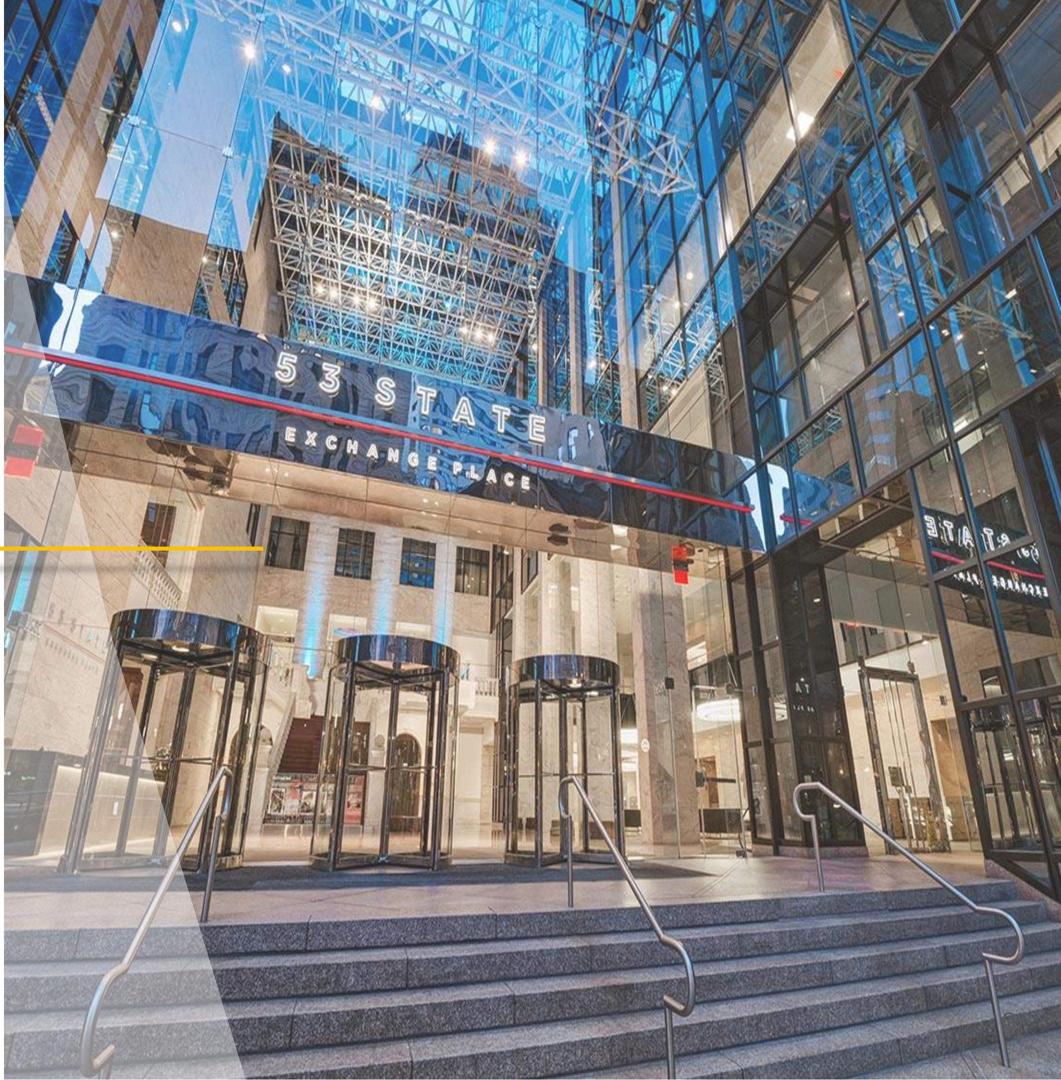
2020

Quality Minus Junk

Advisor: Philippe Wells
Presenters: Zhenyu Chen, Aileen Kate Fu, Sirui Jiang, Keyu Jin, Salma Mohammed



Executive Summary

STATE STREET

1 Stocks' quality measures are time persistent, meaning QMJ (long quality stocks and short junk stocks) is still a potential investment strategy that may generate profit.

2 Quality measures can positively affect stock prices; profitability and growth are the most important characteristics.

3 High quality stocks mostly earn higher average excess returns and alphas than low quality stocks.

4 Firms that are high quality in one quality measure doesn't necessarily to be also high quality in other quality measures.

5 When controlling for the market, size and value factors, the QMJ still have a positive excess return.

6 QMJ has explanatory power on standard factors like SMB, HML, and UMD; when controlled for quality, size effect and value effect appear to be stronger.

Glossary

Quality Minus Junk (QMJ)

The investment strategy that long quality stocks and short junk stocks.

Quality Measures

Quality measures include profitability, growth, safety and payout scores that are calculated from stock fundamentals and market information to quantify stocks' quality characteristics (see **Appendix 2** for detailed calculation).

Quality Score

An aggregated measure of quality characteristics which is the average of profitability, growth, safety, and payout scores.

Standard Factors

Fama-French factors, including MKT, SMB, HML and UMD.

Table of Contents

Introduction

Chiffords S. Asness, Andrea Frazzini and Lasse H. Pedersen, in their paper *Quality Minus Junk* (2013), discussed stock prices, investment strategies and their connections with quality measures and standard factors from 1956 to 2012 across countries. In the paper, authors concluded quality minus junk as a persistent investment strategy.

Adopted *Quality Minus Junk*'s methodologies, this project used the most up-to-date US data from 1995 to 2019 and tried to find **whether the paper's results still hold in a shorter and more recent time frame**, and to provide coding demos to anyone who is interested in this paper.

This report summarized our steps, assumptions, results and conclusions throughout the project. Specifically, we aimed to replicate paper's tables. Due to data insufficiency, we are not able to copy Table 1, 7, and 8. We will also discuss where our results deviate from the paper's results and list our preliminary explanations. Important codes are attached to provide with more detailed technical supports.

Data

1 Source (**Appendix 1**)

- S&P 500
- Jan 1995 – Dec 2019
- Stock Price: end of fiscal year price
- Stock Fundamentals: annual reports
- Market Return: monthly value-weighted return including distributions
- Risk-free Return: monthly 10-year T-bill
- Dataset: CRSP/Compustat Merged

2 Pre-Processing

Used stocks’ fundamental and price data to:

- Construct yearly quality measures and quality score for each stock (**Appendix 2**)
- Construct quality portfolios and calculate Fama-French standard factors (**Appendix 3**)

Table 2 – Procedures
Persistence of Quality Measures

1 Sort Stocks by Quality Measures (table2.R)

- Rank stocks by their quality measures and assign them, from low to high, into ten portfolios: P1 to P10
- The return for each portfolio is value weighted

2 Conduct Time Series Analysis

- Calculate each portfolio's quality measures at time 0 (2000), time 0 + 12M (2001), time 0 + 36M (2003), time 0 + 60M (2005) and time 0 + 120M (2010)
- Calculate QMJ returns (P10 – P1) by directly subtract return of P1 from that of P10
- Results are saved to table2.csv

Table 2 – Results

Persistence of Quality Measures

Paper

Panel A: Long Sample U.S., 1956 - 2012		P1 (Low)	P2	P3	P4	P5	P6	P7	P8	P9	P10 (High)	P10 - P1
Quality	t	-1.38	-0.71	-0.39	-0.15	0.05	0.25	0.46	0.69	1.00	1.56	**2.94**
Quality	t + 12M	-0.60	-0.29	-0.14	0.00	0.14	0.29	0.45	0.63	0.86	1.31	**1.92**
Quality	t + 36M	-0.33	-0.12	-0.05	0.05	0.15	0.27	0.40	0.54	0.74	1.16	**1.49**
Quality	t + 60M	-0.16	-0.02	0.04	0.09	0.16	0.22	0.35	0.46	0.68	1.04	**1.20**
Quality	t + 120M	-0.09	0.00	0.03	0.07	0.09	0.21	0.30	0.38	0.62	0.89	**0.98**
Profit	t + 120M	-0.37	-0.19	-0.10	0.05	0.12	0.18	0.29	0.35	0.59	1.08	**1.44**
Growth	t + 120M	-0.23	-0.19	-0.13	-0.12	-0.10	-0.12	-0.02	0.11	0.11	0.34	**0.57**
Safety	t + 120M	-0.28	-0.15	-0.03	0.08	0.15	0.21	0.35	0.49	0.63	0.67	**0.95**
Payout	t + 120M	0.12	0.29	0.28	0.29	0.38	0.39	0.49	0.49	0.56	0.61	**0.49**

