Advanced Econometrics: TIME SERIES ANALYSIS

TOPIC 6:

Non-Linear Time Series models and Models with changing coefficients

Plan

- Cover basic ideas
- Cover some univariate models
  - Motivations, setups, descriptions
- Mention extensions to multivariate models
- Main Moral of the Story:
  Sound economics should guide choices
  - not bells and whistles

Basic Idea

- Time series behaviour of processes until now
  - Underlying driving processes always linear in parameters
  - Parameters of the DGP were constants (relative to the level of the series, shocks, time)
    - Even if first two moments of data were not

- This can be generalized to cases where
  - Parameters that vary over range of variables/errors
  - Parameters that vary over time
    - structural breaks
  - Parameters that “switch” from one value to another with some unknown probability (Markov Switching Models)
    - Probability of switching can be exogenous or endogenous
  - Parameters that vary endogenously over time
    - The parameters themselves are time series processes

Issues

- Vast array of directions non-linearity could go
  - TAR (1TAR, multi-TAR, band-TAR)
  - STAR (LSTAR, ESTAR, STARCH, E-STARCH)
  - GAR
  - Bi Linear
  - Structural breaks
- Detecting SOME non-linearity is not difficult
  - distinguishing between two specific types is very hard
- Estimation becomes more fragile
  - "Number of alternatives one could consider may increase faster than sample size"
  - Different non-linear models are non-nested
  - Use sound economic reasoning to guide what type is considered

Two broad classes:

1. Direct generalization of ARMA (or VAR) models to incorporate non-linearities into the (single) DGP
   – We will discuss these briefly: not very intuitive
2. Regime switching (or “structural break”) models, where there are more than one “possible version” of the DGP
   – Very intuitive:
      - Actual policy regimes change discretely
         – Monetary aggregate targeting vs inflation targeting
         – Auction based monetary policy vs full allotment corridor system (EU)
         – Fixed vs floating exchange rates
         – The different Basel frameworks on banking regulation
      - Periods of collusion vs no collusion (a micro-economic example)

Two broad classes:

1. Direct generalization of ARMA (or VAR) models to incorporate non-linearities into the (single) DGP
   – We will discuss these briefly: not very intuitive
2. Regime switching (or "structural break") models, where there are more than one "possible version" of the DGP
   – Very intuitive:
      - Actual policy regimes change discretely
         – Monetary aggregate targeting vs inflation targeting
         – Auction based monetary policy vs full allotment corridor system (EU)
         – Fixed vs floating exchange rates
         – The different Basel frameworks on banking regulation
      - Periods of collusion vs no collusion (a micro-economic example)

Direct Generalization of ARMA

- Idea:
  - exact form of non-linearity unknown so approximate via a Taylor polynomial approximation

- Problems:
  - Interpretation
  - Distinguishing between non-linearity vs time varying parameters
  - Dimensionality – unsuited to Multivariate analysis

Detecting non-linear effects

- Against the null of linearity, several tests can be used for detecting some non-linearity, but few can provide a specific structure of non-linearity as an alternative

Tests:

- Generic:
  - Parameter stability tests
    - Strong non-linear effects should make linear coefficients unstable over the sample
  - ACF and Ljung-Box tests of the square of residuals from a linear regression
  - Regression Error Specification Test (RESET)
    - Fit linear model, store residuals and fitted values
    - Regress residuals on exogenous variables and higher order versions of fitted values
    - If linearity is appropriate, coefficients should be zero
    - Conduct an F-test for non-zero coefficients
- Against a specific non-linear alternative:
  - Lagrange Multiplier test against a specific non-linear alternative – see Enders

Non-linear autoregressive models

- Shape of non-linearity unknown
  - GAR(p)
    - Generalized AR(p) – E.g. Quadratic GAR(2)
    - Based on Taylor Approximation
    - Estimation simple (once order chosen)
  - Bilinear
    - Generalized ARMA(p,q)
    - Random AR coefficients with fixed mean
    - Estimation requires Maximum Likelihood approach
- Difficult to interpret in clean, economically consistent way
  - Very hard to make call on which is best

Non-linear autoregressive models

- Simplest statement: $$y_t = f(y_{t-1}) + \varepsilon_t$$
  - Where f is some nonlinear function
- Alternative view:

$$y_t = a_1(y_{t-1}) \cdot y_{t-1} + \varepsilon_t$$

  - I.e. the value of the autoregressive coefficient depends on the level of the lagged value

Non-linear autoregressive models

- Third order Generalized AR(2):

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + a_{12} y_{t-1} y_{t-2} + a_{11} y_{t-1}^2 + a_{22} y_{t-2}^2 + a_{112} y_{t-1}^2 y_{t-2} + a_{122} y_{t-1} y_{t-2}^2 + a_{111} y_{t-1}^3 + a_{222} y_{t-2}^3 + \varepsilon_t$$

- General GAR(p):

$$y_t = a_0 + \sum_{i=1}^{p} a_i y_{t-i} + \sum_{i=1}^{p} \sum_{j=1}^{p} \sum_{k=1}^{r} \sum_{l=1}^{s} a_{ijkl} y_{t-i}^k y_{t-j}^l + \varepsilon_t$$

- The problem of dimensionality is obvious
- Additionally, even/odd powers are likely to be highly correlated – precise estimates difficult
- Testing for non-linearity often uses this formulation, as it is still linear in parameters, so t/F tests are valid
  - Enders gives nuanced details

Non-linear autoregressive models

- Generalizing an ARMA model gives a Bilinear model:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + c_1 y_{t-1}\varepsilon_{t-1} + \varepsilon_t$$

- This can be interpreted as an AR model with a random autoregressive coefficient:

$$y_t = \alpha_0 + (\alpha_1 + c_1\varepsilon_{t-1})y_{t-1} + \varepsilon_t$$

- E.g. if $c_1 > 0$, positive shocks are more persistent than negative shocks

Two broad classes:

1. Direct generalization of ARMA (or VAR) models to incorporate non-linearities into the (single) DGP
   – We will discuss these briefly: not very intuitive
2. Regime switching (or “structural break”) models, where there are more than one “possible version” of the DGP
   – Very intuitive:
     - Actual policy regimes change discretely
       – Monetary aggregate targeting vs inflation targeting
       – Auction based monetary policy vs full allotment corridor system (EU)
       – Fixed vs floating exchange rates
       – Periods of collusion vs no collusion (a micro-economic example. Thank you, Annerie)

Regime Switches and Structural Breaks

- Idea:
  - We are studying a dynamic system of interaction with three economically relevant features:
  1. Continual technological progress
  2. Frictions
     - Transaction/transportation costs may influence the incentives for active arbitrage of small deviations
  3. Occasional sharp shocks/changes
     - Changes in policy may change the dynamic relationship between variables
     - Unemployment may behave differently in recessions vs booms
     - Monetary policy may be more active in recessions than in booms

- Guiding principle:
  - Use economic reasoning on relevant features to suggest parsimonious, interpretable empirical generalizations of time series models

Examples

- Economic examples can suggest different types of non-linearity
  - Behaviour piecewise linear around a focal point
    - Focal point can be a level of the variable (Threshold-AR) or a point in time (Structural Break)
  - Behaviour smooth function of distance from focal point
    - Generalizes TAR models
  - Behaviour function of a discrete set of states
    - Markov Switching models

Options

- We will study the different generalizations of the linear model in the following sequence, using the single equation modelling environment as the simplest source of intuition
  1. Discrete threshold AR models (TAR)
     - Some level of the endogenous variable implies a change of dynamics of the DGP
  2. Smooth threshold AR models (STAR)
     - Dynamics of the DGP depends on the distance from some level of some variable
  3. Structural Breaks
     - At one (or more) points in time, the dynamics of DGP changes permanently
  4. Regime Switching
     - The economy is in one of a few possible states. In each, the dynamics of the DGP is different and it may switch between states randomly
  5. (Continuously) Time varying coefficients
     - The dynamics (as represented by the coefficients) of the DGP evolves continuously and endogenously in a way that can be captured by time series methods

Options

- We will study the different generalizations of the linear model in the following sequence, using the single equation modelling environment as the simplest source of intuition
  1. Discrete threshold AR models (TAR)
     - Some level of the endogenous variable implies a change of dynamics of the DGP
  2. Smooth threshold AR models (STAR)
     - Dynamics of the DGP depends on the distance from some level of some variable
  3. Structural Breaks
     - At one (or more) points in time, the dynamics of DGP changes permanently
  4. Regime Switching
     - The economy is in one of a few possible states. In each, the dynamics of the DGP is different and it may change states randomly
  5. (Continuously) Time varying coefficients
     - The dynamics (as represented by the coefficients) of the DGP evolves continuously and endogenously in a way that can be captured by time series methods

Threshold example

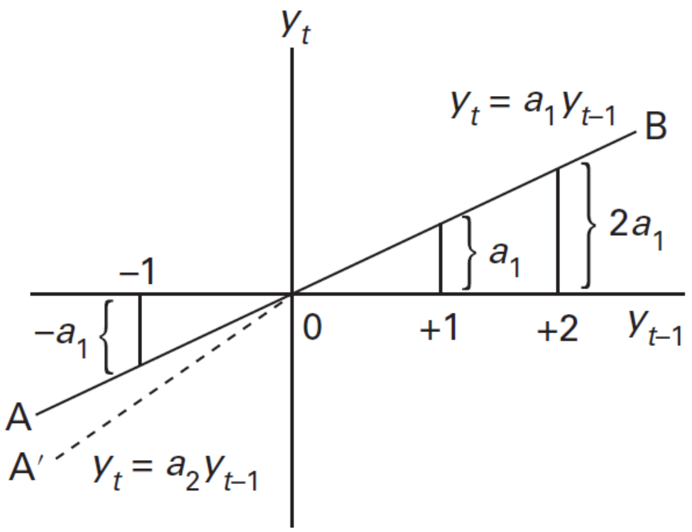
- Unemployment (US) shows clear asymmetric behavior: dynamics in recessions different from in booms
  - Shock: 1% increase in unemployment
- During recessions: Unemployment high, Risks high
  - Large pool of unemployed – risk of remaining unemployed high
  - Shocks to unemployment are likely to be persistent
- During booms: Unemployment low, Risks low
  - Small pool of unemployed – risk of remaining unemployed is low
  - Shocks to unemployment are likely to be less persistent
- Suggests model with single threshold (that defines “recessions” vs “booms”)
  - Different AR(p) model above vs below some threshold (TAR model)
  - Variance?

Threshold AR models

- The simplest version of a TAR model:

$$y_t = \begin{cases} a_1 y_{t-1} + \varepsilon_{1t} \text{ if } y_{t-1} > 0 \\ a_2 y_{t-1} + \varepsilon_{2t} \text{ if } y_{t-1} \le 0 \end{cases}$$

- I.e. the degree of persistence is different depending on the value the lagged term



Threshold AR models

- The simplest version of a TAR model:

$$y_t = \begin{cases} a_1 y_{t-1} + \varepsilon_{1t} \text{ if } y_{t-1} > 0 \\ a_2 y_{t-1} + \varepsilon_{2t} \text{ if } y_{t-1} \leq 0 \end{cases}$$

  - I.e. the degree of persistence is different depending on the value the lagged term
- Estimation is simplest when the variance is equal in the two regimes (i.e. $\mathrm{var}(\varepsilon_{1t}) = \mathrm{var}(\varepsilon_{2t})$)

$$y_t = I_t \left[\alpha_{10} + \sum_{i=1}^{p} \alpha_{1i} y_{t-i} \right] + (1 - I_t) \left[\alpha_{20} + \sum_{i=1}^{r} \alpha_{2i} y_{t-i} \right] + \varepsilon_t$$

  - Where $I_t = 1$ if $y_{t-1} < \tau$ (the threshold), $I_t = 0$ if $y_{t-1} \geq \tau$

