Replication of Fama-French, Hou-Xue-Zhang Factors and Spanning Regression Table

Jiayin Wu[1]

[1]Smith School of Business, Queen's University

1 Methodology

1.1 Data Sources and Preparation

The analysis utilizes monthly stock returns from the Center for Research in Security Prices (CRSP) and accounting data from Compustat. The sample spans January 1967 to December 2024 and includes all NYSE, AMEX, and NASDAQ common stocks (CRSP share codes 10 or 11). Financial firms (SIC 6000–6999) and firms with negative book equity are excluded in q-factor construction.

CRSP return data is adjusted for delisting, and market equity is calculated as the absolute stock price multiplied by shares outstanding. December market equity values are retained for book-to-market ratio computations. Compustat fundamentals are processed to construct book equity (BE), defined as stockholders' equity plus deferred taxes minus preferred stock. Operating profitability (OP) is derived as revenue minus cost of goods sold, selling expenses, and interest, scaled by BE. Cash profitability (CP) subtracts accruals (changes in working capital) from OP. Investment (Inv) is measured as annual asset growth, and return on equity (ROE) uses quarterly income divided by lagged book equity. The CRSP/Compustat Merged (CCM) database links accounting variables to stock returns via `gvkey` (Compustat) and `PERMNO` (CRSP).

1.2 Fama-French 6 Factors

- **HML (Value)**: Independent 2x3 sorts on size (NYSE median market equity) and book-to-market (NYSE 30th/70th percentiles). The factor is the average of small and big stock spreads:

$$\text{HML} = \frac{1}{2} * [(\text{H\_S} - \text{L\_S}) + (\text{H\_B} - \text{L\_B})].$$

- **CMA (Investment)**: 2x3 sorts on size and asset growth. The factor reverses the investment spread:

$$\text{CMA} = \frac{1}{2} * [(\text{L\_S} - \text{H\_S}) + (\text{L\_B} - \text{H\_B})].$$

- **RMW (Operating Profitability)**: 2x3 sorts on size and operating profitability.

- **SMB (Size)**: SMB aggregates size premia from three independent sorts (HML, RMW, and CMA). For each sort, the size premium is calculated as:

$$\text{SMB\_BM} = \frac{1}{3} * \sum_{\text{K=L,N,H}} (\text{K\_S}) - \frac{1}{3} * \sum_{\text{K=L,N,H}} (\text{K\_B}),$$

  with analogous calculations for SMB_OP and SMB_Inv. The composite SMB is the average:

$$\text{SMB} = \frac{1}{3} * (\text{SMB\_BM} + \text{SMB\_OP} + \text{SMB\_Inv}).$$

- **UMD (Momentum)**: Sorted monthly on average 11-month returns (t-12 to t-2).

- **MKT-RF**: Value-weighted market return minus the 1-month Treasury bill rate.

1.3 Hou-Xue-Zhang q-Factors

The q-factors are constructed through independent triple sorting on size, investment-to-assets (I/A), and return on equity (ROE). Portfolios are formed annually in June for size and I/A, and monthly for ROE. Size portfolio groups using NYSE median market equity. I/A and Roe groups using the NYSE breakpoints for the low 30%, middle 40%, and high 30% of the ranked values. Financial firms and firms with negative book equity are excluded. The 2 by 3 by 3 sorting generates 18 portfolios, with value-weighted returns.

