Economics Bulletin

Volume 43, Issue 1

Are unit root tests useful for univariate time series forecasts with different orders of integration? A Monte Carlo study

Adam J. Check
U.S. Bank

Ming Chien Lo
Metropolitan State University

Kwok Ping Tsang
Virginia Tech

Abstract

In this paper, we consider univariate forecasts made when using stationary, near unit root, and unit root data. Like Diebold and Kilian (2000), we conduct a Monte Carlo experiment investigating the usefulness of unit root tests prior to forming univariate forecasts. In our experiment, we consider more than one unit root test and also vary the order of integration in the time series. We find that unit root tests are indeed useful for forecasting, especially when the series has a large number of in-sample observations. However, the choice of unit test matters. Using root mean square error as a criterion for forecast performance, we find that the Philips-Perron test has an edge over the augmented Dickey-Fuller test and the Kwiatkowski–Phillips–Schmidt–Shin test. We recommend practitioners to be mindful of the choice of test, as the KPSS test is the default used in the forecast package in R, following Hyndman and Khandakar (2008), but the Philips-Perron test is available as an option in that package.

The opinions expressed are those of the author and do not necessarily reflect views of the U.S. Bank National Association or any of its affiliated entities. We thank the the referees for the comments on the paper.
Citation: Adam J. Check and Ming Chien Lo and Kwok Ping Tsang, (2023) "Are unit root tests useful for univariate time series forecasts with different orders of integration? A Monte Carlo study", *Economics Bulletin*, Volume 43, Issue 1, pages 203-244
Contact: Adam J. Check - adam.check@usbank.com, Ming Chien Lo - ming.lo@metrostate.edu, Kwok Ping Tsang - byront@vt.edu.
Submitted: February 13, 2022. **Published:** March 30, 2023.

Section 1: Introduction

Unit root tests have been used for decades to determine if a given time-series is stationary.[1] There are several reasons why econometricians want to know if a time-series is stationary. First, stationarity implies mean reversion, a feature that is often implied by economic theory. For example, purchasing power parity implies that the real exchange rate is stationary. Second, distinguishing between stationary and non-stationary data matters for applied time-series analysis. For example, using non-stationary data in linear regression may result in spurious regression (Ventosa-Santaulària (2009)).

Distinguishing between stationary and non-stationary data also matters for forecasters. Forecasting under the assumption of stationarity will lead to different forecasts than those made under the assumption of non-stationarity. Therefore, researchers like Diebold and Kilian (2000) have suggested the use of unit root tests prior to forecasting. They propose a pre-test method in which, prior to forecasting, a series is tested for a unit root. In a Monte-Carlo study, they find that the augmented Dickey-Fuller (ADF) test reliably informs forecasters when the data should be first-differenced. This process was later extended by Hyndman and Khandakar (2008, hereafter HK). HK first use a unit root test— Kwiatkowski–Phillips–Schmidt–Shin (KPSS) by default, or ADF or Philips-Perron (PP) as alternatives—before applying an ARIMA model.

Our paper extends the idea of Diebold and Kilian (2000) to compare *which* unit root test results in better forecasting performance.[2] We conduct a Monte Carlo exercise assuming linear data generating processes (DGP) ranging from $I(0)$ to $I(2)$, and use these simulated series to compare the HK approach against a model averaging approach (AVG) that assumes an I(0), I(1), or I(2) process with equal probability.[3] Our choice to use HK is partially a matter of convenience, as its creators have developed and maintained the popular and widely used *forecast* package in *R*, which includes the HK algorithm. Therefore, our results also provide guidance for the effective use of the *forecast* package.[4]

We find that each of the three unit root tests improve forecast performance over the model average, but not equally so. Additionally, we find that using a unit root test is more valuable when the sample size is large. We also find that the PP test more reliably improves forecast performance compared to the default KPSS used in the *R* package. Therefore, practitioners should be mindful of the choice of unit root test when following HK.

Section 2: The Models

The HK approach is standard, and details can be found in Hyndman and Athanasopoulos (2018). To approximate the experience of most practitioners, we keep the default HK settings from the *R*

---

[1] See Wolters and Hassler (2006) for an overview of the history.

[2] Diebold and Kilian (2000) is more limited; they explicitly try to capture the features of U.S. real GDP, including the presence of a linear time trend.

[3] Therefore, unit root tests are absent in AVG.

[4] To get a sense of the popularity, albeit somewhat unscientifically, we conducted several Google searches on June 15, 2021. Keywords “forecast package” generates 85.6 million results and “forecast package in R” 39.3 million. The specific function that executes the HK algorithm is auto.arima. Keywords “auto.arima” generates 2.7 million.

package, with a few exceptions.[5] The HK algorithm first uses a unit root test to determine the order of integration, and then applies an ARIMA model with the appropriate level of differencing. Based on the results of the unit root test, the HK algorithm fits an ARMA(p,q) model to data that has been differenced 0, 1, or 2 times. Given the results from the unit root test, HK commits fully—that is, a 100% weight—on one of these three levels of differencing. Next, an information criterion is used to select the order of autoregression (p) and moving average (q).

The HK approach in *R* includes three built-in unit root tests: KPSS, ADF and PP. The ADF and PP tests are built around the same basic premise. For simplicity, consider the AR(1) model[6] without trend:

$$y_t = a + \phi y_{t-1} + \varepsilon_t$$

Both the PP and ADF use a t-test to test a null hypothesis of a unit root, $\phi = 1$, with a one-sided stationary alternative, $\phi < 1$.[7] If the residuals from this simple AR(1) model are white noise, then the PP and ADF tests are identical. However, in practice the residuals will never be perfect white noise, so the results of the PP and ADF tests can differ.

In the presence of remaining residual autocorrelation, the PP and ADF tests account for it differently. The PP test never alters the test equation above. Instead, it always uses robust standard errors to adjust the standard error of the point-estimate of ϕ, which changes the value of the t-statistic. In contrast, if the level of remaining residual autocorrelation is statistically significant, the ADF test includes additional lags in the test equation (e.g., it may use an AR(4) if appropriate). After controlling for these extra lags, both the point-estimate and the standard error of ϕ can differ from their values in the simple AR(1) model, changing the value of the t-statistic.

Due to the low power of the ADF test (Hassler and Wolters, 1994; Paparoditis and Politis, 2018), and due to the poor performance of the PP test in small samples (Davidson and McKinnon, 2004), the KPSS test was developed as an alternative. In the KPSS test, the null hypothesis is flipped – the null is that the data is stationary, while the alternative is that it contains a unit root.

Differencing after pre-testing for a unit root is a convenient but only sometimes appropriate method of rendering a non-stationary series stationary.[8] Size distortion, lack of power, presence of outliers or structural breaks, etc., can all affect the results of unit root tests. For example, even on a single simulated AR(1) time-series, we find that the unit root tests can behave erratically as the sample length increases (i.e., as time progresses). One possible alternative to pre-testing, which we explore, is to not rely on any unit root test. Our AVG approach consists of first imposing each level of differencing (0, 1, or 2) and then, conditional on the difference, using the second step of the HK algorithm to determine the order of

---

[5] When forecasters approach real-world data, the default settings are likely to be deployed. Later research can further investigate whether results are robust against changes to the default settings.

[6] In practice, the ADF test is typically applied to an AR(1) model for the differenced data: $\Delta y_t = \mu + \rho \Delta y_{t-1} + \varepsilon_t$ where $\rho = \phi - 1$. This alternative specification is functionally identical and leads to the same p-value.

[7] The critical values under both tests are non-standard.

[8] Differencing the data may result in introducing unnecessary components or misspecification to a model that is trend-stationary. For example, if $y_t = \alpha t + \beta y_{t-1} + \varepsilon_t$, $\Delta y_t = \alpha + \beta \Delta y_{t-1} + \varepsilon_t - \varepsilon_{t-1}$ with a moving average error term. However, in an exploratory exercise, we found that differencing trend-stationary data as a detrending method did not lead to underperformance of the HK algorithm against the AVG even when a trend component is not explicitly added to the unit root tests.

autoregression (p) and moving average (q). Finally, we assign 1/3 weight to each of the forecasts from these three ARIMA models.

Section 3: The Monte Carlo and Out-of-Sample Forecast Design

Subsection 3.1: Data generating processes

We use an AR(1) model as the baseline, since it is a relatively common and simple model that allows for highly persistent I(0) series. Therefore, our first DGP is an AR(1) process that is highly persistent, becoming a unit root process when $b = 1$

$$y_t = a + by_{t-1} + \varepsilon_t. \tag{1}$$

For simplicity, we restrict $a = 0$. We allow $b \in \{0.90, 0.95, 0.975, 0.99, 1.00\}$. Focusing on $b \geq 0.9$ covers the area in which a unit root test may have either size distortion or lack of power. Many macroeconomic series that exhibit slow mean reversion behavior, such as the unemployment rate or the real exchange rate, can be represented by this DGP. When $b = 0.975$ or 0.99, we have a so-called "near unit root" process.

The second DGP assumes a similar structure, but for the first difference,

$$\Delta y_t = a + (b-1)\Delta y_{t-1} + \eta_t \tag{2}$$

which can be written as

$$y_t = a + by_{t-1} - (b-1)y_{t-2} + \eta_t. \tag{3}$$

We restrict $a = 0$ and allow $b \in \{1.50, 2.00\}$. When $b = 2$, the series is $I(2)$. When $b = 1.5$, the series is fractionally integrated between $I(0)$ and $I(2)$. Note that if $b = 1$, both (1) and (3) result in a random walk model without drift.

