Jun Pan
junpan@mit.edu
MIT Sloan School of Management
15.433-15.4331 Financial Markets, Fall 2015
253-3083
E62-624

Classes 4 & 5: Equity in the Cross Section, Part 1

This Version: September 26, 2016

When I presented the timeline of Modern Finance in class, one of the students asked why the blue events (work developed by academics) stopped after the 1970s. Of course, we academics kept writing papers, at an even faster rate. But the 1970s were a great time to do Finance in the academic world. Theoretical papers that are foundation building and trail brazing happened in that era. Since the 1990s, most of the exciting work in Finance happened in the empirical area.

What we are going to cover in this class represents the most influential work in Finance since 1990s. And the intellectual leader is Prof. Eugene Fama, who was awarded a Nobel prize in 2013. The ideas behind the research papers helped inspire and create this fast growing area called quant investing in the late 1990s and early 2000s. More recently, with the growing popularity of factor investing, the mutual fund and ETF world is also incorporating these ideas.

As a PhD student at Stanford GSB in the late 1990s, I didn't know much about the Fama-French factors. I was into my own research at that time. Fortunately, I had to teach 15.433 at MIT Sloan. So it was through having to teach the MBA students at Sloan that I got to learn and admire the work of Prof. Fama and his co-authors.

1 Quant Investing

Both quant investors and stock pickers are interested in generating alpha, but they differ in their approach. To argue which approach, stock picker or quant investing, is better is meaningless, but to find out which one suits you better is extremely important. To quote a recent column by John Authers from the *Financial Times*, "If we do want to try to do better than passive then there are two logical ways to do it. Either we can adopt a tightly disciplined approach designed to exploit persistent market anomalies or factors; or we can focus tightly on a sector or industry and make concentrated bets with high conviction." So on the one end of the spectrum of this alpha generating business is an investor like Warren Buffett, who makes concentrated bets with high conviction; and on the other end are many long/short equity quant funds using quant signals to exploit persistent patterns in the cross-section of stocks. The Warren Buffett path has clearly been admired and traveled by many but there are limited number of success stories. After all, how can you replicate a person's mind? The quant approach, on the other hand, has generated relatively more success stories. This approach serves more as a tool and it is much easier to replicate. Not surprisingly, it ended up as an over-crowded field. In terms of skills, quant investing really does not require a lot of quantitative skills in the traditional sense. It requires a curious and creative mind, love and respect for the data, and some basic programming skills such as regressions and data cleaning, merging, and sorting.

The first observation of quant investing is that even within the US equity markets, there are thousands of stocks to choose from. For a stock picker, it would be a daunting task to cover them all. With the availability of computers and data, it seems obvious that they should develop a systematic approach to search through the data for alpha. This, of course, assumes that the patterns found in the data persist in the near future. This is where quantitative signals come in. The key insight is that such quant signals are useful in separating one group of stocks (high alpha) from another (zero or negative alpha). Potentially, there are two interpretations or reasons as to why such signals might work. First, they help us exploit the mis-pricing in the markets. Second, they represent differences in exposure to certain risk factors (that are unrelated to the market portfolio). Subscribing to the first interpretation, you believe that your alpha comes from market inefficiency. The second line of reasoning leads you to believe that your alpha comes from exposures to certain systematic risk (that is unrelated to the market portfolio). In this case, the alpha's are simply beta's in disguise.

One signature approach of quant investing is forming portfolios. This arises from the desire to be exposed only to the risk (or anomaly) one is interested in. The portfolio approach helps diversify away unwanted idiosyncratic risk. Another signature approach of quant investing in the hedge fund world is the long/short strategy. Again, this arises from the desire to have a razor sharp focus on the target risk factor. The long/short strategy helps take out the unwanted systematic risk (e.g., the market risk). The best place to learn about quant investing is to read carefully the tables in Fama and French (1992, 1993). Afterwards, go to Prof. French's website and play with the data.

The most creative part of quant investing is to come up with signals that could generate alpha, especially those signals that help us identify market inefficiency in the cross-section. Unfortunately, most of the signals used by quant funds have their origin in academic papers and, in my opinion, are not that creative. It either indicates that markets are not that inefficient, or quant funds are not that creative.

The need for more innovation in this area certainly shows up during the recent quant meltdown in August 2007. What we learned from this event is that the quant investing space is very crowded, populated by funds with very similar ideas. The initial success of quant investing in the 1990s attracted many investors, and the quant investing world enjoyed a great rise in the first half of 2000s.[1] It turns out that many quant funds are trading on very similar signals. Prof. Daniel, who was at Goldman during the quant meltdown, wrote an interesting and informative set of slides. The initial trigger was in the sub-prime mortgage market, and it spilled over to investment-grade credit markets shortly thereafter. As multi-strategy hedge funds experiences losses in their illiquid mortgage and credit positions, they liquidated their more liquid assets in the quant investing side to raise cash. As this unwinding took place, many quant investing funds rushed to the door, triggering a 20-sigma move in the quant investing space. Previously unrelated stocks suddenly started to move together during the unwind. If you were not in the quant space, you probably would not have noticed the 20-sigma move. But if you are in the quant space, then most likely your portfolio experienced a 10 to 20 sigma drop over one week.

