Covariate Unit Root Tests with Good Size and Power

Sebastian Fossati*
University of Alberta

This version: May 20, 2011

Abstract

The selection of the truncation lag for covariate unit root tests is analyzed using Monte Carlo simulation. It is shown that standard information criteria such as the BIC or the AIC can result in tests with large size distortions. Modified information criteria can be used to construct tests with good size and power. An empirical illustration is provided.

Keywords: unit root tests, truncation lag, information criteria, vector autoregressions.
JEL Codes: C32, C12, C52.

*Contact: Department of Economics, University of Alberta, Edmonton, AB T6G 2H4, Canada. Email: sfossati@ualberta.ca. Web: http://www.ualberta.ca/~sfossati/. I would like to thank Eric Zivot, Charles Nelson, Chang-Jin Kim, Richard Startz, and Drew Creal for helpful advice and support and Bruce Hansen and Michael Jansson for providing their Matlab code. I would also like to acknowledge financial support from the Buechel Memorial Fellowship and Grover and Creta Ensley Fellowship, which funded part of this research while I was a graduate student at the University of Washington.

1 Introduction

Hansen (1995) initiated a new direction in the search for unit root tests with better power considering a multivariate framework and showing that correlated stationary covariates can be used to obtain tests with higher power than univariate tests. He proposes a unit root test that uses the Dickey-Fuller regression equation and incorporates covariates to increase power (henceforth CADF test). Recent contributions to this literature, however, build on the findings in Elliott et al. (1996). In particular, Elliott and Jansson (2003) derive the power envelope for unit root tests with stationary covariates and propose a point-optimal test (henceforth EJ test) with maximal power against a point alternative. In addition, Pesavento (2006) introduces generalized least squares (GLS) detrending to the CADF test (henceforth CADF-GLS test) and shows that this yields near-optimal tests with small size distortions. The implementation of these tests requires the selection of a truncation lag (k). In this paper, I consider the selection of k using information criteria for the covariate unit root tests proposed in Elliott and Jansson (2003) and Pesavento (2006). These tests have greater power than the CADF test and, hence, the test proposed in Hansen (1995) is not considered here. Using Monte Carlo simulations, I compare the performance of several information criteria for a variety of data generating processes (DGPs). The objective is to provide recommendations on how to construct covariate unit root tests with good size and power.

The truncation lag for the EJ test is selected in a first step where a vector autoregression (VAR) is estimated under the null of non-stationarity (Elliott and Jansson, 2003). Previous research on the selection of the truncation lag for VARs has focused on the accuracy of information criteria to detect the true value of k (Lütkepohl, 1985) or fitting impulse-response functions (Ivanov and Kilian, 2005). In the case of the EJ test, simulation results show that using the Bayesian information criterion (BIC) or the Akaike information criterion (AIC) can result in covariate unit root tests with large size distortions. In contrast, a modified Akaike information criterion (MIC) for cointegration tests introduced in Qu and Perron (2007) yields big size improvements and powerful tests.

In the case of the CADF-GLS test, simulation results show that using the truncation lag selected by the BIC, AIC, or the MIC of Ng and Perron (2001) on the CADF-GLS regression equation can yield unit root tests with large size distortions. On the other hand, selecting the truncation lag on an univariate regression using the MIC of Ng and Perron (2001), as suggested in Pesavento (2006), yields tests with small size distortions. A consequence of this strategy, however, is that the lead and lag orders are constrained to be equal causing a small reduction in power when unnecessary leads and/or lags are added to the regression equation. A simple strategy that removes this constraint yields tests with improved power.

This paper contributes to the literature in several ways. First, it shows that practitioners should pay special attention to the selection of the truncation lag when constructing covariate unit root tests. The results for the EJ and CADF-GLS tests are similar to those found in the univariate literature by Schwert (1989), Ng and Perron (1995), Oke and Lyhagen (1999), and Ng and Perron (2001). Specifically, using the BIC or AIC to select the truncation lag can lead to tests with large size distortions. In two recent empirical applications of covariate unit root tests by Elliott and Pesavento (2006) and Amara and Papell (2006) the authors report selecting the truncation lag using the BIC rule. The large number of rejections in those papers (taken as evidence in favor of the purchasing power parity theory) is potentially biased as a consequence of the use of this rule to select the lag order. Another contribution of this paper is to show that covariate unit root tests with good size and power are available when using the MIC of Qu and Perron (2007) to construct the EJ test and the MIC of Ng and Perron (2001) to construct the CADF-GLS test. In particular, in the case of models with a moving average root close to one, the standard BIC and AIC select lag lengths that are too small resulting in tests with important size distortions. On the other hand, the modified information criteria of Ng and Perron (2001) and Qu and Perron (2007) select higher lag orders, which is necessary for unit root tests to have good size (see Ng and Perron, 1995). In addition, this paper proposes a strategy to select the lead and lag orders for the CADF-GLS test that yields tests with better power. Finally, these results should also prove useful for the covariate unit root tests proposed in Fossati (2009) and Galvao (2009).

The paper proceeds as follows. In the next section, the test statistics and selection rules considered are reviewed. In section 3, the finite sample properties of the tests for different selection rules are analyzed using Monte Carlo simulation. In section 4, an empirical application to standard macroeconomic data shows the relevance of these results. Finally, section 5 concludes.

2 Test Statistics and Selection of the Truncation Lag

2.1 The EJ Test

Following Elliott et al. (2005), consider a data generating process for the $(m+1)$-dimensional vector time series $z_t = (y_t, x_t')'$ of the form

$$y_t = \beta_{y,0} + \beta_{y,1} t + u_{y,t}, \tag{1}$$

$$x_t = \beta_{x,0} + \beta_{x,1} t + u_{x,t}, \tag{2}$$

and

$$A(L) \begin{pmatrix} (1-\rho L) u_{y,t} \\ \Delta u_{x,t} \end{pmatrix} = e_t, \tag{3}$$

where y_t is univariate (and potentially non-stationary) and x_t is an integrated process of dimension $(m \times 1)$ with elements not mutually cointegrated. $A(L)$ is a matrix polynomial in the lag operator of order k with first element equal to the identity matrix and $|A(r)| = 0$ has roots outside the unit circle. The vector e_t is assumed to be $i.i.d.(0, \Sigma)$.

The frequency zero correlation between the shocks to Δx_t and the quasi-differences of y_t can be captured through the parameter R^2. Consider the matrix Ω given by

$$\Omega = \begin{pmatrix} \omega_{yy} & \omega_{yx} \\ \omega_{xy} & \Omega_{xx} \end{pmatrix} = A(1)^{-1} \Sigma A(1)^{-1\prime}.$$

Then, $R^2 = \omega_{yy}^{-1} \omega_{yx} \Omega_{xx}^{-1} \omega_{xy}$ with $0 \leq R^2 < 1$.

