Testing for Nonlinear Cointegration Between Stock Prices and Dividends

George Kapetanios, Andy Snell and Yongcheol Shin

April 30, 2003

1 Introduction

In a seminal paper, Campbell and Shiller(1987) investigate the existence of linear cointegration between aggregate US stock prices and US 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 estimates 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.

In this paper we test for cointegration of asset prices and dividends for eleven stock portfolios allowing for smooth but nonlinear adjustement to equilibrium. The motivation for nonlinearity is the existence of transactions costs via a time varying bid ask spread.

The idea that the bid-ask spread in particular and transactions costs in general will affect the equilibrium expected returns on assets is now well established. Existing literature has focused on either the assets' liquidity costs (see in particular, Amihud and Mendelson, henceforth, AM, JFE, Dec.1986, pp223-249) or the possible adverse selection effects in its market arising from asymmetric information about its fundamental value (Merton, 1987, JOF, pp483-510) rather than transactions costs per se. AM emphasise the role of liquidity costs arising through randomly drawn holding periods of individual agents. Their model predicts that expected returns are an increasing but concave function of spreads. Once incoroporated in an empirical model, they find that allowing for concave effects of the spread drives out the size effect of

Banz(1981). Subsequent papers have given further support to this empirical result (see for example Shen, 1993, Federal Reserve Board of Kansas City mimeo).

In a slightly different vein, Jouini (Journal of Mathematical Economics, December 2000) examines the impact of fixed transaction costs and variable adverse selection costs on the mapping of the no arbitrage condition to the price functional. They find that the no arbitrage condition implies that the price functional (effectively, fundamental prices) will lie within the bid ask spread.[1]. In this paper we examine the extent to which spreads affect the *dynamic* adjustment of asset prices to equilibrium. We draw on standard theoretical models of the spread arising from adverse selection (notably, Glosten and Milgrom, JFE, 1985 and Kyle, Econometrica, 1985) to guide an empirical specification that allows nonlinear adjustment of asset prices to an equilibrium. We discuss how the introduction of additional fixed transaction costs may lead to equilibrium mispricing (where here, equilibrium could be defined by say a no arbitrage condition or by optimal trading rules)..

In order to carry out our empirical investigation of the nonlinear adjustment of asset prices to equilibrium we must first develop the necessary econometric theory to analyse the properties of the relevant models and test statistics. The econometric model we adopt, called the $ESTAR$ model (Exponential Smooth Transition Autoregression), and its corresponding test statistics are developed in the next section. Section 3 outlines a model of asymmetric information in a dealer market in which the bid ask spread adjusts in a nonlinear fashion over time following an information shock. Whilst we cannot derive the exact $ESTAR$ econometric specification from this theory, we argue that the key features of the $ESTAR$ model fit the theory well. Section 4 tests for cointegration against nonlinear $ESTAR$ alternatives for 11 major stock price indices using monthly data since 1974..We find that whilst standard linear Engel Granger tests fail on the whole to reject the null of no cointegration, the nonlinear tests give broad support to the cointegration hypothesis. Impulse response functions estimated under the alternative $ESTAR$ model indicate a very rapid adjustment towards equilibrium (typically <6 months half life) when the shocks are large (i.e. shocks of four

[1] This is in contrast to Eeckhoudt (European Financial Management v5, n3 November 1999 pp323-40) who finds that a monopolistic risk averse market maker may sometimes set a spread that deliberately does not contain his best forecast of the fundamental. However, Eeckhouldt's model is not a full equilibrium model and takes no account of the other side ofthe market.

standard deviations in size) but an extremely slow adjustment rate (typically >5 years half life) when shocks are small.(i.e. one standard deviation in size).

2 Econometric Theory and Tests

Before moving to the theory of transactions costs and nonlinear adjustment to equilibrium in assets markets, we develop the econometric theory that we shall use in the empirical work below.

We are interested in developing tests of cointegration under the alternative of nonlinear adjustment to the linear cointegrating vector. Crucially, we take into account the estimation of the cointegration vector in the testing procedure. We consider the general model

$$\Delta y_t = F(z_{t-1}) + \eta_t \tag{1}$$
$$\Delta \mathbf{x}_t = \mathbf{u}_t \tag{2}$$

where where $z_t = y_t - \boldsymbol{\beta}\mathbf{x}_t$, and where η_t and $\mathbf{u}_t$ are respectively a scalar and a $kx1$ vector of stationary variables whose exact properties will be defined subsequently. We now suggest particular functional forms for the function $F(.)$. We suggest a functional form corresponding to the exponential smooth transition autoregressive ($ESTAR$) models. This functional form is discussed in detail in [**?**] and [**?**]. It is given by

