1. INTRODUCTION:

Understanding how markets price uncertainty is central to asset pricing. I examine whether firms with more volatile historical earnings earn higher future returns by constructing a long–short strategy based on five-year realized EPS dispersion.

Unlike prior studies such as Diether, Malloy, and Scherbina (2002), which use analyst forecast dispersion, I focus on realized fundamentals — specifically, the coefficient of variation in EPS over a five-year window. This backward-looking measure captures persistent uncertainty while smoothing transitory noise, aligning with risk-based insights from Bali, Cakici, and Whitelaw (2011).

Each month, I sort firms into quintiles and take a long position in high-dispersion firms (Q5) and a short position in low-dispersion firms (Q1), with volatility targeting applied to equalize risk across legs. I enforce practical constraints including price filters and winsorization.

The strategy (5-year window) delivers a gross annualized return of 23.2% with a Sharpe ratio of 1.26, and maintains strong performance net of 20 bps costs (20.8% alpha, Sharpe 1.13). Its five-year CAGR of 26.8% outpaces the market's 19.3%. A Fama-French 3-factor regression confirms significance, with a net alpha of 16.2% ($t = 2.00$). The 5-year window offers a practically relevant track record aligned with institutional evaluation horizons, and matches the long-term nature of my 5-year dispersion signal.

By relying on long-run realized dispersion rather than expectations, the strategy captures persistent uncertainty that may be mispriced, supporting both a risk-premium and behavioral underreaction interpretation.

2. ECONOMIC INTUITION:

My results show that stocks with high five-year earnings dispersion consistently outperform more stable peers. This may seem counterintuitive given prior work linking uncertainty to overpricing and underperformance. HoIver, the key distinction is in the signal: unlike forward-looking forecast dispersion, my measure captures realized, persistent volatility in firm fundamentals — a form of uncertainty tied to valuation difficulty.

Risk-Based Explanation:

High earnings dispersion often reflects true economic risk, such as volatile cash flows or macro sensitivity. Investors may demand compensation for holding such firms, and my findings support this view: the strategy delivers strong risk-adjusted returns, positive skewness, and statistically significant alpha. This is especially evident in complex sectors like Electronics and Business Services, consistent with Bali, Cakici, and Whitelaw (2011), who show that idiosyncratic volatility can predict higher returns.

Behavioral Explanation:

Alternatively, high-dispersion stocks may be mispriced due to investor underreaction. These firms are often small, opaque, or difficult to forecast — traits that trigger ambiguity aversion and cause underIighting. This behavioral gap creates persistent alpha, which my strategy exploits through systematic exposure to neglected, high-uncertainty stocks.

Why the Strategy May Persist:

- Uses slow-moving accounting data often ignored by fast-moving quant models
- Targets firms typically screened out for high volatility or uncertainty
- Occupies a distinct niche: neither deep value nor growth, but fundamentally complex firms

Together, these features explain the persistence of the return premium and support the view that long-run earnings dispersion captures a durable, underutilized alpha smyce.

3. DATA & UNIVERSE:

I construct my dataset by merging monthly equity return data from CRSP with firm fundamentals from COMPUSTAT, linked using the CRSP-COMPUSTAT Link Table via PERMNO and GVKEY. All firms with valid data are included, and I apply a $5 price filter to exclude penny stocks and reduce microstructure noise.

Annual earnings per share (EPS) is computed from Compustat net income (niq) and shares outstanding (cshoq), with gaps filled using CRSP's December-end shares (SHROUT). My signal is the five-year coefficient of variation in EPS — the standard deviation divided by the mean over a rolling five-year window. I exclude firms with near-zero average EPS or feIr than five valid observations but do not winsorize the dispersion values.

Firms are retained if they:

- Have five years of valid EPS data,
- Are linked through a primary CRSP-COMPUSTAT key (LINKPRIM = 'P'),
- Have non-missing monthly returns,
- And meet the $5 price threshold.

The resulting dataset includes approximately 3,500 unique firms and over 500,000 firm-month observations, covering a diverse and investable cross-section of U.S. equities. While I do not explicitly correct for survivorship bias, any firm with valid returns and dispersion data is included, reflecting the practical constraints faced by institutional investors.

I conduct two analyses: an academic factor study spanning 1966–2024, using dispersion-sorted portfolios to evaluate long-run return predictability, and a trading strategy backtest over 2019–2024 that incorporates volatility targeting, transaction costs, and real-world constraints. The former isolates the statistical strength of the signal, while the latter tests its practical performance.

4. SIGNAL GENERATION & STATISTICAL CREDIBILITY:

My trading signal is based on realized earnings volatility, which I use as a proxy for uncertainty or disagreement in a firm's fundamental value. The core idea is that firms with volatile historical earnings are harder to value, leading to greater dispersion in investor beliefs or risk premia.

