

Johansson, Martin

Working Paper

TAR models and real exchange rates

Working Paper, No. 2001:21

Provided in Cooperation with:

Department of Economics, School of Economics and Management, Lund University

Suggested Citation: Johansson, Martin (2001) : TAR models and real exchange rates, Working Paper, No. 2001:21, Lund University, School of Economics and Management, Department of Economics, Lund

This Version is available at:

<https://hdl.handle.net/10419/259854>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

TAR models and real exchange rates.

Martin W Johansson*

November 14, 2001

Abstract

The recent past has seen an increased interest in piecewise linear real exchange rate models. By invoking Heckscher's (1916) 'commodity points' it has been argued that a threshold autoregressive (TAR) model should be used to study movements in the real exchange rate. This paper examines the problems of fitting TAR models to real exchange rates. We find that the power of the tests for TAR behavior can be very low for realistic parameter settings. Moreover the confidence intervalls for the threshold parameter are too wide to be used for economic analysis.

Keywords: PPP, real exchange rate, threshold autoregression.

JEL: F3; C5

*Department of Economics, Lund University, P.O. Box 7082, S-220 07 Lund, Sweden. Tel:+46 (0)462227911, fax: +46 (0)462224118, Email: Martin.Johansson@nek.lu.se. This is a (hopefully) improved version of the first paper in my licentiate thesis. I am especially grateful to Johan Lyhagen (Dept of economic statistics, Stocholm school of economics) who was opponent on my licentiate seminar. I am also grateful to Michael Bergman, Göran Hjelm, Peter Jochumzen and other macroseminar participants at the economics department at the University of Lund.

1 Introduction

Over the years a substantial amount of research on exchange rate dynamics has been devoted to testing purchasing power parity (PPP), see Rogoff (1996) for an extensive overview of the literature. In its crudest form PPP suggests that cross-country price differences on the goods market should be eliminated through frictionless trade thus keeping the real exchange rate (i.e. the price ratio expressed in a common currency) stationary over time. This implies that deviations from PPP are transitory by nature. However, most papers investigating the time series properties of real exchange rates have not been able to reject the null of a unit root (i.e. no mean reversion) and where the null has been rejected it has only been in favor of very slow convergence rates¹. Invoking Heckscher (1916) it has been suggested (Pippenger & Goering, 1993) that the slow convergence rates and the non-rejection of a unit-root may arise from the presence of transaction costs. The idea is that as long as transport costs, tariffs and other market frictions put a cost on trade that is greater than the corresponding gain, cross-country price differences may be sustained indefinitely. However, if the price differences become large enough trade turns profitable again and mean reversion kicks in. This suggests that real exchange rate dynamics should be studied using a model that discriminates between large and small deviations from PPP. A time series model that can be used for such purposes is the Self-Exciting Threshold Autoregression (TAR). A TAR model (see section 3 for further details) works by allowing for regime switching parameters depending on the distance of an observation from the mean. As an added benefit the estimated threshold will indirectly give a measure of the size of the transaction cost. Applications of TAR models to exchange rates include Obstfeld & Taylor (1997), O'Connell (1998), Pippenger & Goering (1998)² and Erjnaes & Persson (2000)³. In this paper we will investigate the usefulness of TAR models for modelling real exchange rate dynamics at a more general level. That is, given typical sample sizes and parameter values, what is the ability of the available techniques to find and estimate TAR models?⁴ More specifically:

1. How does one optimally test for TAR behavior in real exchange rates and what is the power of this test using realistic parameter settings?
2. What does the density of the threshold estimator look like? Are the estimates precise enough to draw any conclusions of the size of the transaction cost or are the thresholds allocated more or less by chance?

¹Studies referred to in Rogoff (1996) suggest half-lives ranging from 3 to 5 years.

²Pippenger & Goering (1998) deal with *nominal* exchange rate forecasting.

³Erjnaes & Persson (2000) deal more specifically with 'the law of one price', LOOP but their methodology is relevant to us.

⁴The present paper bears some resemblance to Balke & Fomby (1997) but deals much more specifically with exchange rate determination.

To answer these questions we will estimate TAR models for a data set of quarterly and monthly real exchange rates and then perform Monte Carlo simulations on the estimated models. It should be stressed that the main objective of this paper is *not* to test for TAR behavior in a certain set of real exchange rates, although we will do that as a byproduct. The main objective is to test what good the available econometric techniques are *if* the real exchange rate does follow a TAR data generating process (DGP). The outline of the paper is as follows. The next section reviews the existing literature. The third section deals with TAR models in general but with specific relevance to exchange rates. The fourth and fifth section contain the empirical results. Section six concludes.

2 Previous literature

The first application of TAR models to real exchange rates was conducted by Obstfeld & Taylor (1997; hereafter Obstfeld & Taylor) who investigated the case for TAR adjustment using a set of monthly subgroups of relative CPI from both international and inter-US data. They estimated the parameters by means of maximum likelihood and obtained reasonable estimates. To test if a standard AR model could be rejected in favor of a TAR model they relied on Monte Carlo simulated likelihood ratio (LR) tests. For most of the cases they found that the standard AR model could not be rejected. They comment this:

‘This is perhaps not surprising given the low power of relatively short univariate time series to distinguish between near unit-root alternatives’.

