Boston University Questrom School of Business
MF840 – Spring 2019
Eric Jacquier
Problem Set 2
Due Thursday February 28th in class
Solutions

- **Teams "*across sections*" turn in their homework at the beginning the morning class**
- **To get a check, you need to answer <u>all</u> the questions.**
- **You can type in discussion answers directly in this file**.

You can do this quickly if you use the R help given !

The xx-indus-mon.csv files in the DATA directory contain the monthly returns on xx industries. Go to Ken French's web site to understand what the variables are. I cleaned up the files for you and added $XR_M = R_M - R_F$ and R_F for you in the last two columns of each file.

Use the Risk free rate in the last column to compute excess returns for the industries. It's OK to leave all returns in % ala KF. You can now run the excess-returns regressions:

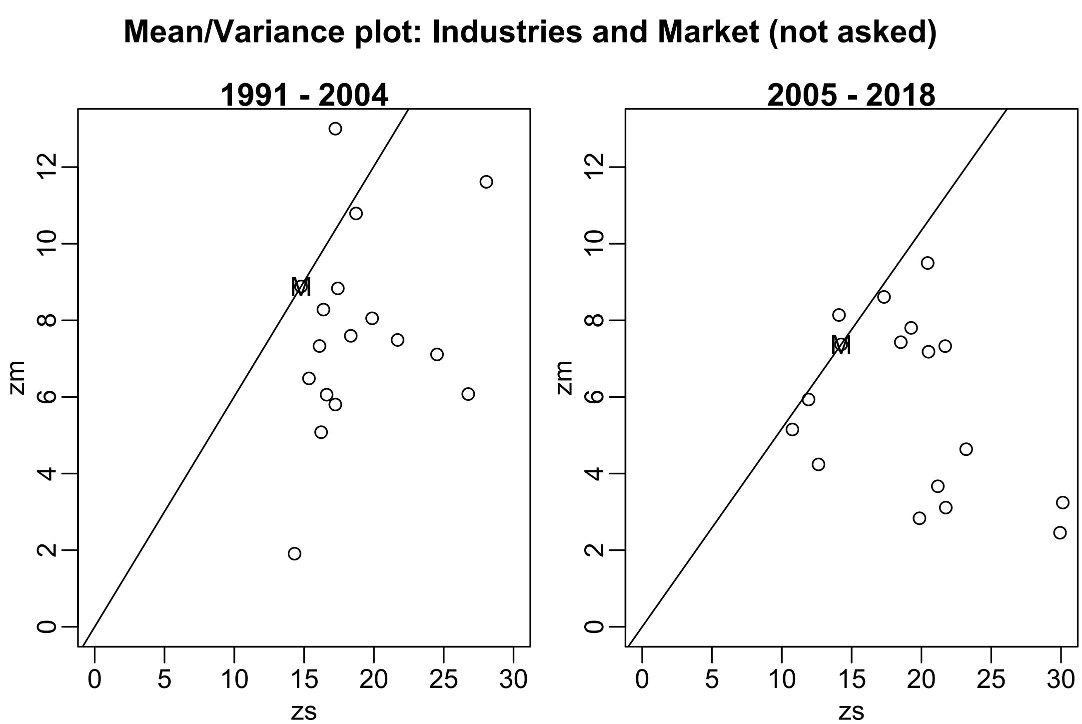
$$XR_{it} = R_{it} - R_{Ft} = \alpha_I + \beta_I \, XR_{Mt}$$

For this entire problem, you can assume that the OLS assumption applies to each regression alone.

<u>Problem 1: Bonferroni and Hotelling corrections:</u>

Use 16indus to fill in Table 1. Use the 60 monthly returns from 200901 to 201312. For each industry, write the slope estimate, the t-statistic for a test of H0: β=1, the excess average return α (aka Jensen's α aka abnormal return).

We are going to look at αs – the abnormal performance. It is tightly related to the mean variance efficiency of the Market Portfolio. For the whole period 26-18, we can plot the 17 portfolios in a (σ, μ) plot.



