BA870: Topics in Financial & Accounting Analytics

Lecture #10 (Thursday, April 21, 2022)

Professor Peter Wysocki

Topics: Financial Market Analytics: Stock Returns & Corporate Event Studies; It is Hard to Predict the Stock Market; Explainabilty

BOSTON UNIVERSITY

What is an Event Study?

- Using financial market data, an event study measures the impact of a specific event on the value of a firm.

- Examples include M&As, earnings announcements, capital issues, macro news announcements, index changes, and other corporate and market actions.

Applications of Event Study?

- Testing Stock Market Trading Strategies:
  - Determine the stock market reaction to an event (on the day, hour, minute, second of the announcement)
    - Are there possible benefits to predicting the likelihood of an event?
  - Determine if stock prices move in a <u>predictable</u> direction after an event (hours, days, weeks, months)

WRDS Query: Step 2

Step 2: Risk model:

- Market-Adjusted Model
- Market Model
- Fama-French Three Factor Model
- Fama-French Plus Momentum

- There are several different risk models. Click on each risk model's name for details,
- Here, we have chosen the default, the Market-Adjusted Model.

Risk Models

The Market-Adjusted Model uses abnormal returns defined in excess of the CRSP Value-weighted market return (assumes market beta of 1 or E(R)=RM(t)).

$$R = R_{t,j} - RM_t$$

Risk Models (cont.)

The Market Model uses abnormal returns defined according to the market model:

$$
\begin{aligned}
AR &= R - E(R) \\
&= R - \left(\alpha + \beta * (R_m - R_f)\right)
\end{aligned}
$$

Risk Models (cont.)

Fama-French Three-Factor Model uses abnormal returns defined with respect to Fama-French 3-factor model:

$$AR = R - (\alpha + \beta_1 \left(R_m - R_f\right) + \beta_2 * SMB + \beta_3 * HML)$$

Risk Models (cont.)

Fama-French Plus Momentum model uses abnormal returns defined with respect to Carhart (1997) model:

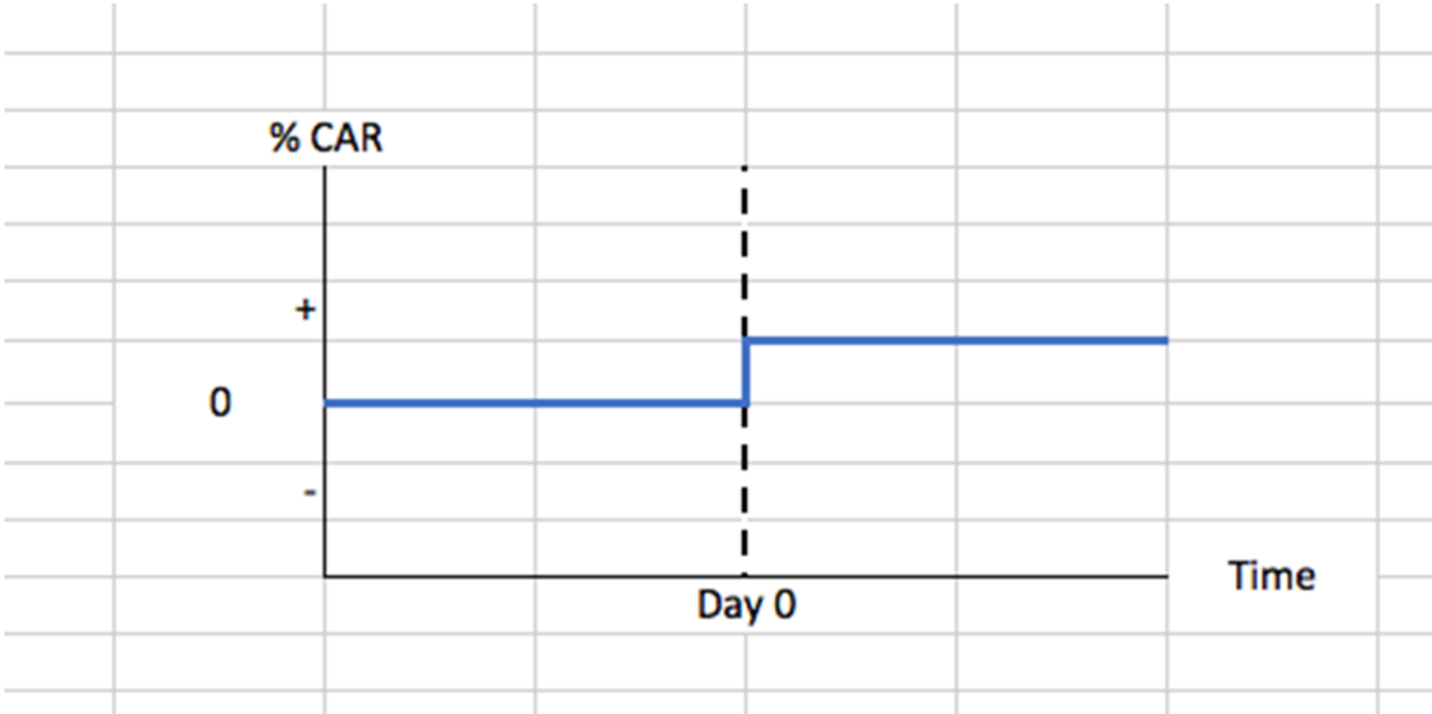
$$AR = R - (\alpha + \beta_1 * \left(R_m - R_f\right) + \beta_2 * SMB + \beta_3 * HML + \beta_4 * MOM\ )$$

Step Three: Analysis of the Data

- Examine the graphical results of the Event Study to begin your analysis.

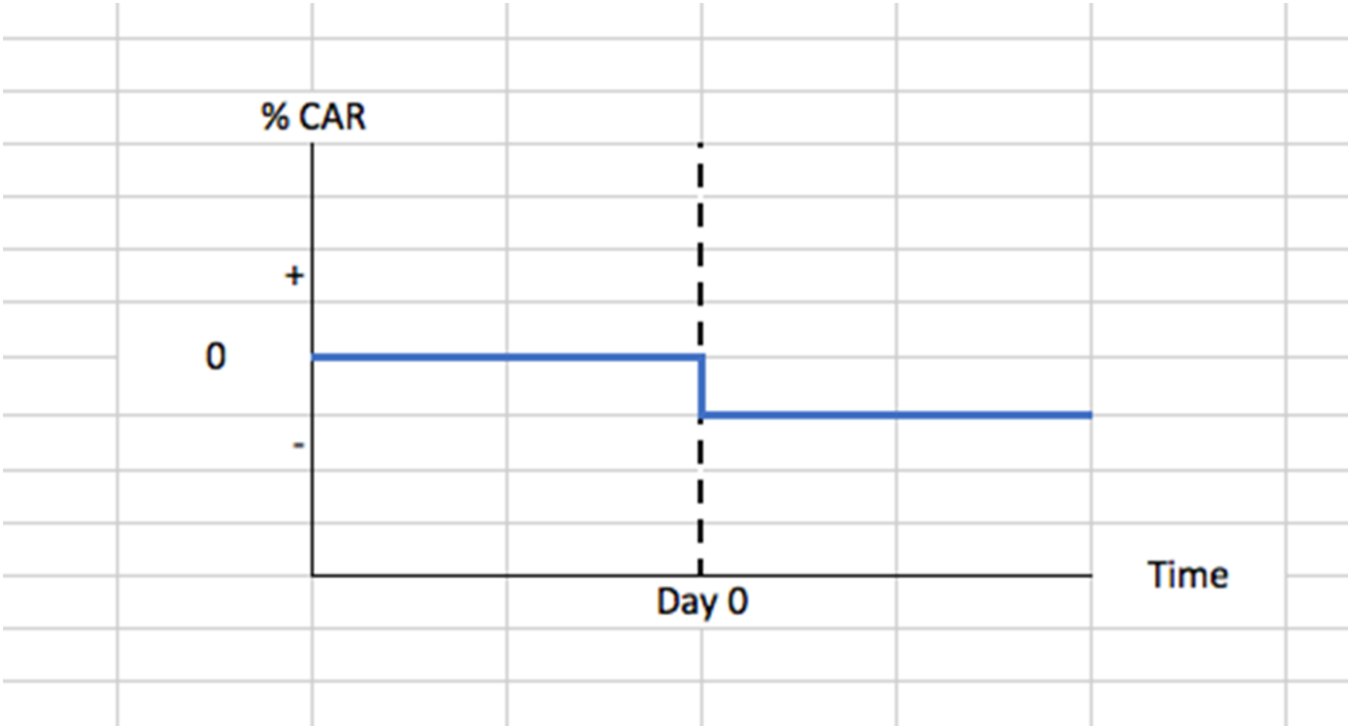
- The top graph shows the mean of Cumulative Abnormal Returns (CAR). This is a simple average of the cumulative abnormal returns.