O’Connell (1998, hereafter O’Connell) takes a less sympathetic tone to TAR modelling of real exchange rates than Obstfeld & Taylor. O’Connell investigates a number of countries using quarterly data often finding that large deviations from PPP are as persistent as small deviations, if not more. However, as noted by O’Connell, the results are not easy to compare. First of all there are differences in the type of data used. Obstfeld & Taylor rely on ratios of different subgroups of consumer price index (CPI) where O’Connell mainly uses a panel of ready-made real exchange rates compiled by *International Financial Statistics* (IFS). Secondly, O’Connell imposes the thresholds rather than estimating them. Thirdly O’Connell and Obstfeld & Taylor use different types of TAR models. A secondary aim of this paper is to find out to what extent the different conclusions in O’Connell and Obstfeld & Taylor can be attributed to the methodologies chosen.

3 Basic TAR models

There are basically two different TAR models to choose from, the EQ-TAR (1) and the Band-TAR (2).

$$\Delta x_t = \begin{cases} \lambda_1 x_{t-1} + e_{1t} & \text{if } x_{t-1} > c \\ \lambda_2 x_{t-1} + e_{2t} & \text{if } |x_{t-1}| < c \\ \lambda_1 x_{t-1} + e_{1t} & \text{if } x_{t-1} < -c \end{cases} \quad (1)$$

$$\Delta x_t = \begin{cases} \lambda_1 (x_{t-1} - c) + e_{1t} & \text{if } x_{t-1} > c \\ \lambda_2 x_{t-1} + e_{2t} & \text{if } |x_{t-1}| < c \\ \lambda_1 (x_{t-1} + c) + e_{1t} & \text{if } x_{t-1} < -c \end{cases} \quad (2)$$

In (1) and (2) x_t is our time series of interest with λ_1 and λ_2 being the adjustment coefficients, c is the threshold separating the two regimes and e_{it} is the noise. As the reader will recognize the TAR models above are generalizations of Dickey-Fuller equations where the strength of the reversion depends on which side of the thresholds the observation is located. The first model (1) exhibits mean reversion towards the mean of the series while (2) exhibits reversion towards the edge of the threshold⁵. Relating to earlier studies Obstfeld & Taylor argue that the Band-TAR makes most sense while O’Connell uses the EQ-TAR (1). We are inclined to support Obstfeld & Taylor on this issue - If the space inside the thresholds represents a zone where trade may be non-profitable there is no reason to believe that traders actively would push the price difference below that level⁶.

3.1 Estimating and testing TAR models

3.1.1 The maximum likelihood estimator

The next question concerns how to estimate a model such as (2). The overriding problem is that if we alter c we alter the set of x_t that belongs to each of the two regimes. Thus we can not use a standard estimation technique such as nonlinear least squares that minimize the residual sum of squares over different parameter values keeping the variable set unaltered. Instead we use the maximum likelihood estimator suggested by Obstfeld & Taylor, see appendix B for details.

3.1.2 The issue of serial correlation

Autocorrelated residuals in autoregressive models is an important issue since neglect may lead to inconsistent parameter estimates. In this paper we choose to follow Peel & Speight

⁵The thresholds are symmetrically placed around the mean of the series with the same adjustment coefficient regardless if the exchange rate is above the upper threshold or below the lower threshold. We decided to stick with these restrictions for two reasons. First of all it is computationally expensive to fit asymmetric TAR models. Secondly, the values of these parameters are (supposedly) governed by transaction costs and since the shipping costs between, say, Germany and Sweden are the same as those between Sweden and Germany there is no reason to assume an asymmetric model.

⁶Another reason for using the Band-TAR instead of the EQ-TAR is that the likelihood function has a unique maximum which facilitates estimation.

(1998) who deal with autocorrelated residuals in TAR models by adding lagged dependent variables on the right hand side, RHS. This means that in the presence of autocorrelation one should estimate (3) rather than (2) which can be seen as a Band-TAR adaptation of the augmented Dickey-Fuller (ADF) test. This method differs substantially from Obstfeld & Taylor who use a computationally expensive grid search procedure to estimate the autoregressive parameter. We find their approach surprising since problems with autocorrelated residuals in the applied econometric literature seldom are solved through Cochrane-Orcutt-type corrections if simpler remedies are at hand. Moreover since (3) performed well in large sample Monte-Carlo experiments we do not see any obvious reason why our approach should be inferior to that of Obstfeld & Taylor's⁷.

