NBER WORKING PAPER SERIES

RANDOM WALK FORECASTS OF STATIONARY PROCESSES HAVE LOW BIAS

Kenneth D. West
Kurt G. Lunsford

Working Paper 34112
http://www.nber.org/papers/w34112

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2025

We thank Todd Clark, James Mitchell, participants in a brownbag seminar at the Federal Reserve Bank of Cleveland, the editor, an associate editor and two anonymous referees for helpful comments. The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Cleveland or the Federal Reserve System. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Random Walk Forecasts of Stationary Processes Have Low Bias
Kenneth D. West and Kurt G. Lunsford
NBER Working Paper No. 34112
August 2025
JEL No. C22, C53, E37, E47

ABSTRACT

We study the use of a misspecified overdifferenced model to forecast the level of a stationary scalar time series. Let x(t) be the series, and let bias be the sample average of a series of forecast errors. Then, the bias of forecasts of x(t) generated by a misspecified overdifferenced ARMA model for Δx(t) will tend to be smaller in magnitude than the bias of forecasts of x(t) generated by a correctly specified model for x(t). Formally, let P be the number of forecasts. The bias from the model for Δx(t) has a variance that is O(1/P^2), while the variance of the bias from the model for x(t) generally is O(1/P). With a driftless random walk as our baseline overdifferenced model, we confirm this theoretical result with simulations and empirical work: random walk bias is generally one-tenth to one-half that of an appropriately specified model fit to levels.

Kenneth D. West
University of Wisconsin - Madison
Department of Economics
and NBER
kdwest@wisc.edu

Kurt G. Lunsford
Federal Reserve Bank of Cleveland
kurt.lunsford@clev.frb.org

1 Introduction

Let forecast bias be the sample average of a series of forecast errors. We study the forecast bias that results when a stationary series is forecast by a misspecified overdifferenced model. The models that we consider are misspecified because they rule out the unit moving average root that results from differencing a stationary series. We document a surprising result: these models tend to have lower forecast bias than do models for the level of the series. Lower bias for the misspecified overdifferenced model holds even if the model for the levels is correctly specified.

Previous literature has evaluated the use of overdifferenced models for forecasting when forecast accuracy is measured not by bias but by the average value of squared forecast errors. Research such as Smith and Yadav (1994), Stock (1996), and Sánchez and Peña (2001) finds that overdifferencing can lower mean squared forecast errors when the series being forecast is very persistent. The intuition is that a misspecified model can perform well if the misspecification (in this case, imposing a unit root) almost nails a hard-to-estimate parameter (in these papers, a fractional integration parameter or an autoregressive root that is almost but not quite one). That is not the intuition for our result, which applies even for series that are not very persistent. Separately, Hendry (2006) proposed overdifferencing to improve forecasts of series subject to parameter shifts. This is again not our environment, which we assume to be stable and stationary.

While bias is typically subsidiary to other forecast evaluation criteria such as mean squared prediction error,[1] it is regularly used to assess the quality of forecasts. It is featured prominently in analyses of survey forecasts (e.g., Croushore (2010)), Congressional Budget Office forecasts (e.g., Farmer, Nakamura, and Steinsson (2024)), and International Monetary Fund Forecasts (e.g., Timmermann (2007), Beaudry and Willems (2022)). It also appears in some comparisons of forecasts of econometric models. A recent example is Bennett and Owyang (2022). However, while we show that forecasts from misspecified overdifferenced models have low bias, our simulations and empirical work also show that these forecasts have high root mean squared prediction errors and high mean absolute prediction errors when the series being forecast is not very persistent. Hence we are not advocating for or against use of misspecified overdifferenced models to forecast stationary series. Rather, our intent is to document how these models affect bias so that researchers and forecasters who use bias can better understand its behavior. In our own research (Lunsford and West, 2024), we were initially puzzled by finding that modeling low-persistence series with a driftless random walk resulted in forecasts with low bias. The present paper is our effort to understand those results.

[1]Petropoulos et al. (2022, p759) provide a list of point forecast evaluation criteria that includes mean squared prediction error and several other criteria based on absolute forecast errors.

A simple example will illustrate why our result holds. Suppose that one is making one-step-ahead forecasts for a stationary univariate time series x_t. The data run up to $t = P + 1$. One predicts first x_2, then x_3, ... and finally $x_{P+1} - P$ predictions in all. For a driftless random walk model, the prediction of x_2 is x_1, the prediction of x_3 is x_2, ..., the prediction of x_{P+1} is x_P. The time series of forecast errors is thus $x_2 - x_1$, $x_3 - x_2$, ..., $x_{P+1} - x_P$. The average value of the forecast error (i.e., bias)[2] is

$$b = \frac{(x_2 - x_1) + (x_3 - x_2) + ... + (x_{P+1} - x_P)}{P}$$

$$= \frac{x_{P+1} - x_1}{P}.$$

Let "var" denote "variance" and "cov" denote "covariance." Clearly

$$Eb = 0, \mathrm{var}(b) = \frac{2\mathrm{var}(x_t) - 2\mathrm{cov}(x_t, x_{t+P})}{P^2} \leq \frac{c}{P^2}, c = 4\mathrm{var}(x_t).$$

That is, bias b has expectation zero, and variance that is $O(1/P^2)$.

In Section 2, we generalize this theoretical result to: allow a forecast horizon $h > 1$, to consider asymptotics both with h fixed (as was implicit in the above analysis) and with h growing with P at a suitable rate, and to allow forecasts from a stationary and invertible ARMA model in the first difference of x_t. The result just given continues to hold, subject to a qualification spelled out below when h grows with P. Our theory is corroborated by simulation results reported in Section 3 and empirical results reported in Section 4. In these two sections, the series are deliberately chosen to not be very persistent, to make for a clear contrast to existing results noted above. Section 5 concludes. The appendix has proofs. An online appendix has some data details and simulation and empirical results omitted from the paper to save space. Throughout, when we refer to an overdifferenced model, it is understood to be a model that imposes a zero mean on Δx_t and is misspecified in the sense that it does not allow a unit moving average root.

2 Theoretical results

We are interested in forecasting either the average or the point in time value of a scalar stationary time series x_t over the next h periods, with the average forecast being $\frac{x_{t+1}+x_{t+2}+...+x_{t+h}}{h}$ and the point in time forecast being x_{t+h}. We focus on forecasting the average, giving below a concise summary of parallel

[2]Here and throughout, we use bias to refer to a sample average.

results for point in time forecasts. The period $t+h$ forecast error is

$$\frac{x_{t+1}+x_{t+2}+...+x_{t+h}}{h} - \text{period } t \text{ forecast of } (\frac{x_{t+1}+x_{t+2}+...+x_{t+h}}{h}). \tag{1}$$

Let us make the innocuous assumption that x_t is observed in period t, i.e., that the period t forecast of x_t is just x_t. Then it follows that the forecast error (1) can be written

$$\frac{h\Delta x_{t+1}+(h-1)\Delta x_{t+2}+...+\Delta x_{t+h}}{h} - \text{period } t \text{ forecast of } (\frac{h\Delta x_{t+1}+(h-1)\Delta x_{t+2}+...+\Delta x_{t+h}}{h}). \tag{2}$$

We assume that a forecast from an invertible ARMA model for Δx_t is used for the period t forecast. Specifically, the forecast satisfies

$$\text{period } t \text{ forecast of } \Delta x_{t+j} = \psi_{j,j}\Delta x_t + \psi_{j,j+1}\Delta x_{t-1} + \psi_{j,j+2}\Delta x_{t-2} + ... + \psi_{j,t+j-1}\Delta x_1, \tag{3}$$

with the $\psi_{j,i}$'s satisfying

$$\text{there is a } c>0 \text{ and } 0 \le \rho < 1 \text{ such that for all } i \ge 1, |\psi_{j,i}| \le c\rho^i \text{ for all } j \le i. \tag{4}$$

Equations (3)-(4) are consistent with use of an ARMA model in Δx_t that assumes that Δx_t has mean zero with ARMA parameters that satisfy the usual stationarity and invertibility conditions. The invertibility condition implies misspecification: $\rho < 1$ rules out a unit moving average root. Because x_t is stationary (per our assumptions below), a correctly specified process for Δx_t will have a unit moving average root.[3] In other words, the model (3)-(4) is not just an overdifferenced version of an invertible ARMA model for the level of x_t.

The random walk model used in the previous section fits into (3) with $\psi_{j,i} = 0$ for all j,i. The $\psi_{j,i}$'s will be nonzero for all j if one uses an ARMA model for Δx_t with a moving average component. For a final example, suppose forecasts are made assuming the AR(1) model $\Delta x_t = \varphi\Delta x_{t-1}+$ unforecastable disturbance, with $|\varphi| < 1$. (Since Δx_t is overdifferenced, it is wrong to assume that the "unforecastable disturbance" really is unforecastable as shown in footnote 3.) Then in (3) we have $\psi_{j,j} = \varphi^j$; $\psi_{j,j+1} = \psi_{j,j+2} = ... = \psi_{j,t+j-1} = 0$; $c = 1$ and $\rho = |\varphi|$. In practice, φ would presumably be an estimate of $\text{cov}(\Delta x_t, \Delta x_{t-1})/\text{var}(\Delta x_t)$.

[3] For example, let x_t be generated by $x_t = \varphi x_{t-1} + u_t$ with $|\varphi| < 1$ and u_t being white noise. Then, differencing yields $\Delta x_t = \varphi\Delta x_{t-1} + \Delta u_t$, which has the moving average component $\Delta u_t = u_t - u_{t-1}$ with a unit root.

The investigator uses the formula (3) to compute the period t forecast in (2). Bias is

$$b = \frac{\sum_{t=1}^{P}[\frac{h\Delta x_{t+1}+(h-1)\Delta x_{t+2}+...+\Delta x_{t+h}}{h} - \text{ period } t \text{ forecast of } (\frac{h\Delta x_{t+1}+(h-1)\Delta x_{t+2}+...+\Delta x_{t+h}}{h})]}{P}. \tag{5}$$

Our theoretical result assumes:

$$x_t \text{ is covariance stationary,} \tag{6}$$

and one of the following three conditions (7), (8) or (9):

$$\sum_{j=0}^{\infty} |\text{cov}(x_t, x_{t-j})| < \infty, \tag{7}$$

$$\text{for some } d,\ 0 < d < 0.5, \text{ and some } c > 0, |\text{cov}(x_t, x_{t-j})| < cj^{2d-1} \text{ for all } j > 0, \tag{8}$$

$$h \text{ is fixed as } P \to \infty. \tag{9}$$

In (7), absolute summability of x_t's autocovariances is weaker than (though consistent with) x_t following an ARMA process. Long memory is allowed in (8), with d the order of fractional integration (Brockwell and Davis (1993, p520)). Under (9), we do not require any condition on the autocovariances of x_t (other than that implied by stationarity). We further discuss our assumptions below.

Our basic theoretical result is:

Theorem 2.1 Define b as in (5). Assume (3), (4), and (6). Then $Eb = 0$ and:

(a)Under (7), var(b) = O(h/P^2).

(b)Under (8), var(b) = O(h^{1+2d}/P^2).

(c)Under (9), var(b) = O$(1/P^2)$.

The proof is in the Appendix. Before discussing the theorem, we present the parallel result for point in time forecasts. As above, assume that the that the period t forecast of x_t is x_t. Then the point in time forecast error is

$$x_{t+h} - \text{period } t \text{ forecast of } x_{t+h} = \tag{10}$$

$$\Delta x_{t+1} + ... + \Delta x_{t+h-1} + \Delta x_{t+h} - \text{period } t \text{ forecast of } (\Delta x_{t+1} + ... + \Delta x_{t+h-1} + \Delta x_{t+h}).$$

Let b_P denote point in time bias:

$$b_P = \frac{\sum_{t=1}^{P} \Delta x_{t+1} + \Delta x_{t+2} + ... + \Delta x_{t+h} - \text{period } t \text{ forecast of } (\Delta x_{t+1} + ... + \Delta x_{t+h-1} + \Delta x_{t+h})}{P}. \quad (11)$$

Theorem 2.2. The statements in Theorem 2.1 continue to hold when bias b is replaced by point in time bias b_P.