DGP's (1) and (2) with the set of possible values of b cover a plausible range for many financial and economic data sets. While economic data has traditionally been assumed to be I(0) or I(1), Caporale, Gil-Alana and Plastun (2019) and Hartl, Tschernig and Weber (2021) both find that series with an order of integration above one are more common than previously thought.

To simulate data from either DGP, for each number of observations N, we simulate $1.1 \times N$ observations y_t according to (1) and (2) respectively and trim off the first 10%. We fix the conditional standard deviation of the processes, by setting the standard deviations of the shocks ε_t and η_t to 1. For both DGPs, we consider $N \in \{50, 200, 500\}$. For each choice of N, we perform an expanding window forecasting exercise and consider various k-period-ahead point-forecasts, where $k \in \{1,3,6,12\}$.[9]

We conduct an expanding window pseudo-out-of-sample forecasting exercise like that of Meese and Rogoff (1983). For any simulated series, we start with an estimation period that starts at the first observation and includes the first 60% of the sample. Then, we form the k-period-ahead point forecast from each model. Next, we add one additional observation and repeat this process. We expand the number of observations for estimation by one at each iteration until all

[9] Because the number of observations is small when $N = 50$, we do not consider $k = 12$ for that case.

data points are exhausted. The set of k-horizon point forecasts are then compared against the actual data using root mean squared errors (RMSE).

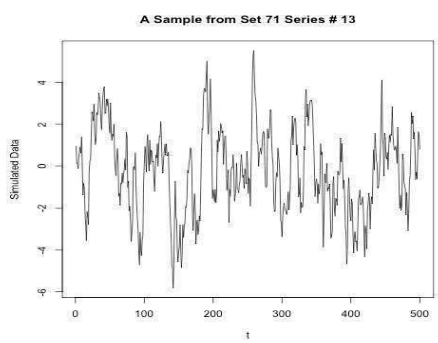
Subsection 3.2: The Monte Carlo design and ARIMA settings

Given the varying values of the $AR(1)$ parameter, b; the sample length, N; and forecast horizon, k, we have 77 parameterizations of the DGPs.[10] For each, 1,000 series are simulated. To get a sense of our simulated data, Figure 1 shows the series generated from DGP (1) and (2) for each possible value of b when $N = 500$. We use the following settings when estimating the ARIMA models:

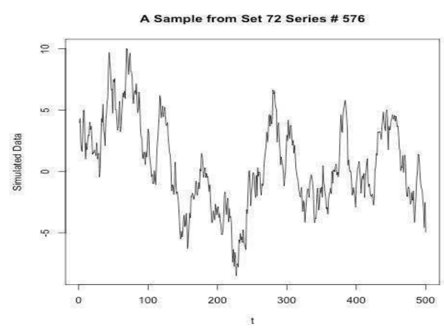
- We use the following default options:
    - The maximum possible order of integration is 2.
    - The level of significance for the unit root tests in auto.arima() is five percent.[11]
    - The information criterion used when selecting the ARMA(p,q) order is the corrected Akaike Information Criterion (AICc).
- Since we know we are using a (possibly integrated) AR(1) model, we set:
    - The combined maximum order of p and q to 3 for auto.arima(); i.e. max.order=3.
    - We do not consider seasonal ARIMA models; i.e. we set seasonal=FALSE.

Figure 1: Sample Simulated Series from AR(1)

b = 0.900

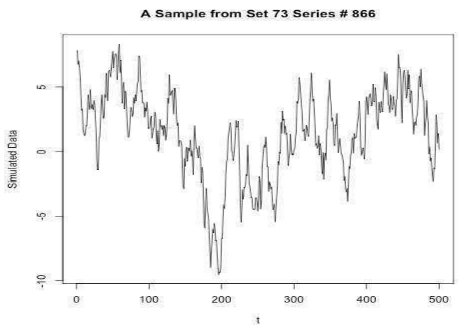


b = 0.950

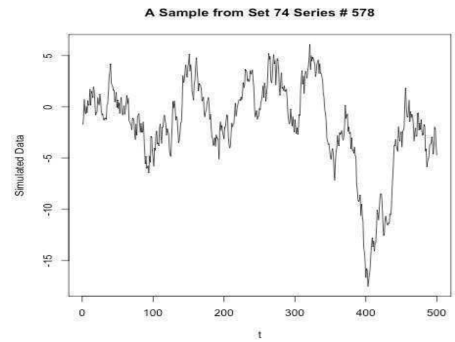


[10] Appendix Table A1 shows the values of the parameters.
[11] Also assumed in Hyndman and Athanasopoulos (2018). In recent years, such a default has been questioned (e.g. Ziliak and McCloskey, 2007; Wasserstein and Lazar, 2016; Amrhein, Greenland and McShane, 2019).

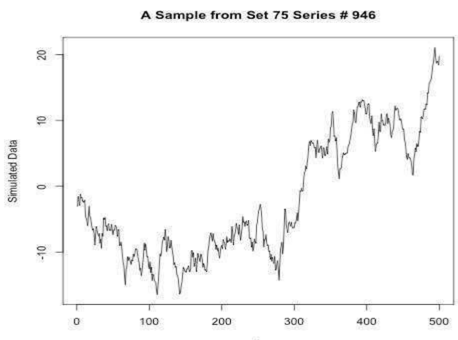
b = 0.975



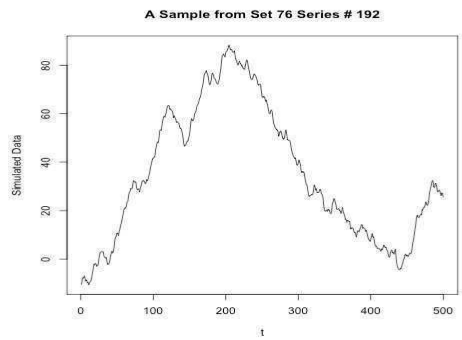
b = 0.990



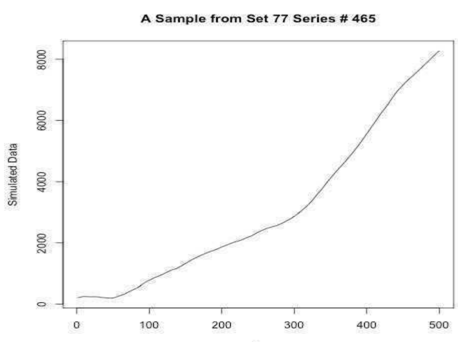
b = 1.000



b = 1.500



b = 2.000



Section 4: Results for Baseline DGP of AR(1)

We summarize the results in two ways. First, we generate aggregate statistics from each of the 77 sets of simulations. After simulating data and forming forecasts for each of the 77 sets, we compare forecast accuracy across four models: (1) HK with the KPSS test (default), (2) HK with the ADF test, (3) HK with the PP test, (4) the model average ("AVG") described in section 2. To measure forecast accuracy, we compute the Monte Carlo mean RMSE from the 1,000 simulations for each model:

$$\overline{RMSE}_m = \frac{\sum_{i=1}^{1000} RMSE_{m,i}}{1000} \tag{4}$$

and rank them to determine the winner in each set, where i is the i-th simulation in the set and m denotes the model. Models resulting in a lower mean RMSE have a better forecast performance.

To get a sense of relative difference in RMSEs, we compute a few additional measures. Since we fix the conditional SD of the error terms in the simulations, the unconditional SD, which also depends on the value of b, differs across different generated time-series. To make it easier to compare *relative* performance across different simulations and forecast horizons, we first compute the relative mean RMSE for each simulation which is given by:

$$Relative\ RMSE = \frac{\overline{RMSE}_m}{\overline{RMSE}_{Winner}} \tag{5}$$

where *Winner* denotes the model with the lowest absolute RMSE (given in equation (4)). Next, for a given experiment, we compute the average of the relative RMSE across each of the 1,000 simulations:

$$Mean\ Ratio = \sum_{i=1}^{1000} \frac{RMSE_{m,i}}{RMSE_{Winner,i}}. \tag{6}$$

Further, we can compute a p-value based on $RMSE_{m,i}/RMSE_{Winner,i}$,

$$p - value = \sum_{i=1}^{1000} I\left(\frac{RMSE_{m,i}}{RMSE_{Winner,i}} \leq 1\right)/1000. \tag{7}$$

The mean ratio and its corresponding p-value may result in different rankings that the mean RMSE given in (4). For example, a model may have a low RMSE in most simulations, but a very high RMSE in a handful. This could result in the model having the best mean ratio or p-value, but a higher mean RMSE compared to alternative models.

The ranking based on (4) and other statistics are reported in detail in Appendix Table A2A. From the table, we observe that:

(i) Using the PP test to determine the order of integration results in the *lowest mean RMSE* in 33 out of the 77 sets, followed by KPSS with 30 and ADF with 12. AVG only has the best mean RMSE in 2 of the test sets. In fact, each of the three HK models considered (i.e., HK with any unit root test) beats AVG in 75 of the 77 sets. Using this criterion, we can conclude that the unit root tests result in better forecast performance than using a simple model average. These results also suggest that using the PP test generates slightly better forecast performance than does using the default KPSS test. These results are summarized in Table 1(a).