Mean CAR Graph



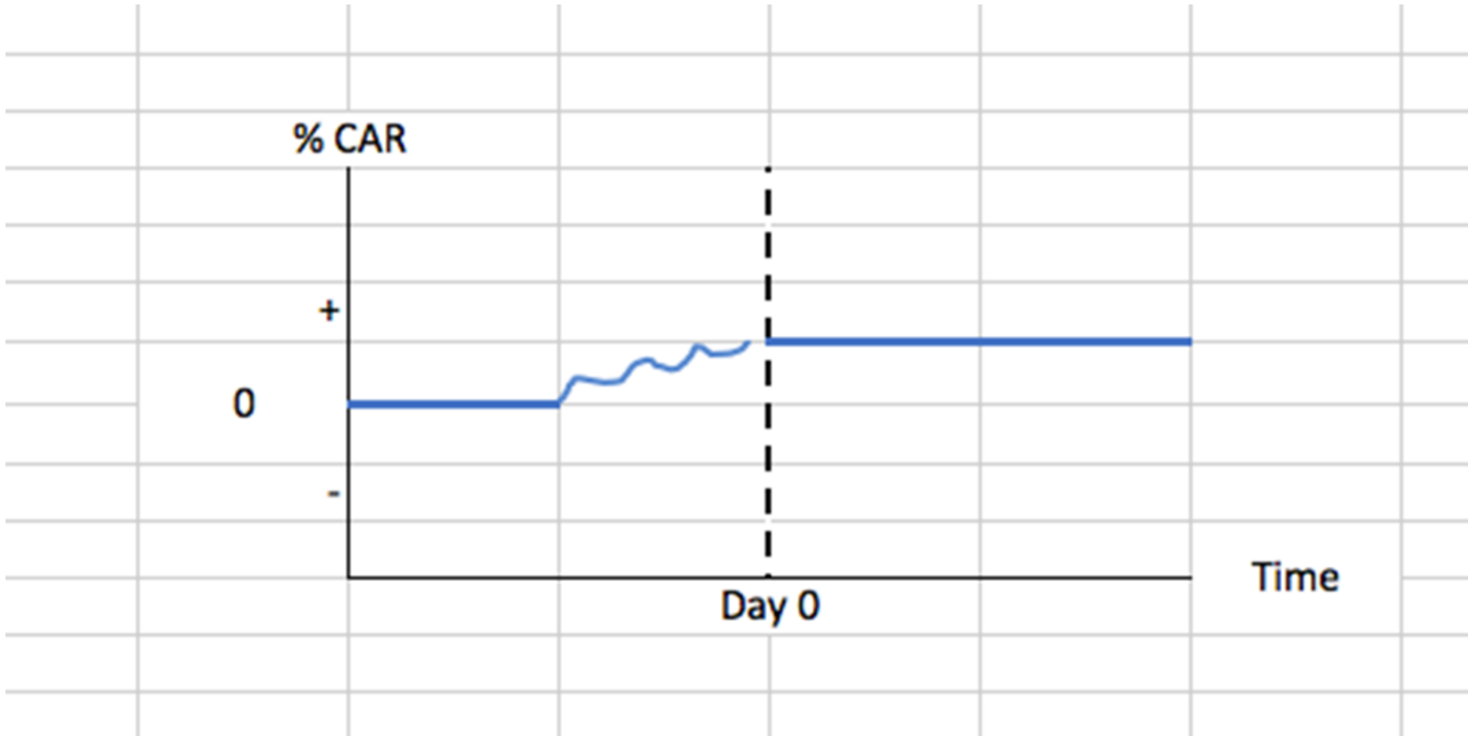
For example: a graph of the mean Cumulative Abnormal Return like this would indicate a positive, one-time impact of the event on the stock price. If no event had occurred, the line would stay near to 0.

Mean CAR Graph (cont.)



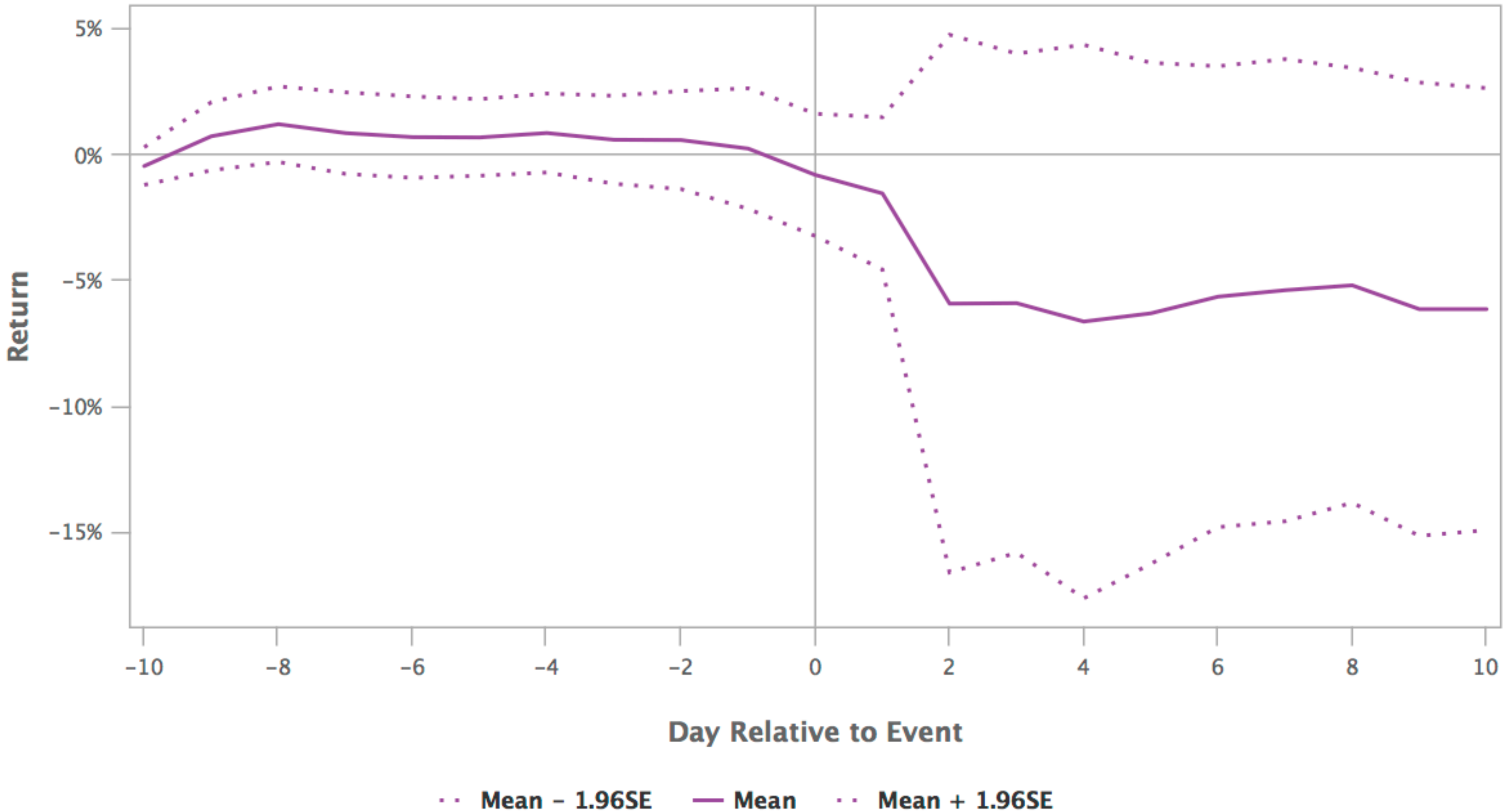
Conversely, a graph of the mean CAR like this would indicate a negative, one-time impact of the event on the stock price.