Comments on Theorems 2.1 and 2.2:

1. Parts (a) and (b) of Theorem 2.1 allow the horizon h to grow with the sample size P, as in Richardson and Stock (1989) or Müller and Watson (2016). Consistency ($b \xrightarrow{p} 0$) follows if var(b) $\to 0$ as $P \to \infty$. To yield var(b) $\to 0$, part (b) requires a slower rate of growth of h than does part (a). Because $d > 0$, the process has long memory, and (7)'s summability condition does not hold. (Note that if x_t is fractionally integrated with $d < 0$, (7) holds (Brockwell and Davis (1993,p 520)) and part (a) of the theorem applies.)

2. Suppose that $h = cP$ for some $c > 0$, as in Müller and Watson (2016). Then $h/P^2 = c/P$ and thus in Theorem 2.1(a), var(b) $= O(1/P)$, as is the case for the variance of bias from a conventional stationary model. According to this metric, then, bias from the overdifferenced model is no less variable than is bias from a stationary model applied to the level of x_t. However, if h/P is small, one can still expect less bias from the forecast from the overdifferenced model. Indeed, our simulations and empirical work indicate that a random walk forecast tends to have lower bias than stationary processes even when h/P is as large as 0.6, although, consistent with our theory, the advantages of a forecast from an overdifferenced model diminish when h/P is large. Under Theorem 2.1(b), var(b) $= O(1/P^{1-2d})$ (because $h = cP \Rightarrow h^{1+2d}/P^2 = c^{1+2d}/P^{1-2d}$). This implies var($b$) $\to 0$ more slowly than $O(1/P)$.

3. We have been assuming a model in first differences. One may wonder about the properties of a forecast based on a model in higher-order integer differences. Let us consider forecasts from the (misspecified) I(2) model $\Delta^2 x_t = u_t$~ iid. Then when h is fixed, as in (2.7), we again obtain var(b) = O($1/P^2$). On the other hand, for forecasts from this I(2) model, if h grows with P, var(b) does not go to zero rapidly, and in fact diverges under (7) when $h = cP$ as in our previous point. This suggests that superior performance will require relatively small values of h/P. Hence our focus on a first difference model.

4. Consistent with the empirical studies cited in the introduction, Theorems (2.1) and (2.2) are about the unconditional mean and variance of bias. Alternatively, researchers may be interested in evaluating

forecasts conditional on information through some period of time t, such as in Clark and McCracken (2013, Sec 4) or Giacomini and White (2006). We leave this as a topic for future research.

5. Our simulations and empirical work will use univariate forecasting models. We note, however, that variance of bias will, in general, be $O(1/P)$ when forecasting from a multivariate stationary model that is not overdifferenced (West (1996))[4]. Hence our theory implies that forecasts from univariate overdifferenced model will tend to have lower bias than will forecasts from a multivariate model.

6. One can ask, which ARMA process for Δx_t yields the smallest var(b)? The answer depends on the particulars of the x_t process. In most of the simulations and empirical work reported here, we simply use a driftless random walk, because it is a commonly used benchmark model.,[5] We have also used a zero mean overdifferenced AR(1) model in place of the random walk, with simulation results reported in our online Appendix and empirical results reported below. We henceforth use b^{RW} to refer to b as defined in (5) when, consistent with a random walk model, the "period t forecast" referenced in (5) is set to zero.

3 Simulation results

To study small sample properties of b^{RW}, we use two sets of simulations. The first set of simulations uses an AR(2) data generating process (DGP) with a sample size of $T+1=160$, and we report a comparison of random walk and AR(2) forecasts. The second set considers the effect of increasing the sample size to see whether Theorem 2.1 is useful in characterizing how var(b^{RW}) grows as the sample size grows. It uses both an AR(2) and a stationary fractionally integrated DGP.

We begin with the first set of simulations. In addition to the AR(2) simulations reported here, we have completed simulations replacing the AR(2) process with an iid process and with a stationary fractionally integrated I(d) process. The simulation results with the iid and I(d) processes are similar to those for the AR(2) process. We report the iid and I(d) simulation results in our online appendix to save space.

In our AR(2) DGP, $x_t = 0.5x_{t-1} - 0.1x_{t-2} + u_t$, with $u_t \sim$ iid $N(0,1)$. The first-order autocorrelation of x_t is about 0.45, which is representative of the first-order autocorrelations of the data studied in the next section. The modulus of both autoregressive roots is about 0.32. Here and in the empirical work in the

[4]West (1996) assumes a certain long run variance is positive definite, an assumption that rules out overdifferencing.

[5]For example, Faust and Wright (2013) use a random walk as a forecasting model for inflation, which is a variable we forecast below. Petropoulos et al. (2022) recommend benchmarking any new forecasting model against a random walk.

next section, we focus on series that are not very persistent because we take it as well-known that overdifferencing can improve forecasts of highly persistent series (see references in the introduction).

We structure the simulations to parallel the quarterly empirical results in the next section, and the two sections together supply complementary details on our setup. The total sample size is $T + 1 = 160$ (40 years of quarterly data). In each simulation sample, we draw initial values from the unconditional distribution of x_t, i.e., from a bivariate $N(0, V)$ distribution in which the diagonals of V are the variance of x_t and the off-diagonals are the first-order autocovariance of x_t.

We use a sample size of $R = 48$ to compute the initial forecasts. That leaves $112 = 160 - 48$ observations for forecasting and for extending the sample used to make forecasts. We compute forecasts for the average value of x over horizons h = 1, 4, 8, 12, 20, and 40. For a given forecast horizon, there are $P = 112 - (h - 1)$ forecasts available for computing the bias. The end dates of the samples used to construct forecasts are $\tau = R, ..., T + 1 - h$. The forecasts use both recursive and rolling estimation samples. (See West (2006), for example, for definitions and illustrations of these two schemes.)

In a sample that ends at date τ, the forecast from the driftless random walk model is x_τ for both the recursive and rolling samples, and for all horizons. For the AR(2) model, ordinary least squares is used to estimate parameters using $\{x_1, ..., x_\tau\}$ (recursive sample) or $\{x_{\tau-R+1}, ..., x_\tau\}$ (rolling sample). The estimated parameters are then used to compute forecasts using the chain rule of forecasting. We also report results from the infeasible AR(2) forecast that relies on population parameters.

We run 1000 simulations. In each simulation, we generate a sample of size 1200 and use the first 160 observations. (The remaining observations will be used for the second set of simulations below.) In each sample, we compute forecast bias, root mean squared prediction error (RMSPE) and mean absolute prediction error (MAPE) for both models. We then compute (1) the absolute value of relative forecast biases $|b_j^{RW}/b_j^{AR}|$, where the subscript j denotes simulation j; (2) relative RMSPEs, $RMSPE_j^{RW}/RMSPE_j^{AR}$; and (3) relative MAPEs, $MAPE_j^{RW}/MAPE_j^{AR}$. We report the median across our 1000 simulations of relative bias, RMSPE and MAPE, as well as the percentage of simulations in which each ratio is less than 1. To keep the presentation smooth, we will generally omit "absolute value" when we reference relative bias: "smaller bias" should be understood to mean "smaller bias in absolute value," for example.

We note that our paper makes no new predictions about relative RMSPEs or MAPEs. We report these to reassure the reader that our setup delivers the expected result that in a stationary environment, the

RMSPE and MAPE of a forecast from a well specified stationary model is less than that of a random walk forecast.[6]

Table 3.1 summarizes results. Each column reports results for a given horizon. For convenience, lines (5) and (6) give the number of predictions P and the ratio h/P.

Lines (2a)-(2c) support our central theoretical prediction. The random walk model delivered smaller bias in 60 percent to 92 percent of the 1000 samples, with the percentage tending to be higher for shorter horizons. Relatively good performance for shorter horizons is consistent with one interpretation of our asymptotic results (see the comments under Theorem 2.2). Note that smaller bias results even when the random walk forecast is compared to the infeasible forecast that relies on population parameters (line (2c)). The value of 0.2 for h=1, population (line (2c)), for example, means that in 500 of our 1000 simulation samples, b^{RW} was at most two-tenths of AR(2) bias.

Lines (3a)-(3c) and (4a)-(4c) present results for relative RMSPE and MAPE. Since the DGP's autoregressive roots are comfortably below one, we expect the AR(2) model to have lower RMSPE and MAPE. And that is indeed the case. With one exception, the "0%" figures in the table are exact, with exactly 0 of the 1000 samples underlying each "0%" producing an RMSPE or MAPE lower for the random walk than for the AR(2) model.[7]

We now turn to our second set of simulations. Our goal here is to consider behavior of var(b^{RW}) as P increases. Theorem 2.1 states that suitably normalized var(b^{RW}) is $O(1)$, with the normalization factor depending on whether one assumes h stays fixed as P grows and whether or not the DGP is fractionally integrated. Simulation evidence that we consider congruent with the $O(1)$ characterization are results in which suitably normalized var(b^{RW}) seems roughly bounded as P increases. Of course "roughly bounded" is not precisely defined, and it is always possible that simulations with still larger values of P than we use would deliver contrary evidence. So our evidence is only suggestive.

Panel A in Table 3.2 reports results for our AR(2) DGP, while panel B reports results when the DGP is the fractionally integrated process $(1-L)^d x_t = u_t$ with d=0.31 and $u_t \sim$iid N(0,1). The value of d was chosen because it implies a first order autocorrelation coefficient of 0.45, which, as noted above, is representative of the data we analyze in the next section.

[6] As shown in the simulations of Stock (1996) and the theory of Sánchez and Peña (2001), if the DGP is a stationary model with a near unit autoregressive root, the random walk model can have lower RMSPE than an estimated stationary model.

[7] The exception: for h=1, recursive, in 1 of the 1000 simulation samples, MAPE^{RW}/MAPE^{AR} was below 1.

We use 1000 simulations and generate 1200 observations per simulation to study four different sample sizes: $T+1$ = 160, 400, 800 and 1200. With 48 observations again used to form estimates for the first prediction, the implied number of predictions for h=1 is given in column (1) of both panels in Table 3.2. Our other horizons are h = 4, 8, 12 and a sample-size dependent horizon $h = 0.2P$. We compute forecasts for the average value of x for each of these horizons. The number of predictions for these horizons shrink as h increases according to the logic illustrated in line (5) of Table 3.1.

As a preliminary, row (2) in each panel presents the value of normalized var(b^{RW}) for P=112. (We discuss the normalization factor below). The values increase as the horizon increases.

Our interest is not in the numerical value as the horizon increases, but the pattern as P increases for given h. Theorem 2.1(c) states that if h is fixed as $P \to \infty$, then var(b^{RW}) $\times P^2$ is bounded. Our intuition was that fixing h at a value of 12 or less while P increases from around 100 to over 1000 is congruent with such asymptotics. The results for h = 1, 4, 8 and 12 (columns (2)-(5) in both panels) confirm this intuition for the AR(2) and fractionally integrated DGPs. Rows (3)-(5) of each panel express var(b^{RW}) $\times P^2$ relative to the value in row (2). The ratios seem roughly bounded as P grows. We interpret that as consistent with our asymptotic theory.

Next, turn to columns (6) and (7), in which h grows with P according to h=0.2P. In both panels, column (6) has the unsurprising result that the fixed h asymptotics do not lead to roughly bounded values of normalized var(b^{RW}). Column (7) in each panel presents var(b^{RW}) normalized according to our theory so that normalized var(b^{RW}) is $O(1)$. (The normalization factors, which are different for the two DGPs, are given in headers to column (7), and follow from remark 2 following Theorem 2.2 above.) Rows (3)-(5) in each panel indicate that as P grows, it does indeed seem that normalized var(b^{RW}) is roughly bounded for both the AR(2) and fractionally integrated DGPs.

