CAMBRIDGE
UNIVERSITY PRESS

Testing for Cointegration in Nonlinear Smooth Transition Error Correction Models
Author(s): George Kapetanios, Yongcheol Shin and Andy Snell
Source: *Econometric Theory*, Vol. 22, No. 2 (Apr., 2006), pp. 279-303
Published by: Cambridge University Press
Stable URL: http://www.jstor.org/stable/4093226
Accessed: 19-01-2018 15:10 UTC

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at http://about.jstor.org/terms

JSTOR

Cambridge University Press is collaborating with JSTOR to digitize, preserve and extend access to *Econometric Theory*

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

Econometric Theory, **22**, 2006, 279–303. Printed in the United States of America.
DOI: 10.1017/S0266466606060129

TESTING FOR COINTEGRATION IN NONLINEAR SMOOTH TRANSITION ERROR CORRECTION MODELS

George Kapetanios
Queen Mary, University of London

Yongcheol Shin
Leeds University Business School

Andy Snell
Edinburgh School of Economics, University of Edinburgh

This paper proposes a new testing procedure to detect the presence of a cointegrating relationship that follows a globally stationary smooth transition process. In the context of nonlinear smooth transition error correction models (ECMs) we provide two simple operational versions of the tests. First, we obtain the associated nonlinear ECM-based tests. Second, we derive the nonlinear analogue of the residual-based test for cointegration in linear models. We derive the asymptotic distributions of the proposed tests. Monte Carlo simulation exercises confirm that our proposed tests have much better power than the linear counterparts against the alternative of a globally stationary nonlinear cointegrating process. In an application to the price-dividend relationship, our test is able to find cointegration, whereas the linear-based tests fail to do so.

1. INTRODUCTION

The joint investigation of nonstationarity and nonlinearity in economics has recently assumed great significance. There has also been increasing concern that the information revealed by the analysis of a linear time series model may be insufficient to give definitive inference on important economic hypotheses. In this regard attention has fallen almost exclusively on regime-switching-type models, though there is no established theory suggesting a unique approach for specifying econometric models that embed various types of change in regimes.

We are grateful to an associate editor, two anonymous referees, Richard Baillie, In Choi, Atanas Christev, Hira Koul, Richard Harris, Cheng Hsiao, Changjin Kim, Joon Park, Peter Schmidt, Yoonjae Whang, Jeff Wooldridge, and seminar participants at University of Edinburgh, Heriot-Watt University, Korea University, University of Leeds, Michigan State University, University of Newcastle, and Sungkunkwan University for their helpful comments. Partial financial support from the ESRC (grant R000223399) is gratefully acknowledged. The usual disclaimer applies. Address correspondence to Yongcheol Shin, Economics Division, Leeds University Business School, Room 1.07, Maurice Keyworth Building, University of Leeds, Leeds LS2 9JT, United Kingdom; e-mail: y.shin@leeds.ac.uk.

© 2006 Cambridge University Press 0266-4666/06 $12.00

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

Estimation of cointegrated models subject to (regime-switching) nonlinear dynamics has been quite prominent. Balke and Fomby (1997) have popularized an estimation method where the dynamics are subject to three-regime threshold cointegration, the case where a process may follow a unit root in a middle regime while at the same time being globally geometrically ergodic. Another popular nonlinear scheme being applied in conjunction with cointegration is based on the smooth transition regression (STR), following Granger and Teräsvirta (1993). Michael, Nobay, and Peel (1997) analyze nonlinearity in the long-run purchasing power parity (PPP) relationship by adopting a null of a unit root for real exchange rates while taking an alternative hypothesis of nonlinear smooth transition autoregressive (STAR) stationarity. Finally, Psaradakis, Sola, and Spagnolo (2004) consider Markov-switching error correction models (ECMs) in which deviations from the long-run equilibrium follow a process that is nonstationary in one state of nature and mean-reverting in the other. These papers argue forcefully that the assumption of linear adjustment is likely to be too limited in various economic situations, particularly where transaction costs and policy interventions are present. For similar modeling approaches see Granger and Swanson (1996), Corradi, Swanson, and White (2000), Choi and Saikkonen (2004), and Saikkonen and Choi (2004).

With regard to statistical inference, there is a large literature proposing tests for unit roots against threshold autoregressive (TAR) alternatives, e.g., Enders and Granger (1998), Caner and Hansen (2001), Bec, Guay, and Guerre (2004), and Kapetanios and Shin (2006). In particular, Kapetanios, Shin, and Snell (2003) develop the testing framework for a linear unit root null against an exponential STAR alternative and show that their proposed test is more powerful than the Dickey–Fuller (1979) test. See also Kiliç (2003) for its bootstrap-based extension. There is also a growing literature analyzing cointegration subject to nonlinear dynamic adjustment, e.g., Escribano and Pfann (1998), Escribano and Mira (2002), Bec and Rahbek (2004), Escribano (2004), and Saikkonen (2005).

However, studies that directly specify and derive the asymptotic properties of cointegration tests against nonlinear alternatives are still relatively few. For example, although Gallagher and Taylor (2001) and Hansen and Seo (2002) conduct cointegration analysis with nonlinear alternatives in mind, they all adopt linear cointegration tests to establish the existence of cointegration and only allow nonlinearity to enter the analysis at the estimation stage once cointegration has been established. Although such linear tests will have power against nonlinear alternatives, it seems far more sensible to use a test in the first stage that is designed to have power against the alternative of interest, namely, nonlinear dynamic adjustment.

In this paper we address this issue by deriving a cointegration test designed to have power against alternatives where the cointegrating error is stationary and follows STR dynamics. We model an equilibrium process where its error correction adjustment is slower when the cointegrating residual is close to zero

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

and where the change in speed of this adjustment process is assumed smooth rather than sharp (as with TAR models). Therefore, we focus on the case in which the EC term follows a globally stationary process under the STR alternative. In particular, our approach is motivated by the implication of economic theory applied to asset arbitrage under noise trading and transaction costs arising from the bid-ask spread. Campbell and Kyle (1993) develop a model of finite-lived traders some of whom misperceive true fundamental asset values and show that the larger are the pricing errors in these models, the larger is the expected degree of arbitrage and hence the speedier is the price response to disequilibrium. Snell and Tonks (2003) show that in an adapted version of the sequential auction model of Kyle (1985) where informed agents who suffer liquidity costs trade with a risk neutral market maker, the prices offered by the market maker follow an ECM with a speed of adjustment parameter that moves smoothly in response to the conditional variance of the disequilibrium price error. To put it more concisely, higher disequilibrium pricing errors are associated with larger ECM adjustment parameters and faster rates of error correction toward equilibrium.

Using a general nonlinear exponential STR (ESTR) ECM framework and following a pragmatic residual-based two-step procedure in the style of Engle and Granger (1987), we propose that a null hypothesis of no cointegration against an alternative of a globally stationary ESTR cointegration be tested directly by examining the significance of the parameter controlling the degree of nonlinearity in the speed of adjustment. We develop two operational test statistics, denoted F_{NEC} (t_{NEC}) and t_{NEG}, respectively, and derive their asymptotic distributions. The F_{NEC} (t_{NEC}) test refers to the F-type (t-type) statistic obtained directly from the nonlinear ESTR error correction regression, whereas the t_{NEG} test is the nonlinear analogue to the Engle and Granger (EG) statistic for linear cointegration.

The small-sample performance of the suggested tests is compared to that of the linear EG and Johansen (1995) tests via Monte Carlo experiments. We find that our proposed nonlinear tests have good size and superior power properties compared to the linear tests. In particular, both the F_{NEC} and t_{NEC} tests are superior to both linear or nonlinear EG tests when the regressors are weakly exogenous in a cointegrating regression. This supports similar findings made in linear models by Kremers, Ericsson, and Dolado (1992), Hansen (1995), Banerjee, Dolado, and Mestre (1998), Arranz and Escribano (2000), and Pesaran, Shin, and Smith (2001) that the EG test loses power relative to ECM-based cointegration tests because of the loss of potentially valuable information from the correlation between the regressors and the underlying disturbances.

We provide an application to investigating the presence of cointegration of asset prices and dividends for eleven stock portfolios (Germany, Belgium, Canada, Denmark, France, Ireland, Italy, Japan, Netherlands, United Kingdom, and United States) allowing for nonlinear STR adjustment to equilibrium. Interestingly, our new tests are able to reject the null of no cointegration in majority

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

cases, whereas the linear EG test rejects only twice. Given the strength of evidence in favor of ESTR cointegration we also estimate adjustment parameters under the alternative, and we find that these estimates are well defined in all cases. We further evaluate the impulse response functions (IRFs) of the error correction term with respect to initial impulses of one to four standard deviation shocks. The striking finding is that the time taken to recover one-half of a one standard deviation shock varies between 5 and 20 years, whereas the time taken to recover one-half of larger shocks varies between just 4 to 18 months. This implies that data periods dominated by extreme volatility may display substantial reversion of prices toward their NPV (net present value) relationship, although in "calmer" times where the error in the NPV relationship takes on smaller values, the process driving it may well look like a unit root.

The plan of the paper is as follows. Section 2 discusses general modeling issues and derives the nonlinear STR error correction models. Section 3 develops the proposed test statistics and derives their asymptotic distributions. Section 4 evaluates the small-sample performance of the proposed tests that take account of the STR nonlinear nature of the alternative. Section 5 presents an empirical application to price-dividend relationships. Section 6 contains some concluding remarks. Mathematical proofs are collected in the Appendix.

2. NONLINEAR STR ERROR CORRECTION MODELS

We begin with the following general nonlinear vector error correction (VEC) model for the $n \times 1$ vector of I(1) stochastic processes, $\mathbf{z}_t$:

$$\Delta \mathbf{z}_t = \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{z}_{t-1} + g(\boldsymbol{\beta}'\mathbf{z}_{t-1}) + \sum_{i=1}^{p} \boldsymbol{\Gamma}_i \Delta \mathbf{z}_{t-i} + \boldsymbol{\varepsilon}_t, \qquad t = 1,2,\dots,T, \tag{2.1}$$

where $\boldsymbol{\alpha}(n \times r)$, $\boldsymbol{\beta}$ $(n \times r)$, and $\boldsymbol{\Gamma}_i$ $(n \times n)$ are parameter matrices with $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ of full column rank and $g : \mathbb{R}^r \to \mathbb{R}^n$ is a nonlinear function. A conventional linear VEC model is obtained by choosing $g(\boldsymbol{\beta}'\mathbf{z}_{t-1}) = \mathbf{0}$ or $g(\boldsymbol{\beta}'\mathbf{z}_{t-1}) = -\boldsymbol{\alpha}\boldsymbol{\mu}'$ where $\boldsymbol{\mu}$ is an $n \times 1$ vector of level parameters.

Recently, analogues of Granger's representation theorem have been studied in the context of the nonlinear VEC model, (2.1), e.g., Escribano and Mira (2002) and Bec and Rahbek (2004). Here we follow Saikkonen (2005) and make the following assumption.

Assumption 1. (i) The errors $\boldsymbol{\varepsilon}_t$ in (2.1) are $iid(\mathbf{0},\boldsymbol{\Sigma})$, with $\boldsymbol{\Sigma}$ being an $n \times n$ positive definite matrix and $E|\boldsymbol{\varepsilon}_t|^{\ell} < \infty$ for some $\ell > 6$. (ii) The distribution of $\boldsymbol{\varepsilon}_t$ in (2.1) is absolutely continuous with respect to the Lebesgue measure and has a density that is bounded away from zero on compact subsets of $\mathbb{R}^n$. (iii) The initial observations $\mathbf{Z}_0 \equiv (\mathbf{z}_{-p},\dots,\mathbf{z}_0)$ are given. (iv) Let $A(z)$ be given by $(1-z)\mathbf{I}_n - \boldsymbol{\alpha}\boldsymbol{\beta}'z - \sum_{i=1}^{p}\boldsymbol{\Gamma}_i(1-z)z^i$. If $\det A(z) = 0$, then $|z| > 1$ or $z = 1$,

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

where the number of unit roots is equal to $n - r$. (v) $g(\bullet)$ is asymptotically no greater than a linear function of $\mathbf{x}_t$.

