Assessment of Stock Performance: Analyzing S&P 100 Companies Using Predictive Regression, Fama-French, and Fama-Macbeth Methods

Advised by Professor Raša Karapandža

Spring 2024

Contents

List of Figures

List of Tables

Assessment of Stock Performance: Analyzing S&P 100 Companies Using Predictive Regression, Fama-French, and Fama-Macbeth Methods

Mai Luo

Spring 2024

Abstract

This investment memo leverages classical and contemporary financial metrics to predict stock performance within the S&P 100 from 2006 to 2016. Specifically, the Debt-to-Equity (D/E) and Price-to-Book (P/B) ratios are employed in the Predictive Regression framework to reflect traditional financial health and valuation perspectives, historically valued for their ability to signal underlying risk and asset valuation discrepancies in equities. In contrast, the Fama-French and Fama-Macbeth models utilize an equal-weighted average of Cash Ratio and Return on Equity (ROE), chosen for their relevance in depicting liquidity and profitability, crucial in assessing operational efficiency and earnings quality in modern financial analysis. This memo aims to validate the selected ratios' predictive power and their relevance in different modeling contexts, enhancing our understanding of financial metrics as robust indicators of future stock performance.

1 Introduction

In the realm of investment strategy, the ability to discern between potential high and low performers is invaluable. This investment memo presents a rigorous analysis of how financial ratios, within different statistical models, predict stock performance over a decade—from July 2006 to June 2016—within the S&P 100. Our investigation is motivated by the historically proven utility of certain financial ratios in forecasting market behavior and firm performance, which may provide a strategic edge in portfolio management.

Our study initiates with Predictive Regression, utilizing the Debt-to-Equity (D/E) and Price-to-Book (P/B) ratios. These ratios are traditionally recognized for their predictive relevance in assessing financial risk and market valuation, respectively. The D/E ratio evaluates a company's financial leverage and associated risk, while the P/B ratio helps identify potentially undervalued stocks that might yield high returns.

In the latter sections, we apply the Fama-French and Fama-Macbeth models using an equal-weighted average of Cash Ratio and Return on Equity (ROE). This choice is predicated on the premise that liquidity (Cash Ratio) and profitability (ROE) are critical to determining a firm's financial health and operational efficiency (to examine if they are indeed the quality stocks that we are looking for). These metrics are especially pertinent in volatile markets, where they may serve as reliable indicators of stability and growth potential.

The integration of these methodologies aims to derive a nuanced understanding of financial indicators as tools for robust investment analysis. By evaluating how these ratios perform across different models, the memo seeks to provide a comprehensive perspective on the viability of these metrics as indicators of future stock performance, thus guiding strategic investment decisions. Through this approach, we expect to uncover insights that validate the effectiveness of combining traditional and modern financial analyses to enhance investment strategies.

2 Data and Methodology

2.1 Data Collection and Selection:

The primary dataset for our analysis is sourced from the WRDS CRSP database. This dataset encompasses monthly returns, identified as 'ret', and delisting returns, denoted as 'dlret', covering the period from July 2006 to July 2016. To accurately compute the final return for each data point, we adopt the following methodological approach:

- Utilize 'ret' when 'dlret' is not available and 'ret' is present.
- Employ 'dlret' exclusively when 'ret' is absent and 'dlret' is accessible.
- Combine both returns using the formula (1+ret)×(1+dlret)-1 when both are available, to account for cumulative effects in periods involving delisting.

For the Predictive Regression model, this dataset includes the Debt-to-Equity (D/E) ratio and Price-to-Book (P/B) ratio, spanning from July 2001 to July 2016. These extended dates account for the requirement of historical lag data to analyze predictive impacts accurately. Additionally, the dataset includes Cash Ratio and Return on

Equity (ROE) for the period from 2005 to 2015. This timeframe allows for the annual reselection of companies based on ratios published one year prior, critical for accurate historical analysis. Data integration for these two financial ratios is handled as follows:

- Use ROE if Cash Ratio is empty and ROE is available
- Use Cash Ratio if ROE is not available and Cash Ratio is available
- Use (0.5*Cash Ratio)+(0.5*ROE) if both ret and dlret are available

This method averages the ratios to mitigate the individual volatility of each ratio and provides a balanced metric that reflects both liquidity and profitability.

The third dataset incorporates the Fama-French three factors, essential for the Fama-French regression analysis. This dataset includes the monthly risk-free rate and the three factors: Market Premium, High Minus Low (HML), and Small Minus Big (SMB) from July 2006 to July 2016.

2.2 Methodology

In our investment analysis, we employ three distinct regression models to analyze stock returns, assess predictive factors, and gauge market dynamics.

1. **Model 1: Predictive Regression**

   The Predictive Regression model evaluates the predictive power of selected variables such as the Debt-to-Equity (D/E) ratio and the Price-to-Book (P/B) ratio on future stock returns. The aim is to determine if these financial metrics can serve as reliable indicators for future performance.

   $$R_t - {R_f}_t = \alpha + \beta_1 \times (D/E)_{t-k} + \beta_2 \times (P/B)_{t-k} + \varepsilon_t \tag{1}$$

   Here, $R_t - {R_f}_t$ represents the excess return of a stock over the risk-free rate at time t, α is the intercept, β_1 andβ_2 are the coefficients for the Debt-to-Equity and Price-to-Book ratios respectively, lagged by k periods, and ε_t is the error term. We will analyze the statistical significance of the coefficients to assess predictive validity.

2. **Model 2: Fama French Regression**

   The Capital Asset Pricing Model (CAPM) is a single-factor model that considers only the market risk factor to predict stock returns. It is defined as:

   $$R_{it} - {R_f}_t = \alpha + \beta_{i1} \times EXM_t + \varepsilon_t \tag{2}$$

   The Fama-French model extends the CAPM by including size and value factors in addition to the market risk factor:

   $$R_{it} - {R_f}_t = \alpha_i + \beta_{i1} \times EXM_t + \beta_{i2} \times HML_t + \beta_{i3} \times SMB_t + \varepsilon_{it} \tag{3}$$

   By incorporating these factors, we expect to better explain the anomalies that the CAPM cannot, such as why smaller companies or those with high book-to-market ratios might yield higher returns and to enhance the predictive accuracy of our stock return forecasts.

