Smooth transition regression

Nicolas Ronderos

February 1, 2015

Abstract

This document explains how to perform testing, estimation and evaluation of smooth transition regression models (STR) using an EViews add-in based on [2]. The testing procedures includes the specific alternatives: LSTR, ESTR, LSTR-D and ESTR-D, transition variables significance and adequacy between STR structures. For the estimation of the STR models the equation or Logl objects are used. The evaluation consists on the usefulness of the transition function in the model, their dynamic properties and diagnostic tests of residuals.

1 STR models

A smooth transition autoregressive regression (STAR) model with $m = p + k + 1$ independent variables, where p are lags of the dependent variable and k are exogenous variables can be written as:

$$y_t = \beta' x_t + (\theta' x_t)F(z_t) + u_t \tag{1.1}$$

Where $u_t \sim iid(0, \sigma^2)$, $\beta = (\beta_0, \beta_1, ..., \beta_m)'$, $\theta = (\theta_0, \theta_1, ..., \theta_m)'$, $x_t = (1, y_{t-1}, ..., y_{t-p}, x_{1t}, ..., x_{kt})'$ and z_t is the transition variable. The function $F(z_t)$ is continuous and might be even or odd.

The odd function $F(z_t) = (1 + exp\{-\gamma(z_t - c)\})^{-1}$ yields the logistic STR model (LSTR), and the even function $F(z_t) = (1 - exp\{-\gamma(z_t - c)^2\})^{-1}$ defines the exponential STR model (ESTR). If one chooses the transition variable to be the residuals of the linear model $y_t = \beta' x_t + \nu_t$ then the STR deviation models are obtained, with the following transition functions $F(z_t) = (1 + exp\{-\gamma(a' \widetilde{\nu}_{t-1})\})^{-1} - \frac{1}{2}$ for the LSTR-D and $F(z_t) = (1 - exp\{-\gamma(a' \widetilde{\nu}^2_{t-1})\})^{-1} - \frac{1}{2}$ for the ESTR-D, where $\sum a_j = 1$, $\widetilde{\nu}_t = (\nu_t, ..., \nu_{t-h+1})$ and $\widetilde{\nu}^2_t = (\nu^2_t, ..., \nu^2_{t-h+1})$.

2 Testing

This section is divided into three subsections, in these will be explained linearity testing against an specific alternative, transition variable significance and testing between STR structures.

2.1 Linearity testing against an specific alternative

These test are of the LM type and can be calculated with the following steps.

1. Regress y_t on x_t by OLS and compute the residuals $\hat{\nu}_t = y_t - \hat{\pi}' x_t$ and their SSR_0.

2. Compute the auxiliary regression by OLS and the SSR. These equations are showed in table 1 for each test, where $\widetilde{x}_t = (y_{t-1}, ..., y_{t-p}, x_{1t}, ..., x_{kt})$ and $\widetilde{x}_{td}$ is some preselected variable or a linear combination of variables to be the transition variable. The value of h set the lags to be tested in the STD-D family, nevertheless a specific lag for the residuals $\hat{\nu}_{t-d}$ could be chosen. It is also showed their null, alternative and the distribution of the statistic of step 3.