Elliott et al. (2005) show that this model has several VAR-type representations. Ignoring the deterministic terms, the model in equations (1)-(3) can be written as

$$A(L)\left(\begin{array}{c}(1-\rho L)y_t,\\ X_t\end{array}\right)=e_t, \tag{4}$$

where $X_t = \Delta x_t$. This representation corresponds to the one examined in Elliott and Jansson (2003). I consider four relevant cases for the deterministic part of the model: (1) $\beta_{y,0} = 0$, $\beta_{y,1} = 0$, and $\beta_{x,1} = 0$; (2) $\beta_{y,1} = 0$, and $\beta_{x,1} = 0$; (3) $\beta_{y,1} = 0$; (4) no restrictions. Note that $\beta_{x,0}$ drops from the model when x_t is differenced. The first case then corresponds to a model with no deterministic terms. The second has no constant or trend in Δx_t but a constant in y_t. The third case includes a constant in Δx_t and a constant in y_t. Case 4 imposes no restrictions. These cases correspond to cases 1-4 considered by Elliott and Jansson (2003). Elliott and Jansson (2003) also consider the case where Δx_t has a constant and time trend (case 5) but this case seems unlikely when x_t is an integrated process.

The test for a unit root is a test of the hypothesis that $\rho = 1$ versus the alternative that $\rho < 1$. Elliott and Jansson (2003) propose a feasible point-optimal test with a test statistic of the form

$$\tilde{\Lambda}(1,\bar{\rho}) = T\left[tr\left(\tilde{\Sigma}(1)^{-1}\tilde{\Sigma}(\bar{\rho})\right)-(m+\bar{\rho})\right], \tag{5}$$

where $\tilde{\Sigma}(r)$ are the covariance matrices of the residuals constructed by detrending the data under the null and the alternative and running $\tilde{A}(L)\tilde{u}_t(r) = \tilde{e}_t(r)$ for $r \in \{1, \bar{\rho}\}$ (see Elliott and Jansson, 2003; Elliott et al., 2005). $\bar{\rho}$ is the value of ρ under the alternative and is given by $\bar{\rho} = 1 + \bar{c}/T$. They recommend using the same values of

$\bar{c}$ as in Elliott et al. (1996) as this corresponds to the worst case scenario, i.e. the covariates have no useful information. As a consequence, $\bar{c} = -7$ for cases 1-3 and $\bar{c} = -13.5$ for case 4. The test rejects for small values of $\tilde{\Lambda}(1, \bar{\rho})$ and critical values reported in Elliott and Jansson (2003) depend on the value of R^2.

A critical step in the construction of unit root tests is the selection of the truncation lag when this is not known. Elliott and Jansson (2003) recommend choosing the number of lags using the BIC on a VAR estimated under the null of $\rho = 1$. The first step then implies running

$$A(L)\Delta z_t = d_t + e_t, \qquad t = k+2, ..., T, \tag{6}$$

where $\Delta z_t = [(1-L)y_t, (1-L)x_t']'$ for $t > 1$ and $\Delta z_1 = [y_1, 0]'$, and d_t includes the deterministic terms. Given that x_t is in first difference, the first observation is treated as in Elliott et al. (2005) and set equal to zero. This VAR is also used to estimate R^2. Let $\hat{\Sigma} = T^{-1}\sum_{t=k+2}^{T} \hat{e}_t\hat{e}_t'$ and $\hat{\Omega} = \hat{A}(1)^{-1}\hat{\Sigma}\hat{A}(1)^{-1\prime}$, then, $\hat{R}^2 = \hat{\omega}_{yy}^{-1}\hat{\omega}_{yx}\hat{\Omega}_{xx}^{-1}\hat{\omega}_{xy}$.

Following Ng and Perron (2005), I estimate each candidate model with the same number of effective observations N and determine the value of k as $k = arg\ min_{k\in[0,k_{max}]}$ $IC(k)$. Selection criteria for VARs have the form

$$IC(k) = ln|\hat{\Sigma}_k| + \frac{C_T}{N}\left(\tau_{Tk}(r) + kn^2\right), \tag{7}$$

where $\hat{\Sigma}_k = N^{-1}\sum_{t=k_{max}+2}^{T} \hat{e}_t\hat{e}_t'$, $N = T - k_{max} - 1$, and $n = m+1$ is the dimension of the VAR in equation (6). As in Elliott et al. (2005) the first observation is dropped in this step only. If $C_T = lnN$ and $\tau_{Tk}(r) = 0$ we have the BIC, and if $C_T = 2$ and $\tau_{Tk}(r) = 0$ we have the AIC.

Information criteria such as the BIC and AIC, however, tend to select overly parsimonious models possibly leading to tests with large size distortions in finite samples. A recent information criterion introduced in Qu and Perron (2007) extends the MIC of Ng and Perron (2001) for cointegration tests and yields tests with improved size properties. Following Elliott et al. (2005), the model in equations (1)-(3) can also be written as

$$\Delta z_t = \Pi z_{t-1} + \sum_{j=1}^{k} \Gamma_j \Delta z_{t-j} + e_t, \qquad t = k+2, ..., T, \tag{8}$$

where deterministic components may be included depending on the underlying process. This is precisely the VAR that needs to be estimated to compute the MIC proposed in Qu and Perron (2007). Consider the problem of testing for r against more than r cointegrating vectors, we can construct the likelihood ratio (LR) test statistic $\tau_{Tk}(r) = -N \sum_{j=r+1}^{n} log(1-\hat{\lambda}_j)$ where $\hat{\lambda}_j$ are the estimated eigenvalues of matrix Π in equation (8). To obtain an Akaike-type MIC for the EJ test I construct the cointegration LR test for zero cointegrating relations, i.e. $r = 0$. Then, if $C_T = 2$ and $\tau_{Tk}(r) = \tau_{Tk}(0)$ we have the MIC proposed in Qu and Perron (2007). A modified BIC can be constructed in a similar way. Note that, under the null, the rank of Π is zero and equation (8) reduces to equation (6). The term $\tau_{Tk}(r)$ provides only a finite sample adjustment.

2.2 The CADF-GLS Test

Pesavento (2006) proposes the CADF-GLS test as an extension of the test proposed in Hansen (1995). The regression equation takes the form

$$\Delta y_t^d = \phi y_{t-1}^d + \sum_{j=-k}^{k} \pi'_{xj} \Delta x_{t-j}^d + \sum_{j=1}^{k} \pi_{yj} \Delta y_{t-j}^d + \epsilon_t, \qquad t = k+2, ..., T-k, \tag{9}$$

where $y_t^d = y_t - d_t' \hat{\beta}^{GLS}$ with $d_t = 0$ for case 1, $d_t = 1$ for cases 2 and 3, and $d_t = (1, t)$

