Testing for observation-dependent regime switching in mixture autoregressive models

Mika Meitz

University of Helsinki

Joint work with: Pentti Saikkonen (University of Helsinki)

Seminar at Keio University, Tokyo

28 May 2019

Outline of the Talk

- Regime-Switching & Mixtures
  - Motivation
  - On Different Types of Regime-Switching Autoregressive (AR) models
    - 'Observation-dependent regime switching'
  - Mixture AR Models
    - Some Example Models
- Testing for Regime Switching ('One Regime vs. Two Regimes')
  - Some Recent Literature
  - Our Set-up – Likelihood Ratio Test
  - 3 Challenges We Face
  - Our Results
    - How we deal with the 3 challenges
    - Asymptotic Distribution
    - Assumptions & Proofs
    - Examples
  - Simulating the asymptotic null distribution / bootstrap
- Conclusions

Motivation

- Regime-switching models are by now rather standard in macroeconomics
  - Expansion / Recession periods of the economy
  - Changes in policy (monetary policy, hawks / doves)
  - Technological innovations leading to higher growth periods
  - Sudden loss of confidence in a country, leading to currency crisis
  - ...
- A fundamental question: Is the regime-switching necessary?
  - We consider statistical tests of “1 regime vs. 2 regimes”
- This is a hard problem and previous literature is scarce
  - The testing problem is non-standard in various ways
  - We focus on simple univariate AR models

Regime-switching AR models – Definition

- General Regime-Switching AR Model (two components, $p = 1$):

$$y_t = s_{t,1}\Big(\phi_0 + \phi_1 y_{t-1} + \sigma_1\varepsilon_t\Big) + s_{t,2}\Big(\varphi_0 + \varphi_1 y_{t-1} + \sigma_2\varepsilon_t\Big)$$

  - y_t is the observed univariate time series of interest
  - $\mathcal{F}_{t-1}$ denotes the σ–algebra generated by past y's
  - $\varepsilon_t \sim IID\, N(0,1)$ (for example), ε_t independent of $\mathcal{F}_{t-1}$
  - $\boldsymbol{s}_t = (s_{t,1}, s_{t,2})$ are unobserved random vectors such that
    - For each t, one component of $\boldsymbol{s}_t$ takes value 1 and the other value 0 with conditional probabilities

$$p_{ijt} = P(s_{t,j} = 1 \mid \mathcal{F}_{t-1}, s_{t-1,i} = 1) \qquad i,j = 1,2$$

    - p_{ijt}: Transition probability that, at time t, regime i will be followed by regime j
  - conditional on $\{\mathcal{F}_{t-1}, \boldsymbol{s}_{t-1}\}$, the $\boldsymbol{s}_t$ and ε_t are independent
- The p_{ijt}'s are (conditional) probabilities that determine which one of the 2 AR components generates the next observation y_t

Regime-switching AR models – 4 Types

- General Regime-Switching AR Model (univariate, two components, $p = 1$):

$$y_t = s_{t,1}\Big(\phi_0 + \phi_1 y_{t-1} + \sigma_1 \varepsilon_t\Big) + s_{t,2}\Big(\varphi_0 + \varphi_1 y_{t-1} + \sigma_2 \varepsilon_t\Big)$$

Regime-switching probabilities $p_{ijt} = P(s_{t,j} = 1 \mid \mathcal{F}_{t-1}, s_{t-1,i} = 1)$. Types:

1) 'Classic' Markov switching AR – e.g. Hamilton (1989 *Econometrica*)
   - dependence on past regime only: $p_{ijt} = P(s_{t,j} = 1 \mid s_{t-1,i} = 1) = p_{ij}$
2) Time-inhomogeneous Markov switching AR – e.g. Filardo (1994 *JBES*)
   - dependence on past regime & past observations: p_{ijt} 'doesn't simplify'
3) 'Basic' Mixture AR – e.g. Wong & Li (2000 *JRSS-B*)
   - no dependence on past regime / observations: $p_{ijt} = P(s_{t,j} = 1) = p_j$
4) Mixture AR with observation-dependent regime switching
   - dependence on past observations only: $p_{ijt} = P(s_{t,j} = 1 \mid \mathcal{F}_{t-1}) = p_{jt}$
   - e.g. Wong & Li (2001 *Biometrika*),
     Kalliovirta, Meitz & Saikkonen (2015 *JTSA*, 2016 *JoE*)
     (also 'mixture-of-experts models' in neural networks / machine learning literature)