Unlike Diether, Malloy, and Scherbina (2002), who use forward-looking analyst forecast dispersion, I rely on backward-looking fundamentals to capture realized uncertainty. Specifically, for each fiscal year, I compute annual EPS from Compustat quarterly net income (niq) and shares outstanding (cshoq), filling missing data using CRSP December-end shares (SHROUT).

I then calculate the coefficient of variation in EPS over a trailing five-year window for each firm:

$$\mathrm{EPS}_{i,t} = \frac{\sum_{q=1}^{4} \mathrm{NIQ}_{i,t,q}}{\mathrm{Shares}_{i,t}}$$

$$\mathrm{Dispersion}_{i,t} = \frac{\sigma\left(\mathrm{EPS}_{i,t-5:t-1}\right)}{\left|\mu\left(\mathrm{EPS}_{i,t-5:t-1}\right)\right|}$$

This ratio captures relative earnings instability. Firms are only included if they have at least five non-missing EPS observations, and I exclude those with near-zero mean EPS to prevent inflated or spurious dispersion values.

Dispersion is computed at the annual frequency using Compustat quarterly net income and shares outstanding. This annual signal is then mapped forward to each month in the same fiscal year, allowing us to form monthly portfolios while keeping the signal fixed within each year. This ensures that the signal is persistent, avoids lookahead bias, and reflects what would be observable to investors at the time.

Low-dispersion firms (e.g., large, stable blue-chip names) exhibit consistent earnings, while high-dispersion firms tend to be speculative or early-stage businesses with unpredictable income streams. This variation helps my strategy distinguish betIen fundamentally stable and uncertain firms.

I conduct additional checks to ensure signal robustness:

- Firms with average EPS below a materiality threshold (e.g., $0.05) are excluded to avoid denominator-driven noise.
- Firms with frequent EPS sign changes (e.g., oscillating betIen profits and losses) are also removed, as they yield unstable dispersion metrics.
- I do not apply winsorization to the dispersion values, preserving the full cross-sectional signal after filtering.

I evaluate the pricing implications of the dispersion signal using CAPM and Fama-French 3-factor regressions across dispersion-sorted portfolios. Over the full sample (1966–2024), CAPM alphas rise from 3.5% (D1) to 13.0% (D5), with a significant long–short alpha of 5.0% ($t = 2.61$) (refer to Appendix A). Under the FF3 model, the alpha remains significant at 5.07% ($t = 3.34$), with some SMB exposure but minimal HML loading (refer to Appendix B). These results confirm that dispersion captures return premia not explained by standard risk factors. In the recent 5-year period (2019–2024), the long–short portfolio earns a CAPM alpha of 10.63% ($t = 2.44$) (refer to Appendix C) and an FF3 alpha of 11.57% ($t = 1.20$) (refer to Appendix D). While the short horizon reduces statistical poIr, the economic magnitude of returns is substantial. The strategy loads positively on SMB (≈ 0.61), reflecting a tilt toward smaller, higher-uncertainty firms. The return distribution is attractive for the recent 5-year period: a mean annual return of 15.6%, volatility of 11.0%, Sharpe ratio of 1.20, and positive skew (0.66) with low kurtosis (0.22).

	MeanReturn(%)	Volatility(%)	Sharpe	Skewness	Kurtosis
D5_minus_D1	15.57	10.98	1.2	0.66	0.22

Overall, dispersion is a robust predictor of returns, with statistically and economically significant performance across periods and models. I also see a monotonic increase in returns across the deciles as I go from D1 to D5 (refer to Appendix F).

5. PORTFOLIO FORMATION:

Each month, I perform a cross-sectional sort of all eligible firms into dispersion quintiles (Q1 to Q5):

- Q1 contains firms with the most stable historical earnings (low dispersion),
- Q5 contains firms with the most volatile earnings (high dispersion).

$$\mathrm{LS}_t = R_t^{Q5} - R_t^{Q1}$$

I then compute equal-Iighted monthly returns for each quintile and focus on the return spread betIen Q5 and Q1. This long–short spread portfolio captures the excess return associated with high versus low earnings dispersion and is updated monthly based on the most recently available dispersion values. While the dispersion values themselves are updated annually, the portfolio is rebalanced monthly as the set of eligible firms evolves, firm rankings shift, and new return data becomes available. As a result, the composition of each quintile can change each month even though the signal input is held constant within the year.

