Financial Econometrics
Assumption and Biases

Darya Yuferova

In this exercise, you investigate the Carhart four-factor model. The model suggest that the exposures of a stock/portfolio to the market, high-minus-low (HML), small-minus-big (SMB), and momentum (WML) factors explain the return of a stock.

Load *Data_for_Cahart_test.Rdata* which includes average industry returns and factor loadings (b = market beta, s = exposure to SMB, h = exposure to HML, w = exposure to momentum) from 2010 to 2016.

Please conduct following analysis:

1. Load and inspect the data
2. Winsorize the data at the 95%. Please create new variables.
3. Make a scatter plot with market beta on the x-axis and return on the y-axis (using winsorized variables). General asset pricing intuition suggests the higher the beta, the higher the return. Does the data support this intuition?
4. Run three regression models explaining return with:
   - The market beta using winsorized variables
   - The Carhart model using the raw variables
   - The Carhart model using the winsorized variables
5. Compute heteroskedasticity robust standard errors for all models. Use *coeftest(..., vcov = hccm)* in which ... indicates the output from *lm(...)*
6. Investigate multicollinearity in model 2 using the Variance Inflation Factor
7. Make a nice stargazer output that uses the heteroskedasticity robust standard errors
8. [Optional] Make a scatter plot with the winsorized returns on the x-axis and the predicted returns from model 3 on the y-axis. Include a 45 degree line. Does the model explain the data well?

The main take away: being able to implement heteroskedasticity robust standard errors, assessing the importance of outliers, and investigating multicollinearity.