Introduction to Algorithmic Trading Strategies
Lecture 9

Quantitative Equity Portfolio Management

Haksun Li
haksun.li@numericalmethod.com
www.numericalmethod.com

Outline

- Alpha
- Factor Models

References

- Ludwig, B. C. and Daehwan, K. Quantitative Equity Portfolio Management: An Active Approach to Portfolio Construction and Management. 2006
- Grinold, R. and Kahn, R. Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Controlling Risk. 1999.

Alpha

Alpha

- α: out-performance
  - the measurement of a portfolio's risk adjusted returns over a reference instrument.
- Benchmark α:
  - $r_p = \alpha + \beta r_B + \varepsilon$
  - $\alpha + \varepsilon$: residual return, the return that is independent of the benchmark
  - βr_B: expected return
- CAPM α:
  - $r_p = \alpha + \beta r_M + \varepsilon$
  - According to CAPM, $\alpha = 0$.
- Multifactor α:
  - $r_p = \alpha + \beta_1 f_1 + \cdots + \beta_K f_K + \varepsilon$
  - Each f_i is a risk factor.
  - $\sum_{i=1}^{K} \beta_i f_i$: expected return

Information Ratio

- $IR = \frac{\alpha^B}{\sigma}$
  - α^B: ex-ante, expected excess return
  - σ: standard deviation of the residuals

The Seven Tenets of QEPM

- Markets are mostly efficient.
- Pure arbitrage opportunities do not exist.
- Quantitative analysis creates statistical arbitrage opportunities.
- Quantitative analysis combines all the available information in an efficient way.
- Quantitative models should be based on sound economic theories.
- Quantitative models should reflect persistent and stable patterns.
- Deviations of a portfolio from the benchmark are justified only if the uncertainty is small enough.

Post Earning Announcement Drift

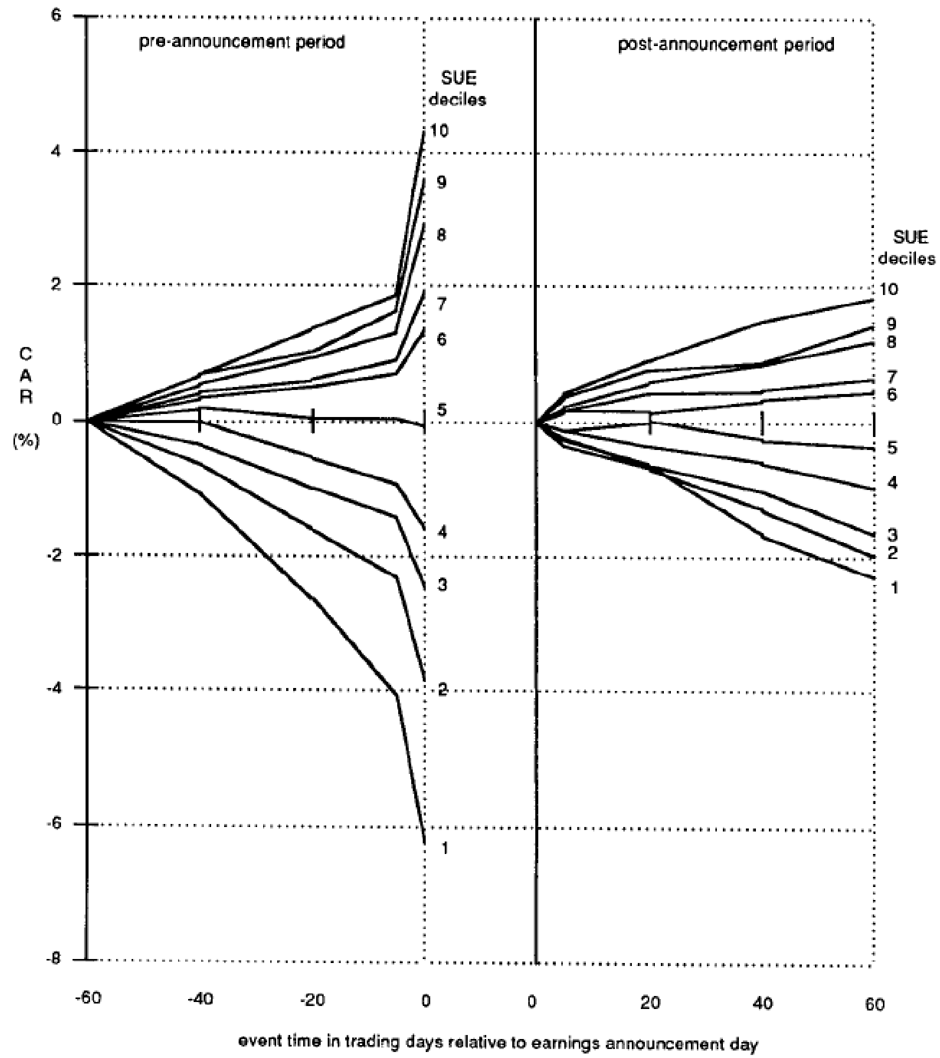
- On a good (bad) earning surprise announcement, the market will continue to react to the news.
- The market takes some time to digest the new information, often weeks or even months.
- We can therefore buy on a good news and take profit when the market finishes consuming the information.
- Our smart filter picks high confident stocks from hundreds of announcements each quarter.

