Lecture 3

Non-linear models

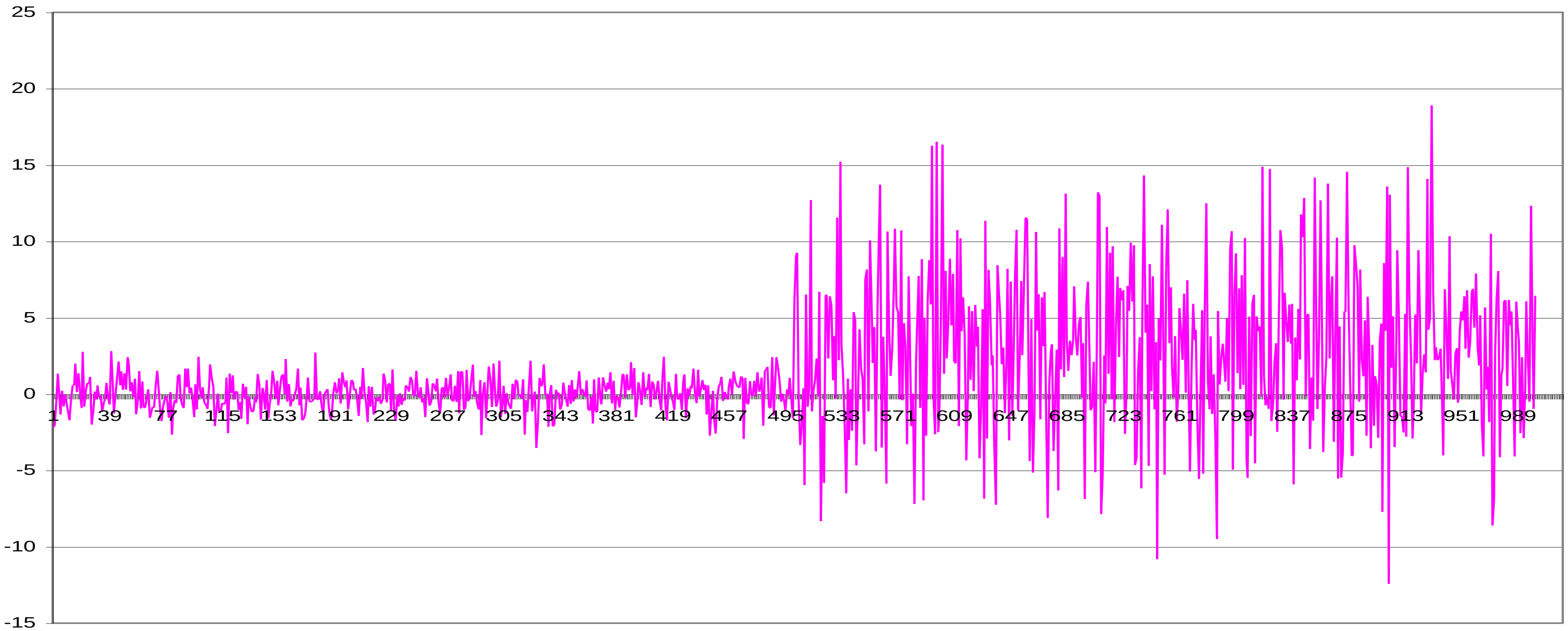
Contents

Non-linear Models

- Motivation: Episodic nature of economic and financial variables. What might cause these fundamental changes in behaviour?
    - Wars
    - Financial panics
    - Significant changes in government policy
    - Changes in market microstructure
    - Changes in market sentiment
    - Market rigidities

- Switches can be one-off single changes or occur frequently back and forth.

Switching Behaviour: A Simple Example for One-off Changes



<u>Dealing with switching variables</u>

We could generalise ARMA models (again) to allow the series, y_t to be drawn from two or more different generating processes at different times. e.g.

$y_t = \mu_1 + \phi_1 y_{t-1} + u_{1t}$ before observation 500 and

$y_t = \mu_2 + \phi_2 y_{t-1} + u_{2t}$ after observation 500

How do we Decide where the Switch or Switches take Place?

