



The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

The Stata Journal (2017)
17, Number 3, pp. 704–722

Response surface models for OLS and GLS detrending-based unit-root tests in nonlinear ESTAR models

Jesús Otero
Facultad de Economía
Universidad del Rosario
Bogotá, Colombia
jesus.otero@urosario.edu.co

Jeremy Smith
Department of Economics
University of Warwick
Coventry, UK
jeremy.smith@warwick.ac.uk

Abstract. In this article, we calculate response surface models for a large range of quantiles of the [Kapetanios, Shin, and Snell \(2003, *Journal of Econometrics* 112: 359–379\)](#) and [Kapetanios and Shin \(2008, *Economics Letters* 100: 377–380\)](#) tests for the null hypothesis of a unit root against the alternative—that the series of interest follows a globally stationary exponential smooth transition autoregressive process. The response surface models allow estimation of finite-sample critical values and approximate p -values for different combinations of the number of observations, T , and the lag order in the test regression, p . The latter can be either specified by the user or optimally selected using a data-dependent procedure. We present the new commands `kssur` and `ksur` and illustrate their use with an empirical example.

Keywords: st0493, `kssur`, `ksur`, unit-root test, nonlinear ESTAR models, Monte Carlo, response surface, critical values, lag length, p -values

1 Introduction

Since [Nelson and Plosser \(1982\)](#), testing whether an economic time series can be best described as an integrated process of order one against order zero—denoted $\sim I(1)$ and $\sim I(0)$, respectively—has become a key stage in empirical macroeconomic analysis. Among the tests proposed in the literature, the [Said and Dickey \(1984\)](#) approach to augmenting the basic [Dickey and Fuller \(1979\)](#) tests, commonly referred to as ADF tests, is one of the most commonly applied procedures. However, a common criticism is that the ADF tests exhibit low power in that there is nonlinearity under the alternative hypothesis.

Developing the unit-root testing framework, [Kapetanios, Shin, and Snell \(2003\)](#), and later [Kapetanios and Shin \(2008\)](#), propose tests, the KSS and KS tests, respectively, of the unit-root null hypothesis against the alternative of a globally stationary exponential smooth transition autoregressive (ESTAR) process. Monte Carlo simulation results show that these tests, based on a nonlinear model, exhibit higher power than the ADF test in the case of relatively more persistent series. To implement the tests, [Kapetanios, Shin, and Snell \(2003\)](#) tabulate asymptotic critical values (CVs) based on $T = 1000$ observations. Subsequent work by [Patterson \(2012\)](#) examines the sensitivity

of the CVs of the KSS test to the sample size by tabulating those for $T = 200$ observations, finding that little variation is observed when the two cases are compared. Neither [Kapetanios, Shin, and Snell \(2003\)](#) nor [Patterson \(2012\)](#) investigate the sensitivity of the CVs to the number of lags of the dependent variable.

This article undertakes an extensive set of Monte Carlo simulations, summarized by response surface regressions, to calculate finite-sample CVs and approximate p -values of the KSS and KS tests. The simulations accommodate processes with zero mean, nonzero mean, and nonzero trend. The lag length of the test regression is allowed to be either fixed or determined endogenously using a data-dependent procedure. The newly developed and available `kssur` and `ksur` commands can easily be used to calculate the KSS and KS test statistics, along with both finite-sample CVs and approximate p -values.

At this point, we emphasize that this article considers only testing. In the ADF test, rejection of the null also means that one has already fit the “correct” model. However, in the current context of the KSS (and KS) tests, a rejection means that one must proceed with another model, that is, the ESTAR model, and not the model used in the testing equations.

This article proceeds as follows: Section 2 provides an overview of the KSS and KS unit-root tests. Section 3 presents the design of the Monte Carlo experiments. Section 4 reports the estimated response surfaces and describes the procedure to estimate the associated approximate p -values. Section 5 describes the `kssur` and `ksur` commands. Section 6 illustrates these commands with an empirical example based on coffee price differentials. Section 7 concludes this article.

2 Theoretical framework

[Kapetanios, Shin, and Snell \(2003\)](#) propose a test for the null hypothesis of a unit root against the alternative of a nonlinear ESTAR process. This test is based on the univariate STAR(1) model

$$y_t = \beta y_{t-1} + \gamma y_{t-1} \Theta(\theta; y_{t-d}) + \varepsilon_t \quad (1)$$

where $t = 1, \dots, T$ is the total number of observations. $\varepsilon_t \sim \text{i.i.d.}(0, \sigma^2)$, β , γ , and σ^2 are unknown parameters and, using the terminology used in the STAR models literature, $\Theta(\theta; y_{t-d})$ is a transition function that is assumed to follow the exponential form

$$\Theta(\theta; y_{t-d}) = 1 - \exp(-\theta y_{t-d}^2) \quad (2)$$

with $\theta \geq 0$ and delay parameter $d \geq 1$. The exponential transition function (2) is a symmetric U-shaped function around zero that is bounded between zero and one. Inserting (2) into (1) yields the ESTAR model

$$\Delta y_t = \phi y_{t-1} + \gamma y_{t-1} \{1 - \exp(-\theta y_{t-d}^2)\} + \varepsilon_t \quad (3)$$

where Δ is the first-difference operator, $\phi = (\beta - 1)$, and $\theta > 0$ determines the speed of transition between regimes. [Kapetanios, Shin, and Snell \(2003\)](#) indicate that the economic application motivating the model they consider can be found in the purchasing

power parity models of [Sercu, Uppal, and Van Hulle \(1995\)](#) as well as Michael, Nobay, and Peel (1997), where the existence of transactions costs gives rise to nonlinear behavior. Here the idea is that the presence of fixed transactions costs reduces the profitability of arbitrage opportunities when the deviations from purchasing power parity are small rather than large. Therefore, the tendency to return to equilibrium of the underlying system would be stronger only for the larger deviations from equilibrium.¹ The type of nonlinear adjustment just described implies that although it is possible to have $\phi \geq 0$, for y_t to be globally stationary, it is necessary that $\gamma < 0$ and $\phi + \gamma < 0$. Under these conditions, y_t may exhibit unit-root or explosive behavior in the corridor regime for small values of y_{t-d}^2 but not for large values of it.