$$F(z_{t-1}) = \lambda(1 - e^{-\theta z_{t-1}^2})z_{t-1}$$

We define the null of no cointegration as

$$H_{02} : \theta = 0$$

If $\boldsymbol{\beta}$ were known then the test could be carried out straightforwardly along the lines suggested by [**?**] and [**?**]. But $\boldsymbol{\beta}$ is neither known nor identified under the null (the Davies problem).Furthermore the null is defined in terms of paremeters that are likewise not identified when it is true (θ for the $ESTAR$ model). We will deal with this manifestation of the Davies problem by constructing an auxiliary regression which will be used for testing the null hypothesis H_{02}. This follows the tradition of testing for linearity versus $ESTAR$ behaviour using an LM test of the significance of the second or third power of the variable contained in $F(.)$ in an auxiliary regression. Because an LM test evaluates the statistic at the null, nuisance parameters that are only present under the alternative vanish and a grid search is not required. Explicitly we use a two stage procedure where in the first stage we estimate $\boldsymbol{\beta}$ and in the second stage we carry out the test using this estimate. Of course, under the null hypothesis, $\boldsymbol{\beta}$ is hard to interpret and its estimate will not tend to a constant but to a random distribution. As we state below and prove in Kapetanios, Shin and Snell (2003) the asymptotic distribution of the test statistic which is in part a function of the asymptotically random elements of $\boldsymbol{\beta}$ is free of nuisance parameters

2.1 STAR Model

For the STAR mechanism we adopt the model

$$\Delta y_t = \sum_{i=0}^{\infty} \boldsymbol{\alpha}_i^{\prime\prime} \Delta \mathbf{x}_{t-i} + \sum_{i=1}^{\infty} \boldsymbol{\gamma}_i \Delta \mathbf{y}_{t-i} + \gamma z_{t-1} \left[1 - \exp\left(-\theta z_{t-1}^2\right)\right] + \epsilon_t \tag{3}$$

$$\Delta \mathbf{x}_t = \sum_{i=1}^{\infty} \mathbf{A}_i \Delta \mathbf{x}_{t-i} + \sum_{i=1}^{\infty} \boldsymbol{\Gamma}_i \Delta \mathbf{y}_{t-i} + \boldsymbol{\varepsilon}_t \tag{4}$$

where $\boldsymbol{\alpha}_i, \mathbf{A}_i, \boldsymbol{\Gamma}_i$ and $\boldsymbol{\gamma}_i$ are $kx1, kxk, kx1$ and $1x1$ parameter arrays, where ϵ_t and $\boldsymbol{\varepsilon}_t$ are (a scalar and a $kx1$ vector respectively) of serially independent errors with finite eighth moments and z_{t-1} is as defined above.We assume that under the null, the VAR defined in (3) and (4) is invertible. We further assume that all processes have zero mean but relax this later to allow for the existence of the usual deterministic elements.Our test directly focuses on a specific parameter, θ, which is zero under the null and positive under the alternative. Hence we test

$$H_0 : \theta = 0, \tag{5}$$

against the alternative

$$H_1 : \theta > 0. \tag{6}$$

Obviously, testing the null hypothesis (5) directly is not feasible, since γ is not identified under the null. See for example Davies (1987). To overcome this problem we follow Luukkonen *et al.* (1988), and derive a t-type test statistic. If we compute a first-order Talyor series approximation to the $ESTAR$ model under the null we get the auxiliary regression

$$\Delta y_t = \delta z_{t-1}^3 + \sum_{i=0}^{\infty} \boldsymbol{\alpha}_i^{\prime\prime} \Delta \mathbf{x}_{t-i} + \sum_{i=1}^{\infty} \gamma_i \Delta \mathbf{y}_{t-i} + \epsilon_t \tag{7}$$

In their test for linear cointegration, Engle and Granger use the residuals from a first stage regression of y_t on $\mathbf{x}_t$ as an "estimate" of z_t. We follow this tradition and estimate (7) by replacing z_{t-1} with $\hat{z}_{t-1}^3$where $\hat{z}_{t-1}$ is the lagged residual $y_{t-1} - \boldsymbol{\beta}^{\prime}\mathbf{x}_{t-1}$. Our test is then just the t-statistic of $\delta = 0$ against $\delta < 0$ [2] To simplify matters, assume for now that the lag polynomials in (3) are finite and that all the relevant lags are included in the auxiliary regression (7). Extension to the infinite order case is discussed below. Our test statistic is then

$$t_{NL} = \delta / s.e.\ \ \delta \quad , \tag{8}$$

where δ is the OLS estimate of δ and $s.e.\ \ \delta$ is the standard error of δ. Our test is motivated by the fact that the auxiliary regression is testing the significance of the score vector of the quasi-likelihood function of the $ESTAR$ model, evaluated at $\theta = 0$. Unlike the case of testing linearity against nonlinearity for the stationary process, the t_{NL} test does not have an asymptotic standard normal distribution.

