Using the Industry-Based Fama-French Model to Evaluate Industry Portfolios

Ossama Elhadary
City University of New York

In this paper I create the industry-based Fama-French 3-factor model by constructing the three Fama-French factors for each industry group separately. I then use this model together with the traditional Fama-French model to evaluate the returns of 41 monthly industry portfolios and I compare the results. Three sets of stock portfolios are used in my analysis: (1)Stocks in a single industry, (2)Stocks in industries in the same industry group, (3) Stocks in a industries not in the same industry group

I show that the modified Fama-French model outperforms the traditional Fama-French model when assessing the performance of those industry portfolios. This is especially true for industry portfolios within small industry groups like mining where the regression R^2 improved by 34.8%, 68.8%, 28% and 1052% for the underlying Mining, Coal, Oil and Gold industries. Actually in 16 out of the 18 non-manufacturing industry portfolios, R^2 improved. These results imply that within each industry group asset space, the new industry-based model tends to be more efficient than the traditional Fama-French model.

Keywords: portfolio, industry, Fama French, Fama

INTRODUCTION

While the Fama-French and its CAPM predecessor models were proven to perform well on the whole market, they tend to underperform when used to evaluate portfolio returns based on individual industries. Hu (2003) for example demonstrated that running the Fama-French regression on the return of 17 industries (table 1) resulted in substantial variation in the calculated adjusted R^2 values (the average=.78, and standard deviation =.13) across the industries. Moskowitz & Grinblatt (1999) also found that a strong and prevalent momentum exists in industries. De Vries (2012) used the CAPM, Fama-French, and Fama-French-Carhart (1997) models, and observed that there are several industries (Manufacturing, High-tech and Other) for which the three models seem to perform above average (in terms of higher R^2). He also observed that the "manufacturing" and "Other" industries are the only ones that produce significant (albeit small) pricing errors.

TABLE 1
ADJUSTED R^2 FROM THE FAMA-FRENCH REGRESSIONS ON THE EXCESS RETURN OF 17 INDUSTRIES

Industry	**Adjusted** R^2	**Industry**	**Adjusted** R^2
Food	0.824	FabPr	0.833
Mines	0.478	Machn	0.895
Oil	0.485	Cars	0.799
Clths	0.793	Trans	0.846
Durbl	0.851	Utils	0.598
Chems	0.856	Rtail	0.792
Cnsum	0.792	Finan	0.832
Cnstr	0.865	Other	0.896
Steel	0.768		

Fama & French (1997) also reported that the estimates of the CAPM and Fama-French 3 factor model for many industries differ by more than two percent and attributed these differences to uncertainty about the "true risk factors and imprecise estimates of period-by-period risk loadings". In these cases, one typically observes a lower R^2, an increase in α, as well as an increase in the standard errors for the cost of equity that render the calculation of the profitability of an investment impossible (Fama-French, 1997).

According to theory, the three Fama-French factors represent fundamental economic risks faced by all firms. I argue though that there are certain localized economic risks that affect certain industries and not others. These risks will contribute to the Fama-French factors but at weights proportional to their respective industry sizes. Imagine for example an economic force that affects the mining industry, but not any of the other industries. Such a force will greatly affect that industry's return but once mixed with other risks and merged into the Fama-French factors, that effect will be greatly diminished. This is especially true of smaller industries compared to larger ones.

In this research I show that a modified Fama-French model with its variables created for each industry group separately outperforms the traditional Fama-French model when assessing the performance of 41 industry portfolios. I thus show that that within each industry group asset space, the new modified model tends to be more efficient than the traditional Fama-French model.

THE INDUSTRY GROUPS

Table 2 lists the nine industry groups used in this research. The Public Administration industry group (SIC Code 9900 to 9999) was dropped because there was a very small number of firms from this industry group in the downloaded data.

TABLE 2
INDUSTRY GROUPS USED IN THIS RESEARCH

Number	Industry Group	SIC Codes
1	Agriculture, Forestry and Fishing	0100-0999
2	Mining	1000-1499
3	Construction	1500-1799
4	Manufacturing	2000-3999
5	Transportation & Communication	4000-4999
6	Wholesale	5000-5199
7	Retail	5200-5999
8	Finance Insurance and Real Estate	6000-6799
9	Services	7000-8999

THE INDUSTRY PORTFOLIOS

In this research, I use 41 monthly value-weighted and equal-weighted industry portfolios downloaded from the Kenneth French web site. Table 3 shows the industries, the associated industry groups, the average number of firms in the portfolio and the average firm size. Note that since the number of firms in each industry portfolio changes month by month, I calculate the average number of firms in industry g using the following equation:

$$\#Firms_g^{Average} = \frac{\sum_{t=1}^{T} \#Firms_{gt}}{T} \quad (1)$$

where T=72 and is the total number of months in the sample, and $\#Firms_{gt}$ is the reported number of firms in month t in industry g.

Similarly, the average firm size varies by month, and accordingly, I calculate the average of that number as follows:

$$FirmSize_g^{Average} = \frac{\sum_{t=1}^{T} AverageFirmSize_{gt}}{T} \quad (2)$$

where T=72 and is the total number of months in the sample, and $AverageFirmSize_{gt}$ is the reported average monthly firm size in month t in industry g.