Imposing $\phi = 0$ and setting $d = 1$ in (3) so that y_t contains a unit root in the corridor regime and the nonlinear effect is assumed to be a function of y_{t-1} , respectively, yields the specific ESTAR model studied by KSS:

$$\Delta y_t = \gamma y_{t-1} \{1 - \exp(-\theta y_{t-1}^2)\} + \varepsilon_t \quad (4)$$

Note that if $\theta = 0$ in (4), then there is a unit root in y_t . Therefore, testing for a unit root in y_t is based on $H_0: \theta = 0$, against the alternative that y_t is a stationary but nonlinear process; that is, $H_a: \theta > 0$. However, given that γ in (4) is not identified under the null, [Kapetanios, Shin, and Snell \(2003\)](#) suggest using a first-order Taylor series approximation of θ around $\theta = 0$ and thus derive the test regression

$$\Delta y_t = \delta y_{t-1}^3 + \text{error}$$

which can be augmented with p lags of Δy_t to allow for the serial correlation

$$\Delta y_t = \delta y_{t-1}^3 + \sum_{j=1}^p \kappa_j \Delta y_{t-j} + \text{error} \quad (5)$$

where the augmentation terms enter in a linear way (as a simplification to obtain more tractable tests).

The null of nonstationarity involves testing $H_0: \delta = 0$ against $H_a: \delta < 0$, for which one calculates the t statistic associated with δ , denoted t_{NL} . [Kapetanios, Shin, and Snell \(2003\)](#) show that the resulting statistic does not have an asymptotic standard normal distribution and so use stochastic simulations to tabulate asymptotic CVs using $T = 1000$ observations with 50,000 replications. Thus far, the analysis assumes that y_t is a zero mean stochastic process. To accommodate stochastic processes with nonzero mean or trend, [Kapetanios, Shin, and Snell \(2003\)](#) recommend demeaning or detrending the data using ordinary least squares (OLS). The resulting statistics for raw, OLS-demeaned, and OLS-detrended data are referred to as t_{NL} , t_{NL}^μ , and t_{NL}^τ , respectively. As for the small-sample properties of the tests, Monte Carlo simulations carried

1. Other instances where the KSS test has been applied because of the (likely) presence of transactions costs and other market rigidities and frictions include [Beechey and Österholm \(2008\)](#) for output and prices, [Gustavsson and Österholm \(2006\)](#) for labor force participation rates, [Zhang \(2013\)](#) for inflation, and [Ghoshray \(2010\)](#) for international coffee prices.

out by [Kapetanios, Shin, and Snell \(2003\)](#) indicate that their tests, based on a nonlinear model, perform best relative to the ADF test in the region of the null hypothesis where the underlying series tends to be relatively more persistent, where the extent of persistence is measured through small values of the parameter θ .

In a subsequent article, [Kapetanios and Shin \(2008\)](#) find that when testing the unit-root null hypothesis against the ESTAR alternative, the generalized least-squares (GLS) detrending (demeaning) procedure advocated by [Elliott, Rothenberg, and Stock \(1996\)](#) achieves further power gains over OLS detrending (demeaning). [Kapetanios and Shin \(2008\)](#) suggest the following two-step procedure for GLS detrending. First, obtain the detrended residuals as $\tilde{y}_t = y_t - \tilde{a} - \tilde{b}t$, where \tilde{a} and \tilde{b} are the respective GLS estimates of a and b obtained from a regression of $y_{\bar{p}} = (y_1, y_2 - \bar{p}y_1, y_3 - \bar{p}y_2, \dots, y_T - \bar{p}y_{T-1})$ on $1_{\bar{p}} = (1, 1 - \bar{p}, 1 - \bar{p}, \dots, 1 - \bar{p})$ and $t_{\bar{p}} = \{1, 2 - \bar{p}, 3 - \bar{p}2, \dots, T - \bar{p}(T - 1)\}$, where $\bar{p} = 1 + \bar{c}/T$. They recommended setting $\bar{c} = -17.5$. According to [Kapetanios and Shin \(2008\)](#), this is the value of \bar{c} for which the asymptotic power of the test under the local alternative is equal to 0.5. Second, the GLS detrending-based unit-root test against the alternative of a ESTAR process, denoted \tilde{t}_{NL}^r , is obtained as the t statistic for testing $H_0: \delta = 0$ in

$$\Delta \tilde{y}_t = \delta \tilde{y}_{t-1}^3 + \sum_{j=1}^p \kappa_j \Delta \tilde{y}_{t-j} + \text{error} \quad (6)$$

where the equation has been augmented with lags of the dependent variable to account for serial correlation. Turning to the case of GLS demeaning, we find we need to modify the preceding steps. Specifically, $t_{\bar{p}}$ is omitted from the first stage, and the relevant value of \bar{p} is calculated using $\bar{c} = -9$, following the recommendation in [Kapetanios and Shin \(2002\)](#). The resulting GLS demeaning-based unit-root test is denoted \tilde{t}_{NL}^μ .

3 Monte Carlo experiments design

Let us consider the following first-order autoregressive process with a unit root

$$y_t = y_{t-1} + e_t \quad (7)$$

where e_t are independently distributed $N(0, 1)$ random variables and $t = 1, \dots, T + 1$. Simulation experiments are carried out for a total of 56 different sample sizes, with $T = 18(2)62, 65(5)100, 110(10)200, 220(20)300, 350(50)500, 600(100)800, 1,000, 1,400$, and 2,000, where, for example, 18(2)62 means that all samples from $T = 18$ up to $T = 62$ increasing in steps of 2 are taken into account, and so on (the same notation is used later when listing significance levels). The time series y_t is generated by setting an initial value $y_{-99} = 0$, and then the first 100 observations are discarded. Each experiment consists of 50,000 Monte Carlo replications. To allow for sampling variability, we repeat the setup just described 50 times so that there will be 50 CVs of the tests for each combination of number of observations T and lag length p . Following [MacKinnon \(1991\)](#), CVs for the KSS and KS tests are calculated at each of 221 significance levels $\{l = 0.0001, 0.0002, 0.0005, 0.001(0.001)0.01, 0.015(0.005)0.990, 0.991(0.001)0.999, 0.9995, 0.9998, \text{ and } 0.9999\}$.