Assumptions 1(i)–(iv) are standard in the cointegration literature, whereas Assumption 1(v) is imposed to deal with nonlinearity of the underlying model (2.1). Under Assumption 1, Saikkonen (2005, Thm. 2) proves that there exists a choice of initial values $\mathbf{z}_{-p},\ldots,\mathbf{z}_0$ such that $\Delta\mathbf{z}_t$ and $\boldsymbol{\beta}'\mathbf{z}_{t-1}$ are strictly stationary.

In this paper we aim to analyze at most one conditional long-run cointegrating relationship between y_t and $\mathbf{x}_t$, and we focus on the conditional modeling of the scalar variable y_t given the k-vector $\mathbf{x}_t$ $(k = n - 1)$ and the past values of $\mathbf{z}_t$ and $\mathbf{Z}_0$, where we decompose $\mathbf{z}_t = (y_t, \mathbf{x}_t')'$. For this we rewrite (2.1) as

$$\Delta\mathbf{z}_t = \boldsymbol{\alpha} u_{t-1} + g(u_{t-1}) + \sum_{i=1}^{p} \boldsymbol{\Gamma}_i \Delta\mathbf{z}_{t-i} + \boldsymbol{\varepsilon}_t, \qquad t = 1,2,\ldots,T, \tag{2.2}$$

where $\boldsymbol{\alpha}$ is an $n \times 1$ vector of adjustment parameters and

$$u_t = y_t - \boldsymbol{\beta}_x'\mathbf{x}_t, \tag{2.3}$$

with $\boldsymbol{\beta}_x$ being a $k \times 1$ vector of cointegrating parameters.

We now make the following assumption.

Assumption 2. (i) The nonlinear function $g(\cdot)$ in (2.2) follows the exponential smooth transition functional form

$$g(u_{t-1}) = -\boldsymbol{\varphi} u_{t-1} e^{-\theta(u_{t-1}-c)^2}, \tag{2.4}$$

where we assume $\theta \geq 0$ for identification purposes and c is a transition parameter. (ii) Partition $\boldsymbol{\alpha} = (\zeta, \boldsymbol{\alpha}_x')'$ and $\boldsymbol{\varphi} = (\gamma, \boldsymbol{\varphi}_x')'$ conformably with $\mathbf{z}_t = (y_t, \mathbf{x}_t')'$. Then, $\boldsymbol{\alpha}_x = \boldsymbol{\varphi}_x = \mathbf{0}$. (iii) There is no cointegration among the k-vector of I(1) variables, $\mathbf{x}_t$. (iv) $\zeta < 0$.

Assumptions 2(ii) and (iii) imply that the process $\mathbf{x}_t$ is weakly exogenous and therefore the parameters of interest in (2.6) are variation-free from the parameters in (2.7), e.g., Pesaran et al. (2001). Assumptions 2(i) and (iv) imply that Assumption 1(v) is satisfied but also ensure that the underlying (nonlinear) error correction mechanism in (2.2) is globally stationary. In practice various functional forms of $g(\bullet)$ can also be considered to allow for the presence of nonlinear adjustment to the error correction mechanism, but we focus on ESTR here.

Next, partitioning $\boldsymbol{\varepsilon}_{\mathrm{t}}$ conformably with $\mathbf{z}_t$ as $\boldsymbol{\varepsilon}_{\mathrm{t}} = (\varepsilon_{yt}, \boldsymbol{\varepsilon}_{xt}')'$ and its variance matrix as $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{yy} & \boldsymbol{\sigma}_{yx} \\ \boldsymbol{\sigma}_{xy} & \boldsymbol{\Sigma}_{xx} \end{pmatrix}$, we may express ε_{yt} conditionally in terms of $\boldsymbol{\varepsilon}_{\mathrm{xt}}$ as

$$\varepsilon_{yt} = \boldsymbol{\sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\varepsilon}_{xt} + e_t, \tag{2.5}$$

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

where $e_t \sim iid(0, \sigma_e^2)$, $\sigma_e^2 \equiv \sigma_{yy} - \boldsymbol{\sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\sigma}_{xy}$, and e_t is uncorrelated with $\boldsymbol{\varepsilon}_{\mathrm{xt}}$ by construction. Substituting (2.5) and (2.4) into (2.2), partitioning $\boldsymbol{\Gamma}_i = (\boldsymbol{\gamma}_{yi}', \boldsymbol{\Gamma}_{xi}')'$, $i = 1, \ldots, p$, defining $\phi = \zeta - \gamma$, and under Assumption 2, we obtain the following conditional exponential smooth transition regression error correction model (STR ECM) for Δy_t and the marginal vector autoregression (VAR) model for $\Delta \mathbf{x}_t$:[1]

$$\Delta y_t = \phi u_{t-1} + \gamma u_{t-1}(1 - e^{-\theta(u_{t-1}-c)^2}) + \boldsymbol{\omega}'\Delta\mathbf{x}_t + \sum_{i=1}^{p} \boldsymbol{\psi}_i' \Delta\mathbf{z}_{t-i} + e_t, \tag{2.6}$$

$$\Delta\mathbf{x}_t = \sum_{i=1}^{p} \boldsymbol{\Gamma}_{xi}\Delta\mathbf{z}_{t-i} + \boldsymbol{\varepsilon}_{xt}, \tag{2.7}$$

where $\boldsymbol{\omega} \equiv \boldsymbol{\Sigma}_{xx}^{-1}\boldsymbol{\sigma}_{xy}$ and $\boldsymbol{\psi}_i' \equiv \boldsymbol{\gamma}_{yi} - \boldsymbol{\omega}'\boldsymbol{\Gamma}_{xi}$, $i = 1, \ldots, p$. We notice that (2.6) is a reparameterization of (2.2) under Assumption 2, which will become more convenient for the construction of the cointegration tests we propose in the next section, and Assumption 2(iv) implies that $\phi + \gamma < 0$. See also Kapetanios et al. (2003).

We call (2.6) the (conditional) nonlinear STR ECM. The representation (2.6) makes economic sense in that many economic models predict that the underlying system tends to display a dampened behavior toward an attractor when it is (sufficiently far) away from it but shows some instability within the locality of that attractor. We note in passing that Assumption 1 is trivially satisfied for our final model specifications (2.6) and (2.7). Therefore, applying Saikkonen's Theorem 2 directly to our model, we may conclude that there exists a choice of initial values $\mathbf{z}_{-p}, \ldots, \mathbf{z}_0$ such that u_{t-1} and $\Delta\mathbf{z}_t$ are strictly stationary.[2]

For a growing literature on the joint analysis of cointegration and STR models see also Gallagher and Taylor (2001), van Dijk, Teräsvirta, and Franses (2002), Choi and Saikkonen (2004), and Saikkonen and Choi (2004). In practice different lag orders for Δy_t and $\Delta\mathbf{x}_t$ in (2.6) can be selected using standard model selection criteria or significance testing procedures without loss of generality or change in the asymptotic analysis, e.g., Ng and Perron (1995).

3. TESTING FOR NONLINEAR COINTEGRATION UNDER STR ECM

To fix ideas for the motivation of the tests, we follow Kapetanios et al. (2003) and impose $\phi = 0$ in (2.6), implying that u_t follows a unit root process in the middle regime; see also Balke and Fomby (1997) in the context of threshold ECMs. Note that for the operational versions of the tests we suggest later we consider both the case $\phi = 0$ and $\phi \neq 0$. It is then straightforward to show that the test of the null of no cointegration against the alternative of globally stationary cointegration can be based on the single parameter θ. Hence, we set the null hypothesis of no cointegration as[3]

$$H_0 : \theta = 0 \tag{3.1}$$

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

against the alternative of nonlinear ESTR cointegration of $H_1 : \theta > 0$, where the positive value of θ determines the stationarity properties of u_t. We note in passing that the null hypothesis implies in terms of the preceding model that $\phi = 0$ and $\theta = 0$. Under the alternative hypothesis u_t follows a nonlinear but globally stationary process as described in Section 2.

We propose a number of operational versions of the cointegration test under the nonlinear STR ECM framework given by (2.6). To this end we follow Engle and Granger (1987) and take a pragmatic residual-based two-step approach. In the first stage, we obtain the residuals, $\hat{u}_t = y_t - \hat{\boldsymbol{\beta}}_x' \mathbf{x}_t$, with $\hat{\boldsymbol{\beta}}_x$ being the ordinary least squares (OLS) estimate of $\boldsymbol{\beta}_x$. In the second stage and to overcome the Davies (1987) problem that γ in (2.6) is not identified under the null, we follow Luukkonen, Saikkonen, and Teräsvirta (1988) and Kapetanios et al. (2003), approximate (2.6) by a first-order Taylor series approximation to $(1 - e^{-\theta(u_{t-1}-c)^2})$, while allowing $\phi \neq 0$ under the alternative hypothesis, and obtain the following auxiliary testing regression:

$$\Delta y_t = \delta_1 \hat{u}_{t-1} + \delta_2 \hat{u}_{t-1}^2 + \delta_3 \hat{u}_{t-1}^3 + \boldsymbol{\omega}' \Delta \mathbf{x}_t + \sum_{i=1}^{p} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + error. \tag{3.2}$$

For this model, we consider an F-type test for $\delta_1 = \delta_2 = \delta_3 = 0$ given by

$$F_{NEC} = \frac{(SSR_0 - SSR_1)/3}{SSR_0/(T-4-p)}, \tag{3.3}$$

where SSR_0 and SSR_1 are the sum of squared residuals obtained from the specification with and without imposing the restrictions $\delta_1 = \delta_2 = \delta_3 = 0$ in (3.2), respectively. It is interesting to note that (3.2), which is called the cubic polynomial nonlinear error correction (NEC) model, has been analyzed by Escribano (1986, 2004) and his coauthors for the estimation of the UK money demand function although it is important to add that these authors did not provide the formal testing framework that we give in this paper.

There are prior theoretical justifications for restricting the switch point, c, to be zero in many economic and financial applications in the ESTR function, in which case we obtain the following (restricted) auxiliary testing regression:

$$\Delta y_t = \delta_1 \hat{u}_{t-1} + \delta_2 \hat{u}_{t-1}^3 + \boldsymbol{\omega}' \Delta \mathbf{x}_t + \sum_{i=1}^{p} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + error \tag{3.4}$$

and obtain the following F-type statistic:

$$F_{NEC}^* = \frac{(SSR_0 - SSR_1)/2}{SSR_0/(T-3-p)}, \tag{3.5}$$

where SSR_0 and SSR_1 are the sum of squared residuals obtained from the specification with and without imposing $\delta_1 = \delta_2 = 0$ in (3.4), respectively.