Threshold AR models

- More generally:

$$y_t = \begin{cases} \alpha_{10} + \alpha_{11} y_{t-1} + \cdots + \alpha_{1p} y_{t-p} + \varepsilon_{1t} \text{ if } y_{t-1} > \tau \\ \alpha_{20} + \alpha_{21} y_{t-1} + \cdots + \alpha_{2r} y_{t-r} + \varepsilon_{2t} \text{ if } y_{t-1} \leq \tau \end{cases}$$

- For some given threshold level τ
- Estimation:
  - Since each component is linear, OLS gives consistent estimates
  - To allow different variances, simply split the sample based on the threshold and estimate the different processes independently
  - Replace values of y with zeros if outside of the regime

Threshold AR models

- Typically, one does not know the threshold τ
- Happily, there is a simple approach that is super consistent for estimating the threshold:
  - For the threshold to be interesting, the process must cross it
  - So we have an obvious set of potential thresholds: the observed values of the series itself!
  - I.e. order observations from small to large
  - For the middle 70% (rule of thumb) of these values estimate a TAR model around each value (candidate threshold)
  - Compare the sum of square residuals – the best fitting model contains the consistent estimate of the threshold

Threshold AR models

Extensions

- Estimating the “delay” parameter (which lag determines the shift between regimes) works exactly the same as estimating the threshold:
  - Estimate all models with different delay parameters and select the best
- If the threshold/regime shift is thought as occurring at a date, the same procedure:
  - Estimate the model for all possible dates of regime shift, pick the best
- More than one threshold
- Smooth transition models

Multiple thresholds in the level of the variable

- The real-world economics of exchange rates provides an intuitive setting for models with more than one threshold

- Exchange Rates:
  - With representative agents (identically informed)
  - Without transaction/information costs,
    - Perfect info, zero transaction costs implies system always in equilibrium
    - All shocks are unpredictable deviations from equilibrium (e.g. exogenous Monetary shocks)
    - Nominal Exchange rates should be a random walks by arbitrage
    - Additionally, nominal exchange rates should be co-integrated with prices
    - Thus, real exchange rate should be stationary, ceteris paribus

- Exchange Rates:
  - – With representative agents (identically informed)
  - – With transaction/information costs
    - Equilibrium unknown, shocks not identifiable
    - Any observed change in exchange rate is
      - – Measurement error, or
      - – A shock that moves system away from equilibrium, or
      - – A correction of a previous shock that moved the system away from equilibrium, or
      - – A structural shocks that moves system to new equilibrium, or
      - – A mixture of the above

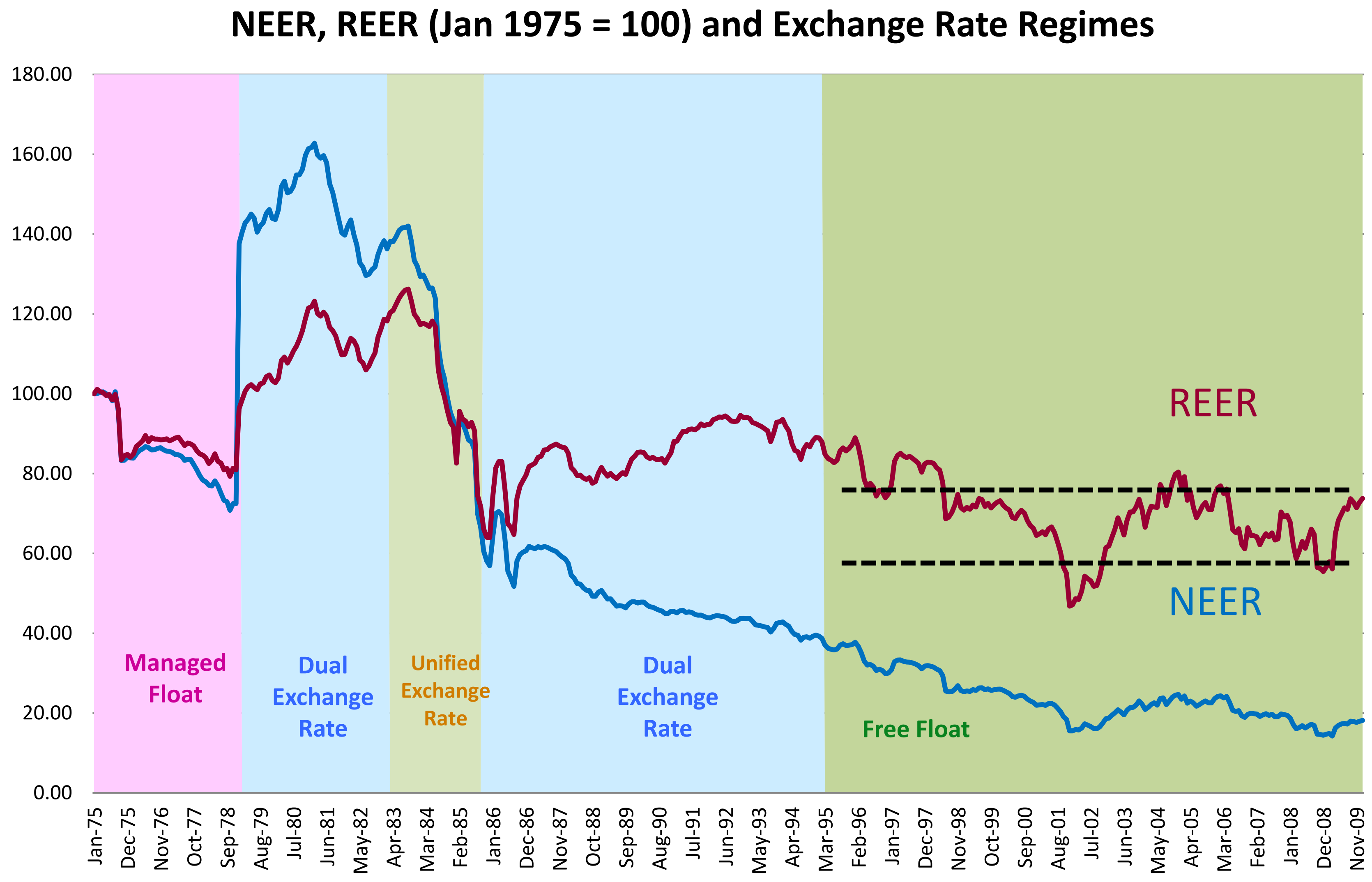
- Exchange Rates with transaction costs
- Suppose economy is known to be in a period of “constant equilibrium real exchange rate”. Consider two situations:
  - Exchange rate near equilibrium
  - Exchange rate far from equilibrium