The number of lagged differences of the dependent variable, p , is set equal to $p = 0, 1, \dots, 8$. For $T \leq 20$, $p \leq 1$ is used; for $22 \leq T \leq 24$, $p \leq 2$ is used; for $26 \leq T \leq 28$, $p \leq 3$ is used; for $30 \leq T \leq 32$, $p \leq 4$ is used; for $34 \leq T \leq 36$, $p \leq 6$ is used; and for $T > 36$, all values of p are used. Overall, there are 456 different pairings of T and p . Furthermore, given that, in practice, the lag order is rarely fixed by the user, we also consider data-dependent procedures to select p based on information criteria such as Akaike and Schwarz, denoted AIC and SIC, respectively. In this approach, the optimal number of lags is determined by varying p in (5), or (6), between $p_{\max} = 8$ and $p_{\min} = 0$ lags and choosing the best model according to the criterion being used. Another procedure to select p involves the often known general-to-specific (GTS) algorithm; see, for example, Hall (1994). The idea of this algorithm is to begin by setting some upper bound on p , say, p_{\max} , where $p_{\max} = 0, 1, 2, \dots, 8$, estimating (5), or (6), with $p = p_{\max}$, and testing the statistical significance of $\kappa_{p_{\max}}$ in that equation. If this coefficient is statistically significant, for example, using significance levels of 5% (denoted GTS₅) or 10% (denoted GTS₁₀), one chooses $p = p_{\max}$. Otherwise, the order of the estimated autoregression is reduced by one until the coefficient on the last included lag is statistically different from zero.

The simulated CVs are then used to fit response surface models at each of the 221 significance levels l . These models are a function of T and p when the number of lags is fixed exogenously and are a function of T and p_{\max} when it is determined using any of the four data-dependent procedures.

The specification of the models follows functional forms chosen by, among others, MacKinnon (1991), Cheung and Lai (1995a,b), and Harvey and van Dijk (2006), in which the CVs are regressed on an intercept term and power functions of $(1/T)$ and (p/T) . The functional form finally chosen is

$$CV_{T,p}^l = \pi_{\infty}^l + \sum_{i=1}^4 \pi_i^l \left(\frac{1}{T}\right)^i + \sum_{i=1}^4 \varrho_i^l \left(\frac{p^i}{T}\right) + \epsilon^l \quad (8)$$

where $CV_{T,p}^l$ is the CV estimate at significance level l , T is the number of observations in Δy_t , and p is the number of lags of the dependent variable included to account for serial correlation.² Notice that the functional form in (8) is such that the larger the T , the weaker the dependence of the CVs on p . In addition to this, as $T \rightarrow \infty$, the intercept term, π_{∞}^l , provides an estimate of the corresponding asymptotic CV.

4 Main results

Before presenting the results of the estimation of the response surfaces, we use the data-generating process outlined in the previous section to carry out a small-scale Monte Carlo simulation to quantify the empirical size of the test statistics when asymptotic

2. Including even higher-powered terms generally yielded coefficients that were not statistically different from 0 at the 1% significance level, nor did it lead to any noticeable increase in the \bar{R}^2 for these models.

CVs are used in small samples and the underlying test regressions are augmented with lags of the dependent variable. In the simulation, the time series of interest y_t is generated under the unit-root null hypothesis according to (7), with e_t being drawn from the standard normal distribution, $t = 1, \dots, T + 1$, and $T = 50, 100$, and 200 . Then, for the KSS tests, we estimate the regression given in (5), while for the KS tests, we estimate (6). To simplify the output of the simulation exercise, we consider only values of $p = 0, 4$, and 8 when the lag order is exogenously fixed and consider $p_{\max} = 0, 4$, and 8 when it is determined using AIC, SIC, GTS₅, and GTS₁₀. Lastly, rejection probabilities are computed at the 5% significance level using the asymptotic CVs provided by Kapetanios, Shin, and Snell (2003) and Kapetanios and Shin (2008), and the number of replications is set at 50,000.

Table 1 summarizes the results of the Monte Carlo simulations. Starting with t_{NL} , t_{NL}^{μ} , and t_{NL}^{τ} , we see that the three KSS tests are correctly sized when $p = 0$ and T is large. However, when p is large and T is small, the tests become significantly undersized when the number of lags is fixed exogenously (for all three cases), and also when using SIC (for t_{NL}^{μ} and t_{NL}^{τ}), but become oversized when applying AIC (for all three cases), GTS₅ (for t_{NL}), and GTS₁₀ (for t_{NL} and t_{NL}^{μ}).

Table 1. Size of KSS and KS unit-root tests for different lag lengths

p/T	Fixed			AIC			SIC			GTS ₅			GTS ₁₀		
	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200
0	0.044	0.047	0.048	0.044	0.047	0.048	t_{NL}			0.044	0.047	0.048	0.044	0.047	0.048
4	0.037	0.043	0.046	0.055	0.054	0.051	0.049	0.049	0.049	0.051	0.052	0.051	0.054	0.053	0.052
8	0.034	0.040	0.043	0.057	0.055	0.052	0.049	0.049	0.049	0.054	0.053	0.051	0.056	0.054	0.052
0	0.047	0.048	0.048	0.047	0.048	0.048	t_{NL}^{μ}			0.047	0.048	0.048	0.047	0.048	0.048
4	0.039	0.042	0.046	0.059	0.055	0.053	0.050	0.048	0.048	0.054	0.052	0.051	0.058	0.054	0.052
8	0.032	0.038	0.044	0.056	0.055	0.053	0.039	0.044	0.047	0.051	0.052	0.052	0.055	0.054	0.053
0	0.048	0.047	0.048	0.048	0.047	0.048	t_{NL}^{τ}			0.048	0.047	0.048	0.048	0.047	0.048
4	0.036	0.040	0.045	0.062	0.056	0.055	0.047	0.045	0.047	0.053	0.052	0.053	0.058	0.055	0.054
8	0.023	0.031	0.040	0.054	0.051	0.053	0.037	0.037	0.042	0.047	0.047	0.051	0.050	0.050	0.052
0	0.110	0.082	0.067	0.110	0.082	0.067	\tilde{t}_{NL}^{μ}			0.110	0.082	0.067	0.110	0.082	0.067
4	0.099	0.076	0.065	0.130	0.093	0.071	0.112	0.084	0.067	0.119	0.089	0.070	0.127	0.092	0.072
8	0.085	0.070	0.062	0.124	0.093	0.072	0.100	0.082	0.067	0.119	0.090	0.072	0.125	0.092	0.073
0	0.099	0.075	0.065	0.099	0.075	0.065	\tilde{t}_{NL}^{τ}			0.099	0.075	0.065	0.099	0.075	0.065
4	0.074	0.069	0.061	0.120	0.094	0.075	0.097	0.079	0.067	0.107	0.088	0.072	0.117	0.093	0.075
8	0.057	0.059	0.058	0.114	0.094	0.075	0.085	0.074	0.065	0.104	0.089	0.073	0.109	0.092	0.075

