Bayesian Shrinkage Estimation of Logistic Smooth Transition Autoregressions

Mario Giacomazzo

November 6, 2019

Background: Regime-Switching Time Series Models

- Purpose of Regime-Switching Time Series Models
  - Regime-Dependent Dynamics
  - Understand Volatility Changes
  - Alternative Approach to Handling Nonstationary and Cyclical Phenomenon
  - State-specific Forecasting
  - Characterization and Classification

Background: Regime-Switching Time Series Models

- Popularized for Economic and Financial Time Series
  - Structural Breaks Between Periods of Recession and Expansion
  - Unemployment Rates (Montgomery et al., 1998; Koop and Potter, 1999; Deschamps, 2008)
  - Gross Domestic Product (Teräsvirta, 1995)
  - Stock Prices (Li and Lam, 1995)
  - Agricultural Prices (Zeng et al., 2011)
  - Extensive Study on 215 Macroeconomic Time Series (Stock and Watson, 1998)
- Applications in Other Areas
  - Epidemiology: Epidemic vs Nonepidemic States (Hsin-Min Lu et al., 2010)
  - Psychology: Mood Changes Under Bipolar Disorder (Hamaker et al., 2010)
  - Traffic Management: Free Flow vs Congested States (Kamarianakis et al., 2010)
  - Ecology: Temperature Changes Due to Climate Changes (Battaglia and Protopapas, 2012)

Motivation

- Model Complexity
  - Determined by Order Parameters
    - Autoregressive (AR)
    - Moving Average (MA)
    - Distributed Lag (DL)
    - Regime-specific Model Orders
  - Frequentist Approach: Select Based on Information Criteria or Forecasting Performance
  - Bayesian Approach: Reversible jump Markov Chain Monte Carlo (RJMCMC) (Green, 1995)
    - AR(P) (Troughton and Godsill, 1997)
    - TAR(P) (Campbell, 2004; Nieto et al., 2013)
    - STAR(P) (Lopes and Salazar, 2006)

Motivation

- Model Complexity (Cont.)
  - Problem
    - Difficult When Multiple Order Parameters Must Be Chosen
    - Often Leads to Inflexible Representations
    - Overfitting Can Still Occur
  - Solution
    - Intentionally Fix Orders to Be Large
    - Restructure Time Series Model as a High Dimensional Linear Regression
    - Apply Penalized Estimation Methods Aimed at Sparse Models

LSTAR Model

- Gaussian LSTAR(P) Model With 2-Regimes
  - Given autoregressive order P, let $\boldsymbol{x}_t' = [y_{t-1}, y_{t-2}, \cdots, y_{t-P}]$, $\boldsymbol{\alpha}' = [\alpha_1, \cdots, \alpha_P]$, and $\boldsymbol{\beta}' = [\beta_1, \cdots, \beta_P]$.

$$y_t = (\mu_\alpha + \boldsymbol{x}_t'\boldsymbol{\alpha})(1 - G(z_t)) + (\mu_\beta + \boldsymbol{x}_t'\boldsymbol{\beta})G(z_t) + \epsilon_t$$

    where $\epsilon_t \sim$ i.i.d. $N(0, \sigma^2)$ and $G(z_t) : \mathbb{R} \to \mathbb{G} \subseteq [0, 1]$.
  - For LSTAR, consider transition function $G(\cdot)$ such that

$$G(z_t, \gamma^*, \delta) = \{1 + \exp[-(\gamma^*/s_Z)(z_t - \delta)]\}^{-1}$$

    - Threshold Parameter δ Determines When Transition Occurs
    - Slope Parameter $\gamma = \gamma^*/s_Z$ Determines the Rate of Transition
    - As $\gamma \to \infty$, $G(z_t, \gamma^*, \delta)$ Becomes a Step Function

LSTAR Model

- Primary Goals
  - Establish the Efficacy of Bayesian Shrinkage Estimation applied to LSTAR
  - Modify Bayesian Shrinkage Priors to Handle Regime-specific Sparsity
  - Allow for Composite Transition Variable to Be Estimated Using Dirichlet Prior