Our online Appendix reports an extensive additional set of simulations. We repeated the simulations underlying Table 3.1 when the stationary DGP and forecasting model were iid, when they were I(d) with d=0.45, and when the AR(2) DGP had GARCH disturbances. We also repeated all of our simulations (Table 3.1, Table 3.2, iid DGP, and I(d) DGP) when the overdifferenced model was a zero mean AR(1) in differences. These simulations delivered strongly congruent results. See the online Appendix for details.

We conclude that the simulations are consistent with the theory proposed in the previous section.

4 Empirical results

This section presents results from pseudo-out-of-sample forecasting. We focus on bias in predicting h period ahead averages (the exclusive focus of the previous section's simulations), with a single concise table giving results for bias in point in time predictions. We first compare the bias of a driftless random walk and either an AR(2) process or an AR process with lag length chosen by BIC, when predicting an h period ahead average (as in Theorem 2.1). We report results for AR(2) here, with the very similar results for BIC-chosen lag length reported in our online appendix. We next present results when a zero mean AR(1) in first differences replaces the random walk. Our final comparison is again random walk versus AR(2) in levels, but now in prediction of the h period ahead point in time observation (as in Theorem 2.2).

To obtain a wide range of data, we relied on groups 1-10 in the quarterly data collection described in McCracken and Ng (2021) to choose 10 series. The series are listed in Table 4.1, along with some summary statistics. We generally use McCracken and Ng's recommendations on how to induce stationarity, for example treating the change but not the level of the unemployment rate as stationary. One exception is that we use CPI inflation following our earlier working paper (Lunsford and West, 2023) instead of the first difference of inflation as in McCracken and Ng (2021). See the online appendix for more details on the data. While there is always some uncertainty about how to induce stationarity, for simplicity of exposition we take the stationarity transformation as given and use "overdifferenced" to refer to a model applied to the difference of the series (for example, applied to the difference of the change in the unemployment rate, i.e., to the second difference of the unemployment rate). The series encompass a range of persistence; we did, however, limit ourselves to series whose estimated first order autocorrelation coefficient was far enough from 1 that supportive results for our overdifferenced models would not plausibly be attributed to overdifferencing removing a near-unit root (see the discussion in the introduction.)

The horizons are h=1, 4, 8, 12, 20, and 40 quarters. Our baseline sample runs 1984:1-2023:4, with data prior to 1996 used only for estimates of the parameters of the AR or overdifferenced AR models. The base for the first prediction thus is 1995:4. The realization for the last prediction was 112 quarters later in 2023:4. For h=4, for example, the first value predicted was the average value over the four quarters 1996:1-1996:4; the second value predicted the average value for the four quarters 1996:2-1997:1, and so on, ending with a final prediction for the average over the four quarters 2023:1-2023:4. This allowed 109 h=4 quarter ahead predictions in all. More generally, for each horizon h the number of predictions P satisfies $P = 112 - (h - 1)$. The various values of h, P and the implied values of h/P are given in lines (5) and (6) of Table 3.1.

An alternative sample used no data past 2019:4, to make sure that the steep fall and rapid recovery of the COVID recession was not skewing results. When we stop the comparison in 2019:4 instead of 2023:4, P falls by 16 for each horizon.

In estimation of the AR(2) model, we used both rolling and recursive sampling schemes as we moved the estimation end point forward. We checked for stationarity of the estimates for each endpoint and when the estimates were explosive we substituted a forecast of a driftless random walk.

Table 4.2 gives detailed results for GDP growth, with our online appendix presenting such details for the other nine variables. Panel A gives results when the sample ends in 2023:4, panel B when the sample ends in 2019:4. Line 2 gives the bias in the random walk forecast (b^{RW}), over the various horizons. The magnitude of the values might be gauged by a comparison to the mean value of GDP growth, which (Table 4.1, row (1), column (2)) is 1.8 percent. In both panels of Table 4.2, and for all horizons but h=40, the absolute value of b^{RW} is a small fraction–1/9 or less–of the mean of 1.8 percent. In our experience in forecasting, that is a good result.

Rows (3a) and (3b) are the test of whether our paper's theoretical results are reflected in the data. In these rows, a value less than 1 means that bias was smaller in absolute value in the random walk forecast. Upon examining the 24 values in these two rows, we see that 21, or 88 percent, of the figures are below 1.[8] The superior performance of the random walk model is not because it is better according to a mean squared or mean absolute error criterion. Of the 24 ratios of RMSPEs in rows (4a) and (4b), only 1, or 4 percent, is less than 1;[9] none of the 24 ratios of MAPEs in rows (5a) and (5b) are less than 1.

Table 4.3 presents summary results for all ten variables. Columns (3)-(8) present forecasting results, with columns (3) and (4) the ones that are central to our paper. In columns (3)-(8), medians and percentages are computed across 24 comparisons. The 24 comparisons are those for 6 horizons × 2 sampling schemes (rolling and recursive) × 2 sample endpoints (2023:4 and 2019:4). For example, the "88%" in row (1), column (4) of Table 4.3 reflects the fact that b^{RW} was lower than b^{AR} in 21 of 24, or 88 percent, of the entries in rows (3a) and (3b) of Table 4.2.

Column (2) in Table 4.3 indicates that the (absolute) value of bias generally was very small–by a factor of 20 or more–than the mean given in Table 4.1. Columns (3) and (4) present our central results, on the ratio

[8] Outliers, such as the value of 6.9 for h=4, last forecast 2023:4 for the recursive scheme are not a result of a large value for b^{RW} but because of an unusually small value for b^{AR}. For that entry, $|b^{RW}| = 0.0049$ and $|b^{AR}| = 0.0007$. In light of the values of $|b^{RW}/b^{AR}|$ for h=4 for the sample that ends in 2019:4, and for the rolling scheme for both samples, the extremely small value of $|b^{AR}| = 0.0007$ in the denominator seems likely a fluke.

[9] The value for h=1 in row (4a) is 0.996, rounded up to 1.0 in the table.

of random walk to AR bias. For nine of the ten series, b^{RW} was usually smaller than b^{AR}, and typically by a substantial magnitude: in column (3), the median value (across 24 comparisons per series) of $|b^{RW}/b^{AR}|$ was 0.5 or less for these nine series. Labor productivity growth is the one exception: the median ratio was greater than 1, and that ratio was less than one in just slightly under half of the 24 comparisons (line (9), columns (3) and (4)). Our online appendix documents that the magnitude of $|b^{RW}|$ for labor productivity growth is similar to $|b^{RW}|$ for the other nine series. However, $|b^{AR}|$ for labor productivity happens to be exceptionally small for some horizons.

We see in columns (5)-(8) that by a mean squared or mean absolute error criterion, the random walk model fares quite poorly compared to the AR model, for all ten data series.

Rows (11) to (13) in Table 4.3 give some statistics aggregated over all ten series. Columns (4) and (3) in row (11) state that in 82 percent of the 240 comparisons (240 = 24 comparisons per series×10 series), b^{RW} was smaller than b^{AR}, with the median value of $|b^{RW}/b^{AR}|$ being 0.3. Rows (12) and (13) give the comparable values when looking at the 160 comparisons for $h \leq 12$ or the 80 comparisons for $h > 12$. These indicate a modest tendency for the benefits of the random walk model to decline with horizon. This is consistent with asymptotics in which h increases along with P (see the discussion in Section 2).

Table 4.4 repeats the analysis in Table 4.3 when the random walk forecast is replaced by a forecast from a zero mean AR(1). Let x_t be one of our series. We used OLS estimates of the model
$\Delta x_t = \rho \Delta x_{t-1}$+disturbance to forecast $\frac{x_{t+1}+x_{t+2}+\ldots+x_{t+h}}{h}$.[10]

The benefits of using an overdifferenced model are slightly less for the differenced AR model than for the random walk, but are qualitatively unchanged. For example, overall, the percentage of forecasts in which the overdifferenced model yielded bias that was smaller in absolute value fell from 82 percent to 79 percent (row (11), column (4)).

Finally, Table 4.5 repeats the analysis in Table 4.3 when, as in Theorem 2.2, the object being forecast is the level x_{t+h} rather than the average $\frac{x_{t+1}+x_{t+2}+\ldots+x_{t+h}}{h}$. We see in columns (3) and (4) that, compared to the forecasts of the average in Table 4.3, the median bias ratio was a little higher, but the percentage of forecast with a bias ratio less than 1 was also a little higher. As in previous tables, the differenced model almost always produced a higher mean squared or mean absolute prediction error (columns (6) and (8)).

[10]To avoid confusion: since x_t is one of our series, it may already be a difference of underlying source data. For our unemployment data, for example, Δx_t is the second difference of unemployment, since we differenced the raw unemployment series to induce stationarity.

In sum, the empirical results are strongly supportive of our theoretical and simulation results. In forecasting a stationary process, bias will generally be more concentrated around 0 for a forecast from a random walk or another overdifferenced model than from a stationary model.

5 Conclusion

Using asymptotic theory, simulations, and empirical estimates, we have shown that if the data are stationary, a misspecified overdifferenced model will tend to have forecast bias more concentrated around zero than will a stationary model. This applies even for the infeasible forecast that relies on population parameters from the stationary process that generated the data. Our empirical work indicates that a substantial reduction in bias can occur in practice; for some aggregate US data, bias from a random walk forecast was generally one-tenth to one-half of that of a plausibly specified stationary model.

6 References

Beaudry, Paul and Tim Willems, 2022, "On the Macroeconomic Consequences of Over-Optimism," *American Economic Journal: Macroeconomics*, 14(1): 38-59. https://doi.org/10.1257/mac.20190332.

Bennett, Julie K. and Michael T. Owyang, 2022, "On the Relative Performance of Inflation Forecasts," Federal Reserve Bank of St. Louis *Review*, 104(2), 131-48. https://doi.org/10.20955/r.104.131-48.

Brockwell, Peter J. and Richard A. Davis, 1993, *Time Series: Theory and Methods*, Springer-Verlag: New York. https://doi.org/10.1007/978-1-4419-0320-4.

Clark, Todd and Michael McCracken, 2013, "Advances in Forecast Evaluation," 1107-1201 in *Handbook of Economic Forecasting*, Vol. 2, G. Elliott and A. Timmermann (eds), Elsevier. https://doi.org/10.1016/B978-0-444-62731-5.00020-8.

Croushore, Dean, 2010, "An Evaluation of Inflation Forecasts from Surveys Using Real-Time Data," *The B.E. Journal of Macroeconomics*, 10:(1) (Contributions), Article10. https://doi.org/10.2202/1935-1690.1677.

Farmer, Leland E., Emi Nakamura, and Jón Steinsson, 2024, "Learning about the Long Run." *Journal of Political Economy*, 132(10): 3334-3377. https://doi.org/10.1086/730207.

Faust, Jon and Jonathan H. Wright, 2013, "Forecasting Inflation," *Handbook of Economic Forecasting Volume 2A*, 2-56. https://doi.org/10.1016/B978-0-444-53683-9.00001-3.

Giacomini, Raffaella and Halbert White, 2006, "Tests of Conditional Predictive Ability," *Econometrica* 74 (6), 1545-1578.https://doi.org/10.1111/j.1468-0262.2006.00718.x

Hendry, David F., 2006, "Robustifying forecasts from equilibrium-correction systems," *Journal of Econometrics* 135 (1–2), 399-426. https://doi.org/10.1016/j.jeconom.2005.07.029.

Lunsford, Kurt G. and Kenneth D. West, 2023, "Random Walk Forecasts of Stationary Processes Have Low Bias," Federal Reserve Bank of Cleveland Working Paper no. 23-18. https://doi.org/10.26509/frbc-wp-202318.

Lunsford, Kurt G. and Kenneth D. West, 2024, "An Empirical Evaluation of Some Long-Horizon Macroeconomic Forecasts," Federal Reserve Bank of Cleveland Working Paper no. 24-20. https://doi.org/10.26509/frbc-wp-202420.

McCracken, Michael W. and Serena Ng, 2021, "FRED-QD: A Quarterly Database for Macroeconomic Research," *Federal Reserve Bank of St. Louis Review*, First Quarter 2021. https://doi.org/10.20955/r.103.1-44.