Theorem 1 *Under the null of a unit root (5) the t_{NL} statistic defined by (8) has the following asymptotic distribution:*

$$t_{NL} \Rightarrow \frac{\int_0^1 B(r)^3 dW(r) dr}{\sqrt{\int_0^1 B(r)^6 dr}}, \tag{9}$$

*where $B(r) = W(r) - \mathbf{W}'(r) \left(\int_0^1 \mathbf{W}(r)\mathbf{W}'(r)dr \right)^{-1} \left(\int_0^1 \mathbf{W}(r)W(r)dr \right)$ and $W(r)$ and $\mathbf{W}(r)$ are respectively scalar and vector standard Brownian motion variates defined on $r \in [0,1]$. Under the alternative hypothesis (6) with the $ESTAR$ model (**??**), the t_{NL} statistic is consistent.*

[2] An LM-type test statistic may be obtained via a similar route. See Granger and Teräsvirta (1993) and Teräsvirta (1994) for more details. The advantage of the t-test over the LM-test is that the t-test deals with one sided alternatives of stationarity explicitly, and thus is expected to be more powerful.

In keeping with the tradition in linear cointegration, we propose a companion test to t_{NL} namely a test which is the analogue to the Engle Granger statistic for linear cointegration. Denoting the t ratio of z_{t-1}^3 from the OLS regression

$$\Delta \hat{z}_t = \delta \hat{z}_{t-1}^3 + \sum_{i=1}^{p(T)} \varphi_i \Delta \hat{z}_{t-i} + error \tag{10}$$

as t_{NLEG}, where $p(T) = O(T^{\frac{1}{3}})$ then

Theorem 2 *Under the null of a unit root (5) t_{NLEG} has the distribution*

For a proof of the above theorems, see Kapetanios, Shin and Snell (2003).

Critical values for the raw data, demeaned data and demeaned and detrended data cases for the two tests are given in Table 1.

3 Application to Asset Pricing in the Presence of Transactions Costs

In this section we overview a simple extension of Kyle's(1985) model of market making in the presence of asymmetric information as a means of illustrating how transaction prices may return to equilibrium following a shock in a nonlinear way. We do not argue that our theory is the most realistic or that it captures all of the markets' features in the context of the actual empirical application. Like all theories it is an abstraction and in our case merely serves to underline the argument that nonlinear rather than llinear adjustment to equilibrium is the mechanism that is likely to be generically relevant in asset markets..

3.1 Adverse selection models of the spread and an ECM for asset prices

Kyle (Econometrica, 1985) and Glosten and Milgrom (JFE, 1985) develop models of the bid ask spread arising purely from adverse selection. Both Glosten and Milgrom (GM) and Kyle assume a risk neutral competitive market maker (and hence one that sets fair prices) facing insiders and noise traders. GM specify unitary buys/sells each time period with a time and state dependent probability that the buy/sell is from an informed trader whilst Kyle characterises the properties of a supply/demand *function* In both models, the spread (depth of market in the case of Kyle) converges to zero as private information is slowly and optimally "released" via the sequential trading of insiders. In a sequential equilibrium, the models lead to the following predictions/market characteristics[3] a) Transaction prices have a martingale property b) The (variance of the) price minus fundamental gap, $(p_t - v)$, declines exponentially to zero c) Trades are uncorrelated d) Insiders act on their inside information in a prolonged "optimal" fashion and the $(p_t - v)$ gap is only closed slowly over time.

To get an explicit ECM for prices, consider Kyle's sequential dealership. Here, the market meets n times within a fixed period and n is allowed to go to infinity. The (forcing) noise trade terms become continuous Brownian motion. In the *tth* market, prices satisfy

$$p_t - v = \frac{1}{2}(p_{t-1} - v) + \lambda_t u_t \tag{11}$$

where $\lambda_t = \frac{1}{2}\sqrt{\frac{\sigma^2_{v|t-1}}{\sigma^2_u}}$ is the depth of market which declines to zero as $p_t -> v$.

In addition to its ECM implications, Kyle's sequential market,also predicts that the measured rate of return (there is no dividend payment) is predicted to be random with an error that has ARCH properties.

In order to implement Kyle's model empirically, we must replace the assumption of a "once and for all draw" of v with a stochastic process (a random walk, say) for v. Under this new scheme, the main features of Kyle's model are little changed. Explicitly we now have that

$$v_t = v_{t-1} + \xi_t$$

where ξ_t is a privately observed shock.The formula in (11) may now be written as

$$p_t - v_t = \frac{1}{2}(p_{t-1} - v_{t-1}) + \varepsilon_t \qquad where \qquad \varepsilon_t = \lambda_t u_t + \xi_t \tag{12}$$

[3]Note that here and henceforth we assume for notational simplicity that the mean of the fundamental is zero. Of course literally this would imply negative fundamental prices with a non-zero probability, which is impossible because of limited liability. A proper interpretation is thererfore that prices and fundamentals are implicitly measured as deviations from a (large) unconditional mean. This will make the probability of negative prices close to zero if the unconditional mean is set sufficiently large.