3. Calculate the statistic $LM = \frac{T(SSR_0 - SSR)}{SSR_0}$.

H_a	*Auxiliary equation*	H_0	*Statistic distribution*
LSTR	$\hat{\nu_t} = \beta^{'} x_t + \sum\sum \varphi_{ij}\tilde{x}_{ti}\tilde{x}_{tj} + \sum\sum\sum \varphi_{ij}\tilde{x}_{ti}\tilde{x}_{tj}\tilde{x}_{tk} + \ldots + e_t$	Quadratic: $\varphi_{ij} = 0, i = 1, \ldots, m, j = i, \ldots, m$	Quadratic: $\chi^2(\frac{m(m+1)}{2})$
LSTR	$\hat{\nu_t} = (\psi_1 + \psi_2\varphi)^{'} x_t + \sum_{i\geq j}^{m} \varphi_{ij}\tilde{x}_{ti}\tilde{x}_{tj} + e_t$	$\varphi_{ij} = 0, i = 1, \ldots, m, j = i, \ldots, m$	$\chi^2(\frac{m(m+1)}{2})$
LSTR	$\hat{\nu_t} = \beta^{'} x_t + \sum_{i\geq j}^{m} \varphi_{ij}\tilde{x}_{ti}\tilde{x}_{tj} + \sum_{j=1}^{m} \psi_j \tilde{x}_{tj}^3 + e_t$	$\varphi_{ij} = 0, i = 1, \ldots, m, j = i, \ldots, m; \psi_j = 0, j = 1, \ldots, m$	$\chi^2(\frac{m(m+1)}{2} + m)$
LSTR	$\hat{\nu_t} = \beta^{'} x_t + \sum_{j=1}^{m} \varphi_{dj}\tilde{x}_{td}\tilde{x}_{tj} + e_t$	$\varphi_{dj} = 0, j = 1, \ldots, m$	$\chi^2(m)$
LSTR	$\hat{\nu_t} = \beta^{'} x_t + \sum_{j=1}^{m}(\varphi_{dj}\tilde{x}_{td}\tilde{x}_{tj} + \psi_{dj}\tilde{x}_{td}^2\tilde{x}_{tj} + \kappa_{dj}\tilde{x}_{td}^3\tilde{x}_{tj}) + e_t$	$\varphi_{dj} = \psi_{dj} = \kappa_{dj} = 0, j = 1, \ldots, m$	$\chi^2(3m)$
ESTR	$\hat{\nu_t} = \beta^{'} x_t + \sum_{i=1}^{m}\sum_{j=1}^{m} \psi_{ij}\tilde{x}_{ti}\tilde{x}_{tj}^2 + \sum_{i\geq j}^{m} \varphi_{ij}\tilde{x}_{ti}\tilde{x}_{tj} + e_t$	$\psi_{ij} = 0, i, j = 1, \ldots, m; \varphi_{ij} = 0, j = 1, \ldots, m$	$\chi^2(m^2 + \frac{m(m+1)}{2})$
ESTR	$\hat{\nu_t} = \beta^{'} x_t + \sum_{i=1}^{m}(\psi_{id}\tilde{x}_{ti}\tilde{x}_{td}^2 + \varphi_{id}\tilde{x}_{ti}\tilde{x}_{td}) + e_t$	$\psi_{id} = \varphi_{id} = 0, i = 1, \ldots, m$	$\chi^2(2m)$
LSTR-D	$\hat{\nu_t} = \beta^{'} x_t + \sum_{i=1}^{m}\sum_{j=1}^{h} \varphi_{ij}\tilde{x}_{ti}\hat{\nu}_{t-j} + e_t$	$\varphi_{ij} = 0, i = 1, \ldots, m, j = 1, \ldots, h$	$\chi^2(mh)$
LSTR-D	$\hat{\nu_t} = \beta^{'} x_t + \sum_{i=1}^{m} \varphi_{id}\tilde{x}_{ti}\hat{\nu}_{t-d} + e_t$	$\varphi_{id} = 0, i = 1, \ldots, m$	$\chi^2(m)$
ESTR-D	$\hat{\nu_t} = \beta^{'} x_t + \sum_{i=1}^{m}\sum_{j=1}^{h} \varphi_{ij}\tilde{x}_{ti}\hat{\nu}_{t-j}^2 + e_t$	$\varphi_{ij} = 0, i = 1, \ldots, m, j = 1, \ldots, h$	$\chi^2(mh)$
ESTR-D	$\hat{\nu_t} = \beta^{'} x_t + \sum_{i=1}^{m} \varphi_{id}\tilde{x}_{ti}\hat{\nu}_{t-d}^2 + e_t$	$\varphi_{id} = 0, i = 1, \ldots, m$	$\chi^2(m)$

