EMPIRICAL TESTING: EXECUTIVE MEMORANDUM

Gayathree Gopi: gg29234
Medha Nalamada: mrn789
Ashley Soto: aps2935
Kennedy Zapalac: kmz459

DATE: Oct 29 2024
SUBJECT: Empirical Testing

Introduction & Objective

Our team has performed an empirical analysis on two periods of portfolio returns: July 1963–June 1993 and September 1994–August 2024. By comparing these two periods against each other and examining them individually, we will assess two commonly used models: the Capital Asset Pricing Model (CAPM) and the Fama-French 3-Factor Model (FF3F). We aim to provide our hedge fund with an overview of the models and their strength in capturing components of systematic risk and predicting returns.

Our data set, acquired from the Ken French Data Library, included 25 portfolios ranked upon size and book-to-market ratio. Our portfolios range across small-cap to large-cap stocks, varying from low book-to-market ratios (growth stocks) to high book-to-market ratios (value stocks) and are titled accordingly. In addition, we have the factor returns for all time periods, including RMRF (market risk premium), SMB (size factor), and HML (value factor). We will perform our analysis for all 25 of these portfolios.

CAPM TEST

CAPM is a strong theoretical framework developed in the 1960s by Sharpe, Litner, Mossin, and Treynor to explain the relationship between expected return and exposure to market risk. The model defines *beta* as a measure of an asset's sensitivity to the market risk premium. With this 1-factor model, we will perform a series of empirical tests to assess the validity of the model. This is important given many concerns have been raised about the validity of CAPM. Many people still use the model to price assets due to its simplicity, but we want you to be aware of the weaknesses of this model should you choose to use it.

Time Series Regression

Our team regressed the asset's historical excess returns against the market's historical excess returns to estimate beta and alpha for CAPM. We also estimated the standard errors and t-statistics for alpha and beta. Exhibit 1 includes our results from the regression for both time periods, and Exhibit 2 displays the time series regression equation.

Beta represents that asset's exposure to systematic risk. If an asset has a beta of 1, its expected return moves in tandem with the market return. If the asset beta is greater than 1, the asset's returns are more sensitive to market risk, and if the asset beta is less than 1, the asset's returns are less sensitive to market risk. We expect all of our betas to be greater than 0 because only the risk-free asset can have a beta of 0 under the CAPM.

Alpha represents excess return above what would be expected by the model. If CAPM is valid and accurately predicts an asset's returns, then we expect the alphas to jointly be 0. This is because the CAPM assumes that every asset's returns can be explained by their exposure to market risk. If an asset's expected returns are not equal to 0 when the market risk premium is 0, then their returns aren't being fully explained by exposure to market risk.

Our formal null hypothesis is that the betas and the alphas are not statistically significantly different from 0. For both time periods, the betas are statistically significant and confirm that all of our assets are risky, as they should be. Some of the alphas appear to be significantly different from 0 ($t > 1.96$), which is our initial indicator that the CAPM may not be accounting for all the factors related to an asset's expected return. However, it's better to perform a joint hypothesis test for this so we don't have a multiple testing problem. You will find our results for this under the "GRS F-Test Section".

Cross-Sectional Regression

The time series regression (TSR) yields beta estimates for each portfolio, which measure each portfolio's sensitivity to market risk. According to the CAPM, you should expect to earn a higher return for exposure to market risk. You can assess whether this expectation is true by running a cross-sectional regression of portfolio expected returns against the betas to estimate λ, which tells us how much return is expected per unit of beta risk. Exhibit 3 displays the equation for cross-sectional regression.

Figure 1: CAPM CSR Results

Time Period	**Lambda (β)**	**Standard Error**	**T-Statistic**	**Significant?**
07/1936 - 06/1993	-0.371056393	0.328054019	-1.131083207	No
09/1994 - 08/2024	-0.438794719	0.229607821	-1.91106172	No

The cross-sectional regression is testing whether the lambdas are not statistically significantly different from 0 (the null hypothesis). We can see that for both time periods, the values are very similar, indicating similar conclusions. For each time period, our t-statistic values show that we fail to reject the null hypothesis (||t||< 1.96). There is no strong evidence that the lambdas are not 0. This means that the assets are not earning a risk premium for exposure to market risk.