- It may be obvious from a plot or from knowledge of the history of the series.
- It can be determined using a model.
- It may occur at fixed intervals as a result of seasonalities.
- A number of different approaches are available, and are described below.

Seasonality in Financial Markets

- If we have quarterly or monthly or even daily data, these may have patterns in.
- Seasonal effects in financial markets have been widely observed and are often termed “calendar anomalies”.
- Examples include day-of-the-week effects, open- or close-of-market effect, January effects, or bank holiday effects.
- These result in statistically significantly different behaviour during some seasons compared with others.
- Their existence is not necessarily inconsistent with the EMH.

Constructing Dummy Variables for Seasonality

- One way to cope with this is the inclusion of dummy variables- e.g. for quarterly data, we could have 4 dummy variables:

$D1_t$ = 1 in Q1 and zero otherwise
$D2_t$ = 1 in Q2 and zero otherwise
$D3_t$ = 1 in Q3 and zero otherwise
$D4_t$ = 1 in Q4 and zero otherwise

- How many dummy variables do we need? We need one less than the "seasonality" of the data. e.g. for quarterly series, consider what happens if we use all 4 dummies

Constructing Quarterly Dummy Variables

	$D1_t$	$D2_t$	$D3_t$	$D4_t$	Sum_t
1986Q1	1	0	0	0	1
Q2	0	1	0	0	1
Q3	0	0	1	0	1
Q4	0	0	0	1	1
1987Q1	1	0	0	0	1
Q2	0	1	0	0	1
Q3	0	0	1	0	1
		etc.			

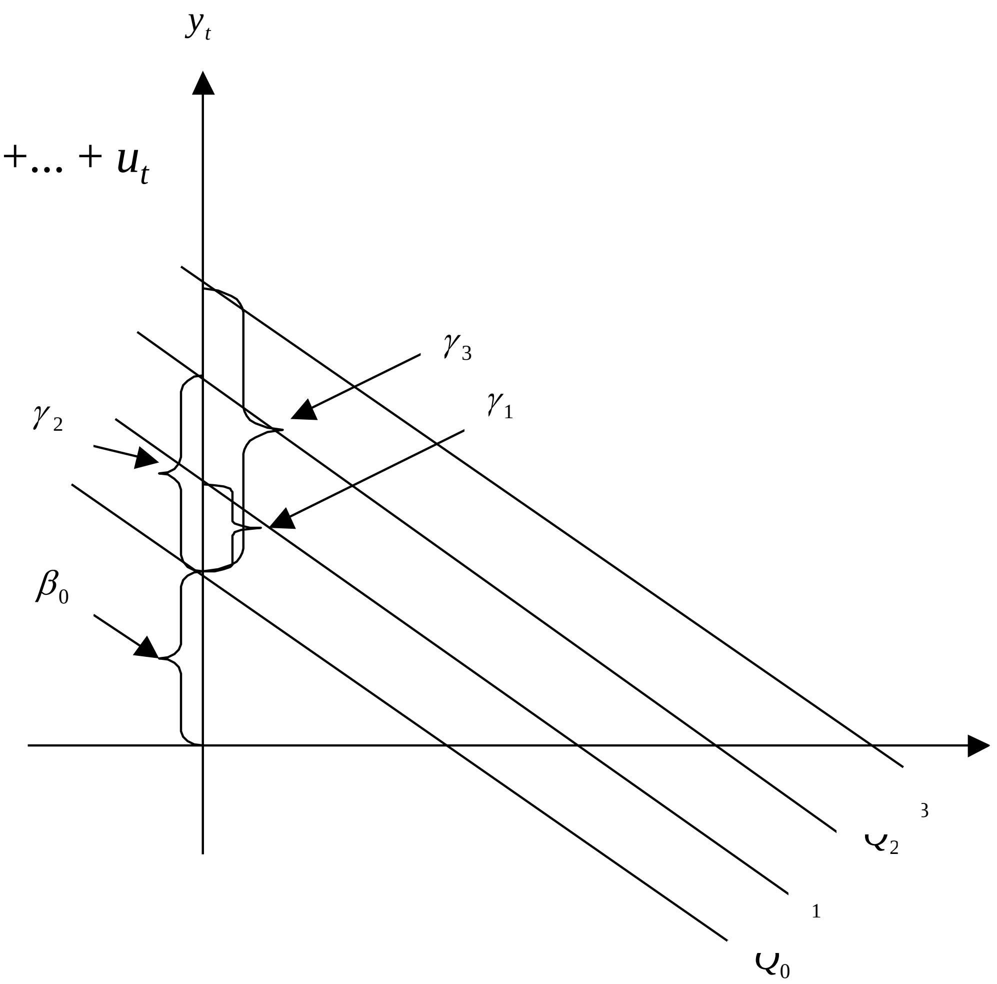
- Problem of multicollinearity, so $(X'X)^{-1}$ does not exist.
- Solution is to just use 3 dummy variables plus the constant or 4 dummies and no constant.