- Exchange Rates with transaction costs
- Suppose economy is known to be in a period of “constant equilibrium real exchange rate”
  - Exchange rate near equilibrium
    - Potential gains smaller than transaction cost = not arbitraged away
    - (Small) Deviations permanent – Random Walk behaviour
  - Exchange rate far from equilibrium
    - Potential gains large, deviations arbitraged away
    - (Large) Deviations temporary – Mean reverting behaviour

- Exchange Rates with transaction costs
  - Exchange rate near “equilibrium”
    - Random Walk behaviour
  - Exchange rate far from “equilibrium”
    - Mean Reverting behaviour
- Suggests model with threshold in terms of a “band”, or two thresholds (band-TAR model)
  - Different AR(p) model within band than outside

$$
\begin{aligned}
s_t &= \bar{s} + a_1(s_{t-1} - \bar{s}) + \varepsilon_t && \text{when } s_{t-1} > \bar{s} + c \\
s_t &= s_{t-1} + \varepsilon_t && \text{when } \bar{s} - c < s_{t-1} \leq \bar{s} + c \\
s_t &= \bar{s} + a_2(s_{t-1} - \bar{s}) + \varepsilon_t && \text{when } s_{t-1} \leq \bar{s} - c
\end{aligned}
$$

Heuristic for Band-TAR



Issues with TAR models

- How many thresholds?
  - What happens when thresholds -> T?
  - Data driven choice best, but risk of data mining, overfitting
- We will consider tests for the number of thresholds under the topic of structural breaks, as they are analytically similar

Options

- We will study the different generalizations of the linear model in the following sequence, using the single equation modelling environment as the simplest source of intuition

1. Discrete threshold AR models (TAR)
   - Some level of the endogenous variable implies a change of dynamics of the DGP
2. Smooth threshold AR models (STAR)
   - Dynamics of the DGP depends on the distance from some level of some variable
3. Structural Breaks
   - At one (or more) points in time, the dynamics of DGP changes permanently
4. Regime Switching
   - The economy is in one of a few possible states. In each, the dynamics of the DGP is different and it may change states randomly
5. (Continuously) Time varying coefficients
   - The dynamics (as represented by the coefficients) of the DGP evolves continuously and endogenously in a way that can be captured by time series methods

Smooth Transition Examples

- Exchange Rates:
  - With Heterogeneous Agents
  - With transaction/information costs
- Simplistic case again:
  myopic but distinct agents
  - Agents have different beliefs AND
  - don’t change their beliefs
- Suppose all believe in the Band-TAR model, but have different beliefs on width of band

Smooth Transition Examples

- Suppose all believe in the Band-TAR model, but have different beliefs on width of band
  - As move further away from equilibrium
  - More traders begin to arbitrage
  - mean reverting behavior becomes stronger
  - Aggregate behaviour smoothly transitions from Random Walk to Mean Reversion as distance from Equilibrium increases
  - Variants: LSTAR, ESTAR (differ in transition shape)

Smooth Transition AR models

- General STAR(1):

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p} + \theta[\beta_0 + \beta_1 y_{t-1} + \cdots + \beta_p y_{t-p}] + \varepsilon_t$$

- Where $\theta \in [0,1]$ is some smooth function of the data

Types:

- LSTAR (logistic)

$$\theta = [1 + \exp(-\gamma(y_{t-1} - c))]^{-1}$$

  - One-sided:

$$y_{t-1} \to -\infty, \theta \to 0 \qquad y_{t-1} \to +\infty, \theta \to 1$$

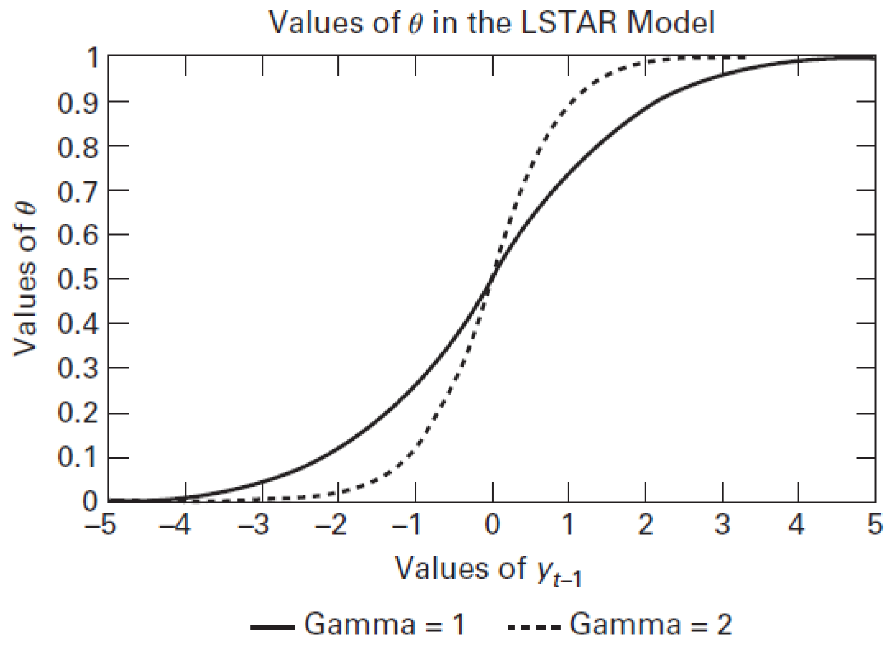
- ESTAR (exponential)

