# Bootstrapping Unit Root Tests with Covariates ${ }^{1}$ 

Yoosoon Chang<br>Department of Economics<br>Indiana University<br>Robin C. Sickles<br>Department of Economics<br>Rice University<br>Wonho Song ${ }^{2}$<br>School of Economics<br>Chung-Ang University


#### Abstract

We consider the bootstrap method for the covariates augmented DickeyFuller (CADF) unit root test suggested in Hansen (1995) which uses related variables to improve the power of univariate unit root tests. It is shown that there are substantial power gains from including correlated covariates. The limit distribution of the CADF test, however, depends on the nuisance parameter that represents the correlation between the equation error and the covariates. Hence, inference based directly on the CADF test is not possible. To provide a valid inferential basis for the CADF test, we propose to use the parametric bootstrap procedure with iid errors to obtain critical values, and establish the asymptotic validity of the bootstrap CADF test. Simulations show that the bootstrap CADF test significantly improves the asymptotic and the finite sample size performances of the CADF test, especially when the covariates are highly correlated with the error. Indeed, the bootstrap CADF test offers drastic power gains over the conventional unit root tests. We apply our testing procedures to the extended Nelson and Plosser data set as well as to postwar annual CPI-based real exchange rates for fifteen OECD countries.


This version: October 21, 2013.
JEL Classification: C12, C15, C22.
Key words and phrases: Unit root tests, covariates, bootstrap consistency.

[^0]
## 1. Introduction

Conventional univariate tests for the presence of unit roots in aggregate economic time series have important implications for the conduct of domestic macro and international economic policy. These tests unfortunately have been plagued by reliance on relatively short time series with relatively low frequencies. Size distortions and low power are well-known problems with conventional testing procedures [see, e.g., Stock (1991), and Campbell and Perron (1991), Domowitz and El-Gamal (2001)]. Current macroeconomic theory provides little in the way of guidance on how to increase the power and moderate size distortions other than by increasing the length of the time series. Reliance on the sort of information that was utilized in conventional empirical macroeconomics before the Lucas critique took hold, namely the information contained in the correlated errors of other overidentified equations in the structural system, has little apparent place in the current unit root testing literature. Even agreement on the candidate set of correlated series has little theoretical basis (Stock and Watson, 1999). The first wide-spread use of univariate tests for the presence of unit roots was carried in the seminal work of Nelson and Plosser (1982) who found that most U.S. macroeconomic time series could be characterized as a univariate time series structure with a unit root. Subsequent empirical analyses have largely confirmed their findings while the literature continues to acknowledge the low power of unit root tests and an implication of this low power, that in finite samples it is almost impossible to discriminate between a unit root process and one which is very close to it.

Clearly the unit root hypothesis has important implications for determining the effect of random shocks on an economic system and the literature has not been silent on the many efforts to overcome the low power of conventional unit root tests. One such contribution was made by Hansen (1995) who noted that conventional univariate unit root tests ignore potentially useful information from related time series and that the inclusion of related stationary covariates in the regression equation may lead to a more precise estimate of the autoregressive coefficient. He proposed to use the covariates augmented Dickey-Fuller (CADF) unit root test rather than conventional univariate unit root tests. He analyzed the asymptotic local power functions for the CADF $t$-statistic and discovered that enormous power gains could be achieved by the inclusion of appropriate covariates. His Monte Carlo study suggested that such gains were also possible in the finite sample power performances of the CADF vis-a-vis conventional ADF test.

Hansen (1995) showed that the limit distribution of the CADF test is dependent on the nuisance parameter that characterizes the correlation between the equation error and the covariates. Therefore, it is not possible to perform valid statistical inference directly using the CADF test. To deal with this inferential difficulty, Hansen (1995) suggested using critical values based on an estimated nuisance parameter. ${ }^{3}$ His two-step procedure can be a practical solution for the implementation of the CADF test. However, relying on the estimated value of the nuisance parameter would introduce additional source of variability.

[^1]In this paper, we apply the parametric bootstrap method with iid errors to the CADF test to deal with the nuisance parameter dependency and to provide a valid basis for inference based on the CADF test. In particular, we show the consistency of the bootstrap CADF test and establish the asymptotic validity of the critical values from the bootstrap distribution of the test. The asymptotic properties of the CADF and bootstrap CADF tests are investigated and the finte sample performances of the CADF tests are compared with various well-known univariate unit root tests. The simulations show that the CADF test based on the two-step procedure suffers from serious size distortions, especially when the covariates are highly correlated with the error, while our bootstrap CADF test significantly improves the asymptotic and the finite sample size performances of the CADF test. Moreover, the bootstrap CADF test offers dramatic power gains over the conventional unit root tests.

As illustrations, we apply our covariate tests and standard unit root tests in a reexamination of the stationarity of U.S. domestic macroeconomic aggregates and international rates of exchange. The former are analyzed with the extended Nelson and Plosser data set. We investigate whether the findings of unit roots in the Nelson and Plosser data set are reversed when the more powerful covariate tests are used. The latter are examined using postwar annual CPI-based real exchange rates for fifteen OECD countries, for which most previous studies failed to reject the null hypothesis of a unit root. We find that our new covariate test rejects the unit root hypothesis in four series in the Nelson and Plosser data set for the period 1930-1972 and six cases for the postwar real exchange rates.

The paper is organized as follows. Section 2 introduces the unit root test with covariates and derives limit theories for the sample tests. Section 3 applies the bootstrap methodology to the sample tests considered in Section 2 and establishes the asymptotic validity of the bootstrap test. Section 3 also provides a discussion of practical issues arising from the implementation of the bootstrap methodology. Section 4 considers asymptotic powers of the bootstrap tests against the local-to-unity models. In Section 5, we conduct simulations to investigate the finite sample performances of the bootstrap CADF test. Empirical applications are presented in Section 6 while Section 7 concludes. All mathematical proofs are provided in the Appendix.

## 2. Unit Root Tests with Covariates

### 2.1 Model and Assumptions

We consider the time series $\left(y_{t}\right)$ given by

$$
\begin{equation*}
\triangle y_{t}=\alpha y_{t-1}+u_{t} \tag{1}
\end{equation*}
$$

for $t=1, \ldots, n$, where $\triangle$ is the usual difference operator. ${ }^{4}$ We let the regression errors $\left(u_{t}\right)$ in the model (1) to be serially correlated, and also allow them to be related to other stationary

[^2]covariates. To define the data generating process for the errors $\left(u_{t}\right)$ more explicitly, let ( $w_{t}$ ) be an $m$-dimensional stationary covariates. It is assumed that the errors ( $u_{t}$ ) are given by a $p$-th order autoregressive exogenous (ARX) process specified as
\[

$$
\begin{equation*}
\alpha(L) u_{t}=\beta(L)^{\prime} w_{t}+\varepsilon_{t} \tag{2}
\end{equation*}
$$

\]

where $L$ is the lag operator, $\alpha(z)=1-\sum_{k=1}^{p} \alpha_{k} z^{k}$ and $\beta(z)=\sum_{k=-r}^{q} \beta_{k} z^{k}$.
We consider the test of the unit root null hypothesis $\alpha=0$ for ( $y_{t}$ ) given as in (1), against the alternative of the stationarity $\alpha<0$. The initial value $y_{0}$ of $\left(y_{t}\right)$ does not affect our subsequent analysis so long as it is stochastically bounded, and therefore we set it at zero for expositional brevity.

Under the null hypothesis of unit root, $\Delta y_{t}=u_{t}$ and we have from (2) that

$$
\begin{equation*}
\Delta y_{t}=\alpha y_{t-1}+\sum_{k=1}^{p} \alpha_{k} \triangle y_{t-k}+\sum_{k=-r}^{q} \beta_{k}^{\prime} w_{t-k}+\varepsilon_{t} \tag{3}
\end{equation*}
$$

which is an autoregression of $\Delta y_{t}$ augmented by its lagged level $y_{t-1}$ and the leads and lags of the $m$ stationary covariates in $\left(w_{t}\right)$. Indeed, the above regression may be viewed as a further augmentation of the usual ADF regression, which is an autoregression of $\triangle y_{t}$ augmented by its lagged level $y_{t-1}$ only. Our test statistics for testing the unit root in $\left(y_{t}\right)$, which are introduced in the next section, will be based on the least squares estimator for $\alpha$ from this CADF regression.

For the subsequent analysis, we also need to be more explicit about the data generating process for the stationary variables $\left(w_{t}\right)$ that are used as covariates. We assume that ( $w_{t}$ ) is generated by an $\operatorname{AR}(\ell)$ process as

$$
\begin{equation*}
\Phi(L) w_{t+r+1}=\eta_{t} \tag{4}
\end{equation*}
$$

where $\Phi(z)=I_{m}-\sum_{k=1}^{\ell} \Phi_{k} z^{k}$.
To define explicitly the correlation between the covariates $\left(w_{t}\right)$ and the series to be tested $\left(y_{t}\right)$, we consider together the innovations $\left(\eta_{t}\right)$ and $\left(\varepsilon_{t}\right)$ that generate, respectively, the covariates $\left(w_{t}\right)$ and the regression error $\left(u_{t}\right)$, which in turn generates $\left(y_{t}\right)$. Define

$$
\xi_{t}=\left(\varepsilon_{t}, \eta_{t}^{\prime}\right)^{\prime}
$$

and denote by $|\cdot|$ the Euclidean norm: for a vector $x=\left(x_{i}\right),|x|^{2}=\sum_{i} x_{i}^{2}$ and for a matrix $A=\left(a_{i j}\right),|A|^{2}=\sum_{i, j} a_{i j}^{2}$. We now lay out assumptions needed for the development of our asymptotic theory.
Assumption 2.1 We assume
(a) Let $\left(\xi_{t}\right)$ be a sequence of iid random variables such that $\mathbf{E} \xi_{t}=0, \mathbf{E} \xi_{t} \xi_{t}^{\prime}=\Sigma>0$ and $\mathbf{E}\left|\xi_{t}\right|^{\gamma}<\infty$ for some $\gamma \geq 4$.
(b) $\alpha(z), \operatorname{det}(\Phi(z)) \neq 0$ for all $|z| \leq 1$.

Here, we assume $\left(\xi_{t}\right)$ to be an iid sequence to make the bootstrap procedure in the next section meaningful. Assumption 2.1 (a) excludes conditional heteroskedasticity (e.g. GARCH)
in all equations in the system including the covariates. It also states that the regression error $\left(\varepsilon_{t}\right)$ in equation (3) is serially uncorrelated and independent of $\left(\eta_{t+k}\right)$ for $k \geq 1$. The condition effectively implies that the regression error $\left(\varepsilon_{t}\right)$ is orthogonal to the lagged differences of the dependent variable ( $\Delta y_{t-1}, \ldots, \Delta y_{t-p}$ ) and the leads and lags of the stationary covariates $\left(w_{t+r}, \ldots, w_{t-q}\right)$, which is necessary for the regression (3) to be consistently estimated via usual least squares estimation. (See Hansen (1995) for more details.)

Under Assumption 2.1 (a), the following invariance principle holds

$$
\begin{equation*}
\frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]} \xi_{t} \rightarrow_{d} B(s) \tag{5}
\end{equation*}
$$

for $s \in[0,1]$ as $n \rightarrow \infty$. The limit process $B=\left(B_{\varepsilon}, B_{\eta}^{\prime}\right)^{\prime}$ is an $(1+m)$-dimensional vector Brownian motion with covariance matrix

$$
\Sigma=\left(\begin{array}{cc}
\sigma_{\varepsilon}^{2} & \sigma_{\varepsilon \eta}  \tag{6}\\
\sigma_{\eta \varepsilon} & \Sigma_{\eta}
\end{array}\right)
$$

The asymptotic behavior of $\left(y_{t}\right)$ is determined by that of $\left(u_{t}\right)$ as shown in model (1), and the latter is dependent upon the limiting behavior of the stationary covariates $\left(w_{t}\right)$ and the innovations $\left(\varepsilon_{t}\right)$ as indicated in the relation (2). We may then derive the limit behavior of $\left(u_{t}\right)$ using the specification given in (2) from those of $\left(\varepsilon_{t}\right)$ and $\left(w_{t}\right)$ as follows:

$$
\frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]} u_{t} \rightarrow_{d} \pi(1)\left(\beta(1)^{\prime} \Psi(1) B_{\eta}(s)+B_{\varepsilon}(s)\right)
$$

as $n \rightarrow \infty$, where $\pi(1)=1 / \alpha(1)$ and $\Psi(1)=\Phi(1)^{-1}$. This is derived in Lemma A. 1 (b) in the Appendix. The variance of the limit process given in the previous equation is easily derived as

$$
\begin{equation*}
\sigma_{u}^{2}=\pi(1)^{2}\left(\beta(1)^{\prime} \Psi(1) \Sigma_{\eta} \Psi(1)^{\prime} \beta(1)+\sigma_{\varepsilon}^{2}+2 \beta(1)^{\prime} \Psi(1) \sigma_{\eta \varepsilon}\right) \tag{7}
\end{equation*}
$$

using the parameters defined in the preceding equations.
Let $z_{t}=\left(\triangle y_{t-1}, \ldots, \Delta y_{t-p}, w_{t+r}^{\prime}, \ldots, w_{t-q}^{\prime}\right)^{\prime}$. We assume
Assumption $2.2 \sigma_{u}^{2}>0$ and $\mathbf{E} z_{t} z_{t}^{\prime}>0$.
The condition $\sigma_{u}^{2}>0$ ensures that the series $\left(y_{t}\right)$ is $\mathrm{I}(1)$ when $\alpha=0$, which is necessary to be able to interpret testing $\alpha=0$ as testing for a unit root in $\left(y_{t}\right)$. The condition $\mathbf{E} z_{t} z_{t}^{\prime}>0$ implies that the stationary regressors in $\left(z_{t}\right)$ are asymptotically linearly independent, which is required along with the condition Assumption 2.1 (a) for the consistency of the LS coefficient estimates for $\left(z_{t}\right)$.

### 2.2 Covariates Augmented Unit Root Tests

To introduce our test statistics more effectively, we first define

$$
A_{n}=\sum_{t=1}^{n} y_{t-1} \varepsilon_{t}-\left(\sum_{t=1}^{n} y_{t-1} z_{t}^{\prime}\right)\left(\sum_{t=1}^{n} z_{t} z_{t}^{\prime}\right)^{-1}\left(\sum_{t=1}^{n} z_{t} \varepsilon_{t}\right)
$$

$$
\begin{aligned}
B_{n} & =\sum_{t=1}^{n} y_{t-1}^{2}-\left(\sum_{t=1}^{n} y_{t-1} z_{t}^{\prime}\right)\left(\sum_{t=1}^{n} z_{t} z_{t}^{\prime}\right)^{-1}\left(\sum_{t=1}^{n} z_{t} y_{t-1}\right) \\
C_{n} & =\sum_{t=1}^{n} \varepsilon_{t}^{2}-\left(\sum_{t=1}^{n} \varepsilon_{t} z_{t}^{\prime}\right)\left(\sum_{t=1}^{n} z_{t} z_{t}^{\prime}\right)^{-1}\left(\sum_{t=1}^{n} z_{t} \varepsilon_{t}\right) .
\end{aligned}
$$

Now we have

$$
\begin{aligned}
\hat{\alpha}_{n} & =A_{n} B_{n}^{-1} \\
\hat{\sigma}_{n}^{2} & =n^{-1}\left(C_{n}-A_{n}^{2} B_{n}^{-1}\right) \\
s\left(\hat{\alpha}_{n}\right)^{2} & =\hat{\sigma}_{n}^{2} B_{n}^{-1}
\end{aligned}
$$

where $\hat{\alpha}_{n}$ is the OLS estimator of $\alpha$ from the covariates augmented regression (3), $\hat{\sigma}_{n}^{2}$ is the usual error variance estimator, and $s\left(\hat{\alpha}_{n}\right)$ is the estimated standard error for $\hat{\alpha}_{n}$. We also let

$$
\begin{equation*}
\hat{\alpha}_{n}(1)=1-\sum_{k=1}^{p} \hat{\alpha}_{k} \tag{8}
\end{equation*}
$$

where $\hat{\alpha}_{k}$ 's are the OLS estimators of $\alpha_{k}$ 's in the CADF regression (3).
The statistics that we will consider in the paper are given by

$$
\begin{align*}
S_{n} & =\frac{n \hat{\alpha}_{n}}{\hat{\alpha}_{n}(1)}  \tag{9}\\
T_{n} & =\frac{\hat{\alpha}_{n}}{s\left(\hat{\alpha}_{n}\right)} . \tag{10}
\end{align*}
$$

Note that $S_{n}$ is a test for the unit root based on the estimated unit root regression coefficient, and $T_{n}$ is the usual $t$-statistics for testing the unit root hypothesis from the CADF regression (3). The test $T_{n}$ is considered in Hansen (1995).

The limit theories for the tests $S_{n}$ and $T_{n}$ are given in
Theorem 2.3 Under the null hypothesis $\alpha=0$, we have as $n \rightarrow \infty$,

$$
\begin{aligned}
& S_{n} \rightarrow_{d} \sigma_{\varepsilon} \frac{\int_{0}^{1} Q(s) d P(s)}{\int_{0}^{1} Q(s)^{2} d s} \\
& T_{n} \rightarrow_{d} \frac{\int_{0}^{1} Q(s) d P(s)}{\left(\int_{0}^{1} Q(s)^{2} d s\right)^{1 / 2}}
\end{aligned}
$$

under Assumptions 2.1 and 2.2, where

$$
Q(s)=\beta(1)^{\prime} \Psi(1) B_{\eta}(s)+B_{\varepsilon}(s)
$$

and $P(s)=B_{\varepsilon}(s) / \sigma_{\varepsilon}$.
The asymptotic distributions are presented explicitly in terms of the Brownian motions $B_{\varepsilon}$ and $B_{\eta}$ via $Q=\beta(1)^{\prime} \Psi(1) B_{\eta}+B_{\varepsilon}$ and $P=B_{\varepsilon} / \sigma_{\varepsilon}$. In this way we can easily relate the asymptotic distributions of the bootstrapped tests, which are developed in the next section, to the limit distributions of the sample statistics given above. Moreover, it is straightforward to show that the null asymptotic distribution for the CADF test $T_{n}$ given in Theorem 2.3 is actually equivalent to the one derived in Hansen (1995, Theorem 3) (thus its derivation is omitted).

The asymptotic distributions for both $S_{n}$ and $T_{n}$ are nonstandard and depend upon the nuisance parameters that characterize the correlation between the covariates and the regression error as shown in Hansen (1995). The limit distributions are therefore basically unknown. Consequently it is impossible to perform valid statistical inference based directly on the CADF tests. As a feasible practical solution, one may simulate critical values for the tests for each value of the nuisance parameter and use its estimated value to obtain the most appropriate critical value available from the tabulated values. ${ }^{5}$ This two-step procedure can be a feasible practical solution for the implementation of the CADF tests as long as the estimators are consistent; in finite samples, however, the variability in the nuisance parameter estimate is not properly taken into account and more careful interpretation of the results is required.

The models with deterministic components can be analyzed similarly. When the time series $\left(x_{t}\right)$ with a nonzero mean is given by

$$
\begin{equation*}
x_{t}=\mu+y_{t} \tag{11}
\end{equation*}
$$

or with a linear time trend

$$
\begin{equation*}
x_{t}=\mu+\delta t+y_{t} \tag{12}
\end{equation*}
$$

where $\left(y_{t}\right)$ is generated as in (1), we may test for the presence of the unit root in the process $\left(y_{t}\right)$ from the CADF regression (3) defined with the fitted values $\left(y_{t}^{\mu}\right)$ or $\left(y_{t}^{\tau}\right)$ obtained from the preliminary regression (12) or (13). The limit theories for the CADF tests given in Theorem 2.3 extend easily to the models with nonzero mean and deterministic trends, and are given similarly with the following demeaned and detrended Brownian motions

$$
Q^{\mu}(s)=Q(s)-\int_{0}^{1} Q(t) d t
$$

and

$$
Q^{\tau}(s)=Q(s)+(6 s-4) \int_{0}^{1} Q(t) d t-(12 s-6) \int_{0}^{1} t Q(t) d t
$$

in the place of the Brownian motion $Q(s)$.

[^3]In the next section, we consider bootstrapping the covariates augmented tests $S_{n}$ and $T_{n}$ to deal with the nuisance parameter dependency problem and to provide a valid basis for inference based on the covariates augmented unit root tests.

## 3. Bootstrap Unit Root Tests with Covariates

In this section, we consider the bootstrap for the covariates augmented unit root tests $S_{n}$ and $T_{n}$ introduced in the previous section. We establish the bootstrap consistency of the tests and show the asymptotic validity of the tests. Throughout the paper, we use the usual notation $*$ to signify the bootstrap samples, and use $\mathbf{P}^{*}$ and $\mathbf{E}^{*}$ respectively to denote the probability and expectation conditional on a realization of the original sample. Various issues arising in practical implementation of the bootstrap methodology are also addressed.

To construct the bootstrap CADF tests, we first generate the bootstrap samples for the $m$-dimensional stationary covariates $\left(w_{t}\right)$ and the series $\left(y_{t}\right)$ to be tested. We begin by constructing the fitted residuals which will be used as the basis for generating the bootstrap samples. We first let $u_{t}=\triangle y_{t}$ and fit the regression

$$
\begin{equation*}
u_{t}=\sum_{k=1}^{p} \tilde{\alpha}_{k} u_{t-k}+\sum_{k=-r}^{q} \tilde{\beta}_{k}^{\prime} w_{t-k}+\tilde{\varepsilon}_{t} \tag{13}
\end{equation*}
$$

by the usual OLS regression. It is important to base the bootstrap sampling on regression (14) with the unit root restriction $\alpha=0$ imposed. The samples generated by regression (3) without the unit root restriction do not behave like unit root processes, and this will render the subsequent bootstrap procedures inconsistent as shown in Basawa et al. (1991).