How Does the Dummy Variable Work?

- It works by changing the intercept.

  Consider the following regression:

  $y_t = \beta_0 + \gamma_1 D1_t + \gamma_2 D2_t + \gamma_3 D3_t + \beta_2 x_{2t} + ... + u_t$



So we have as the constant

$\hat{\beta}_1 + \hat{\gamma}_1$ in the first quarter

$\hat{\beta}_1 + \hat{\gamma}_2$ in the second quarter

$\hat{\beta}_1 + \hat{\gamma}_3$ in the third quarter

$\hat{\beta}_1$ in the fourth quarter

Seasonalities in South East Asian Stock Returns

- Brooks and Persand (2001) examine the evidence for a day-of-the-week effect in five Southeast Asian stock markets: South Korea, Malaysia, the Philippines, Taiwan and Thailand.
- The data, are on a daily close-to-close basis for all weekdays (Mondays to Fridays) falling in the period 31 December 1989 to 19 January 1996 (a total of 1581 observations).
- They use daily dummy variables for the day of the week effects in the regression:

$$r_t = \gamma_1 D1_t + \gamma_2 D2_t + \gamma_3 D3_t + \gamma_4 D4_t + \gamma_5 D5_t + u_t$$

- Then the coefficients can be interpreted as the average return on each day of the week.

Values and Significances of Day of the Week Effects in South East Asian Stock Markets

	South Korea	Thailand	Malaysia	Taiwan	Philippines
Monday	0.49E-3	0.00322	0.00185	0.56E-3	0.00119
	(0.6740)	(3.9804)**	(2.9304)**	(0.4321)	(1.4369)
Tuesday	-0.45E-3	-0.00179	-0.00175	0.00104	-0.97E-4
	(-0.3692)	(-1.6834)	(-2.1258)**	(0.5955)	(-0.0916)
Wednesday	-0.37E-3	-0.00160	0.31E-3	-0.00264	-0.49E-3
	-0.5005)	(-1.5912)	(0.4786)	(-2.107)**	(-0.5637)
Thursday	0.40E-3	0.00100	0.00159	-0.00159	0.92E-3
	(0.5468)	(1.0379)	(2.2886)**	(-1.2724)	(0.8908)
Friday	-0.31E-3	0.52E-3	0.40E-4	0.43E-3	0.00151
	(-0.3998)	(0.5036)	(0.0536)	(0.3123)	(1.7123)

Notes: Coefficients are given in each cell followed by *t*-ratios in parentheses; * and ** denote significance at the 5% and 1% levels respectively. Source: Brooks and Persand (2001).

Slope Dummy Variables

- As well as or instead of intercept dummies, we could also use slope dummies:
- For example, this diagram depicts the use of one dummy – e.g., open and close data.
- In the latter case, we could

define D_t = 1 for open observations

and D_t=0 for close.