Müller, Ulrich K. and Mark W. Watson, 2016, "Measuring Uncertainty about Long-Run Predictions," *Review of Economic Studies* 83 (4), 1711-1740. https://doi.org/10.1093/restud/rdw003.

Petropoulos, Fotios et al., 2022, "Forecasting: theory and practice," *International Journal of Forecasting* 38(3), 705-871. https://doi.org/10.1016/j.ijforecast.2021.11.001.

Richardson, Matthew and James H. Stock, 1989, "Drawing inferences from statistics based on multiyear asset returns," *Journal of Financial Economics* 25(2), 323-348. https://doi.org/10.1016/0304-405X(89)90086-X.

Sanchez, Ismael and Daniel Peña, 2001, "Properties of Predictors in Overdifferenced Nearly Nonstationary Autoregression," *Journal of Time Series Analysis*, 22(1), 45-66. https://doi.org/10.1111/1467-9892.00211.

Smith, Jeremy and Sanjay Yadav, 1994, "Forecasting costs incurred from unit differencing fractionally integrated processes," *International Journal of Forecasting* 10(4), 507-514. https://doi.org/10.1016/0169-2070(94)90019-1.

Stock, James H., 1996, "VAR, Error Correction and Pretest Forecasts at Long Horizons," *Oxford Bulletin of Economics and Statistics*, 58(4), 685-701. https://doi.org/10.1111/j.1468-0084.1996.mp58004006.x.

Timmermann, Allan, 2007, "An Evaluation of the World Economic Outlook Forecasts," IMF Staff Papers, 54(1): 1-33. https://doi.org/10.1057/palgrave.imfsp.9450007.

West, Kenneth D., 1996, "Asymptotic Inference About Predictive Ability," *Econometrica* 64:1067-1084. https://doi.org/10.2307/2171956

West, Kenneth D., 2006, "Forecast Evaluation," 100-134 in *Handbook of Economic Forecasting*, Vol. 1, G. Elliott, C. Granger and A. Timmermann (eds), Amsterdam: Elsevier. https://doi.org/10.1016/S1574-0706(05)01003-7.

7 Appendix

In this Appendix, we prove Theorem 2.1. The proof of Theorem 2.2 is analogous. Throughout, c or c with a subscript is a generic constant that varies from statement to statement. Let

$$\gamma_j = \mathrm{cov}(x_t, x_{t-j})$$

and define

$$\alpha_{th} \equiv h\psi_{1,t} + (h-1)\psi_{2,t+1} + ... + \psi_{h,t+h-1}. \tag{12}$$

Note that $|\alpha_{th}| < c_\alpha h\rho^t$for a constant c_α that does not depend on P:

$$\begin{aligned} |\alpha_{th}| &\leq h|\psi_{1,t}| + (h-1)|\psi_{2,t+1}| + ... + |\psi_{h,t+h-1}| \\ &\leq c\rho^t[h + (h-1)\rho + ... + \rho^{h-1}] \\ &\leq c\rho^t h(1 + \rho + ...\rho^{h-1}) \\ &\leq c_\alpha h\rho^t, c_\alpha \equiv c/(1-\rho). \end{aligned} \tag{13}$$

We turn now to the proof. From (3), we have

$$\text{period } t \text{ forecast of } [h\Delta x_{t+1} + (h-1)\Delta x_{t+2} + ... + \Delta x_{t+h}] = \alpha_{1h}\Delta x_t + \alpha_{2h}\Delta x_{t-1} + ... + \alpha_{th}\Delta x_1. \tag{14}$$

Using (14) in (5) , we have

$$\begin{aligned} b = (\frac{1}{Ph}) \times \{ &[h\Delta x_2 + (h-1)\Delta x_3 + ... + \Delta x_{h+1}] - [\alpha_{1h}\Delta x_1] \\ &+[h\Delta x_3 + (h-1)\Delta x_4 + ... + \Delta x_{h+2}] - [\alpha_{1h}\Delta x_2 + \alpha_{2h}\Delta x_1] \\ &+[h\Delta x_4 + (h-1)\Delta x_5 + ... + \Delta x_{h+3}] - [\alpha_{1h}\Delta x_3 + \alpha_{2h}\Delta x_2 + \alpha_{3h}\Delta x_1] \\ &+... \\ &+[h\Delta x_{P+1} + (h-1)\Delta x_P + ... + \Delta x_{P+h}] - [\alpha_{1h}\Delta x_P + \alpha_{2h}\Delta x_{P-1} + ... + \alpha_{Ph}\Delta x_1]\}. \end{aligned} \tag{15}$$

In (15), sum the terms multiplied by h (yielding $h(x_{P+1} - x_1)$), sum the terms multiplied by $h-1$, ..., sum the terms multiplied by α_1, ..., sum the terms multiplied by α_P. We get

$$\begin{aligned} b = (\frac{1}{Ph}) \times \{ & [h(x_{P+1} - x_1) + (h-1)(x_{P+2} - x_2) + ... + (x_{P+h} - x_h)] \\ & - [\alpha_{1h}(x_P - x_0) + \alpha_{2h}(x_{P-1} - x_0) + ... + \alpha_{Ph}(x_1 - x_0)] \ \} \\ \equiv (\frac{1}{Ph}) & \times (A_1 - A_2 - B_1 + B_2), \\ & A_1 \equiv hx_{P+1} + (h-1)x_{P+2} + ... + x_{P+h}, \\ & A_2 \equiv hx_1 + (h-1)x_2 + ... + x_h, \\ & B_1 \equiv \alpha_{1h}x_P + \alpha_{2h}x_{P-1} + ... + \alpha_{Ph}x_1, \\ & B_2 \equiv (\alpha_{1h} + \alpha_{2h} + ... + +\alpha_{Ph})x_0. \end{aligned} \tag{16}$$

We prove, in order, part (c), part (a), and, last, part (b) of Theorem 2.1. Proof of part (c) of Theorem (2.1): For h fixed, it suffices to show that the variance of each of the four terms is O(1). Under stationarity, this is immediate for A_1 and A_2. For B_1 and B_2: From (13), B_2 =O(1) is also immediate. For B_1: We have var (B_1)= $\sum\sum_{s,t=1}^{P}\alpha_{sh}\alpha_{th}\gamma_{|s-t|}$. Also, $|\alpha_{th}| \leq c\rho^t \Rightarrow \sum\sum_{s,t=1}^{\infty}\alpha_{sh}\alpha_{th} \leq 2c^2/[(1-\rho^2)(1-\rho)]$ $\equiv c_2 \Rightarrow \text{var}(B_1) \leq c_2\gamma_0$.

Proof of part (a) of Theorem (2.1): It suffices to show that the variance of each of A_1, A_2, B_1 and B_2 is O(h^3). Since $|\alpha_{th}| \leq c_\alpha h\rho^t$, $\text{var}(B_2) = O(h^2)$. We will write out the arguments for A_1 and B_1. The argument for A_2 is similar.

We have

$$\begin{aligned} \text{var}(A_1) = & [h^2 + (h-1)^2 + ... + 1]\gamma_0 \\ & + 2[h(h-1) + (h-1)(h-2) + ... + 2]\gamma_1 \\ & + 2[h(h-2) + (h-1)(h-3) + ... + 3]\gamma_2 \\ & + ... + 2h\gamma_{h-1} \\ & \leq 2[h^2 + (h-1)^2 + ... + 1][|\gamma_0| + |\gamma_1| + |\gamma_2| + ... + |\gamma_{h-1}|] \\ & \leq 2(\sum_{j=0}^{\infty}|\gamma_j|)[h^2 + (h-1)^2 + ... + 1] \\ & = O(h^3), \end{aligned} \tag{17}$$

because $\sum_{j=0}^{\infty}|\gamma_j|$ is finite and $h^2 + (h-1)^2 + ... + 1$=O(h^3).

And $\text{var}(B_1) = O(h^2)$, as follows. For c_α defined in (13), we have

$$\begin{aligned}
var(B_1) &= \gamma_0(\alpha_{1h}^2 + \alpha_{2h}^2 + ...\alpha_{Ph}^2) \\
&+ 2\gamma_1(\alpha_{1h}\alpha_{2h} + \alpha_{2h}\alpha_{3h} + ... + \alpha_{P-1,h}\alpha_{Ph}) \\
&+ ... + 2\gamma_{P-1}(\alpha_{1h}\alpha_{Ph}) \\
&\leq \gamma_0(c_\alpha h)^2(\rho^2 + \rho^4 + ... + \rho^{2P}) \\
&+ 2|\gamma_1|(c_\alpha h)^2(\rho^3 + \rho^5 + ... + \rho^{2P-1}) \\
&+ ... + 2|\gamma_{P-1}|(c_\alpha h)^2\rho^{P+1} \\
&\leq 2\gamma_0(c_\alpha h)^2[\frac{\rho^2}{1-\rho^2} + \frac{\rho^3}{1-\rho^2} + ... + \frac{\rho^{P+1}}{1-\rho^2}) \\
&\leq ch^2, c = \frac{2\gamma_0 c_\alpha^2 \rho^2}{(1-\rho^2)(1-\rho)}.
\end{aligned}$$

<u>Proof of part (b) of Theorem 2.1</u>: Under (8), the logic above continues to yield $\text{var}(B_1) = O(h^2)$ and $\text{var}(B_2) = O(h^2)$. For A_1, note first that under (8),

$$\gamma_0 + \sum\nolimits_{j=1}^{h-1}|\gamma_j| \leq \gamma_0 + c\sum\nolimits_{j=1}^{h-1} j^{2d-1} < c\int_0^h j^{2d-1}dj = c\frac{h^{2d}}{2d} = O(h^{2d}).$$

Using this result and the third from last line in (17),

$$\text{var}(A_1) \leq 2[h^2 + (h-1)^2 + ... + 1][\gamma_0| + \sum\nolimits_{j=1}^{h-1}|\gamma_j|] = O(h^{3+2d}).$$

Thus $(\frac{1}{Ph})^2\text{var}(A_1) = O(h^{1+2d}/P^2)$.and part (b) of Theorem 2.1 follows.

8 Tables

Table 3.1

Simulation results on bias, RMSPE and MAPE: random walk relative to AR(2) forecasts

	(1)		(2)		(3)		(4)		(5)		(6)	
(1)	h=1		h=4		h=8		h=12		h=20		h=40	
	med	%<1	med	%<1	med	%<1	med	%<1	med	%<1	med	%<1
(2) $\|b^{RW} / b^{AR}\|$												
(2a) Rolling	0.3	84%	0.3	80%	0.4	77%	0.5	74%	0.5	72%	0.7	67%
(2b) Recursive	0.1	92%	0.2	90%	0.2	87%	0.3	85%	0.3	82%	0.5	72%
(2c) Population	0.2	91%	0.2	88%	0.2	87%	0.3	83%	0.4	75%	0.7	60%
(3) RMSPE^{RW} / RMSPE^{AR}												
(3a) Rolling	1.1	0%	1.5	0%	2.0	0%	2.3	0%	2.8	0%	3.5	0%
(3b) Recursive	1.2	0%	1.6	0%	2.0	0%	2.4	0%	2.9	0%	4.0	0%
(3c) Population	1.2	0%	1.6	0%	2.1	0%	2.6	0%	3.3	0%	4.9	0%
(4) MAPE^{RW} / MAPE^{AR}												
(4a) Rolling	1.1	0%	1.5	0%	2.0	0%	2.2	0%	2.7	0%	3.3	0%
(4b) Recursive	1.2	0%	1.6	0%	2.0	0%	2.4	0%	2.9	0%	3.8	0%
(4c) Population	1.2	0%	1.6	0%	2.1	0%	2.5	0%	3.2	0%	4.6	0%
(5) P	112		109		105		101		93		73	
(6) h/P	0.01		0.04		0.08		0.12		0.22		0.55	

Notes:

1. 1000 simulation samples of size T+1=160 were generated from a stationary AR(2) model. The DGP was: $x_t = 0.5x_{t-1} - 0.1x_{t-2} + u_t$, $u_t \sim$ iid N(0,1). Predictions were made from a driftless random walk model (superscript *RW*) and an AR(2) model (superscript *AR*). In lines labeled "rolling" and "recursive," AR(2) predictions were made using least squares estimates; in lines labeled "population," population parameters were used.