Next, we fit the $\ell$-th order autoregression of $w_{t}$ as

$$
\begin{equation*}
w_{t+r+1}=\tilde{\Phi}_{1, n} w_{t+r}+\cdots+\tilde{\Phi}_{\ell, n} w_{t+r-\ell+1}+\tilde{\eta}_{t} \tag{14}
\end{equation*}
$$

by the usual OLS regression. We may prefer, especially in small samples, to use the YuleWalker method to estimate (15) since it always yields an invertible autoregression, thereby ensuring the stationarity of the process $\left(w_{t}\right)$ [see, e.g., Brockwell and Davis (1991, Sections 8.1 and 8.2 )]. As the sample size increases, however, the problem of noninvertibility in the OLS estimation vanishes a.s., and the two methods become equivalent. Our subsequent results are applicable also for the Yule-Walker method, since it is asymptotically equivalent to the OLS method.

We then generate the $(1+m)$-dimensional bootstrap samples $\left(\xi_{t}^{*}\right), \xi_{t}^{*}=\left(\varepsilon_{t}^{*}, \eta_{t}^{* \prime}\right)^{\prime}$ by resampling from the centered fitted residual vectors $\left(\tilde{\xi}_{t}\right), \tilde{\xi}_{t}=\left(\tilde{\varepsilon}_{t}, \tilde{\eta}_{t}^{\prime}\right)^{\prime}$ where $\left(\tilde{\varepsilon}_{t}\right)$ and $\left(\tilde{\eta}_{t}\right)$ are the fitted residuals from (14) and (15). That is, obtain iid samples $\left(\xi_{t}^{*}\right)$ from the empirical distribution of

$$
\left(\tilde{\xi}_{t}-\frac{1}{n} \sum_{t=1}^{n} \tilde{\xi}_{t}\right)_{t=1}^{n}
$$

The bootstrap samples $\underset{\sim}{\sim}\left(\xi_{t}^{*}\right)$ constructed as such will satisfy $\mathbf{E}^{*} \xi_{t}^{*}=0$ and $\mathbf{E}^{*} \xi_{t}^{*} \xi_{t}^{* \prime}=\tilde{\Sigma}$, where $\tilde{\Sigma}=(1 / n) \sum_{t=1}^{n} \tilde{\xi}_{t} \tilde{\xi}_{t}^{\prime}$. ${ }^{6}$

Next, we generate the bootstrap samples for $\left(w_{t}^{*}\right)$ recursively from $\left(\eta_{t}^{*}\right)$ using the fitted autoregression given by

$$
\begin{equation*}
w_{t+r+1}^{*}=\tilde{\Phi}_{1, n} w_{t+r}^{*}+\cdots+\tilde{\Phi}_{\ell, n} w_{t+r+1-\ell}^{*}+\eta_{t}^{*} \tag{15}
\end{equation*}
$$

with appropriately chosen $\ell$-initial values of $\left(w_{t}^{*}\right)$, where $\tilde{\Phi}_{k}, 1 \leq k \leq \ell$ are the coefficient estimates from the fitted regression (15). Initialization of $\left(w_{t}^{*}\right)$ is unimportant for our subsequent theoretical development, though it may play an important role in finite samples. ${ }^{7}$ Then we obtain $\left(w_{t+r}^{*}, \ldots, w_{t-q}^{*}\right)$ from the sequence $\left(w_{t}^{*}\right)$, and construct the bootstrap samples $\left(v_{t}^{*}\right)$ as

$$
\begin{equation*}
v_{t}^{*}=\sum_{k=-r}^{q} \tilde{\beta}_{k}^{\prime} w_{t-k}^{*}+\varepsilon_{t}^{*} \tag{16}
\end{equation*}
$$

using the LS estimates $\tilde{\beta}_{k},-r \leq k \leq q$ from the fitted regression (14). Then we generate $\left(u_{t}^{*}\right)$ recursively from $\left(v_{t}^{*}\right)$ using the fitted autoregression given by

$$
\begin{equation*}
u_{t}^{*}=\tilde{\alpha}_{1} u_{t-1}^{*}+\cdots+\tilde{\alpha}_{p} u_{t-p}^{*}+v_{t}^{*} \tag{17}
\end{equation*}
$$

with appropriately chosen $p$-initial values of $\left(u_{t}^{*}\right)$, and where $\tilde{\alpha}_{k}, 1 \leq k \leq p$ are the estimates for $\alpha_{k}$ 's from the fitted regression (14).

Finally, we generate $\left(y_{t}^{*}\right)$ from $\left(u_{t}^{*}\right)$ with the null restriction $\alpha=0$ imposed. This is to ensure the nonstationarity of the generated bootstrap samples $\left(y_{t}^{*}\right)$, which is claimed under the null hypothesis, and to make the subsequent bootstrap tests valid. Thus we obtain $\left(y_{t}^{*}\right)$ as

$$
\begin{equation*}
y_{t}^{*}=y_{t-1}^{*}+u_{t}^{*}=y_{0}^{*}+\sum_{k=1}^{t} u_{k}^{*} \tag{18}
\end{equation*}
$$

which also requires initialization $y_{0}^{*}$. An obvious choice would be to use the initial value $y_{0}$ of $\left(y_{t}\right)$, and generate the bootstrap samples $\left(y_{t}^{*}\right)$ conditional on $y_{0}$. As discussed earlier, the choice of initial value may affect the finite sample performance of the bootstrap; however, the effect of the initial value becomes negligible asymptotically as long as it is stochastically bounded. If the mean or linear time trend is maintained as in (12) or (13) and the unit root test is performed using the demeaned or detrended data, the effect of the initial value $y_{0}^{*}$ of

[^4]the bootstrap sample would disappear. We may therefore just set $y_{0}^{*}=0$ for the subsequent development of our theory in this section.

To construct the bootstrapped tests, we consider the following bootstrap version of the covariates augmented regression (3), which was used to construct the sample CADF tests $S_{n}$ and $T_{n}$ in the previous section

$$
\begin{equation*}
\triangle y_{t}^{*}=\alpha y_{t-1}^{*}+\sum_{k=1}^{p} \alpha_{k} \triangle y_{t-k}^{*}+\sum_{k=-r}^{q} \beta_{k}^{\prime} w_{t-k}^{*}+\varepsilon_{t}^{*} \tag{19}
\end{equation*}
$$

We test for the unit root hypothesis $\alpha=0$ in (20) using the bootstrap versions of the CADF tests, defined in (23) and (24) below, that are constructed analogously as their sample counterparts $S_{n}$ and $T_{n}$ defined in (9) and (10).

Similarly as before, we denote by $\hat{\alpha}_{n}^{*}$ and $s\left(\hat{\alpha}_{n}^{*}\right)$ respectively the OLS estimator for $\alpha$ and the standard error for $\hat{\alpha}_{n}^{*}$ obtained from the CADF regression (20) based on the bootstrap samples. To define them more explicitly, we let

$$
z_{t}^{*}=\left(\triangle y_{t-1}^{*}, \ldots, \Delta y_{t-p}^{*}, w_{t+r}^{* \prime}, \ldots, w_{t-q}^{* \prime}\right)^{\prime}
$$

and subsequently define

$$
\begin{aligned}
A_{n}^{*} & =\sum_{t=1}^{n} y_{t-1}^{*} \varepsilon_{t}^{*}-\left(\sum_{t=1}^{n} y_{t-1}^{*} z_{t}^{* \prime}\right)\left(\sum_{t=1}^{n} z_{t}^{*} z_{t}^{* \prime}\right)^{-1}\left(\sum_{t=1}^{n} z_{t}^{*} \varepsilon_{t}^{*}\right) \\
B_{n}^{*} & =\sum_{t=1}^{n} y_{t-1}^{* 2}-\left(\sum_{t=1}^{n} y_{t-1}^{*} z_{t}^{* \prime}\right)\left(\sum_{t=1}^{n} z_{t}^{*} z_{t}^{* \prime}\right)^{-1}\left(\sum_{t=1}^{n} z_{t}^{*} y_{t-1}^{*}\right)
\end{aligned}
$$

and the variance of the bootstrap sample $\left(\varepsilon_{t}^{*}\right)$, which is given by

$$
\begin{equation*}
\tilde{\sigma}_{n}^{2}=\frac{1}{n} \sum_{t=1}^{n}\left(\tilde{\varepsilon}_{t}-\bar{\varepsilon}_{n}\right)^{2}, \tag{20}
\end{equation*}
$$

where $\bar{\varepsilon}_{n}=n^{-1} \sum_{t=1}^{n} \tilde{\varepsilon}_{t}$. Then we may write the OLS estimator of $\alpha$ from the bootstrap CADF regression (20) and its estimated variance as

$$
\begin{aligned}
\hat{\alpha}_{n}^{*} & =A_{n}^{*} B_{n}^{*-1} \\
s\left(\hat{\alpha}_{n}^{*}\right)^{2} & =\tilde{\sigma}_{n}^{2} B_{n}^{*-1} .
\end{aligned}
$$

We also define, accordingly as $\hat{\alpha}_{n}(1)$ introduced in (8),

$$
\begin{equation*}
\tilde{\alpha}_{n}(1)=1-\sum_{k=1}^{p} \tilde{\alpha}_{k}, \tag{21}
\end{equation*}
$$

where $\tilde{\alpha}_{k}$ 's are the estimates for $\alpha_{k}$ 's from the fitted regression (14).

Now we consider the statistics

$$
\begin{align*}
S_{n}^{*} & =\frac{n \hat{\alpha}_{n}^{*}}{\tilde{\alpha}_{n}(1)}  \tag{22}\\
T_{n}^{*} & =\frac{\hat{\alpha}_{n}^{*}}{s\left(\hat{\alpha}_{n}^{*}\right)} \tag{23}
\end{align*}
$$

corresponding to $S_{n}$ and $T_{n}$ introduced in (9) and (10) of the previous section. For the construction of the bootstrap statistics $S_{n}^{*}$ and $T_{n}^{*}$, it is possible to replace $\tilde{\alpha}_{n}(1)$ and $\tilde{\sigma}_{n}^{2}$ with $\hat{\alpha}_{n}^{*}(1)$ and $\hat{\sigma}_{n}^{2 *}$, the bootstrap counterparts to $\hat{\alpha}_{n}(1)$ and $\hat{\sigma}_{n}^{2}$. We can compute $\hat{\alpha}_{n}^{*}(1)$ and $\hat{\sigma}_{n}^{2 *}$ from regression (20) in the same way that their sample counterparts are computed from regression (3). We may indeed show that such replacements do not affect the limiting distributions of the statistics. For the theoretical analysis in the paper, however, we only consider $S_{n}^{*}$ and $T_{n}^{*}$ defined in (23) and (24) for the expositional brevity. ${ }^{8}$

To implement the bootstrap CADF tests, we repeat the bootstrap sampling for the given original sample and obtain $a_{n}^{*}(\lambda)$ and $b_{n}^{*}(\lambda)$ such that

$$
\begin{equation*}
\mathbf{P}^{*}\left\{S_{n}^{*} \leq a_{n}^{*}(\lambda)\right\}=\mathbf{P}^{*}\left\{T_{n}^{*} \leq b_{n}^{*}(\lambda)\right\}=\lambda \tag{24}
\end{equation*}
$$

for any prescribed size level $\lambda$. The bootstrap CADF tests reject the null hypothesis of a unit root if

$$
S_{n} \leq a_{n}^{*}(\lambda), T_{n} \leq b_{n}^{*}(\lambda)
$$

It will now be shown under appropriate conditions that the tests are asymptotically valid, i.e., they have asymptotic size $\lambda$. We do not analyze in the paper the randomness associated with the bootstrap sampling in computing the bootstrap critical values $a_{n}^{*}(\lambda)$ and $b_{n}^{*}(\lambda)$. We simply assume that enough number of bootstrap iterations are carried out to make it negligible. See Andrews and Buchinsky (1999) for a study on the number of bootstrap iterations to achieve the desired level of bootstrap sampling accuracy.

We now introduce the notation $\rightarrow_{d^{*}}$ for bootstrap asymptotics. For a sequence of bootstrapped statistics $\left(Z_{n}^{*}\right)$, we write

$$
Z_{n}^{*} \rightarrow_{d^{*}} Z \text { a.s. }
$$

if the conditional distribution of $\left(Z_{n}^{*}\right)$ weakly converges to that of $Z$ a.s. Here it is assumed that the limiting random variable $Z$ has distribution independent of the original sample realization.

We now present the limit theories for the bootstrap CADF tests $S_{n}^{*}$ and $T_{n}^{*}$.
Theorem 3.3 Under the null hypothesis $\alpha=0$, we have as $n \rightarrow \infty$,

$$
S_{n}^{*} \rightarrow_{d^{*}} \sigma_{\varepsilon} \frac{\int_{0}^{1} Q(s) d P(s)}{\int_{0}^{1} Q(s)^{2} d s} \text { in } \mathbf{P}
$$

[^5]$$
T_{n}^{*} \rightarrow_{d^{*}} \frac{\int_{0}^{1} Q(s) d P(s)}{\left(\int_{0}^{1} Q(s)^{2} d s\right)^{1 / 2}} \text { in } \mathbf{P}
$$
under Assumptions 2.1 and 2.2 where $Q(s)$ and $P(s)$ are defined in Theorem 2.3, and 'in $\mathbf{P}^{\prime}$ signifies the usual convergence in probability.

Theorem 3.3 shows that the bootstrap statistics $S_{n}^{*}$ and $T_{n}^{*}$ have the same null limiting distributions as the corresponding sample statistics $S_{n}$ and $T_{n}$. It implies, in particular, that the bootstrap CADF tests are asymptotically valid.

To discuss the asymptotic validity of the tests using bootstrap critical values, denote by $S$ and $T$ the weak limits of $S_{n}$ and $T_{n}$ respectively, and define $a(\lambda)$ and $b(\lambda)$ to be the asymptotic critical values of the size $\lambda$ tests based on $S_{n}$ and $T_{n}$, i.e.,

$$
\mathbf{P}\{S \leq a(\lambda)\}=\mathbf{P}\{T \leq b(\lambda)\}=\lambda .
$$

Since the distributions of $S$ and $T$ are absolutely continuous with respect to Lebesgue measure, we have from Theorem 3.3

$$
\begin{equation*}
\mathbf{P}^{*}\left\{S_{n}^{*} \leq a(\lambda)\right\}, \mathbf{P}^{*}\left\{T_{n}^{*} \leq b(\lambda)\right\} \rightarrow \lambda \text { a.s. } \tag{25}
\end{equation*}
$$

under Assumptions 2.1 and 2.2, and the results in (26) imply

$$
\left(a_{n}^{*}(\lambda), b_{n}^{*}(\lambda)\right) \rightarrow(a(\lambda), b(\lambda)) \text { a.s. }
$$

where $a_{n}^{*}(\lambda)$ and $b_{n}^{*}(\lambda)$ are the size $\lambda$ bootstrap critical values defined in (25). Consequently, we have under Assumptions 2.1 and 2.2

$$
\mathbf{P}\left\{S_{n} \leq a_{n}^{*}(\lambda)\right\}, \mathbf{P}\left\{T_{n} \leq b_{n}^{*}(\lambda)\right\} \rightarrow \lambda
$$

as $n \rightarrow \infty$, which proves that the bootstrap CADF tests have size $\lambda$ asymptotically.
Our bootstrap theory here easily extends to the tests for a unit root in models with deterministic trends, such as those introduced in (12) or (13). It is straightforward to establish the bootstrap consistency for the CADF tests applied to the demeaned and detrended time series, using the results obtained in this section. The bootstrap CADF tests are therefore valid and applicable also for the models with deterministic trends.

## 4. Asymptotics under Local Alternatives

In this section, we consider local alternatives given by

$$
\begin{equation*}
H_{1}: \alpha=-\frac{c}{n} \tag{26}
\end{equation*}
$$

where $c>0$ is a fixed constant, and let ( $y_{t}$ ) be generated by (1) and (2). The asymptotic theories for the local-to-unity models are now well established [see, e.g., Stock (1994)], and the following limit theories are easily derived from them for our model:

$$
\begin{align*}
& S_{n} \rightarrow_{d} \quad S(c)=-c+\sigma_{\varepsilon} \frac{\int_{0}^{1} Q_{c}(s) d P(s)}{\int_{0}^{1} Q_{c}(s)^{2} d s}  \tag{27}\\
& T_{n} \rightarrow_{d} \quad T(c)=-\frac{c}{\sigma_{\varepsilon}}\left(\int_{0}^{1} Q_{c}(s)^{2} d s\right)^{1 / 2}+\frac{\int_{0}^{1} Q_{c}(s) d P(s)}{\left(\int_{0}^{1} Q_{c}(s)^{2} d s\right)^{1 / 2}} \tag{28}
\end{align*}
$$

where

$$
Q_{c}(s)=Q(s)-c \int_{0}^{1} e^{-c(s-r)} Q(r) d r
$$

is Ornstein-Uhlenbeck process, which may be defined as the solution to the stochastic differential equation $d Q_{c}(s)=-c Q_{c}(s) d s+d Q(s)$, and $Q$ is defined in Theorem 2.3.

Bootstrap theories for the local-to-unity models are established in Park (2003). Here we may follow Park (2003) to obtain the limit theories for the bootstrap statistics $S_{n}^{*}$ and $T_{n}^{*}$ for our local-to-unity models.

Theorem 4.1 Let Assumptions 2.1 and 2.2 hold. Then under the local alternatives (27), we have

$$
\begin{array}{lll}
S_{n}^{*} & \rightarrow_{d^{*}} & S \text { in } \mathbf{P} \\
T_{n}^{*} & \rightarrow_{d^{*}} & T \text { in } \mathbf{P}
\end{array}
$$

as $n \rightarrow \infty$, where $S$ and $T$ are the limiting null distributions of $S_{n}$ and $T_{n}$ given in Theorem 2.3.

We note that the limiting distributions of the bootstrap statistics $S_{n}^{*}$ and $T_{n}^{*}$ under the local alternatives are identical to their limiting distributions under the exact unit root null. This is, however, well expected, since the bootstrap samples are generated under the unit root restriction regardless of the true data generating mechanism, thereby forcing them to behave asymptotically as the unit root processes. Note also that

$$
\begin{equation*}
\mathbf{P}\{S(c) \leq x\}>\mathbf{P}\{S \leq x\}, \mathbf{P}\{T(c) \leq x\}>\mathbf{P}\{T \leq x\} \tag{29}
\end{equation*}
$$

for all $x \in \mathbf{R}$, and therefore we may expect that the unit root testing based on the tests $S_{n}$ and $T_{n}$ have some discriminatory powers against the local-to-unity model. In particular, we have under the alternative of the local-to-unity model,

$$
\mathbf{P}^{*}\left\{S \leq a_{n}^{*}(\lambda)\right\}, \mathbf{P}^{*}\left\{T \leq b_{n}^{*}(\lambda)\right\} \rightarrow_{a . s .} \lambda
$$

as $n \rightarrow \infty$, which in turn gives

$$
\begin{aligned}
\lim _{n \rightarrow \infty} \mathbf{P}^{*}\left\{S_{n} \leq a_{n}^{*}(\lambda)\right\} & =\lim _{n \rightarrow \infty} \mathbf{P}^{*}\left\{S(c) \leq a_{n}^{*}(\lambda)\right\}>\lambda \text { a.s. } \\
\lim _{n \rightarrow \infty} \mathbf{P}^{*}\left\{T_{n} \leq b_{n}^{*}(\lambda)\right\} & =\lim _{n \rightarrow \infty} \mathbf{P}^{*}\left\{T(c) \leq b_{n}^{*}(\lambda)\right\}>\lambda \text { a.s. }
\end{aligned}
$$

due to (28)-(30). Thus, the bootstrap unit root tests have non-trivial powers against the local-to-unity model.

## 5. Simulations

### 5.1 Data Generating Process

In this section, we perform a set of simulations to investigate the performances of the bootstrap tests. For the comparison of the bootstrap tests with other well-known tests, we consider only $T_{n}^{*}$ statistic here. For the simulations, we consider $\left(y_{t}\right)$ given by the unit root model (1) with the error $\left(u_{t}\right)$ generated by

$$
u_{t}=\alpha_{1} u_{t-1}+v_{t},
$$

where the error term $\left(v_{t}\right)$ is given by

$$
\begin{equation*}
v_{t}=\beta w_{t}+\varepsilon_{t} . \tag{30}
\end{equation*}
$$

We model the covariate $\left(w_{t}\right)$ to follow an $\mathrm{AR}(1)$ process as follows:

$$
\begin{equation*}
w_{t+1}=\phi w_{t}+\eta_{t} . \tag{31}
\end{equation*}
$$

The innovations $\left(\xi_{t}\right), \xi_{t}=\left(\varepsilon_{t}, \eta_{t}\right)^{\prime}$ are randomly chosen from iid $\mathbf{N}(0, \Sigma)$, where

$$
\Sigma=\left(\begin{array}{cc}
1 & \sigma_{\varepsilon \eta} \\
\sigma_{\eta \varepsilon} & 1
\end{array}\right)
$$

Under this setup, we have the following covariate augmented ADF regression:

$$
\begin{equation*}
\Delta y_{t}=\alpha y_{t-1}+\alpha_{1} \Delta y_{t-1}+\beta w_{t}+\varepsilon_{t} . \tag{32}
\end{equation*}
$$

The relative merit of constructing a unit root test from the covariate augmented regression depends on the correlation between the error $\left(v_{t}\right), v_{t}=\beta w_{t}+\varepsilon_{t}$ and the covariate $\left(w_{t}\right)$. As can be seen clearly from (31) and (32), the correlation depends on two parameter values, the coefficient $\beta$ on the covariate and the AR coefficient $\phi$ of the covariate. We thus control the degree of correlation between the error $\left(v_{t}\right)$ and the covariate $\left(w_{t}\right)$ through these parameters. The values of $\beta$ and $\phi$ are allowed to vary among $\{-0.8,-0.5,0.5,0.8\}$. The coefficient $\alpha_{1}$ on the lagged difference term is set at 0.2 throughout the simulations. The contemporaneous covariance $\sigma_{\varepsilon \eta}$ is set at 0.4 . For the test of the unit root hypothesis, we set $\alpha=0$ and investigate the sizes in relation to corresponding nominal test sizes. For the powers, we consider $\alpha=-0.10 .{ }^{9}$

[^6]
### 5.2 Asymptotic Properties

In this section, the asymptotic size properties of the CADF and the bootstrap CADF tests are compared. The size properties of the CADF test critically depends on the consistent estimation of the correlation coefficient $\rho^{2}$. Unless $\rho^{2}$ is consistently estimated, the CADF test would be biased even in asymptotics. Therefore, more specifically, we first examine whether the correlation coefficient $\rho^{2}$ is precisely estimated in the CADF test. Next, the performances of the CADF test based on the estimated $\hat{\rho}^{2}$ are evaluated and compared with those of the bootstrap CADF test. To implement the CADF test, sample correlation coefficient $\hat{\rho}^{2}$ is estimated and the asymptotic critical values corresponding to each sample estimate of $\rho^{2}$ are read from Table 1 in Hansen (1995). ${ }^{10}$ The bootstrap CADF test, however, does not depend on the asymptotic critical values and uses bootstrapped critcal values instead. The regression equation for the CADF test is based on the true model and it contains one lagged difference term and the current value of the covariate. The regression equation for covariate is estimated using the $\operatorname{AR}(1)$ model as in (32).