Example: Insys Therapeutics, Inc. (INSY)

Response to Earning Surprise

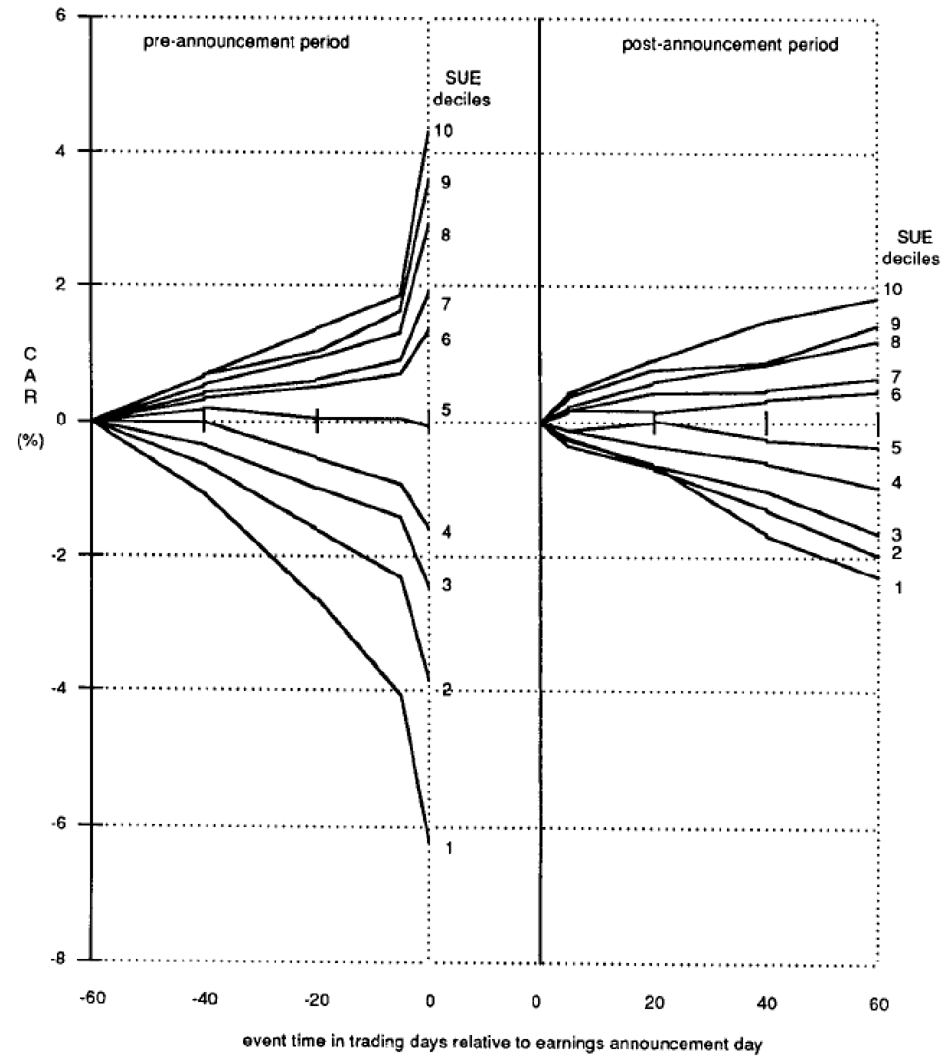
Post-Earning Announcement Drift (PEAD)

- Anomaly discovered by Ball and Brown (1968)
  - The tendency for a stock's cumulative abnormal returns to drift for several weeks (even several months) following positive earnings announcement
  - Studied and confirmed by countless academics in many international markets