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

Finally, under the further assumption that $\phi = 0$ (which is the maintained assumption made in Kapetanios et al., 2003) (3.4) is simplified to

$$\Delta y_t = \delta \hat{u}_{t-1}^3 + \boldsymbol{\omega}' \Delta \mathbf{x}_t + \sum_{i=1}^{p} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + error. \tag{3.6}$$

For this model, we propose a t-type statistic for $\delta = 0$ (no cointegration) against $\delta < 0$ (ESTR cointegration), which is given by

$$t_{NEC} = \frac{\hat{\mathbf{u}}_{-1}^{3\prime} \mathbf{Q}_1 \Delta \mathbf{y}}{\sqrt{\hat{\sigma}_{NEC}^2 \hat{\mathbf{u}}_{-1}^{3\prime} \mathbf{Q}_1 \hat{\mathbf{u}}_{-1}^3}}, \tag{3.7}$$

where $\hat{\sigma}_{NEC}^2 = T^{-1} \sum_{t=1}^{T} (\Delta y_t - \hat{\delta} \hat{u}_{t-1}^3 - \hat{\boldsymbol{\omega}}' \Delta \mathbf{x}_t - \sum_{i=1}^{p} \hat{\boldsymbol{\psi}}_i' \Delta \mathbf{z}_{t-i})^2$, $\hat{\mathbf{u}}_{-1}^3 = (\hat{u}_0^3, \ldots, \hat{u}_{T-1}^3)'$, $\Delta \mathbf{y} = (\Delta y_1, \ldots, \Delta y_T)'$, $\mathbf{Q}_1 = \mathbf{I}_T - \mathbf{S}(\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'$ is the $T \times T$ idempotent matrix, $\mathbf{S} = (\Delta \mathbf{X}, \Delta \mathbf{Z}_{-1}, \ldots, \Delta \mathbf{Z}_{-p})$ is the $T \times (k + p(k+1))$ data matrix with $\Delta \mathbf{X} = (\Delta \mathbf{x}_1, \ldots, \Delta \mathbf{x}_T)'$ and $\Delta \mathbf{Z}_{-i} = (\Delta \mathbf{z}_{1-i}, \ldots, \Delta \mathbf{z}_{T-i})'$, $i = 1, \ldots, p$, and $\hat{\delta}$, $\hat{\boldsymbol{\omega}}$, $\hat{\boldsymbol{\psi}}_i$, $i = 1, \ldots, p$, are the OLS estimates of δ, $\boldsymbol{\omega}$, $\boldsymbol{\psi}_i$ in (3.6).

Alternatively and in keeping with the tradition in linear cointegration, we propose a companion test statistic, one that is the analogue to the EG statistic for linear cointegration, which is obtained by estimating the following approximate regression:

$$\Delta \hat{u}_t = \delta \hat{u}_{t-1}^3 + \sum_{i=1}^{p(T)} \varphi_i \Delta \hat{u}_{t-i} + \xi_{t,p(T)}, \tag{3.8}$$

$$\xi_{t,p(T)} = \xi_t + \sum_{i=p(T)+1}^{\infty} \varphi_i \Delta \hat{u}_{t-i}, \tag{3.9}$$

where $\xi_t = \varphi(L) \Delta \hat{u}_t$ is the error term associated with an infinite autoregressive (AR) representation of $\Delta \hat{u}_t$ and $\varphi(L) = 1 - \sum_{i=1}^{\infty} \varphi_i L^i$, e.g., Phillips and Ouliaris (1990, (A.14)). This is so even if the underlying process $\mathbf{z}_t$ follows a finite-order VAR(p) model with $p \geq 2$. Hence we now need to allow $p(T)$ to tend to grow with T for consistent estimation of (3.8). We follow Phillips and Ouliaris (1990) and impose the condition $p(T) = o(T^{1/3})$,[4] under which we derive the following t-type test for $\delta = 0$ in (3.8):

$$t_{NEG} = \frac{\hat{\mathbf{u}}_{-1}^{3\prime} \mathbf{Q}_2 \Delta \hat{\mathbf{u}}}{\sqrt{\hat{\sigma}_{NEG}^2 \hat{\mathbf{u}}_{-1}^{3\prime} \mathbf{Q}_2 \hat{\mathbf{u}}_{-1}^3}}, \tag{3.10}$$

where $\hat{\sigma}_{NEG}^2 = T^{-1} \sum_{t=1}^{T} (\Delta \hat{u}_t - \tilde{\delta} \hat{u}_{t-1}^3 - \sum_{i=1}^{p} \hat{\varphi}_i \Delta \hat{u}_{t-i})^2$, $\Delta \hat{\mathbf{u}} = (\Delta \hat{u}_1, \ldots, \Delta \hat{u}_T)'$, $\mathbf{Q}_2 = \mathbf{I}_T - \Delta \hat{\mathbf{U}}_{p(T)} (\Delta \hat{\mathbf{U}}_{p(T)}' \Delta \hat{\mathbf{U}}_{p(T)})^{-1} \Delta \hat{\mathbf{U}}_{p(T)}'$, $\Delta \hat{\mathbf{U}}_{p(T)} = (\Delta \hat{\mathbf{u}}_{-1}, \ldots, \Delta \hat{\mathbf{u}}_{-p(T)})$, $\Delta \hat{\mathbf{u}}_{-i} = (\Delta \hat{u}_{1-i}, \ldots, \Delta \hat{u}_{T-i})'$, $i = 1, \ldots, p(T)$, and $\tilde{\delta}$, $\hat{\varphi}_i$ are the OLS estimates of δ and φ_i in (3.8).

The main asymptotic distributional results for all these tests are described in Theorem 3.1 under the following additional assumption.

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

Assumption 3. Partition $\boldsymbol{\Gamma}_{xi} = (\boldsymbol{\gamma}'_{xyi}, \boldsymbol{\Gamma}'_{xxi})'$ in (2.7) conformably with $\mathbf{z}_t = (y_t, \mathbf{x}'_t)'$ and set $\boldsymbol{\gamma}_{xyi} = \mathbf{0}$.

This assumption states that Δy_t does not Granger cause $\Delta \mathbf{x}_t$. Under Assumptions 1(i) and (ii), 2(i) and (ii), and 3, the well-known multivariate invariance principle holds for e_t and $\boldsymbol{\varepsilon}_{xt}$:

$$T^{-1/2} \sum_{t=1}^{[Ta]} (e_t, \boldsymbol{\varepsilon}'_{xt})' \Rightarrow (\sigma_e W(a), \boldsymbol{\Sigma}_{xx}^{1/2} \mathbf{W}_x(a)')', \tag{3.11}$$

where $[Ta]$ is the integer part of Ta, $\Rightarrow$ denotes a weak convergence, and $W(a)$ and $\mathbf{W}_x(a)$, defined on $a \in [0,1]$, are independent scalar and k-vector standard Brownian motions. Without Assumption 3, $W(a)$ and $\mathbf{W}_x(a)$ are no longer independent under the null of no cointegration because the weak exogeneity assumption guarantees only that $\sum_{t=1}^{\infty} E(e_t \Delta \mathbf{x}'_0) = \mathbf{0}$, but not that $\sum_{t=1}^{\infty} E(e_0 \Delta \mathbf{x}'_t) = 0$, and therefore the long-run correlation between the partial sums of e_t and $\Delta \mathbf{x}_t$ is possibly nonzero. See also Banerjee et al. (1998) and Pesaran et al. (2001). At the end of this section we discuss a standard extension to the testing procedures that allows us to relax this assumption without affecting asymptotic results.

THEOREM 3.1. *Consider the conditional nonlinear ESTR ECM, (2.6), and let Assumptions 1–3 hold. Under the null hypothesis of no cointegration, the F_{NEC} statistic, defined by (3.3), has the following asymptotic distribution:*

$$F_{NEC} \Rightarrow \frac{1}{3}\left[\int B\,dW, \int B^2\,dW, \int B^3\,dW\right] \times \begin{bmatrix} \int B^2\,da & \int B^3\,da & \int B^4\,da \\ \int B^3\,da & \int B^4\,da & \int B^5\,da \\ \int B^4\,da & \int B^5\,da & \int B^6\,da \end{bmatrix}^{-1} \begin{bmatrix} \int B dW \\ \int B^2\,dW \\ \int B^3\,dW \end{bmatrix}, \tag{3.12}$$

where B and W are shorthand notations for $B(a) = W(a) - \mathbf{W}_x(a)' (\int_0^1 \mathbf{W}_x(a)\mathbf{W}_x(a)'\,da)^{-1} \times (\int_0^1 \mathbf{W}_x(a) W(a)\,da)$ and $W(a)$ defined on $a \in [0,1]$. Next, imposing $c = 0$ and under the null of no cointegration, the F^*_{NEC} statistic defined by (3.5) has the following asymptotic distribution:

$$F^*_{NEC} \Rightarrow \frac{1}{2}\left[\int B\,dW, \int B^3\,dW\right] \begin{bmatrix} \int B^2\,da & \int B^4\,da \\ \int B^4\,da & \int B^6\,da \end{bmatrix}^{-1} \begin{bmatrix} \int B\,dW \\ \int B^3\,dW \end{bmatrix}. \tag{3.13}$$

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

Further, imposing both $\phi = 0$ and $c = 0$, and under the null of no cointegration (3.1), the t_{NEC} and t_{NEG} statistics defined by (3.7) and (3.10) have the following asymptotic distributions, respectively:[5]

$$t_{NEC} \Rightarrow \frac{\int B^3 \, dW}{\sqrt{\int B^6 \, da}}, \qquad t_{NEG} \Rightarrow \frac{\int B^3 \, dB}{(1 + \boldsymbol{\tau}'\boldsymbol{\tau})\sqrt{\int B^6 \, da}}, \tag{3.14}$$

where $\boldsymbol{\tau} = [\int_0^1 \mathbf{W}_x(a)\mathbf{W}_x(a)' \, da]^{-1}[\int_0^1 \mathbf{W}_x(a)W(a)\, da]$. Under the alternative hypothesis that the cointegrating relationship follows ESTR, all the tests are consistent.

To accommodate deterministic components in (2.3), we extend to consider regressions with an intercept and with an intercept and a linear time trend,

$$y_t = a_0 + \boldsymbol{\beta}_x'\mathbf{x}_t + u_t, \tag{3.15}$$

$$y_t = a_0 + a_1 t + \boldsymbol{\beta}_x'\mathbf{x}_t + u_t. \tag{3.16}$$

Alternatively, we take a simpler but equivalent approach and reexpress (3.15) and (3.16) as

$$y_t^* = \boldsymbol{\beta}_x'\mathbf{x}_t^* + u_t^*, \tag{3.17}$$

$$y_t^+ = \boldsymbol{\beta}_x'\mathbf{x}_t^+ + u_t^+, \tag{3.18}$$

where the superscripts $*$ and $+$ indicate the demeaned and the detrended data, respectively. The respective F_{NEC}, F_{NEC}^*, t_{NEC}, and t_{NEG} statistics are then obtained as follows. First, the appropriate residuals are obtained from either (3.17) or (3.18), and then the corresponding regressions are constructed. For example, we have

$$\Delta y_t^* = \delta_1 \hat{u}_{t-1}^* + \delta_2 \hat{u}_{t-1}^{*2} + \delta_3 \hat{u}_{t-1}^{*3} + \boldsymbol{\omega}'\Delta\mathbf{x}_t^* + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta\mathbf{z}_{t-i}^* + error, \tag{3.19}$$

$$\Delta y_t^+ = \delta_1 \hat{u}_{t-1}^+ + \delta_2 \hat{u}_{t-1}^{+2} + \delta_3 \hat{u}_{t-1}^{+3} + \boldsymbol{\omega}'\Delta\mathbf{x}_t^+ + \sum_{i=1}^{p-1} \boldsymbol{\psi}_i' \Delta\mathbf{z}_{t-i}^+ + error, \tag{3.20}$$

where $\hat{u}_t^* = y_t^* - \boldsymbol{\beta}_x'\mathbf{x}_t^*$ and $\hat{u}_t^+ = y_t^+ - \boldsymbol{\beta}_x'\mathbf{x}_t^+$. Then, the F_{NEC} statistics are obtained as the F-statistic for $\delta_1 = \delta_2 = \delta_3 = 0$ in (3.19) or (3.20), respectively. The corresponding F_{NEC}^*, t_{NEC}, and t_{NEG} statistics are obtained in an analogous way. For the regression with a nonzero intercept, (3.15), it is easily seen that the asymptotic distributions of the F_{NEC}, F_{NEC}^*, t_{NEC}, and t_{NEG} statistics are the same as in Theorem 3.1, except that $W(a)$ and $\mathbf{W}_x(a)$ are replaced by the demeaned Brownian motions, denoted $\widetilde{W}(a)$ and $\widetilde{\mathbf{W}}_x(a)$. Similarly, for the regression with nonzero intercept and nonzero linear trend coefficient, (3.16), the asso-

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

ciated asymptotic distributions are such that $W(a)$ and $\mathbf{W}_x(a)$ are replaced by the demeaned and detrended Brownian motions, denoted $\hat{W}(a)$ and $\hat{\mathbf{W}}_x(a)$. Asymptotic critical values of the F_{NEC}, F^*_{NEC}, t_{NEC}, and t_{NEG} statistics for the preceding three cases have been tabulated for $k = 1,\ldots,5$, via stochastic simulations with $T = 1{,}000$ and 50,000 replications in Table 1.

