COVARIATE-AUGMENTED UNIT ROOT TESTS WITH MIXED-FREQUENCY DATA

BANCO DE PORTUGAL
EUROSYSTEM

Working Papers 2015

Cláudia Duarte

7

7

COVARIATE-AUGMENTED UNIT ROOT TESTS WITH MIXED-FREQUENCY DATA

Working Papers 2015

Cláudia Duarte

June 2015

Please address correspondence to
Banco de Portugal, Economics and Research Department
Av. Almirante Reis 71, 1150-012 Lisboa, Portugal
T +351 213 130 000 | estudos@bportugal.pt

BANCO DE PORTUGAL
EUROSYSTEM

Lisbon, 2015 • www.bportugal.pt

WORKING PAPERS | Lisbon 2015 • Banco de Portugal Av. Almirante Reis, 71 | 1150-012 Lisboa • www.bportugal.pt •
Edition Economics and Research Department • ISBN 978-989-678-351-8 (online) • ISSN 2182-0422 (online)

Covariate-augmented unit root tests with mixed-frequency data

Cláudia Duarte
Banco de Portugal
ISEG-UL

May 2015

Abstract
Unit root tests typically suffer from low power in small samples, which results in not rejecting the null hypothesis as often as they should. This paper tries to tackle this issue by assessing whether it is possible to improve the power performance of covariate-augmented unit root tests, namely the ADF family of tests, by exploiting mixed-frequency data. We use the mixed data sampling (MIDAS) approach to deal with mixed-frequency data. The results from a Monte Carlo exercise indicate that mixed-frequency tests have better power performance than low-frequency tests. The gains from exploiting mixed-frequency data are greater for near-integrated variables. An empirical illustration using the US unemployment rate is presented.

JEL: C12, C15, C22

Keywords: Unit Root, Hypothesis testing, Mixed-frequency data.

Acknowledgements: I would like to thank Paulo Júlio, José Francisco Maria, Carlos Robalo Marques, João Nicolau and Paulo Rodrigues. The opinions expressed in the article are those of the author and do not necessarily coincide with those of Banco de Portugal or the Eurosystem. Any errors and omissions are the sole responsibility of the author.

E-mail: cfduarte@bportugal.pt

1. Introduction

The importance of unit root testing for modelling and forecasting has been well established since the seminal paper by Granger and Newbold (1974), who showed that applying least squares to non-stationary variables can lead to spurious results. In these circumstances, standard errors are biased, the traditional t-ratio significance test does not have the standard limiting distribution and, hence, the analysis of parameter estimates becomes unreliable (Phillips (1986)).

Many studies have focused on unit root tests and their properties (see, for example, Schwert 1989; Stock 1986; and Haldrup and Jansson 2006 for reviews on the topic). The augmented Dickey-Fuller (ADF) test for zero frequency unit roots is the most commonly used procedure, though not necessarily the one with best power performance. This test looses in terms of power to other tests, both asymptotically and in finite samples, especially in the context of near integrated processes.[1] Lower power means that the null hypothesis, being false, is not rejected as often as it should be, leading to the wrong conclusion that the variable is non-stationary.

To overcome this shortcoming, alternative tests exploiting information from covariates have been proposed. In particular, Hansen (1995) generalised the ADF test to include covariates - the CADF test. The intuition is that including a weakly exogenous and stationary variable in the auxiliary test regression may lead to efficiency gains. The performance of covariate-augmented unit root tests depends crucially on the relationship between the variable of interest (dependent variable) and the covariates. The higher the correlation between the variables, the greater the potential power gains. In practice, exploiting these correlations may entail some challenges, especially when the variables involved are sampled at different/mixed frequencies. The typical approach of temporally aggregating high-frequency variables to the same (low) frequency as the variable of interest (e.g., by skip-sampling or computing simple averages) can result in information losses (see Silvestrini and Veredas 2008 for a survey on temporal aggregation and its implications).

This article puts forward a new class of CADF tests that is able to deal with mixed-frequency data.[2] We assess the impact of this extension on size and power performances. In particular, the MI(xed) DA(ta) S(ampling) framework is used to deal with mixed-frequency data. Inspired in the distributed lag

1. The ADF test is a tougher competitor in terms of size. Nevertheless, Perron and Ng (1996) showed that the M-tests originally suggested by Stock (1999) have lower size distortions compared to other unit root tests that are available in the literature. However, for the tests to have good size properties it is essential that an autoregressive spectral density estimator is used as to consistently estimate the long run variance.

2. This article focuses on unit root tests with non-stationarity under the null hypothesis. See Jansson (2004) for a unit root test with covariates where the null hypothesis is stationarity.

models, MIDAS weighting schemes are very flexible, can be quite parsimonious and are able to account for different frequencies (for a brief overview of the main topics related with MIDAS regressions see, for example, Andreou *et al.* 2011). To the best of our knowledge, this is the first application of MIDAS to covariate-augmented unit root testing.

This new mixed-frequency test framework is applied to the well-known CADF test proposed by Hansen (1995) and also to the more recent test proposed by Pesavento (2006), which is a modified version of the CADF test, similar to the GLS generalisation of the ADF test in Elliott *et al.* (1996).

Using a Monte Carlo experiment, we show that mixed-frequency covariate-augmented unit root tests have better power performance than traditional low-frequency tests, while the size performance is similar. Moreover, the performance of mixed-frequency tests improves when variables are near-integrated. These results are robust to the size of the sample, to the lag specification of the test regressions and to different combinations of time frequencies.

The remainder of this article is organised as follows. Section 2 summarises the covariate-augmented unit root tests — CADF and CADF-GLS — as they were initially presented, while Section 3 describes the mixed-frequency approach to unit root testing. Section 4 reports a simulation-based study on the power and size implications of this new approach. Section 5 compares the performance of the alternative approaches for testing the presence of a unit root in the US unemployment rate. Finally, Section 6 concludes.

2. Covariate-augmented unit root tests

This section presents two covariate-augmented unit root tests commonly found in the literature: the CADF test proposed by Hansen (1995) and the CADF-GLS test in Pesavento (2006). For the sake of simplicity, the notation was developed for the case of a single covariate but can be readily extended for multiple covariates.

The common analytical framework is as follows. As in Hansen (1995) and Elliott and Jansson (2003), assume that the variable of interest, Y_t, is the sum of a deterministic component, $d_{Y,t}$, and a stochastic component, $u_{Y,t}$, such as

$$Y_t = d_{Y,t} + u_{Y,t}, \tag{1}$$

where the deterministic component equals $d_{Y,t} = 0$, $d_{Y,t} = \beta_{Y,0}$, or $d_{Y,t} = \beta_{Y,0} + \beta_{Y,1}t$, with t denoting a linear trend. Similarly, the stationary covariate series, X_t, can be expressed as

$$X_t = d_{X,t} + u_{X,t}. \tag{2}$$

Hence, consider the VAR model formulation

$$\begin{bmatrix} Y_t \\ X_t \end{bmatrix} = d_t + u_t, \tag{3}$$

where $d_t = z_t'\beta$, $\beta = [\beta_{Y,0} \;\; \beta_{X,0} \;\; \beta_{Y,1} \;\; \beta_{X,1}]'$,

$$z_t' = \begin{bmatrix} 1 & 0 & t & 0 \\ 0 & 1 & 0 & t \end{bmatrix}, \text{ and } u_t = \begin{bmatrix} u_{Y,t} \\ u_{X,t} \end{bmatrix}.$$

Five different combinations of the deterministic variables can be considered:

Case 1: $\beta_{Y,0} = \beta_{Y,1} = \beta_{X,0} = \beta_{X,1} = 0$;
Case 2: $\beta_{Y,1} = \beta_{X,0} = \beta_{X,1} = 0$;
Case 3: $\beta_{Y,1} = \beta_{X,1} = 0$;
Case 4: $\beta_{X,1} = 0$;
Case 5: No restrictions.

In addition, the stochastic component u_t is expressed as

$$A(L) \begin{bmatrix} (1-\alpha L)u_{Y,t} \\ u_{X,t} \end{bmatrix} = \begin{bmatrix} e_{Y,t} \\ e_{X,t} \end{bmatrix} \tag{4}$$

or,

$$\begin{bmatrix} u_{Y,t} \\ u_{X,t} \end{bmatrix} = \begin{bmatrix} (1-\alpha L) & 0 \\ 0 & 1 \end{bmatrix}^{-1} A^{-1}(L) \begin{bmatrix} e_{Y,t} \\ e_{X,t} \end{bmatrix}$$

where $A(L)$ is a matrix polynomial of order k in the lag operator L. In the following analysis we assume that:

Assumption 1: The roots of $A(L)$ lie outside the unit circle;
Assumption 2: $u_0, u_{-1}, ..., u_{-k}$ are $O_p(1)$;
Assumption 3: $E_{t-1}(e_t) = 0, E_{t-1}(e_t e_t') = \Sigma,$ where Σ is positive definite, and $sup_t E\|e_t\|^{2+\kappa} < \infty,$ for some $\kappa > 0$,

where E_{t-1} denotes the conditional expectation with respect to $e_{t-1}, e_{t-2}, ...,$ and Σ can be expressed as

$$\Sigma = \begin{bmatrix} \sigma_{YY} & \sigma_{YX} \\ \sigma_{YX} & \sigma_{XX} \end{bmatrix}. \tag{5}$$

Assumption 1 is a standard stationarity condition. Assumption 2 implies that the initial values are asymptotically negligible and Assumption 3 implies that e_t satisfies a functional central limit theorem (Phillips 1987). Additionally, let $v_t = [(1-\alpha L)u_{y,t} \;\; u_{x,t}]'$. Note that $v_t = A(L)^{-1}e_t$, with autocovariance function denoted by $\Gamma(k) = E(v_t v_{t+k}')$. Its spectral density at frequency zero (scaled by 2π), denoted by $\Omega = A(1)^{-1}\Sigma A'(1)^{-1}$, is assumed to be bounded away from zero and can be decomposed as

$$\Omega = \begin{bmatrix} \omega_{YY} & \omega_{YX} \\ \omega_{YX} & \omega_{XX} \end{bmatrix}. \quad (6)$$

It is further assumed that

Assumption 4: The autocovariance function of v_t, $\Gamma(k)$, is absolutely summable, $\sum_{j=-\infty}^{+\infty} \|\Gamma(k)\| < \infty$,

where $\|.\|$ is the standard Euclidean norm.

2.1. The CADF test

The covariate augmented Dickey Fuller (CADF) test aims at combining the generally good size properties of the ADF test with higher power due to the inclusion of covariates. To see how let us start by mixing equations (3) and (4) to obtain

$$\begin{bmatrix} \Delta Y_t \\ X_t \end{bmatrix} = z_t'\beta^* + \delta \begin{bmatrix} Y_{t-1} \\ 0 \end{bmatrix} + \begin{bmatrix} v_{Y,t} \\ v_{X,t} \end{bmatrix} \quad (7)$$

where $\delta = \alpha - 1$ and $\beta^* = [-\delta\beta_{Y,0} + (1+\delta)\beta_{Y,1} \quad \beta_{X,0} \quad -\delta\beta_{Y,1} \quad \beta_{X,1}]'$.

Under Assumptions 1 to 4 it is possible to write (Saikkonen 1991 and Brillinger 2001)

$$v_{Y,t} = \sum_{j=0}^{\infty} b_j^* v_{X,t-j} + \eta_t \quad (8)$$

where η_t is a stationary process with zero mean and spectral density at frequency zero (scaled by 2π) equal to $\omega_{\eta\eta} = \omega_{YY} - \omega_{YX}\omega_{XX}^{-1}\omega_{YX}$.[3] Furthermore,

$$E(v_{X,t}\eta_{t+k}) = 0 \quad (9)$$

for any $|k| = 0, 1, 2, ...$, meaning that the right-hand side variables in (8) are orthogonal to the regression error. Given that b_j^* is absolute summable, then (8) can be approximated by

$$v_{Y,t} = \sum_{j=0}^{k} b_j^* v_{X,t-j} + \eta_t = b^*(L)v_{X,t} + \eta_t = b^*(L)(X_t - \beta_{X,0}^* - \beta_{X,1}^* t) + \eta_t \quad (10)$$

where k is large enough so that $b_j^* \approx 0$ for $j > k$, as in Remark 2.1 of Chang and Park (2002), and $b^*(L)$ is a lag polynomial of order k. Combining (7) and

3. For simplicity, it is assumed that the polynomial $b^*(L)$ only includes lags. This is not necessary; see Hansen (1995) for more details.

