A Simple Hybrid Bootstrap Test for Predictive Ability Based on Autoregressions

Ching-Chuan Tsong*

National Chi Nan University

Cheng-Feng Lee**

National Kaohsiung University of Applied Sciences

Biing-Shen Kuo***

National ChengChi University

May 2007

* Department of Economics, National Chi Nan University, Puli 545, Taiwan; Tel: (886 49) 2910 960 ext. 4920, Fax: (886 49) 291-4435, Email: tcc126@ncnu.edu.tw

** Department of Business Administration, National Kaohsiung University of Applied Sciences, No.415, Jiangong Rd., Sanmin District, Kaohsiung City 80778, Taiwan (R.O.C.); Tel: (886 7) 381-4526 ext. 7310, Fax: (886 7) 396-1245, Email: jflee@cc.kuas.edu.tw

*** Department of International Trade, National Chengchi University, Taipei 116, Taiwan; Tel: (886 2) 2939 3091 ext. 81046, Tel: (886 2) 2939 3091 ext. 81029, Fax: (886 2) 2938 7699, Email: bsku@nccu.edu.tw

Abstract

It has been by far well-documented that the extant tests for predictive ability suffer from size distortions. It results from the sensitivity of the tests to the choice of kernel functions in the estimation of the long-run variance. This paper proposes a bootstrap testing procedure using autoregression to overcome this problem. The regression approach transforms testing for the null of equal predictive ability into testing for an zero intercept in the estimated autoregression, which is easy to implement. To respectively account for autocorrelation and to retain heteroskedasticity of unknown from in forecast errors, the proposed resampling scheme combines both autoregressive sieve and wild bootstraps. We establish the bootstrap consistency by showing that the suggested test and the asymptotic counterpart have the same normality limit in distributions. Simulations revealed that our bootstrap testing procedure has a robust size performance to correlated forecast errors with conditional variance of GARCH or SV, in contrast to the existing tests adopting a moving block bootstrap.

JEL classification: C22; E32

Keywords: sieve bootstrap; wild bootstrap; predictive accuracy test, encompassing test

1. Introduction

Prediction has drawn a considerable amount of attention for decades, particularly in the fields of economics and finance. Forecasts of variables are useful, not only to know the future path of the economy but also for choosing the most proper specification among competing empirical models. For these purposes, we require some criteria to evaluate the out-of-sample forecast performances of given models. A common, though informal, criterion for identifying the best model is to select the one with the minimum mean square prediction error (MSPE) or mean absolute error (MAE) (Meese and Rogoff, 1983; Akgiray, 1989) among models. However, merely using MSPE or MAE appears to be inappropriate, since they are all random variables with potentially unknown distributions. We cannot make any statistical inference regarding the realizations without knowing their distributions.

Two formal approaches can be used to evaluate the forecast performance. One option involves adopting the predictive accuracy test proposed by Diebold and Mariano (1995) (henceforth DM) in order to test the null hypothesis that the population means of the economic loss functions for two different models are equal. The alternative option is to apply the forecasting encompassing test (see Chong and Hendry, 1986; Clements and Hendry, 1993) (henceforth ENC) to decide if the preferred model contains more useful information than the other ones. Both tests have the same asymptotic standard normal distribution even though they are constructed differently.

It is well known that the forecast errors generated from many predicting models are usually serially correlated and conditionally heteroskedastic, particularly when the forecast horizon is long. DM and ENC tests deal with the dependence of forecast errors by using a non-parametric method in order to estimate the long-run variance, wherein a suitable choice of kernel function and bandwidth is involved. Using simulations, Clark (1999) pointed out that the results from DM are considerably affected by the selection of different kernels in finite samples. The idea of implementing these tests without non-parametrically calculating the long-run variance is appealing to empirical researchers.

For the purpose of illustration, let us assume that two forecast errors, denoted by e_{1t} and e_{2t}, are obtained from two competing models. Both DM and ENC tests are conducted to test the null hypothesis that the sample mean of series d_t[1] equals zero. By borrowing this idea from the Augmented Dickey-Fuller (ADF) unit root test, we have proposed a new testing procedure for the extant tests in this paper. We have treated series d_t as a covariance stationary process with moving average (MA) properties. Under some general conditions (e.g. Bühlmann, 1995, Lemma 2.1), it can be expressed as an infinite-order autoregressive (AR(∞)) process with possibly conditional heteroskedastic errors. In ideal circumstances, it is feasible to approximate the AR(∞) model with a p-order AR model, where order p should increase with the sample size at a certain rate. The intercept in the AR(p) model, which stands for the population mean of series d_t, is the parameter on which we will focus.

To be specific, we have tested the null hypothesis by first fitting series d_t with the AR(p) model with an intercept. As determined by Akaike Information Criteria (AIC), order p needs to be sufficiently long to account for serial correlation. Thereafter, we have estimated the model using OLS and tested for zero intercept using *t*-statistics (henceforth $AR-t_c$). The proposed $AR-t_c$ statistics, which can be shown to have the same asymptotic standard normal distribution as DM and ENC tests, can be easily used in major standard econometric software.

---

[1] $d_t = e_{1t}{}^2 - e_{2t}{}^2$ and $e_{1t}(e_{1t} - e_{2t})$ for DM and ENC tests, respectively.

Our simulation results reveal that the $AR-t_c$ test outperforms DM and ENC in size and shows a comparable power across a wide variety of cases. However, it continues to encounter a size distortion problem, particularly when the sample size is small and forecast horizon is long. Hence, a new bootstrap procedure, aimed at estimating the finite sample distribution of $AR-t_c$ test, is proposed to cope with such size distortions. Our bootstrap is a hybrid one, in which we have combined the autoregressive sieve bootstrap and wild bootstrap (WB) for different purposes. The AR sieve bootstrap is first used to account for the serial correlations in series d_t, which is accomplished by fitting the series with a finite AR process so that its order increases with the sample size. In turn, the conditional heteroskedasticity of unknown form in d_t is addressed by applying the wild bootstrap to the residuals of the AR sieve regression.

The two bootstrap methods employed in our research have been widely used in the literature regarding this topic. Chang and Park (2003) established the use of the sieve bootstrap for an ADF unit root test with general linear processes. With possible conditional heteroskedasticity of unknown form for autoregressive processes, Goncalves and Kilian (2004) contributed to the study of three easy-to-implement wild bootstrap proposals and demonstrated their asymptotic validity. With the help of simulations, Flachaire (2005) investigated and compared the finite sample performances of wild bootstrap as well as pairs bootstrap. In the research by O'Reilly and Whelany (2005), the wild bootstrap has been applied to generate critical values for testing parameter stability. Kapetanios and Psaradakis (2006) studied the properties of the sieve bootstrap for a class of linear processes that exhibit strong dependence and formulated the asymptotic validity of the sieve bootstrap in the case of sample mean and sample autocovariances. Richard (2006) proposed the use of an MA sieve bootstrap for the ADF test and demonstrated that the test based on this bootstrap was consistent.

A bootstrap procedure remains incomplete if its consistency is not provided, in the sense that the asymptotic distribution of the $AR-t_c$ test is the same as that of its bootstrap counterpart. In this paper, we show that the $AR-t_c$ test based on the AR sieve and wild bootstraps is asymptotically valid. In addition, another bootstrap method for dependent data, called moving block bootstrap (MBB, Künsch, 1989), is also investigated for the sake of comparison. The MBB is implemented by dividing the data into blocks, sampling the blocks randomly with replacement and laying them end-to-end in order to produce the bootstrap sample. We employ MBB in DM and ENC tests and compare their respective performances with that of the bootstrapped $AR-t_c$ test.

Simulation results indicate that despite the conditionally heteroskedastic forecast errors, our bootstrap procedures can effectively reduce the size distortions of the $AR-t_c$ test in various combinations of sample size and forecast horizon. Moreover, the size results of DM and ENC with MBB are apparently inferior to those of the bootstrapped $AR-t_c$ test. As expected, MBB heavily depends on the choice of different block lengths.

This paper proceeds as follows. DM and ENC tests are briefly reviewed in Section 2. We have introduced the $AR-t_c$ test in Section 3 and studied its asymptotic behaviour by deriving the limiting distribution. In Section 4, we have proposed bootstrap procedures for the $AR-t_c$ test and developed the asymptotic distribution of its bootstrap test. Simulation results pertaining to size and power are presented in Section 5. Section 6 concludes the paper, and mathematical proofs are collated in the appendix.

2. The Review of Forecast Tests

In this section, we have briefly reviewed two predictive performance tests—the predictive accuracy test and forecast encompassing test.

With two forecasting models for the same variable y_t, researchers are often concerned about how to discriminate between them in terms of their out-of-sample prediction performance. Diebold and Mariano (1995) pioneered the work of comparing the predictive accuracy of two forecasts in a statistical framework. The asymptotics for DM test are established by assuming the model parameters without estimation. This can be outlined as follows.

Suppose (e_{1t}, e_{2t}), $t = 1, 2, ..., n$ are h-step forecast errors of models 1 and 2, respectively. Taking MSPE as a measure of prediction loss, the loss differential from the two models can be expressed as $d_t = e_{1t}{}^2 - e_{2t}{}^2$. The DM test is implemented to compare whether the population mean of d_t is equal to zero. The null hypothesis of DM test is $E[d_t] = 0$, as against the alternative hypothesis $E[d_t] \neq 0$.

DM test can be formulized as commonly used t-statistics by using the sample mean of d_t to test the null hypothesis as follows.

$$DM = \frac{\bar{d}}{\sqrt{Var(\bar{d})}}, \tag{1}$$

where $\bar{d}$ refers to the sample mean of d_t and $Var(\bar{d})$ is a consistent estimator for the long-run variance of d_t.

The alternative method for forecast evaluation involves the use of ENC test, which is based on the fact that the competing forecasts could have a different information content than the preferred one. Harvey et al. (1998) proposed the use of t-statistics to test for the population covariance of e_{1t} and $(e_{1t} - e_{2t})$. If we define $d_t = e_{1t}(e_{1t} - e_{2t})$, ENC test shows the same form as (1). Under the null hypothesis that the forecast of model 1 encompasses that of model 2, $E[d_t]$ will be less than or equal to 0. However, if the alternative hypothesis that the model 2 forecast contains more information holds true, then $E[d_t]$ should be positive.

As stated earlier, DM and ENC require an estimation of long-run variance for series d_t, since the optimal h-step forecast errors are $(h-1)$-dependent. DM test uses a non-parametric method of estimation, using the uniform kernel equipped with bandwidth $(h-1)$. Its estimator assumes the form $Var(\bar{d}) = \frac{1}{n}\sum_{k=-(h-1)}^{h-1} \hat{\gamma}_k$, where $\hat{\gamma}_k$ denotes the sample auto-covariance of d_t with lag k. Finally, under the null hypothesis, both test statistics follow a limiting standard normal distribution.

3. Autoregressive Testing Procedure

In this section, we have proposed an autoregressive procedure to construct both DM and ENC tests in (1). We will begin with the following example to illustrate the test idea. Supposing series d_t to be expressed as $d_t = c + \eta_t$, where c is a constant and η_t is iid$(0, \sigma^2{}_\eta)$, $t = 1, 2, ..., n$. As explained in Section 2, the extant tests are conducted to verify whether $E(d_t) = 0$, which is equivalent to testing the null hypothesis $H_0 : c = 0$; this can be done using t-statistics in the regression context. Under some general conditions, the t-statistics can be shown to have an asymptotic standard normal distribution.

In cases wherein series d_t is a covariance stationary process with possible serial correlations and conditional heteroskedasticity, we can adopt a parametric approach by fitting d_t with the $AR(p)$ model, where order p is determined by AIC, after which we test the $AR(p)$ model for zero intercept using t-statistics.

The AR testing procedure proposed by us has twofold advantages. First, it can be conducted in major standard software. Second, it prevents the final result from being affected by the selection of kernel function and bandwidth and therefore generates a more robust result than DM and ENC. The literature pertaining to unit root tests also documents that in the estimation of the long-run variance for finite samples, the parametric approach may produce a better performance than the non-parametrical one. Hence, our $AR-t_c$ test, which is a parametric approach, is expected to perform well in this aspect.

In order to obtain the asymptotic distribution for the t-statistics in the AR procedure, we require the following assumptions.