We comment on one further modeling issue. An alternative popular nonlinear STR adjustment scheme to the one favored in this paper is given either by the first-order logistic function, $1 - b/(1 + \exp(-\theta u_{t-1}))$, or by the second-order logistic function, $1 - b/(1 + \exp(-\theta u_{t-1}^2))$, where $b < 1$ is a scaling constant. We note that the resulting t-type tests against the alternative of the

TABLE 1. Asymptotic critical values for alternative statistics

	Case 1			Case 2			Case 3		
k	90%	95%	99%	90%	95%	99%	90%	95%	99%
F_{NEC}									
1	10.00	12.28	16.81	11.79	13.73	17.38	13.95	16.13	19.97
2	11.41	13.22	17.33	12.89	14.87	19.33	15.70	17.83	22.88
3	12.46	14.15	19.64	14.40	16.69	21.81	16.99	19.38	24.71
4	13.97	16.39	21.85	15.77	18.05	23.62	17.83	20.75	25.38
5	15.31	18.20	21.99	17.88	20.84	26.33	19.58	22.24	28.46
F^*_{NEC}									
1	7.34	9.06	12.53	10.13	12.17	16.36	12.83	15.07	19.46
2	9.00	10.83	14.54	11.72	14.09	17.66	14.81	16.96	20.65
3	10.26	12.45	16.96	12.92	15.37	20.07	16.21	18.63	23.66
4	11.65	14.04	19.96	14.99	17.71	22.24	17.21	20.14	25.69
5	14.01	16.21	21.40	16.04	19.03	24.47	19.09	22.03	28.16
t_{NEC}									
1	−2.38	−2.66	−3.35	−2.92	−3.22	−3.78	−3.30	−3.59	−4.17
2	−2.67	−3.01	−3.59	−3.12	−3.43	−4.00	−3.46	−3.79	−4.40
3	−2.95	−3.28	−3.93	−3.32	−3.61	−4.19	−3.62	−3.96	−4.54
4	−3.15	−3.47	−4.14	−3.46	−3.77	−4.38	−3.75	−4.07	−4.70
5	−3.33	−3.67	−4.31	−3.58	−3.92	−4.53	−3.87	−4.20	−4.85
t_{NEG}									
1	−2.59	−2.85	−3.38	−2.98	−3.28	−3.84	−3.41	−3.71	−4.26
2	−3.01	−3.30	−3.89	−3.36	−3.67	−4.23	−3.64	−3.99	−4.53
3	−3.34	−3.66	−4.23	−3.63	−3.93	−4.50	−3.90	−4.18	−4.76
4	−3.65	−3.95	−4.56	−3.90	−4.19	−4.68	−4.09	−4.39	−4.95
5	−3.88	−4.13	−4.75	−4.10	−4.42	−4.97	−4.36	−4.67	−5.23

Notes: Critical values reported in this table are computed via stochastic simulations with $T = 1{,}000$ and 50,000 replications. Case 1 uses the raw data, Case 2 the demeaned data, and Case 3 the detrended data, and k is the number of regressors.

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

second-order logistic function have exactly the same form as in (3.7) and (3.10). It is also easily seen that adopting the first-order Taylor approximation to the first-order logistic terms as we did before gives the regression

$$\Delta y_t = \delta \hat{u}_{t-1}^2 + \boldsymbol{\omega}' \Delta \mathbf{x}_t + \sum_{i=1}^{p} \boldsymbol{\psi}_i' \Delta \mathbf{z}_{t-i} + error, \tag{3.21}$$

and we could develop a t-type test for $\delta = 0$ in (3.21). Though the asymptotics for this case are straightforwardly obtained via a simple extension of the analysis in this paper, the interpretation of the error correction coefficient in the first-order logistic STR error correction model is problematic. By contrast, in the exponential case the interpretation of this coefficient is both natural and intuitive (for a discussion of this point, see Kapetanios et al., 2003). Therefore, we focus on the exponential case.

Finally, we comment on how to relax Assumption 3 in practice. To purge the potentially nonzero long-run correlation between the partial sums of e_t and $\Delta \mathbf{x}_t$, we should augment our conditional model (2.6) with leads of $\Delta \mathbf{x}_t$, as is standard in the cointegration literature; see, e.g., Banerjee et al. (1998) and Saikkonen (1991). Then, following Theorem 4.1 of Saikkonen (1991), it is readily seen that the limit distribution of these ECM-based tests will be identical to those derived in Theorem 3.1, under the further condition that the number of leads grows at a rate less than $T^{1/3}$. We also notice that the asymptotic result for the t_{NEG} test in this situation would be intact too.

4. FINITE-SAMPLE PROPERTIES

In this section we undertake a small-scale Monte Carlo investigation of the finite-sample size and power performance of our proposed t_{NEC}, t_{NEG}, F_{NEC}, and F_{NEC}^* tests in conjunction with the linear cointegration tests of Engle and Granger and Johansen denoted t_{EG} and JOH, respectively.

We consider experiments based on a bivariate ECM similar to that adopted by Arranz and Escribano (2000). We generate the data as follows:

$$\Delta y_t = \lambda \Delta x_t + \gamma u_{t-1}[1 - \exp(-\theta u_{t-1}^2)] + \varepsilon_t, \tag{4.1}$$

$$\Delta x_t = v_t, \qquad u_t = y_t - \beta_x x_t, \tag{4.2}$$

$$\begin{bmatrix} \varepsilon_t \\ v_t \end{bmatrix} \sim iidN\left(\mathbf{0}, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}\right). \tag{4.3}$$

Here we fix $\beta_x = 1$ and $\sigma_1^2 = 1$. In the tabulated experiments we fix $\phi = 0$ but also discuss (rather than tabulate) results for nonzero ϕ. Under the null we set $\theta = 0$, whereas we consider parameter values for $\theta = \{0.01, 0.1, 1\}$ and $\gamma = \{-1, -0.5, -0.3, -0.1\}$ under the alternative. To investigate the impact of the common factor (COMFAC) restriction, $\lambda = 1$ and the impact of different signal-

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

to-noise ratios we consider different parameter values for $\lambda = \{0.5, 1\}$ and $\sigma_2^2 = \{1, 4\}$.

To save space we only report the results for the demeaned case (see (3.17)) with $T = 100$.[6] Table 2 gives results on the size of alternative tests and shows that at this sample size there are no substantial size distortions for any of the tests under consideration. This is convenient as it allows a direct comparison of test powers without the need to worry about size-adjusted critical values.

Turning to the power performance of the tests, which is summarized in Table 3, a general finding is that nonlinear-based tests dominate linear counterparts in terms of power for almost all cases considered. As expected, the power gain of our suggested NEC-based tests over the nonlinear analogue of the EG test becomes substantial where the COMFAC restriction ($\lambda = 1$) does not hold. The result that when the COMFAC restriction is violated the power superiority of t_{NEC} over t_{NEG} increases with the variance of the innovation in x is both intuitive (a higher variance increases the correlation with the regression error) and a reiteration of what is now well known to occur in the linear case.

Next, we focus on the results for $\theta = 0.01$, because this value is found to be close to empirically plausible estimates.[7] For example, when $(\gamma, \lambda, \sigma^2) = (-1, 0.5, 1)$, the powers of the t_{NEC}, F^*_{NEC}, F_{NEC}, and t_{NEG} tests are 48%, 45%, 43%, and 33%, whereas the powers of t_{EG} and JOH are 24% and 26% only. What is surprising however is the relative performance of t_{NEC}, F_{NEC}, and F^*_{NEC}—in terms of power there is little to choose between these three. This occurs despite the fact that the F-statistics are testing more coefficients. Put another way, the implicit alternatives of the F-tests are "larger" than that of t_{NEC} in the sense that they nest approximations to a greater number of alternative nonlinear models than does t_{NEC}. Allowing ϕ and/or c to be nonzero does not alter the general tenor of these results. In sum it seems that moving from t_{NEC} when its implicit alternative is true to portmanteau F-tests (whose alternatives nest the truth as a special case) does not seem to significantly jeopardize power. We would hesitate to draw general conclusions from a simple and single data generating process, but we might cautiously recommend that if there is some doubt as to the nature of the STR alternative and in particular if we do not know for sure that $c = 0$ then the F-type tests should be used.

TABLE 2. Size of alternative tests for demeaned data ($T = 100$)

λ	σ_2^2	t_{EG}	t_{NEG}	t_{NEC}	JOH	F_{NEC}	F^*_{NEC}
0.5	1	0.047	0.042	0.041	0.065	0.043	0.045
0.5	4	0.057	0.041	0.045	0.057	0.044	0.044
1.0	1	0.046	0.042	0.045	0.054	0.043	0.045
1.0	4	0.049	0.047	0.044	0.052	0.047	0.053

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

TABLE 3. Power of alternative tests for demeaned data ($T = 100$)