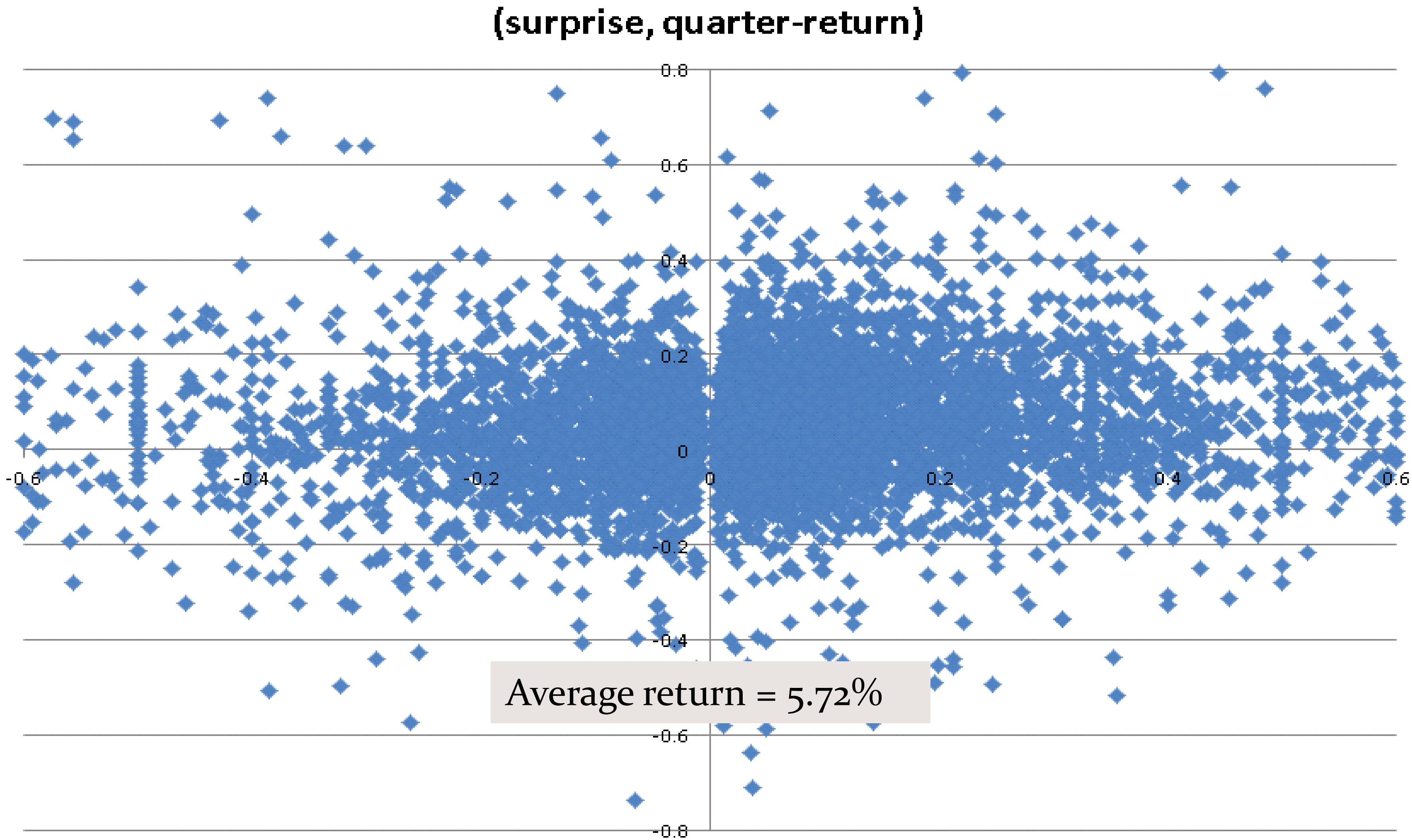


PEAD Trading

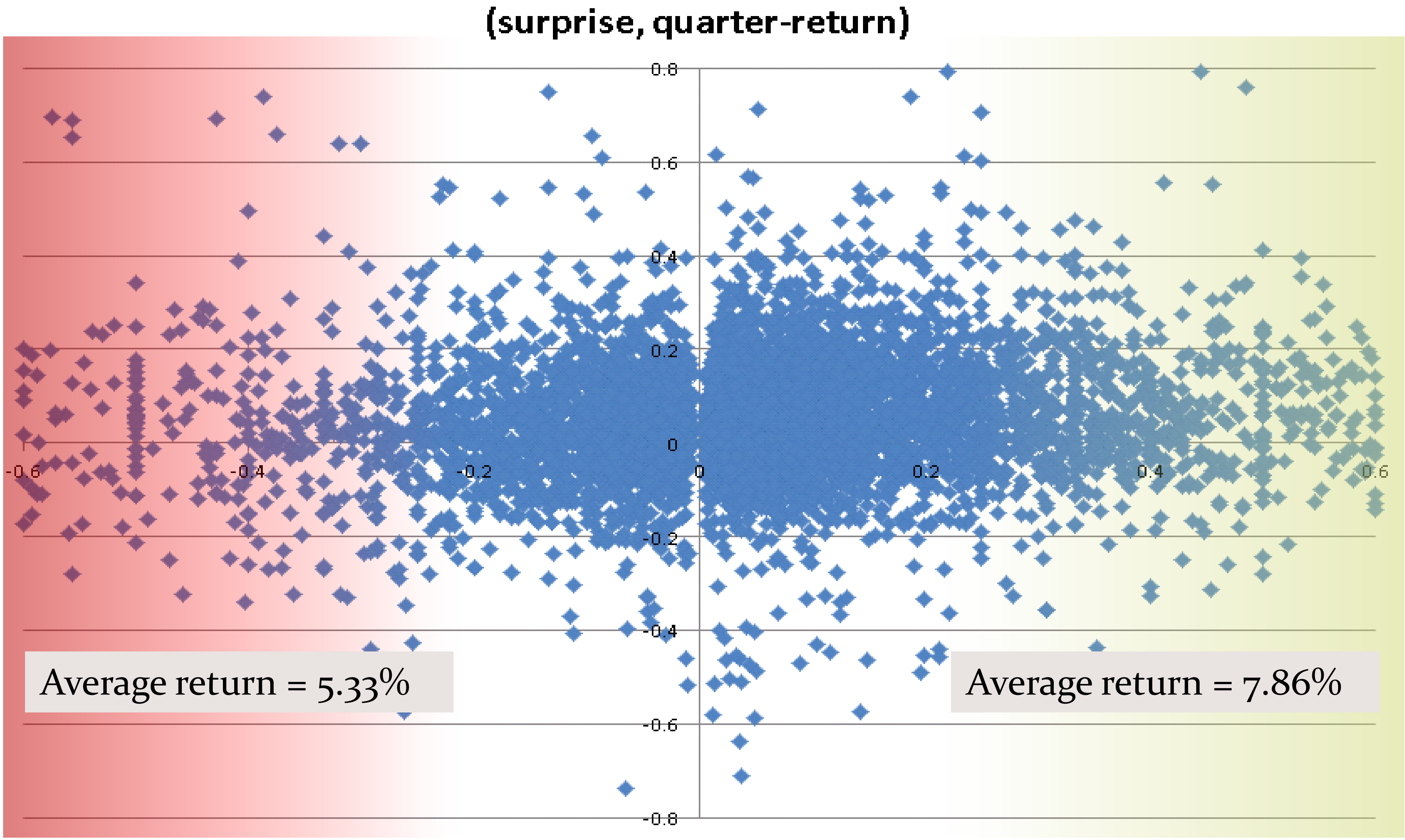
- Fundamental reasons:
  - Under-reaction to earnings announcements
  - Distraction by other announcements
- Trading idea:
  - Long good news
  - Short bad news