LSTAR Model

- Prior Distributions
  - TAR(Geweke and Terui, 1993; Chen and Lee, 1995; Koop and Potter, 1999), STAR(Lubrano, 2000; Lopes and Salazar, 2006; Livingston Jr. and Nur, 2017)
  - $\mu_\alpha \sim \mathcal{N}(\cdot,\cdot)$ and $\mu_\beta \sim \mathcal{N}(\cdot,\cdot)$
  - $\sigma^2 \sim \mathcal{IG}(\cdot,\cdot)$
  - $\gamma^* \sim \mathcal{LN}(\cdot,\cdot)$
  - $\delta \sim \mathrm{U}[q_Z(0.15), q_Z(0.85)]$ where $q_Z(\cdot)$ is the empirical quantile function
  - Bayesian Global-Local Shrinkage Priors for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$

$$\alpha_k, \beta_k | \lambda_k^2, \lambda^2, \sigma^2 \sim N(0, \sigma^2 \lambda^2 \lambda_k^2)$$

$$\lambda_k^2 \sim \pi_{Local}(.) \text{ and } \lambda^2 \sim \pi_{Global}(.)$$

    - LASSO (Park and Casella, 2008; Hans, 2009; Schmidt and Makalic, 2013)
    - Regime-Specific LASSO (Separate Global Tuning Parameters)
    - Variable Selection LASSO (Lykou and Ntzoufras, 2013)
    - Horseshoe (Carvalho et al., 2010; Makalic and Schmidt, 2016)

LSTAR Model

- Transition Variable
  - Change-Point Option: $z_t = t$
  - Exogenous Option: $z_t = x_{t-d}$
  - Endogenous Option: $z_t = y_{t-d}$ (Self-Exciting)
  - Let $\boldsymbol{d}_t' = [y_{t-1}, y_{t-2}, \cdots, y_{t-d_{max}}]$ and $\boldsymbol{\phi}_t' = [\phi_1, \phi_2, \cdots, \phi_{d_{max}}]$. Reparameterize transition variable $z_t = \boldsymbol{\phi}'\boldsymbol{d}_t$.

$$\phi \sim Dir\left(\left[\frac{1}{d_{max}}, \frac{1}{d_{max}}, \cdots, \frac{1}{d_{max}}\right]'\right)$$

    Now, z_t is a weighted average of all considered transition variables. Because the weights are constrained to sum to 1, the prior for δ does not require modification.

  - Advantages
    - Allows for a composite transition variable
    - Estimates a more encompassing LSTAR model.

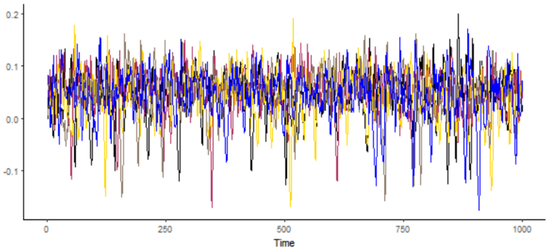
Simulation 1

Well-Behaved LSTAR(2) Used in Lopes and Salazar (2006)

$$
\begin{aligned}
y_t &= (1.8y_{t-1} - 1.06y_{t-2})[1 - G(y_{t-2})] \\
&\quad + (0.02 + 0.9y_{t-1} - 0.265y_{t-2})[G(y_{t-2})] + \epsilon_t \\
&\quad \text{where: } G(y_{t-2}) = \left\{1 + \exp\left[-100(y_{t-2} - 0.02)\right]\right\}^{-1} \\
&\quad \text{and } \epsilon_t \sim \text{i.i.d. } N(0, 0.02^2).
\end{aligned}
$$

