

The forecasting models for retail sales in lectures and tutorials have used the date 2008.5 (2008q3) as the trend break date for the GFC. This assignment will explore methods how this choice of date might be evaluated. The [National Bureau of Economic Research officially dated the GFC recession](#) as lasting from 2007q4 to 2009q2, so we will explore this range of trend break dates.

To carry out this assignment, read in the total retail sales time series and create a `ts` object for quarterly log retail sales from 2000q1 to 2019q4 in the same way as lectures and tutorials.

- Initially consider the *estimation sample* to be **2000q1 to 2018q4**. Consider the time series models

$$Y_t = \beta_0 + \beta_1 \text{Time}_t + \beta_2 \text{TimePostGFC}_t(\tau) + \delta_1 Q_{1,t} + \delta_2 Q_{2,t} + \delta_3 Q_{3,t} + Z_t$$

$$E(Z_t | \mathcal{Z}_{t-1}) = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p}$$

for $p = 0, 1, \dots, 8$ and with

$$\text{TimePostGFC}_t(\tau) = (\text{Time}_t - \tau) \times 1(\text{Time}_t > \tau)$$

where τ represents the GFC break date which may any one of 2007q4, 2008q1, 2008q2, 2008q3, 2008q4, 2009q1, 2009q2.

- Set up R code that will, for each of the 7 possible GFC break dates, use the AICc to select the AR order p . Also include code to check whether in each case the residuals of the chosen model show evidence of autocorrelation.
- Then set up code to further use the AICc to select the preferred value for the GFC break date τ .
- Further develop the code in parts (a) and (b) with an extra loop to implement a *recursive regression* approach to extend the end of the estimation sample to each of 2019q1, 2019q2, 2019q3 respectively (while keeping the start of the estimation sample at 2000q1).

Report the results of this analysis, in particular reporting on:

- The selected AR lag order for each GFC break date and each estimation sample.
 - The selected GFC break date for each estimation sample.
- For each of the four estimation samples and each of the seven possible GFC break dates, use the AICc-selected AR model to calculate one-step-ahead forecasts and forecast errors for log retail sales. For each possible GFC break date, calculate the RMSE of the one-step-ahead forecasts *across* the four estimation samples. (i.e. for each possible GFC break date the Root Mean Squared Error is taken across the four one-step-ahead forecast errors calculated from the four recursive estimation samples.)
Report the RMSE for each of the seven possible GFC break dates. (It is not necessarily to tabulate all of the individual forecasts and forecast errors.)
 - Discuss** how the outcomes of this modelling and forecasting exercise depend on the choice of GFC break date. (You may produce additional statistics or graphs to help inform your discussion, within the constraint of the page limit below.)

Submission:

- Maximum of 1 page:** the reports/discussions indicated in red above.
- In addition, your R script that produced these results. If required, the script must be able to be run by the marker with the results clearly printed out.