(10) yields the regression equation

$$\Delta Y_t = \beta^*_{Y,0} + \beta^*_{Y,1} t + \delta Y_{t-1} + b^*(L)(X_t - \beta^*_{X,0} - \beta^*_{X,1} t) + \eta_t. \qquad (11)$$

This equation resembles the test regression used in Dickey-Fuller (DF) tests, augmented with the covariate X_t. In practice, η_t can be serially correlated. Similarly to ADF tests, (11) can be augmented with lags of the dependent variable, so that the error process is approximately white noise. Letting $\psi(L)\eta_t = \xi_t$, where ξ_t is white noise, the augmented version of (11) can be expressed as

$$\psi(L)\Delta Y_t = \mu_0 + \mu_1 t + \delta^* Y_{t-1} + b(L)(X_t - \beta^*_{X,0} - \beta^*_{X,1} t) + \xi_t. \qquad (12)$$

where $\psi(L)$ is a lag polynomial of order p with all roots lying outside the unit circle, $\mu_0 = \psi(1)\beta^*_{Y,0}$, $\mu_1 = \psi(L)\beta^*_{Y,1}$, $\delta^* = \psi(1)\delta$ and $b(L) = \psi(L)b^*(L)$ is a lag polynomial of order q. Accordingly, $v_{Y,t} = b(L)v_{X,t} + \xi_t$. The long-run covariance matrix between $v_{Y,t}$ and ξ_t is

$$\Phi = \begin{bmatrix} \omega_{YY} & \omega_{Y\xi} \\ \omega_{Y\xi} & \omega_{\xi\xi} \end{bmatrix} \qquad (13)$$

where $\omega_{\xi\xi} = \omega_{YY} - \omega_{YX}\omega^{-1}_{XX}\omega_{YX}$. Define ρ^2 as the long-run zero frequency squared correlation between $v_{Y,t}$ and ξ_t, which can be expressed as

$$\rho^2 = \frac{\omega^2_{Y\xi}}{\omega_{YY}\omega_{\xi\xi}}, \qquad (14)$$

and let Q^2 be the ratio of variances,

$$Q^2 = \frac{\omega_{\xi\xi}}{\omega_{YY}}. \qquad (15)$$

In a well-specified dynamic regression, ξ_t is uncorrelated with X_{t-k} for all k, so $\omega_{Y\xi} = \omega_{\xi\xi}$ and $\rho^2 = Q^2$.

Equation (12) is the covariate-augmented ADF unit root test regression, denoted by CADF(p,q). The null hypothesis for the presence of a zero frequency unit root in Y_t is $\delta^* = 0$, which is tested against the one-sided alternative hypothesis, $\delta^* < 0$. Following Phillips (1987), the asymptotic theory in Hansen (1995) is based on local-to-unity asymptotics, implying that $\delta^* = c/T$, where T represents the sample size and c is a fixed non-centrality parameter. This means that the null hypothesis holds when $c = 0$ and holds locally for $c < 0$ and $T \to \infty$. However, as noted by Hansen (1995), in a fixed sample this representation is merely a reparameterization.

The test regression is estimated by OLS. The test statistic is the t-statistic associated with the estimated coefficient of interest ($\hat{\delta^*}$) and is distributed as,

$$t(\hat{\delta^*}) \Rightarrow (c/Q)\left(\int_0^1 (W_1^c)^2\right)^{1/2} + \rho\frac{\int_0^1 W_1^c dW}{\left(\int_0^1 (W_1^c)^2\right)^{1/2}} + (1-\rho^2)^{1/2}N(0,1) \qquad (16)$$

where W_1^c is an Ornstein-Uhlenbeck process generated by a stochastic differential equation, such as $dW_1^c(r) = cW_1^c + dW(r)$, and W is a standard Brownian motion and the $N(0,1)$ variable is independent of W.[4] Under the null hypothesis, the asymptotic distribution of the t-statistic is a convex linear combination of the Dickey-Fuller (DF) distribution of the univariate unit root tests and the standard Normal

$$t(\hat{\delta^*}) \Rightarrow \rho \frac{\int_0^1 W dW}{\left(\int_0^1 W^2\right)^{1/2}} + (1-\rho^2)^{1/2} N(0,1) \tag{17}$$

where the weights are determined by the nuisance parameter ρ^2. The parameter ρ^2 (or Q^2) can be interpreted as the relative contribution of ξ_t to explain $v_{Y,t}$ at the zero frequency. On the one hand, if $b(L)$ equals zero, $v_{Y,t} = \xi_t$ and $\rho^2 = 1$. In this case the CADF test is equivalent to the typical ADF test. On the other hand, if the importance of ξ_t to explain the zero-frequency movements in $v_{Y,t}$ decreases, then $\rho^2 \to 0$ and the relevant distribution becomes closer to the Normal distribution.[5]

Perhaps more intuitively, one can define a third measure, R^2, such that,

$$R^2 = 1 - Q^2 = \frac{\omega_{YX}\omega_{XX}^{-1}\omega_{YX}}{\omega_{YY}} \tag{18}$$

which accounts for the relative contribution of regressor X_t to explain $v_{Y,t}$ at the zero frequency. If X_t does not contribute at all to explain the variation in $v_{Y,t}$, then $R^2 = 0$. Conversely, if X_t has an increasing contribution to explain $v_{Y,t}$, then $R^2 \to 1$.

Given that the distribution of the test statistic depends on ρ^2, a consistent estimate of this parameter is needed, in order to select the appropriate critical value. According to Hansen (1995), an estimate ($\hat{\rho}^2$) can be obtained in a non-parametric way from

$$\hat{\rho}^2 = \frac{\hat{\omega}_{Y\xi}^2}{\hat{\omega}_{YY}\hat{\omega}_{\xi\xi}} \tag{19}$$

where

$$\hat{\Phi} = \begin{bmatrix} \hat{\omega}_{YY} & \hat{\omega}_{Y\xi} \\ \hat{\omega}_{Y\xi} & \hat{\omega}_{\xi\xi} \end{bmatrix} = \sum_{k=-M}^{M} w(k/M) \frac{1}{T} \sum_{t=1}^{T-m} \hat{\pi}_t \hat{\pi}_{t+k}' \tag{20}$$

and $\hat{\pi}_t = (\hat{v}_{Y,t}, \hat{\xi}_t)'$ are least squares estimates of $\pi_t = (v_{Y,t}, \xi_t)'$ from the appropriate regression model. For example, assuming no intercepts, $\hat{\xi}_t =$

4. The case above presented does not include deterministic variables. When extending to the cases where these variables are included, the structure of (16) remains unchanged except that W_1^c would be appropriately replaced; for more details see Hansen (1995).

5. The case of $\rho^2 = 0$ is excluded, ruling out the situation where the variable of interest is cointegrated with the cumulated stationary covariate (Lupi 2009).

$\hat{\psi}(L)\Delta Y_t - \hat{\delta}^* Y_{t-1} - \hat{b}(L)X_t$ and $\hat{v}_{Y,t} = \hat{b}(L)X_t + \hat{\xi}_t$. The function $w(.)$ is a kernel weight function, such as the Bartlett or Parzen kernels, and M is the bandwidth selected to grow slowly with sample size (Andrews 1991; Jansson 2002). Table 1 in Hansen (1995) presents the relevant asymptotic critical values for a range of ρ^2 values and is made available in Table A.1 of the Appendix.

In theory, the power of unit root tests can be improved by the inclusion of covariates because these contribute to reduce the standard error of the estimate of the autoregressive parameter. Given that δ^* is estimated more precisely, the unit root test for the null hypothesis $H_0 : \delta^* = 0$ will have more power. Greater reductions in the standard error, i.e. higher power, are associated with lower ρ^2 (higher R^2). As analytically shown by Caporale and Pittis (1999), if there is contemporaneous correlation between $v_{Y,t}$ and X_t and Granger causality from $v_{Y,t}$ to X_t, then, in some cases, adding covariates may also lead to an increase in the absolute value of the parameter estimate itself, further enhancing the power of the unit root test.

The discussion above is based on the assumption that the covariates are stationary. The CADF test is no longer valid if X_t is integrated. In case of doubt about the stationarity of the covariates, one should take first differences of these series before proceeding into testing. As discussed in Hansen (1995), this seems to be a sensible approach because over-differencing results in neither significant size distortions nor power loss.

In practice, the presence of correlation between the variable of interest and the covariates, as well as its nature, matter for the performance of the test. In its empirical application, Hansen (1995) concluded that there are important power gains to be obtained from using the CADF test to assess the stationarity of real GNP per capita, industrial production and the unemployment rate for the US. Nevertheless, the CADF test is more prone to size distortions than the ADF test. Caporale and Pittis (1999) performed a similar exercise, analysing a wider set of US macroeconomic series. The authors concluded that the finding of a unit root does not always hold when the more powerful CADF test is used instead of the standard ADF method, although there is evidence of high persistence.

2.2. Tests with GLS demeaning

Recognising that the difficulties with the traditional univariate unit root tests (namely, DF and ADF tests) are associated with inefficient estimates of the deterministic component, Elliott *et al.* (1996) suggested that modifying the estimation of this component could improve their performance. For this purpose, the authors suggested GLS-demeaning/detrending the variable of interest prior to testing for the presence of unit roots (DF-GLS test in Section 2.2.1). Pesavento (2006) proposed a generalisation of the DF-GLS test to include stationary covariates, the so-called CADF-GLS test (Section 2.2.2).

2.2.1. The DF-GLS test. In brief, the DF-GLS test is similar to the ADF-test. The aim of the DF-GLS test is to assess whether $\delta = 0$ (null hypothesis) against the point alternative that $\delta = c/T < 0$. However, in the test regression, instead of using the original Y_t series, the GLS-demeaned/detrended version (Y_t^d) is used. The GLS-demeaned/detrended Y_t^d series is obtained as $Y_t - z_t'\hat{\beta}$ and $\hat{\beta}$ are the coefficient estimates from regressing $Y(\bar{\alpha})$ on $z(\bar{\alpha})$, which are transformed versions of the dependent variable and deterministic variables, respectively. More precisely, $Y(\bar{\alpha}) = (Y_1, Y_2 - \bar{\alpha}Y_1, ..., Y_T - \bar{\alpha}Y_{T-1})$, $z(\bar{\alpha}) = (z_1, z_2 - \bar{\alpha}z_1, ..., z_T - \bar{\alpha}z_{T-1})$ and $\bar{\alpha} = 1 + \bar{c}/T$.

The literature shows that values of $\bar{c}$ associated with an asymptotic power of one half yield tests with power functions tangent to the power envelope at that value, and close to the power envelope over a considerable range of alternative values. The appropriate value of $\bar{c}$ depends on the deterministic specification. Simulation results in Elliott *et al.* (1996) suggest that $\bar{c}$ should equal -13.5 if a trend is included (meaning that when $\bar{c} = -13.5$ the point optimal test is tangent to the power envelope at 0.5 if a constant and trend are estimated), or -7 for constant only (i.e., when $\bar{c}$ equals -7 the point optimal test is tangent to the power envelope at 0.5).

The DF-GLS test statistic is the t-statistic for testing whether $\delta = 0$ in the following regression, without deterministic regressors

$$\Delta Y_t^d = \delta Y_{t-1}^d + a_1 \Delta Y_{t-1}^d + ... + a_p \Delta Y_{t-p}^d + e_t \tag{21}$$

where Δ denotes the first difference, p is the number of lags and e_t is an error term. For constant only, the critical values are those of the conventional DF-tests, when there is no intercept. In the linear trend case, the critical values can be found in Table 1 in Elliott *et al.* (1996).

A related issue concerns the choice of the values assigned to the first observation in the GLS demeaning/detrending procedure. Elliott *et al.* (1996) considered the first observation of the quasi-differenced series as being equal to the first observation in levels (fixed initial observation assumption). Elliott (1999) extended this framework to the case where the initial observation is drawn from its unconditional distribution under the alternative hypothesis. The author concluded that there are differences between the two approaches, but none is the best; the user's choice will depend on the his/her belief as to the correct alternative to be tested. More recently, Westerlund (2015) also assessed the importance of the hypothesis about the first observation in GLS demeaning/detrending. He compared the fixed initial observation assumption (i.e., the first quasi-difference equals the first level) with simply ignoring the first quasi-difference (i.e., equals zero). His results suggest that choosing between these two alternatives matters, and the first observation does not seem to be negligible. Moreover, the first assumption seems to work better.

2.2.2. The CADF-GLS test. Merging Hansen's approach with the GLS demeaning/detrending used in Elliott *et al.* (1996), the CADF-GLS test is constructed by demeaning/detrending each variable (dependent variable and the covariate) according to the assumptions on the deterministic terms, and then estimating a test regression similar to equation (12) but with the demeaned/detrended variables.[6]

Hence, the CADF-GLS test statistic is the t-statistic for testing whether $\delta = 0$ against the point alternative that $\delta = c/T < 0$ in the following regression, without deterministic regressors

$$\Delta Y_t^d = \delta Y_{t-1}^d + \sum_{j=1}^{p} a_j \Delta Y_{t-j}^d + \sum_{j=0}^{q} b_j X_{t-j}^{d*} + e_t \tag{22}$$

where Y_t^d and X_t^{d*} are the demeaned/detrended versions of the original series, p and q are the respective number of lags (chosen by an information criterion, such as the BIC) and e_t is an error term. The Y_t^d series is obtained as described in the previous section. Following Pesavento (2006) and Christopoulos and Léon-Ledesma (2008), given that X_t is assumed to be stationary, X_t^{d*} is obtained by OLS demeaning/detrending the original X_t series, depending on the deterministic component included.

The asymptotic test distribution is

$$t(\hat{\delta})^{GLS} \Rightarrow c\left(\int_0^1 J^2\right)^{1/2} + \frac{\int_0^1 JdW}{\left(\int_0^1 J^2\right)^{1/2}} \tag{23}$$

where J is a Ornstein-Uhlenbeck process, such that

$$J(r) = W(r) + c\int_0^1 e^{(\lambda - s)c} W(s)ds \tag{24}$$

where $W(r) = \sqrt{\frac{R^2}{1-R^2}} W_x(r) + W_y(r)$, $\lambda = (1-c)/(1-c+c^2/3)$, and W_x and W_y are independent standard Brownian motions.[7] In the case of no deterministic variables or only a constant for the dependent variable, (23) is equivalent to (16) and the critical values are those of the CADF test, when there is no intercept. For the other cases, Pesavento (2006) reported asymptotic critical values for a significance level of 5 per cent and different values of R^2 (from 0 to 0.9) that can be found in Table A.2 of the Appendix. The author also refers that she used $\bar{c}$ equal to -7 for cases 1 to 3 and -13.5 for cases 4 and 5, in order to make a reasonable comparison with previous work. The estimate of R^2 is obtained non-parametrically, as in Hansen (1995).