To calculate the portfolio size for industry g, I multiply the average number of firms by the average firm size as follows:

$$PortfolioSize_g = \#Firms_g^{Average} * FirmSize_g^{Average} \quad (3)$$

TABLE 3
THE 41 INDUSTRY PORTFOLIOS USED

Industry Group	Industry	# of Firms	Average Firm Size	% of Industry Group
1- Agriculture, Forestry and Fishing				
	Agriculture	9	5,919.59	100.0%
2 - Mining				
	Mining	18	4,383.69	23.7%
	Coal	9	2,194.71	11.9%
	Oil	146	9,093.30	49.1%
	Gold	9	2,842.61	15.4%
3- Construction				
	Construction	44	1,701.55	100.0%
4 - Manufacturing				
	Food	51	5,982.90	3.5%
	Drugs	323	4,786.96	2.8%
	Chemicals	70	4,795.67	2.8%
	Rubber and Plastic	17	1,474.29	0.9%
	Steel Works Etc	40	1,588.42	0.9%
	Fabricated Products	7	727.72	0.4%
	Machinery	103	3,781.45	2.2%
	Electrical Equipment	55	1,805.63	1.1%
	Autos	52	4,261.74	2.5%
	Aero	17	14,360.41	8.4%
	Ships	8	1,807.17	1.1%
	Guns	9	5,268.02	3.1%
	Chips	191	4,672.04	2.7%
	Paper (Business Supplies)	34	5,980.15	3.5%
	Medical Equipment	116	2,370.53	1.4%
	Soda	10	13,844.75	8.1%
	Beer	10	23,022.72	13.5%
	Smoke	4	53,810.78	31.6%
	Toys	17	1,426.21	0.8%
	Books	18	1,574.13	0.9%
	Household (Consumer Goods)	43	8,637.14	5.1%
	Clothes	36	4,448.18	2.6%
5 - Transportation & Communication				
	Utilities	90	7,108.39	31.2%
	Telecom	94	10,352.56	45.4%

	Transportation	72	5,348.34	23.4%
6 - Wholesale				
	Wholesale	99	2,307.96	100.0%
7 - Retail				
	Retail	163	7,321.37	100.0%
8 - Finance Insurance and Real Estate				
	Banking	472	3,071.19	19.3%
	Insurance	107	6,279.45	39.5%
	Real Estate	20	1,747.36	11.0%
	Fianamce (Trading)	95	4,795.41	30.2%
9 - Services				
	Healthcare	56	1,694.35	12.0%
	Personal Services	40	1,463.08	10.3%
	Business Services	172	1,943.38	13.7%
	Software	267	6,367.87	45.0%
	Entertainment	43	2,671.52	18.9%

Table 3 shows the number and average size of firms in each of the 41 industry portfolios. The table also shows the relative size of the industry portfolios compared to its industry group. One can see for example that on the average, Tobacco companies are larger than all other companies in the sample, while Fabricated Products companies are the smallest. In fact, an average Tobacco company is almost 74 times larger than an average Fabricated Products company. With 267 Software companies, the software industry seems to be the largest industry in the sample.

DATA CLUSTERS AND THE INDUSTRY PORTFOLIOS

To support my decision to use the Industry-based Fama-French model, I explore the clusters within the data and try to show that the clusters are centered on the industry groups used. I create 9 data clusters using the STATA Clustering functionality and then I explore the relationship between the data clusters and the 41 industry portfolios introduced in the previous section.

Table 4 shows the correlations between the value-weighted returns of the 41 industries and the returns of the nine clusters. The results clearly show substantial variation across the clusters and across the industries.

TABLE 4
CORRELATION BETWEEN VALUE-WEIGHTED INDUSTRY PORTFOLIOS AND THE NINE CLUSTERS

Industry Group	Industry	C1	C2	C3	C4	C5	C6	C7	C8	C9
1 Agriculture, Forestry and Fishing	Agriculture	0.19	0.58	0.67	0.31	0.62	0.46	0.67	0.83	0.72
2 Mining	Mines	-.20	0.76	0.89	0.75	0.34	0.62	0.36	0.65	0.79
	Coal	0.19	0.67	0.88	0.78	0.34	0.40	0.65	-.15	0.52
	Oil	0.17	0.71	0.92	0.90	0.53	0.61	0.79	0.86	0.81
	Gold	-.41	-.26	0.19	0.50	0.05	0.10	0.27	0.48	0.36
3 Construction	Construction	0.12	0.85	0.91	0.87	0.82	0.73	0.88	0.76	0.82
4 Manufacturing	Food	0.26	0.69	0.66	0.74	0.48	0.41	0.70	0.60	0.51
	Drugs	0.11	0.49	0.68	0.47	0.66	0.51	0.76	0.84	0.65
	Chemicals	0.19	0.82	0.95	0.95	0.72	0.61	0.88	0.96	0.87
	Rubber and Plastic	0.36	0.71	0.93	0.92	0.76	0.67	0.89	0.78	0.88
	Steel Works	0.24	0.81	0.96	0.85	0.64	0.71	0.84	0.85	0.88
	Fabricated Products	0.36	0.90	0.85	0.89	0.76	0.54	0.75	0.43	0.68
	Machinery	0.10	0.96	0.96	0.96	0.82	0.68	0.93	0.94	0.84
	Electrical	0.18	0.97	0.92	0.91	0.84	0.65	0.93	0.88	0.72
	Autos	0.53	0.87	0.90	0.89	0.80	0.60	0.90	0.91	0.71
	Aero	0.67	0.60	0.89	0.86	0.66	0.60	0.77	0.89	0.80
	Ships	0.58	0.62	0.90	0.92	0.62	0.66	0.83	0.88	0.80
	Guns	0.49	0.85	0.59	0.85	0.47	0.50	0.37	0.53	0.47
	Chips	0.29	0.70	0.89	0.94	0.59	0.60	0.78	0.87	0.90
	Paper (Business Supplies)	0.51	0.82	0.97	0.91	0.75	0.66	0.88	0.93	0.88
	Medical Equipment	0.22	0.60	0.84	0.81	0.71	0.70	0.81	0.84	0.84
	Soda	0.25	0.50	0.36	0.63	0.30	0.19	0.48	0.61	0.20
	Beer	0.35	0.33	0.41	0.70	0.34	0.34	0.62	0.68	0.20
	Smoke	-.13	0.56	0.58	0.48	0.12	0.23	0.40	0.55	0.30
	Toys	0.09	0.67	0.81	0.78	0.52	0.55	0.75	0.90	0.74
	books	0.23	0.94	0.85	0.62	0.82	0.62	0.89	0.96	0.67
	Household (Consumer Goods)	0.35	0.84	0.69	0.73	0.50	0.39	0.49	0.65	0.49
	Clothes	0.23	0.65	0.75	0.87	0.46	0.41	0.38	0.30	0.55
5 Transportation & Communication	Utilities	0.32	-.14	0.56	0.77	0.33	0.44	0.77	0.48	0.26
	Telecommunications	0.14	0.84	0.84	0.94	0.57	0.63	0.69	0.85	0.65
	Transportation	0.66	0.59	0.89	0.97	0.58	0.67	0.75	0.78	0.77
6 Wholesale	Wholesale	0.57	0.96	0.95	0.94	0.85	0.69	0.90	0.95	0.89
7 Retail	Retail	0.46	0.61	0.84	0.86	0.59	0.52	0.82	0.65	0.81
8	Banking	0.14	0.87	0.88	0.87	0.79	0.66	0.73	0.89	0.86
Finance	Insurance	0.22	0.91	0.95	0.77	0.79	0.64	0.89	0.84	0.85