Note: Rejection probabilities at the 5% level. The asymptotic CVs are taken from Kapetanios, Shin, and Snell (2003) and Kapetanios and Shin (2008).

As for the two KS tests, we observe that \tilde{t}_{NL}^{μ} and \tilde{t}_{NL}^{τ} can be substantially oversized, even for $p = 0$ and when T is relatively large.³ The overrejection of the KS tests statistics remains as p increases, regardless of the method used to select the augmentation order of the test regression.

Because of the potential size distortions in both the KSS and KS tests noted in table 1, it is important to have small sample CVs that also address the number of lags included in the model. In table 2, we report OLS estimates of the response surface models based on (8) for 5% significance level; that is, $l = 0.05$, which can be used to obtain CVs for any given T and p . For the KSS tests, a total of 3,315 response surface regressions are estimated (because there are 3 tests multiplied by the 5 criteria, to select p , multiplied by the 221 significance levels). Overall, the chosen functional form performed well. The average coefficient of determination was 0.94; in only 149 (out of 3,315) cases, it was below 0.85. For the KS test, the average coefficient of determination was 0.972; in only 55 (out of 2,210) cases, it was below 0.85. The implied asymptotic 5% CVs reported in table 2 match those reported by Kapetanios, Shin, and Snell (2003) and Kapetanios and Shin (2008), which are in turn based on $T = 1000$ and $p = 0$.

3. Similar findings were reported by Kapetanios and Shin (2008) in table 1.

Table 2. Response surface estimates for the KSS and KS tests at $l = 0.05$

Test	Lags	Intercept	(s.e.)	1/T	1/T ²	1/T ³	1/T ⁴	p/T	p ² /T	p ³ /T	p ⁴ /T	R ²
t_{NL}	Fixed	-2.219	(0.0002)	5.538	-240.7	5639.1	-46315.0	-1.002	0.919	-0.168	0.010	0.926
	AIC	-2.217	(0.0002)	1.808	50.4	-1369.0	8428.6	-3.678	1.128	-0.162	0.008	0.929
	SIC	-2.216	(0.0002)	3.965	-88.7	1193.1	-6391.3	-2.730	0.994	-0.146	0.008	0.743
	GTS ₅	-2.217	(0.0002)	1.481	82.6	-2254.4	15996.9	-1.847	0.456	-0.064	0.003	0.882
	GTS ₁₀	-2.217	(0.0002)	1.776	51.7	-1309.3	7275.2	-2.782	0.738	-0.101	0.005	0.916
t_{NL}^{μ}	Fixed	-2.935	(0.0002)	5.020	-189.3	2403.3	-14360.3	-1.189	0.966	-0.166	0.010	0.934
	AIC	-2.933	(0.0002)	0.857	130.3	-5182.2	44161.3	-4.128	1.242	-0.168	0.009	0.941
	SIC	-2.930	(0.0002)	3.219	1.6	-3689.4	43446.4	-3.268	1.238	-0.162	0.008	0.923
	GTS ₅	-2.932	(0.0002)	0.170	216.0	-7888.8	68362.7	-2.345	0.696	-0.093	0.005	0.808
	GTS ₁₀	-2.933	(0.0002)	0.392	185.0	-6811.5	58198.7	-3.353	0.959	-0.126	0.006	0.915
t_{NL}^{τ}	Fixed	-3.412	(0.0002)	4.963	-265.8	3239.9	-23877.9	-0.835	1.414	-0.248	0.014	0.979
	AIC	-3.412	(0.0003)	0.262	98.1	-5407.4	43853.5	-5.238	1.901	-0.271	0.014	0.931
	SIC	-3.414	(0.0002)	10.547	-642.3	10409.1	-62396.8	-4.278	2.024	-0.285	0.014	0.957
	GTS ₅	-3.414	(0.0003)	1.720	64.3	-6408.0	63425.4	-3.093	1.290	-0.193	0.010	0.907
	GTS ₁₀	-3.414	(0.0003)	-0.115	180.1	-8475.4	74520.1	-4.423	1.647	-0.234	0.012	0.913
\tilde{t}_{NL}^{μ}	Fixed	-2.222	(0.0002)	-22.347	463.6	-8803.9	69057.7	-0.593	0.831	-0.163	0.010	0.992
	AIC	-2.219	(0.0002)	-26.798	786.7	-16054.5	122200.3	-3.683	1.191	-0.175	0.009	0.996
	SIC	-2.218	(0.0002)	-25.288	761.2	-17718.2	149917.3	-2.627	1.135	-0.169	0.009	0.996
	GTS ₅	-2.218	(0.0002)	-27.988	913.9	-19729.3	153080.1	-1.517	0.461	-0.074	0.004	0.995
	GTS ₁₀	-2.218	(0.0002)	-27.460	849.7	-17729.5	135231.4	-2.713	0.804	-0.119	0.006	0.996
\tilde{t}_{NL}^{τ}	Fixed	-2.966	(0.0002)	-17.693	347.2	-6903.5	44876.6	0.162	0.945	-0.196	0.012	0.984
	AIC	-2.966	(0.0003)	-24.320	865.8	-19353.5	142612.0	-4.434	1.459	-0.217	0.012	0.988
	SIC	-2.967	(0.0002)	-16.263	335.9	-9397.4	86400.9	-3.297	1.471	-0.220	0.011	0.995
	GTS ₅	-2.967	(0.0003)	-23.473	890.3	-21982.3	175819.7	-2.091	0.706	-0.114	0.007	0.987
	GTS ₁₀	-2.967	(0.0003)	-24.622	936.9	-22050.9	169449.9	-3.533	1.144	-0.170	0.009	0.986

Note: The standard error of the intercept term appears in parentheses. R^2 denotes the coefficient of determination.