6. Adding to lagged terms, leads can also be included. See Pesavento (2006) for more details.

7. The case above presented does not include deterministic variables. See Pesavento (2006) for extending to the cases where these variables are included.

3. Mixed-frequency covariate-augmented unit root tests

The performance of covariate-augmented unit root tests depends crucially on the relationship between the variable of interest and the covariate used. Economic theory may help in the choice of covariates for unit root testing. However, empirical evidence is not always clear-cut.

In some cases, this may result from ignoring the fact that time series are often analysed at intervals that reflect the timing of data collection. Economic data are sampled at different frequencies. Suppose that the variable of interest and the covariate are described by a high-frequency data generating process (DGP) at a given frequency (e.g., monthly). Moreover, assume that only the covariate is available at that frequency, while the dependent variable is observed at a lower frequency (e.g., annually or quarterly). To deal with this situation the variables are typically temporally aggregated to the same (low) frequency by skip-sampling or averaging.

Regarding the impact of testing for unit roots in the feasible low frequency version of the dependent variable, Granger and Siklos (1995) and Marcellino (1999), among other, showed that zero frequency unit roots are not affected by temporal aggregation. However, temporal aggregation may cause deviations from the test distributions. Assume that a high-frequency variable y_t has the following DGP

$$y_t = \alpha y_{t-1/m} + u_t, \tag{25}$$

where $u_t \sim iid(0, \sigma^2)$ and $E(u_t u_{t-i}) = 0$. Under the null hypothesis of a unit root,

$$\Delta^{1/m} y_t = u_t, \tag{26}$$

where m is the sampling frequency and $\Delta^{1/m}$ is the high-frequency difference operator. For simplicity, assume that $m = 2$. The first difference of the aggregate variable y_t^a resumes to

$$\Delta y_t^a = \begin{bmatrix} \varphi_1 & \varphi_2 \end{bmatrix} \begin{bmatrix} \Delta y_t \\ \Delta y_{t-1/2} \end{bmatrix} = \begin{bmatrix} \varphi_1 & \varphi_2 \end{bmatrix} \left(\begin{bmatrix} u_t \\ u_{t-1/2} \end{bmatrix} + \begin{bmatrix} u_{t-1/2} \\ u_{t-1} \end{bmatrix} \right), \tag{27}$$

where Δ is the low-frequency difference operator and φ_i for $i = 1, ..., m$ represents the aggregation scheme. Now, consider the first-order autocovariance of the aggregated errors under the null hypothesis, denoted as u_t^a. Given that u_t is serially uncorrelated, the autocovariance of the aggregate only contains the product of the terms with the same time subscript. Hence,

$$cov(u_t^a) = \varphi_1 \varphi_2 E(u_{t-1}^2) = \varphi_1 \varphi_2 \sigma^2. \tag{28}$$

As noted by Working (1960), if the aggregation scheme is skip-sampling ($\varphi_1 = 1$ and $\varphi_2 = 0$ for end-of-period sampling or $\varphi_1 = 0$ and $\varphi_2 = 1$ for beginning-of-period sampling) then u_t^a is not serially correlated.

When the aggregation scheme is some kind of averaging (e.g., flat sampling, with $\varphi_1 = \varphi_2 = 1/2$), Working (1960) showed that u_t^a is serially correlated. In this case, the serial correlation affects the limiting distribution for testing unit roots in the aggregate variable. This serial correlation cannot be controlled for by adding lagged terms. Following the same reasoning as in Ghysels and Miller (2013), a first-order moving average (MA) polynomial can be written as

$$\begin{aligned}\begin{bmatrix}\Delta y_t \\ \Delta y_{t-1/2}\end{bmatrix} &= \begin{bmatrix}1 & 1 \\ 0 & 1\end{bmatrix}\begin{bmatrix}u_t \\ u_{t-1/2}\end{bmatrix} + \begin{bmatrix}0 & 0 \\ 1 & -1\end{bmatrix}\begin{bmatrix}1 & 1 \\ 0 & 1\end{bmatrix}\begin{bmatrix}u_{t-1} \\ u_{t-3/2}\end{bmatrix} \\ &= u_t^* + \begin{bmatrix}0 & 0 \\ 1 & -1\end{bmatrix} u_{t-1}^* \\ &= u_t^* + A u_{t-1}^*. \end{aligned} \tag{29}$$

This MA polynomial is not invertible because the matrix A does not fulfil the condition $det(I + Az) \neq 0$ when $|z| \leq 1$. Therefore, this polynomial cannot be well approximated by a high-order autoregressive polynomial.

Regarding the role of the temporally aggregated covariate for enhancing the power of the unit root tests, note that the above-mentioned caveats of temporal aggregation also apply. Focusing on the skip-sampling case (for both the variable of interest and the covariate), the assumption in (8) ensures that in the aggregate equation the regressors are also orthogonal to the regression error.

However, temporal aggregation entails information losses and may reduce the contribution of the covariate to explain the variability of the variable of interest. Assuming that temporal aggregation of the variable of interest is inevitable, this article contributes to the literature by proposing a unit root test of the CADF family that is able to deal with mixed-frequency data. In particular, we assess whether the mixed-frequency approach contributes to the waning of potential distortions in the correlation between the dependent variable and the covariate generated by temporal aggregation. To deal with mixed-frequency data we use the MIDAS framework, which we briefly describe in the next section.

3.1. The MIDAS approach

Introduced by Ghysels *et al.* (2004) and used in, e.g., Ghysels *et al.* (2006) and Ghysels *et al.* (2007), the MIDAS approach provides simple, reduced-form models to approximate more elaborate, though unknown, high-frequency models.

Consider a low-frequency variable Y_t and a high-frequency variable x_t, which has a time frequency m times higher. MIDAS regressions assume that the coefficients associated with the high-frequency variable and its lags are captured by an aggregation lag polynomial $B(L^{1/m})$

$$Y_t = \mu + B(L^{1/m}) x_t^{(m)} + u_t \tag{30}$$

where μ is a constant, $x_t^{(m)}$ is the skip-sampled version of the high-frequency x_t, $B(L^{1/m}) = \sum_{j=0}^{J} B(j)\ L^{j/m}$ is a polynomial of length J in the $L^{1/m}$ lag operator, i.e., $L^{j/m}x_t^{(m)} = x_{t-j/m}^{(m)}$, $B(j)$ is an aggregation weighting scheme, and u_t is a standard *iid* error term. The index j indicates how many high-frequency periods starting from the end of the low-frequency period are taken into account. Note that $B(L^{1/m})x_t^{(m)}$ can be interpreted as a temporally aggregated variable using a more flexible, data-driven weighting scheme compared with commonly used temporal aggregation schemes, such as skip-sampling or averaging. For more details on MIDAS regressions, see Andreou *et al.* (2013), among others.

One crucial assumption about the covariate regards its stationarity. In this case, if x_t is a stationary variable, then $x_t^{(m)}$, its skip-sampled version (or, X_t, the general low-frequency aggregate) also is, as discussed above. Notice that we are assuming that x_t does not display a seasonal behaviour, in order to exclude the possibility that the unit root at the zero frequency can arise because of temporal aggregation of a series which has a unit root at some seasonal frequency (Granger and Siklos 1995).

Regarding the weighting function, there are several possible choices. Ghysels *et al.* (2007) considered two alternatives, both assuming that the weights are determined by a few hyperparameters: the exponential Almon lag and the beta polynomial. Given that these options have nonlinear functional specifications, in both cases MIDAS regressions are estimated using nonlinear least squares.

Alternatively, there is the aggregation scheme underlying the unrestricted MIDAS regressions (U-MIDAS), used in Marcellino and Schumacher (2010), Foroni and Marcellino (2012) and, Foroni *et al.* (2011)

$$Y_t = \mu + B_U(L^{1/m})x_t^{(m)} + u_t, \tag{31}$$

where $B_U(L^{1/m}) = \sum_{j=0}^{J} B_j L^{j/m}$.

Equation (31) can involve a large number of parameters, namely when the difference between the low and the high frequency is large. Hence, large differences in sampling frequencies between the variables are readily penalised in terms of parsimony in U-MIDAS regressions.

This article focuses on a parameterised weighting scheme that can be estimated by OLS, namely the traditional Almon lag polynomial. This aggregation scheme assumes that J lag weights can be related to d linearly estimable underlying parameters, with $d < J$, as follows:

$$B_A(j) = \sum_{i=0}^{d} \theta_i j^i, \quad j = 1, \ldots, J \tag{32}$$

where θ_i, $i = 0, \ldots, d$, denotes the hyperparameters. In the following analysis it is assumed that $d = 2$.[8]

8. This weighting scheme also works in the cases where m is not fixed (e.g., combining monthly with weekly or daily data). In these cases, instead of having one set of weights, we have a different set of weights for each low-frequency period of the sample.

Following the notation in Section 2, the next two sections describe how MIDAS regressions were used to extend covariate-augmented unit root tests to mixed frequency data: Section 3.2 for the mixed-frequency CADF test (M-CADF); and Section 3.3 for the mixed-frequency CADF test with GLS detrending (M-CADF-GLS).

3.2. The M-CADF test

The CADF test can be extended to account for mixed-frequency data as follows

$$\psi(L)\Delta Y_t = d_{Y,t} + \delta Y_{t-1} + B(L^{1/m})(x_t^{(m)} - \beta_{x,0} - \beta_{x,1}t) + \xi_t \tag{33}$$

where $d_{Y,t}$ represents the deterministic component of the dependent variable and ξ_t is white noise. The M-CADF(p, J) test assesses the null hypothesis for the presence of a zero frequency unit root in Y_t ($\delta = 0$), against the alternative hypothesis that Y_t is stationary. Notice that the test regression in (33) simply consists in plugging in high-frequency lags of the covariate in the CADF test regression in (12), instead of the low-frequency lags already included. Considering the Almon MIDAS regression, (33) is estimated by OLS. As in the original CADF test, the test statistic is the t-statistic associated with the estimated $\hat{\delta}$ coefficient and the distribution of the test is as in (17).

3.3. The M-CADF-GLS test

Similarly to (22), the test regression of the M-CADF-GLS(p, J) test is

$$\Delta Y_t^d = \delta Y_{t-1}^d + \sum_{j=1}^{p} a_j \Delta Y_{t-j}^d + B(L^{1/m})x_t^{(m),d*} + e_t \tag{34}$$

where Y_t^d and $x_t^{(m),d*}$ are the demeaned/detrended versions of the original series, as in Section 2.2.2, p is the number of autoregressive lags, $B(L^{1/m})$ is a lag polynomial of order J and e_t is white noise. As in the previous section, the GLS version of the M-CADF test also resumes to replacing the low-frequency lags by the high-frequency lags of the covariate. The M-CADF-GLS(p, j) test statistic is the t-statistic for testing whether $\delta = 0$ against the point alternative that $\delta = c/T < 0$ and the distribution of the test is as in (23). For improving the comparability with the existing literature, the figures for $\bar{c}$ used in this article are -7 for cases 1 to 3 and -13.5 for cases 4 and 5. In addition, the asymptotical critical values are the same as in Pesavento (2006).

4. Monte Carlo simulation

The finite sample size and power performance of the proposed mixed-frequency versions of the covariate-augmented unit root tests is investigated by means of a Monte Carlo simulation exercise. Inspired by Hansen (1995) and Galvão Jr. (2009), this exercise considers the following DGP

$$\begin{aligned} y_t &= d_{y,t} + \alpha y_{t-1} + v_{y,t} \\ x_t &= d_{x,t} + v_{x,t} \end{aligned} \tag{35}$$

where y_t and x_t are both in the same (high) time frequency, $\alpha = 1 + c/T$ and d_t represents the deterministic terms. Four alternatives for c were considered, namely 0, -5, -10 and -15. Regarding the deterministic terms, the leading cases 3 (constant for both variables) and 5 (constant and time trend for both variables) were considered; see Elliott and Jansson (2003) and Juhl and Xiao (2003). The error process $v_t = [v_{y,t}v_{x,t}]'$ is generated by a VARMA model $A(L)v_t = B(L)\xi_t$, where $A(L) = I_2 - AL$, $B(L) = I_2 + BL$,

$$A = \begin{bmatrix} a_1 & a_2 \\ a_2 & a_1 \end{bmatrix}, \qquad B = \begin{bmatrix} b_1 & b_2 \\ b_2 & b_1 \end{bmatrix}, \tag{36}$$

and $\xi_t = [\xi_{y,t}\xi_{x,t}]' \sim N(0, \Sigma)$, where I_2 is a (2×2) identity matrix and Σ is such that the long-run variance matrix of v_t satisfies

$$\Omega = (I_2 - AL)^{-1}(I_2 + BL)\Sigma(I_2 + BL)'(I_2 - AL)^{-1\prime} = \begin{bmatrix} 1 & R \\ R & 1 \end{bmatrix} \tag{37}$$

and, as before, $R^2 \in [0, 1[$. Various values for R^2 are examined, namely $R^2 = \{0.2, 0.5, 0.8\}$. Note that this DGP does not include seasonal features.

Regarding the treatment of the initial condition of the DGP, Hansen (1995) and Galvão Jr. (2009) dropped the first 100 observations to eliminate the start-up effects. However, Müller and Elliott (2003) noted that different initial conditions lead to dramatic changes in the power of unit root tests. In particular, the GLS-family of unit root tests has the best performance when the initial value is near zero. In this article, the initial condition was set at zero.[9]

The low-frequency series Y_t and X_t are obtained by aggregating the generated high-frequency data y_t and x_t. The aggregating scheme considered is skip-sampling (as for stock variables) and $m = 3$, mimicking the combination of quarterly and monthly data. The sample size is set at $T = 100$, and $10,000$ replications are used. In the following sections we present a sensitivity analysis to different m and a larger T, namely $T = 500$.