Surprise vs. Return



Surprise vs. Return

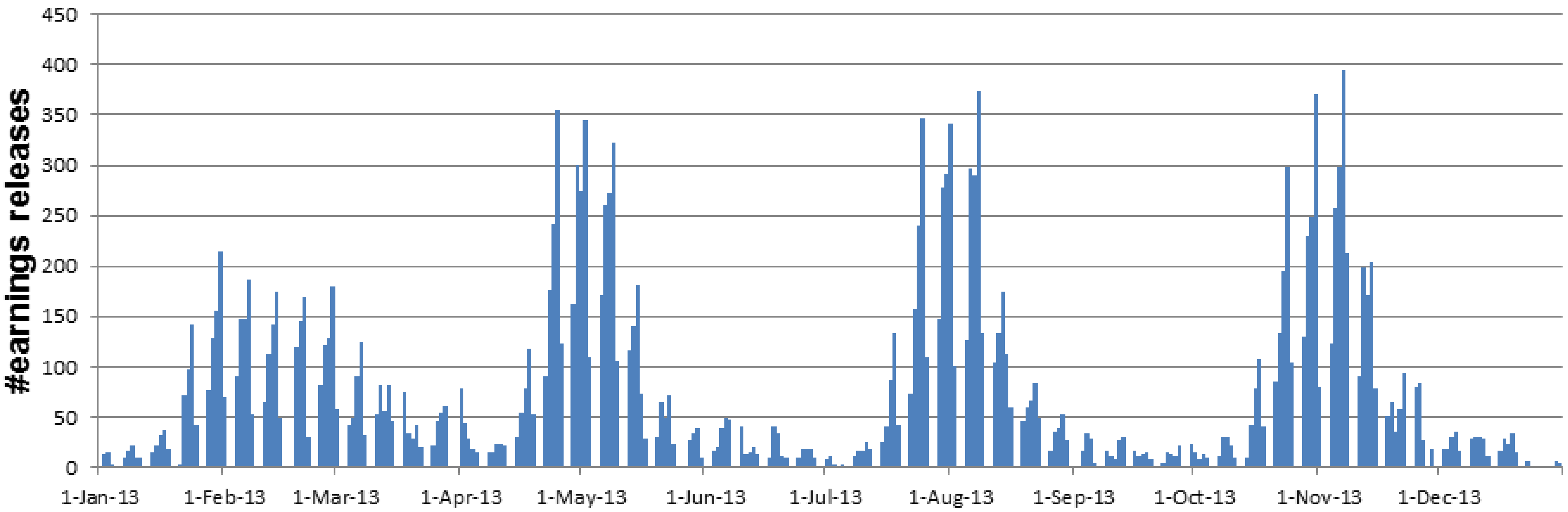


Capital Allocation

- U.S. stocks
  - Quarterly reporting system dictates frequent announcements, hence more trading opportunities
- Diversification
  - No single stock weights more than 10%
- Long-only
- Stop-loss
- No penny stocks