It is easy to show that the composite error in (12) is heteroscedastic with variance $\frac{1}{4}\sigma^2_{v|t-1} + \sigma^2_\xi$ where $\sigma^2_{v|t-1} = \frac{1+\sigma^2_\xi}{2+\sigma^2_\xi}^{t-1}\sigma^2_0$. Clearly, unless the private information shock is heteroscedastic, the error term converges to being homoscedastic.

So far we have derived a standad ECM with fixed coefficients. However the fixed coefficients aspects of the results are non generic in these types of model. As an example, consider the model by Snell and Tonks(Economic Journal 2003) who extend the simple Kyle model to allow for competing but identical traders who suffer liquidity shocks. In their model it is assumed that the insiders' objective in the tth market is

$$U_t = (v_t - p_t)x_t + \varphi(x_t - u_t)^2 \tag{13}$$

where x_t is the trade.of the typical insider and where v_t follows the same random walk process as above [4]. The (Cournot Nash) equilibrium of the model is symmetric with respect to the n traders who submit identical demands and prices satisfy

$$p_t - v_t = \lambda_t\beta_t(p_{t-1} - v_{t-1}) + \lambda_t u_t \tag{14}$$

where $\lambda_t = \frac{\varphi\sigma^2_{t-1}}{\varphi^2\sigma^2_u - \sigma^2_{t-1}}$ and $\beta_t = \frac{n[\varphi^2\sigma^2_u - \sigma^2_{t-1}]}{\varphi[\varphi^2\sigma^2_u + n\sigma^2_{t-1}]}$ so that the ECM coefficien $\lambda_t\beta_t\{= \frac{n\varphi\sigma^2_{t-1}}{\varphi[\varphi^2\sigma^2_u + n\sigma^2_{t-1}]}\}$ is time varying and increases in the conditional variance of the price-fundamentals gap. It is this sort of scheme that motivates the application of an empirical ECM with nonlinear endogenous adjustment but in order to make any such scheme implementable we require an empirical model for fundamental prices.

[4]We should note that objective is somewhat myopic and inconsistent with trading over time. Firstly, traders may wait for the tth market in the prospect of higher gains. Second if we take the interpretation of u as a liquidity shock literally then following a trade in the tth market, the insider should carry forward unsatisfied liquidity demand $(x_t - u_t)$ to the next period to augment u_{t+1} A model with intertemporal dynamic behaviour such as this would be hard to formulate and solve and is beyond the scope of the current paper.

An alternative rationalisation comes from Madhavan who sees the liquiity shock arising from a desire to rebalance portfolios in response to wealth shocks. We could interpret inclusion of the second term in the objective therefore as a simple attempt to model risk aversion and the portfolio rebalancing that it requires rather than as pure noise trade costs. These issues are discussed further in ST.

3.1.1 A Model of the fundamental value of the security

We follow Campbell and Shiller (JPE, 1981) and adopt a net present value relationship to determine the fundamental value of the security. We assume that the fundamental for asset i satisfies the following expected return relationship in long run equilibrium

$$E(r_{it}|t) = r_i \tag{15}$$

(15) is consistent with Campbell (AER, 1992) who, by loglinearising the representative consumer's budget constraint is able to derive a CAPM-type relationship for securities with constant expected equilibrium returns. It is (obviously) also consistent with a one period APT type of framework.[5]

Using the definition of returns we may develop (15) to get a NPV relationship for prices. Writing the definition of returns explicitly in terms of prices and dividends and then solving forwards (ignoring for the moment the possibility of bubbles by invoking the usual transversality condition) gives

$$E(r_t) = E(\frac{p_{t+1} - p_t + d_{t+1}}{pt}) = r => p_t = \sum_{i=1}^{\infty} \delta^i E(d_{t+i}|t) = v_t \tag{16}$$

where $\delta = \frac{1}{1+r}$,d_{t+1} are dividends paid at the end of time $t+1$ (the stock at time t is assumed to be "ex dividend") and where we have dropped the asset index (i) for simplicity. In keeping with much of the literature and with the balance of empirical evidence, we assume that dividends are an exogenous $I(1)$ process. To give an illustration, suppose that dividends follow a random walk with drift τ, then the long run relationship becomes

$$p_t = \frac{\tau}{r} + \frac{1}{r} d_t \tag{17}$$

For more general linear $I(1)$ processes for dividends, we would have to add a stationary error term to (17) in which case it would become a standard linear cointegration relationship.