Other example: Responsiveness of *y* to *x*

increases during the financial crisis,

where *y* is bank interest rate and *x*

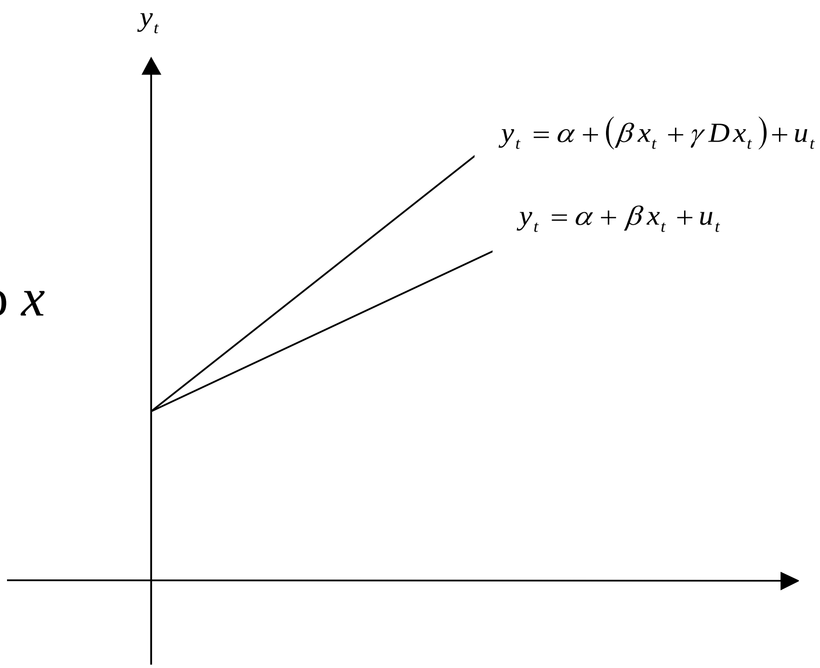
is market rate.

- Such dummies change the slope

but leave the intercept unchanged.

- We could use more slope dummies

or both intercept and slope dummies.



Seasonalities in South East Asian Stock Returns Revisited

- It is possible that the different returns on different days of the week could be a result of different levels of risk on different days.

- To allow for this, Brooks and Persand re-estimate the model allowing for different betas on different days of the week using slope dummies:

$$r_t = (\sum_{i=1}^{5} \alpha_i D_{it} + \beta_i D_{it} RWM_t) + u_t$$

- where D_{it} is the i^{th} dummy variable taking the value 1 for day $t=i$ and zero otherwise, and RWM_t is the return on the world market index

- Now both risk and return are allowed to vary across the days of the week.

Values and Significances of Day of the Week Effects in South East Asian Stock Markets allowing for Time-Varying risks

	Thailand	Malaysia	Taiwan
Monday	0.00322	0.00185	0.544E-3
	(3.3571)**	(2.8025)**	(0.3945)
Tuesday	-0.00114	-0.00122	0.00140
	(-1.1545)	(-1.8172)	(1.0163)
Wednesday	-0.00164	0.25E-3	-0.00263
	(-1.6926)	(0.3711)	(-1.9188)
Thursday	0.00104	0.00157	-0.00166
	(1.0913)	(2.3515)*	(-1.2116)
Friday	0.31E-4	-0.3752	-0.13E-3
	(0.03214)	(-0.5680)	(-0.0976)
Beta-Monday	0.3573	0.5494	0.6330
	(2.1987)*	(4.9284)**	(2.7464)**
Beta-Tuesday	1.0254	0.9822	0.6572
	(8.0035)**	(11.2708)**	(3.7078)**
Beta-Wednesday	0.6040	0.5753	0.3444
	(3.7147)**	(5.1870)**	(1.4856)
Beta-Thursday	0.6662	0.8163	0.6055
	(3.9313)**	(6.9846)**	(2.5146)*
Beta-Friday	0.9124	0.8059	1.0906
	(5.8301)**	(7.4493)**	(4.9294)**

Notes: Coefficients are given in each cell followed by *t*-ratios in parentheses; * and ** denote significance at the 5% and 1% levels respectively. Source: Brooks and Persand (2001).

