Project 2: Executive Memo

Aman Sharma: AS235548, Tom Starkie: TS43286, Lindsay Bartol: LLB2686, Biagio Alessandrello: BJA2372

Methodology

Time Series Regression

For each combination of time period and model, we started by running a time series regressions. The goal of this procedure is to estimate the beta(s) for every portfolio or stock. To do this, we ran regressions of each portfolio's excess returns against the factor premium(s). For CAPM, this regression line would represent the Security Characteristic Line on a graph. For both models, we also estimate alpha, which is the intercept. Finally, we calculate the standard errors, which measure the uncertainty of our estimates, and the t-statistics to determine if these values are statistically different from zero for each portfolio. Here, our null hypothesis is that the alphas are not statistically significantly different from zero, for this would mean our model is inaccurate in pricing the individual portfolios.

In this procedure, our main assumptions are that the regression models are valid and that we have linear relationships, homoskedasticity, and no heterogeneity, meaning that the x variables (the factors) are not correlated with the residuals.

Cross-Sectional Regression

Next, we ran cross-sectional regressions to estimate lambda for each factor and the standard error for each lambda. Our null hypothesis here is that the lambdas are statistically significantly different from zero. In other words, we are assessing if we are rewarded for the risk we are taking on. To do so, we first averaged the returns of each portfolio to find the expected returns. We then ran a regression of these returns against the beta estimations from the time series regression.

While this method gives an estimation of lambdas and their standard error, the standard error is typically underestimated because we assume that the residuals are independent. In reality, there is usually correlation across portfolios due to factors like common industries and market-wide influences. Since residual variation is shared among assets, we have less independent information than assumed, resulting in artificially low standard errors. To obtain a more accurate estimate, we need to run a Fama-MacBeth Cross-Sectional Procedure to better estimate the standard error and assess if the lambdas are statistically different from zero.

Fama MacBeth CS Procedure

Next, we applied the Fama-MacBeth Procedure to more accurately estimate the lambdas and their standard errors, allowing us to judge whether the lambdas are statistically significant. Again, our null hypothesis here is that the lambdas are statistically significantly different from zero. First, we ran a series of cross-sectional regressions of the returns against the betas for each time t to estimate lambda for each factor. We then averaged these lambda estimates over time to find the final lambdas, which represent the factor risk premiums. To calculate the standard error, we took the standard deviation of the lambda estimates and divided it by the square root of the number of time periods.

This approach more accurately captures the standard error by incorporating the variability in lambda estimates over time, rather than relying on a single cross-sectional estimate. This better accounts for cross-sectional correlation because running the regression in each period allows us to capture correlations between assets at that time, reducing the underestimation of the standard error. Ultimately, this procedure provides a more robust standard error estimate and allows us to better assess the statistical significance of the risk premium, which is why we go the extra mile to execute it. However, it should be noted that we are assuming independence of returns from one period to the next.

Time Series Joint Hypothesis Test
Lastly, we executed a time series joint hypothesis test to test the null hypothesis that the alphas of the model are not jointly statistically significantly different from zero. If the alphas are not jointly zero, this indicates that our model is not accurately pricing assets. First, we calculated the residuals of our time series regression model, then de-meaned these residuals and calculated the residual covariance matrix. Next, we computed the f-GRS statistic and compared it to the F critical value. If the f-GRS statistic is greater than the F critical value, we must reject the null hypothesis. Rejecting the null hypothesis means our model is insufficient to explain the returns of the assets, indicating that there is remaining systematic risk that we are not accounting for.

CAPM Across Time

Cross-Sectional Regression (CSR):

Time Period	Lambda - B	S.E. Lambda - b	t-Stat Lambda - b
07/63 - 06/93	-0.37	0.33	-1.13
09/94 - 08/24	-0.44	0.23	-1.91

After running cross-sectional regressions for the CAPM in both time periods, we see that our lambda estimate for the market risk factor is negative. This does not align with the theory of the model, as this would mean that we are getting less return for the risk we are taking on. However, even with the deflated estimates for lambda’s standard error, the t-tests show that these estimates are not statistically significant on a 95% confidence interval.