for case 4. $\hat{\beta}^{GLS}$ is the GLS estimate of β obtained from the least squares (OLS) regression of $y_t(\bar{\rho})$ on $d_t(\bar{\rho})$, where the quasi-differenced series are $y_t(\bar{\rho}) = (1 - \bar{\rho}L)y_t$ for $t > 1$ and $y_1(\bar{\rho}) = y_1$. $d_t(\bar{\rho})$ is obtained similarly. Δx_t^d is the OLS demeaned Δx_t where $\Delta x_t = (1 - L)x_t$ for $t > 2$ and $\Delta x_1 = 0$. Note that x_t is in first difference so the first observation is set equal to zero as in Elliott et al. (2005). The CADF-GLS test statistic is the t-statistic on ϕ. Critical values for the unit root hypothesis are reported in Pesavento (2006) and depend on the value of R^2. Hansen (1995) suggests using a nonparametric estimator of the form

$$\hat{R}^2 = 1 - \left(\frac{\hat{\theta}_{21}^2}{\hat{\theta}_{11}\hat{\theta}_{22}} \right), \tag{10}$$

where

$$\hat{\Theta} = \begin{pmatrix} \hat{\theta}_{11} & \hat{\theta}_{12} \\ \hat{\theta}_{21} & \hat{\theta}_{22} \end{pmatrix} = \sum_{i=-M}^{M} w(i/M) \frac{1}{T} \sum_t \hat{\nu}_{t-i}\hat{\nu}_t', \tag{11}$$

with $\hat{\nu}_t = \left(\hat{\epsilon}_t + \sum_{j=-k}^{k} \hat{\pi}_{xj}' x_{t-j}^d, \hat{\epsilon}_t \right)'$, $w(\cdot)$ is a kernel weight function, e.g. the Bartlett or Parzen kernels, and M is a bandwidth. In this paper, all estimations are performed using the Parzen kernel and a bandwidth determined following Andrews (1991).

For the CADF-GLS test, Pesavento (2006) recommends using the truncation lag selected by the MIC of Ng and Perron (2001) on an univariate regression on the GLS detrended y_t. Another option is to apply the selection rules directly to the output of the CADF-GLS regression equation. The information criteria considered are BIC, AIC, and the MIC of Ng and Perron (2001) on equation (9). Selection criteria for the CADF-GLS regression equation have the form

$$IC(k) = ln\hat{\sigma}_k^2 + \frac{C_T}{N} \left(\tau_T(k) + k(2m+1) \right), \tag{12}$$

where $\hat{\sigma}_k^2 = N^{-1} \sum_{t=k_{max}+2}^{T-k_{max}} \hat{\epsilon}_t^2$, and $N = T - 2k_{max} - 1$. Again, the first observation is treated as in Elliott et al. (2005) and dropped in this step. If $C_T = lnN$ and $\tau_T(k) = 0$ we have the BIC, and if $C_T = 2$ and $\tau_T(k) = 0$ we have the AIC. For the MIC we need $C_T = 2$ and the quantity $\tau_T(k) = (\hat{\sigma}_k^2)^{-1} \hat{\phi}^2 \sum_{t=k_{max}+2}^{T-k_{max}} y_{t-1}^{d2}$.

To avoid confusion, from now on I adopt the following notation. Criteria constructed using the EJ VAR and given by equation (7) are labeled with the subscript e (BIC_e, AIC_e, MIC_e). Criteria constructed using the CADF-GLS regression equation and given by equation (12) are labeled with the subscript c (BIC_c, AIC_c, MIC_c). Finally, criteria constructed using the univariate regression on the GLS detrended y_t as in Ng and Perron (2001) are labeled with the subscript u (BIC_u, AIC_u, MIC_u).

3 Finite Sample Simulations

3.1 Monte Carlo Design

In order to accommodate a wide variety of simulation exercises I assume a very general specification. The setup, which is used throughout, is to consider an error process generated by a VARMA(1,1) model of the form

$$(I_2 - AL) \begin{pmatrix} (1-\rho L)u_{yt} \\ (1-L)u_{xt} \end{pmatrix} = (I_2 + BL)e_t, \tag{13}$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix},$$

$e_t \sim iid\ N(0, \Sigma)$ for $t > 0$, $e_t = 0$ for $t \leq 0$, and

$$\Sigma = (I_2 + B)^{-1}(I_2 - A) \begin{pmatrix} 1 & R \\ R & 1 \end{pmatrix} (I_2 - A)'(I_2 + B)^{-1'}. \tag{14}$$

I consider different degrees of correlation with $R \in \{0, 0.3, 0.5, 0.7, 0.9\}$, and two sample sizes $T \in \{100, 250\}$. In all cases, the lower bound k_{min} is zero and the upper bound is selected using $k_{max} = int[12(T/100)^{1/4}]$. The results are based on 5,000 replications. For each sample an array of unit root test statistics are constructed. These include the ADF-GLS test and the point-optimal (P_T) test of Elliott et al. (1996), the CADF-GLS test, and the EJ test.

Results presented in this paper include five reference models. The first model is a vector white noise process (henceforth VWN) where $a_{ij} = b_{ij} = 0$ for all i and j. The second model is a strongly autocorrelated VAR process (henceforth VAR$_1$) where $a_{11} = a_{22} = 0.8$ and $a_{ij} = 0$ otherwise while $b_{ij} = 0$ for all i and j. The third model is an infinite order VAR process (henceforth VMA$_1$) where $a_{ij} = 0$ for all i and j while $b_{11} = b_{22} = -0.8$ and $b_{ij} = 0$ otherwise. The last two models consider the case where y_t Granger-causes x_t. The fourth model is a VAR process (henceforth VAR$_2$) where $a_{21} = 0.5$, $a_{22} = 0.2$, and $a_{ij} = 0$ otherwise while $b_{ij} = 0$ for all i and j. The fifth model is an infinite order VAR process (henceforth VMA$_2$) where $a_{ij} = 0$ for all i and j while $b_{21} = 0.5$, $b_{22} = 0.2$, and $b_{ij} = 0$ otherwise. Five other models (in terms of the values of a_{ij} and b_{ij}) were considered. Results are qualitatively similar and, hence, not reported.

3.2 Size and Power

I first examine the average number of lags selected by each procedure over the 5,000 simulations. The covariates are non-stationary, and I assume this is known, which implies that the tests are constructed with the covariates in first differences. The results for case 3 and $T = 100$ are presented in Table 1. As we can observe, the BIC-type rules always choose the most parsimonious models and results are consistent for different values of R^2. In the case of the VAR models, the AIC- and MIC-type rules choose values of k in the same range and the selected lag orders are in general appropriate given the DGPs considered here. For the VMA models, however, the BIC- and AIC-type rules select models that appear to be too parsimonious given that the true order of the VAR is infinity. The MIC-type rules, on the other hand, choose more lags. Note, however, that the MIC-type rules choose notably different values of k even for the same model. These selection rules are evaluated in terms of the size and power of the tests next.

[TABLE 1 ABOUT HERE]

Table 2 shows the results for the size of the tests for case 3 and $T = 100$. In general, the tests have inflated sizes and the magnitude of the size distortions depends on the model and the selection rule considered. When the truncation lag is chosen using the BIC-type rules the tests can have very large size distortions. Note that, for a nominal size of 5%, the actual size can be as large as 81.8%. The AIC-type rules yield tests with actual size closer to the nominal size but still large distortions appear in the case of the VMA_1 model. As shown in Ng and Perron (2001) and Qu and Perron (2007), in the case of models with a moving average root close to one, standard information criteria select lag lengths that are too small and can result in tests with important size distortions.