However, this test fails to account for correlations between portfolio returns at each time period t. We anticipate that our standard errors here are understated, so we will use the Fama MacBeth Procedure to account for cross-sectional correlations. This likely will not change our conclusions since the estimate of lambda generally remains the same and the standard errors increase, decreasing the likelihood of significant results.

Fama MacBeth Procedure

Instead of relying on a single cross-sectional regression, the Fama MacBeth Procedure (FMP) performs a series of cross-sectional regressions (expected returns against betas) at each time point, then averages the coefficients (lambdas) across time. The standard error of lambda is calculated by getting the standard error of the lambda estimates for each time period. By using FMP, we can obtain more accurate and reliable standard errors for our estimated risk premiums, helping us make more confident inferences about the significance of the lambdas. Exhibit 4 displays the equation for Fama Macbeth Procedure.

Figure 2: CAPM FMP Results

Time Period	**Lambda (β)**	**Standard Error**	**T-Statistic**	**Significant?**

07/1936 - 06/1993	0.371056393	0.461189584	-0.804563688	No
09/1994 - 08/2024	0.438794719	0.640013171	-0.685602639	No

While our lambda values for either time period did not change, we can see the standard errors have increased from the CSR results, showing that the FMP test did a better job of capturing cross-sectional correlations between portfolios. The t-statistic slightly increased, but our conclusion remains the same. The lambdas are not statistically significantly different from 0 (fail to reject the null hypotheses). Investors are not being compensated for their exposure to market risk, which violates the assumptions of CAPM. This indicates building a portfolio based on market exposure risk may not be wise because you will not earn a risk premium for your exposure to market exposure risk.

Time Series Joint Hypothesis: GRS F-Test

As stated before, the alphas represent the portfolio's return that cannot be explained by a model's factors. All the alphas from CAPM should jointly be 0 or else an asset's exposure to market risk does not fully explain its expected return. The GRS F-Test is used to test the joint hypothesis that all portfolio alphas from the time series regression are equal to zero. For CAPM, with one factor being the market risk premium, the results are as follows:

Figure 3: CAPM GRS F-Test Results

Time Period	$f_{(GRS)}$	**Critical Value**	**Reject/Fail to Reject Null?**
07/1936 - 06/1993	2.401097993	1.539109951	Reject
09/1994 - 08/2024	3.863377190	1.539109951	Reject

The formula used to calculate the F-statistic and critical value is provided in Exhibit 5. We assume that the residual returns are the excess returns minus the predicted returns. For CAPM, the predicted returns for each portfolio i at time t are estimated by $\alpha_i + \beta_i \times RMRF_t$.

The results in Figure 3 show that, for both time periods, the calculated F-statistics exceed the critical value of 1.5391 (significance level of $p < 0.05$). This indicates that the null hypothesis—that all alphas are jointly equal to zero—is rejected in both periods. Since we reject the null hypothesis in each period, this implies that the CAPM model fails to fully explain the returns for the tested portfolios across both time periods. There are systematic excess returns, or non-zero alphas, that the CAPM model cannot account for, suggesting potential for the presence of other factors influencing portfolio returns.

This outcome shows that CAPM–while theoretically sound, simplistic, and useful–may lack additional factors necessary to capture variations in portfolio returns, encouraging us to extend our analysis to include a multi-factor model.

Fama-French 3-Factor Model

The Fama-French 3-Factor Model (FF3F) was developed in the early 1990s by Fama and French and acts as an extension to CAPM. The model adds two additional factors, size and value, for a total of 3 factors used to explain the variability in stock returns. The size factor (SMB) captures the tendency for small-cap stocks to outperform large-cap stocks, and the value factor (HML) captures the tendency for firms with high book-to-market ratios to outperform firms with lower ratios. We can perform the same tests from the CAPM section to observe if our analysis will change with the inclusion of more factors.