Given our model specifications in Section 5.1, $\rho^{2}=\sigma_{v \varepsilon}^{2} /\left(\sigma_{v}^{2} \sigma_{\varepsilon}^{2}\right)$ is calculated as follows:

$$
\begin{aligned}
\sigma_{v}^{2} & =\frac{\beta^{2} \sigma_{\eta}^{2}}{(1-\phi)^{2}}+\sigma_{\varepsilon}^{2}+\frac{2 \beta \sigma_{\eta \varepsilon}}{1-\phi} \\
\sigma_{v \varepsilon} & =\frac{\beta \sigma_{\eta \varepsilon}}{1-\phi}+\sigma_{\varepsilon}^{2}
\end{aligned}
$$

where $\sigma_{\varepsilon}^{2}=1, \sigma_{\eta}^{2}=1$ and $\sigma_{\eta \varepsilon}=0.4$. Then, for the parameters we consider, the true $\rho^{2}$ varies from 0 to 0.950 . Now, we can compare the estimated $\hat{\rho}^{2}$ with true $\rho^{2}$ under the simulation setup as shown in Table 1. The table shows that, in finite samples (for $n=25,50,100$ ), there are large biases in $\hat{\rho}^{2}$ especially when $\rho^{2}$ is low. For example, when the true $\rho^{2}$ is 0.0 , the estimated $\hat{\rho}^{2}$ is 0.206 . Since the CADF test attains potential power gains at low levels of $\rho^{2}$, the size distortions possibly caused by these biases pose serious problems to the use of the test. Moreover, these biases do not seem to vanish even for large $n$ (e.g., $n=3,000$ ) for some parameter values. For example, for $(\beta=-0.5, \phi=0.5)$ and ( $\beta=-0.8, \phi=0.5$ ), the differences between $\rho^{2}$ and $\hat{\rho}^{2}$ are non-negligible. From these results, we conclude that the CADF test may suffer from size distortions that come from the imprecise estimation of $\rho^{2}$.

Next, we examine the size performances of the tests in Table 2. As the sample size increases, the overall size performances of the CADF test improves and the sizes are close to $5 \%$. However, as mentioned above, when $\rho^{2}$ are incorrectly estimated for ( $\beta=-0.5, \phi=$ 0.5 ) and ( $\beta=-0.8, \phi=0.5$ ), the CADF test tends to slightly underreject. Moreover, although we have correct estimate $\hat{\rho}^{2}$ for large $n$, the CADF test still shows large size distortions. For example, for $(\beta=-0.5, \phi=0.8), \hat{\rho}^{2}$ is close to zero, the true value of $\rho^{2}$, but the size of the CADF test is only $1 \%$ (for $n=3,000$ ). Again, this is a serious drawback of the CADF test because the CADF test are the most useful in terms of good power performance at low levels of $\rho^{2}$. Therefore, when $\rho^{2}$ is low, the CADF test, which is based on $\hat{\rho}^{2}$, shows unreliable results even in large samples.

[^7]In contrast, the bootstrap CADF test does not depend on the estimated $\hat{\rho}^{2}$ for choosing critical values and it uses, instead, bootstrapped critical values for the test. As shown in the Table 2, the sizes of the bootstrap CADF test are more stable along various parameter values than the CADF test. In particular, for the parameters that we considered above, the bootstrap CADF test shows good size properties. For example, for $(\beta=-0.5, \phi=0.8)$, the size of the bootstrap CADF test is $5.1 \%$ while that of the CADF test is only $1.4 \%$. The bootstrap CADF test tends to slightly overreject for some parameters such as $(\beta=$ $-0.5, \phi=-0.8$ ), but this is not our concern because it corresponds to the case where $\hat{\rho}^{2}$ is very high and the CADF tests are the least useful in terms of power performance. Based on these experiments, we conclude that the bootstrap CADF test shows more reliable size performances even in large samples than the CADF test.

### 5.3 Finite Sample Properties

The finite sample performances of the bootstrap CADF test are compared with those of the sample CADF test computed from the regression (33) as well as other well-known unit root tests. More specifically, in addition to the CADF test, we also consider another CADF test suggested by Elliott and Jansson (2003) (called the EJ test here). This test is known to have maximal power against a point alternative. Thus, these three tests are all covariate-augmented and the comparisons of their size and power performances would be meaningful.

As a benchmark, we consider the ADF test based on the usual ADF regression. The ADF regression does not include the covariate $\left(w_{t}\right)$ as a regressor, and thus the regression error effectively becomes $v_{t}=\beta w_{t}+\varepsilon_{t}$ with the DGP considered in our simulation setup. Hence, the conventional ADF test suffers from efficiency loss by failing to utilize additional information coming from $\left(w_{t}\right)$. Moreover, in our simulation setup, the effective error $\left(v_{t}\right)$ is obviously serially correlated due to serial correlation in $\left(w_{t}\right)$. This is another source of inefficiency that may be handled by increasing the number of lagged differences, thus whitening the error $\left(v_{t}\right)$.

Choice of lag lengths critically affects the finite sample properties of the tests. To investigate the effects of lag length selection on the finite sample performances of the tests, we use popular lag length selection methods. For the ADF and the CADF tests, AIC was used and, for the EJ test, BIC was used as suggested by Elliott and Jansson (2003). Maximum lag length is set at four for $n=50$ and 100 and two for $n=25$. For the choice of the lead and lag lengths of covariate of the CADF test, maximum lengths are set at four.

The lead and lag lengths of the covariates are chosen in such a way that $\hat{\rho}^{2}$ is minimized. It is worth noting that such choice of covariate does not necessarily induce pretest bias in finite samples. As mentioned above, the consistent estimation of $\rho^{2}$ affects the size properties of the CADF test whereas the choice of the covariates with low $\hat{\rho}^{2}$ is intended to improve powers. In other words, choosing the covariates based on $\hat{\rho}^{2}$ affects the power performance but it does not necessarily introduce pretest bias for size as long as $\rho^{2}$ is consistently estimated. On the other hand, this means that imprecise estimaes of $\rho^{2}$ may cause problems both in size and powers for the CADF test. The size of the bootstrap CADF
test, however, is not affected but the power may be by the imprecise estimates of $\rho^{2}$.
Other well-known univariate unit root tests are also considered. Ng and Perron (2001) argue that the MIC information criterion along with GLS-detrended data yields a set of tests with desirable size and power properties. In light of their argument, we calculate the following tests based on the GLS-detrended data for both the statistic and the spectral density, and select the lag lengths by the MIC, with the lower bound zero and the upper bound given by $\operatorname{int}\left(12(n / 100)^{1 / 4}\right)$. The tests considered are the $\mathrm{Z}_{\alpha}$ test by Phillips and Perron (1988), $\mathrm{M}^{\mathrm{GLS}}$ tests as discussed in Ng and Perron (2001), $\mathrm{DF}^{\mathrm{GLS}}$ test and feasible point optimal test $\left(\mathrm{P}_{\mathrm{t}}\right)$ by Elliott, Rothenberg and Stock (1996), and modified point optimal test $\left(\mathrm{MP}_{\mathrm{t}}\right)$ by Ng and Perron (2001). ${ }^{11}$

All regressions include a fitted intercept, and the results when including a time trend are also provided. Sample sizes of $n=25,50$, and 100 are examined for $5 \%$ nominal size tests. Size-adjusted powers are reported where sizes are controlled by using the finite sample critical values. The reported results are based on 3,000 simulation iterations with the bootstrap critical values computed from 3,000 bootstrap repetitions. Each replication discards the first 100 observations to eliminate start-up effects. The finite sample sizes and powers for the tests are reported in Tables 3-14.

Tables 3-5 show the size results for the tests when only constant is included and Tables $6-8$ show the size results when a time trend is also included. As can be seen clearly, the sample CADF test has quite noticeable size distortions over various parameters especially for small samples $(n=25)$. For example, sizes are higher than $10 \%$ and in some cases it reaches $19 \%$. On one hand, these distortions imply that the limit distribution of the CADF test poorly approximates the finite distribution. On the other hand, this also implies that the choice of lag lengths produces a lot of variability of the estimates and add to the size distortions of the CADF test. The distortions are even larger when a time trend is included. Hence, the CADF test shows severe size distortions especially for small samples and with a time tend.

The boostrap CADF test substantially correct the biases of the CADF test particularly when $\hat{\rho}^{2}$ is low. In Table 3 , for example, when $(\beta=0.500, \phi=0.800)$ for $\hat{\rho}^{2}=0.126$, the size of the CADF test is $11.8 \%$ while that of the bootstrap CADF test is $6.4 \%$. Thus, even in finite samples, the bootstrap CADF test shows reasonably good size performances when $\hat{\rho}^{2}$ is low. This improvement of the size performance is much conspicuous when a time trend is included and for small samples. When the sample size is $n=100$, however, both tests show similar size performances and the size improvements of the bootstrap test almost disappear. When $\hat{\rho}^{2}$ is high, the performances of the bootstrap test also becomes worse. But, as we mentioned above, that is not our concern because the CADF tests usually work best for low $\rho^{2}$.

The size performances of the EJ test is the most unstable among the considered tests. For example, the size distortions of the EJ test for small sample are hugh as high as $38.6 \%$ in Table 4. Moreover, the test becomes more unstable when a time trend is included. Thus, even for $n=100$, sometimes the sizes are over $75 \%$ and in other cases they are very close to

[^8]$5 \%$. Hence, the size performances the EJ test are very unstable across different parameters. Therefore, the EJ test is unrealiable especially when a time trend is included and for small samples.

The size performances of the ADF test are as good as those of the bootstrap CADF test. This implies that ignoring covariates does not significantly affect the size properties of the ADF test. The difference is that the ADF test shows similar size performances across various parameters whereas the bootstrap CADF test shows better size performances when $\hat{\rho}^{2}$ is low than when $\hat{\rho}^{2}$ is high. The other tests shows reasonably good size properties as sample size increases but they show quite unstable results when a time trend is included. Thus, for some parameters, they tend to overreject and, for others, they tend to severely underreject.

In summary, although the null distribution of the CADF test is asymptotically valid, the finite sample performances are poor. On the other hand, the bootstrap test shows good size performance in finite samples and the performances are similar to those of the ADF test especially when $\hat{\rho}^{2}$ is low. Other tests show reasonably good size performances when only a constant is included, but they show unstble results when a time trend is included, over-or underrejecting the null. Therefore, we conclude that only the ADF test and the bootstrap CADF test show reliable and satisfying size performances.

The Tables $9-11$ show the results of powers when a contant is included and the Tables 12-14 show those when a time trend is included. The significant improvement in the finite sample sizes that the bootstrap CADF test offers does not come at the expense of finite sample powers. Indeed, the results in Tables 9-14 show that the bootstrap CADF test offers drastic power gains over the conventional ADF test when $\hat{\rho}^{2}$ is low, where the covariates tests are expected to improve the power properties. The powers of the bootstrap CADF are more than two or three times as large as those of the other tests when $\hat{\rho}^{2}$ is low. Moreover, the powers of the bootstrap CADF test are comparable to those of the CADF test and sometimes even larger than those of the CADF test especially when a time trend is included. The EJ test has the highest nominal power but its size-adjusted power is similar to those of the CADF tests.

As discussed above, the estimate of the correlation coefficient $\rho^{2}$ affects the power performances of the CADF test because the covariates are chosen in a way that minimizes the estimated $\hat{\rho}^{2}$. Since $\hat{\rho}^{2}$ may be imprecisely estimated, the (bootstrap) CADF test may suffer from potential power loss. However, as shown in the simulation results above, the powers of the (bootstrap) CADF tests are much higher than those of any other univariate tests. Hence, we conclude that the effects of $\hat{\rho}^{2}$ on the power performances of the (bootstrap) CADF tests would be minor.

In contrast, the other tests show considerably lower powers compared with those of three covariate-based tests. The powers of the ADF test are lowest among the tests considered and the other tests show similar power performances. In particular, when a time trend is included, the other tests substantially lose powers. These results show that using covariates may bring enormous power gains over other univariate unit root tests.

In summary, the bootstrap CADF test has good size and power properties for all combinations of parameter values and time series dimensions and is robust to the inclusion of
a time trend. From all these observations, we conclude that the bootstrap CADF test has the best size and power properties under our simulation setup.

## 6. Empirical Applications

We next apply our testing procedures to a set of macroeconomic aggregates and real exchange rates. A number of econometric studies have found that standard tests for a unit root, such as the ADF tests (Dickey and Fuller, 1979; Said and Dickey, 1984) and the PP test (Phillips and Perron, 1988), have low power against stationary alternatives in the relatively small samples we consider in this section [see Dejong, Nankervis, Savin, and Whiteman (1992a, 1992b), among others]. This is especially true when a series under investigation is a near-integrated process. Since the low power of the univariate unit root tests is the primary problem, it is important to investigate whether or not the null hypothesis of a unit root is rejected by more powerful tests. Moreover, various univariate unit root tests provide mixed results for a given time series when they do not consistently reject or accept the null hypothesis. In such a situation making a definitive conclusion about the (non)stationarity of the time series may be problematic. Hence, the use of more powerful tests may point to sharper conclusions about the stationarity property of the particular time series.

In our empirical applications, we consider two data sets. The first is the Nelson and Plosser (1982) data set extended by Schotman and Van Dijk (1991). The second is annual CPI-based real exchange rates for fifteen OECD countries from 1950 to 1998. The real exchange rate, $r_{i t}$, for the $i$-th country is computed using the U.S. dollar as numeraire currency. ${ }^{12}$ Real exchange rates are analyzed for Australia, Austria, Belgium, Canada, France, Finland, Germany, Italy, Japan, Luxemburg, the Netherlands, Norway, Spain, Switzerland and the United Kingdom.

The testing strategy is as follows. We use lagged differences of each series for covariates, thus, only stationary covariates will be utilized in our multivariate tests. ${ }^{13}$ Among the candidates for covariates we choose the one which gives us the smallest $\hat{\rho}^{2}$ since this covariate provides the most powerful test, as shown in Section 5. All the regressions include a constant. The lags of the differenced dependent variable are selected using the Akaike Information Criterion (AIC) with the maximum lag length four. ${ }^{14}$ For the CADF tests, current covariate is included and the combinations of past and future covariates are tried up to the lag length four, among which the lag lengths with the smallest $\hat{\rho}^{2}$ are chosen. For the bootstrap tests, we use critical values computed from 5,000 bootstrap iterations.

[^9]
### 6.1 The Nelson and Plosser Data Set

The Nelson and Plosser data set is one of the most widely analyzed macroeconomic aggregate time series data sets. Nelson and Plosser (1982) studied the time series properties of fourteen series and found that all of them, except the unemployment series, were characterized by stochastic nonstationarity. We apply the aforementioned unit root tests to the fourteen time series whose nonstationarity have been questioned. All variables in the data set are measured annually in natural logarithms. The estimated period is 1929-1973 in consideration of the structural breaks in 1929 and 1973 coinciding with the onset of the Great Depression and oil shock [see Perron (1989)]. A time trend is included in the regressions. Table 13 presents the results.

For all cases the values of $\hat{\rho}^{2}$ are lower than 0.09 , thus we should expect, based on our simulation results, more powerful test results with the CADF and the bootstrap CADF tests than with the other tests. With these new tests we can reject the null hypothesis of a unit root for five series (GNP Deflator, Wages, Money Stock, Velocity and S\&P500) by the sample CADF test and three series (GNP Deflator, Money Stock and Velocity) by the bootstrap CADF test. Looking at the other tests, the EJ test rejects for eight series and the $\mathrm{DF}^{\mathrm{GLS}}$ test rejects for two series. The other tests reject the null hypothesis for only one series or none of the series.

The results for other tests are not surprising because the simulation results for $n=50$ with a time trend show that the powers of other tests are very low. Also, for such small $\hat{\rho}^{2}$ as our data set, the EJ test as well as the CADF test tend to severely overreject ${ }^{15}$. On the other hand, the bootstrap CADF test shows reasonable size and power performances. Thus, we may conclude that the results from the EJ test and the CADF test are less reliable and accept the results from the bootstrap CADF test that there are three stationary series in the Nelson and Plosser data set for the considered sample period.

### 6.2 Real Exchange Rates

Another situation in which the covariate tests are especially useful is with certain types of panel data in which cross-sectional correlations between time series are present. If this information can be properly modelled, then it can provide an efficiency gain over univariate methods as pointed out by Hansen (1995). To illustrate these potential efficiency (and power) gains, we analyze annual CPI-based real exchange rates for fifteen OECD countries for the period of 1950-1998, the period including the Bretton Woods system and flexible exchange rate regime.

Movements in real exchange rates are thought to be driven primarily by deviations from purchasing power parity (PPP), and models of exchange rate determination are built on the assumption that the PPP hypothesis holds. There are, however, conflicting empirical evidences. Studies by MacDonald (1996), Frankel and Rose (1996), Oh (1996), Papell (1997), O'Connell (1998), and more recently, Cheung, Lai, and Bergman (2004), Choi (2004) and

[^10]Chortareas and Kapetanios (2009) among others, suggest that the issue is not completely settled. In particular, when considering data for the recent flexible rate experience (1973present), many researchers have been unable to reject the null hypothesis of a unit root [see, e.g., Papell (1997), Papell and Theodoridis (1998), and O'Connell (1998)]. One response to this nonrejection might be that the tests do not encompass a sufficiently long time span to capture the mean reversion necessary to reject the null hypothesis. It is of interest, therefore, if real exchange rates are mean reverting for longer spans of time than the period of recent float regime, using low frequency data. One advantage of using relatively low frequency annual or quarterly data is that it can possibly increase the power of statistical tests for random walk behavior (Shiller and Perron, 1985).

In the estimated equation, time trend is not included because such an inclusion would be theoretically inconsistent with the long-run PPP [see Papell (2002)]. Table 14 shows the results for the real exchange rates. The estimated $\hat{\rho}^{2}$ are lower than 0.15 . As anticipated, the tests using the covariates are able to reject for more countries than other univariate tests, consistent with the increased power of the covariate-based tests. Thus, the EJ test rejects for four countries and the CADF and the bootstrap CADF tests for six countries whereas the other tests reject for only one or three or even none of the countries. In our sample for real exchange rates, it happened that the countries that turned out to be stationary from the CADF test and the bootstrap CADF test coincided with each other. Thus the two tests give us consistent results. Unlike the case of the Nelson and Plosser data set, the EJ test does not reject the null hypothesis more often than the CADF tests. As shown in the simulations, the bootstrap CADF test shows the best size and power properties, and we conclude that the real exchange rates of Australia, Canada, Finland, Italy, the Netherlands and Norway are stationary.

## 7. Conclusion

In this paper, we consider the bootstrap procedure for the covariate augmented DickeyFuller (CADF) unit root test which substantially improves the power of univariate unit root tests. Hansen (1995) originally proposed the CADF test and suggested a two-step procedure to overcome the nuisance parameter dependency problem. Here, we propose bootstrapping the CADF test in order to directly deal with the nuisance parameter dependency and base inferences on the bootstrapped critical values. We also establish the bootstrap consistency of the CADF test and show that the bootstrap CADF test is asymptotically valid.

The asymptotic properties of the CADF and bootstrap CADF tests are investigated and the finte sample performances of the CADF tests are compared with various well-known univariate unit root tests through simulations. The bootstrap CADF test significantly improves the asymptotic and the finite sample size performances of the CADF test, especially when the covariates are highly correlated with the error. Indeed, the bootstrap CADF test offers drastic power gains over the conventional ADF and other univariate tests. As illustrations, we apply the tests to the fourteen macroeconomic time series in the Nelson and Plosser data set for the post-1929 samples as well as to postwar annual CPI-based real
exchange rates for fifteen OECD countries. In contrast to the results from univariate unit root tests, our empirical results show that the null hypothesis of a unit root is rejected for more series by the CADF tests.

## 8. Appendix

Lemma A. 1 Under Assumption 2.1, we have as $n \rightarrow \infty$
(a) $\frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]} w_{t} \rightarrow_{d} \Psi(1) B_{\eta}(s)$
(b) $\frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]} u_{t} \rightarrow_{d} \pi(1)\left(\beta(1)^{\prime} \Psi(1) B_{\eta}(s)+B_{\varepsilon}(s)\right)$
for $s \in[0,1]$, where $\Psi(1)=\Phi(1)^{-1}$ and $\pi(1)=1 / \alpha(1)$.
Proof of Lemma A. 1 To establish the stated results, we use the Beveridge-Nelson (BN) representations for the finite order lag polynomials $\alpha(L), \beta(L)$ and $\Phi(L)$ defined in (2) and (4) and the limit theory from the invariance principle given in (5).

