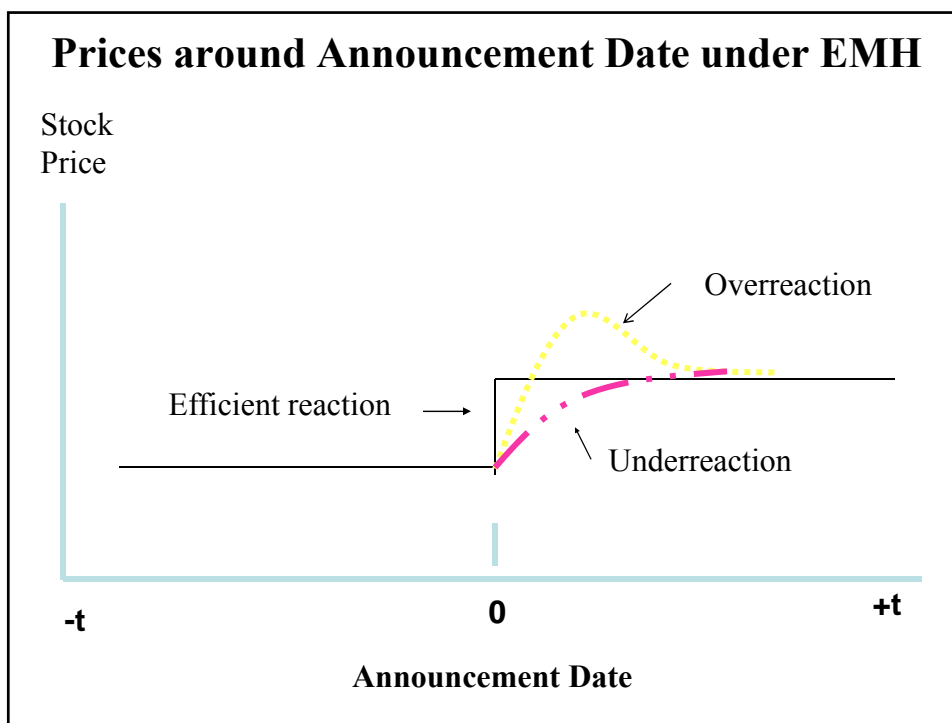


Lecture 6

Event Studies

Event Study Analysis

- Definition: An event study attempts to measure the valuation effects of a corporate event, such as a merger or earnings announcement, by examining the response of the stock price around the announcement of the event.
- One underlying assumption is that the market processes information about the event in an efficient and unbiased manner.
- Thus, we should be able to see the effect of the event on prices.



- The event that affects a firm's valuation may be:
 - 1) within the firm's control, such as the event of the announcement of a stock split.
 - 2) outside the firm's control, such as a macroeconomic announcement that will affect the firm's future operations in some way.
- Various events have been examined:
 - mergers and acquisitions
 - earnings announcements
 - issues of new debt and equity
 - announcements of macroeconomic variables
 - IPO's
 - dividend announcements.
 - etc.

- Technique mainly used in corporate finance (not economics).
- Simple on the surface, but there are a lot of issues.
- Long history in finance:
 - First paper that applies event-studies, as we know them today: Fama, Fisher, Jensen, and Roll (1969) for stock splits.
 - Today, we find thousands of papers using event-study methods.
- This is also known as an event-time analysis to differentiate it from a calendar time analysis.

Classic References

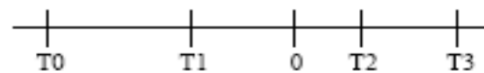
- Brown and Warner (1980, 1985): Short-term performance studies
- Loughran and Ritter (1995): Long-term performance study.
- Barber and Lyon (1997) and Lyon, Barber and Tsai (1999): Long-term performance studies.
- Eckbo, Masulis and Norli (2000) and Mitchell and Stafford (2000): Potential problems with the existing long-term performance studies.
- Ahern (2008), WP: Sample selection and event study estimation.
- Updated Reviews:
 - M.J. Seiler (2004), **Performing Financial Studies: A Methodological Cookbook**. Chapter 13.
 - Kothari and Warner (2006), Econometrics of event studies, Chapter 1 in **Handbook of Corporate Finance: Empirical Corporate Finance**.

