Nonlinearly testing for a unit root in the presence of a break in the mean

Gluschenko, Konstantin

August 2004

Nonlinearly Testing for a Unit Root in the Presence of a Break in the Mean

Konstantin Gluschenko

Institute of Economics and Industrial Engineering,
Siberian Branch of the
Russian Academy of Sciences

pr. Lavrentieva, 17, 630090 Novosibirsk, Russia
glu@nsu.ru

First version: August 2004
This version: September 2005

NONLINEARLY TESTING FOR A UNIT ROOT IN THE PRESENCE OF A BREAK IN THE MEAN

KONSTANTIN GLUSCHENKO

This paper deals with testing a time series with a structural break in its mean for a unit root when the break date is known. A nonlinear (with respect to coefficients) test equation is used, providing asymptotically efficient estimates. Finite-sample and quasi-asymptotic empirical distributions of the unit root test statistics are estimated, comparing them with those associated with the Perron-type equations. Asymptotic distributions of the nonlinear test statistics are found to be the Dickey-Fuller distributions. The nonlinear test proves to have more power than the test based on the linear model.

KEY WORDS: Structural break, Nonlinear regression, Nonstandard distribution.
JEL CODES: C12, C15, C16, C22

1. INTRODUCTION

Since the seminal works of Perron (1989) and Rappoport and Reichlin (1989), the growing literature explores various aspects of a structural break in time series, such as taking account of changes in any or all of a time-series model parameters (intercept, trend parameter, autoregressive parameter, residual variance), unknown break date, multiple breaks, etc. – see, e.g., Hansen (1997, 2001), Bartley *et al*. (2001), Lanne *et al*. (2002), to name a few. Nevertheless, the rather simple case of a single structural break affecting only the mean of a time series at a known date is still of interest, for example, in testing for the law of one price or purchasing power parity.

The most widespread method of testing for stationarity in this case is that put forward by Perron (1990) and Perron and Vogelsang (1992). Many time-series econometrics textbooks describe this method as a standard one. However, there is room for improvements in it. The point is that the test equations used are in fact a linear approximation of "true" ones which are nonlinear with respect to coefficients. Therefore estimates of the autoregressive parameter, being consistent, are asymptotically inefficient. Hence it can be expected that models describing the process under consideration more adequately would provide tests with better properties.

This paper deals with a nonlinear specification of the test equations describing the first-order autoregressive process with a one-time break in the mean of the process, considering equation with and without the intercept term. (For empirical purposes, such a test was used by Gluschenko (2004) for testing the law of one price in the Russian economy.) The *t*-ratio of the autoregression coefficient is used as the unit root test statistic. Finite-sample and quasi-asymptotic empirical distributions of this statistic are estimated, comparing them with those associated with the Perron-type equations.

Like the Perron distributions, the finite-sample distributions of the test statistics based on the nonlinear regressions depend on the time point when the break occurs. In the asymptotic case, contrastingly, they do not. The most interesting property of the nonlinear-test statistics is that their limiting distributions are the Dickey-Fuller ones. The nonlinear test is found to have more power than that based on the Perron-type equations.

The rest of the paper is organized as follows. Section 2 introduces nonlinear test equations. Section 3 presents and discusses estimated distributions of the statistics for testing time series with a break for a unit root, comparing those based on the linear and nonlinear test equations. Section 4 reports results of power experiments. Section 5 contains an empirical application of the nonlinear tests to testing for the law of one price in Russia (relevant time series containing breaks caused by the 1998 financial crisis in the country). Section 6 concludes.

2. THE MODELS

Let us consider a first-order autoregressive process with a break which changes the mean of the process from μ_0 to μ_1 at a known point $t = \theta + 1$:

$$\begin{aligned} y_t &= \mu_0 + (\mu_1 - \mu_0)B_{\theta t} + \nu_t \quad (t = 0, 1,\ldots, T), \\ \nu_t &= (\lambda+1)\nu_{t-1} + \varepsilon_t \quad (t = 1,\ldots, T), \\ \nu_0 &= \xi, \end{aligned} \tag{1}$$

where ε_t is a Gaussian white noise; ξ is either a constant (which can, for example, result in $y_0 = 0$) or a random variable, and $B_{\theta t}$ is a step (level) dummy such that $B_{\theta t} = 0$ if $t \leq \theta$ and 1 otherwise. In the literature, two dummies are commonly used to characterize the break, the pulse dummy and the step one. However, this is superfluous, since the pulse dummy taking the value of 1 if $t = \theta + 1$ and 0 otherwise can be represented as $B_{\theta t} - B_{\theta,t-1}$ or $\Delta B_{\theta t}$, where Δ stands for the first difference operator.

The interest is to distinguish between hypotheses H_0: $\lambda = 0$ against H_1: $\lambda < 0$. In doing so, the t-ratio of λ, $\tau = \hat{\lambda}/\hat{\sigma}_{\lambda}$, will be used as the test statistic. This t-statistic is denoted by τ (with a subscript which indicates belonging to a particular test) in order to underline that it has a distribution differing from the standard t distribution.

Combining the first two equations in (1), a nonlinear model is arrived at: $\Delta y_t = \alpha + \lambda y_{t-1} + \gamma(B_{\theta t} - (\lambda+1)B_{\theta,t-1}) + \varepsilon_t$, or

$$\Delta y_t = \alpha + \lambda y_{t-1} + \gamma B_{\theta t} - \gamma(\lambda+1)B_{\theta,t-1} + \varepsilon_t, \tag{2}$$

where $\alpha \equiv -\lambda\mu_0$ and $\gamma \equiv \mu_1 - \mu_0$; from now on, $t = 1,\ldots,T$.

Supposing the base level of the process to be zero, $\mu_0 = 0$ in (1), we have a model with no intercept:

$$\Delta y_t = \lambda y_{t-1} + \gamma B_{\theta t} - \gamma(\lambda+1)B_{\theta,t-1} + \varepsilon_t. \tag{3}$$

Under the null hypothesis of a unit root, $\lambda = 0$, only a pulse remains in (2) and (3); both regressions degenerate to

$$\Delta y_t = \gamma(B_{\theta t} - B_{\theta,t-1}) + \varepsilon_t. \tag{4}$$

Using the above notations, the specification of the AR(1) model allowing for break in the spirit of Perron (1990) – the Perron-type equation – looks like

$$\Delta y_t = \alpha + \lambda y_{t-1} + \psi B_{\theta t} + \delta \Delta B_{\theta t} + \varepsilon_t, \tag{5}$$

or, with $\gamma \equiv \psi + \delta$,

$$\Delta y_t = \alpha + \lambda y_{t-1} + \gamma B_{\theta t} - \delta B_{\theta,t-1} + \varepsilon_t. \tag{6}$$

In the case of no intercept, the Perron-type model takes the form:

$$\Delta y_t = \lambda y_{t-1} + \gamma B_{\theta t} - \delta B_{\theta,t-1} + \varepsilon_t. \tag{7}$$

It is (6) and (7) – usually, in a form similar to (5) – and their modifications that are commonly used to test for a unit root in time series having a break in their mean, the parameterized null hypothesis coinciding with (4). Note that, since the process under consideration is AR(1), there is no difference between the additive outlier model and the innovational outlier model considered by Perron (1990) and Perron and Vogelsang (1992). Comparing (2) and (3) with (6) and (7), it is seen that the constraint $\delta = \gamma(\lambda + 1)$ is omitted in the latter equations, so making them linear. But the cost of this linearity is some sacrifice in adequate characterization of properties of process (1).

Let us rearrange model (2) so that it has the customary "two-dummy representation:"

$$\Delta y_t = \alpha + \lambda y_{t-1} - \gamma\lambda B_{\theta t} + \gamma(\lambda + 1)\Delta B_{\theta t} + \varepsilon_t. \tag{8}$$

Thus, in fact $\psi = -\gamma\lambda$ and $\delta = \gamma(\lambda + 1)$. And so, coefficients ψ and δ in (5) have no transparent interpretation by themselves, while parameter γ in the nonlinear models has a simple meaning: it is just the "height" of the break, $\mu_1 - \mu_0$. At the same time, it is reasonable to expect that in the extreme cases of no autocorrelation ($\lambda = -1$) and of a unit root ψ and δ, respectively, would represent the height of the break. Given no autocorrelation, the lagged values of break variable $B_{\theta t}$ should not contribute to the current value of the dependent variable. Indeed, with $\lambda = -1$, $\psi = \gamma$ and $\delta = 0$ in (8); and we have a step only: $\Delta y_t = \alpha - y_{t-1} + \gamma B_{\theta t} + \varepsilon_t$, that is, $y_t = \mu_0 + \gamma B_{\theta t} + \varepsilon_t$. But that is not the case when we deal with (5), since ψ and δ are independent of each other and of λ; a combination of a step and a pulse still takes place like in (8) with $-1 < \lambda < 0$. Similarly, provided that $\lambda = 0$, $\psi = 0$ and $\delta = \gamma$ in (8), so producing only a pulse (4), $\Delta y_t = \gamma \Delta B_{\theta t} + \varepsilon_t$. But there is no reason why ψ would vanish in (5) in the unit-root case. Hence, while the parameterization of the null hypothesis as (4) explicitly follows from the nonlinear models themselves, it does not follow from the Perron-type models, being an *ad hoc* one.

Moreover, the absence of the constraint on parameters in the Perron-type models leads to that parameter estimates, while being consistent, are not asymptotically efficient. Hence it can be expected that obtaining efficient estimates with the use of more adequate nonlinear models would provide unit root tests with better properties.

Let the unit root test statistics for (2) and (3) be denoted $\tau_{\mu NL}$ and τ_{0NL}, respectively, and those for (6) and (7) be labeled as $\tau_{\mu P}$ and τ_{0P}. Using known expressions for estimators of λ (to be exact, of $\rho = \lambda + 1$) and its t-ratio, Perron (1990) derives the asymptotic distribution of $\tau_{\mu P}$ under the null hypothesis (and the distribution of τ_{0P} could be similarly derived). But the nonlinearity of test equations (2) and (3) prevents from obtaining such estimators in a closed form; and so, even the asymptotic distributions of $\tau_{\mu NL}$ and τ_{0NL} cannot be derived analytically. Thus, the only way to explore properties of the unit root tests based on these equations is to examine numerical distributions estimated through Monte Carlo simulations.

3. DISTRIBUTIONS OF THE TEST STATISTICS

To derive numerical distributions of the test statistics under the null hypothesis of a unit root, the data generating process used is $y_t = y_{t-1} + \gamma(B_{\theta t} - B_{\theta,t-1}) + \varepsilon_t$, where $\varepsilon_t \sim$ iid

$N(0,1)$. In each simulation, distributions of $\tau_{\mu NL}$, τ_{0NL}, $\tau_{\mu P}$, and τ_{0P} are estimated over the same set of simulated time series; simulations are based on 200,000 replications. For comparison, the distributions of the Dickey-Fuller τ-statistics associated with regressions with and without constant, τ_{μ} and τ_0, respectively, are estimated as well. Models (2) and (3) are estimated through nonlinear least squares, applying the Marquardt algorithm. The EViews built-in procedures have been used for estimating; see Quantitative Micro Software (2004). The accuracy (measured relative to the norm of the vector of scaled regression parameters) is taken equal to 10^{-7}.

Simulations – hence relevant distributions – differ in the set of parameters $\{T, \Theta, \gamma, y_0\}$, where Θ is the "pre-break fraction," $\Theta = \theta/T$. The distributions of $\tau_{\mu NL}$ and τ_{0NL} as well as those of $\tau_{\mu P}$ and τ_{0P} turn out to be independent of parameter γ, and so, the results are reported for the case of $\gamma = 0$. (Thus, the data generating process is in fact a pure random walk, $y_t = y_{t-1} + \varepsilon_t$.) The initial value, y_0, is set either to a constant or to $y_0 \sim N(0, \zeta^2)$; in the latter case, y_0 is distributed independently of $\{\varepsilon_t\}$. The sample size of $T = 10{,}000$ is used to obtain "quasi-asymptotic" distributions. Judging from the fact that the estimated critical values of the Dickey-Fuller statistics and Perron's $\tau_{\mu P}$ for $T = 10{,}000$ prove to be close to their asymptotical values, such a sample size can be believed to be a good "approximation of infinity."