Time Series Regression

Instead of one factor with CAPM, this TSR test will estimate $b\ (RMRF),\ s\ (SMB), and\ h\ (HML)$: the slopes from a multivariate regression, regressing excess returns of our portfolios against all factor returns. By using multivariate regression instead of a univariate regression for each of the factors, we can isolate the effect of each of the factors. Our results for the TSR are provided in Exhibit 6.

Similar to the time series above, our null hypothesis here is that each of the coefficients is not statistically significantly different from 0. The coefficient of b, s, and h generally appear to be statistically significantly different from 0, which means that RMRF, SMB, and HML are related to a portfolio's returns while holding the other factors constant. Alpha is occasionally statistically significantly different from 0. We encounter issues with multiple testing again though, so you should refer to the GRS F-test for a more formal analysis of the alphas.

Cross-Sectional Regression

Following the same intuitions from CAPM, we can perform a CSR test to estimate the lambdas for each of the factors to show the risk premiums associated with each factor in the FF3F model. By regressing the average returns of multiple portfolios on $b,\ s,\ and\ h$, the CSR provides estimates for each lambda, representing the expected return premium per unit of exposure to each factor (market, size, and value). This allows us to examine if the factors contribute statistically significant premiums, meaning investors are compensated for bearing these risks.

Figure 4: FF3F CSR Results

Time Period	λ_b	$\sigma(\lambda_b)$	$t(\lambda_b)$	λ_s	$\sigma(\lambda_s)$	$t(\lambda_s)$	λ_h	$\sigma(\lambda_h)$	$t(\lambda_h)$
07/1936 - 06/1993	-0.09615	0.44640	-0.21540	0.20025	0.05218	3.83759	0.51373	0.06525	7.87311
09/1994 - 08/2024	-0.78153	0.31971	-2.44443	0.03081	0.05571	0.55302	0.13370	0.06863	1.94811

In time period 1, we can see that while the market risk premium still shows no significance, the size and value factors are significant and strongly capture the variability in returns ($||t|| > 1.96$). The size factor has a positive lambda of 0.20025 with a t-statistic of 3.83759, indicating a statistically significant positive premium for exposure to small-cap risk while holding the other factors constant. The value factor also shows a significant lambda of 0.51373 with a very high t-statistic of 7.87311, suggesting a strong premium for exposure to value risk during this period. However, the results in time period 2 offer different conclusions. We can see that the significance of size and value factors decreased by a large magnitude, making them insignificant for the most recent 30 years. This may reflect structural changes in financial markets, such as increased market efficiency, the influence of institutional investors, and the rise of new risk factors that have become more prominent in recent years. The factors could have less impact now since they were discovered and widely disseminated.

FMP Regression

Performing the FMP regression will again help us understand the true standard errors of our factor lambdas and allow us to be more confident in our inferences by accounting for cross-sectional correlations.

Figure 5: FF3F FMP Results

Time Period	λ_b	$\sigma(\lambda_b)$	$t(\lambda_b)$	λ_s	$\sigma(\lambda_s)$	$t(\lambda_s)$	λ_h	$\sigma(\lambda_h)$	$t(\lambda_h)$

07/1936 - 06/1993	-0.09615	0.42951	-0.22387	0.20025	0.15453	1.29583	0.51373	0.13955	3.6813
09/1994 - 08/2024	-0.78153	0.42395	-1.84341	0.03081	0.17734	0.17373	0.13370	0.18285	0.7312

As predicted, the standard errors have increased for most factors, but $\sigma(\lambda_b)$ has actually slightly decreased in time period 1. This anomaly might arise if the variability in factor exposure for the market risk premium remains relatively constant across time periods, which lessens the influence of cross-sectional correlations on its standard error under FMP. Now the only factor with a significant risk premium is the value factor in period 1 ($||t|| > 1.96$). We would be hesitant to recommend building a portfolio based on these factors because you may not earn a risk premium by picking assets that have high market risk exposure, high small-cap risk exposure, and high value risk exposure.