In recent years, the basic ideas in quant investing have found their popularity in the world of mutual funds and ETFs. While the sales pitch in the quant hedge fund world is all about Alpha, now the emphasis is on Beta: smart beta and factor investing. In any case, if you are interested in a career in this area, what we are going to cover in the next few classes is going to be very useful. Coming straight from the original research papers, it is also the gold standard.

2 Forming Portfolios using Quantitative Signals

- **Popular quant signals:** Quant investing uses stock characteristics as signals. Most quant investors believe that their signals help capture the fundamentals that drive alpha. Here is a list of widely adopted quant categories and strategies:
  - Size: The market capitalization = stock price × number of shares outstanding.
  - Valuation: How is the company priced relative to fundamental accounting measure? For this, we have the widely used book-to-market ratio:

$$\texttt{BtM} = \frac{\text{book value of equity}}{\text{market value of equity}}\,.$$

[1] I often infer the popularity of a field from the number of Finance professors (whom I personally know) it managed to attract to switch jobs. In mid-2000s, I observed quite a few.

- Momentum: How has the market responded to the company's changing fortunes? Sample Metric: `price momentum`.
- Profitability: What are the company's profit margins? How efficient are its operations? Sample metric: `earnings-to-sales ratio` and `OP` (operating profitability) in the five-factor model of Fama and French.
- Earnings Quality: Were earnings derived from sustainable sources? Sample metric: `the accruals-to-total-assets`.
- Analysts Sentiment: Are analysts upgrading or downgrading their view of this company? Sample metric: `earnings forecast revisions`.
- Management Impact: How is the company's management employing its capital? Sample metric: `change in shares outstanding` or the `Investment` variable (growth in firm assets) in the five-factor model of Fama and French.

In coming up with the above list, I mostly used the information from Prof. Daniel's slides. In addition, I also listed the two new Fama-French variables, Profitability and Investment, from their recent five-factor model. We will cover size, value, and momentum in detail. For most of the other signals, Googling will lead you to the key research articles behind these strategies.

These signals differ in various ways. Some are momentum signals (e.g., earnings forecast revisions), indicating a slow reaction to information. Some are contrarian signals (e.g., valuation), indicating a reversal in price pattern due to over-reactions in the past. Some are over a long horizon. For example, studies on change in shares outstanding (due to seasoned equity offerings or share repurchase announcements) focus on returns with holding periods of 3 or more years. Some are over a horizon of a few months (e.g., momentum).

- **Sorting stocks into portfolios:** The concept of sorting is pretty straightforward. Of course, there are many details one needs to pay attention to. The best resources are Fama and French (1992), which by now is the gold standard in this area. Prof. French's website also provides a great deal of information. It should be mentioned that Prof. French offers a tremendous service to our profession by making the data available on his website. If I didn't have access to the materials posted on his website, I would have to construct a lot of the tables and plots in this class from scratch.

  In this class, we will first look at univariate sorts (by size or book-to-market) into deciles, and then move on to double sorts (by size and book-to-market) into 5x5. One important convention is that the breakpoints of these sorts are first established by using NYSE stocks only. The main reason is that the stock population in NYSE is more representative.

It is also important to emphasize that sorting is done dynamically. Stock characteristics fluctuate over time. So we need to periodically update this information and re-sort stocks by their new characteristics so that the sorted portfolios contain stocks of the right characteristics. The frequency of sorting depends on the variability of the signals. For example, Fama and French sort their size portfolios once a year, using the market value in June of year t for portfolio returns from July of year t to June of t+1. For the momentum portfolios, however, the signal is the stock's past returns, which are more variable, and the sorting is done at a monthly frequency. More generally, the variability of a signal also affects the portfolio turnover. For a signal such as market cap, the portfolio turnover is low because market cap is relatively stable. By contrast, for a momentum signal such as past stock returns, the portfolio turnover could be quite high. All of these considerations could factor into the execution costs of a strategy.