Part (a) We begin by deriving the BN representation for $\left(w_{t}\right)$ from (4). Let $\Phi(1)=$ $I_{m}-\sum_{k=1}^{\ell} \Phi_{k}$. Then we may easily get

$$
\Phi(1) w_{t}=\eta_{t-r-1}+\sum_{k=1}^{\ell} \sum_{j=k}^{\ell} \Phi_{j}\left(w_{t-k}-w_{t-k+1}\right)
$$

or

$$
\begin{equation*}
w_{t}=\Psi(1) \eta_{t-r-1}+\left(\bar{w}_{t-1}-\bar{w}_{t}\right), \tag{33}
\end{equation*}
$$

where $\Psi(1)=\Phi(1)^{-1}$ and $\bar{w}_{t}=\sum_{k=1}^{\ell} \bar{\Phi}_{k} w_{t-k+1}$, with $\bar{\Phi}_{k}=\Psi(1) \sum_{j=k}^{\ell} \Phi_{j}$. Under our condition in Assumption 2.1, $\left(\bar{w}_{t}\right)$ is well defined both in a.s. and $L^{\gamma}$ sense [see Brockwell and Davis (1991, Proposition 3.1.1)]. Then we have

$$
\sum_{k=1}^{t} w_{k}=\Psi(1) \sum_{k=1}^{t} \eta_{k-r-1}+\left(\bar{w}_{0}-\bar{w}_{t}\right) .
$$

Note that $\left(\bar{w}_{t}\right)$ is stochastically of smaller order of magnitude than the sum $\sum_{k=1}^{t} \eta_{k}$, and hence will become negligible in the limit. Then it follows directly from (5) that

$$
\frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]} w_{t}=\Psi(1) \frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]} \eta_{t-r-1}+\frac{1}{\sqrt{n}}\left(\bar{w}_{0}-\bar{w}_{[n s]}\right) \rightarrow_{d} \Psi(1) B_{\eta}(s)
$$

for $s \in[0,1]$, giving the stated result in part (a).
Part (b) Let $\alpha(1)=1-\sum_{k=1}^{p} \alpha_{k}$. Similarly, we derive the BN representation for ( $u_{t}$ ) from (2) as follows

$$
u_{t}=\frac{1}{\alpha(1)}\left(\beta(L)^{\prime} w_{t}+\varepsilon_{t}\right)+\sum_{k=1}^{p} \frac{\sum_{j=k}^{p} \alpha_{j}}{\alpha(1)}\left(u_{t-k}-u_{t-k+1}\right)
$$

$$
=\pi(1)\left(\beta(L)^{\prime} w_{t}+\varepsilon_{t}\right)+\left(\bar{u}_{t-1}-\bar{u}_{t}\right)
$$

where $\pi(1)=1 / \alpha(1)$ and $\bar{u}_{t}=\sum_{k=1}^{p} \bar{\alpha}_{k} u_{t-k+1}$, with $\bar{\alpha}_{k}=\pi(1) \sum_{j=k}^{p} \alpha_{j}$. The process $\left(\bar{u}_{t}\right)$ is also well defined both in a.s. and $L^{\gamma}$ sense. We may also obtain the BN representation for $\beta(L)^{\prime} w_{t}$ as follows

$$
\begin{aligned}
\beta(L)^{\prime} w_{t} & =\beta(1)^{\prime} w_{t}+\sum_{k=0}^{q-1} \sum_{j=k+1}^{q} \beta_{j}^{\prime}\left(w_{t-k-1}-w_{t-k}\right)+\sum_{k=0}^{r-1} \sum_{j=k+1}^{r} \beta_{-j}^{\prime}\left(w_{t+k+1}-w_{t+k}\right) \\
& =\beta(1)^{\prime} w_{t}+\left(\bar{w}_{t-1}^{+}-\bar{w}_{t}^{+}\right)+\left(\bar{w}_{t}^{-}-\bar{w}_{t-1}^{-}\right)
\end{aligned}
$$

where $\bar{w}_{t}^{+}=\sum_{k=0}^{q-1} \bar{\beta}_{k}^{+\prime} w_{t-k}$ and $\bar{w}_{t}^{-}=\sum_{k=0}^{r-1} \bar{\beta}_{k}^{-\prime} w_{t+k+1}$, with $\bar{\beta}_{k}^{+}=\sum_{j=k+1}^{q} \beta_{j}$ and $\bar{\beta}_{k}^{-}=$ $\sum_{j=k+1}^{r} \beta_{-j}$. The $\left(\bar{w}_{t}^{+}\right)$and $\left(\bar{w}_{t}^{-}\right)$are well defined both in a.s. and $L^{\gamma}$ sense.

Then it follows that

$$
\begin{equation*}
u_{t}=\pi(1)\left(\beta(1)^{\prime} w_{t}+\varepsilon_{t}\right)+\pi(1)\left(\left(\bar{w}_{t-1}^{+}-\bar{w}_{t}^{+}\right)+\left(\bar{w}_{t}^{-}-\bar{w}_{t-1}^{-}\right)\right)+\left(\bar{u}_{t-1}-\bar{u}_{t}\right) \tag{34}
\end{equation*}
$$

and

$$
\frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]} u_{t}=\pi(1)\left(\beta(1)^{\prime} \frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]} w_{t}+\frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]} \varepsilon_{t}\right)+o_{p}(1)
$$

since

$$
\begin{aligned}
\frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]}\left(\bar{u}_{t-1}-\bar{u}_{t}\right) & =\frac{1}{\sqrt{n}}\left(\bar{u}_{0}-\bar{u}_{[n s]}\right)=o_{p}(1) \\
\frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]}\left(\left(\bar{w}_{t-1}^{+}-\bar{w}_{t}^{+}\right)+\left(\bar{w}_{t}^{-}-\bar{w}_{t-1}^{-}\right)\right) & =\frac{1}{\sqrt{n}}\left(\bar{w}_{0}^{+}-\bar{w}_{[n s]}^{+}\right)+\frac{1}{\sqrt{n}}\left(\bar{w}_{[n s]}^{-}-\bar{w}_{0}^{-}\right)=o_{p}(1) .
\end{aligned}
$$

Now the stated result is immediate from the invariance principle given in (5) and the result in part (a).

Lemma A. 2 Under the null hypothesis $\alpha=0$, we have as $n \rightarrow \infty$,
(a) $n^{-2} \sum_{t=1}^{n} y_{t-1}^{2} \rightarrow_{d} \pi(1)^{2} \int_{0}^{1} Q(s)^{2} d s$
(b) $n^{-1} \sum_{t=1}^{n} y_{t-1} \varepsilon_{t} \rightarrow_{d} \sigma_{\varepsilon} \pi(1) \int_{0}^{1} Q(s) d P(s)$
under Assumption 2.1, where $Q$ and $P$ are defined in Theorem 2.3.
Proof of Lemma A. 2 When $\alpha=0$, we have

$$
y_{t}=y_{t-1}+u_{t}=\sum_{k=1}^{t} u_{k}
$$

since $y_{0}=0$. Then it follows from Lemma A. 1 (b) that

$$
\begin{equation*}
\frac{1}{\sqrt{n}} y_{[n s]}=\frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]} u_{t} \rightarrow_{d} \pi(1) Q(s) \tag{35}
\end{equation*}
$$

using the notation introduced in Theorem 2.3. The stated result in part (a) now follows immediately from (36) and the continuous mapping theorem. For part (b), we also have from (36) that

$$
\begin{aligned}
\frac{1}{n} \sum_{t=1}^{n} y_{t-1} \varepsilon_{t} & =\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{y_{t-1}}{\sqrt{n}} \varepsilon_{t} \\
& \rightarrow_{d} \sigma_{\varepsilon} \pi(1) \int_{0}^{1} Q(s) d P(s)
\end{aligned}
$$

as required, where $P$ is the normalized Brownian motion of $B_{\varepsilon}$, i.e., $P=B_{\varepsilon} / \sigma_{\varepsilon}$.
Proof of Theorem 2.3 We have from Lemma 2.1 of Park and Phillips (1989) that

$$
\sum_{t=1}^{n} z_{t} z_{t}^{\prime}=O_{p}(n), \quad \sum_{t=1}^{n} z_{t} \varepsilon_{t}=O_{p}\left(n^{1 / 2}\right), \quad \text { and } \quad \sum_{t=1}^{n} y_{t-1} z_{t}^{\prime}=O_{p}(n) .
$$

Then it follows that

$$
\begin{aligned}
\left|\left(\sum_{t=1}^{n} y_{t-1} z_{t}^{\prime}\right)\left(\sum_{t=1}^{n} z_{t} z_{t}^{\prime}\right)^{-1}\left(\sum_{t=1}^{n} z_{t} \varepsilon_{t}\right)\right| & \leq\left|\sum_{t=1}^{n} y_{t-1} z_{t}^{\prime}\right|\left|\left(\sum_{t=1}^{n} z_{t} z_{t}^{\prime}\right)^{-1}\right|\left|\sum_{t=1}^{n} z_{t} \varepsilon_{t}\right|=O_{p}\left(n^{1 / 2}\right) \\
\left|\left(\sum_{t=1}^{n} y_{t-1} z_{t}^{\prime}\right)\left(\sum_{t=1}^{n} z_{t} z_{t}^{\prime}\right)^{-1}\left(\sum_{t=1}^{n} z_{t} y_{t-1}\right)\right| & \leq\left|\sum_{t=1}^{n} y_{t-1} z_{t}^{\prime}\right|\left|\left(\sum_{t=1}^{n} z_{t} z_{t}^{\prime}\right)^{-1}\right|\left|\sum_{t=1}^{n} z_{t} y_{t-1}\right|=O_{p}(n) \\
\left|\left(\sum_{t=1}^{n} \varepsilon_{t} z_{t}^{\prime}\right)\left(\sum_{t=1}^{n} z_{t} z_{t}^{\prime}\right)^{-1}\left(\sum_{t=1}^{n} z_{t} \varepsilon_{t}\right)\right| & \leq\left|\sum_{t=1}^{n} \varepsilon_{t} z_{t}\right|| |\left(\sum_{t=1}^{n} z_{t} z_{t}^{\prime}\right)^{-1}| | \sum_{t=1}^{n} z_{t} \varepsilon_{t} \mid=o_{p}(n) .
\end{aligned}
$$

Hence,

$$
\begin{aligned}
n^{-1} A_{n} & =n^{-1} \sum_{t=1}^{n} y_{t-1} \varepsilon_{t}+o_{p}(1) \\
n^{-2} B_{n} & =n^{-2} \sum_{t=1}^{n} y_{t-1}^{2}+o_{p}(1) \\
n^{-1} C_{n} & =n^{-1} \sum_{t=1}^{n} \varepsilon_{t}^{2}+o_{p}(1)
\end{aligned}
$$

Under the null, $\alpha=0$ and we have from (9) that

$$
S_{n}=\frac{n B_{n}^{-1} A_{n}}{\hat{\alpha}_{n}(1)}=\frac{1}{\hat{\alpha}_{n}(1)}\left(\frac{n^{-1} \sum_{t=1}^{n} y_{t-1} \varepsilon_{t}}{n^{-2} \sum_{t=1}^{n} y_{t-1}^{2}}\right)+o_{p}(1) \rightarrow_{d} \sigma_{\varepsilon} \frac{\int_{0}^{1} Q(s) d P(s)}{\int_{0}^{1} Q(s)^{2} d s}
$$

as required, due to Lemma A.2. Similarly, the stated limit distribution of $T_{n}$ follows directly from (10) and Lemma A. 2 as

$$
T_{n}=\frac{1}{\hat{\sigma}_{n}}\left(\frac{A_{n}}{B_{n}^{1 / 2}}\right)=\frac{1}{\hat{\sigma}_{n}}\left(\frac{n^{-1} \sum_{t=1}^{n} y_{t-1} \varepsilon_{t}}{\left(n^{-2} \sum_{t=1}^{n} y_{t-1}^{2}\right)^{1 / 2}}\right)+o_{p}(1) \rightarrow_{d} \frac{\int_{0}^{1} Q(s) d P(s)}{\left(\int_{0}^{1} Q(s)^{2} d s\right)^{1 / 2}}
$$

since $\hat{\sigma}_{n}^{2} \rightarrow_{p} \sigma_{\varepsilon}^{2}$ and $\pi(1)=1 / \alpha(1)$.
For the derivation of the limit distributions for the bootstrap CADF tests $S_{n}^{*}$ and $T_{n}^{*}$, we rely on the invariance principles for the bootstrapped samples, which are analogous to those derived for the original samples in Lemma A.1. We will use the symbol $o_{p}^{*}(1)$ to signify the bootstrap convergence in probability. For a sequence of bootstrapped random variables $Z_{n}^{*}$, for instance, $Z_{n}^{*}=o_{p}^{*}(1)$ a.s. and in $\mathbf{P}$ imply respectively that

$$
\mathbf{P}^{*}\left\{\left|Z_{n}^{*}\right|>\delta\right\} \rightarrow 0 \text { a.s. or in } \mathbf{P}
$$

for any $\delta>0$. Similarly, we will use the symbol $O_{p}^{*}(1)$ to denote the bootstrap version of the boundedness in probability. Needless to say, the definitions of $o_{p}^{*}(1)$ and $O_{p}^{*}(1)$ naturally extend to $o_{p}^{*}\left(c_{n}\right)$ and $O_{p}^{*}\left(c_{n}\right)$ for some nonconstant numerical sequence $\left(c_{n}\right)$. Many of well known results for $o_{p}$ and $O_{p}$ extend to $o_{p}^{*}$ and $O_{p}^{*}$ and are provided in Lemma 3.1 of Chang and Park (2003).

For a sequence of bootstrapped statistics $\left(Z_{n}^{*}\right)$ which weakly converges a.s. (or in $\mathbf{P}$ ), it follows that $Z_{n}^{*}=O_{p}^{*}(1)$ a.s. (or in $\mathbf{P}$ ). Moreover, if $Z_{n}^{*} \rightarrow_{d^{*}} Z$ a.s. (or in $\mathbf{P}$ ), then $Z_{n}^{*}+Y_{n}^{*} \rightarrow_{d^{*}} Z$ a.s. (or in $\mathbf{P}$ ) for any $\left(Y_{n}^{*}\right)$ such that $Y_{n}^{*}=o_{p}^{*}(1)$ a.s. (or in $\mathbf{P}$ ). For further discussions on bootstrap asymptotics, the reader is referred to Park (2002) and Chang and Park (2003).

Lemma A. 3 Under Assumption 2.1, we have
(a) $\frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]} w_{t}^{*} \rightarrow_{d^{*}} \Psi(1) B_{\eta}(s)$
(b) $\frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]} u_{t}^{*} \rightarrow_{d^{*}} \pi(1)\left(\beta(1)^{\prime} \Psi(1) B_{\eta}(s)+B_{\varepsilon}(s)\right)$
as $n \rightarrow \infty$.
Proof of Lemma A. 3 Under Assumption 2.1, the following invariance principle for $\left(\xi_{t}^{*}\right)=$ $\left(\varepsilon_{t}^{*}, \eta_{t}^{* \prime}\right)^{\prime}$ holds:

$$
\begin{equation*}
\frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]} \xi_{t}^{*} \rightarrow_{d^{*}} B=\binom{B_{\varepsilon}}{B_{\eta}} \tag{36}
\end{equation*}
$$

as $n \rightarrow \infty$ due to Theorem 3.3 of Chang, Park and Song (2006). As in the proof of Lemma A.1, we use the BN representations for the bootstrapped series $\left(w_{t}^{*}\right),\left(v_{t}^{*}\right)$ and $\left(u_{t}^{*}\right)$ to derive the limit distributions of their partial sum processes.

Part (a) Let $\tilde{\Phi}_{n}(1)=I_{m}-\sum_{k=1}^{\ell} \tilde{\Phi}_{k, n}$, where $\tilde{\Phi}_{k, n}$ 's are the coefficient estimates from the fitted regression (15), and define

$$
\tilde{\Psi}_{n}(1)=\tilde{\Phi}_{n}(1)^{-1}
$$

We may derive the BN representations for $\left(w_{t}^{*}\right)$ from the fitted autoregression (15) as we did for $\left(w_{t}\right)$ in (34) as

$$
\begin{equation*}
w_{t}^{*}=\tilde{\Psi}_{n}(1) \eta_{t-r-1}^{*}+\left(\bar{w}_{t-1}^{*}-\bar{w}_{t}^{*}\right) \tag{37}
\end{equation*}
$$

where $\bar{w}_{t}^{*}=\sum_{k=1}^{\ell}\left(\tilde{\Psi}_{n}(1) \sum_{j=k}^{\ell} \tilde{\Phi}_{j, n}\right) w_{t-k+1}^{*}$. Then the stated result in part (a) follows directly from the invariance principle (37) as

$$
\frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]} w_{t}^{*}=\tilde{\Psi}_{n}(1) \frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]} \eta_{t-r-1}^{*}+\frac{1}{\sqrt{n}}\left(\bar{w}_{0}^{*}-\bar{w}_{[n s]}^{*}\right) \rightarrow_{d^{*}} \Psi(1) B_{\eta}(s)
$$

as $n \rightarrow \infty$, since $\tilde{\Psi}_{n}(1) \rightarrow_{a . s .} \Psi(1)$ and $n^{-1 / 2}\left(\bar{w}_{0}^{*}-\bar{w}_{[n s]}^{*}\right)=o_{p}^{*}(1)$.
Part (b) Define

$$
\tilde{\pi}_{n}(1)=\tilde{\alpha}_{n}(1)^{-1} \quad \text { and } \quad \tilde{\beta}_{n}(1)=\sum_{k=-r}^{q} \tilde{\beta}_{k},
$$

where $\tilde{\alpha}_{n}(1)$ is defined as in (22) and $\tilde{\beta}_{k}$ 's are the estimates from (14). Now we derive the BN representation for ( $u_{t}^{*}$ ) from the fitted regression (14) similarly as in (35) as

$$
\begin{equation*}
u_{t}^{*}=\tilde{\pi}_{n}(1)\left(\tilde{\beta}_{n}(1)^{\prime} w_{t}^{*}+\varepsilon_{t}^{*}\right)+\tilde{\pi}_{n}(1)\left(\left(\tilde{w}_{t-1}^{*+}-\tilde{w}_{t}^{*+}\right)+\left(\tilde{w}_{t}^{*-}-\tilde{w}_{t-1}^{*-}\right)\right)+\left(\bar{u}_{t-1}^{*}-\bar{u}_{t}^{*}\right) \tag{38}
\end{equation*}
$$

where

$$
\begin{aligned}
\bar{u}_{t}^{*} & =\tilde{\pi}_{n}(1) \sum_{k=1}^{p}\left(\sum_{i=k}^{p} \tilde{\alpha}_{i}\right) u_{t-k+1}^{*} \\
\tilde{w}_{t}^{*+} & =\sum_{k=0}^{q-1} \tilde{\beta}_{k}^{+\prime} w_{t-k}^{*} \\
\tilde{w}_{t}^{*-} & =\sum_{k=0}^{r-1} \tilde{\beta}_{k}^{-\prime} w_{t+k+1}^{*}
\end{aligned}
$$

with $\tilde{\beta}_{k}^{+}=\sum_{j=k+1}^{q} \tilde{\beta}_{j}$ and $\tilde{\beta}_{k}^{-}=\sum_{j=k+1}^{r} \tilde{\beta}_{-j}$. Note that $\tilde{\pi}_{n}(1) \rightarrow_{a . s .} \pi(1)$ and $\tilde{\beta}_{n}(1) \rightarrow_{a . s .}$ $\beta(1)$. Then the stated result follows as

$$
\begin{aligned}
\frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]} u_{t}^{*}= & \tilde{\pi}_{n}(1)\left(\tilde{\beta}_{n}(1)^{\prime} \frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]} w_{t}^{*}+\frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]} \varepsilon_{t}^{*}\right)+o_{p}^{*}(1) \\
\rightarrow_{d^{*}} & \pi(1)\left(\beta(1)^{\prime} \Psi(1) B_{\eta}(s)+B_{\varepsilon}(s)\right)
\end{aligned}
$$

due to (37) and the result in part (a).

Lemma A. 4 Under the null hypothesis $\alpha=0$, we have as $n \rightarrow \infty$
(a) $n^{-2} \sum_{t=1}^{n} y_{t-1}^{* 2} \rightarrow_{d^{*}} \pi(1)^{2} \int_{0}^{1} Q(s)^{2} d s$ a.s.
(b) $n^{-1} \sum_{t=1}^{n} y_{t-1}^{*} \varepsilon_{t}^{*} \rightarrow_{d^{*}} \sigma_{\varepsilon} \pi(1) \int_{0}^{1} Q(s) d P(s)$ in $\mathbf{P}$
under Assumption 2.1, where $Q(s)$ and $P(s)$ are defined in Theorem 2.3.
Proof of Lemma A. 4 From (19), we have

$$
\frac{1}{\sqrt{n}} y_{[n s]}^{*}=\frac{1}{\sqrt{n}} \sum_{t=1}^{[n s]} u_{t}^{*}+\frac{y_{0}^{*}}{\sqrt{n}} .
$$

Then the stated result in Part (a) follows directly from Lemma A. 3 (b) and the continuous mapping theorem, since $n^{-1 / 2} y_{0}^{*}=o_{p}^{*}(1)$.

