Empir Econ (2008) 35:591–606
DOI 10.1007/s00181-007-0180-z

ORIGINAL PAPER

Non-linear cointegration and adjustment: an asymmetric exponential smooth-transition model for US interest rates

David G. McMillan

Received: 15 February 2006 / Accepted: 15 September 2007 / Published online: 4 January 2008

Abstract This paper seeks to extend the extant empirical evidence regarding asymmetric adjustment to equilibrium of short and long interest rates. Using an adaptation of the exponential smooth transition model to allow for sign asymmetry in the transition function, we show that equilibrium reversion exhibits two broad characteristics. First, small deviations are random, while large deviations are reverting. Second, deviations that arise when the long rate exceeds the short rate are characterised by quicker reversion than the opposite case. These results are consistent with the effects of arbitrage and central bank intervention. Finally, forecasting exercises support this model over alternate linear and non-linear specifications.

Keywords Asymmetric adjustment · Cointegration · Interest rates

JEL Classification C22 · G12

1 Introduction

An ongoing research theme in financial econometrics is the examination of asymmetric adjustment between pairs of cointegrating series. Most notably, the dynamics of interest rates, and US interest rates in particular, have been a major source of interest.[1] More specifically, Balke and Fomby (1997), Enders and Granger (1998) and

[1] The examination of non-linear cointegration and adjustment has also been extended to other economic settings, see, for example, Granger and Lee (1989) and Escribano and Pfann (1998) who consider adjustment in sales and employment, respectively, Escribano and Granger (1998) who examine gold and silver prices, and Dwyer et al. (1996), Martens et al. (1998) and Tse (2001) who examine equity spot-futures data.

D. G. McMillan (✉)
School of Management, University of St Andrews, The Gateway, North Haugh,
St Andrews KY16 9SS, UK
e-mail: dgm6@st-andrews.ac.uk

Enders and Siklos (2001) use threshold models to examine the adjustment dynamics of long and short term interest rates. While the behaviour of non-US rates has been examined by van Dijk and Franses (2000) and McMillan (2004), both of whom use smooth-transition models for Dutch and UK interest rates, respectively.

The typical view in this extant research is that such asymmetry is motivated through policy-orientated explanations, whereby central banks, which are primarily interested in some inflation target, will respond asymmetrically in periods of rising and falling inflation (or inflationary expectations).[2] The results from the studies noted above generally support this view of asymmetric intervention through reporting that adjustment of the short rate, which is effectively set by the central bank, is quicker when it is exceeded by the long rate, which is indicative of rising future inflation, than in the converse case. Alternatively, such asymmetry can also arise from investors who require different premiums (perhaps due to different preferences for liquidity or different risk considerations) in periods of rising and falling rates. For example, in periods of falling rates, where reinvestment risk is high, investors may be willing to pay a higher premium to hold long-term assets, than in periods of rising rates. As such empirical modelling has generally focussed upon those models that differentiate between periods of rising and falling rates or changes in rates. More specifically, for US rates the momentum threshold model was supported over the standard threshold approach (Enders and Siklos 2001), for Dutch rates the smooth-transition testing procedure selected the logistic version (van Dijk and Franses 2000), while in a competing models exercise for UK data McMillan (2004) also selected a logistic smooth-transition model.

However, the bond market is, of course, an arbitrage market and bonds of different maturity will be linked by an arbitrage relationship, such that we would expect any deviations from an equilibrium position to be arbitraged away. Therefore, it may be expected that pressures on the movements of rates would not only arise from those exerted according to the sign of the equilibrium deviation but also the size of the deviation. More specifically, whereas the sign of any disequilibrium may trigger asymmetric behaviour as noted above, arbitrageurs might be more interested in the size of the disequilibrium. That is, whether the one period return on bonds of different maturity has drifted apart, for example, should the return on a short rate bond exceed that of a long rate bond, then arbitrageurs will buy the short bond and sell the long bond, so pushing up the price of the short bond and pushing down the price of long bond, until their respective one-period returns are equalised subject to any liquidity or other risk premium on the long bond. Furthermore, of course, arbitrageurs' actions may be limited due to the presence of market frictions such as transaction costs, such that the arbitrage activity must be delayed until the benefits from engaging in trade outweigh the costs (Anderson 1997).

Given this, the empirical models noted above are only able to capture one aspect of the potential dynamic influences impacting upon the adjustment of short and long rates to equilibrium. That is, the threshold and logistic smooth-transition models only the capture sign asymmetry, perhaps arising from central bank or asymmetric investor behaviour. While a model such as the exponential smooth-transition model is only

[2] See, for example, Murchison and Siklos (1999) and Svensson (1999).

capable of capturing size non-linearity, perhaps arising from trading cost considerations.[3] Therefore, in this paper we consider a model that is capable of capturing both these influences on the dynamic interaction between short and long rates, the asymmetric exponential-smooth transition (AESTR) model. The remainder of the paper is as follows: Sect. 2 presents the theoretical arguments underlying the paper; Sect. 3 presents the AESTR model; Sect. 4 introduces the data and presents our empirical results, including a forecasting exercise; Sect. 5 concludes.

