Unit Roots and Smooth Transitions: A Replication

Kulaksizoglu, Tamer

4 February 2015

Unit Roots and Smooth Transitions: A Replication

Tamer Kulaksizoglu*

February 4, 2015

Abstract

This paper replicates Leybourne et al. (1998), who propose a Dickey-Fuller type test for unit root that is most appropriate when there is reason to suspect the possibility of deterministic structural change in the series. We find that our replicated results are quite similar to the authors' results. We also make the Ox source code available.

Key words: Dickey-Fuller test; Integrated process; Nonlinear trend; Structural change
JEL classification: C12; C15

1 Introduction

The alternative hypothesis of the standard augmented Dickey-Fuller (ADF) tests is stationarity around a fixed mean or a linear trend. However, another possibility is stationarity around a linear trend with an instantaneous break[1]. Leybourne et al. (1998) broaden this class of alternatives to allow for a smooth transition between trends by developing Dickey-Fuller type tests and exploring their properties. In this paper, we try to replicate their results. Section 2 discusses the smooth transition regression models. Section 3 discusses the tests. Section 4 presents the replicated results. Section 5 concludes the paper.

2 Smooth Transition Regression Models

Leybourne et al. (1998) consider the following three logistic smooth transition regression (STR) models

$$\text{Model A: } y_t = \alpha_1 + \alpha_2 S_t(\gamma, \tau) + v_t$$
$$\text{Model B: } y_t = \alpha_1 + \alpha_2 S_t(\gamma, \tau) + \beta_1 t + v_t$$
$$\text{Model C: } y_t = \alpha_1 + \alpha_2 S_t(\gamma, \tau) + \beta_1 t + \beta_2 t S_t(\gamma, \tau) + v_t$$

*The Ox source code used to obtain the results in this paper can be downloaded from the author's GitHub account at `https://github.com/tamerk/ox-lnv98`.

[1] See Perron (1989), Perron (1990), and Zivot and Andrews (1992).

where y_t is the series to be tested for unit root, t is a time trend, and v_t is a zero-mean I(0) process. $S_t(\gamma, \tau)$ is the logistic transition function, which controls the transition between regimes

$$S_t(\gamma, \tau) = [1 + \exp\{-\gamma(t - \tau T)\}]^{-1}$$

where $\gamma > 0$, $t = 1, \ldots, T$, and T is the sample size. τ is the transition midpoint fraction, which is also called the location parameter. It detemines the timing of the transition midpoint. γ is the speed of transition parameter, which is also called the slope or smoothness parameter in the literature. It determines the speed of the transition between regimes.

The models represent deterministic structural change. In order to see this, rewrite Model C, which nests the other models, as

$$y_t = (\alpha_1 + \alpha_2 S_t(\gamma, \tau)) + (\beta_1 + \beta_2 S_t(\gamma, \tau))\, t + v_t$$

The first part on the right hand-side of the equation is the time-varying intercept and the second part is the time-varying slope with respect to time trend. However, the model is also able to mimic no break when $\gamma = 0$ and instantaneous structural break when $\gamma \to \infty$.

3 Unit Root Tests

The authors consider the following hypotheses

$$\begin{aligned}
&H_0: \; y_t = \mu_t \quad \mu_t = \mu_{t-1} + \varepsilon_t \quad \mu_0 = \psi \\
&H_1: \; \text{Model A, Model B, or Model C} \\
&\quad \text{and} \\
&H_0: \; y_t = \mu_t \quad \mu_t = \kappa + \mu_{t-1} + \varepsilon_t \quad \mu_0 = \psi \\
&H_1: \; \text{Model B or Model C}
\end{aligned}$$

where ε_t is zero-mean $I(0)$ process. Notice that the null hypothesis is unit root but the alternative hypothesis is trend-stationarity. They propose a test statistic that can be calculated in a two-step procedure.

STEP 1. Estimate the models with nonlinear least squares (NLS) and obtain the residuals[2]

$$\begin{aligned}
&\text{Model A: } \hat{v}_t = y_t - \hat{\alpha}_1 - \hat{\alpha}_2 S_t(\hat{\gamma}, \hat{\tau}) \\
&\text{Model B: } \hat{v}_t = y_t - \hat{\alpha}_1 - \hat{\alpha}_2 S_t(\hat{\gamma}, \hat{\tau}) - \hat{\beta}_1 t \\
&\text{Model C: } \hat{v}_t = y_t - \hat{\alpha}_1 - \hat{\alpha}_2 S_t(\hat{\gamma}, \hat{\tau}) - \hat{\beta}_1 t - \hat{\beta}_2 t S_t(\hat{\gamma}, \hat{\tau})
\end{aligned}$$

[2] Note that the equations of Models B and C in Step 1 contain typos in the paper. $\hat{\alpha}_1 t$ should be $\hat{\alpha}_1$ and $\hat{\beta}_1$ should be $\hat{\beta}_1 t$.