Fama-MacBeth Procedure Regression (FMP):

Time Period	Lambda - B	S.E. Lambda - b	t-Stat Lambda - b
07/63 - 06/93	-0.37	0.46	-0.80
09/94 - 08/24	-0.44	0.64	-0.69

Next, we executed the Fama-MacBeth procedure for the CAPM in both time periods in order to get a better estimate of the standard error. For both time periods, the standard error increased, showing that we had initially underestimated it, likely due to cross-sectional correlation. Again, this further solidifies that these lambda estimates for the market risk factor are not statistically significantly different from zero. While this means that we are not necessarily receiving negative returns for the risk we are taking on, we are still apparently not receiving any extra returns. Ultimately, this does align with the theory of CAPM

Gibbons-Ross-Shanken F-Test (GRS F-Test):

Time Period	f-GRS Statistic	Critical Value	Reject or Fail to Reject
07/63 - 06/93	2.40	1.54	Reject the Null Hypothesis
09/94 - 08/24	3.86	1.54	Reject the Null Hypothesis

Lastly, to test if the alphas are jointly statistically significantly different from zero, we ran a GRS f-test in both time periods. In both cases, the f-GRS Statistic was larger than the critical value, and so we rejected the null hypothesis that the alphas are jointly statistically significantly different from zero. This indicates that the CAPM is not accurately pricing these portfolios in either time period and that there are remaining systematic risk factors or anomalies that are unaccounted for by the CAPM.

FF3F Across Time

Cross-Sectional Regression (CSR):

Time Period	Lambda - b	S.E. Lambda-b	t-Stat Lambda-b	Lambda - s	S.E. Lambda- s	t-Stat Lambda- s	Lambda - h	S.E. Lambda- h	t-Stat Lambda- h
07/63 - 06/93	-0.10	0.45	-0.22	0.20	0.05	3.84	0.51	0.07	7.87
09/94 - 08/24	-0.78	0.32	-2.44	0.03	0.06	0.55	0.13	0.07	1.95

Based on cross-sectional analysis of the Fama-French Three Factor Model in the two 30-year time periods, we can make a few interesting observations. From 1963 to 1993, both the s (SML) and h (HML) t-statistics show significance in a 95% confidence interval, but the b (RMRF) does not. This means that we are receiving factor premiums for size and value, but not for market risk. Inversely, in 1994 through 2024, only the lambda b is statistically significant while the lambda s and lambda h are not, meaning that we are not receiving increasing returns for these risk factors. Additionally, it's important to note that the lambda b in both time periods is negative, indicating that we are receiving negative returns for taking on more risk. Next, we executed the Fama-MacBeth procedure in order to more accurately estimate the standard errors and better evaluate whether or not these figures are significant.

Fama-MacBeth Procedure Regression (FMP):

Time Period	Lambda - b	S.E. Lambda - b	t-Stat Lambda - b	Lambda - s	S.E. Lambda - s	t-Stat Lambda - s	Lambda - h	S.E. Lambda - h	t-Stat Lambda - h
07/63 - 06/93	-0.10	0.43	-0.22	0.20	0.15	1.30	0.51	0.14	3.68
09/94 - 08/24	-0.78	0.42	-1.84	0.03	0.18	0.17	0.13	0.18	0.73

Using the FMP model to calculate the metrics, we see that almost all of the standard error estimates increase for the lambdas, with lambda b in the 1963-1993 period being the only exception. As a result of these updated values, only lambda h in the time period 1963 to 1993 is statistically significantly different from zero, showing that we are only receiving higher returns for value risk, but not for the other two types of risk. From, 1994 to 2024, none of the lambdas prove to be statistically significant, showing that we are not receiving higher returns for taking on more risk in relation to any of these factors.