$$\Delta x_t = (I_{out})\lambda_1\Phi(x_{t-1}, c) + (I_{in})\lambda_2x_{t-1} + \sum_{j=1}^i \phi_j\Delta x_{t-j} + e_t \quad (3)$$

$$\Phi(x_{t-1}, c) = x_{t-1} - c \text{ if } x_{t-1} > c \quad (4)$$

$$\Phi(x_{t-1}, c) = x_{t-1} + c \text{ if } x_{t-1} < -c \quad (5)$$

$$I_{out} = 1 \text{ if } |x_{t-1}| > c; \quad \text{zero otherwise} \quad (6)$$

$$I_{in} = 1 \text{ if } |x_{t-1}| < c; \quad \text{zero otherwise} \quad (7)$$

3.2 Tsay's (1989) test

As a pretest for TAR behavior one may use Tsay's (1989) test for linearity. We will here give a brief description of the test. For details the reader is referred to Tsay (1989). Assume a stationary AR(p) process, $q_t = (1 \ q_{t-1} \dots q_{t-p})\beta + a_t$ where a_t is the noise. For each (usable) q_t there exists a set of observations $(1 \ q_{t-1} \dots q_{t-p})$ which we call a case of data for the AR(p) model. We now sort the cases according to the size of the regressor which we suspect may signal a regime switch, in our case q_{t-1} . That is, we place the smallest q_{t-1} first and the largest q_{t-1} last. This gives an arranged autoregression, $q_t^* = (1 \ q_{t-1}^* \dots q_{t-p}^*)\alpha + e_t$. If there are nonlinearities of the TAR type then the α -vector associated with small and large values of q_{t-1}^* should be different from that associated with medium sized q_{t-1}^* . This hypothesis can be checked by testing the arranged autoregression for structural breaks. To perform Tsay's test estimate recursively the arranged autoregression and pick out the recursive residuals, ϵ_t . Then perform the following regression $\epsilon_t = (1 \ q_{t-1}^* \dots q_{t-p}^*)\varpi + e_t$ and test if $\varpi \neq 0$ with an F -test (i.e. test if the prediction failures are related to the size of q_{t-1}) and denote the F -statistic F_{FWD} . To increase the power of the test sort the cases in reversed order (the largest q_{t-1} first and the smallest last), repeat the entire procedure with the reversed series and denote the corresponding F -statistic F_{REW} . Pick out the largest F -statistic of the two and use it as the test statistic. If it is larger than the critical value we reject the null of linearity.

⁷It should be noted that estimating (3) comes with the price of imposing equal variance across the two regimes.

4 The data sets

We will estimate our TAR models on two data sets. Intuitively, as the market friction hypothesis is a theory on arbitrage it seems wise to focus on goods that are arbitragable (i.e. not services). The first data set consists of 94 quarterly observations from 1975:1 to 1998:2 on so called VAD rates for 14 European countries (Austria, Belgium, Switzerland, Denmark, Finland, France, Germany, Ireland, Italy, Netherlands, Norway, Spain, Sweden and UK). The VAD rates are trade weighted real exchange rates based on the price levels for manufactured goods. The VAD rates were also used by O'Connell but as mentioned earlier the methodology used by him is quite different from ours - O'Connell assumed a value for the thresholds, we will estimate them. The second data set consists of 233 monthly observations on trade weighted relative CPI levels for the same 14 countries during 1980:1 to 1999:5. Both data sets were obtained from IFS⁸. Prior to analysis the data was converted into logarithms and demeaned. We also excluded some countries from further analysis. The countries that were sorted out displayed real exchange rates that rose or fell during the whole sample. The reason for this sample selection is that the TAR bands are horizontally placed around the mean of the series and it does not make sense to estimate a TAR if there is a marked trend in the real exchange rate⁹. Thus we concentrate our analysis on 'promising' candidates only. Using this rule we excluded Austria, Belgium, Denmark, Germany and the Netherlands from the VAD rates and Austria from the relative CPI rates. Following standard operating procedures we subjected the (remaining) rates to ADF tests¹⁰. For many of the series we either rejected the null of a unit-root or were rather close to doing so. Yet, performing ADF tests on suspected TAR processes is not, as noted by Pippenger & Goering (1993), uncomplicated as the ADF tests are biased towards non-rejection of a unit-root if the true DGP is of the TAR type. Hence we will not dwell any further on them.

5 Estimation and Monte Carlo simulations

In this section we will estimate two sets of TAR models, one for quarterly data and one for monthly data. The purpose is to see if a likelihood ratio (LR) test rejects the AR model

⁸IFS CD-ROM, sept 1999 issue.

⁹Obstfeld & Taylor used a linear trend to deal with this problem.

¹⁰ADF test on VAD rates (Crit value, 10%: -2.58): Ch (-2.96), Fin (-0.61), Fr (-2.54), Ir (-2.20), It (-1.75), No (-2.51), Sp (-2.63), Sw (-2.42), UK (-1.53). The residuals from the ADF regressions showed no signs of autocorrelation (bar Ch and Fr which were corrected by adding one augmentation term) or heteroscedasticity but we rejected the null of normality for all series.