3. **Model 3: Fama Macbeth Regression**

   The Fama-MacBeth regression employs a two-step procedure to analyze the relationship between excess returns and predictive variables across time:

   $$R^i{}_t - {R_f}_t = \gamma_i + \beta_i \times EXM_t + \varepsilon_{it} \tag{4}$$

   $$R^i{}_t - {R_f}_t = \lambda_0^t + \lambda_1^t \times \beta_i + \lambda_2^t \times ROE_i^t + \lambda_3^t \times Cash_Ratio_i^t + \varepsilon_{it} \tag{5}$$

In the first step, we estimate the coefficients λ^t at each time t using a cross-sectional regression of returns on the betas obtained from previous models and financial ratios such as Balanced ROE and Cash Ratio. The second step involves averaging these coefficients over time to understand their evolution and stability. This model allows us to assess how consistently different factors are priced in the market over time.

By analyzing cross-sectional regressions over multiple periods, we aim to observe the stability of factor loadings and the robustness of financial indicators in predicting returns. We expect this model to reveal which factors consistently contribute to pricing assets in the market, providing crucial insights for long-term investment strategies and risk management. The Fama-MacBeth approach is particularly valuable in testing the temporal validity of our predictive factors and in avoiding the pitfalls of overfitting that can occur in more static models.

2.3 Dataset Checks

2.3.1 Model 1: Predicitve Regression

For the integrity of the analysis, it is imperative that the datasets used for the predictive regression model maintain consistency in terms of company coverage. During our preliminary data checks, we identified discrepancies between the Monthly Returns Dataset and the Financial Ratios Dataset. Specifically, three companies present in the Monthly Returns Dataset were absent from the Financial Ratios Dataset. To maintain dataset integrity, these companies were excluded from the Monthly Returns Dataset to ensure consistency across the data used for analysis.

Following the adjustments for data consistency, the distribution of the remaining data in the Monthly Returns Dataset is as follows:

No.Periods	No.Companies
120	78
35	2
11	2
4	2
70	1
69	1
68	1
63	1
97	1
52	1
44	1
84	1
31	1
21	1
14	1
88	1
59	1

Table 1: Frequency of periods count

Approximately 80% of the companies have complete data for all 120 months, indicating robust coverage for a substantial portion of the dataset. Only 2 companies have only 4 months data points available and another 2 with only 11 months. Most of the other companies that do not have the total 120 months seem to demonstrate good amount of data points in the dataset. This should suffice to be able to proceed.

Given the construction of an equal-weight portfolio of 100 companies, it is crucial to address how missing data impacts the representation of portfolio returns. To provide a transparent view of company availability throughout the study period, we generated monthly distribution tables:

Date	No.Companies available
2006-07	90
2006-08	90
2006-09	89
2006-10	89
2006-11	89
2006-12	87

Table 2: 2006 Monthly count of available companies

Date	No.Companies available
2016-01	86
2016-02	86
2016-03	86
2016-04	86
2016-05	86
2016-06	86
2016-07	86

Table 3: 2016 Monthly count of available companies

Date	No. Companies available	Date	No. Companies available	Date	No. Companies available
2007-01	87	2008-01	86	2009-01	86
2007-02	87	2008-02	86	2009-02	86
2007-03	87	2008-03	86	2009-03	86
2007-04	87	2008-04	86	2009-04	86
2007-05	87	2008-05	86	2009-05	86
2007-06	87	2008-06	86	2009-06	86
2007-07	86	2008-07	86	2009-07	85
2007-08	86	2008-08	86	2009-08	85
2007-09	86	2008-09	86	2009-09	85
2007-10	86	2008-10	86	2009-10	85
2007-11	86	2008-11	86	2009-11	85
2007-12	86	2008-12	86	2009-12	85

Table 4: 2007 Monthly count

Table 5: 2008 Monthly count

Table 6: 2009 Monthly count

Date	No. Companies available	Date	No. Companies available	Date	No. Companies available
2010-01	85	2011-01	85	2012-01	85
2010-02	85	2011-02	85	2012-02	84
2010-03	85	2011-03	85	2012-03	85
2010-04	84	2011-04	86	2012-04	85
2010-05	84	2011-05	86	2012-05	85
2010-06	84	2011-06	86	2012-06	84
2010-07	84	2011-07	85	2012-07	84
2010-08	84	2011-08	85	2012-08	84
2010-09	84	2011-09	85	2012-09	84
2010-10	84	2011-10	85	2012-10	84
2010-11	84	2011-11	85	2012-11	84
2010-12	85	2011-12	85	2012-12	84

Table 7: 2010 Monthly count

Table 8: 2011 Monthly count

Table 9: 2012 Monthly count

Date	No. Companies available	Date	No. Companies available	Date	No. Companies available
2013-01	84	2014-01	84	2015-01	84
2013-02	84	2014-02	84	2015-02	84
2013-03	84	2014-03	84	2015-03	84
2013-04	84	2014-04	84	2015-04	84
2013-05	84	2014-05	84	2015-05	84
2013-06	84	2014-06	84	2015-06	84
2013-07	84	2014-07	84	2015-07	85
2013-08	84	2014-08	84	2015-08	86
2013-09	84	2014-09	83	2015-09	86
2013-10	84	2014-10	84	2015-10	86
2013-11	84	2014-11	84	2015-11	86
2013-12	84	2014-12	84	2015-12	86

Table 10: 2013 Monthly count

Table 11: 2014 Monthly count

Table 12: 2015 Monthly count

As Table 2 - 10 suggest, none of the 120 months has fully 100 companies data entries. Thus Given the variability in the number of available companies across different months, applying a fixed equal weight (1/100) to all companies regardless of their data availability may introduce bias, particularly if missing data is related to factors such as delisting. Instead, we adjust the weighting scheme based on the actual number of companies available each month, thus mitigating the distortion in the portfolio's performance representation that would be caused by treating missing data as zero return.

2.3.2 Model 2: Fama French

The dataset for the Fama-French model analysis incorporates financial ratios critical for determining company performance, such as Cash Ratio and Return on Equity (ROE). Initial data integrity checks reveal that the dataset contains records for 97 unique PERMNOs, which implies some companies are missing from the original set of 100 companies intended for analysis.

For the purpose of our analysis, the absence of some companies in the ranking ratio dataset does not significantly impede our ability to conduct a comprehensive study. Companies missing from the dataset will naturally be excluded from the portfolio analysis. This exclusion is methodologically sound as our focus is on companies for which complete and consistent data is available, ensuring the reliability of our comparisons and conclusions.