In the next three subsections, results obtained are presented and discussed. For brevity, cumulative distribution functions are referred to as simply distributions. Unless otherwise indicated, the results are reported for the baseline case of $y_0 = 0$. (The distributions of $\tau_{\mu P}$ and $\tau_{\mu NL}$ are found to be independent of the initial condition, as they must.)

3.1 *Statistics for Regressions with Constant*

The finite-sample simulation results for the unit root test statistics associated with equations (2) and (6) are reported in Table I. Figure 1 demonstrates the 10% tails of selected distributions.

TABLE I.

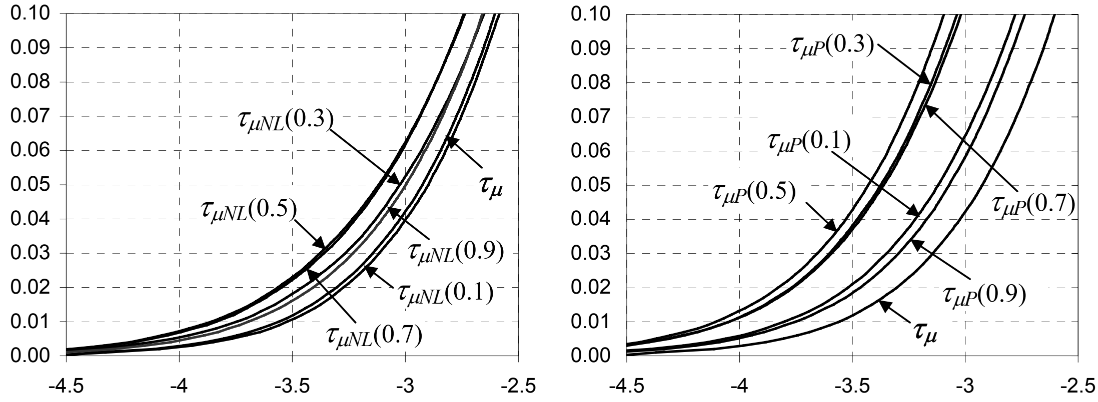
PERCENTAGE POINTS OF THE DISTRIBUTIONS OF $\tau_{\mu NL}$ AND $\tau_{\mu P}$ STATISTICS

Statistic	Θ	Percentage points							
		1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
					T =50				
τ_{μ}		-3.560	-3.216	-2.925	-2.604	-0.403	-0.036	0.284	0.661
		(-3.568)	(-3.213)	(-2.921)	(-2.599)	(-0.406)	(-0.040)	(0.281)	(0.659)
$\tau_{\mu NL}$	0.1	-3.523	-3.186	-2.898	-2.584	-0.389	-0.026	0.290	0.664
$\tau_{\mu P}$		-3.803	-3.426	-3.121	-2.779	-0.491	-0.117	0.204	0.579
$\tau_{\mu NL}$	0.2	-3.642	-3.253	-2.934	-2.592	-0.381	-0.017	0.298	0.667
$\tau_{\mu P}$		-3.964	-3.591	-3.284	-2.938	-0.604	-0.234	0.096	0.474
$\tau_{\mu NL}$	0.3	-3.745	-3.354	-3.025	-2.651	-0.379	-0.012	0.305	0.675
$\tau_{\mu P}$		-4.046	-3.679	-3.374	-3.037	-0.695	-0.326	-0.014	0.370
$\tau_{\mu NL}$	0.4	-3.818	-3.418	-3.084	-2.708	-0.379	-0.006	0.311	0.686
$\tau_{\mu P}$		-4.085	-3.727	-3.425	-3.086	-0.752	-0.387	-0.078	0.284
$\tau_{\mu NL}$	0.5	-3.856	-3.459	-3.118	-2.736	-0.378	-0.006	0.314	0.689
$\tau_{\mu P}$		-4.105	-3.732	-3.431	-3.094	-0.766	-0.410	-0.096	0.257
$\tau_{\mu NL}$	0.6	-3.859	-3.458	-3.124	-2.745	-0.389	-0.007	0.316	0.691
$\tau_{\mu P}$		-4.092	-3.715	-3.410	-3.072	-0.742	-0.382	-0.069	0.320
$\tau_{\mu NL}$	0.7	-3.835	-3.445	-3.109	-2.741	-0.393	-0.015	0.317	0.705
$\tau_{\mu P}$		-4.047	-3.670	-3.362	-3.019	-0.676	-0.309	0.009	0.395
$\tau_{\mu NL}$	0.8	-3.782	-3.410	-3.073	-2.713	-0.402	-0.020	0.312	0.701
$\tau_{\mu P}$		-3.937	-3.570	-3.262	-2.909	-0.587	-0.210	0.114	0.494
$\tau_{\mu NL}$	0.9	-3.694	-3.312	-2.998	-2.659	-0.405	-0.028	0.300	0.691
$\tau_{\mu P}$		-3.752	-3.373	-3.071	-2.736	-0.475	-0.104	0.225	0.610
					T =100				
τ_{μ}		-3.497	-3.166	-2.889	-2.580	-0.421	-0.063	0.253	0.641
		(-3.497)	(-3.167)	(-2.891)	(-2.582)	(-0.423)	(-0.059)	(0.259)	(0.632)
$\tau_{\mu NL}$	0.1	-3.473	-3.151	-2.875	-2.571	-0.415	-0.055	0.257	0.642
$\tau_{\mu P}$		-3.743	-3.390	-3.094	-2.763	-0.523	-0.151	0.170	0.552
$\tau_{\mu NL}$	0.2	-3.518	-3.170	-2.880	-2.569	-0.410	-0.052	0.263	0.649
$\tau_{\mu P}$		-3.872	-3.527	-3.246	-2.919	-0.635	-0.261	0.064	0.445
$\tau_{\mu NL}$	0.3	-3.581	-3.212	-2.914	-2.589	-0.411	-0.051	0.265	0.646
$\tau_{\mu P}$		-3.964	-3.620	-3.330	-3.008	-0.730	-0.361	-0.047	0.334
$\tau_{\mu NL}$	0.4	-3.625	-3.244	-2.945	-2.608	-0.410	-0.050	0.269	0.650
$\tau_{\mu P}$		-4.003	-3.670	-3.375	-3.058	-0.789	-0.427	-0.116	0.247
$\tau_{\mu NL}$	0.5	-3.642	-3.282	-2.968	-2.628	-0.411	-0.048	0.271	0.659
$\tau_{\mu P}$		-3.998	-3.666	-3.393	-3.076	-0.802	-0.452	-0.143	0.219
$\tau_{\mu NL}$	0.6	-3.645	-3.281	-2.976	-2.636	-0.416	-0.050	0.272	0.654
$\tau_{\mu P}$		-3.999	-3.655	-3.376	-3.054	-0.782	-0.430	-0.120	0.246
$\tau_{\mu NL}$	0.7	-3.640	-3.269	-2.968	-2.636	-0.422	-0.051	0.271	0.660
$\tau_{\mu P}$		-3.949	-3.605	-3.323	-3.000	-0.724	-0.361	-0.039	0.340
$\tau_{\mu NL}$	0.8	-3.613	-3.250	-2.947	-2.624	-0.422	-0.054	0.271	0.662
$\tau_{\mu P}$		-3.882	-3.523	-3.231	-2.901	-0.623	-0.258	0.066	0.447
$\tau_{\mu NL}$	0.9	-3.559	-3.215	-2.922	-2.605	-0.425	-0.059	0.266	0.668
$\tau_{\mu P}$		-3.711	-3.362	-3.069	-2.746	-0.517	-0.142	0.177	0.554

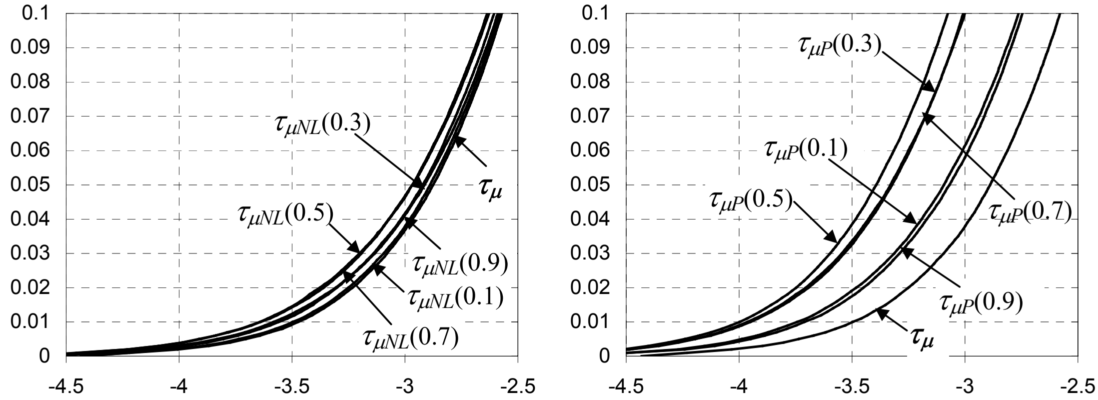
Table I (continued)

					T=150				
τ_{μ}		-3.474	-3.157	-2.891	-2.580	-0.429	-0.068	0.253	0.622
		(-3.474)	(-3.151)	(-2.881)	(-2.577)	(-0.429)	(-0.066)	(0.252)	(0.624)
$\tau_{\mu NL}$	0.1	-3.452	-3.146	-2.881	-2.574	-0.425	-0.065	0.256	0.626
$\tau_{\mu P}$		-3.715	-3.376	-3.086	-2.763	-0.533	-0.166	0.162	0.533
$\tau_{\mu NL}$	0.2	-3.465	-3.152	-2.879	-2.574	-0.422	-0.061	0.260	0.626
$\tau_{\mu P}$		-3.854	-3.511	-3.228	-2.905	-0.645	-0.278	0.046	0.417
$\tau_{\mu NL}$	0.3	-3.520	-3.181	-2.900	-2.587	-0.420	-0.059	0.259	0.626
$\tau_{\mu P}$		-3.931	-3.589	-3.311	-3.000	-0.740	-0.378	-0.053	0.321
$\tau_{\mu NL}$	0.4	-3.551	-3.202	-2.919	-2.599	-0.421	-0.058	0.261	0.634
$\tau_{\mu P}$		-3.953	-3.638	-3.365	-3.057	-0.805	-0.444	-0.125	0.239
$\tau_{\mu NL}$	0.5	-3.557	-3.221	-2.933	-2.607	-0.423	-0.059	0.263	0.634
$\tau_{\mu P}$		-3.952	-3.633	-3.368	-3.066	-0.819	-0.465	-0.156	0.211
$\tau_{\mu NL}$	0.6	-3.570	-3.222	-2.938	-2.613	-0.425	-0.059	0.261	0.634
$\tau_{\mu P}$		-3.959	-3.632	-3.353	-3.046	-0.794	-0.438	-0.125	0.239
$\tau_{\mu NL}$	0.7	-3.567	-3.225	-2.937	-2.612	-0.425	-0.063	0.264	0.634
$\tau_{\mu P}$		-3.913	-3.589	-3.315	-3.002	-0.734	-0.366	-0.054	0.321
$\tau_{\mu NL}$	0.8	-3.549	-3.209	-2.927	-2.606	-0.427	-0.062	0.256	0.632
$\tau_{\mu P}$		-3.848	-3.514	-3.228	-2.904	-0.635	-0.265	0.058	0.432
$\tau_{\mu NL}$	0.9	-3.515	-3.181	-2.911	-2.596	-0.429	-0.066	0.256	0.631
$\tau_{\mu P}$		-3.693	-3.355	-3.072	-2.758	-0.524	-0.162	0.155	0.535
					T=200				
τ_{μ}		-3.462	-3.143	-2.882	-2.583	-0.431	-0.070	0.244	0.610
		(-3.463)	(-3.144)	(-2.876)	(-2.574)	(-0.432)	(-0.069)	(0.249)	(0.620)
$\tau_{\mu NL}$	0.1	-3.452	-3.136	-2.876	-2.578	-0.427	-0.069	0.247	0.611
$\tau_{\mu P}$		-3.699	-3.354	-3.076	-2.764	-0.539	-0.170	0.150	0.520
$\tau_{\mu NL}$	0.2	-3.462	-3.136	-2.874	-2.575	-0.426	-0.065	0.249	0.616
$\tau_{\mu P}$		-3.828	-3.503	-3.226	-2.910	-0.652	-0.281	0.041	0.424
$\tau_{\mu NL}$	0.3	-3.489	-3.154	-2.884	-2.583	-0.426	-0.064	0.251	0.614
$\tau_{\mu P}$		-3.918	-3.589	-3.313	-3.005	-0.748	-0.383	-0.059	0.322
$\tau_{\mu NL}$	0.4	-3.513	-3.175	-2.901	-2.591	-0.425	-0.063	0.253	0.616
$\tau_{\mu P}$		-3.957	-3.633	-3.354	-3.049	-0.800	-0.445	-0.131	0.231
$\tau_{\mu NL}$	0.5	-3.530	-3.188	-2.911	-2.600	-0.428	-0.063	0.252	0.622
$\tau_{\mu P}$		-3.953	-3.634	-3.367	-3.059	-0.824	-0.471	-0.168	0.179
$\tau_{\mu NL}$	0.6	-3.533	-3.192	-2.915	-2.601	-0.429	-0.066	0.250	0.623
$\tau_{\mu P}$		-3.939	-3.623	-3.356	-3.047	-0.806	-0.453	-0.145	0.220
$\tau_{\mu NL}$	0.7	-3.536	-3.191	-2.913	-2.602	-0.432	-0.067	0.251	0.627
$\tau_{\mu P}$		-3.912	-3.588	-3.307	-2.999	-0.743	-0.385	-0.062	0.324
$\tau_{\mu NL}$	0.8	-3.517	-3.180	-2.908	-2.599	-0.432	-0.064	0.249	0.627
$\tau_{\mu P}$		-3.836	-3.497	-3.216	-2.904	-0.640	-0.274	0.053	0.429
$\tau_{\mu NL}$	0.9	-3.490	-3.164	-2.899	-2.593	-0.433	-0.067	0.248	0.618
$\tau_{\mu P}$		-3.689	-3.355	-3.075	-2.759	-0.536	-0.166	0.155	0.542