Notice that the aggregation process affects the value of α tested in the alternative hypothesis. As shown in Pierse and Snell (1995), let y_t be a variable

9. A sensitivity analysis to this assumption was performed and the results indicate that the main result — better power performance of mixed-frequency covariate-augmented unit root tests — does not qualitatively change if a different initial value was considered (e.g., by dropping the initial 100 observations), though the relative performance of the GLS-family of tests is significantly affected by this choice.

generated by the following first-order process

$$y_t = \alpha y_{t-1} + u_t.$$

Then, the m-period aggregated variable denoted as Y_t^a is given by

$$Y_t^a = \alpha^a Y_{t-1}^a + u_t^a,$$

where $\alpha^a = \alpha^m$, whether y_t is a flow or a stock. Hence, the alternative values of α go from 0.95, 0.90 and 0.85, for the high-frequency process, to 0.86, 0.73 and 0.61, respectively, for the low-frequency process. As discussed in Section 3, aggregation also affects the correlation between the variable of interest and the covariate and, thus, affects R^2, though it is much harder to predict the actual impact.

The number of lags is assumed unknown, replicating what happens in practice. The choice of the number of lags is very important for the performance of the test. Choosing a lag order is crucial to find a good enough approximation to the true DGP, which yields unit root tests with size close to the nominal size while retaining acceptable power. The choice of the number of lags is particularly important in the case of (negative) moving average errors.

A common result in the literature is that estimating the number of lags solely by applying the AIC or BIC in a VAR model under the null leads to a very conservative number of lags, which results in noticeable size distortions. An alternative method is the sequential t-test for the significance of the last lag considered, as in Ng and Perron (1995). This procedure has the ability to yield a higher number of lags than the BIC when there are negative moving-average errors and, hence, reduce size distortions. But, the sequential procedure tends to overparameterize in other cases, which also leads to less efficient estimates and subsequently to power losses. Hansen (1995) used ad-hoc rules to choose the number of lags and Pesavento (2006) suggested applying the MAIC approach proposed by Ng and Perron (2001) to an univariate regression (an ADF-type regression) with GLS detrended series for choosing the relevant number of lags for the Y_t variable and, then, using the same number of lags for the covariate.[10]

Fossati (2012) analysed the size and power performance of covariate-augmented unit root tests for different selection procedures of truncation lags. The author showed that the approach in Pesavento (2006) could lead to including too many unnecessary lags and, thus, to size distortions. He suggested applying the MAIC to the output of a CADF test regression and dropped the restriction of using the same number of lags for the Y_t variable and the covariate. Moreover, he considered two alternatives within this unrestricted framework: one with the maximum number of lags given by the rule $int(12(T/100)^{0.25})$, where T is the number of observations (Schwert 1989); and another where the

10. Ng and Perron (2001) also consider a modified version of the BIC, denoted as MBIC, but discounted this alternative due to the superior properties of the MAIC.

maximum number of lags is selected using the procedure in Pesavento (2006). The first has lower size distortions — leading to tests with an almost exact size — but worse power performance than the second, which is not too far off from the results for the restricted version in Pesavento (2006).

In this article we follow the approach suggested by Fossati (2012). However, instead of using the MAIC proposed by Ng and Perron (2001), we use the multivariate version of MAIC in Perron and Qu (2007). The maximum number of lags allowed for the dependent variable was chosen according to the rule in Schwert (1989) and for the covariate we assumed a maximum number of lags equal to 5. This restriction reduces considerably the computation time while not affecting the results. In order to have a meaningful comparison, the same number of lags is used for all tests.[11]

Table 1 shows the range of simulation designs, as well as the number of lags chosen, under the null hypothesis, for each DGP. The values for the initial R^2 refer to the high-frequency processes. Different values were used for the persistence in the high-frequency autoregressive dynamics of the dependent variable, namely $a_1 = \{0.2, 0.5, 0.8\}$, and for the high-frequency moving average dynamics, $b_1 = \{-0.2, -0.5, -0.8\}$.[12] The criterion used to select the number of lags seems to work well, namely in the case of moving average dynamics (DGP 16 to 24), for which the number of lags for the dependent variable is higher than in the case of autoregressive dynamics.

Heteroskedasticity and autocorrelation consistent (HAC) estimates of the elements of the long-run variance-covariance matrix Ω are used. This calculation commonly involves the use of pre-whitening filters based on simple autoregressive models. This procedure may induce bias in the estimation of autoregressive coefficients, which is transmitted to the recolouring filter. To mitigate the potential bias associated with these filters, recursive demeaning/detrending procedures were assessed, as in Taylor (2002), Sul *et al.* (2005) and Rodrigues (2006).

In order to implement the unit root tests we use finite-sample critical values. For $\alpha = 1$, the observed rejection rates of each test were based on critical values from the limiting distribution obtained for the DGP with the simplest dynamics, i.e., DGP 1, 2 and 3 in Table 1. For $\alpha < 1$ the size-adjusted power of the tests was based on critical values estimated from the

11. The number of lags was chosen with the low frequency dataset. For the mixed-frequency tests the same time span of lagged information for the covariate is covered, which corresponds to a different number of lags in the high time frequency.

12. Only negative figures were considered for the moving average dynamics because they represent the most difficult case in terms of size distortions, as the reversion the to the mean is higher. Regarding autoregressive dynamics, other DGP were tested, namely with a higher persistence in the covariate. The results remain qualitatively unchanged. So, for the sake of brevity, those results will not be reported, but are available form the author upon request.

	Simulation design					Median of the lag order							
						$T = 100$				$T = 500$			
	Error structure				Initial	Constant		Trend		Constant		Trend	
	a_1	a_2	b_1	b_2	R^2	Y_t	X_t	Y_t	X_t	Y_t	X_t	Y_t	X_t
1	0	0	0	0	0.20	0	0	1	0	0	0	1	0
2	0	0	0	0	0.50	0	0	1	0	0	0	1	0
3	0	0	0	0	0.80	0	0	1	0	0	0	1	0
4	0.2	0	0	0	0.20	0	0	0	0	1	0	0	0
5	0.2	0	0	0	0.50	0	0	0	0	1	0	0	0
6	0.2	0	0	0	0.80	0	0	0	0	1	1	0	1
7	0.5	0	0	0	0.20	1	0	0	0	2	0	2	0
8	0.5	0	0	0	0.50	1	0	0	0	2	1	2	1
9	0.5	0	0	0	0.80	1	1	0	1	2	1	2	1
10	0.8	0	0	0	0.20	2	0	2	0	2	2	2	2
11	0.8	0	0	0	0.50	2	1	2	1	2	2	2	2
12	0.8	0	0	0	0.80	2	1	2	1	2	2	2	2
13	0.2	0.2	0	0	0.20	0	0	0	0	1	1	0	1
14	0.2	0.2	0	0	0.50	0	1	0	0	1	1	1	1
15	0.2	0.2	0	0	0.80	0	1	0	1	1	1	1	1
16	0	0	-0.2	0	0.20	1	0	1	0	1	0	2	0
17	0	0	-0.2	0	0.50	1	0	1	0	1	1	2	1
18	0	0	-0.2	0	0.80	1	0	1	0	1	1	2	1
19	0	0	-0.5	0	0.20	2	0	3	0	3	1	4	1
20	0	0	-0.5	0	0.50	2	1	3	1	3	1	4	1
21	0	0	-0.5	0	0.80	2	1	3	1	3	1	4	1
22	0	0	-0.8	0	0.20	6	1	6	1	8	1	10	1
23	0	0	-0.8	0	0.50	6	1	6	1	8	2	10	2
24	0	0	-0.8	0	0.80	6	1	6	2	8	2	10	2

TABLE 1. Simulation design and median of the lag order selected by MAIC (Perron and Qu 2007)

simulated data generated under the null hypothesis ($\alpha = 1$).[13] Following the suggestion in Elliott and Jansson (2003), the critical values were interpolated for the estimated figures of R^2.

4.1. Baseline results

Table 2 reports the probability of rejecting the null hypothesis under the unit root case, i.e., the finite sample size for unit root tests considering nominal size 5 per cent.[14]

13. As mentioned by Haug (2002), size-unadjusted power is rather misleading, so those results are not reported.

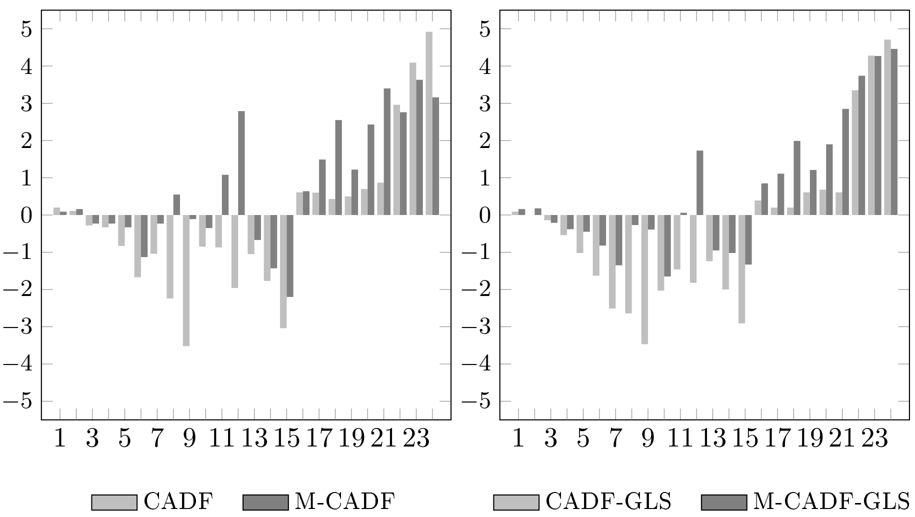
14. The codes were written in Matlab. Some functions were taken from the Econometrics Toolbox by James P. LeSage (http://www.spatial-econometrics.com). The procedure to perform the CADF unit root tests was greatly inspired in the code made available by Bruce E. Hansen (http://www.ssc.wisc.edu/ bhansen/progs/ et_95.html). The MIDAS toolbox was inspired in a code kindly provided by Arthur Sinko.

	Error structure					Constant				Trend			
	a_1	a_2	b_1	b_2	Initial R^2	CADF	CADF GLS	M CADF	M CADF GLS	CADF	CADF GLS	M CADF	M CADF GLS
1	0	0	0	0	0.20	5.2	5.1	5.1	5.2	5.4	5.3	5.3	5.3
2	0	0	0	0	0.50	5.1	5.0	5.2	5.2	4.9	5.1	5.1	5.0
3	0	0	0	0	0.80	4.7	4.9	4.8	4.8	4.6	4.7	4.4	4.9
4	0.2	0	0	0	0.20	4.7	4.5	4.8	4.6	4.2	4.1	4.5	4.3
5	0.2	0	0	0	0.50	4.2	4.0	4.7	4.6	3.5	3.7	4.3	4.0
6	0.2	0	0	0	0.80	3.3	3.4	3.9	4.2	2.7	2.6	3.0	3.6
7	0.5	0	0	0	0.20	4.0	2.5	4.8	3.7	2.4	1.6	3.4	2.7
8	0.5	0	0	0	0.50	2.8	2.4	5.6	4.7	1.5	1.0	4.5	4.2
9	0.5	0	0	0	0.80	1.5	1.5	4.9	4.6	0.5	0.5	4.3	4.3
10	0.8	0	0	0	0.20	4.2	3.0	4.7	3.4	2.6	1.3	3.6	2.1
11	0.8	0	0	0	0.50	4.1	3.5	6.1	5.1	2.5	2.1	6.1	4.8
12	0.8	0	0	0	0.80	3.0	3.2	7.8	6.7	1.4	2.0	7.9	7.6
13	0.2	0.2	0	0	0.20	4.0	3.8	4.3	4.1	3.6	3.3	4.2	4.0
14	0.2	0.2	0	0	0.50	3.2	3.0	3.6	4.0	2.6	2.1	3.8	3.6
15	0.2	0.2	0	0	0.80	2.0	2.1	2.8	3.7	1.1	1.0	2.7	2.7
16	0	0	-0.2	0	0.20	5.6	5.4	5.6	5.9	6.2	6.0	6.3	6.5
17	0	0	-0.2	0	0.50	5.6	5.2	6.5	6.1	6.0	5.9	6.7	7.0
18	0	0	-0.2	0	0.80	5.4	5.2	7.6	7.0	5.6	5.8	8.0	7.9
19	0	0	-0.5	0	0.20	5.5	5.6	6.2	6.2	6.9	6.8	7.2	7.2
20	0	0	-0.5	0	0.50	5.7	5.7	7.4	6.9	7.0	6.8	8.5	8.4
21	0	0	-0.5	0	0.80	5.9	5.6	8.4	7.9	6.9	6.7	9.9	10.1
22	0	0	-0.8	0	0.20	8.0	8.4	7.8	8.7	12.7	13.6	12.5	13.0
23	0	0	-0.8	0	0.50	9.1	9.3	8.6	9.3	14.5	15.2	13.5	14.1
24	0	0	-0.8	0	0.80	9.9	9.7	8.2	9.5	16.2	16.3	13.5	14.6

TABLE 2. Finite sample size for unit root tests considering nominal size of 5 per cent, $T = 100$

Overall, size distortions are larger when there is a linear trend in the regression. The same occurs when b_1 is nonzero, i.e., in the presence of a negative moving average root, which is a result that is commonly found in the unit root literature; see, among others, Schwert (1989). Downward distortions are mainly for stronger and more complex autoregressive dynamics, as also reported in Hansen (1995) and Galvão (2013).