[TABLE 2 ABOUT HERE]

In the case of the EJ test, the MIC_e yields substantial size improvements over the BIC_e and AIC_e recommended in Elliott and Jansson (2003). And while the MIC_u yields tests with larger size distortions in the case of the VAR_2 and the VMA_2 models when R^2 is large, in the case of the VMA_1 model, the MIC_u exhibits notable improvements. Finally, in the case of the VAR_1 model, all selection criteria produce tests with large size distortions, particularly large as R^2 increases. In the case of the CADF-GLS test, when the MIC_c is used to select k the test shows some improvements over the BIC_c and AIC_c but still remains largely over-sized in the case of the VMA_1 model. When the MIC_u is used, the CADF-GLS test exhibits very small size distortions even in the case of the VMA_1 model. The CADF-GLS test constructed using the MIC_u only yields tests with large size distortions when the models exhibit Granger causality and R^2 is large.

I evaluate the size-adjusted power of the tests at $\bar{\rho} = 1 + \bar{c}/T$. For case 3 and $T = 100$, $\bar{c} = -7.0$ and $\bar{\rho} = 0.93$. Results reported in Table 3 show that when the covariates carry no useful information, $R^2 = 0$, the power of the tests is similar to that of their univariate counterparts and below the asymptotic 50%. Important power gains are available, however, as R^2 increases. In the case of the EJ test, the MIC_e yields more powerful tests than the MIC_u. And, except for the case of the VAR_1 model, the EJ test is more powerful than the CADF-GLS test. This difference in power is notably large in the case of the models that exhibit Granger causality. In the case of the VMA_1 model, although power gains are only modest, the improvement compared to the univariate tests is still substantial.

[TABLE 3 ABOUT HERE]

The results presented above correspond only to one deterministic case (3), one sample size ($T = 100$), and five of the ten models considered. The other five models show results that are, in general, better than the ones presented above. With respect to the other deterministic cases, the tests exhibit similar results for all cases. Finally, size distortions and power improve significantly for $T = 250$ but the main results still hold: the BIC-type rules can still yield tests with actual size above 50% while the MIC-type rules yield powerful tests with smaller size distortions. In sum, the CADF-GLS test constructed with the truncation lag selected using the MIC_u shows smaller size distortions than the EJ test for most models considered here. Only in the case of the models where y_t Granger-causes x_t I find important size distortions. Although the EJ test has better power, size distortions can be large when the DGP exhibits strong autocorrelation.

3.3 Leads, Lags, and Power Loss in the CADF-GLS Test

The CADF test of Hansen (1995) requires the selection of a lag order for y_t and lead and lag orders for x_t. An extensive search considering all possible combinations of leads and lags, however, is only feasible when the maximum value of k is low. On the other hand, results in the previous section show that restricting the number of lags can imply tests with large size distortions. To avoid this problem, Pesavento (2006) restricts the lead and lag orders to be the same and looks for the optimal k considering a wide range of values and the MIC_u. A consequence of Pesavento's strategy is that unnecessary leads and/or lags will cause a power reduction in the tests. Aiming to construct tests with good size and better power properties, in this section I remove this restriction and allow the lead and lag orders to be different. The unrestricted regression equation for the CADF-GLS test takes the form

$$\Delta y_t^d = \phi y_{t-1}^d + \sum_{j=-k_1}^{k_2} \pi_{xj}' \Delta x_{t-j}^d + \sum_{j=1}^{k_3} \pi_{yj} \Delta y_{t-j}^d + \epsilon_t, \quad t = \tilde{k}+2, ..., T-k_1, \tag{15}$$

where $\tilde{k} = max(k_2, k_3)$ and the rest is defined as in the previous section. Selection criteria in this case have the form

$$IC(k_1, k_2, k_3) = ln\hat{\sigma}_{k_1,k_2,k_3}^2 + \frac{C_T}{N}\left(\tau_T(k_1, k_2, k_3) + (k_1 + k_2)m + k_3\right), \tag{16}$$

where $\hat{\sigma}_{k_1,k_2,k_3}^2 = N^{-1} \sum_{t=\tilde{k}+2}^{T-k_1} \hat{\epsilon}_t^2$, and $N = T - k_1 - \tilde{k} - 1$. If $C_T = lnN$ and $\tau_T(k_1, k_2, k_3) = 0$ we have the BIC, and if $C_T = 2$ and $\tau_T(k_1, k_2, k_3) = 0$ we have the AIC. For the MIC we need $C_T = 2$ and the quantity $\tau_T(k_1, k_2, k_3) = (\hat{\sigma}_{k_1,k_2,k_3}^2)^{-1} \hat{\phi}^2 \sum_{t=\tilde{k}+2}^{T-k_1} y_{t-1}^{d2}$.

Considering all possible combinations of lead and lag orders can result in a large number of candidate models. As a consequence, I consider two variations of this procedure that limit the total number of models to be evaluated. In the first case, the unrestricted case, I generate all possible combinations of lead and lag orders such that $0 \leq k_i \leq k_{max}$ for $i = 1, 2, 3$ where $k_{max} = int[12(T/100)^{1/4}]$ and use an information criterion to select the optimal orders. I will refer to these selection rules as the unrestricted $\overline{BIC}_c$, $\overline{AIC}_c$, and $\overline{MIC}_c$. In the second case, the restricted case, I generate all possible combinations of lead and lag orders such that $0 \leq k_i \leq k^*$ for $i = 1, 2, 3$ with k^* selected using the MIC_u on an univariate regression and use an information criterion to select the optimal orders. I will refer to these selection rules as the restricted $\overline{BIC}_c$, $\overline{AIC}_c$, and $\overline{MIC}_c$. While methods for selecting the best subset regression models as in Gatu et al. (2007), Hofmann et al. (2007), and Gatu et al. (2008) can be a useful alternative, these methods are not considered here.

Table 4 shows the results for the size of the CADF-GLS test for case 3 and $T = 100$. Using the results of the MIC_u in Table 2 as reference we can see that the $\overline{BIC}_c$ and $\overline{AIC}_c$ yield tests with larger size distortions for the two variations considered here. Using the restricted $\overline{MIC}_c$ results in tests with similar size properties and no clear gains result from this modification. On the other hand, using the unrestricted $\overline{MIC}_c$ yields tests with excellent size properties. In this last case, the actual size is almost equal to the nominal 5% in most cases.

[TABLE 4 ABOUT HERE]

In terms of the size-adjusted power, the unrestricted $\overline{MIC}_c$ generates tests with very low power when compared to the other selection rules. Results presented in Table 5 also show that the restricted $\overline{MIC}_c$ yields more powerful tests when compared to the MIC_u and the unrestricted $\overline{MIC}_c$. These results show that size-adjusted power increases by (about) two percent for most models. As a consequence, restricting the lead and lag orders to be equal, as in Pesavento (2006), results in unnecessary leads and/or lags being estimated and causes a power reduction in the tests. This power reduction, however, does not appear to be large.

