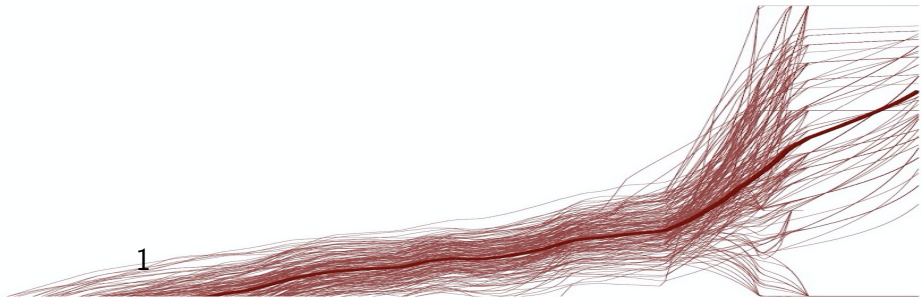
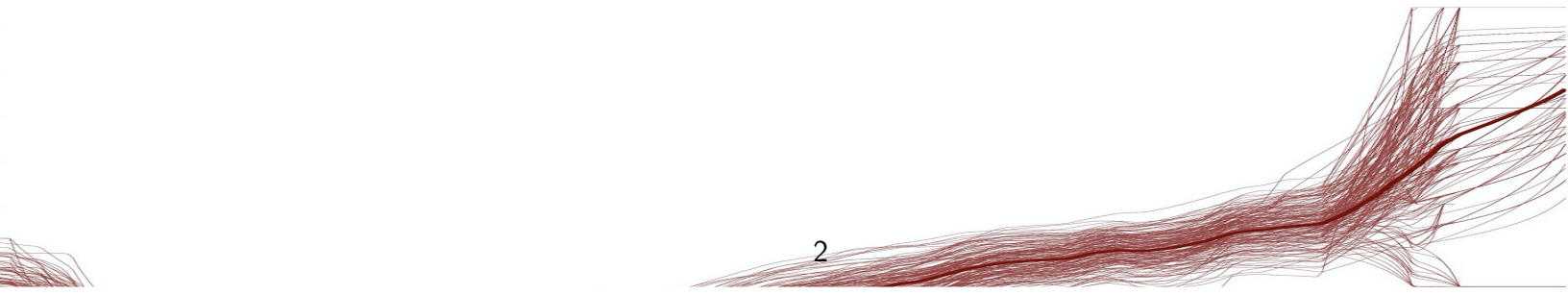
Financial Econometrics
Lab Sessions & Homeworks

Pr. Gaëlle Le Fol
UNIVERSITE PARIS - DAUPHINE
Master 203 - Master 104

September 2025

Table des matières





Preparing the lab classes

Lab classes provide students with the opportunity to apply and investigate theoretical and conceptual knowledge, to develop and code a range of techniques and approaches, and to improve skills in analysing, and interpreting data. However, if you come to the lab class unprepared, the only thing that it will provide you with will be the opportunity to copy code lines.

To make the most of these lab sessions, you must :

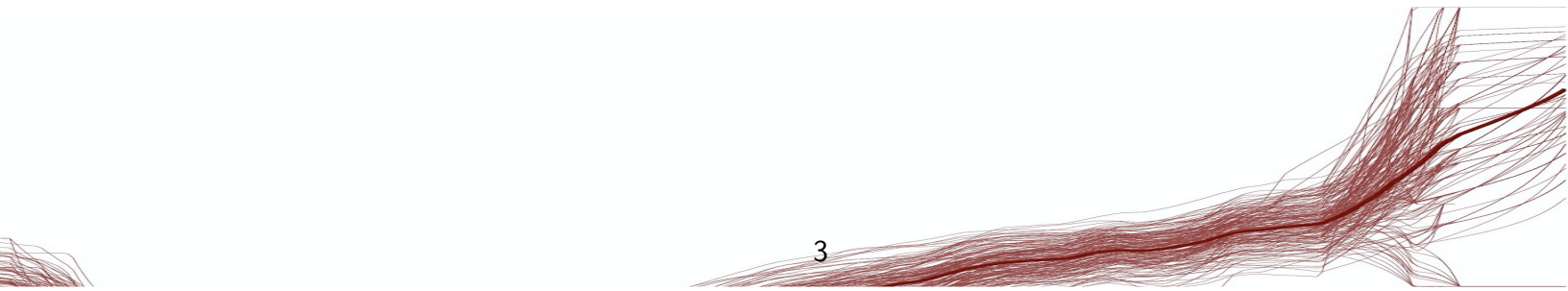
— Read through the Lab class handout and check that you know hwo to use the commands/functions/tools you will need all along the exercises of the day.