Mean CAR Graph (cont.)



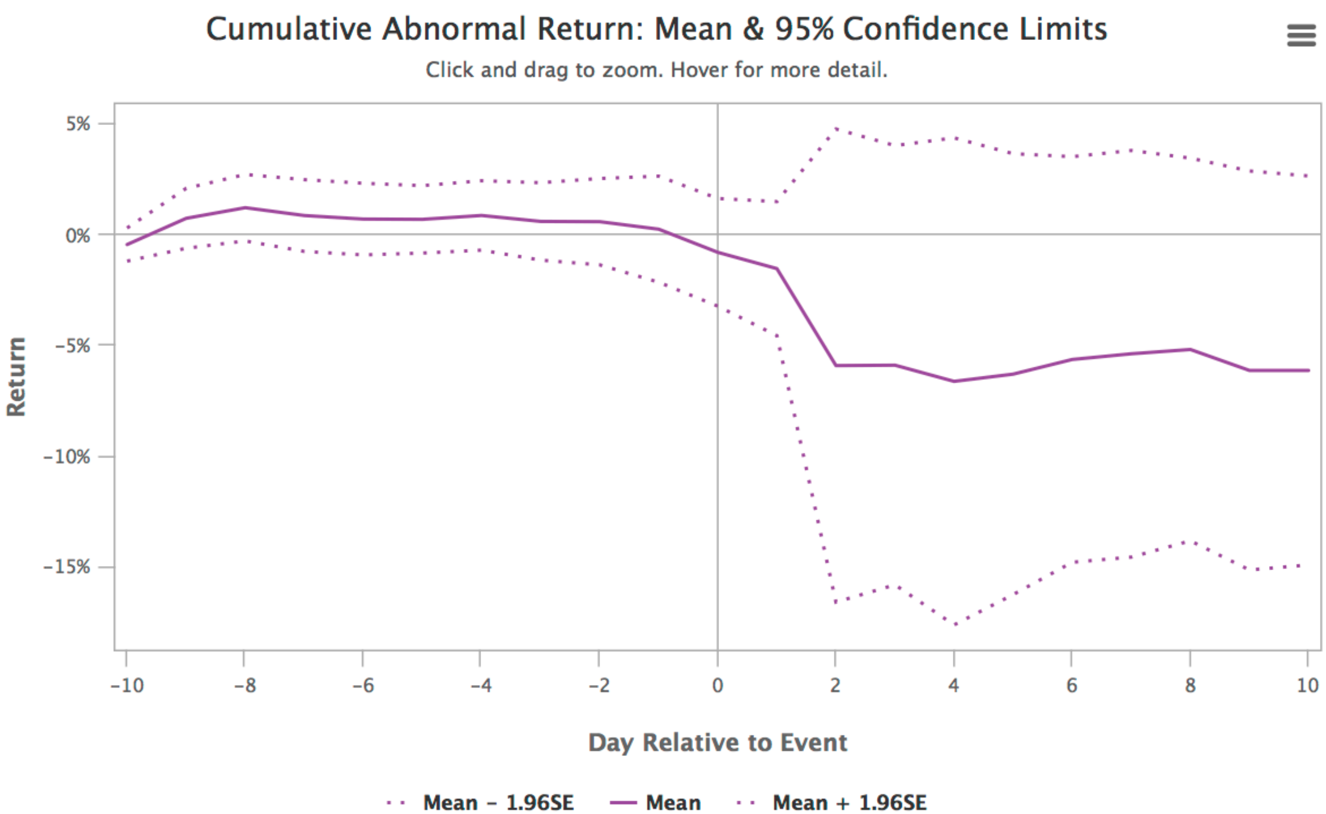
A graph of the mean CAR like this would indicate that news of the event may have preceded the event date, prematurely impacting the stock price in a positive direction.

Cumulative *Average* Abnormal Return and 95% Confidence Limits



BOSTON UNIVERSITY

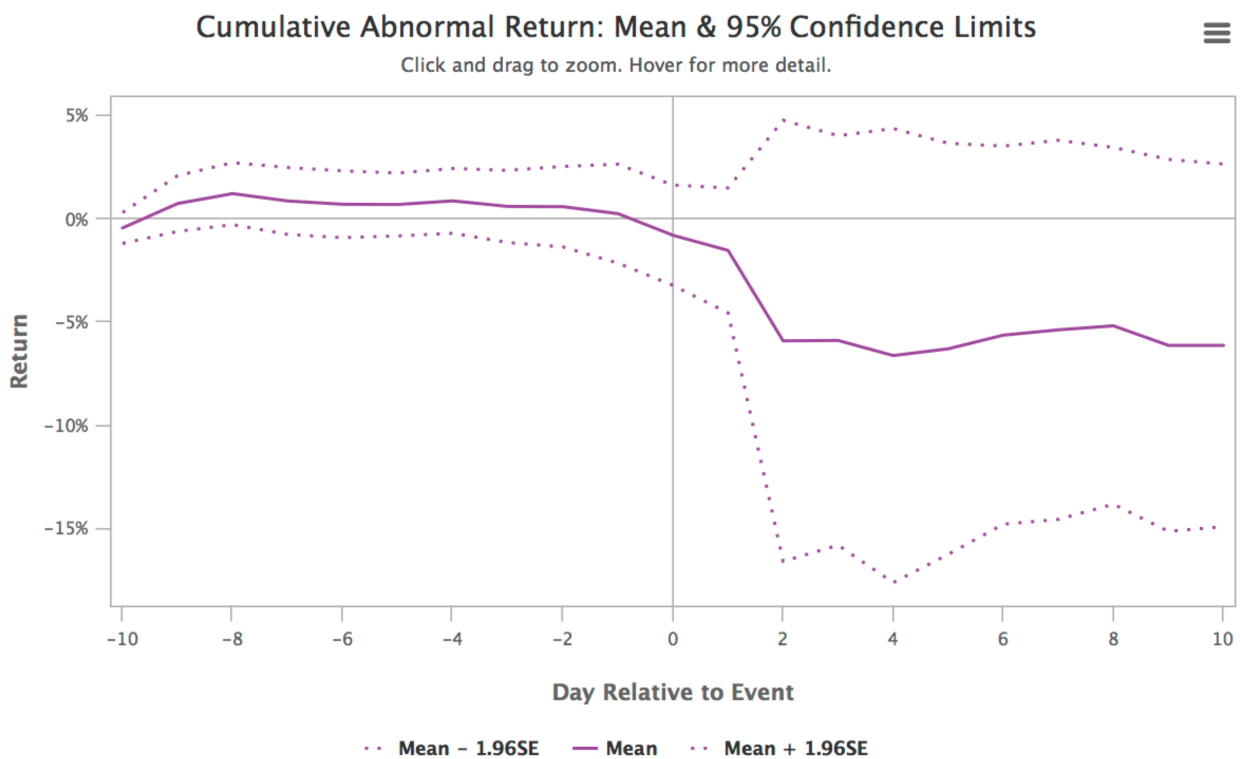
Mean CARs Graph Observations



- Observation 1: The mean Cumulative Abnormal returns for all companies drops by about 5%. This suggests an impact of the event on the returns.