STEP 2. Estimate the following auxiliary regression

$$\Delta\hat{v}_t = \hat{\rho}\hat{v}_{t-1} + \sum_{i=1}^{k}\hat{\delta}_i\Delta\hat{v}_{t-i} + \hat{\eta}_t$$

The ADF test statistic is the t-ratio of the parameter $\hat{\rho}$. The summation part accounts for any stationary dynamics in ε_t. The test statistics are denoted by s_α, $s_{\alpha(\beta)}$, and $s_{\alpha\beta}$ if the residuals come from Model A, Model B, and Model C, respectively.

Estimation of the smooth transition regression models are discussed in Granger and Teräsvirta (1993), Teräsvirta (1998), and Teräsvirta (2004). In general, the estimation is not an easy task. However, observing that the models are linear in $(\alpha_1, \alpha_2, \beta_1, \beta_2)$ once γ and τ are known, Leybourne et al. (1998) propose a "concentrated" NLS estimation[3]. For Model C, the objective function is set to be

$$SSR = \sum_{t=1}^{T}\left(y_t - \hat{\boldsymbol{\pi}}'\hat{\mathbf{x}}_t\right)^2$$

where

$$\hat{\boldsymbol{\pi}} = \left[\hat{\alpha}_1, \hat{\beta}_1, \hat{\alpha}_2, \hat{\beta}_2\right]' = \left(\sum_{t=1}^{T}\hat{\mathbf{x}}_t\hat{\mathbf{x}}_t'\right)^{-1}\sum_{t=1}^{T}\hat{\mathbf{x}}_t y_t$$

and

$$\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_t\left(\hat{\gamma}, \hat{\tau}\right) = \left\{1, t, S_t\left(\hat{\gamma}, \hat{\tau}\right), tS_t\left(\hat{\gamma}, \hat{\tau}\right)\right\}$$

The estimated values $\hat{\gamma}$ and $\hat{\tau}$ are obtained from the optimization algorithm. They use the BFGS optimization algorithm in the OPMUM library of GAUSS 3.1. We use the same optimization algorithm implemented in the `MaxBFGS` function in Ox. The advantage of this procedure is that the parameter space that needs to be searched over is only two-dimensional instead of four through six.

The NLS estimates of γ and τ do not have closed-form solutions. That's why it is difficult to establish the analytical relationship between $\hat{v}_t$ and y_t. Thus the null asymptotic distributions of the test statistics s_α, $s_{\alpha(\beta)}$, and $s_{\alpha\beta}$ are not easily tractable by analytical means. Instead, the authors conduct Monte Carlo simulation experiments to approximate the null distribution of the test statistics.

Before presenting the results, we would like to mention a few details in regard to the calculations. First of all, it is conventional to divide the argument of the exponential function in the logistic transition function by the standard deviation of the transition variable to make it scale-free[4]. Since this is not mentioned in the paper, we do not apply the scaling. Second, it is also conventional to use a

[3] Maugeri (2014) warns that "this procedure needs to be used with caution as it may yield biased and inconsistent estimates, especially when faced with small samples".

[4] See Teräsvirta (2004), page 228.

grid search to find reasonable starting values for γ and τ[5]. However, following the authors, we set the starting values for γ and τ to 1 and 0.5, respectively[6]. Finally, as the authors suggested, we use the Moore-Penrose generalized inverse to estimate the $\hat{\boldsymbol{\pi}}$ vector when the components of$\hat{\mathbf{x}}_t$ becomes linearly dependent due to the fact that the optimization algorithm selects values of $\hat{\gamma}$ and $\hat{\tau}$ that make $S_t(\hat{\gamma}, \hat{\tau})$ constant for all t[7].

4 Experiments and Application

In this section, we present the results of our replications. Since the replications include generating random numbers, exact replication results should not be expected. Instead, we focus on the proximity of the numbers and confirmation of the qualitative results. The experiments examine the finite sample size and power characteristics of the tests. More specifically, Experiment I generates the critical values, Experiment II generates the sizes of the tests, and Experiments III and IV investigate the power properties of the tests.