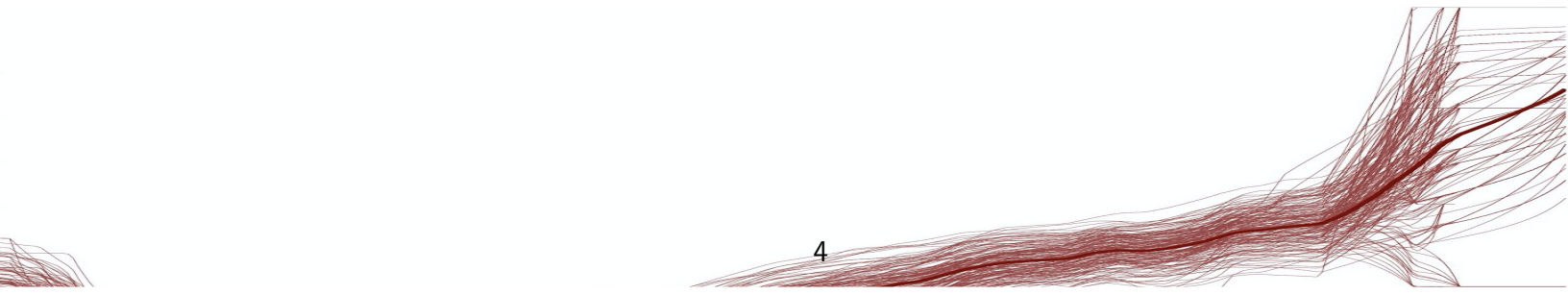
— Revise the part of the course that the lab class is exploring.

— Check with your lab class information to see if you are expected to do any preparatory work before the class.

