

EMPIRICAL EVIDENCE AGAINST CAPM: RELATING ALPHAS AND RETURNS TO BETAS

Mayur Agrawal

Debabrata Mohapatra

Ilya Pollak

Purdue University
School of Electrical and Computer Engineering
West Lafayette, IN 47907, USA

ABSTRACT

One of the consequences of the Capital Asset Pricing Model (CAPM) is that the expected excess return of a financial instrument is proportional to the expected excess market return. The proportionality constant, called the instrument's beta, is the coefficient in the linear least-squares fit of the excess return of the instrument with the excess return of the market. CAPM therefore implies that stocks with larger empirical estimates of beta will tend to produce larger returns. Following the testing procedure from a 2006 study by Grantham [5], we analyze this hypothesis using the stock return data for the S&P 500 constituents from 1965 to 2009. We obtain several statistically significant results inconsistent with the hypothesis.

Index Terms— CAPM, alpha, beta, regression, stock, market, finance, statistical significance

1. BACKGROUND

We empirically analyze the Capital Asset Pricing Model (CAPM) [10, 7, 8] in the context of the US stock market, by constructing and tracking over time several portfolios consisting of shares of large publicly traded US companies.

The *daily return* of a portfolio on trading day n is defined as the difference between the portfolio's closing prices on trading days n and $n - 1$, divided by the closing price on day $n - 1$:

$$\text{return}(n) = \frac{\text{price}(n) - \text{price}(n-1)}{\text{price}(n-1)}.$$

Given a risk-free interest rate r , the *excess return* R_s of a portfolio s is defined as the difference between the portfolio's return and the risk-free rate. We assume that the risk-free rate is zero,¹ and therefore the excess return is the same as the return. For now, we view the return as a random variable, denoting the mean and the variance of R_s by μ_s and λ_s , respectively. We additionally define the market return, R_m , to be the return of a portfolio broadly representative of the overall market. We let its mean and variance be μ_m and λ_m , respectively.

The linear least-squares estimator [1] of R_s based on R_m is:

$$\hat{R}_s = \mu_s + \frac{\lambda_{sm}}{\lambda_m}(R_m - \mu_m),$$

where λ_{sm} is the covariance between R_s and the market return R_m . This estimator can be rewritten as follows:

$$\hat{R}_s = \alpha_s + \beta_s R_m,$$

¹Another widely used proxy for the risk-free rate is the interest rate on the 1-month US Treasury Bills [4].

with

$$\beta_s = \frac{\lambda_{sm}}{\lambda_m} \quad (1)$$

$$\alpha_s = \mu_s - \beta_s \mu_m. \quad (2)$$

These two parameters are very widely used both in the financial industry and literature, and are referred to as the *beta* and *alpha* of the portfolio s . Portfolio s could consist of a single stock. In this case, the parameters defined in Eqs. (1,2) are called the beta and alpha of that stock.

CAPM consists of a number of assumptions on the structure of the market and the preferences of the market participants. One consequence of the model is that, for any portfolio, $\alpha_s = 0$, i.e.,

$$\mu_s = \beta_s \mu_m. \quad (3)$$

Both the assumptions of the model and the derivation of Eq. (3) are laid out in a number of textbooks, e.g., [9]. In the remainder of the paper, we describe several experiments which show that the empirical behavior of the US stock market over the past four decades disagrees with Eq. (3) in a statistically significant manner.

2. CONSTRUCTING THE EXPERIMENTS

There are several difficulties with empirically evaluating Eq. (3), as none of the quantities involved in it are directly observable. Moreover, there is no single, universally accepted definition for one crucial ingredient of that equation: the market return. Since we restrict our studies to the US stock market, we choose to follow a large body of literature (see, e.g., [4, 9] and references therein) in using the return of a broad stock index as a proxy for the market return. Specifically, we use the S&P 500 returns [13] as market returns.