2 Theoretical background

Let $R(t, m)$ be the continuously compounded yield to maturity of a m period pure discount (zero-coupon) bond at time t. In a standard "no arbitrage" framework this implies that:

$$R(t, m) = \frac{1}{m}\left[\sum_{i=1}^{m} E_t\{R(t+i-1,\ 1)\}\right] + L(t, m) \tag{1}$$

where E_t denotes expectations at time t and $L(t, m)$ represent the term premium. Equation (1) can be viewed as the fundamental relationship that states the long rate is a weighed average of current and expected future short rates. Rearranging the equation in terms of the spread we have the usual expectation hypothesis expression:

$$S(t, m, 1) = \frac{1}{m}\left[\sum_{k=1}^{m}\sum_{i=1}^{k} E_t \Delta R(t+i,\ 1)\}\right] + L(t, m) \tag{2}$$

where $S(t, m, 1)$ represents the spread between the m maturity bond and a one-period bond. This equation shows that, assuming yields are integrated of order one and that premia are stationary, the right-hand side of (2) is stationary, which implies the long and short rates are cointegrated with a vector (1,−1). Furthermore, the spread will equal a constant value, here denoted z, which depends upon investor expectations about future one-period bills and their attitude towards risk.

Deviations of the spread from the relationship identified in Eq. (2) therefore represent arbitrage opportunities, such that should long and short rates deviate from this long-run equilibrium position then arbitrageurs will enter the market to ensure reversion to equilibrium. That is, if the spread $S(t, m, 1) > z$ then arbitrageurs will sell short term bonds and buy long term bonds, whilst if $S(t, m, 1) < z$ then arbitrageurs will sell long term bonds and buy short term bonds. Both actions will ensure adjustment towards equilibrium $S(t, m, 1) = z$.

Non-linear dynamics have been introduced into the empirical examination of the spread through the presence of transactions costs. That is, arbitrageurs will not enter

[3] As discussed by Anderson (1997) where individual arbitrageurs face different transaction costs then the market process of adjustment aggregating over all individual traders will appear smooth, rather than abrupt as suggested by Heaviside threshold models.

the market until the gains from engaging in trade outweigh the costs of trade, such that a band of inaction can arise around the equilibrium position. Where c represents transactions costs an arbitrageur will only sell short bonds and buy long bonds if $S(t, m, 1) - z > c$, and conversely sell long bonds and buy short bonds if $S(t, m, 1) - z < -c$. Should the equilibrium deviation lie in the region $-c < S(t, m, 1) < c$ then no arbitrage will take place. This extension, of course, suggests a Heaviside threshold, that is, abrupt reversion around $|S(t, m, 1) - z| > c$ and therefore implicitly assumes that all traders face the same identical transaction costs. However, given that individual investors may face different transactions costs, due to different liquidity constraints, time spent obtaining information and processes the transaction, as well as individual specific fees, commissions and tax liabilities, the threshold is more likely to be blurred. That is, because not all traders will trade at the same time, as we aggregate over individuals it may be more appropriate to assume a smooth-transition between regimes of behaviour. The allowance for a smooth-transition would therefore support the view that as the deviation from equilibrium becomes larger so the adjustment back to equilibrium will become stronger, which is intuitively plausible.

Finally, whilst the above arguments suggest that the speed of reversion differs between small and large deviations from equilibrium, it remains symmetric for positive and negative deviations. However, it is plausible that the adjustment process may depend on whether it is the long or short-term bond that is over/under-valued, although there is no theoretical guidance as to which sign of deviation will engender stronger reversion. That is, investors may require different term premia in periods of rising or falling rates. Related, a further strand of the literature supports the potential for asymmetric reversion due to the actions of the monetary policy authority. Here the belief is that where, for example, central banks have an explicit inflation target, they pay more attention to rising interest rates than to falling rates due to their different implications for inflation. That is, where the movement of long-term rates provides a guide to inflationary expectations, central banks may respond more quickly in adjusting short-term rates to rising long rates (rising inflationary expectations) than falling to long rates.[4] Thus, we may expect the short rate response to be quicker in periods of rising inflation than falling inflation (see for example Murchison and Siklos 1999; Svensson 1999). Furthermore, in the UK context a recent paper (Martin and Milas 2004) has provided evidence that the Bank of England actively pursues an asymmetric policy by responding more to positive deviations from its inflation target than negative deviations despite the symmetrical nature of the target.

3 Asymmetric exponential smooth-transition model

A standard approach to examine the issue of stationarity in a two-variable system is the Engle and Granger (1987) two-step approach. That is, the long-run relationship is estimated as such:

[4] Supportive empirical evidence is provide by Enders and Siklos (2001), van Dijk and Franses (2000) and McMillan (2004).

$$x_{1t} = \beta_0 + \beta_1 x_{2t} + \varepsilon_t \tag{3}$$

where the variables of interest x_{1t} and x_{2t} are both integrated of order one, the coefficients β_0 and β_1 are the estimated cointegrating parameters and ε_t is the disturbance term. This disturbance term is then used to in the following regression to determine the existence of cointegration:

$$\Delta \varepsilon_t = \alpha \varepsilon_{t-1} + u_t \tag{4}$$

where u_t is a white noise error term and the null hypothesis of non-cointegration is rejected if α is statistically significant and lies in the region $-2 < \alpha < 0$. Should this case arise then the residuals in Eq. (3) are zero-mean stationary, implying that Eq. (3) is a long-run attractor, with deviations revert according to α, and hence adjustment is symmetric.

Enders and Granger (1998) and Enders and Siklos (2001) proposed two asymmetric versions of Eq. (4) in order to capture sign asymmetry. First, the threshold (TAR) model:

$$\Delta \varepsilon_t = \rho_1 \varepsilon_{t-1} I_t + \rho_2 \varepsilon_{t-1} (1 - I_t) + u_t \tag{5}$$

where I_t is a Heaviside indicator function such that it is equal to one when ε_{t-1} is greater than zero (or some other, estimated, threshold value), and zero otherwise. Alternatively, they also proposed a second adjustment mechanism the momentum TAR (MTAR) model:

$$\Delta \varepsilon_t = \rho_1 \varepsilon_{t-1} M_t + \rho_2 \varepsilon_{t-1} (1 - M_t) + u_t \tag{6}$$

where M_t is a Heaviside indicator function such that it is equal to one when $\Delta\varepsilon_{t-1}$ is greater than zero (or some other, estimated, threshold value), and zero otherwise. Following the arguments in the two cited papers and the discussion above, where the long rate is indicative of inflationary expectations, we would expect to see quicker reversion to equilibrium following deviations that arise from increases in the long rate. Indeed this is the result reported by Enders and Siklos (2001) and, in the UK context, by McMillan (2004) using a smooth-transition variant.