Look for Profitable Stocks

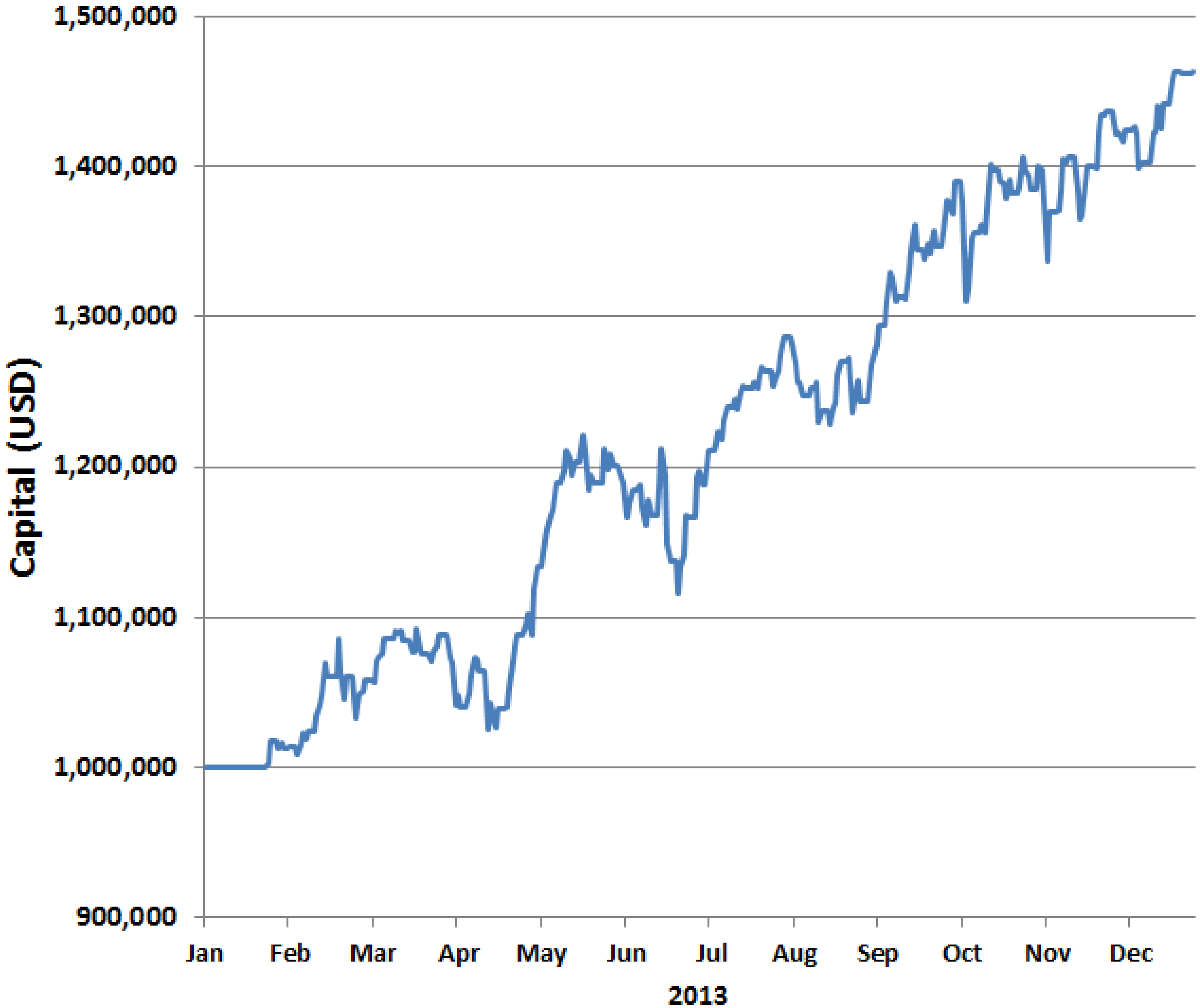
- Over 300 announcements on peak days
- Our Smart Stock Filter picks stocks with higher winning probability



A case for 2013

US Stocks
Profit: 46.36%
#Trade: 103
#Win: 62
#Loss: 41
Max. gain: 91.6%
Max. loss: -10.00%
Gain/trade: 7.80%

Sharpe ratio: 2.95



Historical Returns

- 2008: 11.3%
- 2009: 138.1%
- 2010: 26.2%
- 2011: 8.9%
- 2012: 11.6%
- 2013: 68%
- 2014: 16.5%

Fundamental Law of Active Management

- $r_{it} = \alpha_i + \beta_{i1} f_{1t} + \cdots + \beta_{iK} f_{Kt} + \varepsilon_{it}$

Fundamental Law of Active Management

- With $\mathrm{E}(r_{it})$ and $\mathrm{Cov}(r_{it}, r_{jt})$, we can compute the optimal min-variance portfolio.
  - IR^2 approximately equals the goodness of fit, R^2.
  - BR: number of explanatory factors.
  - IC is the average covariance of predictions vs. factors.
- $IR^2 = \frac{\alpha^B}{\sigma} = \frac{[w' \,\mathrm{E}(r_{T+1})]^2}{w' \,\mathrm{V}(r_{T+1}) w} = R^2 = IC^2 \times \sqrt{BR}$
  - IR: information ratio
  - IC: information coefficient
  - BR: breadth
- Finding more significant/accurate factors increases IC.
- Finding more uncorrelated factors increases BR.

Data Mining

- You can always find a "good" model by trying hard enough – testing enough factors.
- $r_t = \alpha + \beta_1 f_{1t} + \cdots + \beta_{99} f_{99t} + \varepsilon_{it}$
- When there are 100 months of return, the equation will fit perfectly, $R^2 = 1$.
- Forward selection will not work as it will guarantee to pick the most significant variable out of the 100 random factors.
- Community is doing a collective data mining.
  - E.g., company size may not be a factor.

Parameter Estimation

- Sensitivity and stability.
  - Divide the data into groups along the timeline. Estimations should be more or less the same for all sub-groups.
- Uncertainty and confidence.

Factor Models

Expected Return

- Expected Return = Factor Premium x Factor Exposure
- Factor Premium, f_i: how much an investor is willing to pay for each factor. Payoff.
- Factor Exposure, β_i: how sensitive is a stock return to the factor. Exposure to risk.