Consider two-sided 5% tests.		ANSWER HERE
	What is the cut-off for the absolute value of t?	**2.00**
	*Flag every t exceeding the cutoff with a * on its left.*	Bolded in table
Bonferroni:	*What is the Bonferroni cutoff.*	
	Use $\alpha/16 = 0.003125$ $\lvert t_{\alpha/2}(60\text{-}2)\rvert$=	**3.08**
	?	
	Put a * in the Bonferroni column if it exceeds it.	
Hotelling:	*What is the exact Hotelling cut-off,*	**6.35**
	What is the asymptotic cut-off?	**5.13**
	*Use the exact cut-off to put a * in the Hot. Column*	

Questions:

- *Does it look like we can reject the null H_0 that the βs are all 1?*

  Even with the very conservative Hotelling cutoff, we can reject the null that all βs are 1.
  Note though how much more severe the Hotelling cutoff is than Bonferroni
  Note also that using the asymptotic cut-off leads to a misspecified test (wrong size)

- *Does it look like we can reject the null H_0 that theses industries have no abnormal return*

  2 rejections with the uncorrected t, one with Bonferroni, none with Hotelling.
  This is not surprisins given the large imprecision of mean estimates.

Table 1: OLS estimates, monthly excess returns on 16 US industries

Industry	β	t_β	Bonf.	Hot.	α x 12	t_α	Bonf.	Hot.
Food	0.51	**-9.39**	*	*	3.68	1.20		
Mines	1.30	1.73			-13.39	-1.30		
Oil	1.03	0.38			-7.20	-1.51		
Clothes	1.24	1.98			5.53	0.79		
Durables	1.46	**4.47**	*		-1.95	-0.33		
Chems	1.39	**4.46**	*		-1.24	-0.24		
Cnsum	0.58	**-6.22**	*		4.15	1.05		
Constr	1.27	**2.75**			-1.12	-0.20		
Steel	1.76	**6.84**	*	*	-22.89	**-3.50**	*	
FabPr	1.12	1.58			1.30	0.30		
Machn	1.20	**3.01**			-2.43	-0.62		
Auto	1.33	**2.36**			4.27	0.53		
Trans	1.15	**2.71**			-1.34	-0.41		
Utils	0.49	**-6.48**	*	*	-0.49	-0.11		
Retail	0.73	**-4.46**	*		4.64	1.29		
Finance	1.38	**5.48**	*		-9.99	**-2.46**		

Problem 2: SUR and Wald Tests:

This is interesting but not quite satisfactory. You want a YES/NO answer to each question. You know that as these are tests of βs and αs across equations, you need the SUR framework. Let's move to more industries. For Problem 2, use 30indus-mon, same period 2009-2014

While doing your SUR answer these questions

- *Are the OLS βs "correct" or we need to use SUR to compute a different β_{GLS}?*

  We regress all 30 Ys on the same X, the market excess return.
  So we know that the OLS and SUR $\hat{\beta}$s are equal. The OLS $\hat{\beta}$s are "correct".

- *What are the typical (average over the 30) R^2 and stdev($\hat{\beta}$) for these regressions.*

  The average R^2 is **0.69**
  The typical (average) standard error for a large portfolio $\hat{\beta}$ with a 60 monthly return regression is **0.10**

Be sure to not count the 30 correlations of 1 when answering these questions!

ANSWER HERE

- *What is the average correlation of the 30 industries?* **0.68**
- *What is the average correlation of the 30 residuals?* **0.03**

- *Consider these 435 paiwise residual correlations you just estimated, what is their standard deviation*:

  0.25

- *The asymptotic standard deviation of a correlation estimate is – you won't believe !! $1/\sqrt{T}$*
  Compare this with your "realized" correlation of the 435 estimates. With lots of ifs – e.g., if the 120 correlations are independently estimated, answer qualitatively.
  Does it look like all these non-zero correlations could be just due to estimation error?

  The asymptotic standard deviation is $1/\sqrt{60} = \mathbf{0.13}$ when the true correlation is 0

  If the 435 correlation estimates were independent and coming from a true zero correlation, we would expect a zero mean (we get 0.03 which is basically zero) and a 0.13 standard deviation. In fact we get **0.25**, **twice as much**.

  This 0.25 spread seems to indicate that the true correlation is not zero for all pairs i,j. Regressing the 30 industries on the market explained **a lot but not all** their co-movements.

  Interesting question: What if it's the asymptotic $1/\sqrt{T}$ that's wrong?
  You can check this with a simulation. See the R file

a) You wonder if all these industries have the same systematic risk. Do a Hotelling T^2 test of H_0 that the **30 βs** are equal. You know to write your test under H_0 as $R\,\beta = r$.

- What are the dimensions of the matrix R and what elements does it contain?

  The R matrix is 29x30, we only need 29 restrictions. It's a Toeplitz matrix with top row: (1,-1,0…0).