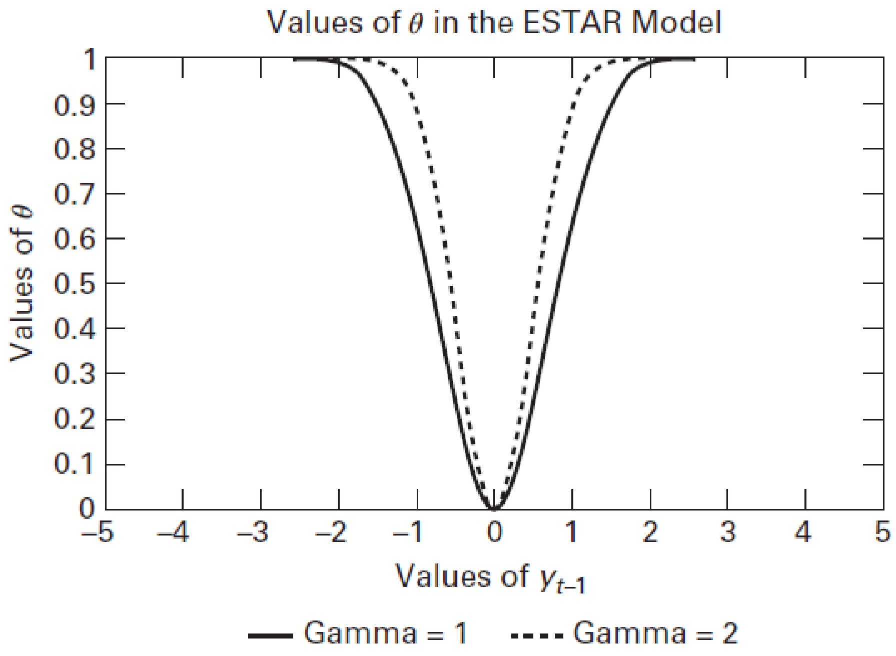
$$\theta = 1 - \exp\left[-\gamma(y_{t-1} - c)^2\right] \quad \gamma > 0.$$

  - Symmetric - smooth version of Band-TAR
  - At $y_{t-1} = c, \theta = 0$, as y_{t-1} moves further away from c, $\theta \to 1$

Smooth Transition AR models

•





Options

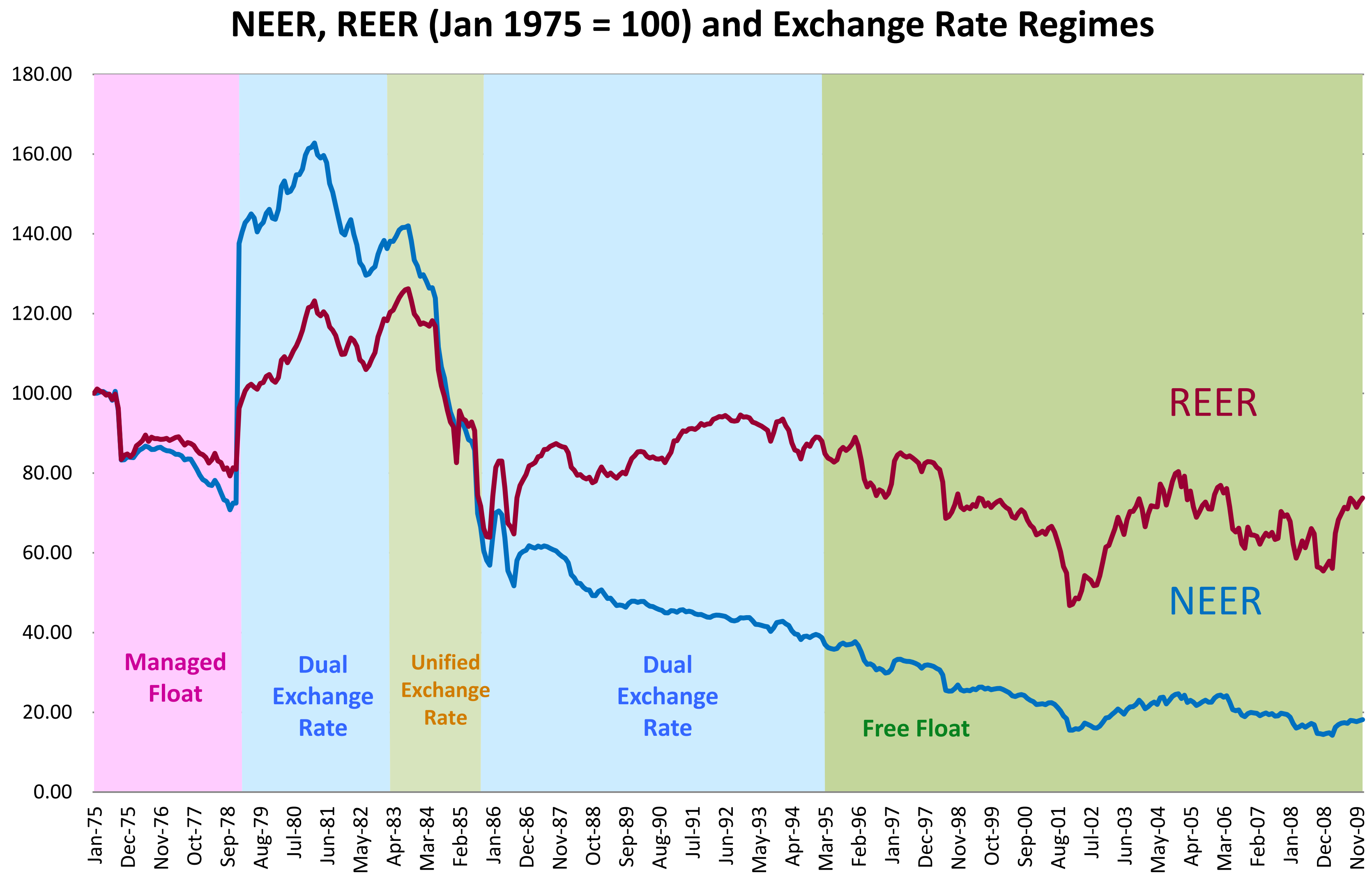
- We will study the different generalizations of the linear model in the following sequence, using the single equation modelling environment as the simplest source of intuition

