

Measuring Stock Market Risk Using Correlation and Linear Regression

Hao Zhang^a

Farmer School of Business, Miami University, Oxford, OH, United States

^aEmail: zhang978@miamioh.edu

Abstract

This report evaluates total and systematic risk for eight stocks-Microsoft, Exxon Mobil, Caterpillar, Johnson & Johnson, McDonald's, Sandisk, Qualcomm, and Procter & Gamble-using the S&P 500 return as the market benchmark. Total risk is summarized by the mean, standard deviation, extrema, and a sample Sharpe ratio calculated with a zero risk-free rate. Systematic exposure is estimated from a simple market-model regression in which each stock return is regressed on the market return; beta measures market sensitivity, while R^2 shows the proportion of return variation explained by the market. In the supplied sample, Sandisk has the highest volatility and beta, Caterpillar has the highest Sharpe ratio, and McDonald's has the largest R^2 . The results indicate that volatility, risk-adjusted performance, and market sensitivity capture different dimensions of risk and should not be interpreted interchangeably. Because the sample period and data source are not documented in the draft dataset, the findings should be viewed as illustrative rather than as current investment recommendations.

Keywords

Stock Market Risk; Systematic Risk; Beta Coefficient; Linear Regression; Sharpe Ratio.

1. Introduction

Risk measurement supports portfolio selection and informed decision-making. Mean-variance analysis established volatility and covariance as central inputs to portfolio risk [1]. The capital asset pricing model subsequently distinguished systematic market exposure from diversifiable firm-specific risk and represented that exposure through beta [2]. This report analyses eight individual stocks using dispersion measures, the Sharpe ratio, and a market-model regression against the S&P 500.

The single-index beta model is useful because it separates market-related variation from firm-specific variation. Nevertheless, beta is an exposure measure rather than a complete forecast of return. Empirical research has tested the beta-return relationship [4], while later multifactor evidence shows that size and value-related factors can also explain return variation [5]. Accordingly, this report uses beta and R^2 as descriptive measures and does not treat them as stand-alone trading signals.

Recent evidence shows that beta estimates can differ across estimation choices and that systematic exposure can vary over time and investment horizons [7-8]. Therefore, the static beta estimates reported here should be interpreted as sample-specific measures rather than permanent firm characteristics.

The remainder of the report contains three sections. Section 2 introduces the linear regression model. Section 3 presents and discusses the descriptive statistics, beta estimates, and R^2 results. Section 4 summarizes the findings, limitations, and directions for further analysis.

2. Linear Regression

Simple linear regression estimates the relationship between a dependent variable and one independent variable. Following standard econometric notation [6], the market model used here regresses each stock return on the S&P 500 return:

$$R_{i,t} = \alpha_i + \beta_i R_{m,t} + \varepsilon_{i,t}.$$

Here, $R_{i,t}$ is stock i 's return, $R_{m,t}$ is the market return, α_i is the intercept, β_i is the slope measuring market sensitivity, and $\varepsilon_{i,t}$ is the stock-specific disturbance. In a simple regression with variation in the market return, the zero conditional mean assumption supports unbiased OLS coefficient estimates. Homoskedasticity and no serial correlation are additional assumptions for the conventional standard-error formulas, while approximate normality is relevant to exact small-sample inference [6]. The relevant disturbance assumptions are summarized below:

- 1) Zero conditional mean: $E(\varepsilon_{i,t} | R_{m,t}) = 0$.
- 2) Homoskedasticity: $\text{Var}(\varepsilon_{i,t} | R_{m,t}) = \sigma^2$.
- 3) No serial correlation: $\text{Cov}(\varepsilon_{i,t}, \varepsilon_{i,s}) = 0$ for $t \neq s$.
- 4) Approximate normality of $\varepsilon_{i,t}$ is used for exact small-sample inference.