By failing to prepare, you are preparing to fail
Benjamin Franklin

Preparation time is never wasted time
Anonymous

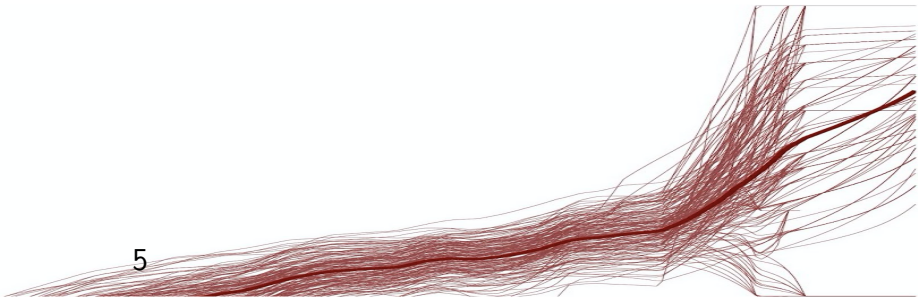


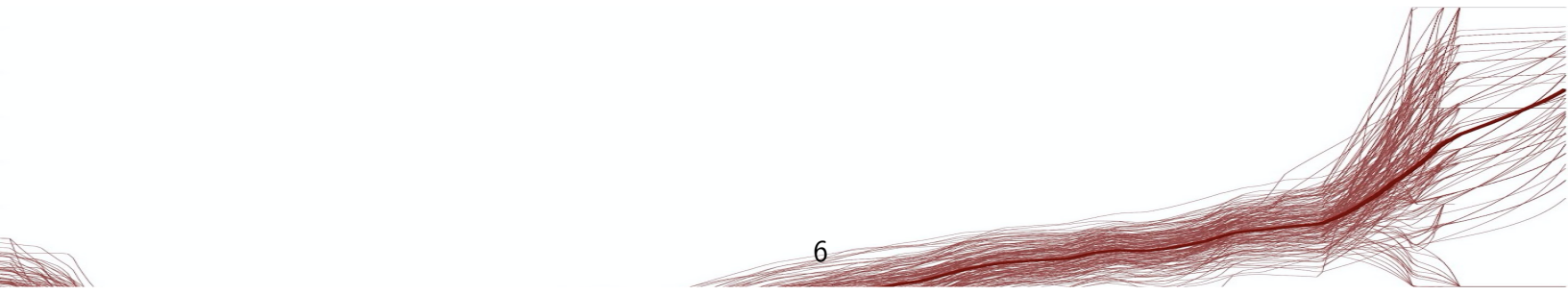


Catching up with Gretl : preparatory work

This is a catch-up class that will help you getting use to Gretl and revising the notions seen in chapters 1 (Introduction) to 4 (Simple linear regression with Gretl). Before starting this catch-up class,

1. If Gretl is not installed, download and install it from http ://gretl.sourceforge.net/ ;
2. Make sure that you know the 4 first chapters ;
3. Read "A (very) short introduction to Gretl using scripts"
4. Read the statement of the the catchup lab and find out what are the equations, the statistical methods and tests that will be useful ;
5. find out what are the commands and functions that will be needed using the help command or menu.





This is a catch-up class that will help you getting use to Gretl (or R) and revising the notions seen in chapters 1 (Introduction) to 4 (Simple linear regression with Gretl). Before starting this catch-up class,

1. make sure that you know the 4 chapters before coming to the lab class ;
2. revise the functions and commands that will be needed :
   (a) What does each command and function ?
   (b) What is the syntax to use it ?
3. read the statement of the lab class and find out what are the equations that will be useful ;
4. explore the commands and functions that will be needed and are listed in the table below.

Gretl Programming		
Commands	Functions	
append	$coeff	abs
dataset	$df	cov
genr	$ess	delete
gnuplot	$ncoeff	diag
lags	$nobs	inv
ols	$rsq	log
open	$stderr	mean
outfile	$uhat	nobs
pvalue	$yhat	pvalue
print		rows
printf		sqrt
rename		sum
scatters		var
setobs		
smpl		
store		
summary	Types	
workdir	matrix	scalar

R Programming		
abline	length	read_excel
abs	lines	read.table
as.matrix	lm	save
as.vector	mean	setwd
attach	merge	sqrt
cbind	ncol	sum
coef	nrow	summary
data.frame	plot	t.test
colnames	plot.ts	var
diag	pt	view
Packages		
ggplot2	lmtest	readxl
tseries	zoo	

Preparatory work - Analysis of the link between the CAC40 and some other indices

Daily return data for the CAC40, the EuroStoXX50, the Dow Jones and the S&P500 are stored in an Excel file called "DataPrepWK.xlsx" or in four separated files called "CAC40.csv" "EuroStoxx.csv", "DowJones.csv" and "SandP500.csv". We are trying to figure out if the evolution of the CAC40 index depends on the evolutions of the EuroStoxx, the Dow Jones or the S&P500.

1. Look at the Excel file described above and check the data formats : What is name of the spreadsheet where the data are stored ? What is the line from which we are loading the data ? What is the type of data ? How many observations do we have in each file ? What are the beginning and ending dates ? What is the frequency of the observations ?

File	CAC40	DowJones	EuroStoxx50	SandP500
Spreadsheet				
Cell NB				
Data type				
Begin date				
End date				
Freq.				

2. Start a new script file. by writing some comments (#) to describe your script file. Propose some code to clear all the existing data and variables from Gretl memory (command **clear**). Declare a working directory (command **workdir**) "/Users/Yourname/.../myworkingdir" and save the script as "GretlCatchUp".
3. Load the EuroStoxx series (command **open**) into Gretl (or R), define the date format (command **setobs**), rename the price as `PEURO` (command **rename**) and save the file (command **store**) as a Gretl or an R database, called "DailyEURO".
4. Do the same for the CAC40 series.
5. From the smalest data base, append (command **append**) the data of the other data base. What happen if you do the reverse (from the largest data base, append the other data base) ?
6. Compute the return of both time series using "(-1)" after the name of the variable you want to lag by one observation or the command **lags**. Save the series into a Gretl (or R) database called "Indexes".
7. Compute some summary statistics of the returns (command **summary**).