The ranking ratios are extracted annually at the end of December to coincide with fiscal year-end reports for most companies, providing a standardized reference point for annual performance analysis. This approach aligns with common academic practices for ensuring data relevance and temporal consistency in financial analysis. Additionally, the data is updated quarterly, but the choice of December data ensures we capture end-of-year financial standings, which are often used in annual reporting and analysis.

Further analysis identified 21 companies with incomplete data records throughout the dataset period. Table 13 lists these companies, noting the gaps which prevent them from being included in some years' analyses:

Company ID	Note
12079	Incomplete Date Range
12369	Incomplete Date Range
12622	Incomplete Date Range
13267	Incomplete Date Range
14040	Incomplete Date Range
14276	Incomplete Date Range
14461	Incomplete Date Range
14882	Incomplete Date Range
15408	Incomplete Date Range
20220	Incomplete Date Range
22840	Incomplete Date Range
39087	Incomplete Date Range
39917	Incomplete Date Range
42024	Incomplete Date Range
47255	Incomplete Date Range
60097	Incomplete Date Range
76171	Incomplete Date Range
76656	Incomplete Date Range
77481	Incomplete Date Range
82621	Incomplete Date Range
83332	Incomplete Date Range

Table 13: List of companies with incomplete date ranges

Post data sorting and ranking, portfolios are constructed based on the financial health indicators of the available companies. Tables showing an example of top and bottom portfolio companies illustrate how rankings are derived based on composite scores from ROE and cash ratios. The tables provide a snapshot of the data used to make empirical assessments of company performance within the market.

PERMNO	roe	cash_ratio	TICKER	year-month	final_ratio	rank
82621	-0.498	4.793	MS	2005-12-01	2.1475	1
15579	0.165	2.505	TXN	2005-12-01	1.335	2
10107	0.235	2.303	MSFT	2005-12-01	1.269	3
59328	0.215	1.833	INTC	2005-12-01	1.024	4
83332		0.735	LU	2005-12-01	0.735	5
40416	0.98	0.488	AVP	2005-12-01	0.734	6
14461	0.369	1.031	SNSTA	2005-12-01	0.7	7
22111	0.274	1.085	JNJ	2005-12-01	0.6795	8
10147	0.111	1.129	EMC	2005-12-01	0.62	9
14008	0.17	1.046	AMGN	2005-12-01	0.608	10
22752	0.261	0.947	MRK	2005-12-01	0.604	11
10104	0.295	0.908	ORCL	2005-12-01	0.6015	12
43123	0.568	0.548	ATI	2005-12-01	0.558	13
60097	0.175	0.94	MDT	2005-12-01	0.5575	14
76076	0.235	0.845	CSCO	2005-12-01	0.54	15
76656	0.033	1.005	MEDI	2005-12-01	0.519	16
11308	0.318	0.606	KO	2005-12-01	0.462	17
87447	0.189	0.734	UPS	2005-12-01	0.4615	18

Table 14: E.g. Top Portfolio Companies

PERMNO	roe	cash_ratio	TICKER	year-month	final_ratio	rank
10890	-1.154	0.268	UIS	2005-12	-0.443	90
12079	-0.122		GM	2005-12	-0.122	89
76226	-0.347	0.107	VIA	2005-12	-0.12	88
66093	0.068	0.029	SBC	2005-12	0.0485	87
47896	0.069		JPM	2005-12	0.069	86
39917	0.074		WY	2005-12	0.074	85
25785	0.076		F	2005-12	0.076	84
79323	0.084		ALL	2005-12	0.084	83
24643	0.099	0.07	AA	2005-12	0.0845	82
77481	-0.142	0.314	EP	2005-12	0.086	81
24942	0.08	0.097	RTN	2005-12	0.0885	80
18411	0.103	0.084	SO	2005-12	0.0935	79
89525	0.017	0.194	CMCSA	2005-12	0.1055	78
38156	0.047	0.174	WMB	2005-12	0.1105	77
21776	0.146	0.078	EXC	2005-12	0.112	76
47255	0.214	0.022	LB	2005-12	0.118	75
65875	0.121	0.124	VZ	2005-12	0.1225	74
40125	0.081	0.166	CSC	2005-12	0.1235	73

Table 15: E.g. Bottom Portfolio Companies

2.3.3 Model 3: Fama Macbeth

The Fama MacBeth regression utilizes a comprehensive dataset comprising financial ratios, market premium factors from the Fama-French three factors, and monthly returns data. This approach leverages data already employed in earlier models, ensuring consistency across analyses and maximizing the informational value of the data collected.

PERMNO	γ_i	β_i	R^2
10104	0.003813	1.076694	0.494234
10107	0.004652	0.985897	0.377019
10145	0.004121	1.244979	0.714822
10147	0.003444	1.096644	0.396664
10890	-0.016379	3.282059	0.464214
11308	0.005122	0.522602	0.266978
11703	0.000614	1.465309	0.626291
11850	0.001992	0.535697	0.260819
12052	0.002815	1.107359	0.557151
12060	-0.004111	1.431529	0.599233

Table 16: Snapshot of 100 Companies' Betas to Excess Return Explanation

3 Results and discussions

3.1 Predicative Regressions

$$R_t - R_{f_t} = \alpha + \beta_1 \times (D/E)_{t-k} + \beta_2 \times (P/B)_{t-k} + \varepsilon_t \tag{6}$$

Horizon	D/E Ratio	P/B ratio	t(D/E)	t(P/B ratio)	R^2	$\sigma[E_t(R^e)]$	$\frac{\sigma[E_t(R^e)]}{E(R^e)}$
1mon	-0.000570	-0.996514	-0.053123	-1.290784	0.014762	0.647108	0.769745
3mon	-0.014698	-1.056721	-1.387642	-1.381086	0.038976	1.051491	1.250765
6mon	0.001856	-1.546009*	0.174681	-2.012969	0.033633	0.976772	1.161885
1yr	-0.032440**	-0.672362	-3.117293	-0.789067	0.094512	1.637387	1.94697
2yr	0.011447	-0.805573	1.035574	-0.753779	0.010786	0.553149	0.657980
3yr	0.006483	-0.487018	0.590395	-0.485919	0.003915	0.333238	0.396392
4yr	-0.001246	-1.356643	-0.114386	-1.372475	0.017776	0.710105	0.844681
5yr	0.002879	-0.008883	0.265582	-0.009522	0.000618	0.132420	0.157516

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table 17: Predicative Regression Results

The predictive regression analysis underscores the challenges in using D/E and P/B ratios as standalone predictors of stock returns. While there are glimpses of predictive power at specific intervals (notably at 6 months for P/B and 1 year for D/E), the overall effectiveness across various horizons is limited.