- **Size and BtM sorted portfolios:** The size-sorted deciles are useful in our understanding of the overall size distribution of stocks listed on the three US exchanges. Using the 2015 number, we see that the average market cap is a mere $111 millions for stocks in decile 1, which contains 1362 stocks. By contrast, decile 10 has only 173 stocks. Given that the breakpoints are determined by NYSE stocks, this implies that most of the AMEX and Nasdaq stocks fall into the smaller decides. For stocks in decile 10, the average market cap is close to $84 billions. Of course, this is still no comparison to those mega-large stocks such as Google ($427B), Apple ($661B), or Amazon ($244).

  The book-to-market sorted deciles give us a sense of how much the equity value of a firm differ from its book value. For some stocks, equity investors value the stocks to the extent that they are willing to pay much more than the existing book value of its equity. As a result, the market value of the equity takes into account the firm's future growth component, which is not reflected in the firm's current book value. Such growth stocks are of low book-to-market ratio and show up in the lower deciles. For example, using the 2015 number, the average book-to-market ratio of stocks in decile 1 is 0.095: for each dollar in market value, the book value is only 0.095. Or, for each dollar in the book value, the market is willing to pay 1/0.095=10.5 dollars. You can imagine that Google was once a growth stock. Back in 2006, Google had a book-to-market ratio of 0.04 and its market cap was $107B. Right now, its price-to-book is 3.84 according to Yahoo Finance. So its book-to-market ratio has gone up quite a bit in the past 10 years to it current level of 0.26.

At the other end of the spectrum are stocks with very high book-to-market ratio. These stocks, usually referred to as value stocks, have a depressed market value. Using the 2015 number, we see that stocks in our decile 10 have an average book-to-market ratio of 1.339.[2] Basically, investors are not willing to pay the full book value for the stock. For example, back in 2006, the book-to-market ratio of GM was 1.28. For each dollar in book value, investors are only willing to pay only 1/1.28=0.78 dollar in the stock market. On the morning of June 1, 2009, GM filed for bankruptcy protection. In general, firms have a high book-to-market ratio prior to filing for bankruptcy, but this does not mean that high book-to-market firms are bankruptcy firms.

At this point, it is worthwhile to emphasize again that sorting is done dynamically. For example, back in 2006, GM, with its book-to-market value of 1.28, showed up in the book-to-market decide 10. After its filing for bankruptcy protection, GM dropped out of the sample. As of today (September 14, 2015), according to Yahoo Finance, GM has a price-to-book ratio of 1.37, indicating a book-to-market ratio of 1/1.37=0.7299. So now GM shows up in decile 7 or 8.

3 Testing the CAPM using Fama-French 25 Portfolios

Let's start with the regression:

$$R_t^i - r_f = \alpha_i + \beta_i \left(R_t^M - r_f \right) + \epsilon_t^i \,, \tag{1}$$

where R_t^i is the month-t return of a portfolio i. Recall that testing the CAPM pricing equation is equivalent to testing whether or not α_i is significantly different from zero. If we can find many portfolios with large α's, then the CAPM will be in trouble. Indeed, this is at the heart of what we are going to do.

- **Use the CAPM beta:** We use the famous Fama-French 25 portfolios to test the CAPM. For each portfolio i, we run the regression in Equation (**??**) to obtain its β_i. After obtaining an estimator for the market risk premium λ^M, we calculate the risk premium for portfolio i according to the CAPM: $\beta_i \, \lambda^M$. We call this number the risk premium predicted by the CAPM. At the same time, we use the realized returns of portfolio i to estimate the risk premium directly. We call this number the risk premium

[2] The reported average BtM is value weighted. That is, within each decile, we value-weight each stock's book-to-market ratio by its size to calculate the average BtM for the decile.

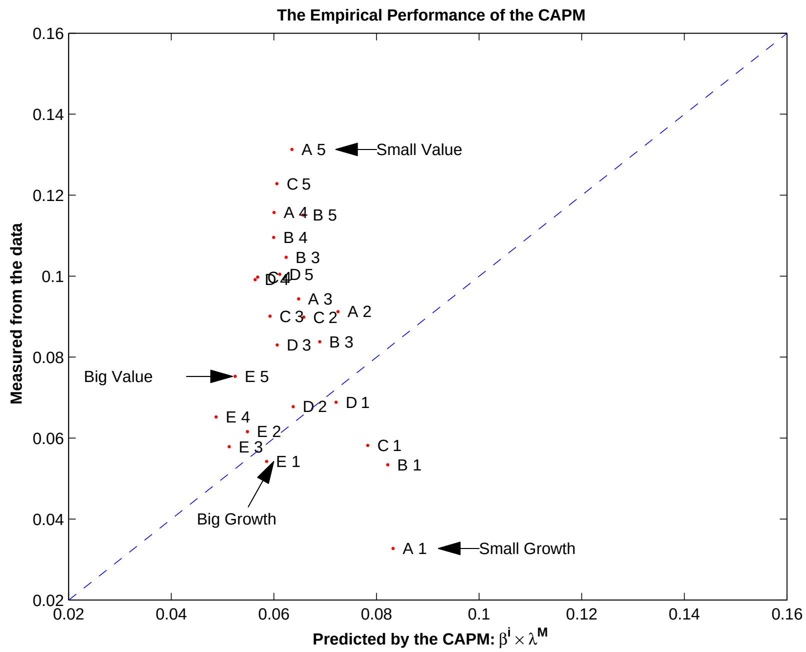

Figure 1: The empirical performance of the CAPM, using the Fama-French 25 portfolios. For each portfolio i, its risk premium measured from the data is plotted against that predicted by the CAPM.