The strategy's long–short construction reveals a distinct tilt in both size and sector exposure. Based on boxplot analysis, firms in the highest dispersion quintile (Q5) are disproportionately smaller, with earnings instability decreasing monotonically across size quintiles. This reflects a natural size effect: smaller firms tend to exhibit greater earnings volatility (refer to Appendix G). On the industry front, the long leg (Q5) is concentrated in Public Administration, Oil & Gas, and Holding Companies, while the short leg (Q1) is heavily Iighted toward Banks, Utilities, and Insurance — sectors typically associated with stable, regulated earnings (refer to Appendix H). The strategy therefore exploits a contrast betIen high-uncertainty, often undercovered firms and low-volatility, Ill-established institutions, aligning with the broader interpretation of dispersion as a proxy for informational complexity or mispricing.

In my trading implementation (2019–2024), I build upon this basic construction by applying:

- Volatility targeting: Each leg of the strategy is scaled to maintain a constant annualized volatility (target = 20%), improving risk control.
- Transaction cost adjustments: A conservative 20 basis points per month is deducted to reflect realistic execution costs.
- Winsorization: I apply light winsorization (at the 0.5% level in each tail) to the long–short return series to mitigate the impact of extreme outliers, which can arise due to monthly return spikes in thinly traded stocks.
- Monthly rebalancing: Returns and exposures are refreshed monthly to reflect updated CRSP return data, while keeping the dispersion signal fixed for the fiscal year.

This results in a robust and implementable return stream that mirrors what a systematic investor could have realistically achieved, with risk budgeting and cost awareness embedded into portfolio construction.

6. PERFORMANCE EVALUATION:

The dispersion-based trading strategy delivers robust performance from 2019 to 2024, compounding at 26.76% annually compared to the market's 19.28%. After accounting for 20 bps/month transaction costs and applying volatility targeting, the strategy achieves a compounding annual return of 23.80%, net alpha of 20.76%, volatility of 18.42%, and a Sharpe ratio of 1.13. Its return profile is positively skeId (0.72) with low kurtosis (0.26), reflecting favorable upside risk and limited tail exposure.

- **CAGR (Strategy - Gross): 26.76%**
- **CAGR (Strategy - Net of Costs): 23.80%**
- **CAGR (Market): 19.28%**

	Annualized Alpha (%)	Volatility (%)	Sharpe Ratio	Skewness	Kurtosis
VolTargeted WML	23.16	18.42	1.26	0.72	0.26
Net of Costs VolTargeted WML	20.76	18.42	1.13	0.72	0.26

Factor regressions confirm that this alpha is not explained by standard risk premia: the strategy earns an FF3 alpha of 16.21% (t = 2.00) with minimal loading on market or value and modest exposure to SMB. This supports the view that dispersion captures a distinct, priced risk or persistent mispricing. The factor regression alpha for the net of cost strategy also yields a strong 13.81% with modest t-statistic of 1.70 (refer to Appendix J).

The strategy targets high-uncertainty firms and avoids traditional style overlaps. Risk control is built in via volatility scaling, and transaction costs are explicitly modeled. Robustness is validated through a dedicated out-of-sample 5-year window, net-of-cost metrics, and factor model validation — all of which are implemented directly in code. These safeguards ensure the signal's practical viability and reduce the likelihood of overfitting.

7. LIMITATIONS:

While the strategy demonstrates strong performance, several limitations remain. First, it requires five years of clean EPS data, which systematically excludes younger or newly listed firms. Second, as a backward-looking measure, the dispersion signal may be slow to reflect rapid changes in fundamentals. Although I incorporate trading cost adjustments in the final implementation, the underlying signal may still rely on less liquid, volatile stocks that can be more expensive to trade in practice. Finally, delisted firms may be underrepresented due to incomplete return histories, introducing potential survivorship bias into the backtest.

Nonetheless, the strategy offers a compelling alternative to traditional factors by targeting long-run earnings uncertainty—a poIrful and persistent signal that is underexploited, economically intuitive, and delivers strong outperformance even after accounting for costs and risks.

8. REFERENCES:

Bali, Cakici, and Whitelaw (2011):

Bali, T. G., Cakici, N., & Whitelaw, R. F. (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. Jmynal of Financial Economics, 99(2), 427–446. https://doi.org/10.1016/j.jfineco.2010.08.014

Diether, Malloy, and Scherbina (2002):

Diether, K. B., Malloy, C. J., & Scherbina, A. (2002). Differences of opinion and the cross section of stock returns. Jmynal of Finance, 57(5), 2113–2141. https://doi.org/10.1111/1540-6261.00490