Insurance and	Real estate	0.38	0.70	0.93	0.96	0.74	0.72	0.90	0.88	0.88
Real Estate	Finance	0.27	0.88	0.94	0.91	0.83	0.70	0.85	0.96	0.87
9 Services	Health	-.04	0.66	0.77	0.86	0.51	0.72	0.49	0.38	0.79
	Personal Services	0.51	0.84	0.88	0.83	0.67	0.67	0.93	0.61	0.84
	Business Services	0.52	0.85	0.98	0.96	0.81	0.72	0.93	0.95	0.96
	Software	0.12	0.75	0.88	0.87	0.63	0.62	0.75	0.88	0.91
	Entertainment	0.12	0.54	0.83	0.82	0.60	0.61	0.63	0.89	0.81

In table 5, I summarize the results of table 4 by listing the average, median, highest, and lowest correlations per cluster as well as the industry that has with the highest return correlation with the cluster return.

TABLE 5
SUMMARY OF CORRELATIONS PER CLUSTER

Cluster	Mean	Median	Highest	Lowest	Industry with Highest Correlation
Cluster 1	0.26	0.24	0.67	-0.41	Aero
Cluster 2	0.69	0.71	0.97	-0.26	Electrical
Cluster 3	0.81	0.88	0.98	0.19	Business Services
Cluster 4	0.81	0.86	0.97	0.31	Transportation
Cluster 5	0.60	0.63	0.85	0.05	Wholesale
Cluster 6	0.57	0.62	0.73	0.10	Construction
Cluster 7	0.73	0.77	0.93	0.27	Electrical
Cluster 8	0.74	0.84	0.96	-0.15	Finance
Cluster 9	0.71	0.79	0.96	0.20	Business Services

The results above show that the coefficients of the correlations between the industries' value weighted returns and the returns of the clusters vary widely. The average of the correlation coefficient for the nine clusters vary between .26 for cluster 1 and .81 for clusters 3 and 4. The median on the other hand varies between .24 for cluster 1, and .88 for cluster 3. The results also show that the returns of certain industries seem to consistently correlate very highly with the clusters returns. The return of the Electrical Equipment industry for example has the highest correlation (out of all 41 industries) with clusters 2 and 7. Similarly, the return of the Business Services industry had the highest correlation with the returns of clusters 3 and 9. For the returns of clusters 1, 4, 5, 6, and 8, the highest correlations were with the returns of the Aircraft, Transportation, Wholesale, Construction, and Finance industries respectively. I then repeat the same exercise with 41 equal-weighted portfolios and show the results in table 6.