1. Discrete threshold AR models (TAR)
   - Some level of the endogenous variable implies a change of dynamics of the DGP
2. Smooth threshold AR models (STAR)
   - Dynamics of the DGP depends on the distance from some level of some variable
3. Structural Breaks
   - At one (or more) points in time, the dynamics of DGP changes permanently
4. Regime Switching
   - The economy is in one of a few possible states. In each, the dynamics of the DGP is different and it may change states randomly
5. (Continuously) Time varying coefficients
   - The dynamics (as represented by the coefficients) of the DGP evolves continuously and endogenously in a way that can be captured by time series methods

Structural Breaks

- Sometimes known events change the structure of interaction/behavior
- This suggests that the DGP before and after the event may be different
- At some date τ the DGP changes
- Problem:
  - "official dates" need not be the economically relevant dates
  - Economic agents may anticipate changes

exchange rate regimes



Structural Breaks

- Sometimes known, exogenous events change the structure of interaction/behavior
  - This suggests that the DGP before and after the event may be different

- Detecting structural breaks:
  - Known exogenous break:
    - Chow breakpoint test
  - Unknown/Endogenous break(s):
    - Supremum Test
    - Generalized implementation: Bai and Perron (2003)

Detecting a break: single known break point

- Suppose we have a sample of T time periods and we know that the DGP changes at some date t_m
- The Chow break-point test proceeds as follows:
  - Estimate the model with all data, and store the SSR
  - Estimate the model with data up to t_m: $t = [1, t_m]$, store SSR_1
  - Estimate the model with data after t_m: $t = [t_{m+1}, t_T]$, store SSR_2

$$y_t = a_0(1) + a_1(1)y_{t-1} + \cdots + a_p(1)y_{t-p} + \varepsilon_t + \beta_1(1)\varepsilon_{t-1} + \cdots + \beta_q(1)\varepsilon_{t-q}$$
$$\text{using } t_1, \ldots, t_m$$
$$y_t = a_0(2) + a_1(2)y_{t-1} + \cdots + a_p(2)y_{t-p} + \varepsilon_t + \beta_1(2)\varepsilon_{t-1} + \cdots + \beta_q(2)\varepsilon_{t-q}$$
$$\text{using } t_{m+1}, \ldots, t_T$$

Detecting a break: single known break point

- The F-statistic for no difference before/after t_m is given by:

$$F = \frac{(\mathrm{SSR} - \mathrm{SSR}_1 - \mathrm{SSR}_2)/n}{(\mathrm{SSR}_1 + \mathrm{SSR}_2)/(T - 2n)}$$

- Intuitively:
  - If the breakpoint is irrelevant, the individual models should be identical to the full sample model, so we expect: $SSR = SSR_1 + SSR_2$, or that the F-statistic should be zero
  - If the F-statistic is larger than a critical value, we reject the hypothesis that the two separate models are the same as the full-sample model and conclude that there is a significant structural break at t_m

Detecting a single unknown break-point

- Typically we do not know when the break occurs
- We can extend the Chow test as follows (Andrews 1993):
  – Let the DGP change at some unknown date t^*:

$$y_t = \begin{cases} \beta_{01} + \sum_{i=1}^{k} \beta_{k1} z_t + \varepsilon_t \ \text{ if } \ t < t^* \\ \beta_{02} + \sum_{i=1}^{k} \beta_{k2} z_t + \varepsilon_t \ \text{ if } \ t \geq t^* \end{cases}$$

Detecting a single unknown break-point

- Typically we do not know when the break occurs
- We can extend the Chow test as follows (Andrews 1993):
  – Let the DGP change at some unknown date t^*:
  – Define $D_t = 1$ if $t \geq t^*, 0$ otherwise:

$$y_t = \alpha_{01} + \sum_{i=1}^{k} \beta_{k1} z_t + D_t \left(\gamma_0 + \sum_{i=1}^{k} \gamma_k z_t \right) + \varepsilon_t$$

  – So that $\beta_{k2} = \beta_{k1} + \gamma_k$

Detecting a single unknown break-point

$$y_t = \alpha_{01} + \sum_{i=1}^{k} \beta_{k1} z_t + D_t \left(\gamma_0 + \sum_{i=1}^{k} \gamma_k z_t \right) + \varepsilon_t$$

- Estimate the model without the break, store the squared residuals SSR_r
- For each possible date of the break (rule of thumb: excluding the first and last 15% of the time period), estimate the model with the break, store SSR_u
- Select the model with the lowest SSR_u
- Equivalently, this is the model with the highest F-statistic:

$$F = \frac{\frac{SSR_r - SSR_u}{k+1}}{\frac{SSR_u}{T - 2(k+1)}}$$

Testing whether break-points are significant

- The above algorithm gives the best estimate of the break-point, t^*
  - **If** there is a break-point, this estimate is super-consistent.
  - However: if there isn't one, the estimate is not identified
- Moreover, the F-statistic of the best model no longer has an F distribution
  - Why?

Testing whether break-points are significant

- The above algorithm gives the best estimate of the break-point, t^*
  - **If** there is a break-point, this estimate is super-consistent.
  - However: if there isn't one, the estimate is not identified
- Moreover, the F-statistic of the best model no longer has an F distribution
  - Why? Because we searched over several models for it! It is not a single random draw from the F-distribution
  - What we need is the distribution of the maximum (or supremum) of a set of draws of the F-distribution of the null
  - This has been tabulated by Andrews (1993) and generalized by Bai and Perron (1998).

