# Non-Linear Models
# Examples

Laurent Ferrara

CEF, Ljubljana
Nov. 2019

# Plan

# Simulation of a SETAR model

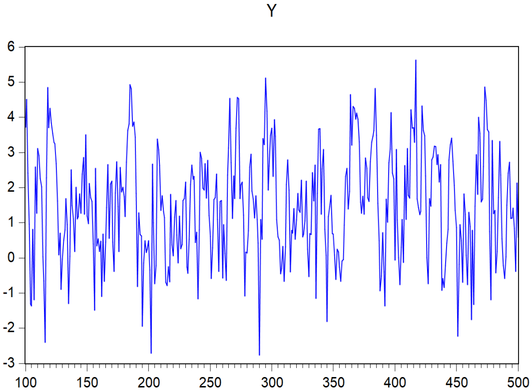
- We simulate the following SETAR(2,1):

$$\begin{aligned} y_t &= (1 - I(y_{t-1} > 1))(0.5 - 0.3y_{t-1}) \\ &\quad + I(y_{t-1} > 1)(0.1 + 0.8y_{t-1}) + \varepsilon_t, \end{aligned} \tag{1}$$

  with unit variance for $\varepsilon_t$ (i.e. $\sigma_\varepsilon^2 = 1$)

- We generate 600 observations, but only the last 500 are considered to avoid dependence to starting values (see Figure)

# Simulated series for SETAR model



# Testing TAR models

- We apply the Hansen (1997) procedure based on Bootstrap replications.
- The 95% acceptance region of $F_s$ for the null of linearity is $[1.89, 10.51]$.
- As the outcome of the test is 54.92, we strongly reject the null of linearity, i.e. the series is supposed to have been generated by a non-linear process

# TAR estimation

We proceeed as follows to estimate the SETAR

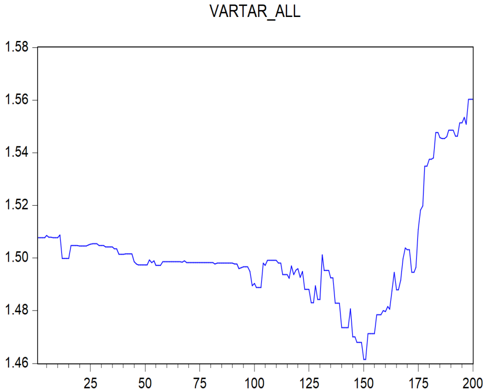
1. Prepare a grid of $M = 200$ possible values for the threshold $c \in [-0.05, 1.5]$
2. For each $i = 1, \ldots, M$, generate a dummy $D_i$ such that $D_i = I(y_{t-1} > c_i)$
3. Using OLS estimate the following model for each $c_i$

$$\begin{aligned} y_t &= (1 - D_i)(\phi_{1,0} + \phi_{1,1} y_{t-1}) \\ &+ D_i(\phi_{2,0} + \phi_{2,1} y_{t-1}) + \varepsilon_t, \end{aligned} \tag{2}$$

4. For each $i$ and resulting OLS estimates $\hat{\phi}^i$, let's compute the variance error term $\hat{\sigma}_t^2(c_i)$
5. Search for $c^* = Arg \min \hat{\sigma}_t^2(c_i)$

# Variance of residuals for various thresholds $c_i$

We choose $c^* = 1$, corresponding to $c_{150}$



# Estimated parameters for $c^* = 1$

Dependent Variable: Y
Method: Least Squares
Date: 08/17/17 Time: 15:53
Sample (adjusted): 101 500
Included observations: 400 after adjustments

| Variable | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| C | 0.440694 | 0.099080 | 4.447860 | 0.0000 |
| DUMMY_C(-1) | 0.025795 | 0.222073 | 0.116157 | 0.9076 |
| Y(-1)*DUMMY_C(-1) | 0.669082 | 0.074419 | 8.990683 | 0.0000 |
| Y(-1)*(1-DUMMY_C(-1)) | -0.288351 | 0.124134 | -2.322899 | 0.0207 |

| | | | |
|---|---|---|---|
| R-squared | 0.396122 | Mean dependent var | 1.486710 |
| Adjusted R-squared | 0.391547 | S.D. dependent var | 1.557600 |
| S.E. of regression | 1.214981 | Akaike info criterion | 3.237283 |
| Sum squared resid | 584.5665 | Schwarz criterion | 3.277198 |
| Log likelihood | -643.4566 | Hannan-Quinn criter. | 3.253090 |
| F-statistic | 86.58721 | Durbin-Watson stat | 1.981115 |
| Prob(F-statistic) | 0.000000 | | |