γ	θ	λ	σ_2^2	t_{EG}	t_{NEG}	t_{NEC}	JOH	F_{NEC}	F_{NEC}^*
−1.0	0.01	0.5	1	0.239	0.333	0.483	0.264	0.433	0.451
−1.0	0.01	0.5	4	0.302	0.476	0.791	0.681	0.839	0.841
−1.0	0.01	1	1	0.224	0.288	0.304	0.159	0.290	0.306
−1.0	0.01	1	4	0.224	0.310	0.322	0.165	0.270	0.306
−1.0	0.1	0.5	1	1.0	0.997	0.996	0.998	1.0	1.0
−1.0	0.1	0.5	4	1.0	0.999	1.0	1.0	1.0	1.0
−1.0	0.1	1	1	1.0	0.997	0.998	0.947	1.0	1.0
−1.0	0.1	1	4	0.999	0.998	0.998	0.965	1.0	1.0
−1.0	1	0.5	1	1.0	1.0	1.0	1.0	1.0	1.0
−1.0	1	0.5	4	1.0	1.0	1.0	1.0	1.0	1.0
−1.0	1	1	1	1.0	1.0	1.0	1.0	1.0	1.0
−1.0	1	1	4	1.0	1.0	1.0	1.0	1.0	1.0
−0.5	0.01	0.5	1	0.126	0.159	0.245	0.142	0.214	0.228
−0.5	0.01	0.5	4	0.131	0.203	0.487	0.392	0.521	0.497
−0.5	0.01	1	1	0.128	0.151	0.176	0.113	0.141	0.164
−0.5	0.01	1	4	0.154	0.164	0.189	0.118	0.144	0.178
−0.5	0.1	0.5	1	0.878	0.858	0.932	0.789	0.959	0.973
−0.5	0.1	0.5	4	0.973	0.906	0.990	0.995	1.0	1.0
−0.5	0.1	1	1	0.831	0.833	0.855	0.496	0.850	0.903
−0.5	0.1	1	4	0.852	0.836	0.843	0.541	0.875	0.906
−0.5	1	0.5	1	1.0	0.972	0.988	1.0	1.0	1.0
−0.5	1	0.5	4	1.0	0.975	1.0	1.0	1.0	1.0
−0.5	1	1	1	1.0	0.970	0.973	0.994	1.0	1.0
−0.5	1	1	4	1.0	0.971	0.973	0.990	1.0	1.0
−0.3	0.01	0.5	1	0.090	0.109	0.136	0.088	0.122	0.138
−0.3	0.01	0.5	4	0.083	0.099	0.305	0.233	0.325	0.301
−0.3	0.01	1	1	0.110	0.121	0.123	0.061	0.103	0.126
−0.3	0.01	1	4	0.120	0.116	0.121	0.088	0.107	0.130
−0.3	0.1	0.5	1	0.489	0.535	0.675	0.385	0.692	0.737
−0.3	0.1	0.5	4	0.536	0.550	0.896	0.884	0.940	0.952
−0.3	0.1	1	1	0.450	0.504	0.521	0.201	0.488	0.536
−0.3	0.1	1	4	0.482	0.540	0.557	0.246	0.517	0.586
−0.3	1	0.5	1	0.961	0.714	0.850	0.878	0.968	0.982
−0.3	1	0.5	4	0.932	0.663	0.954	0.996	0.999	1.0
−0.3	1	1	1	0.955	0.740	0.760	0.693	0.896	0.947
−0.3	1	1	4	0.971	0.753	0.773	0.696	0.925	0.954
−0.1	0.01	0.5	1	0.057	0.068	0.079	0.057	0.063	0.080
−0.1	0.01	0.5	4	0.065	0.066	0.129	0.138	0.141	0.139
−0.1	0.01	1	1	0.068	0.071	0.068	0.037	0.077	0.075
−0.1	0.01	1	4	0.061	0.078	0.077	0.049	0.058	0.070
−0.1	0.1	0.5	1	0.121	0.123	0.167	0.103	0.156	0.181
−0.1	0.1	0.5	4	0.089	0.108	0.260	0.263	0.280	0.297
−0.1	0.1	1	1	0.131	0.115	0.132	0.076	0.120	0.155
−0.1	0.1	1	4	0.135	0.139	0.139	0.070	0.110	0.143
−0.1	1	0.5	1	0.189	0.164	0.222	0.155	0.220	0.267
−0.1	1	0.5	4	0.140	0.117	0.310	0.345	0.375	0.415
−0.1	1	1	1	0.224	0.194	0.200	0.093	0.190	0.232
−0.1	1	1	4	0.205	0.165	0.180	0.105	0.170	0.212

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

5. EMPIRICAL APPLICATION TO ASSET PRICING IN THE PRESENCE OF TRANSACTIONS COSTS

In a seminal paper, Campbell and Shiller (1987) investigate the existence of linear cointegration between aggregate U.S. stock prices and U.S. dividends, as predicted by a simple equilibrium model of constant expected asset returns. Their results were ambiguous. A null hypothesis of no (linear) cointegration was marginally rejected in their data, but the implied estimate of long-run asset returns was implausible. Imposing a more credible long-run return caused non-rejection of the null of no cointegration. Subsequent literature has met with similar mixed results, e.g., Campbell and Shiller (1988), Froot and Obstfeld (1991), and Cuthbertson, Hayes, and Nitzsche (1997).

In this section we apply our proposed tests for cointegration to asset prices and dividends for 11 stock portfolios allowing for nonlinear adjustment to equilibrium of the STR variety. The theoretical motivation for STR nonlinearity has already been given in the introduction, and we reemphasize that theoretical reasoning suggests that it is STR adjustment rather than threshold adjustment that is likely to arise in our current application.[8] Gallagher and Taylor (2001) have applied the STR-based ECM to the U.S. log dividend-price ratio and find that the speed of reversion of the U.S. log dividend-price ratio toward the fundamental equilibrium follows a nonlinear STR adjustment and is an increasing function of degree of mispricing. In their study, however, they adopt linear cointegration tests such as the Engle and Granger (1987) and Johansen (1995) tests.

We collected monthly data from January 1974 to November 2002 on end period real prices and within period real dividend yields for value weighted market portfolio indexes of stocks traded on the main exchanges of the following 11 countries: Germany, Belgium, Canada, Denmark, France, Ireland, Italy, Japan, Netherlands, United Kingdom, and United States. A dividend series was constructed as the product of dividend yield and prices. Although not presented here, augmented Dickey–Fuller (ADF) tests give overwhelming support to the hypothesis that all variates are I(1).

As alluded to before, we test for the existence of a linear cointegrating relationship between dividends and prices of the form

$$p_t = \beta d_t + u_t. \tag{5.1}$$

The existence of bid-ask spreads discussed previously motivates the following nonlinear dynamic STR ECM specification:

$$\Delta p_t = \gamma(1 - e^{-\theta u_{t-1}^2})u_{t-1} + \alpha \Delta x_t + \varepsilon_t, \tag{5.2}$$

where $u_{t-1} = p_{t-1} - \beta d_{t-1}$ and ε_t is a (possibly autocorrelated) error term that captures other microstructure effects such as specific kinds of noise trading and dealer inventory control mechanisms whereby the adjustment of prices to elicit

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

inventory-correcting trades generates autocorrelated price movements, e.g., Snell and Tonks (1998).

We computed three tests. The first two, t_{EG} and t_{NEG}, are the linear EG test and its nonlinear counterpart. The third, t_{NEC}, is the t-ratio on $\hat{u}_{t-1}^3$ in the STR ECM formulation (see (3.6)) where $\hat{u}_{t-1}$ is the residual from the first stage (spurious under the null) regression of p_t on d_t. The price and dividend series appeared to have an upward trend so that all series were demeaned and detrended before use.[9] We estimated the appropriate auxiliary regressions for $p = 12$ and then dropped all insignificant lags in a general-to-specific modeling.[10] Finally, we note that there is overwhelming prior theoretical justification for restricting the switch point, c, to be zero. As a result we do not compute F-type tests that test additional linear and quadratic terms.

The results for the tests are in columns 2–4 in Table 4. Looking at the results we see that viewed through the "eyes" of linear cointegration tests there is little support for the hypothesis that dividends and prices move together in the long run, with only two of the t_{EG} tests rejecting the null, albeit at the 1% level. Furthermore, none of the remaining nine t_{EG} statistics are significant even at the 10% level. By contrast the nonlinear t_{NEG} test rejects in seven out of eleven stock markets with four of these rejections occurring at the 1% level. A further three statistics are quite close to the 10% critical value. The success in rejecting the null of no cointegration is less marked for t_{NEC}, with only five rejections at standard significance levels, although three of these reject also at the 1%. A further two t_{NEC} statistics are quite close to the 10% critical value. If the alternative is really true then we could interpret this finding as being somewhat at odds with the Monte Carlo evidence presented before, which generally shows

TABLE 4. Cointegration tests and estimates of the ESTR parameter for asset prices and dividends

Country	t_{EG}	t_{NEG}	t_{NEC}	$\hat{\theta}$	t_θ
Germany	−2.62	−3.68*	−4.69**	0.011	3.40
Belgium	−2.06	−4.69**	−4.89**	0.007	3.79
Canada	−1.73	−3.05	−1.53	0.008	3.26
Denmark	−5.04**	−4.82**	−4.73**	0.009	3.75
France	−3.27	−3.07	−0.99	0.007	2.83
Ireland	−2.72	−3.76*	−3.73*	0.017	3.85
Italy	−1.81	−5.50**	−3.04	0.010	3.51
Japan	−2.44	−2.46	−2.11	0.009	2.86
Netherlands	−4.94**	−7.79**	−1.90	0.015	6.18
United Kingdom	−2.43	−3.28	−3.22	0.008	2.78
United States	−2.10	−3.90*	−3.70*	0.007	3.76

Notes: The sample period in all cases is March 1974 to November 2002. *(**) denotes significance at the 5% (1%) level.

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

that t_{NEC} has more power than t_{NEG}. One possible explanation is that prior empirical investigation of the data suggested that the COMFAC restriction may hold here. If true, a bivariate ECM would lack parsimony compared with the univariate specification, and this may lead to a loss in power when moving from t_{NEG} to t_{NEC}.

Given the strength of evidence against the null and support for the alternative we could obtain estimates of adjustment parameters under the alternative. Focusing on the univariate model we obtained nonlinear least squares estimates of θ from the alternative ESTR model,

$$\Delta \hat{u}_t = -\{1 - \exp(-\theta \hat{u}_{t-1}^2)\}\hat{u}_{t-1} + \sum_{i=1}^{12} \varphi_i \Delta \hat{u}_{t-i} + \xi_t, \tag{5.3}$$

where the model has been specialized compared with the general ESTR considered previously by imposing a unit coefficient on γ. Early attempts to estimate γ jointly with θ foundered on severe identification problems, and our nonlinear algorithm failed to converge in most cases—hence the specialization. Under the alternative (and estimation of (5.3) only makes sense if the alternative is true), θ is scale dependent. To clarify its interpretation and to facilitate numerical convergence, we normalized the $\hat{u}_t$ series to have unit sample variance (a procedure that only makes sense under the alternative). The results for $\hat{\theta}$ and its t-statistic are given in Table 4. Although we cannot interpret the t-statistic as a significance from zero test (for obvious reasons) we refer to it as "significant" if an asymptotic 95% confidence interval around the estimate excludes zero. We see that $\hat{\theta}$ is "significant" in all cases and varies between 0.007 and 0.017.

To get a feel for what such values imply, Figure 1 plots IRFs for the error correction term for initial impulses of one, two, three, and four standard deviation shocks, respectively. To be explicit, we take the univariate specification in (5.3) and perturb its disturbance by one to four standard deviation shocks. We trace the time path implied by (5.3) following each of the four shocks and plot them on a single IRF diagram.[11] For completeness and comparison we compare the corresponding IRF with that obtained from the estimated linear models. The striking finding is the length of time taken to recover from small shocks. In particular the time taken to recover one-half of a one standard deviation shock varies between 5 and 20 years. By contrast, the time taken to recover one-half of a large shock (such as three or four standard deviations) is comparable to that of the linear case and varies between just 4 and 18 months. This implies that data periods dominated by extreme volatility may display substantial reversion of prices toward their NPV relationship but in "calmer" times, where the error in the NPV relationship takes on smaller values, the process driving it may well look like a unit root. This suggests that in practice the ESTR and threshold autoregressive (TAR) models may not be too dissimilar in terms of overall inference in any given sample.