Event Study Design

- The steps for an event study are as follows:
 - Event Definition
 - Selection Criteria
 - Normal and Abnormal Return Measurement
 - Estimation Procedure
 - Testing Procedure
 - Empirical Results
 - Interpretation

Time-line

- The time-line for a typical event study is shown below in event time:



- The interval T0-T1 is the estimation period
- The interval T1-T2 is the event window
- Time 0 is the event date in calendar time
- The interval T2-T3 is the post-event window
- There is often a gap between the estimation and event periods

- Issues with the Time-line:
 - Defintion of an event: We have to define what an event is. It must be unexpected. Also, we must know the exact date of the event. Dating is always a problem (WSJ is not a good source -leakage).
 - Frequency of the event study: We have to decide how fast the information is incorporated into prices. We cannot look at yearly returns. We can't look at 10-seconds returns. People usually look at daily, weekly or monthly returns.
 - Sample Selection: We have to decide what is the universe of companies in the sample.
 - Horizon of the event study: If markets are efficient, we should consider short horizons –i.e., a few days. However, people have looked at long-horizons. Event studies can be categorized by horizon:
 - Short horizon (from 1-month before to 1-month after the event)
 - Long horizon (up to 5 years after the event).

Short and long horizon studies have different goals:

- Short horizon studies: how fast information gets into prices.
- Long horizon studies: Argument for inefficiency or for different expected returns (or a confusing combination of both)

Models for measuring normal performance

- We can always decompose a return as:

$$R_{i,t} = E[R_{i,t} | X_t] + \xi_{i,t},$$

where X_t is the conditioning information at time t :

- In event studies, $\xi_{i,t}$ is called the “abnormal” return.
- Q: Why abnormal? It is assumed that the unexplained part is due to some “abnormal” event that is not captured by the model.
- In a sense, we want to get close to a natural experiment.
 - There is an exogenous (unanticipated) shock that affects some stocks.
 - We want to compare the returns of those stocks around the announcement to others that are not affected.

- Definition of “Normal” Returns: We need a benchmark (control group) against which to judge the impact of returns.
 - There is a huge literature on this topic.
 - From the CAPM/APT literature, we know that what drives expected stock returns is not exactly clear.
 - This is precisely what we need to do in event studies: We need to specify expected returns (we just call them “normal” returns).
 - Note that if we are looking at short horizon studies, we can assume that expected returns do not change. No problem, here.
 - If we are looking at long horizons, we know that expected returns change. Big problem. We have to be careful.
 - In long horizon studies, the specification of expected returns makes a huge difference, because small errors are cumulated. There is no easy way out of this problem.

Statistical or economic models for normal returns?

- Statistical models of returns are derived purely from statistical assumptions about the behavior of returns -i.e., multivariate normality.
 - Multivariate normality produces two popular models:
 - 1) constant mean return model and
 - 2) the market model.

Note: If normality is incorrect, we still have a least squared interpretation for the estimates.

Actually, we only need to assume stable distributions. See Owen and Rabinovitch (1983).

- Economic models apply restrictions to a statistical model that result from assumed behavior motivated by theory -i.e., CAPM, APT.
 - If the restrictions are true, we can calculate more precise measures of abnormal returns.
 - CLM: “there seems to be no good reason to use an economic model.”

Popular Statistical Models

Constant mean return model

- For each asset i , the constant mean return model assumes that asset returns are given by:

$$R_{i,t} = E[R_{i,t} | X_t] + \xi_{i,t}, \text{ where}$$

$$E[R_{i,t} | X_t] = \mu,$$

$$E[\xi_{i,t}] = 0 \text{ and } \text{Var}[\xi_{i,t}] = \sigma_{\xi,i}^2$$

- Brown and Warner (1980, 1985) find that the simple mean returns model often yields results similar to those of more sophisticated models because the variance of abnormal returns is not reduced much by choosing a more sophisticated model.