2. For horizons h given in row (1), the number of predictions P in a given sample is given in row (5). In a given sample, P predictions (row (5)) were used to compute bias b (row (2)), root mean squared prediction error (RMSPE, row (3)) and mean absolute prediction error (MAPE, row (4)). Here and throughout, bias is the sample average of the difference between realization and prediction, RMSPE is the square root of the average of squared forecast errors and MAPE is the average of the absolute value of forecast error.

3. In rows (2)-(4), the "med" column reports the median value, across the 1000 samples, of the indicated statistic. For all three ratios (bias, RMSPE, MAPE), a value less than one indicates that the random walk model usually produced a lower value than did the AR(2) model. The "%<1" column reports the percentage of the 1000 samples in which the ratio was less than 1.

4. See text for additional details.

Table 3.2

Simulation results on normalized var(b^{RW}), various P

A. AR(2) DGP

(1) P for h=1	(2)	(3)	(4) var(b^{RW})×P^2	(5)	(6)	(7) var(b^{RW})×P
(1)	h=1	h=4	h=8	h=12	h=0.2P	h=0.2P
(2) 112	3	8	15	23	35	0.37
(3) 352	1.01	1.00	1.08	1.01	3.17	1.01
(4) 752	1.10	0.97	1.05	1.07	7.19	1.08
(5) 1152	1.04	1.05	1.04	1.01	10.75	1.05

B. $I(d)$ DGP

(1) P for h=1	(2)	(3)	(4) var(b^{RW})×P^2	(5)	(6)	(7) var(b^{RW})×$P^{(1\text{-}2d)}$
(1)	h=1	h=4	h=8	h=12	h=0.2P	h=0.2P
(2) 112	3	9	21	35	67	0.04
(3) 352	0.91	1.03	1.06	1.08	6.27	0.99
(4) 752	1.02	1.09	1.13	1.15	20.33	0.94
(5) 1152	0.9	1.17	1.14	1.13	39.61	0.92

Notes:

1. The simulation results are based on 1000 samples of size T+1=1200. In panel A, the DGP was the stationary AR(2) model described in notes to Table 3.1. In panel B, the DGP was the I(d) model $(1\text{-}L)^d x_t = u_t$, d=0.31, u_t ~ iid N(0,1). As in Table 3.1, h is the horizon and P the number of predictions. The h=0.2P values for sample sizes T+1=160, 400, 800, and 1200 are h=19, 59, 126, and 193 with corresponding values of P=94, 294, 627, and 960. See notes to Table 3.1.

2. var(b^{RW}) is the variance of random walk bias across 1000 simulation samples. Row (2) in each panel gives the normalized value of var(b^{RW}) when P=112, with the normalization factor being either P^2, P or $P^{(1\text{-}2d)}$ as given in the headers in the table. In rows (3)-(5) of each panel, columns (2)-(7) express normalized var(b^{RW}) relative to the value in row (2).

3. Our asymptotic theory concludes that if h is held fixed as $P\rightarrow\infty$, var(b^{RW})×P^2 = O(1). If h=cP, var(b^{RW})×P = O(1) for the DGP in panel A, var(b^{RW})×$P^{(1\text{-}2d)}$ = O(1) for the DGP in panel B.

Table 4.1

Data summary

(1)	(2) mean	(3) standard deviation	(4) first order autocorrelation
(1) Per capita GDP growth	1.8	2.3	0.4
(2) CPI inflation	2.6	1.9	0.3
(3) Growth in household assets	6.1	6.2	0.2
(4) Growth in housing starts	-0.6	27.6	0.2
(5) Change in interest rate on 10 year treasuries	-0.1	0.4	0.3
(6) Inventory growth	2.6	3.4	0.6
(7) Industrial production growth	1.9	4.8	0.7
(8) Growth in M3	5.5	2.8	0.6
(9) Labor productivity growth	2.0	2.5	0.1
(10) Change in unemployment rate	0.0	0.3	0.7

Notes

1. Data are quarterly. The statistics in columns (2)-(4) were computed from a sample running 1984:1-2019:4.

Table 4.2

Bias, RMSPE and MAPE: random walk relative to AR, per capita GDP growth
Forecasts of h period ahead averages

	A: Last forecast 2023:4						B: Last forecast 2019:4					
(1)	h=1	h=4	h=8	h=12	h=20	h=40	h=1	h=4	h=8	h=12	h=20	h=40
(2) b^{RW}	0.0	0.0	-0.1	0.0	-0.2	-0.3	0.0	0.0	0.0	-0.1	-0.2	-0.4
(3) $\|b^{RW} / b^{AR}\|$												
(3a) Rolling	0.0	0.0	0.4	0.0	1.1	0.6	0.1	0.0	2.0	0.5	0.6	0.5
(3b) Recursive	0.0	6.8	0.7	0.0	0.4	0.4	0.0	0.0	0.2	0.2	0.4	0.5
(4) RMSPE^{RW} / RMSPE^{AR}												
(4a) Rolling	1.0	1.3	2.0	2.5	2.0	2.5	1.2	1.4	1.6	1.7	2.1	2.5
(4b) Recursive	1.1	1.8	2.5	3.1	2.4	2.7	1.2	1.5	1.7	1.9	2.4	2.7
(5) MAPE^{RW} / MAPE^{AR}												
(5a) Rolling	1.1	1.5	1.8	2.0	1.9	2.0	1.3	1.4	1.5	1.6	2.1	2.0
(5b) Recursive	1.2	1.7	2.1	2.4	2.5	2.4	1.3	1.5	1.7	1.9	2.5	2.1

Notes:

1. b^{RW} is bias in predictions of the average value of GDP growth over the horizons indicated in row 1, from a model that assumes that GDP growth follows a driftless random walk. All forecasting exercises are pseudo-out-of-sample.

2. For a given horizon h, the first forecast is for average GDP growth from 1996:1 through the next h-1 quarters. In panel A, the last forecasted observation is for average GDP growth over h quarters ending in 2023:4; the number of forecasts and forecast errors P is 112-(h-1) (see line (5) in Table 3.1). In panel B, the date of the first forecast is the same but the ending quarter is 2019:4; P is smaller by 16, for each horizon.

3. b^{AR} is bias in forecasts computed from a series of estimates of an AR(2) model. In the regression relying on the first sample used to estimate the AR model, the left-hand-side variable spans the 48 quarters from 1984:1 to 1995:4. For h=1, the left-hand-side in the last estimation sample in panel A runs from 2011:4 to 2023:3 (rolling) or 1984:1 to 2023:3 (recursive); the dates in panel B are 2007:4-2019:3 and 1984:1-2019:3. End dates for h>1 are shifted back to accommodate the longer horizon (e.g., for h=4, the final recursive sample in panel A runs 1984:1-2022:4).

4. Rows (3a) and (3b) give the absolute value of the ratio of the bias from the random walk forecast to the bias from the AR forecast. Rows (4a) and (4b) give the ratio of root mean squared prediction errors (RMSPEs). Rows (5a) and (5b) give the ratio of mean absolute prediction errors (MAPEs). For all three ratios (bias, RMSPE, MAPE), a value less than one indicates that the random walk model produced a lower value than did the AR(2) model.

Table 4.3

Bias, RMSPE and MAPE: random walk relative to AR
Forecasts of h period ahead averages

(1)	(2) median $\lvert b^{RW}\rvert$	(3) bias ratio median	(4) bias ratio %<1	(5) RMSPE ratio median	(6) RMSPE ratio %<1	(7) MAPE ratio median	(8) MAPE ratio %<1
(1) Per capita GDP growth	0.1	0.4	88%	2.0	4%	1.9	0%
(2) CPI inflation	0.1	0.1	92%	1.9	0%	1.6	0%
(3) Growth in HH assets	0.1	0.3	88%	1.9	0%	2.0	0%
(4) Growth in housing starts	0.3	0.3	83%	1.8	0%	1.7	0%
(5) Change in 10Y rate	0.0	0.2	100%	3.4	0%	3.4	0%
(6) Inventory growth	0.1	0.2	83%	1.4	0%	1.4	0%
(7) IP growth	0.2	0.5	79%	1.3	8%	1.5	8%
(8) Growth in M3	0.0	0.2	96%	1.7	4%	1.6	4%
(9) Labor prod growth	0.1	1.7	42%	2.2	0%	1.9	0%
(10) Chg in unemployment rate	0.0	0.1	71%	1.2	17%	1.2	17%
(11) All 10 series, all horizons	0.0	0.3	82%	1.8	3%	1.7	3%
(12) All 10 series, $h \leq 12$	0.0	0.2	86%	1.5	5%	1.5	4%
(13) All 10 series, $h>12$	0.2	0.4	75%	2.4	0%	2.1	0%

Notes:

1. For per capita GDP growth (row 1), columns (2)-(8) are computed from the values in Table 4.2, as follows:

a. In column (2), "median $|b^{RW}|$" is median absolute value bias for the random walk model, computed from the absolute values of the 12 entries in row (2) of Table 4.2.

b. Column (3) presents the median value of the absolute value of the ratio of the bias of random walk model to the bias of the AR model (median of $|b^{RW}/b^{AR}|$), computed from the 24 values of rows (3a) and (3b) of Table 4.2.

c. Column (4) presents the percentage of forecast comparisons in which bias for the random walk model is less than that for the AR(2) model in absolute value, again computed from the 24 values in rows (3a) and (3b) of Table 4.2.

d. Columns (5) and (6) are comparable values for ratio of root mean squared prediction error $RMSPE^{RW}$ / $RMSPE^{AR}$, while columns (7) and (8) are comparable values for ratio of mean absolute prediction error $MAPE^{RW}$ / $MAPE^{AR}$. These are computed from the 24 values in rows (4a) and (4b), or rows (5a) and (5b), in Table 4.2.

2. The values in rows (2)-(10) are computed analogously to those for GDP growth. For series other than GDP growth, the online appendix gives the underlying values comparable to those in Table 4.2.

3. For all three ratios (bias, RMSPE, MAPE), a value less than one indicates that the random walk model produced a lower value than did the AR(2) model.

Table 4.4

Bias, RMSPE and MAPE: differenced AR relative to AR
Forecasts of h period ahead averages

(1)	(2) median $\lvert b^{DAR}\rvert$	(3) bias ratio median	(4) bias ratio %<1	(5) RMSPE ratio median	(6) RMSPE ratio %<1	(7) MAPE ratio median	(8) MAPE ratio %<1
(1) Per capita GDP growth	0.1	0.1	92%	1.6	0%	1.4	0%
(2) CPI inflation	0.2	0.5	75%	1.7	0%	1.7	0%
(3) Growth in HH assets	0.1	0.3	92%	1.6	0%	1.6	0%
(4) Growth in housing starts	0.2	0.2	88%	1.5	0%	1.4	0%
(5) Change in 10Y rate	0.0	0.2	100%	3.0	0%	3.0	0%
(6) Inventory growth	0.1	0.3	88%	1.4	0%	1.4	0%
(7) IP growth	0.3	0.6	67%	1.3	0%	1.6	0%
(8) Growth in M3	0.0	0.1	88%	1.7	0%	1.6	0%
(9) Labor prod growth	0.1	2.1	38%	1.7	0%	1.4	0%
(10) Chg in unemployment rate	0.0	0.7	63%	1.2	8%	1.3	4%
(11) All 10 series, all horizons	0.0	0.3	79%	1.6	1%	1.5	0%
(12) All 10 series, $h \leq 12$	0.1	0.3	81%	1.4	1%	1.3	1%
(13) All 10 series, $h>12$	0.2	0.4	75%	2.1	0%	1.9	0%

Notes:

1. This table differs from the previous table in that random walk forecasts were replaced by forecasts from least squares estimates of the model $\Delta x_t=\rho\Delta x_{t-1}$+disturbance. Here, x_t is one of the variables listed in column (1). In column (2), b^{DAR} is bias from that model, where "DAR" stands for "differenced AR." The numerators of the RMSPE and MAPE ratios also come from the differenced AR model.