Assumption 1: Suppose $d_t = \mu_d + \psi(L)\varepsilon_t$, where μ_d is a constant, $\psi(L) = \sum_{j=0}^{\infty} \pi_j L^j$, $\pi_0 = 1$ and possesses the following properties:

1. $(\varepsilon_t, \Im_t)$ is a martingale difference sequence with some filtration $(\Im_t)$, satisfying (a) $E(\varepsilon_t^2) = \sigma^2$, (b) $n^{-1}\sum_{t=1}^{n} \varepsilon_t^2 = \sigma^2 + o_p(1)$ and (c) the existence of some $r \ge 4$ and $K > 0$ such that $E|\varepsilon_t|^r < K$.
2. $\forall |L| \le 1$, $\psi(L) \ne 0$ and $\exists s > 2$ such that $\sum_{j=0}^{\infty} j^s |\pi_j| < \infty$.

Assumption 1 allows d_t to be generated by means of a general linear process, including a finite order ARMA as a special case with π_k decaying geometrically; the assumption also ensures that d_t is covariance stationary. It should also be noted that Assumption 1(1) sets ε_t as a martingale difference sequence, which includes a variety of second order stationary ARCH and GARCH models that are widely used in empirical finance. Assumption 1(2) is commonly seen in literature; it can be satisfied by assuming appropriate mixing conditions, allowing for some types of serial correlations. Under Assumption 1(2), according to Bühlmann (1995), d_t can be written as follows.

$$d_t = \mu_d + \alpha(L)(d_t - \mu_d) + \varepsilon_t \, , \tag{2}$$

$$= \mu_d + \sum_{i=1}^{\infty} \alpha_i (d_{t-i} - \mu_d) + \varepsilon_t , \tag{3}$$

$$\equiv c + \sum_{i=1}^{\infty} \alpha_i d_{t-i} + \varepsilon_t , \tag{4}$$

where $1 - \alpha(L) = 1/\psi(L)$ and $\alpha(L) \neq 1$ for all $|L| \leq 1$. Moreover, there exists $s \geq 1$ such that $\sum_{j=1}^{\infty} j^s |\alpha_j| < \infty$ and $c = \mu_d (1 - \sum_{i=1}^{\infty} \alpha_i)$ in (4). Since $\sum_{i=1}^{\infty} \alpha_i \neq 1$, testing the null hypothesis of zero mean of d_t is equivalent to testing the following hypothesis.

$$H_0 : c = 0 . \tag{5}$$

As in the ADF test, we can approximate (4) by $AR(p)$ with order p increasing at a certain rate as the sample size increases. In empirical applications, LS is used to estimate an approximating $AR(p)$ model, and we obtain

$$d_t = \hat{c} + \sum_{k=1}^{p} \hat{\alpha}_{p,k} d_{t-k} + \hat{\varepsilon}_{p,t} , \tag{6}$$

where $\hat{c}$ and $\hat{\alpha}_{p,k}$ for $k = 1, 2, \ldots, p$ are the LS estimators, and $\hat{\varepsilon}_{p,t}$ is the LS residual. Correspondingly, the test statistics, denoted by $AR - t_c$, are formalized as follows.

$$AR - t_c = \frac{\hat{c}}{SE(\hat{c})} , \tag{7}$$

where $SE(\hat{c})$ is the standard error of $\hat{c}$.

In addition to Assumption 1, we also require the following assumption to obtain the asymptotics of $AR - t_c$.

Assumption 2: Let $p \to \infty$ and $p = o(n^{1/3})$ as $n \to \infty$.

Assumption 2 states that when approximating d_t with $AR(p)$, the order p should increase with n at a smaller rate than $n^{1/3}$. It should be noted that we do not impose any lower bound on the divergence rate for the order p. This condition is weaker than that in Said and Dickey (1984), in which they assume the lower rate of p to be $n^{1/r}$

for some $r > 0$ in order to ensure the accuracy of the approximation. Assumption 2 is merely required in theory; it does not indicate how to choose p in practical terms. Two common information criteria, such as AIC or BIC, are recommended. In such a forecasting context, however, Shibata (1980) claims that AIC may be a better choice than BIC because the former results in a more efficient asymptotic order estimate for some infinite-order autoregressive processes. Further, as shown by Park (2002), if we select the order p with AIC, the order p satisfies the condition $p = o(n^{-(1+s)})$ a.s. , and Assumption 2 holds if $s > 2$. Therefore, throughout this paper, we have used AIC to select the order p. The limit distribution of $AR - t_c$ is stated in Theorem 1.

Theorem 1: Let Assumption 1 and 2 hold. Under the null hypothesis (5), we have

$$AR - t_c \xrightarrow{d} N(0,1) . \tag{8}$$

Theorem 1 shows that the asymptotic distribution of $AR - t_c$ is standard normal, similar to those of DM and ENC, even when the error pertaining to d_t is subjected to a martingale difference sequence.

4. Bootstrap Asymptotics

As in DM and ENC tests, $AR - t_c$ suffers from size distortions at small samples with long forecast horizons. In this section, we intend to propose an appropriate bootstrap procedure, after taking into account both the dependence and heteroskedasticity of unknown form in the forecast errors, to estimate the empirical distribution of our test. Following this, we have provided a proof of bootstrap consistency to ensure the correctness of this procedure. Throughout this paper, the notation * denotes the bootstrap samples. Moreover, P^* and E^* denote the probability and expectation, respectively, that are conditional on the realization of the original sample. In this section, we will describe our resampling schemes in detail.

1. Suppose a sample $\{d_t\}_{t=1}^{n}$ is considered. Use LS to estimate the $AR(p)$ model with p chosen by AIC and obtain the LS residuals as follows.

$$\hat{\varepsilon}_{p,t} = d_t - \hat{c} - \sum_{k=1}^{p} \hat{\alpha}_{p,k} d_{t-k} \text{ , for } t = p+1, p+2, ..., n, \tag{9}$$

where $\hat{c}$ and $\hat{\alpha}_{p,k}$ for $k = 1, 2, \dots, p$ are LS estimates.

2. Generate WB residual according to

$$\hat{\varepsilon}^*_{p,t} = \hat{\varepsilon}_{p,t}\eta_t \text{, for } t = p+1, p+2, ..., n, \tag{10}$$

where η_t is any iid(0,1) random variable with a bounded fourth moment with respect to the probability measure P^*. We set η_t to be iid $N(0,1)$ in the follow simulations.

3. Generate a bootstrap sample d^*_t using autoregression,

$$d^*_t = \sum_{k=1}^{p} \hat{\alpha}_{p,k} d_{t-k} + \hat{\varepsilon}^*_{p,t} \text{, for } t = 1, 2, \dots, n, \tag{11}$$

where each $\hat{\alpha}_{p,k}$ is the LS estimate in (9) and $[d_1, ..., d_p]'$ serve as the initial values of d^*_t.

4. Compute the bootstrap counterpart of $AR - t_c$, denoted by $AR - t^*_c$, by regressing d^*_t on $[1, x'_{p,t}]$, where $x_{p,t} = [d_{t-1}, d_{t-2}, ..., d_{t-p}]'$.

5. Repeat Steps 2 to 4 NB times.

6. Compute the empirical distribution function (edf) of NB values of $AR - t^*_c$ and use this empirical distribution function as an approximation of the cumulative distribution function (cdf) of the bootstrap null distribution for the test statistics.

7. Make an inference based on the bootstrap critical value.

We will now discuss the bootstrap procedure. Step 1 involves an estimation of the

$AR(p)$ in order to account for the autocorrelation in d_t. It is also important to note that we have included an intercept in the model regardless of whether the data is from the null or alternative hypothesis. Step 2 generates WB innovations by multiplying the LS residuals from Step 1 with an iid N(0,1) series, which enables us to mimic the heteroskedasticity of unknown form possibly existing in the d_t series. This step is important for the sieve bootstrap to work efficiently in the next step, since the conditional heteroskedasticity in the original LS residuals does not meet the sieve bootstrap procedure's requirement of independent re-sampling with replacement. With the help of WB, and in accordance with Wu's (1986) claim, we may succeed in repairing the unknown structure of heteroskedasticity in the residuals, which would lead to good performance as regards size.

Step 3, suggested by Kreiss (1997), is called the fixed-design wild bootstrap, wherein a bootstrap sample d_t^* is generated by maintaining the regressors at the fixed value of their sample values in each re-sampling. It should be noted that in this step, we have imposed the condition of zero intercept on the generation of the bootstrap sample in order to be consistent with the null hypothesis of $AR-t_c$. In Step 4, the fixed-design regression is conducted by regressing the bootstrap sample on the fixed regressors; the bootstrap counterpart of $AR-t_c$ is subsequently calculated. Goncalves and Kilian (2003) have proved that the fixed-regressor WB is less restrictive than the conventional recursive bootstrap. In addition, we found that the sieve autoregression with fixed regressors can also successfully account for the serial correlation. Therefore, in this paper, we have merely considered the fixed-regressors scheme in Steps 3 and 4.

In theory, it is necessary to prove that the distribution of our bootstrap test is equivalent to the corresponding limit distribution of the original asymptotic test. This is because bootstrap distribution is used to approximate the unknown finite-sample distribution of $AR-t_c$. Therefore, if the bootstrap procedure is correct, the large-sample bootstrap distribution should be very 'close' to the limit distribution. Theorem 2 provides a proof for bootstrap consistency that ensures the correctness of our bootstrap procedure.

Theorem 2: Under Assumptions 1, 2 and 3 and the null hypothesis (6), we have

$$AR-t_c^{*} \xrightarrow{d*} N(0,1) \quad \text{in Prob.} \tag{12}$$

As shown in Theorem 2, the bootstrapped $AR-t_c^{*}$ has an asymptotic standard normal distribution. Additionally, the convergence holds for almost each d_t with respect to probability. As compared with Theorem 1, $AR-t_c$ and $AR-t_c^{*}$ have the same asymptotic distribution; this ensures the first-order asymptotic validity for the proposed bootstrap procedure.

Remark 1: Moving Block Bootstrap. It is well known that MBB is an alternative method of analysing dependent data, in which the bootstrap is implemented by dividing the data into blocks and randomly sampling the blocks with replacement. We will now describe it briefly. Let $B_i=(d_i,d_{i+1},...,d_{i+b-1})$ be the block of b consecutive observations starting from d_i, for $i=1, 2, \dots, n-b+1$. With these $n-b+1$ overlapping blocks, the bootstrap sample d_t^{*} can be obtained by randomly re-sampling n/b blocks with replacement and laying them end-to-end in the order in which they are sampled. The bootstrap version of DM and ENC tests can be easily obtained by using the standard bootstrap procedure. Some points regarding MBB need to be mentioned in this context. First, the testing results from MBB are sensitive to the choice of block length b. Although Hall and Jing (1996) have shown that although the optimal block length should be proportionate to $n^{1/4}$ for a one-sided test (e.g. ENC) and $n^{1/5}$ for a two-sided test (e.g. DM), this is not readily applicable in practice. In the following text, we will explore whether the choice of different block lengths has a severe impact on DM and ENC tests. Second, as implemented in Step 3, while conducting MBB, the bootstrap sample d_t^* continues to be generated with the imposition of the null hypothesis, without which the bootstrap test would have no power. This can be done purely by applying the MBB procedure to the de-mean series $\bar{d}_t$, instead of the original d_t series.

Remark 2: Throughout this paper, we have derived the asymptotics for the $AR-t_c$ test and its bootstrap version under the assumption that no estimation errors occur in the forecast errors. This technique was first utilised by DM (1995) in their test, after which Harevey, Leybourne and Newbold (1997, 1998) and Harvey and Newbold (2000) followed suit. In contrast, numerous papers in the related literature establish their predictive tests while taking into account parameter estimation errors (see West, 1996; West and McCracken, 1998; McCracken, 2000, 2006; Clark and McCracken, 2001). With the help of this assumption, we can pay more attention to the development of bootstrap asymptotics for our test. However, we are aware that this assumption is slightly stricter and is likely to lead to a limited empirical utility for the test. This is because in most applications, it is usually necessary to obtain the parameter estimation of a model before making the forecast. However, this limitation of our test is alleviated, as pointed out by West (2005), that the estimation errors can be ignored if the ratio of the numbers of out-of-sample observations to that of in-sample observations is less than 0.1.