[5] The CAPM ,APT and other pricing models would appear to be inconsistent with the risk neutrality of the previous section. However if we assume that the inside information in Kyle's model is idiosyncratic and that portfolios are initially well diversified (e.g. the market maker operates in several stocks) then the risk associated with the inside information would not effect the objectives of the agents.in the model.

In general however prices move as a result of dealers' pro-active inventory control as well as a result of adverse selection so that this simple specification above excludes a multitude of market microstructure effects. For example Snell and Tonks(1996,1998) and Madhavan and Smidt (1993) develop and estimate structural models which imply that price quotes are autocorrelated as an optimal response to the existence of both inventory control and asymmetric information. These effects could be thought of as inducing a pricing error that is non zero and possibly autocorrelated for all t of the form.

$$\varepsilon_t = p_t - v_t \tag{18}$$

Finally it would be natural to add fixed dealership costs per trade. Such costs would be impossible to incorporate in a rigorous fashion in our model. Nonetheless we note that on their own, they would lead to a fixed spread within which there would be no arbitrage pressure on prices to return to equilibrium at all. Insensitivity to equilibrium in a fixed range around equilibrium describes threshold stationary behaviour or TAR for short rather than the $ESTAR$ stationarity that we adopt. Interestingly, Taylor(2002) has shown via numerical simulations that when one aggregates over several TAR processes each with different thresholds the resulting process looks much more like a smooth transition AR process than a TAR. This gives yet further justification for our adoptrion of the $ESTAR$ model for testing stationarity.

Below we apply the nonlinear adjustment model to monthly data on large diversified stock market portfolios and some discussion about this is in order.It may at first seem implausible that market microstructure effects associated with adverse selection should be drawn out over a number of months. Most models of market microstructure deal with hourly rather than monthly data. Clearly the sort of trade-revealed information that moves prices on an hour by hour basis is not the sort we have in mind here. We give two examples of the kinds of process we have in mind. One such process could arise when large financial institutions take controlling or at least influential stakes in the firms whose shares they own. Once acquired, a large stake gives incentives to both monitor and improve the performance of such firms over quite a period of time. The success of such a scheme could well be captured by a large permanent rise in (say) the monthly mean of the dividend process. If the share price falls sufficiently far below the fundamental value in terms of the NPV of dividends, then the institution will attempt to offload its stock.

Given that its stake however, it would be unable to make large sales without severely depressing the selling price - this is particularly true for sales outside the normal market size for which quoted prices in dealership markets are no longer valid. The optimal rate of trade may well be described by the model given above. A second example could arise from the process whereby an analyst undertakes extensive research into companies. The analyst purchases shares in those firms it discovers are undervalued at a sufficiently slow rate to maximise its return. Again the model above could well describe this process which could be drawn out over a series of months. Finally it may be that analysis of a portfolio rather than a single stock may mask any nonlinear adjustment that is occuring in individual stocks via the averaging process. We fully intend to extend the current empirical work to several individual stocks within a single stock market.in future work.

3.2 Empirical application of the ESTAR model to prices and dividends from eleven major world stock markets.

We collected monthly data from January 1974 to November 2002 on end period prices and within period dividend yields for value weighted market portfolio indices of stocks traded on the main exchanges of the following eleven countries:- Germany,Belgium,Canada, Denmark, France,Ireland, Italy, Japan, Netherlands, UK and the US. A dividend series was constructed as the product of dividend yield and prices. As alluded to above we test for the existence of a linear cointegrating relationship between dividends and prices of the form

$$p_t = \beta d_t + u_t \tag{19}$$

The adverse selection models of the spread discussed above motivate the specification of a nonlinear dynamic adjustment mechanism such as the $ESTAR$ model giving the nonlinear ECM model

$$\begin{aligned} \Delta p_t &= \lambda(1 - e^{-\theta z_{t-1}^2}) z_{t-1} + \eta_t && (20) \\ z_{t-1} &= p_{t-1} - \beta d_{t-1} && (21) \\ \Delta d_t &= \mathbf{u}_t && (22) \end{aligned}$$

where η_t is a (possibly autocorrelated) error term which captures other microstructure effects such as specific kinds of noise trading (e.g. index trading) and dealer inventory control mechanisms (whereby the adjustment of prices to elicit inventry-correcting trades generates autocorrelated price movements - see for example Madhavan and Schmidt, 1992 and Snell and Tonks, 1998). Although not presented here, simple ADF tests from an initial data analysis give broad support to the hypothesis that all variates are I(1).

3.2.1 Cointegration tests against the ESTAR alternatives

In this section we test for cointegration of asset prices and dividends for eleven stock portfolios allowing for nonlinear adjustment to equilibrium of the $ESTAR$ variety. The motivation for nonlinearity is the existence of transactions costs via a bid ask spread that varies over stocks. At first we might expect that transactions costs which arise from a fixed spread might motivate the consideration of price adjustment mechanism of the $SETAR$ variety. However our data consists of prices and dividends averaged over a widely diversified portfolio of stocks and it has been shown in numerical simulations that this aggregation process leads to a specification that is better approximated by an $ESTAR$ rather than a $SETAR$ model (see for example Taylor and Sarno, 2003).