Our Results

Panel A: Long Sample U.S., 1995-2019		P1 (Low)	P2	P3	P4	P5	P6	P7	P8	P9	P10 (High)	P10-P1
Quality	t	-0.35	-0.16	0.01	0.11	0.14	0.16	0.16	0.27	0.35	0.76	1.11
Quality	t + 12M	-0.27	-0.16	-0.02	-0.02	0.00	0.06	0.11	0.14	0.17	0.33	0.59
Quality	t + 36M	-0.33	-0.14	-0.10	-0.06	-0.01	0.08	0.12	0.16	0.20	0.38	0.71
Quality	t + 60M	-0.37	-0.18	-0.10	-0.06	-0.02	0.05	0.10	0.17	0.25	0.54	0.91
Quality	t + 120M	-0.36	-0.19	-0.13	-0.09	-0.05	0.00	0.07	0.14	0.24	0.63	0.99
Profit	t + 120M	-0.66	-0.24	-0.21	-0.08	0.02	0.13	0.34	0.36	0.54	0.77	1.44
Growth	t + 120M	-0.23	-0.19	-0.17	-0.16	-0.04	-0.13	-0.07	0.12	0.32	1.40	1.63
Safety	t + 120M	-0.35	-0.34	-0.20	-0.06	-0.12	-0.07	-0.02	0.04	0.13	0.28	0.62
Payout	t + 120M	-0.20	0.01	0.05	-0.08	-0.04	0.08	0.03	0.03	-0.03	0.09	0.28

Table 2 – Analysis

Persistence of Quality Measures

Paper's results **Stock's quality is a persistent characteristic**

Stocks that were profitable, growing, safe and well managed currently are likely to keep these characteristics in the future. This makes quality-minus-junk (QMJ) strategy plausible as the strategy requires quality stocks stay quality and junk stocks stay junk during the investment period.

Our results **Stock's quality measures and quality score are persistent**

- Quality measures and quality scores are persistent for each portfolio throughout the period
- Quality portfolios have higher quality scores

Table 3 – Procedures

The Price of Quality

1 Data Preparation (data_table3.R)

- Retrieve quality scores calculated in section 1 – Quality.csv
- Import extra fundamental data, including stock price, common shares outstanding, shareholder's equity from SP500_Safety1.csv
- Calculate market to book ratio, market to equity ratio with fundamental data
- For each stock, form a panel dataset of quality scores, stock price and calculated ratios and record them to data_table3.csv

2 Run Regressions (table3.R)

- With all data from data_table3.csv, run linear time-series regressions
- Panel A (1), standardized market to book ratio over quality score
- Panel A (2), standardized market to book ratio over quality score, size and lagged return
- Panel B (1)-(4), standardized market to book ratio over respectively profitability, growth, safety and payout scores
- Results saved to table3.csv

Note: We ignored regressions with Industry FE and Country FE because they are not specifically explained in paper, but the code is written in a way that can be easily modified once they are available.

Table 3 – Results
The Price of Quality

Paper (Panel A)

Long Sample (U.S., 1956 – 2012)

	(1)	(2)
Quality	**0.32** (22.47)	**0.19** (15.94)
Size		**0.31** (19.19)
Ret(t-12,t)		**0.27** (21.36)
Industry FE Country FE	No	No
Average R2	0.12	0.31

Our Results

	Long Sample (U.S., 1995-2019)	
	(1)	**(2)**
Quality	**0.12** (3.562)	**0.11** (3.424)
Size		0.01 (1.226)
Ret(t-12, t)		0.001 (0.747)
Industry FE Country FE	No	No
Average R2	0.002	0.003

Table 3 – Results, Cont.

The Price of Quality

Paper (Panel B)

	Long Sample (U.S., 1956 – 2012)			
	(1)	(2)	(3)	(4)
Profitability	**0.41** (26.19)			
Growth		**0.38** (31.18)		
Safety			**0.14** (9.95)	
Payout				**-0.10** -(11.11)
Industry FE Country FE	No	No	No	No
Average R2	0.18	0.15	0.03	0.01

Our Results

	Long Sample (U.S., 1995-2019)			
	(1)	**(2)**	**(3)**	**(4)**
Profitability	**0.19** (10.189)			
Growth		0.001 (0.658)		
Safety			-0.08 (-2.966)	
Payout				-0.01 (-0.715)
Industry FE Country FE	No	No	No	No
Average R2	0.02	0	0.002	0

Table 3 – Analysis
The Price of Quality

Paper's results **Stock price is positively related to quality score**

The paper suggests that quality score can be used to predict stock price (measured by market to book ratio). Specifically, stock price is mainly determined by profitability and growth scores. Safety score is less important and payout score has negative impact on stock price.

Our results **Similar results but with lower explanatory power**

- Quality, profitability and growth scores positively affect stock price; safety and payout scores negatively affect stock price.
- Opposite result for safety score, but we believe this is acceptable because explanatory power was relatively low in paper itself (low R2 in Panel B, regression 3).
- **Our replicated regressions' R^2 are low**

  <u>Hypothesis:</u> since our data has a relatively shorter time frame and includes two large financial crises, it is hard to find clear linear relationships between stock prices and quality scores. This might be resolved if we expand data to a longer time frame.

Table 4 – Procedures

Quality Sorted Portfolio

1 Data Preparation (3_table4.R)

- Retrieve portfolio returns from Table 2; quality scores from data_table3.csv; 10-year T-Bill (Risk-free) from DJ_Safety3.csv; market return from DJ30_Market.csv
- Calculate company's market value by multiplying common shares outstanding by stock price

2 Calculate Excess Return and Standard Factors

- Excess Return: calculate the excess return of each portfolio by subtracting annualized risk-free return from the value-weighted portfolio returns
- Use all available data to construct MKT, SMB, HML and UMD factors based on the methodology described in the paper (see Appendix 3 for detailed calculation methods)
- Results saved to 2_data_table4.csv

Table 4 – Procedures, Cont.

Quality Sorted Portfolio

3 For Each Portfolio, Run Regressions

- <u>CAPM:</u> Run linear regression on portfolio excess return over the year based on the CAPM model (excess return ~ MKT) and save the coefficients and p-value
- <u>Fama-French 3-Factor:</u> Run linear regression on portfolio excess return over the years based on Fama and French three-factor models (excess return ~ MKT+SMB+HML) and save the coefficients and p-value
- <u>Fama-French 4-Factor:</u> Run linear regression on portfolio excess return over the years based on Fama and French four-factor models (excess return ~ MKT+SMB+HML+UMD) and save the coefficients and p-value
- Beta, Sharpe Ratio, information ratio and adjusted R-Squared are also calculated.
- Results are saved to table4.csv.