In most cases the size differences between the two sets of tests — mixed- and low-frequency — are not substantial. Figure 1 shows the difference between the finite sample size of the unit root tests and the nominal size of 5 per cent for each DGP.[15] When downward distortions exist, they tend to be less marked for mixed-frequency tests. In the case of strong negative moving average dynamics the upward size distortions are also smaller for mixed-frequency tests than for the low-frequency ones. However, when the moving average parameter is smaller (in absolute terms) the opposite happens.



(A) OLS demeaned tests

(B) GLS demeaned tests

FIGURE 1: Finite sample size distortions vis-à-vis a nominal size of 5 per cent, $T = 100$, Constant only

Tables 3 and 4 report the empirical rejection frequency of the null hypothesis under the alternative, i.e., the power of the unit root tests. Recall that the power is size-adjusted.

15. The figure shows the figures for the case of only a constant term; results are similar if also a time trend is included.

	Error structure					Estimated R^2				$\bar{c}=-5$				$\bar{c}=-10$				$\bar{c}=-15$			
	a_1	a_2	b_1	b_2	Initial R^2	CADF	CADF GLS	M CADF	M CADF GLS	CADF	CADF GLS	M CADF	M CADF GLS	CADF	CADF GLS	M CADF	M CADF GLS	CADF	CADF GLS	M CADF	M CADF GLS
1	0	0	0	0	0.20	0.10	0.10	0.23	0.23	35.6	64.4	44.9	70.3	60.2	75.0	67.2	79.9	67.2	77.0	73.4	81.1
2	0	0	0	0	0.50	0.20	0.20	0.50	0.49	42.7	69.7	71.5	85.0	66.3	79.7	85.8	90.7	73.5	81.9	90.2	91.4
3	0	0	0	0	0.80	0.30	0.31	0.78	0.76	50.5	75.4	95.3	96.7	73.2	84.4	98.8	98.2	80.7	86.4	99.4	97.8
4	0.2	0	0	0	0.20	0.13	0.13	0.24	0.23	39.6	70.4	47.5	73.9	65.0	79.9	70.3	82.7	70.8	81.6	75.2	83.8
5	0.2	0	0	0	0.50	0.27	0.27	0.51	0.50	51.1	78.3	74.7	88.0	73.3	86.3	87.6	92.7	79.3	87.6	91.1	93.2
6	0.2	0	0	0	0.80	0.41	0.42	0.79	0.78	64.6	84.9	96.1	97.5	82.3	91.7	99.1	98.6	87.8	92.7	99.5	98.5
7	0.5	0	0	0	0.20	0.18	0.19	0.25	0.25	38.9	81.2	41.6	78.6	74.5	89.9	74.0	88.8	79.4	90.8	79.7	89.8
8	0.5	0	0	0	0.50	0.41	0.42	0.52	0.52	65.2	90.0	76.6	91.0	86.7	95.5	89.6	95.6	89.8	95.9	92.0	95.9
9	0.5	0	0	0	0.80	0.64	0.64	0.80	0.79	86.0	95.7	97.0	98.6	95.7	98.4	99.1	99.4	97.5	98.6	99.5	99.3
10	0.8	0	0	0	0.20	0.22	0.26	0.28	0.28	17.3	58.8	19.8	57.4	40.6	86.1	37.9	82.8	63.9	93.0	55.7	91.1
11	0.8	0	0	0	0.50	0.45	0.54	0.54	0.54	56.6	83.6	60.0	82.4	80.9	96.1	72.6	92.9	88.2	97.5	76.1	95.4
12	0.8	0	0	0	0.80	0.73	0.80	0.80	0.80	91.4	97.4	91.3	96.4	98.3	99.7	96.3	98.8	99.0	99.8	96.7	98.6
13	0.2	0.2	0	0	0.20	0.12	0.13	0.25	0.24	40.0	73.7	45.4	75.6	66.3	81.6	68.8	83.0	71.4	82.8	73.6	84.4
14	0.2	0.2	0	0	0.50	0.31	0.31	0.52	0.52	56.7	82.9	75.2	88.7	77.6	89.4	86.7	93.4	81.6	90.5	90.2	94.0
15	0.2	0.2	0	0	0.80	0.53	0.53	0.81	0.80	76.7	91.7	96.4	97.9	90.1	96.0	98.9	99.2	93.7	96.5	99.5	99.1
16	0	0	-0.2	0	0.20	0.08	0.08	0.23	0.23	32.5	59.7	42.0	66.3	56.2	71.0	64.2	76.5	64.6	73.7	71.4	77.9
17	0	0	-0.2	0	0.50	0.16	0.15	0.49	0.48	36.3	63.7	66.8	81.6	60.9	75.0	83.1	87.7	69.5	77.5	88.1	88.1
18	0	0	-0.2	0	0.80	0.23	0.22	0.77	0.75	41.0	67.8	93.6	95.2	65.7	78.6	98.3	97.0	75.0	81.1	99.2	96.7
19	0	0	-0.5	0	0.20	0.07	0.07	0.22	0.22	30.8	50.7	38.3	57.1	54.2	62.7	61.3	66.9	63.4	65.9	69.5	69.4
20	0	0	-0.5	0	0.50	0.12	0.11	0.49	0.47	32.3	52.9	61.0	72.3	57.3	65.1	79.9	79.6	67.7	68.6	85.5	80.1
21	0	0	-0.5	0	0.80	0.17	0.15	0.78	0.75	34.3	54.7	88.9	89.4	60.8	67.3	97.0	92.5	72.3	71.0	98.5	91.8
22	0	0	-0.8	0	0.20	0.10	0.09	0.21	0.19	32.2	32.2	37.1	34.1	57.2	49.7	61.7	51.4	67.2	58.3	71.5	59.6
23	0	0	-0.8	0	0.50	0.19	0.17	0.47	0.44	35.2	33.5	51.5	43.0	61.6	51.8	75.5	59.0	72.8	60.8	82.5	65.9
24	0	0	-0.8	0	0.80	0.28	0.24	0.77	0.74	38.0	34.6	78.3	60.4	67.5	53.7	92.2	71.1	80.2	63.9	95.8	75.0

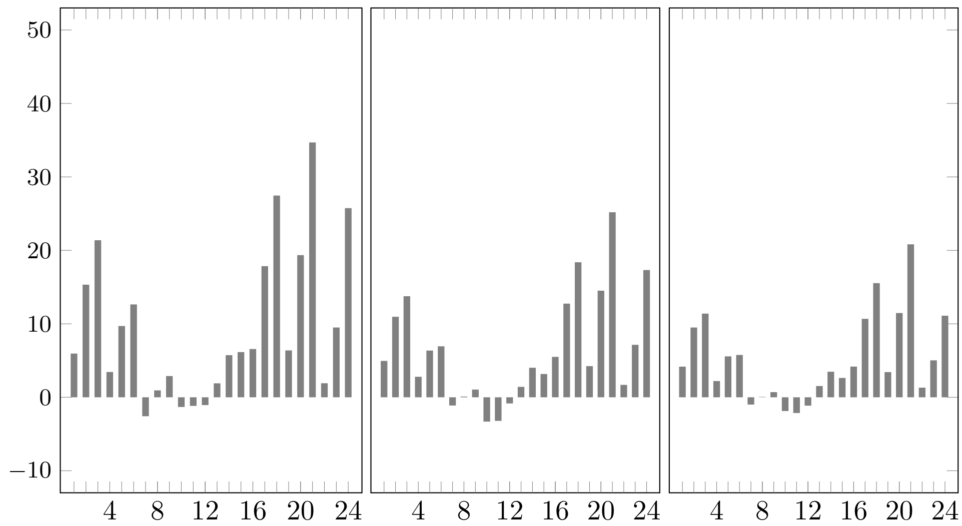
TABLE 3. Size-adjusted power of unit root tests, Constant only, $T = 100$

	Error structure					Estimated R^2				$\bar{c}=-5$				$\bar{c}=-10$				$\bar{c}=-15$			
	a_1	a_2	b_1	b_2	Initial R^2	CADF	CADF GLS	M CADF	M CADF GLS	CADF	CADF GLS	M CADF	M CADF GLS	CADF	CADF GLS	M CADF	M CADF GLS	CADF	CADF GLS	M CADF	M CADF GLS
1	0	0	0	0	0.20	0.09	0.10	0.20	0.23	29.2	45.4	36.9	52.7	59.1	69.4	64.9	74.3	68.1	74.9	72.9	79.0
2	0	0	0	0	0.50	0.19	0.20	0.42	0.49	35.2	51.4	63.3	74.2	65.2	74.3	82.9	88.3	73.4	79.7	87.6	90.9
3	0	0	0	0	0.80	0.28	0.31	0.62	0.76	42.2	58.7	92.0	94.1	71.2	79.6	97.7	98.1	79.3	84.9	98.9	98.8
4	0.2	0	0	0	0.20	0.12	0.13	0.20	0.23	32.6	51.6	39.8	57.3	65.6	75.3	69.6	78.7	72.8	79.0	76.1	82.1
5	0.2	0	0	0	0.50	0.25	0.27	0.43	0.50	43.9	61.7	67.9	78.6	73.5	81.8	85.4	90.8	79.6	85.7	89.2	92.8
6	0.2	0	0	0	0.80	0.38	0.42	0.63	0.78	56.8	72.6	93.5	95.6	81.7	88.4	98.1	98.8	86.6	91.8	99.0	99.3
7	0.5	0	0	0	0.20	0.16	0.19	0.21	0.25	34.6	62.4	34.6	57.3	75.0	87.5	73.1	85.7	83.7	89.4	82.9	88.5
8	0.5	0	0	0	0.50	0.37	0.42	0.42	0.52	59.9	78.5	71.4	81.9	87.3	94.0	89.5	93.9	90.7	95.4	92.4	95.2
9	0.5	0	0	0	0.80	0.59	0.64	0.62	0.79	82.2	91.0	96.0	97.1	95.1	98.2	98.7	99.2	97.1	99.0	99.3	99.5
10	0.8	0	0	0	0.20	0.19	0.27	0.23	0.28	15.8	34.2	16.3	30.7	35.6	70.2	30.5	57.1	57.8	88.3	43.9	77.4
11	0.8	0	0	0	0.50	0.44	0.55	0.39	0.53	54.9	68.4	57.3	67.9	78.6	92.0	67.6	81.2	85.9	96.1	64.6	84.4
12	0.8	0	0	0	0.80	0.72	0.81	0.50	0.79	89.8	93.3	92.2	93.0	98.1	99.3	96.3	97.2	98.8	99.6	94.9	96.1
13	0.2	0.2	0	0	0.20	0.11	0.13	0.21	0.24	32.2	54.7	36.5	56.2	67.5	77.4	69.7	78.2	73.9	80.6	75.3	81.3
14	0.2	0.2	0	0	0.50	0.28	0.31	0.45	0.51	48.3	68.5	67.6	79.2	77.6	85.8	85.2	90.7	81.8	88.3	88.2	92.7
15	0.2	0.2	0	0	0.80	0.49	0.53	0.68	0.79	70.2	83.7	93.9	96.0	89.3	94.9	97.8	98.8	92.2	96.7	98.8	99.3
16	0	0	-0.2	0	0.20	0.08	0.08	0.20	0.22	26.0	40.9	33.6	46.2	53.7	64.5	60.2	69.0	65.0	71.5	69.9	75.3
17	0	0	-0.2	0	0.50	0.15	0.15	0.41	0.48	29.5	44.1	57.2	67.1	58.2	68.2	79.6	84.2	69.2	75.3	85.8	87.8
18	0	0	-0.2	0	0.80	0.21	0.22	0.61	0.75	32.7	47.9	88.7	90.9	62.8	72.2	97.0	97.2	73.7	79.3	98.5	98.0
19	0	0	-0.5	0	0.20	0.07	0.07	0.19	0.22	23.1	34.0	29.8	39.4	49.7	57.6	55.9	62.3	62.3	66.1	66.6	69.5
20	0	0	-0.5	0	0.50	0.12	0.11	0.40	0.47	24.4	35.8	50.7	56.8	52.6	59.9	75.4	77.0	65.6	69.3	82.8	81.5
21	0	0	-0.5	0	0.80	0.16	0.16	0.63	0.75	25.1	37.1	82.4	83.9	55.0	62.6	94.5	93.7	68.8	72.1	97.3	95.2
22	0	0	-0.8	0	0.20	0.09	0.09	0.18	0.19	24.1	24.2	26.9	26.2	51.0	48.5	54.5	50.7	63.7	60.4	66.1	62.0
23	0	0	-0.8	0	0.50	0.18	0.17	0.39	0.44	26.1	25.9	35.6	34.1	54.8	51.2	63.9	58.7	68.2	64.0	74.4	68.5
24	0	0	-0.8	0	0.80	0.27	0.25	0.63	0.74	27.8	27.2	62.1	55.6	59.2	54.3	84.0	75.2	74.2	68.4	89.2	81.0

TABLE 4. Size-adjusted power of unit root tests, Time trend included, $T = 100$

As expected, when a time trend is included in the test regression power is in general lower than when only a constant is present. Moreover, the tests with GLS detrending tend to have a better power performance than the tests with OLS detrending. This is true whether only a constant is considered or a time trend is also included. As reported in Hansen (1995), the power of the tests increases as more correlated covariates are included in the unit root test regressions.

