\widehat{\phi}^{OLS}$. We save the median of the 1000 simulations for each of the models with parameter $(\phi - \delta_i)$. Finally, we find the median-unbiased estimator $\widehat{\phi}^{BC}$ as $\phi - \delta_j$ such that this latter term delivers the minimum value of $\left|m(\phi - \delta_i) - \widehat{\phi}^{OLS}\right|$.

In Figure 1A-B, we show the estimation of the median bias with both Gaussian and fat-tails innovations for different choices of the sample size T. With a small T, bias in absolute value is higher, and the curve across persistence is steeper than with a greater T. In both cases, with Gaussian and fat-tails innovations and for any given T, the bias is directly related to persistence. In Figure 1A-B the lowest bias in absolute term is 0.01. It is achieved for the following combination of persistence and sample size: $(\min\{\phi\}, \max\{T\})$. The lowest value is obtained with the BC estimator when the process is generated with Gaussian innovations. The highest bias in absolute terms is 0.097, as reported in Figure 1A-B (obtained with the combination of OLS and fat-tails). Interestingly, the combination of persistence and sample size that delivers this result is given by $(\max\{\phi\}, \min\{T\})$.

For our simulations we consider the following levels of persistence: $\phi = \{0.90; 0.95; 0,975; 0.99\}$. When we compute the OLS and the BC estimators we use rolling windows of a fixed size T.

We explore the robustness of our results using three different choices of $T = \{50; 100; 200\}$. The forecasting horizons are $h = \{1; 12; 24; 36; 48; 96; 120\}$, referring to the common case of series at a monthly frequency. We fix the total number of h-step-ahead forecasts to 2000, made in an indirect iterative manner. The constant term in our true DGP is set to unity, the variance of the stochastic term is also set to $\sigma_\varepsilon^2 = 1$ (for both kind of distributions) and the total number of simulations is 10000. We construct forecast combinations as follows:

$$\begin{aligned} \mathsf{C1} \quad &: \quad y_t^{C1}(h) = \lambda_1 \cdot y_t^{RW}(h) + (1-\lambda_1) \cdot y_t^{BC}(h), \\ \mathsf{C2} \quad &: \quad y_t^{C2}(h) = \lambda_2 \cdot y_t^{BC}(h) + (1-\lambda_2) \cdot y_t^{OLS}(h), \end{aligned} \tag{3}$$

where $\lambda_1 = \lambda_2 = 0.5$, following the suggestion of Stock and Watson (2004), and $y_t^M(h)$ denotes an h-step-ahead forecast constructed at time t using method $M \in \{BC; OLS; RW; C1; C2\}$. The predictive comparison is based on the RMSFE statistic:

$$RMSFE(h) = \left[\frac{1}{T} \sum_{t=1}^{T} (y_{t+h} - y_t^M(h))^2 \right]^{\frac{1}{2}}, \tag{4}$$

Specifically, we base our conclusions on the RMSFE-ratio between the BC forecast acting as a pivot, and the competing forecast $M' \in \{OLS; RW; C1; C2\}$: $RMSFE^{BC}/RMSFE^{M'}$, for any given horizon. Hence, values above unity imply a worse performance of BC-based forecasts.

In Figure 1C-D we show the variance ratio defined as $\mathbb{V}(\widehat{\phi}^{BC})/\mathbb{V}(\widehat{\phi}^{OLS})$. Notice that a similar profile is found with both kinds of innovations. $\mathbb{V}(\widehat{\phi}^{BC})$ is always lower than $\mathbb{V}(\widehat{\phi}^{OLS})$, in a ratio ranging around 0.880–0.905. The variance of the OLS estimator is the highest with ϕ=0.99, suggesting better relative predictive results for BC.

3.2 Simulation results

The predictive results are presented in Table 1. We focus on three aspects of them: the relative performance of BC versus OLS; same for BC versus RW; and the behaviour of our combination strategies.