TABLE 6
CORRELATION BETWEEN EQUAL-WEIGHTED INDUSTRY PORTFOLIOS AND THE NINE CLUSTERS

Industry Group	Industry	C1	C2	C3	C4	C5	C6	C7	C8	C9
1 Agriculture										
	Agriculture	0.10	0.81	0.88	0.35	0.72	0.62	0.83	0.67	0.70
2 Mining										
	Mines	-.18	0.53	0.82	0.91	0.44	0.57	0.62	0.70	0.84
	Coal	0.52	0.59	0.87	0.88	0.05	0.28	0.37	-0.23	0.53
	Oil	-.03	0.51	0.94	0.77	0.45	0.59	0.83	0.69	0.91
	Gold	-.46	-.60	0.67	0.73	0.01	0.41	0.34	0.04	0.55
3 Construction										
	Construction	0.31	0.87	0.92	0.85	0.84	0.74	0.88	0.68	0.90
4 Manufacturing										
	Food	0.42	0.88	0.80	0.92	0.79	0.66	0.79	0.93	0.66
	Drugs	-.05	0.18	0.87	0.86	0.83	0.68	0.79	0.82	0.87
	Chemicals	0.05	0.92	0.92	0.87	0.82	0.68	0.93	0.96	0.92
	Rubber and Plastic	0.45	0.63	0.95	0.97	0.70	0.71	0.89	0.82	0.90
	Steel Works	0.11	0.85	0.96	0.94	0.69	0.72	0.85	0.75	0.91
	Fabricated Products	0.44	0.84	0.91	0.86	0.74	0.69	0.85	0.64	0.83
	Machinery	0.38	0.94	0.97	0.96	0.88	0.72	0.95	0.83	0.88
	Electrical	0.40	0.83	0.89	0.90	0.84	0.72	0.90	0.91	0.86
	Autos	0.50	0.96	0.92	0.89	0.90	0.71	0.95	0.92	0.88
	Aero	0.63	0.74	0.88	0.79	0.78	0.64	0.88	0.90	0.84
	Ships	0.62	0.73	0.89	0.91	0.66	0.68	0.82	0.93	0.77
	Guns	-.17	-.09	0.77	0.59	0.51	0.60	0.74	0.49	0.61
	Chips	0.37	0.82	0.94	0.98	0.91	0.72	0.94	0.76	0.91
	Paper (Business Supplies)	0.60	0.37	0.95	0.91	0.74	0.80	0.83	0.95	0.91
	Medical Equipment	0.45	0.49	0.87	0.91	0.88	0.73	0.82	0.70	0.90
	Soda	-.11	0.42	0.73	0.68	0.46	0.52	0.62	0.61	0.49
	Beer	0.45	0.01	0.74	0.19	0.38	0.47	0.13	0.61	0.68
	Smoke	-.14	0.46	0.45	0.55	0.21	0.32	0.52	0.57	0.50
	Toys	0.43	0.62	0.71	0.87	0.47	0.49	0.55	0.57	0.65
	books	0.40	0.78	0.74	0.71	0.84	0.61	0.91	0.97	0.59
	Household (Consumer Goods)	0.23	0.97	0.95	0.89	0.77	0.70	0.70	0.92	0.89
	Clothes	0.53	0.56	0.85	0.87	0.55	0.54	0.81	-.05	0.69
5 Transportation & Communication										
	Utilities &	0.38	0.14	0.76	0.87	0.46	0.53	0.89	0.49	0.40
	Telecommunication	0.02	0.90	0.95	0.93	0.63	0.77	0.85	0.79	0.86

	Transportation	0.56	0.79	0.90	0.97	0.80	0.74	0.93	0.95	0.80
6 Wholesale										
	Wholesale	0.37	0.83	0.98	0.95	0.88	0.74	0.94	0.93	0.92
7 Retail										
	Retail	0.54	0.76	0.90	0.94	0.75	0.65	0.94	0.86	0.91
8 Finance Insurance and Real Estate										
	Banking	0.12	0.84	0.8	0.65	0.84	0.65	0.86	0.77	0.84
	Insurance	0.50	0.84	0.97	0.94	0.81	0.71	0.92	0.82	0.95
	Real estate	0.33	0.89	0.85	0.91	0.60	0.58	0.68	0.96	0.42
	Finance	0.34	0.89	0.95	0.93	0.90	0.77	0.90	0.98	0.93
9 Services										
	Health	0.30	0.65	0.79	0.86	0.73	0.75	0.73	0.85	0.75
	Personal Services	0.56	0.79	0.89	0.89	0.73	0.70	0.79	0.86	0.83
	Business Services	0.43	0.79	0.96	0.86	0.92	0.79	0.87	0.96	0.95
	Software	0.15	0.76	0.94	0.97	0.92	0.75	0.81	0.94	0.96
	Entertainment	0.38	0.69	0.89	0.82	0.72	0.68	0.77	0.87	0.78

In table 7, I again summarize the results of table 6 by listing the average, median, highest, and lowest correlations per cluster as well as the industry group that has return with the highest correlation with the cluster return.

TABLE 7
SUMMARY OF CORRELATIONS PER CLUSTER

Cluster	Mean	Median	Highest	Lowest	Highest Correlation Industry
Cluster 1	0.29	0.38	0.63	-0.46	Aero
Cluster 2	0.65	0.77	0.97	-0.60	Household (Consumer Goods)
Cluster 3	0.86	0.89	0.98	0.45	Wholesale
Cluster 4	0.83	0.88	0.98	0.19	Chips
Cluster 5	0.68	0.74	0.92	0.01	Business Services
Cluster 6	0.65	0.68	0.80	0.28	Paper (Business Supplies)
Cluster 7	0.78	0.83	0.95	0.13	Autos
Cluster 8	0.74	0.82	0.98	-0.23	Finance
Cluster 9	0.78	0.84	0.96	0.4	Software

The above table again shows large variations in the correlation coefficients across all the clusters. Cluster 1 continues to have the lowest average and median correlations while cluster 3 continues to have the highest mean average and median. The main difference between table 5 and 7, is in the industries with the highest correlations with the clusters. As shown in the table the Aircraft industry continues to have the highest correlation with returns of cluster 1, but all the other ones have changed.