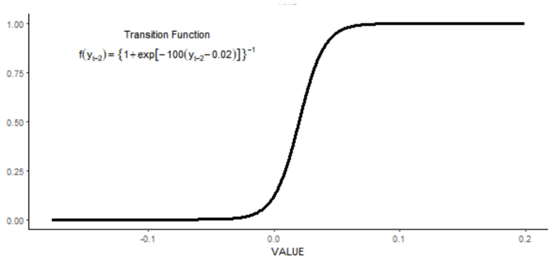
Simulation 1

Figure: Ten Random Replications



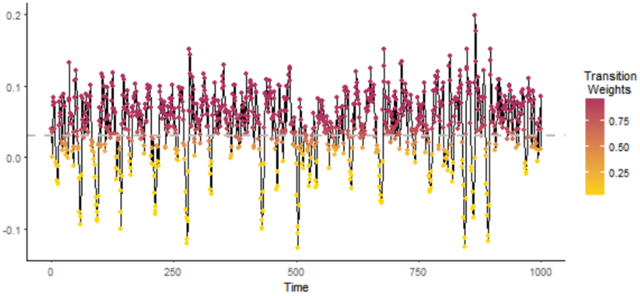
Simulation 1

Figure: Transition Function



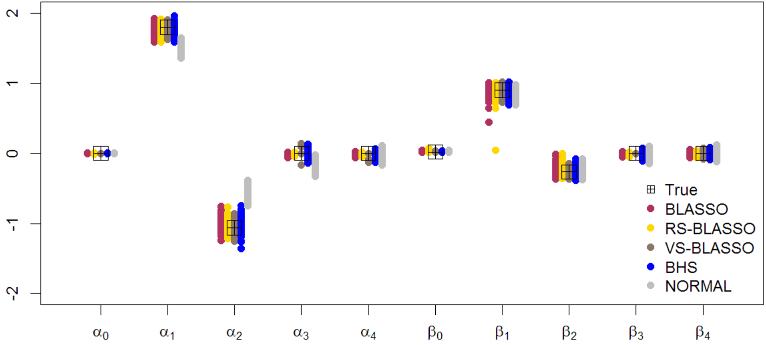
Simulation 1

Figure: Illustration of Regime-switching Behavior



Simulation 1

Figure: Posterior Estimates from 100 Replications



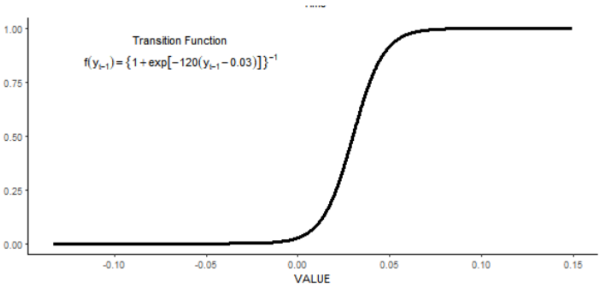
Simulation 2

LSTAR(2) With Regime-Specific Sparsity

$$
\begin{aligned}
y_t &= (-0.7y_{t-3})[1 - G(y_{t-1})] \\
&\quad + (0.06 + 0.4y_{t-1} - 0.35y_{t-2} + 0.2y_{t-3})[G(y_{t-1})] + \epsilon_t \\
&\text{where: } G(y_{t-1}) = \Big\{ 1 + \exp\big[- 120(y_{t-1} - 0.03) \big] \Big\}^{-1} \\
&\text{and } \epsilon_t \sim \text{i.i.d. } N(0, 0.02^2)
\end{aligned}
$$

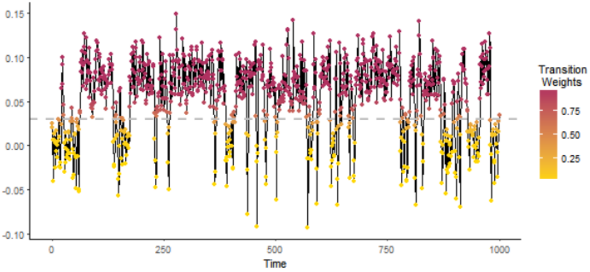
Simulation 2

Figure: Transition Function



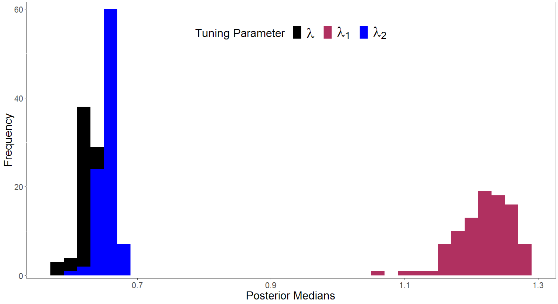
Simulation 2

Figure: Illustration of Regime-switching Behavior



Simulation 2

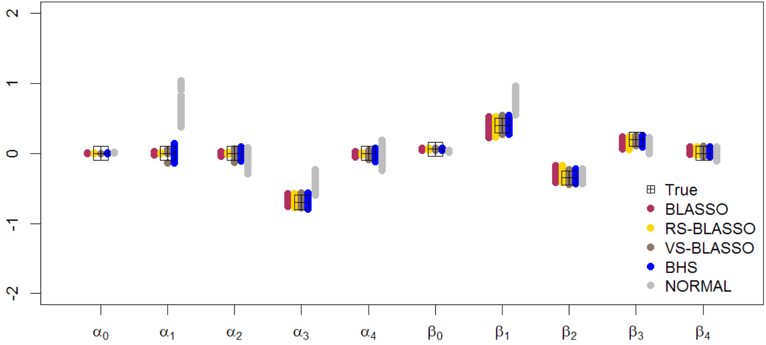
Figure: Posterior Means of Shrinkage Parameters from BLASSO(λ) and RS-BLASSO (λ_1,λ_2)



For BLASSO, 75% of replications converged compared to 94% for RS-BLASSO.