3.2.1 BC versus OLS

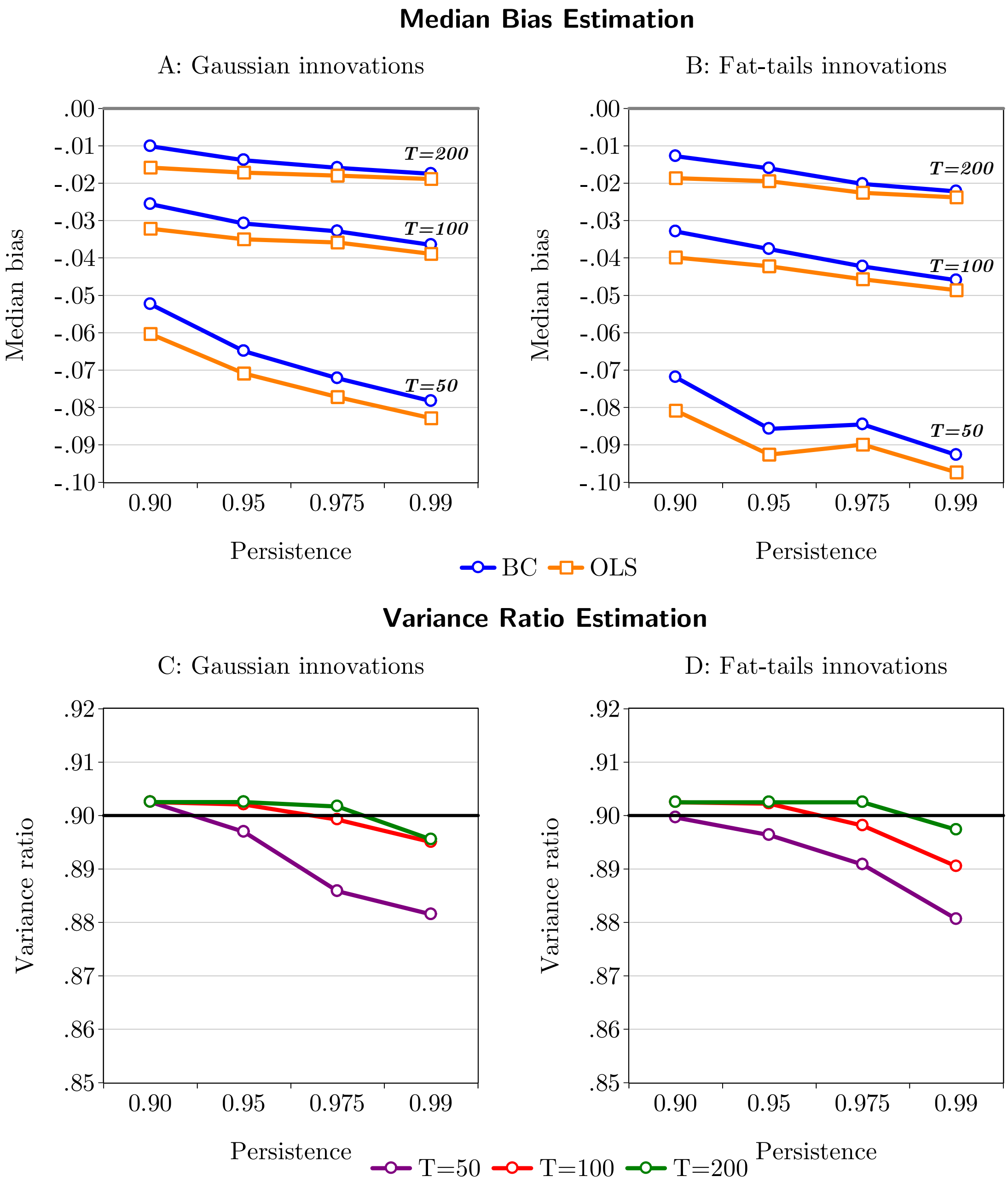
The performance of BC is outstanding when compared to OLS forecasts, as most of the corresponding ratios are below unity (around 95%). These predictive gains show a tendency to increase with both the persistence level and the forecasting horizon. On the contrary, the predictive gains tend to decrease as the sample size increases, especially at long horizons. We notice that OLS seem an inappropriate alternative for h>36 when the sample size is low (T=50), and the level of persistence is high. In these cases Table 1 shows ratios that are negligible. These results are robust to the source of innovations, either Gaussian or displaying fat tails. To have a quantitative flavour of the predictive gains, notice that for 24-periods-ahead the RMSFE provided by the BC approach achieves a half of that from OLS. It is also remarkable that in the least favourable case for BC, forecasts generated using OLS display only a small gain of 1% in RMSFE.