Table 1: Auxiliary equations

2.2 Transition variable significance

To test if a variable is statistically significant enough to be the transition variable the next equation must be calculated:

$$\hat{\nu_t} = \beta_0^{'} x_t + \beta_1^{'} x_t x_{td} + \beta_2^{'} x_t x_{td}^2 + \beta_3^{'} x_t x_{td}^3 + \eta_t \tag{2.1}$$

Where $\hat{\nu_t}$ are the residuals of the linear model. The null of the test is $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$ and if rejected for several choices of x_{td} one must select the one with the smallest p-value or the biggest statistical. As the transition variable could be the residuals of the linear model this step decides between the STR or the STR-D family. If this procedure do not find any candidate to be the transition variable then linearity will not be rejected, nevertheless if linearty is rejected when it is true the estimation of the nonlinear model may either fail or may not give sensible results.

2.3 Testing between STR structures

This test decides between LSTR or ESTR structure, it is based on (2) and in the sequence of nested test showed below.

$$H_0 : \beta_3 = 0 \tag{2.2}$$

$$H_0 : \beta_2 = 0 \mid \beta_3 = 0 \tag{2.3}$$

$$H_0 : \beta_1 = 0 \mid \beta_3 = \beta_2 = 0 \tag{2.4}$$

The selection rule stated that if (4) has the smallest p-value select the ESTR models, otherwise choose the LSTR family. It is important to mention that the test is subject to the transition variable chosen.

An ESTR model can be approximated by a LSTR if the data of the transition variable usually lies above the value of the transition parameter c. Also a LSTR model can be approximated by a ESTR if its transition function does not grow too rapidly around c (γ small), therefore no ESTR model offers a reasonable approximation if the transition from one regime to the other is quick, like a TAR model.

3 Estimation

The estimation could be performed with non-linear least squares or maximum likelihood using a optimization procedure, both options are available in the add-in. It is assumed normality for the log likelihood function, therefore the function is:

$$ln(l) = -\frac{1}{2}ln(2\pi) - ln(\sigma) - \frac{1}{2}(\frac{u_t}{\sigma})^2 \tag{3.1}$$

Where u_t is replaced by the definition given in (1). When the user choose one transition variable the function $F(z_t)$ is normalized in the exponential function by $\hat{\sigma}(z_t)$ in the LSTR and LSTR-D case and by $\hat{\sigma^2}(z_t)$ in the ESTR and ESTR-D case, this is done for rescaling the slope parameter of the transition function γ, given that its value could be relatively higher than the other parameters. The rescaling allows to choose the initial value for the parameter in the optimization procedure more easily, a usual starting value is one.

It is important to mention that in the add-in the estimation is performed for the full model i.e. without imposing suitable exclusion restrictions of the form $\beta_i = 0, \theta_j = 0$ and $\beta_i = -\theta_j$ which are recommended for model parsimony.

4 Evaluation

The first check is to ensure that the parameter estimates seems reasonable, and in the transition function this means evaluate the c parameter, if this is well outside the range of the transition variable it indicates misspecification.

Then the residuals have to be checked for autocorrelation, ARCH effect and remained unexplained nonlinearities[1] if the tests evidence the existence of any of this it may indicate misspecification and the model should be reformulated.

Last if the nonlinear model performs significantly worse than the linear one, the specification should be reconsidered. The SSR of the two models can be used for this, if the SSR of the nonlinear model is bigger than the linear is an indication of misspecification.

5 The add-in

The add-in could be used via point and clicking from a series object (the dependent variable) or via command line, the command are showed in table 2.

series_name.star(options)

Option	Command
Linear and changing variables	*Variables=list of regressors*
Transition variables	*Transition=list of transition variables*
Expansion order	*exp=integer between 1 and 6 (default 3)*
Lags for STR-D models	*lags=integer (default 2)*
Specific lag for STR-D models	*s_lag=integer list (default 2)*
Estimation object	*logl (default equation)*
LSTR estimation	*lstr*
ESTR estimation	*estr*
LSTR-D estimation	*lstr_d*
ESTR-D estimation	*estr_d*
Starting values vector	*sv=vector object*
Evaluation of the transition function	*evaluation*