Detecting multiple break-points

- Bai and Perron (2003) gives a practical guide to empirical implementation of the more general test when there can be $n \geq 1$ breaks. They provide:
  - An efficient algorithm
  - Allow for pure or partial breaks (only some of the coefficients change)
  - Estimating the number of breaks
  - Confidence intervals for the break dates
  - Testing under very general assumptions on data and errors

$$y_t = \alpha_{01} + \sum_{i=1}^{k} \beta_{k1} z_t + \sum_{j=1}^{n} D_{jt} \left(\gamma_0 + \sum_{i=1}^{k} \gamma_k z_t \right) + \varepsilon_t$$

Detecting multiple break-points

- Bai and Perron (2003)
  - Two algorithms:
    - test for all possible number of break points against the null of none

q	$k = 1$	$k = 2$	$k = 5$	UDmax
1	9.63	8.78	6.69	10.17
2	12.89	11.60	9.12	13.27
3	15.37	13.84	11.15	16.82

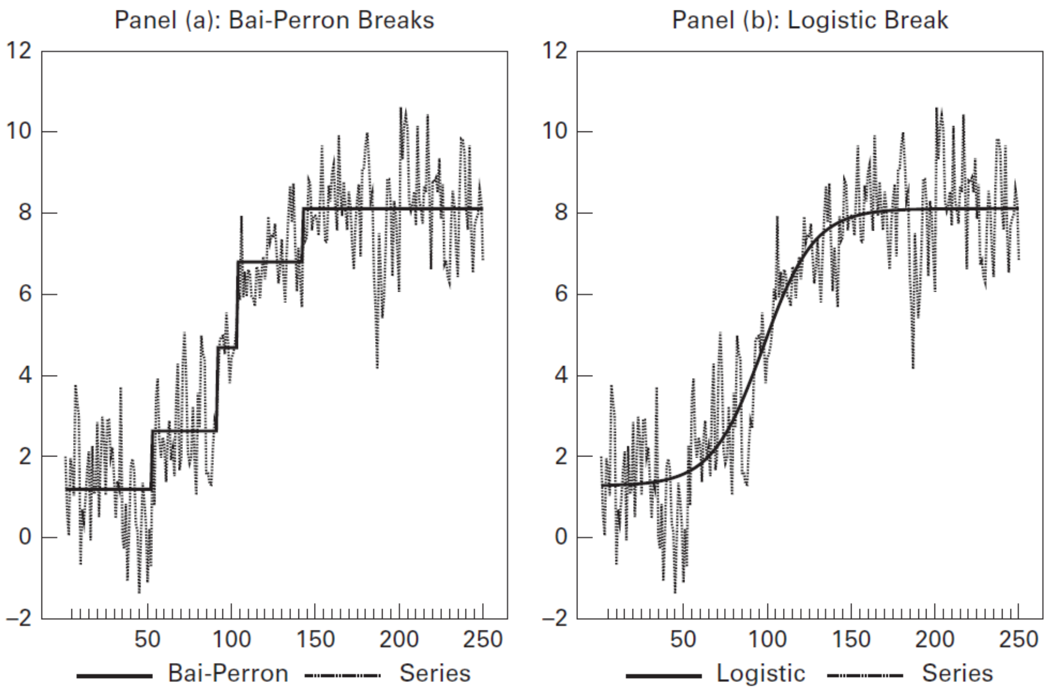
    - Test sequentially: 1 break against null of none, 2 breaks against the null of 1…

q	$\ell = 0$	$\ell = 1$	$\ell = 2$	$\ell = 4$
1	9.63	11.14	12.16	13.45
2	12.89	14.50	15.42	16.61
3	15.37	17.15	17.97	19.23

    - Use SBC to select number of breaks (or economic reasoning)
    - q is the number of coefficients allowed to differ across regimes, k the number of breaks and the values in the tables are the 95% confidence critical values.

Smooth (gradually) temporal breaks

- The Bai and Perron (2003) algorithm works for discrete breaks that happen from one period to the next
- An alternative view would be that a structural break occurs gradually/smoothly over a number of periods
- To estimate such a non-linear, smooth break, one can use the Logistic Break model
  - There are two regimes with coefficient vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.
  - A transition parameter θ that is a logistic function with centrality parameter t^* that has to be estimated
  - t^* is the "midpoint" between the two regimes



$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \cdots + \alpha_p y_{t-p} + \theta[\beta_0 + \beta_1 y_{t-1} + \cdots + \beta_p y_{t-p}] + \varepsilon_t$$

where

$$\theta = [1 + \exp(-\gamma(t - t^*))]^{-1}$$

Testing for a unit root if there is a structural break

- Zivot and Andrews (2002) present a test that extends the ADF test to allow for an unknown structural break in a similar way (building on Perron (1988,1989))
  - Structural breaks may look like unit root non-stationarity
  - ADF test likely to test a unit root when there is a strong structural break
  - Thus: pick the break-point that is most against the hypothesis
  - They derive the asymptotic distribution of such an “infimum-t” statistic and simulate for small sample corrections

Options

- We will study the different generalizations of the linear model in the following sequence, using the single equation modelling environment as the simplest source of intuition

1. Discrete threshold AR models (TAR)
   - Some level of the endogenous variable implies a change of dynamics of the DGP
2. Smooth threshold AR models (STAR)
   - Dynamics of the DGP depends on the distance from some level of some variable
3. Structural Breaks
   - At one (or more) points in time, the dynamics of DGP changes permanently
4. Regime Switching
   - The economy is in one of a few possible states. In each, the dynamics of the DGP is different and it may change states randomly
5. (Continuously) Time varying coefficients
   - The dynamics (as represented by the coefficients) of the DGP evolves continuously and endogenously in a way that can be captured by time series methods

Markov Switching Models

- Also more than one regime, but switching between them is now stochastic

- Transitions between them do not depend deterministically on time or values of variables
  - Random shifts = need probability description
  - probabilities must be estimated
  - Maximum Likelihood, Bayesian methods
  - Persistence now depends on autocorrelation within regimes and transition probabilities

- Well established algorithms, even for VARs
  - beware of overly parameterised models on small data sets

Markov Switching Models

$$y_t = a_{10} + a_{11}y_{t-1} + \varepsilon_{1t} \quad \text{if } s_t = 1$$
$$y_t = a_{20} + a_{21}y_{t-1} + \varepsilon_{2t} \quad \text{if } s_t = 2$$

With transition probabilities:

$$p_{11} = \Pr(s_{t+1} = 1 | s_t = 1)$$
$$(1 - p_{11}) = \Pr(s_{t+1} = 2 | s_t = 1)$$
$$p_{22} = \Pr(s_{t+1} = 2 | s_t = 2)$$
$$(1 - p_{22}) = \Pr(s_{t+1} = 1 | s_t = 2)$$

This yields *unconditional* probabilities of different regimes:

$$p_1 = (1 - p_{22})/(2 - p_{11} - p_{22})$$
$$p_2 = (1 - p_{11})/(2 - p_{11} - p_{22})$$

Markov Switching Models

Extensions allow the transition probabilities to be functions of explanatory variables (i.e. time varying, endogenous)

Based on estimating the probabilities as logistic functions of data

Filardo, Andrew J. 1994. “Business-Cycle Phases and Their Transitional Dynamics.” *Journal of Business & Economic Statistics* 12 (3): 299–308.