- **Calculation**:

$$R_{\text{ME},t} = \frac{1}{9}\left(\sum_{j=1}^{3}\sum_{k=1}^{3} R_{Sjk,t}\right) - \frac{1}{9}\left(\sum_{j=1}^{3}\sum_{k=1}^{3} R_{Bjk,t}\right),$$

  where $R_{Sjk,t}$ and $R_{Bjk,t}$ are value-weighted returns of small and big portfolios in month t.

$$R_{\text{I/A},t} = \frac{1}{6}\left(\sum_{i=1}^{2}\sum_{k=1}^{3} R_{iLk,t}\right) - \frac{1}{6}\left(\sum_{i=1}^{2}\sum_{k=1}^{3} R_{iHk,t}\right).$$

$$R_{\text{Roe},t} = \frac{1}{6}\left(\sum_{i=1}^{2}\sum_{j=1}^{3} R_{ijH,t}\right) - \frac{1}{6}\left(\sum_{i=1}^{2}\sum_{j=1}^{3} R_{ijL,t}\right).$$

2 Replication of Spanning Regression Table

2.1 Methodology

The spanning regression table evaluates the incremental explanatory power of the q-factors relative to the Fama-French 6-factor model (FF6) and vice versa. The table has three columns: (1) Average Returns, (2) 6-Factor Alphas, and (3) q-Factor Alphas. Below I detail the methodology for each column:

- **Column 1 Average Returns:**
  For each factor ($R_{\text{I/A}}$, R_{Roe}, HML, CMA, RMW, UMD), the average monthly return is calculated using a time-series regression of the factor's returns on a constant:

  $$R_t = \gamma + \epsilon_t,$$

  where γ represents the average return. Newey-West standard errors adjust for autocorrelation and heteroskedasticity. The t-statistic tests whether γ is significantly different from zero.

- **Column 2 6-Factor Alphas:**
  The q-factors ($R_{\text{I/A}}$, R_{Roe}) are regressed on the FF6 factors (MKT, HML, SMB, CMA, RMW, UMD):

  $$R_{\text{q},t} = \alpha + \beta_1 MKT_t + \beta_2 HML_t + \beta_3 SMB_t + \beta_4 CMA_t + \beta_5 RMW_t + \beta_6 UMD_t + \epsilon_t.$$

  The intercept α measures abnormal returns unexplained by the FF6 model.

- **Column 3 q-Factor Alphas**
  Conversely, FF6 factors (HML, CMA, RMW, UMD) are regressed on the q-factors (R_{MKT}, R_{ME}, $R_{\text{I/A}}$, R_{Roe}):

  $$R_{\text{FF6},t} = \alpha + \beta_1 R_{\text{MKT},t} + \beta_2 R_{\text{ME},t} + \beta_3 R_{\text{I/A},t} + \beta_4 R_{\text{Roe},t} + \epsilon_t.$$

  The intercept α measures abnormal returns unexplained by the q-factor model.

TABLE I: 1/1967–12/2024: The q-alphas of HML, CMA, RMW, and UMD = 0 ($p = 0.34$); the 6-factor alphas of $R_{I/A}$ and $R_{\text{Roe}} = 0$ ($p = 0.00$)

	Average returns	**6-factor alphas**	**q-factor alphas**
The investment factor, $R_{I/A}$	0.34	0.05	
	(3.50)	(1.41)	
The Roe factor, R_{Roe}	0.57	0.25	
	(5.41)	(4.09)	
HML	0.26		-0.02
	(1.73)		(-0.18)
CMA	0.26		0.04
	(2.88)		(1.37)
RMW	0.33		0.01
	(3.33)		(0.13)
UMD	0.54		0.22
	(3.24)		(1.04)

2.2 Interpretation

Table I underscores the distinct explanatory power of the q-factor model relative to the Fama-French 6-factor. The results reveal that the q-factors systematically account for return anomalies that elude the FF6 model, while the reverse does not hold. This asymmetry highlights the q-model's broader economic intuition.

Three findings stand out. First, the significant 6-factor alphas for $R_{I/A}$ ($\alpha = 0.05$, $t = 1.41$) and R_{Roe} ($\alpha = 0.25$, $t = 4.09$) indicate that the q-factors capture persistent anomalies unexplained by size, value, or momentum. Second, the inability of the FF6 model to explain q-factors (GRS test $p = 0.00$) suggests structural limitations in its specification. Third, the q-model's ability to subsume FF6 factors (GRS test $p = 0.34$)—particularly the near-zero alphas for HML ($\alpha = -0.02$) and RMW ($\alpha = 0.01$)—implies that value and profitability effects are better articulated through I/A and Roe spreads.