3.2.2 BC versus RW

The performance of the BC forecasts is in most cases better than the RW performance. In this case, more than two thirds of the corresponding ratios in Table 1 are below 1 (71%). These predictive gains of BC show a tendency to increase with both the forecasting horizon and the sample size. The predictive gains of BC, however, show a tendency to decrease with the persistence level. These results are robust to the distribution of innovations. Differing from the previous comparison with

OLS, BC forecasts are clearly outperformed by RW forecasts when the level of persistence is high. For instance, in the least favourable BC case, RW-forecasts display a gain of 22% in RMSFE. Therefore, RW results in a valid benchmark at high levels of persistence, indicating that in this setup misspecification is less harmful than parameter uncertainty.

Figure 1: Median bias and relative variance of the estimators (*)



(*) A-B: Median bias is defined as: $\widehat{\phi}^i - \phi$, with i=$\{BC;OLS\}$ for different sample sizes.
C-D: Variance ratio is defined as $\mathbb{V}(\widehat{\phi}^{BC})/\mathbb{V}(\widehat{\phi}^{OLS})$. Number of replications: 10000.
Source: Authors' elaboration.

3.2.3 Combinations performance

Table 1 shows that, roughly speaking, C1 dominates C2: the combination between BC and RW displays a lower RMSFE than the combination between OLS and RW. This result, plus the fact that BC quasi-dominates OLS forecasts, opens the question about the best forecasting strategy between BC, RW, and C1. Table 1 helps to produce a rule-of-thumb-type of answer as we clearly see that for levels of persistence between 0.90-0.95, a good strategy is the construction of forecasts using BC (it outperforms RW and C1 in 89% of our simulations). Similarly, as the persistence increases, the use of C1 would provide a good and relatively robust performance (it outperforms RW and BC in 52% of our simulations). This result is similar to that of Hansen (2010).

It is important to keep in mind that according to Table 1, BC and C1 do not *always* provide the most accurate forecasts. Nevertheless, when they are outperformed by either one of the other two best forecasting strategies it is often by just a small margin.

4 Concluding remarks

We analyse the multihorizon forecasting performance of several strategies to generate forecasts for the stationary AR(1) model in a near-unity context. We focus on the Andrews' (1993) exact median-unbiased estimator–BC–, the simple OLS estimator–OLS–, and the artificial imposition of a unit-root in the autoregressive coefficient, which is equivalent to generate forecasts following a driftless Random Walk. In addition, we explore the performance of two pairwise combinations between these individual strategies: the equally-weighted combination of BC and RW (labelled C1), and the equally-weighted combination of OLS and BC (labelled C2). For completeness, we analyse the behaviour of these forecasts with both Gaussian and fat-tails innovations.

Via Monte Carlo simulations we find that when forecasts are generated using Andrews' (1993) bias correction, they typically outperform forecasts generated using the standard OLS estimator. When compared to driftless RW-based forecasts, results are less clear cut as RW-based forecasts perform fairly well when the persistence level of the AR(1) es very high. Interestingly, the combination between RW and BC-based forecasts outperforms competing strategies in many cases, in line with the findings in Hansen (2010). As a rule of thumb, we propose the use of BC-based forecasts when the level of persistence is medium-high, and the use of an equally-weighted combination between RW and BC forecasts when the persistence of the process is very high.

References

1. Andrews, D.W.K., 1993, "Exactly Median-Unbiased Estimation of First Order Autoregressive/Unit Root Models," *Econometrica* **61**(1): 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**(2): 187-204.

3. Bates, J.M. and C.W.J. Granger, 1969, "The Combination of Forecasts," *Operational Research Quarterly* **20**(4):451-468.

4. Clements, M.P. and D.F. Hendry, 2001, "Forecasting with Difference Stationary and Trend Stationary Models," *Econometrics Journal* **4**: S1-S19.