Table 4 – Results

Quality Sorted Portfolio

Paper

Panel A: Long Sample U.S. , 1956 - 2012	P1 (Low)	P2	P3	P4	P5	P6	P7	P8	P9	P10 (High)	H-L
Excess return	0.15	0.36	0.38	**0.39**	**0.45**	**0.45**	**0.57**	**0.47**	**0.58**	**0.61**	**0.47**
	(0.55)	(1.56)	(1.90)	(2.04)	(2.51)	(2.60)	(3.42)	(2.75)	(3.48)	(3.68)	(2.80)
CAPM alpha	**-0.53**	**-0.24**	**-0.15**	**-0.12**	-0.02	-0.01	**0.13**	0.01	**0.14**	**0.18**	**0.71**
	(-4.62)	(-2.85)	(-2.25)	(-2.01)	(-0.33)	(-0.18)	(2.41)	(0.23)	(2.71)	(2.86)	(4.92)
3-factor alpha	**-0.67**	**-0.38**	**-0.25**	**-0.21**	-0.08	-0.06	**0.12**	0.01	**0.16**	**0.29**	**0.97**
	(-7.83)	(-5.47)	(-4.47)	(-4.11)	(-1.44)	(-1.09)	(2.26)	(0.12)	(3.37)	(5.24)	(9.02)
4-factor alpha	**-0.56**	**-0.42**	**-0.26**	**-0.29**	**-0.14**	**-0.12**	0.04	-0.05	**0.19**	**0.41**	**0.97**
	(-6.24)	(-5.73)	(-4.26)	(-5.39)	(-2.37)	(-2.22)	(0.68)	(-1.08)	(3.62)	(7.10)	(8.55)
Beta	1.28	1.22	1.08	1.09	1.03	1.01	0.97	1.00	0.95	0.90	-0.38
Sharpe Ratio	0.07	0.21	0.25	0.27	0.33	0.35	0.46	0.37	0.46	0.49	0.37
Information Ratio	-0.90	-0.82	-0.61	-0.77	-0.34	-0.32	0.10	-0.15	0.52	1.02	1.23
Adjusted R2	0.90	0.91	0.92	0.93	0.90	0.91	0.91	0.93	0.92	0.90	0.60

Our Results

U.S., 1995-2019	P1 (Low)	P2	P3	P4	P5	P6	P7	P8	P9	P10 (High)	H-L
Excess return	-0.037	0.008	0.043	0.050	0.070	0.071	0.106	0.064	0.060	0.772	0.769
CAPM alpha	-0.063	-0.010	0.023	0.034	0.050	0.050	0.083	0.037	0.045	0.719	0.746
CAPM alpha p-value	0.006	0.545	0.290	0.180	0.004	0.034	0.011	0.004	0.017	0.285	0.238
3-factor alpha	-0.069	-0.033	0.010	0.021	0.034	0.023	0.069	0.027	0.049	0.291	0.360
3-factor alpha p-value	0.015	0.138	0.738	0.536	0.084	0.123	0.013	0.065	0.051	0.040	0.020
4-factor alpha	0.068	0.035	0.162	0.170	0.059	0.088	-0.003	0.004	0.073	-1.008	-1.076
4-factor alpha p-value	0.327	0.588	0.051	0.087	0.314	0.040	0.972	0.940	0.316	0.000	0.000
Beta	1.002	0.959	1.041	0.846	0.829	0.847	0.929	0.823	0.788	2.109	1.190
Sharpe ratio	-0.155	0.035	0.173	0.234	0.358	0.337	0.438	0.330	0.319	0.278	0.289
Information ratio	0.999	0.546	2.101	1.808	1.029	2.222	-0.035	0.085	1.024	-6.366	-8.081
Adjusted R2	0.896	0.898	0.878	0.755	0.889	0.955	0.883	0.939	0.814	0.996	0.997

STATE STREET

Table 4 – Analysis
Quality Sorted Portfolio

Paper's results

- High quality stocks earn higher average returns than low quality stocks (reject the null hypothesis of no difference in their average returns)
- When controlling for market risk and other factor exposures, the alpha of high-quality stocks is larger, since high-quality stocks have lower market and factor exposures than low-quality stocks
- QMJ portfolio earns average abnormal returns ranging from 71 to 97 basis points per month

Paper's results

- Consistent with the paper's results that high-quality stocks mostly earn higher average excess returns and alphas than low quality stocks

STATE STREET.

Table 5 – Procedures

Quality Minus Junk: Correlations

1 Data Preparation (table5.R)

- Retrieve quality scores from data_table3.csv; quality factors from section 4's output (1_mkt_excessret.csv, 1_smb_hml.csv, 1_umd.csv); Risk-free returns from SP500_Safety3.csv; common shares outstanding from SP500_Safety1.csv
- Calculate company's market value by multiplying common shares outstanding by stock price
- Excess returns are calculated by subtracting annualized risk-free returns from annual stock returns
- Abnormal returns are calculated as the sum of the alpha and residuals of the Fama-Franch four-factor model for each stock

Table 5 – Procedures, Cont.

Quality Minus Junk: Correlations

2 QMJ Factors Formulation

- Rank all stocks by size and sort them into 2 portfolios – small and big (follows the SMB construction method)
- Then for each size portfolio, rank stocks by each quality measures (quality, profitability, growth, safety and payout) and sort them into 3 quality portfolios -quality, middle and junk (the sorting criteria for QMJ is quality score)
- QMJ returns are calculated by 1/2(small quality + big quality) - 1/2(small junk + big junk) with stock returns; for the upper part of Panel A, use stocks' excess returns; for the lower part of Panel A, use stocks' abnormal returns

3 Create correlation matrix

- Use the results generated from step 2 to create correlation matrix
- Results are saved to table5_return_matrix.csv and table5_abnormalReturn_matrix.csv

Table 5 – Results

Quality Minus Junk: Correlations

Panel A: Long Sample (U.S. , 1956 - 2012)	QMJ	Profitability	Safety	Growth	Payout
			Returns		
QMJ	1.00				
Profitability	0.82	1.00			
Safety	0.88	0.64	1.00		
Growth	0.24	0.52	0.15	1.00	
Payout	0.69	0.35	0.53	-0.34	1.00
			Abnormal Returns (4-factor)		
QMJ	1.00				
Profitability	0.82	1.00			
Safety	0.72	0.43	1.00		
Growth	0.42	0.49	0.18	1.00	
Payout	0.62	0.44	0.30	-0.06	1.00

Paper

Panel A: Long Sample (U.S., 1995-2019)	QMJ	Profitability	Safety	Growth	Payout
			Returns		
QMJ	1.00				
Profitability	1.00	1.00			
Safety	0.14	0.12	1.00		
Growth	1.00	1.00	0.11	1.00	
Payout	1.00	1.00	0.11	1.00	1.00
			Abnormal Returns (4-factors)		
QMJ	1.00				
Profitability	0.45	1.00			
Safety	-0.03	-0.07	1.00		
Growth	0.69	0.40	0.32	1.00	
Payout	0.96	0.48	-0.15	0.63	1.00

Our Result

Table 5 – Analysis

Quality Minus Junk: Correlations

Paper’s results

- All of the pairwise correlations among the quality components are positive, except the correlation between growth and payout. The negative correlation reflects that higher payout is naturally associated with lower growth.

- The average pairwise correlation among the quality components is 0.40 in the US and 0.38 for abnormal returns.
- Firms that are high quality in one quality measure tend to also be high quality in other quality measures.

Our results

- Firms that are high quality in one quality measure doesn’t necessarily to be also high quality in other quality measures. It is fine that the quality measures are not related to one another. Though having a few negative correlations, most of the correlations in our result table are still positive.

Table 6 – Procedures

Quality Minus Junk: Returns

1 Data Preparation

- Merge QMJ factor data (data_table5_cleaned.csv) with other Fama-French standard factors created for Table 4 (2_data_table4.csv)
- Have 5 factors (MKT, SMB, HML, UMD, and QMJ) for each year

2 Run Regression

- Take QMJ, Profitability, Safety, Growth, Payout each as the dependent variable and run regression with respect to the 1-, 3-, and 4-factor models
- i.e. for QMJ 1-factor model, QMJ ~ MKT;
  for QMJ 3-factor model, QMJ ~ MKT+SMB+HML;
  for QMJ 4-factor model, QMJ ~ MKT+SMB+HML+UMD

3 Complete the Remaining Parts of Table 6

- Calculate excess return as the average returns for each target factor over the time length of the data
- Calculate Alpha as the intercept of the regression
- Calculate Sharpe Ratio, Information Ratio, and Adjusted R^2 by definition
- Results are saved to table6.csv