Time Series Joint Hypothesis: GRS F-Test

For the FF3F model, we must still test the significance of the alphas to determine whether the model fully explains the returns of the portfolios. If the alphas are significantly different from zero, it suggests that there are additional sources of return not captured by the three factors. We will now use a different estimate of predicted returns to calculate the residuals for portfolio i in time period t:

$\alpha_i + b_i \times RMRF_t + s_i \times SMB_t + h_i \times HML_t$.

After accounting for all the factors and subtracting from the excess returns, we follow the formula in Exhibit 5 to perform the test

Figure 6: FF3F GRS F-Test Results

Time Period	$f_{(GRS)}$	**Critical Value**	**Reject/Fail to Reject Null?**
07/1936 - 06/1993	1.623394774	1.539308705	Reject
09/1994 - 08/2024	3.854349803	1.539308705	Reject

Our results indicate that the F-statistic is greater than the critical value at the significance level of $p < 0.05$ for both time periods, which means that we reject the null hypothesis that the alphas are jointly statistically significantly different from 0. Similar to CAPM, this rejection suggests that the Fama-French 3-Factor Model does not fully capture the returns of the portfolios across both time periods, as there are statistically significant alphas that remain unexplained by the model.

Python Tool

Our team has also created an Empirical Finance Tool in Python to perform all of the tests mentioned above. The tool queries the user for the data file (must include portfolio returns and factor data), and also asks them to provide a date range in YYYYMM format. We included input validation to make sure the data file exists, ensuring the number of periods is a positive integer, and that the start and end dates are formatted as expected. The program was written in an object-oriented method so all of the data is represented as attributes within a class. There are two different classes for CAPM and the FF3F model, with different functions to perform TSR, CSR, FMP, and the GRS F-Test. The model accounts for multiple time periods, so the user can select as many periods as are required. Afterwards, all the tests will be performed automatically. Finally, the tool writes to Excel so the user can view the results to assess the validity and usefulness of CAPM and Fama-French 3-Factor Model for predicting an asset's expected return.

Analysis and Conclusions

Based on our comprehensive analysis of both the CAPM and Fama-French 3-Factor Model (FF3F) across two distinct time periods, our team concluded that both models exhibit limitations in fully explaining portfolio returns, especially in the more recent period. The GRS F-Test results for both CAPM and FF3F indicate that we must reject the null hypothesis that the alphas are jointly equal to zero across both periods, suggesting that neither model sufficiently captures all the factors influencing portfolio returns. This finding indicates the presence of systematic excess returns that remain unexplained by the factors in each model.

Significant constraints were found in the CAPM results, as indicated by statistically significant non-zero alphas, suggesting that market risk is not a sufficient explanation for returns. The lambda for CAPM from the Fama Macbeth Procedure also wasn't significant, indicating that investors would not earn a risk premium by investing in assets with higher exposure to market risk. By adding the size and value factors, the FF3F model enhances CAPM by lowering some of the prior period's unexplained returns, or alphas. The FF3F model, however, is also inadequate in the more recent era, as the null hypothesis for the alphas is still rejected, suggesting that there may be other elements besides market, size, and value required to account for the variability in returns. The lambdas from the Fama Macbeth Procedure are also insignificant for most of the factors during both time periods, except for HML in time period 1, indicating that you would not earn a risk premium for acquiring assets with exposure to market risk, low market capitalization risk, or high book-to-market value risk.

While the FF3F model provides a more comprehensive framework than CAPM, the persistence of significant alphas in both models suggests that neither fully captures the sources of return variation across portfolios, particularly in recent decades. These findings emphasize the need for continued development and refinement of asset pricing models. Models have developed over time to include additional factors such as profitability, momentum, or investment, to better align with modern market dynamics and more accurately explain the cross-section of returns. In further empirical explorations, with more resources, our team can build additional analytical tools and perform tests using larger factor models, such as the Fama-French 5-Factor or the Carhart 4-Factor Model, to dive deeper into the sources of return variation. In doing so, we could enhance the explanatory power of our models and better inform investment strategies for our clients.