[TABLE 5 ABOUT HERE]

In sum, while it is possible to construct tests with almost exact size by using the unrestricted $\overline{MIC}_c$, this improvement comes at the expense of very large power reductions. On the other hand, the restricted $\overline{MIC}_c$ only exhibits small improvements over the MIC_u and, as a consequence, forcing the lead and lag orders to be the same does not result in an important power reduction in the tests.

4 Empirical Illustration

To illustrate the relevance of the results presented above, I apply the covariate unit root tests considered here to the United States inflation rate. Some possible covariates suggested by standard macroeconomic theory are: the unemployment rate, output, and the nominal interest rate. Inflation (π) is the annualized quarterly inflation rate as 400 times the logged difference of successive quarters of the GDP deflator. Similarly, output (y) is 400 times the logged real GDP such that Δy is the annualized quarterly growth rate. The unemployment rate (u) is given by the average civilian unemployment rate of the quarter. The interest rate (i) is the average secondary market 3-Month Treasury Bill rate of the quarter. All series are taken from FRED −FED St. Louis− for the period 1959.I − 2010.III. Figure 1 plots the aforementioned series together with their first differences. Table 6 presents univariate unit root tests constructed using the truncation lag determined by the MIC_u as in Ng and Perron (2001). Note that the selected number of lags differs widely for the series considered. The tests fail to reject the unit root hypothesis for inflation, output, and the interest rate. The evidence for the unemployment rate is mixed. All variables appear to be non-stationary or at least highly persistent so, following Hansen (1995), the covariates should be included in first differences.

[FIGURE 1 ABOUT HERE]

[TABLE 6 ABOUT HERE]

Table 7 presents the covariate unit root tests for the inflation rate. Because the tests require that the covariates, in levels, are not cointegrated, the covariates are included one at a time. The tests are constructed for deterministic case 3 as the inflation rate does not exhibit a clear trend and the covariates are included in first differences. First I construct the tests using the truncation lag selected using the BIC_e for the EJ test and the BIC_c for the CADF-GLS test (top panel in Table 7). In the case of the CADF-GLS test, the lead and lag orders are restricted to be the equal as in equation (9). We can see that the BIC-type rules choose very parsimonious models with lead and lag orders in the 1-3 range. The results are consistent for both tests and the unit root hypothesis is not rejected when the change in the unemployment rate and output growth are used as covariates. When the change in the interest rate is used as covariate both tests reject the unit root hypothesis at the 5% level. The change in the interest rate yields the tests with higher value of R^2 and, therefore, is expected to generate the most powerful test. Practitioners unaware of the findings presented in this paper could be inclined to interpret the results in Table 7 as evidence of a stationary inflation rate. Failure to reject the null hypothesis of a unit root in the inflation rate when using other covariates or the univariate tests could be due to low power of the available tests, while the more powerful tests (those with higher R^2) reject the unit root hypothesis. Now, the unit root hypothesis can also be rejected due to large size distortions and this distortions can be specially large when the tests are constructed using BIC-type rules to select the truncation lag. When the rules suggested here (MIC_e for the EJ test and restricted $\overline{MIC}_c$ for the CADF-GLS test) are used to determine the lead and lag orders, the unit root hypothesis is not rejected for all covariates and we can conclude that the inflation rate appears to be I(1) (bottom panel in Table 7). As expected, these rules select models with higher lead and/or lag orders than those selected by the BIC-type rules.

[TABLE 7 ABOUT HERE]

5 Conclusion

This paper considered the selection of the truncation lag using information criteria for the covariate unit root tests proposed in Elliott and Jansson (2003) and Pesavento (2006). The focus was on the construction of tests with good size and power and the results are similar to those found in the univariate literature. In particular, an overly parsimonious model may lead to tests with large size distortions, while an over parameterized model may lead to tests with low power. Information criteria such as the BIC or the AIC tend to select truncation lags that are too small for some of the models considered, leading to tests with large size distortions. The MIC for cointegration tests proposed in Qu and Perron (2007) and the MIC for unit root tests proposed in Ng and Perron (2001), however, yield covariate unit root tests with good size and power. In the case of the CADF-GLS test, the results also showed that forcing the lead and lag orders to be equal, as in Pesavento (2006), can result in a small power reduction.

References

Amara, J., Papell, D.H., 2006. Testing for purchasing power parity using stationary covariates. Applied Financial Economics 16, 29-39.

Andrews, D.W.K., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59(3), 817-858.

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

Elliott, G., Jansson, M., Pesavento, E., 2005. Optimal Power for Testing Potential Cointegrating Vectors With Known Parameters for Nonstationarity. Journal of Business & Economic Statistics 23 (1), 34-48.

Elliott, G., Pesavento, E., 2006. On the failure of purchasing power parity for bilateral exchange rates after 1973. Journal of Money, Credit, and Banking 38 (6), 1405-1430.

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

Fossati, S., 2009. Unit root testing with stationary covariates and a structural break in the trend function. Working Paper.

Galvao, A., 2009. Unit root quantile autoregression testing using covariates. Journal of Econometrics 152, 165-178.

Gatu, C., Yanev, P.I., Kontoghiorghes, E.J., 2007. A graph approach to generate all possible regression submodels. Computational Statistics & Data Analysis 52, 799-815.

Gatu, C., Kontoghiorghes, E.J., Gilli, M., Winker, P., 2008. An efficient branch-and-bound strategy for subset vector autoregressive model selection. Journal of Economic Dynamics & Control 32, 1949-1963.

Hansen, B.E., 1995. Rethinking the univariate approach to unit root testing: using covariates to increase power. Econometric Theory 11, 1148-1172.

Hofmann, M., Gatu, C., Kontoghiorghes, E.J., 2007. Efficient algorithms for computing the best subset regression models for large-scale problems. Computational Statistics & Data Analysis 52, 16-29.

Ivanov, V., Kilian, L., 2005. A practitioner's guide to lag order selection for VAR impulse response analysis. Studies in Nonlinear Dynamics & Econometrics 9(1), Article 2.

Lütkepohl, H., 1985. Comparison of Criteria for Estimating the Order of a Vector Autoregressive Process. Journal of Time Series Analysis 6, 35-52.

Ng, S., Perron, P., 1995. Unit root tests in arma models with data-dependent methods for selection of the truncation lag. Journal of the American Statistical Association 90, 268-281.

Ng, S., Perron, P., 2001. Lag length selection and the construction of unit root tests with good size and power. Econometrica 69 (6), 1519-1554.

Ng, S., Perron, P., 2005. A note on the selection of time series models. Oxford Bulletin of Economics and Statistics 67, 115-134.

Oke, T., Lyhagen, J., 1999. Small-sample properties of some tests for unit root with data-based choice of the degree of augmentation. Computational Statistics & Data Analysis 30, 457-469.