We use the following procedure to empirically estimate the alpha and beta of any portfolio s during a time interval of N trading days. We first use the N observed daily returns to estimate the expectation μ_s as the empirical mean of the returns $r_s(1), \dots, r_s(N)$:

$$\hat{\mu}_s = \frac{1}{N} \sum_{n=1}^N r_s(n).$$

We form the empirical covariance of the portfolio returns $r_s(1), \dots, r_s(N)$ and the market returns $r_m(1), \dots, r_m(N)$,

$$\hat{\lambda}_{sm} = \frac{1}{N-1} \sum_{n=1}^N (r_s(n) - \hat{\mu}_s)(r_m(n) - \hat{\mu}_m),$$

where $\hat{\mu}_m$ is the empirical mean return for the market portfolio. The empirical market variance is:

$$\hat{\lambda}_m = \frac{1}{N-1} \sum_{n=1}^N (r_m(n) - \hat{\mu}_m)^2.$$

These are then used to construct estimates of β_s and α_s :

$$\hat{\beta}_s = \frac{\hat{\lambda}_{sm}}{\hat{\lambda}_m} \quad (4)$$

$$\hat{\alpha}_s = \hat{\mu}_s - \hat{\beta}_s \hat{\mu}_m. \quad (5)$$

Eq. (3) suggests the following testing strategy, developed in [5]:

1. On day d , compute $\hat{\beta}_s$ for all stocks that are members of S&P 500 as of that day.
2. Sort all the stocks by their estimated betas, from the smallest to the largest, and partition them into q quantiles.
3. Form q quantile portfolios and analyze the portfolio returns for days $d + 1, \dots, d + K$.
4. Increment d by K and go to Step 1. Iterate until the end of the available data is reached.

In our simulations, we use $N = 5$ years, $K = 1$ year, and $q = 10$. If the stock market is consistent with Eq. (3), then we would observe the following:

- (1) The average returns during days $d + 1, d + 2, \dots, d + K$ for the large-beta quantile portfolios would generally be larger than the average returns for the small-beta quantile portfolios. For example, the data must be consistent with the following consequence of Eq. (3):²

$$\beta_H \mu_L - \beta_L \mu_H = 0, \quad (6)$$

where β_L and β_H are the betas for lowest-beta and highest-beta quantile portfolios, respectively, and where μ_L and μ_H are their respective expected returns.

- (2) The alphas of all the quantile portfolios estimated from days $d + 1, d + 2, \dots, d + K$ would not be different from zero by a statistically significant amount.

In fact, as shown in the remainder of the paper, we observe the opposite: both Eq. (6) and the hypothesis that alpha is always zero for all the portfolios can be rejected with a very high level of confidence.

3. EXPERIMENTAL RESULTS

We use the return data for the S&P 500 constituent stocks and for the S&P 500 index itself, from January 1, 1965 until December 31, 2009. The data is obtained from CRSP [3] through WRDS [14].

On the first trading day of each year, we take all the S&P 500 constituent stocks and use the five preceding years to estimate the alpha and beta for each stock. We then form ten decile portfolios based on the estimated betas. For example, the lowest-decile portfolio contains those 10% of the S&P 500 constituents that have the lowest estimated betas. We allocate the same dollar amount to each stock within a portfolio. We then calculate the daily returns of each portfolio over the following year. For example, during the year 1970 we calculate the daily returns of ten portfolios formed on the basis of betas estimated over the years 1965–1969.

We perform this procedure for all years from 1970 until 2009.³ Note that this simulation can be regarded as an investment strategy: we construct the ten decile portfolios at the beginning of a year, invest in each of them, and track the returns for one year. After one

²Eq. (6) is obtained from Eq. (3) because, according to Eq. (3), both $\beta_H \mu_L$ and $\beta_L \mu_H$ are equal to $\beta_L \beta_H \mu_m$.

³Due to our five-year training window, this requires data over 1965–2009.

year, we re-estimate all the betas, rebalance the ten portfolios accordingly, hold the new portfolios for one more year, etc.⁴

The average annual returns of the ten decile portfolios over the entire 40-year testing period are shown in Fig. 1(a). The average annual S&P 500 return over the same period is 9.9%. To explain how these numbers were computed, we take an example: the lowest-beta decile portfolio whose average annual return is 13.2%. This means that someone who invested \$1 into the lowest-beta portfolio on January 1, 1970 and kept fully reinvesting into the lowest-beta portfolio every year, would end up with $\$1 \cdot (1.132)^{40} = \142.51 on December 31, 2009.

The weight of each company in the S&P 500 index is based on the company's market capitalization [13]. Since our portfolios are equal-weighted, it is reasonable to also compare our portfolios' performance with that of an equal-weighted S&P 500 portfolio. The equal-weighted S&P 500 return for day n is the arithmetic average of the daily returns on day n of all the constituents of S&P 500. The average annual equal-weighted S&P 500 return over our 40-year testing period is 13.1%.

The plot in Fig. 1(a) shows a trend from large returns for low-beta portfolios to small returns for high-beta portfolios. The four lowest-beta portfolios outperform the equal-weighted S&P 500 index, and the remaining six portfolios underperform it. This trend has been pointed out in the literature—for example, in [2] (p. 297) and [5]—and goes squarely against Eq. (6). In the next section, we analyze the statistical significance of this finding.