5. Monte Carlo Analysis

5.1 Experimental Design

This section is devoted to investigating the finite-sample performance of our proposed test. Our simulation featured the use of bootstrap methods to provide a good control over size for DM, ENC and $AR-t_c$ tests. As a comparison, we applied MBB to both DM and ENC tests, denoted by DM* and ENC*. The sample size and forecast horizons were chosen to be $n=(16,\ 32,\ 64,\ 128,\ 256)$ and $h=(1,\ 2,\dots,\ 8)$. Rejection frequencies based on 5% asymptotic and bootstrap critical values were used to evaluate the test performance. Monte Carlo replications were 5,000 for the results of DM, ENC and $AR-t_c$; however, while conducting the bootstrap tests, both Monte Carlo and bootstrap replications were set to be 1,000 for the computer time saved. Meanwhile, the optimal lag p for $AR-t_c$ was chosen using AIC, with the maximum order set as 5, while the block length of MBB was set to be $[c_1 n^{1/5}]$ for DM* and $[c_1 n^{1/4}]$ for ENC*, wherein $c_1=1$, 3. In contrast, the long-run variance of DM and ENC was non-parametrically estimated by adopting the Bartlett kernel with the bandwidth set to h-1; there was no pre-whitening. The experimental designs were borrowed from Harvey et al. (1997, 1998), with some extensions to the GARCH-type and SV-type innovations driving the forecast error processes. GARCH and SV models are widely used in empirical finance due to their ability to account for some stylized facts in financial series, such as fat tail and variance clustering.

As seen in Harvey et al. (1997), the forecast errors (e_{1t}, e_{2t}) for DM, $AR-t_c$, $AR-t_c^*$ and DM* were generated by

$$\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{k} \end{bmatrix} \begin{bmatrix} \upsilon_{1t} \\ \upsilon_{2t} \end{bmatrix}. \tag{13}$$

While computing the type I error, we set $k=1$ to meet the null hypothesis of

$E(d_t)=0$. However, k was selected for each sample size in order to obtain powers that allowed meaningful comparisons. Hence, we set k = 2, 1.5, 1.375, 1.25 and 1.1875 to correspond to $n =$ 16, 32, 64, 128 and 256, respectively.

The experimental design for the forecast encompassing test was the same as in Harvey, Leybourne and Newbold (1998). To generate the forecast errors (e_{1t}, e_{2t}), let the covariance matrix be $R=\begin{bmatrix} 1 & \delta \\ \delta & \omega \end{bmatrix}$, $\text{cov}(e_{1t}, e_{2t})=\delta$ and $\text{var}(e_{2t})=\omega>\delta^2$. Thereafter, the forecast error can be generated by

$$\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ \delta & \sqrt{\omega-\delta^2} \end{bmatrix}\begin{bmatrix} \upsilon_{1t} \\ \upsilon_{2t} \end{bmatrix}, \tag{14}$$

where the first matrix on the right hand side of the equal sign is the Choleski factor for R. Since the value of ω is not likely to affect the simulation result without the loss of generality, ω is set to be 5. We let $\delta=1$ and 0.5 for size and power computations, respectively. Harvey, et al. (1998) pointed out that for any sample size, the powers of the tests depend solely on the single parameter $k=\sqrt{\omega-\delta^2}\big/1-\delta$. For any sample size, we choose k to be such that the size-adjusted powers are approximately 30%. In such a situation, the k of 3, 4.25, 6.25, 9, 12.75 must correspond to the n of 16, ... , 256, respectively.

To allow (e_{1t}, e_{2t}) in (13) and (14) to have MA property, we assume

$$\upsilon_{i,t}=\sum_{l=0}^{h-1}\pi_l\varepsilon_{i,t-l}, \quad i=1, 2, \quad h=1, 2, \ldots, 8, \tag{15}$$

with $(\pi_0, \pi_1, ..., \pi_7)=(1, 0.1, \text{-}\,0.1, 0.2, \text{-}\,0.2, 0.3, \text{-}\,0.3, 0.4)$. In addition, three types of $\varepsilon_{i,t}$ are considered below.

1. The standard normal distribution. That is, the innovations $\varepsilon_{i,t}$ are generated from N(0,1).

2. The GARCH process. That is,

$$\varepsilon_{i,t} = u_{i,t}\sqrt{h_{i,t}}\ , \tag{16}$$

$$h_{i,t} = w + ah_{i,t-1} + b\varepsilon_{i,t-1}^2\ , \tag{17}$$

where $u_{i,t}$ is standard normal, and $(w,a,b) = (0.15, 0.13, 0.2)$.

3. The SV process. To be specific,

$$\varepsilon_{i,t} = u_{i,t}\exp^{0.5\alpha_{i,t}}\ , \tag{18}$$

$$\alpha_{i,t} = \phi\alpha_{i,t-1} + \xi_{i,t-1}\ , \tag{19}$$

where $u_{i,t}$ and $\xi_{i,t-1}$ are $N(0,1)$, assumed to be serially uncorrelated and independent of each other. Moreover, ϕ is set to be 0.5.

The GARCH and SV specifications help us to investigate the usefulness of the $AR - t_c$ test, when applied to financial time series data.

5.2 Simulation Results

All the simulation results are displayed in Tables 1-1 to 4-3. The results presented in Tables 1-1 to 1-3 are of the predictive accuracy tests with the forecast errors generated by MA, GARCH and SV processes, respectively. We have also presented the size and power results based on 5% asymptotic and bootstrap critical values. It should be noted that the asymptotic powers reported are size-adjusted. Similarly, the results of the encompassing tests are presented in Tables 2-1 to 2-3. Moreover, in Tables 3-1 to 3-3 and 4-1 to 4-3, we have displayed the dependence of DM*or ENC* on the choice of block length to illustrate the superiority of sieve and wild bootstraps.

A number of conclusions emerge from the above-mentioned simulations. In Table 1-1, when forecast errors are homoskedastic, DM and $AR - t_c$ are revealed to be subject to substantial size distortions that increase along with h for a fixed n.

Moreover, as the sample sizes increase, each attains an actual size that is closer to the 5% nominal size. For example, for combinations of $(n, h) = (16, 1)$ and $(16, 8)$—which are of practical interest—the actual sizes for DM and $AR - t_c$ are 0.067, 0.363 and 0.349, 0.398, respectively. Moreover, we observe that for a fixed $h = 1$, as n increases from 16 to 256, the respective sizes of $AR - t_c$ are 0.349, 0.135, 0.083, 0.068 and 0.059, which confirms the validity of Theorem 1. The same observations can be made for the other horizons h as well.

Thereafter, we have adopted a bootstrap method to overcome the size distortions of the $AR - t_c$ test. From the entries labelled under 'bootstrap' in Table 1-1, two observations emerge. First, the bootstrap test exhibits an excellent control over the empirical sizes for cases where the sample size is small ($n = 16$) and a long horizon ($h = 8$), with size numbers of 0.349 and 0.060 corresponding to $AR - t_c$ and $AR - t_c^*$, respectively. Overall, the sizes for $AR - t_c^*$ are between 0.044 and 0.082, and they steadily vary, in contrast with those of $AR - t_c$ and DM. Second, the $AR - t_c^*$ test has bootstrap powers that are comparable or higher in some cases, as compared to its asymptotic size-adjusted counterpart. Additionally, the powers for both $AR - t_c$ and $AR - t_c^*$ increase as the sample sizes increase.

Similar observations regarding the size and power can also be made from Tables 1-2 and 1-3, where the forecast errors are generated from the process that reveals a moderately persistent, conditional hetroskedasticity. Both $AR - t_c$ and DM tests show a poor size performance when the sample size is varied between $n = 16$ and 32, along with longer horizons. Despite this, our bootstrap test shows an excellent performance in size and power. For example in Table 1-2, when $n = 16$, the sizes of $AR - t_c$ vary between 0.361 and 0.451. In contrast, the bootstrap sizes of $AR - t_c^*$

range between 0.053 and 0.073.

The results displayed in Tables 2-1 to 2-3 correspond to the encompassing tests. In general, the findings from in Tables 1-1 to 1-3 continue to hold for Tables 2-1 to 2-3, where moderate size distortions persist for the asymptotic tests, irrespective of the DGPs of the forecast errors. As expected, our proposed bootstrap scheme shows a good control over the size performance of $AR-t_c$ and exhibits a comparable bootstrap power in comparison with the asymptotic one.

To further investigate the manner in which the sizes of DM and ENC tests are affected by the choice of block length in MBB, we have changed them using the rule $b=c_1 n^{1/5}$ for DM and $b=c_1 n^{1/4}$ for ENC, with c_1 set to be 1 or 3. In addition, we have also considered three kinds of forecast errors. The results are presented in Tables 3-1 to 4-3. First, it is evident that on the whole, DM* and ENC* produce less accurate sizes than $AR-t_c^*$. This is particularly obvious when forecast errors are subject to a moderate persistent SV process in Table 3-3, where the sizes for DM* are not as good as expected even when $n=256$. Second, a change in block length has a substantial impact on the size performance of both tests. Considering all cases, on an average, assuming a shorter length ($c_1=1$) in MBB appears to produce a better size for both tests. However, these results cannot be treated as a practical rule solely on the basis of our limited simulations. Based on these facts pertaining to MBB, in this paper, we have recommended a practical guideline for the evaluation of forecast performance by combining the $AR-t_c$ test with sieve and wild bootstraps.

6. Application

In this section, using data from Taiwan, we will illustrate the practical use of our $AR-t_c$ test as an application to exchange-rate forecasting. The data set comprises monthly observations of spot and 1-, 2-, 3-, 4- and 6-month forward NTD/USD exchange rates for the period from 1992:1 to 2003:12, which makes a total of 114 observations for each series. The data is sourced from the AREMOS database of the Ministry of Education, Taiwan.

In accordance with the no-parameter estimation error assumption, we will attempt the *h*-horizon forecast of the change in nominal NTD/USD spot rates by directly calculating the difference between the *h*-month forward rate and the spot rate, $h = 1, 2, 3, 4, 6$. This is implemented based on the theory that the forward exchange rate would be an unbiased predictor of the corresponding future spot rate, provided the forward market efficiency hypothesis holds. We have also chosen a random walk model as a benchmark comparison, the forecast of which is constant at 0. The results are presented in Table 5.

The evidence in the second column of Table 5 reveals that as regards predictive accuracy, there is a substantial difference between the asymptotic test and the bootstrap test. The bootstrap *p*-values of $AR-t_c$ universally reject the null hypothesis at the 5% level of significance, which indicates that the forward market forecast is more accurate than the random walk one. The rejection of the null hypothesis, however, is only found in two out of five cases for the asymptotic $AR-t_c$ test. Similar but weak results are obtained from DM test, wherein the superiority of the forward market occurs for almost all the cases of the bootstrap DM test, but rarely for the DM test. In addition, the results in the third column of Table 5 show that no matter which test is conducted, both the encompassing tests fail to reject the null hypothesis, indicating that the forward rate model contains more useful information than the base model. This finding reinforces the conclusion that the forward market model outperforms the random walk in predicting the change in the NTD/USD spot rate.

7. Conclusions

In this paper, we have developed a simple and computationally feasible method for discriminating between two competing forecasting models, using the assumption that the forecast errors may be serially correlated and heteroskedastic. The proposed method relies on an autoregressive model that deals with the serial correlation by adding extra lags, which enables us to overcome the difficulty in choosing an appropriate kernel function for the estimation of long-run variance. In addition, it can be used to formulate both DM and ENC tests. Our Theorem 1 shows that the $AR-t_c$ test has an asymptotic standard normal distribution. Not surprisingly, the test is over-sized for small sample sizes or long forecast horizons. In order to overcome the problem regarding small samples, we proposed a hybrid bootstrap procedure consisting of an autoregressive sieve and WB. We also provided the proof of consistency for our hybrid procedure. A simulation study was performed to identify the finite sample properties of our method, and the results were compared with those of two other tests adopting MBB. We found that our bootstrap method performs well in controlling the size distortions, as compared to the other two tests. In conclusion, we demonstrated the empirical relevance of our proposed test by applying it to forecasting the exchange rate.

Appendix: Mathematical Proofs

In what follows, we use the notation $\|\cdot\|$ to signify the usual Euclidean norm. We define $\| x \| = \sqrt{x_1^2 + \cdots + x_n^2}$ for a p-vector x, and let $\|A\| = \max_x \|Ax\| / \|x\|$ for a $p \times p$ matrix A. Plus, any positive number is denoted by K without ambiguity.