Table 2: Command line

The variables list is mandatory for testing and estimation, this is equivalent to the $\widetilde{x}_t$ vector[2] and its components are used for some tests (the ones that no require a transition variable) of section 2.1 they are also used for tests of sections 2.2 and 2.3 where every component in the list are assumed to be the transition variable x_{td}. The transition list could be a single variable (which is recommended) or a list of variables, where their weights are assumed to be equal for estimation. This list is optional for testing, given that tests of section 2.1 that use a transition variable are not calculated and tests of

[1]This can be checked using the BDS test or the tests of section 2 applied to the residuals of the estimated STR model.
[2]The constant is always included in testing and estimation.

sections 2.2 and 2.3 use the components of $\widetilde{x}_t$ as the transition variables. And is mandatory for estimation of the LSTR and ESTR models.

The expansion order integer is used for the first auxiliary equation test in table 1. The lags integer are used for the residual lags h and for use this lags to be x_{td} in the tests showed in section 2.2 and 2.3. The specific lag could be an integer or a list of integers, which are used for the tests of section 2.1 being $\hat{\nu}_{t-d}$ and for the estimation procedure of the STD-D models.

The option logl is used to select the logl object as the estimation tool if it is not included then the equation object is used. If the options LSTR, ESTR, LSTR-D and ESTR-D are included the respective model are estimated, for this an already created vector object of starting values must be included and this must be at least of the size of the total parameters of the model. Last, the evaluation option is used to obtain the evaluation of the c parameter for every estimated model.

6 Examples

Two examples will be showed using the add-in, one on simulated data an the other on the percentage change of inflation. The first was simulated as the following LSTR model:

$$z_t = 1 + 2x_t + x_{t-1} + 0.5x_{t-2} - (2x_t + x_{t-1} + 0.5x_{t-2})/(1 + e^{-4(x_{t-1}+1.7)}) + e_t \tag{6.1}$$

where x_t and e_t are $nid(0,1)$. As can be seen when $x_{t-1} \gg -1.7$ the series behaves like a linear model, otherwise the series is described as a white noise with nonzero mean. In figure 6.1 the series is plotted for 10000 observations.

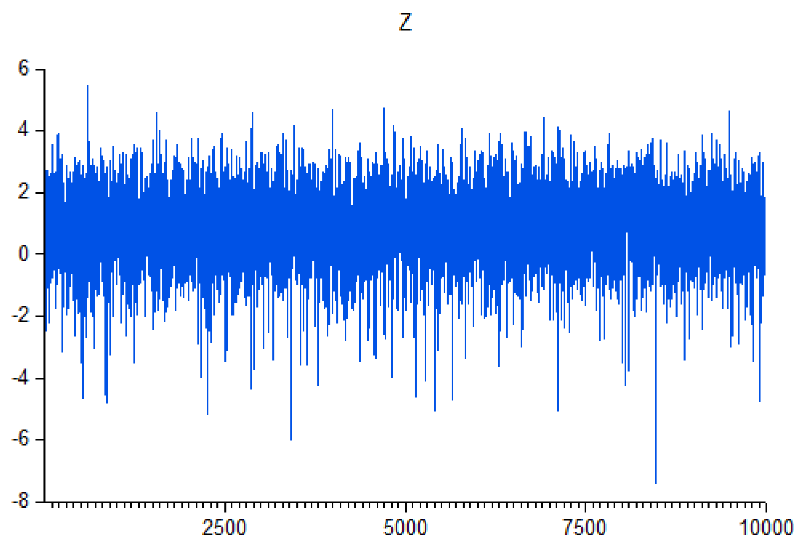

Figure 6.1: z_t

To obtain the STR tests, estimation and evaluation the following commands were used, where the linear model was chosen taking into account (7). All the starting values are assumed to be 1.

vector(11) sv=1

z.star(variables=x x(-1) x(-2),transition=x(-1),lstr,evaluation,sv=sv)

The output is an spool called *str_tests* that contain the tests of sections 2.1, 2.2 and 2.3 with his inputs left in their default values, an equation and a series object with the estimated transition function. The tests are showed in table 3, as can be seen the null of linearity is rejected against LSTR, some of the other tests also reject linearty, given the considerable power of this against LSTR. The tests of transition variable selection all reject the null, with the biggest statistic for x_{t-1} the true transition variable. Last, when the transition variable is x_{t-1} the suggested structure is LSTR.