In different market settings a further adjustment process has been identified, rather than adjustment occurring due to the sign of the disequilibrium, non-linear adjustment may arise according to the size of the disequilibrium. For example, in the context of the cost-of-carry model between equity index spot and futures Dwyer et al. (1996) and Brooks and Garrett (2002) have proposed that due to the presence of transaction costs deviations from the equilibrium relationship will only be arbitraged away once such deviations have become sufficiently large so the benefits from engaging in trade out way the costs. Further, Tse (2001) argues that given different constraints faced by different arbitrageurs a smooth-transition back to equilibrium is more appropriate when aggregating over the market than the Heaviside switching of the above models. Therefore, the exponential smooth-transition (ESTR) model has become a popular tool when examining such dynamic behaviour:

$$\Delta \varepsilon_t = \delta \varepsilon_{t-1} \left(1 - \exp\left(-\gamma \varepsilon_{t-1}^2\right)\right) + \nu_t ; \qquad \gamma > 0 \tag{7}$$

where the parameter γ measures the speed of transition between the outer and inner regimes that switch symmetrically around the attractor point of zero with ε_{t-1}.[5] Thus, this model implies the dynamics of the middle ground differ from the dynamics of larger deviations. More specifically, the transition function, $F_t = (1 - \exp(-\gamma x_{t-1}^2))$, varies from zero to one, such that when $F_t = 0$ the process is governed by an integrated process at ε_{t-1}, while when $F_t = 1$ the process exhibits reversion to its long-run attractor. Finally, if $\delta = 0$ then the process is integrated.

As noted in the Introduction, the adjustment of short and long rates back to a long-run attractor point may contain both of these dynamic influences, that is, both sign asymmetry and size non-linearity. Therefore, we extend the ESTR model in Eq. (7) by including the asymmetric sign related terms from Eqs. (5) and (6) to allow for both types of adjustment to occur, hence we have two asymmetric ESTR (AESTR) models, one which permits threshold behaviour in the transition function (AESTR-TAR) and one which permits momentum threshold behaviour in the transition function (AESTR-MTAR):

$$\Delta\varepsilon_t = \delta\varepsilon_{t-1} \left(1 - \exp\left(-\gamma_1\varepsilon_{t-1}^2 I_t - \gamma_2\varepsilon_{t-1}^2 1 - I_t\right)\right) + \nu_t ; \qquad \gamma_1, \gamma_2 > 0 \tag{8}$$

$$\Delta\varepsilon_t = \delta\varepsilon_{t-1} \left(1 - \exp\left(-\gamma_1\varepsilon_{t-1}^2 M_t - \gamma_2\varepsilon_{t-1}^2 1 - M_t\right)\right) + \nu_t; \qquad \gamma_1, \gamma_2 > 0 \tag{9}$$

where the two indicator functions, I_t and M_t are defined as above.

For stationarity of the ε_t term (i.e. cointegration between the short and long rates) in the TAR and MTAR models we require the ρ_1 and ρ_2 coefficients to be negative and jointly significantly different from zero. Whilst for there to be asymmetric adjustment we require that $\rho_1 \neq \rho_2$. Similarly, with regard to the ESTR models, for stationarity we require the parameter δ to be significantly negative, while for asymmetric adjustment to be present we require the speed of adjustment parameters to be statistically different, i.e. $\gamma_1 \neq \gamma_2$. As is well known, in the context of thresholds that are not identified under the null hypothesis (of non-stationarity) the usual critical values for rejecting the null are no longer valid (see, for example, Andrews and Ploberger 1994; Hansen 1996). For the TAR and MTAR models Enders and Siklos (2001) provide critical values via Monte Carlo experiments. Therefore, in order to obtain usable critical values for the ESTR models we also conducted a Monte Carlo experiment whereby 50,000 pairs of independent random walk processes of various observation lengths were generated. In addition, we also compute critical values based upon a bootstrapping procedure, whereby the original data is resampled (50,000 times), this latter procedure allows the critical values to be computed in the context of nuisance parameters such as lags of the dependent variable to account for serial correlation. In both cases the ESTR and asymmetric-ESTR models of Eqs. (3) and (7), (8) or (9) were then estimated and the

[5] As with the TAR models the attractor point can be estimated and take a non-zero value.

Table 1 Critical values for stationarity in non-linear models

No. of obs	Exponential smooth transition (ESTR) model			Asymmetric exponential smooth transition: threshold (AESTR-TAR) model			Asymmetric exponential smooth transition: momentum threshold (AESTR-MTAR) model		
	Significance level			Significance level			Significance level		
	1%	5%	10%	1%	5%	10%	1%	5%	10%
50	−3.39	−2.84	−2.55	−3.38	−2.82	−2.54	−3.40	−2.83	−2.55
100	−3.20	−2.64	−2.35	−3.15	−2.62	−2.33	−3.20	−2.63	−2.35
250	−2.94	−2.36	−2.07	−2.91	−2.34	−2.04	−2.92	−2.34	−2.05
500	−2.81	−2.16	−1.85	−2.77	−2.15	−1.83	−2.79	−2.15	−1.83
Bootstrap	−5.91	−4.91	−4.38	−6.18	−5.18	−4.63	−6.43	−5.44	−4.91

Entries are computed critical values for stationarity for each model that is defined in Sect. 3, Eqs. (7)–(9)

critical values for statistical significance at the 1, 5 and 10% levels are presented in Table 1.