- *What is in the vector r:*

  We test. $R\beta = r$, where r is a vector of zeros

- *Write the formula for the Wald Test:*

  Following the note: $(R\hat{\beta})'\ (R\,V(\hat{\beta})\,R')^{-1}\,R\hat{\beta}$

- *What Wald test statistic do you find:* **$T^2 = 303$**

- *What are the asymptotic distribution and 5% level cutoff for the test.*

  Asymptotic **$\chi^2(29)$** cutoff **43**

- *Reject or Accept (use the asymptotic cutoff)?* **Oh Yes !!**

- *Given what you know of the distribution of the Hotelling T^2 statistic in exact sample, does the asymptotic cutoff for the Wald test look appropriate?*

  Hotelling exact cutoff is: (60-1)/(60-29) 29 F(29, 60-29)
  1.9 5% point is: **53.21** **101**

  Asymptotic cutoff totally inappropriate. It does not change the result though.

b) *This multivariate regressions is well-known as the* ***1-pass method*** *to test and Asset Pricing Model: take a bunch of portfolios, estimate their βs and αs simultaneously.*
CAPM says the αs should be zero. Accept or Reject, done!
You want to do this for these 30 industries. Do a Hotelling T2 that all the αs are zero.

$T^2 = 85$

Accept /Reject: Yes with $\chi^2(30)_{0.95} = 44$, but that's not right.
No with exact distribution, cutoff is **108**

Problem 3: Sorting on Estimates

You are interested in the 2-pass approach because it has a predictive content and is used both for testing the models and by practitioners to control risk while attempting to forecast returns . This is why you are really excited about the upcoming lecture on Thursday Feb. 28th, where you will learn more about it. You heard a lot has to do with estimating risk from one period and using it for the next … or something like that.

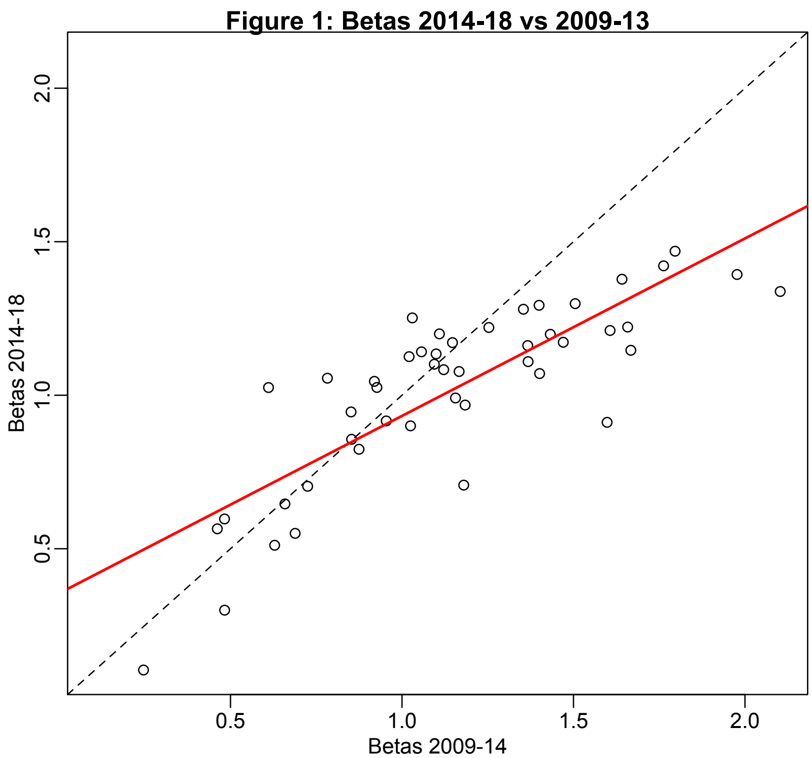
You wonder if βs are stable from period to period. There are two issues, estimates are noisy but maybe true βs move around. You know one thing: Since ϵ ▯ $N(0,\sigma^2 I)$, the βs are estimated with no bias. So there should not be a bias from one period to the next …. unless true βs vary with time in a systematic manner.

*You are going to use **47indus** for this experiment.*

a) Consider for period 1: 200901-201312   and period 2: 201401-201812. Estimate the β vector for each period. In Figure 1 plot β_2 vs β_1. Regress the 47 β_2s against their first period counterparts. Write the estimates below:

$$\hat{\beta}_2 = \mathbf{0.35} + \mathbf{0.58}\,\hat{\beta}_1$$

Add the regression line to your Figure 1, as well as the 45 degree line.