Factors

- Stock specific/fundamental factors: PE, PB, PS, D/E, size, momentum, volume, earning surprise, analyst rating changes
- Market/economic factors: GDP, inflation, unemployment, interest rate, survey-based indices, …
- Factor exposure: exposure to a risk
- Factor premium: the premium/price/fair return/dollar value the market places of 1 unit of risk (factor exposure)
- average stock return = factor exposure * factor premium

Fundamental Model

- $r_i = \alpha_i + \beta_{i1} f_1 + \cdots + \beta_{iK} f_K + \varepsilon_i$
  - r_i can be replaced by $r_i{}^* = r_i - r_f$ because the r_f portion is not awarded by taking risks.
  - r_i: monthly returns, for example.
- Factors are fundamental factors such as P/E, size.
- The β_{ij} are directly observable (from accounting reports).
  - factor dependent; stock dependent; time independent
- The premiums are estimated by cross-sectional/panel regression.
  - factor dependent; stock independent; time independent

Risk

- total risk = non-diversifiable risk + diversifiable risk
- $\mathrm{V}(r_i) = \mathrm{V}\left(\beta_i{}' f\right) + \mathrm{V}(\varepsilon_i)$
- $= \beta_i{}' \, \mathrm{V}(f) \, \beta_i + \mathrm{V}(\varepsilon_i)$
- Non-diversifiable risk comes from the randomness of the premiums.
- Diversifiable risk comes from the unexplained risk of the model.

Estimation of Factor Premiums

- Returns of N stocks over T periods: $\{(r_{11}, \dots, r_{N1})\}$, ..., $\{(r_{1T}, \dots, r_{NT})\}$.
- Factor exposures known: $\{(\beta_{11}, \dots, \beta_{N1})\}$, ..., $\{(\beta_{1T}, \dots, \beta_{NT})\}$.
- Model: $r_{it} = \boldsymbol{\beta_{it}}'\boldsymbol{f} + \varepsilon_{it}$
- Panel:
  - $r_{11} = \beta_{11,1}f_1 + \cdots + \beta_{11,K}f_K + \varepsilon_{11}, \dots$
  - $r_{1T} = \beta_{1T,1}f_1 + \cdots + \beta_{1T,K}f_K + \varepsilon_{1T}, \dots$
  - $r_{N1} = \beta_{N1,1}f_1 + \cdots + \beta_{N1,K}f_K + \varepsilon_{N1}, \dots$
  - $r_{NT} = \beta_{NT,1}f_1 + \cdots + \beta_{NT,K}f_K + \varepsilon_{NT}, \dots$

OLS

- $y_i = \beta_0 + x_{1i}\beta_1 + x_{2i}\beta_2 + \cdots + +x_{ki}\beta_k + \varepsilon_i,\ i = 1,2,3,\ldots,N$
- Objective:
  - Minimize the sum of squared residuals.
- Assumptions:
  - No (perfect) collinearity: $|\mathrm{Cov}(x_i, x_j)| \neq 1$
    - $X'X$ cannot be inverted.
  - Regressors are exogenous: $\mathrm{E}(\varepsilon_i \mid x_{1i}, \ldots, x_{ki}) = 0$.
  - Uncorrelated errors: $\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0$.
  - Homoskedasticity: $\mathrm{V}ar(\varepsilon_i) = \mathrm{V}ar(y_i \mid x_{1i}, \ldots, x_{ki}) = \sigma^2$.
- Properties:
  - Consistent: $\hat{\beta} \to \beta$.
  - Unbiased: $\mathrm{E}(\hat{\beta}) = \beta$.
  - Minimum variance
  - Errors are normally distributed. ε_i is $N(0, \sigma^2)$.
  - $\hat{\beta}$ are MLE.

OLS Estimation

- RSS = $S(b) = \sum_{i=1}^{n}\left(y_i - x_i'b\right)^2 = (y - Xb)'(y - Xb)$
- $\hat{\beta} = \operatorname{argmin} S(b) = \left(\sum_{i=1}^{n}(x_i - \overline{x_i})(x_i - \overline{x_i})'\right)^{-1}\left(\sum_{i=1}^{n}(x_i - \overline{x_i})(y - \bar{y})\right) = (X'X)^{-1}X'y$
  - E.g., apply first order condition on $S(b)$
  - $\hat{\beta}_i = \frac{\text{Cov}(x_i, y)}{\text{Var}(x_i)}$
- $\hat{y} = X\hat{\beta}$
- $\hat{\varepsilon} = y - \hat{y} = y - X\hat{\beta}$
- $s^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n-p}$, OLS estimator for σ^2, unbiased
- $\hat{\sigma}^2 = \frac{\hat{\varepsilon}'\hat{\varepsilon}}{n}$, MLE estimator for σ^2, biased but minimizes the mean squared error of the estimator.
- $R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2}$

OLS Finite Sample Properties

- $\mathrm{E}\left[\hat{\beta} \mid X\right] = \beta$
- $\mathrm{E}\left[s^2 \mid X\right] = \sigma^2$
- $\mathrm{Var}\left[\hat{\beta} \mid X\right] = \sigma^2 (X'X)^{-1}$
  - $\hat{\beta} \sim N\left(\beta, \sigma^2 (X'X)^{-1}\right)$
- $\widehat{\mathrm{s.e.}}\left(\hat{\beta}_j\right) = \sqrt{s^2 (X'X)^{-1}{}_{jj}}$
- $\mathrm{Cov}\left[\hat{\beta}, \hat{\varepsilon} \mid X\right] = 0$