Tests on Z. H0:Linearty.
Transition variable: x(-1)
Ha:LSTR

	LM statistic	P-value
Unknown transition variable: Expansion order 3	1359.577	0.000
Unknown transition variable	790.439	0.000
Unknown transition variable: Cubic expantion	1007.601	0.000
Transition variable	787.356	0.000
Transition variable: Cubic expantion	1452.982	0.000

Transition variable: x(-1)
Ha:ESTR

	LM statistic	P-value
Unknown transition variable	1358.562	0.000
Transition variable: Cubic expantion	1346.689	0.000

Transition variable: Deviation from a linear path
Ha:LSTR-D or bilinearty

	LM statistic	P-value
Unknown residual lag: 2	7.907	0.245
Specific residual lag: 2	5.907	0.116

Transition variable: Deviation from a linear path
Ha:ESTR-D or bilinearty

	LM statistic	P-value
Unknown residual lag: 2	68.113	0.000
Specific residual lag: 2	6.500	0.090

Transition variable tests on: x(-1) x x(-2) and 2 residual lags
H0: The transition variable is not significant (STR or STR-D)

Transition variable (STR)	F-statistic	P-value	Residual lag (STR-D)	F-statistic	P-value
x(-1)	1697.864	0.000	1	79.447	0.000
x	305.077	0.000	2	14.895	0.000
x(-2)	57.757	0.000			

Structure tests on: x(-1) x x(-2) and 2 residual lags
Choice: LSTR or ESTR at 5% of significance.

Transition variable (STR)	Structure	Residual lag (STR-D)	Structure
x(-1)	LSTR	1	ESTR
x	ESTR	2	ESTR
x(-2)	ESTR		

Table 3: Tests

The estimated model is showed in table 4, as can be seen the estimated coefficients are close to the true values, is important to note that c_5 is not statistically different from zero. The add-in does not handle the kind of restrictions mentioned in section 3, therefore the estimated model fails the parsimony criteria given that $c_2 = -c_6$, $c_3 = -c_7$ and $c_4 = -c_5$ whereby the model could be estimated with least three parameters.

Dependent Variable: DEP
Method: Least Squares
Date: 01/29/15 Time: 14:41
Sample (adjusted): 4 10000
Included observations: 9997 after adjustments
Estimation settings: tol= 0.01000
Initial Values: C(1)=1.00000, C(2)=1.00000, C(3)=1.00000, C(4)=1.00000, C(5)=1.00000, C(6)=1.00000, C(7)=1.00000, C(8)=1.00000, C(9)=1.00000, C(10)=1.00000
Convergence achieved after 19 iterations
DEP=C(1)+C(2)*X+C(3)*X(-1)+C(4)*X(-2)+(C(5)+C(6)*X+C(7)*X(-1)+C(8)*X(-2))/(1+@EXP(-C(9)*(X(-1)-C(10))/1.11740413109537))

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	1.145643	0.391239	2.928246	0.0034
C(2)	1.974189	0.113012	17.46879	0.0000
C(3)	1.073681	0.156447	6.862928	0.0000
C(4)	0.453037	0.059793	7.576761	0.0000
C(5)	-0.136823	0.396389	-0.345174	0.7300
C(6)	-1.974996	0.115146	-17.15211	0.0000
C(7)	-1.093584	0.151293	-7.228260	0.0000
C(8)	-0.461285	0.062443	-7.387317	0.0000
C(9)	4.266833	0.352662	12.09894	0.0000
C(10)	-1.696779	0.042413	-40.00629	0.0000

R-squared	0.194014	Mean dependent var	0.901908
Adjusted R-squared	0.193288	S.D. dependent var	1.117404
S.E. of regression	1.003620	Akaike info criterion	2.846105
Sum squared resid	10059.45	Schwarz criterion	2.853317
Log likelihood	-14216.25	Hannan-Quinn criter.	2.848546
F-statistic	267.1156	Durbin-Watson stat	1.991387
Prob(F-statistic)	0.000000		

Table 4: LSTR Model

The evaluation consist on the utility of the transition function i.e. if this activates or not, in figure is showed the periods where the transition functions activates and their frequency, as can be seen the series mostly behaves like a white noise porcess. The BDS, the ARCH test and the Jarque-Bera test were applied to the residuals, and they conclude that this are *iid* with no ARCH effect and *nid*.