9. APPENDIX:

Appendix A: CAPM Regression Results (1966–2024)

	alpha(ann %)	t(alpha)	p(alpha)	beta_Mkt	t(beta_Mkt)	p(beta_Mkt)	R2_adj
Portfolio							
D1	3.53	2.80	0.01	0.84	28.02	0.0	80.54
D2	3.42	2.66	0.01	1.00	32.47	0.0	85.27
D3	5.80	4.01	0.00	1.11	35.40	0.0	83.44
D4	9.43	5.55	0.00	1.20	35.62	0.0	80.69
D5	13.03	6.03	0.00	1.29	28.89	0.0	75.88
D5_minus_D1	5.01	2.61	0.01	0.46	11.79	0.0	32.81

Appendix B: Fama-French 3-Factor Regression Results (1966–2024)

	alpha (ann %)	t(alpha)	p(alpha)	beta_Mkt	t(beta_Mkt)	p(beta_Mkt)	beta_SMB	t(beta_SMB)	p(beta_SMB)	beta_HML	t(beta_HML)	p(beta_HML)	R2_adj
Portfolio													
D1	1.81	2.03	0.04	0.83	38.38	0.0	0.29	7.00	0.0	0.38	9.67	0.00	90.37
D2	1.96	2.30	0.02	0.95	48.29	0.0	0.43	9.65	0.0	0.30	7.44	0.00	94.34
D3	4.47	4.84	0.00	1.02	64.40	0.0	0.59	16.95	0.0	0.25	7.88	0.00	94.71
D4	8.23	7.46	0.00	1.07	55.74	0.0	0.73	19.24	0.0	0.20	8.44	0.00	94.11
D5	11.36	9.08	0.00	1.14	45.42	0.0	0.90	17.01	0.0	0.30	7.22	0.00	92.76
D5_minus_D1	5.07	3.34	0.00	0.32	9.49	0.0	0.61	12.78	0.0	-0.09	-1.41	0.16	60.34

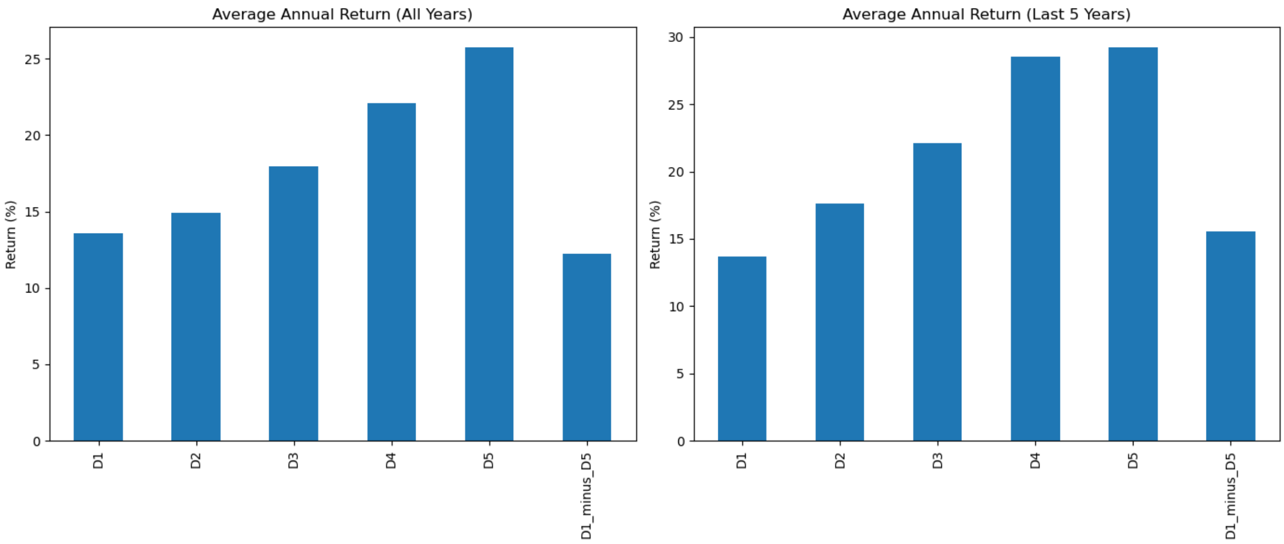
Appendix C: CAPM Regression Results (2019–2024)

	alpha(ann %)	t(alpha)	p(alpha)	beta_Mkt	t(beta_Mkt)	p(beta_Mkt)	R2_adj
Portfolio							
D1	-5.94	-0.92	0.36	0.99	11.21	0.0	78.25
D2	-3.07	-0.56	0.57	1.14	13.50	0.0	85.54
D3	-1.00	-0.20	0.84	1.23	15.11	0.0	86.81
D4	4.07	0.84	0.40	1.30	16.22	0.0	87.77
D5	7.07	1.09	0.27	1.37	12.42	0.0	83.64
D5_minus_D1	10.63	2.44	0.01	0.38	6.64	0.0	42.58