Observation-dependent reg-switching – Motivation

- Why should regime switching probabilities depend on observed data?
  - Allows one to associate changes in regime to observable economic variables
  - Makes interpretation of regime switches easier
- Hamilton's (2016) Handbook of Macroeconomics chapter "Macroeconomic regimes and regime shifts" begins with

  "*Many economic time series exhibit dramatic <u>breaks associated with events</u> such as economic recessions, financial panics, and currency crises. <u>Such changes in regime</u> may arise from tipping points or other nonlinear dynamics and are core to some of the most important questions in macroeconomics.*"
- This paper: Simple univariate AR's – complicated enough
  - Extensions? Regime-switching reduced-form VAR's or Structural VAR's.

Mixture AR (MAR) Models

- General Mixture AR Model (univariate, two components, $p = 1$):

$$y_t = s_t\Big(\phi_0 + \phi_1 y_{t-1} + \sigma_1\varepsilon_t\Big) + (1 - s_t)\Big(\varphi_0 + \varphi_1 y_{t-1} + \sigma_2\varepsilon_t\Big)$$

  - $\varepsilon_t \sim IID\, N(0, 1)$, ε_t independent of $\mathcal{F}_{t-1}$
  - s_t (unobserved) Bernoulli (1 / 0) random variables with

$$P(s_t = 1 \mid \mathcal{F}_{t-1}) = \alpha_t \qquad (\alpha_t \in (0, 1) \text{ function of } y_{t-j},\, j > 0)$$

  - conditional on $\mathcal{F}_{t-1}$, the s_t and ε_t are independent

- "y_t generated by 2 AR components with probabilities α_t and $1 - \alpha_t$"
- Different models $\longleftrightarrow$ Different specifications for α_t
- The 'basic' mixture AR model – Wong & Li (2000 *JRSS-B*)

$$\alpha_t = \alpha \qquad \text{with } \alpha \in (0, 1) \text{ a constant}$$

Mixture AR (MAR) Models

- Logistic MAR (LMAR) model – Wong & Li (2001 *Biometrika*)

$$\alpha_t = \alpha_t(\alpha_0, \alpha_1) = \frac{\exp(\alpha_0 + \alpha_1 y_{t-1})}{1 + \exp(\alpha_0 + \alpha_1 y_{t-1})} \qquad (\alpha_0, \alpha_1 \text{ parameters})$$

- Gaussian MAR (GMAR) model
  – Kalliovirta, Meitz & Saikkonen (2015 *JTSA*, 2016 *JoE*)

$$\alpha_t = \alpha_t(\alpha, \phi, \varphi) = \frac{\alpha \, \mathsf{n}_1(y_{t-1}; \phi)}{\underbrace{\alpha \, \mathsf{n}_1(y_{t-1}; \phi)}_{\substack{\text{density of } N\left(\frac{\phi_0}{1-\phi_1}, \frac{\sigma_1^2}{1-\phi_1^2}\right) \\ \text{evaluated at } y_{t-1}}} + (1-\alpha) \underbrace{\mathsf{n}_1(y_{t-1}; \varphi)}_{\substack{\text{density of } N\left(\frac{\varphi_0}{1-\varphi_1}, \frac{\sigma_2^2}{1-\varphi_1^2}\right) \\ \text{evaluated at } y_{t-1}}}}$$

$\alpha \in (0,1)$ a parameter
$\phi = (\phi_0, \phi_1, \sigma_1^2)$
$\varphi = (\varphi_0, \varphi_1, \sigma_2^2)$

  - This choice of α_t leads to nice properties

Testing for Regime Switching

- In all regime switching models, a crucial question is:

One Regime or Two Regimes ?

- What do we do in this paper?

Study the appropriate Likelihood Ratio (LR) test in mixture AR models with observation-dependent regime switching.