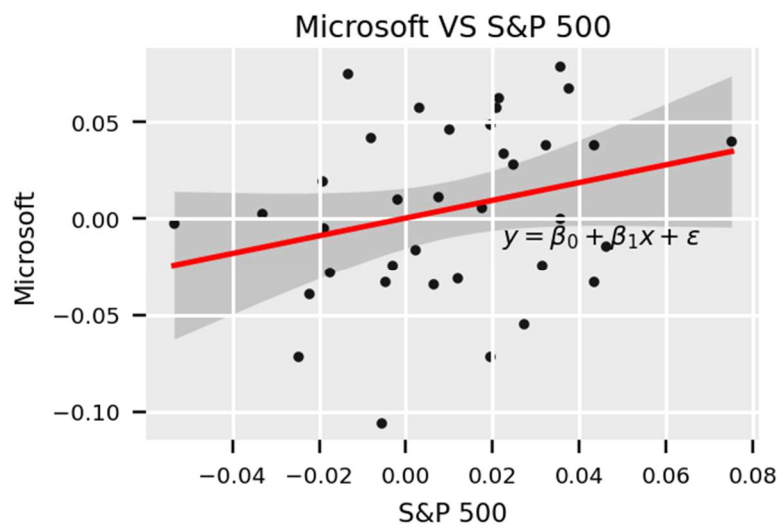


Figure 1. Simple Linear Regression of Microsoft Returns on S&P 500 Returns

The fitted line summarizes the direction and magnitude of the relationship. In Figure 1, the generic coefficients β_0 and β_1 correspond to α_i and β_i , respectively, in the market-model notation used above. Residuals capture variation not explained by the market, and R^2 measures goodness of fit. These diagnostics should be interpreted together rather than as interchangeable measures of risk.

3. Measuring Stock Market Risk Using Linear Regression

The supplied figures and summary statistics are internally consistent with 36 paired return observations for each stock and the S&P 500 benchmark. However, the source file does not identify the observation dates, return frequency, data provider, price-adjustment convention, or precise return-construction method. Accordingly, the analysis is descriptive and sample-specific.

3.1. Descriptive Statistics of Individual Stocks

This subsection reports the mean, standard deviation, maximum, minimum, and Sharpe ratio for each stock. The Sharpe ratio is defined as $(\text{mean return} - \text{risk-free rate}) / \text{standard deviation}$ [3]. To match the supplied calculations, the risk-free rate is set to $R_0 = 0$; therefore, the values are sample ratios and are not annualized.

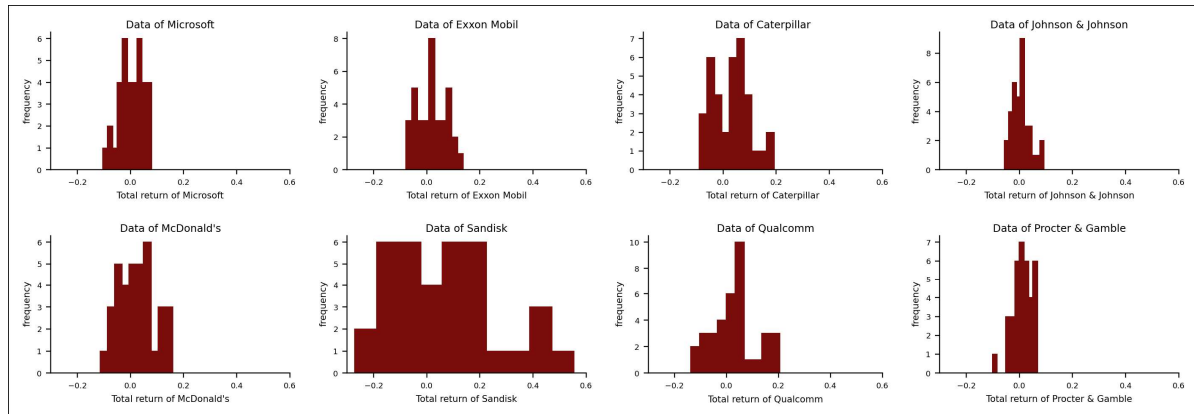


Figure 2. Sample Return Distributions for the Eight Stocks

Table 1. Descriptive Statistics for the Eight Stocks

Stocks	Mean	Standard deviation	Maximum	Minimum	Sharpe ratio
Microsoft	0.0050256	0.0453716	0.08883	-0.08201	0.110765325
Exxon Mobil	0.0166372	0.0553401	0.23217	-0.11646	0.300635525
Caterpillar	0.0300972	0.0685633	0.21850	-0.10060	0.438969536
Johnson & Johnson	0.0052956	0.034866	0.10334	-0.05917	0.151884357
McDonald's	0.0244742	0.0680964	0.18260	-0.11440	0.359405196
Sandisk	0.0692619	0.1953974	0.50170	-0.28330	0.354466846
Qualcomm	0.0283603	0.0861869	0.21060	-0.12170	0.329055808
Procter & Gamble	0.0105889	0.0370667	0.08783	-0.05365	0.285671506