In the vast majority of cases, the differences in size-adjusted power between the mixed- and the low-frequency unit root tests are positive, meaning that the power of mixed-frequency unit root tests is higher than that of low-frequency tests. This is true for both tests with OLS or GLS demeaning. The power differences between the low- and mixed-frequency with GLS demeaning are presented in Figure 2.



(A) $\bar{c} = -5$ (B) $\bar{c} = -10$ (C) $\bar{c} = -15$

FIGURE 2: Differences in size-adjusted power between mixed- and low-frequency tests with GLS demeaning for $c = -5$, -10 and -15, Constant only, $T = 100$

While mixed-frequency regressions tend to capture the original correlation between the dependent variable and the covariate, this correlation is hampered, in most cases, by time aggregation. Hence, the actual (and estimated) R^2 for the low-frequency tests may differ significantly from the initial R^2 of the time disaggregated DGP, as shown in the top panel of Table 5.[16]

16. Table 5 shows the results for the DGP with the simplest dynamics, i.e., DGP 1, 2 and 3 in Table 1. Moreover, it covers the case of including a constant only. Results are qualitative similar if a time trend is also included and are available from the author upon request.

	Initial R^2					
	OLS detrending			GLS detrending		
	0.2	0.5	0.8	0.2	0.5	0.8
Estimated R^2						
Low-frequency tests						
$T = 100$	0.11	0.23	0.34	0.12	0.25	0.38
$T = 500$	0.09	0.20	0.32	0.10	0.23	0.37
Mixed-frequency tests						
$T = 100$	0.23	0.50	0.78	0.23	0.49	0.77
$T = 500$	0.20	0.48	0.77	0.20	0.48	0.77
Estimated δ						
Low-frequency tests						
$T = 100$	-0.046	-0.041	-0.036	-0.019	-0.016	-0.014
$T = 500$	-0.010	-0.009	-0.008	-0.003	-0.003	-0.002
Mixed-frequency tests						
$T = 100$	-0.041	-0.027	-0.013	-0.016	-0.008	0.000
$T = 500$	-0.009	-0.006	-0.003	-0.003	-0.001	0.000
Finite-sample critical values						
Low-frequency tests						
$T = 100$	-2.61	-2.57	-2.52	-1.81	-1.80	-1.77
$T = 500$	-2.67	-2.64	-2.60	-1.80	-1.77	-1.73
memo: Asymptotical critical values						
	-2.81	-2.75	-2.70	-1.91	-1.85	-1.78
Mixed-frequency tests						
$T = 100$	-2.62	-2.58	-2.48	-1.84	-1.75	-1.55
$T = 500$	-2.69	-2.61	-2.49	-1.80	-1.74	-1.70
memo: Asymptotical critical values						
	-2.75	-2.60	-2.32	-1.87	-1.72	-1.52

TABLE 5. Estimated R^2 and δ parameters and critical values, Constant only

Notes: All estimates were obtained from the DGP with the simplest dynamics, i.e., DGP 1, 2 and 3 in Table 1. The estimated R^2 and δ correspond to the average values over the replications. The asymptotical critical values are interpolated for the estimated R^2.

This information loss penalises the power performance of the low-frequency tests. Time aggregation has a milder impact on the correlation between the aggregated data when strong autoregressive dynamics is in place. In this scenario, the performance of the mixed-frequency tests is similar or, in some cases, slightly worse than the performance of the low-frequency tests.

Adding to greater efficiency, the power performance of tests that exploit mixed-frequency data also benefits from less parameter bias (middle panel in Table 5). As expected, the average estimate of the parameter of interest, δ, is closer to zero for tests with GLS demeaning and with higher R^2. In addition, the mixed-frequency estimates are also closer to zero than their low-frequency counterparts.

The bottom panel of Table 5 shows the finite-sample and asymptotic critical values for each test. The asymptotic values were obtained from a simulation exercise with 1,500 observations and 60,000 replications. Separate exercises were performed for the low- and mixed-frequency approaches, delivering similar results. The results are also similar to the values collected from the original papers. The asymptotic critical values presented in the table are interpolated for the values of R^2. Finite-sample critical values converge to their asymptotic values, though at a lower pace in the case of GLS demeaning. For comparable values of R^2 the mixed-frequency critical values are closer to their asymptotic value than the low frequency ones. This result is underpinned by the greater efficiency and less biased estimates of mixed-frequency tests.

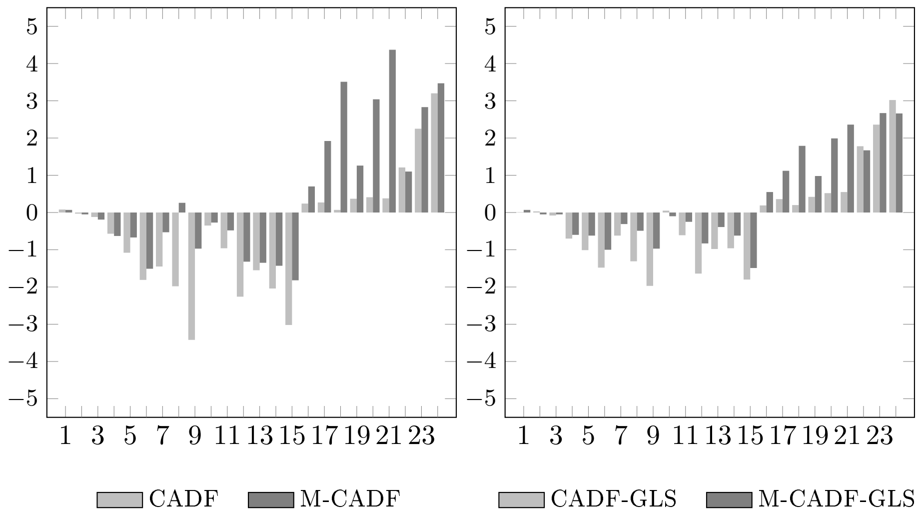
On average, the power gains from taking on board mixed-frequency data are quite substantial, reaching 9.6, 6.4 and 5.2 per cent for $c = -5$, -10 and -15, respectively. Recall that the results above mentioned correspond to $\alpha = 0.95$, 0.90 and 0.85 on the disaggregated process, meaning that time aggregation with $m = 3$ leads to $\alpha = 0.86$, 0.73 and 0.61, respectively. The power gains increase as the alternative hypothesis are more demanding, i.e., are closer to the unit root. This results from losses in power performance, which are smaller for the mixed-frequency tests than for the low-frequency ones.

4.2. The case of a larger sample size

Now consider a sample with 500 observations. Recall that the values of $\bar{c}$, corresponding to $\bar{\alpha} = 0.99$, 0.98 and 0.97, are for the disaggregated processes. Time aggregation with $m = 3$ leads to $\bar{\alpha} = 0.97$, 0.94 and 0.91, respectively. Not only do we consider a larger sample but the alternative hypothesis is more demanding, being closer to the unit root case.

In general, there are less size distortions using this larger sample.[17] Again, as in the case of the sample with 100 observations, when downward distortions exist, they tend to be less marked for mixed-frequency tests (Figure 3). Though upward biased, the performance of the mixed-frequency tests is more favourable

17. The results presented in this section refer to the case of only a constant included in the test regressions. The results are qualitatively similar when also a time trend is considered and are available from the author upon request.



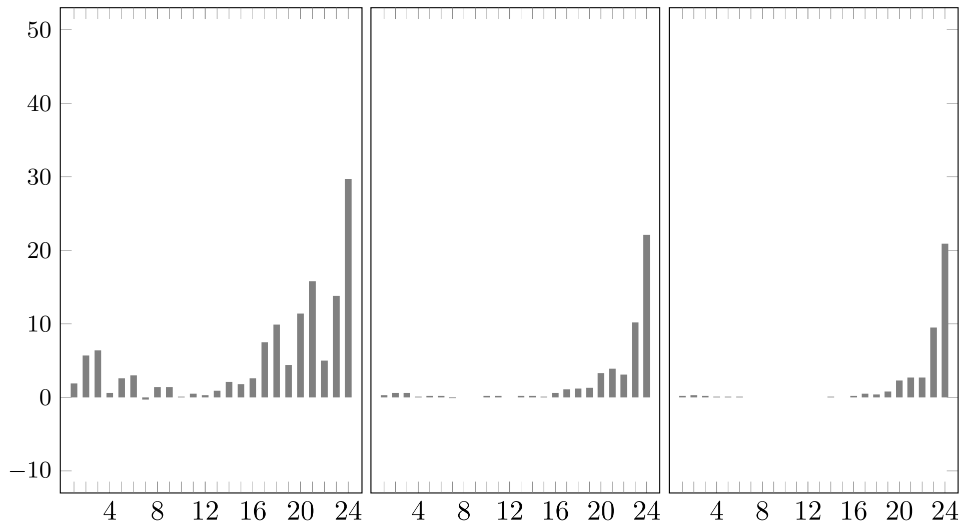
(A) OLS demeaned tests (B) GLS demeaned tests

FIGURE 3: Finite sample size distortions vis-à-vis a nominal size of 5 per cent, $T = 500$

with strong moving average dynamics. This is not the case when the moving average coefficient is smaller (in absolute terms).

Regardless of the DGP, power is always higher in the sample with 500 observations. Moreover, size-adjusted power of mixed-frequency tests is at least as high as the one of low-frequency tests. This is true for both tests with OLS or GLS demeaning. The power differences between the low- and mixed-frequency tests, for the GLS case, are presented in Figure 4. In particular, size-adjusted power of mixed-frequency tests is substantially higher for processes with moving average dynamics.

Mixed-frequency tests also deal better with near-integration. On average, the power gains from exploiting mixed-frequency data increase as c increases, from 1.7 to 2.1 and 5.4 per cent for $c = -15$, -10 and -5, respectively. Hence, as the series becomes near-integrated, the advantage of exploiting mixed-frequency information increases progressively.



(A) $\bar{c} = -5$ (B) $\bar{c} = -10$ (C) $\bar{c} = -15$

FIGURE 4: Differences in size-adjusted power between mixed- and low-frequency tests with GLS demeaning for $c = -5$, -10 and -15, Constant only, $T = 500$

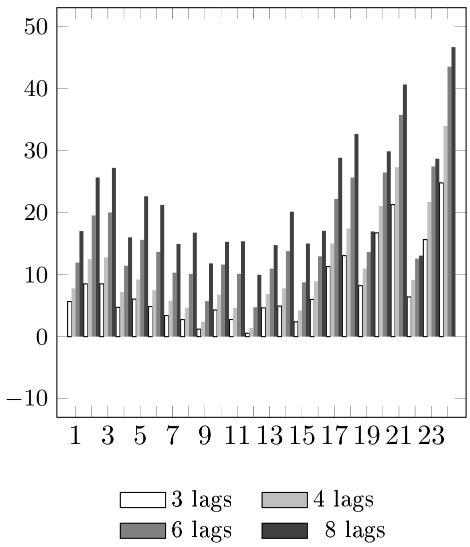
4.3. Different lags

In this section we present a sensitivity analysis to the choice of the truncation lag. Figure 5 shows the differences between the size-adjusted power of mixed- and low-frequency tests with an ad-hoc number of lags included in the test regression (3, 4, 6 and 8 lags, respectively).[18] The same number of lags is used in each replication and for each variable (the dependent variable and the covariate). These results illustrate the case of $\bar{c} = -5$, only a constant added to the test regression and of GLS demeaning.[19] A positive bar means that the size-adjusted power of mixed-frequency tests is higher than the one of low-frequency tests.

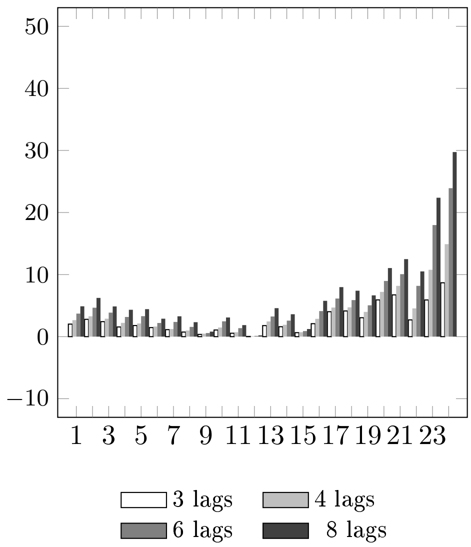
In spite of high costs in terms of size (as expected), this exercise shows that regardless of the particular choice of lags, the mixed-frequency tests tend to outperform the low-frequency ones. This is due to the fact that exploiting mixed-frequency data enables us to capture the (stronger) underlying correlation between the dependent variable and the covariate.

18. The number of lags refers to the low frequency. For the mixed-frequency tests the same time span of lagged information of the covariate is covered, which corresponds to a different number of lags in the high time frequency.

19. Results for other $\bar{c}$, including a time trend and OLS demeaning are not qualitatively different and are available from the author upon request.