Negative coefficients at shorter horizons (like at 1-3 months) might indicate that higher leverage is perceived as risky by investors (or shareholders). High leverage can signal higher default risk as the company is embarking on more projects, leading to lower stock prices, possibly serving as a mean of transferring the wealth of shareholders to the creditors.

Positive coefficients at certain periods (like at 2 years and 3 years) could reflect a different market condition where leverage is seen as beneficial from relatively a long-term perspective as it might indicate the company's expansions to facilitate growth and thus potentially increasing expected returns. Investors might view leveraged growth opportunities positively at a relatively long-run stance, leading to higher stock prices for leveraged companies if they managed to sustain the debt services.

A negative coefficient generally indicates that higher P/B ratios (which might suggest overvaluation relative to book values) are associated with lower future returns. This consistent pattern suggests that the market tends to correct overvalued stocks by adjusting their prices downward, aligning more closely with their fundamental book values over time.

This observation aligns with value investing principles, where stocks with lower P/B ratios are considered undervalued and thus more likely to offer better returns. The negative relationship indicates that investors might do well to focus on companies with lower P/B ratios, expecting them to yield higher returns as their market prices adjust to reflect their book values.

Nevertheless, the significant volatility in the predictions and low R^2 values across most horizons further indicate that stock returns are influenced by a complex array of factors beyond the scope of simple financial ratios. Simply by looking at R^2 with the same amount of independent variables and the same dependent variable would tell us the 1-year horizon is probably the optimal model to leverage. Yet the $\sigma[E_t(R^e)]$ has its lowest value at the longest horizon of 5 years, similarly for $\frac{\sigma[E_t(R^e)]}{E(R^e)}$. This suggests that the model's predicted returns are relatively stable and close to the actual average returns, indicating a more reliable and stable predictive performance than a high value which indicates that the model's predictions are not only volatile but also significantly dispersed relative to the average level of returns actually realized. This stability can be particularly valuable in environments where investors seek consistency and are risk-averse.

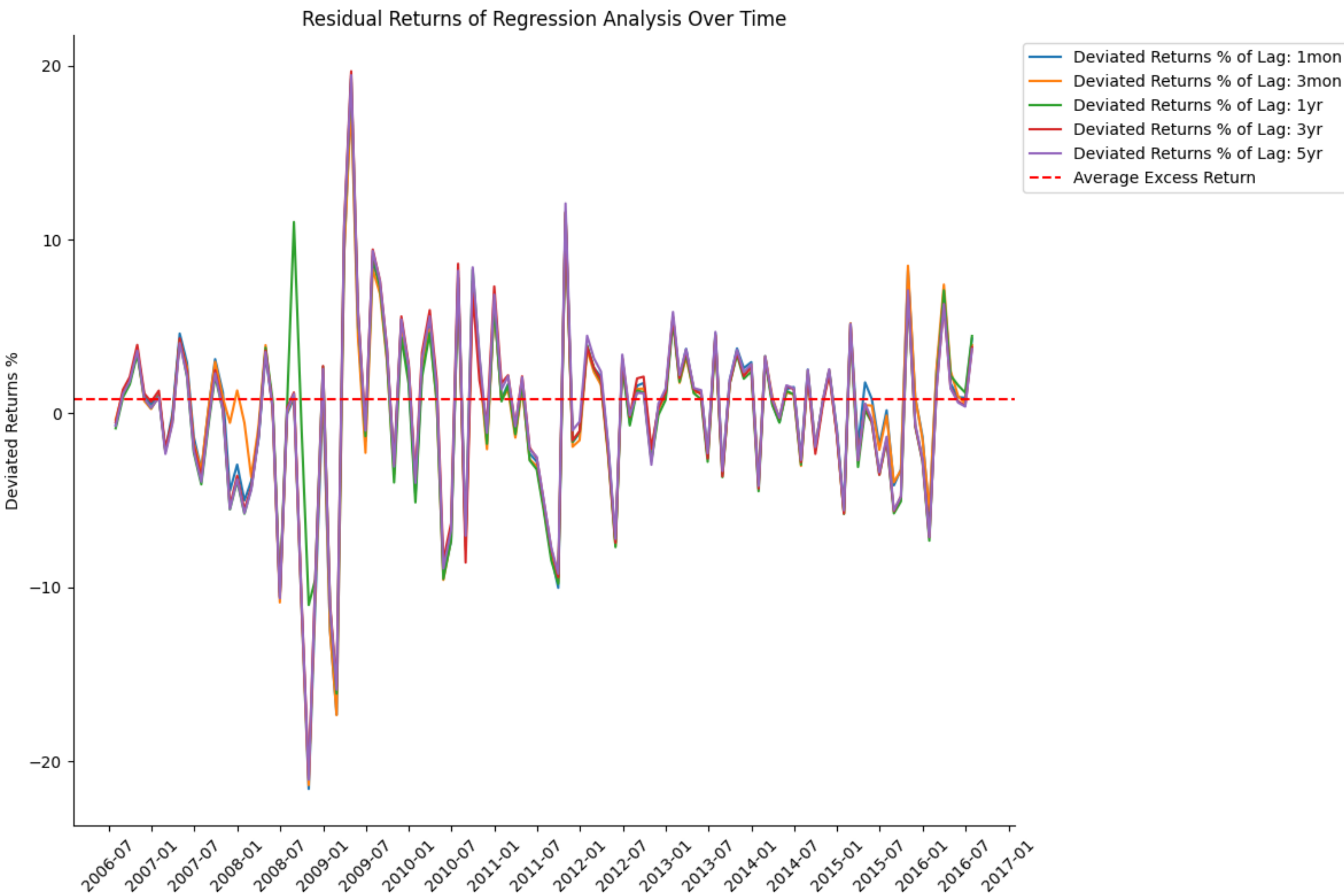

Figure 1: Residuals_Plots

The residuals shown in the chart represent the deviation of the predicted returns from the actual excess returns over time. The various lines (1 month, 3 months, 1 year, 3 years, 5 years) reflect how predictions at different lag periods differ from the actual market performance.

The residuals generally fluctuate around the zero line, which represents the average excess return. This mean-reverting behavior indicates that while the model may overestimate or underestimate returns in certain periods, it tends to balance out over time. However, there are periods where the residuals are persistently above or below zero, suggesting phases where the model systematically deviates from actual returns.

