**Exercise 1**

The aim of this exercise is for you to design your own simulation study. I won’t specify exactly what it will be on, but it could be to do with some of the issues we have talked about so far in this course, for example:

* The difference between maximum likelihood and restricted maximum likelihood
* The bias associated with different sample sizes at level 2

Or it could be something else in statistical modelling, such as:

* The effect of omitting important variables from your model
* How correlated variables have to be in order to reduce statistical power

Or something else entirely.

The aim, is to run a simulation study using either Stata or R.

You can work individually, or in pairs. If you do not have some proficiency in Stata or R, I would recommend pairing up with someone who does.

You should follow the following steps:

1. Decide what your end goal is. This might be a graphic, or a table, or just a set of numbers. Draw (with pen and paper) how you think such a figure might look, and keep the drawing until tomorrow.
2. Decide your data generating process, and write code to generate one set of it in Stata or R. Consider what parameters you will be varying
3. Decide the analysis you will be conducting, and code this in R using the dataset you have generated
4. Run a simulation based on just that one dataset (that is, at this stage, do not run multiple datasets.

You may need to ask for help, especially if you are unfamiliar with Stata and R. This is completely fine – just call me over.

**Exercise 2**

We will pick up from where we left off in the first practical. The aim is to use the techniques available in R and Stata to automate multiple runs of the simulations that we designed yesterday. The final aim will be to generate a graphic along the lines of what you produced conceptually yesterday.

You should follow the following steps

1. Check you are still happy with your ‘end-goal’ produced yesterday. This can change as research progresses, so feel free to adjust it if you need to
2. Write R code to simulate many dataset, fit models to each oe, and extract the parameters/quantities of interest.
3. Calculate derived quantities of interest, e.g. RMSE, relative bias, optimism, etc.
4. Generate graphics and/or tables as per step 1.