BOSTON UNIVERSITY

Mean CARs Graph Observations (cont.)



- Observation 2: The dotted lines above and below the solid return line represent a plus/minus 2-standard deviation confidence interval.

Information and Value

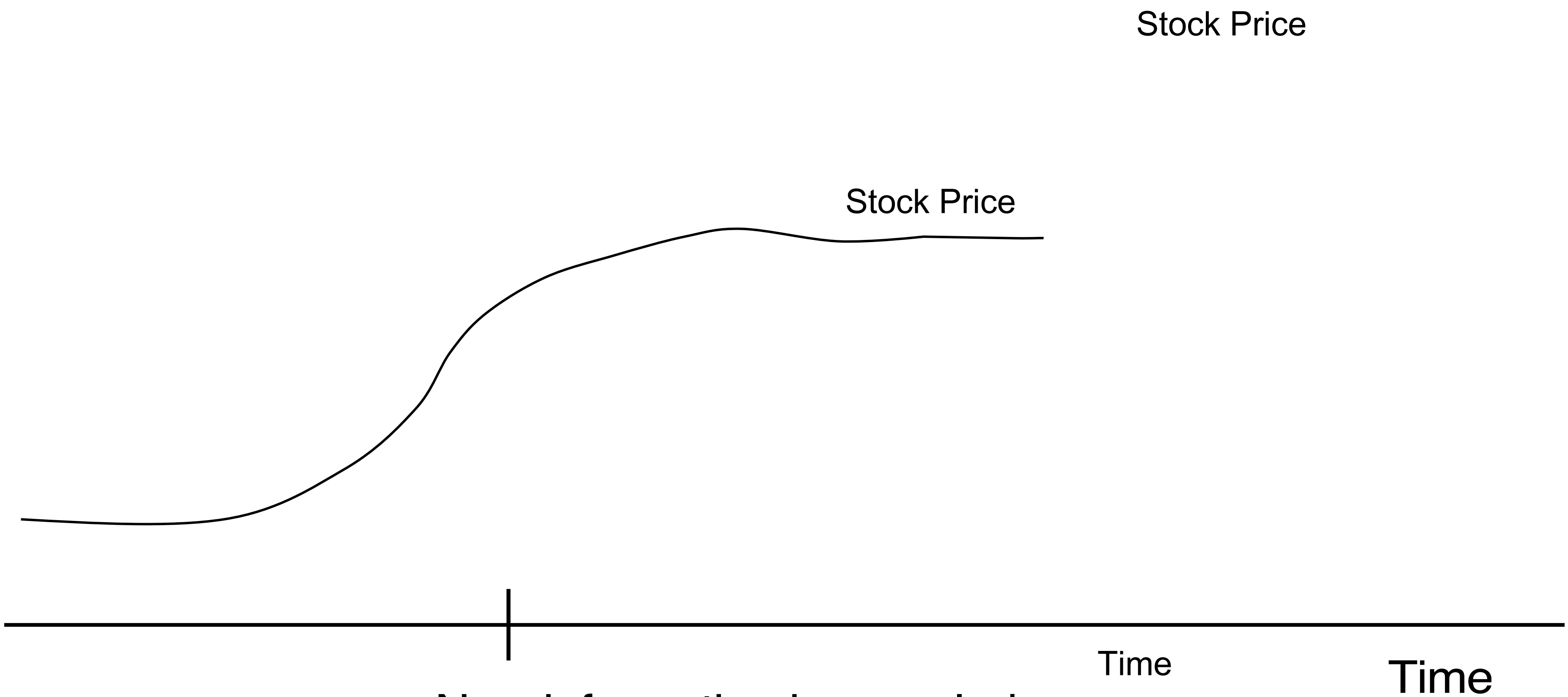
- Investors attempt to assess the value of an asset based upon the information that they have about that asset at that point in time.
- At the same time, different investors will arrive at different assessments of value for the same asset

–Because the information they have is different

–Because they have different ways of processing the same information

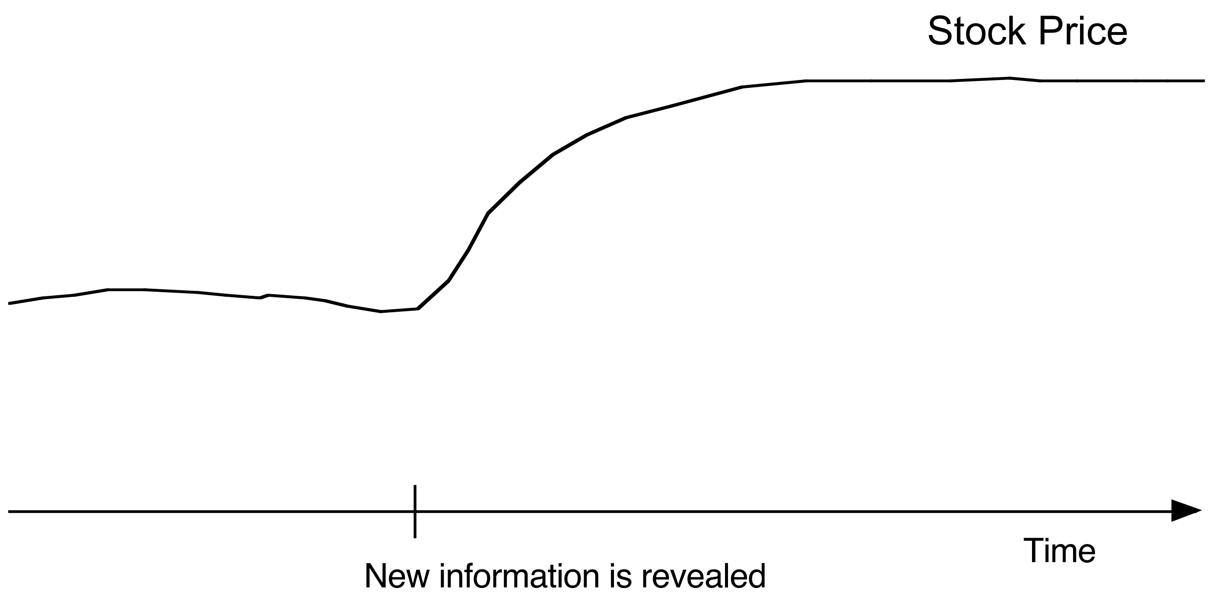
- The price is determined by demand and supply.