Simulation 2

Figure: Posterior Estimates from 100 Replications



Simulation 3

Let $\boldsymbol{d}_t = [y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}]'$ and $\phi' = [\phi_1, \phi_2, \phi_3, \phi_4]$.

$$y_t = (-0.6y_{t-3})[1 - G(y_{t-1})] + (0.02 + 0.75y_{t-3})[G(y_{t-1})] + \epsilon_t$$

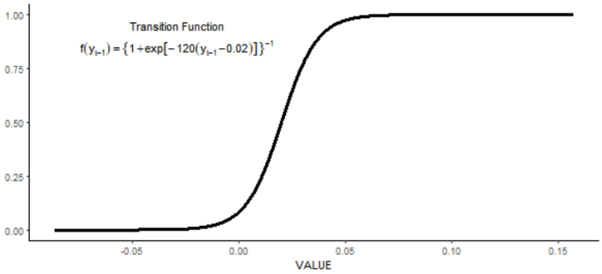
$$\text{where: } G(y_{t-1}) = \left\{1 + \exp\left[-120(\phi' \boldsymbol{d}_t - 0.02)\right]\right\}^{-1}$$

$$\text{and } \epsilon_t \sim \text{i.i.d. } N(0, 0.02^2)$$

Under prior $\phi \sim Dir([0.25, 0.25, 0.25, 0.25]')$, we conduct posterior sampling for three different threshold variables $\{z_{1,t}, z_{2,t}, z_{3,t}\}$ defined through ϕ. BHS priors are used for autoregressive coefficients.

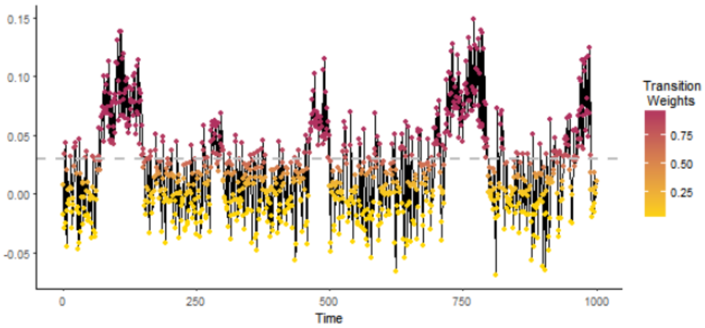
Simulation 3

Figure: Transition Function



Simulation 3

Figure: Illustration of Regime-switching Behavior

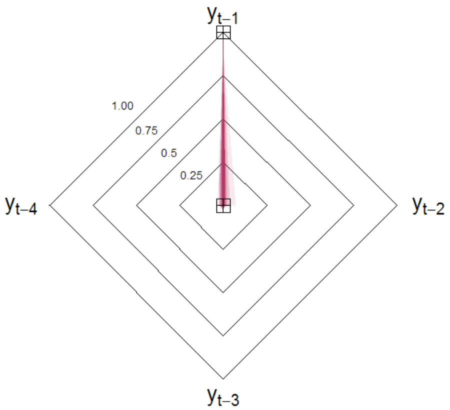


Simulation 3

- Bayesian Selection of the Threshold Variable (Scenario 1)

  Consider $z_{1,t} = y_{t-1} = [1, 0, 0, 0]\boldsymbol{d}_t$.

Figure: Posterior Means of ϕ from 100 Replications

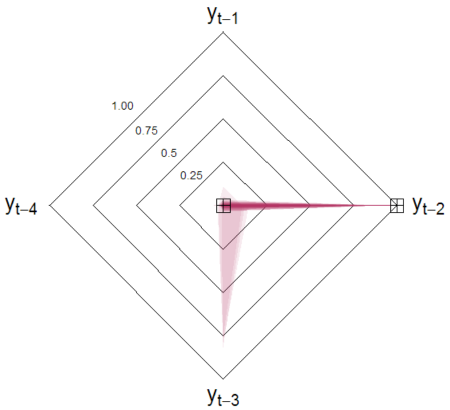


Simulation 3

- Bayesian Selection of the Threshold Variable (Scenario 2)

  Consider $z_{2,t} = y_{t-2} = [0, 1, 0, 0]\boldsymbol{d}_t$.