Significant spikes or drops in the residuals might indicate that specific market events or conditions (such as financial crises, abrupt economic changes, or significant corporate actions from 2008-2010) have a profound impact that the model fails to anticipate. These outliers are critical as they might suggest the need for incorporating additional dynamic factors or adjusting the model to better handle anomalies.

From the chart's analysis, while still exhibiting deviations from zero, the residuals appear slightly less erratic from 2013. This could imply that financial ratios are better suited for capturing trends over steady or stable economic periods, aligning more closely with fundamental investment theses that play out over extended time frames.

3.2 Fama French Regressions

		EXM	
	α_i	β_{i_1}	R^2
Long Top Portfolio	0.0032^{*}	0.9249^{***}	0.849
	(0.002)	(0.036)	
Short Bottom Portfolio	-0.0021	-1.3099^{***}	0.755
	(0.003)	(0.068)	
Long-Short Portfolio	0.0019	-0.3874^{***}	0.155
	(0.004)	(0.083)	

standard errors in ()
$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table 18: Regression Results for Portfolio Models on Single Factor

Long Top Portfolio shows a significant positive alpha and a beta close to 1 ($\beta_{i_1} = 0.9249$), suggesting that it not only tracks the market closely with marginally small risk exposures but also achieves slightly higher returns than expected based solely on market movements. This could imply effective management or a selection of securities that slightly outperform the market on a regular basis.

Short Bottom Portfolio has a higher beta than -1 ($\beta_{i_1} = -1.3099$)), indicating an opposite while greater volatility and higher market risk compared to the top portfolio. The lack of significant alpha suggests that the additional risk taken by this portfolio does not yield additional returns above the market, pointing to inefficiencies or higher exposure to systematic risk without corresponding returns. Thus when shorting, this portfolio introduces much more loss than the market at the time.

Long-Short Portfolio demonstrates a negative beta, suggesting it moves in opposition to the market. This is typical for a market-neutral strategy aiming to hedge out market risk. The low R^2 and non-significant alpha indicate that while the portfolio is somewhat effective at hedging market risk, it does not generate significant positive returns above the market on its own.

From this standalone structure, taking the long-short position and long top portfolio appears to be a systematically correct decision.

		EXM	***SMB***	***HML***	
	α_i	β_{i_1}	β_{i_2}	β_{i_3}	R^2
Long Top Portfolio	0.0026	0.9817^{***}	-0.1191	-0.1877^{**}	0.862
	(0.002)	(0.039)	(0.075)	(0.063)	
Short Bottom Portfolio	-0.0043	-1.1590^{***}	-0.1573	-0.6474^{***}	0.813
	(0.003)	(0.068)	(0.131)	(0.111)	
Long-Short Portfolio	-0.0010	-0.1787^{*}	-0.2811	-0.8360^{***}	0.390
	(0.003)	(0.080)	(0.155)	(0.131)	

standard errors in ()
$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table 19: Regression Results for Portfolio Models on Three Factors

Long Top Portfolio in the three-factor model maintains a high correlation with the market and shows a slight aversion to value (negative HML) and small size (negative SMB) stocks, possibly indicating a tilt towards growth and large-cap stocks.

Short Bottom Portfolio exhibits negative alpha which remains consistent with the Single Facor Model analysis, confirming that this portfolio, on average, underperforms the risk-free rate when the market is static. This could be indicative of the costs inherent in maintaining short positions or possibly poor timing in entering or exiting these positions, with consistently higher β amplifying both potential gains and losses. The negative coefficients on both SMB and HML factors reinforce that the portfolio is positioned to benefit from the underperformance of small-cap and high book-to-market (value) stocks. The considerable negative loading on HML suggests a particular expectation that value stocks will perform poorly, a stance that could capitalize on market phases where growth stocks outperform value stocks. With the heavy negative exposure to HML, monitoring economic indicators that affect the performance of value stocks becomes crucial. Shifts in investor sentiment towards value investing or macroeconomic changes favoring value stocks could significantly impact the portfolio's performance.

Long-Short Portfolio distinctly positions itself against the market and value stocks, with significantly negative loadings on two of the three factors. This strategy likely aims to capitalize on the underperformance of high-risk stocks or could be structurally designed to profit from downturns in market segments heavily weighted in value stocks. But similarly as it demonstrates significantly negative exposure to HML, monitoring the value stocks would become much more essential.