Information and Stock Prices: Base Case - Efficient Market

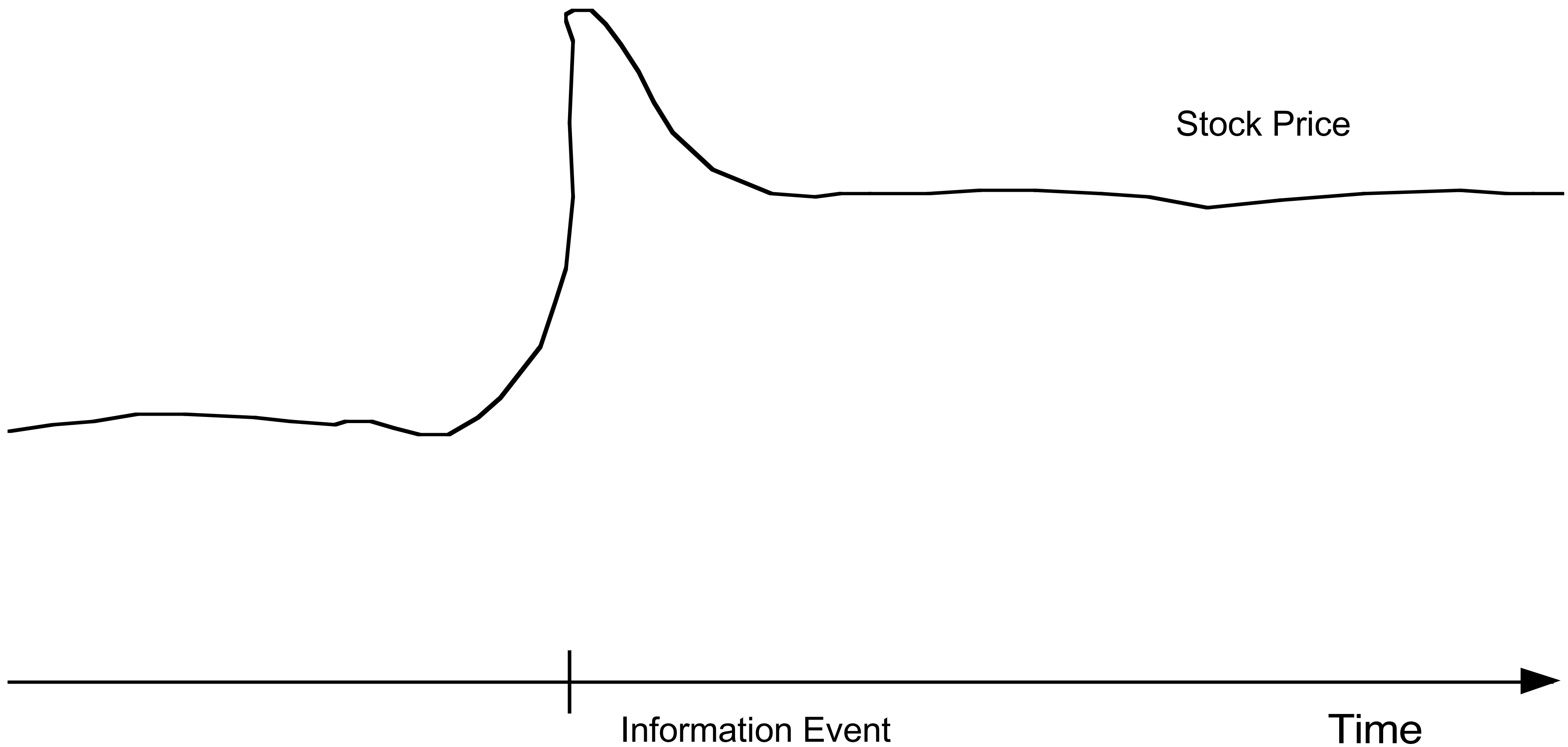


New information is revealed

A Slow Learning Market…

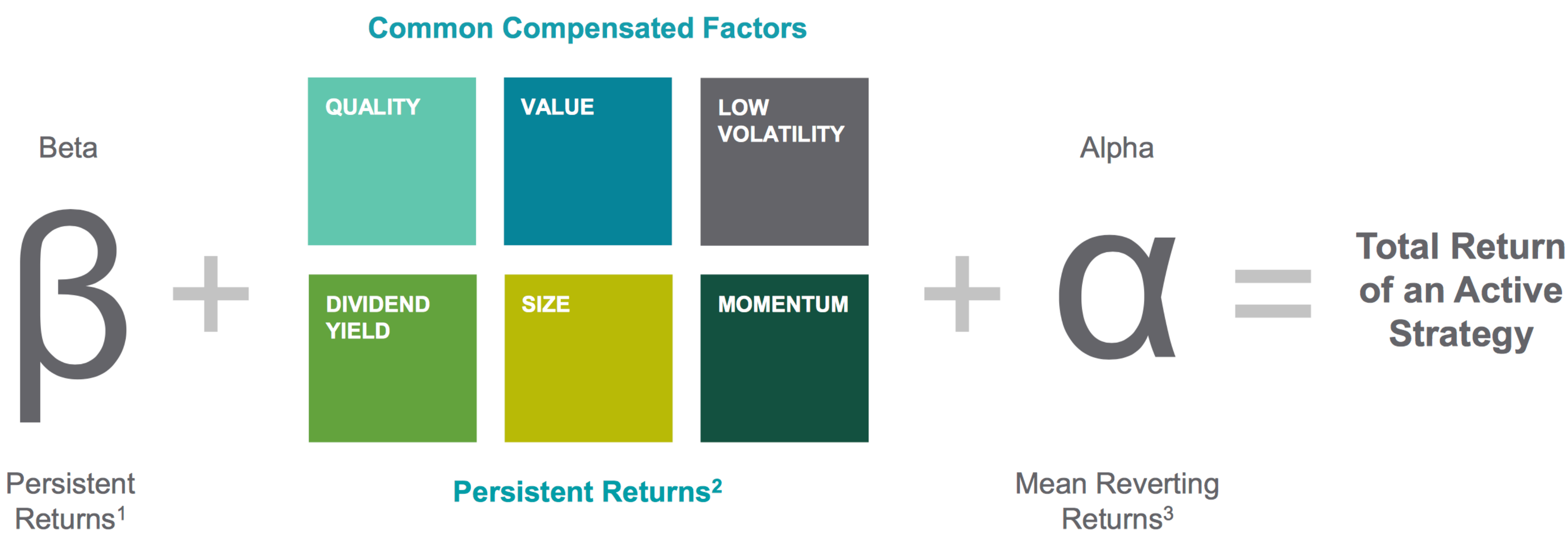


An Overreacting Market



Which Risk Factors Matter? Where to Find "Alpha"?

Performance Composition



Most Common Risk Factors - Research

- Size (**SMB** = small firms minus large firms): Banz (1981), Fama and French (1992), Fama and French (1993), Van Dijk (2011), Asness et al. (2018) and Astakhov, Havranek, and Novak (2019).

- Value (**HML** = high minus low: undervalued minus `growth' firms): Fama and French (1992), Fama and French (1993), C. S. Asness, Moskowitz, and Pedersen (2013).

- Momentum (**WML** = winners minus loser): Jegadeesh and Titman (1993), Carhart (1997) and C. S. Asness, Moskowitz, and Pedersen (2013). The winners are the assets that have experienced the highest returns over the last year (sometimes the computation of the return is truncated to omit the last month). Cross-sectional momentum is linked, but not equivalent, to time-series momentum (trend following), see e.g., Moskowitz, Ooi, and Pedersen (2012) and Lempérière et al. (2014). Momentum is also related to contrarian movements that occur both at higher and lower frequencies (short-term and long-term reversals), see Luo, Subrahmanyam, and Titman (2020).

- Profitability (**RMW** = robust minus weak profits): Fama and French (2015), Bouchaud et al. (2019). In the former reference, profitability is measured as (revenues - (cost and expenses))/equity.

- Investment (**CMA** = conservative minus aggressive): Fama and French (2015), Hou, Xue, and Zhang (2015). Investment is measured via the growth of total assets (divided by total assets). Aggressive firms are those that experience the largest growth in assets.