4 Data and empirical results

Monthly end of period observations for the US Federal Funds rate and yield on the 10-year Treasury bond over the sample period 1954:07–2004:08 were obtained from the Federal Reserve web-site. The series are chosen as US rates have been widely analysed in the empirical finance literature, and for asymmetries in particular, such that the results reported here can be viewed in comparison with those previously reported (e.g. Balke and Fomby 1997; Enders and Granger 1998; Enders and Siklos 2001). Preliminary unit root test results are not reported, although they do indicate the presence of a unit root in each series.

Using the linear Engle and Granger (1987) two-step method, we test for cointegration. The estimated cointegrating vector, Eq. (3), is reported below, while the cointegrating residual unit root test, Eq. (4), is reported in the first column of Table 2:

$$r_{st} = \underset{(-9.01)}{-1.6586} + \underset{(3.72)}{1.1146} r_{lt} + \varepsilon_t. \tag{10}$$

where r_{st} and r_{lt} represent the values of the short and long rate, respectively, while the figures in parenthesis are heteroscedasticity-robust t-statistics. The results of the cointegrating unit root test, with sufficient lags, four, to ensure white noise residuals in the test equation, support the presence of a single common stochastic trend, and hence, cointegration between the respective series.

To examine for potential asymmetry in the relationship between the short and long rate we estimate the TAR and MTAR models of Eqs. (5) and (6). As noted above, we could utilise the long-run attractor of zero as the threshold point, which is perhaps

Table 2 Model coefficient estimates and specification tests

Coeff's	Linear	TAR	MTAR	ESTR	ESTR	AESTR-TAR	AESTR-TAR	AESTR-MTAR	AESTR-MTAR
ρ_1	−0.052(−3.76)	−0.07(−3.86)	−0.117(−5.09)	−0.274(−6.36)	−0.344(−7.29)	−0.357(−7.14)	−0.343(−7.55)	−0.283(−6.64)	−0.325(−7.09)
ρ_2	–	−0.030(−1.55)	−0.019(−1.17)						
γ_1	–			0.351(10.96)	0.110(7.29)	0.272(10.93)	0.081(5.38)	0.296(7.82)	0.092(4.77)
γ_2	–					0.459(11.61)	0.176(5.89)	0.448(8.16)	0.192(5.11)
L_1	0.300(7.42)	0.303(7.49)	0.341(8.18)	0.321(8.06)	0.334(8.46)	0.332(8.41)	0.339(8.58)	0.309(7.62)	0.326(78.14)
L_2	−0.189(−4.49)	−0.186(−4.41)	−0.181(−4.31)	−0.180(−4.37)	−0.170(−4.16)	−0.167(−4.09)	−0.169(−4.17)	−0.172(−4.18)	−0.171(−4.16)
L_3	0.070(1.67)	0.072(1.72)	0.075(1.81)	0.067(1.65)	0.069(1.70)	0.070(1.73)	0.069(1.71)	0.065(1.58)	0.068(1.67)
L_4	−0.158(−3.89)	−0.158(−3.88)	−0.163(−4.05)	−0.181(−4.57)	−0.183(−4.68)	−0.180(−4.59)	−0.185(−4.75)	−0.168(−4.25)	−0.177(−4.50)
τ	–	1.01	0.26	0	0.84	0	0.20	0	0.36
F		2.40	12.41	–	–	17.50	8.84	5.43	2.86
AIC	1.48132	1.48062	1.46389	1.44125	1.41684	1.41976	1.41166	1.43574	1.42559
BIC	1.51811	1.52476	1.50803	1.48539	1.46834	1.47125	1.46315	1.48724	1.47708
R^2	0.134	0.136	0.151	0.170	0.191	0.188	0.195	0.176	0.184

For model description see Sect. 2. Numbers in parentheses are t-values economically valid in the present context. However, there is no reason to unequivocally believe that the threshold will coincide with the attractor point, therefore following Chan (1993) we search over the potential thresholds, that is, the middle 70% of the ordered data (ε_t in the TAR model and $\Delta\varepsilon_t$ in the MTAR model), to find the value that minimises the sum of squared residuals in the regressions of (5) and (6). With the thresholds so determined, the estimated models are presented in the second and third columns of Table 2, where τ defines the threshold. As noted in Enders and Siklos (2001) to ensure stationarity of the TAR and MTAR models both coefficient estimates must be negative and jointly significantly different from zero. As can be seen from Table 2 both parameters do indeed indicate reversion, while an F-test statistic for the null hypothesis that $\rho_2 = \rho_1 = 0$ is 8.28 and 13.40 for the TAR and MTAR models, respectively. As noted in Sect. 3 standard critical values cannot be used in the context of threshold models, however, using the values derived by Enders and Siklos these tests do indeed indicate reversion, and therefore cointegration. To test whether asymmetric reversion is supported by the data a second F-test is performed, with the null that $\rho_2 = \rho_1$. The resultant test statistic is reported in Table 2 in the F row and for which conventional critical values can be used. This test statistic suggests that the TAR model does not support asymmetric adjustment, as the null hypothesis of symmetry cannot be rejected. However, the MTAR model does support asymmetric adjustment, although the coefficient results appear to suggest that reversion to equilibrium is quicker for positive changes in the error-correction term.

Whilst the TAR and MTAR models capture sign asymmetry, as noted above where an arbitrage relationship exists we may also expect to observe size non-linearity. The ESTR model of Eq. (7) is thus reported in Table 2 including a model that imposes a threshold of zero (column four) and one that estimates the threshold (column 5). In the ESTR model, stationarity is governed by the significance (and negativity) of the parameter δ. As can be seen from both versions of the ESTR model this parameter is indeed both negative and statistically significant (using the critical values presented in Table 1), supporting non-linear reversion to equilibrium arising when deviations from equilibrium become sufficiently large.