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 indices of stocks traded on the main exchanges of the following eleven countries: Germany, Belgium, Canada, Denmark, France, Ireland, Italy, Japan, Netherlands, UK and US. A dividend series was constructed as the product of dividend yield and prices. Although not presented here, simple ADF tests from an initial data analysis give overwhelming support to the hypothesis that all variates are I(1).

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

$$p_t = \beta d_t + u_t. \tag{23}$$

The existence of bid ask spreads discussed above motivate the specification of a nonlinear dynamic adjustment mechanism such as the $ESTAR$ model giving the following nonlinear STAR-ECM model:

$$\Delta p_t = \gamma \left(1 - e^{-\theta u_{t-1}^2}\right) u_{t-1} + \alpha \Delta x_t + \varepsilon_t, \tag{24}$$

where $u_{t-1} = p_{t-1} - \beta d_{t-1}$ and ε_t is a (possibly autocorrelated) error term which captures other microstructure effects such as specific kinds of noise trading (*e.g.* index trading) and dealer inventory control mechanisms whereby the adjustment of prices to elicit inventory-correcting trades generates autocorrelated price movements, see for example Snell and Tonks (1998).

We computed three tests. The first two, t_{EG} and t_{NLEG} are the linear Engel-Granger test and its nonlinear counterpart. The third, t_{NLECM} is the t-ratio on $\hat{u}_{t-1}^3$ in the STAR-ECM formulation 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.[6] We estimated the appropriate auxiliary regressions for $p = 12$ and then dropped all insignificant lags in a single round of general to specific modelling.[7]

The results for the tests are in columns 2 to 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 2 of the t_{EG} tests rejecting the null, albeit at the 1% level. Furthermore, none of the remaining 9 t_{EG} statistics are significant even at the 10% level. By contrast the nonlinear t_{NLEG} test rejects in 7 out of 11 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_{NLECM} with only 5 rejections at standard significance levels although three of these reject also at the 1%. A further two t_{NLECM} statistics are quite close to the 10% critical value.[8]

Table 4 about here

[6] The issue of whether or not stock prices and dividends contain a deterministic time trend in the long run is contentious (see for example Shiller, 2000). However there is a clearly discernible trend in both dividends and prices in our data hence we detrend and demean. It is comforting to note that if we do not detrend but only demean, the results are qualitatively almost identical.

[7] We should note that although further exploration revealed some significant lags beyond 12th order, the test statistics were not in general very sensitive to the choice of lag length.

[8] If the alternative is really true then we could interpret this finding as being somewhat at odds with the Monte Carlo evidence, which generally shows that t_{NLECM} has more power than t_{NLEGDF}. However, there is good prior reason to believe that dividends are weakly exogenous in the system. If true, a bivariate ECM would lack parsimony compared with the univariate specification and this may have lead to a loss in power.

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 $ESTAR$ model,

$$\Delta\hat{u}_t = -\left(1 - \exp\left(-\theta\hat{u}_{t-1}^2\right)\right)\hat{u}_{t-1} + \sum_{i=1}^{12}\varphi_i\Delta\hat{u}_{t-i} + \xi_t. \tag{25}$$

The model has been specialised compared with the general $ESTAR$ considered above 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 specialisation. Under the alternative (and estimation of (25) only makes sense if the alternative is true), θ is scale dependent. To clarify its interpretation and to facilitate numerical convergence, we normalised the $\hat{u}_t$ series to have unit sample variance (a procedure which only makes sense under the alternative). We also used demeaned rather than demeaned and detrended data which is an appropriate procedure, asymptotically, 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 .007 and .017.

To get a feel for what such values imply, Figure 1 below plots impulse response functions (irfs) for the error correction term for initial impulses of 1, 2, 3, 4 standard deviation shocks, respectively. For completeness and comparison we compare the corresponding irf with that obtained from the estimated linear models. The striking thing about the graphs 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 five and twenty years. By contrast, the time taken to recover one half of a large shock (such as 3 or 4 standard deviations) is comparable to that of the linear case and varies between just 4 to 18 months. This implies that data periods dominated by extreme volatility may display substantial reversion of prices towards 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 $ESTAR$ and $SETAR$ models may not be too dissimilar in terms of overall inference in any given (finite) sample.