Pesavento, E., 2006. Near-optimal unit root tests with stationary covariates with better finite sample size. Working Paper.

Qu, Z., Perron, P., 2007. A Modified Information Criterion for Cointegration Tests based on a VAR Approximation. Econometric Theory 23, 638-685.

Schwert, G.W., 1989. Tests for unit roots, a Monte Carlo investigation. Journal of Business and Economic Statistics 7, 147-160.

Table 1: Average values of the selected truncation lag for case 3 and $T = 100$.

Model	k	R^2	BIC_u	MIC_u	BIC_e	AIC_e	MIC_e	BIC_c	AIC_c	MIC_c
VWN	0	0.00	0.05	0.84	0.00	0.27	0.53	0.01	0.97	0.74
		0.09	0.06	0.82	0.00	0.26	0.52	0.01	1.01	0.74
		0.25	0.05	0.81	0.00	0.26	0.53	0.01	0.91	0.69
		0.49	0.05	0.79	0.00	0.26	0.54	0.01	0.95	0.70
		0.81	0.06	0.83	0.00	0.25	0.52	0.01	1.04	0.98
VAR_1	1	0.00	1.05	1.82	1.00	1.31	1.63	1.01	2.24	1.85
		0.09	1.05	1.81	1.00	1.32	1.63	1.01	2.26	1.85
		0.25	1.05	1.81	1.00	1.29	1.62	1.01	2.15	1.78
		0.49	1.06	1.79	1.00	1.32	1.63	1.01	2.24	1.83
		0.81	1.05	1.87	1.00	1.31	1.62	1.01	2.28	1.94
VMA_1	∞	0.00	1.83	6.36	1.93	4.62	6.43	0.50	3.24	5.36
		0.09	1.87	6.33	1.92	4.60	6.36	0.57	3.51	5.22
		0.25	1.84	6.30	1.91	4.58	6.45	0.66	3.73	5.12
		0.49	1.83	6.35	1.92	4.63	6.43	0.86	4.06	4.76
		0.81	1.81	6.38	1.92	4.58	6.45	1.18	4.38	3.94
VAR_2	1	0.00	0.05	0.84	0.81	1.29	1.56	0.88	2.23	1.92
		0.09	0.06	0.82	0.94	1.29	1.55	0.97	2.24	1.92
		0.25	0.05	0.81	0.99	1.30	1.58	1.00	2.15	1.84
		0.49	0.05	0.79	1.00	1.30	1.59	1.01	2.11	1.82
		0.81	0.06	0.83	1.00	1.29	1.58	1.30	3.16	2.97
VMA_2	∞	0.00	0.05	0.84	0.78	1.37	1.60	0.87	2.13	1.86
		0.09	0.06	0.82	0.92	1.39	1.62	0.96	2.16	1.87
		0.25	0.05	0.81	0.99	1.40	1.66	1.01	2.22	1.93
		0.49	0.05	0.79	1.01	1.47	1.70	1.09	2.62	2.24
		0.81	0.06	0.83	1.01	1.51	1.70	1.69	3.90	3.64

Note: k is the order of the VAR representation for each VARMA model. For example, the VMA_1 model is an invertible VARMA(0,1) and can be represented as an infinite VAR.

Table 2: Size of the tests for case 3 and $T = 100$.

		P_T	EJ					ADF-GLS	CADF-GLS				
Model	R^2	MIC_u	BIC_e	AIC_e	MIC_e	BIC_u	MIC_u	MIC_u	BIC_c	AIC_c	MIC_c	BIC_u	MIC_u
VWN	0.00	0.038	0.062	0.063	0.059	0.064	0.053	0.057	0.078	0.089	0.067	0.079	0.058
	0.09	-	0.054	0.059	0.054	0.059	0.051	-	0.069	0.080	0.060	0.073	0.055
	0.25	-	0.059	0.060	0.059	0.061	0.057	-	0.079	0.086	0.068	0.080	0.067
	0.49	-	0.055	0.057	0.050	0.056	0.052	-	0.075	0.083	0.064	0.075	0.067
	0.81	-	0.049	0.055	0.049	0.055	0.054	-	0.062	0.069	0.055	0.062	0.061
VAR_1	0.00	0.065	0.070	0.075	0.082	0.072	0.083	0.053	0.084	0.097	0.066	0.086	0.065
	0.09	-	0.069	0.079	0.081	0.070	0.084	-	0.082	0.102	0.066	0.081	0.065
	0.25	-	0.096	0.103	0.104	0.099	0.107	-	0.094	0.108	0.080	0.095	0.076
	0.49	-	0.146	0.155	0.153	0.152	0.157	-	0.087	0.104	0.075	0.087	0.076
	0.81	-	0.316	0.316	0.303	0.318	0.316	-	0.066	0.081	0.058	0.066	0.062
VMA_1	0.00	0.038	0.643	0.294	0.169	0.575	0.123	0.107	0.818	0.481	0.161	0.547	0.087
	0.09	-	0.620	0.278	0.168	0.553	0.121	-	0.794	0.444	0.161	0.525	0.091
	0.25	-	0.575	0.259	0.158	0.527	0.113	-	0.752	0.391	0.146	0.494	0.085
	0.49	-	0.530	0.242	0.146	0.503	0.114	-	0.663	0.313	0.134	0.450	0.083
	0.81	-	0.588	0.296	0.194	0.533	0.146	-	0.368	0.165	0.083	0.307	0.074
VAR_2	0.00	0.038	0.074	0.081	0.069	0.019	0.025	0.057	0.083	0.102	0.070	0.024	0.027
	0.09	-	0.074	0.077	0.069	0.026	0.031	-	0.078	0.093	0.063	0.033	0.032
	0.25	-	0.082	0.083	0.075	0.042	0.042	-	0.084	0.094	0.072	0.054	0.049
	0.49	-	0.072	0.074	0.067	0.072	0.066	-	0.070	0.082	0.060	0.093	0.077
	0.81	-	0.063	0.068	0.057	0.165	0.137	-	0.059	0.079	0.049	0.214	0.170
VMA_2	0.00	0.038	0.070	0.080	0.068	0.027	0.030	0.057	0.097	0.097	0.078	0.034	0.032
	0.09	-	0.065	0.069	0.060	0.037	0.039	-	0.084	0.084	0.064	0.050	0.044
	0.25	-	0.068	0.075	0.068	0.066	0.058	-	0.083	0.083	0.067	0.083	0.070
	0.49	-	0.057	0.064	0.057	0.097	0.082	-	0.070	0.070	0.059	0.123	0.099
	0.81	-	0.055	0.067	0.056	0.172	0.142	-	0.069	0.069	0.051	0.229	0.180

Table 3: Size-adjusted power of the tests for case 3 and $T = 100$.