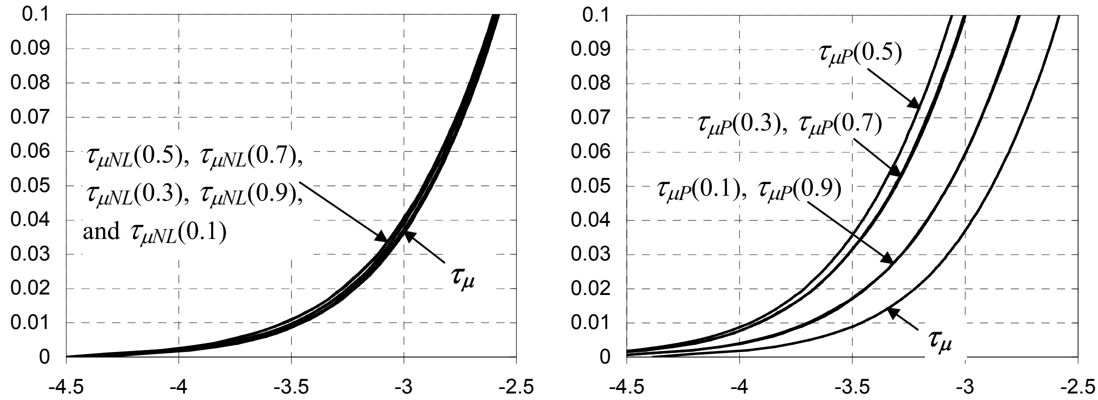
Note: MacKinnon's (1996) values of τ_{μ} are in parentheses.



$T = 50$



$T = 100$



$T = 200$

FIGURE 1. – Left-hand tails of the cumulative distribution functions of $\tau_{\mu NL}(\Theta)$ and $\tau_{\mu P}(\Theta)$.

The asymptotic distribution of $\tau_{\mu P}$, as is seen from Formula (8) in Perron (1990), depends on Θ. Three its properties are of interest: (i) the distribution is symmetric around the point $\Theta = 0.5$; (ii) when Θ tends to either 0 or 1, the distribution of $\tau_{\mu P}(\Theta)$ tends to the Dickey-Fuller distribution of τ_{μ}. Therefore, (iii) the family of distributions of $\tau_{\mu P}(\Theta)$ is bounded by the distribution of $\tau_{\mu P}(0.5)$ at the left, and by the Dickey-Fuller distribution of τ_{μ} at the right. As numerical experiments evidence, properties (ii) and (iii) hold for finite-sample distributions as well. However, property (i), the symmetry, does not. A distribution for smaller of Θ and $1 - \Theta$ is shifted to the left; and the farther Θ from 0.5, the wider the gap between the distributions of $\tau_{\mu P}(\Theta)$ and $\tau_{\mu P}(1 - \Theta)$. Nonetheless, if the sample is not too small, the difference is minor and may be neglected (say, for samples of size 100 and more); with $T = 200$, test sizes for Θ and $1 - \Theta$ practically coincide. The spread between the extreme critical values, $\tau_{\mu P}(0.5)$ and τ_{μ}, is near-constant across T (with $T \in [50, 200]$); for the 10% size of the test, it equals approximately 0.49.

Perron (1990) reports finite-sample critical values of $\tau_{\mu P}$ for $T = 50$, 100, and 200 in his Table 4. The values presented in Table I are in good agreement with them. Since he has used 5,000 replications, the data from Table I can be regarded as more accurate estimates of this Perron statistic, being based on 200,000 replications. Besides, it contains some new information regarding $\tau_{\mu P}$: evidence of splitting the finite-sample distributions for Θ and $1 - \Theta$.

Although having some similar features, the behavior of the distribution of $\tau_{\mu NL}$ is sufficiently different. (While exploring behavior of the distributions under consideration, many simulations have been performed for various Θ and T in addition to those presented in this paper. A number of conclusions are based not only on the reported simulations, but also on these additional, not reported ones.) In finite samples, the distribution also depends on Θ; but it seems to not possess even approximate symmetry around $\Theta = 0.5$ or some other point. The family of distributions of $\tau_{\mu NL}(\Theta)$ lies between the distributions of $\tau_{\mu NL}(\Theta_L)$ and $\tau_{\mu NL}(\Theta_R)$ that bound it at the left and right, respectively. However, this differs from property (iii). The value of Θ_L is approximately 0.5 to 0.6. It is impossible to give more exact figure, since, for example, the distributions for $\Theta = 0.5$ and $\Theta = 0.6$ are very close and can be hardly distinguished from one another. The value of Θ_R is roughly 0.1 to 0.15; and the relevant distribution lies a bit to the right of the Dickey-Fuller finite-sample distribution. At the same time, property (ii) still holds. Thus, the distribution of $\tau_{\mu NL}(\Theta)$ behaves somehow strangely when Θ changes from Θ_L to 0: it is shifting to the right up to the distribution of $\tau_{\mu NL}(\Theta_R)$; and then, while Θ decreases further from Θ_R to 0, it is moving backwards to the Dickey-Fuller distribution.

In fact, property (ii) is not an asymptotic one. Given $\Theta = 1$, $B_{\Theta t}$ is identically zero for all $t \leq T$. Therefore both (2) and (6) degenerate to the Dickey-Fuller test equation with constant, $\Delta y_t = \alpha + \lambda y_{t-1} + \varepsilon_t$. Hence, it is of no surprise that that the distributions of $\tau_{\mu P}$ and $\tau_{\mu NL}$ tend to that of τ_{μ} when $\Theta \to 1$. Dealing with $\Theta = 0$, there is a small friction in that $B_{\Theta t} = 1$ for $1 \leq t \leq T$, but $B_{\Theta t} = 0$ for $t = 0$. For the most part, the case of $t = 1$ can be uniformed, incorporating the "nuisance term" of the equation at $t = 1$ into the initial condition. However, for the sake of simplicity – at a minor sacrifice of rigorism – $B_{\Theta t}$ is taken to be identically 1 for all $0 \leq t \leq T$ if $\Theta = 0$. Then (6) becomes $\Delta y_t = (\alpha + \gamma - \delta) + \lambda y_{t-1} + \varepsilon_t$, i.e., again the Dickey-Fuller test equation with constant; similarly, (2) turns into $\Delta y_t = -\lambda\mu_1 + \lambda y_{t-1} + \varepsilon_t$. Thus, it is to be expected that when $\Theta \to 0$, the distributions of $\tau_{\mu P}$ and $\tau_{\mu NL}$ also tend to that of τ_{μ}., as in the case of $\Theta \to 1$.

Worthy of mention is the fact that the critical values of the nonlinear test for a given size of the test are smaller – in absolute value – than the Perron critical values. And so, one could expect the nonlinear test to have more power. One more feature is much smaller difference in $\tau_{\mu NL}$ across adjacent Θs than the difference in $\tau_{\mu P}$. This implies that using critical values for a tabulated Θ instead of an actual one (say, for $\Theta = 0.3$ rater than for 0.35) yields much smaller distortions of p-values in the nonlinear test than while applying the linear one.

But the most remarkable feature of the distribution of $\tau_{\mu NL}$ is its behavior with increasing sample size. The spread between extreme critical values, $\tau_{\mu NL}(\Theta_L)$ and $\tau_{\mu NL}(\Theta_R)$, rapidly diminishes: for the 10% size of the test, it is equal to 0.16 with $T = 50$, to 0.07 with $T = 100$, to 0.04 with $T = 150$, and to 0.03 with $T = 200$. Moreover, both $\tau_{\mu NL}(\Theta_L)$ and $\tau_{\mu NL}(\Theta_R)$ tend to τ_μ when T rises. Hence, if a sample is not too small, one could use Dickey-Fuller's critical values to test model (2) for a unit root. In doing so, e.g., with T = 100, the p-value would be understated, at the worst, by about 1.5 percent points.

The (quasi-)asymptotic critical values of $\tau_{\mu NL}$ and $\tau_{\mu P}$ are tabulated in Table II (regarding $\tau_{\mu P}$, their values are close to those reported in Table 4 of Perron, 1990); entire distributions (for selected Θ) are plotted in Figure 2. The spread between $\tau_{\mu P}(0.5)$ and τ_μ diminishes only slightly, to 0.48 for the 10% size of the test as compared to 0.49 with T = 50. As for $\tau_{\mu NL}(\Theta_L)$ and $\tau_{\mu NL}(\Theta_R)$, the spread between them vanishes at all, both bounds coinciding with τ_μ. Thus, the asymptotic distribution of $\tau_{\mu NL}$ is the Dickey-Fuller distribution of τ_μ, not depending on Θ.

TABLE II.

PERCENTAGE POINTS OF THE QUASI-ASYMPTOTIC DISTRIBUTIONS OF $\tau_{\mu NL}$ AND $\tau_{\mu P}$

Statistic	Θ	Percentage points							
		1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
τ_μ		-3.434	-3.131	-2.871	-2.572	-0.441	-0.079	0.236	0.598
		(-3.430)	(-3.122)	(-2.861)	(-2.567)	(-0.440)	(-0.078)	(0.238)	(0.607)
$\tau_{\mu NL}$	0.1	-3.435	-3.132	-2.872	-2.572	-0.441	-0.079	0.237	0.599
$\tau_{\mu P}$		-3.670	-3.348	-3.075	-2.766	-0.548	-0.184	0.139	0.517
$\tau_{\mu NL}$	0.2	-3.438	-3.132	-2.871	-2.573	-0.441	-0.079	0.237	0.598
$\tau_{\mu P}$		-3.789	-3.474	-3.209	-2.898	-0.662	-0.296	0.029	0.401
$\tau_{\mu NL}$	0.3	-3.434	-3.132	-2.870	-2.572	-0.441	-0.078	0.237	0.598
$\tau_{\mu P}$		-3.866	-3.555	-3.296	-2.991	-0.758	-0.395	-0.077	0.286
$\tau_{\mu NL}$	0.4	-3.437	-3.133	-2.873	-2.572	-0.441	-0.079	0.236	0.598
$\tau_{\mu P}$		-3.895	-3.595	-3.332	-3.037	-0.828	-0.478	-0.168	0.197
$\tau_{\mu NL}$	0.5	-3.436	-3.132	-2.871	-2.573	-0.441	-0.079	0.236	0.598
$\tau_{\mu P}$		-3.910	-3.607	-3.345	-3.051	-0.842	-0.501	-0.191	0.173
$\tau_{\mu NL}$	0.6	-3.436	-3.133	-2.871	-2.572	-0.441	-0.079	0.235	0.596
$\tau_{\mu P}$		-3.895	-3.590	-3.331	-3.032	-0.827	-0.471	-0.163	0.199
$\tau_{\mu NL}$	0.7	-3.434	-3.133	-2.871	-2.573	-0.441	-0.078	0.237	0.596
$\tau_{\mu P}$		-3.867	-3.553	-3.289	-2.991	-0.760	-0.393	-0.077	0.297
$\tau_{\mu NL}$	0.8	-3.435	-3.131	-2.873	-2.572	-0.441	-0.080	0.236	0.598
$\tau_{\mu P}$		-3.794	-3.478	-3.209	-2.898	-0.662	-0.286	0.036	0.410
$\tau_{\mu NL}$	0.9	-3.436	-3.132	-2.872	-2.572	-0.441	-0.078	0.237	0.598
$\tau_{\mu P}$		-3.662	-3.345	-3.074	-2.763	-0.548	-0.181	0.140	0.510

Note: Asymptotic MacKinnon's (1996) values of τ_μ are in parentheses; they differ from the finite-sample values for T=10,000 by less than 0.001.