8. Do a scatter plot of the two indexes. Give the evolution of the two indices over the entire sample period (commands **gnuplot** or **scatters**).
9. Convert the daily series into monthly series (command **dataset**), delete the missing observation (if any) (command **smpl**) and save the (new) monthly database[1]. We call it "MIndexes".

R 9. Construct a dummy variable that is equal to 1 if we change month and use the `subset` function to extract the last index value of each month.

10. Create a graph showing the evolution of the two monthly index returns over a sub-period of your choice from the sample.
11. Compute the sum (function `sum`) of
    (a) the EurostoXX50 return and its square (use "^2"),
    (b) the CAC40 return and its square,
    (c) the CAC40-EurostoXX50 return cross product,

    and print the results (command **print** or **printf**).
12. From the previous sums, compute the
    (a) empirical means, variances and covariance (control that you did good by using the functions `mean`, `var`, `cov`),
    (b) estimated parameters,
    (c) associated standard errors (using the function `sqrt`), and print the results.
13. Estimate the simple regression model (command **ols**), and check the values calculated in the previous question. Save the regression table into a file (command **outfile ... end outfile**) called "IndexRegResults.txt".
14. Test the hypothesis that the coefficient associated with the constant is zero.
    (a) Compute the test statistics (using the `abs` function).
    (b) Using the `pvalue` function or the **pvalue** command, get the pvalue of the test and interpret the result. Can we use the fonction $pvalue instead ? Do it if the answer is yes and explain otherwise.

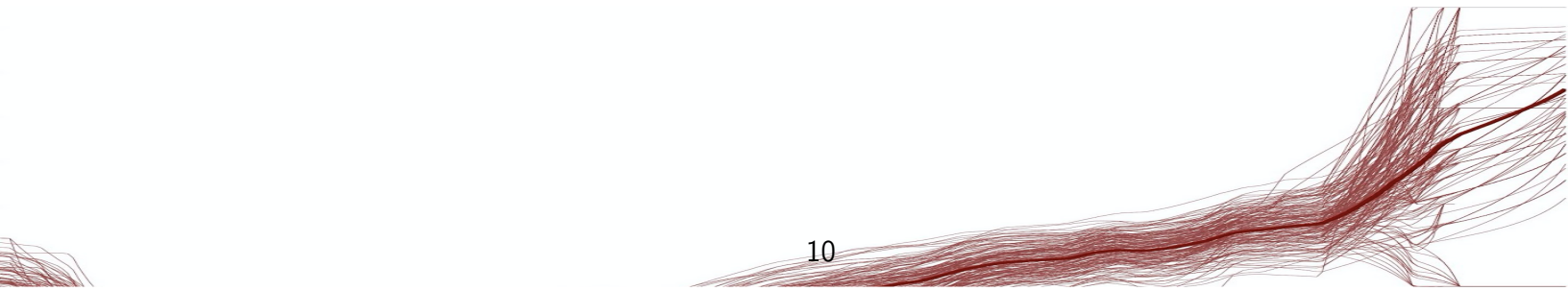
    (R b) Using the `pt` command, get the pvalue of the test and interpret the result. Can we use the `t.test` command ? Do it if the answer is yes and explain otherwise.

    (c) Verify the results of the previous question in the simple regression table.
15. Do again questions 12 to 13 in matrix format and use the `matrix`, and `scalar` declarations as well as the functions `diag`, `inv`, `rows`, `sqrt`.
16. What would you say about the link between the CAC40 and the Eurostoxx ? What about the link between the CAC40 and the Dow Jones and CAC40 and the S&P500 ?
17. Conclusion.

TAKE AWAY

In order to end that preparatory work, propose a pitch : what was this preparatory work about ? What did you learn ? What was the interest (if any) of that work for a market finance student ?

1. Do not forget to check that you are getting the 'right' monthly returns, for example by checking the values you obtain for the observations corresponding to the first two and last monthly returns of the two series.