- Low `risk' (sometimes: **BAB** = betting against beta): Ang et al. (2006), Baker, Bradley, and Wurgler (2011), Frazzini and Pedersen (2014), Boloorforoosh et al. (2020), Baker, Hoeyer, and Wurgler (2020)and Asness et al. (2020). In this case, the computation of risk changes from one article to the other (simple volatility, market beta, idiosyncratic volatility, etc.).

Link with Machine Learning

- Large increase in data availability: temptation to determine if future returns can be predicted (“alpha”) from the abundance of attributes available at the firm level.
- Classical data like accounting ratios and alternative data (such as sentiment) can be used in conjunction with Machine Learning.
- Using a large set of predictor variables (financial, accounting, textual, alternative), goal is to predict a future performance (both earnings and stock returns).
- This is HARD!

BOSTON UNIVERSITY

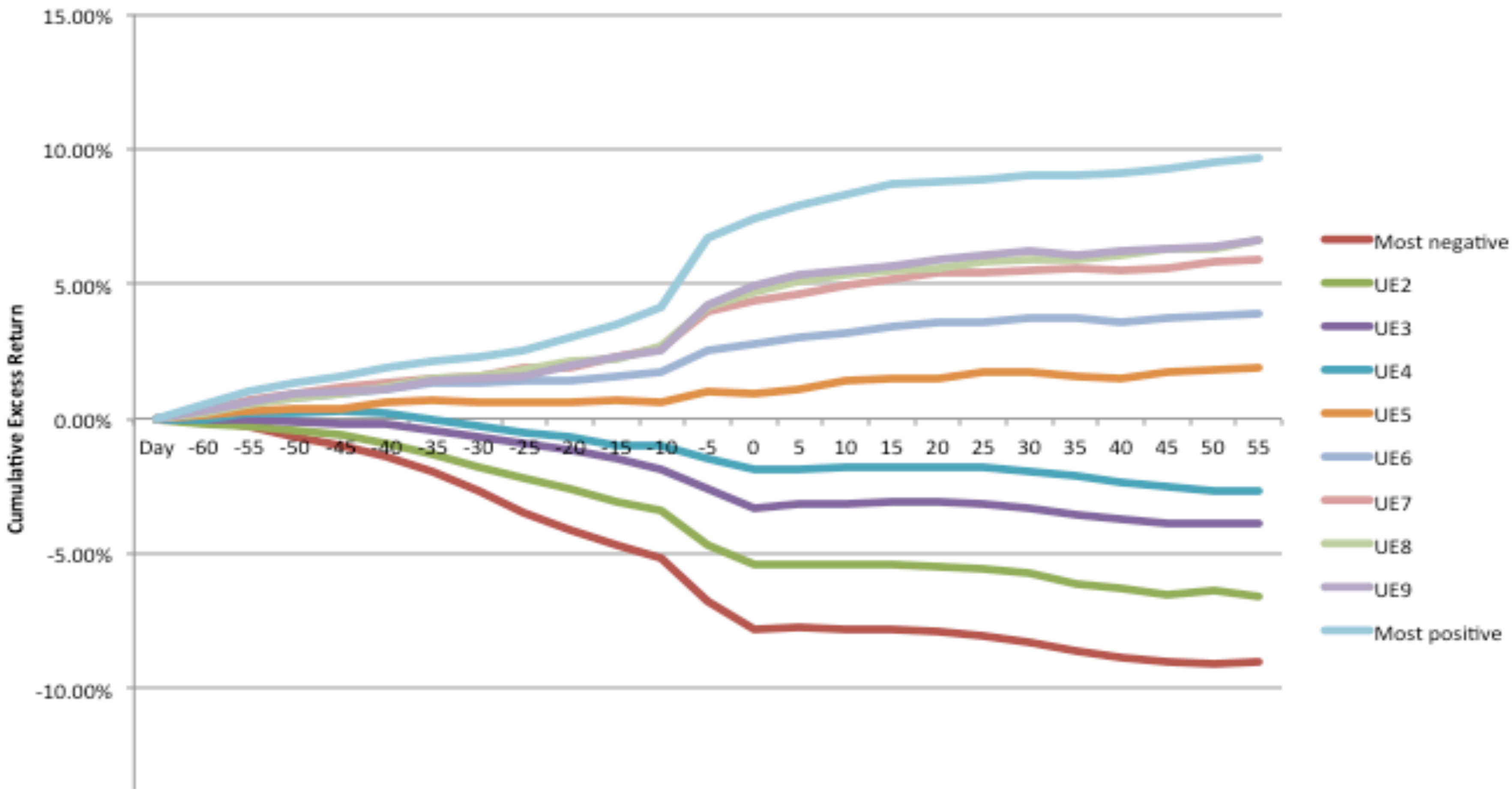
Trading on Public Information

- There is substantial information that comes out about stocks. Some of the information comes from the firm - earnings and dividend announcements, acquisitions and other news - and some comes from competitors.
- Prices generally react to this information.

Example: Earnings Announcements

- Every quarter (in the U.S) and less frequently elsewhere, firms report their earnings for the most recent period.
  - When firms make earnings announcements, they convey information to financial markets about their current and future prospects. The magnitude of the information, and the size of the market reaction, should depend upon how much the earnings report exceeds or falls short of investor expectations.
  - In an efficient market, there should be an instantaneous reaction to the earnings report, if it contains surprising information, and prices should increase following positive surprises and down following negative surprises.