CONSTRUCTING THE INDUSTRY-BASED FAMA-FRENCH MODEL

In the industry-based Fama-French model, I recreate the three Fama-French variables for each industry group separately. For this I downloaded the daily stocks data from CRSP for all NYSE, AMEX, and NASDAQ from January 2010 to December 2015. Quarterly Capitalization and Book to Market Ratio used in this research were downloaded from Compustat data. Like previous researches, I dropped stocks with price less than $5 to exclude Penny stocks. I then grouped all stocks according to their industry group's corresponding SIC codes (see Table 2) resulting in nine groups of stocks representing the nine industry groups. The traditional Fama-French 3-factor model can be written as:

$$Excess\ Stock\ Return = \alpha_{ff} + \beta_{mktrf} Mktrf_{ff} + \beta_{smb} SMB_{ff} + \beta_{hml} HML_{ff} + \epsilon_{\mathrm{ff}} \tag{4}$$

where $Mktrf_{ff}$, SMB_{ff}, and HML_{ff} are the Fama and French benchmark factors and the excess return is the difference between the stock return and the risk-free return.

While my industry-based Fama-French model can be written as:

$$Excess\ Stock\ Return = \alpha_{Ind} + \beta_{mktrf} Mktrf_{ind} + \beta_{smb} SMB_{ind} + \beta Hml_{hml} HML_{ind} + \epsilon_{\mathrm{ind}} \tag{5}$$

where $Mktrf_{ind}$, SMB_{ind}, and HML_{ind} are the author-created industry-based Fama and French benchmark factors and the excess return is the difference between the stock return and the risk-free return.

$Mktrf_{ind}$ is the first factor in the industry-based model and is the excess return on the industry-group (the difference between the value-weighted return on all NYSE, AMEX, and NASDAQ stocks in the industry group and the risk-free rate). To calculate this factor, I calculate the total market capitalization of stocks in industry group G (see equation 6) first.

$$TotalCap_t^G = \frac{1}{K} \sum_{i=1}^{K} mkvaltq_{it} \tag{6}$$

where $\mathrm{mkvaltq}_{\mathbf{it}}$ is stock i's (in industry group G) quarterly market capitalization at time t and K is the number of stocks in industry group G. Now I can calculate $Mktrf_{Ind}$ according to equation 7.

$$MKTRF_{Ind_t^G} = \left(\frac{1}{K * TotalCap_t^G} \sum_{i=1}^{K} return_{it} * \mathrm{mkvaltq}_{it}\right) - rf_t \tag{7}$$

where $MKTRF_{Ind_t^G}$ is the monthly excess return on industry group G at time t, $\mathrm{return}_{\mathrm{it}}$ is the daily return of stock i in industry group G at time t, $\mathrm{mkvaltq}_{\mathrm{it}}$ is stock i's quarterly market capitalization at time t, K is the number of stocks in industry group G, and rf_{t} is the risk free interest rate on the one month Treasury bill at time t.

To calculate the two other industry-based Fama and French factors, I start by creating six daily portfolios in each industry group. Like Fama and French (1993), the portfolios (S/L, S/M, S/H, B/L, B/M, and B/H) are formed from the intersection of the two Market Equity (ME) and the three Book Equity (BE)/ME groups (see figure 1). I then calculate the daily value-weighted returns on the six portfolios. Fama and French explained that they used value weighted returns in order to minimize variance since return variances are negatively related to size, and also because "using value weighted components results in mimicking portfolios that capture the different return behaviors of small and big stocks. Or high- and low-BE/ME stocks, in a way that corresponds to realistic investment opportunities".

SMB_{ind} (Small minus Big) is the second variable in the industry-based model and is the daily average return in each industry group on three small portfolios minus the average return on three big portfolios in each industry group.

$$SMB_{ind} = \left(\frac{1}{3}\right) * \{(SmallValue + SMallNeutral + SmallGrowth) - (BigValue + BighNuetral + BigGrowth)\} \quad (8)$$

HML_{ind} (High minus Low) is the third variable in the industry-based model and is the daily average return in each industry group on two value portfolios minus the average return on two growth portfolios:

$$HML_{ind} = \left(\frac{1}{2}\right) * (SmallValue + BigValue) - \left(\frac{1}{2}\right) * (SmallGrowth + BighGrowth) \quad (9)$$

USING THE INDUSTRY-BASED MODEL

In this section, I demonstrate the benefit of using the industry-based Fama-French model to evaluate various portfolios. Throughout this exercise, I use 41 monthly weighted industry portfolios that I downloaded from the Kenneth French web site. Three cases are considered:

1- Single industry portfolio
2- Portfolio formed of multiple industries in the same industry group
3- Portfolio formed of multiple industries not in the same industry group

Single Industry Portfolio

In this section, I run both the traditional Fama-French regression (equation 4) and the industry-based Fama-French regression (equation 5) on 41 monthly value-weighted individual industry portfolios downloaded from the Kenneth French web site. I repeat the same procedure with the monthly equal-weighted portfolios, and do not observe major changes in the results. I compare the two R^2 values from the two regression equations by defining the $DiffR_g^2$ variable as follows:

$$DiffR_g^2 = \frac{100}{R_{ff-g}^2} * (R_{ff-g}^2 - R_{ind-g}^2) \quad (10)$$

where R_{ff-g}^2 is the R^2 value obtained from regressing industry g excess return on the Fama-French factors, and R_{ind-g}^2 is the R^2 value obtained from regressing industry g excess return on the Fama-French industry-based factors.