STATE STREET

Table 6 – Results

Quality Minus Junk: Returns

	Panel A: Long Sample (U.S. , 1956 - 2012)				
	QMJ	Profitability	Safety	Growth	Payout
Excess Returns	**0.40**	**0.27**	**0.23**	0.12	**0.31**
	(4.38)	(3.81)	(2.06)	(1.63)	(3.37)
CAPM-alpha	**0.55**	**0.33**	**0.42**	0.08	**0.46**
	(7.27)	(4.78)	(4.76)	(1.06)	(6.10)
3-factor alpha	**0.68**	**0.45**	**0.59**	**0.20**	**0.43**
	(11.10)	(7.82)	(8.68)	(3.32)	(6.86)
4-factor alpha	**0.66**	**0.53**	**0.57**	**0.38**	**0.21**
	(10.20)	(8.71)	(7.97)	(6.13)	(3.43)
MKT	**-0.25**	**-0.11**	**-0.34**	**0.05**	**-0.20**
	(-17.02)	(-8.08)	(-20.77)	(3.35)	(-14.47)
SMB	**-0.38**	**-0.21**	**-0.41**	**-0.05**	**-0.30**
	(-17.50)	(-10.21)	(-17.00)	(-2.53)	(-14.82)
HML	**-0.12**	**-0.28**	**-0.23**	**-0.44**	**0.39**
	(-5.03)	(-12.16)	(-8.50)	(-18.81)	(16.68)
UMD	0.02	**-0.07**	0.01	**-0.17**	**0.21**
	(0.82)	(-3.80)	(0.64)	(-8.55)	(10.79)
Sharpe Ratio	0.58	0.51	0.27	0.22	0.45
Information Ratio	1.46	1.25	1.14	0.88	0.49
Adjusted R2	0.57	0.37	0.63	0.40	0.60

Paper

	Panel A: Long Sample (U.S., 1995-2019)				
	QMJ	Profitability	Safety	Growth	Payout
Excess Returns	0.30	0.00	0.07	0.24	0.25
CAPM-alpha	**0.29**	**0.00**	**0.07**	0.23	0.24
3-factor alpha	**0.15**	**0.00**	**0.08**	0.06	0.09
4-factor alpha	**-0.52**	**-0.10**	**-0.05**	-0.47	-0.44
MKT	0.02	-0.20	0.17	-0.07	0.03
SMB	-1.08	**-0.29**	-0.35	-0.70	-0.72
HML	0.49	-0.05	0.11	0.37	0.47
UMD	1.02	**0.15**	0.21	0.81	0.82
Sharpe Ratio	0.30	-0.03	0.64	0.24	0.25
Information Ratio	-4.80	-1.75	-0.60	-3.93	-4.55
Adjusted R2	0.99	0.49	0.18	0.98	0.99

Our Results

Table 6 – Analysis
Quality Minus Junk: Returns

Paper's results

- Each quality factor delivers a statistically significant positive excess returns and alphas with respect to the 1-, 3- and 4-factor models
- QMJ factor performs best in all factor models
- Safety has the most market exposure (i.e. largest negative MKT value)
- Value factor (HML) is negative in QMJ

Our Results

- In terms of the 4-factor model, each quality factor delivers a negative alpha. Moreover, profitability factor delivers a near to zero excess return in 1- and 3- factor models
- QMJ factors performs best in 1- and 3- factor models, however, performs worst in terms of 4-factor model
- Profitability has the most market exposure (i.e. largest negative MKT value)
- Value factor (HML) is positive in all quality factors, except for profitability

Note: More analysis explaining about the possible reasons for the difference in Table 6 can be found in the conclusion and limitation part.

Table 9 – Procedures (table9.R)
Quality Minus Junk: Returns

1 Data Preparation

- Merge QMJ factor data (data_table5_cleaned.csv) with other Fama-French standard factors created for Table 4 (2_data_table4.csv)
- Have 5 factors (MKT, SMB, HML, UMD, and QMJ) for each year

2 Run Regression

- Take SMB, HML, and UMD each as the dependent variable and run regressions against remaining standard factors, with and without QMJ factors
- i.e. for SMB, SMB ~ MKT+HML+UMD (+ QMJ)

3 Complete the Remaining Parts of Table 9

- Calculate excess return as the average returns for each target factor over the time length of the data
- Calculate Alpha as the intercept of the regression
- Calculate Sharpe Ratio, Information Ratio, and Adjusted R^2 by definition
- Results are saved to table9.csv

Table 9 – Results
Pricing HML, SMB, and UMD

	Panel A: Long Sample (U.S. , 1956 - 2012)					
Left-hand side	SMB	SMB	HML	HML	UMD	UMD
Excess Returns	**0.28**	**0.28**	**0.34**	**0.34**	**0.70**	**0.70**
	(2.54)	(2.54)	(2.66)	(2.66)	(4.52)	(4.52)
Alpha	0.13	**0.64**	**0.77**	**0.94**	**1.05**	**1.01**
	(1.16)	(6.39)	(8.01)	(9.35)	(9.11)	(8.05)
MKT	**0.19**	**-0.08**	**-0.16**	**-0.23**	**-0.20**	**-0.18**
	(7.38)	(-3.06)	(-7.04)	(-8.74)	(-7.39)	(-5.61)
SMB			**0.08**	-0.03	0.04	0.07
			(2.34)	(-0.86)	(1.05)	(1.33)
HML	**0.10**	-0.03			**-0.81**	**-0.80**
	(2.34)	(-0.86)			(-23.24)	(-22.10)
UMD	0.04	0.04	**-0.55**	**-0.53**		
	(1.05)	(1.33)	(-23.24)	(-22.10)		
QMJ		**-0.83**		**-0.29**		0.06
		(-17.50)		(-5.03)		(0.82)
Sharpe Ratio	0.34	0.34	0.35	0.35	0.60	0.60
Information Ratio	0.17	0.96	1.10	1.36	1.23	1.18
Adjusted R2	0.07	0.36	0.45	0.47	0.46	0.46

Paper

	Panel A: Long Sample (U.S. , 1995 - 2019)					
Left-hand side	SMB	SMB	HML	HML	UMD	UMD
Excess Returns	0.37	0.37	-0.22	-0.22	1.30	1.30
Alpha	**-0.36**	**-0.42**	-0.27	**0.61**	**0.65**	**0.54**
MKT	0.14	0.08	-0.42	-0.17	-0.24	-0.07
SMB			-0.13	**1.41**	**1.77**	**1.22**
HML	-0.02	**0.23**			0.03	**-0.36**
UMD	**0.56**	**0.77**	0.08	**-1.35**		
QMJ		**-0.49**		**1.35**		**0.74**
Sharpe Ratio	0.28	0.28	-0.69	-0.69	0.55	0.55
Information Ratio	-3.43	-5.84	-0.87	3.41	3.47	5.88
Adjusted R2	0.99	1.00	-0.09	0.61	0.99	1.00

Our Results

Table 9 – Analysis
Pricing HML, SMB, and UMD

1 Result Comparison

Factors		Paper Reported	Our Result
SMB	Excess Return	Modest but significant	Modest
	QMJ	Large and negative	Modest and negative
HML	Alpha	Controlling for QMJ significantly increases alpha	Controlling for QMJ significantly increases alpha
	QMJ	Negative	Positive
UMD	Alpha	Controlling for QMJ lowers alpha, but still significant	Controlling for QMJ lowers alpha, but still significant
	QMJ	Positive	Positive

2 Comments on QMJ Factor

- The paper reported:
  - SMB: **negative** loading on QMJ as small stocks are junky relative to big stocks
  - HML: **negative** loading on QMJ as cheap stocks usually have lower qualities; controlling for QMJ increases the alpha of HML, strengthening the value effect
  - UMD: **positive** loading on QMJ as high return stocks generally have higher qualities
- Our Result:
  - SML: Corresponds with the paper and have practical sense
  - HML: Opposite to the paper, partly due to the time length difference of the data
  - UMD: Corresponds with the paper and have practical sense

Limitations and Conclusions

As discussed, the results from our data are **mostly consistent** with *Quality Minus Junk*'s results. Due to our time and resource constraints, our project has some limitations which can be improved in the future.

Insufficient Number of Stocks We do not have access to cloud services, thus only S&P500 stocks are used in this project to expedite our computation. Moreover, when calculating the size factor (SMB), small (last 20%) and big (top 20%) stocks only account for 40% in total of the 500 stocks, making our dataset even smaller. We believe S&P500 data is sufficient to reflect the market trend, but it is not perfect when compared to over 3000 stocks used in the paper.

Short Time Frame Due to the same cloud computing resource limitation, we focused on a shorter period of 25 years (1995-2019). The paper's 56-year time frame has in general an uptrend in economy, while our shorter time frame consists of two major financial crisis, possibly leading to biased regression results.