(ii) The p-values among the top three models in any given set is generally high. When $N = 500$, the RMSEs from the last place model is consistently larger than that from the winner. Because AVG results in the worst forecast performance in 64 out of 77 test sets, we can conclude – perhaps unsurprisingly – that the unit root tests are especially useful in large samples.

(iii) We observe larger gaps in relative RMSE (5) and mean ratio (6) between models when the order of integration increases. Therefore, it seems that as the order of integration increases, it becomes more important to identify the presence of integration accurately.

As an alternative form of summary, we analyze the 77 outcomes using a multinomial logit model where the dependent variable is the winning model (KPSS, ADF, PP, and model averaging) and the independent variables are the sample size, horizon, and the integration order. Here, the winning model is not determined by lowest mean RMSE, which is an aggregate statistic from the 1,000 simulations. Instead, the winning model is the one that garners the largest number of wins (i.e. lowest RMSE) out of the 1,000 simulations in each of the 77 test sets. Winning counts calculated this way are reported in Table 1(b). Defining the best model this way results in model rankings that are similar to those obtained when using mean RMSE, but results in an even larger number of wins by PP (37) over KPSS (23).

Without loss of generality, we use KPSS as the baseline model for the logit analysis. The results are reported in Table 2(a), and they are consistent with our earlier findings. First, when we control for the sample size, the forecast horizon and the order of integration, PP is frequently superior to KPSS. In the second row, we see that the estimate on the order of integration is relatively large, at -1.051. While the p-value is also high (above 10%), the large point estimate suggests that the advantage from using the PP test (rather than the KPSS test) may decline somewhat when the order of integration is high.

Second, the ADF test also has a large and significant constant term. However, with an estimate of -3.766, the integration term offsets the positive constant (3.987) when the order of integration is one, and more than offsets it when the order of integration is 1.5 or 2.

Finally, our model average of I(0), I(1), and I(2) models, i.e. AVG, is clearly inferior to the default KPSS test. Even though the constant term is estimated at a positive value of 3.715, it is offset by the negative estimate for the parameter for N—even when $N = 50$ which is the smallest in our Monte Carlo, the product of -0.111 and 50 is -5.55, more than offsetting the positive constant. When N is 200 or 500, the negative impact would become even larger. This result once again implies that using unit root tests to determine the order of integration is superior to using a simple model average. This is increasingly true as the sample size increases.

Table 1: Number of Sets Won by Models

(a) Baseline AR(1)

	KPSS	ADF	PP	AVG
(1) Lowest mean RMSE among sets	30	12	33	2
(2) Most wins in each set	23	14	37	3

These tables report the number of sets won by a model out of the 77 sets. We count the number of wins in two ways. Method (1) is based on mean RMSE in (4). A model wins a set when it has the lowest mean RMSE. Detailed results under method (1) can be found in Appendix Table A2A. For example, for set #1 (N=50, k=1, b=0.9), PP has the lowest mean RMSE. Method (2) is simply based on the winning counts out of the 1,000 simulations. A model wins a set when it has the largest number of wins (lowest RMSE) out of the 1,000 simulations. For example, for set #1 (N=50, k=1, b=0.9), AVG has the largest number of wins in terms of RMSE. It is possible to have different winning model in the same set when a different method is used.

(b) ARMA Specification

	KPSS	ADF	PP	AVG
(1) Lowest mean RMSE among sets	24	16	27	10
(2) Most wins in each set	25	15	30	7

(c) Structural Breaks

	KPSS	ADF	PP	AVG
(1) Lowest mean RMSE among sets	1	2	5	3
(2) Most wins in each set	0	0	0	11

Table 2: Summary Results from a Multinomial LOGIT Regression

(a) Baseline AR(1)

	Constant	Sample Size N	Horizon k	Coefficient b
ADF	3.987*** (1.200)	-0.002 (0.019)	0.047 (0.095)	-3.766*** (0.656)
PP	2.266** (1.081)	-0.002 (0.002)	0.030 (0.071)	-1.051 (0.676)
AVG	3.715** (1.638)	-0.111*** (0.007)	0.024 (0.336)	1.076 (0.971)
Pseudo R^2	0.118			

Robust standard errors reported in parentheses. *** p-value below 1% with null hypothesis that the parameter is equal to zero.

(b) ARMA Specification

	Constant	Sample Size N	Horizon k	Coefficient b
ADF	2.876** (1.145)	0.002 (0.019)	0.029 (0.092)	-3.906*** (0.681)
PP	0.875 (1.038)	0.001 (0.002)	0.003 (0.070)	-0.705 (0.687)
AVG	14.424*** (4.887)	-0.125*** (0.007)	-2.021** (1.027)	-2.721 (2.826)
Pseudo R^2	0.241			

Robust standard errors reported in parentheses. *** p-value below 1% with null hypothesis that the parameter is equal to zero.

Section 5: Deviations from the Baseline

Subsection 5.1: $ARMA(p, q)$

We extend our DGP by allowing higher order of AR and MA terms such that the ARMA(p,q) where p can vary from 1 to 3 and q from 0 to 3. Our program randomly chooses a value for p and q in each simulation. In each simulation, we ensure that the values of the AR coefficients are set appropriately so that they imply the order of integration specified for each set. For the auto.arima procedure used when assessing model performance, we set max.p=3 and max.q=3.

Comparing Table 1(a) and 1(b), unit root testing is still superior to simple model averaging, but less so than in the baseline case. When examining the details in Appendix Table A2B, we find that AVG improves its performance when both the sample size (N) and the forecast horizon (k) are small. This observation is supported by our subsequent analysis. In Table 2(b), we replicate our Table 2(a) by conducting a logit regression. The results for AR(1) and ARMA are mostly the same except that when the forecast horizon (k) is small, AVG's performance improves. Among the unit root tests, PP still dominates over KPSS and ADF.

Subsection 5.2: Structural breaks

Real-world data may be more complicated than the linear ARMA processes studied above. Here, we attempt to offer some insights on what may happen in the presence of structural breaks. We have chosen to allow for structural breaks in each series with some randomization:

- We vary the number of breaks depending on series length. If N=50 we allow one break; if N=200, we allow either one or two breaks; if N=500, we allow one, two or three breaks.
- There are 11 sets of Monte Carlo designs with varying N (50, 200, 500) and horizon (1, 3, 6, 12)—for N=50, we do not include cases of k=12.
- Break points are chosen randomly. If there is more than one break, there must be at least 15 data points between break points.
- In each stationary regime, the sum of the AR parameters ranges from 0.9000 (less persistent) to 0.9999 (highly persistent), and we allow for one random walk regime with 80% probability.

Our AVG forecast is composed of equally weighted forecasts from I(0) and I(1) models only, since we only consider breaks in equation (1). We also do not conduct any logit analysis since the number of sets has been reduced to 11 and therefore the degree of freedom is poor.

Table 1(c) reports the winning counts and Appendix Table A3 reports the details. In this case, the results under the two forecast scores do not agree. Using lowest mean RMSE as a criterion, the HK algorithm tends to dominate. However, when using winning counts within a set, AVG wins in all 11 sets. Upon closer examination, we find that these results are driven by a few extremely large outlier RMSEs in AVG, giving this model a larger mean RMSE even though its RMSE is smallest in a majority of simulations. Table 3 shows that the maximum RMSE of AVG in each set is 1.87 to 7.71 times larger than the maximum RMSE of the winner model.

Table 3: Maximum RMSE for Each Model and AVG's Outlier RMSE

N	Horizon	KPSS	ADF	PP	AVG	Relative Difference	# of Outliers
50	1	1.725485	**1.722987**	2.759139	3.222901	1.87	12
50	3	5.004424	4.808418	**4.808418**	28.317883	5.89	6
50	6	9.648729	8.374953	8.374953	**14.935339**		
200	1	1.933442	1.463256	**1.470875**	6.663352	4.53	22
200	3	**3.785853**	3.785853	3.785853	13.633259	3.60	14
200	6	9.832408	9.832408	9.832408	**19.789169**		
200	12	18.0332	18.0332	18.0332	**20.18657**		
500	1	1.374264	1.374264	**1.374264**	10.592159	7.71	39
500	3	6.943965	**5.585818**	5.594647	12.241989	2.19	13
500	6	7.350329	8.143947	**7.90799**	21.423266	2.71	15
500	12	15.77235	15.77235	**15.77235**	87.94463	5.58	1

Column 2-4 reports the maximum RMSE out of the 1,000 simulations in each set.
Bold numbers indicate the winning model (i.e. Winner) when the lowest mean RMSE is used.
Relative difference is calculated from 1 + ($RMSE_{AVG}$ – $RMSE_{Winner}$)/$RMSE_{Winner}$.
The last column reports the number of outlier RMSEs in AVG that exceeds the maximum RMSE of the winner.

Section 6: Empirical Data

We have chosen three exchange rate series and four CPI inflation rates to investigate model performance on real-world economic series. The DGP of these real-world series is unknown, but these series may be reasonably approximated by a highly persistent AR process, an integrated process, or a persistent process with changes in regime (possibly moving into RW regimes), as studied in our Monte Carlo experiments. To investigate the performance of the HK algorithm and simple model averaging, we perform a pseudo-out-of-sample exercise like that performed in our Monte Carlo exercises.