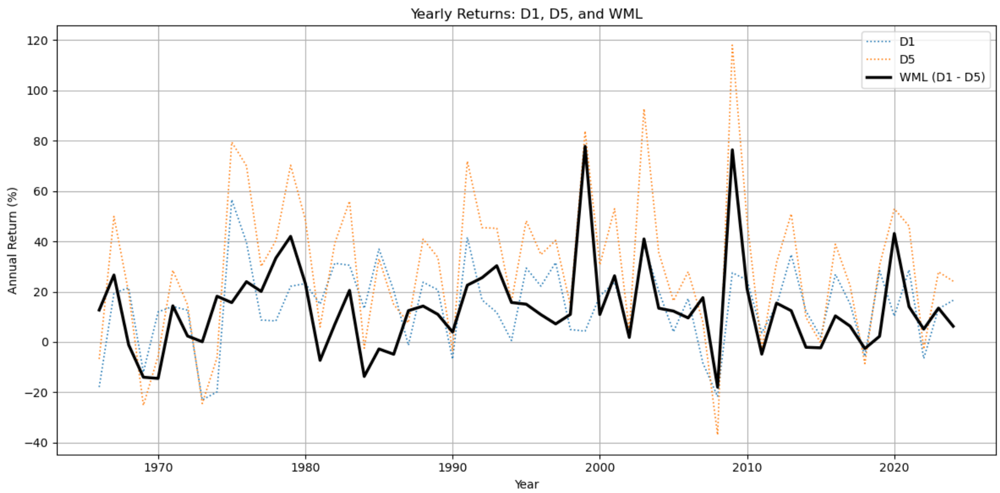
Appendix D: Fama-French 3-Factor Regression Results (2019–2024)

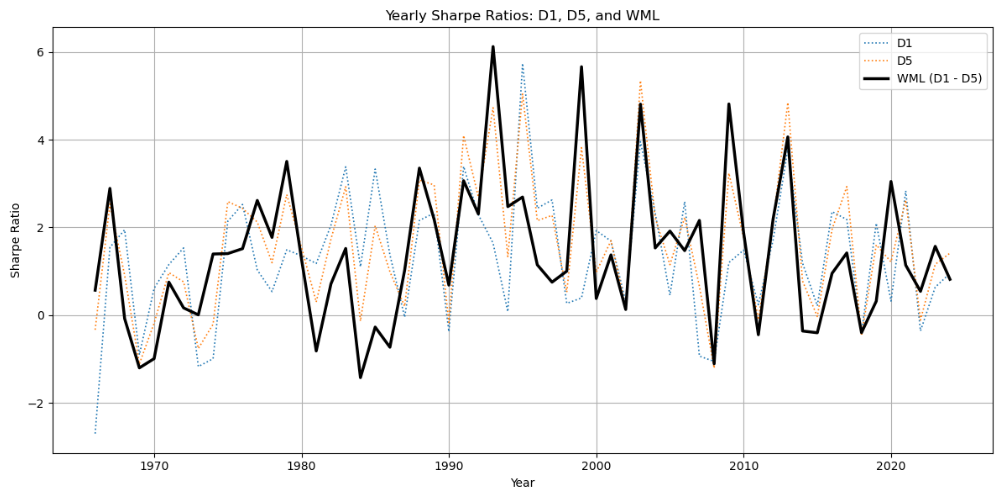
	alpha(ann %)	t(alpha)	p(alpha)	beta_Mkt	t(beta_Mkt)	p(beta_Mkt)	beta_SMB	t(beta_SMB)	p(beta_SMB)	beta_HML	t(beta_HML)	p(beta_HML)	R2_adj
Portfolio													
D1	1.16	0.66	0.51	0.82	22.06	0.0	0.61	7.75	0.00	0.40	8.81	0.00	96.38
D2	3.49	2.10	0.04	0.98	29.45	0.0	0.64	9.33	0.00	0.31	8.63	0.00	97.43
D3	5.75	4.81	0.00	1.05	35.31	0.0	0.80	19.52	0.00	0.19	8.40	0.00	97.96
D4	10.43	4.70	0.00	1.13	29.74	0.0	0.74	9.55	0.00	0.19	4.87	0.00	96.56
D5	15.04	4.02	0.00	1.16	28.22	0.0	0.80	7.58	0.00	0.36	6.63	0.00	95.41
D5_minus_D1	11.57	2.44	0.01	0.35	4.99	0.0	0.19	1.20	0.23	-0.04	-0.46	0.64	43.02