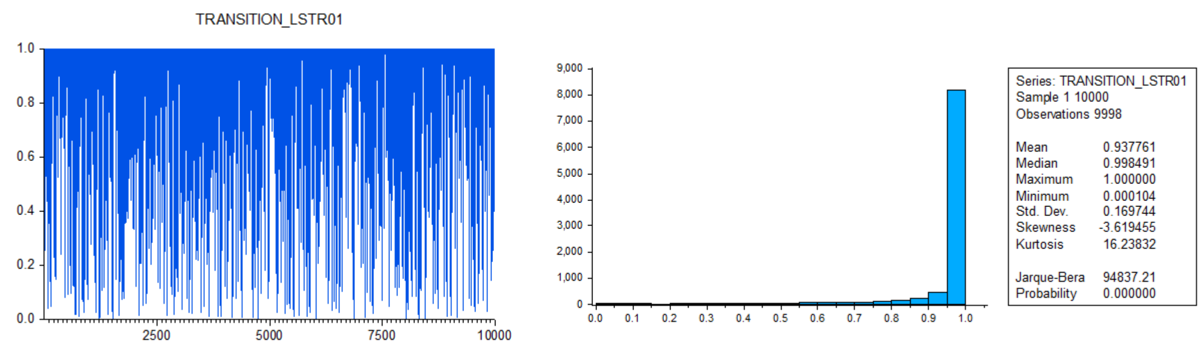

Figure 6.2: Transition function

For the second example an STR model was fitted for USA inflation (m/m)[3]. The following commands were used to obtain the tests, the estimation and evaluation, the linear model was chosen using the autocorrelation function.

vector(21) sv=1

p.star(variables=p(-1) p(-2) p(-3) p(-4) p(-5) p(-6) p(-7) p(-12),transition=p(-1),lags=12,s_lag=12,lstr,sv=sv,evaluation)

The tests suggest an LSTR model with the first lag of inflation as the transition variable. In table 5 is showed the estimated model.

Dependent Variable: P
Method: Least Squares
Date: 02/01/15 Time: 17:27
Sample (adjusted): 1948M02 2014M12
Included observations: 803 after adjustments
Estimation settings: tol= 0.01000, derivs=accurate numeric
Initial Values: C(1)=1.00000, C(2)=1.00000, C(3)=1.00000, C(4)=1.00000, C(5)=1.00000, C(6)=1.00000, C(7)=1.00000, C(8)=1.00000, C(9)=1.00000, C(10)=1.00000, C(11)=1.00000, C(12)=1.00000, C(13)=1.00000, C(14)=1.00000, C(15)=1.00000, C(16)=1.00000, C(17)=1.00000, C(18)=1.00000, C(19)=1.00000, C(20)=1.00000
Convergence achieved after 43 iterations
P=C(1)+C(2)*P(-1)+C(3)*P(-2)+C(4)*P(-3)+C(5)*P(-4)+C(6)*P(-5)+C(7)*P(-6)+C(8)*P(-7)+C(9)*P(-12)+(C(10)+C(11)*P(-1)+C(12)*P(-2)+C(13)*P(-3)+C(14)*P(-4)+C(15)*P(-5)+C(16)*P(-6)+C(17)*P(-7)+C(18)*P(-12))/(1+@EXP(-C(19)*(P(-1)-C(20))/0.349768978295261))