measured from the data. We now have 25 pairs of numbers, each pair corresponds to one of the Fama-French 25 portfolios.

Figure **??** plots the 25 pairs of numbers: the risk premium measured from the data (y-axis) and the risk premium according to the CAPM (x-axis). I understand that both pairs of numbers are noisy because they are estimated from the data. But let's use them for now, and we will come back to a proper test later. Also, although the estimations and regressions are all done using monthly returns, I annualized the risk premium (by multiplying the monthly risk premiums by 12) for ease of communication.

Now let's come back to Figure **??**. If the CAPM works well, then these 25 dots should line up pretty nicely along the 45-degree line: data and model in agreement. In practice, however, most of these dots are clustered together along the x-axis dimension and spread out along the y-axis dimension. Recall that plotted along the x-axis is model-implied risk premium: $\beta_i \lambda^M$. So effectively, the "clustering" implies that most of the 25 portfolios have very similar β. Moreover, given that the market risk premium λ^M is close to 6% per year, this implies that most of the 25 portfolios have a β that is very close to one. On the other hand, the wide variation along the y-axis dimension indicates that those portfolios in fact perform very differently in reality: some perform well with a high risk premium, while some perform poorly with a low risk premium.

Overall, Figure **??** is not good news for the CAPM. Instead of the predicted relation between risk premium and beta, we find a wide range of risk premiums for portfolios that are very similar in beta. At this point, you might ask how much we should trust those 25 pairs of numbers, which are estimated with noise. So let's do our test more properly. Recall that the key to the test is Alpha. We can estimate the alpha's from the plot by measuring the vertical distance between each dot and 45-degree line. From the plot, we see that some portfolios (mostly value stocks) have positive alpha's, while other portfolios (mostly growth stocks) have negative alpha's.

- **Use the CAPM alpha:** Recall that the alpha's can be easily obtained from the regression in Equation (**??**). Table **??** reports, for the 25 portfolios, the CAPM α, β, and the adjusted R-squared from the 25 regressions. For those portfolios with statistically significant α's, I print the number in bold. The t-stat's of the α's are also reported in the table. According to the CAPM, all of the α's should be indistinguishable from zero. But we have quite a few portfolios with statistically significant α's. Moreover, there is a pattern to it. For example, for small stocks in group A, moving along the book-to-market dimension, the portfolio α's turned from negative (and statistical significant) to positive. The same pattern repeated for all size groups. For large stocks in group E, none of the α's are statistically significant, but you can see the magnitude of α's increasing as we move the book-to-market from low to high.

  To jointly test the statistical significance of those 25 α's, we can use the GRS test, named after Gibbons, Ross, and Shanken (1989). It is actually a very cool test. It maps the joint α test to how inefficient the market portfolio R^M is. If you recall, the CAPM tells us that the market portfolio is the tangent portfolio sitting at the mean-variance frontier with the highest Sharpe ratio. By being able to construct portfolios with positive alpha's, the story breaks down: the market portfolio is no longer the mean-variance efficient portfolio.

- **The importance of the CAPM:** Before closing this section, I would like to emphasize one more time that this test result does not hurt the importance of the CAPM model in Finance. In fact, without the model, we will not even know where and how to start the test. Moreover, the later development, including the Fama-French models we will see, always includes the market portfolio in the test. Indeed, the CAPM model serves as the foundation for all models to come. I have yet to see one model without the market portfolio in it. Finally, the main insight of the CAPM remains: there are

Table 1: The Fama-French 25 Portfolios in the CAPM and the Fama-French Three-Factor Model. All α's are reported in annualized terms (x12). Statistically significant α's are reported in bold. Monthly data from January 1962 to July 2015.