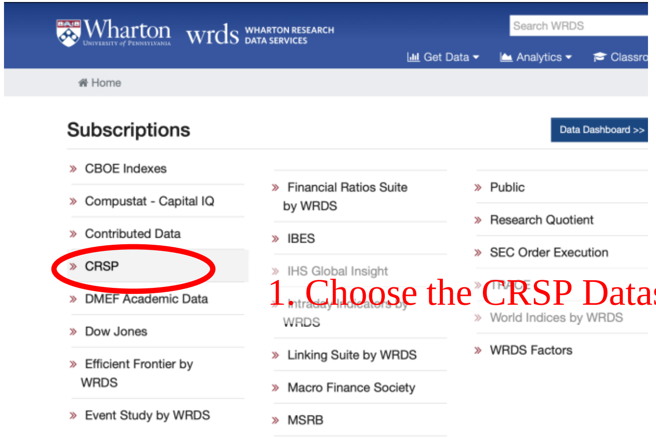
S&P500 Stock List Change S&P500 as of April 1st, 2020 is used for our stock selection. With the company list constantly changing in the real world, we are not able to adjust the changes every year with our current capabilities.

Units of Our Data Not Confirmed Units are not specified in our downloaded data. We assume data units are consistent when downloading from a same database, but there may still be inconsistencies that would cause wrong results.

Reference

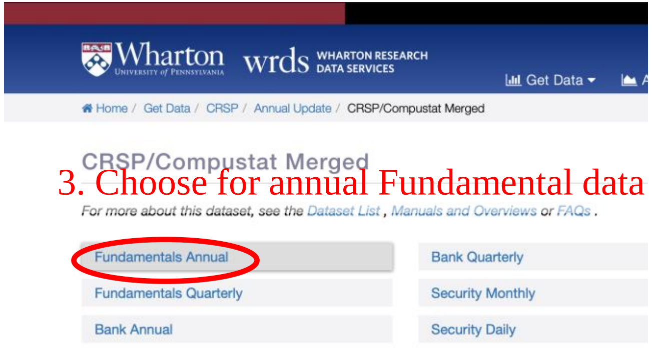
- Asness, Clifford S., Andrea Frazzini, and Lasse Heje Pedersen. "Quality minus junk." *Review of Accounting Studies*24.1 (2019): 34-112.
- Fama, Eugene F., and Kenneth R. French. “Common risk factors in the returns on stocks and bonds.” *Journal of Economics* (1993).
- Fischer, Serverin. “Quality Minus Junk in the Swiss Stock Market.” University of Zurich, Bachelor Thesis in Banking and Finance (2014).

Appendix 1 – Instruction to Download Data

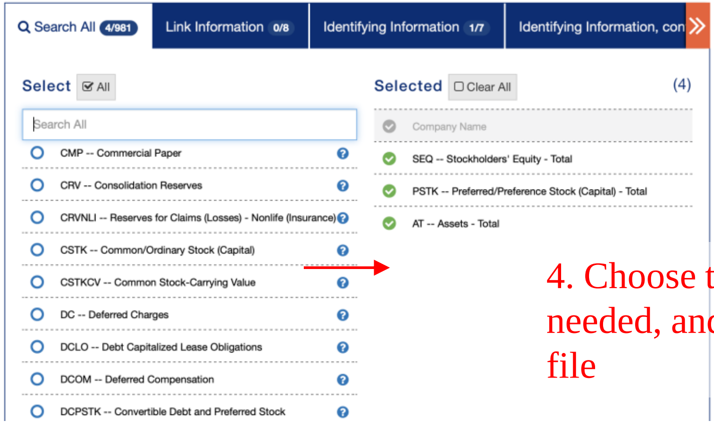


1. Choose the CRSP Dataset

2. Choose for the CRSP/Compustat Merged dataset to get fundamental data



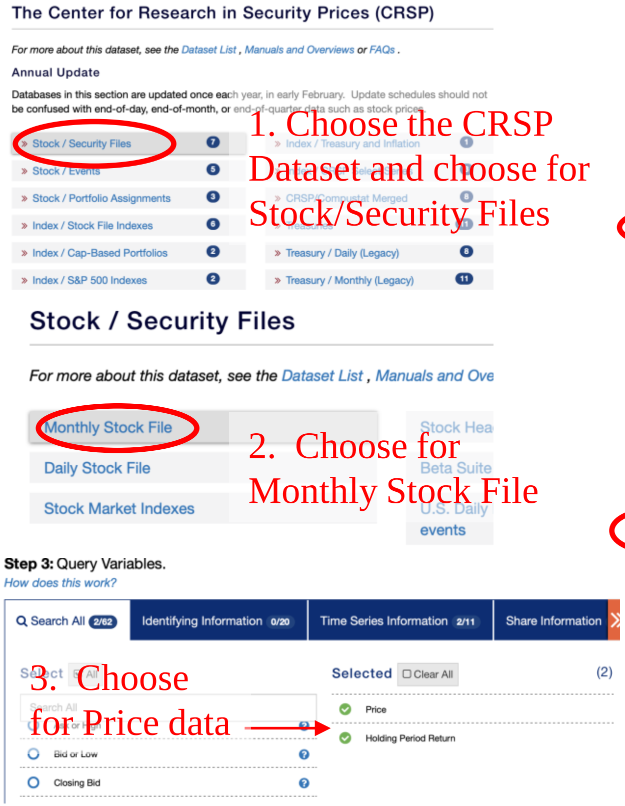
3. Choose for annual Fundamental data



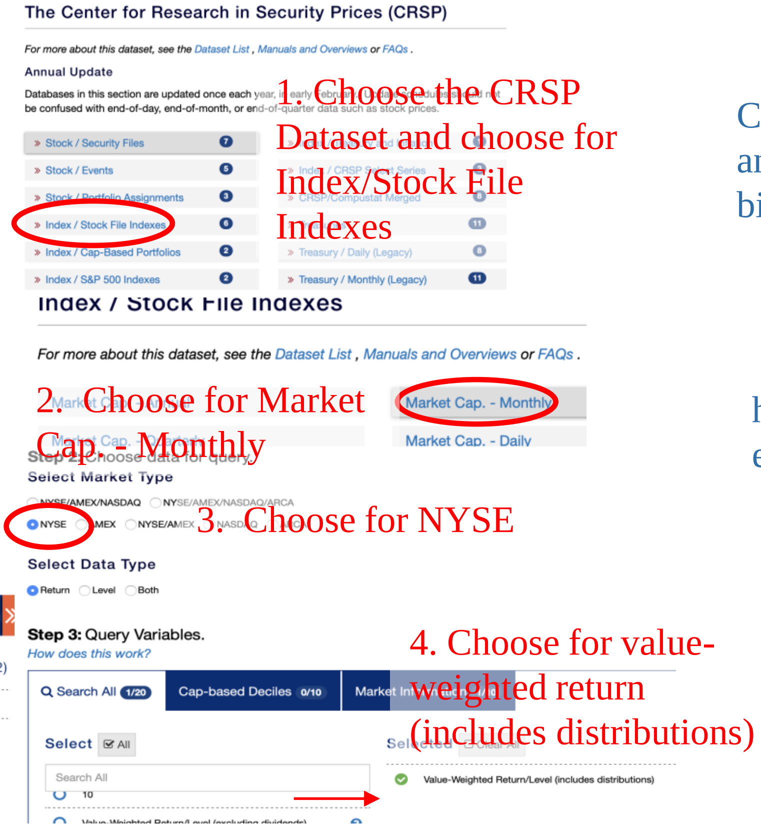
4. Choose the items needed, and export to csv file

Appendix 1 – Instruction to Download Data, Cont.

Stock Price



Market Return



Risk-Free Return

CRSP → Index/ Treasury and Inflation → 10-year T-bill Return

Inflation/ CPI

https://fred.stlouisfed.org/series/CPIAUCSL

Appendix 2 – Construct Quality Score

This appendix guides you through how to build the variable components of a quality score.