As expected, the residuals of the estimated response surfaces exhibit heteroskedasticity. Thus, to assess the robustness of the OLS results, we also considered estimation using the generalized method of moments (GMM) procedure described in [MacKinnon \(1994, 1996\)](#). For the purposes of our simulation exercise, the GMM procedure amounts to averaging the CVs across the 50 replications for each combination of T and p and scaling all the variables in (8) by the standard error in these replications. Then the resulting equation using the rescaled variables is estimated by OLS. This GMM procedure yields very similar results to those obtained when using OLS and are therefore not reported here.

To assist the reader, table 3 reports the estimated $l = 0.05$ CVs for selected values of T and p based on the response function estimates of table 2.⁴ The corresponding asymptotic CVs tabulated by [Kapetanios, Shin, and Snell \(2003\)](#) and [Kapetanios and Shin \(2008\)](#) are also included for comparison. Interestingly, the implied CVs exhibit dependence on the method used to select the lag length. In some cases, the differences may be noticeable, especially when T is small and p is large. In particular, for a given T , the implied CVs decrease (in an absolute sense) in p when the augmentation order is fixed by the user, while they increase (in an absolute sense) in p_{\max} when it is optimally determined using any data-dependent procedures that are being considered. Similar findings were obtained by [Harvey and van Dijk \(2006\)](#) when estimating response surfaces for the [Hylleberg et al. \(1990\)](#) test for seasonal unit roots with quarterly data.

4. Qualitatively similar results are observed for $l = 0.01, 0.10$, although these are not reported here to save space.

Table 3. Lag order and $l = 0.05$ finite-sample CVs of the KSS and KS tests

p/T	Fixed				AIC				SIC				GTS ₅				GTS ₁₀			
	50	100	200	500	50	100	200	500	50	100	200	500	50	100	200	500	50	100	200	500
0	-2.17	-2.18	-2.20	-2.17	-2.20	-2.21	-2.21	-2.16	t_{NL}	-2.18	-2.20	-2.17	-2.20	-2.21	-2.21	-2.17	-2.20	-2.21	-2.22	-2.22
2	-2.16	-2.18	-2.19	-2.25	-2.24	-2.23	-2.21	-2.21	-2.21	-2.21	-2.21	-2.22	-2.22	-2.22	-2.22	-2.24	-2.23	-2.23	-2.22	-2.22
4	-2.12	-2.16	-2.18	-2.27	-2.24	-2.23	-2.21	-2.21	-2.21	-2.21	-2.21	-2.24	-2.23	-2.23	-2.22	-2.26	-2.24	-2.24	-2.23	-2.23
0	-2.89	-2.90	-2.91	-2.90	-2.92	-2.93	-2.89	-2.89	t_{NL}^{μ}	-2.90	-2.91	-2.89	-2.92	-2.93	-2.93	-2.90	-2.92	-2.93	-2.93	-2.93
2	-2.89	-2.90	-2.91	-2.99	-2.96	-2.95	-2.94	-2.94	-2.93	-2.93	-2.93	-2.95	-2.94	-2.94	-2.94	-2.97	-2.95	-2.95	-2.95	-2.95
4	-2.84	-2.88	-2.90	-3.00	-2.97	-2.95	-2.92	-2.92	-2.92	-2.92	-2.92	-2.95	-2.95	-2.95	-2.94	-2.99	-2.96	-2.96	-2.95	-2.95
0	-3.40	-3.39	-3.39	-3.40	-3.40	-3.41	-3.39	-3.39	t_{NL}^{τ}	-3.36	-3.38	-3.39	-3.40	-3.40	-3.40	-3.40	-3.40	-3.41	-3.41	-3.40
2	-3.35	-3.36	-3.38	-3.50	-3.45	-3.43	-3.44	-3.44	-3.39	-3.39	-3.39	-3.44	-3.42	-3.42	-3.42	-3.48	-3.44	-3.44	-3.43	-3.43
4	-3.25	-3.31	-3.36	-3.49	-3.45	-3.43	-3.37	-3.37	-3.36	-3.36	-3.37	-3.42	-3.41	-3.41	-3.41	-3.46	-3.44	-3.44	-3.43	-3.43
0	-2.54	-2.41	-2.32	-2.55	-2.42	-2.34	-2.54	-2.41	\tilde{t}_{NL}^{μ}	-2.41	-2.33	-2.55	-2.42	-2.34	-2.34	-2.55	-2.42	-2.34	-2.34	-2.21
2	-2.52	-2.40	-2.32	-2.63	-2.46	-2.35	-2.58	-2.43	-2.43	-2.43	-2.34	-2.58	-2.44	-2.35	-2.35	-2.61	-2.45	-2.35	-2.35	-2.35
4	-2.48	-2.38	-2.31	-2.64	-2.47	-2.36	-2.56	-2.42	-2.42	-2.42	-2.33	-2.59	-2.45	-2.35	-2.35	-2.63	-2.46	-2.36	-2.36	-2.36
0	-3.23	-3.11	-3.05	-3.24	-3.14	-3.07	-3.22	-3.10	\tilde{t}_{NL}^{τ}	-3.10	-3.04	-3.23	-3.13	-3.06	-3.06	-3.23	-3.14	-3.07	-2.93	-2.93
2	-3.17	-3.09	-3.03	-3.33	-3.19	-3.09	-3.27	-3.13	-3.13	-3.13	-3.05	-3.27	-3.15	-3.08	-3.08	-3.31	-3.18	-3.09	-3.09	-3.09
4	-3.10	-3.05	-3.01	-3.34	-3.19	-3.09	-3.24	-3.11	-3.11	-3.11	-3.05	-3.28	-3.16	-3.08	-3.08	-3.32	-3.18	-3.09	-3.09	-3.09

Note: For $T = 1000$, the CVs of t_{NL} , t_{NL}^{μ} and t_{NL}^{τ} are from Kapetanios, Shin, and Snell (2003); for \tilde{t}_{NL}^{μ} and \tilde{t}_{NL}^{τ} , they are taken from Kapetanios and Shin (2008).