APPENDIX: EXHIBITS

Exhibit 1: CAPM Time Series Regression Results

Time Series Regression	Beta	1.4228695	1.2453349	1.1546382	1.0725501	1.1010283	1.4270472	1.2330666
	SE Beta	0.0507483	0.0440197	0.0399108	0.0395659	0.0448566	0.0383339	0.0326002
	t of Beta	28.037783	28.290372	28.9305	27.107933	24.545526	37.226802	37.823928
	Alpha	-0.321399	0.1751476	0.2785508	0.4649415	0.6255859	-0.230348	0.1325658
	SE Alpha	0.2290186	0.1986538	0.1801106	0.1785544	0.2024303	0.1729944	0.1471192
	t of Alpha	-1.403375	0.8816725	1.5465539	2.6039212	3.0903774	-1.331532	0.9010779

Time Series Regression	Beta	1.1139423	1.0308712	1.1255602	1.3553795	1.1620848	1.0329253	0.9743327	1.0675926	1.2237926
	SE Beta	0.0308318	0.0292908	0.0367347	0.0298628	0.0251575	0.0244711	0.0246247	0.0342901	0.0223022
	t of Beta	36.1297	35.194406	30.640224	45.386931	46.192391	42.209951	39.567286	31.134145	54.873265
	Alpha	0.4146401	0.5188814	0.6118569	-0.143105	0.214629	0.28337	0.475195	0.5716176	-0.063471
	SE Alpha	0.1391386	0.1321844	0.1657777	0.1347657	0.1135316	0.1104342	0.1111272	0.1547455	0.100646
	t of Alpha	2.9800505	3.9254367	3.6908275	-1.061876	1.8904781	2.5659635	4.2761363	3.6939219	-0.630633

Time Series Regression	Beta	1.1276595	1.0349675	0.9755001	1.0789706	1.0044612	0.981674	0.8646886	0.8361395	0.8688941
	SE Beta	0.0195673	0.0204947	0.0244539	0.0318738	0.0201565	0.0166115	0.0215918	0.0226989	0.0324279
	t of Beta	57.629847	50.499286	39.891346	33.85131	49.833106	59.095952	40.046999	36.836118	26.794644
	Alpha	-0.053479	0.2013275	0.4218721	0.4928321	-0.097902	-0.027886	0.0144032	0.1960591	0.2714099
	SE Alpha	0.0883039	0.0924892	0.1103565	0.1438413	0.090963	0.0749651	0.0974404	0.1024364	0.1463417
	t of Alpha	-0.605622	2.1767687	3.8228111	3.4262221	-1.07628	-0.371985	0.1478157	1.9139593	1.854631

Exhibit 2: Time Series Regression Equation

$$R_t^{ei} = a_i + \beta_i f_t + \varepsilon_{it} \quad t = 1 \ldots T,\ \forall i$$

Exhibit 3: Cross-Sectional Regression Equation

$$E(R_t^{ei}) = (\gamma) + \beta_i \lambda_t + \alpha_i \quad i = 1 \ldots N,\ \forall t$$

Exhibit 4: Fama Macbeth Procedure Regression Equation

$$R_t^{ei} = \beta_i \lambda_t + \alpha_{it} \quad i = 1 \ldots N,\ \forall t$$

Exhibit 5: GRS F-statistics and critical value calculations

$$f_{GRS} = \left[\frac{T-N-K}{N}\right]\left[\frac{\alpha' \Sigma_{\varepsilon}^{-1} \alpha}{\mu_f \Sigma_f^{-1} \mu_f}\right] \sim F(N, T-N-K)$$

Critical value = F.INV.RT(0.05,N,T-N-K)