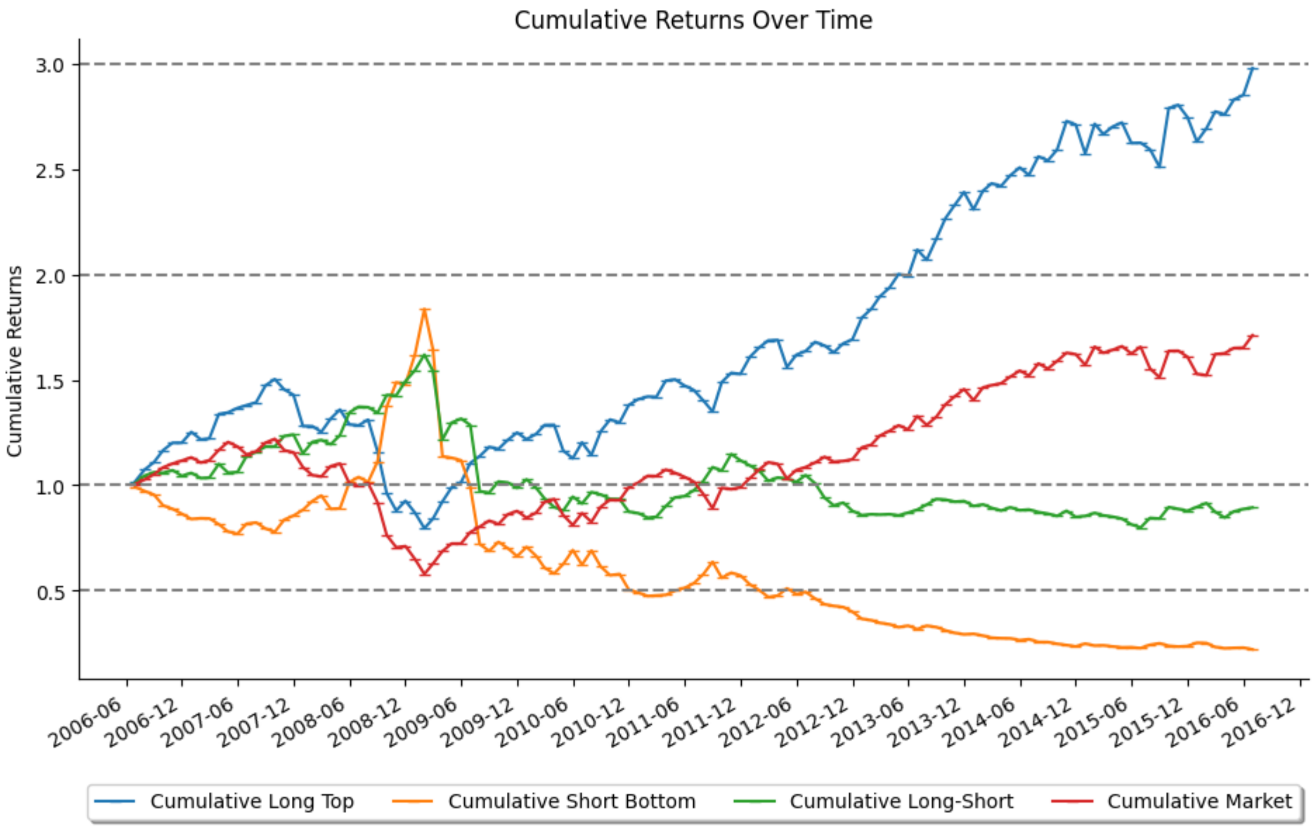

Figure 2: Cumulative Returns Comparison Against the Market

Plotting the cumulative returns of the previous three strategies provides further colorful details of the performances regarding of each of these portfolios.

Long Top Portfolio (Blue Line): This portfolio consistently outperforms the Market throughout the observed period, particularly post-2012, where it shows a significant upward trajectory as indicated by its ability to recover quickly from downturns and continue to grow. This aligns with the regression results indicating a positive alpha and a market-related beta close to 1, suggesting that this portfolio not only tracks the market closely but also captures additional returns, possibly due to effective stock selection and management strategies. It also suggests a resilient strategy likely focused on growth or quality stocks, as previously hypothesized.

Short Bottom Portfolio (Orange Line): The Short Bottom Portfolio exhibits significant volatility, particularly noticeable during the financial crisis around 2008 and the recovery period that follows. During the 2008 crisis, while the market and most other portfolios were experiencing significant downturns, the Short Bottom Portfolio's returns surged, peaking sharply. This is indicative of the portfolio effectively capitalizing on the poor performance of the bottom assets through its short positions. The significant negative beta of this portfolio, as highlighted by the regression results, supports this behavior, indicating an inverse relationship with the market movements. During the market downturn, such a strategy would naturally yield substantial returns. The subsequent sharp decline in the Short Bottom Portfolio's performance post-crisis suggests a rapid reversal in the conditions that initially favored the short positions, likely as the market began to recover.

Long-Short Portfolio (Green Line): The Long-Short Portfolio's performance is relatively flat and does not show substantial growth over time, which is consistent with the low R^2 and non-significant alpha reported in the regression results. This indicates that while the portfolio effectively hedges against market risks, it does not contribute significantly to returns beyond that. Notably, it is largely protective during the 2008 crisis, underlining its effectiveness in capital protection during market downturns. The relatively stable performance after that time period, unaffected by major market swings, highlights its role in risk mitigation rather than return generation.

Market (Red Line): The Market trajectory provides a benchmark, showing substantial growth with expected market cycles. The performance of the Top and Bottom Portfolios relative to this benchmark is telling of their respective strategies' effectiveness

The graph corroborates the detailed findings from the regression analyses provided earlier. The Top Portfolio's ability to consistently outperform the Market underscores its superior stock selection and strategy implementation. The Short Bottom Portfolio's struggles highlight a need for strategic overhaul to avoid low-performing investments and improve its risk-return profile. The Long-Short Portfolio's performance, while not impressive in terms of returns, demonstrates its utility in risk management, particularly valuable during volatile market periods.

3.3 Fama Macbeth Regression

we performed 120 separate monthly regressions across a decade, each examining the relationship between financial performance metrics and subsequent stock returns. This method isolates the individual effects of variables across time, providing a clear view of how each predictor contributes to return variations. The analysis focuses on three key coefficients: λ_1, λ_2, and λ_3 which represent different dimensions of financial data impacting stock performance.

		β_i	**Balanced Ratio**	
Date	${\lambda_1}^t$	${\lambda_2}^t$	${\lambda_3}^t$	R^2
2006-07-01	0.000518	-0.035188***	0.101792**	0.282331
2006-08-01	-0.012061	0.006553	0.063407*	0.067157
2006-09-01	0.019640	0.000981	0.006469	0.001331
2006-10-01	0.046585**	0.011765	-0.029417	0.058081
2006-11-01	0.005206	0.011998	0.008889	0.075400
2006-12-01	0.021904	-0.004761	-0.034730	0.036287
2007-01-01	0.001192	0.009027	0.026236	0.026717
2007-02-01	-0.000806	0.006245	-0.060090*	0.092420
2007-03-01	-0.011036	0.013873	0.004541	0.054133
2007-04-01	-0.033679*	0.043775***	0.100833**	0.346524

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table 20: Snapshot of Fama Macbeth Regressions

The significance and values of these coefficients varied substantially over the observed period, suggesting fluctuating influences on stock returns. For instance:

In July 2006, λ_2 is significantly negative, suggesting a period where higher leverage or risk metrics negatively impacted returns.

By April 2007, λ_2 exhibits a significant positive relationship with returns, indicating a shift in market dynamics or investor sentiment where financial risk is potentially rewarded. This variation highlights the contextual sensitivity of financial ratios and their impact on stock performance.