4.1 Experiment 1

The first experiment in this paper replicates Table I in Leybourne et al. (1998), which shows the critical values of the test statistics. Those critical values will be used later to calculate the empirical sizes and powers of the tests. The null DGP is pure random walk

$$y_t = \mu_t, \quad \mu_t = \mu_{t-1} + \varepsilon_t, \quad \mu_0 = 0, \quad \varepsilon_t \sim NID(0,1)$$

As a precaution against possible convergence failures, we set the number of replications for the null distribution to 25,000 but used the first 20,000 in critical value calculations. We also set $k = 0$ in the ADF tests. Table 1 shows the replicated critical values[8] of the null distributions of the test statistics s_α, $s_{\alpha(\beta)}$, and $s_{\alpha\beta}$ at 0.10, 0.05, and 0.01 significance levels for various sample sizes. Overall, the replicated critical values are quite close to the authors' critcal values. The minimum absolute difference is 0.001 and the maximum is 0.151.

[5]See Teräsvirta (2004), page 228. This grid search is also used in the econometric software `JMulTi`, which can be downloaded at the URL `http://www.jmulti.de/`.

[6]The authors claim that the solutions at convergence were not found to be sensitive to these choices.

[7]They suggest applying the standard ADF test when this happens and also when the convergence is very slow.

[8]Leybourne et al. (1998) check the robustness of the simulated critical values by also generating random walks with innovations drawn from $\chi^2(1) - 1$ and $t(6)$ distributions and find that the empirical sizes of all the three tests are close to the nominal sizes using the critical values in Table I of their paper for all samples. We did not conduct a similar experiment. However, the interested reader can conduct it easily by replacing the `rann` function in the class `SimRW` with `ranchi` and `rant` functions and then run the script file `Table_01.ox` for each change.

Table 1: Null Critical Values for Unit Root Tests

	s_α			$s_{\alpha(\beta)}$			$s_{\alpha\beta}$		
T	0.10	0.05	0.01	0.10	0.05	0.01	0.10	0.05	0.01
25	-4.308	-4.737	-5.717	-5.066	-5.529	-6.554	-5.624	-6.098	-7.218
50	-4.039	-4.405	-5.122	-4.634	-4.982	-5.768	-5.025	-5.396	-6.107
100	-3.922	-4.235	-4.841	-4.468	-4.775	-5.389	-4.774	-5.098	-5.742
200	-3.858	-4.174	-4.749	-4.373	-4.676	-5.265	-4.657	-4.978	-5.586
500	-3.834	-4.125	-4.716	-4.326	-4.610	-5.124	-4.582	-4.854	-5.381

Note: Nominal sizes 0.10, 0.05, and 0.01.

Table 2: Empirical Sizes of the Test s_α for ARIMA(1,1,0) Processes

		T=100			T=200		
φ	k	0.10	0.05	0.01	0.10	0.05	0.01
0.0	0	0.112	0.060	0.015	0.107	0.051	0.015
0.0	1	0.102	0.058	0.014	0.099	0.051	0.010
0.0	4	0.078	0.037	0.008	0.090	0.044	0.008
-0.4	0	0.602	0.485	0.300	0.631	0.530	0.331
-0.4	1	0.100	0.048	0.008	0.104	0.056	0.011
-0.4	4	0.068	0.034	0.005	0.093	0.046	0.007
0.4	0	0.008	0.004	0.001	0.006	0.004	0.002
0.4	1	0.094	0.045	0.009	0.093	0.047	0.011
0.4	4	0.079	0.046	0.009	0.093	0.047	0.006
-0.8	0	0.982	0.971	0.931	0.991	0.983	0.959
-0.8	1	0.094	0.047	0.011	0.103	0.055	0.014
-0.8	4	0.069	0.029	0.004	0.082	0.040	0.013
0.8	0	0.017	0.012	0.006	0.027	0.018	0.006
0.8	1	0.100	0.054	0.016	0.104	0.057	0.012
0.8	4	0.074	0.039	0.008	0.089	0.039	0.009

Note: Nominal sizes 0.10, 0.05, and 0.01.