Market Reaction to Earnings Reports

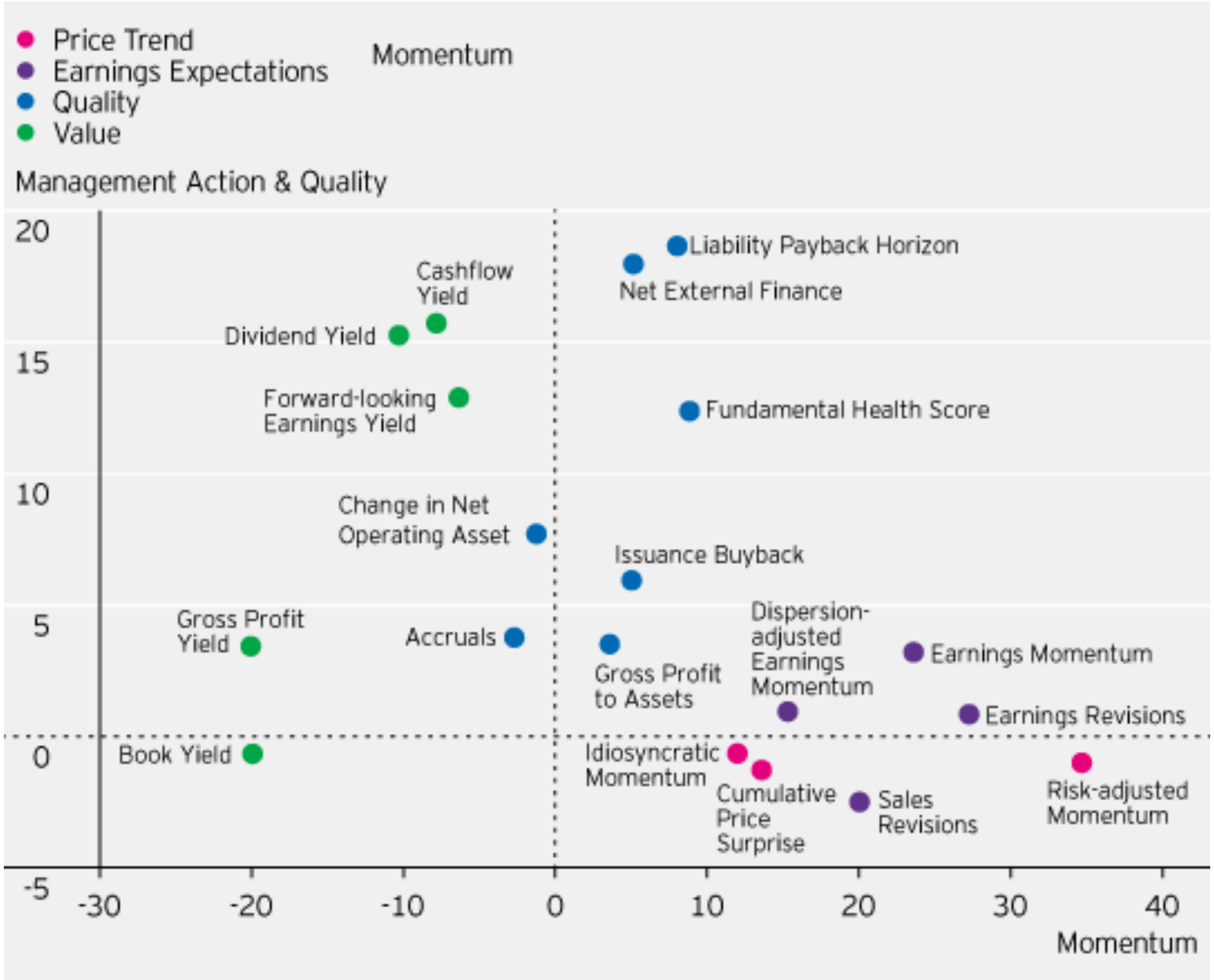


BOSTON UNIVERSITY

Can you make money of earnings announcements?

- One strategy is to buy stocks that report large positive earnings surprises, hoping to benefit from the drift. The evidence indicates that across all stocks, the potential for excess returns from buying after earnings announcements is very small.
  - You can concentrate only on earnings announcements made by smaller, less liquid companies where the drift is more pronounced. In addition, you can try to direct your money towards companies with higher quality earnings surprises by avoiding firms with large accruals.
  - Your biggest payoff is in investing in companies before large positive earnings surprises. You may be able to use a combination of quantitative techniques (time series models that forecast next quarter's earnings based upon historical earnings) and trading volume (insiders do create blips in the volume) to try to detect these firms. Even if you are right only 55% of the time, you should be able to post high excess returns.

Which Information Signals Create Alpha?



BOSTON UNIVERSITY

In closing..

- If you can trade on private information, you can get ahead of market movements. Unfortunately, trading on private information is usually illegal.
- You can trade on information after it has been made public, taking advantage of systematic errors that markets make in how they react. If markets overreact, you will buy after bad news and sell after good news. If markets under react, you will buy after good news and sell after bad news.
- The evidence suggests that there is some price drift after news announcements, suggesting a slow learning market. The drift is small and you have to react quickly and have low transactions cost to earn excess returns.

What Is General Goal of Investment Managers?

Try to Maximize the “Sharpe Ratio” of their investment portfolio

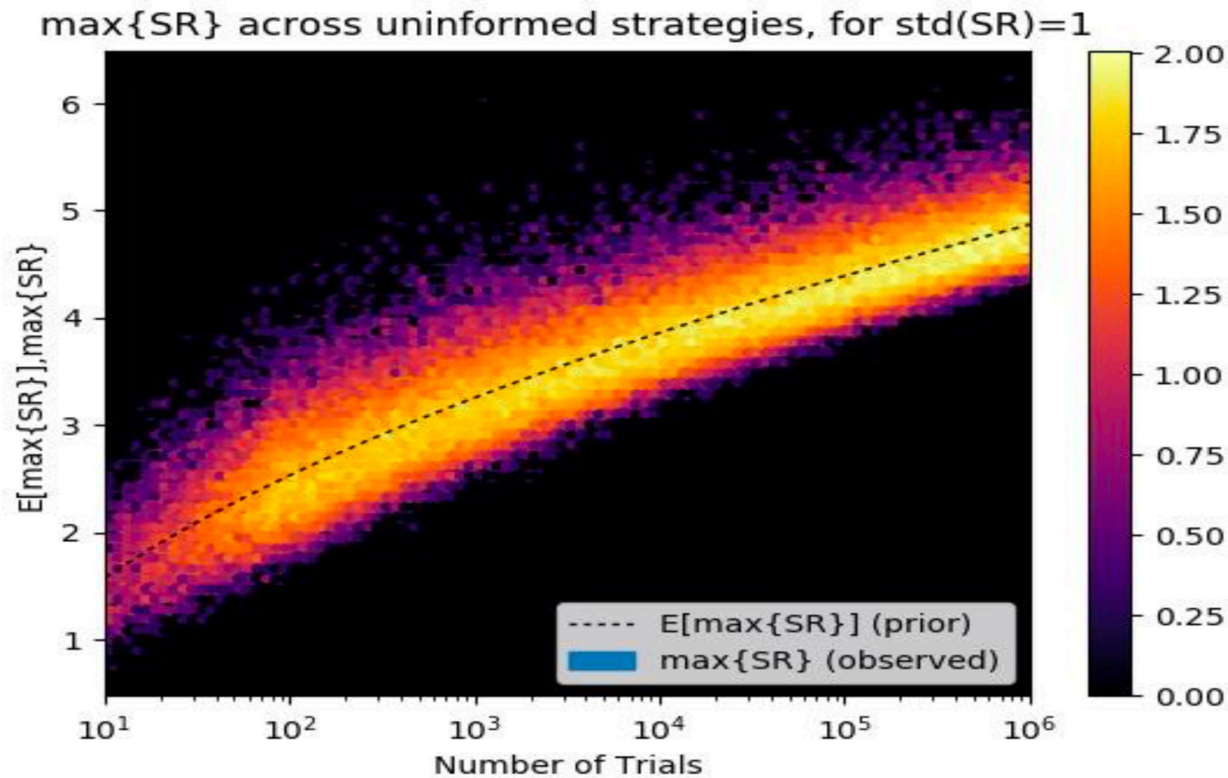
$$\text{Sharpe Ratio Formula} = \frac{\text{Return of Portfolio} - \text{Risk Free Rate}}{\text{Standard Deviation of Portfolio's Excess Return}}$$

A Major Part of Finding and Assessing Trading Strategies: Backtesting

- "Backtesting" a quest to find a profitable investing strategy (that is valid and not a fluke). Two recent research warn against the **perils of data snooping**.
- Fabozzi and Prado (2018): State that only strategies that "seem to work" make it to the public, while thousands (at least) have been tested.
  - Picking the pleasing outlier (the only strategy that seemed to work) is likely to generate disappointment when switching to real trading.
  - Worst types are **false positives** whereby strategies are found (often by cherrypicking) to outperform in one very particular setting, but will likely fail in live implementation.
- Backtesting is more complicated than it seems and it is easy to make small mistakes that lead to *apparently* good portfolio policies (that ultimately fail).

BOSTON UNIVERSITY

The Most Important Plot In Finance



The y-axis displays the distribution of the maximum Sharpe ratios (max{SR}) for a given number of trials (x-axis). A lighter color indicates a higher probability of obtaining that result, and the dash-line indicates the expected value. For example, after only 1,000 independent backtests, the expected maximum Sharpe ratio (E[max{SR}]) is 3.26, even if the true Sharpe ratio of the strategy is zero!

The reason is *Backtest Overfitting*: When selection bias (picking the best result) takes place under multiple testing (running many alternative configurations), that backtest is likely to be a false discovery. **Most quantitative firms invest in false discoveries**.