The results are reported in Table 4. The successes of the HK algorithm are mixed. The HK algorithm outperforms the AVG clearly in three series (Canadian-U.S. dollar exchange rates, yen-dollar exchange rates and Canada inflation rates) and moderately in two (pound-dollar exchange rates and U.K. inflation rates). However, AVG dominates in the other two series (U.S. and Japan inflation rates). Given the relative performance of simple model averaging in the Monte Carlo exercises, this suggests that there may be structural breaks in these latter two series, but more research should be conducted on this topic.

Table 4: Results from Empirical Data

Horizon	KPSS	ADF	PP	AVG
		CAD-USD Exchange Rate		
1	**0.026349**	**0.026349**	**0.026349**	0.034443
3	**0.044612**	**0.044612**	**0.044612**	0.056078
6	**0.063801**	**0.063801**	**0.063801**	0.069279
12	0.086554	0.086554	0.086554	**0.084905**
		JPY-USD Exchange Rate		
1	**0.026099**	**0.026099**	**0.026099**	0.035481
3	**0.048439**	**0.048439**	**0.048439**	0.118566
6	**0.071906**	**0.071906**	**0.071906**	0.200592
12	**0.102003**	**0.102003**	**0.102003**	0.208055
		GBP-USD Exchange Rate		
1	**0.025968**	0.026100	0.026032	0.025998
3	**0.047362**	0.047910	0.047719	0.047436
6	0.072141	0.073685	0.072980	**0.072055**
12	0.096422	0.100591	0.099227	**0.096105**
		U.S. CPI Inflation Rate		
1	0.415851	0.415851	0.415851	**0.414161**
3	1.033206	1.033206	1.033206	**1.019044**
6	1.547637	1.547637	1.547637	**1.500432**
12	2.331053	2.331053	2.331053	**2.193044**
		Canada CPI Inflation Rate		
1	**0.448788**	**0.448788**	**0.448788**	0.902221
3	**0.811803**	**0.811803**	**0.811803**	1.107497
6	**1.129578**	**1.129578**	**1.129578**	1.343380
12	**1.609507**	**1.609507**	**1.609507**	1.741674
		Japan CPI Inflation Rate		
1	**0.326511**	**0.326511**	**0.326511**	0.332093
3	0.663813	0.663813	0.663813	**0.658711**
6	0.968618	0.968618	0.968618	**0.954919**
12	1.509350	1.509350	1.509350	**1.471567**
		U.K. CPI Inflation Rate		
1	0.270144	0.270144	0.270144	**0.270081**
3	**0.533370**	**0.533370**	**0.533370**	0.534177
6	**0.885464**	**0.885464**	**0.885464**	0.885666
12	1.487474	1.487474	1.487474	**1.485325**

RMSEs reported here. The smallest ones at a specific forecast horizon are in bold. If two or more unit root tests result in the same detection of unit root or its absence at every step, their RMSEs will be identical.

Section 7: Conclusion

We use the *forecast* package in R, specifically the auto.arima() function, to examine the forecasting performance of various approaches to unit-root testing in a Monte Carlo exercise in which the data has varying degrees of persistence and integration. We are particularly interested in the performance of the Hyndman-Khandakar algorithm, which uses unit root tests to determine the appropriate number of differences to ensure stationarity before selecting an ARMA model. We find that both the HK algorithm and the unit root tests improve forecasting performance relative to a simple model average across stationary, I(1), and I(2) models. Our results are qualitatively similar when the data are simulated from an AR(1) model or from more complicated ARMA(p,q) models.

Our results constitute an important contribution to this literature. Previously, Diebold and Kilian (2000) examined the performance of the ADF test for forecasting in a relatively narrow context. Our study includes other unit root tests as well as a model-averaging alternative that represents ignorance of the test results. We also generalize the Monte Carlo exercise by considering different orders of integration, as well as near unit root processes. Our results imply that the Philips-Perron test may generate better forecast than the KPSS or ADF tests when the DGP is linear and the errors are normally distributed.

For forecast practitioners, our results suggest that relying solely on the KPSS test may result in worse forecast performance for near-unit root, I(1), fractionally integrated, or I(2) processes, and that using the PP test for these types of series would lead to an increase in expected forecast accuracy. Of course, while the patterns of integration listed above capture many real-world time-series, many other real-world series feature nonlinearities or outliers in addition to (or in lieu of) these patterns of integration. The Monte Carlo exercise using data simulated from a structural break specification is not as conclusive. While the HK algorithm performs well when using mean RMSE, model averaging has a lower RMSE on a larger fraction of simulated series. Finally, when applying these methods to empirical data, the success of the HK algorithm is apparent but not overly dominating. We leave the study of the performance of unit root tests and the HK algorithm in these more complicated environments to future research.

References

Amrhein, V., S. Greenland and B. McShane (2019) "Scientists Rise Up Against Statistical Significance" *Nature* **567**, 305-307.

Caporale, G.M., L. Gil-Alana and A. Plastun (2019). "Long Memory and Data Frequency in Financial Markets" *Journal of Statistical Computation and Simulation* **89**, 1763-1779.

Davidson, R. and J.G. MacKinnon (2004) *Econometric Theory and Methods* Oxford University Press: New York.

Diebold, F.X. and L. Kilian (2000) "Unit Root Tests Are Useful For Selecting Forecasting Models" *Journal of Business and Economic Statistics* **18**, 265-273.

Hartl, T., R. Tschernig and E. Weber (2021) "Solving the Unobserved Components Puzzle: A Fractional Approach to Measuring the Business Cycle" University of Regensburg Working Paper.

Hassler, U. and J. Wolters (1994) “On the Power of Unit Root Tests Against Fractional Alternatives” *Economics Letters* **45**, 1-5.

Hyndman, R.J. and G. Athanasopoulos (2018) *Forecasting: Principles and Practice* (2nd Edition) https://otexts.com/fpp2/.

Hyndman, R.J. and Y. Khandakar (2008) “Automatic Time Series Forecasting: The forecast Package for R” *Journal of Statistical Software* **27**. https://www.jstatsoft.org

Meese, R.A. and K. Rogoff (1983) “Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?” *Journal of International Economics* **14**, 3-24.

Paparoditis, E. and D.N. Politis (2018) “The Asymptotic Size and Power of the Augmented Dickey–Fuller Test for a Unit Root” *Econometric Review* **37**, 955-973.

Wasserstein, R.L. and N.L. Lazar (2016) “The ASA Statement on p-Values: Context, Process, and Purpose” *The American Statistician* **70**, p.129-133.

Wolters, J. and U. Hassler (2006) “Unit Root Testing” *Allgemeines Statistisches Archiv* **50**, 43-58.

Ventosa-Santaulària, D. (2009) “Surious Regression” *Journal of Probability and Statistics* DOI:10.1155/2009/802975

Ziliiak, S.T. and D.N. McCloskey, D.N. (2007) *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives* The University of Michigan Press: Ann Arbor.

APPENDIX

Table A1: Our Monte Carlo Design

The constant a is set to 0 and the standard deviation of the error term is set to 1 for all sets

Set #	N (Sample Size)	Horizon	b values
1-7	50	1	0.9,0.95,0.975,0.99,1,1.5,2
8-14	50	3	
15-21	50	6	
22-28	200	1	
29-35	200	3	
36-42	200	6	
43-49	200	12	
50-56	500	1	
57-63	500	3	
64-70	500	6	
71-77	500	12	

Table A2A: Summary Statistics for Baseline AR(1)

Set #1: N=50 | Horizon=1 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	PP	AVG	KPSS	ADF
Relative RMSE	1	1.003455	1.007998	1.008445
Mean Ratio	1	1.004195	1.008908	1.008584
p-value	--	0.432	0.61	0.745

Set #2: N=50 | Horizon=1 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	PP	AVG	ADF	KPSS
Relative RMSE	1	1.00739	1.011537	1.015381
Mean Ratio	1	1.008005	1.011393	1.016112
p-value	--	0.383	0.736	0.575

Set #3: N=50 | Horizon=1 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	PP	AVG	ADF	KPSS
Relative RMSE	1	1.010008	1.010488	1.018991
Mean Ratio	1	1.011175	1.010755	1.019539
p-value	--	0.378	0.777	0.596

Set #4: N=50 | Horizon=1 | b=0.99				
Rank	1st	2nd	3rd	4th

Model	PP	AVG	ADF	KPSS
Relative RMSE	1	1.00991	1.010072	1.018807
Mean Ratio	1	1.011323	1.010149	1.019439
p-value	--	0.393	0.769	0.596

Set #5: N=50 | Horizon=1 | b=1				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	AVG	KPSS
Relative RMSE	1	1.008752	1.01594	1.016253
Mean Ratio	1	1.009168	1.017499	1.016387
p-value	--	0.767	0.421	0.623

Set #6: N=50 | Horizon=1 | b=1.5				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.006233	1.006234	1.178944
Mean Ratio	1	1.013271	1.012991	1.187369
p-value	--	0.509	0.647	0.573