5. Frisch, R. and F.V. Waugh, 1933, "Partial Time Regressions as Compared with Individual Trends," *Econometrica* **1**(4): 387-401.

6. Gospodinov, N., 2002, "Median Unbiased Forecasts for Highly Persistent Autoregressive Processes," *Journal of Econometrics* **111**(1): 85-101.

7. Hansen, B.E., 1999, "The Grid Bootstrap and the Autoregressive Model," *Review of Economics and Statistics* **81**(4): 594-607.

8. Hansen, B.E., 2010, "Averaging Estimators for Autoregressions with a Near Unit Root," *Journal of Econometrics* **158**(1): 142-155.

9. Kendall, M.G., 1954, "Note on Bias in the Estimation of Autocorrelation," *Biometrika* **41**(3-4): 403-404.

10. Kim, J.H., 2003, "Forecasting Autoregressive Time Series with Bias-corrected Parameter Estimators," *International Journal of Forecasting* **19**(4): 493-502.

11. Kim, H. and N. Durmaz, 2012, "Bias Correction and Out-of-sample Forecast Accuracy," *International Journal of Forecasting* **28**(3): 575-586.

12. Lovell, M., 2008, "A Simple Proof of the FWL Theorem," *Journal of Economic Education* **39**(1): 88-91.

13. Maekawa, K., 1987, "Finite Sample Properties of Several Predictors from an Autoregressive Model," *Econometric Theory* **3**: 359-370.

14. Marriott, F.H.C. and J.A. Pope, 1954, "Bias in the Estimation of Autocorrelations," *Biometrika* **41**(3-4): 393-403.

15. Mikusheva, A., 2007, "Uniform Inference in Autoregressive Models," *Econometrica* **75**(5): 1411-1452.

16. Orcutt, G.H. and H.S. Winokur, 1969, "First Order Autoregression: Inference, Estimation, and Prediction," *Econometrica* **37**(1): 1-14.

17. Patterson, K.D., 2000, "Bias Reduction in Autoregressive Models," *Economics Letters* **68**(2): 135-141.

18. Pincheira, P. and C.A. Medel, 2012, "*Forecasting Inflation with a Random Walk*," Working Paper 669, Central Bank of Chile.

19. Pincheira, P. and C.A. Medel, 2015, "*Forecasting Inflation with a Random Walk*," *manuscript*, [available upon request].

20. Rachev, S., B. Racheva-Iotova, and S. Stoyanov, 2010, "Capturing Fat Tails," *Risk Magazine* **May**: 72-77.

21. Roy, A. and W.A. Fuller, 2001, "Estimation for Autoregressive Time Series with a Root Near One," *Journal of Business and Economic Statistics* **19**(4): 482-493.

22. Shaman, P. and R.A. Stine, 1988, "The Bias of Autoregressive Coefficient Estimators," *Journal of the American Statistical Association* **83**(403): 842-848.

23. So, B.S. and D.W. Shin, 1999, "Recursive Mean Adjustment in Time-series Inferences," *Statistics and Probability Letters* **43**(1): 65-73.

24. Stine, R.A. and P. Shaman, 1989, "A Fixed Point Characterization for Bias of Autoregressive Estimators," *Annals of Statistics* **17**(3): 1275-1284.

25. Stock, J.H. and M.W. Watson, 2004, "Combination Forecasts of Output Growth in a Seven-country Data Set," *Journal of Forecasting* **23**: 405-430.

26. Withers, C.S. and S. Nadarajah, 2011, "Estimates of Low Bias for the Multivariate Normal," *Statistics and Probability Letters* **81**(11): 1635-1647.

27. Yu, J., 2012, "Bias in the Estimation of the Mean Reversion Parameter in Continuous Time Models," *Journal of Econometrics* **169**(1): 114-122.