Given that there is evidence of both disequilibrium size non-linearity and sign asymmetry, we now consider the AESTR models that are capable of capturing both types of identified behaviour. The AESTR-TAR and AESTR-MTAR models are presented in the last four columns of Table 2, again two versions are estimated one which imposes a threshold of zero and one in which the threshold is estimated. In these models the main analytical interest arises first from the parameter determining stationarity and second, whether there is asymmetry in reversion. As with the ESTR model above, the parameter δ is both negative and statistically significant for both versions of the AESTR-TAR and AESTR-MTAR models. With regard to the parameters γ_1 and γ_2 which determine the speeds of adjustment to equilibrium for positive and negative deviations, an F-test of the null hypothesis of symmetry, i.e. $\gamma_1 = \gamma_2$, is presented in Table 2 and for which all versions of the AESTR model reject this null (although only at the 10% level for the AESTR-MTAR model where the threshold is estimated). Therefore, this model supports not only non-linear adjustment between small and large equilibrium deviations, but also asymmetric adjustment between positive and negative deviations. Furthermore, for both the TAR and MTAR versions of the AESTR model

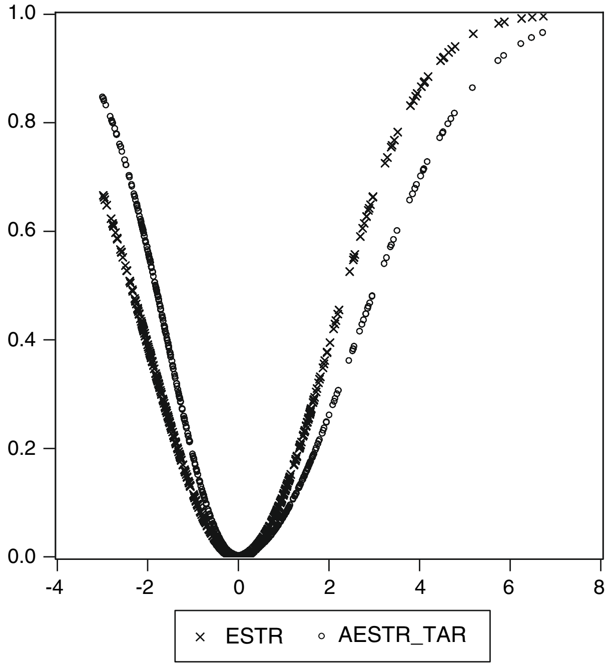

Fig. 1 Transition functions for ESTR and AESTR-TAR

reversion is quicker when the long rate (or changes in the long rate) exceeds the short (or changes in the short rate).

4.1 Model evaluation and forecasting

In order to select a preferred model, we also present in Table 2 some standard specification tests, namely, the adjusted R^2, and the AIC, BIC model selection metrics. On all three of these measures the AESTR-TAR model is selected. To better illustrate these results, we present in Figs. 1, 2 and 3 the estimated transition functions. To aid comparison of the three transition functions they are plotted from the estimated models where zero was imposed as the threshold value. Further, the functions are plotted against each other, that is, in Fig. 1 both the ESTR and AESTR-TAR functions are presented, in Fig. 2 the ESTR and AESTR-MTAR functions are plotted, while in Fig. 3 the AESTR functions are plotted. From these figures we can observe that when the deviation from equilibrium are negative (i.e. long rates exceed short rates) then the ESTR model understates the speed of reversion compared to the AESTR models, whilst when the equilibrium deviation is positive then the ESTR model overstates the speed of reversion. Comparing the asymmetric ESTR models it can be observed that when deviations and changes in equilibrium deviations are negative the speed of reversion in the two models is similar, whilst similarly when deviations are positive and the change in deviations are positive the two models have similar speeds of