2. For all three ratios (bias, RMSPE, MAPE), a value less than one indicates that the differenced AR(1) model produced a lower value than did the AR(2) model.

3. See notes to Table 4.3.

Table 4.5

Bias, RMSPE and MAPE: random walk relative to AR
Forecasts of point in time observations h periods ahead

(1)	(2) median $\lvert b_P^{RW}\rvert$	(3) bias ratio median	(4) bias ratio %<1	(5) RMSPE ratio median	(6) RMSPE ratio %<1	(7) MAPE ratio median	(8) MAPE ratio %<1
(1) Per capita GDP growth	0.1	0.2	88%	1.3	0%	1.3	0%
(2) CPI inflation	0.1	0.4	92%	1.4	4%	1.4	0%
(3) Growth in HH assets	0.2	0.3	83%	1.3	0%	1.4	0%
(4) Growth in housing starts	0.5	0.4	83%	1.3	0%	1.4	0%
(5) Change in 10Y rate	0.0	0.3	96%	1.4	0%	1.3	0%
(6) Inventory growth	0.1	0.3	88%	1.3	0%	1.4	0%
(7) IP growth	0.5	0.6	83%	1.2	8%	1.4	8%
(8) Growth in M3	0.1	0.2	96%	1.3	4%	1.3	4%
(9) Labor prod growth	0.2	1.3	42%	1.3	0%	1.4	0%
(10) Chg in unemployment rate	0.0	0.2	79%	1.1	17%	1.2	8%
(11) All 10 series, all horizons	0.0	0.4	83%	1.3	3%	1.4	2%
(12) All 10 series, $h\leq12$	0.0	0.2	86%	1.3	5%	1.4	3%
(13) All 10 series, $h>12$	0.3	0.5	78%	1.3	0%	1.4	0%

Notes:

1. This table differs from Table 4.3 in that the object forecasted is the point in time realization h periods ahead rather than the average value over the next h periods. In column (2), b_P^{RW} is the bias that results when a random walk model is used to make that forecast.

2. For all three ratios (bias, RMSPE, MAPE), a value less than one indicates that the random walk model produced a lower value than did the AR(2) model.

3. See notes to Table 4.3.

9 Online Appendix

Random Walk Forecasts of Stationary Processes Have Low Bias

Online Appendix

Kurt G. Lunsford
Federal Reserve Bank of Cleveland

Kenneth D. West
University of Wisconsin

July 2023; revised January 2025

This online appendix presents data details and simulation and empirical results omitted from the paper to save space.

1. Data:

a. FRED mnemonics for our data series:

(1) Per capita GDP growth	A939RX0Q048SBEA
(2) CPI inflation	CPIAUCSL
(3) Growth in household assets	TABSHNO
(4) Growth in housing starts	HOUST
(5) Change in interest rate on 10 year treasuries	GS10
(6) Inventory growth	INVCQRMTSPL
(7) Industrial production growth	INDPRO
(8) Growth in M3	MABMM301USM189S
(9) Labor productivity growth	OPHPBS
(10) Change in unemployment rate	UNRATE

Series that are available monthly on FRED were averaged to produce our quarterly series.

b. The initial version of this paper used four of these 10 series: per capita GDP growth, CPI inflation, labor productivity growth, and M3 growth. As stated in the text, when we expanded to 10 series, the additional six were chosen from six different groups in McCracken and Ng (2021–henceforth MN), with us generally following MN. "Generally" because, for continuity, we continued to use the four initial series, which implied the following slight departures from MN: (a)per capita real GDP growth instead of real GDP growth as our series from MN's group 1; (b)stationarity induced in log CPI with first rather than second differences (MN's group 6); (c)growth in nominal M3 rather than growth in real M2 (MN's group 9).

2. Tables A3.1a and A3.1b present simulation results analogous to those presented in Table 3.1, but with a different DGP and a different stationary forecasting model.

   a. Table A3.1a: the DGP is $x_t = u_t \sim$ iid N(0,1); the stationary model for comparison with the random walk model is iid possibly around a nonzero mean. That is, the stationary forecast is $x_{t+h}=\bar{x}$ for all h, where $\bar{x}$ is the sample mean from the relevant rolling or recursive sample. The population forecast is $x_{t+h}=0$ for all h.

   b. Table 3.2B and Table A3.1b: the DGP is $(1\text{-}L)^d x_t = u_t \sim$ iid N(0,1), with d=0.31.

      i. As in Chung and Baillie (1993), data in a given simulation sample were drawn from an N(0,V) DGP where V is the variance-covariance matrix of a (1200×1) vector of x's. For example, the diagonals of V were set to the variance of x_t (i.e., to $\Gamma(1\text{-}2d)/\Gamma^2(1\text{-}d)$, where Γ is the gamma function and, again, d=0.31). Off-diagonals were set in accordance with the autocovariances of x_t. See Brockwell and Davis (1993, p 522) for formulas. All 1200 observations were used in Table 3.2B and the first 160 were used in Table A3.1b.

      ii. The technique of Geweke and Porter-Hudak (1983) was used to estimate d, setting the estimate to -0.49 or 0.49 if the estimate fell outside the interval (-0.49,0.49). Forecasts were made using the autoregressive representation of x_t. The autoregression was truncated at lag 48 (rolling samples) or at the first observation in the sample (recursive samples). As in our AR(2) forecasts, h-period-ahead-forecasts were made using the chain rule of forecasting. For example, for rolling samples, forecasts were based on the AR(48)

$$x_t-\bar{x} = \hat{\varphi}_1(x_{t-1}-\bar{x}) + \hat{\varphi}_2(x_{t-2}-\bar{x}) + \ldots + \hat{\varphi}_{48}(x_{t-48}-\bar{x}).$$

      Here $\bar{x}$ is the sample mean, $\hat{\varphi}_1=\hat{d}$, $\hat{\varphi}_2=-\hat{d}(\hat{d}\text{-}1)/2!$, ... and $\hat{d}$ is the estimate of d. Population forecasts were made from this AR representation using population values, i.e., setting $\hat{d}$=0.31 and $\bar{x}$=0 (=population mean of x).

3. Tables A3.1(c)-A3.1(e) are analogous to Tables 3.1, 3.1(a) and 3.1(b), except that the RW model is replaced by an overdifferenced AR(1) model.

4. Table A3.1(f) is analogous to Table 3.1, except that the disturbance to the AR(2) follows a GARCH process: $u_t = \sigma_t v_t$, v_t~iid N(0,1), $\sigma_t^2 = \omega+\alpha\sigma_{t-1}^2+\beta u_{t-1}^2$ with ω=0.25, α=0.25 and β=0.50.

5. Table A3.2 is analogous to Table 3.2, except that the RW model is replaced by an overdifferenced AR(1) model.

6. Tables A4.2-(2) through A4.2-(10) present results analogous to those presented for GDP growth in Table 4.2, for series numbered (2)-(10) in Table 4.1.

7. Table A4.3-BIC presents results analogous to Table 4.3, except that lag length was chosen by BIC with a maximum lag length of 4. (Table 4.3 relied on an AR(2) model for its results.)

Additional references:

Chung, Ching-fan and Richard T. Baillie, 1993, “Small Sample Bias in Conditional Sum-of-Squares Estimators of Fractionally Integrated ARMA Models,” *Empirical Economics* 18, 791-806. https://doi.org/10.1007/BF01205422

Geweke, John and Susan Porter-Hudak, 1983, “The Estimation and Application of Long Memory Time Series Models,” *Journal of Time Series Analysis* 4(4), 221-238. https://doi.org/10.1111/j.1467-9892.1983.tb00371.x

Table A3.1a

Simulation results on bias, RMSPE and MAPE, random walk relative to iid forecasts, iid DGP

(1)	(1) h=1 med	%<1	(2) h=4 med	%<1	(3) h=8 med	%<1	(4) h=12 med	%<1	(5) h=20 med	%<1	(6) h=40 med	%<1
(2) $\lvert b^{RW} / b^{iid}\rvert$												
(2a) Rolling	0.3	85%	0.3	81%	0.4	77%	0.4	76%	0.5	74%	0.7	67%
(2b) Recursive	0.1	91%	0.2	90%	0.2	87%	0.3	87%	0.3	81%	0.6	71%
(2c) Population	0.1	92%	0.2	88%	0.2	86%	0.3	83%	0.4	77%	0.7	61%
(3) RMSPE^{RW} / RMSPE^{iid}												
(3a) Rolling	1.4	0%	2.2	0%	2.8	0%	3.3	0%	4.1	0%	5.3	0%
(3b) Recursive	1.4	0%	2.2	0%	2.9	0%	3.5	0%	4.3	0%	5.9	0%
(3c) Population	1.4	0%	2.2	0%	3.0	0%	3.6	0%	4.7	0%	7.2	0%
(4) MAPE^{RW} / MAPE^{iid}												
(4a) Rolling	1.4	0%	2.2	0%	2.8	0%	3.3	0%	4.0	0%	5.0	0%
(4b) Recursive	1.4	0%	2.2	0%	2.9	0%	3.4	0%	4.2	0%	5.6	0%
(4c) Population	1.4	0%	2.2	0%	3.0	0%	3.6	0%	4.6	0%	6.7	0%
(5) P	112		109		105		101		93		73	
(6) h/P	0.01		0.04		0.08		0.12		0.22		0.55	

Notes: See notes to Table 3.1. This differs from Table 3.1 in that the DGP and stationary forecasting model are iid rather than AR(2).

Table A3.1b

Simulation results on bias, RMSPE and MAPE, random walk relative to I(d) forecasts, I(d) DGP

(1)	(1) h=1 med	%<1	(2) h=4 med	%<1	(3) h=8 med	%<1	(4) h=12 med	%<1	(5) h=20 med	%<1	(6) h=40 med	%<1
(2) $\lvert b^{RW} / b^{I(d)}\rvert$												
(2a) Rolling	0.2	90%	0.3	86%	0.4	82%	0.4	82%	0.5	79%	0.7	72%
(2b) Recursive	0.1	93%	0.2	90%	0.2	88%	0.3	86%	0.4	83%	0.6	72%
(2c) Population	0.1	91%	0.2	88%	0.2	85%	0.3	83%	0.4	79%	0.6	67%
(3) RMSPE^{RW} / $\mathrm{RMSPE}^{I(d)}$												
(3a) Rolling	1.2	4%	1.4	0%	1.7	0%	1.8	0%	2.0	0%	2.3	0%
(3b) Recursive	1.2	5%	1.5	0%	1.7	0%	1.8	0%	2.0	0%	2.4	0%
(3c) Population	1.2	0%	1.5	0%	1.8	0%	1.9	0%	2.2	0%	2.7	0%
(4) MAPE^{RW} / $\mathrm{MAPE}^{I(d)}$												
(4a) Rolling	1.2	3%	1.4	0%	1.6	0%	1.8	0%	2.0	0%	2.2	0%
(4b) Recursive	1.2	4%	1.5	0%	1.7	0%	1.8	0%	2.0	0%	2.3	0%
(4c) Population	1.2	0%	1.5	0%	1.8	0%	1.9	0%	2.1	0%	2.6	0%
(5) P	112		109		105		101		93		73	
(6) h/P	0.01		0.04		0.08		0.12		0.22		0.55	

Notes: See notes to Table 3.1. This differs from Table 3.1 in that the DGP and stationary forecasting model are I(d) rather than AR(2).