Exhibit 6: FF3F Time Series Regression Results

Time Series Regression	b	1.030324	0.966742	0.941189	0.895363	0.953588	1.094237	1.023533	0.964902	0.964755	1.068907	1.098788	1.024534	0.973393
	SE b	0.025892	0.019444	0.015784	0.014909	0.016302	0.020067	0.016699	0.015956	0.015133	0.016538	0.018417	0.017505	0.017343
	t of b	39.7934	49.7194	59.62746	60.05598	58.49501	54.52788	61.29154	60.47123	63.75127	64.63383	59.66055	58.52674	56.12639
	s	1.427296	1.282514	1.154731	1.116126	1.19992	1.010462	0.924105	0.841258	0.69927	0.852803	0.703787	0.63181	0.532179
	SE s	0.037924	0.02848	0.02312	0.021837	0.023878	0.029393	0.02446	0.023371	0.022166	0.024223	0.026976	0.02564	0.025402
	t of s	37.63583	45.03269	49.94604	51.11175	50.25287	34.37777	37.78077	35.99524	31.54767	35.20622	26.08945	24.64136	20.95015
	h	-0.291606	0.096255	0.267346	0.399173	0.637196	-0.471797	0.026844	0.226012	0.467056	0.685498	-0.44841	0.045936	0.310918
	SE h	0.042023	0.031558	0.025619	0.024197	0.026459	0.03257	0.027104	0.025898	0.024562	0.026842	0.029892	0.028412	0.028148
	t of h	-6.939149	3.050078	10.43557	16.4965	24.08264	-14.48557	0.990433	8.727115	19.01578	25.53872	-15.00106	1.616778	11.04583
	Alpha	-0.410215	-0.1011	-0.069238	0.051601	0.06714	-0.147736	-0.043174	0.144136	0.140957	0.087411	-0.019879	0.079353	0.020168
	SE Alpha	0.104763	0.078674	0.063867	0.060324	0.065961	0.081196	0.067569	0.064562	0.061231	0.066915	0.07452	0.07083	0.070172
	t of Alpha	-3.915663	-1.285052	-1.084095	0.855403	1.017882	-1.819494	-0.638959	2.232508	2.302048	1.306304	-0.266766	1.120338	0.287414

Time Series Regression	b	0.972156	1.065102	1.061578	1.073257	1.046238	1.0349	1.147252	0.957252	1.026705	0.974119	0.994748	1.040928
	SE b	0.016276	0.020554	0.018322	0.020176	0.019552	0.020102	0.024359	0.016326	0.017676	0.021624	0.017889	0.026766
	t of b	59.72833	51.81898	57.94	53.1946	53.50995	51.48226	47.09702	58.63212	58.08317	45.04874	55.60813	38.89008
	s	0.444889	0.64474	0.296996	0.269401	0.223695	0.222944	0.351628	-0.193746	-0.197818	-0.283069	-0.196077	-0.034377
	SE s	0.02384	0.030106	0.026836	0.029552	0.028638	0.029444	0.035679	0.023913	0.025891	0.031672	0.026201	0.039204
	t of s	18.66152	21.41572	11.06694	9.116204	7.811079	7.571917	9.855258	-8.101996	-7.640492	-8.937459	-7.483477	-0.876862
	h	0.489923	0.713214	-0.449887	0.04012	0.3061	0.537856	0.725537	-0.446099	-0.00489	0.210448	0.545977	0.792754
	SE h	0.026417	0.03336	0.029737	0.032746	0.031734	0.032626	0.039536	0.026498	0.028689	0.035096	0.029034	0.043442
	t of h	18.54581	21.37915	-15.12876	1.225162	9.645832	16.48537	18.3513	-16.83502	-0.170457	5.996364	18.80497	18.24862
	Alpha	0.129057	0.06823	0.131446	-0.122418	-0.00548	0.088126	0.033762	0.180447	0.009264	-0.051659	-0.069127	-0.157256
	SE Alpha	0.065857	0.083166	0.074134	0.081636	0.079112	0.081336	0.098562	0.066059	0.071522	0.087493	0.07238	0.108299
	t of Alpha	1.959671	0.820408	1.773085	-1.499561	-0.069272	1.083482	0.342541	2.731591	0.129531	-0.590432	-0.95505	-1.452048

Exhibit 7: ChatGPT Prompts and Output

- Task: Input validation
  - Prompt and output
- Task: Writing dataframes to Excel
  - Prompt and output
- Task: Skipping rows while reading in data file
  - Prompt and output
- Task: Converting to series from dictionaries and lists
  - Prompt and output
- Task: Accessing the correct parameters after regression
  - Prompt and output
- Task: Matrix multiplication syntaxes
  - Prompt and output
- Used ChatGPT for basic brush-ups on object-oriented programming syntaxes