Finally, the various KSS and KS test statistics and finite-sample CVs can be used to compute p -values. For this, we follow MacKinnon (1994, 1996) by estimating the regression

$$\Phi^{-1}(l) = \lambda_0^l + \lambda_1^l \widehat{CV}^l + \lambda_2^l \left(\widehat{CV}^l \right)^2 + v^l \quad (9)$$

where Φ^{-1} is the inverse of the cumulative standard normal distribution at each of the 221 quantiles and \widehat{CV}^l is the fitted value from (8) at the l quantile. As in Harvey and van Dijk (2006), (9) is estimated by OLS using 15 observations, made up of the actual quantile and the 7 quantile observations on either side of the desired quantile.⁵ Denoting the calculated test statistic as S , we estimate its associated approximate probability value as

$$p\text{-value} = \Phi \left(\widehat{\lambda}_0^l + \widehat{\lambda}_1^l S + \widehat{\lambda}_2^l S^2 \right)$$

where $\widehat{\lambda}_0^l$, $\widehat{\lambda}_1^l$, and $\widehat{\lambda}_2^l$ are the OLS parameter estimates from (9). The `kssur` and `ksur` commands, which will be presented in the next section, also compute the approximate p -value using the procedure just described.

5 The `kssur` and `ksur` commands

The `kssur` and `ksur` commands calculate both the KSS and KS test statistics along with their associated finite-sample CVs for $l = 0.01, 0.05$, and 0.10 . The commands also calculate the approximate p -values.⁶ The estimation of CVs and approximate p -values permits different combinations of the number of observations, T , and the lag order in the test regression, p , where the latter can be either specified by the user or optimally selected using a data-dependent procedure.

5.1 Syntax

Before using the `kssur` and `ksur` commands, one must `tsset` the data, much like as with all other Stata time-series commands. The syntax for the `kssur` command is

```
kssur varname [if] [in], [noprint maxlag(integer) noconstant trend]
```

The syntax for the `ksur` command is

```
ksur varname [if] [in], [noprint maxlag(integer) trend]
```

Here note that *varname* may not contain gaps.

5. For $l \leq 0.004$ and $l \geq 0.996$, we use the actual quantile and the 14 observations closest to the desired quantile because there will not be 7 on either side.

6. Our codes benefited greatly from `dfgls`, developed by Baum and Sperling (2000). Baum (2016) was also a particularly useful reference.

5.2 Options

The `kssur` command allows the following options:

`noprnt` specifies that the results are to be returned but not printed.

`maxlag(integer)` sets the number of lags to be included in the test regression to account for residual serial correlation. The default is that `kssur` sets the number of lags following [Schwert \(1989\)](#), with the formula $\text{maxlag}() = \text{integer}\{12(T/100)^{0.25}\}$, where T is the total number of observations.

`noconstant` and `trend` specify the modeling of intercepts and trends. Use `noconstant` when *varname* is a zero mean stochastic process, in which case [Kapetanios, Shin, and Snell \(2003\)](#) recommend using the raw data. Use `trend` when *varname* is a nonzero trend stochastic process, in which case [Kapetanios, Shin, and Snell \(2003\)](#) recommend detrending the data using OLS. By default, `kssur` considers *varname* to be a nonzero mean stochastic process, in which case [Kapetanios, Shin, and Snell \(2003\)](#) recommend demeaning the data using OLS.

For the command `ksur`, the options `noprnt` and `maxlag()` work the same as before. For the option that specifies the modeling of intercepts and trends; use `trend` when *varname* is a nonzero trend stochastic process, in which case [Kapetanios and Shin \(2008\)](#) recommend detrending the data using GLS. By default, `ksur` considers *varname* to be a nonzero mean stochastic process, in which case [Kapetanios and Shin \(2008\)](#) recommend demeaning the data using GLS.

5.3 Stored results

`kssur` stores the following in `r()`:

Scalars

<code>r(N)</code>	number of observations	<code>r(maxp)</code>	last time period used in the
<code>r(minp)</code>	first time period used in the test regression		test regression

Macros

<code>r(varname)</code>	variable name	<code>r(tsfmt)</code>	time series format of the
<code>r(treat)</code>	either <code>rawdata</code> , <code>demeaned</code> , or <code>detrended</code>		time variable

Matrices

<code>r(results)</code>	results statistics table
-------------------------	--------------------------

In turn, `ksur` stores the following in `r()`:

Scalars			
<code>r(N)</code>	number of observations	<code>r(maxp)</code>	last time period used in the
<code>r(minp)</code>	first time period used in the test regression		test regression
Macros			
<code>r(varname)</code>	variable name	<code>r(tsfmt)</code>	time series format of the
<code>r(treat)</code>	either <code>demeaned</code> or <code>detrended</code>		time variable
Matrices			
<code>r(results)</code>	results statistics table		

6 Empirical application

The idea of market integration has often been assessed by examining the validity of the law of one price (LOOP), either by testing whether the prices of identical products traded in different locations are the same once they are converted to a common currency, as in the absolute version of the law, or by testing whether price discrepancies can be best described as stationary processes, as in the relative version of the law; see, for example, [Froot and Rogoff \(1995\)](#). Thus, testing the validity of the LOOP in its relative version is essentially a question of whether price differentials contain a unit root.

This section illustrates the use of the `kssur` and `ksur` commands. We draw on [Ghoshray \(2010\)](#), who tests the validity of the LOOP among the monthly price series of the four best-known types of coffee, namely, unwashed Arabicas (mainly coffee from Brazil), Colombian mild Arabicas (mainly coffee from Colombia), other mild Arabicas (mainly coffee from other Latin American countries), and Robusta coffee (mainly coffee grown in African countries and Southeast Asia). The coffee prices, denoted `br`, `co`, `om`, and `ro`, respectively, are considered after applying the logarithmic transformation. The sample period runs from 1990m1 to 2004m1, that is, a total of 169 time observations for each series, and the data were downloaded from the website of the International Coffee Organization at <http://www.ico.org>. We select the same sample period used by [Ghoshray \(2010\)](#) so we can replicate and supplement the results reported in that article.