ADF test on Rel. CPI (Crit value, 10%: -2.57): Bel (-2.88), Ch (-2.30), Dk (-1.49), Fin (-1.18), Fr (-3.00), Ger (-2.91), Ir (-2.59), It (-1.78), Nd (-3.51), No (-2.41), Sp (-1.54), Sw (-2.03), UK (-2.19). The residuals from the ADF regressions showed no signs of autocorrelation (after adding one augmentation term) or heteroscedasticity but we rejected the null of normality for all series.

in favor of a TAR model. We will then use the estimated TAR parameters to generate pseudo-data in the next section to investigate the power of the LR test and the distribution of the threshold parameter. Table 1 and 2 report the results from the TAR estimation on actual data. The p-values for the LR tests were calculated using bootstrap techniques, see appendix C for details. The results from the LR tests are rather disappointing for the market friction hypothesis. For the quarterly VAD rates we are only able to reject the null of a standard AR process in one case, Spain, albeit with rather strange parameter values. Turning to the monthly rates, table 2, we are able to reject the AR null in two cases, Italy and Norway. However, when we estimated the models we set as a precondition that at least 10% of the observations should be allocated in one of the regimes. Looking at the rightmost column in table 2 we see that the likelihood function was still increasing for Norway and Italy when estimation was broken off. This feeds the suspicion that the above mentioned rejections are driven by outliers. Even though we seldom reject the AR null it is interesting to note that for many countries the adjustment speed is considerably slower (or even non-existent) in the inner regime compared to that of the outer regime, thus providing some tentative support for the market friction hypothesis. This result is divergent from that reached by O’Connell and hints that the parameter estimates indeed are sensitive to the chosen methodology. We can also compare our estimated threshold levels for the quarterly VAD rates to the fixed threshold ($\bar{c} = 0.1$)¹¹ employed by O’Connell. The mean of our threshold estimates for the VAD rates is 0.043 with the highest being that of Ireland’s, 0.065. Hence our estimated thresholds are considerably smaller than that used by O’Connell. This can to a large extent be attributed to the fact that many of the exchange rates used here (i.e. the promising cases) seldom deviate enough from the mean to end up on the outer side of $\bar{c} = 0.1$. This puts some doubt on the validity of the fixed threshold employed by O’Connell.

5.1 Power of the LR tests

In the previous section we found that not so few of the parameter estimates were consistent with the kind of nonlinear adjustment predicted by the market friction hypothesis, but formal testing only occasionally ‘accepted’ the TAR model. This raises the question of the power of the LR test - that is, given that the data is in fact generated by a TAR process, what are our chances of detecting it using a LR test? To investigate this we performed some Monte Carlo experiments. The strategy of these experiments is as follows:

1. Generate a TAR model with the same number of observations as the empirical series (i.e. 93 or 231) and estimate the model under both the AR null and the TAR alternative and save the parameter estimates and calculate the LR.

¹¹Why is $\bar{c} = 0.1$? O’Connell quotes Rogoff (1996) who in turn quotes IMF’s *Direction of Trade Statistics* (dec. 1994).

Table 1: TAR vs. AR null, VAD rates

Country	eq.	\hat{c}	$\hat{\lambda}_{out}$	$\hat{\lambda}_{in}$	LR	p[LR]	$\frac{T_{out}}{T}$
Ch	3	0.060	-1.602	-0.195	4.671	0.49	0.14
Fin	3	0.047	-0.046	0.208	1.566	0.98	0.72
Fr	2	0.028	-0.236	-0.078	11.85	0.25	0.74
Ir	2	0.065	-0.302	-0.171	3.467	0.96	0.18
It	2	0.057	-0.112	0.014	17.49	0.15	0.41
No	2	0.039	-0.334	-0.164	6.074	0.76	0.33
Sp	2	0.010	-0.126	3.597	31.94	0.05	0.81
Sw	2	0.064	-0.315	0.169	7.881	0.72	0.46
UK	2	0.021	-0.053	1.923	6.824	0.60	0.85

Notes: The second column shows if the TAR was estimated using Obstfeld & Taylor's estimator (2) or our augmented version (3). The next three columns show the parameter estimates. The last three columns report the LR ratio (2500 bootstrap replications), corresponding p-value and the proportion of observations in the outer regime.

Table 2: TAR vs. AR null, rel.CPI rates

Country	eq.	\hat{c}	$\hat{\lambda}_{out}$	$\hat{\lambda}_{in}$	LR	p[LR]	$\frac{T_{out}}{T}$
Bel	3	0.022	-0.044	-0.162	3.616	0.64	0.59
Ch	3	0.016	-0.040	0.536	6.075	0.27	0.85
Dk	3	0.023	-0.024	-0.185	4.655	0.48	0.73
Fin	3	0.164	-0.334	-0.009	4.374	0.53	0.10
Fr	3	0.004	-0.052	0.958	2.275	0.90	0.90
Ger	3	0.068	-0.226	-0.051	5.859	0.29	0.15
Ir	3	0.052	-0.135	-0.042	3.308	0.75	0.29
It	3	0.138	-0.463	0.000	20.00	<0.01	0.10
Nd	3	0.038	-0.145	-0.062	2.234	0.90	0.23
No	3	0.063	-0.899	-0.011	19.38	<0.01	0.10
Sp	3	0.147	-0.137	-0.020	2.574	0.85	0.17
Sw	3	0.132	-0.463	-0.013	5.200	0.39	0.10
UK	3	0.059	-0.062	-0.079	2.681	0.86	0.61

Notes: See table 1.

2. Generate 500 replications of the above estimated AR null with the same number of observations as the empirical series (i.e. 93 or 231) and estimate under both hypotheses and pick out the 90th and 95th percentile and use these as critical values. Compare the simulated critical values to the actual LR from step one. If the actual LR is greater than the critical values we reject the null for a given significance level.
3. Repeat step 1 and 2 500 times and calculate how often we reject the AR null. If we reject the AR null for a great proportion of the replications test is said to have good power.