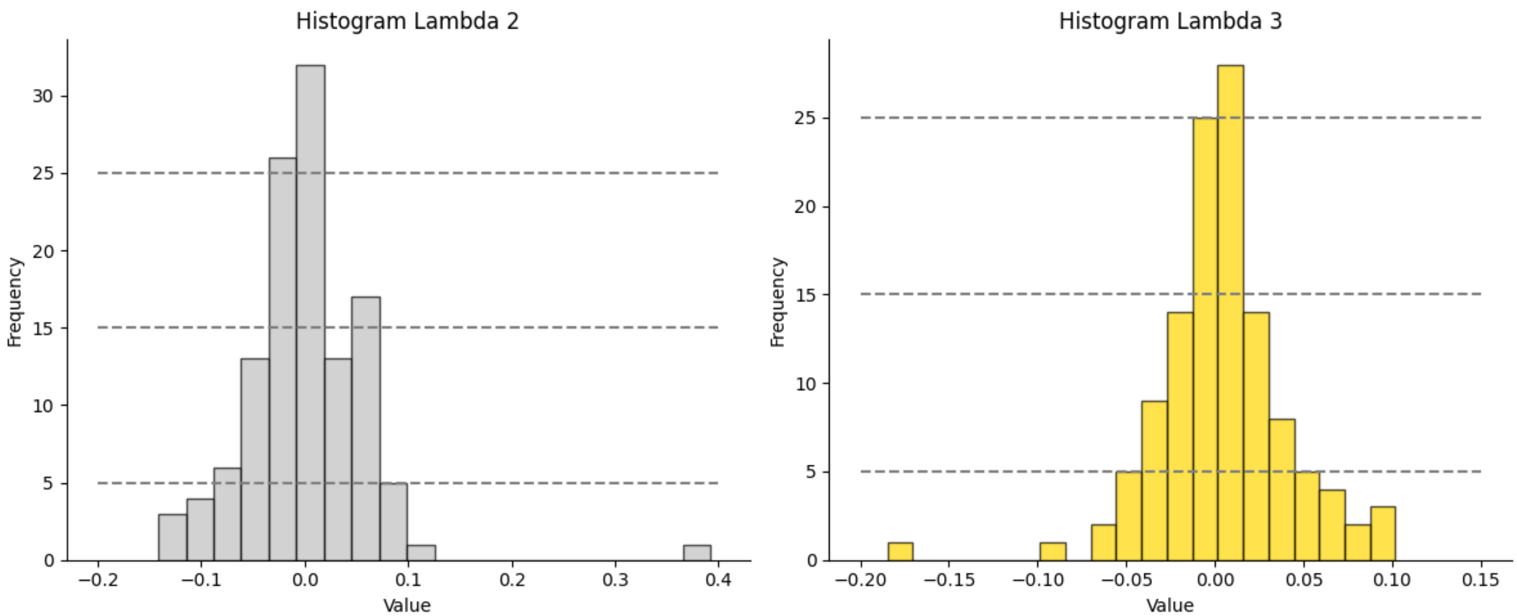

Figure 3: λ Distribution

The histograms provided (Lambda 2 and Lambda 3) elucidate the distribution of these coefficients over multiple periods:

The distribution of λ_2 is somewhat balanced around zero indicating that the influence of the related risk sensitivity metric can cause minimal impacts on returns.

λ_3 shows a more consistent clustering around zero if looking the frequency, suggesting a generally stable but quite limited impact on stock returns.

The table of statistically significant lambda values for each month further supports this:

No.$\lambda_2 > 0$	No.$\lambda_3 > 0$	No.P-value $\lambda_2 <= 0.05$	No.P-value $\lambda_3 <= 0.05$
61	65	66	21

Table 21: λ Distribution

The number of months with positive values for λ_2 and λ_3 suggests a relatively strong consistent direction in their impact, pointing to their conditional relevance. A relatively small number of months where λ_2 and λ_3 p-values are below 0.05 indicates that while these metrics can be significant, compared to the risk exposure metric, the financial metrics selected do not provide consistent predictive power across all periods.

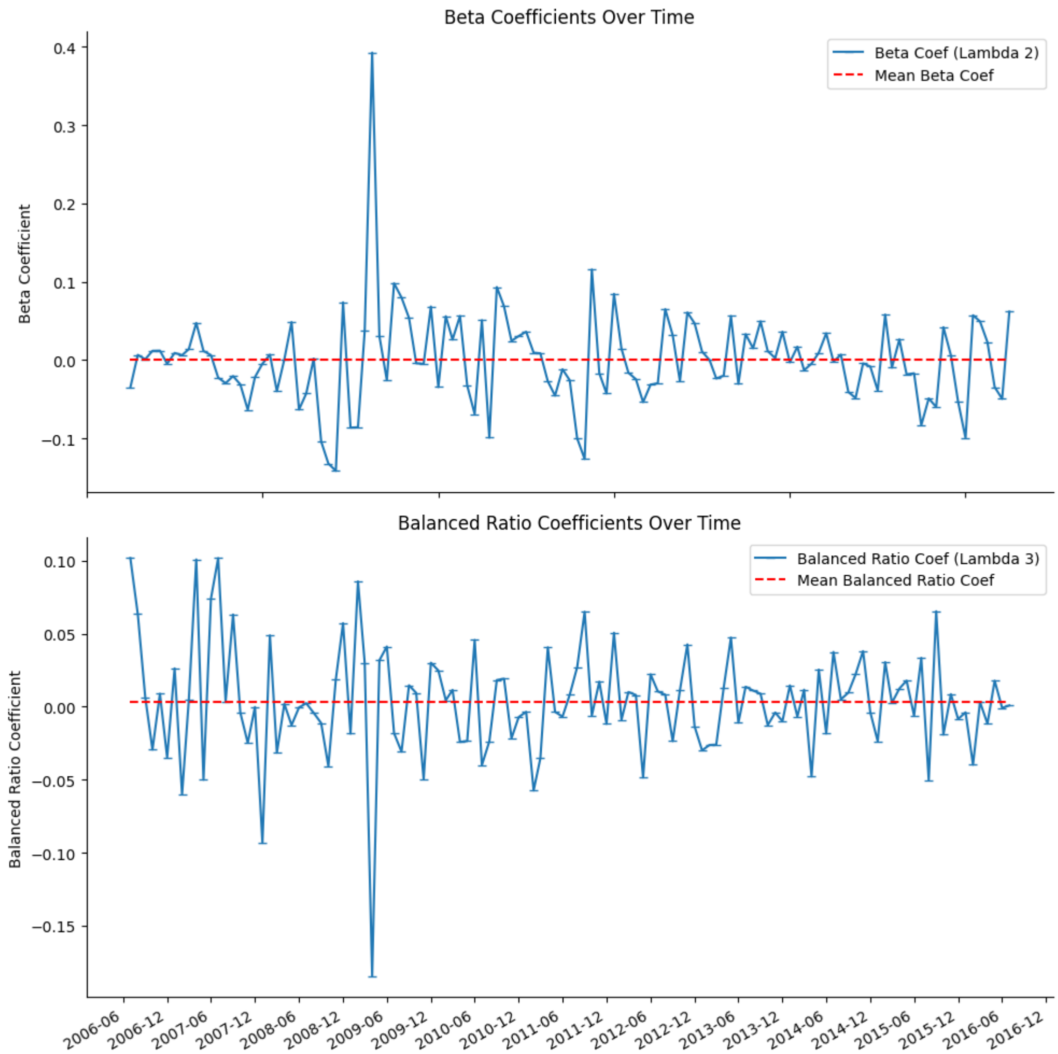

Figure 4: λ Distribution

The charts provided display the fluctuating values of λ_2 (Beta Coefficient) and λ_3 (Balanced Ratio Coefficient) over the period from 2006 to mid-2016, along with their respective mean values.