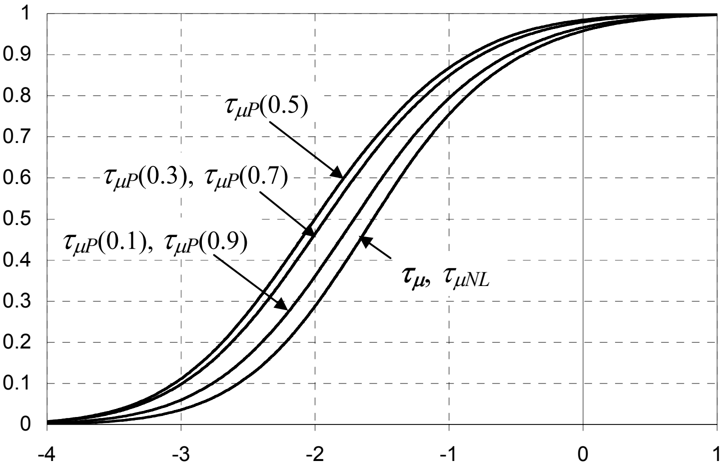


FIGURE 2. – Quasi-asymptotic cumulative distribution functions of $\tau_{\mu NL}$ and $\tau_{\mu P}(\Theta)$.

3.2 *Statistics for Regressions with no Constant Term*

The finite-sample simulation results for the unit root test statistics associated with equations (3) and (7) are reported in Table III. Figure 3 demonstrates the 10% tails of selected distributions.

This time, the distribution of the Perron statistic has no symmetry about any Θ even in the asymptotic case; at a given p-value, τ_{0P} changes monotonically over Θ. The family of distributions of $\tau_{0P}(\Theta)$ is bounded by the distribution of $\tau_{0P}(0)$ at the left and by the distribution of $\tau_{0P}(1)$ at the right. As easily seen, these are the Dickey-Fuller distributions of τ_{μ} and τ_0, respectively. Indeed, given $\Theta = 1$, (7) degenerates to the Dickey-Fuller test equation without constant, $\Delta y_t = \lambda y_{t\text{-}1} + \varepsilon_t$. Provided that $\Theta = 0$, (7) becomes the Dickey-Fuller test equation with constant, $\Delta y_t = (\gamma - \delta) + \lambda y_{t\text{-}1} + \varepsilon_t$. Thus, the spread between the extreme critical values of the τ_{0P} statistic (across Θ) is that between τ_{μ} and τ_0. It is approximately twice as large as in the case of $\tau_{\mu P}$. For example, with the 10% size of the test, the spread is 0.986 if $T = 50$, and 0.957 if $T = 200$ (the asymptotic value is 0.950).

TABLE III.

PERCENTAGE POINTS OF THE DISTRIBUTIONS OF τ_{0NL} AND τ_{0P} STATISTICS

Statistic	Θ	Percentage points							
		1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
					T =50				
τ_0		-2.619	-2.259	-1.956	-1.616	0.900	1.309	1.665	2.086
		(-2.612)	(-2.249)	(-1.948)	(-1.613)	(0.906)	(1.309)	(1.660)	(2.073)
τ_{0NL}	0.1	-3.194	-2.778	-2.402	-1.946	0.918	1.320	1.671	2.073
τ_{0P}		-3.540	-3.180	-2.885	-2.562	-0.364	-0.002	0.314	0.688
τ_{0NL}	0.2	-3.172	-2.744	-2.361	-1.901	0.916	1.328	1.667	2.087
τ_{0P}		-3.500	-3.145	-2.848	-2.514	-0.171	0.193	0.514	0.900
τ_{0NL}	0.3	-3.149	-2.716	-2.328	-1.866	0.919	1.326	1.673	2.091
τ_{0P}		-3.459	-3.093	-2.797	-2.457	0.052	0.449	0.794	1.210
τ_{0NL}	0.4	-3.099	-2.670	-2.279	-1.819	0.920	1.325	1.678	2.097
τ_{0P}		-3.413	-3.055	-2.746	-2.392	0.267	0.683	1.044	1.470
τ_{0NL}	0.5	-3.060	-2.616	-2.228	-1.776	0.918	1.330	1.687	2.099
τ_{0P}		-3.348	-2.983	-2.663	-2.313	0.438	0.865	1.224	1.663
τ_{0NL}	0.6	-2.999	-2.550	-2.170	-1.739	0.922	1.335	1.694	2.116
τ_{0P}		-3.277	-2.895	-2.573	-2.209	0.592	1.013	1.369	1.794
τ_{0NL}	0.7	-2.917	-2.478	-2.109	-1.703	0.921	1.336	1.696	2.129
τ_{0P}		-3.165	-2.785	-2.448	-2.077	0.709	1.122	1.488	1.903
τ_{0NL}	0.8	-2.831	-2.402	-2.052	-1.672	0.917	1.331	1.696	2.117
τ_{0P}		-3.034	-2.623	-2.295	-1.919	0.793	1.207	1.565	1.988
τ_{0NL}	0.9	-2.703	-2.317	-2.001	-1.644	0.913	1.323	1.686	2.116
τ_{0P}		-2.789	-2.413	-2.096	-1.732	0.866	1.276	1.636	2.060
					T =100				
τ_0		-2.582	-2.240	-1.948	-1.613	0.893	1.297	1.638	2.038
		(-2.588)	(-2.238)	(-1.944)	(-1.615)	(0.897)	(1.296)	(1.641)	(2.043)
τ_{0NL}	0.1	-2.985	-2.546	-2.182	-1.744	0.900	1.303	1.644	2.041
τ_{0P}		-3.466	-3.133	-2.854	-2.544	-0.372	-0.017	0.297	0.669
τ_{0NL}	0.2	-2.947	-2.519	-2.150	-1.725	0.902	1.306	1.644	2.042
τ_{0P}		-3.430	-3.097	-2.814	-2.497	-0.190	0.175	0.513	0.896
τ_{0NL}	0.3	-2.918	-2.486	-2.118	-1.704	0.903	1.305	1.648	2.046
τ_{0P}		-3.404	-3.056	-2.771	-2.445	0.038	0.435	0.782	1.186
τ_{0NL}	0.4	-2.880	-2.455	-2.089	-1.687	0.904	1.306	1.651	2.053
τ_{0P}		-3.359	-3.012	-2.722	-2.387	0.245	0.656	1.014	1.424
τ_{0NL}	0.5	-2.829	-2.409	-2.058	-1.671	0.901	1.308	1.653	2.052
τ_{0P}		-3.286	-2.952	-2.652	-2.304	0.422	0.842	1.207	1.635
τ_{0NL}	0.6	-2.790	-2.372	-2.039	-1.662	0.904	1.306	1.654	2.062
τ_{0P}		-3.225	-2.869	-2.559	-2.207	0.568	0.992	1.353	1.768
τ_{0NL}	0.7	-2.738	-2.329	-2.007	-1.648	0.901	1.309	1.658	2.065
τ_{0P}		-3.133	-2.764	-2.446	-2.081	0.683	1.099	1.465	1.871
τ_{0NL}	0.8	-2.676	-2.299	-1.988	-1.637	0.899	1.309	1.652	2.056
τ_{0P}		-2.988	-2.613	-2.297	-1.935	0.779	1.186	1.540	1.949
τ_{0NL}	0.9	-2.629	-2.272	-1.969	-1.623	0.898	1.305	1.660	2.058
τ_{0P}		-2.807	-2.426	-2.112	-1.757	0.847	1.245	1.590	1.996

Table III (continued)

					T=150				
τ_0		-2.570	-2.229	-1.938	-1.614	0.904	1.300	1.648	2.041
		(-2.581)	(-2.235)	(-1.943)	(-1.615)	(0.894)	(1.292)	(1.635)	(2.034)
τ_{0NL}	0.1	-2.861	-2.439	-2.079	-1.685	0.909	1.309	1.653	2.040
τ_{0P}		-3.430	-3.107	-2.838	-2.534	-0.373	-0.015	0.296	0.655
τ_{0NL}	0.2	-2.823	-2.412	-2.056	-1.672	0.909	1.306	1.651	2.037
τ_{0P}		-3.407	-3.080	-2.801	-2.487	-0.188	0.177	0.496	0.874
τ_{0NL}	0.3	-2.794	-2.384	-2.039	-1.667	0.909	1.302	1.648	2.048
τ_{0P}		-3.362	-3.038	-2.759	-2.441	0.033	0.424	0.773	1.183
τ_{0NL}	0.4	-2.764	-2.350	-2.014	-1.655	0.907	1.306	1.653	2.049
τ_{0P}		-3.331	-2.994	-2.705	-2.380	0.236	0.651	1.004	1.430
τ_{0NL}	0.5	-2.737	-2.338	-2.009	-1.648	0.907	1.309	1.654	2.048
τ_{0P}		-3.289	-2.945	-2.645	-2.304	0.415	0.837	1.206	1.621
τ_{0NL}	0.6	-2.698	-2.312	-1.989	-1.642	0.908	1.306	1.656	2.055
τ_{0P}		-3.199	-2.855	-2.558	-2.204	0.558	0.980	1.341	1.752
τ_{0NL}	0.7	-2.666	-2.289	-1.973	-1.634	0.910	1.311	1.656	2.060
τ_{0P}		-3.116	-2.751	-2.442	-2.082	0.677	1.099	1.457	1.877
τ_{0NL}	0.8	-2.627	-2.265	-1.963	-1.630	0.908	1.307	1.660	2.058
τ_{0P}		-2.977	-2.612	-2.297	-1.936	0.767	1.176	1.530	1.927
τ_{0NL}	0.9	-2.598	-2.243	-1.950	-1.621	0.907	1.308	1.651	2.048
τ_{0P}		-2.801	-2.434	-2.118	-1.767	0.848	1.255	1.598	1.990
					T=200				
τ_0		-2.572	-2.226	-1.939	-1.612	0.899	1.295	1.641	2.037
		(-2.577)	(-2.233)	(-1.942)	(-1.616)	(0.892)	(1.290)	(1.632)	(2.029)
τ_{0NL}	0.1	-2.787	-2.377	-2.033	-1.663	0.904	1.297	1.642	2.035
τ_{0P}		-3.440	-3.112	-2.839	-2.537	-0.377	-0.015	0.293	0.663
τ_{0NL}	0.2	-2.765	-2.347	-2.010	-1.657	0.904	1.300	1.638	2.040
τ_{0P}		-3.396	-3.072	-2.801	-2.488	-0.189	0.175	0.500	0.881
τ_{0NL}	0.3	-2.733	-2.336	-2.002	-1.648	0.904	1.300	1.642	2.038
τ_{0P}		-3.367	-3.033	-2.760	-2.440	0.033	0.431	0.770	1.172
τ_{0NL}	0.4	-2.713	-2.317	-1.994	-1.643	0.904	1.300	1.644	2.040
τ_{0P}		-3.325	-2.993	-2.708	-2.379	0.239	0.652	1.014	1.416
τ_{0NL}	0.5	-2.681	-2.299	-1.984	-1.640	0.905	1.301	1.646	2.041
τ_{0P}		-3.271	-2.924	-2.641	-2.303	0.411	0.827	1.192	1.611
τ_{0NL}	0.6	-2.657	-2.285	-1.977	-1.636	0.906	1.299	1.651	2.045
τ_{0P}		-3.201	-2.856	-2.559	-2.209	0.553	0.971	1.330	1.741
τ_{0NL}	0.7	-2.638	-2.266	-1.962	-1.627	0.904	1.301	1.646	2.041
τ_{0P}		-3.116	-2.758	-2.451	-2.086	0.671	1.081	1.430	1.839
τ_{0NL}	0.8	-2.621	-2.254	-1.954	-1.622	0.901	1.302	1.649	2.041
τ_{0P}		-2.996	-2.625	-2.304	-1.941	0.759	1.174	1.528	1.926
τ_{0NL}	0.9	-2.593	-2.237	-1.947	-1.620	0.901	1.296	1.642	2.048
τ_{0P}		-2.800	-2.436	-2.120	-1.775	0.835	1.238	1.595	2.002

Note: MacKinnon's (1996) values of τ_0 are in parentheses.