Gibbons-Ross-Shanken F-Test (GRS F-Test):

Time Period	f-GRS Statistic	Critical Value	Reject or Fail to Reject
07/63 - 06/93	1.62	1.54	Reject the Null
09/94 - 08/24	3.85	1.54	Reject the Null

After running a GRS f-test, we can see that the f-GRS statistic is larger than the critical value in both time periods. As a result, we can conclude that in both periods, the alphas are jointly statistically significantly different from zero. Ultimately, there are underlying risk factors or anomalies that we have not been able to capture with the Fama-French Three Factor Model.

However, when we run the model around the time Fama and French released their paper, the model can partially predict returns. We can see its higher accuracy by looking at how close te f-GRS statistic is to the critical value. In the end, there are still significant alphas, but they are less apparent than in the later period. In the last 30 years, the FF3F model's predictability power has diminished, leaving even more unexplained risk on the table.

Conclusions & Insights

To conclude our analysis as mentioned above, both models, CAPM and Fama French 3 Factor, show a rejection of the null hypothesis that the alphas are jointly zero, for both time periods (1963-1993 & 1994-2024). For CAPM, the lambdas representing MRP were found to be insignificant in both time periods with both CSR and FMP tests. For FF3F, on the CSR and FMP tests some of the metrics were significant which means that they are able to explain some of the risk of the portfolios. In CSR, the lambdas for the size premium and the value premium were significant while the market risk premium was not in 1963-1993, but in 1994-1924, only market risk premium was significant while size premium and value were not. In FMP, only the lambda for value premium is significant in 1963-1993 while all the other metrics were not significant for both time periods. Ultimately, for the period 1963 to 1993, the Fama French Three Factor Model has the model predictive power. However, for the time period 1994 to 2024, both models are about equally as weak.

We believe that the Fama French Model was not able to explain the returns 1994-2024 time period mostly due to the fact that the factors have been priced into the assets since the release of the initial paper, thus decreasing the excess returns that they can predict. The inclusion of additional factors like momentum, volatility, sales growth and others published later on, have successfully increased the performance of the models and helped at capturing more of that risk that is unexplained. Going forward, we would like to try models that include these factors as well. Additionally, we hypothesize that changes in the markets, such as stock market crashes like the 2008 Real Estate Crash and Covid in 2020, could have significantly impacted the model's performance in the recent time period. This change in patterns is quite visible from the figure F1, and F2 in the Appendix. The increase in returns across value and size is very obvious in the time period 1963 to 1993, but it is not distinguishable 1994 to 2024. Ultimately, these two models that we have evaluated do not truly explain returns in the present day. Going forward, we expect that risk factors and anomalies will continue to evolve as we become more and more globalized, causing continuous shifts throughout time.

Appendix

Figures

F.1. Value & Size Total Excess Returns 1963-1993

		Value Quintiles				
	E (r) - r_f	LoBM	BM2	BM3	BM4	HiBM
Size Quintiles	SMALL	27%	69%	76%	91%	108%
	ME2	36%	64%	88%	95%	108%
	ME3	42%	70%	71%	88%	101%
	ME4	44%	41%	63%	83%	94%
	BIG	32%	38%	37%	54%	63%

F.2. Value & Size Total Excess Returns 1994-2024

		Value Quintiles				
	E (r) - r_f	LoBM	BM2	BM3	BM4	HiBM
Size Quintiles	SMALL	33%	85%	80%	103%	83%
	ME2	85%	86%	85%	86%	75%
	ME3	79%	83%	80%	89%	97%
	ME4	87%	90%	81%	86%	84%
	BIG	90%	76%	81%	55%	76%

"We attest that this assignment has been completed in accordance with the academic integrity protocols and honor codes of the McCombs School of Business and the University of Texas. We relied on no unauthorized resources (including material submitted or produced by other students—past or present, individuals not listed in this document, internet message boards, solution manuals, or other forms of outside assistance). Any assistance from AI tools (e.g., ChatGPT) has been appropriately cited. This work product is the sole result of our efforts and is being submitted as such. Aman Sharma: AS235548, Tom Starkie: TS43286, Lindsay Bartol: LLB2686, Biagio Alessandrello: BJA2372."