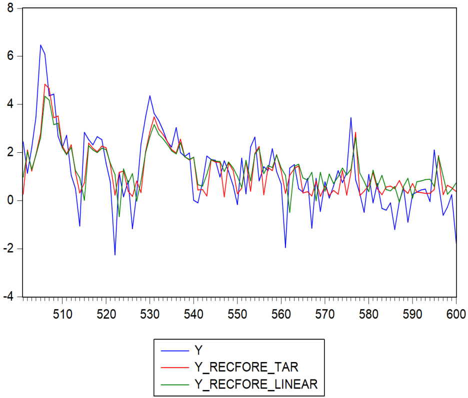
# Forecasting TAR models

- The $h = 1$-step-ahead forecasts are given by:

$$\begin{aligned}\hat{y}_T(1) &= (1 - I(y_T > c^*))(\hat{\phi}_{1,0} + \hat{\phi}_{1,1} y_T) \\ &\quad + I(y_T > c^*)(\hat{\phi}_{2,0} + \hat{\phi}_{2,1} y_T) + \varepsilon_t, \qquad (3)\end{aligned}$$

- RMSFE with TAR model is equal to 1.08 compared to 1.23 with a simple AR(1) linear model

# Forecasts


8
6
4
2
0
-2
-4
510
520
530
540
550
560
570
580
590
600
Y
Y_RECFORE_TAR
Y_RECFORE_LINEAR

# Simulation of a STAR model

- We simulate the following STAR(2,1):

$$y_t = (2 - 0.9 y_{t-1}) + G(y_{t-1}, \gamma, c)(3 - 0.8 y_{t-1}) + \varepsilon_t,$$

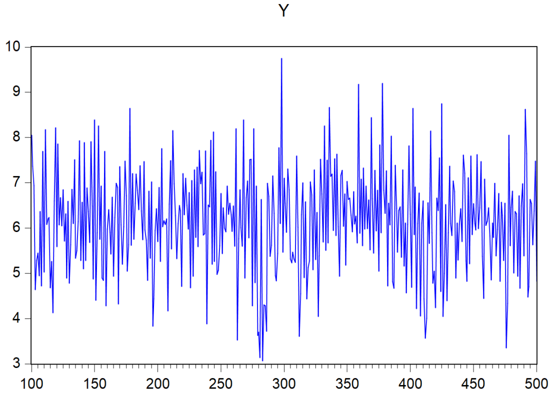
  with unit variance for $\varepsilon_t$ (i.e. $\sigma_\varepsilon^2 = 1$) and

$$G(y_{t-1}, \gamma, c) = \frac{1}{1 + \exp(-\gamma(y_{t-1} - c))}$$

  with $\gamma = 5$ and $c = 6$

- Be careful to the interpretation of parameters with this specification, in low and high regimes

# Simulated series for STAR model



# Testing TAR models

- We implement the procedure by Luukkonen, Saikkonen and Terasvirta (1988)
- We use an F-test to test for $\delta = 0$ in the following augmented model:

$$y_t = \eta_1 + \eta_2 \tilde{x}_t + \delta \tilde{x}_t q_t + u_t,$$

  where $\tilde{x}_t = y_{t-1}$ and $q_t = \gamma/4(y_{t-1} - c)$
- By simplicity we use the true values of $\gamma$ and $c$
- As $\delta$ is a scalar in this specific case, the F-test boils down to a Student test for $\delta$ (see Table)

# Estimated regression for the STAR test

We reject the null of linearity at very low risk $\alpha$

Dependent Variable: Y
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 08/17/17 Time: 16:38
Sample: 100 500
Included observations: 401
Y=ETA(1)+ETA(2)*Y(-1)+DELTA(1)*Y(-1)*TAYLOR

| | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| ETA(1) | 2.809797 | 1.171045 | 2.399393 | 0.0169 |
| ETA(2) | 0.586302 | 0.201058 | 2.916086 | 0.0037 |
| DELTA(1) | -0.055042 | 0.014356 | -3.834067 | 0.0001 |

| | | | |
|---|---|---|---|
| R-squared | 0.060748 | Mean dependent var | 6.150589 |
| Adjusted R-squared | 0.056028 | S.D. dependent var | 1.148882 |
| S.E. of regression | 1.116233 | Akaike info criterion | 3.065250 |
| Sum squared resid | 495.8986 | Schwarz criterion | 3.095130 |
| Log likelihood | -611.5825 | Hannan-Quinn criter. | 3.077081 |
| F-statistic | 12.87068 | Durbin-Watson stat | 1.962244 |
| Prob(F-statistic) | 0.000004 | | |