	CAPM				The FF Three Factor Model					
Portfolio	α (%)	t-stat	β	R2 (%)	α (%)	t-stat	β	s	h	R2 (%)
A1	**-5.05**	-2.19	1.41	62.81	**-5.32**	-4.69	1.06	1.38	-0.29	91.25
A2	1.88	0.95	1.23	63.50	-0.10	-0.13	0.96	1.30	0.04	94.20
A3	2.95	1.80	1.10	66.77	-0.09	-0.15	0.92	1.10	0.28	95.20
A4	**5.57**	3.46	1.02	64.34	**1.65**	2.57	0.89	1.03	0.46	94.48
A5	**6.78**	3.82	1.08	62.38	**1.49**	2.21	0.98	1.09	0.70	94.71
B1	-2.88	-1.68	1.39	74.93	**-2.09**	-2.73	1.11	0.99	-0.39	95.10
B3	1.49	1.08	1.17	76.33	-0.42	-0.62	1.01	0.87	0.13	94.35
B3	**4.23**	3.27	1.06	75.07	1.07	1.63	0.97	0.77	0.39	93.74
B4	**4.96**	3.78	1.02	73.07	0.89	1.43	0.97	0.73	0.56	94.15
B5	**4.94**	3.06	1.11	68.12	-0.66	-1.00	1.08	0.87	0.81	94.77
C1	-2.01	-1.41	1.33	79.58	-0.60	-0.84	1.09	0.73	-0.44	95.03
C2	**2.40**	2.23	1.12	82.83	0.67	0.85	1.04	0.53	0.18	91.10
C3	**3.08**	2.83	1.00	79.31	0.08	0.10	0.99	0.44	0.44	89.73
C4	**4.29**	3.68	0.96	75.49	0.38	0.50	1.00	0.40	0.62	90.18
C5	**6.22**	4.31	1.03	69.61	1.23	1.44	1.06	0.55	0.77	89.58
D1	-0.32	-0.30	1.22	85.24	**1.46**	2.05	1.06	0.38	-0.42	93.73
D2	0.40	0.45	1.08	86.89	-1.03	-1.25	1.08	0.22	0.21	89.15
D3	**2.24**	2.21	1.03	82.26	-0.44	-0.52	1.08	0.18	0.45	88.26
D4	**4.28**	3.96	0.96	77.91	0.85	1.09	1.02	0.22	0.57	88.79
D5	**3.94**	2.81	1.04	71.14	-0.84	-0.89	1.14	0.25	0.81	87.45
E1	-0.43	-0.56	0.99	88.52	**1.88**	3.44	0.98	-0.24	-0.36	94.19
E2	0.68	0.91	0.93	87.53	0.47	0.71	0.99	-0.22	0.09	90.24
E3	0.66	0.70	0.87	79.61	-0.65	-0.83	0.97	-0.23	0.30	86.20
E4	1.65	1.50	0.83	71.88	**-1.38**	-2.03	0.98	-0.20	0.60	89.41
E5	2.28	1.57	0.89	62.79	-1.76	-1.65	1.05	-0.08	0.76	80.48

undiversifiable risks in the market, and you get rewarded for bearing this kind of risk. There is no reward for holding diversifiable risk. In fact, it is this insight that prompts people to locate new risk factors. The one-factor structure of the CAPM might not work very well with the data, and the new multi-factor models might work better. But all builds on this insight of locating systematic risk factors.

4 The Fama-French Three Factor Model

Our test of the CAPM informs us on how the CAPM failed to price the Fama-French 25 portfolios: value stocks outperform growth stocks, and small stocks outperform big stocks. In the one-factor model of CAPM, the only risk factor is the market portfolio and the only measure of risk is beta. There is no additional role for size or value in the model. So the logical next step is to build a model that incorporates these two factors. This is what Fama and French did in their 1993 paper by introducing the SMB and HML factors.

- **The Fama and French factors:** In order to construct the factors, Fama and French use a coarser double sort. Along the size dimension, stocks are sort into two groups: small or big. Along the value dimension, stocks are sort into three groups with 30% in value, 40% in neutral, and 30% in growth. Because these portfolios are to be used to construct factors, one would like to have them as diversified as possible. A coarser sort would allow each bin to have more stocks and therefore improve diversification. Using the 6 (2x3) portfolios, the SMB and HML factors are constructed as

  – SMB (Small Minus Big):

  $$R^{\text{SMB}} = R^{\text{small}} - R^{\text{big}},$$

  where R^{small}=1/3 (small value + small neutral + small growth) and $R^{\text{big}} = 1/3$ (big value + big neutral + big growth)

  – HML (High Minus Low):

  $$R^{\text{HML}} = R^{\text{value}} - R^{\text{growth}},$$

  where R^{value}=1/2 (small value + big value) and R^{growth}=1/2 (small growth + big growth).

  As you can see, the factors are constructed by a long/short strategy. For example, the HML factor involves buying value stocks and selling growth stocks. The motivation behind such a factor is to help investors focus on the targeted risk factor, which is the difference between values and growth stocks. Any unwanted risks are taken out from the factor: the portfolio approach diversifies away the idiosyncratic risk in individual stocks and the long/short strategy hedges out the exposure to the market risk.