4.2 Experiment 2

The second experiment replicates Table II, which shows the empirical size of the test s_α. The null DGP is a more general I(1) process

$$y_t = \mu_t, \quad \Delta\mu_t = \varphi\Delta\mu_{t-1} + \varepsilon_t, \quad \mu_0 = 0, \quad \varepsilon_t \sim NID\,(0,1)$$

The purpose of the experiment is to see whether the null critical values are robust to more general I(1) null DGPs. Notice that y_t follows an ARIMA(1,1,0) process for $\varphi \neq 0$. The number of replications is 2,500, again as a possible precaution against convergence failures, but we used the first 2,000 in the table. The replicated results are shown in Table 2, which are remarkably close to those in the authors' paper. The maximum absolute difference is 0.026.

Table 3: Empirical Powers

		T=100				T=200			
		s_α		τ_τ		s_α		τ_τ	
φ	k	0.10	0.05	0.10	0.05	0.10	0.05	0.10	0.05
0.9	0	0.167	0.090	0.312	0.172	0.526	0.337	0.798	0.627
0.9	4	0.078	0.038	0.203	0.107	0.318	0.166	0.609	0.432
0.8	0	0.552	0.359	0.815	0.644	0.993	0.961	1.000	0.999
0.8	4	0.198	0.104	0.468	0.300	0.771	0.578	0.961	0.881
0.7	0	0.894	0.765	0.991	0.952	1.000	1.000	1.000	1.000
0.7	4	0.329	0.192	0.671	0.481	0.953	0.872	0.996	0.986

Note: Nominal sizes 0.10 and 0.05.

4.3 Experiment 3

The third experiment replicates Table III, which shows the empirical power of the test s_α. The null DGP is a stationary AR(1) process

$$y_t = \mu_t, \quad \mu_t = \varphi\mu_{t-1} + \varepsilon_t, \quad \mu_0 = 0, \quad \varepsilon_t \sim NID\,(0,1)$$

where $\varphi < 1$. The number of replications is 2,500 but we used the first 2,000 in the table. By way of comparison, the powers of the standard ADF test with trend and intercept, denoted by τ_τ (a natural competitor to s_α), is also included[9]. Once again, our replicated values are close to the author's. The maximum absolute difference is 0.049.

4.4 Experiment 4

The fourth experiment replicates Table IV, which also shows the empirical power of the test s_α. When the DGP is a stationary AR(1) process, the standard ADF test has more power than the s_α test. The fourth experiment generates data from Model A with first-order autoregressive innovations and conducts the same tests. The DGP is

$$y_t = 1 + 10S_t\,(\gamma,\tau) + \mu_t, \quad \mu_t = 0.8\mu_{t-1} + \varepsilon_t$$

where $\mu_0 = 0$ and $\varepsilon_t \sim NID\,(0,1)$. The number of replications is 2,600 but we used the first 2,000 in the table. The experiment is repeated for various values of γ and τ. Table 4 shows the results of the power simulations. Again, our replicated values are very close to the author's. We find that the maximum absolute difference is 0.09.

4.5 Application

Leybourne et al. (1998) apply the most general form of their test procedure $s_{\alpha\beta}$ and the standard ADF test τ_τ to the U.S. data set first analyzed by Nelson

[9]We use the critical values obtained from EViews for sample sizes 100 and 200 and zero lag length.

Table 4: Empirical Powers of a Smooth Transition and ADF Test

			T=100				T=200			
			s_α		τ_τ		s_α		τ_τ	
γ	τ	k	0.10	0.05	0.10	0.05	0.10	0.05	0.10	0.05
0.01	0.5	0	0.576	0.394	0.827	0.654	0.978	0.935	1.000	0.999
0.01	0.5	4	0.217	0.112	0.482	0.293	0.717	0.540	0.952	0.880
0.01	0.2	0	0.590	0.394	0.826	0.649	0.986	0.948	1.000	0.999
0.01	0.2	4	0.216	0.107	0.463	0.295	0.727	0.550	0.954	0.885
0.10	0.5	0	0.659	0.482	0.497	0.306	0.993	0.972	0.740	0.487
0.10	0.5	4	0.365	0.225	0.203	0.110	0.862	0.703	0.170	0.064
0.10	0.2	0	0.602	0.420	0.207	0.092	0.974	0.947	0.369	0.160
0.10	0.2	4	0.255	0.149	0.062	0.028	0.789	0.637	0.047	0.014
0.50	0.5	0	0.563	0.380	0.015	0.003	0.991	0.950	0.134	0.039
0.50	0.5	4	0.223	0.127	0.012	0.005	0.738	0.554	0.025	0.008
0.50	0.2	0	0.522	0.336	0.002	0.001	0.981	0.940	0.023	0.004
0.50	0.2	4	0.167	0.091	0.001	0.000	0.729	0.551	0.005	0.001
1.00	0.5	0	0.511	0.331	0.008	0.003	0.989	0.949	0.109	0.025
1.00	0.5	4	0.199	0.107	0.019	0.005	0.717	0.524	0.060	0.018
1.00	0.2	0	0.476	0.309	0.001	0.000	0.969	0.926	0.015	0.001
1.00	0.2	4	0.150	0.071	0.003	0.000	0.713	0.531	0.007	0.001
5.00	0.5	0	0.474	0.299	0.012	0.002	0.991	0.945	0.177	0.049
5.00	0.5	4	0.146	0.077	0.022	0.005	0.720	0.519	0.061	0.017
5.00	0.2	0	0.451	0.283	0.002	0.001	0.975	0.933	0.027	0.004
5.00	0.2	4	0.123	0.060	0.003	0.001	0.692	0.503	0.013	0.002