Set #7: N=50 | Horizon=1 | b=2				
Rank	1st	2nd	3rd	4th
Model	KPSS	PP	ADF	AVG
Relative RMSE	1	1.283366	1.725714	18.013551
Mean Ratio	1	1.826307	2.447705	25.531011
p-value	--	0.783	0.779	0.055

Set #8: N=50 | Horizon=3 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	PP	KPSS	ADF	AVG
Relative RMSE	1	1.014197	1.02857	1.029255
Mean Ratio	1	1.016384	1.02886	1.03785
p-value	--	0.634	0.766	0.388

Set #9: N=50 | Horizon=3 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.024744	1.025368	1.026721
Mean Ratio	1	1.025463	1.025392	1.034516
p-value	--	0.779	0.605	0.363

Set #10: N=50 | Horizon=3 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	PP	AVG	ADF	KPSS

Relative RMSE	1	1.020786	1.028209	1.033499
Mean Ratio	1	1.027397	1.028273	1.034006
p-value	--	0.381	0.779	0.631

Set #11: N=50 | Horizon=3 | b=0.99				
Rank	1st	2nd	3rd	4th
Model	PP	AVG	ADF	KPSS
Relative RMSE	1	1.020928	1.030178	1.037587
Mean Ratio	1	1.026761	1.030633	1.038427
p-value	--	0.385	0.762	0.615

Set #12: N=50 | Horizon=3 | b=1				
Rank	1st	2nd	3rd	4th
Model	PP	AVG	ADF	KPSS
Relative RMSE	1	1.018256	1.026702	1.029974
Mean Ratio	1	1.025036	1.02768	1.031179
p-value	--	0.397	0.776	0.65

Set #13: N=50 | Horizon=3 | b=1.5				
Rank	1st	2nd	3rd	4th
Model	AVG	KPSS	PP	ADF
Relative RMSE	1	1.021028	1.035996	1.075434
Mean Ratio	1	1.036793	1.048971	1.090565
p-value	--	0.441	0.393	0.293

Set #14: N=50 | Horizon=3 | b=2				
Rank	1st	2nd	3rd	4th
Model	KPSS	PP	ADF	AVG
Relative RMSE	1	1.04513	1.217764	6.691965
Mean Ratio	1	1.160229	1.329423	7.722637
p-value	--	0.772	0.768	0.1

Set #15: N=50 | Horizon=6 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	KPSS	PP	ADF	AVG
Relative RMSE	1	1.008203	1.067561	1.106316
Mean Ratio	1	1.032556	1.097093	1.170762
p-value	--	0.632	0.562	0.306

Set #16: N=50 | Horizon=6 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	PP	KPSS	ADF	AVG
Relative RMSE	1	1.030387	1.071168	1.083631

Mean Ratio	1	1.036245	1.078841	1.122675
p-value	--	0.652	0.749	0.308

Set #17: N=50 | Horizon=6 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	PP	KPSS	AVG	ADF
Relative RMSE	1	1.033409	1.050179	1.052459
Mean Ratio	1	1.037028	1.080774	1.060603
p-value	--	0.649	0.358	0.753

Set #18: N=50 | Horizon=6 | b=0.99				
Rank	1st	2nd	3rd	4th
Model	PP	KPSS	AVG	ADF
Relative RMSE	1	1.034117	1.054336	1.066677
Mean Ratio	1	1.041564	1.087769	1.07436
p-value	--	0.679	0.334	0.772

Set #19: N=50 | Horizon=6 | b=1				
Rank	1st	2nd	3rd	4th
Model	PP	AVG	KPSS	ADF
Relative RMSE	1	1.036764	1.039198	1.055246
Mean Ratio	1	1.070475	1.039129	1.059355
p-value	--	0.371	0.689	0.793

Set #20: N=50 | Horizon=6 | b=1.5				
Rank	1st	2nd	3rd	4th
Model	AVG	KPSS	PP	ADF
Relative RMSE	1	1.077021	1.125792	1.191732
Mean Ratio	1	1.08255	1.126867	1.199453
p-value	--	0.42	0.363	0.279

Set #21: N=50 | Horizon=6 | b=2				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.06112	1.061748	3.385059
Mean Ratio	1	1.079697	1.109213	3.782588
p-value	--	0.854	0.662	0.166

Set #22: N=200 | Horizon=1 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.001331	1.003139	1.018882
Mean Ratio	1	1.001364	1.003239	1.019073

p-value	--	0.578	0.477	0.233

Set #23: N=200 | Horizon=1 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	KPSS	PP	ADF	AVG
Relative RMSE	1	1.000231	1.000367	1.012075
Mean Ratio	1	1.000257	1.000415	1.01217
p-value	--	0.665	0.664	0.26

Set #24: N=200 | Horizon=1 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.000136	1.000477	1.012846
Mean Ratio	1	1.000147	1.000538	1.012959
p-value	--	0.861	0.763	0.271

Set #25: N=200 | Horizon=1 | b=0.99				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.000212	1.00131	1.02109
Mean Ratio	1	1.000223	1.001347	1.021117
p-value	--	0.897	0.734	0.253

Set #26: N=200 | Horizon=1 | b=1				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1.002107	1.004663	1.095962
Mean Ratio	1	1.002178	1.004744	1.097244
p-value	--	0.892	0.893	0.243

Set #27: N=200 | Horizon=1 | b=1.5				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1.009022	1.015799	1.497889
Mean Ratio	1	1.009265	1.016105	1.497451
p-value	--	0.839	0.867	0.325

Set #28: N=200 | Horizon=1 | b=2				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	17.514322	22.3747	206.219115
Mean Ratio	1	17.661532	22.452207	207.663668
p-value	--	0.76	0.753	0.034

Set #29: N=200 | Horizon=3 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.004018	1.01004	1.072952
Mean Ratio	1	1.004356	1.011122	1.076298
p-value	--	0.562	0.445	0.13

Set #30: N=200 | Horizon=3 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	KPSS	PP	ADF	AVG
Relative RMSE	1	1.003267	1.003627	1.045195
Mean Ratio	1	1.003422	1.003875	1.047766
p-value	--	0.655	0.662	0.2

Set #31: N=200 | Horizon=3 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.000433	1.001184	1.040103
Mean Ratio	1	1.000425	1.001253	1.041947
p-value	--	0.857	0.769	0.216

Set #32: N=200 | Horizon=3 | b=0.99				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.000898	1.004076	1.049168
Mean Ratio	1	1.000855	1.004568	1.051306
p-value	--	0.887	0.727	0.225

Set #33: N=200 | Horizon=3 | b=1				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1.001348	1.002966	1.166385
Mean Ratio	1	1.002148	1.003567	1.169493
p-value	--	0.871	0.865	0.233

Set #34: N=200 | Horizon=3 | b=1.5				
Rank	1st	2nd	3rd	4th
Model	ADF	KPSS	PP	AVG
Relative RMSE	1	1.000855	1.002552	1.30796
Mean Ratio	1	1.005708	1.002836	1.3132
p-value	--	0.589	0.943	0.251

Set #35: N=200 | Horizon=3 | b=2				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	3.936626	5.165677	58.140174
Mean Ratio	1	3.955202	5.270147	60.395285
p-value	--	0.809	0.792	0.046

Set #36: N=200 | Horizon=6 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.007786	1.018622	1.177918
Mean Ratio	1	1.008572	1.022926	1.195531
p-value	--	0.573	0.442	0.086

Set #37: N=200 | Horizon=6 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1.002809	1.003512	1.108704
Mean Ratio	1	1.004235	1.004793	1.118641
p-value	--	0.669	0.663	0.112

Set #38: N=200 | Horizon=6 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.000654	1.004005	1.08688
Mean Ratio	1	1.000687	1.004904	1.096328
p-value	--	0.871	0.76	0.165

Set #39: N=200 | Horizon=6 | b=0.99				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.000472	1.00808	1.094193
Mean Ratio	1	1.000564	1.008029	1.102115
p-value	--	0.892	0.725	0.176

Set #40: N=200 | Horizon=6 | b=1				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.001265	1.004112	1.172762
Mean Ratio	1	1.002015	1.005701	1.184114
p-value	--	0.92	0.753	0.192

Set #41: N=200 | Horizon=6 | b=1.5

Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.000495	1.01563	1.209008
Mean Ratio	1	1.0025	1.018147	1.214299
p-value	--	0.913	0.609	0.258

Set #42: N=200 | Horizon=6 | b=2				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	2.434667	2.483053	24.943324
Mean Ratio	1	2.442281	2.559736	25.547177
p-value	--	0.81	0.792	0.04

Set #43: N=200 | Horizon=12 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.012226	1.030499	1.433694
Mean Ratio	1	1.01438	1.040967	1.490443
p-value	--	0.571	0.443	0.02

Set #44: N=200 | Horizon=12 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	KPSS	PP	ADF	AVG
Relative RMSE	1	1.004164	1.004608	1.253157
Mean Ratio	1	1.006346	1.006662	1.291357
p-value	--	0.699	0.694	0.068

Set #45: N=200 | Horizon=12 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.000549	1.005934	1.18933
Mean Ratio	1	1.000678	1.006546	1.218356
p-value	--	0.868	0.76	0.103

Set #46: N=200 | Horizon=12 | b=0.99				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.000701	1.013709	1.151638
Mean Ratio	1	1.001004	1.014006	1.180478
p-value	--	0.909	0.733	0.166