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

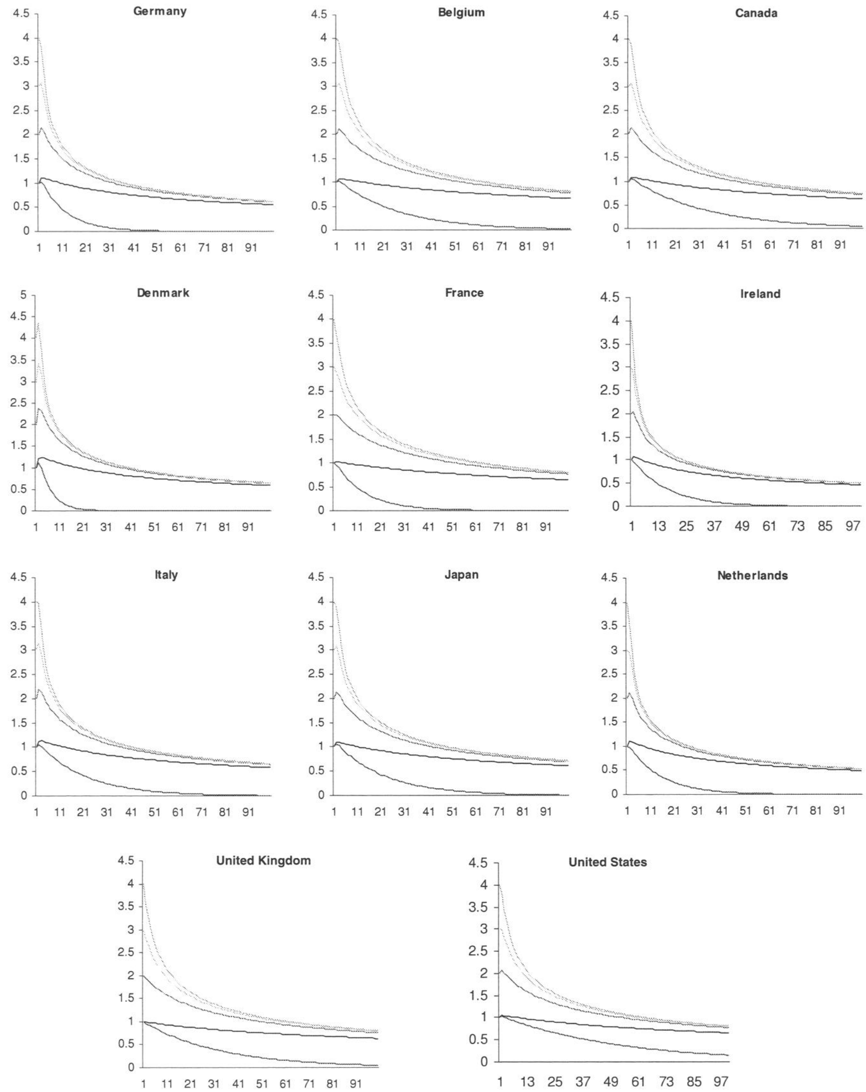

FIGURE 1. Impulse responses of error correction terms to k $(= 1,2,3,4)$ standard deviation shock for various countries.

6. CONCLUDING REMARKS

The investigation of nonstationarity in conjunction with nonlinear autoregressive modeling has recently assumed a prominent role in econometric study. It is clear that misclassifying a stable nonlinear process as nonstationary can be

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

misleading both in impulse response and forecasting analysis. Similarly, not allowing for the presence of cointegration when the speed of adjustment varies with the position of the system as in the case of nonlinear cointegration can be equally misleading. In this paper we have proposed a simple direct cointegration test procedure that is designed to have power against STR nonlinear error correction specifications. Our proposed tests are shown to have better power than the linear cointegration tests that ignore the nonlinear nature of the alternative. An empirical application clearly shows the potential of our approach. Unlike linear cointegration tests, the nonlinear tests find substantial evidence of cointegration in stock price and dividend systems. Further research to develop similar tests for alternative nonlinear models is currently under way.

NOTES

1. The exponential transition function in (2.6) is bounded between zero and 1, i.e., $F:\mathbb{R} \rightarrow [0,1]$ has the properties $F(0) = 0$; $\lim_{x\rightarrow\pm\infty} F(x) = 1$, and is symmetrically U-shaped around zero.
2. Kapetanios et al. (2003) show that the $u_t(= y_t - \boldsymbol{\beta}_x'\mathbf{x}_t)$ are geometrically ergodic so long as $\phi + \gamma < 0$ in the univariate context where the cointegrating parameters $\boldsymbol{\beta}_x$ are assumed given.
3. The null and alternative are expressed alternatively as $H_0: \theta\gamma = 0$ and $H_1: \theta\gamma < 0$. This formulation clearly shows the lack of identification of either θ or γ when the other is zero. It also makes it clearer why we will develop t- or F-type tests that do not depend on the value of the unidentified parameter.
4. This upper bound follows from Berk (1974), who shows that this bound is sufficient for the second sample moment matrix to converge to its population counterpart in norm. Therefore, for consistent estimation it is sufficient to impose this upper bound. See also Said and Dickey (1984).
5. It is worth noting that the asymptotic distribution of the t_{NEC} statistic is free of any nuisance parameters as proved in the Appendix. In the literature it is well established that the limiting null distribution of an ECM t-test depends on the signal-to-noise ratio, though this dependency is due to the use of known cointegrating parameters. It has been proved in linear models that an ECM t-test obtained using the residuals is free of any nuisance parameters, and a formal proof of this is available upon request.
6. Overall simulation results for the different sample sizes and for the detrended data are qualitatively similar to the main findings reported in the paper and are available upon request.
7. Notice in our application that the estimates of θ (which we obtain under the constraint that $\gamma = -1$ for convenience; see also Taylor, Peel, and Sarno, 2001) indeed range between 0.007 and 0.017. In general, it is hard to get an exact definition of "small" and "large" θ because it is not a scale-free parameter. But it is easily seen that given the values of σ^2 and γ, as θ grows, the u_t becomes less persistent, where we note that the term $e^{-\theta u_{t-1}^2}$ here measures the root of the series at time t. For example, we find via simulation that for $\sigma^2 = 1$ and $\gamma = -1$ with $T = 1{,}000$, the average sample values of $e^{-\theta u_{t-1}^2}$ are 0.95, 0.83, and 0.34 for $\theta = 0.01$, 0.1, and 1, respectively.
8. There is also some simulation evidence that shows that even if bid-ask spreads are fixed, aggregation across assets with different fixed spreads will in fact lead to STR-type rather than TAR-type dynamic adjustment.
9. The issue of whether or not stock prices and dividends contain a deterministic time trend is contentious. However, there is a clearly discernible trend in both dividends and prices in our data, and hence we detrend. It is comforting to note that if we only demean, the results are qualitatively almost identical.
10. We should note that although further exploration revealed some significant lags beyond twelfth order, the test statistics were not in general very sensitive to the choice of lag length.

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

11. Alternatively, it would be interesting to further investigate the IRF analysis in the context of nonlinear STR ECMs in which both conditional and marginal models (2.6) and (2.7) should be jointly considered. But this is beyond the current study. For further discussions see van Dijk et al. (2002).

REFERENCES

Arranz, M.A. & A. Escribano (2000) Cointegration testing in the presence of structural breaks. *Oxford Bulletin of Economics and Statistics* 62, 23–52.

Balke, N.S. & T.B. Fomby (1997) Threshold cointegration. *International Economic Review* 38, 627–645.

Banerjee, A., J.J. Dolado, & R. Mestre (1998) ECM tests for cointegration in a single equation framework. *Journal of Time Series Analysis* 19, 267–283.

Bec, F., A. Guay, & E. Guerre (2004) Adaptive Consistent Unit Root Tests Based on Autoregressive Threshold Model. Manuscript, Crest-ENSAE.

Bec, F. & A. Rahbek (2004) Vector equilibrium correction models with nonlinear discontinuous adjustments. *Econometrics* 7, 628–651.

Berk, K.N. (1974) Consistent autoregressive spectral estimates. *Annals of Statistics* 2, 489–502.

Campbell, J.Y. & A. Kyle (1993) Smart money, noise trading and stock price behavior. *Review of Economic Studies* 60, 1–34.

Campbell, J.Y. & R.J. Shiller (1987) Cointegration and tests of present value models. *Journal of Political Economy* 95, 1062–1088.

Campbell, J.Y. & R.J. Shiller (1988) Stock prices, earnings and expected dividends. *Journal of Finance* 43, 661–676.

Caner, M. & B.E. Hansen (2001) Threshold autoregression with a unit root. *Econometrica* 69, 1555–1596.

Choi, I. & P. Saikkonen (2004) Testing linearity in cointegrating smooth transition regressions. *Econometrics Journal* 7, 341–365.

Corradi, V., N.R. Swanson, & H. White (2000) Testing for stationarity-ergodicity and for comovement between nonlinear discrete time Markov processes. *Journal of Econometrics* 96, 39–73.

Cuthbertson, K., S. Hayes, & D. Nitzsche (1997) The behavior of the UK stock prices and returns: Is the market efficient? *Economic Journal* 107, 986–1008.

Davies, R.B. (1987) Hypothesis testing when a nuisance parameter is present under the alternative. *Biometrika* 74, 33–43.

Dickey, D.A. & 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.

Enders, W. & C.W.J. Granger (1998) Unit root test and asymmetric adjustment with an example using the term structure of interest rates. *Journal of Business & Economic Statistics* 16, 304–311.

Engle, R. & C.W.J. Granger (1987) Cointegration and error correction: Representation, estimation and testing. *Econometrica* 55, 251–276.

Escribano, A. (1986) Identification and modelling of economic relationships in a growing economy. Ph.D Dissertation, University of California, San Diego.

Escribano, A. (2004) Nonlinear error correction: The case of money demand in the UK (1878–2000). *Macroeconomic Dynamics* 8, 76–116.

Escribano, A. & S. Mira (2002) Nonlinear error correction models. *Journal of Time Series Analysis* 23, 509–522.

Escribano, A. & G. Pfann (1998) Nonlinear error correction, asymmetric adjustment and cointegration. *Economic Modelling* 15, 197–216.

Froot, K. & M. Obstfeld (1991) Intrinsic bubbles: The case of stock prices. *American Economic Review* 81, 1189–1214.

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

Gallagher, L.A. & M.P. Taylor (2001) Risky arbitrage, limits of arbitrage and nonlinear adjustment in the dividend-price ratio. *Economic Inquiry* 39, 524–536.

Granger, C.W.J. & N.R. Swanson (1996) Further studies in the study of cointegrated variables. *Oxford Bulletin of Economics and Statistics* 58, 537–553.

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

Hansen, B.E. (1995) Rethinking the univariate approach to unit root tests: How to use covariates to increase power. *Econometric Theory* 11, 1148–1171.

Hansen, B.E. & B. Seo (2002) Testing for two-regime threshold cointegration in vector error correction models. *Journal of Econometrics* 110, 293–318.

Johansen, S. (1995) *Likelihood Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press.

Kapetanios, G. & Y. Shin (2006) Unit root tests in three-regime SETAR models. *Econometrics*, in press.

Kapetanios, G., Y. Shin, & A. Snell (2003) Testing for a unit root in the nonlinear STAR framework. *Journal of Econometrics* 112, 359–379.

Kiliç, R. (2003) A Testing Procedure for a Unit Root in the STAR Model. Manuscript, Georgia Institute of Technology.

Kremers, J.J.M., K.R. Ericsson, & J.J. Dolado (1992) The power of cointegration tests. *Oxford Bulletin of Economics and Statistics* 54, 325–348.

Kyle, A.S. (1985) Continuous auctions and insider trading. *Econometrica* 53, 1315–1335.

Luukkonen, R., P. Saikkonen, & T. Teräsvirta (1988) Testing linearity against smooth transition autoregressive models. *Biometrika* 75, 491–499.

Michael, P., R.A. Nobay, & D.A. Peel (1997) Transactions costs and nonlinear adjustment in real exchange rates: An empirical investigation. *Journal of Political Economy* 105, 862–879.

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

Park, J.Y. & P.C.B. Phillips (2001) Nonlinear regressions with integrated time series. *Econometrica* 69, 117–161.

Pesaran, M.H., Y. Shin & R.J. Smith (2001) Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics* 16, 289–326.

Phillips, P.C.B. & S. Ouliaris (1990) Asymptotic properties of residual based tests of cointegration. *Econometrica* 58, 165–193.

Phillips, P.C.B. & V. Solo (1992) Asymptotics for linear processes. *Annals of Statistics* 20, 971–1001.

Psaradakis, Z., M. Sola, & F. Spagnolo (2004) On Markov error correction models with an application to stock prices and dividends. *Journal of Applied Econometrics* 19, 69–88.

