Computer Tutorial 2: Nonlinear Time Series Model

Question 1: The file SP500.xlsx contains montly observations on the S&P500 stock market index from January 1959 through August 2016.
a) Construct a new variable which is the log difference of the stock price index. Graph the stock price index and its log difference. Use the log difference as your dependent variable, y_t, for the remaining parts of this question.
b) Estimate an AR(1) model for y_t. Calculate information criteria for this model.
c) Estimate a homoskedastic TAR model for y_t using $z_t = y_{t-1}$ as threshold trigger and $\tau = 0$ as the threshold. Calculate information criteria for this model.
d) Estimate a heterokedastic TAR model for y_t using $z_t = y_{t-1}$ as threshold trigger and $\tau = 0$ as the threshold. Calculate information criteria for this model.
e) Discuss your results from parts b) through d). Is there evidence of threshold nonlinearity in this variable? If so, which parameters are changing across regimes?
f) Extra question (if time permits or you are ambitious): It is possible that y_t will have different behaviour depending on whether stock market volatility is high or low. Investigate this possibity using a TAR model for y_t and $z_t = (y_{t-1} - \overline{y})^2$ and $\tau = \overline{z}$. Note: $\overline{y}$ denotes the sample mean of y_t and z_t can be interpreted as an approximate measure of stock market volatility last month.

Question 2: In the lecture I went through an empirical example using a two regime Markov switching AR(1) model for US GDP growth (see the lecture slides for data definition, the data set is called US_rgdp_change.xls). I compared a homoskedatic variant of the model to a heteroskedastic one where the two regimes had different error variances and found strong evidence in favour of the latter.
a) Carry out a simiar model selection exercise including a wider range of Markov switching models including: i) models with 3 regimes, ii) models with long lag lengths and iii) models with regime switching only in the error variance (i.e. the AR coefficients and intercept are the same in all regimes). What is your preferred model?
b) For your preferred model, plot estimates of the probabilities that each time period falls in each regime. In the lecture slides, for the heteroeskedastic model, I was able to interpret the two regimes as reflecting high and low variances. Can you offer a similar interpretation of the regimes for your preferred model? If not, how do you interpret your regimes?