Applied Regression Analysis
Portfolio Problem and The Market Model

Carlos M. Carvalho
The University of Chicago Booth School of Business
http://faculty.chicagobooth.edu/carlos.carvalho/teaching/
Building Portfolios

Let’s assume we are considering 3 investment opportunities

1. IBM stocks
2. ALCOA stocks
3. Treasury Bonds (T-bill)

How should we start thinking about this problem?
Building Portfolios

Let’s first learn about the characteristics of each option by assuming the following models:

- $IBM \sim N(\mu_I, \sigma^2_I)$
- $ALCOA \sim N(\mu_A, \sigma^2_A)$

and

- The return on the T-bill is 3%

After observing some return data we can came up with estimates for the means and variances describing the behavior of these stocks.
Building Portfolios

The data “tells” us that...

<table>
<thead>
<tr>
<th>IBM</th>
<th>ALCOA</th>
<th>T-bill</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\mu}_I = 12.5$</td>
<td>$\hat{\mu}_A = 14.9$</td>
<td>$\mu_{T\text{bill}} = 3$</td>
</tr>
<tr>
<td>$\hat{\sigma}_I = 10.5$</td>
<td>$\hat{\sigma}_A = 14.0$</td>
<td>$\sigma_{T\text{bill}} = 0$</td>
</tr>
</tbody>
</table>

$\text{corr}(\text{IBM, ALCOA}) = 0.33$

Now what?? What should I do?
Building Portfolios

▶ How about combining these options? Is that a good idea? Is it good to have all your eggs in the same basket? Why?

▶ What if I place half of my money in ALCOA and the other half on T-bills...

▶ Remember that:

\[
\begin{align*}
E(aX + bY) &= aE(X) + bE(Y) \\
Var(aX + bY) &= a^2 Var(X) + b^2 Var(Y) + 2ab \times Cov(X, Y)
\end{align*}
\]
Building Portfolios

- So, by using what we know about the means and variances we get to:

\[
\hat{\mu}_P = 0.5\hat{\mu}_A + 0.5\mu_{Tbill}
\]

\[
\hat{\sigma}^2_P = 0.5^2\hat{\sigma}^2_A + 0.5^2 \cdot 0 + 2 \cdot 0.5 \cdot 0.5 \cdot 0
\]

- \(\hat{\mu}_P\) and \(\hat{\sigma}^2_P\) refer to the estimated mean and variance of our portfolio

- What are we assuming here?
Building Portfolios

▶ What happens if we change the proportions...
Building Portfolios

- What about investing in IBM and ALCOA?

How much more complicated this gets if I am choosing between 100 stocks?
The Market Model

- Empirical version of the CAPM
- Frequently used application of the regression model
- Provides us with a simple way to estimate covariances between all stocks
- Is often used as a benchmark to gauge performance of portfolio managers (Windsor fund example in Section 1)
The Market Model

The model takes the following form:

\[ R_i = \alpha + \beta R_M + \epsilon \quad \text{with} \quad \epsilon \sim N(0, \sigma_i^2) \]

where \( R_i \) is the return on stock \( i \), and \( R_M \) is the return on the market index. That implies:

- \( E(R_i) = \alpha + \beta E(R_M) \)
- \( Var(R_i) = \beta_i^2 Var(R_M) + \sigma_i^2 \)
- \( Cov(R_i, R_j) = \beta_i \beta_j Var(R_M) \)

What about 100 stocks? All we need to do is to run 100 regressions and “choose” our portfolio!!
The Market Model - Portfolio Implications

Let $R_p$ be the return on a portfolio with $N$ securities so that

$$R_p = \sum_{i=1}^{N} w_i R_i$$

where $w_i$ is the weight assigned to security $i$. That implies:

$$E(R_p) = \sum_{i=1}^{N} w_i E(R_i)$$

$$= \sum_{i=1}^{N} w_i \alpha_i + \sum_{i=1}^{N} w_i \beta_i E(R_M)$$
The Market Model - Portfolio Implications

...and

\[
\text{Var}(R_p) = \sum_{i=1}^{N} w_i^2 \text{Var}(R_i) + \sum_{i}^{N} \sum_{j\neq i}^{N} w_i w_j \text{Cov}(R_i, R_j)
\]

\[
= \sum_{i=1}^{N} w_i^2 \beta_i^2 \text{Var}(R_M) + \sum_{i=1}^{N} w_i^2 \sigma_i^2 + \sum_{i}^{N} \sum_{j\neq i}^{N} w_i w_j \beta_i \beta_j \text{Var}(R_M)
\]

\[
= \sum_{i}^{N} \sum_{j}^{N} w_i w_j \beta_i \beta_j \text{Var}(R_M) + \sum_{i=1}^{N} w_i^2 \sigma_i^2
\]

\[
= \left( \sum_{i=1}^{N} w_i \beta_i \right) \left( \sum_{j=1}^{N} w_j \beta_j \right) \text{Var}(R_M) + \sum_{i=1}^{N} w_i^2 \sigma_i^2
\]

\[
= \beta_p^2 \text{Var}(R_M) + \sum_{i=1}^{N} w_i^2 \sigma_i^2
\]
The Market Model - Portfolio Implications