We begin by loading the dataset and declaring that it has a time-series format:

```
. use http://www2.warwick.ac.uk/fac/soc/economics/staff/jsmith/
> research/coffeedata
. tsset date
    time variable:  date, 1990m1 to 2004m1
        delta: 1 month
```

Next, let us say that we would like to test whether the price differential between `br` and `co`, that is, `brco`, contains a unit root, against the alternative that it is a globally stationary ESTAR process. Given that `brco` has a nonzero mean, the relevant KSS statistic is based on OLS-demeaned data, which are implemented using the default `constant` option. Setting $p = 3$ lags, we see that the results of applying the `kssur` command are

```
. kssur brco, maxlag(3)
Kapetanios, Shin & Snell (2003) test results for 1990m5 - 2004m1
Variable name: brco
Ho: Unit root
Ha: Stationary nonlinear ESTAR model
OLS demeaned data
```

Criteria	Lags	KSS stat.	p-value	1% cv	5% cv	10% cv
FIXED	3	-2.279	0.211	-3.454	-2.903	-2.625
AIC	1	-2.072	0.320	-3.525	-2.954	-2.667
SIC	0	-2.467	0.148	-3.484	-2.926	-2.644
GTS05	0	-2.467	0.151	-3.508	-2.942	-2.657
GTS10	1	-2.072	0.319	-3.520	-2.951	-2.665

In turn, for the KS test, the relevant statistic is based on GLS-demeaned data, which can be implemented as follows:

```
. ksur brco, maxlag(3)
Kapetanios & Shin (2008) test results for 1990m5 - 2004m1
Variable name: brco
Ho: Unit root
Stationary nonlinear ESTAR model
GLS demeaned data
```

Criteria	Lags	KS stat.	p-value	1% cv	5% cv	10% cv
FIXED	3	-2.035	0.102	-2.892	-2.327	-2.045
AIC	1	-1.871	0.157	-2.970	-2.379	-2.085
SIC	0	-2.289	0.059	-2.930	-2.354	-2.065
GTS05	0	-2.289	0.061	-2.953	-2.368	-2.075
GTS10	1	-1.871	0.156	-2.965	-2.376	-2.082

Table 4 summarizes the results of applying the `kssur` and `ksur` commands, as already illustrated, to the six price differentials that can be constructed with the variables `br`, `co`, `om`, and `ro`. This table also reports the results of performing the ADF test when a constant is included in the test regression. For the ADF test, the p -values are those estimated by [MacKinnon \(1996\)](#).⁷

7. [Ghoshray \(2010\)](#) in table 2 computes the ADF and KSS unit-root tests and performs inference at the 1% and 5% significance levels. Finite sample CVs from [MacKinnon \(1996\)](#) are used for the ADF test, while asymptotic CVs from [Kapetanios, Shin, and Snell \(2003\)](#) are used for the KSS test.

Table 4. Application of the ADF, KSS, and KS tests to coffee price differentials

Test	Series	Fixed		AIC		SIC		GTS ₅		GTS ₁₀	
		<i>t</i> stat.	<i>p</i> -value	<i>t</i> stat.	<i>p</i> -value	<i>t</i> stat.	<i>p</i> -value	<i>t</i> stat.	<i>p</i> -value	<i>t</i> stat.	<i>p</i> -value
ADF	brco	-2.619	[0.091]								
	brom	-2.938	[0.043]								
	brro	-2.540	[0.108]								
	coom	-2.049	[0.266]								
	coro	-1.923	[0.321]								
	omro	-1.843	[0.359]								
KSS	brco	-2.279	[0.211]	-2.072	[0.320]	-2.467	[0.148]	-2.467	[0.151]	-2.072	[0.319]
	brom	-3.064	[0.032]	-2.932	[0.053]	-2.932	[0.049]	-2.932	[0.051]	-2.932	[0.053]
	brro	-2.843	[0.058]	-2.938	[0.052]	-2.938	[0.048]	-2.938	[0.051]	-2.938	[0.052]
	coom	-2.858	[0.056]	-2.955	[0.050]	-2.955	[0.046]	-2.955	[0.048]	-2.955	[0.050]
	coro	-4.270	[0.001]	-4.169	[0.001]	-3.662	[0.006]	-4.169	[0.001]	-4.169	[0.001]
	omro	-2.723	[0.079]	-2.898	[0.058]	-2.898	[0.054]	-2.898	[0.056]	-2.898	[0.057]
KS	brco	-2.035	[0.102]	-1.871	[0.157]	-2.289	[0.059]	-2.289	[0.061]	-1.871	[0.156]
	brom	-2.756	[0.015]	-2.692	[0.022]	-2.692	[0.020]	-2.692	[0.021]	-2.692	[0.022]
	brro	-2.563	[0.026]	-2.668	[0.024]	-2.668	[0.022]	-2.668	[0.023]	-2.668	[0.023]
	coom	-2.884	[0.010]	-2.980	[0.010]	-2.980	[0.009]	-2.980	[0.009]	-2.980	[0.010]
	coro	-4.039	[0.000]	-3.956	[0.000]	-3.491	[0.002]	-3.956	[0.000]	-3.956	[0.000]
	omro	-2.494	[0.032]	-2.668	[0.024]	-2.668	[0.022]	-2.668	[0.023]	-2.668	[0.023]

Interestingly, the results of the ADF test indicate that at the 10% significance level, there is evidence supporting the LOOP in only 2 out of 6 cases, namely, for `brco` and `brom`. In contrast, using the same 10% significance level, the results of the KSS and KS tests indicate that the LOOP holds in 5 out of 6 cases. Comparing the results of the KSS and KS tests across the different criteria, both at the 10% significance level, we observe that inference is not affected in the overwhelming majority of the cases; the exception is when the KS test is applied to `brco`, in which case the unit-root null is rejected when using SIC and GTS₅ but not for the other criteria. Using a 5% significance level, we see that several test statistics are close to their corresponding CV, so inference may change. In summary, the results of the KSS and KS tests tend to be more supportive of the LOOP. Ghoshray (2010) argues that the presence of quality differences and transaction costs may be some of the factors behind the nonlinear behavior.