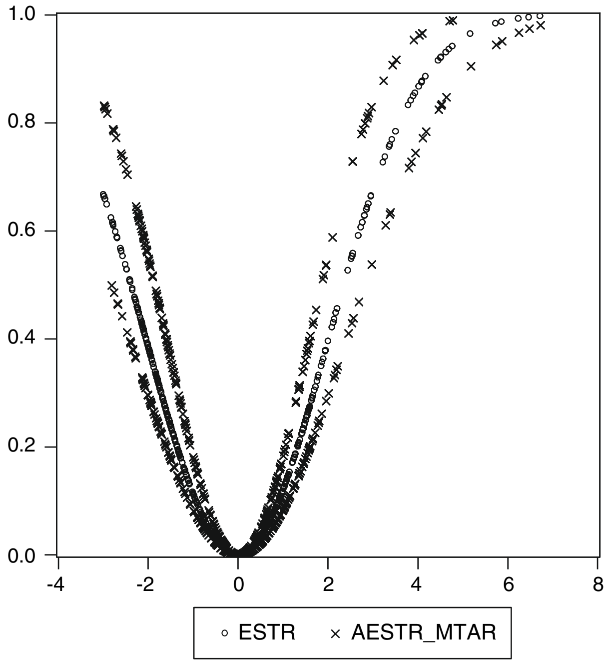

Fig. 2 Transition functions for ESTR and AESTR-MTAR

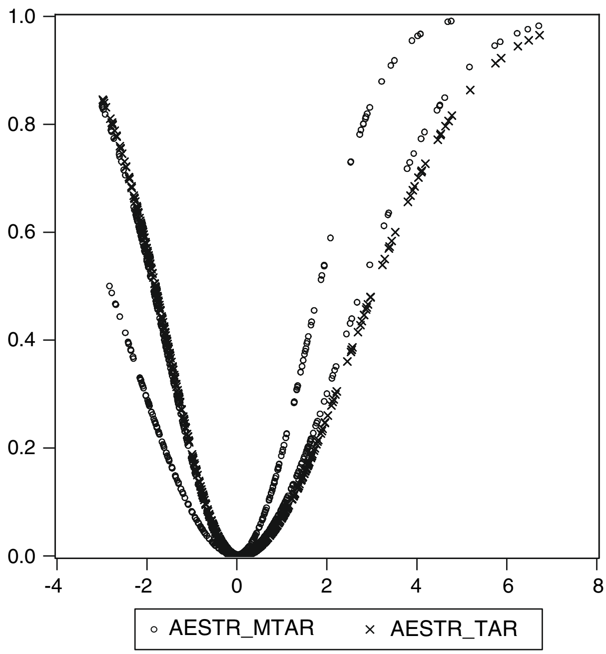

Fig. 3 Transition functions for AESTR-TAR and AESTR-MTAR

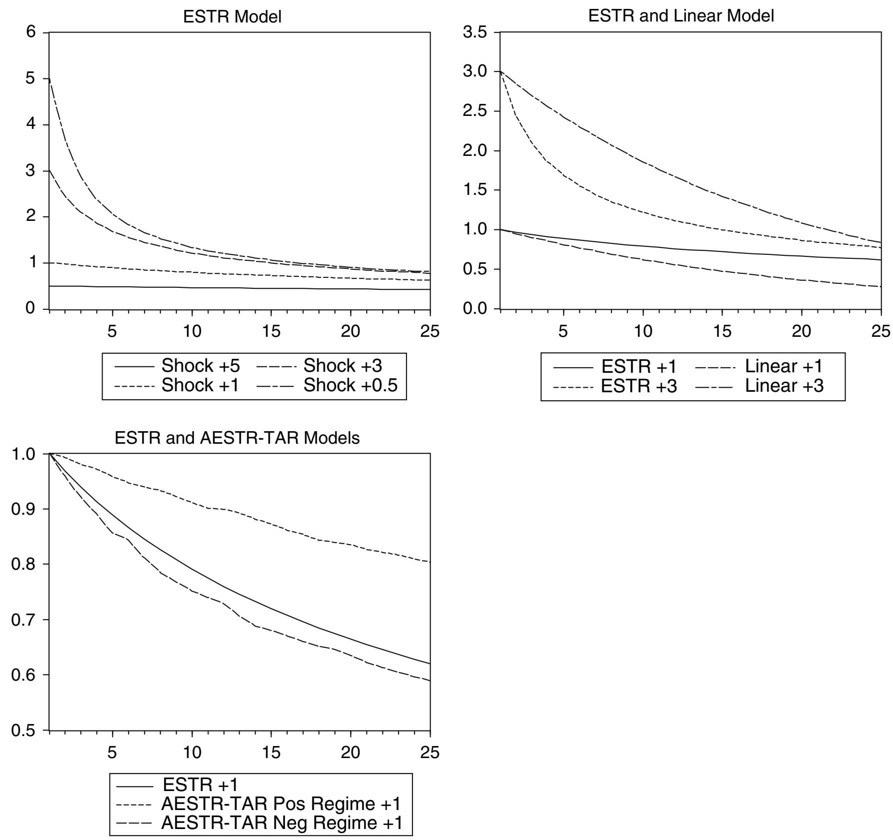

Fig. 4 Impulse response functions

reversion. As a final illustration of the dynamics of these models we present in Fig. 4 some impulse response functions for the linear and ESTR models. The top panel show the response of the ESTR model to shocks of different sizes, the pertinent observation that can be made here is that the larger the shock the quicker is the initial speed of adjustment. That is, the response to small shocks is quite muted, while the effect of large shocks dissipates comparatively quickly. This highlights one of the dynamics characteristics of the ESTR model in that larger shocks engender quicker reversion compared to small shocks. This latter point is further illustrated in the second panel, where the ESTR model is compared to the linear model for two different shocks. Here we can see that the response in the linear setting is the same regardless of the size of the shock. Finally, the lower panel presents the effects of a positive shock on the ESTR model and the AESTR-TAR model according to whether the initial disequilibrium is positive or negative. Here we can note the response to a shock is much slower if the process is initially in the positive regime.

To further analyse the usefulness of our models to market agents or policy makers, we conduct a forecasting exercise for our competing models to determine whether the in-sample specification results are repeated for the out-of-sample data. In order

Table 3 Forecast results

	Linear	TAR	MTAR	ESTR	AEST-TAR	AESTR-MTAR	LSTR	MLSTR
RMSE	0.5784	0.5764	0.5755	0.5784	0.5637	0.5769	0.5835	0.5780
SR	0.56	0.60	0.58	0.56	0.61	0.58	0.60	0.56
Trade	0.07	0.11	0.09	0.07	0.12	0.09	0.09	0.08