A very common way of characterizing the performance of portfolios and securities is by calculating their *Sharpe ratio* [11, 12], i.e., the mean excess return divided by the standard deviation of the excess return. The annualized Sharpe ratios for the ten decile portfolios over the period 1970–2009 are shown in Fig. 1(b). The annualized Sharpe ratios for the S&P 500 and equal-weighted S&P 500 portfolios over the same period are 0.64 and 0.80, respectively.

Fig. 1(c) shows the time series plots during 1970–2009 of the total values for the lowest-beta decile portfolio and the highest-beta decile portfolio. They briefly converge during the end of the dot-com bubble in 1999–2000; however, apart from that and a few other brief periods, the lowest-beta portfolio consistently outperforms the highest-beta portfolio.

4. STATISTICAL SIGNIFICANCE OF THE RESULTS

4.1. Returns

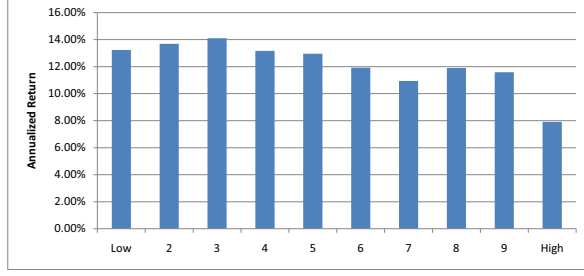
We first test Eq. (6). We let portfolio L be the lowest-beta decile portfolio—i.e., the portfolio that, on any day during a calendar year, contains the stocks whose empirical beta estimated from the preceding five calendar years, is in the lowest decile. We similarly define H , the highest-beta decile portfolio. As described in the previous section, the composition of both portfolios remains unchanged during each calendar year.

We let $R_L(n)$ and $R_H(n)$ be the daily returns on trading day n of portfolios L and H , respectively. We let β_L and β_H be their respective betas, and we let μ_L and μ_H be their respective expected returns. For each trading day n during our 40-year investment period, we form the following quantity:

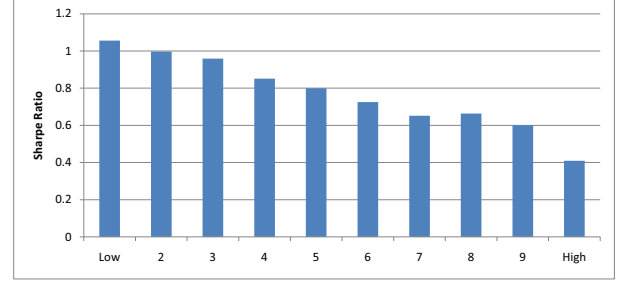
$$R(n) \equiv \beta_H R_L(n) - \beta_L R_H(n).$$

We assume that $R(n)$'s are independent and identically distributed with mean μ and standard deviation σ . Note that their mean μ is

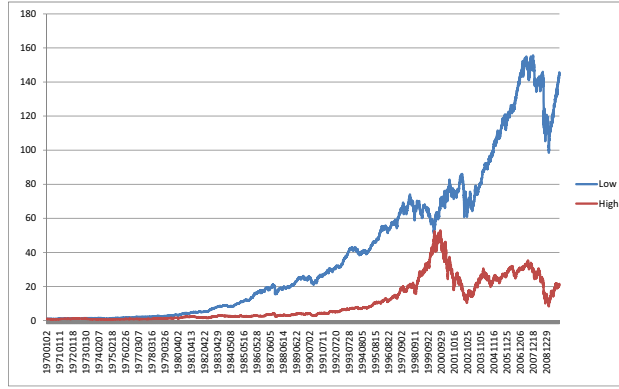
⁴Since the trading costs associated with rebalancing only once per year are very small, we ignore the trading costs in all our experiments and analysis.



(a) Annual returns for beta decile portfolios, 1970–2009



(b) Sharpe ratios for beta decile portfolios, 1970–2009



(c) Highest-beta and lowest-beta portfolio values, 1970–2009

Fig. 1. (a) Average annual returns for the ten beta decile portfolios, 1970–2009. The average annual S&P 500 and equal-weighted S&P 500 returns over the same period are 9.9% and 13.1%, respectively. (b) Annualized Sharpe ratios for the ten beta decile portfolios, 1970–2009. The annualized Sharpe ratios for the S&P 500 and equal-weighted S&P 500 portfolios over the same period are 0.64 and 0.80, respectively. (c) Time series plots of the values for the highest-beta portfolio and the lowest-beta portfolio during 1970–2009.

equal to $\beta_H \mu_L - \beta_L \mu_H$, which is the left-hand side of Eq. (6). Therefore, verifying Eq. (6) is equivalent to testing the hypothesis $\mu = 0$.