As you might notice, SMB and HML are not totally orthogonal to the market risk. The beta's of small stocks are usually higher than those of big stocks. As a result, SMB has a slightly positive beta (around 0.20). The beta's of growth stocks are usually higher than those of value stocks. As a result, HML has a slightly negative beta (around -0.2). A purist would not like this. Nevertheless, by choosing to form their factors using such a simple long/short strategy, Fama and French seem to value simplicity and intuitiveness over perfection. I would have done the same thing given the cost and benefit.

- **The three-factor regression:** Now we are ready to run the following regression for our 25 portfolios:

$$R_t^i - r_f = \alpha_i + \beta_i \left(R_t^M - r_f\right) + s_i \, R_t^{\text{SMB}} + h_i \, R^{\text{HML}} + \epsilon_t^i \tag{2}$$

Notice that we've put the two new factors in the regression and label the corresponding slope coefficients to be s and h. If you like, you can think of them as the "beta" on SMB and HML.

Since the new factors are slightly correlated with the existing factor, $R^M - r_f$, the β in the current regression is no longer the CAPM beta. In fact, using Table **??**, we can see that the β from this new regression are slightly different from the CAPM β. The benefit of using SMB and HML is that they are very simple to construct and also very intuitive. As we will work in a regression framework for the three factor model, having slightly correlated factors is not a problem at all. Just be careful with the interpretation of the new β.

Table **??** also reports the values for the SMB beta s and the HML beta h. As we move along the size dimension from group A to E, the estimated numbers for s move from positive to negative. Likewise, as we move along the value dimension, from group 1 to 5, the estimated numbers for h move from negative to positive. It tells us that indeed there is commonality in movement among small stocks that is different from large stocks. Regressing returns of small stocks on the SMB factor picks up this comovement. Similarly, values stocks comove together in ways that are different from growth stocks. Hence the HML factor. Overall, these regression results tell us that size and value are not simply characteristics. Putting small stocks together against big stocks actually forms a factor. Likewise, putting value stocks together against growth stocks forms another factor.

Comparing the R-squared numbers in the one-factor regression with the three-factor regression also tells a similar story. For example, the R-squared's for small stocks (in Groups A) are around 60% in the one-factor regression. In the three-factor regression, the R-squared's increase to around 90%. Of course, having two more factors always improve the R-squared, but not by this much. This is telling us that the SMB factor is picking additional commonality in small stocks.

- **The Fama-French three-factor model:** Borrowing from the CAPM, the pricing relation of the three-factor model is pretty straightforward:

  $$E(R_t^i) - r_f = \beta_i \left(E\left(R_t^M\right) - r_f \right) + s_i \, E\left(R_t^{\text{SMB}}\right) + h_i \, E\left(R_t^{\text{HML}}\right) \, .$$

  The risk premiums for the three factors can be estimated using the historical data. Overall, the size premium is somewhat weak in recent periods. The estimated size premium is 3.20% with a t-stat of 1.68. So it is not really significant. The value premium is stronger: 5.15% with a t-stat of 2.78. For the same sample period from 1962 to 2014, the market risk premium is 6.46% with a t-stat of 2.64.

  The empirical performance of the Fama-French three-factor model is plotted in Figure **??**. Comparing this plot against the one for the CAPM, we can see a clear improvement. By now, we are not surprised that it would work. In the three-factor model, small stocks have a positive factor loading on SMB and are compensated for this exposure. So the model-predicted risk premium is higher than that in the CAPM, where only beta matters. Likewise, value stocks have a positive factor loading on HML and are compensated for this exposure. As a result, in Figure **??**, the dots for those portfolios in groups 4 and 5 move horizontally to the right, while those in group 1 move horizontally to the left. So effectively, by having the two added dimensions along size and value, the model performs better.

- **Use the Fama-French three factor model:**

  The three-factor model can be used as a benchmark model to evaluate the performance of fund managers. For example, you can put Peter Lynch's performance on the left hand side and regress it against the three factors. You can investigate his exposures to the factors and evaluate how much of his performance derives from such exposures.