Appendix E: Monotonic Increase in Decile Average Returns

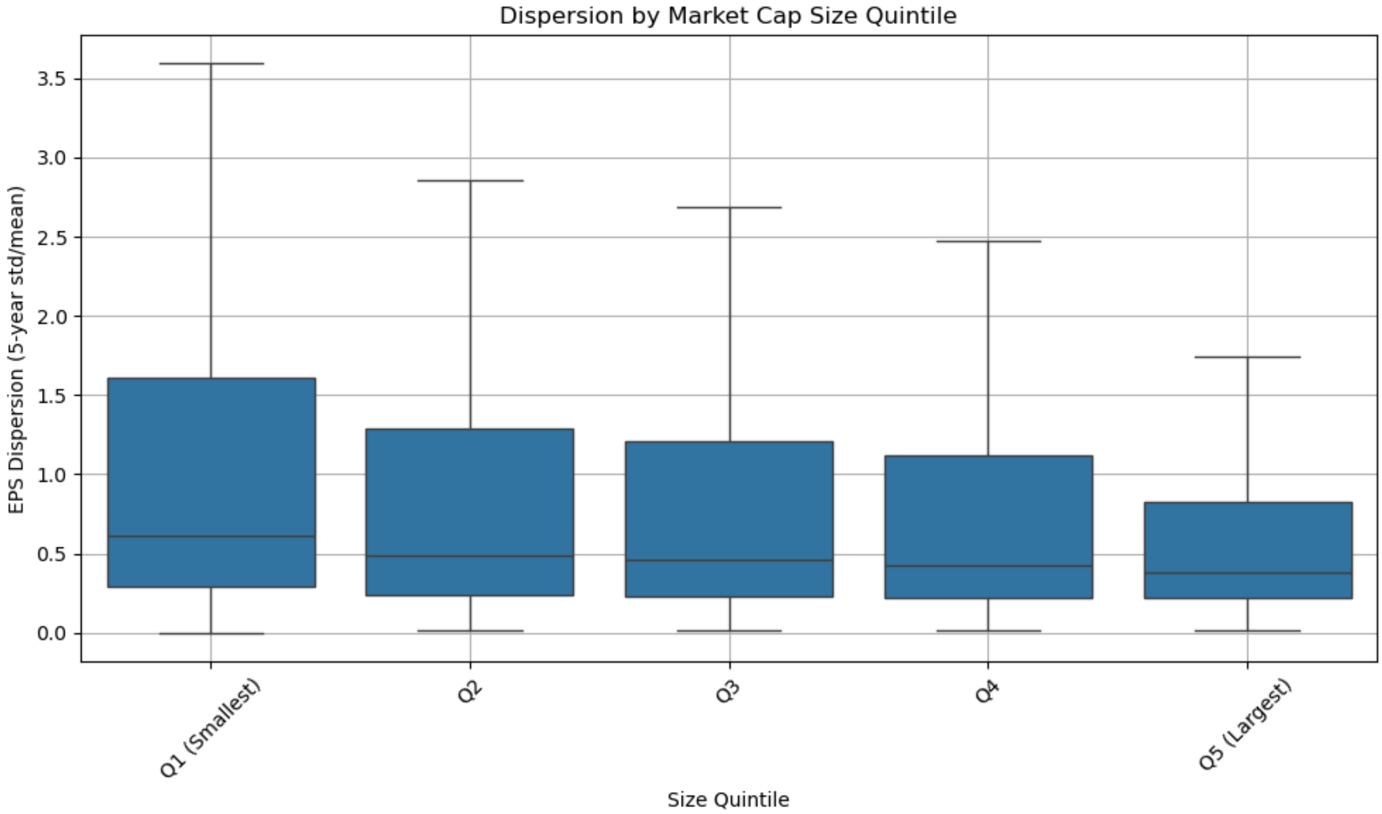


Appendix F: Year to Year Returns and Sharpe



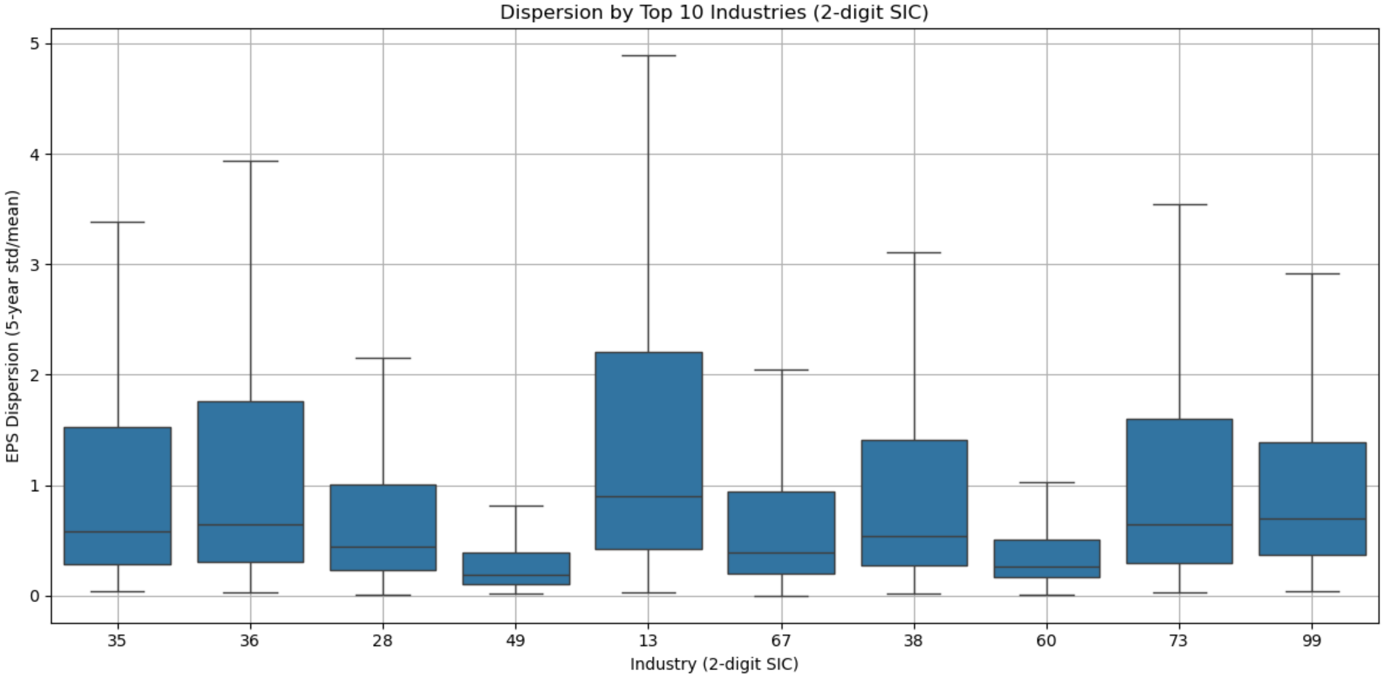


Appendix G: Dispersion by Market Size



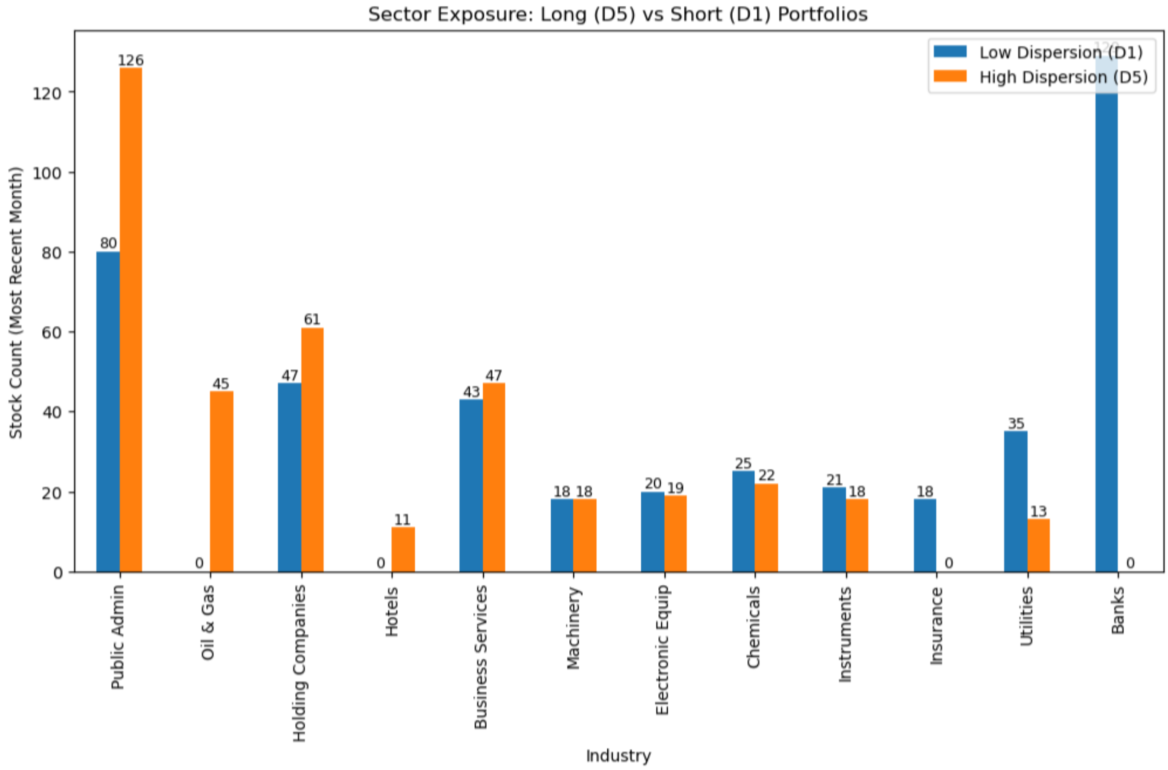
size_bin	count	mean	std	min	25%	50%	75%	max
Q1 (Smallest)	375812.0	5.876830	235.007587	0.000585	0.294877	0.612945	1.613727	38700.718389
Q2	375806.0	3.887383	152.487961	0.011415	0.236628	0.487522	1.285044	40549.493615
Q3	375809.0	5.260507	368.882978	0.013609	0.229603	0.459880	1.212518	77253.245861
Q4	375809.0	3.104585	187.682778	0.011364	0.220749	0.427120	1.120336	77253.245861
Q5 (Largest)	375809.0	2.480984	132.285248	0.011184	0.216234	0.384657	0.826734	28451.879076