# Steps for STAR estimation

1. Prepare a grid of $M = 200$ possible values for the threshold $c$ and a grid of $K = 200$ values for the smoothing parameter $\gamma$
2. For each $i = 1, \ldots, M$, $j = 1, \ldots, K$ compute the function:

$$G(y_{t-1}, \gamma_i, c_i) = \frac{1}{1 + \exp(-\gamma_i(y_{t-1} - c_j))}$$

3. Using OLS estimate the following model linear model We simulate the following STAR(2,1):

$$y_t = (\phi_{01} + \phi_{11} y_{t-1}) + G(y_{t-1}, \gamma_i, c_j)(\phi_{02} + \phi_{12} y_{t-1}) + \varepsilon_t,$$

4. For each $i, j$ and resulting estimates $\hat{\phi}$ compute the variance error term $\hat{\sigma}_t^2(\gamma_i, c_j)$
5. Search for $(\gamma^*, c^*) = \underset{(\gamma_i, c_j)_{i,j}}{Arg} \min \hat{\sigma}_t^2(\gamma_i, c_j)$

# Estimation results for STAR

We get $c^* = 6$ and $\gamma^* = 8.6$ and:

Dependent Variable: Y
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 08/17/17 Time: 16:38
Sample: 100 500
Included observations: 401
Y=ALPHA(1)+ALPHA(2)*Y(-1)+ (BETA(1)-BETA(2)*Y(-1)) * G_EST

| | Coefficient | Std. Error | t-Statistic | Prob. |
|---|---|---|---|---|
| ALPHA(1) | 1.970040 | 0.654466 | 3.010149 | 0.0028 |
| ALPHA(2) | 0.910297 | 0.129081 | 7.052157 | 0.0000 |
| BETA(1) | 3.781540 | 0.893495 | 4.232300 | 0.0000 |
| BETA(2) | 0.909919 | 0.149807 | 6.073961 | 0.0000 |

| | | | |
|---|---|---|---|
| R-squared | 0.190065 | Mean dependent var | 6.150589 |
| Adjusted R-squared | 0.183944 | S.D. dependent var | 1.148882 |
| S.E. of regression | 1.037852 | Akaike info criterion | 2.922108 |
| Sum squared resid | 427.6229 | Schwarz criterion | 2.961948 |
| Log likelihood | -581.8826 | Hannan-Quinn criter. | 2.937883 |
| F-statistic | 31.05418 | Durbin-Watson stat | 2.006318 |
| Prob(F-statistic) | 0.000000 | | |

# Steps for STAR estimation

- As an alternative approach, we can use a Maximum Likelihood estimation
- We need to assume the Gaussianity of the error term
- In such a case, the difficulty is find correct starting values for the maximization algorithm
- Especially, the value of $\gamma$ is very difficult to estimate as it is very sensitive to starting values

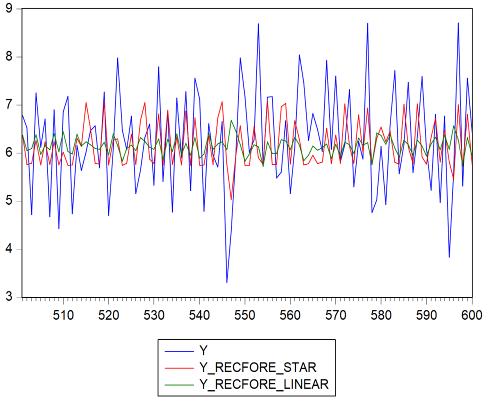
# Forecasting TAR models

- The $h = 1$-step-ahead forecasts are given by:

$$\hat{y}_T(1) = (\hat{\phi}_{01} + \hat{\phi}_{11} y_T) + G(y_T, \gamma^*, c^*)(\hat{\phi}_{02} + \hat{\phi}_{12} y_T)$$