Data preparation: From CRSP/Compustat Merged, obtain relevant annual fundamental data; Monthly stock price/return data for beta calculation; Monthly 10-Y T-bill return

Profitability Averaging z-scores of GPOA, ROE, ROA, CFOA, GMAR and ACC

- GPOA = (revenue – costs of goods sold) / total assets
- ROE = (income before extraordinary items) / (stockholder's equity – preferred stocks equity)
- ROA = (income before extraordinary items) / total assets
- CFOA = (net income + depreciation – changes in working capital – capital expenditures) / total assets
- GMAR = (revenue – costs of goods sold) / total sales
- ACC = (depreciation – changes in working capital) / total assets
- Z-score formula: minus mean and then divided by standard deviation

Growth Averaging z-scores of five-year growth in GPOA, ROE, ROA, CFOA, GMAR and ACC

- GP = revenue – costs of goods sold
- BE = stockholder's equity – preferred stocks equity
- CF = net income + depreciation – changes in working capital – capital expenditures
- MWCPD = depreciation – changes in working capital

Appendix 2 – Construct Quality Score, Cont.

Safety Averaging z-scores of Beta, LEV, IDIOVOL, O, Z, EVOL

- Compute market Beta by running regression for each company in each year
- LEV = - (long term debt + short term debt + minority interest + preferred stock) / total debt
- Compute Ohlson's O-Score and Altman's Z-Score following the formulas
- Obtain IDIOVOL from CRSP and then annualize it
- EVOL can be obtained from CRSP directly

Payout Averaging z-scores of DISS, EISS, NPOP

- DISS = -log(one-year % change in total debt)
- EISS = -log(one-year % change in number of shares)
- NPOP = the sum of total net payout over the past 5 years

Quality Averaging z-scores of Profitability, Growth, Safety, Payout

Appendix 2 – Construct Quality Score – Key codes

Z-score Construction

Section1/profit&growth.R

This code is an example of computing the profitability z-score

```r
###Profitability
attach(Profit)
#calculate each variables
Profit$gpoa<-(revt-cogs)/at
Profit$roe<-ib/(seq-pstk)
Profit$roa<-ib/at
Profit$cfoa<-(ni+dp-chwc-capx)/at
Profit$gmar<-(revt-cogs)/revt
Profit$acc<-(dp-chwc)/at
Profit$cfoa[is.na(Profit$cfoa)]

#calculate mean
Profit<-Profit %>% group_by(fyear) %>% mutate(avg_gpoa=mean(gpoa),avg_roe=mean(roe),avg_roa=mean(roa),
                                       avg_cfoa=mean(cfoa),avg_gmar=mean(gmar),avg_acc=mean(acc))
#calculate standard deviation
Profit<-Profit %>% group_by(fyear) %>% mutate(std_gpoa=sd(gpoa),std_roe=sd(roe),std_roa=sd(roa),
                                          std_cfoa=sd(cfoa),std_gmar=sd(gmar),std_acc=sd(acc))
#calculate z scores
attach(Profit)
Profit$z_gpoa= (gpoa-avg_gpoa)/std_gpoa
Profit$z_roe= (roe-avg_roe)/std_roe
Profit$z_roa= (roa-avg_roa)/std_roa
Profit$z_cfoa= (cfoa-avg_cfoa)/std_cfoa
Profit$z_gmar= (gmar-avg_gmar)/std_gmar
Profit$z_acc= (acc-avg_acc)/std_acc

#Calculate the profitability score
Profit <-Profit %>% group_by(fyear) %>% mutate(z_profit=(z_gpoa+z_roe+z_roa+z_cfoa+z_gmar+z_acc)/6)
```

Appendix 3 – Constructing Quality Portfolios and Standard Factors

Quality portfolios P1 (lowest quality) to P10 (highest quality) are used in table 2 and 4 to analyze quality measures and returns across stocks with different quality characters.

Quality Portfolios

- Split stocks into 10 quality portfolios by quality scores, P10 as the highest quality portfolio and P1 as the lowest quality portfolio
- Each quality portfolio is value-weighted

Standard factors (Fama-French factors) MKT, SMB, HML and UMD are used in table 3, 6, 9 to analyze stock returns and quality measures. These factors can be downloaded directly from AQR website (link), but in this project we built them with CRSP/Compustat data.

MKT Market premium

- Value-weighted return on all available stocks minus risk free (we used 10-year Treasury rates)

SMB Size factor, small minus big

- Sort stocks by market value of equity, top 20% are big stocks and bottom 20% are small stocks
- Within each category, split into 3 value-weighted portfolios by book-to-market ratios – value, natural and growth; stocks with low book-to-market ratios are considered growth stocks

Appendix 3 – Constructing Quality Portfolios and Standard Factors, Cont.

SMB Size factor, small minus big (Cont.)

- SMB is the average return on the 3 small portfolios minus the average return on 3 big portfolios: SMB = 1/3(small value + small natural + small growth) - 1/3 (big value + big natural + big growth)

HML Value factor, high minus low

- Same portfolio construction as SMB but only focus on the average return on the value and growth portfolios: HML = 1/2(small value + big value) - 1/2 (small growth + big growth)

UMD Momentum factor, up minus down

- Sort stocks by market value of equity, top 20% are big stocks and bottom 20% are small stocks
- Within each category, split into 3 value-weighted portfolios by stock returns – high, median, low
- UMD is the average return on the 2 high return portfolios minus the average return on 2 low return portfolios: UMD = 1/2(small high + big high) - 1/2 (small low + big low)

Note: The paper uses median NYSE market equity as size breakpoint. However, we followed State Street's instructions – regard top 20% as big stocks and bottom stocks. We tried both and found the results would be closer to the paper's result with State Street's method.

Appendix 3 – Key Codes for Quality Portfolios

```r
create_portfolio <- function(dataset, by='me', num_subsets=10, portfolio_initio='P') {
  dataset$portfolio <- NULL # new "portfolio column to record portfolio
  count_entry <- 1 # row index for the n-th entry
  dataset <- dataset %>% group_by(fyear) %>%
    mutate(rank=rank(z_quality)) %>%
    arrange(fyear,z_quality) %>%
    mutate(cum_me=cumsum(!!as.name(by)))
  uniqueYear <- as.vector(unique(dataset$fyear))
  head(dataset)
  for (year in uniqueYear) {
    count_subsetEntry <- 1 # row index for the n-th entry in subset
    count_portfolio <- 1 # initalize portfolio index: 1 means P1
    sum_portfolio <- 0 # initalize cumulative market value
    subset <- dataset %>% filter(fyear == year)
    subset_length <- length(subset$GVKEY)
    mkt_cap <- sum(subset$me)
    portfolio_cap <- 1/num_subsets*mkt_cap
    for (i in c(1:subset_length)) {
      sum_portfolio <- sum_portfolio + subset[i, by]
      if (sum_portfolio > count_portfolio*portfolio_cap) {
        # if cumulative market value > portfolio boundary, then move on to the next portfolio
        count_portfolio <- count_portfolio + 1
      }
      dataset[count_entry, 'portfolio'] <- paste0(portfolio_initio, count_portfolio)
      count_entry <- count_entry + 1
      count_subsetEntry <- count_subsetEntry + 1
    }
  }
  dataset <- as.data.frame(dataset)
  dataset
}
```

Quality Portfolio Construction

section4/1_mkt_excessret.R

This code constructs quality portfolios and calculates the MKT factor

Function create_portfolio(dataset, by='me', num_subsets=10, portfolio_initio='P')
dataset: the dataset to be use passed
by: sort by; default 'me' (market equality)
num_subsets: number of portfolios to be created; default=10
portfolio_initio: determines how the portfolios are named; default='P' (the outputs would be P1, P2, etc.)