We use Eq. (4) over a calendar year to obtain the estimates $\hat{\beta}_L$ and $\hat{\beta}_H$ of the betas of portfolios L and H during that calendar year. We then form the sample mean M and sample standard deviation Σ of $R(n)$'s over the entire 40-year investment period. From our assumptions that $R(n)$'s are independent and identically distributed, it follows that the mean of M is μ and the standard deviation of M is $\sigma/\sqrt{N}$, where $N = 10097$ is the total number of trading days during the 40 years. We estimate the standard deviation of M as $\Sigma/\sqrt{N}$. The specific realizations for our data are: $M = 0.000524$, $\Sigma = 0.00968$, $\Sigma/\sqrt{N} = 0.0000963$. Conditioned on $\mu = 0$, the expectation of M is zero, and therefore the t -statistic is

$$\frac{M}{\Sigma/\sqrt{N}} = \frac{0.000524}{0.0000963} = 5.43,$$

which is a very compelling piece of evidence for rejecting the hypothesis $\mu = 0$. For example, if we assume that the t -statistic follows the t -distribution with $N - 1$ degrees of freedom, we can reject the hypothesis at significance level 2.88×10^{-8} , assuming that the alternative hypothesis is $\mu > 0$.

We can convert the preceding discussion into an implementable investment strategy by computing the estimates $\hat{\beta}_L$ and $\hat{\beta}_H$ for every trading day n over the preceding five calendar years instead of during the current one. In other words, for a trading day n which falls in a calendar year y , we estimate the betas using the data from January 1 of year $y - 5$ until December 31 of year $y - 1$. The quantity

$$P(n) = \frac{\hat{\beta}_H R_L(n) - \hat{\beta}_L R_H(n)}{\hat{\beta}_H + \hat{\beta}_L}$$

is then the profit on day n from entering into the following positions on day $n - 1$: buying $\$ \hat{\beta}_H / (\hat{\beta}_H + \hat{\beta}_L)$ worth of portfolio L and selling short $\$ \hat{\beta}_L / (\hat{\beta}_H + \hat{\beta}_L)$ worth of portfolio H . The time se-

ries of $P(n)$, over the entire investment period, is the profit stream of a strategy that makes an investment in this manner every trading day and liquidates after each trading day. Note that the initial gross investment (i.e., the dollar amount of long positions plus the dollar amount of short positions) for this strategy for each trading day is always \$1. The sample mean and sample standard deviation of $P(n)$ are 0.000316 and 0.0051, which produces a t -statistic of 6.21, Sharpe ratio of 0.98, and correlation coefficient with S&P 500 returns of 0.22 and with equal-weighted S&P 500 returns of 0.26.⁵ We therefore have constructed a strategy which has very low correlation with the market yet produces returns which are above zero by a statistically significant margin, and which in fact has a better Sharpe ratio than the market.

4.2. Alphas

In order to evaluate the statistical significance of our alpha estimates, we make the following modeling assumptions, which are somewhat different from those made in the preceding sections. Suppose that we are evaluating the statistical significance over a time window of duration D trading days, say, days $1, 2, \dots, D$. We assume that the market returns for all these days are deterministic and that we have the following relationship between the returns of the q -th quantile portfolio and the market returns:

$$r_q(n) = \alpha_q + \beta_q r_m(n) + u_q(n), \text{ for } n = 1, \dots, D, \quad (7)$$

where $u_q(1), \dots, u_q(D)$ are independent and identically distributed zero-mean Gaussian random variables with variance λ_q . It can then be shown [6] that Eqs. (4,5) are unbiased estimates of β_q and α_q , respectively, and that the covariance matrix of the vector $(\hat{\alpha}_q, \hat{\beta}_q)^T$ is:

$$\lambda_q (X^T X)^{-1}, \quad (8)$$

where X is a $D \times 2$ matrix whose first column consists of all 1's and whose second column is the D -dimensional vector of market returns, $(r_m(1), \dots, r_m(D))^T$. An unbiased estimate of λ_q is obtained as follows [6]:

$$\hat{\lambda}_q = \frac{1}{D-2} \sum_{n=1}^D (r_q(n) - \hat{\alpha}_q - \hat{\beta}_q r_m(n))^2.$$

Using this to replace λ_q in Eq. (8) gives the following standard error for the estimate $\hat{\alpha}_q$:

$$\hat{\sigma}_{\hat{\alpha}} = \sqrt{\hat{\lambda}_q x_{11}},$$

where x_{11} is the first diagonal entry of the matrix $(X^T X)^{-1}$.