- RMSFE with STAR is equal to 0.95 compared to 1.04 with a simple AR(1) linear model

# Forecasts


9
8
7
6
5
4
3
510
520
530
540
550
560
570
580
590
600
Y
Y_RECFORE_STAR
Y_RECFORE_LINEAR

# Simulation of a MS model

- We simulate the following MS-AR(2) model:

$$y_t = \begin{cases} -0.2 + 1.2y_{t-1} - 0.3y_{t-2} + \varepsilon_t, & \text{if} \quad S_t = 1 \\ 0.2 + 0.4y_{t-1} - 0.3y_{t-2} + \varepsilon_t, & \text{if} \quad S_t = 2 \end{cases}$$

  with unit variance for $\varepsilon_t$ (i.e. $\sigma_\varepsilon^2 = 1$)

- We generate 600 observations, but only the last 500 are considered to avoid dependence to starting values (see Figure)

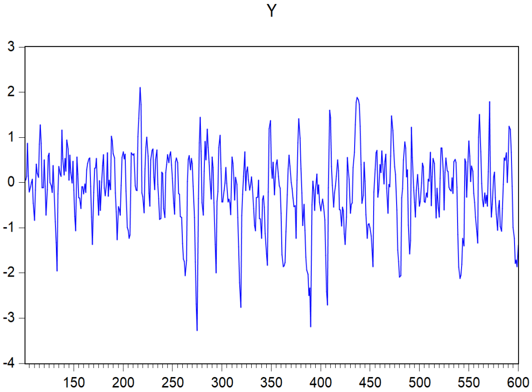
# Simulation of a MS model

- We generate the Markov chain state variable supposed to be here a 2-regime Markov chain with the transition matrix:

$$\eta = \begin{pmatrix} 0.83 & 0.25 \\ 0.17 & 0.75 \end{pmatrix}$$

- Conditionally on the generated state variable $S_t$, we simulate the series with different AR(2) dynamics

# Simulated series for MS-AR(2) model



# Estimated MS-AR(2) model

Dependent Variable: Y
Method: Markov Switching Regression (BFGS / Marquardt steps)
Date: 08/18/17 Time: 11:19
Sample: 101 500
Included observations: 400
Number of states: 2
Initial probabilities obtained from ergodic solution
Standard errors & covariance computed using observed Hessian
Random search: 25 starting values with 10 iterations using 1 standard deviation (rng=kn, seed=412318124)
Convergence achieved after 9 iterations

| Variable | Coefficient | Std. Error | z-Statistic | Prob. |
|---|---|---|---|---|
| | | Regime 1 | | |
| C | 0.167649 | 0.049243 | 3.404489 | 0.0007 |
| AR(1) | 0.535570 | 0.059784 | 8.958425 | 0.0000 |
| | | Regime 2 | | |
| C | -0.126784 | 0.067382 | -1.881567 | 0.0599 |
| AR(1) | 1.391337 | 0.049237 | 28.25784 | 0.0000 |
| | | Common | | |
| AR(2) | -0.403547 | 0.044506 | -9.067239 | 0.0000 |
| LOG(SIGMA) | -0.956207 | 0.048050 | -19.90013 | 0.0000 |
| | | Transition Matrix Parameters | | |
| P11-C | 0.408276 | 0.452953 | 0.901367 | 0.3674 |
| P21-C | -0.603580 | 0.254304 | -2.373456 | 0.0176 |

| | | | |
|---|---|---|---|
| Mean dependent var | -0.141756 | S.D. dependent var | 0.856265 |
| S.E. of regression | 0.570283 | Sum squared resid | 128.1378 |
| Durbin-Watson stat | 2.027593 | Log likelihood | -282.8282 |
| Akaike info criterion | 1.454141 | Schwarz criterion | 1.533970 |
| Hannan-Quinn criter. | 1.485755 | | |

| | | |
|---|---|---|
| 1: Inverted AR Roots | .27+.58i | .27-.58i |
| 2: Inverted AR Roots | .98 | .41 |

# Estimated MS-AR(2) model

Equation: EQ_MARKOV
Date: 08/18/17 Time: 11:19
Transition summary: Constant Markov transition probabilities and expected durations
Sample: 101 500
Included observations: 400