Function logics

- Create subsets by year
- For each year, sort by market equity; calculate portfolio's market cap threshold = 1/# of portfolios*total mkt cap; calculate each stock's weight in its portfolio = stock mkt cap/portfolio mkt cap
- Once a portfolio is full (portfolio mkt cap > threshold) then move on to the next portfolio

Appendix 3 – Key Codes for Standard Factors

Part 1

```r
#sort stocks by first size then value
reg <- data %>%
  group_by(fyear) %>%
  mutate(size_rank=rank(me), small_threshold=max(size_rank)*0.2, big_threshold=max(size_rank)*0.8, bm_rank=rank(bm)) %>%
  arrange(fyear, me, bm)

#rank them and sort them into 6 portfolios by ranks
reg_s <- reg %>% filter(size_rank <= small_threshold)
reg_b <- reg %>% filter(size_rank >= big_threshold)
reg_s[, "portfolio"] <- bin_data(reg_s$bm_rank, bins=3, binType = "explicit")
reg_b[, "portfolio"] <- bin_data(reg_b$bm_rank, bins=3, binType = "explicit")
reg_s$portfolio<-as.factor(reg_s$portfolio)
reg_b$portfolio<-as.factor(reg_b$portfolio)
levels(reg_s$portfolio)<-c("S1","S2","S3")
levels(reg_b$portfolio)<-c("B1","B2","B3")
reg_smb <- rbind(reg_s, reg_b)
```

Part 2

```r
# calculate SMB and HML
uniqueYear <- as.vector(unique(ret$fyear))
smbMatrix <- matrix(NA, nrow=length(uniqueYear), ncol=3)
count = 1
for (y in uniqueYear) {
  subdata <- ret %>% filter(fyear == y) %>% arrange(portfolio)
  if (length(unique(subdata$portfolio)) == 6) { # years that only contains 6 complete sub-portfolios s1-s3, b1-b3
    s1 <- subdata %>% filter(portfolio == 'S1')  # small growth
    s2 <- subdata %>% filter(portfolio == 'S2')  # small netural
    s3 <- subdata %>% filter(portfolio == 'S3')  # small value
    b1 <- subdata %>% filter(portfolio == 'B1')  # big growth
    b2 <- subdata %>% filter(portfolio == 'B2')  # big netural
    b3 <- subdata %>% filter(portfolio == 'B3')  # big value
    smb <- 1/3*sum(s1[,'weighted_ret']) + 1/3*sum(s2[,'weighted_ret']) + 1/3*sum(s3[,'weighted_ret']) -
          1/3*sum(b1[,'weighted_ret']) - 1/3*sum(b2[,'weighted_ret']) - 1/3*sum(b3[,'weighted_ret']) # small - big
    hml <- 1/2*sum(rbind(s3,b3)[,'weighted_ret']) - 1/2*sum(rbind(s1,b1)[,'weighted_ret'])
    smbMatrix[count, 1] <- y
    smbMatrix[count, 2] <- smb
    smbMatrix[count, 3] <- hml
  }
  count <- count + 1
}
```

SMB & HML

section4/1_smb_hml.R

Similar for UMD: *section4/1_umd.R*

This code calculates SMB and HML factors

Part 1 logics

- For each year, split stocks into big and small stocks (top 20% and bottom 20%)
- First sort by size and then sort by market equity

Part 2 logics

- Initialize a matrix to save results; row = number of years
- Evaluate whether 6 portfolios are complete for each year (may have missing portfolios in the first few years)
- Calculate SMB and UMD factors

Appendix – Table 2

Part 1

```
# create portfolio function
create_portfolio <- function(dataset, by='me', num_subsets=10, portfolio_initio='P')
```

Part 2

```
table2 <- table2 %>% group_by(fyear, portfolio) %>% mutate(me_portfolio=sum(me))
attach(table2)
table2$weight <- me/me_portfolio
table2$weighted_quality <- weight*z_quality
table2$weighted_profit <- weight*z_profit
table2$weighted_growth <- weight*z_growth
table2$weighted_safety <- weight*z_safety
table2$weighted_payout <- weight*z_payout
```

Part 3

```
colnames(table2_output) <- c('P1', 'P10', 'P2', 'P3', 'P4', 'P5', 'P6', 'P7', 'P8', 'P9')
rownames(table2_output) <- c('Quality(t)', 'Quality(t+1Y)', 'Quality(t+3Y)', 'Quality(t+5Y)','Quality(t+10Y)',
                             'Profit(t+10Y)', 'Growth(t+10Y)', 'Safety(t+10Y)', 'Payout(t+10Y)')
table2_output <- as.data.frame(table2_output) # convert to dataframe so as to perform 'select' reorder in the next line
table2_output <- table2_output %>% select(P1, P2, P3, P4, P5, P6, P7, P8, P9, P10)
```

Table 2

section2/table2.R

This code constructs table 2

Part 1 logics
Similar to the create_portfolio() function in Appendix 2 – Key Codes for Quality Portfolios

Part 2 logics
Because quality portfolios are value-weighted, portfolios' quality scores = sum(individual stock weights in the portfolio * that stock's quality scores)

Part 3 logics
Rename column names and row names; Note that in R "P10" is ranked after "P1", so need to reorganize the output table

Appendix – Table 3

Part 1

```
delete_year <- unique(data$fyear)[1:2]
data <- data %>% filter(fyear %notin% delete_year)
```

Part 2

```
reg3.1.2<- lm(z_mb ~ z_quality + z_me + retLag)
summary(reg3.1.2)
```

Table 3

section3/data_table3.R -> prepare data for table 3

section3/table3.R -> construct table 3

Part 1 logics

- Filter out first 2 available years, because lagging variables create missing values
- %notin% function provided in mefa4 package

Part 2 logics

Sample code for creating regressions

Appendix – Table 4

```
for (p in uniquePort) {
  subdata <- data %>% filter(portfolio==p)
  # 1. excess return
  table[1,count] <- mean(subdata$excess_ret)

  # 2. excess return significance
  # 3. CAPM alpha
  lm_capm <- lm(data=subdata, excess_ret ~ mkt)
  table[3,count] <- coef(lm_capm)[1]

  # 4. CAPM alpha significance
  table[4,count] <- summary(lm_capm)$coefficients[1,4]

  # 5. 3-factor alpha
  lm_3factor <- lm(data=subdata, excess_ret ~ mkt+smb+hml)
  table[5,count] <- coef(lm_3factor)[1]

  # 6. 3-factor alpha significance
  table[6,count] <- summary(lm_3factor)$coefficients[1,4]

  # 7. 4-factor alpha
  lm_4factor <- lm(data=subdata, excess_ret ~ mkt+smb+hml+umd)
  table[7,count] <- coef(lm_4factor)[1]

  # 8. 4-factor alpha significance
  table[8,count] <- summary(lm_4factor)$coefficients[1,4]

  # 9. beta
  table[9,count] <- coef(lm_capm)[2]

  # 10. sharpe ratio
  table[10,count] <- mean(subdata$excess_ret)/sd(subdata$excess_ret)

  # 11. information ratio
  table[11,count] <- coef(lm_4factor)[1]/sd(lm_4factor$residuals)

  # 12. adjusted r2
  table[12,count] <-summary(lm_4factor)$adj.r.squared
  count <- count + 1
}
```

Table 4

This table mainly uses the standard factors created before.

 # Create table with portfolio names, P1 to P10, H-L
uniquePort <- c('P1', 'P2', 'P3', 'P4', 'P5', 'P6', 'P7', 'P8', 'P9', 'P10', 'H-L')
table <- as.data.frame(matrix(NA, nrow=12, ncol=11))
colnames(table) <- uniquePort

Logics

- Run regressions with different factors.
- Alpha is known as the intercept of the regression results.
- P-value is used here instead of t-statistics.
- Sharpe ratio = average of excess return / standard deviation of excess return.
- Information ratio = four-factor alpha(intercept) / standard deviation of the estimated residuals
- Fill in the table with the results gotten from the regressions with a for loop function.