Said, S.E. & D.A. Dickey (1984) Testing for unit roots in ARMA models of unknown order. *Biometrika* 71, 599–607.

Saikkonen, P. (1991) Asymptotically efficient estimation of cointegrating regressions. *Econometric Theory* 7, 1–21.

Saikkonen, P. (2005) Stability results for nonlinear error correction models. *Journal of Econometrics* 127, 69–81.

Saikkonen, P. & I. Choi (2004) Cointegrating smooth transition regressions. *Econometric Theory* 20, 301–340.

Snell, A. & I. Tonks (1998) Testing for asymmetric information and inventory control effects in market maker behavior on the London Stock Exchange. *Journal of Empirical Finance* 5, 1–25.

Snell, A. & I. Tonks (2003) A theoretical analysis of institutional investors' trading costs in auction and dealer markets. *Economic Journal* 113, 576–598.

Taylor, M.P., D.A. Peel, & L. Sarno (2001) Nonlinearity in real exchange rates: Towards a solution of the purchasing power parity puzzles. *International Economic Review* 42, 1015–1042.

van Dijk, D., T. Teräsvirta, & P.H. Franses (2002) Smooth transition autoregressive models—A survey of recent developments. *Econometric Reviews* 21, 1–47.

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

APPENDIX: Proof of Theorem 3.1

Under the null $\delta_1 = \delta_2 = \delta_3 = 0$, the F_{NEC} statistic given by (3.3) can be rewritten as

$$F_{NEC} = \frac{1}{3}\frac{\mathbf{e}'\mathbf{Q}_1\hat{\mathbf{U}}_{-1}^{(3)}(\hat{\mathbf{U}}_{-1}^{(3)\prime}\mathbf{Q}_1\hat{\mathbf{U}}_{-1}^{(3)})^{-1}\hat{\mathbf{U}}_{-1}^{(3)\prime}\mathbf{Q}_1\mathbf{e}}{\hat{\sigma}_{FNEC}^2}, \tag{A.1}$$

where $\mathbf{e} = (e_1,\dots,e_T)'$, $\hat{\mathbf{U}}_{-1}^{(3)} = (\hat{\mathbf{u}}_0^{(3)},\dots,\hat{\mathbf{u}}_{T-1}^{(3)})'$, $\hat{\mathbf{u}}_t^{(3)} = (\hat{u}_t,\hat{u}_t^2,\hat{u}_t^3)'$, $\mathbf{Q}_1$ is the $T\times T$ idempotent matrix defined following (3.7), and $\hat{\sigma}_{FNEC}^2 = T^{-1}\sum_{t=1}^{T}(\Delta y_t - \hat{\delta}_1\hat{u}_{t-1} - \hat{\delta}_2\hat{u}_{t-1}^2 - \hat{\delta}_3\hat{u}_{t-1}^3 - \hat{\boldsymbol{\omega}}'\Delta\mathbf{x}_t - \sum_{i=1}^{p}\hat{\boldsymbol{\psi}}_i'\Delta\mathbf{z}_{t-i})^2$ is the estimated variance with $\hat{\delta}_1$, $\hat{\delta}_2$, $\hat{\delta}_3$, $\hat{\boldsymbol{\omega}}$, and $\hat{\boldsymbol{\psi}}_i$ being the OLS estimators obtained from (3.2).

Notice under the null that (2.6) can be written alternatively as the following autoregressive distributed lag (ARDL) specification:

$$\phi(L)\Delta y_t = \boldsymbol{\theta}(L)\Delta\mathbf{x}_t + e_t, \tag{A.2}$$

where $\phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$ and $\boldsymbol{\theta}(L) = \boldsymbol{\theta}_0 + \boldsymbol{\theta}_1 L + \cdots + \boldsymbol{\theta}_p L^p$. Decomposing $\boldsymbol{\psi}_i' = (\psi_{yi},\boldsymbol{\psi}_{xi}')$ in (2.6), and comparing (A.2) with (2.6), we find that $\phi_i = \psi_{yi}$, $i = 1,\dots,p$, $\boldsymbol{\theta}_0 = \boldsymbol{\omega}'$, and $\boldsymbol{\theta}_i = \boldsymbol{\psi}_{xi}'$, $i = 1,\dots,p$. Using the Beveridge–Nelson (BN) decomposition (e.g., Phillips and Solo, 1992) we have

$$\phi(L) = \phi(1) + (1-L)\phi^*(L); \qquad \boldsymbol{\theta}(L) = \boldsymbol{\theta}(1) + (1-L)\boldsymbol{\theta}^*(L), \tag{A.3}$$

where $\phi^*(L) = 1 - \phi_1^* L - \cdots - \phi_{p-1}^* L^{p-1}$ and $\boldsymbol{\theta}^*(L) = \boldsymbol{\theta}_0^* + \boldsymbol{\theta}_1^* L + \cdots + \boldsymbol{\theta}_{p-1}^* L^{p-1}$. Summing both sides of (A.2) from 1 to t and using (A.3), we obtain

$$\phi(1)y_t = \boldsymbol{\theta}(1)x_t + s_{et} + \phi(1)y_0 - \boldsymbol{\theta}(1)x_0 + \phi^*(L)\Delta y_t + \boldsymbol{\theta}^*(L)\Delta\mathbf{x}_t, \tag{A.4}$$

where $s_{et} = \sum_{j=1}^{t} e_j$. Under Assumptions 1(iii), 2(i) and (ii), and 3 and using (A.4), then it is easily seen that the residual $\hat{u}_t$ obtained from the spurious regression of y_t on $\mathbf{x}_t$ becomes

$$\hat{u}_t = \frac{1}{\phi(1)}\{s_{et} - \mathbf{x}_t'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{s}_e\} + O_p(1),$$

where $\mathbf{s}_e = (s_{e1},s_{e2},\dots,s_{eT})'$ and $\mathbf{X} = (\mathbf{x}_1,\dots,\mathbf{x}_T)'$. Therefore, using (3.11) we have

$$T^{-1/2}s_{et} = T^{-1/2}\sum_{j=1}^{[Tr]} e_j \Rightarrow \sigma_e W(a), \tag{A.5}$$

$$T^{-1/2}\hat{u}_t = \frac{1}{\phi(1)}\left\{T^{-1/2}s_{et} - (T^{-1/2}\mathbf{x}_t)'\left(\frac{\mathbf{X}'\mathbf{X}}{T^2}\right)^{-1}\frac{\mathbf{X}'\mathbf{s}_e}{T^2}\right\} + o_p(1) \Rightarrow \varpi_{11}B(a), \tag{A.6}$$

where $\varpi_{11} = \sigma_e/\phi(1)$ and $B(a)$ is defined in Theorem 3.1.

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

Next, noting that the $\hat{u}_{t-1}^i$, $i = 1,2,\dots$, are a regular transformation of $\hat{u}_{t-1}$ in the sense of Park and Phillips (2001), we can apply Lemma 2.1 of their paper and obtain

$$T^{-((i+2)/2)} \sum_{t=1}^{T} \hat{u}_{t-1}^i = T^{-1} \sum_{t=1}^{T} (T^{-1/2}\hat{u}_{t-1})^i \Rightarrow \varpi_{11}^i \int B(a)^i \, da,$$

$$T^{-((i+1)/2)} \sum_{t=1}^{T} \hat{u}_{t-1}^i e_t = T^{-1/2} \sum_{t=1}^{T} (T^{-1/2}\hat{u}_{t-1})^i e_t \Rightarrow \sigma_e \varpi_{11}^i \int_0^1 B(a)^i \, dW(a),$$

$$T^{-((i+1)/2)} \sum_{t=1}^{T} \hat{u}_{t-1}^i \mathbf{s}_t = T^{-1/2} \sum_{t=1}^{T} (T^{-1/2}\hat{u}_{t-1})^i \mathbf{s}_t = O_p(1),$$

where $\mathbf{s}_t$ are the tth row of $\mathbf{S}$ and $\mathbf{S}$ is the $T \times (k + p(k+1))$ I(0) data matrix defined following (3.7). Using the preceding results, and defining the scaling matrix $\mathbf{D}_T = \operatorname{diag}(T^{-1}, T^{-3/2}, T^{-2})$, it is readily seen that

$$\begin{aligned}
\mathbf{D}_T \hat{\mathbf{U}}_{-1}^{(3)\prime} \mathbf{Q}_1 \mathbf{e} &= \mathbf{D}_T \hat{\mathbf{U}}_{-1}^{(3)\prime} \mathbf{e} + T^{-1/2} (\mathbf{D}_T \hat{\mathbf{U}}_{-1}^{(3)\prime} \mathbf{S})(T^{-1}\mathbf{S}'\mathbf{S})^{-1}(T^{-1/2}\mathbf{S}'\mathbf{e}) \\
&= \mathbf{D}_T \hat{\mathbf{U}}_{-1}^{(3)\prime} \mathbf{e} + o_p(1) \Rightarrow \sigma_e \begin{bmatrix} \varpi_{11} \int B \, dW \\ \varpi_{11}^2 \int B^2 \, dW \\ \varpi_{11}^3 \int B^3 \, dW \end{bmatrix},
\end{aligned} \tag{A.7}$$

$$\begin{aligned}
\mathbf{D}_T \hat{\mathbf{U}}_{-1}^{(3)\prime} \mathbf{Q}_1 \hat{\mathbf{U}}_{-1}^{(3)} \mathbf{D}_T &= \mathbf{D}_T \hat{\mathbf{U}}_{-1}^{(3)\prime} \hat{\mathbf{U}}_{-1}^{(3)} \mathbf{D}_T + T^{-1} (\mathbf{D}_T \hat{\mathbf{U}}_{-1}^{(3)\prime} \mathbf{S})(T^{-1}\mathbf{S}'\mathbf{S})^{-1}(\mathbf{S}'\hat{\mathbf{U}}_{-1}^{(3)} \mathbf{D}_T) \\
&= \mathbf{D}_T \hat{\mathbf{U}}_{-1}^{(3)\prime} \hat{\mathbf{U}}_{-1}^{(3)} \mathbf{D}_T + o_p(1) \\
&\Rightarrow \begin{bmatrix} \varpi_{11}^2 \int B(a)^2 \, da & \varpi_{11}^3 \int B(a)^3 \, da & \varpi_{11}^4 \int B(a)^4 \, da \\ \varpi_{11}^3 \int B(a)^3 \, da & \varpi_{11}^4 \int B(a)^4 \, da & \varpi_{11}^5 \int B(a)^5 \, da \\ \varpi_{11}^4 \int B(a)^4 \, da & \varpi_{11}^5 \int B(a)^5 \, da & \varpi_{11}^6 \int B(a)^6 \, da \end{bmatrix}.
\end{aligned} \tag{A.8}$$

It is also straightforward to show under the null that

$$\begin{aligned}
\hat{\sigma}_{FNEC}^2 &= T^{-1} \sum_{t=1}^{T} \left(e_t - \hat{\delta}_1 \hat{u}_{t-1} - \hat{\delta}_2 \hat{u}_{t-1}^2 - \hat{\delta}_3 \hat{u}_{t-1}^3 - (\hat{\boldsymbol{\omega}} - \boldsymbol{\omega})' \Delta \mathbf{x}_t - \sum_{i=1}^{p} (\hat{\boldsymbol{\psi}}_i - \boldsymbol{\psi}_i)' \Delta \mathbf{z}_{t-i} \right)^2 \\
&= T^{-1} \sum_{t=1}^{T} e_t^2 + o_p(1) \xrightarrow{p} \sigma_e^2,
\end{aligned} \tag{A.9}$$

where we have used $T^{1+((i-1)/2)}\hat{\delta}_i = O_p(1)$, $i = 1,2,3$, $\sqrt{T}(\hat{\boldsymbol{\omega}} - \boldsymbol{\omega}) = O_p(1)$, and $\sqrt{T}(\hat{\boldsymbol{\psi}}_i - \boldsymbol{\psi}_i) = O_p(1)$. Using (A.7)–(A.9) in (A.1), we obtain the required asymptotic null distribution of the F_{NEC} test.