Lemma 1: Under Assumption 1 and 2 along with the null hypothesis of (5), we have for large n

$$(a) \quad \sum_{k=p+1}^{\infty} |a_k| = o(p^{-s})$$

$$(b) \quad \frac{1}{n}\sum_{t=1}^{n} \varepsilon_{p,t}^2 = \frac{1}{n}\sum_{t=1}^{n} \varepsilon_t^2 + o_p(p^{-s})$$

$$(c) \quad \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \varepsilon_{p,t} = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \varepsilon_t + o_p(1)$$

$$(d) \quad \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \varepsilon_{p,t} \xrightarrow{d} N(0, \sigma^2)$$

Proof of Lemma 1 For part (a) and (b) respectively, see Chang and Park (2002) p.434 and Lemma 3.1.(c).

From Chang and Park (2002) p.441, we have

$$\sum_{t=1}^{n} (\varepsilon_{p,t} - \varepsilon_t) = \sum_{k=p+1}^{\infty} \pi_{p,k} \sum_{t=1}^{n} \varepsilon_{t-k} = o_p(n^{1/2} p^{-s})$$

where $\varepsilon_{p,t} = \varepsilon_t + \sum_{k=p+1}^{\infty} \alpha_k d_{t-k}$. Therefore part (c) is done. Under Assumption 1 by CLT for martingale difference sequence (see, e.g., White (1984), Corollary 5.25), we have $\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \varepsilon_t \xrightarrow{d} N(0, \sigma^2)$. Along with (c), we have (d).

Lemma 2: Let $x_{p,t} = [d_{t-1}, d_{t-2}, \cdots, d_{t-p}]'$. Under the same conditions of Lemma 1, we have for large n

$$(a) \quad \left\| \left(\frac{1}{n}\sum_{t=1}^{n} x_{p,t} x_{p,t}'\right)^{-1} \right\| = O_p(1)$$

$$(b) \quad \left\| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} x_{p,t} \varepsilon_{p,t} \right\| = o_p(n^{1/2} p^{-1/2})$$

$$(c) \quad \left\| \frac{1}{n} \sum_{t=1}^{n} x_{p,t} \right\| = O_p(n^{-1/2} p^{1/2})$$

Proof of Lemma 2 For (a) and (b), refer to (a) and (c) of Lemma 3.2 respectively in Chang and Park (2002). To show the result of part (c), first note that we have

$$E(\sum_{t=1}^{n} (d_{t-i} d_{t-j} - \Gamma_{i-j}))^2 = O(n)$$

uniformly in i and j. See Change and Park (2002) p.443. This implies $E(\frac{1}{n} \sum_{t=1}^{n} d_{t-i})^2 = O(n^{-1})$. Therefore, we have $\left\| \frac{1}{n} \sum_{t=1}^{n} x_{p,t} \right\| = O_p(n^{-1/2} p^{1/2})$.

Lemma 3: Let $\hat{c}$ and $\hat{\beta}_p = [\hat{\alpha}_{p,1}, \hat{\alpha}_{p,2}, \cdots, \hat{\alpha}_{p,p}]'$ denote the LS estimators of c and $\beta_p = [\alpha_1, \alpha_2, \cdots, \alpha_p]'$ respectively in (6). Also, let $\hat{\varepsilon}_{p,t}$ be LS residuals. Define $x_{p,t} = [d_{t-1}, d_{t-2}, \cdots, d_{t-p}]'$ and $\hat{\sigma}_n^2 = n^{-1} \sum_{t=1}^{n} \hat{\varepsilon}_{p,t}^2$. Under the same conditions of Lemma 1, we have

$$(a) \quad \sqrt{n}\hat{c} = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \varepsilon_{p,t} + o_p(1)$$

$$(b) \quad \left\| \hat{\beta}_p - \beta_p \right\| = o_p(p^{-1/2})$$

$$(c) \quad \hat{\sigma}_n^2 = \sigma^2 + o_p(1)$$

$$(d) \quad \sqrt{n}\hat{c} \xrightarrow{d} N(0, \sigma^2)$$

Proof of Lemma 3 We first prove (d). Lemma 3(d) holds by Lemma 3(a) and Lemma 1(d).

Let $F = diag(n^{-1/2}, n^{-1/2} I_p)$. Since $\hat{c}$ and $\hat{\beta}_p$ respectively in (6), we have

$$[\sqrt{n}\hat{c}, \sqrt{n}(\hat{\beta}_p - \beta_p)']' = (\sum_{t=1}^{n} F[1, x'_{p,t}]'[1, x'_{p,t}]F)^{-1} (\sum_{t=1}^{n} F[1, x'_{p,t}]' \varepsilon_{p,t}) . \tag{E1}$$

Define A_n, B_n and C_n as follows.

$$A_n = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \varepsilon_{p,t} - (\frac{1}{n} \sum_{t=1}^{n} x'_{p,t})(\frac{1}{n} \sum_{t=1}^{n} x_{p,t} x'_{p,t})^{-1} (\frac{1}{\sqrt{n}} \sum_{t=1}^{n} x_{p,t} \varepsilon_{p,t})$$

$$B_n = 1 - (\frac{1}{n} \sum_{t=1}^{n} x'_{p,t})(\frac{1}{n} \sum_{t=1}^{n} x_{p,t} x'_{p,t})^{-1} (\frac{1}{n} \sum_{t=1}^{n} x_{p,t})$$

$$C_n = (\frac{1}{n}\sum_{t=1}^{n} x_{p,t}x'_{p,t})^{-1}(\frac{1}{\sqrt{n}}\sum_{t=1}^{n} x_{p,t}\varepsilon_{p,t}) - (\frac{1}{n}\sum_{t=1}^{n} x_{p,t})(\frac{1}{n}\sum_{t=1}^{n} x_{p,t}x'_{p,t})^{-1}(\frac{1}{\sqrt{n}}\sum_{t=1}^{n} \varepsilon_{p,t})$$

From (E1), the following equations hold.

$$\sqrt{n}\hat{c} = \frac{A_n}{B_n}$$

$$\sqrt{n}(\hat{\beta}_p - \beta_p) = \frac{C_n}{B_n}$$

With Lemma 1(d) and Lemma 2, we have the following results.

$$B_n = 1 + o_p(1) \tag{20}$$

$$\sqrt{n}\hat{c} = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \varepsilon_{p,t} + o_p(1) \tag{21}$$

$$\left\|\hat{\beta}_p - \beta_p\right\| = o_p(p^{-1/2}) \tag{22}$$

Therefore, we have part (a) and (b). To show the result of part (c), plug $\hat{\varepsilon}_{p,t} = \varepsilon_{p,t} - \hat{c} - (\hat{\beta}_p - \beta_p)' x_{p,t}$ into $\hat{\sigma}_n^2 = n^{-1}\sum_{t=1}^{n} \hat{\varepsilon}_{p,t}^2$. And along with Lemma 2, Lemma 3(a), (b) and (d), we have

$$\frac{1}{n}\sum_{t=1}^{n} \hat{\varepsilon}_{p,t}^2 = \frac{1}{n}\sum_{t=1}^{n} \varepsilon_{p,t}^2 + o_p(p^{-1}). \tag{23}$$

Clearly, with (23) and Lemma 1(b), we have

$$\begin{aligned}
\left|\frac{1}{n}\sum_{t=1}^{n} \hat{\varepsilon}_{p,t}^2 - \frac{1}{n}\sum_{t=1}^{n} \varepsilon_t^2\right| &\leq \left|\frac{1}{n}\sum_{t=1}^{n} \hat{\varepsilon}_{p,t}^2 - \frac{1}{n}\sum_{t=1}^{n} \varepsilon_{p,t}^2\right| + \left|\frac{1}{n}\sum_{t=1}^{n} \varepsilon_{p,t}^2 - \frac{1}{n}\sum_{t=1}^{n} \varepsilon_t^2\right| \\
&= o_p(p^{-1}) + o_p(p^{-s}) \\
&= o_p(p^{-1}). \qquad \text{(E2)}
\end{aligned}$$

Hence the proof of Lemma 3(c) is complete by (E2) and Assumption 1(b).

Proof of Theorem 1 First note that, by its definition, $AR - t_c$ can be written as

$$\begin{aligned}
AR - t_c &\equiv \frac{\hat{c}}{SE(\hat{c})} \\
&= \frac{\sqrt{n}\hat{c}}{\hat{\sigma}_n\sqrt{B_n}}
\end{aligned}$$

Then all together with Lemma 3(c), Lemma 3(d) and (21), the proof of Theorem 1 is complete.

Lemma 4: Let $\hat{c}^*$ and $\hat{\beta}_p^* = [\hat{\alpha}_{p,1}^*, \hat{\alpha}_{p,2}^*, \cdots, \hat{\alpha}_{p,p}^*]'$ denote the LS bootstrap estimators and $\hat{\upsilon}_{p,t}^*$ be the residuals in Step 4 of the proposed bootstrap procedure. Define $\hat{\sigma}_n^{*2} = n^{-1}\sum_{t=1}^n \hat{\upsilon}_{p,t}^{*2}$. Plus, $\hat{\beta}_p$ and $x_{p,t}$ are defined in Lemma 3. Under the same conditions of Lemma 1, we have

(a) $\left\| \frac{1}{\sqrt{n}} \sum_{t=1}^n x_{p,t} \hat{\varepsilon}_{p,t}^* \right\| = O_{p*}(p^{1/2})$ in Prob.

(b) $\sqrt{n}\hat{c}^* = \frac{1}{\sqrt{n}} \sum_{t=1}^n \hat{\varepsilon}_{p,t}^* + o_{p*}(1)$ in Prob.

(c) $\left\| \hat{\beta}_p^* - \hat{\beta}_p \right\| = O_{p*}(n^{-1/2} p^{1/2})$ in Prob.

(d) $\hat{\sigma}_n^{*2} = \sigma^2 + o_{p*}(1)$ in Prob.

(e) $\frac{1}{\sqrt{n}} \sum_{t=1}^n \hat{\varepsilon}_{p,t}^* \xrightarrow{d*} N(0, \sigma^2)$ in Prob.

Proof of Lemma 4 To show part (a), first note that

$$
\begin{aligned}
E^*\left(\left\| \frac{1}{\sqrt{n}} \sum_{t=1}^n x_{p,t} \hat{\varepsilon}_{p,t}^* \right\|^2 \right) &= E^*\left(\frac{1}{n} \sum_{t=1}^n \sum_{s=1}^n x'_{p,t} x_{p,s} \hat{\varepsilon}_{p,t}^* \hat{\varepsilon}_{p,s}^*\right) \\
&= \frac{1}{n} \sum_{t=1}^n \sum_{s=1}^n E^*(x'_{p,t} x_{p,s} \hat{\varepsilon}_{p,t} \eta_t \hat{\varepsilon}_{p,s} \eta_s) \\
&= \frac{1}{n} \sum_{t=1}^n \sum_{s=1}^n x'_{p,t} x_{p,s} \hat{\varepsilon}_{p,t} \hat{\varepsilon}_{p,s} E^*(\eta_t \eta_s) \\
&= \frac{1}{n} \sum_{t=1}^n x'_{p,t} x_{p,s} \hat{\varepsilon}_{p,t}^2 \qquad \text{(E3)} \\
&= \frac{1}{n} \sum_{t=1}^n \left| x'_{p,t} x_{p,s} \hat{\varepsilon}_{p,t}^2 \right| \\
&\le KO(p). \qquad \text{(E4)}
\end{aligned}
$$

Clearly, if both (E3) and (E4) hold, part (a) is complete. The equation (E3) holds since

$$
E^*(\eta_t \eta_s) = \begin{cases} 1 & if \quad t = s \\ 0 & if \quad t \neq s \end{cases}
$$

by the definition of η_t. Now prove (E4). Bye the Cauchy-Schwartz inequality, we have that

$$
\frac{1}{n} \left| \sum_{t=1}^n x'_{p,t} x_{p,t} \hat{\varepsilon}_{p,t}^2 \right| \le \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n \left\| x_{p,t} \right\|^4\right)^{1/2} \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n \hat{\varepsilon}_{p,t}^4\right)^{1/2} \equiv B_1 + B_2 .
$$

Note that $B_1 = O_p(p^2)$, because $E|d_t|^r$ is bounded uniformly in t. Now show that $B_2 = O_p(1)$. Note that $\hat{\varepsilon}_{p,t} = \varepsilon_t + \sum_{j=1}^{\infty} \alpha_j d_{t-j} - \hat{c} - (\hat{\beta}_p - \beta_p)' x_{p,t}$. By the c_γ-inequality, we have that

$$\begin{aligned} \frac{1}{n}\sum_{t=1}^{n} \hat{\varepsilon}_{p,t}^4 &\leq K(n^{-1}\sum_{t=1}^{n} \varepsilon_t^4 + n^{-1}\sum_{t=1}^{n}(\sum_{j=p+1}^{\infty} \alpha_j d_{t-j})^4 + \hat{c}^4 + n^{-1}\sum_{t=1}^{n}((\hat{\beta}_p - \beta_p)' x_{p,t})^4) \\ &\equiv A_1 + A_2 + A_3 + A_4 . \end{aligned}$$

By Assumption 1, we have $A_1 = O_p(1)$. By Minkowski's inequality, we have $A_2 = o_p(p^{-4s})$. We have $A_3 = O_p(n^{-2})$, by Lemma 4(a). Also, $A_4 = o_p(1)$ because $\|\hat{\beta}_p - \beta_p\|^4 = o_p(p^{-2})$ (Lemma 4(c)) and $E\|x_{p,t}\|^4 = O(p^2)$. Therefore, $B_1 = O_p(1)$.