7 Conclusions

In this article, we fit response surface models for the CVs of the KSS and KS unit-root tests. The models were fit as a function of the number of observations, T , and the lags of the dependent variable in the test regressions, p , for 221 significance levels. The lag length can be determined either exogenously by the user or endogenously using a data-dependent procedure. The results suggest that the method used to select the order of the augmentation affects the finite-sample CVs.

The new commands `kssur` and `ksur` can easily be used to calculate the KSS and KS test statistics, along with both finite-sample CVs and approximate p -values. As an empirical application, we illustrated the new commands by examining whether the LOOP holds among the prices of four types of coffee. We found that in some instances, inference may depend upon the procedure used to select the number of lags that are included in the test regressions.

8 Acknowledgments

Jesús Otero acknowledges financial support received from the Fondo de Investigaciones de la Universidad del Rosario through project DVG-144. We thank Christopher F. Baum, Jeisson Cárdenas, and Sergio Montoya for their comments and suggestions on developing the Stata commands `kssur` and `ksur`. We also thank an anonymous referee for detailed comments and suggestions on an earlier version of the paper. The usual disclaimer applies.

9 References

- Baum, C. F. 2016. *An Introduction to Stata Programming*. 2nd ed. College Station, TX: Stata Press.
- Baum, C. F., and R. Sperling. 2000. `dfgls`: Stata module to compute Dickey–Fuller/GLS unit root test. Statistical Software Components S410001, Department of Economics, Boston College. <https://ideas.repec.org/c/boc/bocode/s410001.html>.

- Beechey, M., and P. Österholm. 2008. Revisiting the uncertain unit root in GDP and CPI: Testing for non-linear trend reversion. *Economics Letters* 100: 221–223.
- Cheung, Y.-W., and K. S. Lai. 1995a. Lag order and critical values of the augmented Dickey–Fuller test. *Journal of Business and Economic Statistics* 13: 277–280.
- . 1995b. Lag order and critical values of a modified Dickey–Fuller test. *Oxford Bulletin of Economics and Statistics* 57: 411–419.
- 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.
- Elliott, G., T. J. Rothenberg, and J. H. Stock. 1996. Efficient tests for an autoregressive unit root. *Econometrica* 64: 813–836.
- Froot, K., and K. Rogoff. 1995. Perspectives on PPP and long-run real exchange rates. In *Handbook of International Economics*, ed. G. M. Grossman and K. Rogoff, vol. 3, 1647–1688. Amsterdam: North Holland.
- Ghoshray, A. 2010. The extent of the world coffee market. *Bulletin of Economic Research* 62: 97–107.
- Gustavsson, M., and P. Österholm. 2006. The informational value of unemployment statistics: A note on the time series properties of participation rates. *Economics Letters* 92: 428–433.
- Hall, A. 1994. Testing for a unit root in time series with pretest data-based model selection. *Journal of Business and Economic Statistics* 12: 461–470.
- Harvey, D. I., and D. van Dijk. 2006. Sample size, lag order and critical values of seasonal unit root tests. *Computational Statistics and Data Analysis* 50: 2734–2751.
- Hylleberg, S., R. F. Engle, C. W. J. Granger, and B. S. Yoo. 1990. Seasonal integration and cointegration. *Journal of Econometrics* 44: 215–238.
- Kapetanios, G., and Y. Shin. 2002. GLS detrending for nonlinear unit root tests. Queen Mary University of London, School of Economics and Finance, Working Paper No. 472. <https://ideas.repec.org/p/qmw/qmwecw/wp472.html>.
- . 2008. GLS detrending-based unit root tests in nonlinear STAR and SETAR models. *Economics Letters* 100: 377–380.
- Kapetanios, G., Y. Shin, and A. Snell. 2003. Testing for a unit root in the nonlinear STAR framework. *Journal of Econometrics* 112: 359–379.
- MacKinnon, J. G. 1991. Critical values for cointegration tests. In *Long-Run Economic Relationships: Readings in Cointegration*, ed. R. F. Engle and C. W. J. Granger, 267–276. Oxford: Oxford University Press.

- . 1994. Approximate asymptotic distribution functions for unit-root and cointegration tests. *Journal of Business and Economic Statistics* 12: 167–176.
- . 1996. Numerical distribution functions for unit root and cointegration tests. *Journal of Applied Econometrics* 11: 601–618.
- Michael, P., A. R. Nobay, and D. A. Peel. 1997. Transactions costs and nonlinear adjustment in real exchange rates; An empirical investigation. *Journal of Political Economy* 105: 862–879.
- Nelson, C. R., and C. I. Plosser. 1982. Trends and random walks in macroeconomic time series: Some evidence and implications. *Journal of Monetary Economics* 10: 139–162.
- Patterson, K. 2012. *Unit Root Tests in Time Series. Volume 2: Extensions and Developments*. Chippingham: Palgrave Macmillan.
- Said, S. E., and D. A. Dickey. 1984. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika* 71: 599–607.
- Schwert, G. W. 1989. Tests for unit roots: A Monte Carlo investigation. *Journal of Business and Economic Statistics* 7: 147–159.
- Sercu, P., R. Uppal, and C. Van Hulle. 1995. The exchange rate in the presence of transaction costs: Implications for tests of purchasing power parity. *Journal of Finance* 50: 1309–1319.
- Zhang, L. 2013. Revisiting the empirics of inflation in China: A smooth transition error correction approach. *Economics Letters* 119: 68–71.

About the authors

Jesús Otero is a professor of economics at the Universidad del Rosario in Colombia. He received his PhD in economics from Warwick University in the United Kingdom. His field of study is applied time-series econometrics.

Jeremy Smith is a professor of economics at the University of Warwick. He received his PhD in economics from the Australian National University. His field of study is applied time-series econometrics.