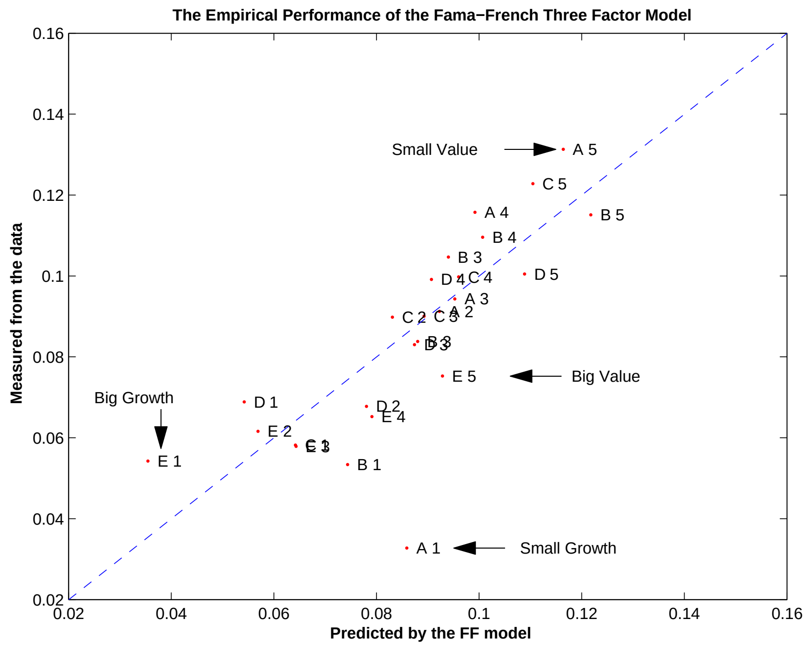

Figure 2: The empirical performance of the Fama-French three-factor model, using the Fama-French 25 portfolios. For each portfolio i, its risk premium measured from the data is plotted against that predicted by the model.

The alpha from the regression tells you the magnitude of his performance that cannot be explain by the three factors.

A fund manager might have a pretty nice looking CAPM alpha, but when evaluate his performance against the three factor model, his three-factor alpha might be insignificant. This implies that most of his CAPM alpha in fact comes from exposures to the size or value factor. This is what people mean when they say "beta in disguise." For this fund manager, his CAPM alpha actually comes from a beta exposure to a previously unknown risk factor called size or value. Maybe this is why as this quant investing approach moves into the world of mutual funds and ETFs, people are not selling them as alpha's anymore. Instead, they are emphasizing on beta and risk factors.

Appendix

A On Running Multivariate Regression

Many students do not like the fact that SMB and HML are not orthogonal to the market portfolio. For example, using annual data from 1962 to 2014, let's regress SMB on the market:

$$R_t^{SMB} = \alpha^{\text{SMB}} + \beta^{\text{SMB}}\left(R_t^M - r_f\right) + \epsilon_t \,.$$

We have a CAPM beta of 0.22, indicating that small stocks on average have a slightly higher beta than large stocks. Run the same regression using the HML factor:

$$R_t^{HML} = \alpha^{\text{HML}} + \beta^{\text{HML}}\left(R_t^M - r_f\right) + \epsilon_t \,.$$

you get a CAPM beta of -0.21, indicating that growth stocks on average have a slightly higher beta than value stocks.

How does this affect our multivariate regression in Equation (**??**)? The only real effect is that the beta in the three-factor regression is no longer the CAPM beta. Other than this, I cannot think of any significant "damage" of having a factor that is slightly correlated with the market. Of course, Fama and French form their factors using this long/short strategy exactly to take out the market component. As in many situations, simplicity is preferred. In this case, it is really more simple and intuitive to use SMB and HML. If the cost is not being able to read the CAPM beta directly from the three-factor regression, then it is an acceptable cost.

Recall that we call $E(R^{\text{SMB}})$ and $E(R^{\text{HML}})$ the value and size premiums. If we want to be really careful, we should call them the average returns of SMB and HML. The alpha of the above regression, α^{SMB} gives us the true performance of SMB: 1.76% with a t-stat of 0.91. And α^{HML} is 6.51% with a t-stat of 3.44. So indeed, the size premium is small and insignificant for the period from 1962 to 2014, while the value premium is pretty strong.

Also, in the above regressions, you will never put $R_t^{\text{SMB}} - r_f$ on the left hand size. This is because R_t^{SMB} is already a long/short portfolio. If you really want, you could do

$$R_t^{\text{small}} - r_f = \alpha^{\text{small}} + \beta^{\text{small}}\left(R_t^M - r_f\right) + \epsilon_t \,,$$

or

$$R_t^{\text{big}} - r_f = \alpha^{\text{big}} + \beta^{\text{big}}\left(R_t^M - r_f\right) + \epsilon_t \,.$$

Moreover, you notice that $\alpha^{\text{SMB}} = \alpha^{\text{small}} - \alpha^{\text{big}}$ and similarly for beta.