Appendix – Table 5

Part 1

```
sort_portfolio <- function (by='z_quality') {
  # first sort by size and then desired variable (e.g. z_quality)
  stocks_small <- data_table5 %>% group_by(fyear) %>% filter(size_rank <= junk_threshold) %>% mutate(rank=rank(!!as.name(by)))
  stocks_big <- data_table5 %>% group_by(fyear) %>% filter(size_rank >= quality_threshold) %>% mutate(rank=rank(!!as.name(by)))
  #then, for each size, sort
  stocks_small[, "portfolio"] <- bin_data(stocks_small$rank, bins=3, binType = "explicit")
  stocks_big[, "portfolio"] <- bin_data(stocks_big$rank, bins=3, binType = "explicit")
  #set each portfolio into factor
  stocks_small$portfolio<-as.factor(stocks_small$portfolio)
  stocks_big$portfolio<-as.factor(stocks_big$portfolio)
  #then rename
  levels(stocks_small$portfolio)<-c("S1","S2","S3")
  levels(stocks_big$portfolio)<-c("B1","B2","B3")
  sorted_data <- rbind(stocks_small, stocks_big)
  # keep variables for table 5
  sorted_data  <- sorted_data  %>%
    dplyr::select(fyear, GVKEY, ret.excess, ret.abnormal, portfolio) %>%
    arrange(fyear, portfolio)
  # function return
  sorted_data
}
```

Part 2

```
portfolio_return <- function (sorted_data) {
  uniqueYear <- as.vector(unique(sorted_data$fyear))
  yearlyDataMatrix <- matrix(NA, nrow=length(uniqueYear), ncol=3)
  count = 1
  for (y in uniqueYear) {
    subdata <- sorted_data %>% filter(fyear == y) %>% arrange(portfolio)
    # calculate qmj based on returns
    excess.ret <- 0.5*mean(subdata[subdata$portfolio=='B3',]$ret.excess) + 0.5*mean(subdata[subdata$portfolio=='S3',]$ret.excess) -
      0.5*mean(subdata[subdata$portfolio=='B1',]$ret.excess) - 0.5*mean(subdata[subdata$portfolio=='S1',]$ret.excess)
    abnormal.ret <- 0.5*mean(subdata[subdata$portfolio=='B3',]$ret.abnormal) + 0.5*mean(subdata[subdata$portfolio=='S3',]$ret.abnormal)
      0.5*mean(subdata[subdata$portfolio=='B1',]$ret.abnormal) - 0.5*mean(subdata[subdata$portfolio=='S1',]$ret.abnormal)
    yearlyDataMatrix[count, 1] <- y
    yearlyDataMatrix[count, 2] <- excess.ret
    yearlyDataMatrix[count, 3] <- abnormal.ret
    count <- count + 1
  }
  # rename matrix columns
  colnames(yearlyDataMatrix) <- c('fyear', 'excess_return', 'abnormal_return')
  # function return
  yearlyDataMatrix
}
```

Table 5

Section5/table5.R

This code creates size portfolios; calculate portfolio excess returns sorted by quality, profitability, safety, growth and payout scores; construct correlation matrix

Part 1 logics

Function sort_portfolio(by='z_quality')

by: sort by; default 'z_quality' (z-score for quality)

- Create 6 portfolios – small quality, small normal, small junk, big quality, big normal, big junk by assigned z-score
- Specific logics similar to Part 1 in Appendix 3 – Key Codes for Standard Factors

Part 2 logics

Function portfolio_return(sorted_data)

sorted data: output from sort_portfolio()

- Calculate returns and abnormal returns for that year with QMJ strategy by ½(small quality + big quality) – ½(small junk + big junk)
- Specific logics similar to Part 2 in Appendix 3 – Key Codes for Standard Factors

Appendix – Table 5, Cont.

Part 3

```
portfolio_sorted_data <- sort_portfolio(by = 'z_quality')
return_table <- portfolio_return(sorted_data = portfolio_sorted_data)
colnames(return_table) <- c('fyear', 'return_z_quality', 'abnormal_return_z_quality')
# and then sort portfolio by profitability, safety, growth and payout, and merge them into return_table
sort_by <- c('z_profit', 'z_growth', 'z_safety', 'z_payout')
for (i in sort_by) {
  portfolio_sorted_data <- sort_portfolio(by = i)
  return_table_new <- portfolio_return(sorted_data = portfolio_sorted_data)
  colnames(return_table_new) <- c('fyear', paste0('excess_return_', i), paste0('abnormal_return_', i))
  return_table <- merge(return_table, return_table_new, by='fyear')
}
```

Part 4

```
# correlation matrix
table5_ret <- cor(return_table[,c(2,4,8,6,10)])
colnames(table5_ret) <- c('QMJ', 'Profitability', 'Safety', 'Growth', 'Payout')
rownames(table5_ret) <- c('QMJ', 'Profitability', 'Safety', 'Growth', 'Payout')
print(table5_ret)
```

Part 3 logics

- Applied the two functions in part 1 and 2
- Generate yearly QMJ strategy returns and abnormal returns with z-scores for quality, profitability, safety, growth and payout

Part 4 logics

- Select relevant columns to construct matrices for returns and abnormal returns

Appendix – Table 6

```r
#run regressions and create table 6
### QMJ
table_qmj <- as.data.frame(matrix(NA, nrow=11, ncol=1))
colnames(table_qmj) <- c('QMJ')
rownames(table_qmj) <-c('Excess Returns','CAPM-alpha','3-factor-alpha',
                        '4-factor-alpha','MKT','SMB','HML','UMD',
                        'Sharpe Ratio','Information Ratio','Adjusted R2')
#excess return
table_qmj[1,1]<-mean(data2$QMJ) #should we substract rf or not?

#run regression
capm_qmj<-lm(data=data2,QMJ~mkt)
f3_qmj<-lm(data=data2,QMJ~mkt+smb+hml)
f4_qmj<-lm(data=data2,QMJ~mkt+smb+hml+umd)

#put them into the table
table_qmj[2,1]<-coef(capm_qmj)[1] #capm-alpha
table_qmj[3,1]<-coef(f3_qmj)[1] #3factor-alpha
#4factor coefficients
for (c in (4:8)) {
  table_qmj[c,1] <- coef(f4_qmj)[c-3]
}

#sharpe ratio
table_qmj[9,1] <-  mean(data2$QMJ)/sd(data2$QMJ)

#information ratio
table_qmj[10,1] <- coef(f4_qmj)[1]/sd(f4_qmj$residuals)

#Adjusted R2
table_qmj[11,1] <- summary(f4_qmj)$adj.r.squared
```

Table 6

Section6/table6.R

This code is an example of running the QMJ regressions to get the first column of Table 6.

- Create a blank table and name the columns and rows
- 1st row: excess return is the average excess return of the QMJ factor over the years
- Run regressions for 1-, 3- and 4-factor models
- Store the alphas and 4-factor coefficients into the table
- Sharp ratio = mean(QMJ return) / std(QMJ return)
- *Note: the QMJ return itself has already substracted the risk-free rate*
- Information ratio of 4-factor model = (excess return from 4-factor model) / standard deviation

Appendix – Table 9

```r
#run regression and create table 9
###smb
table_smb <- as.data.frame(matrix(NA, nrow=9, ncol=2))
colnames(table_smb) <- c('SMB','SMB2')
rownames(table_smb) <-c('Excess Returns','Alpha',
                        'MKT','HML','UMD','QMJ',
                        'Sharpe Ratio','Information Ratio',
                        'Adjusted R2')
#excess return
table_smb[1,1]<-mean(data2$smb) #should we substract rf or not?
table_smb[1,2]<-mean(data2$smb)

#run regression
lm_smb<-lm(data=data2,smb~.-avg_rf-qmj)
lm_smb2<-lm(data=data2,smb~.-avg_rf)

#put them into the table
for (c in (2:9)) {
  table_smb[c,1] <- coef(lm_smb)[c-1]
  table_smb[c,2] <-coef(lm_smb2)[c-1]
}
#sharpe ratio
table_smb[7,1] <-  mean(data2$smb)/sd(data2$smb)
table_smb[7,2] <-  mean(data2$smb)/sd(data2$smb)

#information ratio
table_smb[8,1] <- coef(lm_smb)[1]/sd(lm_smb$residuals) # alpha of lm_smb/sd.residual
table_smb[8,2] <- coef(lm_smb2)[1]/sd(lm_smb2$residuals) # alpha of lm_smb2/sd.residual

#Adjusted R2
table_smb[9,1] <- summary(lm_smb)$adj.r.squared
table_smb[9,2] <- summary(lm_smb2)$adj.r.squared
```

Table 9

Section9/table9.R

This code is an example of running the SMB regressions to get the first two columns of Table 9.

The basic logic is similar to that of Table 6. The difference is that moving QMJ from the left of the equation to the right.

- Create a blank table and name the columns and rows
- 1st row: excess return is the average excess return of the SMB factor over the years
- Run regressions with and without QMJ factor
- Store the alphas and coefficients into the table
- Sharp ratio = mean(SMB return) / std(SMB return)
- *Note: the SMB return itself has already substracted the risk-free rate*
- Information ratio of 4-factor model with QMJ = (excess return from 4-factor model) / standard deviation