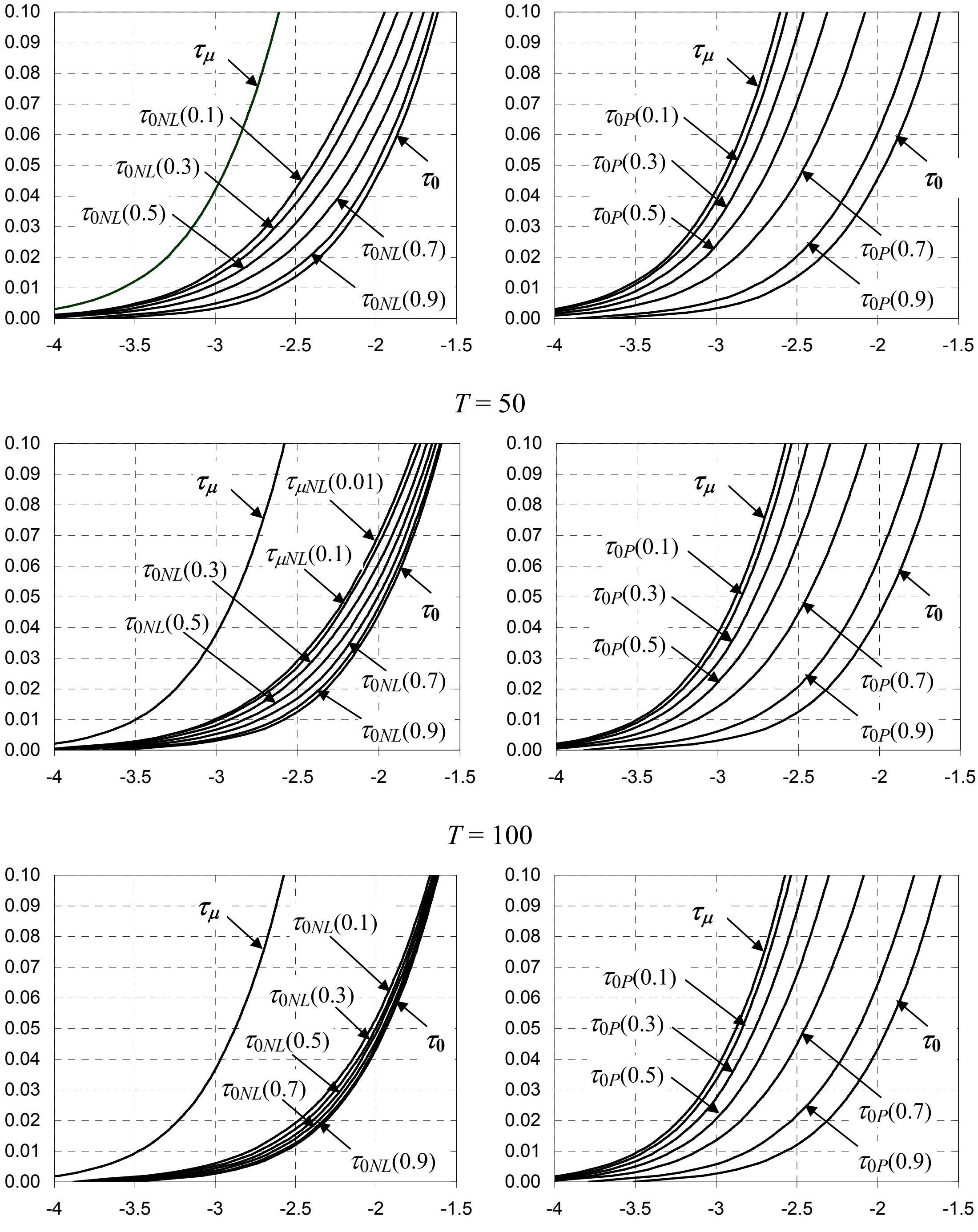


FIGURE 3. – Left-hand tails of the cumulative distribution functions of $\tau_{0NL}(\Theta)$ and $\tau_{0P}(\Theta)$.

The behavior of the distribution of the nonlinear-test statistic is similar in that τ_{0NL} changes monotonically over Θ (at a given test size), and that the family of distributions of $\tau_{0NL}(\Theta)$ is bounded by the distribution of $\tau_{0NL}(1) = \tau_0$ at the right.

A striking feature is that the left-hand bound is not the Dickey-Fuller distribution of τ_μ. Seemingly, it would be, since (3) becomes $\Delta y_t = -\lambda\gamma + \lambda y_{t-1} + \varepsilon_t$ with $\Theta = 0$ (and thus, $\tau_{0NL}(0) = \tau_\mu$). But as seen from Figure 3, the distributions of $\tau_{0NL}(0.1)$ for each sample size are much nearer to the distributions of τ_0 than to those of τ_μ. Further diminishing Θ does not change the pattern; distributions of $\tau_{0NL}(1/T)$ are rather close to the distributions of $\tau_{0NL}(0.1)$. As an example, a non-tabulated distribution of $\tau_{0NL}(0.01)$ is plotted in Figure 3 for $T = 100$. Apparently, the distribution of $\tau_{0NL}(\Theta)$ as a function of Θ is discontinuous, having a jump at $\Theta = 0$. Thus, the left bound of the family of distributions of $\tau_{0NL}(\Theta)$ is some limit of the distribution of $\tau_{0NL}(\Theta)$ with $\Theta \to 0$, the critical values of which for a given size of the test are much smaller (in absolute value) than τ_μ. As a result, the spread between extreme critical values, for $\Theta = 1/T$ and $\Theta = 1$, of the nonlinear test is much smaller than that for the case of τ_{0P}.

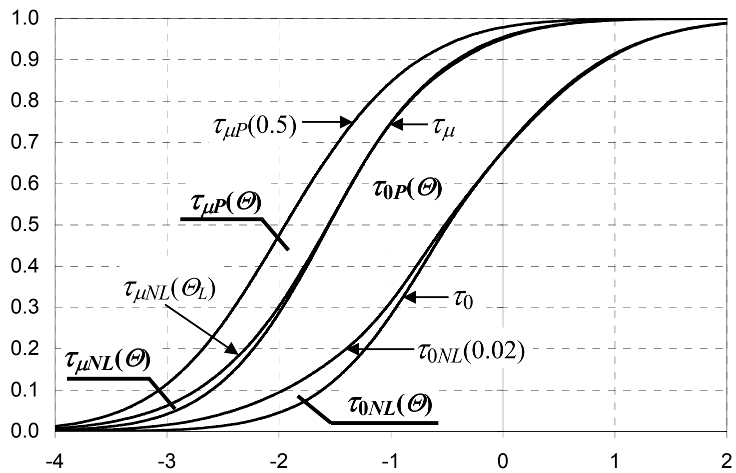


FIGURE 4. – Areas of the cumulative distribution functions of $\tau_{0NL}(\Theta)$ and $\tau_{0P}(\Theta)$, and $\tau_{\mu NL}(\Theta)$ and $\tau_{\mu P}(\Theta)$; $T = 50$.

To give a general idea of the difference between the families of distributions of the Perron and nonlinear-test statistics, Figure 4 demonstrates the areas of these families for samples of size 50; the cases of regressions with and without constant term are combined in the figure. The distributions of $\tau_{0P}(\Theta)$ fill the whole area between lines τ_0 and τ_μ, while the distributions of $\tau_{0NL}(\Theta)$ fill only a small portion of this area in its right-hand part; at p-values more than approximately 0.5, the distributions of $\tau_{0NL}(\Theta)$ for all Θ practically coincide with the Dickey-Fuller distribution of τ_0. (For $\tau_{\mu NL}(\Theta)$ and $\tau_{\mu P}(\Theta)$, the pattern is qualitatively similar.)

TABLE IV.

PERCENTAGE POINTS OF THE QUASI-ASYMPTOTIC DISTRIBUTIONS OF τ_{0NL} AND τ_{0P}

Statistic	Θ	Percentage points							
		1.0%	2.5%	5.0%	10.0%	90.0%	95.0%	97.5%	99.0%
τ_0		-2.572	-2.230	-1.950	-1.623	0.889	1.279	1.622	2.009
		(-2.565)	(-2.227)	(-1.941)	(-1.617)	(0.888)	(1.284)	(1.624)	(2.015)
τ_{0NL}	0.1	-2.578	-2.230	-1.951	-1.624	0.889	1.280	1.621	2.009
τ_{0P}		-3.388	-3.094	-2.826	-2.529	-0.389	-0.036	0.264	0.641
τ_{0NL}	0.2	-2.578	-2.232	-1.952	-1.625	0.890	1.279	1.621	2.009
τ_{0P}		-3.370	-3.063	-2.791	-2.488	-0.211	0.157	0.476	0.843
τ_{0NL}	0.3	-2.573	-2.232	-1.948	-1.624	0.889	1.279	1.622	2.008
τ_{0P}		-3.342	-3.024	-2.755	-2.440	0.010	0.404	0.752	1.163
τ_{0NL}	0.4	-2.573	-2.232	-1.950	-1.624	0.889	1.279	1.621	2.008
τ_{0P}		-3.307	-2.980	-2.704	-2.384	0.215	0.624	0.980	1.391
τ_{0NL}	0.5	-2.575	-2.230	-1.949	-1.624	0.889	1.278	1.622	2.008
τ_{0P}		-3.253	-2.928	-2.641	-2.307	0.393	0.819	1.177	1.596
τ_{0NL}	0.6	-2.574	-2.232	-1.950	-1.624	0.889	1.279	1.622	2.009
τ_{0P}		-3.188	-2.860	-2.560	-2.212	0.540	0.955	1.321	1.735
τ_{0NL}	0.7	-2.574	-2.230	-1.950	-1.623	0.888	1.280	1.622	2.009
τ_{0P}		-3.117	-2.758	-2.452	-2.100	0.659	1.072	1.428	1.832
τ_{0NL}	0.8	-2.575	-2.230	-1.950	-1.624	0.889	1.278	1.620	2.011
τ_{0P}		-2.981	-2.623	-2.315	-1.960	0.756	1.162	1.505	1.892
τ_{0NL}	0.9	-2.572	-2.230	-1.949	-1.624	0.889	1.279	1.622	2.010
τ_{0P}		-2.813	-2.452	-2.148	-1.798	0.828	1.228	1.572	1.963

Note: Asymptotic MacKinnon's (1996) values of τ_0 are in parentheses; they differ from the finite-sample values for T=10,000 by less than 0.001.

The spread between extreme critical values, $\tau_{0NL}(1/T)$ and $\tau_{0NL}(1)$, rapidly diminishes with increasing sample size. For the 10% size of the test, the spread is equal to 0.35 with $T = 50$, to 0.15 with $T = 100$, to 0.09 with $T = 150$, and to 0.06 with $T = 200$. As the right-hand bound of the distribution of $\tau_{0NL}(\Theta)$ is always the distribution of τ_0, this implies that $\tau_{0NL}(1/T)$ tends to τ_0 when T rises (hence, so do $\tau_{0NL}(\Theta)$ for all Θ). Then the left and right bounds of the family of distributions of $\tau_{0NL}(\Theta)$ should eventually join. The quasi-asymptotic simulations reported in Table IV and Figure 5 suggest that to be the case. Thus, the asymptotic distribution of τ_{0NL} – not depending on Θ – is the Dickey-Fuller distribution of τ_0.