Threshold Autoregressive (TAR) Models

- Intuition: a variable is specified to follow different autoregressive processes in different regimes, with movements between regimes governed by an observed variable.
- The model is

$$y_t = \begin{cases} \mu_1 + \phi_1 y_{t-1} + u_{1t} \; if \; s_{t-k} < r \\ \mu_2 + \phi_2 y_{t-1} + u_{2t} \; if \; s_{t-k} \geq r \end{cases}$$

- But what is s_{t-k}? It is the state determining variable and it can be any variable which is thought to make y_t shift from one regime to another.

SETAR model

- If $k = 0$, it is the current value of the state-determining variable that influences the regime that y is in at time t.
- The simplest case is where $s_{t-k} = y_{t-k}$ we then have a self-exciting TAR, or a SETAR. The model is

$$y_t = \begin{cases} \mu_1 + \phi_1 y_{t-1} + u_{1t} \; if \; y_{t-k} < r \\ \mu_2 + \phi_2 y_{t-1} + u_{2t} \; if \; y_{t-k} \geq r \end{cases}$$

- We could of course have more than one lag in each regime (and the number of lags in each need not be the same).
- Under the TAR model approach, unlike the Markov switching model, the transitions between regimes are discrete.

SETAR model (cont.)

- Note that you can rewrite this model

$$y_t = \begin{cases} \mu_1 + \phi_1 y_{t-1} + u_{1t} \; if \; y_{t-k} < r \\ \mu_2 + \phi_2 y_{t-1} + u_{2t} \; if \; y_{t-k} \geq r \end{cases}$$

alternatively to

$$y_t = (\mu_1 + \Phi_1 * y_{t-1})(1 - I[y_{t-k} > r]) + (\mu_2 + \Phi_2 * y_{t-1})(1 - I[y_{t-k} < r]) + \varepsilon_t$$

Where I[A] is a indicator function with I[A]=1 if the event A occurs and I[A]=0 otherwise (it is like dummy variable).

Threshold estimation

- In some applications the threshold is known
- For example, inflation above or below inflation target, or positive vs. negative GDP growth in business cycle analysis
- If the threshold is unknown, it must be estimated
- Grid search is used, e.g. different values for thresholds are tried
- Threshold, which gives the best fit, is chosen

STAR

- SETAR model assumes that the border between two specific regimes is given by a specific value of threshold variable
- More gradual transition between the different regimes can be obtained by replacing the indicator function I[A] by a continuous function $G(y_{t-1};\gamma,c)$, which changes from 0 to 1 as y_{t-1} increases.
- In consequence, we receive a Smooth Transition AR (STAR)

STAR model (cont.)

- *Instead of* $y_t = (\mu_1 + \Phi_1 * y_{t-1})(1 - I[y_{t-k} > r]) + (\mu_2 + \Phi_2 * y_{t-1})(1 - I[y_{t-k} < r]) + \varepsilon_t$
- STAR model is given as
- $y_t = (\mu_1 + \Phi_1 * y_{t-1})(1 - G(y_{t-1}; \gamma, c) > r]) + (\mu_2 + \Phi_2 * y_{t-1})(1 - G(y_{t-1}; \gamma, c) < r]) + \varepsilon_t$
- A popular choice for so-called transition function $G(y_{t-1}; \gamma, c)$ is the logistic function $G(y_{t-1}; \gamma, c) = 1/(1+\exp(-\gamma[y_{t-1} - c]))$
- c is interpreted as the regime threshold
- γ *d*etermines the smoothness of the change in the value of logistic function

Simple example of TAR model

Testing asymmetry of monetary policy rule

- Many researchers estimate the following regression called the monetary policy rule (or Taylor rule):
- *Interest rate = a + b*(inflation-inflation target)+c*(output-potential output)+residual*
- May be the coefficient b can differ according to if (inflation-inflation target)>0 or (inflation-inflation target)<0
- If (inflation-inflation target)>0, central bankers could be more concerned about anchoring inflation expectations and treat target asymmetrically
- Solution: Estimate TAR model and test, if b differs according to the sign of inflation-inflation target.