To prove Part (b), we first write the partial sum of $\left(u_{t}^{*}\right)$ explicitly using the BN decomposition given in (39) as

$$
\begin{aligned}
\sum_{k=1}^{t} u_{k}^{*}= & \tilde{\pi}_{n}(1) \sum_{k=1}^{t}\left(\tilde{\beta}_{n}(1)^{\prime} w_{k}^{*}+\varepsilon_{k}^{*}\right)+\tilde{\pi}_{n}(1)\left(\left(\tilde{w}_{0}^{*+}-\tilde{w}_{t}^{*+}\right)+\left(\tilde{w}_{t}^{*-}-\tilde{w}_{0}^{*-}\right)\right) \\
& +\left(\bar{u}_{0}^{*}-\bar{u}_{t}^{*}\right)
\end{aligned}
$$

Then, we use this to write

$$
\begin{align*}
\frac{1}{n} \sum_{t=1}^{n} y_{t-1}^{*} \varepsilon_{t}^{*} & =\frac{1}{n} \sum_{t=1}^{n}\left(\sum_{k=1}^{t-1} u_{k}^{*}+y_{0}^{*}\right) \varepsilon_{t}^{*} \\
& =R_{1 n}+R_{2 n}+R_{3 n}+R_{4 n}+R_{5 n}+R_{6 n} \tag{39}
\end{align*}
$$

where

$$
\begin{aligned}
& R_{1 n}=\tilde{\pi}_{n}(1) \frac{1}{n} \sum_{t=1}^{n}\left(\sum_{k=1}^{t-1} \tilde{\beta}_{n}(1)^{\prime} w_{k}^{*}\right) \varepsilon_{t}^{*} \\
& R_{2 n}=\tilde{\pi}_{n}(1) \frac{1}{n} \sum_{t=1}^{n}\left(\sum_{k=1}^{t-1} \varepsilon_{k}^{*}\right) \varepsilon_{t}^{*} \\
& R_{3 n}=\tilde{\pi}_{n}(1) \frac{1}{n} \sum_{t=1}^{n}\left(\tilde{w}_{0}^{*+}-\tilde{w}_{t-1}^{*+}\right) \varepsilon_{t}^{*} \\
& R_{4 n}=\tilde{\pi}_{n}(1) \frac{1}{n} \sum_{t=1}^{n}\left(\tilde{w}_{t-1}^{*-}-\tilde{w}_{0}^{*-}\right) \varepsilon_{t}^{*} \\
& R_{5 n}=\frac{1}{n} \sum_{t=1}^{n}\left(\bar{u}_{0}^{*}-\bar{u}_{t-1}^{*}\right) \varepsilon_{t}^{*} \\
& R_{6 n}=y_{0}^{*} \frac{1}{n} \sum_{t=1}^{n} \varepsilon_{t}^{*} .
\end{aligned}
$$

To study the limit behavior of $R_{1 n}$, we use the BN decomposition of ( $w_{t}^{*}$ ) given in (38) to write it more explicitly as

$$
\begin{align*}
R_{1 n} & =\tilde{\pi}_{n}(1) \frac{1}{n} \sum_{t=1}^{n}\left(\sum_{k=1}^{t-1} \tilde{\beta}_{n}(1)^{\prime}\left(\tilde{\Psi}_{n}(1) \eta_{k-r-1}^{*}+\left(\bar{w}_{k-1}^{*}-\bar{w}_{k}^{*}\right)\right)\right) \varepsilon_{t}^{*} \\
& =R_{1 n}^{a}+R_{1 n}^{b} \tag{40}
\end{align*}
$$

where

$$
\begin{aligned}
R_{1 n}^{a} & =\tilde{\pi}_{n}(1) \frac{1}{n} \sum_{t=1}^{n}\left(\sum_{k=1}^{t-1} \tilde{\beta}_{n}(1)^{\prime} \tilde{\Psi}_{n}(1) \eta_{k-r-1}^{*}\right) \varepsilon_{t}^{*} \\
R_{1 n}^{b} & =\tilde{\pi}_{n}(1) \frac{1}{n} \sum_{t=1}^{n} \tilde{\beta}_{n}(1)^{\prime}\left(\bar{w}_{0}^{*}-\bar{w}_{t-1}^{*}\right) \varepsilon_{t}^{*} .
\end{aligned}
$$

For the first part $R_{1 n}^{a}$, we note from the invariance principle given in (37) that

$$
\left.\begin{array}{c}
\frac{1}{\sqrt{n}} \sum_{k=1}^{[n s]} \tilde{\beta}_{n}(1)^{\prime} \tilde{\Psi}_{n}(1) \eta_{k-r-1}^{*} \\
\rightarrow_{d^{*}} \tag{41}
\end{array}\right](1)^{\prime} \Psi(1) B_{\eta}(s) \text { a.s. }
$$

where $s$ is such that $[n s]=t-1$. Then it follows that

$$
\begin{equation*}
R_{1 n}^{a} \rightarrow_{d^{*}} \pi(1) \int_{0}^{1}\left(\beta(1)^{\prime} \Psi(1) B_{\eta}(s)\right) d B_{\varepsilon}(s) \text { in } \mathbf{P} \tag{42}
\end{equation*}
$$

due to Kurtz and Protter (1991).
For $R_{1 n}^{b}$, we first note that

$$
\begin{equation*}
\mathbf{E}^{*}\left(\frac{1}{n} \sum_{t=1}^{n} \varepsilon_{t}^{*}\right)^{2}=\frac{1}{n} \mathbf{E}^{*} \varepsilon_{t}^{* 2}=\frac{1}{n}\left(\frac{1}{n} \sum_{t=1}^{n}\left(\tilde{\varepsilon}_{t}-\bar{\varepsilon}_{n}\right)\right)^{2} \rightarrow_{a . s} 0 \tag{43}
\end{equation*}
$$

due to Assumption 2.1. Then, it follows from (4), Assumption 2.1 and (44) that

$$
\begin{align*}
\mathbf{E}^{*}\left(\frac{1}{n} \sum_{t=1}^{n} \bar{w}_{t-1}^{*} \varepsilon_{t}^{*}\right)^{2} & =\frac{1}{n} \mathbf{E}^{*}\left(\bar{w}_{t-1}^{*} \varepsilon_{t}^{*}\right)^{2}=\frac{1}{n} \mathbf{E}^{*}\left(\mathbf{E}^{*}\left(\bar{w}_{t-1}^{*} \varepsilon_{t}^{*}\right)^{2} \mid \mathcal{F}_{t-1}\right) \\
& =\frac{1}{n} \mathbf{E}^{*} \bar{w}_{t-1}^{* 2} \mathbf{E}^{*} \varepsilon_{t}^{* 2} \rightarrow_{a . s .} 0 \tag{44}
\end{align*}
$$

and that

$$
\begin{equation*}
\mathbf{E}^{*}\left(\frac{1}{n} \sum_{t=1}^{n} \bar{w}_{0}^{*} \varepsilon_{t}^{*}\right)^{2}=\bar{w}_{0}^{* 2} \frac{1}{n} \mathbf{E}^{*} \varepsilon_{t}^{* 2} \rightarrow_{a . s .} 0 \tag{45}
\end{equation*}
$$

The above results imply that

$$
\frac{1}{n} \sum_{t=1}^{n} \bar{w}_{t-1}^{*} \varepsilon_{t}^{*}, \frac{1}{n} \sum_{t=1}^{n} \bar{w}_{0}^{*} \varepsilon_{t}^{*}=o_{p}^{*}(1)
$$

which in turn implies $R_{1 n}^{b}=o_{p}^{*}(1)$. Now we may deduce from (41) and (43) that

$$
\begin{equation*}
R_{1 n} \rightarrow_{d^{*}} \pi(1) \int_{0}^{1}\left(\beta(1)^{\prime} \Psi(1) B_{\eta}(s)\right) d B_{\varepsilon}(s) \text { in } \mathbf{P} . \tag{46}
\end{equation*}
$$

From (42), we have

$$
\begin{equation*}
R_{2 n} \rightarrow_{d^{*}} \pi(1) \int_{0}^{1} B_{\varepsilon}(s) d B_{\varepsilon}(s) \text { in } \mathbf{P} \tag{47}
\end{equation*}
$$

again due to Kurtz and Protter (1991).
As in (45), we may similarly show that

$$
\frac{1}{n} \sum_{t=1}^{n} \tilde{w}_{t-1}^{*+} \varepsilon_{t}^{*}, \frac{1}{n} \sum_{t=1}^{n} \tilde{w}_{t-1}^{*-} \varepsilon_{t}^{*}, \frac{1}{n} \sum_{t=1}^{n} u_{t-1}^{*} \varepsilon_{t}^{*}=o_{p}^{*}(1)
$$

and as in (46) we may also show that

$$
\tilde{w}_{0}^{*+} \frac{1}{n} \sum_{t=1}^{n} \varepsilon_{t}^{*}, \tilde{w}_{0}^{*-} \frac{1}{n} \sum_{t=1}^{n} \varepsilon_{t}^{*}, \bar{u}_{0}^{*} \frac{1}{n} \sum_{t=1}^{n} \varepsilon_{t}^{*}, y_{0}^{*} \frac{1}{n} \sum_{t=1}^{n} \varepsilon_{t}^{*}=o_{p}^{*}(1) .
$$

Then we may deduce that

$$
R_{3 n}, R_{4 n}, R_{5 n}, R_{6 n}=o_{p}^{*}(1)
$$

which together with the results in (40), (47) and (48) proves the stated result.
Proof of Theorem 3.3 The stochastic orders for the bootstrap sample moments appearing in the definitions of the bootstrap test $S_{n}^{*}$ and $T_{n}^{*}$ are easily obtained

$$
\left|\left(\sum_{t=1}^{n} z_{t}^{*} z_{t}^{* \prime}\right)^{-1}\right|=O_{p}^{*}\left(n^{-1}\right), \quad\left|\sum_{t=1}^{n} z_{t}^{*} \varepsilon_{t}^{*}\right|=O_{p}^{*}\left(n^{1 / 2}\right) \text { and }\left|\sum_{t=1}^{n} y_{t-1}^{*} z_{t}^{* \prime}\right|=O_{p}^{*}(n)
$$

from the results in Lemma 3 of Chang and Park (2003). Then we have

$$
\begin{array}{r}
\left|\left(\sum_{t=1}^{n} y_{t-1}^{*} z_{t}^{* \prime}\right)\left(\sum_{t=1}^{n} z_{t}^{*} z_{t}^{* \prime}\right)^{-1}\left(\sum_{t=1}^{n} z_{t}^{*} \varepsilon_{t}^{*}\right)\right| \leq\left|\sum_{t=1}^{n} y_{t-1}^{*} z_{t}^{* \prime}\right|\left|\left(\sum_{t=1}^{n} z_{t}^{*} z_{t}^{* \prime}\right)^{-1}\right|\left|\sum_{t=1}^{n} z_{t}^{*} \varepsilon_{t}^{*}\right|=O_{p}^{*}\left(n^{1 / 2}\right) \\
\left|\left(\sum_{t=1}^{n} y_{t-1}^{*} z_{t}^{* \prime}\right)\left(\sum_{t=1}^{n} z_{t}^{*} z_{t}^{* \prime}\right)^{-1}\left(\sum_{t=1}^{n} z_{t}^{*} y_{t-1}^{*}\right)\right| \leq\left|\sum_{t=1}^{n} y_{t-1}^{*} z_{t}^{* \prime}\right|\left|\left(\sum_{t=1}^{n} z_{t}^{*} z_{t}^{* \prime}\right)^{-1}\right|\left|\sum_{t=1}^{n} z_{t}^{*} y_{t-1}^{*}\right|=O_{p}^{*}(n)
\end{array}
$$

and consequently

$$
\begin{aligned}
n^{-1} A_{n}^{*} & =n^{-1} \sum_{t=1}^{n} y_{t-1}^{*} \varepsilon_{t}^{*}+o_{p}^{*}(1) \\
n^{-2} B_{n}^{*} & =n^{-2} \sum_{t=1}^{n} y_{t-1}^{* 2}+o_{p}^{*}(1)
\end{aligned}
$$

Note that $\tilde{\alpha}_{n}(1) \rightarrow_{a . s .} \alpha(1)$ and $\tilde{\sigma}_{n}^{2} \rightarrow_{a . s .} \sigma_{\varepsilon}^{2}$. Then it follows from the definitions of $S_{n}^{*}$ and $T_{n}^{*}$, given in (23) and (24), and the results in the previous equation that

$$
\begin{aligned}
& S_{n}^{*}=\frac{n B_{n}^{*-1} A_{n}^{*}}{\tilde{\alpha}_{n}(1)}=\frac{1}{\tilde{\alpha}_{n}(1)}\left(\frac{n^{-1} \sum_{t=1}^{n} y_{t-1}^{*} \varepsilon_{t}^{*}}{n^{-2} \sum_{t=1}^{n} y_{t-1}^{* 2}}\right)+o_{p}^{*}(1) \\
& T_{n}^{*}=\frac{1}{\tilde{\sigma}_{n}}\left(\frac{A_{n}^{*}}{B_{n}^{* 1 / 2}}\right)=\frac{1}{\tilde{\sigma}_{n}}\left(\frac{n^{-1} \sum_{t=1}^{n} y_{t-1}^{*} \varepsilon_{t}^{*}}{\left(n^{-2} \sum_{t=1}^{n} y_{t-1}^{* 2}\right)^{1 / 2}}\right)+o_{p}^{*}(1) .
\end{aligned}
$$

Now the stated limit theories for $S_{n}^{*}$ and $T_{n}^{*}$ follow immediately from Lemma A.4.

Proof of Theorem 4.1 Given the results in Lemmas A. 3 and A.4, the proof is essentially identical to the proof of Theorem 4.1 in Park (2003).

## 9. References

Andrews, D.W.K. (1991), "Heteroskedasticity and autocorrelation consistent covariance matrix estimation," Econometrica, 59, 817-858.

Andrews, D.W.K. and M. Buchinsky (1999), "A three-step method for choosing the number of bootstrap repetitions," Econometrica, 68, 23-52.

Basawa, I.V., A.K. Mallik, W.P. McCormick, J.H. Reeves and R.L. Taylor (1991), "Bootstrapping unstable first-order autoregressive processes," Annals of Statistics, 19, 10981101.

Berk, K.N. (1974), "Consistent autoregressive spectral estimates," The Annals of Statistics, 15, 1050-1063.

Brockwell, P.J. and R.A. Davis (1991), Time Series: Theory and Methods, Springer-Verlag: New York.

Campbell, J.Y. and P. Perron (1991), "Pitfalls and Opportunities: What Macroeconomists Should Know about Unit Roots," NBER Macroeconomics Annual, 6, 141-201.

Chang, Y. and J.Y. Park (2002), "On the asymptotics of ADF tests for unit roots," Econometric Reviews, 21, 431-447.

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.

Chang, Y., J.Y. Park and K.M. Song (2006), "Bootstrapping cointegrating regressions," Journal of Econometrics, 133, 703-739.

Cheung, Y., K.S. Lai, and M. Bergman (2004), "Dissecting the PPP puzzle: the unconventional roles of nominal exchange rate and price adjustments," Journal of International Economics, 64, 135-150.

Choi, C. (2004), "Searching for evidence of long-run PPP from a post-Bretton Woods panel: separating the wheat from the chaff," Journal of International Money and Finance, 23, 1159-1186.

Chortareas, G., and G. Kapetanios (2009), "Getting PPP right: Identifying mean-reverting real exchange rates in panels," Journal of Banking and Finance, 33, 390-404.

DeJong, D.N., J.C. Nankervis, N.E. Savin, and C.H. Whiteman (1992a), "The power problems of unit root tests in time series with autoregressive errors," Journal of Econometrics, 53, 323-343.

DeJong, D.N., J.C. Nankervis, N.E. Savin, and C.H. Whiteman (1992b), "Integration versus trend stationarity in time series," Econometrica, 60, 423-433.

Dickey, D.A. and W.A. Fuller (1979), "Distribution of estimators for autoregressive time series with a unit root," Journal of the American Statistical Association, 74, 427-431.

Domowitz, I., and M. El-Gamal (2001), "A consistent nonparametric test of ergodicity for time series with applications," Journal of Econometrics, 102, 365-398.

Elliott, G. and M. Jansson (2003), "Testing for unit roots with stationary covariates," Journal of Econometrics, 115, 75-89.

Elliott, G., T.J. Rothenberg, and J.H. Stock (1996), "Efficient tests for an autoregressive unit root," Econometrica, 64, 813-836.

Frankel, J. and A. Rose (1996), "A panel project on purchasing power parity: Mean reversion within and between countries," Journal of International Economics, 40, 209-224.

Hansen, B.E. (1995), "Rethinking the univariate approach to unit root testing: Using covariates to increase power," Econometric Theory, 11, 1148-1171.

Kurtz, T.G. and P. Protter (1991), "Weak limit theorems for stochastic integrals and stochastic differential equations," The Annals of Probability, 19, 1035-1070.

MacDonald, R. (1996), "Panel unit root tests and real exchange rates," Economics Letters, 50, 7-11.

Nelson, C.R. and C. Plosser (1982), "Trends and random walks in macroeconomic time series: Some evidence and implications," Journal of Monetary Economics, 10, 139162.

Ng, S. and P. Perron (2001), "Lag length selection and constructing unit root tests with good size and power," Econometrica, 69, 1519-1554.

O'Connell, P.G.J. (1998), "The overvaluation of purchasing power parity," Journal of International Economics, 44, 1-20.

Oh, K.Y. (1996), "Purchasing power parity and unit root tests using panel data," Journal of International Money and Finance, 15, 405-418.

Papell, D. (1997), "Searching for stationarity: Purchasing power parity under the current float," Journal of International Economics, 43, 313-323.

Papell, D. (2002), "The great appreciation, the great depreciation, and the purchasing power parity hypothesis," Journal of International Economics, 57, 51-82.

Papell, D., and H., Theodoridis (1998), "Increasing evidence of purchasing power parity over the current float," Journal of International Money and Finance, 17, 41-50.

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

Park, J.Y. (2003), "Bootstrap unit root tests," Econometrica, 71, 1845-1895.
Park, J.Y., and P.C.B. Phillips (1989), "Statistical Inference in Regressions with Integrated Processes: Part 2," Econometric Theory, 5, 95-131.

Perron, P. (1989), "The Great Crash, the oil price shock, and the unit root hypothesis," Econometrica, 57, 1361-1401.

Phillips, P.C.B. and P. Perron (1988), "Testing for a unit root in time series regression," Biometrika, 75, 335-346.

Said, S.E. and D.A. Dickey (1984), "Testing for unit roots in autoregressive-moving average models of unknown order," Biometrika, 71, 599-608.

Saikkonen, P. (1991), "Asymptotically efficient estimation of cointegration regressions," Econometric Theory, 7, 1-21.

Shiller, R. and P. Perron (1985), "Testing the random walk hypothesis: Power versus frequency of observation," Economics Letters, 18, 381-386.

Schotman, P.C. and H.K. van Dijk (1991), "On Bayesian routes to unit roots," Journal of Applied Econometrics, 6, 387-401.

Stock, J.H. (1991), "Confidence intervals for the largest autoregressive root in U.S. macroeconomic time series," Journal of Monetary Economics, 28, 435-459.

Stock, J.H. (1994), "Unit roots, structural breaks and trends," Handbook of Econometrics, Vol. 4, 2739-2841.

Stock, J. and M. Watson (1999), "Business cycle fluctuations in U.S. macroeconomic time series," Handbook of Macroeconomics, edited by J. Taylor and M. Woodford, North Holland, Vol. 1a, 3-64.

Table 1. Estimates of $\rho^{2}$ for Various $n(\alpha=1)$

| $\beta$ | $\phi$ | 25 | 50 | 100 | 1,000 | 3,000 | true $\rho^{2}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0.800 | 0.800 | 0.359 | 0.320 | 0.306 | 0.307 | 0.314 | 0.335 |
| 0.500 | 0.800 | 0.492 | 0.453 | 0.428 | 0.413 | 0.416 | 0.432 |
| -0.500 | 0.800 | 0.302 | 0.243 | 0.206 | 0.103 | 0.070 | 0.000 |
| -0.800 | 0.800 | 0.139 | 0.068 | 0.036 | 0.003 | 0.002 | 0.026 |
| 0.800 | 0.500 | 0.522 | 0.524 | 0.530 | 0.547 | 0.550 | 0.556 |
| 0.500 | 0.500 | 0.671 | 0.688 | 0.691 | 0.696 | 0.696 | 0.700 |
| -0.500 | 0.500 | 0.595 | 0.615 | 0.635 | 0.614 | 0.569 | 0.300 |
| -0.800 | 0.500 | 0.327 | 0.329 | 0.322 | 0.241 | 0.193 | 0.057 |
| 0.800 | -0.500 | 0.685 | 0.722 | 0.743 | 0.816 | 0.836 | 0.860 |
| 0.500 | -0.500 | 0.816 | 0.833 | 0.844 | 0.879 | 0.905 | 0.932 |
| -0.500 | -0.500 | 0.781 | 0.831 | 0.850 | 0.878 | 0.882 | 0.890 |
| -0.800 | -0.500 | 0.608 | 0.642 | 0.669 | 0.702 | 0.710 | 0.722 |
| 0.800 | -0.800 | 0.733 | 0.816 | 0.850 | 0.876 | 0.886 | 0.893 |
| 0.500 | -0.800 | 0.831 | 0.881 | 0.913 | 0.938 | 0.944 | 0.950 |
| -0.500 | -0.800 | 0.798 | 0.862 | 0.891 | 0.907 | 0.918 | 0.924 |
| -0.800 | -0.800 | 0.625 | 0.702 | 0.745 | 0.771 | 0.790 | 0.803 |
| 0.000 | 0.000 | 0.932 | 0.974 | 0.988 | 0.999 | 1.000 | 1.000 |

Note: The results for $n=25,50$, and 100 are based on 3,000 simulation iterations and those for $n=1,000$ and 3,000 are based on 1,000 simulation iterations.