B Matlab Code

Code 1: Test the models using the Fama-French 25 portfolios

```
n_model=input('which model? Market (1); FF 3 Factor (2)');

load FF_Factors.txt;
start_time=196201;
end_time=201606;
FF_Factors=FF_Factors(FF_Factors(:,1)>=start_time & FF_Factors(:,1)<=
    end_time,:);
Market=FF_Factors(:,2)/100;
SMB=FF_Factors(:,3)/100;
HML=FF_Factors(:,4)/100;
RF=FF_Factors(:,5)/100;

switch n_model,
  case 1, X=Market;
  case 2, X=[Market SMB HML];
end

load FF_Portfolio_25.txt;
FF_Portfolio_25=FF_Portfolio_25(FF_Portfolio_25(:,1)>=start_time &
    FF_Portfolio_25(:,1)<=end_time,:);
n_Portfolio=size(FF_Portfolio_25,2)-1;
Portfolio=FF_Portfolio_25(:,2:end)/100-kron(RF,ones(1,n_Portfolio));
Name=['A1';'A2';'A3';'A4';'A5'; ...
      'B1';'B3';'B3';'B4';'B5'; ...
      'C1';'C2';'C3';'C4';'C5'; ...
      'D1';'D2';'D3';'D4';'D5'; ...
      'E1';'E2';'E3';'E4';'E5'];

% output for alpha, beta, and R2 tables
if n_model==1, beta_CAPM=[]; alpha_CAPM=[]; R2_CAPM=[]; end
if n_model==2, beta_FF3=[]; alpha_FF3=[]; R2_FF3=[]; end
for i=1:n_Portfolio,
```

```
    [b,R2]=Reg_OLS(Portfolio(:,i),X);
    switch n_model,
    case 1,
      R2_CAPM=[R2_CAPM; R2*100];
      beta_CAPM=[beta_CAPM; b(1,2:end)];
      alpha_CAPM=[alpha_CAPM; [b(1,1)*12*100 b(3,1)]];
    case 2,
      R2_FF3=[R2_FF3; R2*100];
      beta_FF3=[beta_FF3; b(1,2:end)];
      alpha_FF3=[alpha_FF3; [b(1,1)*12*100 b(3,1)]];
    end
end

if n_model==1, beta_out=beta_CAPM; else, beta_out=beta_FF3; end;
Y=mean(Portfolio)';
Y_fitted=beta_out*mean(X)';

figure(n_model); plot(Y_fitted*12,Y*12,'r.')
switch n_model,
case 1, title('\bf The Empirical Performance of the CAPM');
case 2, title('\bf The Empirical Performance of the Fama-French Three Factor
     Model');
end
axis([0.02 0.14 0.02 0.14])
hold on;
for k=1:n_Portfolio
     text(Y_fitted(k)*12+0.001,Y(k)*12,[' ' char(Name(k,1))]);
     text(Y_fitted(k)*12+0.004,Y(k)*12,[' ' char(Name(k,2))]);
     if n_model == 1,
       if k==5, text(0.080,Y(k)*12,'Small Value'); arrow([0.080,Y(k)
           *12],[0.072,Y(k)*12]);end
       if k==1, text(0.10,Y(k)*12,'Small Growth'); arrow([0.10,Y(k)
           *12],[0.092,Y(k)*12]);end
       if k==25, text(0.023,Y(k)*12,'Big Value'); arrow([0.043, Y(k)
           *12],[0.052,Y(k)*12]); end
       if k==21, text(0.045,0.040,'Big Growth'); arrow([0.055 0.043],[0.06,Y(
           k)*12]); end
     else
```

```
        if k==1, text(0.108,Y(k)*12,'Small Growth'); arrow([0.105,Y(k)
           *12],[0.095,Y(k)*12]);end
        if k==5, text(0.083,Y(k)*12,'Small Value'); arrow([0.105,Y(k)
           *12],[0.115,Y(k)*12]);end
        if k==25, text(0.118,Y(k)*12,'Big Value'); arrow([0.116, Y(k)
           *12],[0.106,Y(k)*12]); end
        if k==21, text(0.025,0.07,'Big Growth'); arrow([0.038 0.067],[0.038,Y(
           k)*12+0.003]); end
    end
end
if n_model == 1,
 xlabel(['\bf Predicted by the CAPM: \beta^i \times \lambda^M'] );
else,
 xlabel(['\bf Predicted by the FF model'] );
end
ylabel('\bf Measured from the data')
hold on
plot([0.02 0.16],[0.02 0.16],'b--')
hold off
```

Code 2: My OLS Regression Function

```
function [out,R2]=Reg_OLS(Y,X)

A=[ones(length(Y),1) X];
b=inv(A'*A)*(A'*Y);
Eps=Y-A*b;
SE=sqrt(diag(inv(A'*A)*var(Eps)));

out=[b'; SE'; (b./SE)'];
R2=1-var(Eps)/var(Y);
% use this if need adjusted R2: adj_R2=R2-(1-R2)*size(X,2)/(size(X
   ,1)-size(X,2)-1);
```