The calculations outlined above are quite time consuming and since there are six parameters to control, T , c , λ_{out} , λ_{in} , $(\sigma_{out}, \sigma_{in})$ or (σ, ϕ) , large scale exercises are ruled out. Rather we decided to examine the power for a limited number of 'reasonable' parameter sets. The parameter sets considered are listed in table 3 and 4. The parameters in A1, A2, A4, B1, B2, B3 and B5 are identical to the estimated parameters for France, Ireland and Norway in table 1 and Germany, Norway, Ireland and Spain in table 2. A3 and A5 are identical to Italy and Sweden except for the fact that we have restricted the estimated explosive like behavior for the inner regime to a ordinary random walk. The same goes for B4.

In table 3 and 4 we find that the power is greatly diminished when we raise the threshold and/or decrease difference between the adjustment coefficients for the respective regimes which makes sense. The power of the LR test is not that bad when operating under what can be considered favorable conditions - in the case of A3 (i.e. Italy) the power is 82% when testing on the 10%. However, alarmingly, for many cases the power of the test does not differ much from its size! Taken together it is our opinion that the power of the LR test often is low enough to cast doubt on its usefulness in a real world situation. It may be noted that the Tsay test nowhere performs better than the LR test. This is perhaps not surprising as the Tsay test is a test against general TAR behavior rather than against a specific alternative hypothesis.

5.2 The density of the threshold parameter

In this section we will examine with what precision the thresholds are estimated. Before we present the evidence one should perhaps ask the question why we are interested in this parameter. The threshold is the parameter of most interest since it has an implicit interpretation as the size of the transaction cost. Hence it must be measured with some precision - if it is badly measured or even allocated by chance there is little point in trying to estimate it. In such case we could be better of using some informed prior on the size of transaction cost. To gauge the precision we calculated 90% confidence intervals for the

Table 3: Power of the LR test for VAD rates

DGP→	A1(Fr)	A2(Ir)	A3(it)	A4(No)	A5(Sw)
T	93	93	93	93	93
c	0.028	0.065	0.057	0.039	0.065
λ_{out}	-0.24	-0.30	-0.11	-0.33	-0.32
λ_{in}	-0.08	-0.17	0	-0.16	0
σ_{out}	0.016	0.026	0.033	0.027	0.036
σ_{in}	0.026	0.022	0.018	0.020	0.026
$asy[\frac{T_{out}}{T}]$	0.64	0.17	0.37	0.32	0.34
LR power 10%	0.73	0.13	0.83	0.35	0.48
LR power 5%	0.63	0.07	0.76	0.23	0.37
Tsay power 10%	0.24	0.08	0.35	0.19	0.47
Tsay power 5%	0.16	0.04	0.22	0.09	0.34

Notes: The parameters in table 3 are described in connection to equation (2) and (3) and $asy[\frac{T_{out}}{T}]$ denotes the proportion of observations that asymptotically would be allocated to the outer regime.

Table 4: Power of the LR test for rel. CPI rates

DGP→	B1(Ger)	B2(No)	B3(Ir)	B4(It)	B5(Sp)
T	231	231	231	231	231
c	0.068	0.063	0.052	0.138	0.147
λ_{out}	-0.23	-0.90	-0.14	-0.46	-0.14
λ_{in}	-0.05	-0.01	-0.04	0	-0.02
ϕ	0.32	0.26	0.30	0.37	0.34
σ	0.008	0.009	0.011	0.012	0.011
$asy[\frac{T_{out}}{T}]$	0.04	0.10	0.27	0.12	0.04
LR power 10%	0.10	0.76	0.17	0.64	0.10
LR power 5%	0.06	0.67	0.10	0.53	0.05

Notes: See table 3. We did not calculate the power of the Tsay test since table 3 showed it to be inferior to the LR test.

Table 5: Properties of the threshold estimator

DGP↓	c	\hat{c}	$s_{(\hat{c})}$	LR power 10%	CI 90%
A1(Fr)	0.028	0.028	0.011	0.73	0.012; 0.051
A2(Ir)	0.065	0.040	0.028	0.13	0.005; 0.090
A3(It)	0.057	0.058	0.021	0.83	0.024; 0.088
A4(No)	0.039	0.037	0.019	0.35	0.006; 0.070
A5(Sw)	0.065	0.062	0.027	0.48	0.013; 0.106
B1(Ger)	0.068	0.025	0.018	0.10	0.004; 0.063
B2(No)	0.063	0.054	0.015	0.76	0.014; 0.066
B3(Ir)	0.052	0.040	0.015	0.17	0.007; 0.075
B4(It)	0.138	0.113	0.040	0.64	0.022; 0.146
B5(Sp)	0.147	0.052	0.040	0.10	0.009; 0.144

Notes: c denotes the actual threshold while \hat{c} and $s_{(\hat{c})}$ denotes the mean of estimated threshold and its standard deviation. The next column replicates the power estimates from table 3 and 4. The last column reports the 90% confidence interval for \hat{c} .

threshold estimates¹² together with some descriptive statistics, see table 5. According to table 5 most of the estimators come rather close to the true value, although in some low powered cases (f.x. A2, B1 and B5) the underestimation is too large for comfort. More important however are the confidence intervals according to which the point estimates are very imprecise. Take the high-powered case of A3 (i.e. Italy) where we are told that the true threshold lies between 0.02 and 0.09. Put in other words: the transaction cost is between 2% and 9% of the price. This margin of error is likely to be too large to be of practical use. To give visual support to the descriptive statistics in table 5 we estimated the kernel density functions¹³, see appendix A. As can be seen the density functions are funny looking creatures. Many are highly skewed and/or display two maxima, one global and one local. The last property is too conspicuous and frequent to be discarded as a random occurrence.