Lab session 1 : CAPM estimations and tests

This Lab class corresponds to chapters 1 (Introduction) to 4 (Simple linear regression with Gretl) + chapters 7 (heteroscedasticity and serial correlation) to end (OLS : adaptation and beyond) appendices. Before coming to the Lab class,

1. make sure that you know the corresponding chapters and appendices ;
2. revise the functions and commands that will be needed :
   (a) What does each command and function ?
   (b) What is the syntax to use it ?
3. read the statement of the lab class and find out what are the equations that will be useful ;
4. explore the "new" commands and functions that will be needed and are listed in the table below.

Gretl Programming		
Commands	Functions	
boxplot	$coeff	diff
chow	$df	freq
dataset	$dw	log
loop	$dwpval	loop
modtest	$ess	mean
normtest	$rsq	nobs
printf	$sigma	pvalue
qlrtest	$stderr	quantile
quantreg	$uhat	round
reset		sum
restrict		
setobs		
smpl		

R Programming		
AIC	ks.test	qqPlot
BIC	length	quantile
bgtest	linearHypothesis	resettest
boxplot	lm	residuals
coeftest	match	rq
Fstats	pchisq	sctest
gqtest	pf	shapiro.test
grubbs.test	pt	
Packages & Libraries		
ggplot2	quantreg	strucchange
lmtest	sandwich	whitestrap
outliers	skedastic	zoo

The objective of this lab session is to determine whether the monthly returns of a particular stock can be explained by the variation of the market risk premium in line with the Capital Asset Pricing Model (CAPM) using the simple linear regression model. We will first estimate the CAPM model using Ordinary Least Squares (OLS). We will then test for the main OLS hypotheses and discuss the implications. We will also run some statistical inference tests. Finally, we will propose some corrections and/or some alternative estimators when needed.

Exercise 1 : Estimating the CAPM model using OLS

The CAPM (Capital Asset Pricing Model) is a theory (highly stylized) stipulating that the average return of any asset is explained by its relative sensitivity to a market portfolio (a portfolio composed of all assets in the world). This market sensitivity is the only risk in the portfolio, called β. We thus obtain the following model :

$$\mathbb{E}[R_{i,t}] = R_f + \beta_i \left(\mathbb{E}[R_{m,t}] - R_f\right), \tag{1}$$

where $\mathbb{E}[R_{i,t}]$ is the expected return of stock i at date t, R_f is the risk free rate, β_i is the sensitivity and $\mathbb{E}[R_{m,t}]$ is the expected market return at date t.

For the application, you have access to US monthly stock prices (in bps) for Ford (FORD), GE Aerospace (GE), Microsoft (MSFT), and Oracle (ORACLE), the S&P500 index (SANDP) and the 3-Month Treasury Bill Market Rate, from January, 2002 to February, 2018.

1. Give the estimable model derived from the CAPM theory and justify your choice.
2. Look at the database called "`capm.xlsx`" under Excel before loading it into Gretl (Python or R). Keep the names of the variables as in the Excel file.
3. Do a quick research on these stocks, what are the characteristics of these companies/stocks ? What about the S&P500 index ?
4. Load the data you need into Gretl (Python or R).
5. Construct r_f the risk free rate.
6. Build the variables of the excess return to the risk free rate for the market and for the stocks.
7. Propose two interesting graphes of the excess return of the stock and the market.
8. Estimate the simple linear regression model that comes from the CAPM (Capital Asset Pricing Model).
   (a) What should we get for β_0 and β_1 ? Explain.
   (b) What do you think of the quality of the regression ?
   (c) Conclusion.

Exercise 2 : Checking the statistical adequacy of the simple linear model

We now want to test if the simple linear regression model used for our CAPM estimation is statistically adequate.

1. What are the necessary hypotheses of the classical linear regression model ?
2. In each case, propose a test in Gretl (Python or R).
3. What are the validated hypotheses ? What are the implications ?
4. How would you test in Gretl (Python or R) that the beta coefficient is 1 ? Can we do it ? Explain.
5. Conclusion.

