# MFIN 4328
# Applied Financial Econometrics
# Assignment 2

Prof. Riccardo Sabbatucci

**Due**: March 9, 2022

*****Please submit your solution in a standalone PDF saved as "YourStudentID_YourFirstName_YourLastName" report that contains the code and output from your R script. You could either do this by clicking "-File-Knit Document-PDF" in R Studio or using R markdown following instructions at this link *****

**Total points: 20**

1. (5 points) For this exercise, use the data in "pension.dta". The equation of interest is a linear probability model

$$pira = \beta_0 + \beta_1 p401k + \beta_2 inc + \beta_3 inc^2 + \beta_4 age + \beta_5 age^2 + u \tag{1}$$

where $pira$ is a binary (e.g., 0/1) variable equal to 1 if the person has an IRA account. The goal is to test whether there is a tradeoff between participating in a 401(k) plan and having an individual retirement account (IRA). Therefore, we want to estimate $\beta_1$

describe the coefficients and what they mean economically

(a) Estimate the equation (1) by OLS and discuss the estimated effect of $p401k$

(b) For the purposes of estimating the ceteris paribus tradeoff between participation in two different types of retirement savings plans, what might be a problem with ordinary least squares (OLS)?

(c) The variable $e401k$ is a binary variable equal to one if a worker is *elegible* to participate in a 401(k) plan. Explain what is required for $e401k$ to be a valid IV for $p401k$. Do these assumptions seem reasonable?

(d) Estimate the reduced form for $p401k$ and verify that $e401k$ has significant partial correlation with $p401k$. Since the reduced form is also a linear probability model, use a heteroskedasticity-robust standard errors (e.g., White S.E.s).

(e) Now, estimate the structural equation by instrumental variable (IV) and compare the estimate of $\beta_1$ with the OLS estimate.

2. (5 points) For this exercise, use the data in "loans.dta".

(a) Estimate a probit model of *approve* on *white*. Find the estimated probability of loan approval for both whites and nonwhites.

(b) Estimate a simple linear probability model of *approve* on *white*. How do these estimates compare to the ones in point (a)?

(c) Now, add the variables *hrat*, *obrat*, *loanprc*, *unem*, *male*, *married*, *dep*, *sch*, *cosign*, *chist*, *pubrec*, *mortlat*1, *mortlat*2, and *vr* to the probit model. Re-estimate the model. Is there statistically significant evidence of discrimination against non-whites now?

(d) Estimate the model in part (c) by logit. Compare the coefficient on *white* to the probit estimate.

3. (5 points) Two researchers are asked to perform a time series analysis of the equal-weighted index return series (EWRET) from January 2, 1997 to November 30, 2011, denoted $x_t$. They determine the appropriate model by using an estimation of the information criteria for each ARMA model order from (0,0) to (5,5). Researcher A uses Akaike's information criterion (AIC), Researcher B uses Schwarz's Bayesian information criterion (BIC). Obtain the data from the sheet "stock_daily" of the file "data_assignment2.xlsx" and compute the log (or continuously compounded) return series

   (a) Plot the sample ACF and the sample PACF, up to 22 lags, of the log return series and comment on their patterns. comment on the patterns

   (b) Find the optimal model estimated by each researcher, produce parameter estimates with their standard errors and check the fitted model.

   (c) Use the fitted ARMA models to compute 1-step to 21-step ahead forecasts at the forecast origin November 30, 2011. Show the prediction plots including the two standard error limits of the forecasts and the actual observed values of the (EWRET) return series. Compare and comment on the forecasting accuracy of the models using the MSE and the MAE of the forecasts.

   (d) Keeping the residuals series from the estimations performed in (b) by Researcher B, form the absolute residuals and squared residuals series

      i. Compute the variance, skewness, excess kurtosis, minimum and maximum of the absolute residuals and squared residuals series.

      ii. Plot the autocorrelogram (ACF) and the partial autocorrelogram (PACF) of the absolute residuals and squared residuals series.

      iii. Test whether the squared residuals are stationary or have a unit root.

      iv. Build an ARMA model for the squared residuals series and check the fitted model (for this part, use the BIC criterion).

4. (5 points) Consider the monthly U.S. 1-year and 5-year Treasury rates from January 1980 to December 2016. Data can be found in the "rates" sheet of the file "data_assignment2.xlsx". Define the vector $x_t = \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix}$ where $x_{1,t}$ is the 1-year Treasury rate and $x_{2,t}$ is the 5-year Treasury rate.

   (a) Estimate a VAR(2) model for the bivariate interest rate series. Write down the fitted model.

(b) Compute and plot the impulse response function of the fitted VAR model. You should compute 12-periods impulse responses.