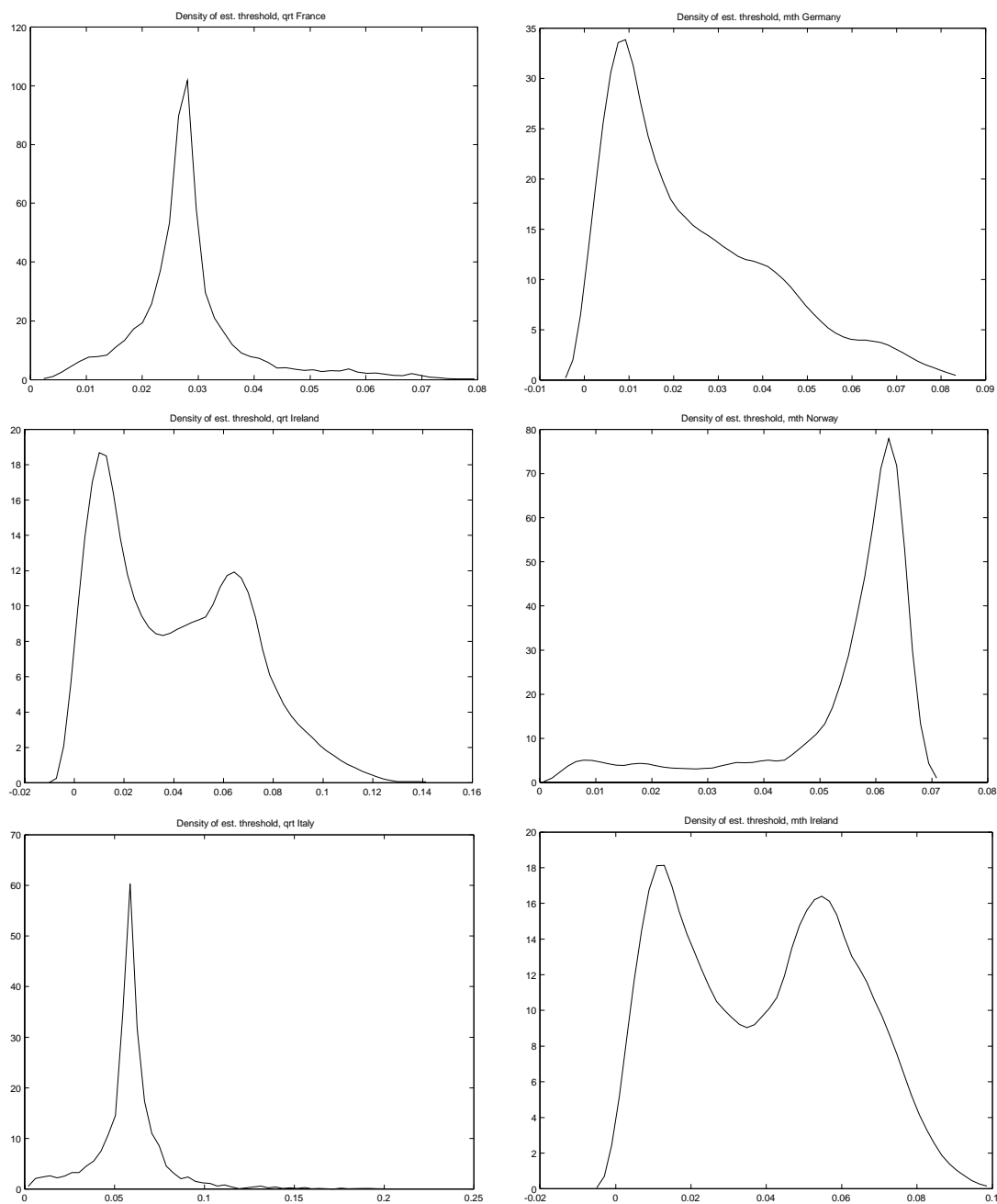
¹²To find the distribution we generated 2500 series from each parameter setting (using the DGPs from table 3 and 4) and estimated the thresholds by means of grid search. The confidence intervals are calculated using the percentile method and are not symmetric around the mean. Moreover it should be noted that it is pointless to pose the question "is the estimate significantly different from zero?", since the maximum likelihood estimator operates under that assumption.