		P_T	EJ					ADF-GLS	CADF-GLS				
Model	R^2	MIC_u	BIC_e	AIC_e	MIC_e	BIC_u	MIC_u	MIC_u	BIC_c	AIC_c	MIC_c	BIC_u	MIC_u
VWN	0.00	0.327	0.478	0.463	0.423	0.479	0.380	0.394	0.530	0.498	0.448	0.530	0.401
	0.09	-	0.533	0.509	0.472	0.528	0.443	-	0.578	0.543	0.498	0.572	0.456
	0.25	-	0.645	0.626	0.582	0.643	0.559	-	0.689	0.659	0.605	0.685	0.574
	0.49	-	0.792	0.768	0.715	0.789	0.708	-	0.807	0.769	0.716	0.804	0.722
	0.81	-	0.988	0.966	0.902	0.982	0.936	-	0.960	0.915	0.828	0.958	0.902
VAR_1	0.00	0.315	0.295	0.289	0.277	0.298	0.258	0.310	0.389	0.379	0.318	0.387	0.309
	0.09	-	0.315	0.304	0.293	0.318	0.280	-	0.428	0.403	0.349	0.429	0.351
	0.25	-	0.356	0.356	0.345	0.360	0.335	-	0.503	0.480	0.418	0.503	0.432
	0.49	-	0.412	0.411	0.394	0.410	0.395	-	0.617	0.582	0.514	0.616	0.545
	0.81	-	0.405	0.413	0.397	0.410	0.423	-	0.815	0.759	0.652	0.812	0.758
VMA_1	0.00	0.118	0.325	0.347	0.266	0.343	0.216	0.199	0.200	0.296	0.174	0.341	0.153
	0.09	-	0.342	0.375	0.285	0.356	0.236	-	0.219	0.320	0.184	0.349	0.165
	0.25	-	0.384	0.430	0.336	0.386	0.296	-	0.250	0.354	0.228	0.384	0.210
	0.49	-	0.431	0.486	0.415	0.412	0.384	-	0.314	0.404	0.300	0.406	0.275
	0.81	-	0.438	0.605	0.572	0.456	0.586	-	0.540	0.543	0.451	0.539	0.409
VAR_2	0.00	0.327	0.423	0.476	0.432	0.169	0.198	0.394	0.459	0.459	0.377	0.190	0.204
	0.09	-	0.513	0.517	0.475	0.174	0.221	-	0.530	0.500	0.424	0.237	0.245
	0.25	-	0.626	0.602	0.553	0.266	0.309	-	0.630	0.594	0.522	0.357	0.360
	0.49	-	0.766	0.736	0.679	0.537	0.540	-	0.743	0.696	0.617	0.632	0.595
	0.81	-	0.961	0.932	0.871	0.826	0.827	-	0.900	0.822	0.715	0.817	0.796
VMA_2	0.00	0.327	0.418	0.469	0.428	0.219	0.226	0.394	0.519	0.490	0.402	0.263	0.252
	0.09	-	0.485	0.489	0.463	0.289	0.286	-	0.559	0.516	0.450	0.355	0.323
	0.25	-	0.591	0.574	0.537	0.437	0.420	-	0.634	0.585	0.515	0.510	0.460
	0.49	-	0.740	0.720	0.673	0.705	0.653	-	0.719	0.678	0.586	0.759	0.683
	0.81	-	0.969	0.931	0.875	0.854	0.841	-	0.893	0.801	0.686	0.812	0.793

Note: Power evaluated at $\bar{\rho} = 0.93$ ($\bar{\rho} = 1 + \bar{c}/T$ and $\bar{c} = -7.0$).

Table 4: Size of the CADF-GLS test for case 3 and $T = 100$.

Model	R^2	MIC_u	Restricted			Unrestricted		
			$\overline{BIC}_c$	$\overline{AIC}_c$	$\overline{MIC}_c$	$\overline{BIC}_c$	$\overline{AIC}_c$	$\overline{MIC}_c$
VWN	0.00	0.058	0.077	0.075	0.061	0.080	0.101	0.049
	0.09	0.055	0.070	0.066	0.056	0.073	0.096	0.042
	0.25	0.067	0.079	0.075	0.065	0.081	0.105	0.051
	0.49	0.067	0.076	0.076	0.066	0.079	0.101	0.047
	0.81	0.061	0.062	0.064	0.058	0.063	0.087	0.044
VAR_1	0.00	0.065	0.081	0.080	0.061	0.080	0.111	0.042
	0.09	0.065	0.094	0.084	0.065	0.099	0.124	0.049
	0.25	0.076	0.092	0.090	0.074	0.097	0.125	0.052
	0.49	0.076	0.090	0.087	0.076	0.093	0.121	0.051
	0.81	0.062	0.065	0.065	0.057	0.070	0.103	0.049
VMA_1	0.00	0.087	0.578	0.323	0.094	0.612	0.336	0.078
	0.09	0.091	0.560	0.317	0.089	0.597	0.317	0.072
	0.25	0.085	0.536	0.291	0.081	0.563	0.297	0.063
	0.49	0.083	0.473	0.240	0.072	0.500	0.242	0.052
	0.81	0.074	0.255	0.130	0.045	0.273	0.142	0.036
VAR_2	0.00	0.027	0.033	0.034	0.023	0.068	0.112	0.036
	0.09	0.032	0.035	0.038	0.026	0.056	0.105	0.033
	0.25	0.049	0.056	0.057	0.047	0.073	0.100	0.045
	0.49	0.077	0.085	0.084	0.077	0.082	0.103	0.046
	0.81	0.170	0.181	0.176	0.170	0.095	0.099	0.046
VMA_2	0.00	0.032	0.042	0.043	0.030	0.075	0.113	0.035
	0.09	0.044	0.050	0.052	0.040	0.061	0.105	0.036
	0.25	0.070	0.077	0.077	0.067	0.074	0.100	0.048
	0.49	0.099	0.109	0.107	0.100	0.081	0.103	0.043
	0.81	0.180	0.186	0.183	0.179	0.078	0.098	0.046

Table 5: Size-adjusted power of the CADF-GLS test for case 3 and $T = 100$.