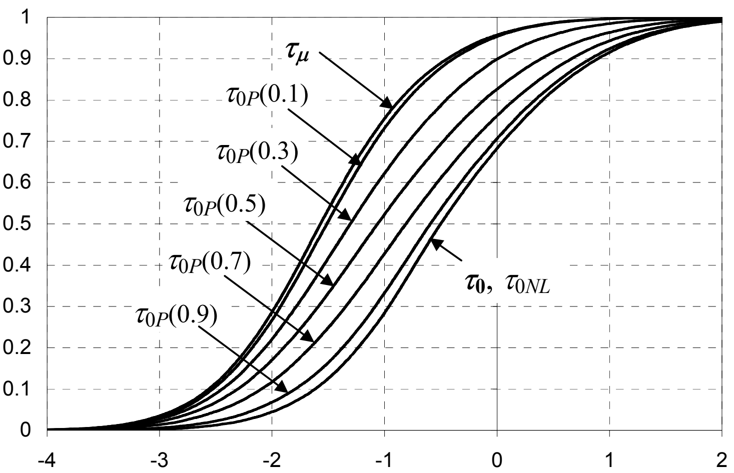


FIGURE 5. – Quasi-asymptotic cumulative distribution functions of τ_{0NL} and $\tau_{0P}(\Theta)$.

A well-known unpleasant feature of the unit-root τ-statistics for the case of no constant is their dependence on the initial condition in finite samples. The first mention of this fact seems to date back to Dickey and Fuller (1979). Evans and Savin (1981) studied the effect of the initial condition in detail, finding the distribution of $\hat{\lambda}+1$ to be affected by the value of y_0/σ; and Phillips (1987) explained their results theoretically. Recently, Müller and Elliott (2003) have analyzed the role of the initial condition in testing for unit roots, including the Dickey-Fuller τ-tests. Based on this literature, it can be expected that the statistics τ_{0NL} and τ_{0P} are also affected by the initial condition.

Monte Carlo experiments corroborate this expectation. Figure 6 demonstrates how the distributions of τ_{0NL} and τ_{0P} depend on the initial condition for the cases of fixed and random y_0. Since $\{\varepsilon_t\}$ are drawn from $N(0,1)$, y_0 and ζ can be deemed as measured in the units of σ, the standard deviation of the innovations.

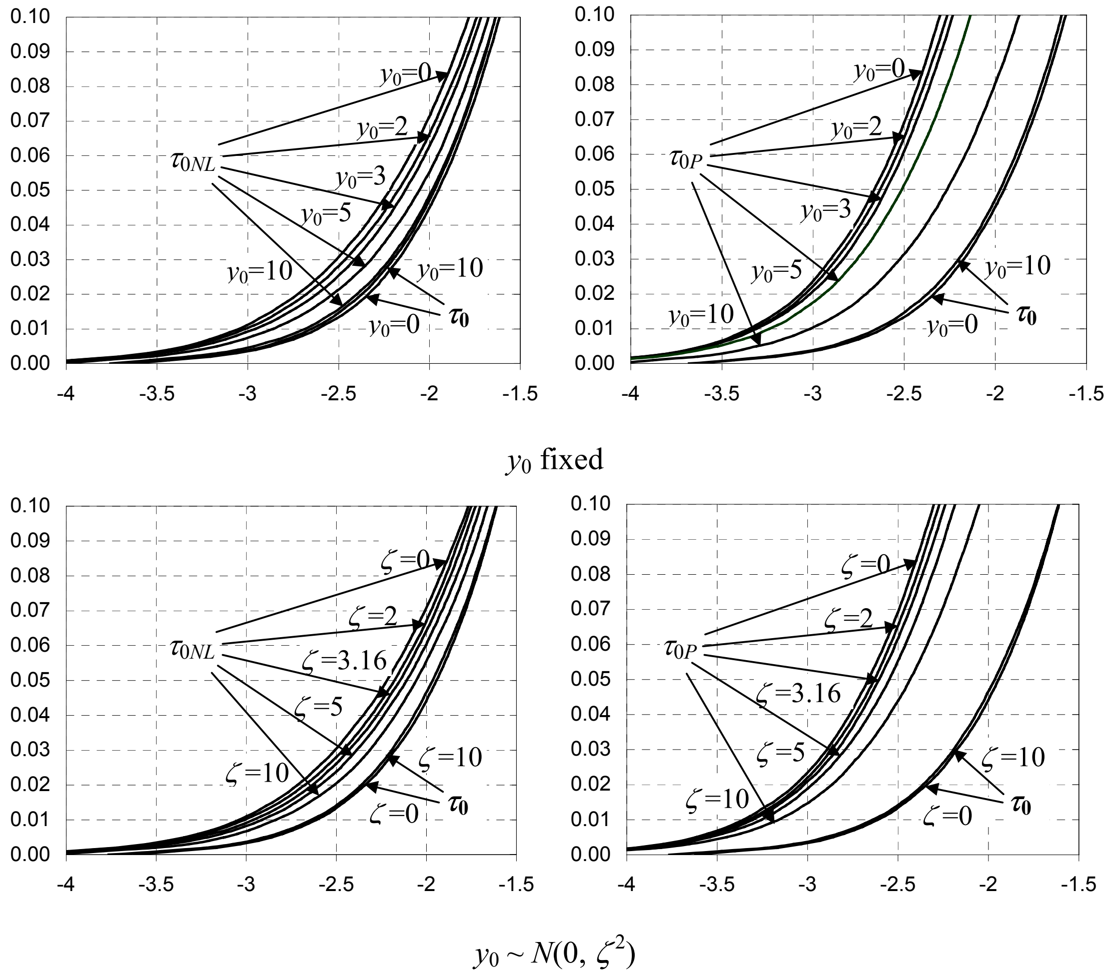


FIGURE 6. – Effect of the initial condition on distributions of τ_{0NL} and τ_{0P}; $T = 50$; $\Theta = 0.5$.

At low significance levels, 20% and smaller, the Dickey-Fuller distribution of τ_0 depends only slightly on the initial condition. Even with $y_0 = 10$, or equal standard deviation of zero-mean random y_0, the changes in the Dickey-Fuller critical values are so minor that can be neglected. By contrast, both τ_{0NL} and τ_{0P} are rather sensitive to altering the initial condition. Size distortions – below the nominal levels – are symmetric around $y_0 = 0$ and become pronounced (at samples of size 50) beginning with values of

$|y_0|$ equaling about three standard deviations of the errors. With $y_0 = 10$, the distribution of τ_{0NL} turns out to be very close to the Dickey-Fuller distribution of τ_0. (In fact, tracing its behavior for further significance levels, it is close to the distribution of τ_0 for $y_0 = 10$.) As seen from Figure 6, the impact of the standard deviation of random y_0 is qualitatively similar, although quantitatively it is somehow weaker. Thus, when the initial value is random and has a non-zero mean, the size of the tests will be affected by a combined impact of the (average) magnitude of the initial value and its variance.

3.3 *Reversed Breaks*

In empirical applications, it is sometimes convenient to characterize a process as that having a reversed break, that is, $B'_{\theta t} = 1$ if $t \leq \theta$ and 0 otherwise, or $B'_{\theta t} = 1 - B_{\theta t}$. Let the relevant τ-statistics be designated as $\tau'_{\mu NL}$, τ'_{0NL}, $\tau'_{\mu P}$, and τ'_{0P}.

It is intuitively clear that when the case at hands is a switching between two arbitrary levels μ_0 and μ_1, it does not matter which of them is taken as a base one; and so, the tests for a unit root would be invariant to such a choice. Indeed, replacing $B_{\theta t}$ by $1 - B_{\theta t}$ in (2), we have $\Delta y_t = (\alpha - \lambda\gamma) + \lambda y_{t-1} - \gamma B_{\theta t} + \gamma(\lambda+1)B_{\theta,t-1} + \varepsilon_t$, which is equivalent (from the viewpoint of its structure) to the original equation. (Note that $\alpha - \lambda\gamma = -\lambda\mu_1$.) Similarly, (6) with the reversed break comes to $\Delta y_t = (\alpha + \gamma - \delta) + \lambda y_{t-1} - \gamma B_{\theta t} + \delta B_{\theta,t-1} + \varepsilon_t$. Hence, the distributions of $\tau'_{\mu NL}$ and $\tau'_{\mu P}$ coincide with the distributions of $\tau_{\mu NL}$ and $\tau_{\mu P}$, respectively.

The pattern is different if there is the singled-out base level $\mu = 0$, i.e., when model (3) or (7) is dealt with. Expressed in terms of $B_{\theta t}$, these models with the reversed break come to forms similar to the respective models with intercept term, (2) and (6), however, a restriction being imposed on the intercept. For (3), we have $\Delta y_t = \lambda y_{t-1} + \gamma B'_{\theta t} - \gamma(\lambda+1)B'_{\theta,t-1} + \varepsilon_t = -\lambda\gamma + \lambda y_{t-1} - \gamma B_{\theta t} + \gamma(\lambda+1)B_{\theta,t-1} + \varepsilon_t$, which up to signs is (2) but the nonlinear restriction $\alpha = -\lambda\gamma$. Such a restriction turns out to be linear, $\alpha = \gamma - \delta$, for Perron-type model (7) with the reversed break: $\Delta y_t = \lambda y_{t-1} + \gamma B'_{\theta t} - \delta B'_{\theta,t-1} + \varepsilon_t = (\gamma - \delta) + \lambda y_{t-1} - \gamma B_{\theta t} + \delta B_{\theta,t-1} + \varepsilon_t$. As a result, the distributions of τ'_{0NL} and τ'_{0P} differ from the distributions of τ_{0NL} and τ_{0P}. Their left-hand tails (for $T = 50$) are plotted in Figure 7.

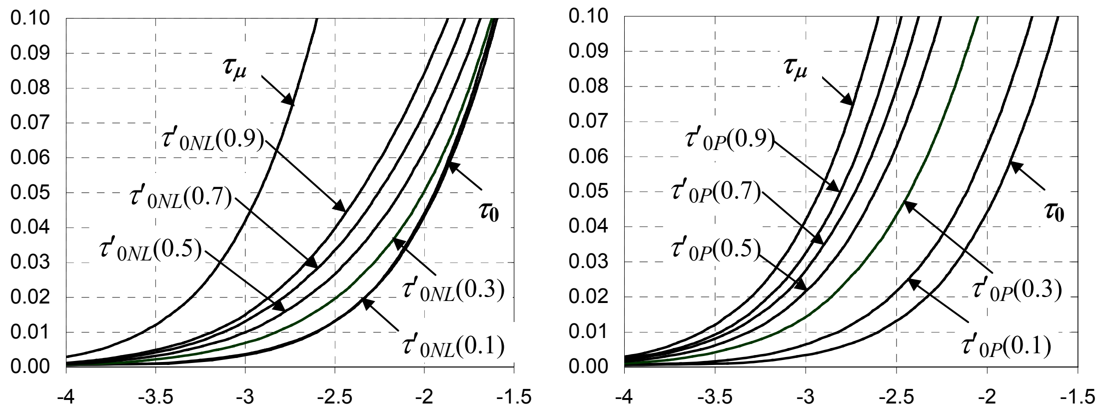


FIGURE 7. – Left-hand tails of the cumulative distribution functions of $\tau'_{0NL}(\Theta)$ and $\tau'_{0P}(\Theta)$; $T = 50$.