It is worth noting that both λ_2 and λ_3 have their mean values approaching zero. This suggests that, on average, the risk sensitivity metric and financial metrics corresponding to λ_2 and λ_3 do not consistently add to or detract from stock returns when averaged over the entire period. The proximity of the mean to zero despite high volatility indicates that the positive and negative impacts might offset each other over time.

The mean's statistical test returns:

	λ_2	λ_3
t-stats	0.02007880549138559	0.9633950378940588
p-value	0.9840138552718114	0.3372866952705279

Table 22: E(λ) Statistical Test

The lack of statistical significance in the mean values of both λ_2 andλ_3 underscores the primary conclusion of our Fama-Macbeth regression analysis: while specific financial metrics can exhibit transient influence on stock returns, their overall predictive power is limited when generalized across diverse market conditions and times. These findings highlight the importance of situational analysis and the potential risks of relying on these metrics without consideration of broader economic indicators or market trends. Investors should be cautious in employing these financial metrics as standalone predictors within their investment strategies.

Year-Month	Unique Companies Count
2006-07	74
2006-08	74
2006-09	73
2006-10	73
2006-11	73
2006-12	71

Table 23: 2006

Year-Month	Unique Companies Count
2016-01	68
2016-02	65
2016-03	64
2016-04	64
2016-05	67
2016-06	67
2016-07	67

Table 24: 2016

Year-Month	Unique Companies Count
2007-01	71
2007-02	70
2007-03	70
2007-04	70
2007-05	71
2007-06	70
2007-07	70
2007-08	70
2007-09	70
2007-10	70
2007-11	70
2007-12	70

Table 25: 2007

Year-Month	Unique Companies Count
2008-01	70
2008-02	69
2008-03	69
2008-04	69
2008-05	70
2008-06	70
2008-07	70
2008-08	70
2008-09	70
2008-10	70
2008-11	70
2008-12	70

Table 26: 2008

Year-Month	Unique Companies Count
2009-01	70
2009-02	68
2009-03	68
2009-04	68
2009-05	70
2009-06	70
2009-07	70
2009-08	70
2009-09	70
2009-10	70
2009-11	70
2009-12	70

Table 27: 2009

Year-Month	Unique Companies Count
2010-01	70
2010-02	68
2010-03	67
2010-04	67
2010-05	68
2010-06	68
2010-07	68
2010-08	68
2010-09	68
2010-10	68
2010-11	68
2010-12	68

Table 28: 2010

Year-Month	Unique Companies Count
2011-01	68
2011-02	68
2011-03	68
2011-04	68
2011-05	69
2011-06	68
2011-07	68
2011-08	68
2011-09	68
2011-10	68
2011-11	67
2011-12	67

Table 29: 2011

Year-Month	Unique Companies Count
2012-01	67
2012-02	66
2012-03	66
2012-04	66
2012-05	68
2012-06	67
2012-07	67
2012-08	67
2012-09	67
2012-10	67
2012-11	67
2012-12	67

Table 30: 2012

Year-Month	Unique Companies Count
2013-01	67
2013-02	66
2013-03	65
2013-04	65
2013-05	66
2013-06	66
2013-07	66
2013-08	65
2013-09	65
2013-10	65
2013-11	66
2013-12	65

Table 31: 2013

Year-Month	Unique Companies Count
2014-01	65
2014-02	64
2014-03	64
2014-04	64
2014-05	65
2014-06	65
2014-07	65
2014-08	65
2014-09	64
2014-10	64
2014-11	64
2014-12	64

Table 32: 2014

Year-Month	Unique Companies Count
2015-01	64
2015-02	63
2015-03	62
2015-04	62
2015-05	63
2015-06	64
2015-07	64
2015-08	67
2015-09	67
2015-10	67
2015-11	68
2015-12	68

Table 33: 2015

Here significant challenges were encountered in applying more complex statistical tests such as the joint test of residuals using a covariance matrix, mainly due to inconsistencies in the data sets related to varying company counts each month.

4 Conlusion

This investment memo critically evaluates the predictive power of financial ratios within the S&P 100 from 2006 to 2016 using Predictive Regression, Fama-French, and Fama-Macbeth models, which provides an analysis of the effectiveness of Debt-to-Equity (D/E) and Price-to-Book (P/B) ratios, alongside a composite of Cash Ratio and Return on Equity (ROE), in formulating robust investment strategies.

Our analysis suggests that D/E and P/B ratios hold some predictive merit over extended periods, despite minor deviations from projected returns. However, their overall efficacy is constrained, both economically and statistically, particularly for the D/E ratio. This variability underscores the potential risks of relying solely on these metrics for long-term investment decisions. Notably, a five-year horizon emerges as a moderately reliable benchmark, deviating slightly from expected returns.

In terms of portfolio construction, the combination of Cash Ratio (as a liquidity measure) and ROE (as an indicator of operational efficiency) appears effective for ranking and selecting firms from the S&P 100. Annually investing in top-ranked firms within an equal-weight portfolio yielded significant returns relative to the market, although shorting the lowest-ranked stocks proved less beneficial. This discrepancy may indicate that the singular importance of these two ratios in investment decision-making could be overestimated. Moreover, the Long-Short portfolio strategy, while providing a hedge against market risk, does not adequately incorporate factors beyond these two ratios, leading to suboptimal outcomes in realizing downside returns.

The combination of Cash Ratio and ROE shows a more balanced but still limited predictive capability as per the Fama-Macbeth analyses. The lack of statistical significance in the lambda values from the Fama-Macbeth regressions indicates that these financial metrics, while useful under certain conditions, should not be the sole basis for investment decisions. Instead, they should be part of a broader, more diversified analysis strategy as our statistical analysis indicated that neither set of the market risk exposure nor financial metrics consistently predicts stock returns, as evidenced by their overall performance across various tests.