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	0.206203	2.271453	0.090780	0.9277
C(2)	1.667761	5.033838	0.331310	0.7405
C(3)	-8.508802	40.08428	-0.212273	0.8319
C(4)	5.521311	25.07838	0.220162	0.8258
C(5)	-3.209005	14.88447	-0.215594	0.8294
C(6)	3.408657	15.63067	0.218075	0.8274
C(7)	-9.120590	42.19294	-0.216164	0.8289
C(8)	2.623587	11.85673	0.221274	0.8249
C(9)	-3.062350	14.93822	-0.205001	0.8376
C(10)	-0.125546	2.332879	-0.053816	0.9571
C(11)	-1.323244	5.063612	-0.261324	0.7939
C(12)	8.695896	40.12453	0.216723	0.8285
C(13)	-5.552982	25.10226	-0.221214	0.8250
C(14)	3.303032	14.89779	0.221713	0.8246
C(15)	-3.369821	15.66084	-0.215175	0.8297
C(16)	9.281682	42.23285	0.219774	0.8261
C(17)	-2.546786	11.87050	-0.214547	0.8302
C(18)	3.004962	14.96162	0.200845	0.8409
C(19)	0.956209	0.505607	1.891210	0.0590
C(20)	-1.392776	2.471484	-0.563538	0.5732

R-squared	0.471634	Mean dependent var	0.287382
Adjusted R-squared	0.458813	S.D. dependent var	0.337413
S.E. of regression	0.248219	Akaike info criterion	0.075584
Sum squared resid	48.24287	Schwarz criterion	0.192355
Log likelihood	-10.34679	Hannan-Quinn criter.	0.120434
F-statistic	36.78560	Durbin-Watson stat	2.044373
Prob(F-statistic)	0.000000		

Table 5: LSTR inflation

[3] The ADF test were applied to the series using the algorithm proposed by [1], see http://forums.eviews.com/viewtopic.php?f=15&t=11133 for the program code.

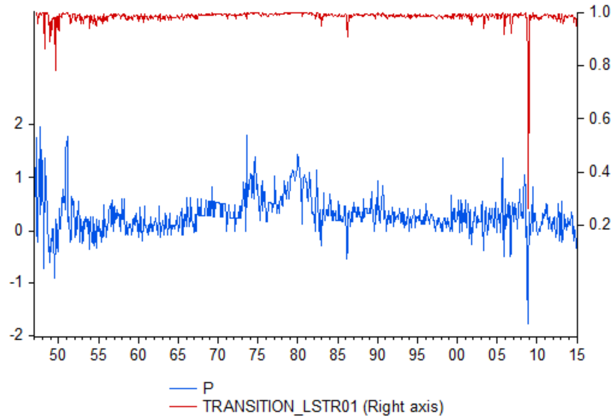

Figure 6.3: Transition function

The estimation states smooth transition regimes given the low value of γ (0.95), if $p_{t-1} \gg -1.39$ then the process is described by the coefficients showed above with $F(p_{t-1}) = 1$ and $F(p_{t-1}) = 0$ otherwise. This dynamics are plotted in figure 6.3 together with p_t, it can be seen that the process is placed in its lower regime when the deflation is high, the prominent case occurs in December 2008 (where $p_{t-1} = -1.8\%$), another intermediate cases also occur. It is important to mention that the SSR of the linear model is larger than the SSR of the LSTR.

When the process is at its upper regime their dynamics are ruled close by an AR(1) model $p_t = 0.08 + 0.34p_{t-1} + e_t$, therefore is a stable process[4]. For example when $F(p_{t-1}) = 0.25$, like in December of 2008 their dynamics are ruled by a explosive process, therefore the economy prices moves from high deflation to inflation very aggressively, at least until -1.39, then growth stabilizes. The roots for this different regimes can be calculated, nevertheless the know whether it is stable or unstable is enough.

References

[1] Walter Enders. *Applied econometric time series.* John Wiley & Sons, 2008.

[2] Clive WJ Granger and Timo Terasvirta. *Modelling non-linear economic relationships.* Oxford University Press, 1993.

[4]As in the previous example, the parsimony criteria could be reached by imposing some restrictions.