Figure 1. Impulse Response Functions for the Disequilibrium Error

Table1. Asymptotic Critical Values of the t_{NLEGDF} and t_{NLECM} Statistics

t_{NLEGDF}									
	Case 1			Case 2			Case 3		
k	90%	95%	99%	90%	95%	99%	90%	95%	99%
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

t_{NLECM}									
	Case 1			Case 2			Case 3		
k	90%	95%	99%	90%	95%	99%	90%	95%	99%
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

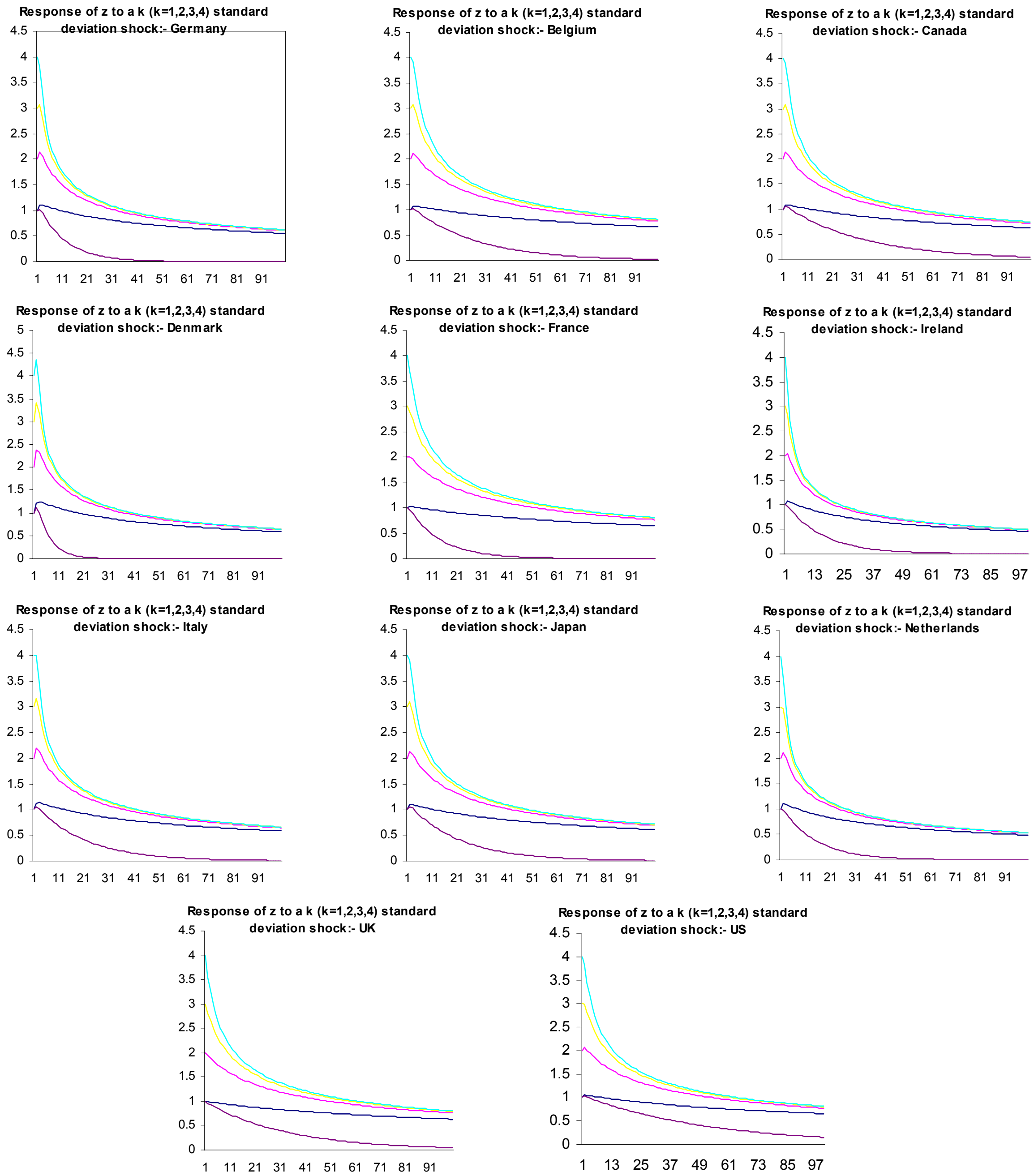
Response of z to a k (k=1,2,3,4) standard deviation shock:- Germany
Response of z to a k (k=1,2,3,4) standard deviation shock:- Belgium
Response of z to a k (k=1,2,3,4) standard deviation shock:- Canada
Response of z to a k (k=1,2,3,4) standard deviation shock:- Denmark
Response of z to a k (k=1,2,3,4) standard deviation shock:- France
Response of z to a k (k=1,2,3,4) standard deviation shock:- Ireland
Response of z to a k (k=1,2,3,4) standard deviation shock:- Italy
Response of z to a k (k=1,2,3,4) standard deviation shock:- Japan
Response of z to a k (k=1,2,3,4) standard deviation shock:- Netherlands
Response of z to a k (k=1,2,3,4) standard deviation shock:- UK
Response of z to a k (k=1,2,3,4) standard deviation shock:- US

Figure 1:

Andrews, D.W.K. and W. Ploberger (1994), "Optimal Tests when a Nuisance Parameter is Present only under the Alternative," *Econometrica*, 62, 1383-1414.