Entries are for the root mean squared error (RMSE) statistic, the success ratio (SR) statistic and returns from a trading strategy based upon the estimated models

to conduct this exercise we re-estimated the models over the time period from the beginning of the sample until 1969:12, with the remainder of the sample used in the forecasting exercise. We initially conduct a recursive exercise where one-step ahead forecasts are constructed. That is, we estimated each model over the period 1954:7–1969:12 and forecast the value of $\Delta\varepsilon_t$ for 1970:1, we then roll the end of the sample forward and estimated each model over the period 1954:7–1970:1 and obtained the forecast value for 1970:2, and continue this process until the end of the sample is reached. Thus, at each point in time we are using the data available to market participants and that the parameters are updated. To evaluate these forecasts we used three metrics, first, the standard root mean square error, second, the success ratio, that is, how many times the forecast value is of the same sign as the actual value, and third a "de facto" trading rule for yield changes, where at each point in time, if the forecasted value of $\Delta\varepsilon_t$ is positive then traders expect the short rate to increase relative to the long rate, whilst if the forecast value of $\Delta\varepsilon_t$ is negative then traders expect the long rate to increase relative to the short rate. We then allow traders to earn this change in the yield differential, by, for example in the latter scenario, borrowing at the short rate and lending at the long rate at the beginning of the one month period, with the position then held constant or "locked-in" over the period, before being reversed. These latter two metrics are often argued to be preferable in the context of financial assets as sign information is viewed as more important than overall (size) forecastability. The results are presented in Table 3, where we also consider two additional non-linear models that have been examined in the literature, namely the logistic STR (LSTR) model favoured by van Dijk and Franses (2000) and McMillan (2004) and the modified LSTR (MLSTR) model considered in the context of exchange rates (Sollis et al. 2002), but not previously interest rates.[6]

The results from this exercise again support the AESTR-TAR model as it achieves the lowest RMSE and both the highest success ratio and yield differential trading rule (for interest, the average monthly change in the yield differential is 0.01). Aside from the AESTR-TAR model, both the TAR and MTAR model appear to perform well, whilst in common with van Dijk and Franses (2000) and McMillan (2004) the LSTR model performs reasonably well. This analysis raises two further considerations, first, rather than the change in the spread being the variable of interest, we also want to see whether these models can forecast the spread itself, and second, although the AESTR-TAR model is preferred on the basis of our forecast metrics the gain appears marginal.

[6] More details on these models, including estimation results, are available upon request.

Therefore, in the first instance we use the monthly recursive one-step ahead forecasts to examine the forecastability of the spread itself. For clarity we also focus our attention here on a subset of the models, namely the linear, TAR, MTAR and AESTR-TAR models that are the supported models in the previous analysis. The RMSE for these four models are 0.6084, 0.6233, 0.6265 and 0.6042, respectively, again supporting, albeit marginally, the AESTR-TAR model. For the success ratio and trading rule (which is now defined where traders can earn the yield differential rather than the change in the yield differential, that is, traders are able to swap between yields) results are less unequivocal. Specifically, the value of these two metrics, respectively, are the same at 0.93 and 1.61 for the non-linear models and only slightly lower at 0.92 and 1.59 for the linear model, respectively, (of interest, the average yield differential is 0.99). Again, this suggests only a marginal forecast improvement by the non-linear models.

Finally, these results are perhaps not surprising given that, at each step, we are only forecasting one period ahead. Greater disparity between the models may occur as we forecast further into the future. Therefore, we reconduct the recursive exercise for the spread but only re-estimate the models every five years, that is, instead of only forecasting the next period we are forecasting five years ahead. However, as is well-know, the multi-step forecasting of non-linear processes depends upon assumptions made about the error term. More specifically, where the one-step ahead forecast is given by (for simplicity in an AR(1) framework): $x^f_{t+1} = E(x_{t+1}|I_t) = G(x_t; \beta)$ where $G(.)$ is the non-linear function and E is the expected value, conditional on the information set at time t, I_t. The second-step ahead forecast is given by: $x^f_{t+2} = E(x_{t+2}|I_t) = E(G(x_{t+1}; \beta|I_t)$. Given that linear and non-linear conditional expectations cannot be interchanged, that is $E(G(.)) \neq G(E(.))$, the relationship between the one-step ahead and two-step ahead forecasts is given by: $x^f_{t+2} = E(G(G(x_t; \beta) + \nu_{t+1}; \beta)|I_t) = E(G(x^f_{t+1} + \nu_{t+1}; \beta)|I_t)$.

Therefore, in order to obtain multi-step forecasts we approximate the conditional expectation in $x^f_{t+2} = E(G(G(x_t; \beta) + \nu_{t+1}; \beta)|I_t) = E(G(x^f_{t+1} + \nu_{t+1}; \beta)|I_t)$ through the use of Monte Carlo simulation. That is, the two-step ahead Monte Carlo forecast is given by: $x^f_{t+2} = 1/k\Sigma^k_{i=1} G(x^f_{t+1} + \nu_i; \beta)|I_t)$, where k is the number of repetitions. Previous research presented by Lin and Granger (1994) and Clements and Smith (1997) has suggested that the Monte Carlo approach in obtaining multi-step ahead forecasts in a non-linear setting is favourable compared to alternate approaches that attempt to derive the conditional expectation directly (see also, Brown and Mariano 1989). The results from using $k = 1000$ each period for the TAR, MTAR and AESTR-TAR models are presented in Table 4. These results show the superior forecasting performance of the nonlinear models over the linear model, with the AESTR-TAR model again providing the principle forecasting performance.[7] In addition to the forecast results we also present in Table 4 some interval estimates to determine how accurately the forecast models can describe the distribution of the actual series. For the non-linear

[7] Of interest, the multi-step forecasting performance for the change in the yield is, for the linear model, 0.6109, 0.54 and 0.03 for the RMSE, success ratio and switching rule, for the TAR model, 0.5030, 0.59, 0.08, respectively, for the MTAR model, 0.5038, 0.58, 0.08, respectively, and for the AESTR-TAR model 0.4871, 0.63, 0.10, respectively.