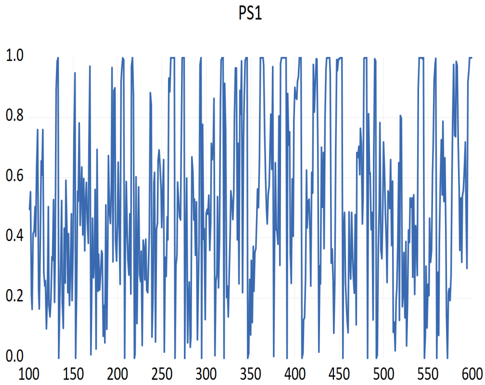
Constant transition probabilities:
P(i, k) = P(s(t) = k | s(t-1) = i)
(row = i / column = j)

| | 1 | 2 |
|---|---|---|
| 1 | 0.600675 | 0.399325 |
| 2 | 0.353525 | 0.646475 |

Constant expected durations:

| | 1 | 2 |
|---|---|---|
| | 2.504223 | 2.828655 |

# Smoothed probability of being in regime 1



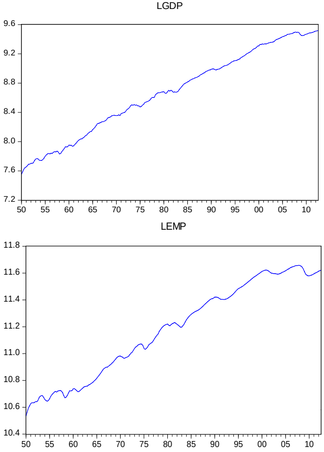
# Forecasting with MS

- Weak gains as regards forecasting accuracy

| Model | RMSFE | DM pval |
|---|---|---|
| AR(2) | 0.555088 | . |
| MSAR | 0.549637 | 0.248488 |

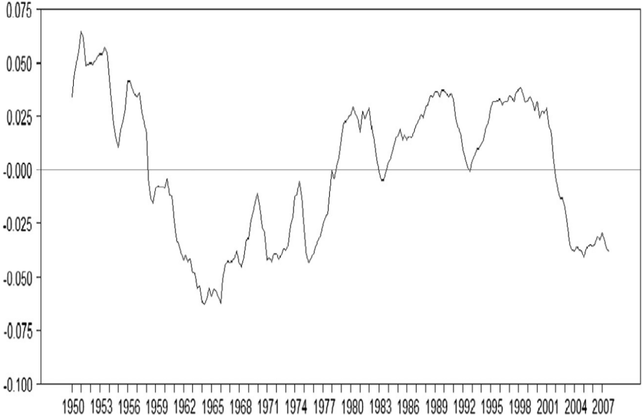
- Results in line with the literature (see Ferrara, Marcellino, Mogliani, IJF, 2015): univariate MS do not perform very well in terms of forecasting accuracy when compared to linear AR models
- MS models are more useful for turning point detection

# US log-GDP and log-Employment (1950q1 - 2012q3)



# Residuals from the long-run relationship

We assume that residuals are $I(0)$ but non-linear



# A Smooth-Transition ECM (STECM)

- The STECM model is given by:

$$\begin{aligned} \Delta lemp_t &= \alpha_0 + \alpha_1 \Delta lemp_{t-1} + \alpha_2 \Delta lgdp_t + \alpha_3 \Delta lgdp_{t-1} + \alpha_4 u_{t-} \\ &+ (\alpha'_0 + \alpha'_1 \Delta lemp_{t-1} + \alpha'_2 \Delta lgdp_t + \alpha'_3 \Delta lgdp_{t-1} \\ &+ \alpha'_4 u_{t-1}) \times G(Z_t, \gamma, c) + \varepsilon_t \end{aligned}$$

- The transition variable $Z_t$ is a proxy for the business cycle, defined by

$$Z_t = \frac{\Delta lgdp_{t-2} + \Delta lgdp_{t-3}}{2}$$

# Estimated STECM

**Table 1**
Estimates for the non-linear error-correction model.

| | Regime 1 | Regime 2 |
|---|---|---|
| Constant | $-0.0071^{***}$ | $0.0067^{***}$ |
| $\Delta LEMP_{t-1}$ | 0.2149 | $0.3741^{***}$ |
| $\Delta LGDP_t$ | $0.6439^{***}$ | $-0.3361^{***}$ |
| $\Delta LGDP_{t-1}$ | $0.1932^{*}$ | $-0.1976$ |
| $ECT_{t-1}$ | $-0.0193^{**}$ | 0.0070 |
| $\hat{\gamma}$ | 5.1624 | |
| $\hat{c}$ | 0.0001 | |

In order to obtain the relevant coefficients in Regime 2, add the Regime 2 coefficient to the corresponding Regime 1 coefficient.
ECT denotes error-correction term.
*** Significant at the 1% level.
** Significant at the 5% level.
* Significant at the 10% level.

