14.33 Event Studies: Stacking data and statistical power **Fall 2006**
TA: Maisy Wong

Goal: Test if there's an impact of event A on the stock returns of company 1, r_1

Data: We have historical stock returns of company 1 and 2, r_1 and r_2, and market wide index, r_m. Company 1 and company 2 are in the same industry. We have this data from 19900102 to 19991231. Event A happened on 19950414. We will use a 1-day event window.

Table 1: Dataset (wide)

date	r_1	r_2	r_m
19900102	$r_{1,19900102}$	$r_{2,19900102}$	$r_{m,19900102}$
…	…	…	…
19991231	$r_{1,19991231}$	$r_{2, 19991231}$	$r_{m, 19991231}$

Set-up:

The model we are going to use is:

$$r_{1t} = \alpha + \gamma_1\ 1(\text{date}=19950414) + \beta_1 r_{mt} + e_{1t} \qquad (1)$$

Our model says, on April 14th, 1995, $r_{1,19950414} = \alpha + \gamma_1 + \beta_1 r_{m,19950414}$

For other dates, $r_{1t} = \alpha + \beta_1 r_{mt}$

α is the excess return on company 1. But, on 19950414, the excess return is $\alpha + \gamma_1$. If there were no other events on April 14th, we interpret the impact of the event as γ_1. Note that we have N obs and we are estimating 3 parameters (α, γ_1 and β_1). Here, the degrees of freedom is N-3. (N is the number of trading days in the dataset)

Statistical power:

Suppose we ran the regression and found that the coefficient γ_1 was not significant (ie. we failed to reject the null hypothesis, γ_1=0). This could mean either:

- The event has no impact
- Our test has no statistical power (the test fails to reject even if the null is false).

Recall, the formula of the t-statistic is $\frac{\hat{\gamma}}{se(\hat{\gamma})}$. We could fail to reject if the numerator is small or the denominator is large. One way to get more statistical power is to reduce the standard errors of our estimate. How do we reduce standard errors? We can bring in more information.

Step-wise, what we need to do to improve statistical power is:

I. Bring in extra information (this will be in the form of some hypothesis/restrictions on the regression equation)
II. Test if we can use the extra information (ie. if we can reject the hypothesis)
III. If the test is ok, we can go ahead and incorporate the extra information in our model

I. Extra information:

Suppose we know that company 1 and company 2 are in the same industry and we think they should be affected equally by event A. Then, we can use this information to help us.

II. Testing (stacking data):

It is tempting to run two separate regressions, $r_{it} = \alpha + \gamma_i \; 1(\text{date}=19950414) + \beta_i \, r_{mt} + e_{1t}$ for i=1,2 and see if $\hat{\gamma}_1 = \hat{\gamma}_2$. However, we can't test the hypothesis that $\gamma_1 = \gamma_2$ because the dependent variables are different. So, the two regressions are not comparable. Instead of running 2 regressions that estimate 3 parameters each, what we can do is stack the data and run 1 regression estimating 6 parameters.

$$r_t = \alpha_1 1(\text{comp1}) + \gamma_1 \; 1(\text{date}=19950414)*1(\text{comp1}) + \beta_1 r_{mt}*1(\text{comp1}) + \alpha_2 1(\text{comp2}) + \gamma_2 \; 1(\text{date}=19950414)*1(\text{comp2}) + \beta_2 r_{mt}*1(\text{comp2}) + e_t \quad (2)$$

where 1(compi) is the dummy for company i. What the model says is:

For company 1, $r_t = \alpha_1 + \gamma_1 + \beta_1 r_{mt}$ on April 14th 1995, and $r_t = \alpha_1 + \beta_1 r_{mt}$. Likewise for company 2. Note that instead of dropping 1 of the company dummies, we impose that there is no constant.[1] This way, we get to estimate the excess returns of both companies.

Notice that the dependent variables in (1) and (2) are different and there are a few new variables. To operationalize (2), we need to:

i. Stack r_1 and r_2 into one column, r (see Table 2). Likewise for r_m. The STATA command for stacking is reshape
ii. Create event dummies
iii. Create company dummies, C1 and C2

[1] Suppose the model was $r_t = \alpha_0 + \alpha_1 1(\text{comp1}) + \gamma_1 1(\text{date}=19950414)*1(\text{comp1}) + \beta_1 r_{mt}*1(\text{comp1}) + \alpha_2 1(\text{comp2}) + \gamma_2 \; 1(\text{date}=19950414)*1(\text{comp2}) + \beta_2 r_{mt}*1(\text{comp2}) + e_t$ where α_0 is a constant. Then, we would run into multicollinearity problems, from α_0, α_1 and α_2 (because α_0, α_1 and α_2 would be linear combinations of each other). The solution is typically to drop one of them. Here, we drop α_0.

iv. Create interaction variables for the event dummy and both company dummies as well as interaction variables for r_m and both company dummies

Table 2: Dataset (long)

date	r	r_m	C1	C2	Event dummy	C1* Event	C2* Event	C1* r_m	C2* r_m
19900102	$r_{1,19900102}$	$r_{m,19900102}$	1	0	0	0	0	$r_{m,19900102}$	0
…	…	…	1	0	0	0	0	…	0
19950414	…	…	1	0	1	1	0	…	0
…	…	…	1	0	0	0	0	…	0
19991231	$r_{1,19991231}$	$r_{m, 19991231}$	1	0	0	0	0	$r_{m, 19991231}$	0
19900102	$r_{2,19900102}$	$r_{m,19900102}$	0	1	0	0	0	0	$r_{m,19900102}$
…	…	…	0	1	0	0	0	0	…
19950414	…	…	0	1	1	0	1	0	…
…	…	…	0	1	0	0	0	0	…
19991231	$r_{2, 19991231}$	$r_{m, 19991231}$	0	1	0	0	0	0	$r_{m, 19991231}$

Now, we are ready to run regression (2), with r as the dependent variable, the 2 company dummies and the 4 interaction variables as the independent variables. Notice that now we have 2N obs and we are estimating 6 parameters. So, we have 2N-6 degrees of freedom which is no different than estimating two separate regressions with N-3 degrees of freedom.

The thing that buys us more statistical power is the information that company 1 and company 2 may be affected equally by the event ($\gamma_1 = \gamma_2$). If this is true, then we only need to estimate α_1, α_2, γ_1, β_1 and β_2. And we get γ_2 using the information that $\gamma_1 = \gamma_2$. Now, our degrees of freedom is 2N-5.

In STATA, we need to run regression (2). Then, test the hypothesis, $\gamma_1 - \gamma_2 = 0$.

III. Constrained regression:

If we fail to reject the null hypothesis that $\gamma_1 - \gamma_2 = 0$. Then, we can try to run regression (2) constraining $\gamma_1 = \gamma_2$. The STATA command is cnsreg. If we rejected the hypothesis, then we have no grounds to run the constrained regression.