Qualitatively, the pattern resembles that of τ_{0NL} and τ_{0P} with Θ changed to $1 - \Theta$. It may appear from this that the distributions of $\tau'_{0NL}(\Theta)$ and $\tau'_{0P}(\Theta)$ are the same as the distributions of $\tau_{0NL}(1 - \Theta)$ and $\tau_{0P}(1 - \Theta)$. However, that is not the case. The distributions for the reversed breaks are shifted a bit towards the distribution of τ_0. For example, the distribution of $\tau'_{0NL}(0.1)$ almost coincides with it, while the distribution of $\tau_{0NL}(0.9)$ does not. At the 10% size of the test and $T = 50$, $\tau'_{0NL}(0.98) = -1.892$ and $\tau_{0NL}(0.02) = -1.957$. For the Perron statistic, the shift increases with Θ: given the 10% size of the test and $T = 50$, $\tau'_{0P}(0.9) = -2.473$ and $\tau_{0P}(0.1) = -2.562$, while $\tau'_{0P}(0.1)$ and $\tau_{0P}(0.9)$ are close to one another. Hence it follows that when reversed breaks are dealt with and the model does not include the constant term, one needs critical values differing from those tabulated in Table III. To save space, these values are not reported.

4. POWER SIMULATIONS

To explore and compare power properties of the tests, the data generating process is based on (1). Specifically, the series are generated as $y_t = \alpha + (\lambda + 1)y_{t-1} + \gamma(B_{\theta t} - (\lambda+1)B_{\theta,t-1}) + \varepsilon_t$; 200,000 replications are used with $\varepsilon_t \sim$ iid $N(0,1)$. The nominal size of the tests is 5%; the finite-sample critical values are taken from Tables I and III. The power is assessed at $\lambda = -0.2, -0.1, -0.05$, and -0.01; $\lambda = 0$ is used to evaluate the size of the tests. The constant is specified as $\alpha = 0$, 0.5, and 1 (with $\lambda = 0$, α is always zero). Table V reports results for $T = 100$, $\Theta = 0.5$, $y_0 = 0$, and $\gamma = 0$, 1, 2, 3, and 5.

TABLE V.

POWER OF SIZE 0.05 TESTS FOR T = 100 AND Θ = 0.5

	$\lambda = -0.2$			$\lambda = -0.1$			$\lambda = -0.05$			$\lambda = -0.01$			$\lambda = 0$		
Test:	τ_0	τ_{0NL}	τ_{0P}	τ_0	τ_{0NL}	τ_{0P}	τ_0	τ_{0NL}	τ_{0P}	τ_0	τ_{0NL}	τ_{0P}	τ_0	τ_{0NL}	τ_{0P}
							$\alpha = 0$								
γ=0	0.998	0.989	0.937	0.771	0.700	0.438	0.324	0.298	0.172	0.080	0.078	0.067	0.051	0.050	0.051
γ=1	0.986	0.989	0.937	0.698	0.700	0.436	0.304	0.298	0.172	0.080	0.079	0.068	0.050	0.050	0.051
γ=2	0.861	0.989	0.938	0.507	0.701	0.439	0.254	0.299	0.174	0.078	0.078	0.067	0.051	0.050	0.051
γ=3	0.473	0.989	0.936	0.286	0.701	0.440	0.189	0.298	0.174	0.076	0.078	0.069	0.051	0.050	0.052
γ=5	0.009	0.990	0.937	0.037	0.701	0.438	0.074	0.297	0.173	0.067	0.078	0.069	0.051	0.051	0.052
Test:	τ_μ	$\tau_{\mu NL}$	$\tau_{\mu P}$	τ_μ	$\tau_{\mu NL}$	$\tau_{\mu P}$	τ_μ	$\tau_{\mu NL}$	$\tau_{\mu P}$	τ_μ	$\tau_{\mu NL}$	$\tau_{\mu P}$	τ_μ	$\tau_{\mu NL}$	$\tau_{\mu P}$
							$\alpha = 0.5$								
γ=0	0.903	0.870	0.720	0.429	0.397	0.267	0.226	0.186	0.122	0.065	0.050	0.021	0.051	0.051	0.051
γ=1	0.834	0.868	0.720	0.382	0.397	0.267	0.202	0.187	0.124	0.056	0.050	0.021	0.050	0.050	0.051
γ=2	0.637	0.868	0.718	0.298	0.396	0.266	0.167	0.185	0.121	0.048	0.050	0.021	0.048	0.050	0.050
γ=3	0.367	0.869	0.719	0.206	0.397	0.265	0.130	0.185	0.122	0.038	0.050	0.020	0.046	0.051	0.051
γ=5	0.041	0.867	0.719	0.069	0.396	0.265	0.068	0.184	0.121	0.021	0.050	0.022	0.038	0.050	0.050
							$\alpha = 1$								
γ=0	0.965	0.936	0.829	0.758	0.675	0.458	0.677	0.593	0.253	0.206	0.179	0.025	0.050	0.050	0.051
γ=1	0.928	0.935	0.829	0.705	0.672	0.459	0.634	0.592	0.253	0.192	0.178	0.026	0.050	0.050	0.051
γ=2	0.811	0.936	0.828	0.620	0.674	0.458	0.581	0.591	0.253	0.177	0.179	0.026	0.049	0.052	0.051
γ=3	0.593	0.936	0.828	0.511	0.672	0.458	0.520	0.591	0.253	0.156	0.180	0.025	0.045	0.050	0.050
γ=5	0.136	0.936	0.830	0.271	0.673	0.457	0.367	0.591	0.251	0.104	0.179	0.025	0.038	0.049	0.050

As seen from Table V (as well as from Tables VI and VII below), the nonlinear test always dominates the linear one. Naturally, when there is no break in series ($\gamma = 0$), the Dickey-Fuller test proves to be more powerful than both tests allowing for break, since the respective models describe the process more adequately than models with a break. But the power of both nonlinear and linear tests is invariant to the value of γ, while the power of the Dickey-Fuller test decreases as γ rises. Thus, there is some value of γ, at which the tests allowing for break start dominating the Dickey-Fuller test. With T =100, this is a value exceeding one for the nonlinear test, or, in terms of the standard deviation of the errors, σ, $\gamma > \sigma$. For the linear test to exhibit advantage over the Dickey-Fuller test, γ should exceed at least two standard deviations of the residuals. (For example, when $\lambda = -0.05$ and $\alpha = 1$, the τ_μ-test is more powerful than the $\tau_{\mu P}$-test even with $\gamma = 5$.)

The power properties of the tests allowing for break are in certain respects akin to those of the Dickey-Fuller test. The power of the $\tau_{\mu NL}$-test and $\tau_{\mu P}$-test increases with the increase in the absolute value of intercept, like the power of the τ_μ-test. And, in common with the Dickey-Fuller test the tests allowing for break have low power in the neighborhood of the unit root. For example, if $\lambda = -0.01$, the power of all the no-intercept tests is about 7% to 8% (given $T = 100$); it is even smaller for test with intercept when $\alpha = 0.5$. (Note that in this case the $\tau_{\mu P}$-test displays bias, having power less than the size for both values of α.)

TABLE VI.

POWER OF SIZE 0.05 TESTS FOR T = 100, AND Θ = 0.1 AND 0.9

Test	$\lambda = -0.2$		$\lambda = -0.1$		$\lambda = -0.05$		$\lambda = -0.01$		$\lambda = 0$	
	$\Theta = 0.1$	$\Theta = 0.9$	$\Theta = 0.1$	$\Theta = 0.9$	$\Theta = 0.1$	$\Theta = 0.9$	$\Theta = 0.1$	$\Theta = 0.9$	$\Theta = 0.1$	$\Theta = 0.9$
τ_{0NL}	0.984	0.996	0.649	0.750	0.276	0.312	0.076	0.078	0.049	0.050
τ_{0P}	0.867	0.994	0.321	0.691	0.121	0.280	0.059	0.076	0.050	0.050
$\tau_{\mu NL}$	0.958	0.953	0.748	0.731	0.674	0.647	0.207	0.189	0.050	0.049
$\tau_{\mu P}$	0.846	0.932	0.357	0.669	0.229	0.556	0.090	0.123	0.050	0.050

Note: For τ_{0NL} and τ_{0P}, $\alpha = 0$; for $\tau_{\mu NL}$ and $\tau_{\mu P}$, $\alpha = 1$; $\gamma = 0$; $y_0 = 0$.

Reporting power estimates for $\Theta = 0.1$ and $\Theta = 0.9$, Table VI in combination with Table V illustrates the behavior of the test powers in relation to Θ. Like the null distributions, the distributions under alternatives depend on Θ. Therefore, an effect of Θ on the test powers is determined by relative position of corresponding null and alternative distributions. In brief, it results in that the power of both τ_{0NL}-test and τ_{0P}-test increases with the rise of Θ. The reason is that the width of both families of alternative distributions is rather small, and so, horizontal distances between critical values for different values of Θ at a given test size play much more important role than vertical distances between alternative distributions for different Θ.

In general, a similar reason explains the power behavior of the $\tau_{\mu NL}$-test. The power of this test has a minimum at about $\Theta = 0.5$; and it increases as Θ moves away from 0.5 in either direction. (Recall that, at a given test size, critical values of the $\tau_{\mu NL}$-test move to the left when Θ changes from 0 to about 0.5 to 0.6; and then – as Θ increases to 1 – they move to the right.) The power of the $\tau_{\mu P}$-test demonstrates an unlike pattern: it increases with increasing Θ, as in the case of the τ_{0P}-test, but for a quite different reason. The width of the family of alternative distributions is comparable with the width of the family of null distributions of $\tau_{\mu P}(\Theta)$. At the same time, there is no symmetry about $\Theta = 0.5$ in the family of alternative distributions. Although there are shifts to the right when Θ changes from about 0.7 to 1, they are small: e.g., the alternative distribution for $\Theta = 0.9$ lies between distributions for $\Theta = 0.5$ and $\Theta = 0.4$. Therefore, though critical values of the $\tau_{\mu P}$-test for, say, $\Theta = 0.1$ and $\Theta = 0.9$ are almost equal, the vertical distance between corresponding alternative distributions at a given critical value turns out to be considerable, so causing the rise in power for greater Θ.

An expected feature is the rise in power of the τ_{0NL}-test and τ_{0P}-test when y_0 deviates from zero either deterministically or randomly. The power increases with increasing (in absolute value) y_0 if it is fixed, or its mean or/and variance if y_0 is random. Along with this, in accordance with Figure 6, the non-zero initial condition causes size distortions (below the nominal level) in the τ_{0NL}-test and τ_{0P}-test that become noticeable starting with values of (fixed) y_0 as large as about three times the standard deviation of the residuals. Since the null distributions of the test statistics associated with equations including intercept, $\tau_{\mu NL}$ and $\tau_{\mu P}$, are invariant to the initial condition, its altering does not change the sizes of the relevant tests. The power behavior of these tests in relation to y_0 is somehow more involved, being similar to that of the Dickey-Fuller τ_{μ}-test. Depending on a particular combination of values of α and y_0, the power can either rise or fall.

Following Evans and Savin (1984), the initial condition can be incorporated into the intercept. Subtracting y_0 from both sides of (2) or (6) gives the transformed model $\Delta z_t = (\alpha + \lambda y_0) + \lambda z_{t\text{-}1} + \gamma B_{\Theta t} - \delta B_{\Theta,t\text{-}1} + \varepsilon_t$, where $z_t = y_t - y_0$ (hence, $z_0 = 0$); dealing with the $\tau_{\mu NL}$-test, $\delta = \gamma(\lambda + 1)$. Thus, the original equation with an arbitrary initial condition is equivalent to the equation with identically zero initial condition and changed intercept. Given $y_0 = 0$, the power of $\tau_{\mu NL}$-test and $\tau_{\mu P}$-test increase with increasing $|\alpha|$. In a like manner, when $y_0 \neq 0$, the power increases with increasing $|\alpha + \lambda y_0|$. Thus, the power of the tests as a function of y_0 is symmetric around $y_0 = -\alpha/\lambda$, being minimal at this point. As compared to the case of $y_0 = 0$, the power falls if $y_0 \in (0, -2\alpha/\lambda)$, and rises otherwise (excluding the equality points 0 and $-2\alpha/\lambda$). For example, with $\alpha = 1$ and $\lambda = -0.1$, negative initial values and values exceeding 20 increase the power of the tests, while positive values below 20 decrease it.

At last, Table VII demonstrates the rise in the power of the nonlinear and linear tests with sample size increasing from 50 to 200. This range covers sample sizes typical in economic applications. The fact engages attention that all three tests involving constant are biased at $\lambda = -0.01$ with $T = 50$. At the same time, their power becomes satisfactory at this point with $T = 200$, except for the $\tau_{\mu P}$-test, the power of which still remains rather low.