Set #47: N=200 | Horizon=12 | b=1				
Rank	1st	2nd	3rd	4th

Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.000393	1.015724	1.165693
Mean Ratio	1	1.000726	1.015536	1.197695
p-value	--	0.925	0.755	0.204

Set #48: N=200 | Horizon=12 | b=1.5				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.001336	1.031803	1.087317
Mean Ratio	1	1.00237	1.034229	1.109796
p-value	--	0.924	0.615	0.337

Set #49: N=200 | Horizon=12 | b=2				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1.402899	1.61796	9.476538
Mean Ratio	1	1.401325	1.620361	9.822624
p-value	--	0.799	0.761	0.07

Set #50: N=500 | Horizon=1 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.00006	1.010647	1.035855
Mean Ratio	1	1.00006	1.010673	1.035888
p-value	--	0.982	0.203	0.028

Set #51: N=500 | Horizon=1 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.000488	1.005006	1.025376
Mean Ratio	1	1.000489	1.005034	1.025434
p-value	--	0.645	0.287	0.064

Set #52: N=500 | Horizon=1 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	KPSS	PP	ADF	AVG
Relative RMSE	1	1.000027	1.000101	1.021098
Mean Ratio	1	1.000042	1.000117	1.021096
p-value	--	0.693	0.688	0.062

Set #53: N=500 | Horizon=1 | b=0.99				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG

Relative RMSE	1	1.002059	1.002396	1.046412
Mean Ratio	1	1.002083	1.00246	1.046487
p-value	--	0.81	0.8	0.059

Set #54: N=500 | Horizon=1 | b=1				
Rank	1st	2nd	3rd	4th
Model	KPSS	PP	ADF	AVG
Relative RMSE	1	1.02052	1.020749	1.190873
Mean Ratio	1	1.02067	1.020787	1.191166
p-value	--	0.88	0.881	0.042

Set #55: N=500 | Horizon=1 | b=1.5				
Rank	1st	2nd	3rd	4th
Model	KPSS	PP	ADF	AVG
Relative RMSE	1	1.031386	1.040838	1.52338
Mean Ratio	1	1.030264	1.038921	1.521829
p-value	--	0.89	0.854	0.086

Set #56: N=500 | Horizon=1 | b=2				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	89.086028	163.058565	910.918427
Mean Ratio	1	88.671228	163.740244	913.171557
p-value	--	0.766	0.68	0.009

Set #57: N=500 | Horizon=3 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.000023	1.029878	1.129021
Mean Ratio	1	1.000024	1.02982	1.130358
p-value	--	0.989	0.202	0.008

Set #58: N=500 | Horizon=3 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.000696	1.014545	1.086868
Mean Ratio	1	1.000738	1.014802	1.088264
p-value	--	0.649	0.256	0.024

Set #59: N=500 | Horizon=3 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	KPSS	PP	ADF	AVG
Relative RMSE	1	1.000316	1.000982	1.068573

Mean Ratio	1	1.000433	1.001083	1.069678
p-value	--	0.692	0.665	0.036

Set #60: N=500 | Horizon=3 | b=0.99				
Rank	1st	2nd	3rd	4th
Model	KPSS	PP	ADF	AVG
Relative RMSE	1	1.001153	1.001526	1.114964
Mean Ratio	1	1.001078	1.001408	1.116442
p-value	--	0.828	0.831	0.038

Set #61: N=500 | Horizon=3 | b=1				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1.016588	1.02037	1.551227
Mean Ratio	1	1.017735	1.021877	1.5525
p-value	--	0.917	0.905	0.023

Set #62: N=500 | Horizon=3 | b=1.5				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1.026642	1.032241	1.773309
Mean Ratio	1	1.028144	1.033487	1.776353
p-value	--	0.859	0.9	0.072

Set #63: N=500 | Horizon=3 | b=2				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	27.614636	54.011441	252.923501
Mean Ratio	1	27.503982	54.042266	253.687137
p-value	--	0.781	0.693	0

Set #64: N=500 | Horizon=6 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.00023	1.055766	1.291237
Mean Ratio	1	1.000225	1.056004	1.296852
p-value	--	0.98	0.201	0.002

Set #65: N=500 | Horizon=6 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.002137	1.029398	1.19066
Mean Ratio	1	1.002111	1.030853	1.197066

p-value	--	0.655	0.264	0.013

Set #66: N=500 | Horizon=6 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.000427	1.000494	1.139901
Mean Ratio	1	1.000458	1.000977	1.145083
p-value	--	0.713	0.681	0.023

Set #67: N=500 | Horizon=6 | b=0.99				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1.001171	1.001277	1.190894
Mean Ratio	1	1.001409	1.001525	1.195424
p-value	--	0.842	0.834	0.043

Set #68: N=500 | Horizon=6 | b=1				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1.013072	1.01569	1.612353
Mean Ratio	1	1.01485	1.017507	1.614163
p-value	--	0.882	0.877	0.026

Set #69: N=500 | Horizon=6 | b=1.5				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1.008913	1.038041	1.808439
Mean Ratio	1	1.009044	1.038015	1.813686
p-value	--	0.878	0.906	0.102

Set #70: N=500 | Horizon=6 | b=2				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	7.192764	18.427572	92.550488
Mean Ratio	1	7.220018	18.69335	93.7945
p-value	--	0.784	0.665	0.005

Set #71: N=500 | Horizon=12 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.00088	1.103167	1.699199
Mean Ratio	1	1.000835	1.104449	1.719374
p-value	--	0.973	0.178	0

Set #72: N=500 | Horizon=12 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.003469	1.054778	1.439386
Mean Ratio	1	1.00358	1.058229	1.459342
p-value	--	0.652	0.249	0.003

Set #73: N=500 | Horizon=12 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.000988	1.004957	1.300484
Mean Ratio	1	1.001117	1.00706	1.31694
p-value	--	0.708	0.646	0.01

Set #74: N=500 | Horizon=12 | b=0.99				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.001016	1.001308	1.261561
Mean Ratio	1	1.00103	1.001746	1.276623
p-value	--	0.853	0.826	0.024

Set #75: N=500 | Horizon=12 | b=1				
Rank	1st	2nd	3rd	4th
Model	KPSS	PP	ADF	AVG
Relative RMSE	1	1.000381	1.000912	1.46093
Mean Ratio	1	1.002492	1.003604	1.477239
p-value	--	0.884	0.882	0.048

Set #76: N=500 | Horizon=12 | b=1.5				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.001179	1.008082	1.391161
Mean Ratio	1	1.001191	1.009287	1.399122
p-value	--	0.966	0.663	0.127

Set #77: N=500 | Horizon=12 | b=2				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	4.809006	8.634597	37.886865
Mean Ratio	1	4.799286	8.71227	38.656192
p-value	--	0.77	0.697	0.007

Appendix Table A2B: Summary Statistics for ARMA Specification

Set #1: N=50 | Horizon=1 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	AVG	KPSS	PP	ADF
Relative RMSE	1	1.011048	1.012487	1.023303
Mean Ratio	1	1.012702	1.013881	1.024678
p-value	--	0.467	0.462	0.412

Set #2: N=50 | Horizon=1 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	AVG	PP	KPSS	ADF
Relative RMSE	1	1.010098	1.016091	1.020707
Mean Ratio	1	1.012092	1.018282	1.023141
p-value	--	0.452	0.436	0.386

Set #3: N=50 | Horizon=1 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	PP	AVG	ADF	KPSS
Relative RMSE	1	1.005069	1.010025	1.021417
Mean Ratio	1	1.007062	1.010866	1.023152
p-value	--	0.518	0.695	0.639

Set #4: N=50 | Horizon=1 | b=0.99				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.009598	1.010329	1.043436
Mean Ratio	1	1.010383	1.010698	1.044304
p-value	--	0.711	0.648	0.517

Set #5: N=50 | Horizon=1 | b=1				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.00778	1.012554	1.071254
Mean Ratio	1	1.008794	1.013779	1.07467
p-value	--	0.721	0.641	0.492

Set #6: N=50 | Horizon=1 | b=1.5				
Rank	1st	2nd	3rd	4th
Model	PP	KPSS	ADF	AVG
Relative RMSE	1	1.037415	1.046247	1.170455
Mean Ratio	1	1.017646	1.022198	1.18018
p-value	--	0.658	0.585	0.609

Set #7: N=50 | Horizon=1 | b=2				
Rank	1st	2nd	3rd	4th
Model	PP	KPSS	ADF	AVG
Relative RMSE	1	1.000021	2.985397	4.284066
Mean Ratio	1	1.933221	29.036966	33.477276
p-value	--	0.703	0.704	0.042

Set #8: N=50 | Horizon=3 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	KPSS	AVG	PP	ADF
Relative RMSE	1	1.018785	1.02043	1.054325
Mean Ratio	1	1.039073	1.025855	1.055447
p-value	--	0.437	0.598	0.533

Set #9: N=50 | Horizon=3 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	AVG	PP	KPSS	ADF
Relative RMSE	1	1.000383	1.008046	1.047664
Mean Ratio	1	0.995736	1.006059	1.038079
p-value	--	0.558	0.533	0.454

Set #10: N=50 | Horizon=3 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	AVG	PP	KPSS	ADF
Relative RMSE	1	1.007942	1.029485	1.050003
Mean Ratio	1	1.00285	1.023245	1.041624
p-value	--	0.54	0.495	0.441