Appendix H: Dispersion by Industries



	Industry	D1_Count	D5_Count	Net_Long_Tilt
11	Public Admin	80.0	126.0	46.0
0	Oil & Gas	0.0	45.0	45.0
8	Holding Companies	47.0	61.0	14.0
9	Hotels	0.0	11.0	11.0
10	Business Services	43.0	47.0	4.0
2	Machinery	18.0	18.0	0.0
3	Electronic Equip	20.0	19.0	-1.0
1	Chemicals	25.0	22.0	-3.0
4	Instruments	21.0	18.0	-3.0
7	Insurance	18.0	0.0	-18.0
5	Utilities	35.0	13.0	-22.0
6	Banks	129.0	0.0	-129.0

Appendix I: Sector Exposure



Appendix J: Fama French Regression Results (2019-2024 Trading Strategy Implementation)

FF3 Gross Strategy Alpha: 16.21%, t-stat: 2.00

OLS Regression Results

Dep. Variable:	ExcessRet	R-squared:	0.299
Model:	OLS	Adj. R-squared:	0.262
Method:	Least Squares	F-statistic:	9.870
Date:	Sun, 01 Jun 2025	Prob (F-statistic):	2.53e-05
Time:	19:25:43	Log-Likelihood:	102.09
No. Observations:	60	AIC:	-196.2
Df Residuals:	56	BIC:	-187.8
Df Model:	3		
Covariance Type:	HAC		

	coef	std err	z	P>|z|	[0.025	0.975]
const	0.0135	0.007	1.996	0.046	0.000	0.027
Mkt-RF	0.4212	0.088	4.807	0.000	0.249	0.593
SMB	0.4384	0.305	1.436	0.151	-0.160	1.037
HML	-0.0578	0.145	-0.398	0.691	-0.343	0.227

Omnibus:	0.231	Durbin-Watson:	1.553
Prob(Omnibus):	0.891	Jarque-Bera (JB):	0.247
Skew:	-0.136	Prob(JB):	0.884
Kurtosis:	2.842	Cond. No.	34.0

FF3 Net of Cost Strategy Alpha: 13.81%, t-stat: 1.70

```
                            OLS Regression Results
==============================================================================
Dep. Variable:           NetExcessRet   R-squared:                       0.299
Model:                            OLS   Adj. R-squared:                  0.262
Method:                 Least Squares   F-statistic:                     9.870
Date:                Sun, 01 Jun 2025   Prob (F-statistic):           2.53e-05
Time:                        20:31:23   Log-Likelihood:                 102.09
No. Observations:                  60   AIC:                            -196.2
Df Residuals:                      56   BIC:                            -187.8
Df Model:                           3
Covariance Type:                  HAC
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.0115      0.007      1.701      0.089      -0.002       0.025
Mkt-RF         0.4212      0.088      4.807      0.000       0.249       0.593
SMB            0.4384      0.305      1.436      0.151      -0.160       1.037
HML           -0.0578      0.145     -0.398      0.691      -0.343       0.227
==============================================================================
Omnibus:                        0.231   Durbin-Watson:                   1.553
Prob(Omnibus):                  0.891   Jarque-Bera (JB):                0.247
Skew:                          -0.136   Prob(JB):                        0.884
Kurtosis:                       2.842   Cond. No.                         34.0
==============================================================================
```