Now we turn to part (e). We apply Lyapunov's theorem to prove part (e). Define $X_{nt}^* = \hat{\varepsilon}_{p,t}^* / (\sqrt{n}\sigma)$ and $S_n^* = \sum_{t=1}^{n} X_{nt}^*$. Clearly, X_{nt}^* is an independent sequence of random variables by construction. Plus, $E^* X_{nt}^* = 0$ and $\sigma_{nt}^{*2} \equiv E^* X_{nt}^{*2} = \hat{\varepsilon}_{p,t}^2 / (n\sigma^2)$ for all t. Note also that $s_n^{*2} \equiv \sum_{t=1}^{n} \sigma_{nt}^{*2} = (n\sigma^2)^{-1} \sum_{t=1}^{n} \hat{\varepsilon}_{p,t}^2 = 1 + o_p(1)$ by Lemma 1(b). To apply Lyapunov's theorem, we need to verify the following condition

$$\sum_{t=1}^{n} E^* |X_{nt}^*|^{2r} \longrightarrow 0 \text{ in Prob. For some } r > 1$$

Choose $r = 2$, we have $\sum_{t=1}^{n} E^* |X_{nt}^*|^4 = n^{-2} \sum_{t=1}^{n} \hat{\varepsilon}_{p,t}^4$. Recall that we have $n^{-1}\sum_{t=1}^{n} \hat{\varepsilon}_{p,t}^4 = O_p(1)$ in the proof of Lemma 4(a). Clearly, this condition is satisfied with $r = 2$. By Lyapunov's theorem, we have

$$\begin{aligned} \frac{S_n^*}{s_n^*} &= (\frac{\sigma}{\hat{\sigma}_n})(\frac{n^{-1/2}\sum_{t=1}^{n} \hat{\varepsilon}_{p,t}^4}{\sigma}) \\ &\xrightarrow{d*} N(0,1) \text{ in Prob.} \end{aligned}$$

Therefore, $n^{-1/2}\sum_{t=1}^{n} \hat{\varepsilon}_{p,t}^4 \xrightarrow{d*} N(0, \sigma^2)$ in Prob. by Lemma 3(c).

To show the results in part (b) and (c), we define

$$A_n^* = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\hat{\varepsilon}_{p,t}^* - (\frac{1}{n}\sum_{t=1}^{n}x'_{p,t})(\frac{1}{n}\sum_{t=1}^{n}x_{p,t}x'_{p,t})^{-1}(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}x_{p,t}\hat{\varepsilon}_{p,t}^*)$$

$$B_n^* = 1-(\frac{1}{n}\sum_{t=1}^{n}x'_{p,t})(\frac{1}{n}\sum_{t=1}^{n}x_{p,t}x'_{p,t})^{-1}(\frac{1}{n}\sum_{t=1}^{n}x_{p,t}) = B_n$$

$$C_n^* = (\frac{1}{n}\sum_{t=1}^{n}x_{p,t}x'_{p,t})^{-1}(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}x_{p,t}\hat{\varepsilon}_{p,t}^*) - (\frac{1}{n}\sum_{t=1}^{n}x_{p,t})(\frac{1}{n}\sum_{t=1}^{n}x_{p,t}x'_{p,t})^{-1}(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\hat{\varepsilon}_{p,t}^*)$$

Then by the definition of $\hat{c}^*$ and $\hat{\beta}_p^*$, we have

$$\sqrt{n}\hat{c}^* = \frac{A_n^*}{B_n^*}$$

$$\sqrt{n}(\hat{\beta}_p^* - \hat{\beta}_p) = \frac{C_n^*}{B_n^*}$$

By Lemma 2(a), 2(c) and Lemma 4(a), 4(e) we have

$$A_n^* = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\hat{\varepsilon}_{p,t}^* + o_{p^*}(1) \text{ in Prob.}$$

$$B_n^* = 1 + o_{p^*}(1) \text{ in Prob.}$$

$$C_n^* = O_{p^*}(p^{1/2}) \text{ in Prob.}$$

Therefore, the results of (b) and (c) hold.

To prove (d), first note that $\hat{\sigma}_n^{*2} = n^{-1}\sum_{t=1}^{n}\hat{\upsilon}_{p,t}^{*2}$ where $\hat{\upsilon}_{p,t}^{*2} = \hat{\varepsilon}_{p,t}^* - \hat{c}^* - (\hat{\beta}_p^* - \hat{\beta}_p)'x_{p,t}$. Expand the above equation and along with Lemma 4, we can get

$$\hat{\sigma}_n^{*2} = \frac{1}{n}\sum_{t=1}^{n}\hat{\varepsilon}_{p,t}^{*2} + o_{p^*}(1) \text{ in Prob.}$$

If we can show

$$\frac{1}{n}\sum_{t=1}^{n}\hat{\varepsilon}_{p,t}^{*2} = \hat{\sigma}_n^2 + o_{p^*}(1) \text{ in Prob.,} \qquad \text{(E5)}$$

then part (e) is complete by Lemma 3(c). Now, we prove (E5). Let

$$\bar{X}_n^* = n^{-1}\sum_{t=1}^{n}\hat{\varepsilon}_{p,t}^{*2} - n^{-1}\sum_{t=1}^{n}\hat{\varepsilon}_{p,t}^{2} \equiv \frac{1}{n}\sum_{t=1}^{n}X_t^*$$

where $X_t^* = \hat{\varepsilon}_{p,t}^2(\eta_t^2 - 1)$. Clearly, $E^*(X_t^*) = 0$ for all t and $V^*(X_t^*) = K\hat{\varepsilon}_{p,t}^4$. With $\frac{1}{n}\sum_{t=1}^{n}\hat{\varepsilon}_{p,t}^4 = O_p(1)$ at hand, we have

$$
\begin{aligned}
\frac{1}{n}\bar{\sigma}^{*2} &\equiv \frac{1}{n^2}\sum_{t=1}^{n} V^*(X_t^*) \\
&= K\frac{1}{n^2}\sum_{t=1}^{n} \hat{\varepsilon}_{p,t}^4 \\
&\to 0 \quad \text{in Prob.}
\end{aligned}
$$

By Chebyechev's weak law of large numbers, $\bar{X}_n^* = o_{p^*}(1)$ in Prob.. Therefore, (E5) is done.

Proof of Theorem 2 First note that $AR-t_c^*$ can be expressed as

$$
AR-t_c^* = \frac{\sqrt{n}\hat{c}^*}{\hat{\sigma}_n^{*2}\sqrt{B_n}}.
$$

With (20), Lemma 4(b), 4(d) and 4(e) all together, Theorem 2 is trivial.

Table 1-1: Small sample performance of DM and $AR-t_c$ -- MA($h-1$) error

			$n=16$		$n=32$		$n=64$		$n=128$		$n=256$	
			$AR-t_c$	DM	$AR-t_c$	DM	$AR-t_c$	DM	$AR-t_c$	DM	$AR-t_c$	DM
$h=1$	asymptotic	size	0.349	0.067	0.135	0.059	0.083	0.05	0.068	0.053	0.059	0.048
		power	0.076	0.226	0.122	0.178	0.185	0.248	0.209	0.239	0.249	0.277
	bootstrap	size	0.060		0.064		0.055		0.055		0.054	
		power	0.076		0.129		0.175		0.227		0.277	
$h=2$	asymptotic	size	0.337	0.123	0.129	0.086	0.083	0.07	0.069	0.059	0.058	0.056
		power	0.072	0.161	0.117	0.158	0.194	0.229	0.209	0.231	0.247	0.253
	bootstrap	size	0.053		0.061		0.066		0.054		0.056	
		power	0.070		0.122		0.179		0.226		0.274	
$h=3$	asymptotic	size	0.351	0.177	0.138	0.125	0.083	0.083	0.065	0.063	0.059	0.056
		power	0.073	0.133	0.107	0.131	0.189	0.21	0.208	0.217	0.252	0.259
	bootstrap	size	0.058		0.059		0.051		0.046		0.048	
		power	0.073		0.120		0.172		0.205		0.259	
$h=4$	asymptotic	size	0.36	0.21	0.14	0.148	0.088	0.099	0.069	0.074	0.064	0.06
		power	0.071	0.128	0.122	0.118	0.17	0.168	0.196	0.206	0.227	0.236
	bootstrap	size	0.058		0.052		0.061		0.052		0.044	
		power	0.069		0.110		0.159		0.187		0.241	
$h=5$	asymptotic	size	0.357	0.245	0.139	0.167	0.093	0.113	0.075	0.079	0.065	0.064
		power	0.071	0.115	0.12	0.115	0.176	0.152	0.175	0.176	0.216	0.215
	bootstrap	size	0.064		0.072		0.056		0.046		0.059	
		power	0.076		0.123		0.168		0.195		0.243	
$h=6$	asymptotic	size	0.367	0.27	0.145	0.181	0.1	0.12	0.086	0.089	0.072	0.065
		power	0.072	0.101	0.107	0.103	0.167	0.146	0.148	0.156	0.198	0.199
	bootstrap	size	0.052		0.058		0.062		0.057		0.060	
		power	0.079		0.115		0.155		0.189		0.214	
$h=7$	asymptotic	size	0.393	0.321	0.172	0.206	0.112	0.134	0.09	0.09	0.079	0.065
		power	0.065	0.089	0.097	0.093	0.141	0.13	0.149	0.148	0.174	0.178
	bootstrap	size	0.072		0.073		0.072		0.063		0.067	
		power	0.065		0.113		0.169		0.173		0.205	
$h=8$	asymptotic	size	0.398	0.363	0.177	0.221	0.12	0.143	0.103	0.091	0.089	0.07
		power	0.065	0.094	0.09	0.084	0.123	0.113	0.128	0.139	0.148	0.17
	bootstrap	size	0.072		0.082		0.082		0.067		0.075	
		power	0.076		0.135		0.165		0.174		0.217	

Note:

1. The data generation process (DGP) is $\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{k} \end{bmatrix} \begin{bmatrix} \upsilon_{1t} \\ \upsilon_{2t} \end{bmatrix}$, where $\upsilon_{i,t} = \sum_{l=0}^{h-1} \pi_l \varepsilon_{i,t-l}$, $i=1,2$. $h=1$, …, 8 with $(\pi_0, \pi_1, ..., \pi_7) = (1, 0.1, -0.1, 0.2, -0.2, 0.3, -0.3, 0.4)$, and $\varepsilon_{i,t} \overset{iid}{\sim} N(0,1)$. Also, we set $k=1$ for size, and $k = 2$, 1.5, 1.375, 1.25, 1.1875 for power, respectively.
2. For other information, please refer to section 5.1.