Table 1: RMSFE-ratio between BC forecast and candidates, $M' \in \{OLS;RW;C1;C2\}$ (*)

		Gaussian innovations							*Fat-tails innovations*						
									T=50						
	h:	*1*	*12*	*24*	*36*	*48*	*96*	*120*	*1*	*12*	*24*	*36*	*48*	*96*	*120*
ϕ=0.90	OLS	1.00	0.96	0.94	0.91	0.86	0.62	0.53	0.99	0.95	0.90	0.82	0.64	0.08	0.02
	RW	0.99	0.84	0.74	0.74	0.72	0.65	0.69	0.99	0.79	0.68	0.71	0.65	0.69	0.68
	C1	1.02	1.01	0.96	0.96	0.97	0.90	0.94	1.02	0.98	0.92	0.94	0.89	0.95	0.94
	C2	1.02	0.97	0.90	0.90	0.87	0.72	0.72	1.01	0.93	0.86	0.84	0.74	0.23	0.06
ϕ=0.95	OLS	0.99	0.93	0.79	0.61	0.35	0.01	0.00	1.00	0.91	0.71	0.49	0.27	0.00	0.00
	RW	1.04	0.97	0.82	0.78	0.76	0.77	0.80	1.06	0.96	0.87	0.79	0.76	0.80	0.79
	C1	1.04	1.07	0.99	0.95	0.94	0.96	0.98	1.05	1.06	1.01	0.97	0.95	0.97	0.97
	C2	1.04	1.01	0.87	0.77	0.60	0.02	0.00	1.05	0.99	0.85	0.70	0.53	0.02	0.00
ϕ=0.975	OLS	1.00	0.86	0.49	0.12	0.02	0.00	0.00	1.00	0.92	0.76	0.57	0.34	0.00	0.00
	RW	1.04	1.14	1.04	0.96	0.94	0.86	0.86	1.08	1.12	1.04	0.97	0.90	0.87	0.86
	C1	1.04	1.14	1.09	1.04	1.03	0.98	0.98	1.05	1.13	1.10	1.05	1.01	0.99	0.98
	C2	1.03	1.06	0.81	0.35	0.06	0.00	0.00	1.05	1.07	0.95	0.81	0.63	0.02	0.00
ϕ=0.99	OLS	0.98	0.90	0.71	0.46	0.23	0.00	0.00	1.00	0.84	0.42	0.09	0.01	0.00	0.00
	RW	1.06	1.22	1.14	1.14	1.10	1.04	1.02	1.06	1.20	1.14	1.10	1.06	1.00	1.00
	C1	1.05	1.16	1.13	1.12	1.10	1.05	1.05	1.04	1.15	1.12	1.10	1.08	1.04	1.03
	C2	1.03	1.10	0.97	0.80	0.53	0.01	0.00	1.04	1.08	0.80	0.31	0.04	0.00	0.00
									T=100						
ϕ=0.90	OLS	1.00	0.99	0.98	0.98	0.98	0.99	0.98	1.00	0.99	0.98	0.99	0.97	0.97	0.99
	RW	1.01	0.73	0.65	0.60	0.60	0.60	0.58	0.99	0.76	0.68	0.65	0.62	0.58	0.60
	C1	1.03	0.94	0.90	0.87	0.87	0.88	0.86	1.01	0.96	0.92	0.90	0.89	0.87	0.88
	C2	1.02	0.92	0.88	0.84	0.84	0.85	0.83	1.01	0.95	0.90	0.88	0.86	0.83	0.85
ϕ=0.95	OLS	1.00	0.97	0.93	0.83	0.75	0.74	0.48	1.00	0.99	0.96	0.95	0.94	0.90	0.90
	RW	0.99	0.89	0.81	0.74	0.71	0.70	0.69	1.01	0.90	0.77	0.73	0.69	0.64	0.64
	C1	1.01	1.02	0.99	0.95	0.94	0.95	0.94	1.02	1.03	0.97	0.95	0.93	0.91	0.91
	C2	1.00	0.99	0.94	0.86	0.82	0.77	0.50	1.02	1.01	0.93	0.90	0.88	0.83	0.82
ϕ=0.975	OLS	1.00	0.99	0.96	0.93	0.90	0.82	0.80	1.00	0.96	0.91	0.85	0.78	0.53	0.38
	RW	1.02	1.09	1.01	0.94	0.89	0.82	0.82	1.02	1.00	0.94	0.90	0.84	0.78	0.76
	C1	1.02	1.10	1.08	1.04	1.02	0.98	0.98	1.02	1.06	1.05	1.03	0.99	0.96	0.94
	C2	1.02	1.08	1.04	0.98	0.94	0.86	0.85	1.02	1.03	0.98	0.93	0.86	0.72	0.61
ϕ=0.99	OLS	1.00	0.98	0.97	0.92	0.89	0.63	0.48	1.00	0.94	0.86	0.76	0.67	0.31	0.16
	RW	1.03	1.17	1.18	1.13	1.13	1.01	0.97	1.03	1.14	1.08	1.03	1.01	0.99	0.96
	C1	1.03	1.14	1.15	1.13	1.13	1.07	1.04	1.03	1.12	1.10	1.08	1.07	1.05	1.03
	C2	1.02	1.11	1.12	1.07	1.06	0.87	0.76	1.02	1.07	1.01	0.94	0.88	0.60	0.41