Table A3.1c

Simulation results on bias, RMSPE and MAPE, overdifferenced AR(1) relative to AR(2) forecasts, , AR(2) DGP

(1)	(1) h=1 med	%<1	(2) h=4 med	%<1	(3) h=8 med	%<1	(4) h=12 med	%<1	(5) h=20 med	%<1	(6) h=40 med	%<1
(2) $\lvert b^{DAR} / b^{AR} \rvert$												
(2a) Rolling	0.3	83%	0.3	80%	0.4	76%	0.5	74%	0.5	72%	0.7	67%
(2b) Recursive	0.2	90%	0.2	90%	0.2	87%	0.3	85%	0.3	82%	0.5	72%
(2c) Population	0.2	90%	0.2	88%	0.3	86%	0.3	83%	0.4	75%	0.7	60%
(3) RMSPE^{DAR} / RMSPE^{AR}												
(3a) Rolling	1.1	0%	1.5	0%	1.9	0%	2.1	0%	2.6	0%	3.3	0%
(3b) Recursive	1.1	0%	1.5	0%	1.9	0%	2.2	0%	2.7	0%	3.7	0%
(3c) Population	1.2	0%	1.6	0%	2.0	0%	2.4	0%	3.1	0%	4.5	0%
(4) MAPE^{DAR} / MAPE^{AR}												
(4a) Rolling	1.1	1%	1.5	0%	1.8	0%	2.1	0%	2.5	0%	3.1	0%
(4b) Recursive	1.1	0%	1.5	0%	1.9	0%	2.2	0%	2.7	0%	3.5	0%
(4c) Population	1.2	0%	1.6	0%	2.0	0%	2.4	0%	2.9	0%	4.3	0%
(5) P	112		109		105		101		93		73	
(6) h/P	0.01		0.04		0.08		0.12		0.22		0.55	

Notes: See notes to Table 3.1. This differs from Table 3.1 in that the overdifferenced model is an AR(1) rather than a random walk.

Table A3.1d

Simulation results on bias, RMSPE and MAPE, overdifferenced AR(1) relative to iid forecasts, iid DGP

(1)	(1) h=1 med	%<1	(2) h=4 med	%<1	(3) h=8 med	%<1	(4) h=12 med	%<1	(5) h=20 med	%<1	(6) h=40 med	%<1
(2) $\lvert b^{DAR} / b^{iid}\rvert$												
(2a) Rolling	0.3	84%	0.4	80%	0.4	78%	0.4	76%	0.5	73%	0.7	67%
(2b) Recursive	0.1	91%	0.2	90%	0.2	87%	0.3	86%	0.3	81%	0.6	72%
(2c) Population	0.1	90%	0.2	88%	0.2	86%	0.3	83%	0.4	76%	0.7	60%
(3) RMSPE^{DAR} / RMSPE^{iid}												
(3a) Rolling	1.2	0%	1.7	0%	2.2	0%	2.6	0%	3.1	0%	4.0	0%
(3b) Recursive	1.2	0%	1.7	0%	2.3	0%	2.7	0%	3.3	0%	4.4	0%
(3c) Population	1.2	0%	1.8	0%	2.3	0%	2.8	0%	3.6	0%	5.4	0%
(4) MAPE^{DAR} / MAPE^{iid}												
(4a) Rolling	1.2	0%	1.7	0%	2.2	0%	2.5	0%	3.0	0%	3.8	0%
(4b) Recursive	1.2	0%	1.7	0%	2.2	0%	2.6	0%	3.1	0%	4.2	0%
(4c) Population	1.2	0%	1.8	0%	2.3	0%	2.8	0%	3.5	0%	5.1	0%
(5) P	112		109		105		101		93		73	
(6) h/P	0.01		0.04		0.08		0.12		0.22		0.55	

Notes: See notes to Table 3.1. This differs from Table 3.1 in that the overdifferenced model is an AR(1) rather than a random walk, and the DGP and stationary forecasting model are iid.

Table A3.1e

Simulation results on bias, RMSPE and MAPE, overdifferenced AR(1) relative to I(d) forecasts, I(d) DGP

(1)	(1) h=1 med	%<1	(2) h=4 med	%<1	(3) h=8 med	%<1	(4) h=12 med	%<1	(5) h=20 med	%<1	(6) h=40 med	%<1
(2) $\lvert b^{DAR} / b^{I(d)}\rvert$												
(2a) Rolling	0.2	89%	0.3	85%	0.4	82%	0.4	82%	0.5	79%	0.7	72%
(2b) Recursive	0.1	93%	0.2	89%	0.3	87%	0.3	86%	0.4	82%	0.6	73%
(2c) Population	0.2	88%	0.2	87%	0.3	84%	0.3	82%	0.4	78%	0.7	67%
(3) RMSPE^{DAR} / RMSPE$^{I(d)}$												
(3a) Rolling	1.1	19%	1.3	1%	1.4	0%	1.6	0%	1.7	0%	2.0	0%
(3b) Recursive	1.1	15%	1.3	1%	1.5	0%	1.6	0%	1.8	0%	2.1	0%
(3c) Population	1.1	0%	1.4	0%	1.5	0%	1.7	0%	1.9	0%	2.4	1%
(4) MAPE^{DAR} / MAPE$^{I(d)}$												
(4a) Rolling	1.1	18%	1.3	0%	1.4	0%	1.5	0%	1.7	0%	1.9	1%
(4b) Recursive	1.1	15%	1.3	2%	1.5	0%	1.6	0%	1.7	0%	2.0	1%
(4c) Population	1.1	0%	1.4	0%	1.5	0%	1.7	0%	1.9	1%	2.2	2%
(5) P	112		109		105		101		93		73	
(6) h/P	0.01		0.04		0.08		0.12		0.22		0.55	

Notes: See notes to Table 3.1. This differs from Table 3.1 in that the overdifferenced model is an AR(1) rather than a random walk, and the DGP and stationary forecasting model are I(d).

Table A3.1f

Simulation results on bias, RMSPE and MAPE, random walk relative to AR(2) forecasts, GARCH DGP

	(1)		(2)		(3)		(4)		(5)		(6)	
(1)	h=1		h=4		h=8		h=12		h=20		h=40	
	med	%<1	med	%<1	med	%<1	med	%<1	med	%<1	med	%<1
(2) $\lvert b^{RW} / b^{AR}\rvert$												
(2a) Rolling	0.3	86%	0.3	81%	0.4	78%	0.4	75%	0.6	71%	0.7	64%
(2b) Recursive	0.2	90%	0.2	89%	0.2	86%	0.3	83%	0.4	80%	0.6	70%
(2c) Population	0.2	91%	0.2	89%	0.2	84%	0.3	82%	0.4	75%	0.7	63%
(3) RMSPE^{RW} / RMSPE^{AR}												
(3a) Rolling	1.1	0%	1.6	0%	2.0	0%	2.3	0%	2.8	0%	3.5	0%
(3b) Recursive	1.2	0%	1.6	0%	2.1	0%	2.4	0%	3.0	0%	4.0	0%
(3c) Population	1.2	0%	1.7	0%	2.2	0%	2.6	0%	3.4	0%	4.9	0%
(4) MAPE^{RW} / MAPE^{AR}												
(4a) Rolling	1.1	0%	1.5	0%	1.9	0%	2.2	0%	2.7	0%	3.3	0%
(4b) Recursive	1.1	0%	1.6	0%	2.0	0%	2.4	0%	2.9	0%	3.7	0%
(4c) Population	1.2	0%	1.7	0%	2.2	0%	2.6	0%	3.2	0%	4.5	0%
(5) P	112		109		105		101		93		73	
(6) h/P	0.01		0.04		0.08		0.12		0.22		0.55	

Notes: See notes to Table 3.1. This differs from Table 3.1 in that the DGP has GARCH disturbances.

Table A3.2

Simulation results on normalized var(b^{DAR}), various P

A. AR(2) DGP

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	P for h=1			var(b^{DAR})×P^2			var(b^{DAR})×P
(1)		h=1	h=4	h=8	h=12	h=0.2P	h=0.2P
(2)	112	3	9	16	23	35	0.38
(3)	352	1.03	1.00	1.08	1.00	3.13	1.00
(4)	752	1.09	0.96	1.05	1.06	7.08	1.06
(5)	1152	1.05	1.05	1.04	1.01	10.60	1.04

B. I(d) DGP

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	P for h=1			var(b^{DAR})×P^2			var(b^{DAR})×$P^{(1-2d)}$
(1)		h=1	h=4	h=8	h=12	h=0.2P	h=0.2P
(2)	112	4	10	23	38	70	0.04
(3)	352	0.94	1.04	1.06	1.08	6.08	0.96
(4)	752	1.05	1.09	1.14	1.15	19.56	0.90
(5)	1152	0.99	1.17	1.15	1.12	38.01	0.88

See notes to Table 3.2. This differs from Table 3.2 in that the overdifferenced model is an AR(1) rather than a random walk.

Table A4.2-(2) CPI inflation

Bias, RMSPE and MAPE: random walk relative to AR
Forecasts of h period ahead averages

		A: Last forecast 2023:4						B: Last forecast 2019:4					
(1)		h=1	h=4	h=8	h=12	h=20	h=40	h=1	h=4	h=8	h=12	h=20	h=40
(2)	b^{RW}	0.0	0.0	0.1	0.2	0.1	-0.1	0.0	0.0	0.0	0.0	-0.1	-0.3
(3)	$\lvert b^{RW}/b^{AR}\rvert$												
(3a)	Rolling	0.0	0.1	3.0	1.8	0.4	0.3	0.0	0.1	0.1	0.1	0.2	0.4
(3b)	Recursive	0.0	0.0	0.3	0.4	0.2	0.2	0.0	0.0	0.0	0.0	0.2	0.3
(4)	$RMSPE^{RW}$ / $RMSPE^{AR}$												
(4a)	Rolling	1.1	1.4	1.6	1.9	2.4	3.3	1.2	1.9	2.2	2.6	2.9	3.4
(4b)	Recursive	1.2	1.5	1.6	1.9	2.1	2.3	1.2	1.7	1.9	2.0	2.1	2.4
(5)	$MAPE^{RW}$ / $MAPE^{AR}$												
(5a)	Rolling	1.1	1.4	1.6	1.8	1.9	2.6	1.2	1.6	2.0	2.1	2.1	2.6
(5b)	Recursive	1.2	1.3	1.5	1.6	1.6	1.8	1.2	1.4	1.6	1.7	1.6	1.7

See notes to Table 4.2.

Table A4.2-(3) Growth in household assets

Bias, RMSPE and MAPE: random walk relative to AR
Forecasts of h period ahead averages

		A: Last forecast 2023:4						B: Last forecast 2019:4					
(1)		h=1	h=4	h=8	h=12	h=20	h=40	h=1	h=4	h=8	h=12	h=20	h=40
(2)	b^{RW}	0.0	0.0	-0.3	-0.1	0.1	-0.4	0.0	0.1	-0.2	-0.1	-0.5	-0.8
(3)	$\lvert b^{RW}/b^{AR}\rvert$												
(3a)	Rolling	0.0	0.0	2.2	0.4	0.4	0.4	1.7	1.4	0.5	0.2	0.5	0.4
(3b)	Recursive	0.0	0.0	0.4	0.1	0.1	0.2	0.0	0.1	0.1	0.1	0.3	0.3
(4)	$RMSPE^{RW}$ / $RMSPE^{AR}$												
(4a)	Rolling	1.2	1.5	1.8	2.0	2.3	3.0	1.2	1.4	1.6	1.8	2.4	3.1
(4b)	Recursive	1.3	1.6	1.9	2.2	2.5	3.0	1.3	1.5	1.6	1.9	2.5	2.9
(5)	$MAPE^{RW}$ / $MAPE^{AR}$												
(5a)	Rolling	1.3	1.5	1.8	2.0	2.3	2.7	1.3	1.5	1.7	2.0	2.6	2.9
(5b)	Recursive	1.3	1.6	1.9	2.4	2.8	2.9	1.3	1.6	1.8	2.2	2.9	2.6

See notes to Table 4.2.