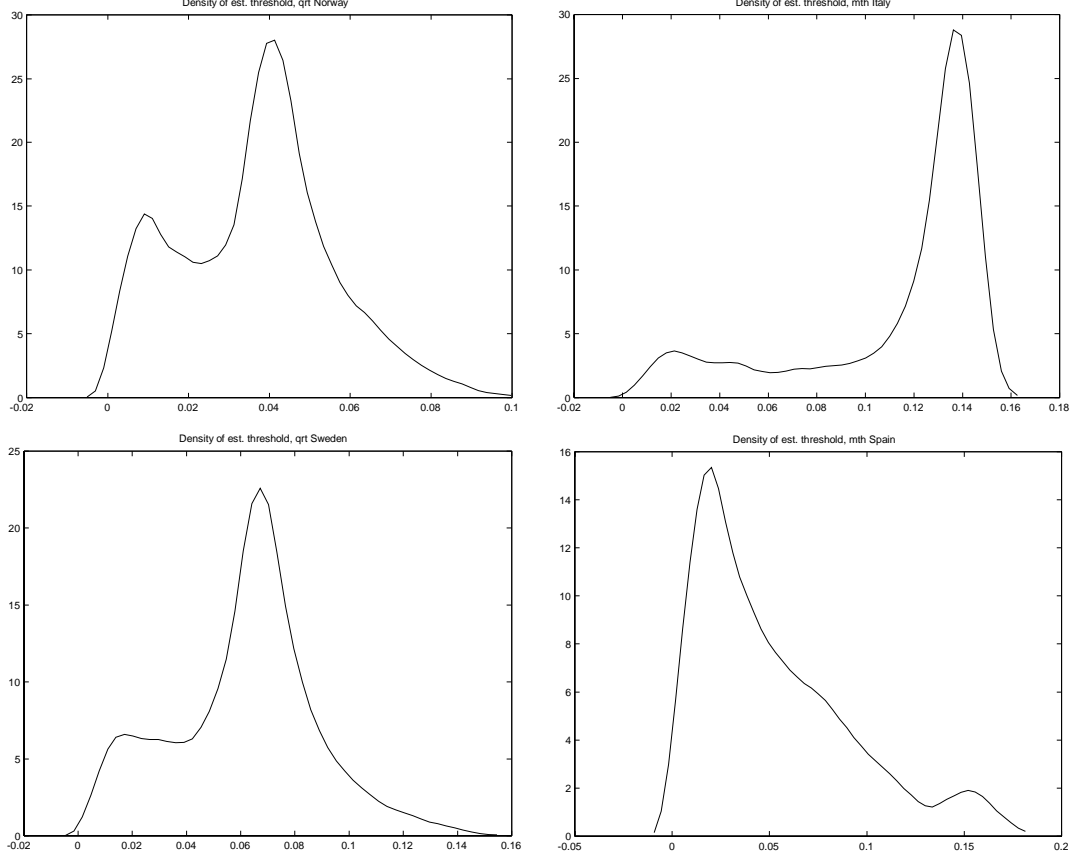
¹³We used the Epanechnikov kernel together with Silverman's (1986) "rule-of-thumb" bandwidth, p. 47. The reader may note that the plots give the impression that the threshold sometimes is less than zero. This is of course not the case but a byproduct of the graphical smoother.

6 Conclusions

Recently a series of papers (Obstfeld & Taylor (1997), O’Connell (1998) and Erjnaes & Persson (2000)) have revived Heckscher’s (1916) notion of arbitrage bands. The underlying idea is that transaction costs make cross-border arbitrage unprofitable for small price differences meaning that small deviations from PPP can be sustained indefinitely. However, once the cross-border price difference exceeds the transaction cost the real exchange rate mean reverts. This implies that the real exchange rate should display stronger mean reversion for large deviations from the mean than for small. This in turn suggests that the real exchange rate may follow a threshold autoregressive (TAR) process. We investigated the power of a (log) likelihood ratio (LR) test used to test the TAR alternative against the AR null. The data generating process was modelled on multilateral real exchange rates culled from *International Financial Statistics*. First we fitted TAR models to the real exchange rates but found that the AR null could seldom be rejected. We then used the estimated parameters to generate pseudo-data and examined the power of the LR test and found that it often was far too low for comfort. We also looked at the empirical distribution of the threshold estimate. We invariably found that the confidence intervals associated with the point estimates were too large to be given a reliable economic interpretation. Thus the lesson to be drawn from this paper is that while it may seem tempting from a theoretical point of view to analyze real exchange rates using TAR models the attempts are likely to founder on the low power of the LR test and a low precision on the part of the threshold estimate.