Table 1 shows that Sandisk has the largest standard deviation (0.1953974) and is therefore the most volatile stock in the sample. Johnson & Johnson has the smallest standard deviation (0.034866), indicating the lowest total volatility. Caterpillar has the highest Sharpe ratio (0.438969536), which represents the strongest risk-adjusted performance under the maintained assumption $R_0 = 0$. These rankings describe the supplied sample and should not be generalized beyond it without a documented sample period and data source.

3.2. Evaluating Market Exposure and Correlation

Table 2. Estimated Beta Coefficients and Intercepts

Stocks	Beta (β)	Intercept (α)
Microsoft	0.4583	0.0004
Exxon Mobil	0.7309	0.00926
Caterpillar	1.4932	0.015
Johnson & Johnson	0.0088	0.00521
McDonald's	1.5031	0.0093
Sandisk	2.6048	0.043
Qualcomm	1.4139	0.0141
Procter & Gamble	0.5065	0.00548

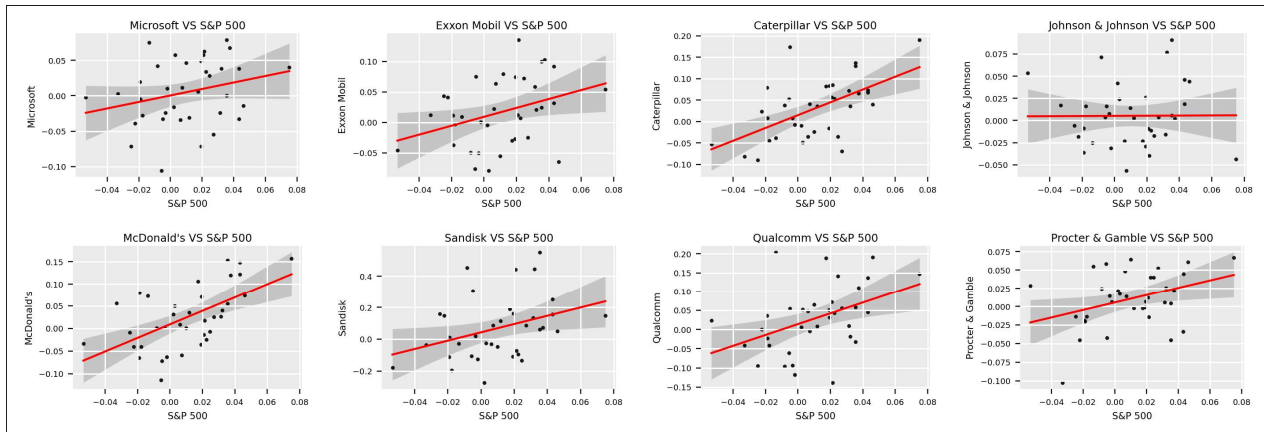


Figure 3. Stock Returns Regressed on S&P 500 Returns

Section 3.1 considers total volatility; this subsection examines systematic market exposure. Figure 3 plots stock returns against S&P 500 returns, and Table 2 reports the estimated intercepts and beta coefficients. Within the capital asset pricing framework, beta reflects covariance with the market scaled by market variance [2].

Because the underlying observations are unavailable, standard errors, t statistics, p values, confidence intervals, and residual diagnostics cannot be reliably recomputed or independently verified. The coefficient estimates are therefore interpreted descriptively rather than as formal evidence of statistical significance.

Table 2 shows that Sandisk has the largest beta (2.6048), indicating the greatest sensitivity to market movements in the sample. A one-percentage-point change in the market return is associated with an estimated 2.6048-percentage-point change in Sandisk's return, on average. Johnson & Johnson has the smallest beta (0.0088), suggesting very limited market sensitivity. Beta does not, however, guarantee superior returns in rising markets or protection in falling markets; it measures exposure, not realized performance [4].