3 Interpretation of Risk-Adjusted Performance

Table II reports the annualized returns, volatilities, and Sharpe ratios for both the Fama-French and q-factors, revealing three key patterns in factor performance:

First, the Hou-Xue-Zhang q-factors demonstrate superior risk-adjusted returns compared to traditional Fama-French factors. The profitability factor R_{ROE} achieves the highest Sharpe

TABLE II: Sharpe Ratios of Factors

Factor	Annualized Return (%)	Annualized Volatility (%)	Sharpe Ratio
MKT-RF	7.51	15.78	0.48
HML	3.07	10.85	0.28
SMB_O	1.72	10.65	0.16
RMW_O	3.97	7.85	0.51
CMA	3.16	7.01	0.45
UMD	6.48	14.21	0.46
R_ME	2.64	10.82	0.24
R_IA	4.09	7.43	0.55
R_ROE	6.85	9.12	0.75

ratio (0.75), delivering 6.85% annualized return with moderate volatility (9.12%). This outperforms even the market factor (MKT-RF, Sharpe Ratio = 0.48), suggesting profitability measures better compensate for risk than broad market exposure.

Second, the investment factor $R_{\mathrm{I/A}}$ shows remarkable efficiency with a Sharpe ratio of 0.55 - higher than all Fama-French factors. This aligns with the spanning regression results, where $R_{\mathrm{I/A}}$ exhibited significant alpha in the 6-factor model, confirming its unique pricing information.

Third, traditional value (HML, Sharpe Ratio = 0.28) and size (SMB_O, Sharpe Ratio = 0.16) factors show the weakest performance. Their low Sharpe ratios - less than half of R_{ROE}'s - suggest these anomalies have decayed or become less economically relevant in the modern period. The momentum factor (UMD) remains competitive (Sharpe Ratio = 0.46).

These results complement the spanning regression analysis: factors with higher Sharpe ratios (R_{ROE}, $R_{\mathrm{I/A}}$) are precisely those that generate significant alphas in competing models, while low-Sharpe factors (HML, SMB_O) become statistically redundant in the q-factor framework.

References

Ball, R., Gerakos, J., Linnainmaa, J. T., and Nikolaev, V. (2016). Accruals, cash flows, and operating profitability in the cross section of stock returns. *Journal of Financial Economics*, 121(1):28–45.

Ball, R., Gerakos, J., Linnainmaa, J. T., and Nikolaev, V. V. (2015). Deflating profitability. *Journal of Financial Economics*, 117(2):225–248.

Eugene, F. and French, K. (1992). The cross-section of expected stock returns. *Journal of finance*, 47(2):427–465.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1):3–56.

Fama, E. F. and French, K. R. (1995). Size and book-to-market factors in earnings and returns. *The journal of finance*, 50(1):131–155.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of financial economics*, 116(1):1–22.

Fama, E. F. and French, K. R. (2018). Choosing factors. *Journal of financial economics*, 128(2):234–252.

Gibbons, M. R., Ross, S. A., and Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica: Journal of the Econometric Society*, pages 1121–1152.

Hou, K., Mo, H., Xue, C., and Zhang, L. (2019). Which factors? *Review of Finance*, 23(1):1–35.

Hou, K., Xue, C., and Zhang, L. (2015). Digesting anomalies: An investment approach. *The Review of Financial Studies*, 28(3):650–705.

Hou, K., Xue, C., and Zhang, L. (2020). Replicating anomalies. *The Review of financial studies*, 33(5):2019–2133.

Hou, K., Xue, C., and Zhang, L. (2024). Technical document: Testing portfolios.

Zhang, L. (2019). Q-factors and investment capm. Technical report, National Bureau of Economic Research.