(A) $T = 100$



(B) $T = 500$

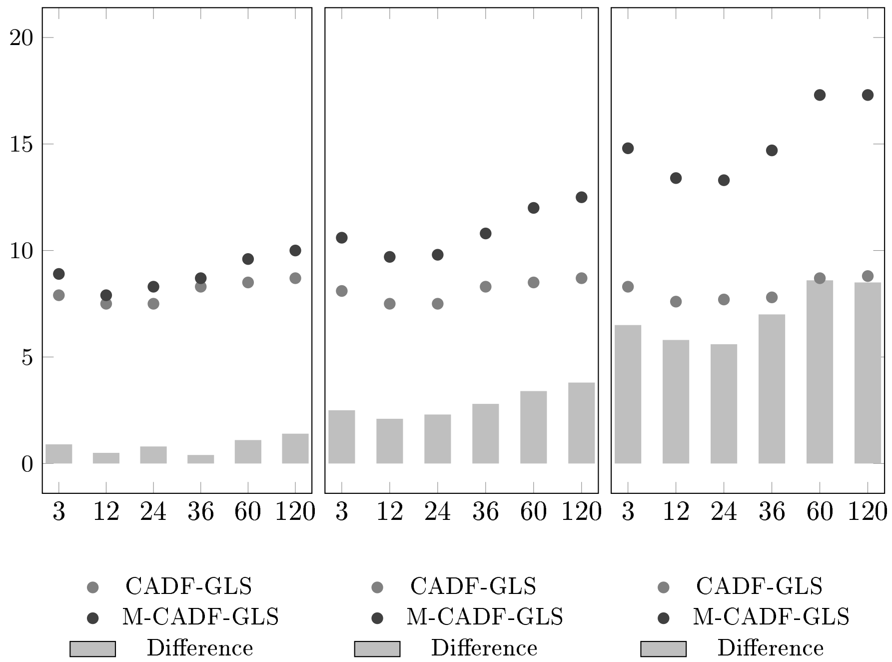
FIGURE 5: Differences in size-adjusted power between mixed- and low-frequency tests with GLS demeaning for $c = -5$ by number of lags included in the test regression, Constant only

In addition, note that the gains tend to increase with the number of lags included. The performance of the low-frequency tests is more severely affected by the inclusion of unnecessary lags. In contrast, the mixed-frequency tests are better able to deal with this kind of misspecification issues, because the weights of the high-frequency lags are data-driven. Hence, in case of great uncertainty about the choice of the truncation lag (as is typically the case in empirical applications), using the mixed-frequency framework may contribute to reduce the impact of potential misspecification in the power of the unit root tests.

4.4. Different time frequencies

To assess the impact of different combinations of time frequencies, alternative figures for m are considered. In addition to $m = 3$, now I will also consider $m = 12$, 24, 36, 60 and 120.

Hence, to have a meaningful comparison across different m, we simulate samples with $T = 100$ and 500 observations such that the aggregate figure for the first-order autoregressive parameter equals 0.95 and 0.99, respectively, regardless of m. Recall that for doing this we need to adjust the values of c accordingly.



(A) $R^2 = 0.2$ (B) $R^2 = 0.5$ (C) $R^2 = 0.8$

Figure 6: Size-adjusted power of mixed- and low-frequency tests for different m, with GLS demeaning, aggregate $\alpha = 0.99$, Constant only, $T = 500$

Figure 6 shows the results in terms of power performance of low- and mixed-frequency unit root tests, with GLS demeaning, using a sample of 500 observations.[20]

Once again, the mixed-frequency test shows better power performance for all m considered in a situation of near-integration. The mixed-frequency test is at least as good as the low-frequency test and, in many cases, substantially better, especially when R^2 is higher. The gains from using the mixed-frequency test are fairly stable across different figures for m, for each R^2. However, as noted before, the gains tend to increase significantly with R^2.

5. An application to the US unemployment rate

There is a rich discussion in the literature about the order of integration of the unemployment rate. Initial contributions by Phelps (1967) and Friedman

20. Results for OLS demeaning or with the smaller sample ($T = 100$) are qualitatively similar and were omitted for the sake of brevity. All results are available from the author upon request.

(1968) described movements in the unemployment rate as fluctuations around a natural rate, which would be generally defined as the equilibrium rate. Given that temporary shocks would have only temporary effects, these traditional theories imply that the unemployment rate is level or, perhaps, trend stationary, evolving around the natural level.[21]

In contrast, Blanchard and Summers (1986, 1988) resort to the concept of hysteresis, meaning that unemployment rates depend sensitively on the shocks an economy experienced in the past, and eventually the unemployment rate should exhibit a unit root. Unifying both strands, the structuralist theories of unemployment assumes that most shocks cause temporary movements of the unemployment rate around the natural rate, but some shocks can cause permanent changes in the natural rate. Hence, the unemployment rate would be stationary around a natural rate, which itself could be subject to structural breaks (for a brief discussion, see Phelps 1995).

The purpose of this exercise is to illustrate the used of mixed-frequency unit root tests, providing additional evidence on the persistence of US unemployment rate. We apply the above described covariate-augmented unit root tests — both low- and mixed-frequency tests — to assess whether US unemployment rate has a unit root.[22] We exploit insights provided by the correlation between the variable of interest and the continued jobless claims.

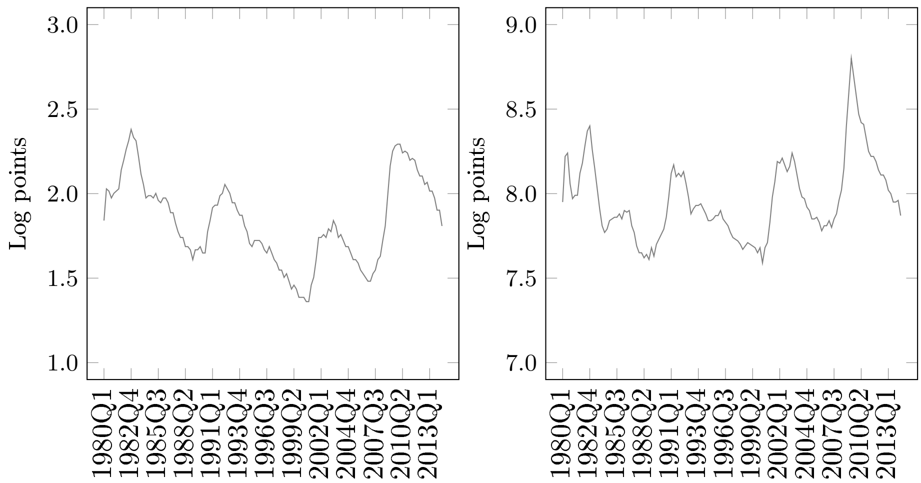
The overall unemployment rate (in logs) is a monthly series taken from the US Bureau of Labor Statistics and is seasonally adjusted. The covariate is the continued jobless claims, also expressed in logs, which is a weekly series. The continued jobless claims are released by the US Department of Labor and are also seasonally adjusted. The stationarity of the covariate was confirmed by univariate unit root tests, namely ADF and ADF-GLS tests.[23] The data cover the period from January 1980 to June 2014. For the low-frequency version of the covariate we use the beginning of the period value. Figure 7 shows the series used, in quarterly frequency (the common frequency) and in logs.

In order to choose the truncation lag of the variables, we used the multivariate version of MAIC in Perron and Qu (2007), applied to the output of an unrestricted version of the CADF test regression (for more details, see

21. There is no consensus in the literature about including or not a trend when modelling unemployment, existing an ongoing discussion based on sample-driven and theoretical arguments.

22. The difference between the unemployment rate and the natural rate of unemployment is also often analysed. In this exercise we focused on the level of the unemployment rate. Estimating a natural rate of unemployment is beyond the scope of this article. Simply using an estimate of the natural rate of unemployment collected elsewhere (e.g., the Congressional Budget Office estimates a quarterly natural rate of unemployment, which is made available by the St. Louis Federal Reserve Bank) would lead to bias on hypothesis testing, as shown by Murphy and Topel (1985).

23. The tests were performed for weekly, monthly and quarterly frequencies and, as expected, the result was always the same.



(A) Unemployment rate (B) Continued jobless claims

FIGURE 7: Data on US unemployment

section 4). Tables 6 and 7 present the results for the unit root tests to the US unemployment rate. These tables provide the estimates of δ, the t-statistics and the estimated R^2 for the covariate-augmented regressions.

Sample: 1980Q1 to 2014Q2						
Constant only						
	ADF	ADF-GLS	CADF	CADF-GLS	M-CADF	M-CADF-GLS
δ	-0.009	-0.009	-0.024	-0.024	-0.022	-0.022
t-statistic	-1.882	-1.879	$-3.439**$	$-3.428*$	$-3.309**$	$-3.293*$
R^2			0.42	0.43	0.96	0.95
Time trend included						
	ADF	ADF-GLS	CADF	CADF-GLS	M-CADF	M-CADF-GLS
δ	-0.009	-0.009	-0.041	-0.027	-0.041	-0.025
t-statistic	-1.876	-1.896	$-4.626**$	$-3.657*$	$-4.861**$	$-3.574*$
R^2			0.39	0.43	0.96	0.96

TABLE 6. Unit root tests for US monthly unemployment rate, using jobless claims as covariate

Note: For the low-frequency covariate-augmented unit root tests the frequency of the covariate equals the frequency of the dependent variable. For the mixed-frequency tests, the covariate has a weekly frequency. * significant at a 5 per cent asymptotic level. ** significant at a 1 per cent asymptotic level. For the covariate-augmented GLS family of tests, Pesavento (2006) only presents 5 per cent asymptotic significance levels.

Sample: 1980Q1 to 2006Q4						
Constant only						
	ADF	ADF-GLS	CADF	CADF-GLS	M-CADF	M-CADF-GLS
δ	−0.007	−0.007	−0.015	−0.012	−0.014	−0.010
t-statistic	−1.217	−1.089	−1.989	−1.628	−1.888	−1.406
R^2			0.44	0.44	0.95	0.92
Time trend included						
	ADF	ADF-GLS	CADF	CADF-GLS	M-CADF	M-CADF-GLS
δ	−0.018	−0.009	−0.088	−0.035	−0.078	−0.032
t-statistic	−1.913	−1.209	−5.147 * *	−3.148*	−4.787 * *	−3.043*
R^2			0.29	0.36	0.87	0.88

TABLE 7. Unit root tests for US monthly unemployment rate, using jobless claims as covariate

Note: For the low-frequency covariate-augmented unit root tests the frequency of the covariate equals the frequency of the dependent variable. For the mixed-frequency tests, the covariate has a weekly frequency. * significant at a 5 per cent asymptotic level. ** significant at a 1 per cent asymptotic level. For the covariate-augmented GLS family of tests, Pesavento (2006) only presents 5 per cent asymptotic significance levels.

We consider two different samples, one from 1980 Q1 to 2014 Q2 in Table 6 and another excluding the great recession period, from 1980 Q1 to 2006 Q4, in Table 7. By looking at Figure 7 one can intuitively see that it seems to be relevant to include a time trend in the test regression for the level of the unemployment rate, namely in the shorter sample. Therefore, for each sample, the top panel in Tables 6 and 7 shows results for only including a constant and the bottom panel shows the results for also including a time trend.

In the shorter sample, all tests for the level of the unemployment rate with only a constant included in the test regression agree in not rejecting the null hypothesis, suggesting that the series is I(1). The conclusions from the univariate tests remain unchanged when a time trend is included. However, all covariate-augmented tests agree in rejecting the null hypothesis, suggesting that the unemployment rate is trend stationary in that sample.

When using the longer sample, univariate unit root tests continue to suggest that the level of US unemployment is not stationary. In contrast, all covariate-augmented tests reject the null hypothesis, whether or not a time trend is included. Hence, the level of unemployment rate seems to be stationary, though highly persistent, as shown by the small values of $\hat{\delta}$. Notice that in all cases, the estimates of R^2 are higher when the mixed-frequency approach is used, suggesting that in this case combining information with different time frequencies allows us to take better advantage of the covariate-augmented framework of unit root tests.

6. Conclusion

Unit root tests typically have low power, especially in near-integrated cases, which results in the over-acceptance of the unit root null. This paper tries to tackle this issue by merging two strands of the literature. In particular, we try to improve the power performance of CADF unit root tests by exploiting mixed-frequency data. The results of a simulation exercise show that there is room for improvement.

Since Hansen (1995), covariate-augmented unit root tests have been proposed has a more powerful version of traditional univariate unit root tests, such as the ADF tests. The main idea is that using a stationary covariate, which is well correlated with the variable of interest, in the test regression contributes to increase the precision of the estimates of the test statistic and, hence, to increase the power of the test.

In this article we assume as main premise that temporal aggregation of the variable of interest is unavoidable. It is well known that time aggregation and the sampling frequency do not affect the long-run properties of time series, namely the presence of unit roots, but may have severe consequences for the correlation between the dependent variable and the covariate.

To exploit the advantages of combining data with different time frequencies — a dependent variable in a lower frequency than the covariate — we use the MIDAS technique. This technique uses data-driven aggregation weights. Monte Carlo experiments show that: (i) mixed-frequency covariate-augmented unit root tests have a better power performance than traditional low-frequency tests; and that (ii) mixed-frequency tests are particularly advantageous when we are in the presence of near-integrated variables. The results are robust to the size of the sample, to the lag specification of the test regression and to different combinations of time frequencies.

Applying the unit root tests — both low- and mixed-frequency — to the US unemployment rate, we found evidence that the unemployment rate is stationary, though highly persistent.

References

Andreou, Elena, Eric Ghysels, and Andros Kourtellos (2011). "Forecasting with mixed-frequency data." In *The Oxford Handbook of Economic Forecasting*, edited by Michael Clements and David Hendry, chap. 8, pp. 225–267. Oxford University Press.

Andreou, Elena, Eric Ghysels, and Andros Kourtellos (2013). "Should macroeconomic forecasters use daily financial data and how?" *Journal of Business and Economic Statistics*, 2(31), 240–251.

Andrews, Donald W K (1991). "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation." *Econometrica*, 59(3), 817–58.

Blanchard, Olivier J. and Lawrence H. Summers (1986). "Hysteresis And The European Unemployment Problem." In *NBER Macroeconomics Annual 1986, Volume 1*, NBER Chapters, pp. 15–90. National Bureau of Economic Research, Inc.