Options

- We will study the different generalizations of the linear model in the following sequence, using the single equation modelling environment as the simplest source of intuition

1. Discrete threshold AR models (TAR)
   - Some level of the endogenous variable implies a change of dynamics of the DGP
2. Smooth threshold AR models (STAR)
   - Dynamics of the DGP depends on the distance from some level of some variable
3. Structural Breaks
   - At one (or more) points in time, the dynamics of DGP changes permanently
4. Regime Switching
   - The economy is in one of a few possible states. In each, the dynamics of the DGP is different and it may change states randomly
5. (Continuously) Time varying coefficients
   - The dynamics (as represented by the coefficients) of the DGP evolves continuously and endogenously in a way that can be captured by time series methods

Yet Other Versions - Multivariate

- VAR with Time varying coefficients
  - Primiceri provides a Bayesian estimation algorithm
    - Stationary VAR (locally)
    - coefficients follow a driftless random walk
  - Primiceri, G.E., 2005. Time varying structural vector autoregressions and monetary policy. *The Review of Economic Studies*, *72*(3), pp.821-852.
  - Del Negro, M. and Primiceri, G.E., 2015. Time varying structural vector autoregressions and monetary policy: a corrigendum. *The review of economic studies*, *82*(4), pp.1342-1345.

- VECM with time varying coefficients
  - Constants of CI-vector: Lacerda et al (2008)
  - All coefficients – not seen yet

Interpreting estimation results

- Non-linear specifications, structural breaks or regime shifts means that the impulse responses are no-longer unique.
  - They will depend on level of the series/size of the shock in a non-linear AR
  - They will differ before and after structural breaks
  - They will differ across different regimes
  - They will evolve over time in a time-varying coefficient model

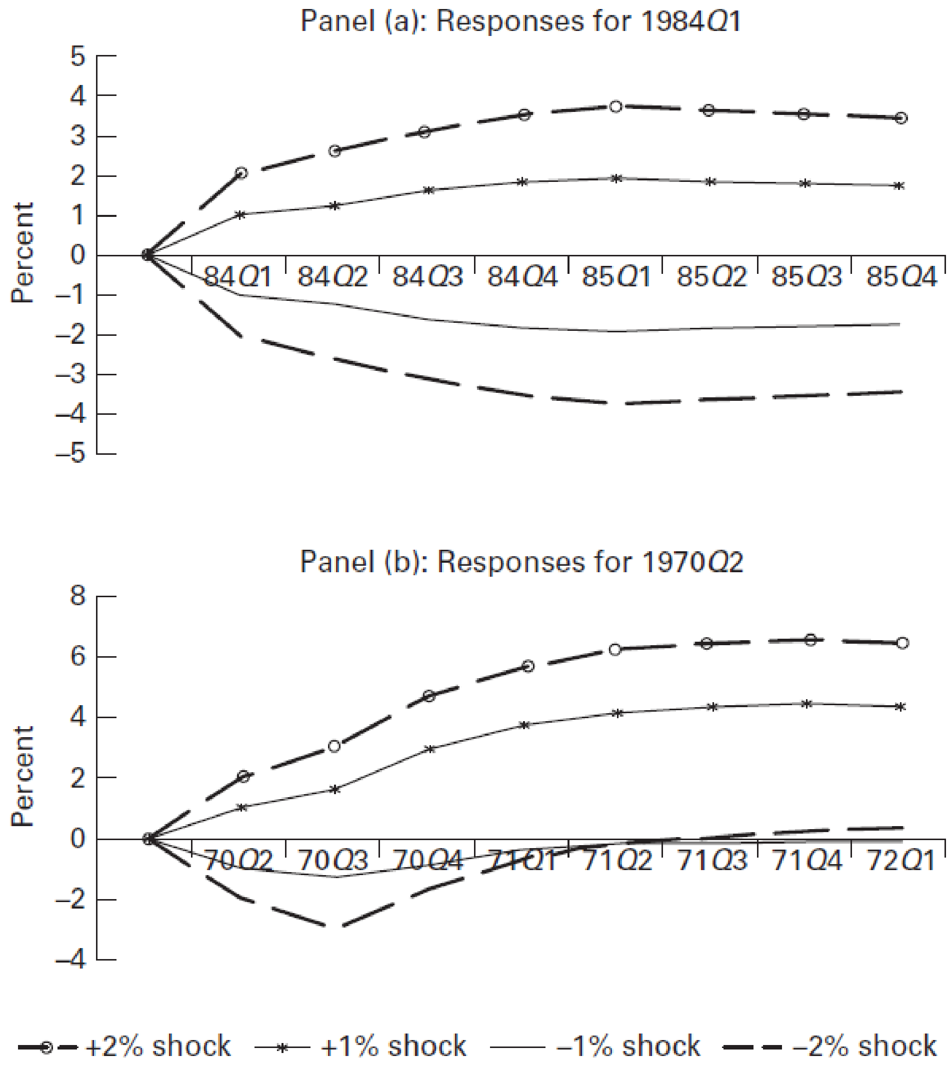


FIGURE 7.11 Impulse Responses for Two Histories

To STARCH or not to STARCH?

- More precisely: when should we consider non-linearities or structural breaks?

- Economic Theory should be primary guide
  - What feature of the underlying structure of the economy should/may induce non-linear behaviour?
  - Use “best understanding” of real world cause/mechanism of non-linearity to determine “most appropriate” non-linear model
  - Always compare to nearest linear version, best alternative non-linear specification
  - Test