Market model (MM) (the most popular in practice)

- For each asset i , the MM assumes that asset returns are given by:

$$R_{i,t} = E[R_{i,t} | X_t] + \xi_{i,t}, \text{ where}$$

$$E[R_{i,t} | X_t] = \alpha_i + \beta_i R_{m,t},$$

$$E[\xi_{i,t}] = 0 \text{ and } \text{Var}[\xi_{i,t}] = \sigma_{\xi,i}^2$$

- In this model $R_{m,t}$ is the return on the market portfolio, and the model's linear specification follows from an assumed joint normality of returns.
 - Usually a broad-based stock index is used as the market portfolio (S&P 500 or the CRSP EW or CRSP VW).
 - When $\beta_i = 0$, we have the constant mean return model.
 - The MM improves over the constant mean return model: we remove from $\xi_{i,t}$ changes related to the return on the market portfolio.
 - The benefit of using the MM depends on the R^2 of the regression. (The higher the R^2 , the greater the power to detect abnormal performance.)

Other Models for Expected Returns

- **CAPM**

$$E[R_{i,t}|X_t] - r_{f,t} = \beta_i (E[R_{m,t}|X_t] - r_{f,t}),$$

- **Fama and French (1993) (FF) 3 factor model**

$$E[R_{i,t}|X_t] - r_{f,t} = a_i + b_{1i}(E[R_m|X_t] - r_{f,t}) + b_{2i}SML_t + b_{3i}HML_t$$

SML: returns on small (Size) portfolio minus returns on big portfolio

HML: returns on high (B/M) portfolio minus returns on low portfolio

Note: More factors can be easily added to this ad-hoc model, for example, a momentum factor –see, Carhart (1997) .

- **Sorts**

- Suppose that there are two factors that affect returns: Size and (B/M). We do not know whether there is a stable or linear relationship as the one specified in the FF model.

- What to do.

- Sort all returns in the universe (CRSP) into 10 deciles according to size.
- Conditional on size, sort returns into ten deciles according to BM. (This gives us 100 portfolios.)
- Compute the average return of the 100 portfolios for each period. This gives us the expected returns of stocks given the characteristics.
- For each stock in the event study:
 - 1) Find in what size decile they belong.
 - 2) Then, find in what B/M decile they belong.
 - 3) Compare the return of the stock to the corresponding portfolio return.
 - 4) Deviations are called “abnormal” return.

- Fact: Sorts give more conservative results. If we use the FF method, we tend to find huge abnormal returns, while with the sorts, we do not.
- Note:
 - Results change if we sort first by B/M and then size (not good).
 - Results change if we sort according to other characteristics.
 - Data-mining problem: We choose a sorting method that works after many trials. Out of 100 trials, there must be five that works, at the 5% level. Pre-testing problem, again!

Estimation of Abnormal Returns

- There are several ways of estimating “abnormal” returns. In addition to specifying expected returns, there is the issue of how to cumulate returns. There are two methods:

1. Cumulated Abnormal Returns (CARs)

$$AR_{i,t} = R_{i,t} - E[R_{i,t} | X_t]$$

$$CAR_{t,t+K}^i = \sum_k AR_{i,t+k}$$

Note: CARs are like prices -they are prices if we have log returns.

- If we fix K; we can compute the variance of the CAR. Then, under certain conditions:

$$CAR_{t,t+K}^i \sim N(0, \sigma_{i,t+k}^2)$$

- Sometimes we are looking at only several categories ($j = 1, \dots, J$) (IPO and non-IPO firms). Suppose there are N_1, \dots, N_j firms in each category.