# Estimated STECM: Proba of being in the low regime



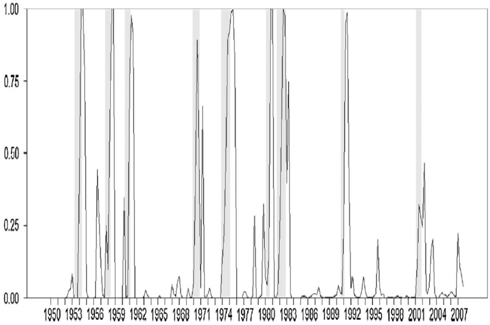



**Fig. 3.** Estimated transition function and US recessions until 2007Q4. Sources: authors' calculations, and NBER for the dating of recessions. The gray bands represent US recessions.

# Estimated STECM: Conditional forecasts

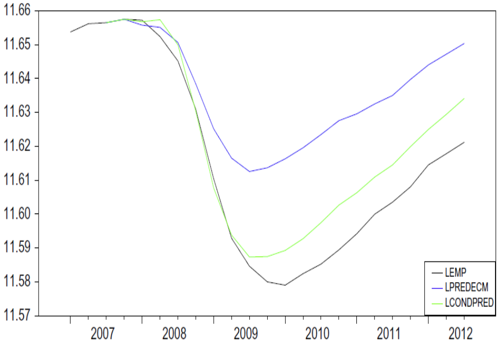



**Fig. 6.** Conditional forecasts of log-employment stemming from linear and non-linear error-correction models. Note: Conditional forecasts stemming from the linear ECM (blue) and the non-linear ECM (green). Observed values are presented in the dark line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

# Application: Building an economic cycle indicator

- In Darne and Ferrara (2011, Oxford Bulletin of Eco and Stat) we put forward a cyclical indicator to assess acceleration and deceleration phases in the euro area using MS models
- We define what we mean by *acceleration cyle* and we provide a dating chronology
- We estimate various MS models on opinion survey series in order to build a real-time indicator for the peaks and troughs of this cycle

# Application: Defining the acceleration cycle

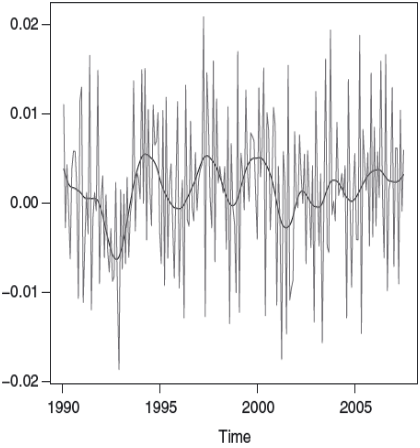



Figure 1. Monthly IPI growth rate and smoothed IPI growth rate obtained by removing high frequencies with a low-pass filter (Jan. 1990–Sep. 2007)

# Application: Dating the acceleration cycle

TABLE 1

*Turning points chronologies for acceleration cycles for GDP and IPI, as well as the proposed dating (third column)*

| | *GDP* | *IPI* | *Dating* | *Duration* | *Amplitude* | *Excess* |
|---|---|---|---|---|---|---|
| Peak | 1987 Q2 | NA | 1987 M5 | | | |
| Trough | 1993 Q1 | 1992 M11 | 1993 M1 | 68 | 2.4 | −0.18 |
| Peak | 1994 Q1 | 1994 M3 | 1994 M3 | 14 | 1.6 | 0.10 |
| Trough | 1996 Q1 | 1996 M2 | 1996 M2 | 21 | 0.9 | −0.04 |
| Peak | 1997 Q2 | 1997 M4 | 1997 M4 | 16 | 1.2 | −0.14 |
| Trough | 1998 Q4 | 1998 M10 | 1998 M10 | 15 | 1.0 | 0.07 |
| Peak | 1999 Q3 | 1999 M9 | 1999 M9 | 11 | 1.0 | 0.03 |
| Trough | 2001 Q3 | 2001 M9 | 2001 M9 | 24 | 1.2 | −0.09 |
| Peak | 2002 Q2 | 2002 M4 | 2002 M4 | 7 | 0.4 | −0.05 |
| Trough | 2003 Q2 | 2003 M4 | 2003 M4 | 12 | 0.4 | 0.04 |
| Peak | 2004 Q1 | 2003 M10 | 2004 M1 | 9 | 0.6 | 0.17 |
| Trough | 2004 Q4 | 2004 M11 | 2004 M11 | 10 | 0.3 | 0.04 |
| Peak | 2006 Q2 | 2006 M4 | 2006 M4 | 17 | 0.7 | −0.05 |