Note: Nominal sizes 0.10 and 0.05.

Table 5: Empirical Application

	T	k	$s_{\alpha\beta}$	τ_τ
US Series				
Real GNP	62	2	-4.31	-2.94
Nominal GNP	62	1	-3.18	-2.32
Per capita real GNP	62	1	-4.46	-3.05
Industrial production	111	7	-4.57	-2.67
Employment	81	1	-3.70	-3.12
Unemployment	81	1	-4.55	-3.92
GNP deflator	82	1	-3.40	-2.52
Consumer prices	111	5	-3.26	-2.37
Wages	71	1	-3.45	-2.52
Real wages	71	1	-4.19	-3.05
Money stock	82	1	-3.42	-3.08
Velocity	102	0	-3.28	-1.66
Bond yield	71	6	-5.95	-0.19
SP 500	100	1	-5.26	-2.65

and Plosser (1982). The data set contains 14 annual macroeconomic series with the numbers of observations ranging from 62 to 111. They present the results in Table V, which is replicated in Table 5[10]. As can be seen from the table, except for the unemployment series, we get very close results. The maximum absolute difference for the unemployment series is 0.86 for $s_{\alpha\beta}$ and 0.66 for τ_τ. The maximum absolute difference for the series excluding the unemployment series is 0.15 for for the $s_{\alpha\beta}$ test and 0.13 for the τ_τ test. The difference in the results for the standard ADF test τ_τ is a bit puzzling since the test is pretty standard. The reasons for these differences might be a slightly different data set[11], software issue, or just typo.

5 Conclusion

In this paper, we tried to replicate Leybourne et al. (1998), who propose an auxiliary test for the augmented Dickey-Fuller test that is most appropriate when there is reason to suspect the possibility of structural change in the series. With a few exceptions, we find that our replicated results are quite similar to the authors' results, which is a testimony to the authors' careful econometric analysis. We also make the Ox source code available to allow for others to double-check our results.

[10] They also apply the same tests to U.K. consols and industrial production data but we don't have these data sets so we cannot replicate their results.

[11] We obtained the data set from the following link: `http://www.ventosa-santaularia.com/NP_database.html`.

References

Granger, C. W. J. and Teräsvirta, T. (1993). *Modelling Nonlinear Economic Relationships*. Oxford University Press.

Leybourne, S., Newbold, P., and Vougas, D. (1998). Unit roots and smooth transitions. *Journal of Time Series Analysis*, 19(1):83–97.

Maugeri, N. (2014). Some pitfalls in smooth transition models estimation: A Monte Carlo study. *Computational Economics*, 44(3):339–378.

Nelson, C. R. and Plosser, C. I. (1982). Trends and random walks in macroeconomic time series: Some evidence and implications. *Journal of Monetary Economics*, 10:139–162.

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

Perron, P. (1990). Testing for a unit root in a time series with a changing mean. *Journal of Business and Economic Statistics*, 8:153–162.

Teräsvirta, T. (1998). *Handbook of Applied Economic Statistics*, chapter Modelling Economic Relationships with Smooth Transition Regressions, pages 229–246. Dekker, New York.

Teräsvirta, T. (2004). *Applied Time Series Econometrics*, chapter Smooth Transition Regression Modelling, pages 222–242. Cambridge University Press, Cambridge.

Zivot, E. and Andrews, D. W. K. (1992). Further evidence on the Great Crash, the oil price shock and the unit root hypothesis. *Journal of Business and Economic Statistics*, 10:251–270.