3.3. Explanatory Power of Market Returns

To assess how much of each stock's variation is associated with the market, the analysis decomposes total variation into the total sum of squares (SST) and the explained sum of squares (ESS). $R^2 = \text{ESS}/\text{SST}$ gives the share of variation in the stock return explained by the S&P 500 regression.

Table 3. Explained Variation and R^2 for the Eight Stocks

Stocks	SST	ESS	$R^2 = \text{ESS}/\text{SST}$
Microsoft	0.07205037	0.005093961	0.0707
Exxon Mobil	0.10718843	0.012959082	0.1209
Caterpillar	0.16453241	0.054098258	0.3288
Johnson & Johnson	0.04254733	0	0
McDonald's	0.16229919	0.054824666	0.3378
Sandisk	1.33630504	0.164632781	0.1232
Qualcomm	0.25998636	0.048513455	0.1866
Procter & Gamble	0.04808791	0.006227384	0.1295

The coefficient of determination is defined as:

$$R^2 = \text{ESS} / \text{SST}, \quad 0 \leq R^2 \leq 1.$$

where:

$$ESS = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2, \quad SST = \sum_{i=1}^n (y_i - \bar{y})^2.$$

Table 3 shows that McDonald's has the highest R^2 (0.3378), so the market explains 33.78% of its return variation within this sample. Johnson & Johnson has the lowest reported R^2 (approximately 0), indicating that the fitted market model explains almost none of its variation. For Microsoft, $R^2 = 0.0707$ means that 7.07% of return variation is explained by the market, while the remainder is associated with omitted influences and idiosyncratic variation. A low R^2 does not by itself imply low total risk or poor performance.

In a simple regression with an intercept, the coefficient of determination equals the square of the Pearson correlation coefficient:

$$R^2 = r^2, \quad r = \text{sign}(\beta)\sqrt{R^2}.$$

Because all estimated betas in Table 2 are positive, $r = \sqrt{R^2}$ for the reported regressions.

Thus, R^2 close to 1 would indicate a strong linear association between a stock and the S&P 500, whereas R^2 close to 0 indicates little linear explanatory power. This statistic describes association within the sample and does not establish a causal relationship.

4. Conclusion

This study compares total risk, risk-adjusted performance, and systematic market exposure for eight stocks. In the supplied sample, Sandisk has the highest standard deviation and beta, Caterpillar has the highest Sharpe ratio under $R_0 = 0$, Microsoft has the lowest Sharpe ratio, McDonald's has the highest R^2 , and Johnson & Johnson has the lowest R^2 . These results demonstrate that volatility, the Sharpe ratio, beta, and R^2 answer different questions and should not be treated as substitutes. The market model is intentionally parsimonious, but it cannot capture firm-specific news, nonlinear relationships, changing betas, or additional common factors documented in multifactor asset-pricing research. Future work should document the data source, observation dates, and return frequency; use an observed risk-free rate; test for heteroskedasticity and serial correlation and report appropriate robust standard errors; and compare static beta estimates with rolling, time-varying, and multifactor specifications.

References

- [1] Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91. <https://doi.org/10.1111/j.1540-6261.1952.tb01525.x>
- [2] Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3), 425–442. <https://doi.org/10.1111/j.1540-6261.1964.tb02865.x>
- [3] Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business*, 39(1), 119–138. <https://doi.org/10.1086/294846>
- [4] Fama, E. F., & Macbeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3), 607–636. <https://doi.org/10.1086/260061>
- [5] Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3–56. [https://doi.org/10.1016/0304-405X\(93\)90023-5](https://doi.org/10.1016/0304-405X(93)90023-5)
- [6] Wooldridge, J. M. (2020). *Introductory econometrics: A modern approach* (7th ed.). Cengage.

- [7] Hong, J., Yu, X., Xiao, W., et al. (2022). The dispersion of beta estimates and the investors' heterogeneous beliefs: Evidence from the stock market in China. *International Review of Economics & Finance*, 79, 540–550. <https://doi.org/10.1016/j.iref.2022.02.025>
- [8] Mestre, R. (2023). Stock profiling using time-frequency-varying systematic risk measure. *Financial Innovation*, 9, 52. <https://doi.org/10.1186/s40854-023-00457-7>