Table 1-2: Small sample performance of DM and $AR-t_c$ -- MA($h-1$)+GARCH(1,1) error

			$n=16$		$n=32$		$n=64$		$n=128$		$n=256$	
			$AR-t_c$	DM	$AR-t_c$	DM	$AR-t_c$	DM	$AR-t_c$	DM	$AR-t_c$	DM
$h=1$	asymptotic	size	0.368	0.134	0.147	0.139	0.092	0.141	0.073	0.146	0.062	0.146
		power	0.072	0.15	0.093	0.126	0.132	0.156	0.136	0.139	0.151	0.157
	bootstrap	size	0.067		0.068		0.066		0.069		0.066	
		power	0.074		0.100		0.134		0.169		0.187	
$h=2$	asymptotic	size	0.361	0.148	0.143	0.109	0.089	0.097	0.071	0.089	0.06	0.083
		power	0.069	0.118	0.089	0.118	0.134	0.156	0.127	0.145	0.148	0.149
	bootstrap	size	0.053		0.075		0.076		0.071		0.075	
		power	0.070		0.106		0.153		0.163		0.186	
$h=3$	asymptotic	size	0.374	0.184	0.149	0.127	0.094	0.087	0.071	0.075	0.062	0.068
		power	0.061	0.111	0.095	0.11	0.131	0.148	0.13	0.138	0.145	0.159
	bootstrap	size	0.054		0.062		0.058		0.062		0.067	
		power	0.074		0.105		0.139		0.154		0.178	
$h=4$	asymptotic	size	0.379	0.219	0.154	0.148	0.094	0.096	0.072	0.074	0.061	0.065
		power	0.062	0.097	0.094	0.102	0.108	0.127	0.131	0.146	0.14	0.144
	bootstrap	size	0.061		0.056		0.066		0.064		0.050	
		power	0.067		0.094		0.122		0.156		0.166	
$h=5$	asymptotic	size	0.382	0.257	0.152	0.166	0.101	0.111	0.074	0.077	0.064	0.067
		power	0.065	0.096	0.1	0.1	0.112	0.108	0.121	0.122	0.136	0.133
	bootstrap	size	0.062		0.087		0.065		0.061		0.070	
		power	0.075		0.108		0.136		0.159		0.165	
$h=6$	asymptotic	size	0.389	0.293	0.16	0.188	0.105	0.122	0.083	0.084	0.066	0.067
		power	0.067	0.088	0.087	0.083	0.119	0.116	0.107	0.112	0.13	0.139
	bootstrap	size	0.057		0.071		0.080		0.068		0.067	
		power	0.070		0.108		0.126		0.149		0.157	
$h=7$	asymptotic	size	0.409	0.347	0.179	0.21	0.116	0.131	0.087	0.088	0.07	0.069
		power	0.063	0.076	0.084	0.079	0.103	0.109	0.104	0.114	0.12	0.123
	bootstrap	size	0.073		0.076		0.084		0.070		0.070	
		power	0.071		0.101		0.150		0.143		0.154	
$h=8$	asymptotic	size	0.415	0.394	0.188	0.232	0.129	0.142	0.099	0.092	0.082	0.073
		power	0.057	0.082	0.079	0.075	0.096	0.095	0.101	0.11	0.118	0.123
	bootstrap	size	0.061		0.088		0.086		0.074		0.069	
		power	0.083		0.128		0.156		0.152		0.169	

Note:

1. The DGP is similar to Note 1 in Table 1-1, except the innovations are generated by $\varepsilon_{i,t}=u_{i,t}\sqrt{h_{i,t}}$, $h_{i,t}=w+ah_{i,t-1}+b\varepsilon_{i,t-1}^2$, where $u_{i,t}$ is standard normal, and $(w,a,b)=(0.15,0.13,0.2)$.
2. For other information, please refer to section 5.1.

Table 1-3: Small sample performance of DM and $AR-t_c$ - MA($h-1$) +SV error

			$n=16$		$n=32$		$n=64$		$n=128$		$n=256$	
			$AR-t_c$	DM	$AR-t_c$	DM	$AR-t_c$	DM	$AR-t_c$	DM	$AR-t_c$	DM
$h=1$	asymptotic	size	0.317	0.094	0.104	0.091	0.076	0.092	0.057	0.087	0.058	0.089
		power	0.052	0.088	0.071	0.078	0.079	0.092	0.087	0.086	0.086	0.092
	bootstrap	size	0.043		0.056		0.056		0.057		0.062	
		power	0.051		0.069		0.096		0.094		0.114	
$h=2$	asymptotic	size	0.314	0.125	0.103	0.084	0.072	0.075	0.058	0.061	0.056	0.064
		power	0.054	0.083	0.067	0.084	0.081	0.091	0.091	0.094	0.094	0.096
	bootstrap	size	0.047		0.045		0.053		0.071		0.073	
		power	0.057		0.067		0.086		0.102		0.108	
$h=3$	asymptotic	size	0.316	0.158	0.103	0.1	0.073	0.074	0.07	0.067	0.06	0.061
		power	0.054	0.077	0.068	0.076	0.08	0.082	0.08	0.08	0.082	0.087
	bootstrap	size	0.055		0.054		0.059		0.062		0.067	
		power	0.059		0.071		0.096		0.088		0.102	
$h=4$	asymptotic	size	0.317	0.214	0.113	0.126	0.082	0.084	0.065	0.062	0.063	0.059
		power	0.057	0.07	0.068	0.068	0.08	0.081	0.087	0.086	0.087	0.088
	bootstrap	size	0.045		0.051		0.066		0.061		0.076	
		power	0.046		0.068		0.100		0.100		0.114	
$h=5$	asymptotic	size	0.327	0.249	0.117	0.144	0.086	0.096	0.07	0.062	0.065	0.058
		power	0.059	0.07	0.066	0.07	0.083	0.083	0.078	0.08	0.089	0.088
	bootstrap	size	0.048		0.060		0.054		0.057		0.084	
		power	0.049		0.078		0.084		0.103		0.125	
$h=6$	asymptotic	size	0.352	0.3	0.139	0.178	0.097	0.103	0.079	0.072	0.07	0.062
		power	0.057	0.064	0.067	0.064	0.082	0.079	0.077	0.082	0.083	0.081
	bootstrap	size	0.054		0.059		0.073		0.087		0.070	
		power	0.051		0.083		0.100		0.118		0.097	
$h=7$	asymptotic	size	0.376	0.36	0.157	0.199	0.118	0.121	0.096	0.077	0.076	0.055
		power	0.055	0.068	0.064	0.067	0.077	0.077	0.077	0.081	0.082	0.094
	bootstrap	size	0.054		0.075		0.067		0.083		0.089	
		power	0.062		0.086		0.104		0.125		0.132	
$h=8$	asymptotic	size	0.392	0.409	0.191	0.225	0.132	0.131	0.116	0.081	0.099	0.065
		power	0.055	0.063	0.063	0.062	0.074	0.073	0.077	0.076	0.075	0.081
	bootstrap	size	0.073		0.092		0.112		0.100		0.119	
		power	0.081		0.108		0.141		0.140		0.158	

Note:

1. The DGP is similar to Note 1 in Table 1-1, except the innovations are generated by $\varepsilon_{i,t}=u_{i,t}\exp^{0.5\alpha_{i,t}}$, $\alpha_{i,t}=\phi\alpha_{i,t-1}+\xi_{i,t-1}$, where $u_{i,t}$ and $\xi_{i,t-1}$ are $N(0,1)$, and ϕ is set to be 0.5.
2. For other information, please refer to section 5.1.

Table 2-1: Small sample performance of ENC and $AR-t_c$ -- MA($h-1$) error

			$n=16$		$n=32$		$n=64$		$n=128$		$n=256$	
			$AR-t_c$	DM	$AR-t_c$	DM	$AR-t_c$	DM	$AR-t_c$	DM	$AR-t_c$	DM
$h=1$	asymptotic	size	0.217	0.068	0.109	0.061	0.073	0.054	0.061	0.052	0.054	0.051
		power	0.136	0.301	0.232	0.331	0.28	0.346	0.324	0.345	0.337	0.345
	bootstrap	size	0.045		0.067		0.057		0.041		0.050	
		power	0.117		0.244		0.290		0.310		0.349	
$h=2$	asymptotic	size	0.221	0.095	0.105	0.078	0.072	0.064	0.061	0.057	0.057	0.054
		power	0.122	0.254	0.224	0.304	0.281	0.328	0.315	0.332	0.323	0.331
	bootstrap	size	0.063		0.068		0.054		0.054		0.065	
		power	0.106		0.233		0.286		0.314		0.347	
$h=3$	asymptotic	size	0.219	0.117	0.109	0.093	0.073	0.072	0.061	0.061	0.058	0.057
		power	0.125	0.221	0.229	0.27	0.283	0.304	0.309	0.32	0.317	0.321
	bootstrap	size	0.065		0.053		0.067		0.054		0.069	
		power	0.104		0.228		0.287		0.318		0.348	
$h=4$	asymptotic	size	0.221	0.139	0.11	0.106	0.076	0.08	0.063	0.063	0.056	0.053
		power	0.12	0.207	0.224	0.25	0.281	0.282	0.297	0.308	0.312	0.326
	bootstrap	size	0.059		0.059		0.056		0.061		0.065	
		power	0.108		0.206		0.271		0.292		0.312	
$h=5$	asymptotic	size	0.23	0.155	0.111	0.121	0.078	0.09	0.067	0.07	0.062	0.059
		power	0.115	0.195	0.218	0.235	0.258	0.271	0.28	0.286	0.289	0.293
	bootstrap	size	0.053		0.067		0.057		0.049		0.047	
		power	0.095		0.219		0.248		0.283		0.282	
$h=6$	asymptotic	size	0.231	0.177	0.118	0.125	0.082	0.091	0.068	0.068	0.067	0.065
		power	0.108	0.184	0.21	0.219	0.24	0.244	0.258	0.269	0.251	0.257
	bootstrap	size	0.043		0.067		0.060		0.049		0.061	
		power	0.104		0.227		0.250		0.292		0.270	
$h=7$	asymptotic	size	0.238	0.2	0.126	0.141	0.09	0.097	0.074	0.074	0.068	0.062
		power	0.116	0.173	0.197	0.192	0.218	0.221	0.243	0.236	0.245	0.244
	bootstrap	size	0.058		0.065		0.055		0.058		0.058	
		power	0.090		0.197		0.247		0.262		0.247	
$h=8$	asymptotic	size	0.246	0.219	0.138	0.145	0.095	0.101	0.082	0.076	0.079	0.067
		power	0.117	0.163	0.182	0.198	0.21	0.21	0.212	0.208	0.215	0.222
	bootstrap	size	0.052		0.081		0.078		0.060		0.075	
		power	0.104		0.227		0.212		0.260		0.277	

Note:

1. The data generation process (DGP) is $\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \delta & \sqrt{\omega-\delta^2} \end{bmatrix} \begin{bmatrix} \upsilon_{1t} \\ \upsilon_{2t} \end{bmatrix}$, where $\upsilon_{i,t} = \sum_{l=0}^{h-1} \pi_l \varepsilon_{i,t-l}$, $i=1,2$. $h=1$, …, 8 with $(\pi_0, \pi_1, ..., \pi_7) = (1, 0.1, -0.1, 0.2, -0.2, 0.3, -0.3, 0.4)$, and $\varepsilon_{i,t} \overset{iid}{\sim} N(0,1)$. Also, $\omega = 5$ and $\delta = 1, 0.5$ for size and power.
2. For other information, please refer to section 5.1.