Table A4.2-(4) Growth in housing starts

Bias, RMSPE and MAPE: random walk relative to AR
Forecasts of h period ahead averages

		A: Last forecast 2023:4						B: Last forecast 2019:4					
(1)		h=1	h=4	h=8	h=12	h=20	h=40	h=1	h=4	h=8	h=12	h=20	h=40
(2)	b^{RW}	0.2	-0.1	-0.6	-0.4	0.5	0.4	0.3	0.3	0.0	-0.2	0.2	2.5
(3)	$\|b^{RW}/b^{AR}\|$												
(3a)	Rolling	1.4	0.1	0.4	0.2	0.3	1.6	0.3	0.2	0.0	0.2	0.4	0.7
(3b)	Recursive	0.2	0.1	0.5	0.3	0.5	4.7	0.3	0.4	0.0	0.4	3.5	0.8
(4)	$RMSPE^{RW}$ / $RMSPE^{AR}$												
(4a)	Rolling	1.2	1.6	1.7	1.8	1.8	2.5	1.1	1.3	1.4	1.5	1.8	2.8
(4b)	Recursive	1.3	1.8	2.0	2.2	2.3	3.5	1.2	1.5	1.6	1.8	2.2	3.9
(5)	$MAPE^{RW}$ / $MAPE^{AR}$												
(5a)	Rolling	1.2	1.5	1.6	1.7	1.7	2.2	1.1	1.4	1.4	1.5	1.7	2.3
(5b)	Recursive	1.2	1.7	1.8	1.9	2.2	2.9	1.2	1.5	1.6	1.7	2.1	3.1

See notes to Table 4.2.

Table A4.2-(5) Change in interest rate on 10 year treasuries

Bias, RMSPE and MAPE: random walk relative to AR
Forecasts of h period ahead averages

		A: Last forecast 2023:4						B: Last forecast 2019:4					
(1)		h=1	h=4	h=8	h=12	h=20	h=40	h=1	h=4	h=8	h=12	h=20	h=40
(2)	b^{RW}	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
(3)	$\|b^{RW}/b^{AR}\|$												
(3a)	Rolling	0.2	0.0	0.6	0.8	0.2	0.4	0.3	0.6	0.1	0.3	0.3	0.0
(3b)	Recursive	0.1	0.0	0.3	0.4	0.1	0.2	0.1	0.3	0.0	0.1	0.1	0.0
(4)	$RMSPE^{RW}$ / $RMSPE^{AR}$												
(4a)	Rolling	1.3	1.9	2.7	3.7	5.6	7.3	1.3	2.2	3.5	4.7	6.0	9.7
(4b)	Recursive	1.3	1.9	2.5	3.4	5.1	6.3	1.3	2.1	3.2	4.3	5.7	7.6
(5)	$MAPE^{RW}$ / $MAPE^{AR}$												
(5a)	Rolling	1.3	1.9	2.7	4.0	5.6	8.1	1.3	2.2	3.2	4.7	6.1	10.6
(5b)	Recursive	1.3	1.9	2.5	3.6	5.0	6.1	1.3	2.1	3.0	4.1	5.6	7.3

See notes to Table 4.2.

Table A4.2-(6) Inventory growth

Bias, RMSPE and MAPE: random walk relative to AR
Forecasts of h period ahead averages

		A: Last forecast 2023:4						B: Last forecast 2019:4					
(1)		h=1	h=4	h=8	h=12	h=20	h=40	h=1	h=4	h=8	h=12	h=20	h=40
(2)	b^{RW}	0.0	0.0	0.1	-0.1	-0.3	-0.1	0.0	0.1	0.0	-0.1	-0.1	0.2
(3)	$\|b^{RW}/b^{AR}\|$												
(3a)	Rolling	0.0	0.1	0.7	1.6	2.1	0.5	0.2	0.3	0.3	1.2	1.2	0.2
(3b)	Recursive	0.0	0.1	0.4	0.1	0.4	0.2	0.1	0.4	0.2	0.2	0.2	0.2
(4)	RMSPE^{RW} / RMSPE^{AR}												
(4a)	Rolling	1.1	1.2	1.4	1.5	1.7	2.2	1.0	1.1	1.2	1.4	1.7	2.2
(4b)	Recursive	1.1	1.3	1.6	1.8	2.2	2.9	1.1	1.2	1.4	1.6	2.2	2.8
(5)	MAPE^{RW} / MAPE^{AR}												
(5a)	Rolling	1.0	1.3	1.3	1.4	1.6	2.0	1.0	1.2	1.2	1.3	1.6	2.1
(5b)	Recursive	1.1	1.4	1.5	1.7	2.0	2.6	1.1	1.3	1.4	1.6	1.9	2.3

See notes to Table 4.2.

Table A4.2-(7) Industrial production growth

Bias, RMSPE and MAPE: random walk relative to AR
Forecasts of h period ahead averages

		A: Last forecast 2023:4						B: Last forecast 2019:4					
(1)		h=1	h=4	h=8	h=12	h=20	h=40	h=1	h=4	h=8	h=12	h=20	h=40
(2)	b^{RW}	0.0	-0.1	-0.2	-0.2	-0.8	-1.0	-0.1	-0.2	-0.2	-0.3	-0.9	-0.5
(3)	$\|b^{RW}/b^{AR}\|$												
(3a)	Rolling	0.0	0.3	1.6	1.1	1.0	1.0	0.7	1.8	1.0	0.7	1.1	0.4
(3b)	Recursive	0.2	0.2	0.2	0.2	0.5	0.6	0.2	0.3	0.2	0.2	0.6	0.3
(4)	RMSPE^{RW} / RMSPE^{AR}												
(4a)	Rolling	0.9	1.2	1.5	1.7	1.3	1.2	1.0	1.1	1.2	1.2	1.3	1.1
(4b)	Recursive	1.0	1.4	1.9	2.4	2.2	2.6	1.0	1.2	1.5	1.8	2.3	2.7
(5)	MAPE^{RW} / MAPE^{AR}												
(5a)	Rolling	0.9	1.2	1.5	1.7	1.8	1.7	0.9	1.2	1.3	1.4	1.7	1.5
(5b)	Recursive	1.0	1.3	1.7	1.9	1.9	2.0	1.0	1.3	1.6	1.7	2.0	2.1

See notes to Table 4.2.

Table A4.2-(8) Growth in M3

Bias, RMSPE and MAPE: random walk relative to AR
Forecasts of h period ahead averages

		A: Last forecast 2023:4						B: Last forecast 2019:4					
(1)		h=1	h=4	h=8	h=12	h=20	h=40	h=1	h=4	h=8	h=12	h=20	h=40
(2)	b^{RW}	0.0	-0.2	-0.2	0.1	0.6	0.2	0.0	0.0	0.0	0.0	0.0	0.0
(3)	$\lvert b^{RW}/b^{AR}\rvert$												
(3a)	Rolling	0.4	1.9	0.4	0.1	0.6	0.2	0.3	0.2	0.1	0.1	0.0	0.0
(3b)	Recursive	0.3	0.4	0.2	0.1	0.5	0.2	0.2	0.1	0.0	0.1	0.0	0.0
(4)	RMSPE^{RW} / RMSPE^{AR}												
(4a)	Rolling	1.0	1.1	1.5	1.7	1.7	2.0	1.1	1.4	1.7	1.8	2.0	1.9
(4b)	Recursive	1.1	1.3	1.7	2.1	1.7	2.0	1.1	1.4	1.8	2.0	2.2	2.4
(5)	MAPE^{RW} / MAPE^{AR}												
(5a)	Rolling	1.0	1.1	1.5	1.6	1.7	1.8	1.1	1.3	1.6	1.8	1.9	1.7
(5b)	Recursive	1.0	1.2	1.5	1.7	1.6	1.7	1.1	1.3	1.6	1.8	1.9	2.0

See notes to Table 4.2.

Table A4.2-(9) Labor productivity growth

Bias, RMSPE and MAPE: random walk relative to AR
Forecasts of h period ahead averages

		A: Last forecast 2023:4						B: Last forecast 2019:4					
(1)		h=1	h=4	h=8	h=12	h=20	h=40	h=1	h=4	h=8	h=12	h=20	h=40
(2)	b^{RW}	0.0	0.0	-0.2	-0.2	0.0	-0.5	0.0	0.0	-0.1	-0.1	-0.2	-0.8
(3)	$\lvert b^{RW}/b^{AR}\rvert$												
(3a)	Rolling	1.8	0.5	2.5	4.1	0.4	1.2	17.0	0.0	0.7	0.7	0.9	3.0
(3b)	Recursive	0.7	0.5	4.0	108.7	1.7	2.6	0.1	0.1	2.4	8.4	8.0	18.4
(4)	RMSPE^{RW} / RMSPE^{AR}												
(4a)	Rolling	1.3	1.8	2.2	2.5	2.2	2.5	1.3	1.8	2.0	2.1	2.2	2.6
(4b)	Recursive	1.3	1.9	2.3	2.7	2.6	3.3	1.3	1.9	2.1	2.3	2.5	3.4
(5)	MAPE^{RW} / MAPE^{AR}												
(5a)	Rolling	1.3	1.8	1.8	2.1	2.0	2.2	1.3	1.8	1.9	1.9	2.0	2.3
(5b)	Recursive	1.4	1.8	1.9	2.2	2.5	2.8	1.4	1.9	1.9	2.0	2.4	3.0

See notes to Table 4.2.

Table A4.2-(10) Change in unemployment rate

Bias, RMSPE and MAPE: random walk relative to AR
Forecasts of *h* period ahead averages

		A: Last forecast 2023:4						B: Last forecast 2019:4					
(1)		h=1	h=4	h=8	h=12	h=20	h=40	h=1	h=4	h=8	h=12	h=20	h=40
(2)	b^{RW}	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-0.1
(3)	$\lvert b^{RW}/b^{AR}\rvert$												
(3a)	Rolling	0.0	0.0	0.0	0.1	1.2	0.4	0.1	0.0	0.0	0.1	0.1	3.2
(3b)	Recursive	0.0	0.1	0.1	12.1	2.1	4.2	0.4	0.1	0.0	0.3	3.1	1.8
(4)	RMSPE^{RW} / RMSPE^{AR}												
(4a)	Rolling	0.8	0.7	0.7	0.7	1.2	1.3	1.0	1.0	1.0	1.1	1.2	1.3
(4b)	Recursive	1.0	1.3	1.5	1.7	1.9	2.8	1.1	1.1	1.3	1.5	1.9	2.9
(5)	MAPE^{RW} / MAPE^{AR}												
(5a)	Rolling	0.9	0.9	1.0	1.0	1.3	1.6	1.1	1.1	1.1	1.1	1.3	1.7
(5b)	Recursive	1.0	1.2	1.3	1.5	1.6	2.4	1.1	1.1	1.2	1.3	1.6	2.7

See notes to Table 4.2.

Table A4.3-BIC

Bias and RMSPE, random walk relative to AR forecasts: summary statistics
Lag length chosen by BIC

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	median	bias ratio		RMSPE ratio		MAPE ratio	
	$\|b^{RW}\|$	median	%<1	median	%<1	median	%<1
(1) Per capita GDP growth	0.1	0.3	92%	1.8	0%	1.7	0%
(2) CPI inflation	0.1	0.2	92%	1.9	0%	1.6	0%
(3) Growth in HH assets	0.1	0.2	88%	1.7	0%	1.8	0%
(4) Growth in housing starts	0.3	0.4	79%	1.8	0%	1.7	0%
(5) Change in 10Y rate	0.0	0.2	100%	3.4	0%	3.3	0%
(6) Inventory growth	0.1	0.3	79%	1.6	0%	1.5	0%
(7) IP growth	0.2	0.5	79%	1.4	4%	1.5	4%
(8) Growth in M3	0.0	0.1	92%	1.7	4%	1.6	4%
(9) Labor prod growth	0.1	1.2	46%	2.2	0%	1.9	0%
(10) Chg in unemployment rate	0.0	0.3	67%	1.2	4%	1.2	4%
(11) All 10 series, all horizons	0.0	0.3	81%	1.8	1%	1.7	1%
(12) All 10 series, $h \leq 12$	0.0	0.2	84%	1.5	2%	1.5	2%
(13) All 10 series, $h>12$	0.2	0.4	75%	2.4	0%	2.1	0%

See notes to Table 4.3. This differs from Table 4.3 in that lag length is not set to 2 but instead is chosen by BIC, with a maximum lag length of 4.