*Notes:* Durations are in months and amplitudes are in points of percentage. NA stands for non-available information.

# Application: Data

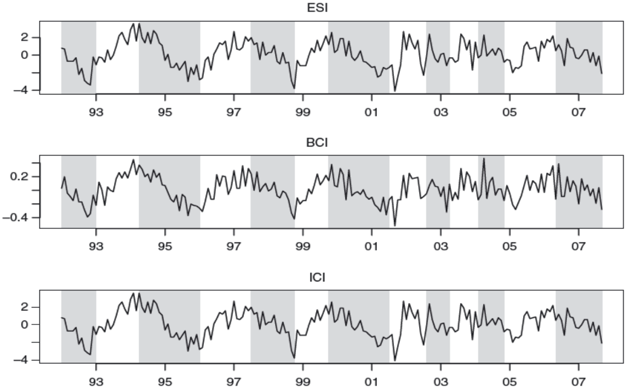



Figure 2. ESI, BCI and ICI (in first differences) and phases of deceleration in the euro area (shaded areas) over the period January 1992–September 2007

# Application: Estimation of MS models

Estimation of simple MS models with 2 regimes (low: $S_t = 1$ and high: $S_t = 2$) and without AR lags:

TABLE 9

*Parameter estimates of the various multivariate MS models over the period Jan. 1992–Sep. 2007 (except for PMIs since 1998 only). Durations D of each regime are expressed in months*

| | $\mu(S_t=1)$ | $\mu(S_t=2)$ | $\sigma_\varepsilon$ | $p_{11}$ | $p_{22}$ | $D(S_t=1)$ | $D(S_t=2)$ |
|---|---|---|---|---|---|---|---|
| Confidence Indexes | | | | | | | |
| ESI | −1.01 | 1.77 | 1.57 | 0.93 | 0.88 | 14 | 8 |
| BCI | −0.12 | 0.21 | 0.17 | 0.93 | 0.87 | 14 | 8 |
| ICI | −0.89 | 1.90 | 1.32 | 0.94 | 0.86 | 16 | 7 |

# Application: Constructing the acceleration cycle index

- Starting from estimated filtered probabilies of being in regime 1 or 2 we compute the index as:

$$I_t = P(S_t = 2|F_t) - P(S_t = 1|F_t)$$

- Thus $I_t \in [-1, +1]$.
- When $I_t$ is close to 1, the economy accelerates and when $I_t$ is close to -1 the economy decelerates
- When $I_t \sim 0$, we interpret this as a stagnating growth (at any level)

# Application: Empirical results

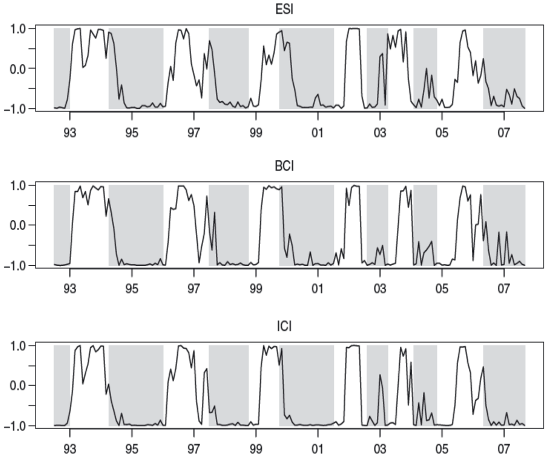



Figure 5. Probabilistic indicators stemming from an MS model applied to ESI (top), BCI (middle) and ICI (bottom) (Jul. 1992–Sep. 2007)