Starting with industry group 2, there are 4 industry portfolios in this group, and table 8 shows the results of the regressions in terms of α_{ff}, α_{ind}, $R_{ff}^2, R_{ind}^2, and\ DiffR^2$. The results clearly show a substantial improvement in the R^2 value when using the industry-based Fama-French model compared to the traditional Fama-French model. The improvement was dramatic for the gold industry for example where the R^2 value increased by 1052%. Looking at the alphas, one can see that the four industry portfolios performed as well as the industry group (judging from the fact that α_{ind} was insignificantly different from zero). This is in stark contrast with the conclusion we get by looking at α_{ff}, which is negative for 3 of the industry portfolios, implying that the whole industry group did not perform as good as the market.

TABLE 8
REGRESSION R^2 AND ALPHAS FOR INDUSTRY GROUP 2

Industry	α	R^2_{ff}	α_{ind}	R^2_{ind}	$Diff R^2_g$
Mining	1.82*	0.50	0.14***	0.67	34.75
Coal	-4.31*	0.28	-1.54***	0.47	68.82
Oil	-0.68**	0.69	0.52***	0.89	28.14
Gold	-1.07***	-0.03	1.10***	0.28	1051.72

* Significant at the 5% level
** Significant at the 10% level
*** Insignificant

For industry group 3 (construction), there is only one industry portfolio in this group, and the improvement in R^2 value was 30.9% (see Table 9). On the other hand, the values of alpha were almost the same (α_{ff} was -.38% significant only at the 10% confidence level, while α_{ind} was -.39% and significant at the 5% significance level).

TABLE 9
REGRESSION R^2 AND ALPHAS FOR INDUSTRY GROUP 3

Industry	α	R^2_{ff}	α_{ind}	R^2_{ind}	$Diff R^2_g$
Construction	-0.38**	0.76	-0.39*	0.99	30.91

* Significant at the 5% level
** Significant at the 10% level
*** Insignificant

For industry group 5 (Transportation and communications), there were three industry portfolios in this group, and in the three cases, the R^2 value improved (see Table 10). The most dramatic example was for the utilities industry where R^2 increased from .245% to .7% (186% improvement). Alphas were insignificant in 2 industries, but for the utilities industry, α_{ind} jumped to .81% and was significant).

TABLE 10
REGRESSION R^2 AND ALPHAS FOR INDUSTRY GROUP 5

Industry	α	R^2_{ff}	α_{ind}	R^2_{ind}	$Diff R^2_g$
Utilities	***	0.25	0.81*	0.70	185.71%
Telecom	***	0.77	***	0.87	13.82%
Transportation	***	0.72	***	0.80	10.77%

* Significant at the 5% level
** Significant at the 10% level
*** Insignificant

Industry groups 6 (Wholesale), and 7 (Retail) were represented by one industry portfolio each. Not surprisingly, we also see an improvement in the value of R^2 (10% and 38% respectively) and insignificant alphas (see Table 11).

TABLE 11
REGRESSION R^2 AND ALPHAS FOR INDUSTRY GROUPS 6 AND 7

Industry	α	R^2_{ff}	α_{ind}	R^2_{ind}	$DiffR^2_g$
Wholesale	***	0.87	***	0.95	9.79%
Retail	***	0.71	***	0.98	38.32%

* Significant at the 5% level
** Significant at the 10% level
*** Insignificant

Industry group 8 (Finance, Insurance, and Real Estate) had 4 industry portfolios, and in 3 of them, R^2 increased (see Table 12), albeit by a modest percentage compared to the previous industries. The situation with alphas was mixed as α_{ind} was significant at the 5% and 10% confidence levels in two of the industries (while α_{ff} was significant at the 5% and 10% conference levels). Overall, α_{ind} seems to be higher than α_{ff} for this industry group.

TABLE 12
REGRESSION R^2 AND ALPHAS FOR INDUSTRY GROUP 8

Industry	α	R^2_{ff}	α_{ind}	R^2_{ind}	$DiffR^2_g$
Banking	***	0.84	0.36**	0.92	10.13%
Insurance	0.37**	0.83	0.44*	0.86	3.74%
Real Estate	***	0.82	***	0.80	-2.80%
Finance (Trading)	-0.55	0.87	***	0.89	2.41%

* Significant at the 5% level
** Significant at the 10% level
*** Insignificant

With 22 industry portfolios, Industry group 4 (manufacturing) is by far the most represented in the list of Kenneth French's industry portfolios. These industries range from food, to guns, to ships, to clothes. Given the great variety in this industry group, it is not surprising that in half the industries R^2_{ff} was larger than R^2_{ind}, and in the other half, it was smaller (see Table 13). In fact, the average of the $DiffR^2$ variable across the 23 industries is merely .45%, which again shows that, R^2_{ff} and $R^2_{ind}$are very close. These results are expected though given that this industry group is the largest industry group representing 39% of the total market (in capitalization). This shows that there is probably little difference between the traditional Fama-French factors and the industry-based Fama-French factors for this industry group (note that every industry group had its own unique Fama-French factors). Also given the breadth of the scope of this industry group, it is hard to imagine industry forces that would affect only industries in this group and not others outside it.