Testing for Regime Switching – Literature

- Testing for Markov-switching type regime switching
  - Hansen (1992 *JAE*), Garcia (1998 *IER*): Early discussions of LR test
  - Cho & White (2007 *Econometrica*):
    - LR test for a mixture model to test for Markov-switching type regime switching
  - Carrasco, Hu & Ploberger (2014 *Econometrica*):
    - 'information matrix type test', asymptotically optimal against Markov switching
  - Recent working papers: Qu & Zhuo (2017), Kasahara & Shimotsu (2017):
    - LR test for regime switching in Markov switching models.
- Testing for mixture type regime switching
  - Extensive literature for case of independent observations without regressors
  - Zhu & Zhang (2004 *JRSS-B*, 2006 *JMVA*), Kasahara & Shimotsu (2015 *JASA*):
    - with regressors, no dependent data. LR tests for regime switching (+ other things)
- Testing for observation-dependent regime switching
  - Previous literature almost non-existent
  - Jeffries (1998 UMaryland PhD thesis): LR test in a specific first-order model
  - Shen & He (2015 *JASA*): An 'expectation maximization test'

Testing: The basic set-up (here with $p = 1$)

- Model under alternative a general Mixture AR model:

$$y_t = s_t\Big(\phi_0 + \phi_1 y_{t-1} + \sigma_1\varepsilon_t\Big) + (1 - s_t)\Big(\varphi_0 + \varphi_1 y_{t-1} + \sigma_2\varepsilon_t\Big)$$
$$P(s_t = 1 \mid \mathcal{F}_{t-1}) = \alpha_t(\alpha, \phi, \varphi)$$
$$\alpha_t(\alpha, \phi, \varphi) \in (0, 1) \text{ and } \sigma(y_{t-1})\text{–measurable}$$
$$\varepsilon_t \sim IID\, N(0, 1)$$
$$\text{Parameters: } \alpha, \;\; \phi = (\phi_0, \phi_1, \sigma_1^2), \;\; \varphi = (\varphi_0, \varphi_1, \sigma_2^2)$$

- Model under null: Gaussian AR Model
- Null to be tested: $\phi = \varphi$
- Test to be used: Likelihood ratio (LR) test

Testing: The log-likelihood

- Basic equation defining Mixture AR model:

$$y_t = s_t\Big(\phi_0 + \phi_1 y_{t-1} + \sigma_1\varepsilon_t\Big) + (1 - s_t)\Big(\varphi_0 + \varphi_1 y_{t-1} + \sigma_2\varepsilon_t\Big)$$

- Conditional density function of y_t given its past:

$$f(y_t \mid \mathcal{F}_{t-1}) = \alpha_t f_t(\phi) + (1-\alpha_t) f_t(\varphi)$$

density of $N(\phi_0 + \phi_1 y_{t-1}, \sigma_1^2)$ evaluated at y_t

density of $N(\varphi_0 + \varphi_1 y_{t-1}, \sigma_2^2)$ evaluated at y_t

  - mixture of two normal densities with mixing weights α_t and $1-\alpha_t$
- The (per observation conditional) log-likelihood:

$$l_t(\alpha, \phi, \varphi) = \log[\alpha_t f_t(\phi) + (1-\alpha_t) f_t(\varphi)]$$

where $\alpha_t = \alpha_t(\alpha, \phi, \varphi)$ depends on parameters α, ϕ, φ and y_{t-1}

Challenge 1. Unidentified parameters

- If $\phi = \varphi$, the parameter α is not identified
  - Two AR components identical $\Rightarrow$ model is the same regardless of α
  - in terms of the likelihood:

$$l_t(\alpha, \phi, \phi) = \log[\alpha_t f_t(\phi) + (1 - \alpha_t) f_t(\phi)] = \underbrace{\log[f_t(\phi)]}_{\text{does not depend on } \alpha!}$$

  so $\alpha_t = \alpha_t(\alpha, \phi, \varphi)$ and thus α vanishes

- This is the classical 'unidentified parameters under the null' problem
  - Davies (1977, 1987 *Biometrika*), Hansen (1996 *Econometrica*)
  - Solution: Use a sup LR test with an appropriate asymptotic distribution