From Marcos Lopez De Prado: https://ssrn.com/abstract=3447398

BOSTON UNIVERSITY

Problems with “Data Mining” to Find Investment Strategies

- With Backtesting, you can find strategies that work in the past (High Sharpe Ratio).
- But High Past Sharpe Ratio is typically unrelated to Future Sharpe Ratio!

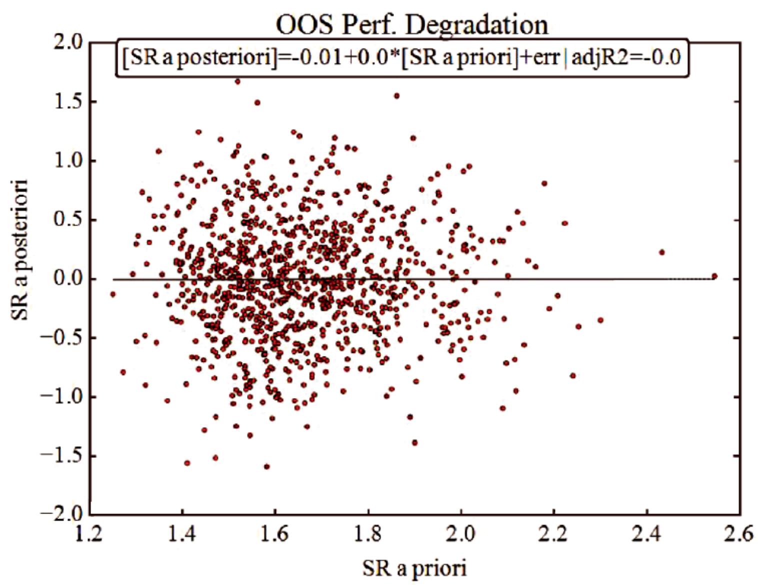


Figure 5. Performance degradation after introducing strategy selection in absence of compensation effects.

Problems with “Data Mining” to Find Investment Strategies

- If investment returns have autoregressive properties, then it even worse.
- Searching for high past Sharpe ratios will lead to negative performance of a strategy in the future!

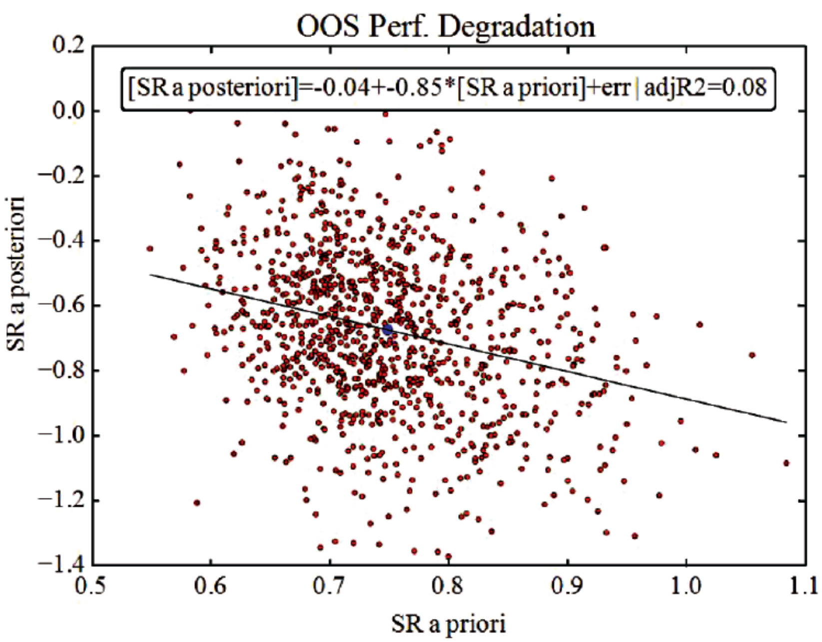


Figure 7. Performance degradation as a result of strategy selection under compensation effects (first-order serial correlation).

a global constraint. If we rerun the previous Monte Carlo experiment, this time for the autoregressive process with $\mu = 0, \sigma = 1, \varphi = 0.995$, and plot the pairs of performance IS vs. OOS, we obtain Figure 7.

The p-values associated with the intercept and the IS performance (SR a priori) are respectively 0.4513 and 0, confirming that the negative linear relation between IS and OOS Sharpe ratios is again statistically significant. Such serial correlation is a well-known statistical feature, present in the performance of most hedge fund strategies. Proposition 5 is proved in the appendix.

Proposition 5. *Given two alternative configurations (A and B) of the same model, where* $\sigma_{IS}^{A} = \sigma_{OOS}^{A} = \sigma_{IS}^{B} = \sigma_{OOS}^{B}$ *and the performance series follows the same first-order autoregressive stationary process,*

$$SR_{IS}^{A} > SR_{IS}^{B} \Leftrightarrow SR_{OOS}^{A} < SR_{OOS}^{B}. \tag{13}$$

Proposition 5 reaches the same conclusion as Proposition 3 (a compensation effect) without requiring a global constraint.

BOSTON UNIVERSITY

Big Topic: Explainable/Interpretable Machine Learning

- Most straightforward path to **explainability** is developing a simple, explainable model.
  - OLS regression models are considered relatively interpretable, especially when regression coefficients have clear economic meaning.
  - Other models: (small) decision trees or decision rules.
- But, other models such as neural networks have very high predictive power! Can we make these "black boxes" interpretable?
  - New approaches: Attempt to 'reverse engineer' the workings of a complex model.
  - Not necessarily try to explain the model itself, but to highlight its salient features.
- There are six ways to do so:

Explainable/Interpretable Machine Learning – 6 ways

1. One way to reverse engineering a complex ML model is to construct a simpler model — such as a regression model or small decision tree — that **approximates the workings** of the complex one.
   - This is called ‘surrogate model’. We would refer to this as a ‘global’ model, as it tries to explain the workings of a complex model **for all input data**.
2. Another global approach: “Feature Importances”
   - Explainability technique developed for Random Forest ML models.
   - Provides the relative importance of features for all input data (global level).
   - Estimates how much the model prediction variance changes due to the exclusion of individual features.
   - But does not capture feature interactions well.

Explainable/Interpretable Machine Learning – Prediction Variance

OLS Example of Prediction Variance – OLS

Main Model:

Y = a + b1*X1 + b2*X2 + e

<u>Calculate Adjusted R2: R2(Main)</u>

Nested Sub-Models:

Y = a + b1*X1 + e

<u>Calculate Adjusted R2: R2(Drop X2)</u>

Y = a + b2*X2 + e

<u>Calculate Adjusted R2: R2(Drop X1)</u>

Explainable/Interpretable Machine Learning – 6 ways

3. Build one or several local surrogate models. Local surrogate models approximate the complex model's predictions on selected sub-sections of the data.
   - One example of a local approach is the Local Interpretable Model-Agnostic Explanation (LIME) method.

4. Build instance-based explanations. This approach does not build a 'model' (global or local). Rather, it provides explanations on a prediction-by-prediction basis.
   - Answers questions like 'what were the driving factors in the case of individual A'?
   - Examples: Shapley values & Individual Conditional Expectations.
   - Advantage of approach is that it can capture feature-interactions.

Explainable/Interpretable Machine Learning – 6 ways

5. Partial Dependence Plots (PDPs) show the impact of one or two variables have on the predictive outcome.

   - These are very useful tools to display the non-linearities and other complexities in the underlying ML model.
   - 'Partial' in the sense that can only display one or two features at a time. Do not consider interactions in the display on features impact on predictive outcomes.

6. Combine various elements of the above 5 approaches