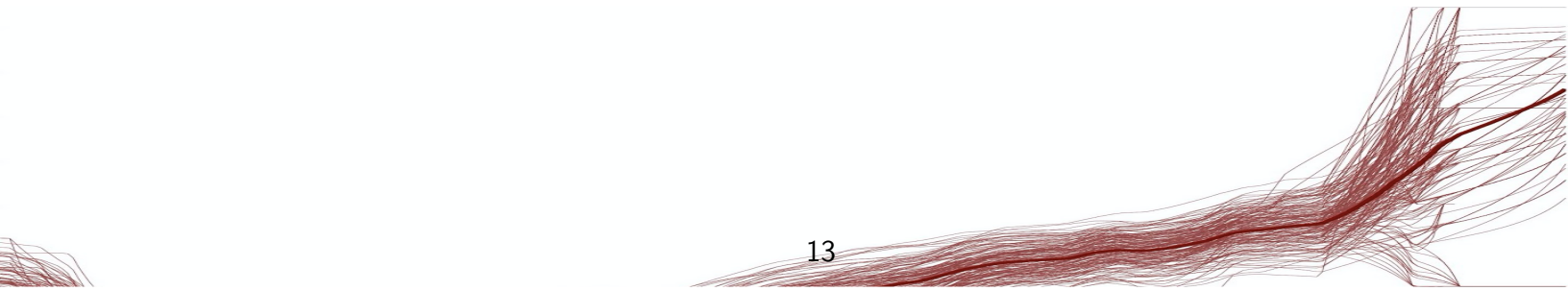
Exercise 3 : Correcting the problems

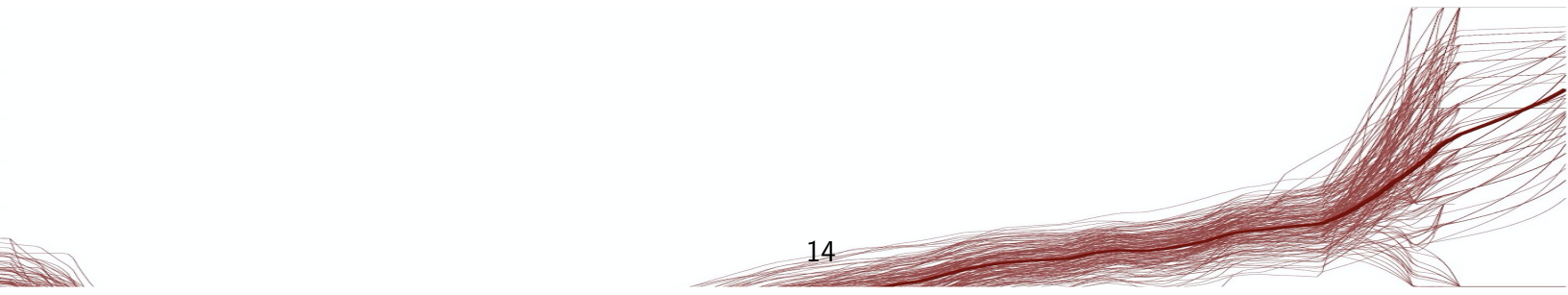
In the previous exercise we did test for the OLS hypotheses.

1. We first explore how to deal with these problems, and in each case, you are asked to propose a program in Gretl (Python or R).
   (a) How would you correct for autocorrelation and/or heteroscedasticity problems ?
   (b) How would you correct for normality problems ?
   (c) How would you correct for parameter stability problems ?
   (d) How would you correct for the presence of outliers ?
   (e) How would you correct for the presence of functional form ?
2. We decide to see what we can learn from using the quantile regression technics on our CAPM model.
   (a) What is the difference between OLS and quantile regression in our case ?
   (b) For the ten deciles $(0.1, \cdots, 0.9)$, run the quantile regression associated with the CAPM.
   (c) Interpret the (new) results and compare to the OLS regression results.
3. Conclusion.

TAKE AWAY

In order to end that LAB session, propose a pitch of LAB 1 : what was LAB 1 about ? What did you learn ? What was the interest (if any) of that LAB for a market finance student ?





The three golden rules of econometrics are "test, test, and test"
Robert Hendry

Doing applied work involves a synthesis of various elements. You must be clear about why you are doing it [...]. You must understand the characteristics of the data you are using and appreciate their weaknesses. You must use theory to provide a model of the process that may have generated the data. You must know the statistical methods, which rely on probability theory, to summarise the data, e.g. in estimates. You must be able to use the software [...]. You must be able to interpret the statistics or estimates in terms of your original purpose and the theory.
Ron Smith