- Then, the CAR for each category is:

$$\overline{CAR}_{t,t+K}^j = \frac{1}{N_j} \sum_{n=1}^{N_j} CAR_{t,t+k}^n$$

The advantage of aggregating across assets is immediately clear:

$$\overline{CAR}_{t,t+K}^j \sim N \left(0, \frac{1}{N_j^2} \sum \sigma_{i,t,t+K}^2 \right)$$

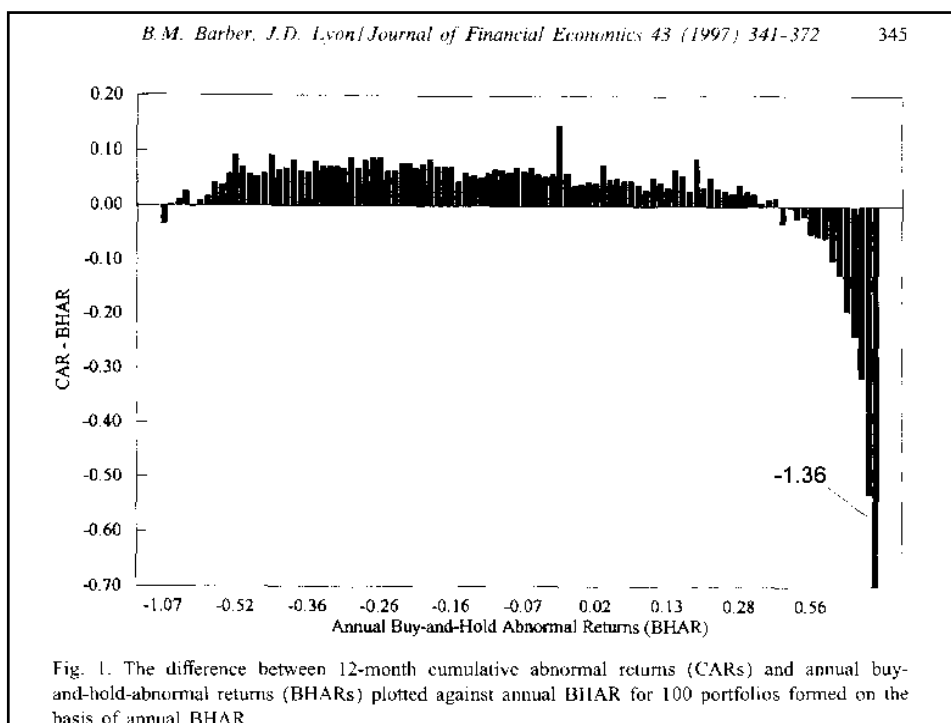
2. Buy and Hold Abnormal Returns (BHAR)

- Several economists (Ritter (1991), Barber and Lyons (1997), Lyons et al. (1999) have argued that CARs are not appealing on economic grounds. BL propose to use buy-and-hold returns, defined as:

$$\begin{aligned} AR_{i,t} &= R_{i,t} - E[R_{i,t} | X_t] \\ BHAR_{t,t+K}^i &= \prod_k (1 + AR_{i,t+k}) \end{aligned}$$

- Difference between CAR and BHAR: arithmetic versus geometric sums.

- As in the case of CARs, we can aggregate BHAR. The variance also is reduced for the same reasons.



- Barber and Lyons (1997) relate BHAR and CAR in a regression:

$$\text{BHAR}_t = -0.013 + 1.041 \text{ CAR}_t + e_t$$

CARs are a biased predictor of long-run BHAR. (There is a measurement bias.)

- Question: Which method to use: BHAR or CAR?
 - For short horizons, both are very similar.
 - For long horizons, BHAR seems conceptually better.
 - BHAR tend to be right skewed (bounded from below!)

Testing

- Null Hypothesis: Event has no impact on returns –i.e., no abnormal mean returns, unusual return volatility, etc.
- The focus is usually on mean returns.

- **Parametric Test.**

Traditional t-statistics (or variations of them) are used:

$$t_{CAR} = \overline{CAR}_{it} / (\sigma(CAR_{it}) / \sqrt{n})$$

or

$$t_{BHAR} = \overline{BHAR}_{it} / (\sigma(BHAR_{it}) / \sqrt{n})$$

- Appealing to a CLT, a standard normal is used for both tests.

- **Non-Parametric Tests**

Advantage: Free of specific assumptions about return distribution.