Appendix A: Kernel density plots





Appendix B: The maximum likelihood estimator

1. Estimate the AR null, $\Delta x_t = \lambda x_{t-1} + e_t$, and calculate the log likelihood function:

$$L_N = -\frac{1}{2} \sum_t (\ln(2\pi) + \ln(\sigma^2) + \frac{e_t^2}{\sigma^2}).$$
2. Pick an upper and lower limit of c , say the 10th and 90th percentile of $|x_t|$. Divide the difference in s equally spaced intervals. This gives s number of potential thresholds, $c_1 \dots c_s$.
3. Estimate the Band-TAR, (2), by imposing c_1 and calculate the log likelihood function: $L_{A1} = -\frac{1}{2} [\sum_{I_{in}(x_{t-1})=1} (\ln(2\pi) + \ln(\sigma_2^2) + \frac{e_{2t}^2}{\sigma_2^2}) + \sum_{I_{out}(x_{t-1})=1} (\ln(2\pi) + \ln(\sigma_1^2) + \frac{e_{1t}^2}{\sigma_1^2})]$, where $I_{in}(x_{t-1}) = I(|x_{t-1}| \leq c)$ and $I_{out}(x_{t-1}) = I(|x_{t-1}| > c)$ are indicator functions.
4. Repeat 3 for $c_2 \dots c_s$.
5. Calculate the log likelihood ratio, $LR_i = 2(L_{Ai} - L_N)$, with $i = 1, \dots, s$, find the maximum LR_i and pick out the associated parameter estimates.

6. Finally, compare the maximum LR_i with a critical value to see if the AR null can be rejected.

Appendix C: Bootstrapping the LR test

Obstfeld & Taylor derived the critical values for their LR tests by means of Monte Carlo simulation. A potential problem with this method is that the simulated shocks are (usually) drawn from a normal distribution. In our case this would be problematic since we always rejected normality, see table 1 and table 2. In an attempt to make a more valid inference we decided to bootstrap the LR statistic rather than rely on Monte Carlo results. The bootstrap is no panacea, but at least we avoid making unrealistic assumption on the error term. Below we describe the bootstrap procedure we used when we calculated the critical values. For simplicity we will use the TAR model without any augmentation terms as an example, but the procedure is in principle identical for the other model.

1. Estimate the AR null, $q_t = aq_{t-1} + e_t$ on actual data $q = [q_0 \ q_1 \dots q_T]'$ and save the residuals $\hat{e} = [\hat{e}_1 \ \hat{e}_2 \dots \hat{e}_T]'$ and the estimated parameter, \hat{a} .
2. Following Johnson & DiNardo (1997, p. 367) rescale the estimated residuals. The rescaled residuals are $v = [v_1 \ v_2 \dots v_T]'$.
3. Recursively generate $\tilde{q}_t = \hat{a}\tilde{q}_{t-1} + v_t$, where $\tilde{q}_0 = q_0$ and v_t is drawn randomly with replacement from v .
4. Estimate the generated series under both the AR null and the TAR alternative and calculate the log likelihood ratio, LR_j .
5. Repeat 3 and 4 n times where n is the number of bootstrap replications.
6. Pick out the 90th and the 95th percentile of $LR = [LR_1 \dots LR_n]$ and use these as critical values.

References

- Balke N & Fomby F [1997] '*Threshold cointegration*', International economic review, **38:3**, 627-645.
- Chen S-L & Wu J-L [2000] '*A re-examination of purchasing power parity in Japan and Taiwan*', Journal of macroeconomics, **22**, 271-284.
- Davutyan N & Pippenger J [1990] '*Testing purchasing power parity: Some evidence of the effects of transaction costs*', Econometric reviews, **9:2**, 211-240.
- Erjraes M & Persson K.G [2000] '*Market integration and transport costs in France*

- 1825-1903: A threshold error correction approach to the law of one price*,
Explorations in economic history, **37**, 149-173.
- Hansen D [1997] '*Testing for linearity*', University of Wisconsin.
- Heckscher E.F [1916] '*Växelkursens grundval vid pappersmyntfot*',
Ekonomisk tidskrift, **18**, 309-312.
- Johnston J & DiNardo J [1997], '*Econometric methods*', 4th ed., McGraw-Hill.
- Michael P, Nobay R & Peel D [1997] '*Transaction costs and nonlinear adjustment in real exchange rates: an empirical investigation* ', Journal of political economy, **105:4**, 862-879.
- Obstfeld M & Taylor A [1997] '*Nonlinear aspects of goods-market arbitrage and adjustment: Heckscher's commodity points revisited*', Journal of the Japanese and international economies, **11**, 441-479.
- O'Connell P.G.J [1998] '*Market frictions and real exchange rates*', Journal of international money and finance , **17**, 71-95.
- Peel D.A & Speight A.E.H [1998] '*The nonlinear time series properties of unemployment rates: some further evidence*', Applied economics, **30**, 287-294.
- Pippenger M.K & Goering G.E [1993] '*A note on the empirical power of unit root tests under threshold processes*', Oxford bulletin of economics and statistics, **55**, 473-481.
- Pippenger M.K & Goering G.E [1998] '*Exchange rate forecasting: Results from a threshold autoregressive model*', Open economies review, **9**, 157-170.
- Rogoff K [1996] '*The purchasing power parity puzzle*', Journal of economic literature, **34**, 647-668
- Silverman B.W [1986] '*Density estimation for statistics and data analysis*', Chapman & Hall.
- Taylor M & Peel D [2000] '*Nonlinear adjustment, long-run equilibrium and exchange rate fundamentals*', Journal of international money and finance, **19**, 33-53.
- Tong H [1990] '*Non-linear time series - A dynamical approach*', Clarendon Press Oxford
- Tsay R.S [1989] '*Testing and modeling threshold autoregressive processes*', Journal of the American statistical association, **84**, 231-240.