(*) Figures below one imply a better performance of the BC forecast; highlighted in shaded cells.

Source: Authors' elaboration.

Table 1: RMSFE-ratio between BC forecast and candidates, $M' \in \{OLS;RW;C1;C2\}$ (*) (...continued)

		Gaussian innovations							*Fat-tails innovations*						
		T=200													
	h:	*1*	*12*	*24*	*36*	*48*	*96*	*120*	*1*	*12*	*24*	*36*	*48*	*96*	*120*
ϕ=0.90	OLS	1.00	1.00	0.99	0.99	0.99	1.00	0.99	1.00	1.00	1.00	1.00	1.00	1.01	1.00
	RW	0.95	0.67	0.59	0.53	0.56	0.54	0.54	0.95	0.69	0.59	0.56	0.54	0.57	0.54
	C1	0.99	0.89	0.87	0.82	0.84	0.83	0.84	0.99	0.92	0.87	0.85	0.83	0.87	0.83
	C2	0.99	0.89	0.86	0.80	0.82	0.82	0.82	0.99	0.91	0.86	0.84	0.82	0.86	0.82
ϕ=0.95	OLS	1.00	0.99	1.00	1.00	0.99	0.99	0.98	1.00	0.99	0.99	0.99	0.98	0.98	0.99
	RW	0.96	0.87	0.76	0.69	0.65	0.60	0.62	0.98	0.85	0.71	0.69	0.63	0.60	0.63
	C1	0.99	1.00	0.95	0.92	0.90	0.88	0.89	1.00	0.99	0.93	0.93	0.89	0.88	0.91
	C2	0.99	0.99	0.94	0.90	0.88	0.85	0.86	1.00	0.98	0.91	0.91	0.87	0.85	0.88
ϕ=0.975	OLS	1.00	0.99	0.97	0.96	0.95	0.93	0.90	1.00	0.99	0.98	0.97	0.96	0.92	0.92
	RW	1.01	0.97	0.89	0.82	0.78	0.72	0.68	1.00	0.90	0.82	0.80	0.76	0.63	0.61
	C1	1.01	1.03	1.01	0.98	0.96	0.94	0.92	1.01	1.00	0.98	0.97	0.96	0.87	0.86
	C2	1.01	1.02	0.98	0.94	0.92	0.88	0.85	1.01	0.99	0.95	0.94	0.92	0.81	0.80
ϕ=0.99	OLS	1.00	0.99	0.96	0.94	0.92	0.83	0.77	1.00	0.99	0.99	0.97	0.94	0.82	0.74
	RW	1.02	1.09	1.09	1.05	1.01	0.94	0.92	1.02	1.09	1.07	1.05	1.03	0.98	0.91
	C1	1.02	1.08	1.10	1.09	1.08	1.04	1.03	1.02	1.09	1.10	1.09	1.09	1.07	1.03
	C2	1.01	1.07	1.07	1.04	1.01	0.93	0.89	1.01	1.07	1.08	1.06	1.04	0.95	0.87

(*) Figures below one imply a better performance of the BC forecast; highlighted in shaded cells.

Source: Authors' elaboration.