			Restricted			Unrestricted		
Model	R^2	MIC_u	$\overline{BIC}_c$	$\overline{AIC}_c$	$\overline{MIC}_c$	$\overline{BIC}_c$	$\overline{AIC}_c$	$\overline{MIC}_c$
VWN	0.00	0.401	0.527	0.481	0.414	0.531	0.476	0.315
	0.09	0.456	0.574	0.537	0.474	0.576	0.510	0.372
	0.25	0.574	0.681	0.649	0.599	0.686	0.596	0.442
	0.49	0.722	0.801	0.771	0.738	0.800	0.699	0.535
	0.81	0.902	0.957	0.938	0.921	0.954	0.849	0.669
VAR_1	0.00	0.309	0.402	0.380	0.312	0.407	0.389	0.254
	0.09	0.351	0.450	0.408	0.358	0.456	0.413	0.262
	0.25	0.432	0.518	0.491	0.444	0.527	0.462	0.310
	0.49	0.545	0.619	0.594	0.556	0.621	0.545	0.375
	0.81	0.758	0.823	0.806	0.777	0.816	0.702	0.488
VMA_1	0.00	0.153	0.328	0.325	0.174	0.304	0.302	0.125
	0.09	0.165	0.340	0.322	0.173	0.316	0.315	0.135
	0.25	0.210	0.369	0.346	0.209	0.355	0.331	0.163
	0.49	0.275	0.409	0.391	0.244	0.389	0.369	0.196
	0.81	0.409	0.592	0.539	0.380	0.508	0.490	0.285
VAR_2	0.00	0.204	0.241	0.243	0.190	0.385	0.422	0.231
	0.09	0.245	0.301	0.293	0.252	0.458	0.460	0.293
	0.25	0.360	0.427	0.409	0.371	0.608	0.559	0.387
	0.49	0.595	0.676	0.650	0.612	0.792	0.690	0.512
	0.81	0.796	0.835	0.823	0.803	0.910	0.798	0.585
VMA_2	0.00	0.252	0.300	0.296	0.238	0.429	0.442	0.255
	0.09	0.323	0.384	0.372	0.328	0.477	0.464	0.305
	0.25	0.460	0.532	0.513	0.473	0.616	0.561	0.390
	0.49	0.683	0.760	0.731	0.695	0.784	0.686	0.504
	0.81	0.793	0.838	0.821	0.805	0.925	0.795	0.584

Note: Power evaluated at $\bar{\rho} = 0.93$ ($\bar{\rho} = 1 + \bar{c}/T$ and $\bar{c} = -7.0$).

Table 6: Univariate unit root tests.

	ADF	ADF-GLS	P_T	MP_T	$\hat{\rho}$	Lags	Trend
π	-1.950	-1.571	5.362	5.012	0.957	3	c
u	-1.981	-2.081*	3.021*	2.992*	0.970	12	c
y	-2.253	-1.363	16.239	14.876	0.985	2	c, t
i	-1.429	-1.210	7.295	7.022	0.983	13	c

Note: The number of lags for each specification is chosen using the MIC_u by Ng and Perron (2001). c denotes constant and t denotes time trend. * denotes rejection at 5% level.

Table 7: Covariate unit root tests for Inflation.

		EJ			CADF-GLS		
		Δu	Δy	Δi	Δu	Δy	Δi
BIC_e / BIC_c	Statistic	3.712	3.623	2.875*	-1.446	-1.589	-2.041*
	R^2	0.200	0.049	0.426	0.115	0.042	0.163
	Lags	1	1	3	[2,2,2]	[2,2,2]	[2,2,2]
MIC_e / $\overline{MIC}_c$	Statistic	6.144	4.844	4.082	-1.070	-1.144	-1.522
	R^2	0.009	0.119	0.388	0.149	0.053	0.180
	Lags	9	4	9	[1,1,3]	[0,3,3]	[2,0,3]

Note: The top panel tests are constructed using the BIC_e and BIC_c. The bottom panel tests are constructed using the MIC_e and the restricted $\overline{MIC}_c$. * denotes rejection at 5% level.

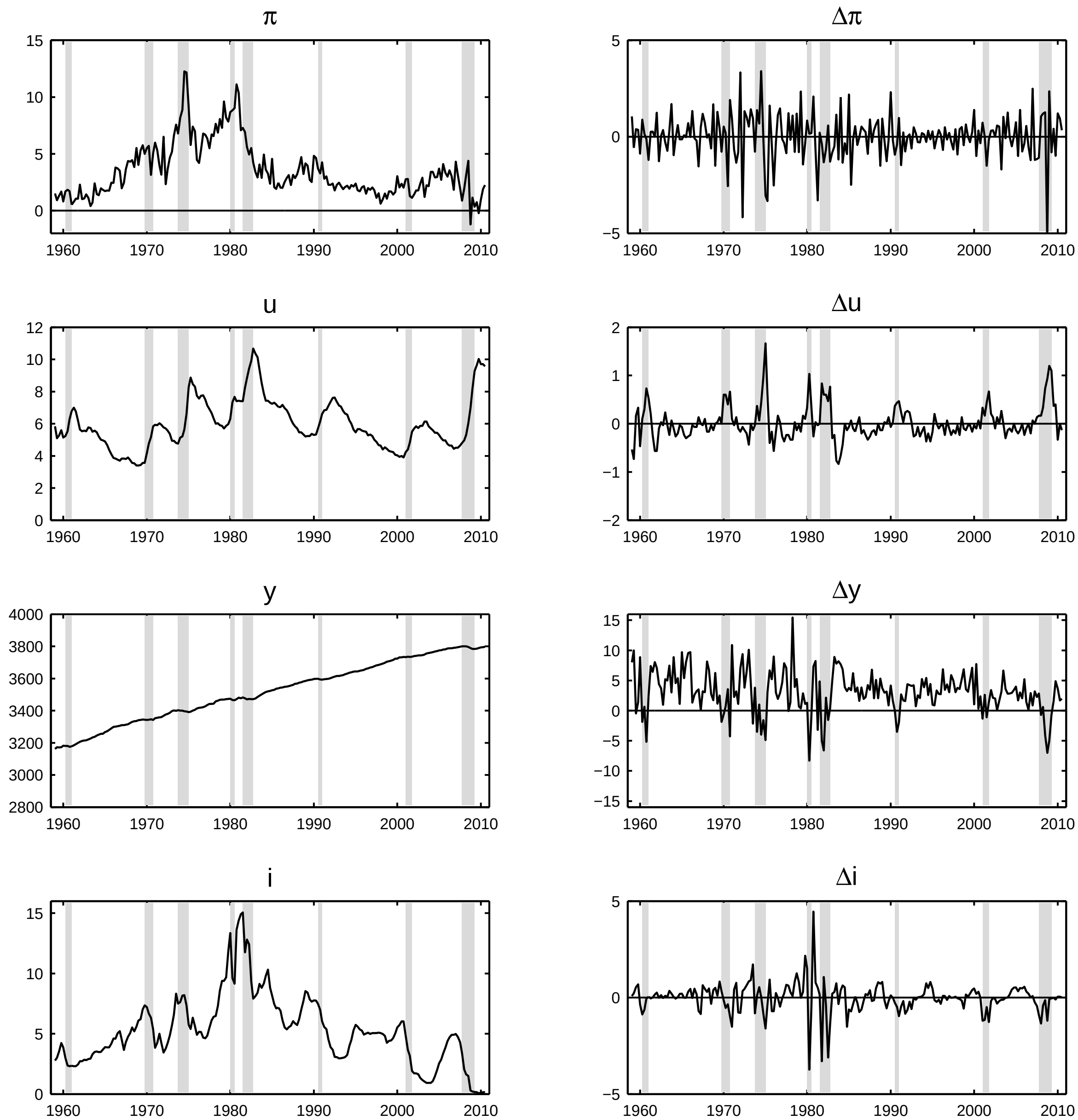

Figure 1: Macroeconomic variables and their first differences. The inflation rate (π), the unemployment rate (u), output (y), and the interest rate (i). FRED IDs: GDPCTPI, UNRATE, GDPC96, and TB3MS. Shaded areas denote NBER recession dates.