Challenge 2. Singular Information Matrix

- Fisher Information Matrix is Singular (but ϕ, φ locally identifiable)
- Illustrate with MAR model: $\alpha_t = \alpha$
- Scores of ϕ & φ are linearly dependent under the null ($\phi = \varphi = \phi^*$):

$$\nabla_\phi l_t(\alpha, \phi^*, \phi^*) = \alpha \frac{\nabla f_t(\phi^*)}{f_t(\phi^*)} \text{ and } \nabla_\varphi l_t(\alpha, \phi^*, \phi^*) = (1-\alpha) \frac{\nabla f_t(\phi^*)}{f_t(\phi^*)}$$

- Fisher information matrix is singular! Trouble!
- Moreover, due to properties of the normal density, linear dependencies also among higher-order derivatives. More trouble!
  - for example $\frac{\nabla_{\sigma^2} f_t(\phi^*)}{f_t(\phi^*)}$ and $\frac{\nabla^2_{\phi_0} f_t(\phi^*)}{f_t(\phi^*)}$ are linearly dependent

Challenge 2. Singular Information Matrix

- Solution? Carefully constructed reparameterization(s) inspired by Rotnitzky, Cox, Bottai & Robins (2000 *Bernoulli*)
  - also used by Kasahara & Shimotsu (2015 *JASA*)
- Intuitive idea?
  - If scores of original parameters are linearly dependent, reparameterize in such a way that the resulting new scores are orthogonal.
  - Due to rank deficiency of the information matrix, some new scores are now necessarily zero.
  - In a Taylor expansion of the reparameterized log-likelihood, the second derivative term now provides the first (nontrivial) local approximation for those (reparameterized) parameters that have zero scores
  - LR test derivation based on a quadratic approximation of the log-likelihood in terms of the reparameterized parameters ($=$ higher-order expansion in terms of original parameters)
  - Rates of convergence vary, depending on degree of first nonzero derivative

Challenge 3. Parameters on the Boundary

- The (reparameterized) parameter vector under the null hypothesis lies on the boundary of the permitted parameter space.
- Furthermore, both the (reparameterized) parameter and its parameter space depend on the unidentified nuisance parameters
- Parameters on the boundary have been considered, e.g., in
  - Andrews (1999, 2001 *Econometrica*)
  - Silvapulle & Sen (2005 *Wiley book*)
- Solution? Approximate the parameter space by a cone

Dealing with Challenges 1–3

- Dealing with Challenges 1 (unidentification) and 3 (boundary):
  - Andrews (1999, 2001 *Econometrica*)
- Our problem: Adding Challenge 2 (singularity)
- Broadly speaking, we analyze
  - the necessary reparameterizations to deal with the singularity
  - a quadratic approximation in terms of the reparameterized parameters
- Our results slightly adapt/extend results in Andrews (1999, 2001 *Econometrica*) and Zhu & Zhang (2006 *JMVA*)

Assumptions & Proofs

- Assumptions?
  - $\alpha_t(\alpha, \phi, \varphi) \in (0, 1)$, $\sigma(y_{t-1}, \ldots, y_{t-p})$–measurable, α vector-valued, $\alpha_t(\alpha, \cdot, \cdot)$ sufficiently differentiable
  - MLE in the Mixture AR model consistent (uniformly over α)
  - Suitable and 'smooth' reparameterization can be found to handle singularities (case-by-case)
  - Quadratic approximation in the reparameterized parameters satisfies conditions similar to Andrews (2001 *Econometrica*)
  - All assumptions verified for LMAR and (a version of) GMAR
- Proofs?
  - Slight adaptations of Andrews (1999, 2001 *Econometrica*) and Zhu & Zhang (2006 *JMVA*)
    - Reparameterized parameter and its parameter space depend on α
    - This parameter space 'uniformly (over α) approximated' by a non-convex cone (not depending on α)

Asymptotic Distribution

- Asymptotic null distribution of sup-LR test $LR = \sup_\alpha LR(\alpha)$?
- Of the Andrews 'sup-of-chi-bar-square' – type:

$$\sup_\alpha \left\{ Z_\alpha' V_\alpha^{-1} Z_\alpha - \inf_{\boldsymbol{\lambda} \in \Lambda} \left\{ (\boldsymbol{\lambda} - Z_\alpha)' V_\alpha^{-1} (\boldsymbol{\lambda} - Z_\alpha) \right\} \right\}$$

$$\text{with } Z_\alpha \sim N(0, V_\alpha) \text{ and } \Lambda \text{ a cone}$$

- Distribution application-specific, cannot be tabulated.
  - Simulating the asymptotic null distribution / bootstrap – in a moment.