Table 2. Asymptotic Sizes $(\alpha=1)$

|  |  | $n=1,000$ |  |  | $n=3,000$ |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\beta$ | $\phi$ | CADF | BCADF | $\hat{\rho}^{2}$ | CADF | BCADF | $\hat{\rho}^{2}$ | true $\rho^{2}$ |
| 0.800 | 0.800 | 0.062 | 0.062 | 0.307 | 0.064 | 0.068 | 0.314 | 0.335 |
| 0.500 | 0.800 | 0.059 | 0.063 | 0.413 | 0.057 | 0.057 | 0.416 | 0.432 |
| -0.500 | 0.800 | 0.018 | 0.051 | 0.103 | 0.014 | 0.051 | 0.070 | 0.000 |
| -0.800 | 0.800 | 0.010 | 0.053 | 0.003 | 0.007 | 0.047 | 0.002 | 0.026 |
| 0.800 | 0.500 | 0.052 | 0.056 | 0.547 | 0.058 | 0.064 | 0.550 | 0.556 |
| 0.500 | 0.500 | 0.053 | 0.056 | 0.696 | 0.057 | 0.062 | 0.696 | 0.700 |
| -0.500 | 0.500 | 0.038 | 0.068 | 0.614 | 0.030 | 0.057 | 0.569 | 0.300 |
| -0.800 | 0.500 | 0.023 | 0.051 | 0.241 | 0.036 | 0.063 | 0.193 | 0.057 |
| 0.800 | -0.500 | 0.058 | 0.058 | 0.816 | 0.051 | 0.054 | 0.836 | 0.860 |
| 0.500 | -0.500 | 0.063 | 0.058 | 0.879 | 0.060 | 0.062 | 0.905 | 0.932 |
| -0.500 | -0.500 | 0.057 | 0.063 | 0.878 | 0.059 | 0.064 | 0.882 | 0.890 |
| -0.800 | -0.500 | 0.054 | 0.062 | 0.702 | 0.054 | 0.060 | 0.710 | 0.722 |
| 0.800 | -0.800 | 0.043 | 0.047 | 0.876 | 0.043 | 0.047 | 0.886 | 0.893 |
| 0.500 | -0.800 | 0.058 | 0.062 | 0.938 | 0.041 | 0.050 | 0.944 | 0.950 |
| -0.500 | -0.800 | 0.065 | 0.074 | 0.907 | 0.061 | 0.070 | 0.918 | 0.924 |
| -0.800 | -0.800 | 0.041 | 0.046 | 0.771 | 0.060 | 0.071 | 0.790 | 0.803 |
| 0.000 | 0.000 | 0.049 | 0.048 | 0.999 | 0.049 | 0.058 | 1.000 | 1.000 |

Note: The results are based on 1,000 simulation iterations with the bootstrap critical values computed from 1,000 bootstrap repetitions.
Table 3. Finite Sample Sizes $(\alpha=1, n=25)$

| $\beta$ | $\phi$ | ADF | $\mathrm{Z}_{\alpha}$ | $\mathrm{MZ}_{\alpha}$ | $\mathrm{MZ}_{\mathrm{t}}$ | MSB | $\mathrm{P}_{\mathrm{t}}$ | $\mathrm{MP}_{\mathrm{t}}$ | $\mathrm{DF}^{\mathrm{GLS}}$ | EJ | CADF | BCADF | $\rho^{2}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0.800 | 0.800 | 0.103 | 0.248 | 0.238 | 0.220 | 0.246 | 0.189 | 0.217 | 0.093 | 0.363 | 0.092 | 0.058 | 0.080 |
| 0.500 | 0.800 | 0.092 | 0.151 | 0.144 | 0.132 | 0.156 | 0.110 | 0.127 | 0.045 | 0.385 | 0.118 | 0.064 | 0.126 |
| -0.500 | 0.800 | 0.084 | 0.079 | 0.068 | 0.064 | 0.073 | 0.056 | 0.062 | 0.029 | 0.381 | 0.120 | 0.069 | 0.127 |
| -0.800 | 0.800 | 0.086 | 0.136 | 0.131 | 0.120 | 0.141 | 0.093 | 0.117 | 0.040 | 0.380 | 0.059 | 0.047 | 0.053 |
| 0.800 | 0.500 | 0.115 | 0.137 | 0.129 | 0.119 | 0.133 | 0.102 | 0.114 | 0.066 | 0.371 | 0.107 | 0.068 | 0.134 |
| 0.500 | 0.500 | 0.092 | 0.091 | 0.079 | 0.073 | 0.086 | 0.066 | 0.071 | 0.049 | 0.349 | 0.126 | 0.079 | 0.217 |
| -0.500 | 0.500 | 0.106 | 0.059 | 0.037 | 0.033 | 0.039 | 0.031 | 0.032 | 0.070 | 0.255 | 0.191 | 0.105 | 0.297 |
| -0.800 | 0.500 | 0.087 | 0.058 | 0.044 | 0.042 | 0.048 | 0.039 | 0.041 | 0.054 | 0.291 | 0.119 | 0.077 | 0.162 |
| 0.800 | -0.500 | 0.084 | 0.070 | 0.041 | 0.038 | 0.046 | 0.035 | 0.037 | 0.072 | 0.240 | 0.117 | 0.090 | 0.283 |
| 0.500 | -0.500 | 0.091 | 0.066 | 0.045 | 0.039 | 0.047 | 0.037 | 0.038 | 0.065 | 0.230 | 0.126 | 0.089 | 0.368 |
| -0.500 | -0.500 | 0.094 | 0.101 | 0.052 | 0.046 | 0.054 | 0.041 | 0.045 | 0.110 | 0.199 | 0.168 | 0.121 | 0.499 |
| -0.800 | -0.500 | 0.083 | 0.114 | 0.049 | 0.046 | 0.049 | 0.045 | 0.046 | 0.116 | 0.200 | 0.158 | 0.109 | 0.413 |
| 0.800 | -0.800 | 0.093 | 0.109 | 0.056 | 0.050 | 0.056 | 0.045 | 0.051 | 0.097 | 0.218 | 0.120 | 0.087 | 0.292 |
| 0.500 | -0.800 | 0.092 | 0.076 | 0.044 | 0.041 | 0.049 | 0.036 | 0.042 | 0.078 | 0.194 | 0.127 | 0.091 | 0.380 |
| -0.500 | -0.800 | 0.088 | 0.109 | 0.056 | 0.051 | 0.058 | 0.046 | 0.051 | 0.094 | 0.191 | 0.145 | 0.099 | 0.497 |
| -0.800 | -0.800 | 0.081 | 0.150 | 0.054 | 0.049 | 0.056 | 0.044 | 0.048 | 0.097 | 0.181 | 0.155 | 0.108 | 0.418 |
| 0.000 | 0.000 | 0.096 | 0.063 | 0.038 | 0.034 | 0.039 | 0.030 | 0.034 | 0.068 | 0.213 | 0.164 | 0.110 | 0.468 |

[^11]Table 4. Finite Sample Sizes $(\alpha=1, n=50)$

| $\beta$ | $\phi$ | ADF | $\mathrm{Z}_{\alpha}$ | $\mathrm{MZ}_{\alpha}$ | $\mathrm{MZ}_{\mathrm{t}}$ | MSB | $\mathrm{P}_{\mathrm{t}}$ | $\mathrm{MP}_{\mathrm{t}}$ | DF $^{\text {GLS }}$ | EJ | CADF | BCADF | $\rho^{2}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0.800 | 0.800 | 0.094 | 0.114 | 0.115 | 0.103 | 0.118 | 0.084 | 0.100 | 0.058 | 0.354 | 0.073 | 0.065 | 0.036 |
| 0.500 | 0.800 | 0.081 | 0.077 | 0.078 | 0.068 | 0.081 | 0.053 | 0.067 | 0.037 | 0.386 | 0.092 | 0.079 | 0.094 |
| -0.500 | 0.800 | 0.068 | 0.044 | 0.045 | 0.042 | 0.049 | 0.035 | 0.041 | 0.021 | 0.372 | 0.088 | 0.083 | 0.125 |
| -0.800 | 0.800 | 0.072 | 0.067 | 0.068 | 0.057 | 0.071 | 0.045 | 0.054 | 0.027 | 0.378 | 0.045 | 0.060 | 0.026 |
| 0.800 | 0.500 | 0.100 | 0.097 | 0.100 | 0.086 | 0.093 | 0.069 | 0.085 | 0.070 | 0.325 | 0.086 | 0.069 | 0.101 |
| 0.500 | 0.500 | 0.087 | 0.065 | 0.067 | 0.060 | 0.066 | 0.046 | 0.058 | 0.049 | 0.287 | 0.122 | 0.093 | 0.205 |
| -0.500 | 0.500 | 0.080 | 0.039 | 0.035 | 0.031 | 0.030 | 0.026 | 0.031 | 0.039 | 0.191 | 0.136 | 0.104 | 0.381 |
| -0.800 | 0.500 | 0.081 | 0.042 | 0.037 | 0.031 | 0.037 | 0.024 | 0.030 | 0.035 | 0.272 | 0.097 | 0.087 | 0.189 |
| 0.800 | -0.500 | 0.085 | 0.040 | 0.034 | 0.029 | 0.032 | 0.024 | 0.030 | 0.040 | 0.194 | 0.114 | 0.094 | 0.315 |
| 0.500 | -0.500 | 0.091 | 0.045 | 0.041 | 0.036 | 0.043 | 0.029 | 0.034 | 0.042 | 0.177 | 0.116 | 0.103 | 0.408 |
| -0.500 | -0.500 | 0.082 | 0.054 | 0.039 | 0.037 | 0.037 | 0.030 | 0.037 | 0.054 | 0.141 | 0.128 | 0.106 | 0.607 |
| -0.800 | -0.500 | 0.073 | 0.067 | 0.045 | 0.040 | 0.042 | 0.033 | 0.040 | 0.061 | 0.148 | 0.132 | 0.117 | 0.507 |
| 0.800 | -0.800 | 0.093 | 0.062 | 0.049 | 0.041 | 0.046 | 0.034 | 0.041 | 0.047 | 0.177 | 0.106 | 0.094 | 0.341 |
| 0.500 | -0.800 | 0.081 | 0.051 | 0.041 | 0.036 | 0.040 | 0.030 | 0.036 | 0.051 | 0.165 | 0.122 | 0.110 | 0.430 |
| -0.500 | -0.800 | 0.087 | 0.069 | 0.050 | 0.045 | 0.047 | 0.038 | 0.045 | 0.049 | 0.144 | 0.121 | 0.116 | 0.619 |
| -0.800 | -0.800 | 0.082 | 0.095 | 0.050 | 0.046 | 0.047 | 0.037 | 0.044 | 0.046 | 0.137 | 0.132 | 0.117 | 0.539 |
| 0.000 | 0.000 | 0.085 | 0.041 | 0.035 | 0.031 | 0.034 | 0.029 | 0.032 | 0.040 | 0.159 | 0.139 | 0.111 | 0.560 |

Note: Only a constant is included.
Table 5. Finite Sample Sizes $(\alpha=1, n=100)$

| $\beta$ | $\phi$ | ADF | $\mathrm{Z}_{\alpha}$ | $\mathrm{MZ}_{\alpha}$ | $\mathrm{MZ}_{\mathrm{t}}$ | MSB | $\mathrm{P}_{\mathrm{t}}$ | $\mathrm{MP}_{\mathrm{t}}$ | $\mathrm{DF} \mathrm{GLS}^{2}$ | EJ | CADF | BCADF | $\rho^{2}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0.800 | 0.800 | 0.071 | 0.060 | 0.065 | 0.056 | 0.065 | 0.048 | 0.054 | 0.038 | 0.134 | 0.053 | 0.068 | 0.032 |
| 0.500 | 0.800 | 0.058 | 0.046 | 0.051 | 0.045 | 0.048 | 0.036 | 0.045 | 0.033 | 0.204 | 0.075 | 0.075 | 0.102 |
| -0.500 | 0.800 | 0.056 | 0.029 | 0.030 | 0.028 | 0.031 | 0.025 | 0.027 | 0.021 | 0.189 | 0.036 | 0.062 | 0.148 |
| -0.800 | 0.800 | 0.057 | 0.039 | 0.042 | 0.034 | 0.042 | 0.029 | 0.034 | 0.026 | 0.179 | 0.025 | 0.054 | 0.019 |
| 0.800 | 0.500 | 0.069 | 0.069 | 0.074 | 0.064 | 0.069 | 0.050 | 0.059 | 0.054 | 0.158 | 0.072 | 0.071 | 0.123 |
| 0.500 | 0.500 | 0.073 | 0.059 | 0.061 | 0.057 | 0.053 | 0.045 | 0.054 | 0.049 | 0.136 | 0.088 | 0.087 | 0.268 |
| -0.500 | 0.500 | 0.073 | 0.039 | 0.039 | 0.036 | 0.036 | 0.029 | 0.035 | 0.036 | 0.087 | 0.085 | 0.090 | 0.518 |
| -0.800 | 0.500 | 0.071 | 0.046 | 0.047 | 0.045 | 0.044 | 0.036 | 0.044 | 0.044 | 0.127 | 0.050 | 0.073 | 0.237 |
| 0.800 | -0.500 | 0.068 | 0.039 | 0.038 | 0.035 | 0.033 | 0.029 | 0.034 | 0.038 | 0.098 | 0.079 | 0.079 | 0.411 |
| 0.500 | -0.500 | 0.079 | 0.050 | 0.049 | 0.044 | 0.043 | 0.039 | 0.044 | 0.049 | 0.095 | 0.083 | 0.084 | 0.516 |
| -0.500 | -0.500 | 0.063 | 0.049 | 0.046 | 0.042 | 0.043 | 0.036 | 0.041 | 0.049 | 0.092 | 0.091 | 0.097 | 0.770 |
| -0.800 | -0.500 | 0.064 | 0.051 | 0.043 | 0.041 | 0.042 | 0.036 | 0.041 | 0.043 | 0.090 | 0.087 | 0.091 | 0.614 |
| 0.800 | -0.800 | 0.069 | 0.054 | 0.052 | 0.044 | 0.051 | 0.039 | 0.043 | 0.038 | 0.084 | 0.081 | 0.086 | 0.431 |
| 0.500 | -0.800 | 0.073 | 0.049 | 0.048 | 0.044 | 0.044 | 0.037 | 0.043 | 0.042 | 0.093 | 0.086 | 0.092 | 0.546 |
| -0.500 | -0.800 | 0.068 | 0.061 | 0.059 | 0.052 | 0.057 | 0.042 | 0.050 | 0.049 | 0.088 | 0.087 | 0.093 | 0.794 |
| -0.800 | -0.800 | 0.076 | 0.067 | 0.053 | 0.048 | 0.047 | 0.040 | 0.048 | 0.044 | 0.079 | 0.092 | 0.095 | 0.681 |
| 0.000 | 0.000 | 0.064 | 0.043 | 0.044 | 0.037 | 0.042 | 0.034 | 0.038 | 0.041 | 0.094 | 0.081 | 0.080 | 0.726 |

[^12]Table 6. Finite Sample Sizes $(\alpha=1.0, n=25)$

| $\beta$ | $\phi$ | ADF | $\mathrm{Z}_{\alpha}$ | $\mathrm{MZ}_{\alpha}$ | $\mathrm{MZ}_{\mathrm{t}}$ | MSB | $\mathrm{P}_{\mathrm{t}}$ | $\mathrm{MP}_{\mathrm{t}}$ | $\mathrm{DF}^{\mathrm{GLS}}$ | EJ | CADF | BCADF | $\rho^{2}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0.800 | 0.800 | 0.120 | 0.191 | 0.182 | 0.172 | 0.194 | 0.171 | 0.166 | 0.050 | 0.590 | 0.141 | 0.057 | 0.065 |
| 0.500 | 0.800 | 0.094 | 0.092 | 0.087 | 0.079 | 0.096 | 0.080 | 0.075 | 0.017 | 0.528 | 0.205 | 0.083 | 0.100 |
| -0.500 | 0.800 | 0.096 | 0.026 | 0.024 | 0.023 | 0.027 | 0.022 | 0.022 | 0.023 | 0.448 | 0.211 | 0.087 | 0.088 |
| -0.800 | 0.800 | 0.086 | 0.082 | 0.075 | 0.069 | 0.085 | 0.070 | 0.064 | 0.019 | 0.554 | 0.104 | 0.057 | 0.043 |
| 0.800 | 0.500 | 0.143 | 0.075 | 0.072 | 0.068 | 0.076 | 0.068 | 0.066 | 0.025 | 0.621 | 0.163 | 0.073 | 0.111 |
| 0.500 | 0.500 | 0.119 | 0.038 | 0.035 | 0.033 | 0.041 | 0.033 | 0.031 | 0.021 | 0.473 | 0.203 | 0.085 | 0.178 |
| -0.500 | 0.500 | 0.118 | 0.020 | 0.011 | 0.010 | 0.013 | 0.011 | 0.009 | 0.062 | 0.263 | 0.279 | 0.119 | 0.226 |
| -0.800 | 0.500 | 0.118 | 0.022 | 0.014 | 0.012 | 0.016 | 0.013 | 0.012 | 0.050 | 0.436 | 0.170 | 0.086 | 0.129 |
| 0.800 | -0.500 | 0.103 | 0.020 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.070 | 0.370 | 0.169 | 0.093 | 0.254 |
| 0.500 | -0.500 | 0.115 | 0.018 | 0.010 | 0.009 | 0.010 | 0.009 | 0.008 | 0.055 | 0.282 | 0.184 | 0.105 | 0.328 |
| -0.500 | -0.500 | 0.104 | 0.044 | 0.007 | 0.007 | 0.008 | 0.006 | 0.007 | 0.105 | 0.093 | 0.243 | 0.132 | 0.425 |
| -0.800 | -0.500 | 0.090 | 0.064 | 0.007 | 0.007 | 0.007 | 0.007 | 0.007 | 0.139 | 0.096 | 0.235 | 0.126 | 0.351 |
| 0.800 | -0.800 | 0.110 | 0.054 | 0.013 | 0.012 | 0.014 | 0.012 | 0.012 | 0.101 | 0.324 | 0.166 | 0.089 | 0.260 |
| 0.500 | -0.800 | 0.114 | 0.030 | 0.011 | 0.011 | 0.011 | 0.011 | 0.011 | 0.079 | 0.224 | 0.180 | 0.096 | 0.340 |
| -0.500 | -0.800 | 0.106 | 0.069 | 0.014 | 0.014 | 0.015 | 0.014 | 0.014 | 0.099 | 0.083 | 0.215 | 0.122 | 0.435 |
| -0.800 | -0.800 | 0.094 | 0.114 | 0.011 | 0.010 | 0.011 | 0.011 | 0.010 | 0.100 | 0.085 | 0.223 | 0.123 | 0.360 |
| 0.000 | 0.000 | 0.115 | 0.019 | 0.008 | 0.007 | 0.011 | 0.008 | 0.007 | 0.060 | 0.169 | 0.239 | 0.109 | 0.393 |

Note: A constant and a time trend are included.
Table 7. Finite Sample Sizes ( $\alpha=1.0, n=50$ )

| $\beta$ | $\phi$ | ADF | $\mathrm{Z}_{\alpha}$ | $\mathrm{MZ}_{\alpha}$ | $\mathrm{MZ}_{\mathrm{t}}$ | MSB | $\mathrm{P}_{\mathrm{t}}$ | $\mathrm{MP}_{\mathrm{t}}$ | $\mathrm{DF}^{\mathrm{GLS}}$ | EJ | CADF | BCADF | $\rho^{2}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0.800 | 0.800 | 0.117 | 0.125 | 0.121 | 0.116 | 0.132 | 0.106 | 0.114 | 0.036 | 0.767 | 0.102 | 0.073 | 0.032 |
| 0.500 | 0.800 | 0.081 | 0.054 | 0.052 | 0.047 | 0.059 | 0.041 | 0.043 | 0.012 | 0.686 | 0.136 | 0.082 | 0.080 |
| -0.500 | 0.800 | 0.075 | 0.021 | 0.020 | 0.018 | 0.023 | 0.018 | 0.017 | 0.007 | 0.579 | 0.125 | 0.085 | 0.102 |
| -0.800 | 0.800 | 0.083 | 0.044 | 0.042 | 0.038 | 0.051 | 0.034 | 0.036 | 0.015 | 0.756 | 0.055 | 0.061 | 0.022 |
| 0.800 | 0.500 | 0.123 | 0.076 | 0.071 | 0.064 | 0.079 | 0.062 | 0.064 | 0.039 | 0.663 | 0.122 | 0.075 | 0.090 |
| 0.500 | 0.500 | 0.115 | 0.035 | 0.031 | 0.030 | 0.036 | 0.027 | 0.028 | 0.019 | 0.512 | 0.177 | 0.105 | 0.180 |
| -0.500 | 0.500 | 0.108 | 0.012 | 0.003 | 0.003 | 0.003 | 0.004 | 0.003 | 0.024 | 0.237 | 0.221 | 0.129 | 0.334 |
| -0.800 | 0.500 | 0.103 | 0.012 | 0.008 | 0.008 | 0.010 | 0.007 | 0.008 | 0.012 | 0.456 | 0.131 | 0.099 | 0.164 |
| 0.800 | -0.500 | 0.109 | 0.017 | 0.005 | 0.005 | 0.005 | 0.005 | 0.005 | 0.025 | 0.321 | 0.160 | 0.109 | 0.298 |
| 0.500 | -0.500 | 0.100 | 0.013 | 0.007 | 0.007 | 0.009 | 0.006 | 0.007 | 0.019 | 0.236 | 0.144 | 0.098 | 0.386 |
| -0.500 | -0.500 | 0.096 | 0.038 | 0.010 | 0.009 | 0.011 | 0.009 | 0.009 | 0.053 | 0.080 | 0.191 | 0.155 | 0.563 |
| -0.800 | -0.500 | 0.096 | 0.046 | 0.010 | 0.011 | 0.011 | 0.011 | 0.011 | 0.049 | 0.098 | 0.200 | 0.155 | 0.473 |
| 0.800 | -0.800 | 0.106 | 0.035 | 0.012 | 0.012 | 0.013 | 0.011 | 0.012 | 0.031 | 0.265 | 0.145 | 0.106 | 0.321 |
| 0.500 | -0.800 | 0.098 | 0.023 | 0.009 | 0.008 | 0.011 | 0.008 | 0.008 | 0.031 | 0.219 | 0.162 | 0.122 | 0.405 |
| -0.500 | -0.800 | 0.114 | 0.040 | 0.015 | 0.014 | 0.016 | 0.013 | 0.014 | 0.036 | 0.083 | 0.186 | 0.154 | 0.574 |
| -0.800 | -0.800 | 0.100 | 0.071 | 0.016 | 0.015 | 0.017 | 0.014 | 0.015 | 0.024 | 0.075 | 0.188 | 0.149 | 0.501 |
| 0.000 | 0.000 | 0.107 | 0.012 | 0.005 | 0.005 | 0.006 | 0.004 | 0.005 | 0.020 | 0.141 | 0.197 | 0.127 | 0.512 |