Set #11: N=50 | Horizon=3 | b=0.99				
Rank	1st	2nd	3rd	4th
Model	PP	AVG	ADF	KPSS
Relative RMSE	1	1.004179	1.023231	1.024276
Mean Ratio	1	1.018281	1.023019	1.024659
p-value	--	0.448	0.73	0.665

Set #12: N=50 | Horizon=3 | b=1				
Rank	1st	2nd	3rd	4th
Model	PP	KPSS	AVG	ADF
Relative RMSE	1	1.02333	1.031936	1.050286
Mean Ratio	1	1.024844	1.043693	1.049981
p-value	--	0.67	0.415	0.683

Set #13: N=50 | Horizon=3 | b=1.5				
Rank	1st	2nd	3rd	4th
Model	AVG	KPSS	ADF	PP
Relative RMSE	1	1.019751	1.078221	1.109534
Mean Ratio	1	1.031107	1.089852	1.101612
p-value	--	0.441	0.289	0.3

Set #14: N=50 | Horizon=3 | b=2				
Rank	1st	2nd	3rd	4th
Model	PP	KPSS	AVG	ADF
Relative RMSE	1	1.000003	1.601731	1.653814
Mean Ratio	1	1.253395	8.331	6.977538
p-value	--	0.677	0.087	0.714

Set #15: N=50 | Horizon=6 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	KPSS	PP	AVG	ADF
Relative RMSE	1	1.054277	1.102511	1.167934
Mean Ratio	1	1.07511	1.173499	1.186033
p-value	--	0.553	0.322	0.484

Set #16: N=50 | Horizon=6 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	KPSS	PP	AVG	ADF
Relative RMSE	1	1.01734	1.03479	1.107172
Mean Ratio	1	1.026739	1.103845	1.114116
p-value	--	0.676	0.402	0.583

Set #17: N=50 | Horizon=6 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	AVG	KPSS	PP	ADF
Relative RMSE	1	1.007213	1.008295	1.10424
Mean Ratio	1	0.995764	0.982569	1.069477
p-value	--	0.562	0.582	0.486

Set #18: N=50 | Horizon=6 | b=0.99				
Rank	1st	2nd	3rd	4th
Model	PP	KPSS	AVG	ADF
Relative RMSE	1	1.014966	1.018921	1.081541
Mean Ratio	1	1.025669	1.082831	1.096729
p-value	--	0.667	0.387	0.69

Set #19: N=50 | Horizon=6 | b=1				

Rank	1st	2nd	3rd	4th
Model	PP	KPSS	AVG	ADF
Relative RMSE	1	1.008941	1.01556	1.085246
Mean Ratio	1	1.023155	1.07453	1.0985
p-value	--	0.721	0.401	0.693

Set #20: N=50 | Horizon=6 | b=1.5				
Rank	1st	2nd	3rd	4th
Model	KPSS	AVG	ADF	PP
Relative RMSE	1	1.00529	1.121816	1.284569
Mean Ratio	1	1.035063	1.192859	1.234813
p-value	--	0.502	0.444	0.522

Set #21: N=50 | Horizon=6 | b=2				
Rank	1st	2nd	3rd	4th
Model	PP	KPSS	AVG	ADF
Relative RMSE	1	1.000005	1.169426	1.278283
Mean Ratio	1	1.138399	3.722727	2.54275
p-value	--	0.705	0.149	0.743

Set #22: N=200 | Horizon=1 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.000012	1.002443	1.009309
Mean Ratio	1	1.000134	1.002655	1.009667
p-value	--	0.647	0.504	0.348

Set #23: N=200 | Horizon=1 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	ADF	KPSS	PP	AVG
Relative RMSE	1	1.000076	1.001046	1.004341
Mean Ratio	1	1.000184	1.00108	1.0046
p-value	--	0.672	0.692	0.403

Set #24: N=200 | Horizon=1 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.000562	1.004758	1.00573
Mean Ratio	1	1.000554	1.005119	1.00603
p-value	--	0.829	0.733	0.409

Set #25: N=200 | Horizon=1 | b=0.99				
Rank	1st	2nd	3rd	4th

Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.000218	1.001169	1.018333
Mean Ratio	1	1.000196	1.001179	1.01858
p-value	--	0.889	0.747	0.373

Set #26: N=200 | Horizon=1 | b=1				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.00018	1.002229	1.192736
Mean Ratio	1	1.000227	1.002236	1.200115
p-value	--	0.907	0.711	0.268

Set #27: N=200 | Horizon=1 | b=1.5				
Rank	1st	2nd	3rd	4th
Model	AVG	PP	KPSS	ADF
Relative RMSE	1	2.177239	2.177339	2.177848
Mean Ratio	1	0.83016	0.835098	0.835695
p-value	--	0.773	0.726	0.738

Set #28: N=200 | Horizon=1 | b=2				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1	1	3.249204
Mean Ratio	1	448.848002	1.904812	348.637237
p-value	--	0.649	0.82	0.023

Set #29: N=200 | Horizon=3 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.000865	1.005072	1.051334
Mean Ratio	1	0.99931	1.006207	1.054955
p-value	--	0.642	0.512	0.237

Set #30: N=200 | Horizon=3 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1.00068	1.002581	1.031577
Mean Ratio	1	1.001111	1.002117	1.035254
p-value	--	0.666	0.645	0.252

Set #31: N=200 | Horizon=3 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG

Relative RMSE	1	1.000488	1.001478	1.024889
Mean Ratio	1	1.000729	1.002216	1.028576
p-value	--	0.759	0.739	0.299

Set #32: N=200 | Horizon=3 | b=0.99				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.000083	1.005801	1.051716
Mean Ratio	1	1.000786	1.005324	1.052053
p-value	--	0.86	0.705	0.264

Set #33: N=200 | Horizon=3 | b=1				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.000735	1.001838	1.297906
Mean Ratio	1	1.000667	1.003913	1.253386
p-value	--	0.9	0.742	0.195

Set #34: N=200 | Horizon=3 | b=1.5				
Rank	1st	2nd	3rd	4th
Model	AVG	KPSS	ADF	PP
Relative RMSE	1	1.082137	1.082137	1.422378
Mean Ratio	1	0.815108	0.815639	0.807727
p-value	--	0.766	0.806	0.83

Set #35: N=200 | Horizon=3 | b=2				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1	1	1.63204
Mean Ratio	1	81.38863	1.188207	76.518175
p-value	--	0.663	0.808	0.025

Set #36: N=200 | Horizon=6 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.003368	1.011274	1.146428
Mean Ratio	1	0.999883	1.014226	1.16171
p-value	--	0.669	0.512	0.119

Set #37: N=200 | Horizon=6 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1.001756	1.009515	1.08026

Mean Ratio	1	1.001835	1.007121	1.092965
p-value	--	0.682	0.636	0.174

Set #38: N=200 | Horizon=6 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.001285	1.002656	1.062141
Mean Ratio	1	1.001931	1.003647	1.077501
p-value	--	0.819	0.729	0.215

Set #39: N=200 | Horizon=6 | b=0.99				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.000399	1.007229	1.063976
Mean Ratio	1	1.001995	1.006531	1.076915
p-value	--	0.851	0.729	0.221

Set #40: N=200 | Horizon=6 | b=1				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.000317	1.013806	1.179148
Mean Ratio	1	0.999987	1.011163	1.180176
p-value	--	0.909	0.699	0.214

Set #41: N=200 | Horizon=6 | b=1.5				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.000004	1.000005	1.785351
Mean Ratio	1	0.992764	1.021713	1.260858
p-value	--	0.928	0.588	0.206

Set #42: N=200 | Horizon=6 | b=2				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1	1	1.121228
Mean Ratio	1	24.557153	1.155671	28.694526
p-value	--	0.667	0.803	0.033

Set #43: N=200 | Horizon=12 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.00748	1.016533	1.432524
Mean Ratio	1	1.000864	1.026203	1.472025

p-value	--	0.665	0.505	0.014

Set #44: N=200 | Horizon=12 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1.008108	1.009921	1.219489
Mean Ratio	1	1.00966	1.009014	1.257142
p-value	--	0.657	0.651	0.067

Set #45: N=200 | Horizon=12 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1.000825	1.004537	1.145873
Mean Ratio	1	1.002721	1.005155	1.189848
p-value	--	0.762	0.758	0.136

Set #46: N=200 | Horizon=12 | b=0.99				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.003348	1.016446	1.109711
Mean Ratio	1	1.000748	1.014383	1.155547
p-value	--	0.888	0.741	0.178

Set #47: N=200 | Horizon=12 | b=1				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.001456	1.023113	1.14641
Mean Ratio	1	0.999511	1.016272	1.182806
p-value	--	0.906	0.744	0.194

Set #48: N=200 | Horizon=12 | b=1.5				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.000206	1.000629	3.529125
Mean Ratio	1	1.014252	1.041451	1.258368
p-value	--	0.877	0.618	0.231

Set #49: N=200 | Horizon=12 | b=2				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1	1	1.013483
Mean Ratio	1	6.237203	1.04339	10.193978
p-value	--	0.682	0.827	0.062

Set #50: N=500 | Horizon=1 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.000297	1.009685	1.025704
Mean Ratio	1	1.000297	1.009691	1.025749
p-value	--	0.927	0.242	0.054