Now, assume that \( w_i = \frac{1}{N} \), i.e., we place equal amount of money in each security...

\[
\begin{align*}
\text{Var}(R_p) &= \beta^2_p \text{Var}(R_M) + \sum_{i=1}^{N} w_i^2 \sigma_i^2 \\
&= \beta^2_p \text{Var}(R_M) + \sum_{i=1}^{N} \frac{1}{N} \sigma_i^2 \\
&= \beta^2_p \text{Var}(R_M) + \frac{1}{N} \sum_{i=1}^{N} \frac{1}{N} \sigma_i^2 \\
&= \beta^2_p \text{Var}(R_M) + \frac{1}{N} \bar{\sigma}^2
\end{align*}
\]
The Market Model - Portfolio Implications

as $N$ grows, i.e., we add more and more securities to our portfolio,

$$
\frac{1}{N} \bar{\sigma}^2 \to 0
$$

and therefore,

$$
\text{Var}(R_p) = \beta_p^2 \text{Var}(R_M)
$$

- The red term is what we call the non-diversifiable risk!
- The blue term is diversifiable, i.e., we can make it go away by spreading our wealth into many stocks.
- The Market Model (CAPM) implies that there exists only ONE SOURCE of risk in the market!
Testing the CAPM

Let’s try to “test” weather or not the CAPM works in reality... in order to so we will focus on the returns of 25 portfolios constructed by sorting firms by their market cap and book-to-price ratio.

If the CAPM is true, all the \( \alpha \)’s have to be zero and we should find no other “systematic” source of common variation in these portfolios... How can we evaluate these implications?

Let’s first run a “market model regression” for each portfolio... that gives us estimates of \( \alpha \)’s and \( \beta \)’s for each security.
Testing the CAPM

This plot shows the estimated $\alpha$ along with its 95% confidence interval for each regression... Does it look good?
Testing the CAPM

Also, if the CAPM is correct, \( E(R_i) = \beta_i E(R_M) \), right? We can look at this by plotting, for each portfolio, the \( \bar{R}_i \) vs the estimated \( \hat{\beta}_i \bar{R}_M \).
Testing the CAPM

Coefficients:

|                | Estimate | Std. Error | t value | Pr(>|t|) |
|----------------|----------|------------|---------|----------|
| (Intercept)    | 1.0187   | 0.3275     | 3.111   | 0.00492  ** |
| Fit0           | -0.7651  | 0.7207     | -1.062  | 0.29942  |

---

Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Residual standard error: 0.2308 on 23 degrees of freedom
Multiple R-squared: 0.04671, Adjusted R-squared: 0.005267
F-statistic: 1.127 on 1 and 23 DF,  p-value: 0.2994

> confint(model0)

<table>
<thead>
<tr>
<th></th>
<th>2.5 %</th>
<th>97.5 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>0.3412927</td>
<td>1.6961865</td>
</tr>
<tr>
<td>Fit0</td>
<td>-2.2560502</td>
<td>0.7257703</td>
</tr>
</tbody>
</table>
Testing the CAPM

Maybe there's is some systematic variation that we are ignoring in the returns... something else might be responsible for the variability in these portfolios...

![Graph showing scatter plot with Size (B/M fixed) on the x-axis, Beta*Rm on the y-axis, and Rbar values for different points. The plot includes a line of best fit.](image-url)
Testing the CAPM
Testing the CAPM

- It looks like there’s a systematic association between the “size” of a firm and its average returns... also true for book-to-price...

- These are known as the size effect and value effect... small firms tend to earn returns too high for their $\beta$’s and the same is true for firms with high book-to-price ratio!
Fama and French proposed an extension to the CAPM to incorporate these effects... they basically added 2 factors that capture the common source of variation related to size and book-to-price... We now have the following regressions:

$$ R_i = \alpha_i + \beta_i R_M + \beta_i^{SMB} SMB + \beta_i^{HML} HML + \epsilon $$

You can think one this model as a generalization to the market model that implies 3 different sources of risk... all else can be diversified away!
Fama-French 3 Factors

The results...
Fama-French 3 Factors

Much nicer, right?!
Fama-French 3 Factors

Coefficients:

|                | Estimate | Std. Error | t value | Pr(>|t|) |
|----------------|----------|------------|---------|----------|
| (Intercept)    | 0.02243  | 0.11313    | 0.198   | 0.845    |
| Fit            | 0.97062  | 0.16251    | 5.973   | 4.33e-06 *** |

---

Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Residual standard error: 0.148 on 23 degrees of freedom
Multiple R-squared: 0.608, Adjusted R-squared: 0.5909
F-statistic: 35.67 on 1 and 23 DF,  p-value: 4.333e-06

> confint(model1)

<table>
<thead>
<tr>
<th></th>
<th>2.5 %</th>
<th>97.5 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-0.2115938</td>
<td>0.2564524</td>
</tr>
<tr>
<td>Fit</td>
<td>0.6344295</td>
<td>1.3068013</td>
</tr>
</tbody>
</table>
Fama-French 3 Factors

And, we have no more systematic variation...
Fama-French 3 Factors

And, we have no more systematic variation...