[^13]Table 8. Finite Sample Sizes ( $\alpha=1.0, n=100$ )

| $\beta$ | $\phi$ | ADF | $\mathrm{Z}_{\alpha}$ | $\mathrm{MZ}_{\alpha}$ | $\mathrm{MZ}_{\mathrm{t}}$ | MSB | $\mathrm{P}_{\mathrm{t}}$ | $\mathrm{MP}_{\mathrm{t}}$ | $\mathrm{DF}^{\mathrm{GLS}}$ | EJ | CADF | BCADF | $\rho^{2}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0.800 | 0.800 | 0.076 | 0.064 | 0.062 | 0.057 | 0.073 | 0.051 | 0.054 | 0.032 | 0.756 | 0.065 | 0.068 | 0.029 |
| 0.500 | 0.800 | 0.056 | 0.026 | 0.026 | 0.024 | 0.028 | 0.021 | 0.023 | 0.012 | 0.722 | 0.095 | 0.077 | 0.092 |
| -0.500 | 0.800 | 0.047 | 0.014 | 0.014 | 0.013 | 0.016 | 0.012 | 0.013 | 0.006 | 0.526 | 0.046 | 0.075 | 0.134 |
| -0.800 | 0.800 | 0.053 | 0.021 | 0.020 | 0.018 | 0.024 | 0.017 | 0.017 | 0.009 | 0.766 | 0.020 | 0.053 | 0.017 |
| 0.800 | 0.500 | 0.087 | 0.063 | 0.060 | 0.058 | 0.062 | 0.057 | 0.057 | 0.051 | 0.554 | 0.098 | 0.075 | 0.116 |
| 0.500 | 0.500 | 0.079 | 0.043 | 0.041 | 0.040 | 0.044 | 0.038 | 0.039 | 0.034 | 0.338 | 0.103 | 0.081 | 0.258 |
| -0.500 | 0.500 | 0.088 | 0.017 | 0.012 | 0.012 | 0.012 | 0.011 | 0.012 | 0.018 | 0.110 | 0.102 | 0.101 | 0.508 |
| -0.800 | 0.500 | 0.071 | 0.015 | 0.013 | 0.013 | 0.015 | 0.011 | 0.012 | 0.015 | 0.292 | 0.061 | 0.080 | 0.224 |
| 0.800 | -0.500 | 0.079 | 0.014 | 0.009 | 0.008 | 0.009 | 0.007 | 0.008 | 0.014 | 0.155 | 0.103 | 0.092 | 0.408 |
| 0.500 | -0.500 | 0.089 | 0.018 | 0.012 | 0.011 | 0.013 | 0.012 | 0.012 | 0.019 | 0.121 | 0.103 | 0.098 | 0.513 |
| -0.500 | -0.500 | 0.075 | 0.033 | 0.017 | 0.018 | 0.019 | 0.018 | 0.019 | 0.032 | 0.050 | 0.109 | 0.111 | 0.763 |
| -0.800 | -0.500 | 0.073 | 0.037 | 0.017 | 0.018 | 0.019 | 0.019 | 0.018 | 0.028 | 0.060 | 0.107 | 0.112 | 0.607 |
| 0.800 | -0.800 | 0.079 | 0.035 | 0.025 | 0.023 | 0.026 | 0.025 | 0.024 | 0.021 | 0.117 | 0.090 | 0.084 | 0.426 |
| 0.500 | -0.800 | 0.089 | 0.021 | 0.013 | 0.014 | 0.015 | 0.014 | 0.014 | 0.017 | 0.103 | 0.097 | 0.096 | 0.542 |
| -0.500 | -0.800 | 0.082 | 0.045 | 0.029 | 0.028 | 0.031 | 0.028 | 0.027 | 0.026 | 0.050 | 0.099 | 0.105 | 0.785 |
| -0.800 | -0.800 | 0.077 | 0.057 | 0.022 | 0.020 | 0.023 | 0.019 | 0.021 | 0.016 | 0.056 | 0.106 | 0.112 | 0.672 |
| 0.000 | 0.000 | 0.075 | 0.014 | 0.009 | 0.009 | 0.010 | 0.008 | 0.009 | 0.017 | 0.067 | 0.099 | 0.094 | 0.717 |

[^14]Table 9. Finite Sample Powers ( $\alpha=0.9, n=25$ )

| $\beta$ | $\phi$ | ADF | $\mathrm{Z}_{\alpha}$ | $\mathrm{MZ}_{\alpha}$ | $\mathrm{MZ}_{\mathrm{t}}$ | MSB | $\mathrm{P}_{\mathrm{t}}$ | $\mathrm{MP}_{\mathrm{t}}$ | $\mathrm{DF}^{\mathrm{GLS}}$ | EJ | CADF | BCADF | $\rho^{2}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0.800 | 0.800 | 0.049 | 0.020 | 0.020 | 0.020 | 0.020 | 0.022 | 0.020 | 0.062 | 0.251 | 0.325 | 0.297 | 0.080 |
| 0.500 | 0.800 | 0.052 | 0.026 | 0.027 | 0.027 | 0.027 | 0.034 | 0.030 | 0.059 | 0.180 | 0.198 | 0.198 | 0.129 |
| -0.500 | 0.800 | 0.059 | 0.035 | 0.034 | 0.036 | 0.030 | 0.038 | 0.037 | 0.083 | 0.152 | 0.137 | 0.186 | 0.123 |
| -0.800 | 0.800 | 0.058 | 0.027 | 0.027 | 0.027 | 0.028 | 0.027 | 0.026 | 0.062 | 0.214 | 0.339 | 0.275 | 0.049 |
| 0.800 | 0.500 | 0.051 | 0.031 | 0.033 | 0.033 | 0.029 | 0.041 | 0.038 | 0.063 | 0.177 | 0.201 | 0.217 | 0.139 |
| 0.500 | 0.500 | 0.070 | 0.035 | 0.036 | 0.040 | 0.035 | 0.045 | 0.043 | 0.072 | 0.158 | 0.108 | 0.157 | 0.222 |
| -0.500 | 0.500 | 0.052 | 0.064 | 0.063 | 0.065 | 0.058 | 0.071 | 0.066 | 0.097 | 0.085 | 0.067 | 0.141 | 0.297 |
| -0.800 | 0.500 | 0.061 | 0.053 | 0.052 | 0.048 | 0.041 | 0.052 | 0.049 | 0.082 | 0.132 | 0.115 | 0.173 | 0.164 |
| 0.800 | -0.500 | 0.066 | 0.054 | 0.051 | 0.053 | 0.044 | 0.062 | 0.053 | 0.086 | 0.145 | 0.100 | 0.152 | 0.282 |
| 0.500 | -0.500 | 0.055 | 0.052 | 0.053 | 0.053 | 0.043 | 0.059 | 0.053 | 0.081 | 0.098 | 0.066 | 0.131 | 0.368 |
| -0.500 | -0.500 | 0.059 | 0.061 | 0.052 | 0.055 | 0.043 | 0.063 | 0.055 | 0.077 | 0.073 | 0.063 | 0.128 | 0.493 |
| -0.800 | -0.500 | 0.061 | 0.072 | 0.074 | 0.076 | 0.064 | 0.072 | 0.077 | 0.104 | 0.066 | 0.064 | 0.134 | 0.412 |
| 0.800 | -0.800 | 0.053 | 0.049 | 0.042 | 0.044 | 0.043 | 0.050 | 0.044 | 0.075 | 0.106 | 0.092 | 0.141 | 0.288 |
| 0.500 | -0.800 | 0.061 | 0.067 | 0.061 | 0.065 | 0.051 | 0.072 | 0.067 | 0.095 | 0.117 | 0.080 | 0.145 | 0.378 |
| -0.500 | -0.800 | 0.060 | 0.065 | 0.052 | 0.054 | 0.051 | 0.065 | 0.059 | 0.102 | 0.066 | 0.054 | 0.116 | 0.507 |
| -0.800 | -0.800 | 0.077 | 0.086 | 0.059 | 0.059 | 0.052 | 0.062 | 0.061 | 0.075 | 0.072 | 0.069 | 0.131 | 0.419 |
| 0.000 | 0.000 | 0.057 | 0.060 | 0.055 | 0.050 | 0.054 | 0.056 | 0.052 | 0.087 | 0.078 | 0.060 | 0.121 | 0.475 |

Note: Only a constant is included.
Table 10. Finite Sample Powers ( $\alpha=0.9, n=50$ )

| $\beta$ | $\phi$ | ADF | $\mathrm{Z}_{\alpha}$ | $\mathrm{MZ}_{\alpha}$ | $\mathrm{MZ}_{\mathrm{t}}$ | MSB | $\mathrm{P}_{\mathrm{t}}$ | $\mathrm{MP}_{\mathrm{t}}$ | $\mathrm{DF}^{\mathrm{GLS}}$ | EJ | CADF | $\mathrm{BCADF}^{2}$ |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0.800 | 0.800 | 0.070 | 0.067 | 0.069 | 0.072 | 0.060 | 0.088 | 0.080 | 0.133 | 0.638 | 0.776 | 0.717 | 0.036 |
| 0.500 | 0.800 | 0.076 | 0.064 | 0.064 | 0.072 | 0.054 | 0.082 | 0.077 | 0.122 | 0.520 | 0.532 | 0.532 | 0.092 |
| -0.500 | 0.800 | 0.068 | 0.076 | 0.072 | 0.073 | 0.059 | 0.085 | 0.074 | 0.132 | 0.411 | 0.435 | 0.506 | 0.124 |
| -0.800 | 0.800 | 0.073 | 0.064 | 0.067 | 0.076 | 0.058 | 0.078 | 0.079 | 0.122 | 0.596 | 0.771 | 0.722 | 0.029 |
| 0.800 | 0.500 | 0.086 | 0.104 | 0.108 | 0.112 | 0.090 | 0.126 | 0.113 | 0.124 | 0.465 | 0.543 | 0.548 | 0.100 |
| 0.500 | 0.500 | 0.096 | 0.121 | 0.119 | 0.127 | 0.105 | 0.125 | 0.124 | 0.139 | 0.335 | 0.308 | 0.404 | 0.204 |
| -0.500 | 0.500 | 0.114 | 0.154 | 0.140 | 0.143 | 0.128 | 0.142 | 0.143 | 0.174 | 0.156 | 0.148 | 0.255 | 0.382 |
| -0.800 | 0.500 | 0.114 | 0.124 | 0.123 | 0.124 | 0.109 | 0.125 | 0.124 | 0.154 | 0.317 | 0.293 | 0.396 | 0.190 |
| 0.800 | -0.500 | 0.094 | 0.138 | 0.136 | 0.141 | 0.120 | 0.157 | 0.144 | 0.174 | 0.275 | 0.211 | 0.306 | 0.310 |
| 0.500 | -0.500 | 0.095 | 0.116 | 0.111 | 0.120 | 0.094 | 0.117 | 0.119 | 0.147 | 0.190 | 0.170 | 0.269 | 0.397 |
| -0.500 | -0.500 | 0.091 | 0.156 | 0.150 | 0.148 | 0.138 | 0.147 | 0.150 | 0.168 | 0.094 | 0.109 | 0.203 | 0.599 |
| -0.800 | -0.500 | 0.097 | 0.156 | 0.146 | 0.148 | 0.126 | 0.149 | 0.151 | 0.159 | 0.090 | 0.099 | 0.197 | 0.509 |
| 0.800 | -0.800 | 0.098 | 0.132 | 0.121 | 0.129 | 0.101 | 0.134 | 0.135 | 0.170 | 0.227 | 0.202 | 0.289 | 0.339 |
| 0.500 | -0.800 | 0.083 | 0.135 | 0.135 | 0.140 | 0.118 | 0.148 | 0.143 | 0.153 | 0.157 | 0.111 | 0.221 | 0.432 |
| -0.500 | -0.800 | 0.106 | 0.140 | 0.117 | 0.121 | 0.105 | 0.126 | 0.123 | 0.166 | 0.092 | 0.093 | 0.193 | 0.620 |
| -0.800 | -0.800 | 0.101 | 0.124 | 0.086 | 0.095 | 0.078 | 0.103 | 0.096 | 0.143 | 0.093 | 0.086 | 0.191 | 0.538 |
| 0.000 | 0.000 | 0.094 | 0.150 | 0.138 | 0.139 | 0.130 | 0.149 | 0.138 | 0.181 | 0.117 | 0.101 | 0.195 | 0.556 |
| Note: Only a constant is included. |  |  |  |  |  |  |  |  |  |  |  |  |  |

Table 11. Finite Sample Powers $(\alpha=0.9, n=100)$

| $\beta$ | $\phi$ | ADF | $\mathrm{Z}_{\alpha}$ | $\mathrm{MZ}_{\alpha}$ | $\mathrm{MZ}_{\mathrm{t}}$ | MSB | $\mathrm{P}_{\mathrm{t}}$ | $\mathrm{MP}_{\mathrm{t}}$ | $\mathrm{DF}^{\mathrm{GLS}}$ | EJ | CADF | BCADF | $\rho^{2}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0.800 | 0.800 | 0.129 | 0.233 | 0.234 | 0.252 | 0.207 | 0.245 | 0.255 | 0.282 | 0.996 | 0.996 | 0.995 | 0.039 |
| 0.500 | 0.800 | 0.130 | 0.209 | 0.209 | 0.221 | 0.197 | 0.224 | 0.223 | 0.252 | 0.956 | 0.965 | 0.960 | 0.106 |
| -0.500 | 0.800 | 0.138 | 0.312 | 0.309 | 0.311 | 0.276 | 0.318 | 0.315 | 0.371 | 0.920 | 0.913 | 0.905 | 0.150 |
| -0.800 | 0.800 | 0.154 | 0.228 | 0.228 | 0.250 | 0.221 | 0.264 | 0.251 | 0.296 | 0.992 | 0.989 | 0.986 | 0.020 |
| 0.800 | 0.500 | 0.249 | 0.305 | 0.300 | 0.320 | 0.275 | 0.324 | 0.328 | 0.336 | 0.940 | 0.955 | 0.957 | 0.130 |
| 0.500 | 0.500 | 0.192 | 0.319 | 0.315 | 0.319 | 0.306 | 0.331 | 0.324 | 0.348 | 0.793 | 0.796 | 0.836 | 0.268 |
| -0.500 | 0.500 | 0.245 | 0.369 | 0.363 | 0.392 | 0.342 | 0.380 | 0.391 | 0.415 | 0.502 | 0.376 | 0.474 | 0.514 |
| -0.800 | 0.500 | 0.225 | 0.305 | 0.293 | 0.303 | 0.270 | 0.298 | 0.301 | 0.324 | 0.765 | 0.760 | 0.759 | 0.236 |
| 0.800 | -0.500 | 0.226 | 0.354 | 0.350 | 0.353 | 0.321 | 0.353 | 0.350 | 0.397 | 0.639 | 0.594 | 0.652 | 0.404 |
| 0.500 | -0.500 | 0.222 | 0.286 | 0.281 | 0.297 | 0.267 | 0.276 | 0.297 | 0.305 | 0.528 | 0.477 | 0.570 | 0.510 |
| -0.500 | -0.500 | 0.253 | 0.371 | 0.356 | 0.363 | 0.329 | 0.350 | 0.369 | 0.371 | 0.294 | 0.223 | 0.327 | 0.770 |
| -0.800 | -0.500 | 0.243 | 0.389 | 0.361 | 0.373 | 0.335 | 0.357 | 0.372 | 0.401 | 0.327 | 0.223 | 0.320 | 0.614 |
| 0.800 | -0.800 | 0.252 | 0.336 | 0.304 | 0.333 | 0.268 | 0.326 | 0.338 | 0.379 | 0.620 | 0.542 | 0.616 | 0.425 |
| 0.500 | -0.800 | 0.232 | 0.335 | 0.310 | 0.327 | 0.298 | 0.320 | 0.332 | 0.358 | 0.471 | 0.403 | 0.527 | 0.538 |
| -0.500 | -0.800 | 0.262 | 0.304 | 0.275 | 0.282 | 0.238 | 0.293 | 0.298 | 0.310 | 0.265 | 0.221 | 0.299 | 0.793 |
| -0.800 | -0.800 | 0.236 | 0.349 | 0.266 | 0.277 | 0.248 | 0.287 | 0.276 | 0.334 | 0.329 | 0.186 | 0.280 | 0.682 |
| 0.000 | 0.000 | 0.272 | 0.328 | 0.318 | 0.344 | 0.275 | 0.330 | 0.344 | 0.361 | 0.334 | 0.311 | 0.376 | 0.722 |
| Note: Only a constant is included. |  |  |  |  |  |  |  |  |  |  |  |  |  |

Table 12. Finite Sample Powers ( $\alpha=0.9, n=25$ )

| $\beta$ | $\phi$ | ADF | $\mathrm{Z}_{\alpha}$ | $\mathrm{MZ}_{\alpha}$ | $\mathrm{MZ}_{\mathrm{t}}$ | MSB | $\mathrm{P}_{\mathrm{t}}$ | $\mathrm{MP}_{\mathrm{t}}$ | $\mathrm{DF}^{\mathrm{GLS}}$ | EJ | CADF | BCADF | $\rho^{2}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0.800 | 0.800 | 0.061 | 0.032 | 0.033 | 0.033 | 0.033 | 0.033 | 0.035 | 0.039 | 0.122 | 0.132 | 0.145 | 0.068 |
| 0.500 | 0.800 | 0.060 | 0.034 | 0.035 | 0.033 | 0.035 | 0.034 | 0.032 | 0.044 | 0.089 | 0.087 | 0.130 | 0.104 |
| -0.500 | 0.800 | 0.053 | 0.050 | 0.046 | 0.046 | 0.041 | 0.045 | 0.045 | 0.063 | 0.085 | 0.056 | 0.120 | 0.087 |
| -0.800 | 0.800 | 0.054 | 0.029 | 0.029 | 0.030 | 0.029 | 0.030 | 0.030 | 0.055 | 0.094 | 0.114 | 0.128 | 0.041 |
| 0.800 | 0.500 | 0.049 | 0.035 | 0.037 | 0.036 | 0.037 | 0.037 | 0.037 | 0.052 | 0.078 | 0.103 | 0.131 | 0.118 |
| 0.500 | 0.500 | 0.046 | 0.044 | 0.042 | 0.044 | 0.044 | 0.044 | 0.044 | 0.062 | 0.078 | 0.062 | 0.114 | 0.183 |
| -0.500 | 0.500 | 0.053 | 0.053 | 0.054 | 0.056 | 0.052 | 0.061 | 0.054 | 0.058 | 0.067 | 0.061 | 0.135 | 0.225 |
| -0.800 | 0.500 | 0.054 | 0.043 | 0.039 | 0.040 | 0.037 | 0.038 | 0.041 | 0.053 | 0.081 | 0.063 | 0.120 | 0.132 |
| 0.800 | -0.500 | 0.052 | 0.055 | 0.056 | 0.061 | 0.055 | 0.058 | 0.058 | 0.053 | 0.081 | 0.067 | 0.118 | 0.251 |
| 0.500 | -0.500 | 0.055 | 0.058 | 0.053 | 0.060 | 0.049 | 0.056 | 0.058 | 0.060 | 0.072 | 0.059 | 0.123 | 0.323 |
| -0.500 | -0.500 | 0.054 | 0.061 | 0.060 | 0.056 | 0.055 | 0.054 | 0.055 | 0.072 | 0.057 | 0.049 | 0.138 | 0.425 |
| -0.800 | -0.500 | 0.051 | 0.068 | 0.070 | 0.063 | 0.063 | 0.062 | 0.065 | 0.069 | 0.057 | 0.049 | 0.131 | 0.352 |
| 0.800 | -0.800 | 0.047 | 0.048 | 0.042 | 0.044 | 0.043 | 0.044 | 0.041 | 0.054 | 0.058 | 0.058 | 0.111 | 0.256 |
| 0.500 | -0.800 | 0.058 | 0.047 | 0.050 | 0.053 | 0.048 | 0.050 | 0.052 | 0.057 | 0.081 | 0.073 | 0.119 | 0.331 |
| -0.500 | -0.800 | 0.058 | 0.056 | 0.052 | 0.057 | 0.051 | 0.061 | 0.057 | 0.062 | 0.061 | 0.054 | 0.123 | 0.439 |
| -0.800 | -0.800 | 0.070 | 0.074 | 0.053 | 0.054 | 0.053 | 0.053 | 0.054 | 0.063 | 0.053 | 0.055 | 0.131 | 0.364 |
| 0.000 | 0.000 | 0.048 | 0.050 | 0.052 | 0.055 | 0.048 | 0.051 | 0.056 | 0.056 | 0.065 | 0.054 | 0.116 | 0.397 |