Blanchard, Olivier Jean and Lawrence H Summers (1988). "Beyond the Natural Rate Hypothesis." *American Economic Review*, 78(2), 182–87.

Brillinger, D.R. (2001). *Time Series: Data Analysis and Theory*. Classics in Applied Mathematics, Society for Industrial and Applied Mathematics.

Caporale, Guglielmo Maria and Nikitas Pittis (1999). "Unit Root Testing Using Covariates: Some Theory and Evidence." *Oxford Bulletin of Economics and Statistics*, 61(4), 583–95.

Chang, Yoosoon and Joon Y. Park (2002). "On the asymptotics of ADF tests for unit roots." *Econometric Reviews*, 21(4), 431–447.

Christopoulos, Dimitris K. and Miguel A. Léon-Ledesma (2008). "Time-series output convergence tests and stationary covariates." *Economics Letters*, 101(3), 297–299.

Elliott, Graham (1999). "Efficient Tests for a Unit Root When the Initial Observation is Drawn From Its Unconditional Distribution." *International Economic Review*, 40(3), 767–784.

Elliott, Graham and Michael Jansson (2003). "Testing for unit roots with stationary covariates." *Journal of Econometrics*, 115(1), 75–89.

Elliott, Graham, Thomas J Rothenberg, and James H Stock (1996). "Efficient Tests for an Autoregressive Unit Root." *Econometrica*, 64(4), 813–36.

Foroni, Claudia and Massimiliano Marcellino (2012). "A Comparison of Mixed Frequency Approaches for Modelling Euro Area Macroeconomic Variables." Economics Working Papers ECO2012/07, European University Institute.

Foroni, Claudia, Massimiliano Marcellino, and Christian Schumacher (2011). "U-MIDAS: MIDAS regressions with unrestricted lag polynomials." Discussion Paper Series 1: Economic Studies 2011,35, Deutsche Bundesbank, Research Centre.

Fossati, Sebastian (2012). "Covariate unit root tests with good size and power." *Computational Statistics & Data Analysis*, 56(11), 3070–3079.

Friedman, Milton (1968). "The Role of Monetary Policy." *The American Economic Review*, 58(1), pp. 1–17.

Galvão, Ana Beatriz (2013). "Changes in predictive ability with mixed frequency data." *International Journal of Forecasting*, 29(3), 395–410.

Galvão Jr., Antonio F. (2009). "Unit root quantile autoregression testing using covariates." *Journal of Econometrics*, 152(2), 165–178.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2006). "Predicting volatility: getting the most out of return data sampled at different frequencies." *Journal of Econometrics*, 131, 59–95.

Ghysels, E., A. Sinko, and R. Valkanov (2007). "MIDAS regressions: Further results and new directions." *Econometric Reviews*, 26(1), 53–90.

Ghysels, Eric and J. Isaac Miller (2013). "Testing for Cointegration with Temporally Aggregated and Mixed-frequency Time Series." Working Papers 1307, Department of Economics, University of Missouri.

Ghysels, Eric, Pedro Santa-Clara, and Rossen Valkanov (2004). "The MIDAS touch: Mixed data sampling regression models." CIRANO Working Papers 2004s-20, CIRANO.

Granger, Clive and P. Newbold (1974). "Spurious regressions in econometrics." *Journal of Econometrics*, 2, 111–120.

Granger, Clive and Pierre Siklos (1995). "Systematic sampling, temporal aggregation, seasonal adjustment, and cointegration theory and evidence." *Journal of Econometrics*, 66, 357–369.

Haldrup, Niels and Michael Jansson (2006). "Improving Size and Power in Unit Root Testing." In *Palgrave Handbook of Econometrics*, vol. 2: Applied Econometrics, edited by Terence C. Mills and Kerry Patterson, chap. 7, pp. 252–277. Palgrave Macmillan.

Hansen, Bruce E. (1995). "Rethinking the Univariate Approach to Unit Root Testing: Using Covariates to Increase Power." *Econometric Theory*, 11(05), 1148–1171.

Haug, Alfred A (2002). "Temporal Aggregation and the Power of Cointegration Tests: A Monte Carlo Study." *Oxford Bulletin of Economics and Statistics*, 64(4), 399–412.

Jansson, Michael (2002). "Consistent Covariance Matrix Estimation For Linear Processes." *Econometric Theory*, 18(06), 1449–1459.

Jansson, Michael (2004). "Stationarity Testing With Covariates." *Econometric Theory*, 20(01), 56–94.

Juhl, Ted and Zhijie Xiao (2003). "Power functions and envelopes for unit root tests." *Econometric Theory*, 19, 240–253.

Lupi, Claudio (2009). "Unit Root CADF Testing with R." *Journal of Statistical Software*, 32(i02).

Marcellino, Massimiliano (1999). "Consequences of temporal aggregation in empirical analysis." *Journal of Business and Economic Statistics*, 17(1), 129–136.

Marcellino, Massimiliano and Christian Schumacher (2010). "Factor MIDAS for Nowcasting and Forecasting with Ragged-Edge Data: A Model Comparison for German GDP." *Oxford Bulletin of Economics and Statistics*, 72(4), 518–550.

Müller, Ulrich K. and Graham Elliott (2003). "Tests for Unit Roots and the Initial Condition." *Econometrica*, 71(4), 1269–1286.

Murphy, Kevin and Robert Topel (1985). "Estimation and Inference in Two-Step Econometric Models." *Journal of Business and Economic Statistics*, 3(4), 370–379.

Ng, Serena and Pierre Perron (1995). "Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag." *Journal of the American Statistical Association*, 90(429), pp. 268–281.

Ng, Serena and Pierre Perron (2001). "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power." *Econometrica*, 69(6), 1519–1554.

Perron, Pierre and Serena Ng (1996). "Useful Modifications to Some Unit Root Tests with Dependent Errors and Their Local Asymptotic Properties." *Review of Economic Studies*, 63(3), 435–63.

Perron, Pierre and Zhongjun Qu (2007). "A simple modification to improve the finite sample properties of Ng and Perron's unit root tests." *Economics Letters*, 94(1), 12–19.

Pesavento, Elena (2006). "Near-Optimal Unit Root Tests with Stationary Covariates with Better Finite Sample Size." Economics Working Papers ECO2006/18, European University Institute.

Phelps, Edmund S. (1967). "Phillips Curves, Expectations of Inflation and Optimal Unemployment over Time." *Economica*, 34(135), pp. 254–281.

Phelps, Edmund S. (1995). "The Structuralist Theory of Employment." *The American Economic Review*, 85(2), pp. 226–231.

Phillips, P. C. B. (1986). "Understanding spurious regressions in econometrics." *Journal of Econometrics*, 33(3), 311–340.

Phillips, P. C. B. (1987). "Towards a unified asymptotic theory for autoregression." *Biometrika*, 74(3), 535–547.

Pierse, R. G. and A. J. Snell (1995). "Temporal aggregation and the power of tests for a unit root." *Journal of Econometrics*, 65(2), 333–345.

Rodrigues, Paulo M.M. (2006). "Properties of recursive trend-adjusted unit root tests." *Economics Letters*, 91(3), 413–419.

Saikkonen, Pentti (1991). "Asymptotically Efficient Estimation of Cointegration Regressions." *Econometric Theory*, 7(01), 1–21.

Schwert, G. William (1989). "Tests for Unit Roots: A Monte Carlo Investigation." *Journal of Business and Economic Statistics*, 7(2), 147–159.

Silvestrini, Andrea and David Veredas (2008). "Temporal Aggregation Of Univariate And Multivariate Time Series Models: A Survey." *Journal of Economic Surveys*, 22(3), 458–497.

Stock, James H. (1986). "Unit roots, structural breaks and trends." In *Handbook of Econometrics*, *Handbook of Econometrics*, vol. 4, edited by R. F. Engle and D. McFadden, chap. 46, pp. 2739–2841. Elsevier.

Stock, James H. (1999). "A Class of Tests for Integration and Cointegration." In *Cointe- gration, Causality, and Forecasting: A Festschrift for Clive W. J. Granger*, edited by R. F. Engle and H. White, chap. 6, pp. 135–167. Oxford University Press.

Sul, Donggyu, Peter C. B. Phillips, and Chi-Young Choi (2005). "Prewhitening Bias in HAC Estimation." *Oxford Bulletin of Economics and Statistics*, 67(4), 517–546.

Taylor, A. M. Robert (2002). "Regression-Based Unit Root Tests with Recursive Mean Adjustment for Seasonal and Nonseasonal Time Series." *Journal of Business and Economic Statistics*, 20(2), 269–281.

Westerlund, Joakim (2015). "On the Importance of the First Observation in GLS Detrending in Unit Root Testing." *Oxford Bulletin of Economics and Statistics*, 1(77), 152–161.

Working, Holbrook (1960). "Note on the Correlation of First Differences of Averages in a Random Chain." *Econometrica*, 28(4), pp. 916–918.

Appendix

	Standard			Demeaned			Detrended		
ρ^2	1%	5%	10%	1%	5%	10%	1%	5%	10%
1	-2.57	-1.94	-1.62	-3.43	-2.86	-2.57	-3.96	-3.41	-3.13
0.9	-2.57	-1.94	-1.61	-3.39	-2.81	-2.50	-3.88	-3.33	-3.04
0.8	-2.57	-1.94	-1.6	-3.36	-2.75	-2.46	-3.83	-3.27	-2.97
0.7	-2.55	-1.93	-1.59	-3.30	-2.72	-2.41	-3.76	-3.18	-2.87
0.6	-2.55	-1.90	-1.56	-3.24	-2.64	-2.32	-3.68	-3.10	-2.78
0.5	-2.55	-1.89	-1.54	-3.19	-2.58	-2.25	-3.60	-2.99	-2.67
0.4	-2.55	-1.89	-1.53	-3.14	-2.51	-2.17	-3.49	-2.87	-2.53
0.3	-2.52	-1.85	-1.51	-3.06	-2.40	-2.06	-3.37	-2.73	-2.38
0.2	-2.49	-1.82	-1.46	-2.91	-2.28	-1.92	-3.19	-2.55	-2.20
0.1	-2.46	-1.78	-1.42	-2.78	-2.12	-1.75	-2.97	-2.31	-1.95

TABLE A.1. Asymptotic critical values for CADF t-statistics

Note: Following Hansen (1995). The critical values were calculated from 60.000 draws generated from samples of size 1.000 with *iid* Gaussian inovations.

R^2	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Cases 1, 2	-1.948	-1.939	-1.929	-1.918	-1.905	-1.881	-1.864	-1.839	-1.818	-1.773
Case 3	-1.948	-1.909	-1.866	-1.812	-1.760	-1.707	-1.647	-1.579	-1.497	-1.405
Case 4	-2.836	-2.786	-2.738	-2.688	-2.628	-2.568	-2.498	-2.418	-2.343	-2.315
Case 5	-2.835	-2.780	-2.730	-2.664	-2.586	-2.497	-2.401	-2.286	-2.152	-2.017

TABLE A.2. Asymptotic critical values for the CADF-GLS test

Note: Following Pesavento (2006). The critical values were computed using 60 000 replications of samples with 1000 observations, drawn using *iid* Gaussian innovations. The critical values reported are for tests of size 5%, with $\bar{c}$ equal to -7 for cases 1, 2 and 3 and to -13.5 for cases 4 and 5.

WORKING PAPERS

2013

01|13 Macroeconomic forecasting using low-frequency filters
João Valle e Azevedo | Ana Pereira

02|13 Everything you always wanted to know about sex discrimination
Ana Rute Cardoso | Paulo Guimarães | Pedro Portugal

03|13 Is there a role for domestic demand pressure on export performance?
Paulo Soares Esteves | António Rua

04|13 Ageing and fiscal sustainability in a small euro area economy
Gabriela Castro | José R. Maria | Ricardo Mourinho Félix | Cláudia Rodrigues Braz

05|13 Mind the gap! The relative wages of immigrants in the Portuguese labour market
Sónia Cabral | Cláudia Duarte

06|13 Foreign direct investment and institutional reform: Evidence and an application to Portugal
Paulo Júlio | Ricardo Pinheiro-Alves | José Tavares

07|13 Monetary policy shocks: We got news!
Sandra Gomes | Nikolay Iskrev | Caterina Mendicino

08|13 Competition in the Portuguese Economy: Estimated price-cost margins under imperfect labour markets
João Amador | Ana Cristina Soares

09|13 The sources of wage variation: a three-way high-dimensional fixed effects regression model
Sonia Torres | Pedro Portugal | John T. Addison | Paulo Guimarães

10|13 The output effects of (non-separable) government consumption at the zero lower bound
Valerio Ercolani | João Valle e Azevedo

11|13 Fiscal multipliers in a small euro area economy: How big can they get in crisis times?
Gabriela Castro | Ricardo M. Felix | Paulo Julio | Jose R. Maria

12|13 Survey evidence on price and wage rigidities in Portugal
Fernando Martins

13|13 Characterizing economic growth paths based on new structural change tests
Nuno Sobreira | Luis C. Nunes | Paulo M. M. Rodrigues

14|13 Catastrophic job destruction
Anabela Carneiro | Pedro Portugal | José Varejão

15|13 Output effects of a measure of tax shocks based on changes in legislation for Portugal
Manuel Coutinho Pereira | Lara Wemans

16|13 Inside PESSOA - A detailed description of the model
Vanda Almeida | Gabriela Castro | Ricardo M. Félix | Paulo Júlio | José R. Maria

17|13 Macroprudential regulation and macroeconomic activity
Sudipto Karmakar

18|13 Bank capital and lending: An analysis of commercial banks in the United States
Sudipto Karmakar | Junghwan Mok

2014

2015