Intuition: Let $p = P(CAR_i \geq 0)$, then under the usual event studies hypothesis, we have $H_0: p \leq 0.5$ against $H_1: p > 0.5$. (Note if distribution of CAR_i is not symmetric, we need to adjust the formulation of p .)

Popular Tests: Sign Test (assumes symmetry in returns) and Rank Test (allows for non-symmetry in returns). See Corrado (1989).

- Example: Sign Test

Let N_+ be the number of firms with $CAR > 0$, and N the total number of firms in the sample. Then, H_0 can be tested using

$$J = [(N_+/N) - 0.5] \sqrt{2N} \sim^A N(0,1)$$

Usually, non-parametric tests are used as a check of the parametric tests.

Econometric Problems

There are many econometric problems in event studies. The problems can be divided into two categories:

- (i) Misspecifications of expected returns (wrong inference due to bias in the estimates of abnormal returns).
- (ii) Non-random sample, leading to non-normal distributions (wrong inference due to standard error calculations).

Bootstrap

- Non-normality (actually, an unknown distribution $F_0(x)$!) and limited number of observations can be a serious problem for testing. Bootstrapping can help to make inferences with good size.
- *Bootstrapping*: We estimate properties of an estimator by measuring those properties by sampling from an approximating distribution. Choice for an approximate distribution: the empirical distribution -EDF or $F_n(x)$.

Note: The EDF is a good choice. The Glivenko-Cantelli theorem that $F_n(x) \rightarrow F_0(x)$ as $n \rightarrow \infty$ uniformly over x almost surely.

Then, we can get approximate critical and p-values for an estimator by sampling from the EDF. This is the essential nature of the “bootstrap.”

- If the sample is a good indicator of the population, the bootstrap works.

- Under mild regularity conditions, the bootstrap yields an approximation to the distribution of an estimator or test statistic

The distribution is at least as accurate as the approximation obtained from first-order asymptotic theory.

Note: The bootstrap substitutes computation for mathematical analysis if calculating the asymptotic distribution of an estimator is difficult.

- Advantages: Simplicity, no parametric assumptions needed. Consistent estimates.
- Disadvantage: It's an approximation, though asymptotically consistent. In small or bad samples, the bootstrap will not work well.
- Keep in mind the key bootstrap analogy:
 The population is to the sample
 as the sample is to the bootstrap samples

• Intuition

We are interested in the behavior of the sample statistic W under H_0 . The CDF of W depends on the unknown $F_\theta(x)$.

The bootstrap estimates the CDF of W by carrying out a Monte Carlo simulation in which random samples are drawn from $F_n(x)$ (*We treat the sample as if it were the population.*)

We take B random samples with replacement from the sample data set. (*We are creating pseudo-samples.*)

Each of these random samples will be the same size n as the original set, the samples will be similar, but not be the same. Thus, each re-sample randomly departs from the original sample. (*We treat the pseudo-samples as realizations from the true population.*)

We calculate the statistic W under H_0 for each resample. W will vary slightly from the original sample statistic.

• Now, we have a relative frequency histogram of W under H_0 , which will help us to understand its distribution.

• **Basic Bootstrap Setup**

The bootstrap method consists of five basic steps:

- (1) Get a sample of data size n from a population.
- (2) Take a random sample of size n with replacement from the sample.
- (3) Calculate the statistic of interest W under H_0 for the random sample.
- (4) Repeat steps (2) and (3) a large number B times (say, 10,000).
- (5) Create a relative frequency histogram of the B statistics of interest W under H_0 (all estimated W have the same probability.)

As B approaches infinity the bootstrap estimate of the statistic of interest will approach the population statistic.

• *Simple Example: Outlier detection*

We have a random (iid) sample $\{Y_1, \dots, Y_n\}$. We observe a high y_0 .

Q: How far into the tail of the typical distribution is y_0 ?

The null hypothesis is y_0 belongs in the distribution.

With unknown μ and s , we use the statistic:

$$T = \frac{Y_0 - \bar{Y}}{s}, \text{ where } s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2}$$

The distribution of the statistic depends on the distribution of the RVs: \bar{Y} and s , which we assume unknown. (Or given a small sample size, CLT results may provide very poor finite sample approximation.)