References

[1] Bec. F., M. B. Salem and M. Carrasco (2001), "Tests of Unit-Root versus Threshold Specification with an Application to the PPP," unpublished manuscript, University of Rochester.

[2] Balke, N.S. and T.B. Fomby (1997), "Threshold Cointegration," *International Economic Review*, 38, 627-645.

[3] Berben, R. and D. van Dijk (1999), "Unit Root Tests and Asymmetric Adjustment: A Reassessment," unpublished manuscript, Erasmus University Rotterdam.

[4] Caner, M. and B.E. Hansen (2001), "Threshold Autoregression with a Unit Root," *Econometrica*, 69, 1555-1596.

[5] Chan, K.S., J.D. Petrucelli, H. Tong and S.W. Woolford (1985), "A Multiple Threshold AR(1) Model," *Journal of Applied Probability*, 22, 267-279.

[6] Davidson J. (1994), *Stochastic Limit Theory*, Oxford University Press: Oxford.

[7] Davies, R.B. (1987), "Hypothesis Testing When a Nuisance Parameter is Present Under the Alternative," *Biometrika*, 74, 33-43.

[8] 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.

[9] Enders, W. and C.W.J. Granger (1998), "Unit Root Test and Asymmetric Adjustment with an Example Using the Term Structure of Interest Rates," *Journal of Business and Economic Statistics*, 16, 304-311.

[10] Engle, R. and C.W.J. Granger (1987), "Cointegration and Error Correction: Representation, Estimation and Testing," *Econometrica*, 55, 251-276.

[11] Feller, W. (1957), *An Introduction to Probability Theory and its Applications*. Wiley: New York.

[12] Gonzalez, M. and J. Gonzalo (1998), "Threshold Unit Root Models," *Universidad Carlos III de Madrid, Discussion Paper.*

[13] Hansen, B.E. (1996), "Inference when a Nuisance Parameter is not Identified under the Null Hypothesis," *Econometrica*, 64, 414–430.

[14] Hicks, J.R. (1950), *A Contribution to the Theory of the Trade Cycle*. Oxford University Press: Oxford.

[15] Kapetanios, G. (1999), *Essays on the Econometric Analysis of Threshold Models*. Unpublished Ph.D. Thesis, University of Cambridge.

[16] Kapetanios, G., A. Snell and Y. Shin (2002), "Testing for a Unit Root in the Nonlinear STAR Framework," forthcoming in *Journal of Econometrics*.

[17] Kapetanios, G., A. Snell and Y. Shin (2002), "Testing for Nonlinear Cointegration" Edinburgh University mimeo.

[18] Karatzas I. and S.E. Shreve (1992), *Brownian Motion and Stochastic Calculus*, Springer-Verlag: New York.

[19] Koop, G., M.H. Pesaran and S. Potter (1996), "Impulse Response Analysis in Nonlinear Multivariate Models," *Journal of Econometrics*, 74, 119-147.

[20] Kurtz, T.G. and P. Potter (1991), "Weak Limit Theorems for Stochastic Integrals and Stochastic Difference Equations," *Annals of Probability*, 19, 1035-1070.

[21] Lo, M.C. and E. Zivot (2001), "Threshold Cointegration and Nonlinear Adjustment to the Law of One Price," forthcoming in *Macroeconomic Dynamics*.

[22] Michael, P., R.A. 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.

[23] Pesaran, M.H. and S.M. Potter (1997), "A Floor and Ceiling Model of US Output," *Journal of Economic Dynamics and Control*, 21, 661-695.

[24] Pippenger, M.K. and G.E. Goering (1993), "A Note on the Empirical Power of Unit Root Tests under Threshold Processes," *Oxford Bulletin of Economics and Statistics*, 55, 473-481.

[25] Phillips, P.C.B. and P. Perron (1988), "Testing for a Unit Root in Time Series Regression," *Biometrika*, 75, 335-346.

[26] 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.

[27] Tong, H. (1990), *Nonlinear Time Series: A Dynamical System Approach*, Oxford University Press: Oxford.

[28] Tsay, R.S. (1998), "Testing and Modelling Multivariate Threshold Models," *Journal of the American Statistical Association*, 94, 1188-1202.

[29] Tweedie, R.L. (1975), "Sufficient Conditions for Ergodicity and Recurrence of Markov on a General State Space," *Stochastic Processes Appl.*, 3, 385-403.