Table 2-2: Small sample performance of ENC and $AR-t_c$ -- MA($h-1$)+GARCH(1,1) error

			$n=16$		$n=32$		$n=64$		$n=128$		$n=256$	
			$AR-t_c$	DM	$AR-t_c$	DM	$AR-t_c$	DM	$AR-t_c$	DM	$AR-t_c$	DM
$h=1$	asymptotic	size	0.216	0.069	0.106	0.064	0.071	0.055	0.061	0.052	0.053	0.051
		power	0.129	0.311	0.233	0.336	0.289	0.346	0.319	0.347	0.332	0.342
	bootstrap	size	0.040		0.065		0.054		0.044		0.044	
		power	0.111		0.237		0.285		0.318		0.356	
$h=2$	asymptotic	size	0.218	0.091	0.105	0.08	0.071	0.063	0.062	0.057	0.059	0.057
		power	0.125	0.267	0.229	0.312	0.29	0.336	0.312	0.338	0.321	0.335
	bootstrap	size	0.049		0.066		0.062		0.047		0.064	
		power	0.125		0.234		0.292		0.320		0.341	
$h=3$	asymptotic	size	0.22	0.115	0.109	0.093	0.076	0.072	0.06	0.06	0.058	0.058
		power	0.123	0.222	0.228	0.274	0.285	0.313	0.315	0.325	0.315	0.323
	bootstrap	size	0.056		0.056		0.059		0.050		0.063	
		power	0.106		0.225		0.293		0.322		0.344	
$h=4$	asymptotic	size	0.22	0.136	0.11	0.106	0.076	0.081	0.062	0.064	0.057	0.054
		power	0.127	0.214	0.216	0.267	0.282	0.288	0.296	0.311	0.311	0.323
	bootstrap	size	0.056		0.065		0.060		0.059		0.063	
		power	0.116		0.197		0.264		0.287		0.311	
$h=5$	asymptotic	size	0.227	0.155	0.11	0.118	0.078	0.089	0.068	0.07	0.061	0.06
		power	0.119	0.203	0.213	0.242	0.268	0.265	0.286	0.287	0.289	0.294
	bootstrap	size	0.051		0.073		0.056		0.049		0.046	
		power	0.102		0.233		0.251		0.281		0.279	
$h=6$	asymptotic	size	0.228	0.175	0.115	0.127	0.08	0.089	0.068	0.067	0.065	0.066
		power	0.119	0.186	0.212	0.218	0.241	0.253	0.263	0.273	0.26	0.258
	bootstrap	size	0.034		0.067		0.064		0.058		0.052	
		power	0.096		0.230		0.245		0.284		0.281	
$h=7$	asymptotic	size	0.238	0.198	0.122	0.137	0.089	0.096	0.072	0.072	0.067	0.063
		power	0.115	0.177	0.202	0.195	0.224	0.227	0.242	0.245	0.243	0.246
	bootstrap	size	0.058		0.062		0.056		0.060		0.060	
		power	0.095		0.189		0.248		0.262		0.257	
$h=8$	asymptotic	size	0.248	0.216	0.134	0.144	0.094	0.1	0.083	0.077	0.078	0.066
		power	0.114	0.158	0.189	0.196	0.212	0.218	0.211	0.211	0.211	0.219
	bootstrap	size	0.055		0.079		0.069		0.062		0.074	
		power	0.110		0.243		0.211		0.257		0.279	

Note:

1. The DGP is similar to Note 1 in Table 2-1, except the innovations are generated by $\varepsilon_{i,t}=u_{i,t}\sqrt{h_{i,t}}$, $h_{i,t}=w+ah_{i,t-1}+b\varepsilon_{i,t-1}^2$, where $u_{i,t}$ is standard normal, and $(w,a,b)=(0.15,0.13,0.2)$.
2. For other information, please refer to section 5.1.

Table 2-3: Small sample performance of ENC and $AR-t_c$ -- MA($h-1$) +SV error

			$n=16$		$n=32$		$n=64$		$n=128$		$n=256$	
			$AR-t_c$	DM	$AR-t_c$	DM	$AR-t_c$	DM	$AR-t_c$	DM	$AR-t_c$	DM
$h=1$	asymptotic	size	0.205	0.061	0.082	0.054	0.061	0.05	0.054	0.05	0.053	0.052
		power	0.117	0.309	0.259	0.351	0.325	0.371	0.339	0.36	0.367	0.37
	bootstrap	size	0.043		0.052		0.048		0.049		0.045	
		power	0.106		0.219		0.338		0.316		0.348	
$h=2$	asymptotic	size	0.194	0.088	0.088	0.069	0.063	0.06	0.056	0.054	0.055	0.055
		power	0.129	0.272	0.253	0.332	0.32	0.356	0.343	0.357	0.341	0.35
	bootstrap	size	0.061		0.062		0.055		0.048		0.059	
		power	0.094		0.226		0.309		0.360		0.393	
$h=3$	asymptotic	size	0.198	0.115	0.089	0.086	0.066	0.071	0.054	0.059	0.055	0.055
		power	0.124	0.214	0.24	0.283	0.301	0.323	0.339	0.343	0.351	0.355
	bootstrap	size	0.051		0.054		0.037		0.056		0.049	
		power	0.106		0.221		0.278		0.345		0.344	
$h=4$	asymptotic	size	0.21	0.143	0.087	0.094	0.071	0.079	0.062	0.066	0.058	0.056
		power	0.123	0.206	0.257	0.265	0.293	0.295	0.302	0.309	0.323	0.327
	bootstrap	size	0.043		0.064		0.059		0.060		0.053	
		power	0.114		0.240		0.296		0.349		0.324	
$h=5$	asymptotic	size	0.212	0.159	0.096	0.109	0.07	0.079	0.06	0.061	0.055	0.056
		power	0.124	0.198	0.238	0.243	0.278	0.286	0.308	0.316	0.325	0.315
	bootstrap	size	0.069		0.049		0.058		0.055		0.065	
		power	0.101		0.200		0.279		0.321		0.325	
$h=6$	asymptotic	size	0.223	0.177	0.105	0.123	0.079	0.092	0.062	0.064	0.063	0.062
		power	0.118	0.188	0.213	0.228	0.258	0.256	0.28	0.288	0.265	0.276
	bootstrap	size	0.061		0.065		0.058		0.053		0.046	
		power	0.098		0.234		0.266		0.302		0.286	
$h=7$	asymptotic	size	0.227	0.205	0.117	0.135	0.087	0.094	0.07	0.067	0.07	0.062
		power	0.107	0.175	0.204	0.207	0.245	0.243	0.255	0.256	0.252	0.259
	bootstrap	size	0.055		0.064		0.052		0.064		0.058	
		power	0.105		0.210		0.250		0.286		0.271	
$h=8$	asymptotic	size	0.237	0.232	0.124	0.141	0.093	0.102	0.083	0.075	0.078	0.066
		power	0.1	0.163	0.186	0.192	0.212	0.207	0.219	0.229	0.216	0.225
	bootstrap	size	0.064		0.067		0.084		0.068		0.068	
		power	0.106		0.238		0.240		0.275		0.287	

Note:

1. The DGP is similar to Note 1 in Table 2-1, except the innovations are generated by $\varepsilon_{i,t} = u_{i,t} \exp^{0.5\alpha_{i,t}}$, $\alpha_{i,t} = \phi\alpha_{i,t-1} + \xi_{i,t-1}$, where $u_{i,t}$ and $\xi_{i,t-1}$ are $N(0,1)$, and ϕ is set to be 0.5.
2. For other information, please refer to section 5.1.

Table 3-1: Comparisons of different block lengths ($b = c_1 n^{1/5}$) for DM*—MA($h-1$) error

		$n=16$		$n=32$		$n=64$		$n=128$		$n=256$	
		$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$
$h=1$	size	0.071	0.167	0.076	0.121	0.068	0.098	0.059	0.080	0.054	0.061
	power	0.266	0.353	0.234	0.254	0.259	0.280	0.248	0.273	0.267	0.298
$h=2$	size	0.070	0.162	0.085	0.122	0.067	0.084	0.066	0.075	0.060	0.058
	power	0.236	0.336	0.211	0.262	0.255	0.263	0.256	0.241	0.263	0.268
$h=3$	size	0.060	0.171	0.084	0.122	0.067	0.074	0.060	0.065	0.056	0.053
	power	0.233	0.346	0.215	0.239	0.240	0.264	0.243	0.267	0.274	0.264
$h=4$	size	0.083	0.160	0.083	0.125	0.075	0.086	0.072	0.073	0.057	0.067
	power	0.238	0.328	0.215	0.251	0.248	0.247	0.232	0.230	0.261	0.243
$h=5$	size	0.091	0.192	0.084	0.118	0.079	0.078	0.074	0.072	0.066	0.068
	power	0.261	0.340	0.228	0.230	0.236	0.241	0.242	0.236	0.274	0.259
$h=6$	size	0.096	0.173	0.091	0.132	0.083	0.088	0.082	0.079	0.065	0.074
	power	0.264	0.329	0.214	0.250	0.243	0.239	0.241	0.229	0.243	0.260
$h=7$	size	0.094	0.191	0.108	0.136	0.093	0.108	0.083	0.091	0.075	0.075
	power	0.244	0.347	0.216	0.243	0.237	0.253	0.228	0.228	0.234	0.235
$h=8$	size	0.106	0.192	0.110	0.156	0.105	0.109	0.090	0.111	0.084	0.085
	power	0.243	0.347	0.227	0.243	0.226	0.234	0.220	0.214	0.226	0.233

Note: see Table 1-1.

Table 3-2: Comparisons of different block lengths ($b = c_1 n^{1/5}$) for DM*—MA($h-1$)+GARCH(1,1) error

		$n=16$		$n=32$		$n=64$		$n=128$		$n=256$	
		$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$
$h=1$	size	0.092	0.166	0.090	0.124	0.101	0.110	0.076	0.095	0.073	0.075
	power	0.223	0.297	0.196	0.216	0.223	0.226	0.200	0.206	0.205	0.209
$h=2$	size	0.080	0.168	0.101	0.129	0.090	0.108	0.088	0.086	0.081	0.070
	power	0.204	0.278	0.180	0.234	0.213	0.210	0.209	0.177	0.186	0.180
$h=3$	size	0.077	0.175	0.102	0.130	0.081	0.089	0.086	0.081	0.072	0.072
	power	0.195	0.289	0.189	0.205	0.199	0.207	0.197	0.205	0.203	0.199
$h=4$	size	0.103	0.172	0.100	0.127	0.089	0.105	0.084	0.091	0.070	0.082
	power	0.208	0.277	0.182	0.205	0.208	0.206	0.191	0.186	0.190	0.192
$h=5$	size	0.110	0.201	0.094	0.132	0.095	0.089	0.086	0.085	0.072	0.079
	power	0.230	0.285	0.201	0.203	0.195	0.189	0.195	0.191	0.211	0.194
$h=6$	size	0.106	0.179	0.101	0.146	0.111	0.107	0.098	0.090	0.085	0.084
	power	0.241	0.299	0.203	0.212	0.216	0.210	0.214	0.189	0.200	0.204
$h=7$	size	0.109	0.192	0.124	0.145	0.115	0.129	0.111	0.108	0.094	0.081
	power	0.230	0.312	0.205	0.215	0.210	0.227	0.216	0.201	0.199	0.198
$h=8$	size	0.126	0.206	0.129	0.167	0.126	0.125	0.120	0.132	0.094	0.100
	power	0.237	0.326	0.225	0.230	0.221	0.217	0.203	0.200	0.210	0.203

Note: see Table 1-2.

Table 3-3: Comparisons of different block lengths ($b = c_1 n^{1/5}$) for DM*—MA($h-1$)+SV error

		$n=16$		$n=32$		$n=64$		$n=128$		$n=256$	
		$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$
$h=1$	size	0.129	0.216	0.191	0.232	0.300	0.281	0.413	0.418	0.618	0.599
	power	0.262	0.350	0.349	0.365	0.500	0.473	0.613	0.599	0.796	0.775
$h=2$	size	0.132	0.221	0.198	0.234	0.286	0.295	0.444	0.434	0.609	0.612
	power	0.256	0.326	0.347	0.368	0.502	0.490	0.628	0.622	0.783	0.789
$h=3$	size	0.144	0.225	0.198	0.236	0.287	0.297	0.419	0.414	0.621	0.604
	power	0.272	0.331	0.342	0.349	0.497	0.488	0.624	0.617	0.794	0.774
$h=4$	size	0.142	0.225	0.216	0.232	0.295	0.301	0.428	0.438	0.616	0.585
	power	0.270	0.343	0.359	0.364	0.504	0.473	0.622	0.635	0.785	0.764
$h=5$	size	0.163	0.235	0.222	0.228	0.318	0.295	0.444	0.400	0.606	0.616
	power	0.289	0.358	0.375	0.349	0.497	0.475	0.642	0.607	0.768	0.790
$h=6$	size	0.180	0.256	0.213	0.247	0.315	0.298	0.466	0.431	0.605	0.608
	power	0.318	0.387	0.369	0.381	0.495	0.472	0.667	0.627	0.772	0.765
$h=7$	size	0.196	0.280	0.257	0.258	0.352	0.314	0.469	0.432	0.614	0.586
	power	0.345	0.411	0.415	0.385	0.529	0.502	0.655	0.615	0.782	0.773
$h=8$	size	0.208	0.297	0.276	0.278	0.343	0.330	0.460	0.429	0.624	0.588
	power	0.357	0.424	0.427	0.421	0.523	0.527	0.637	0.620	0.799	0.757

Note: see Table 1-3.