Inferences can be made based on the following “simulation”:

Simulate 1000's of values of $T = (Y_0 - \bar{Y}) / s$ as follows:

1. Select a sample $Y_{11}, \dots, Y_{n1}, Y_{01}$ at random from the observed data Y_1, \dots, Y_n ;

let $T_1 = (Y_{01} - \bar{Y}_1) / s_1$, where \bar{Y}_1, s_1 are computed from Y_{11}, \dots, Y_{n1} .

2. Select a sample $Y_{12}, \dots, Y_{n2}, Y_{02}$ at random from the observed data Y_1, \dots, Y_n ;

let $T_2 = (Y_{02} - \bar{Y}_2) / s_2$, where \bar{Y}_2, s_2 are computed from Y_{12}, \dots, Y_{n2} .

...

B. Select a sample $Y_{1B}, \dots, Y_{nB}, Y_{0B}$ at random from the observed data Y_1, \dots, Y_n ;

let $T_B = (Y_{0B} - \bar{Y}_B) / s_B$, where \bar{Y}_B, s_B are computed from Y_{1B}, \dots, Y_{nB} .

Use the simulated data T_1, \dots, T_B to determine critical and p-values.

• *Example: Mutual Fund performance*

We have N returns from a Mutual Fund. We want to estimate Jensen's alpha (α_i).

Model for expected performance: Market (CAPM) Model

$$E[R_{i,t} - r_f | X_t] = \alpha_i + \beta_i (R_{m,t} - r_f)$$

H_0 (no abnormal returns): $\alpha_i = 0$. Calculate standard t-test, say t.

Bootstrap to establish a p-value for the t-statistic:

(1) Estimate $R_{i,t} - r_f = \alpha_i + \beta_i (R_{m,t} - r_f) + \varepsilon_{it}$ (Keep (a_i , b_i and e_{it}))

(2) Generate the simulated data under H_0 , using the residuals only:

$$R_{i,t} - r_f = 0 + b_i (R_{m,t} - r_f) + e_{it} \quad (\text{Do this N times.})$$

(3) Estimate the model using the generated data.

$$(R_{i,t} - r_f) = a_1 + b_1 (R_{m,t} - r_f) + \xi_{it} \quad (\text{Calculate t-test, say } \tau^*)$$

(4) Repeat (2) and (3) B Times. The bootstrap p-value of the t-test is the proportion of B samples yielding a $\tau^* \geq t$.

Note: This is a *conditional bootstrap*, since we are conditioning on $R_{m,t}$. We are really bootstrapping the marginal distribution of $R_i|R_m$.

In many situations, it is possible to do an *unconditional bootstrap*. In step 2), we can draw the residuals (e_i) and the explanatory variable (x_i).

The unconditional bootstrap is more conservative. It can have the interpretation of a test for spurious correlation.

(In a framework with dependent x_i 's, like in time series cases, the unconditional bootstrap will not work well.)

• *Bootstrap Inference for event studies: Steps*

- (1) Obtain an event sample of size n . Say $n=2,000$.
- (2) Draw a sample of 2,000 observations with replacement from the returns assuming no event.
- (3) Computations: Calculate expected returns, abnormal returns, CARs.
- (4) Repeat (2) and (3) a large number of times, say 5,000. We have 5,000 CARs computed under the null of no effect. We generate a distribution.
- (5) Compare the computed CAR from the event study to the bootstrapped distribution. If in the tails, reject the null of no effect.

- Ignoring serial dependence in the draws can severely flawed inferences: If the bootstrap is applied to dependent observation, the re-sampled data will not preserve the properties of the original data set.

Event samples are not always random samples. Voluntary event firms are unlikely independent observations. In particular, major corporate events cluster through time by industry:

=> positive cross-correlation of abnormal returns (standard errors need to be corrected. Traditional t-tests tend to overstate).

- Mitchell and Stafford (2000) find no abnormal returns of SEO, mergers and share repurchases, once they correct for the positive cross-correlations of event firm abnormal returns, while a standard bootstrap finds significant abnormal returns.