OLS on Panel Data

- $r_{it} = \boldsymbol{\beta_{it}}'\boldsymbol{f} + \varepsilon_{it}$
- $r_{it} - \bar{r} = \left(\boldsymbol{\beta_{it}} - \bar{\beta}\right)'\boldsymbol{f} + \varepsilon_{it}$
  - $\bar{r} = \frac{1}{NT}\sum_{i=1}^{T}\sum_{i=1}^{N} r_{it}$
  - $\bar{\beta} = \frac{1}{NT}\sum_{i=1}^{T}\sum_{i=1}^{N} \beta_{it}$
- $\tilde{r}_{it} = \boldsymbol{\widetilde{\beta}_{it}}'\boldsymbol{f} + \varepsilon_{it}$
  - $\tilde{r}_{it} = r_{it} - \bar{r}$
  - $\tilde{\beta}_{it} = \mathrm{\beta_{it}} - \bar{\beta}$
- $\hat{f} = \left[\sum_{i=1}^{T}\sum_{i=1}^{N}\left(\tilde{\beta}_{it}\right)\left(\tilde{\beta}_{it}\right)'\right]^{-1}\sum_{i=1}^{T}\sum_{i=1}^{N}\left(\tilde{\beta}_{it}\right)\left(\tilde{r}_{it}\right)$
- $\mathrm{Var}\left(\hat{f}\right) = \hat{\sigma}^2\left[\sum_{i=1}^{T}\sum_{i=1}^{N}\left(\tilde{\beta}_{it}\right)\left(\tilde{\beta}_{it}\right)'\right]^{-1}$
- $\hat{\sigma}^2 = \frac{1}{NT}\sum_{i=1}^{T}\sum_{i=1}^{N}\left(\tilde{r}_{it} - \left(\tilde{\beta}_{it}\right)'\hat{f}\right)^2$

Minimum Absolute Deviation (MAD)

- OLS is sensitive to outliners.
- MAD: minimize the sum of absolute value of residuals.
  - Hence outliners, without being squared, have much less effects.

Generalized Least Square (GLS)

- Heteroskedasticity: ε_{it} are most likely difference across stocks. Different stocks have different variances.
- Model: $Y = X\beta + \varepsilon$, $\mathrm{E}[\varepsilon | X] = 0$, $\mathrm{Var}[\varepsilon | X] = \Omega$
- Fitting: $\hat{\beta} = \mathrm{argmin}(Y - Xb)'\Omega'(Y - Xb)$
- Solution: $\hat{\beta} = \left(X'\Omega^{-1}X\right)'X'\Omega^{-1}Y$
- The GLS estimator is unbiased, consistent, efficient, and asymptotically normal.
  - $\sqrt{n}(\hat{\beta} - \beta) \rightarrow \mathcal{N}\left(0, \left(X'\Omega^{-1}X\right)^{-1}\right)$

GLS on Panel Data

- $\hat{f} = \frac{\sum_{t=1}^{T}\sum_{i=1}^{N}\left(\frac{\beta_{it}}{\hat{\sigma}_i}-\tilde{\beta}\right)\left(\frac{r_{it}}{\hat{\sigma}_i}-\tilde{r}\right)}{\sum_{t=1}^{T}\sum_{i=1}^{N}\left(\frac{r_{it}}{\hat{\sigma}_i}-\tilde{r}\right)^2}$
- $\tilde{\beta} = \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}\frac{\beta_{it}}{\hat{\sigma}_i}$
- $\tilde{r} = \frac{1}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}\frac{r_{it}}{\hat{\sigma}_i}$
- $\hat{\sigma}_i^2 = \frac{1}{T}\sum_{t=1}^{T}\left(r_{it} - {\beta_{it}}'\hat{f}\right)^2$

Robustness Check

- Stability over sub-periods.
- Confidence within periods.

Economic Model

- $r_{it} = \alpha_i + \beta_{i1} f_{1t} + \cdots + \beta_{iK} f_{Kt} + \varepsilon_i$
- Factors are economic factors such as GDP growth, inflation.
- The exposures β_{ij} are not observable and are determined by time series regression from the factor premiums f_{jt}.
  - factor dependent; stock dependent; time independent
- The premiums are market statistics.
  - factor dependent; stock independent; **time dependent**

Stock Return

- $r_{it} = \alpha_i + \beta_{i1} f_{1t} + \cdots + \beta_{iK} f_{Kt} + \varepsilon_{it}$
- Note, we use the time t available information to compute the factor premiums, e.g., latest inflation rate.

Economic/Market Premiums

- Inflation, unemployment, risk free rate, …
- Simply copy and paste from newspapers.
- A premium is a linear function of the true, unexpected, and unobservable part of the factor, e.g., the rewarding portion of inflation.
  - This is not to say that the market is rewarding 5% to a 5% inflation rate, but a linear transformation of the 5%.