TABLE 13
REGRESSION R^2 AND ALPHAS FOR INDUSTRY GROUP 4

Industry	α	R^2_{ff}	α_{ind}	R^2_{ind}	$DiffR^2_g$
Food	0.02*	0.47	0.61**	0.44	-5.59%
Drugs	0.51**	0.62	***	0.63	2.60%
Chemicals	***	0.81	***	0.86	5.65%
Rubber and Plastic	***	0.75	***	0.74	-2.12%

Steel Works Etc.	-1.60*	0.76	***	0.83	9.21%
Fabricated Products	***	0.58	***	0.55	-5.38%
Machinery	***	0.82	0.63**	0.87	6.59%
Electrical Equipment	***	0.80	***	0.82	3.39%
Autos	***	0.74	***	0.74	0.14%
Aero	***	0.76	***	0.79	3.70%
Ships	***	0.67	***	0.66	-1.64%
Guns	1.08**	0.36	***	0.39	8.06%
Chips	***	0.79	***	0.81	3.56%
Paper (Business Supplies)	***	0.86	***	0.87	0.58%
Medical Equipment	***	0.70	***	0.69	-2.14%
Soda	0.76**	0.30	***	0.29	-4.68%
Beer	0.67*	0.42	***	0.42	-1.66%
Smoke	0.57*	0.53	***	0.51	-2.84%
Toys	***	0.51	***	0.55	7.66%
Books	***	0.70	***	0.65	-7.03%
Household (Consumer Goods)	0.13*	0.58	***	0.56	-3.77%
Clothes	***	0.47	***	0.45	-4.22%

* Significant at the 5% level
** Significant at the 10% level
*** Insignificant

For industry group 9, four industries were used and in three of them, $R^2_{ind}$was higher than R^2_{ff} (see Table 14). The improvement though improved from .28% to 14.7%. Only in the case of the healthcare industry was R^2_{ind} smaller than R^2_{ff}. The percentage difference was modest though at -2.91%.

TABLE 14
REGRESSION R^2 AND ALPHAS FOR INDUSTRY GROUP 9

Industry	α	R^2_{ff}	α_{ind}	R^2_{ind}	$DiffR^2_g$
Healthcare	***	0.55	***	0.53	-2.91%
Personal Services	-0.80**	0.71	***	0.71	0.28%
Software	***	0.84	***	0.96	14.68%
Entertainment	***	0.58	***	0.64	11.07%

* Significant at the 5% level
** Significant at the 10% level
*** Insignificant

Overall, though, one can easily conclude, that in the case of individual industry portfolios, the industry-based Fama-French equation seems to fit the data better than the traditional Fama-French equation. It is now time to see though if the same applies to mixed portfolio. Clearly, two distinct cases exist: 1) the industries are in the same industry group, and 2) the industries are in separate industry groups.

Portfolio Formed of Industries in the Same Industry Group

Since I am forming portfolio of industries in the same industry group, I only need to sum the returns of those industries that constitute my portfolios, and, and then run equations 4 and 5 on the excess returns of those portfolios. For illustration purpose, I choose the Entertainment and Software industries, which are both in the Services industry group. The portfolio returns are calculated as follows:

$$Return_{Entertainment+SW} = Return_{Entertainment} + Return_{SW} \tag{11}$$

where $Return_{Entertainment+SW}$ is the return of the mixed portfolio, $Return_{SW}$ is the return of the Software industry portfolio, and $Return_{Entertainment}$ is the return of the Entertainment industry portfolio.

The results from the regression equations 4 and 5 (Table 15) suggest an improvement of 14% in R^2. Note though that as previously shown, when I regressed the two industries portfolios separately R^2 improved by 14.7%, and 11.1% respectively (with an average of 12.8%). This again shows the benefit of using the industry-based Fama-French model in analyzing portfolio returns.

TABLE 15
REGRESSION R2 AND ALPHAS FOR MIXED PORTFOLIOS

Industry	α	R^2_{ff}	α_{ind}	R^2_{ind}	$DiffR^2_g$
Software + Entertainment	***	0.75	***	0.9	0.14%
Coal + Software	-4.4*	0.48	***	0.6	0.31%

* Significant at the 5% level
** Significant at the 10% level
*** Insignificant

Portfolio Formed of Industries in Different Industry Groups

In this section, I form a portfolio of stocks from the Coal (Mining industry group) and Software (Services industry group) with weights equivalent to the sizes of the corresponding industries. The portfolio returns are computed according to the following equations:

$$Return_{Coal+SW} = \frac{PortfolioSize_{Coal}*Return_{Coal}+PortfolioSize_{SW}*Return_{SW}}{PortfolioSize_{Coal}+PortfolioSize_{SW}} \tag{12}$$

where $Return_{Coal+SW}$ is the return of the mixed portfolio, $Return_{Coal}$ is the return of the Coal industry portfolio, $Return_{SW}$ is the return of the Software industry portfolio, $PortfolioSize_{Coal}$ is the size of the Coal industry portfolio (refer to Table 10.), and $PortfolioSize_{SW}$ is the size of the Software industry Portfolio.

For the industry-based equation in this case, clearly some manipulation is needed since I have two sets of industry factors and not one. As an example, I choose the Coal and Software industries for this analysis. The former industry is in the Mining industry group while the latter is again in the Service industry group. The three equations below show how I create the weighted average of the industry-based Fama-French factors for those two industries.

$$MKTRF_{mixed} = \frac{PortfolioSize_{Coal}*MKTRF_{Mining}+PortfolioSize_{SW}*MKTRF_{Services}}{PortfolioSize_{Coal}+PortfolioSize_{SW}} \tag{13}$$

where $PortfolioSize_{Coal}$ is the size of the Coal industry portfolio, $PortfolioSize_{SW}$ is the size of the Software industry Portfolio, $MKTRF_{Mining}$ is the first of the three Fama-French factors and is the value-weighted return on the mining industry group minus the risk free rate, and $MKTRF_{SW}$ is the value-weighted return on the Services industry group minus the risk free rate.
Similarly,

$$SMB_{mixed} = \frac{PortfolioSize_{Coal}*SMB_{Mining}+PortfolioSize_{SW}*SMB_{Services}}{PortfolioSize_{Coal}+PortfolioSize_{SW}} \tag{14}$$

where SMB_{Mining} is the second industry based Fama-French factor and is the average return on three small portfolios minus the average return on three big portfolios for the Mining industry group, and of course $SMB_{Services}$ is the same but for the Services industry group.

$$HML_{mixed} = \frac{PortfolioSize_{Coal}*HML_{Mining}+PortfolioSize_{SW}*HML_{Services}}{PortfolioSize_{Coal}+PortfolioSize_{SW}} \quad (15)$$

where HML_{Mining} is the second industry based Fama-French factor and is the average return on two value portfolios minus the average return on two growth portfolios for the Mining industry group, and of course $HML_{Services}$ is the same but for the Services industry group.