- *Block bootstrap* (accounts for dependences in returns or residuals):

- Build blocks that are long enough to remove most of the dependency structure.
- Recall that the dependence does not mean that the observations are independent when separated by more than 2 observations.
- Suppose we have, $n=200$ observations, we can break it up into 20 non-overlapping blocks of 10 each, then draw the blocks at random, with replacement. (Overlapping blocks can also be used, but less accurate. Hall, Horowitz, and Jing (1995).)

There is a method called “Tapered block bootstrap” to smooth the effects of boundaries between neighbored blocks.

- *Sieve Bootstrap* (dependent data with iid innovations): AR(q) process.

$$X_t = \sum_j \rho_j X_{t-j} + \varepsilon_t, \quad E[\varepsilon_t] = 0, \varepsilon_t \sim \text{iid}$$

- The parameters \mathbf{p} can be estimated by OLS or Yule-Walker equations (get \mathbf{p}). Get residuals, \mathbf{e} .
- The AR order q is usually estimated using AIC.
- Bootstrap resamples $\{X_t^*, t > p\}$ drawing ε_t with replacement. Create X_t^* using the AR process:

$$X_t^* = \sum_j \rho_j X_{t-j}^* + \varepsilon_t.$$

(In practice, we set $X_{t-p+1}^* = \dots = X_0^* = \text{Mean of } X \text{ -or zero if } X \text{ demeaned. It's also common to demean the residuals } \mathbf{e}.)$

References:

General Bootstrap: Efron (1979, 1982, 2003), Wu (1986).

Block Bootstrap: Politis and Romano (1994), Paparoditis and Politis (2001, 2002).

Sieve (Time Series) Bootstrap: Franke and Kreiss (1992).

Review Book: Horowitz (2001) in Handbook of Econometrics, Vol. 5.
Efron and Tibshirani (1993).

Good reference (Time Series): Berkowitz and Kilian, *Econometric Reviews* (2000).

Explaining CARs

- Once “abnormal” returns are found, it is common to try to explain CAR, or find whether CAR are correlated with some economic variables. That is, we want to run the regression:

$$CAR_{i,t} = \alpha + \delta X_{i,t} + v_{i,t}$$

usually, $X_{i,t}$ are firm characteristics.

- We can run this regression in the cross-section.
- Main problem with this regression. The OLS assumption $E[X_{i,t} v_{i,t}] = 0$ might be violated. => Inconsistent estimates!
(Endogeneity is the usual suspect in the cases where the event is within the firm’s control.)

- Consider the following:

$$AR_{i,t} = R_{i,t} - \alpha - \delta R_{m,t} - \phi X_{i,t}$$

where X are some firm characteristics.

People use the X characteristics to correctly anticipate returns (or events). For example, if banks are announcing write-offs, you might expect that the banks that have not announced them, will do so.

(Moreover, banks will try to make the announcement at a time that is most favorable. This introduces a truncation bias –it introduces a non-linearity in the model.)

- If we include X in the regression, then

$$E[AR_{i,t} | X_{i,t}] = 0.$$

It makes no sense to run: $AR_{i,t} = \alpha + \gamma X_{i,t} + \varepsilon_{i,t}$

- If we exclude X ; when it should be included, then we get (suppose that $R_{m,t}$ and $X_{i,t}$ are not correlated):

$$\begin{aligned} AR_{i,t}^* &= R_{i,t} - \alpha - \delta R_{m,t} \\ &= AR_{i,t} + \varphi X_{i,t} \end{aligned}$$

Now, we run: $AR_{i,t}^* = \alpha + \gamma X_{i,t} + \varepsilon_{i,t}^*$

Then,

$$E[\varepsilon_{i,t}^*, X_{i,t}] = E[\varepsilon_{i,t}, X_{i,t}] + \varphi E[X_{i,t}, X_{i,t}] \neq 0!$$

- Prabhala (1997) explains why this problem might not be too severe.

Q: What might be a potential solution?