Examples: LMAR, GMAR

- General theory formulated using high-level assumptions
- All details for two example models
  - LMAR model of Wong & Li (2001 *Biometrika*)
  - GMAR model of Kalliovirta, Meitz & Saikkonen (2015 *JTSA*)
    - with the restriction $\phi_0 = \varphi_0$
- The two examples are rather different:

	LMAR	GMAR
Unidentified parameters?	✓	✓
...that enter the asymptotic distribution of LR test?	✓	
Parameters on boundary?		✓
Singularities?		✓
Degree of Taylor expansion?	2	4

Simulating the asymptotic distribution / bootstrap

The asymptotic null distribution is not simple – two options:

- Simulating the asymptotic null distribution
  - Simulating asymptotically exact draws from the asymptotic null distribution is easy (details omitted here)
    $\rightarrow$ p–values or critical values of the asymptotic null distribution
  - Used by Hansen (1996 *Econometrica*) and Andrews (2001 *Econometrica*)
  - Computationally rather easy
    - Requires computer-intensive but 'simple' calculations, takes only a few seconds
- Bootstrap
  - Based on simulations, restricted parametric bootstrap works very well
  - Computationally quite heavy
    - Requires repeated estimation of the mixture model under the alternative

Next slide: Some Monte Carlo simulations (more in the paper)

Simulation: Size & Power

- Empirical rejection frequencies of tests against LMAR and GMAR.
- DGP's 1–2 (size), AR(1): $y_t = 0.6\, y_{t-1} + \varepsilon_t$ and $y_t = 0.9\, y_{t-1} + \varepsilon_t$
- DGP 3 (power), GMAR(1): $y_t = s_t(0.2\, y_{t-1} + \varepsilon_t) + (1 - s_t)(0.8\, y_{t-1} + \varepsilon_t)$
- DGP 4 (power), LMAR(1): $y_t = s_t(0.5 y_{t-1} + \varepsilon_t) + (1 - s_t)(0.5 y_{t-1} + \sqrt{3}\varepsilon_t)$
- always $\varepsilon_t \sim N(0,1)$; details of $P(s_t = 1 \mid \mathcal{F}_{t-1}) = \alpha_t$ omitted for brevity

		Asymptotic null distribution				Parametric bootstrap			
		LMAR LRT		GMAR LRT		LMAR LRT		GMAR LRT	
DGP	T	10%	5%	10%	5%	10%	5%	10%	5%
1	250	0.12	0.07	0.12	0.06	0.10	0.05	0.11	0.05
	500	0.09	0.05	0.11	0.06	0.08	0.04	0.10	0.05
2	250	0.11	0.05	0.13	0.07	0.11	0.05	0.10	0.05
	500	0.10	0.05	0.12	0.06	0.11	0.05	0.11	0.06
3	250	0.24	0.16	0.62	0.49	0.22	0.13	0.57	0.42
	500	0.32	0.22	0.84	0.76	0.32	0.21	0.82	0.71
4	250	0.89	0.83	0.62	0.49	0.88	0.79	0.58	0.46
	500	0.99	0.98	0.82	0.74	0.99	0.98	0.83	0.72

Conclusions

- We analyze the LR test for regime switching in mixture AR models with observation dependent regime switching
  - Previous literature minimal
- We present results that cover various types of observation-dependence
  - Examples: LMAR, GMAR
- Challenges in obtaining the asymptotic distribution
  - unidentification, singularity, boundary
  - proofs adapt existing results
- Simulations show size & power properties are good

Thank you!