Set #51: N=500 | Horizon=1 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.000397	1.005291	1.015827
Mean Ratio	1	1.000405	1.005338	1.015862
p-value	--	0.692	0.305	0.149

Set #52: N=500 | Horizon=1 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	ADF	KPSS	PP	AVG
Relative RMSE	1	1.000364	1.000623	1.009371
Mean Ratio	1	1.000381	1.000634	1.009422
p-value	--	0.635	0.642	0.229

Set #53: N=500 | Horizon=1 | b=0.99				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1.00016	1.00023	1.031876
Mean Ratio	1	1.000179	1.000258	1.032067
p-value	--	0.792	0.788	0.174

Set #54: N=500 | Horizon=1 | b=1				
Rank	1st	2nd	3rd	4th
Model	KPSS	PP	ADF	AVG
Relative RMSE	1	1.012753	1.013735	1.424707
Mean Ratio	1	1.012767	1.013687	1.423536
p-value	--	0.864	0.88	0.078

Set #55: N=500 | Horizon=1 | b=1.5				
Rank	1st	2nd	3rd	4th
Model	AVG	KPSS	PP	ADF
Relative RMSE	1	2.437439	2.437439	2.437439
Mean Ratio	1	0.725396	0.738947	0.74987
p-value	--	0.927	0.926	0.904

Set #56: N=500 | Horizon=1 | b=2				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1	1	3.783315
Mean Ratio	1	3203.22064	81.894367	1930.85619
p-value	--	0.64	0.734	0.004

Set #57: N=500 | Horizon=3 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.000736	1.029617	1.095795
Mean Ratio	1	1.000808	1.029282	1.095958
p-value	--	0.927	0.213	0.019

Set #58: N=500 | Horizon=3 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.000699	1.014909	1.055393
Mean Ratio	1	1.001267	1.014744	1.057032
p-value	--	0.667	0.268	0.078

Set #59: N=500 | Horizon=3 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	ADF	KPSS	PP	AVG
Relative RMSE	1	1.000989	1.001467	1.035496
Mean Ratio	1	1.000809	1.001048	1.038446
p-value	--	0.651	0.632	0.171

Set #60: N=500 | Horizon=3 | b=0.99				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1.00059	1.000923	1.107239
Mean Ratio	1	1.000754	1.001086	1.107975
p-value	--	0.792	0.807	0.107

Set #61: N=500 | Horizon=3 | b=1				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1.017207	1.020338	1.897285
Mean Ratio	1	1.016055	1.018289	1.83566
p-value	--	0.85	0.847	0.05

Set #62: N=500 | Horizon=3 | b=1.5

Rank	1st	2nd	3rd	4th
Model	AVG	PP	KPSS	ADF
Relative RMSE	1	1.432621	1.432621	1.432621
Mean Ratio	1	0.648507	0.64579	0.656442
p-value	--	0.95	0.949	0.934

Set #63: N=500 | Horizon=3 | b=2				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1	1	1.496086
Mean Ratio	1	347.869817	17.318701	348.361738
p-value	--	0.621	0.719	0.007

Set #64: N=500 | Horizon=6 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.001907	1.060076	1.232837
Mean Ratio	1	1.001989	1.059124	1.234361
p-value	--	0.92	0.193	0.004

Set #65: N=500 | Horizon=6 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.002626	1.030204	1.134291
Mean Ratio	1	1.003574	1.030516	1.141389
p-value	--	0.66	0.255	0.036

Set #66: N=500 | Horizon=6 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.003931	1.004519	1.082319
Mean Ratio	1	1.002169	1.004364	1.090966
p-value	--	0.644	0.625	0.084

Set #67: N=500 | Horizon=6 | b=0.99				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1.00066	1.000768	1.111767
Mean Ratio	1	1.00117	1.001458	1.123623
p-value	--	0.792	0.795	0.073

Set #68: N=500 | Horizon=6 | b=1				
Rank	1st	2nd	3rd	4th

Model	KPSS	PP	ADF	AVG
Relative RMSE	1	1.023648	1.02399	1.622477
Mean Ratio	1	1.014661	1.014612	1.584208
p-value	--	0.863	0.874	0.04

Set #69: N=500 | Horizon=6 | b=1.5				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1	1	1.971821
Mean Ratio	1	1.023004	1.009196	1.736087
p-value	--	0.826	0.682	0.07

Set #70: N=500 | Horizon=6 | b=2				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1	1	1.105456
Mean Ratio	1	94.52397	7.575951	110.550226
p-value	--	0.646	0.742	0.007

Set #71: N=500 | Horizon=12 | b=0.9				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.002589	1.100208	1.570159
Mean Ratio	1	1.002656	1.098883	1.576123
p-value	--	0.933	0.22	0

Set #72: N=500 | Horizon=12 | b=0.95				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.003349	1.057328	1.335846
Mean Ratio	1	1.006081	1.060347	1.351585
p-value	--	0.681	0.258	0.01

Set #73: N=500 | Horizon=12 | b=0.975				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.008518	1.009808	1.191495
Mean Ratio	1	1.004064	1.009952	1.214116
p-value	--	0.663	0.608	0.038

Set #74: N=500 | Horizon=12 | b=0.99				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG

Relative RMSE	1	1.002403	1.00597	1.156584
Mean Ratio	1	1.003612	1.005537	1.18671
p-value	--	0.804	0.788	0.062

Set #75: N=500 | Horizon=12 | b=1				
Rank	1st	2nd	3rd	4th
Model	PP	KPSS	ADF	AVG
Relative RMSE	1	1.000031	1.002534	1.420532
Mean Ratio	1	1.00461	1.001484	1.412913
p-value	--	0.756	0.912	0.058

Set #76: N=500 | Horizon=12 | b=1.5				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1	1	2.776953
Mean Ratio	1	1.016203	1.032556	1.438811
p-value	--	0.827	0.689	0.107

Set #77: N=500 | Horizon=12 | b=2				
Rank	1st	2nd	3rd	4th
Model	KPSS	ADF	PP	AVG
Relative RMSE	1	1	1	1.015814
Mean Ratio	1	30.534417	3.571497	39.535414
p-value	--	0.624	0.733	0.009

Appendix Table A3: Summary Statistics for Structural Break Specification

Structural Break Set #1: N=50 | Horizon=1				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	AVG	KPSS
Relative RMSE	1	1.002246	1.006487	1.009547
Mean Ratio	1	1.003191	1.006538	1.010206
p-value	--	0.873	0.59	0.648

Structural Break Set #2: N=50 | Horizon=3				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	AVG	KPSS
Relative RMSE	1	1.00188	1.009671	1.022781
Mean Ratio	1	1.003316	1.003943	1.025416
p-value	--	0.881	0.642	0.662

Structural Break Set #3: N=50 | Horizon=6				
Rank	1st	2nd	3rd	4th
Model	AVG	ADF	PP	KPSS
Relative RMSE	1	1.048804	1.050604	1.055673
Mean Ratio	1	1.068547	1.071279	1.067635
p-value	--	0.306	0.294	0.303

Structural Break Set #4: N=200 | Horizon=1				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.00022	1.000941	1.024906
Mean Ratio	1	1.000296	1.000988	1.02546
p-value	--	0.783	0.723	0.707

Structural Break Set #5: N=200 | Horizon=3				
Rank	1st	2nd	3rd	4th
Model	KPSS	PP	ADF	AVG
Relative RMSE	1	1.000355	1.000392	1.02087
Mean Ratio	1	1.000721	1.000785	1.014743
p-value	--	0.722	0.744	0.773

Structural Break Set #6: N=200 | Horizon=6				
Rank	1st	2nd	3rd	4th
Model	AVG	KPSS	ADF	PP
Relative RMSE	1	1.006472	1.006927	1.007616
Mean Ratio	1	1.025099	1.02577	1.026194
p-value	--	0.226	0.243	0.237

Structural Break Set #7: N=200 | Horizon=12				
Rank	1st	2nd	3rd	4th
Model	AVG	ADF	KPSS	PP
Relative RMSE	1	1.048606	1.05015	1.052569
Mean Ratio	1	1.054629	1.057867	1.058038
p-value	--	0.223	0.215	0.215

Structural Break Set #8: N=500 | Horizon=1				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.000169	1.002964	1.05001
Mean Ratio	1	1.000193	1.003039	1.047894
p-value	--	0.726	0.516	0.59

Structural Break Set #9: N=500 | Horizon=3				
Rank	1st	2nd	3rd	4th
Model	ADF	PP	KPSS	AVG
Relative RMSE	1	1.000114	1.004966	1.095034
Mean Ratio	1	0.999823	1.00648	1.086249
p-value	--	0.794	0.54	0.594

Structural Break Set #10: N=500 | Horizon=6				
Rank		2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.000508	1.010043	1.094365
Mean Ratio	1	1.000925	1.012903	1.089891
p-value	--	0.756	0.543	0.607

Structural Break Set #11: N=500 | Horizon=12				
Rank	1st	2nd	3rd	4th
Model	PP	ADF	KPSS	AVG
Relative RMSE	1	1.00109	1.022514	1.040824
Mean Ratio	1	1.002365	1.029266	1.023581
p-value	--	0.731	0.502	0.604