[^15]Table 13. Finite Sample Powers ( $\alpha=0.9, n=50$ )

| $\beta$ | $\phi$ | ADF | $\mathrm{Z}_{\alpha}$ | $\mathrm{MZ}_{\alpha}$ | $\mathrm{MZ}_{\mathrm{t}}$ | MSB | $\mathrm{P}_{\mathrm{t}}$ | $\mathrm{MP}_{\mathrm{t}}$ | DF $^{\text {GLS }}$ | EJ | CADF | BCADF | $\rho^{2}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0.800 | 0.800 | 0.058 | 0.036 | 0.036 | 0.035 | 0.036 | 0.037 | 0.036 | 0.084 | 0.362 | 0.458 | 0.481 | 0.032 |
| 0.500 | 0.800 | 0.065 | 0.041 | 0.039 | 0.044 | 0.040 | 0.044 | 0.044 | 0.080 | 0.231 | 0.299 | 0.326 | 0.075 |
| -0.500 | 0.800 | 0.070 | 0.046 | 0.041 | 0.049 | 0.035 | 0.049 | 0.050 | 0.076 | 0.202 | 0.182 | 0.308 | 0.098 |
| -0.800 | 0.800 | 0.050 | 0.044 | 0.043 | 0.044 | 0.040 | 0.044 | 0.044 | 0.068 | 0.309 | 0.506 | 0.494 | 0.023 |
| 0.800 | 0.500 | 0.061 | 0.064 | 0.064 | 0.064 | 0.056 | 0.070 | 0.070 | 0.080 | 0.221 | 0.308 | 0.346 | 0.084 |
| 0.500 | 0.500 | 0.064 | 0.063 | 0.063 | 0.064 | 0.059 | 0.071 | 0.064 | 0.072 | 0.169 | 0.157 | 0.260 | 0.175 |
| -0.500 | 0.500 | 0.069 | 0.091 | 0.086 | 0.090 | 0.086 | 0.088 | 0.088 | 0.097 | 0.094 | 0.071 | 0.198 | 0.335 |
| -0.800 | 0.500 | 0.079 | 0.093 | 0.080 | 0.085 | 0.077 | 0.084 | 0.083 | 0.104 | 0.182 | 0.171 | 0.294 | 0.161 |
| 0.800 | -0.500 | 0.076 | 0.095 | 0.083 | 0.089 | 0.080 | 0.088 | 0.087 | 0.098 | 0.158 | 0.134 | 0.230 | 0.287 |
| 0.500 | -0.500 | 0.070 | 0.072 | 0.067 | 0.072 | 0.057 | 0.070 | 0.073 | 0.088 | 0.117 | 0.114 | 0.200 | 0.368 |
| -0.500 | -0.500 | 0.071 | 0.063 | 0.064 | 0.065 | 0.063 | 0.068 | 0.064 | 0.079 | 0.061 | 0.065 | 0.182 | 0.559 |
| -0.800 | -0.500 | 0.062 | 0.089 | 0.085 | 0.085 | 0.089 | 0.088 | 0.084 | 0.094 | 0.060 | 0.061 | 0.183 | 0.474 |
| 0.800 | -0.800 | 0.073 | 0.085 | 0.078 | 0.080 | 0.077 | 0.088 | 0.079 | 0.104 | 0.133 | 0.122 | 0.210 | 0.316 |
| 0.500 | -0.800 | 0.065 | 0.090 | 0.076 | 0.084 | 0.080 | 0.084 | 0.085 | 0.099 | 0.092 | 0.089 | 0.187 | 0.402 |
| -0.500 | -0.800 | 0.069 | 0.075 | 0.065 | 0.065 | 0.070 | 0.067 | 0.066 | 0.077 | 0.073 | 0.062 | 0.185 | 0.576 |
| -0.800 | -0.800 | 0.082 | 0.099 | 0.058 | 0.058 | 0.055 | 0.058 | 0.058 | 0.080 | 0.075 | 0.064 | 0.182 | 0.502 |
| 0.000 | 0.000 | 0.078 | 0.083 | 0.079 | 0.080 | 0.073 | 0.084 | 0.082 | 0.095 | 0.092 | 0.080 | 0.172 | 0.507 |

[^16]Table 14. Finite Sample Powers ( $\alpha=0.9, n=100$ )

| $\beta$ | $\phi$ | ADF | $\mathrm{Z}_{\alpha}$ | $\mathrm{MZ}_{\alpha}$ | $\mathrm{MZ}_{\mathrm{t}}$ | MSB | $\mathrm{P}_{\mathrm{t}}$ | $\mathrm{MP}_{\mathrm{t}}$ | $\mathrm{DF}^{\mathrm{GLS}}$ | EJ | CADF | BCADF | $\rho^{2}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0.800 | 0.800 | 0.089 | 0.088 | 0.087 | 0.092 | 0.080 | 0.104 | 0.097 | 0.149 | 0.943 | 0.989 | 0.985 | 0.036 |
| 0.500 | 0.800 | 0.107 | 0.121 | 0.117 | 0.132 | 0.111 | 0.128 | 0.131 | 0.167 | 0.813 | 0.911 | 0.906 | 0.096 |
| -0.500 | 0.800 | 0.107 | 0.158 | 0.153 | 0.155 | 0.148 | 0.157 | 0.156 | 0.202 | 0.722 | 0.821 | 0.837 | 0.133 |
| -0.800 | 0.800 | 0.110 | 0.126 | 0.130 | 0.137 | 0.118 | 0.143 | 0.137 | 0.184 | 0.923 | 0.974 | 0.967 | 0.017 |
| 0.800 | 0.500 | 0.135 | 0.144 | 0.143 | 0.150 | 0.138 | 0.149 | 0.152 | 0.160 | 0.775 | 0.879 | 0.893 | 0.121 |
| 0.500 | 0.500 | 0.125 | 0.160 | 0.160 | 0.163 | 0.149 | 0.164 | 0.163 | 0.176 | 0.575 | 0.657 | 0.716 | 0.252 |
| -0.500 | 0.500 | 0.146 | 0.202 | 0.184 | 0.182 | 0.177 | 0.184 | 0.180 | 0.225 | 0.284 | 0.235 | 0.380 | 0.498 |
| -0.800 | 0.500 | 0.133 | 0.157 | 0.144 | 0.150 | 0.142 | 0.149 | 0.153 | 0.176 | 0.537 | 0.590 | 0.665 | 0.219 |
| 0.800 | -0.500 | 0.134 | 0.180 | 0.166 | 0.183 | 0.157 | 0.183 | 0.185 | 0.204 | 0.454 | 0.433 | 0.522 | 0.396 |
| 0.500 | -0.500 | 0.137 | 0.138 | 0.122 | 0.122 | 0.117 | 0.131 | 0.125 | 0.161 | 0.347 | 0.330 | 0.447 | 0.501 |
| -0.500 | -0.500 | 0.144 | 0.209 | 0.198 | 0.197 | 0.177 | 0.204 | 0.195 | 0.225 | 0.159 | 0.135 | 0.255 | 0.761 |
| -0.800 | -0.500 | 0.131 | 0.197 | 0.180 | 0.188 | 0.175 | 0.193 | 0.187 | 0.205 | 0.160 | 0.137 | 0.258 | 0.605 |
| 0.800 | -0.800 | 0.150 | 0.179 | 0.157 | 0.158 | 0.158 | 0.168 | 0.161 | 0.187 | 0.429 | 0.408 | 0.491 | 0.416 |
| 0.500 | -0.800 | 0.142 | 0.194 | 0.174 | 0.181 | 0.176 | 0.184 | 0.185 | 0.204 | 0.329 | 0.300 | 0.413 | 0.529 |
| -0.500 | -0.800 | 0.153 | 0.165 | 0.125 | 0.132 | 0.124 | 0.145 | 0.130 | 0.162 | 0.161 | 0.119 | 0.226 | 0.783 |
| -0.800 | -0.800 | 0.159 | 0.172 | 0.112 | 0.111 | 0.110 | 0.113 | 0.111 | 0.173 | 0.148 | 0.117 | 0.215 | 0.673 |
| 0.000 | 0.000 | 0.159 | 0.178 | 0.164 | 0.161 | 0.157 | 0.160 | 0.162 | 0.202 | 0.214 | 0.183 | 0.287 | 0.711 |

[^17]Table 15. Tests for a Unit Root in the Extended Nelson-Plosser Data

| Series | ADF | $\mathrm{Z}_{\alpha}$ | $\mathrm{MZ}_{\alpha}$ | $\mathrm{MZ}_{\mathrm{t}}$ | MSB | $\mathrm{P}_{\mathrm{t}}$ | $\mathrm{MP}_{\mathrm{t}}$ | $\mathrm{DF}^{\mathrm{GLS}}$ | EJ | CADF | BCADF | $\rho^{2}$ | Covariate |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Real GNP | -3.201 | -7.634 | -6.631 | -1.821 | 0.275 | 14.624 | 13.742 | -2.096 | 61.159 | 1.504 | 1.504 | 0.000 | Real pc GNP |
| Nominal GNP | -2.170 | -17.091 | -15.509 | -2.771 | 0.179 | 7.724 | 5.958 | $-3.190^{*}$ | 26.853 | 0.397 | 0.397 | 0.000 | Money Stock |
| Real per capita GNP | -3.107 | -6.300 | -5.724 | -1.691 | 0.295 | 15.814 | 15.919 | -1.861 | $26.041^{*}$ | -1.863 | -1.863 | 0.000 | Ind. Prod. |
| Industrial Production | -2.942 | -10.919 | -9.365 | -2.163 | 0.231 | 9.847 | 9.733 | -2.522 | $30.366^{*}$ | -0.363 | -0.363 | 0.000 | Money Stock |
| Employment | -2.776 | -6.463 | -5.780 | -1.698 | 0.294 | 16.081 | 15.762 | -1.899 | $36.418^{*}$ | -1.764 | -1.764 | 0.000 | Unemployment |
| Unemployment | -2.733 | -6.150 | -5.681 | -1.666 | 0.293 | 15.454 | 15.998 | -1.804 | $15.785^{*}$ | 0.454 | 0.454 | 0.000 | Employment |
| GNP Deflator | -2.529 | -6.982 | -5.569 | -1.628 | 0.292 | 21.494 | 16.257 | -2.060 | $41.839^{*}$ | $-2.782^{*}$ | -2.782 | 0.000 | Nominal GNP |
| Consumer Prices | -1.852 | -6.158 | -4.584 | -1.462 | 0.319 | 26.806 | 19.495 | -1.986 | 52.776 | -1.411 | -1.411 | 0.000 | GNP Deflator |
| Wages | -1.908 | -5.574 | -4.134 | -1.431 | 0.346 | 29.116 | 21.971 | -2.059 | 29.348 | $-4.618^{*}$ | $-4.618^{*}$ | 0.003 | S\&P 500 |
| Real Wages | -3.122 | -6.471 | -5.898 | -1.675 | 0.284 | 15.174 | 15.389 | -1.838 | $7.231^{*}$ | -0.500 | -0.500 | 0.019 | Real GNP |
| Money Stock | -2.779 | $-18.486^{*}$ | $-17.981^{*}$ | $-2.990^{*}$ | $0.166^{*}$ | 5.513 | $5.118^{*}$ | $-3.301^{*}$ | $8.562^{*}$ | $-3.795^{*}$ | $-3.795^{*}$ | 0.000 | Nominal GNP |
| Velocity | -1.990 | -10.621 | -8.645 | -2.030 | 0.235 | 12.400 | 10.709 | -2.494 | $4.185^{*}$ | $-4.094^{*}$ | $-4.094^{*}$ | 0.002 | Nominal GNP |
| Bond Yield | -2.375 | -2.348 | -1.534 | -0.717 | 0.467 | 52.164 | 44.501 | -1.098 | 21.649 | -1.537 | -1.537 | 0.085 | Nominal GNP |
| S\&P 500 | -1.657 | -7.078 | -4.912 | -1.558 | 0.317 | 27.302 | 18.499 | -2.245 | 15.962 | $-2.422^{*}$ | -2.422 | 0.008 | Money Stock |
| No of Rejections | 0.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.000 | 1.000 | 2.000 | 8.000 | 5.000 | 3.000 |  |  |
| Note: A constant and a time trend are included. * indicates rejection at the $5 \%$ | level. Bootstrap repetition is 5000. |  |  |  |  |  |  |  |  |  |  |  |  |

Table 16. Tests for a Unit Root in Real Exchange Rates

| Series | ADF | $\mathrm{Z}_{\alpha}$ | $\mathrm{MZ}_{\alpha}$ | $\mathrm{MZ}_{\mathrm{t}}$ | MSB | $\mathrm{P}_{\mathrm{t}}$ | $\mathrm{MP}_{\mathrm{t}}$ | $\mathrm{DF}^{\mathrm{GLS}}$ | EJ | CADF | BCADF | $\rho^{2}$ | Covariate |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Australia | -2.438 | -2.936 | -2.453 | -1.105 | 0.451 | 15.568 | 9.975 | -1.028 | $14.680^{*}$ | $-2.823^{*}$ | $-2.823^{*}$ | 0.056 | Japan |
| Austria | -1.514 | -3.726 | -3.519 | -1.215 | 0.345 | 8.288 | 6.944 | -1.192 | 22.543 | -0.929 | -0.929 | 0.000 | Switzerlands |
| Belgium | -2.609 | -4.585 | -4.260 | -1.432 | 0.336 | 6.478 | 5.792 | -1.541 | 67.093 | -1.223 | -1.223 | 0.000 | Luxemburg |
| Canada | -1.111 | -3.529 | -3.453 | -0.855 | 0.247 | 7.204 | 6.956 | -0.475 | 7.982 | $-3.825^{*}$ | $-3.825^{*}$ | 0.143 | Norway |
| France | $-2.901^{*}$ | $-8.935^{*}$ | -8.027 | $-1.996^{*}$ | 0.249 | 3.249 | $3.079^{*}$ | $-2.222^{*}$ | $9.322^{*}$ | 0.109 | 0.109 | 0.002 | Luxemburg |
| Finland | $-3.109^{*}$ | -3.544 | -2.854 | -1.189 | 0.417 | 8.337 | 8.567 | -1.317 | $5.284^{*}$ | $-2.573^{*}$ | $-2.573^{*}$ | 0.100 | Japan |
| Germany | -2.063 | -3.018 | -2.736 | -1.093 | 0.400 | 10.750 | 8.691 | -1.206 | 21.379 | $-1.167^{*}$ | -1.167 | 0.000 | Luxemburg |
| Italy | -2.606 | -5.503 | -4.924 | -1.527 | 0.310 | 6.866 | 5.074 | -1.707 | $5.182^{*}$ | $-2.557^{*}$ | $-2.557^{*}$ | 0.080 | Spain |
| Japan | -1.335 | -1.072 | -0.756 | -0.418 | 0.553 | 25.829 | 18.680 | -0.600 | 48.553 | 0.189 | 0.189 | 0.000 | Austria |
| Luxemburg | $-2.950^{*}$ | -6.690 | -6.177 | -1.754 | 0.284 | 4.053 | 3.976 | -1.900 | 43.304 | 0.444 | 0.444 | 0.001 | Belgium |
| the Netherlands | -1.848 | -3.144 | -2.842 | -1.097 | 0.386 | 11.387 | 8.350 | -1.225 | 24.131 | $-3.425^{*}$ | $-3.425^{*}$ | 0.000 | Switzerlands |
| Norway | -1.753 | -2.166 | -1.784 | -0.846 | 0.474 | 18.617 | 12.373 | -0.941 | 24.792 | $-3.402^{*}$ | $-3.402^{*}$ | 0.002 | Japan |
| Spain | -2.076 | -1.781 | -1.536 | -0.759 | 0.494 | 16.872 | 13.677 | -0.558 | 18.400 | -1.909 | -1.909 | 0.032 | Japan |
| Switzerlands | -1.313 | -1.403 | -1.188 | -0.594 | 0.500 | 18.616 | 15.049 | -0.581 | 40.468 | 2.280 | 2.280 | 0.000 | the Netherlands |
| the United Kingdom | -2.273 | -1.915 | -1.397 | -0.566 | 0.405 | 16.643 | 11.759 | -0.586 | 13.492 | -1.333 | -1.333 | 0.117 | Italy |
| No of rejections | 3.000 | 1.000 | 0.000 | 1.000 | 0.000 | 0.000 | 1.000 | 1.000 | 4.000 | 6.000 | 6.000 |  |  |

Note: Only a constant is included. * indicates rejection at the $5 \%$ level. Bootstrap repetition is 5000 .


[^0]:    ${ }^{1}$ We are grateful to Joon Park, Bill Brown, and David Papell for helpful discussions and comments. Chang gratefully acknowledges the financial support by the National Science Foundation through grant number SES-0453069/0730152.
    ${ }^{2}$ Corresponding author. Address correspondence to Wonho Song, School of Economics, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul, South Korea, or to whsong@cau.ac.kr.

[^1]:    ${ }^{3}$ Table 1 in Hansen (1995) provides asymptotic critical values for the CADF $t$-statistic for values of the nuisance parameter in steps of 0.1 via simulations. For intermediate values of the nuisance parameter, critical values are selected by interpolation.

[^2]:    ${ }^{4}$ We start with the simple model without the deterministic components to effectively deliver the essence of the theory. The models with the deterministic components will be considered at the end of this section.

[^3]:    ${ }^{5}$ Noting that the null limit distribution of the CADF $t$-test depends only on the correlation coefficient $\rho^{2}$, Hansen (1995, Table 1, p.1155) provides the asymptotic critical values for the CADF $t$-test for values of $\rho^{2}$ from 0.1 to 1 in steps of 0.1 . The estimate for $\rho^{2}$ is constructed as $\hat{\rho}^{2}=\hat{\sigma}_{v \varepsilon}^{2} / \hat{\sigma}_{v}^{2} \hat{\sigma}_{\varepsilon}^{2}$, where $v_{t}=\beta(L)^{\prime} w_{t}+\varepsilon_{t}$, and $\hat{\sigma}_{v \varepsilon}, \hat{\sigma}_{v}^{2}$ and $\hat{\sigma}_{\varepsilon}^{2}$ are consistent nonparametric estimators of the corresponding parameters.

[^4]:    ${ }^{6}$ Alternatively, we may resample $\left(\varepsilon_{t}^{*}\right)$ and $\left(\eta_{t}^{*}\right)$ separately from the $\left(\tilde{\varepsilon}_{t}\right)$ and $\left(\tilde{\eta}_{t}\right)$ for $t=1, \ldots, n$. In this case, however, preserving the original correlation structure needs more care. We basically need to prewhiten $\left(\tilde{\varepsilon}_{t}\right)$ and $\left(\tilde{\eta}_{t}\right)$ before resampling, and then recolor the resamples to recover the correlation structure. More specifically, we first prewhiten $\left(\tilde{\varepsilon}_{t}\right)$ and $\left(\tilde{\eta}_{t}\right)$ by premultiplying $\tilde{\Sigma}^{-1 / 2}$ to $\tilde{\xi}_{t}=\left(\tilde{\varepsilon}_{t}, \tilde{\eta}_{t}^{\prime}\right)^{\prime}$, for $t=1, \ldots, n$. Next, generate $\xi_{t}^{*}=\left(\varepsilon_{t}^{*}, \eta_{t}^{* \prime}\right)^{\prime}$ by resampling from the prewhitened $\left(\tilde{\varepsilon}_{t}\right)$ and $\left(\tilde{\eta}_{t}\right)$ and subsequently recoloring them by premultiplying $\tilde{\Sigma}^{1 / 2}$ to restore the original dependence structure.
    ${ }^{7}$ We may use the first $\ell$-values of $\left(w_{t}\right)$ as the initial values of $\left(w_{t}^{*}\right)$. The bootstrap samples $\left(w_{t}^{*}\right)$ generated as such may not be stationary processes. Alternatively, we may generate a larger number, say $n+M$, of $\left(w_{t}^{*}\right)$ and discard first $M$-values of $\left(w_{t}^{*}\right)$. This will ensure that $\left(w_{t}^{*}\right)$ become more stationary. In this case the initialization becomes unimportant, and we may therefore simply choose zeros for the initial values.

[^5]:    ${ }^{8}$ In the simulation and empirical applications, $\hat{\sigma}_{n}^{2 *}$ is used to compute $T_{n}^{*}$ instead of $\tilde{\sigma}_{n}^{2}$. This is because Chang and Park (2003, p.392) show in the simulations that $\hat{\sigma}_{n}^{2 *}$ performs slightly better than $\tilde{\sigma}_{n}^{2}$.

[^6]:    ${ }^{9}$ Here we use the simple terms "size" and "power" to mean "Type I error" and "rejection probability under the alternative hypothesis", respectively.

[^7]:    ${ }^{10}$ Sample estimates of $\rho^{2}$ are calculated using the Parzen kernel and Andrews' (1991) automatic bandwidth.

[^8]:    ${ }^{11}$ We thank Elliott and Jansson, and Ng and Perron for sharing their codes with us.

[^9]:    ${ }^{12}$ The real exchange rate is calculated as $r_{i t}=\log \left(e_{i t} p_{i t}^{*} / p_{i t}\right)$, where $e_{i t}, p_{i t}^{*}$, and $p_{i t}$ denote respectively nominal spot exchange rate for the $i$-th country, the CPI for the U.S. and the CPI for the $i$-th country.
    ${ }^{13}$ Stock and Watson (1999) note that current theoretical literatures in macroeconomics provide neither intuition nor guidance on which covariates are candidates for our CADF and bootstrap CADF tests other than on the basis of stationarity.
    ${ }^{14}$ The Bayesian Information Criterion (BIC) gives almost similar results.

[^10]:    ${ }^{15} \hat{\rho}^{2}$ is low when $\phi$ is positive according to the simulation results. When we calculate the estimates $\phi$ of $\mathrm{AR}(1)$ lags of potential covariates, they all take large positive numbers.

[^11]:    Note: Only a constant is included.

[^12]:    Note: Only a constant is included.

[^13]:    Note: A constant and a time trend are included.

[^14]:    Note: A constant and a time trend are included.

[^15]:    Note: A constant and a time trend are included.

[^16]:    Note: A constant and a time trend are included.

[^17]:    Note: A constant and a time trend are included.