TABLE VII.

POWER OF SIZE 0.05 TESTS FOR SAMPLES OF SIZES 50 AND 200; $\Theta = 0.5$

Test	$\lambda = -0.2$		$\lambda = -0.1$		$\lambda = -0.05$		$\lambda = -0.01$		$\lambda = 0$	
	T=50	T = 200	T=50	T = 200	T=50	T = 200	T=50	T = 200	T=50	T = 200
τ_0	0.777	1.000	0.325	0.998	0.143	0.761	0.061	0.119	0.049	0.050
τ_{0NL}	0.645	1.000	0.268	0.993	0.130	0.728	0.061	0.119	0.050	0.050
τ_{0P}	0.436	1.000	0.167	0.934	0.093	0.435	0.057	0.087	0.050	0.051
τ_μ	0.559	1.000	0.380	0.992	0.279	0.958	0.039	0.823	0.050	0.051
$\tau_{\mu NL}$	0.449	1.000	0.239	0.986	0.173	0.944	0.024	0.813	0.050	0.052
$\tau_{\mu P}$	0.336	1.000	0.163	0.917	0.072	0.698	0.009	0.157	0.050	0.051

Note: For τ_0, τ_{0NL}, and τ_{0P}, $\alpha = 0$; for τ_μ, $\tau_{\mu NL}$, and $\tau_{\mu P}$, $\alpha = 1$; $\gamma = 0$; $y_0 = 0$.

5. EMPIRICAL APPLICATION

In this section, the proposed nonlinear test is applied to analyzing goods market integration in Russia, using data from Gluschenko (2004). A country's regional market is deemed as integrated with the national market if the law of one price holds, that is, if $P_{rt} \equiv \ln(p_{rt}/p_{0t}) = 0$, where p_{rt} is the price of a good in region r, p_{0t} is its price for the country as a whole, and t indexes months. Certainly, there may be a persistent (equilibrium) nonzero price difference, implying $P_{rt} = a$. The trouble is that this a can reflect both the effect of "natural," irremovable market frictions compatible with the notion of integration (e.g., transportation costs), and the effect of artificial, transient ones impeding market integration (e.g., regional protectionism). But in the context of a univariate model, there is no way to separate these effects. That is why the strict price parity is adopted as the benchmark of integration, any deterministic difference in prices being interpreted as an indication of non-integration. Then the problem is to test whether $P_{rt} = 0$ holds over $t = 0,\ldots,T$, that is, to test $\{P_{rt}\}$ for stationarity around zero level, using a model with no constant term.

The time span under consideration is 1994:01 through 2000:12 (thus, $T = 83$). In August 1998, a financial crisis occurred in Russia, causing breaks in price trajectories for some regions of the country. Since the interest is to reveal whether a regional market is integrated in recent years rather then in the pre-crisis time, the break is modeled by the reversed break dummy, $B'_{\theta t}$ (thus, the break is assumed to be parity-directed). The price is represented by the cost of a basket of 25 staples. Of all available 75 regional time series, only four are considered here, labeled for brevity as A, B, C, and D (these are the Ryazan Oblast, Ulyanovsk Oblast, Lipetsk Oblast, and Khabarovsk Krai); they illustrate four possible cases: stationarity/non-stationarity and significant/insignificant break.

The timing of the break is not uniform across regions, varying from 1998:08 through 1999:02 (in terms of θ, from 54 to 60; in terms of Θ, from 0.65 to 0.72). The break point is found by estimating regression (3) for $\theta = 54,\ldots,60$, and choosing θ that yields the least sum of squared residuals. The standard AR(1) model $\Delta y_t = \lambda y_{t-1} + \varepsilon_t$ is used as an alternative specification for cases of insignificant break. For comparison, Perron-type model (7) is also estimated (using the same method of finding the break points).

Dealing with actual processes, there is no the assurance that they are AR(1). Therefore, to eliminate effects of possible serial correlation from the residuals, the

Phillips (1987) transformation is applied, $Z_t(\tau_{0NL})$, $Z_t(\tau_{0P})$, and $Z_t(\tau_0)$, using the Newey-West (1994) automatic bandwidth selection technique with the Bartlett spectral kernel. (Regarding the standard AR(1) model, this is just the Phillips-Perron test.) Certainly, it is not known whether such a method is applicable to the tests allowing for break. Nonetheless, a number of experiments give an impression that it does work.

The estimation results are tabulated in Table VIII. For model (3), τ means τ_{0NL}; τ is τ_{0P} for (7); and τ is τ_0 for AR(1); σ stands for the standard deviation of the residuals; *p*-values for γ and δ are those of the ordinary *t*-test.

TABLE VIII.

EMPIRICAL RESULTS

Model	λ	(s.e.)	τ	*p*-value	$Z_t(\tau)$	*p*-value	γ	(s.e.)	*p*-value	δ	(s.e.)	*p*-value	σ
Series A; $\theta + 1$ = 1998:09													
(3)	–0.153	(0.060)	–2.548	0.022	–2.701	0.017	–0.083	(0.019)	0.000				0.025
(7)	–0.161	(0.060)	–2.698	0.050	–2.698	0.050	–0.055	(0.026)	0.036	0.038	(0.025)	0.133	0.025
AR(1)	–0.036	(0.030)	–1.188	0.213	–1.237	0.197							0.026
Series B; $\theta + 1$ = 1998:12													
(3)	–0.032	(0.015)	–2.058	0.055	–2.386	0.031	0.053	(0.058)	0.363				0.058
(7)	–0.088	(0.034)	–2.609	0.062	–2.653	0.057	0.054	(0.057)	0.349	–0.080	(0.058)	0.168	0.057
AR(1)	–0.035	(0.017)	–2.088	0.036	–2.381	0.018							0.058
Series C; $\theta + 1$ = 1998:09													
(3)	–0.019	(0.016)	–1.210	0.209	–1.405	0.155	0.084	(0.036)	0.024				0.036
(7)	–0.089	(0.039)	–2.307	0.109	–2.149	0.144	0.082	(0.036)	0.024	–0.098	(0.036)	0.008	0.035
AR(1)	–0.026	(0.021)	–1.219	0.203	–1.090	0.248							0.037
Series D; $\theta + 1$ = 1998:12													
(3)	–0.025	(0.022)	–1.141	0.233	–0.964	0.298	0.068	(0.046)	0.139				0.045
(7)	–0.094	(0.049)	–1.893	0.221	–1.283	0.436	0.070	(0.045)	0.124	–0.043	(0.046)	0.352	0.045
AR(1)	–0.021	(0.018)	–1.142	0.229	–1.093	0.247							0.045

Series A cannot be recognized as stationary while using the Dickey-Fuller and Phillips-Perron tests. But the unit root is easily rejected at the 5% level if the structural break is allowed for. The break is highly significant, having the height of 3.3σ. The negative sign of γ indicates that the financial crisis has caused a rise in the relative price in this region, so drawing it to the average Russian value. For series B, the break turns out to be insignificant (while the unit root is rejected at the 10% level). Hence, the ordinary AR(1) model is more adequate for this case. It rejects the unit root at the 5% level. Series C demonstrates the case when the unit root cannot be rejected and the break is significant. Thus, we have a non-stationary process with break. At last, series D has both the insignificant break and the unit root when analyzed with the use of models allowing for break; the Dickey-Fuller and Phillips-Perron tests evidence non-stationarity of this series.

In Table VIII, the nonlinear and Perron-type models yield quantitatively similar patterns. But in general, their results markedly disagree. Of all the 75 series having been analyzed, the inference regarding the unit root differs for 14 ones, or 19%. The inference as to the significance of the break is opposite for 29% of series. The Perron-type and nonlinear models has the minimal sum of squared residuals in different points in time for 45% of series. Though, 20% of the break dates are adjacent; thus, there is 25% of a considerable disagreement regarding the choice of them.

6. CONCLUDING REMARKS

In this paper it is shown that the Perron-type equations which are commonly used to test time series with a structural break for a unit root are, in fact, a linear approximation of nonlinear models. Using the latter more adequate models, a test with better properties is obtained, more power among them. While the asymptotic null distribution of the Perron test statistic depends on the time of break, the nonlinear-test statistics share the same asymptotic distribution for all possible break points. Moreover, this distribution proves to be the Dickey-Fuller one.

This story resembles that of innovation sequences other than AR(1) while testing series without break. Seemingly, dealing with, say, AR(d), we should have a special distribution for each d. But if the properties of the regression errors are properly accounted for by inserting additional lags like in the augmented Dickey-Fuller test, or by applying the Phillips' (1987) transformation of the t-ratio like in the Phillips-Perron test, the Dickey-Fuller distributions turn out to be valid. Similarly, properly taking into account the "deterministic shock" makes the Dickey-Fuller distributions valid for this case.

Two directions of further research are obvious. The first is examining more general innovation sequences than AR(1) in order to reveal whether the Phillips transformation of the test statistic (or inserting additional lags) is applicable for the nonlinear test equations.

Preliminary results suggest that this may be the case, but more extensive numerical experiments are required for a confident conclusion. The second direction is studying trend-and-break time series. Since breaks of three types are possible in this case (a change in intercept, a change in slope of the trend function, and the combination of both changes), it is interesting how the nonlinear-test statistic differs across these types. Judging from the fact that the asymptotic distributions of the Perron statistic for $\Theta = 0$ and $\Theta = 1$ coincide with the Dickey-Fuller distribution associated with regression having the intercept and trend, the nonlinear test statistic may be expected to have the latter (limiting) distribution for all three types of the break.

ACKNOWLEDGEMENTS

I would like to express my appreciation to David N. DeJong, Pierre Perron, and Alexander Tsyplakov for helpful comments and suggestions.

REFERENCES

Bartley, W.A., J. Lee, and M.C. Strazicich (2001), “Testing the null of cointegration in the presence of a structural break,” *Economics Letters*, 73, 315–323.

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

Evans, G.B.A., and N.E. Savin (1981), “Testing for Unit Roots: 1,” *Econometrica*, 49, 753-779.

———— (1984), “Testing for Unit Roots: 2,” *Econometrica*, 52, 1241-1270.

Gluschenko, K. (2004), “Integration of the Russian Market: Empirical Analysis,” *EERC Working paper* No. 04/06.

Hansen B.E. (1997), “Approximate Asymptotic P Values for Structural-Change Tests,” *Journal of Business and Economic Statistics*, 15, 60-67.

———— (2001), “The New Econometrics of Structural Change: Dating Breaks in U.S. labor productivity,” *Journal of Economic Perspectives*, 15, 117-128.

Lanne, M., H. Lütkepohl and P. Saikkonen (2002), “Comparison of Unit Root Tests for Time Series With Level Shifts,” *Journal of Time Series Analysis*, 23 (6), 667-685.

MacKinnon, J. G. (1996), “Numerical Distribution Functions for Unit Root and Cointegration Tests,” *Journal of Applied Econometrics*, 11, 601-618.

Müller, U.K., and G. Elliott (2003), "Tests for Unit Roots and the Initial Condition," *Econometrica*, 71, 1269-1286.

Newey, W. K., and West, K. D. (1994), "Automatic Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61, 631-653.

Perron, P. (1989), "The Great Crash, the Oil Price Shock, and The Unit Root Hypothesis," *Econometrica*, 57, 1361-1401.

———— (1990), "Testing for a Unit Root in a Time Series with a Changing Mean," *Journal of Business and Economics Statistics*, 8, 153-162.

Perron, P. and T.J. Vogelsang (1992), "Nonstationarity and Level Shifts with an Application to Purchasing Power Parity," *Journal of Business and Economics Statistics*, 10, 301-320.

Phillips, P.C.B. (1987), "Time Series Regression with a Unit Root," *Econometrica*, 55, 277-301.

Rappoport, P. and L. Reichlin (1989), "Segmented Trends and Non-Stationary Time Series," *Economic Journal*, 99 (395), 168-177.

Quantitative Micro Software, LLC (2004), *EViews 5 User's Guide*.

The EViews programs for estimating and testing AR(1) with break as well as for obtaining critical values and exploring power properties of the tests can be found on the author's home page http: //econom.nsu.ru/users/gluschenko/econometrics.htm .