Figure: Posterior Means of ϕ from 100 Replications

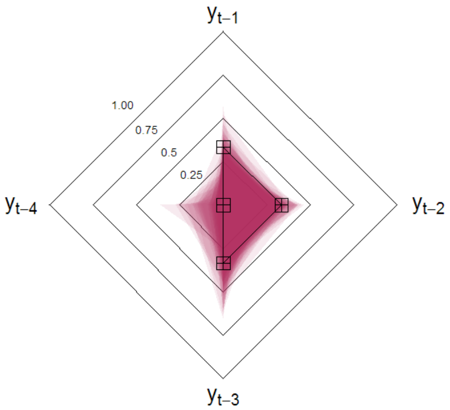


Simulation 3

- Bayesian Selection of the Threshold Variable (Scenario 3)

Consider $z_{3,t} = \frac{y_{t-1}+y_{t-2}+y_{t-3}}{3} = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0]\boldsymbol{d}_t$.

Figure: Posterior Means of ϕ from 100 Replications



Classic Application

- Application to Annual Sunspot Numbers
  - Textbook Example for Nonlinear Models Since Granger (1957)
  - Gathered and Updated by the World Data Center SILSO, Royal Observatory of Belgium, Brussels
  - Square Root Transformation: $y_t = 2[\sqrt{1 + x_t} - 1]$ (Ghaddar and Tong, 1981)
  - In Teräsvirta et al. (2010), LSTAR Outperformed Other AR, TAR,STAR, and Aritificial Neural Net (AR-NN) Models. Sparsity Achieved Via Stepwise Frequentist Procedure Using AIC
  - Training Period (1700-1979) and Testing Period (1980-2006)

Classic Application

- Application to Annual Sunspot Numbers (Cont.)
  - Terasvirta's Best LSTAR Model (F_T)
    LSTAR(10) Model With $d = 2$
  - Frequentist Estimation of Full Saturated LSTAR(10) (F_S)
  - BHS Estimated Linear Model AR(10) (B_L)
  - BHS Estimated LSTAR(10) with $d = 2$ (B_2)
  - BHS Estimated LSTAR(10) Applying Dirichlet Prior (B_D)
  - BHS Estimated LSTAR(10) with $d = 3$ (B_3)

Classic Application

- Application to Annual Sunspot Numbers (Cont.)

  Compare Models on RMSFE(h) for Horizons $h \in \{1, 2, 3, 4, 5\}$

  Bootstrap Method Used for Multi-step Ahead Forecasts for 1980-2006

Model	Horizon 1	2	3	4	5
F_T	1.42	2	2.36	2.51	2.35
F_S	1.86	3.21	3.7	3.63	3.16
B_L	1.73	2.3	2.54	2.53	2.56
B_2	1.42	1.96	2.29	2.19	2.19
B_D	1.77	2.83	3.38	3.5	3.29
B_3	1.86	3.11	3.58	3.62	3.58

New Application

- Application to Daily Maximum Water Temperatures
  - Data Used From 31 Rivers in Spain
  - Models Estimated to Forecast Daily Maximum Water Temperature Using Previously Known Daily Maximum Water Temperatures and Daily Maximum Air Temperatures
  - Combination of BHS and Dirichlet Priors for Estimation of Linear and Nonlinear Models Under Assumption $P = 6$
  - Horizon Specific Models Targeting 3-step and 7-step Ahead Forecasts
  - Nonlinear Models Improved Forecasting Accuracy for a Couple of Rivers (Details Provided in Paper)

Conclusion

- Contribution and Novelty
  - Efficacy of Bayesian Regularization for Nonlinear Time Series Models
  - Use of the Dirichlet Prior Leads to More Encompassing LSTAR Specification
  - Regime-Specific Tuning Parameters Influences Convergence in MCMC
  - Detailed **R** Code Provided for Reproducibility
- Feedback from *International Journal of Forecasting*
  - Focus on Dirchlet Priors for Estimating Transition Variable
  - Better Forecasting Application
  - Consider Density Forecasts Along with Point Forecasts
- Current Work
  - Provide Forecast Results in Simulation Studies
  - Given Classic LSTAR Process, How Do Forecasts from Methods Using Dirichlet Priors Compare to Scenario when the Delay Parameter is Known?
  - Given LSTAR with Composite Transition Variable, What is the Effect of Choosing the Wrong Delay Parameter?

References I

Battaglia, F. and Protopapas, M. K. (2012). An analysis of global warming in the alpine region based on nonlinear nonstationary time series models. *Statistical Methods & Applications*, 21(3):315–334.