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

Under the alternative hypothesis, F_{NEC} can be expressed as

$$F_{NEC} = \frac{1}{3\hat{\sigma}^2_{FNEC}} \left\{ \begin{array}{c} \boldsymbol{\delta}'\hat{\mathbf{U}}^{(3)\prime}_{-1}\mathbf{Q}_1\hat{\mathbf{U}}^{(3)}_{-1}\boldsymbol{\delta} + \mathbf{e}^{*\prime}\mathbf{Q}_1\hat{\mathbf{U}}^{(3)}_{-1}(\hat{\mathbf{U}}^{(3)\prime}_{-1}\mathbf{Q}_1\hat{\mathbf{U}}^{(3)}_{-1})^{-1}\hat{\mathbf{U}}^{(3)\prime}_{-1}\mathbf{Q}_1\mathbf{e}^* \\ + 2\boldsymbol{\delta}'\mathbf{U}^{(3)}_{-1}\mathbf{Q}_1\hat{\mathbf{U}}^{(3)}_{-1}(\hat{\mathbf{U}}^{(3)\prime}_{-1}\mathbf{Q}_1\hat{\mathbf{U}}^{(3)}_{-1})^{-1}\hat{\mathbf{U}}^{(3)\prime}_{-1}\mathbf{Q}_1\mathbf{e}^* \end{array} \right\}, \tag{A.10}$$

where $\boldsymbol{\delta} = (\delta_1, \delta_2, \delta_3)'$, $\mathbf{e}^* = (e_1^*, \ldots, e_T^*)'$, $e_t^* = e_t + (\mathbf{u}_t^{(3)} - \hat{\mathbf{u}}_t^{(3)})'\boldsymbol{\delta}$, $\mathbf{U}^{(3)}_{-1} = (\mathbf{u}_0^{(3)}, \ldots, \mathbf{u}_{T-1}^{(3)})'$, and $\mathbf{u}_t^{(3)} = (u_t, u_t^2, u_t^3)'$. Because u_t, $\hat{u}_t$, u_t^2, $\hat{u}_t^2$ u_t^3, $\hat{u}_{t-1}^3$ are all I(0) and $\hat{\boldsymbol{\beta}}_x - \boldsymbol{\beta}_x = O_p(T^{-1})$ under the alternative, it is easily seen that

$$T^{-1}\hat{\mathbf{U}}^{(3)\prime}_{-1}\mathbf{Q}_1\hat{\mathbf{U}}^{(3)}_{-1} = O_p(1), \qquad T^{-1/2}\hat{\mathbf{U}}^{(3)\prime}_{-1}\mathbf{Q}_1\mathbf{e} = O_p(1), \tag{A.11}$$

where we used that u_{t-1}, $\hat{u}_{t-1}$ and all lagged I(0) regressors in $\mathbf{S}$ are uncorrelated with e_t. Next, tedious but straightforward computations can be used to show that

$$T^{-1/2}\hat{\mathbf{U}}^{(3)\prime}_{-1}\mathbf{Q}_1\mathbf{e}^* = T^{-1/2}\hat{\mathbf{U}}^{(3)\prime}_{-1}\mathbf{Q}_1\mathbf{e} + o_p(1) = O_p(1);$$

$$T^{-1}\mathbf{U}^{(3)}_{-1}\mathbf{Q}_1\hat{\mathbf{U}}^{(3)}_{-1} = O_p(1). \tag{A.12}$$

Also, we have

$$\begin{aligned}\hat{\sigma}^2_{FNEC} &= T^{-1}\sum_{t=1}^{T}\left[e_t^* - (\hat{\boldsymbol{\delta}} - \boldsymbol{\delta})'\hat{\mathbf{u}}^{(3)}_{t-1} - (\hat{\boldsymbol{\omega}} - \boldsymbol{\omega})'\Delta\mathbf{x}_t - \sum_{i=1}^{p}(\hat{\boldsymbol{\psi}}_i - \boldsymbol{\psi}_i)'\Delta\mathbf{z}_{t-i}\right]^2 \\ &= T^{-1}\sum_{t=1}^{T}e_t^{*2} + o_p(1) = O_p(1),\end{aligned} \tag{A.13}$$

where we used $\hat{\boldsymbol{\delta}} - \boldsymbol{\delta} = o_p(1)$, $\hat{\boldsymbol{\omega}} - \boldsymbol{\omega} = O_p(T^{-1/2})$, and $\hat{\boldsymbol{\psi}}_i - \boldsymbol{\psi}_i = O_p(T^{-1/2})$. Therefore, using (A.11)–(A.13) in (A.10), we obtain

$$F_{NEC} = \frac{1}{3\hat{\sigma}^2_{FNEC}}\boldsymbol{\delta}'\hat{\mathbf{U}}^{(3)\prime}_{-1}\mathbf{Q}_1\hat{\mathbf{U}}^{(3)}_{-1}\boldsymbol{\delta} + O_p(T^{1/2}) + O_p(1) = O_p(T), \tag{A.14}$$

which clearly shows that the F_{NEC} statistic diverges to infinity as $T \to \infty$.

It is trivial to derive the asymptotic distributions of the F^*_{NEC} and t_{NEC} tests and prove their consistency as they are special cases of the F_{NEC} test.

We move on to prove results for the t_{NEG} test. Note that the AR representation of $\Delta\hat{u}_t$ is given by $\xi_t = \varphi(L)\Delta\hat{u}_t$, where $\varphi(L) = 1 - \sum_{i=1}^{\infty}\varphi_i L^i$; see (3.9). The t_{NEG} statistic can be written as

$$t_{NEG} = \frac{\hat{\mathbf{u}}^{3\prime}_{-1}\mathbf{Q}_2\Delta\hat{\mathbf{u}}}{\sqrt{\hat{\sigma}^2_{NEG}\hat{\mathbf{u}}^{3\prime}_{-1}\mathbf{Q}_2\hat{\mathbf{u}}^3_{-1}}} = \frac{T^{-2}\hat{\mathbf{u}}^{3\prime}_{-1}\mathbf{Q}_2\Delta\hat{\mathbf{u}}}{\sqrt{\hat{\sigma}^2_{NEG}(T^{-4}\hat{\mathbf{u}}^{3\prime}_{-1}\mathbf{Q}_2\hat{\mathbf{u}}^3_{-1})}}, \tag{A.15}$$

where $\mathbf{Q}_2$ is defined following (3.10). Following Phillips and Ouliaris (1990), it is straightforward to show that

$$T^{-1/2}\sum_{t=1}^{[Ta]}\xi_t \Rightarrow \varphi(1)\varpi_{11}B(a), \qquad \frac{1}{T}\sum_{t=1}^{T}\Delta\hat{u}_t^2 \Rightarrow \varpi_{11}^2(1 + \boldsymbol{\tau}'\boldsymbol{\tau}), \tag{A.16}$$

$$\frac{1}{T}\sum_{t=1}^{T}\xi_t^2 \Rightarrow [\varphi(1)]^2\varpi_{11}^2(1 + \boldsymbol{\tau}'\boldsymbol{\tau}), \tag{A.17}$$

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms

where $\boldsymbol{\tau}$ is defined in Theorem 3.1. We also have

$$T^{-4}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_2\hat{\mathbf{u}}_{-1}^{3} = T^{-4}\hat{\mathbf{u}}_{-1}^{3\prime}\hat{\mathbf{u}}_{-1}^{3} + o_p(1) \Rightarrow \varpi_{11}^{6}\int_0^1 B(a)^6\,da, \tag{A.18}$$

$$T^{-2}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_2\Delta\hat{\mathbf{u}} = T^{-2}\hat{\mathbf{u}}_{-1}^{3\prime}\boldsymbol{\xi} + o_p(1) \Rightarrow \varphi(1)\varpi_{11}^{4}\int_0^1 B(a)^3\,dB(a), \tag{A.19}$$

where $\boldsymbol{\xi} = (\xi_1,\ldots,\xi_T)'$ and the equality in (A.18) and (A.19) follows from the straightforward adaptation of the Said and Dickey (1984) approach as employed in (A.16) of Phillips and Ouliaris (1990) and

$$\hat{\sigma}_{NEG}^2 = T^{-1}\sum_{t=1}^{T}\left(\Delta\hat{u}_t - \tilde{\delta}\hat{u}_{t-1}^3 - \sum_{i=1}^{p(T)}(\hat{\varphi}_i - \varphi_i)\Delta\hat{u}_{t-i}\right)^2 = T^{-1}\sum_{t=1}^{T}\xi_t^2 + o_p(1)$$

$$\Rightarrow [\varphi(1)]^2\varpi_{11}^2(1 + \boldsymbol{\tau}'\boldsymbol{\tau}), \tag{A.20}$$

where we used (A.17), $\tilde{\delta} = O_p(T^{-2})$, and $(\hat{\varphi}_i - \varphi_i) = O_p(T^{-1/2})$. Using (A.18)–(A.20) in (A.15), we obtain the required results for the asymptotic distribution of the t_{NEG} test.

Next, under the alternative hypothesis, t_{NEG} can be expressed as

$$t_{NEG} = \delta\sqrt{\frac{\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_2\hat{\mathbf{u}}_{-1}^{3}}{\hat{\sigma}_{NEG}^2}} + \frac{\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_2\Delta\hat{\mathbf{u}}^*}{\sqrt{\hat{\sigma}_{NEG}^2\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_2\hat{\mathbf{u}}_{-1}^{3}}}, \tag{A.21}$$

where $\Delta\hat{\mathbf{u}}^* = (\Delta\hat{u}_1^*,\ldots,\Delta\hat{u}_T^*)'$ and $\Delta\hat{u}_t^* = \Delta\hat{u}_t + \delta(u_t^{(3)} - \hat{u}_t^{(3)})$. As before, it is straightforward to show that

$$T^{-1}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_2\hat{\mathbf{u}}_{-1}^{3} = O_p(1), \qquad T^{-1/2}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_1\Delta\hat{\mathbf{u}}^* = O_p(1), \tag{A.22}$$

$$\begin{aligned}\hat{\sigma}_{NEG}^2 &= T^{-1}\sum_{t=1}^{T}\left(\Delta\hat{u}_t - (\tilde{\delta} - \delta)\hat{u}_{t-1}^3 - \sum_{i=1}^{p(T)}(\hat{\varphi}_i - \varphi_i)\Delta\hat{u}_{t-i}\right)^2 \\ &= T^{-1}\sum_{t=1}^{T}\xi_t^2 + o_p(1) = O_p(1),\end{aligned} \tag{A.23}$$

where we used $\tilde{\delta} - \delta = O_p(T^{-1/2})$ and $\hat{\varphi}_i - \varphi_i = O_p(T^{-1/2})$. Using (A.22) and (A.23) in (A.21), we have

$$t_{NEG} = \sqrt{T}\delta\sqrt{\frac{T^{-1}\hat{\mathbf{u}}_{-1}^{3\prime}\mathbf{Q}_2\hat{\mathbf{u}}_{-1}^{3}}{\hat{\sigma}_{NEG}^2}} + O_p(1) = O_p(\sqrt{T}), \tag{A.24}$$

which shows that the t_{NEG} statistic diverges to negative infinity as $T \to \infty$. ■

This content downloaded from 130.192.113.29 on Fri, 19 Jan 2018 15:10:59 UTC
All use subject to http://about.jstor.org/terms