Table 4 Multi-Step Forecast Results

	RMSE	SR	Trade	Quantiles				
				0.1	0.25	0.5	0.75	0.9
Linear	1.9489	0.39	−0.69	−0.09	0.06	0.34	1.02	1.56
TAR	0.2472	0.95	1.63	−0.71	−0.20	−0.002	0.06	0.40
MTAR	0.2460	0.96	1.64	−0.62	−0.20	−0.02	0.07	0.44
AESTR-TAR	0.2390	0.98	1.68	−1.79	−1.08	−0.35	1.20	2.48
Actual values	–	–	–	−2.13	−1.63	−0.46	0.73	1.95

Entries are for the root mean squared error (RMSE) statistic, the success ratio (SR) statistic and returns from a trading strategy based upon the estimated models. Quantiles are the values of each quantile of the actual and forecast series

models these quantile value are determined by the appropriate simulated realisation from the ordered set of 1,000 simulations. These results again support the view that the non-linear models provide a better description of the data than the linear model. In particular, the distribution implied by the linear model is too narrow, as well as having a mean of the wrong sign. Furthermore, of the non-linear models considered, the distribution implied by AESTR-TAR model clearly is the closest match to the actual distribution. Thus, the use of these non-linear models and the AESTR-TAR model in particular, are well suited to capture medium term variations in the interest rate spread.

5 Summary and conclusion

There has been an increased interest in the dynamics behaviour of interest rate series, in particular, whether they exhibit reversion to some long-run equilibrium and whether that reversion is asymmetric. Extant empirical evidence has suggested that short and long rates respond asymmetrically according to the sign of the disequilibrium, or the sign of change in disequilibrium. The standard rationale for which concerns the actions of monetary policy authorities that respond asymmetrically to the expected course of future inflation. Nevertheless, the interaction between short and long rates also gives rise to possible arbitrage trading, such that should rates deviate from their long-run paths by a sufficient amount then market agents will act to restore rates back to their long-run value. As has been reported in other arbitrage markets (e.g. between spot and futures prices), non-linearity arises as deviations from equilibrium must become sufficiently large for arbitrage traders to profit before they enter the market, as such a random walk inner band and a reverting outer band occurs for price deviations. Further, due to different constraints that arbitrageurs may face the process of reversion is smooth.

This paper has, in turn, examined US short and long rates for sign asymmetry, size non-linearity, and presented a new adaptation of the smooth-transition model that allows both for these effects simultaneously. The results presented here support a long-run relationship between short and long rates where small deviations are characterised by random walk behaviour, while larger deviation exhibit reversion. Further, large negative deviations (when the long rate exceeds the short rate) are restored quicker than positive deviations. Finally, a series of forecasting exercises supports this model over an alternative linear and non-linear models, especially when forecasting over the medium term, which may be of particular interest to market participants and policy authorities.

References

Anderson HM (1997) Transaction costs and non-linear adjustment towards equilibrium in the US treasury bill market. Oxf Bull Econ Stat 59:465–484

Andrews D, Ploberger W (1994) Optimal tests when a nuisance parameter is present only under the alternative. Econometrica 62:1383–1414

Balke NS, Fomby T (1997) Threshold cointegration. Int Econ Rev 38:627–643

Brooks C, Garrett I (2002) Can we explain the dynamics of the UK FTSE 100 stock and stock index futures market? Appl Financ Econ 12:25–31

Brown BW, Mariano R (1989) Measures of deterministic prediction bias in nonlinear models. Int Econ Rev 30:667–684

Chan KS (1993) Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model. Ann Stat 21:520–533

Clements MP, Smith J (1997) The performance of alternative forecasting methods for SETAR models. Int J Forecast 13:463–475

Dwyer GP Jr, Locke P, Yu W (1996) Index arbitrage and nonlinear dynamics between the S&P 500 futures and cash. Rev Financ Stud 9:301–332

Enders W, Granger CWJ (1998) Unit-roots tests and asymmetric adjustment with an example using the term structure of interest rates. J Bus Econ Stat 16:304–311

Enders W, Siklos PL (2001) Cointegration and threshold adjustment. J Bus Econ Stat 19:166–176

Engle RF, Granger CWJ (1987) Cointegration and error-correction: representation, estimation and testing. Econometrica 55:251–276

Escribano A, Granger CWJ (1998) Investigating the relationship between gold and silver prices. J Forecast 17:81–107

Escribano A, Pfann GA (1998) Nonlinear error-correction, asymmetric adjustment and cointegration. Econ Model 15:197–216

Granger CWJ, Lee TW (1989) Investigation of production, sales and inventory relationships using multicointegration and nonsymmetric error correction models. J Appl Econom 4:S145–S159

Hansen B (1996) Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica 64:413–430

Lin J-L, Granger CWJ (1994) Forecasting from non-linear models in practice. J Forecast 13:1–10

Martens M, Kofman P, Vorst TCF (1998) A threshold error-correction model for intraday futures and index returns. J Appl Econom 13:245–263

Martin C, Milas C (2004) Modelling monetary policy: inflation targetting in practise. Economica 71:209–221

McMillan DG (2004) Non-linear error-correction: evidence for UK interest rates. Manch Sch 72:626–640

Murchison S, Siklos PL (1999) Central bank reaction functions in inflation and non-inflation targeting economies. Wilfrid Laurier University. Working Paper

Sollis R, Leybourne SJ, Newbold P (2002) Tests for symmetric and asymmetric nonlinear mean reversion in real exchange rates. J Money Credit Bank 34:686–700

Svensson LEO (1999) Inflation targeting as a monetary policy rule. J Monet Econ 43:607–654

Tse Y (2001) Index arbitrage with heterogeneous investors: a smooth transition error-correction analysis. J Bank Financ 25:1829–1855

van Dijk D, Franses PH (2000) Non-linear error-correction models for interest rates in the Netherlands. Barnett WA, Hendry DF, Hylleberg S, Teräsvirta T, Tjøstheim D, Würtz A (eds) Nonlinear econometric modelling. In: Proceedings of he 6th (EC)2 meeting. Cambridge University Press, Cambridge, pp 203–227