- *Using the regression. For $\widehat{\beta}_1$s = (0.5, 1, 1.5) what $\widehat{\beta}_2$ do you forecast in the next 60-month period?*

  0.5 → 0.35 + 0.58 * 0.5 = **0.64**
  1 → 0.35 + 0.58 * 1 = **0.93**
  0.84 → 0.84
  1.5 → 0.35 + 0.58 * 1.5 = **1.22**

- *Is there a bias here, which way, for what ranges of values of $\widehat{\beta}_1$?*

  High betas in period one go down in period 2
  Low betas in period one go up in period 2
  Mid range betas seem to be unbiased!

- *What could be the reason?*

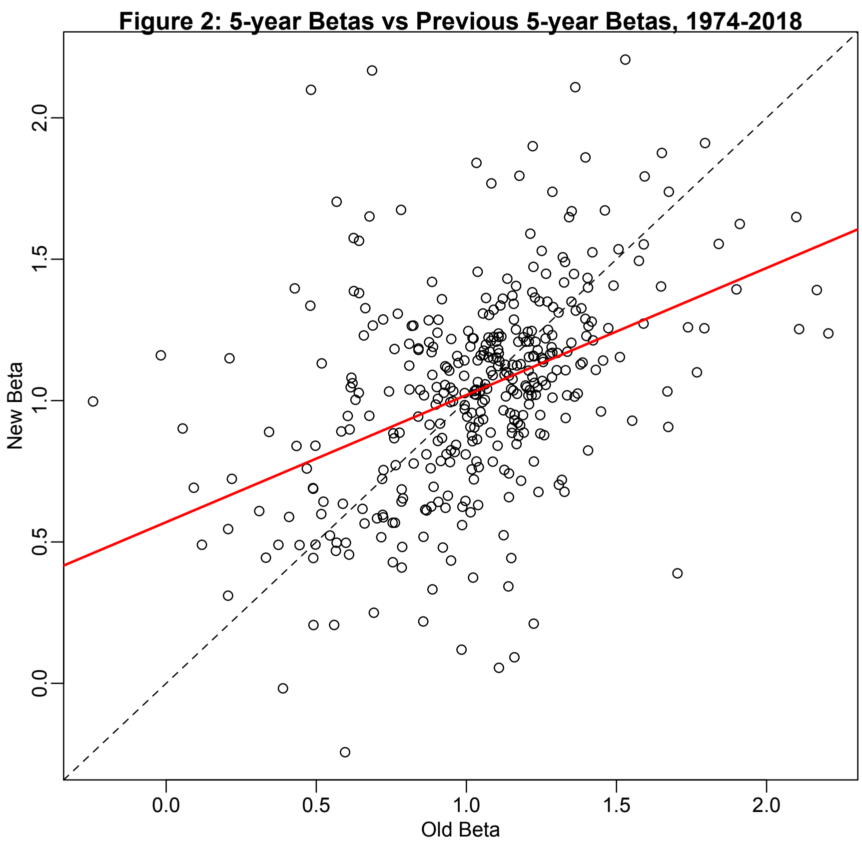
  Could it be specific to this period?

  Or is it (in fact it is !) due to the fact that high betas likely have >0 estimation errors and low betas likely have negative estimation error. The next period, it will be different betas that get >0 or <0 estimation error

b) This is puzzling! You wonder if it is specific to 2009-13, 2013-18. So you repeat the whole thing starting in 197401. This allows you to have exactly 9 periods of 5 years. Estimate the βs for each 5-year period. Then create two vectors: Oldbet and Newbet. Oldbet will have the βs estimated during the first 8 periods, Newbet the β estimates from periods 2 to 9. Each has 47x8 estimates to be sure. Then just redo the plot in Figure 2 and the regression.

$$\hat{\beta}_2 = \mathbf{0.57} + \mathbf{0.45}\,\hat{\beta}_1$$

$1.5 \rightarrow 0.57 + 0.45 * 1.5 = 1.24$
$0.5 \rightarrow 0.57 + 0.45 * 0.5 = 0.79$
$1 \rightarrow 1.02$



Any change? No, same phenomenon, it was not due to the 2009-2018 period

- *What could be the reason for this?*

Sorting on estimates, implicit or explicit, is dangerous ! Whenever we sort a vector of estimates, the positive estimation errors are likely to be in the top and the negative estimation errors in the bottom.

To be followed !