As shown in Table 15, using the industry-based Fama-French to analyze the performance of the mixed portfolio resulted in a 31% improvement in the value of the regression R^2. One can also see that the value of alpha under the traditional Fama-French model was -4.41% implying that the portfolio was performing below the market, while alpha under the industry-based Fama-French model was not statistically different from zero, implying that the portfolio performed as well as the industries that comprised it.

The above 2 examples show that there is some evidence to suggest that also in the case of mixed portfolios, the industry-based Fama-French equation fits the data better than the traditional Fama-French equation with a clear improvement in the regression R^2value. Note though that I only used two examples of mixed portfolios in this research, and there are many other combinations of industries that can be used to form portfolios. Overall, though, I believe it is beneficial to use the industry-based Fama-French model to in conjunction with the traditional Fama-French equation to evaluate the performance of investment portfolios.

CONCLUSION AND DISCUSSION

In this paper I analyze the returns of 41 monthly value-weighted industry portfolios (downloaded from the Kenneth French web site) using both the Fama-French and the industry-based Fama-French models. I repeat the same exercise for monthly equal-weighted industry portfolios and do not observe any major changes in the results. I consider three cases in my analysis:

1- Single industry portfolio
2- Portfolio formed of industries in the same industry group, and
3- Portfolio formed of industries not in the same industry group

My results suggest that R^2 consistently improves (except for the Manufacturing industry group). The regression R^2 for the Gold industry for example, increased by 1052% (from .3% to 28%). Similarly, the regression R^2 for the agriculture, construction, and coal industries increased by 95%, 69% and 32% respectively. As a matter of fact, for the 18 non-manufacturing industry portfolios, the regression R^2 increased in 16, and slightly decreased in only 2.

The manufacturing industry group though is unique as it is the largest and most diverse industry group within my sample with 22 industries that represents 39% of the total market in market capitalization (within the sample). This industry group covers diverse industries like ships, guns, food, etc. which do not always have a lot in common. It is thus not surprising that the differences in R^2 was positive in half the industries (R^2_{ind} is larger than R^2_{ff}) and negative in the other half (R^2_{ff} was larger than R^2_{ind}). In fact, the average value of the DiffR^2 variable across the 22 industries is merely .45% which again shows that R^2_{ff} and R^2_{ind} are very close. These results are expected given that there is no large difference between the traditional Fama-French factors and the industry-based Fama-French factors for this industry group (Elhadary, 2019).

Although the improvement in the regression R^2 seems conclusive for portfolios formed with stocks in the same industry group (except for the manufacturing industry group), the results are only suggestive with portfolios made up of stocks in multiple industries. This is primarily because I only used 2 examples of mixed portfolios in this research, and there are clearly many other combinations of industries that can be used to create portfolios. In the 2 examples though R^2 did improve when I used the industry-based model versus the traditional Fama-French model. Overall, though, I believe it is beneficial to use the industry-based Fama-French model to in conjunction with the traditional Fama-French equation to evaluate the performance of investment portfolios.

REFERENCES

Ang, A., Hodrick, R.J., Xing, Y., & Zhang, X. (2006). The Cross-Section of Volatility and Expected Returns. *The Journal of Finance*, *61*(1), 259–299.

Asgharian, C., & Master, H. (2017). *Testing the Capm and the Fama-French 3-Factor Model on U.S. High-Tech Stocks* [Online]. Lund University School of Economics and Management. Retrieved from http://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=8910802&fileOId=8910803

Asness, C.S., Friedman, J.A., Krail, R.J., & Liew, J.M. (2000). Style timing: Value versus growth. *The Journal of Portfolio Management*, *26*(3), 50-60.

de Vries, M. (2012, July 6). *Asset Pricing Models and Industry Sorted Portfolios*. (U.v. Tilburg, Producer) Retrieved from arno.uvt.nl/show.cgi?fid=129542

Elhadary, O. (2019). *New Factor Structure Models and Idiosyncratic Volatility* (Doctoral dissertation). City University of New York.

Fama, E.F., & French, K.R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, *47*(2), 427–465.

Fama, E.F., & French, K.R. (1992). The Cross-section of Expected Stock Returns.pdf. *The Journal of Finance*, *47*(2).

Fama, E.F., & 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., & French, K.R. (1996). Multifactor Explanations of Asset Pricing Anomalies. *The Journal of Finance*, *51*(1), 55–84.

Fama, E.F., & MacBeth, J.D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*, *81*(3), 607–636.

Hu, O. (2003). *Forecasting ability of the Fama and French three-factor model - Implications for Capital Budgeting*.

Moskowitz, T.J., & Grinblatt, M. (1999). Do industries explain momentum? *The Journal of Finance*, *54*(4), 1249-1290.