Table 4-1: Comparisons of different block lengths ($b = c_1 n^{1/4}$) for ENC*—MA($h-1$) error

		$n=16$		$n=32$		$n=64$		$n=128$		$n=256$	
		$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$
$h=1$	size	0.079	0.124	0.069	0.092	0.062	0.082	0.060	0.061	0.056	0.059
	power	0.374	0.434	0.372	0.396	0.370	0.366	0.341	0.372	0.348	0.355
$h=2$	size	0.076	0.138	0.059	0.097	0.061	0.086	0.064	0.079	0.050	0.059
	power	0.352	0.451	0.372	0.402	0.362	0.385	0.358	0.380	0.343	0.368
$h=3$	size	0.076	0.134	0.076	0.097	0.066	0.075	0.061	0.069	0.053	0.056
	power	0.347	0.433	0.380	0.404	0.378	0.385	0.337	0.355	0.363	0.351
$h=4$	size	0.077	0.141	0.069	0.110	0.065	0.087	0.066	0.060	0.060	0.054
	power	0.358	0.424	0.381	0.390	0.355	0.370	0.347	0.347	0.332	0.340
$h=5$	size	0.082	0.135	0.077	0.106	0.060	0.081	0.062	0.069	0.066	0.070
	power	0.345	0.438	0.370	0.390	0.335	0.369	0.344	0.336	0.321	0.348
$h=6$	size	0.087	0.124	0.085	0.103	0.077	0.091	0.068	0.073	0.069	0.065
	power	0.377	0.404	0.361	0.392	0.339	0.359	0.324	0.308	0.319	0.290
$h=7$	size	0.098	0.132	0.090	0.105	0.080	0.096	0.078	0.076	0.072	0.063
	power	0.362	0.420	0.370	0.367	0.335	0.351	0.328	0.316	0.318	0.293
$h=8$	size	0.105	0.148	0.095	0.123	0.084	0.092	0.087	0.075	0.087	0.075
	power	0.367	0.426	0.362	0.375	0.334	0.335	0.314	0.298	0.293	0.269

Note: see Table 2-1.

Table 4-2: Comparisons of different block lengths ($b = c_1 n^{1/4}$) for ENC*—MA($h-1$)+GARCH(1,1) error

		$n=16$		$n=32$		$n=64$		$n=128$		$n=256$	
		$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$
$h=1$	size	0.078	0.131	0.070	0.100	0.065	0.079	0.065	0.063	0.054	0.059
	power	0.367	0.443	0.391	0.414	0.378	0.386	0.363	0.376	0.360	0.357
$h=2$	size	0.074	0.141	0.058	0.099	0.061	0.086	0.070	0.083	0.058	0.063
	power	0.355	0.444	0.390	0.409	0.375	0.393	0.366	0.393	0.349	0.372
$h=3$	size	0.079	0.138	0.078	0.100	0.065	0.083	0.065	0.069	0.057	0.057
	power	0.352	0.427	0.385	0.411	0.398	0.407	0.348	0.373	0.370	0.368
$h=4$	size	0.076	0.143	0.066	0.111	0.071	0.088	0.067	0.068	0.064	0.060
	power	0.355	0.426	0.397	0.397	0.366	0.391	0.357	0.357	0.341	0.353
$h=5$	size	0.080	0.140	0.073	0.109	0.064	0.086	0.068	0.072	0.068	0.072
	power	0.350	0.436	0.373	0.400	0.349	0.382	0.354	0.351	0.337	0.360
$h=6$	size	0.085	0.116	0.089	0.108	0.071	0.095	0.074	0.076	0.072	0.060
	power	0.375	0.395	0.368	0.395	0.343	0.367	0.333	0.326	0.328	0.298
$h=7$	size	0.099	0.138	0.086	0.102	0.074	0.101	0.077	0.084	0.081	0.068
	power	0.357	0.421	0.375	0.373	0.339	0.363	0.342	0.332	0.331	0.299
$h=8$	size	0.106	0.159	0.100	0.122	0.086	0.098	0.092	0.075	0.086	0.079
	power	0.377	0.433	0.361	0.387	0.342	0.343	0.326	0.301	0.294	0.278

Note : see Table 2-2.

Table 4-3: Comparisons of different block lengths ($b = c_1 n^{1/4}$) for ENC*—MA($h-1$)+SV error

		$n=16$		$n=32$		$n=64$		$n=128$		$n=256$	
		$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$	$c_1=1$	$c_1=3$
$h=1$	size	0.079	0.143	0.067	0.120	0.082	0.081	0.081	0.096	0.065	0.067
	power	0.294	0.374	0.317	0.365	0.327	0.334	0.316	0.314	0.289	0.291
$h=2$	size	0.081	0.144	0.086	0.103	0.082	0.103	0.080	0.079	0.064	0.072
	power	0.300	0.379	0.318	0.347	0.339	0.330	0.310	0.315	0.308	0.299
$h=3$	size	0.085	0.145	0.080	0.117	0.082	0.094	0.083	0.090	0.073	0.088
	power	0.279	0.367	0.337	0.359	0.306	0.320	0.308	0.323	0.283	0.294
$h=4$	size	0.085	0.147	0.090	0.114	0.076	0.089	0.069	0.076	0.067	0.070
	power	0.304	0.377	0.322	0.343	0.325	0.328	0.283	0.297	0.284	0.276
$h=5$	size	0.091	0.145	0.101	0.100	0.085	0.097	0.077	0.086	0.084	0.066
	power	0.300	0.366	0.320	0.330	0.309	0.323	0.298	0.314	0.291	0.266
$h=6$	size	0.103	0.142	0.103	0.112	0.088	0.090	0.086	0.086	0.079	0.067
	power	0.296	0.363	0.332	0.344	0.315	0.294	0.274	0.278	0.281	0.257
$h=7$	size	0.105	0.147	0.087	0.114	0.101	0.096	0.081	0.092	0.081	0.077
	power	0.317	0.362	0.296	0.324	0.300	0.285	0.269	0.281	0.246	0.240
$h=8$	size	0.105	0.156	0.105	0.111	0.092	0.109	0.090	0.087	0.090	0.072
	power	0.309	0.369	0.317	0.311	0.289	0.280	0.261	0.262	0.233	0.229

Note: see Table 2-3.

Table 5: Out-of-sample predictive performance for NTD/USD spot rate

	$AR-t_c$	DM	$AR-t_c$	ENC
$h=1$	-2.009(4)[1] [0.022, 0.021][2]	-2.487 [0.006, 0.001]	0.914(4) [0.180, 0.198]	0.496 [0.310, 0.227]
$h=2$	-1.867(3) [0.031, 0.014]	-1.546 [0.061, 0.041]	0.327(4) [0.372, 0.357]	0.145 [0.442, 0.486]
$h=3$	-1.535(3) [0.062, 0.031]	-1.596 [0.055, 0.028]	-0.589(3) [0.278, 0.279]	-0.471 [0.319, 0.294]
$h=4$	-1.328(1) [0.092, 0.043]	-1.259 [0.104, 0.076]	-0.357(3) [0.361, 0.374]	-0.220 [0.413, 0.373]
$h=6$	-1.430(2) [0.076, 0.034]	-1.687 [0.046, 0.018]	-0.934(4) [0.175, 0.115]	-0.975 [0.165, 0.151]

Notes:

1. The numbers in parentheses are the optimal lag for $AR-t_c$ test.
2. The left numbers in square brackets are asymptotic *p*-values, and the right ones in square brackets are bootstrap *p*-values. The bootstrap replications are 1000.
3. The forward market forecast error is denoted by e_{1t}, and the random walk forecast error is denoted by e_{2t}. The long-run variances in DM and ENC tests are estimated using the Bartlett kernel with a bandwidth set to be $h-1$, and the maximum lag for $AR-t_c$ is 5. Also, the block length in MBB is determined by $n^{1/5}$ for DM and $n^{1/4}$ for ENC individually, where $n=114$.

Reference

Akgiray, V. (1989), “Conditional heteroscedasticity in time series of stock returns: evidence and forecasts”, *Journal of Business*, 62, 55-80.

Bühlmann, P. (1995), ``Moving-average representation of autoregressive approximations", *Stochastic Processes and their Applications*, 60, 331-342.

Chang, Y. and J.Y. Park (2003), ``A sieve bootstrap for the test of a unit root", *Journal of Time Series Analysis*, 24, 379-400.

Chong, Y. Y., and Hendry, D. F. (1986), “Econometric Evaluation of Linear Macro-Economic Models,” *Review of Economic Studies*, 53, 671-690.

Clark, T.E., (1999), ``Finite-sample properties of tests for equal forecast accuracy", *Journal of Forecasting*, 18, 489-504.

Clark, T. E. and McCracken, M. W., (2001) "Tests of equal forecast accuracy and encompassing for nested models," *Journal of Econometrics*, 105, 85-110.

Clements, M. P. and Hendry, D. F. (1993), “On the Limitations of Comparing Mean Square Forecast Errors”, *Journal of Forecasting* 12, 617-637.

Diebold, F.X., and Mariano, R.S. (1995), ``Comparing predictive accuracy," *Journal of Business and Economic Statistics*, 13, 253-263.

Flachaire E. (2005), ``Bootstrapping heteroskedastic regression models: wild bootstrap vs. pairs bootstrap," *Computational Statistics & Data Analysis*, 49. 361-376.

Goncalves, S., and L. Kilian (2004), ``Bootstrapping autoregressions with conditional heteroskedasticity of unknown form’’, *Journal of Econometrics*, 123, 89-120.

Hall, P. and B.Y. Jing (1996), ``On sample reuse methods for dependent data," *Journal of Royal Statistical Society Series B*, 58, 727-737.

Harvey, D.I., Leybourne, S.J., and Newbold, P. (1997), ``Testing the equality of prediction mean squared errors," *International Journal of Forecasting*, 13, 281-291.

Harvey, D. I., Leybourne, S. J., and Newbold, P. (1998), “Tests for Forecast Encompassing,” *Journal of Business and Economic Statistics*, 16, 254-259.

Harvey, D.I. and Newbold, P. (2000), “Tests for multiple forecast encompassing” *Journal of*

Applied Econometrics, 15: 471-482

Horowitz, J.L. (2001), ``The bootstrap," *Handbook of Econometrics*, 5, 3159-3228.

Kapetanios G. and Psaradakis Z. (2006), ``Sieve bootstrap for strongly dependent stationary processes, " Working Paper No. 552, Queen Mary, University of London.

Künsch, H. R. (1989), The jackknife and the bootstrap for general stationary observations,Annals of Statistics, 17, 1217-1241.

Kress, J. P. (1997), ``Asymptotic properties of residual bootstrap for autoregressions, Manuscript, Institute for Mathematical Stochastics,'' Technical University of Braunschweig, Germany.

McCracken, M. W. (2000). ``Robust out–of–sample inference,'' *Journal of Econometrics*, 99:195-223.

McCracken, M. W. (2006). ``Asymptotics for out–of–sample tests of causality,". *Journal of Econometrics* (forthcoming)

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.

O'Reilly and Whelany (2005) ``Testing Parameter Stability:A Wild Bootstrap Approach" working paper.

Park, J.Y. (2002), ``An invariance principle for sieve bootstrap in time series", *Econometric Theory*, 18, 469-490.

Richard P. (2006) ``A Moving Average Sieve Bootstrap for Unit Root Tests", working paper.

Said, S. and Dickey, D. A. (1984), "Testing for Unit Roots in Autoregressive Moving-Average Models of Unknown Order," *Biometrika*, 71, 599-607.

Shibata, R. (1980), ``Asymptotically efficient selection of the order of the model for estimating parameters of a linear process,'' *Annals of Statistics*, 8, 147-164.

West, K. D. (1996). ``Asymptotic inference about predictive ability,'' *Econometrica* 64:1067-1084.

West, K. D.and McCracken, M. W. (1998): "Regression-Based Tests of Predictive Ability", *International Economic Review*, 39, 817-840.

West, K.D., (2005), Forecast evaluation. Manuscript for Handbook of Economic Forecasting, University of Wisconsin.

White H. (1984) *Asymptotic Theory For Econometricians*. New York: Academic Press

Wu, C. F. J. (1986), ``Jackknife, bootstrap and other resampling methods in regression analysis,'' *The Annals of Statistics*, 14, 1261-1295.