Fundamental/Technical/Analyst Premiums

- P/E, P/B
- Momentum
- Rating change

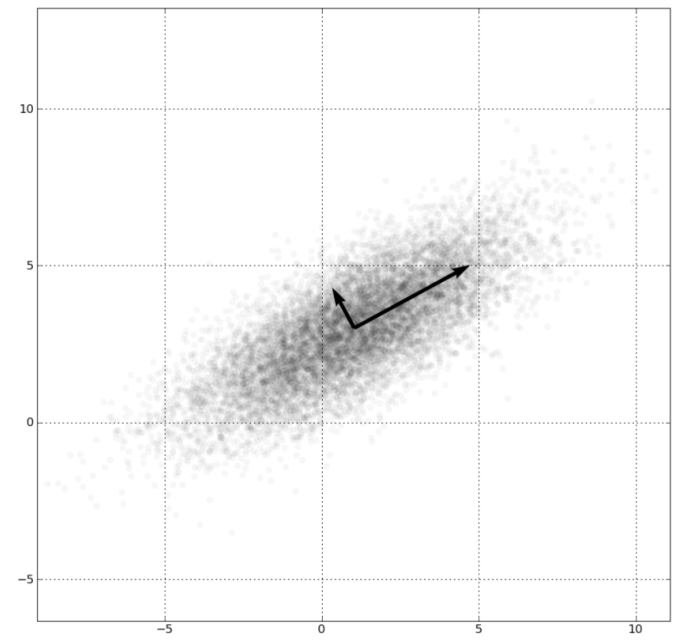
Zero-Investment Portfolio

- For each time t,
- For each factor k, set the upper and lower cutoff points, $\overline{x_k}$ and $\underline{x_k}$.
- Divide the stocks into three groups.
  - High group: $x_{ikt} > \overline{x_k}$
  - Low group: $x_{ikt} < \overline{x_k}$
  - Others
- Factor premium is the expected return to the zero-investment position that put \$1 into the high group and short \$1 in the low group.
- $f_{kt} = \mathrm{E}(r_t \,|x_{kt} > \overline{x_k}) - \mathrm{E}(r_t|x_{kt} < \overline{x_k})$
  - The expectation is taken across stocks.

Statistical Factors

Principal Component Analysis

- Intuition: fit the data to a high N-dimensional space. Find and save the important dimensions (and hence filter out the not-so-important ones).



Principal Component Analysis (Math)

- N stocks' returns over T period.
  - $r_t = (r_{1t}, \ldots, r_{Nt})'$
  - $R = \{r_1, \ldots, r_T\}$
- Find the (sample) covariance matrix for R.
  - $\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}\left(r_t r_t' - \bar{r}\overline{r'}\right)$
  - $\bar{r} = \frac{1}{T}\sum_{t=1}^{T} r_t$
- Diagonalization: $Q'\hat{\Sigma}Q = D$.
  - D are the (all) positive eigenvalues in descending orders.
- Pick the K eigenvectors q_1, …, q_K that correspond to the largest K eigenvalues.
- Statistical factors are: $f_{i,t} = q_i' r_t$.

Forecasting Premiums

- $\mathrm{E}(f_{T+1}) = \hat{\gamma}_0 + \hat{\gamma}_1 f_T + \cdots + +\hat{\gamma}_L f_{T-L+1}$
- Many other ways…

Factor Exposure (Standard Approach)

- Once the premiums are known, we can apply OLS to estimate the factor exposures, β_i.
  - $r_{it} = \beta_i' f_t + \varepsilon_{it}$

Factor Exposure (Mergers)

- $\hat{\beta}_{AB} = \frac{s_A}{s_A + s_B}\hat{\beta}_A + \frac{s_B}{s_A + s_B}\hat{\beta}_B$
- s_A: pre-merger market capitalization of firm A
- s_B: pre-merger market capitalization of firm B

Factor Exposure (Characteristic Matching)

- For a newly IPO company, there is not enough data to compute factor exposure directly.
- We use the factor exposures of M similar firms.
- To identify the M similar firms,
  - We choose L company characteristics;
  - Compute the z-score of those L characteristics for a group of firms $\{z_i = (z_{i1}, \dots, z_{iL})\}$as well as those of the new company, z $= (z_1, \dots, z_L)$.
  - Set a threshold, ε.
  - The similar firms are those with smaller distances. That is,
    - $\|z - z_i\| < \varepsilon$.
- $\hat{\beta} = \frac{1}{M}(\hat{\beta}_1 + \cdots + \hat{\beta}_M)$

Model Comparisons

Fundamental Model

Factor exposure directly observed.

Factor premium from panel regression.

Expected return = exposure * premium.

Factor exposure from time series regression.

Factor premiums either directly observed, zero-cost portfolio, or PCA.

Expected return = exposure * premium.

Economic Model

Stock Ranking

Z-Score

- Standardization of factors: $z_i = \frac{\beta_i - \mu}{\sigma}$

Quintiles Method

- To test whether a factor (or a strategy) is significant in generating alpha…
- Rank/sort the stocks in a universe by the factor.
- Divide them into 5 groups (20% each).
- Portfolio formed each quarter over the test period. Each portfolio is hold for 12 months.
  - Number of portfolios in each quintile for the test period = test period (in years)*4*(size of universe/5).
- Compute average returns for each quintile.
- Factor significant if
  - top first quintile significantly outperformed the universe
  - the bottom fifth quintile significantly underperformed
  - the outperformance/underperformance was consistent over time