Campbell, E. P. (2004). Bayesian selection of threshold autoregressive models. *Journal of time series analysis.*, 25(4):467–482.

Carvalho, C. M., Polson, N. G., and Scott, J. G. (2010). The horseshoe estimator for sparse signals. *Biometrika*, 97(2):465–480.

Chen, C. W. and Lee, J. C. (1995). Bayesian inference of threshold autoregressive models. *Journal of Time Series Analysis*, 16(5):483–492.

Deschamps, P. J. (2008). Comparing smooth transition and Markov switching autoregressive models of US unemployment. *Journal of Applied Econometrics*, 23(4):435–462.

Geweke, J. and Terui, N. (1993). Bayesian threshold autoregressive models for nonlinear time series. *Journal of Time Series Analysis*, 14(5):441–454.

Ghaddar, D. and Tong, H. (1981). Data transformation and self-exciting threshold autoregression. *Applied Statistics*, 30(3).

Granger, C. W. J. (1957). A statistical model for sunspot activity. *Astrophysical Journal*, 126:152.

Green, P. J. (1995). Reversible jump markov chain monte carlo computation and bayesian model determination.

Hamaker, E. L., Grasman, R. P. P. P., and Kamphuis, J. H. (2010). Regime-switching models to study psychological processes. American Psychological Association,.

Hans, C. (2009). Bayesian lasso regression. *Biometrika*, 96(4):835–845.

References II

Hsin-Min Lu, D., Zeng, D., and Hsinchun Chen, D. (2010). Prospective infectious disease outbreak detection using markov switching models. *Knowledge and Data Engineering, IEEE Transactions on*, 22(4):565–577.

Kamarianakis, Y., Gao, H. O., and Prastacos, P. (2010). Characterizing regimes in daily cycles of urban traffic using smooth-transition regressions. *Transportation Research Part C: Emerging Technologies*, 18(5):821–840.

Koop, G. and Potter, S. M. (1999). Dynamic asymmetries in us unemployment. *Journal of Business & Economic Statistics*, 17(3):298–312.

Li, W. and Lam, K. (1995). Modelling asymmetry in stock returns by a threshold autoregressive conditional heteroscedastic model. *The Statistician*, pages 333–341.

Livingston Jr., G. and Nur, D. (2017). Bayesian inference for smooth transition autoregressive (star) model: A prior sensitivity analysis. *Communications in Statistics - Simulation and Computation*, 46(7):5440–5461.

Lopes, H. F. and Salazar, E. (2006). Bayesian model uncertainty in smooth transition autoregressions. *Journal of Time Series Analysis*, 27(1):99–117.

Lubrano, M. (2000). Bayesian analysis of nonlinear time series models with a threshold.

Lykou, A. and Ntzoufras, I. (2013). On bayesian lasso variable selection and the specification of the shrinkage parameter. *Statistics and Computing*, 23(3):361–390.

Makalic, E. and Schmidt, D. F. (2016). A simple sampler for the horseshoe estimator. *IEEE Signal Processing Letters*, 23(1):179–182.

Montgomery, A. L., Zarnowitz, V., Tsay, R. S., and Tiao, G. C. (1998). Forecasting the u.s. unemployment rate. *Journal of the American Statistical Association*, 93(442):478–493.

References III

Nieto, F. H., Zhang, H., and Li, W. (2013). Using the reversible jump mcmc procedure for identifying and estimating univariate tar models. *Communications in Statistics - Simulation and Computation*, 42(4):814–840.

Park, T. and Casella, G. (2008). The bayesian lasso.

Schmidt, D. F. and Makalic, E. (2013). Estimation of stationary autoregressive models with the bayesian lasso. *Journal of Time Series Analysis*, 34(5):517–531.

Stock, J. H. and Watson, M. W. (1998). A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series.

Teräsvirta, T. (1995). Modelling nonlinearity in us gross national product 1889–1987. *Empirical Economics*, 20(4):577–597.

Teräsvirta, T., Tjøstheim, D., and Granger, C. W. J. (2010). *Modelling nonlinear economic time series*. Oxford University Press Oxford.

Troughton, P. T. and Godsill, S. J. (1997). A reversible jump sampler for autoregressive time series, employing full conditionals to achieve efficient model space moves.

Zeng, J.-H., Lee, C.-C., and Chang, C.-P. (2011). Are fruit and vegetable prices non-linear stationary? evidence from smooth transition autoregressive models. *Economics Bulletin*, 31(1):189–207.