We test, for the portfolios L and H , whether their alpha estimates are different from zero by a statistically significant margin. Our null hypothesis, for each portfolio, is $\alpha_q = 0$, and therefore the t -statistic is:

$$T_q = \frac{\hat{\alpha}_q}{\sqrt{\hat{\lambda}_q x_{11}}}.$$

The alternative hypothesis for portfolio L is $\alpha_L > 0$. The t -test at significance level α is to reject the null hypothesis if $T_L > \Psi_{D-2}^{-1}(1 - \alpha)$, where Ψ_{D-2} is the t -CDF with $D - 2$ degrees of freedom. We perform this test for each calendar year during the 40-year testing period. In eight out of the 40 years (1970,

1975, 1976, 1979, 1981, 1982, 1985, 1993) the null hypothesis is rejected at significance level 1%. In 1976 and 1982, it can be rejected at significance level 0.01%.

The alternative hypothesis for portfolio H is $\alpha_H < 0$. The t -test at significance level α is to reject the null hypothesis if $T_H < \Psi_{D-2}^{-1}(\alpha)$. In five out of the 40 years (1970, 1973, 1984, 1985, 1986), the null hypothesis is rejected at significance level 1%. In 1985, it can be rejected at significance level 0.001%. These findings, like the findings in the previous subsection, contradict Eq. (3) and therefore contradict CAPM.

5. ACKNOWLEDGMENTS

This work grew out of a semester project done by the first two authors as part of a financial engineering course taught by the third author at Purdue's School of Electrical and Computer Engineering in Spring 2010. We would like to acknowledge Purdue University Libraries for subscribing to CRSP and other financial databases through WRDS and thus enabling us to access the data used in this research.

6. REFERENCES

- [1] D. P. Bertsekas and J. N. Tsitsiklis. *Introduction to Probability*. Athena Scientific, Belmont, MA, 2008.
- [2] Z. Bodie, A. Kane, and A. J. Marcus. *Investments*, 8th Edition. McGraw-Hill/Irwin, 2009.
- [3] ©2010 CRSP®, Center for Research in Security Prices. Booth School of Business, The University of Chicago. Used with permission. All rights reserved. www.crsp.chicagobooth.edu
- [4] E. F. Fama and K. R. French. "Common risk factors in the returns on stocks and bonds", *Journal of Financial Economics*, vol. 33, no. 1, pp. 3–56, Feb. 1993.
- [5] J. Grantham. "The nature of risk III: The role of value and the premium for excitement or speculation", *GMO Letters to the Investment Committee IX*, Oct. 2006.
- [6] J. D. Hamilton. *Time Series Analysis*. Princeton University Press, 1994.
- [7] J. Lintner. "The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets", *Review of Economics and Statistics*, vol. 47, no. 1, pp. 13–37, Feb. 1965.
- [8] J. Mossin. "Equilibrium in a capital asset market", *Econometrica*, vol. 34, no. 4, pp. 768–783, Oct. 1966.
- [9] D. Ruppert. *Statistics and Finance: An Introduction*. Springer, 2004.
- [10] W. F. Sharpe. "Capital asset prices: A theory of market equilibrium under conditions of risk", *Journal of Finance*, vol. 19, no. 3, pp. 425–442, Sep. 1964.
- [11] W. F. Sharpe. "Mutual fund performance", *Journal of Business*, vol. 39, no. S1, pp. 119–138, 1966.
- [12] W. F. Sharpe. "The Sharpe ratio", *Journal of Portfolio Management*, vol. 21, no. 1, pp. 49–58, Fall 1994.
- [13] "S&P US Indices Methodology", Standard & Poor's, March 2006, www2.standardandpoors.com/spf/pdf/index/US_Methodology.pdf.
- [14] Wharton Research Data Services (WRDS) was used in preparing this paper. This service and the data available thereon constitute valuable intellectual property and trade secrets of WRDS and its third-party suppliers.

⁵Note, however, that we are still ignoring the trading costs which are, for this strategy, non-negligible, since it rebalances every day.