Threshold Models: Estimation Issues

- Estimation of parameters in the context of threshold models is complex.
- Quantities to be determined include the number of regimes, the threshold variable, the threshold variable lag, the value of the threshold, and the coefficients for each regime.
- We cannot estimate all of these at the same time, so some are usually specified *a priori* based on theory or intuition and the others estimated conditional upon them. E.g., set $k = 1$, $J = 2$, r may not require estimation, etc.

Threshold Models: Estimation Issues (cont.)

- The lag length for each regime can be determined using an information criterion conditional upon a specified threshold variable and fixed threshold value. For example Tong (1990) proposes a modified version of AIC:

$$AIC(p_1, p_2) = T_1 \ln\hat{\sigma}_1^2 + T_2 \ln\hat{\sigma}_2^2 + 2(p_1 + 1) + 2(p_2 + 1)$$

  where T_1 and T_2 are the number of observations in regimes 1 and 2 respectively, p_1 and p_2 are the lag lengths, and $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ are the residual variances.
- Estimation of the autoregressive coefficients can then be achieved using nonlinear least squares (NLS).

An Example of a SETAR Model for the French franc / German mark Exchange Rate

- From Chappell *et al*., 1996, Journal of Forecasting
- The study used daily data from 1/5/90 - 30 / 3/ 92.
- Both the FRF & DEM were then in the ERM which allowed for "managed floating".
- Can use a SETAR to allow for different types of behaviour according to whether the exchange rate is close to the ERM boundary. Currencies are allowed to move up to ±2.25% either side of their central parity in the ERM.
- This would suggest the use of the 2-threshold (3-state) SETAR. This did not work as the DEM was never a weak currency then.
- Ceiling in the ERM corresponded to 5.8376 (log of FRF per 100 DEM).
- The first 450 observations are used for model estimation, with the remaining 50 being retained for out of sample forecasting.
- Forecasts are then produced using the threshold model, the SETAR model with 2 thresholds, a random walk and an AR(2)

Estimated FRF-DEM Regime Switching Model and Out of Sample Forecast Accuracies

Model	For Regime	Number of observations
$\hat{E}_t = 0.0222 + 0.9962E_{t-1}$ (0.0458) (0.0079)	$E_{t\text{-}1} < 5.8306$	344
$\hat{E}_t = 0.3486 + 0.4394E_{t-1} + 0.3057E_{t-2} + 0.1951E_{t-3}$ (0.2391) (0.0889) (0.1098) (0.0866)	$E_{t\text{-}1} \geq 5.8306$	103

Source: Chappell, Padmore,and Mistry (1996). Reprinted with permission of John Wiley and Sons.

Panel A: Mean Squared Forecast Error	Steps Ahead				
	1	2	3	5	10
Random Walk	1.84E-07	3.49E-07	4.33E-07	8.03E-07	1.83E-06
AR(2)	3.96E-07	1.19E-06	2.33E-06	6.15E-06	2.19E-05
One threshold SETAR	1.80E-07	2.96E-07	3.63E-07	5.41E-07	5.34E-07
Two threshold SETAR	1.80E-07	2.96E-07	3.63E-07	5.74E-07	5.61E-07
Panel B: Median Squared Forecast Error					
Random Walk	7.80E-08	1.04E-07	2.21E-07	2.49E-07	1.00E-06
AR(2)	2.29E-07	9.00E-07	1.77E-06	5.34E-06	1.37E-05
One threshold SETAR	9.33E-08	1.22E-07	1.57E-07	2.42E-07	2.34E-07
Two threshold SETAR	1.02E-07	1.22E-07	1.87E-07	2.57E-07	2.45E-07

Source: Chappell, Padmore,and Mistry (1996). Reprinted with permission of John Wiley and Sons.