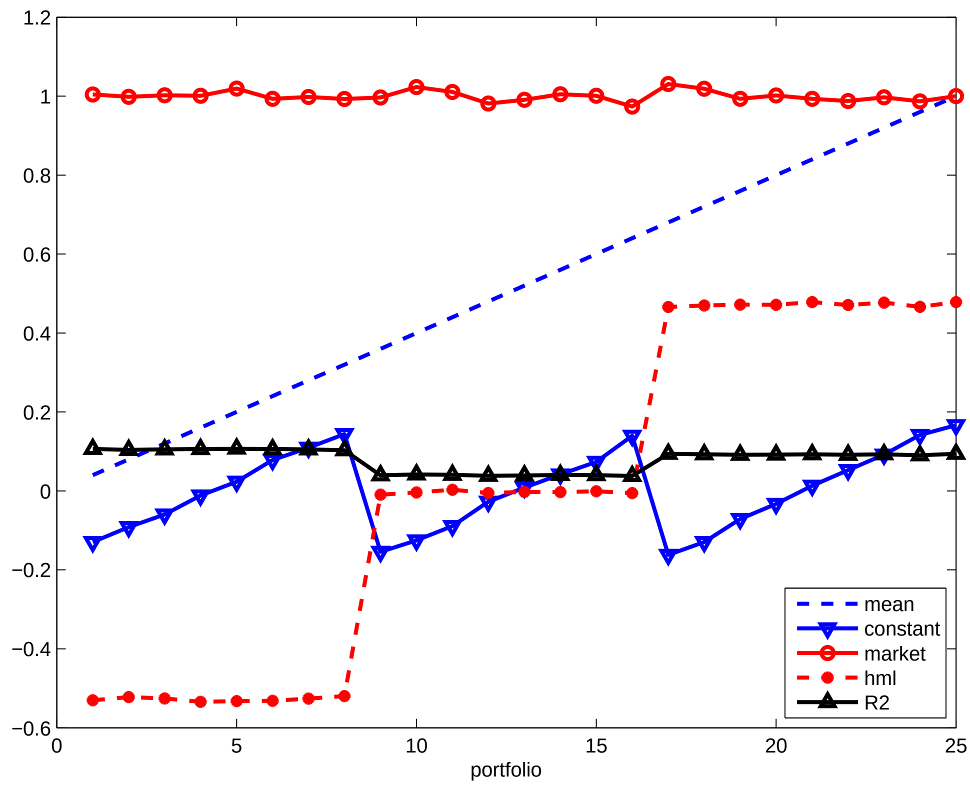
Problem set 3 answers

1. (a) Make a plot of average return $E(R^{ei})$, α_i, b_i, h_i and R^2 versus i. (You can also make tables, but it's easier to see in a plot.)

Answer:



(b) Explain the patterns in b_i and h_i and R^2. Why aren't they zero with uncorrelated assets?

Answer: the market betas are all one, just like FF. The hml betas are all -0.53 for portfolios included in the L portfolio, +0.53 for the portfolios included in the H portfolio, and zero for the portfolios between H and L. This should give you a clue what's going on – the betas are there when the same portfolio is included in both the left and right hand sides. The R^2 are all tiny compared to FF.

You can show this behavior in equations, though I thought the simulation would be more vivid. Suppose you run portfolio

$$R_t^{e1} = a + b\left[\frac{1}{N}\left(R_t^{e1} + R_t^{e2} + R_t^{e3} + ... + R_t^{eN}\right)\right] + \varepsilon_t$$

you can work out

$$\begin{aligned}
\hat{b} &= \frac{cov(x,y)}{var(x)} = \frac{cov\left\{\left[\frac{1}{N}\left(R_t^{e1} + R_t^{e2} + R_t^{e3} + ... + R_t^{eN}\right)\right], R_t^{e1}\right\}}{var\left[\frac{1}{N}\left(R_t^{e1} + R_t^{e2} + R_t^{e3} + ... + R_t^{eN}\right)\right]} \\
&= \frac{\frac{1}{N}var(R_t^{e1})}{\frac{1}{N^2}var(R^{e1}) + var(R^{e2}) .. + var(R^{eN})} = 1 \\
R^2 &= \frac{var[bx_t]}{var(y_t)} = \frac{var\left[\frac{1}{N}\left(R_t^{e1} + R_t^{e2} + R_t^{e3} + ... + R_t^{eN}\right)\right]}{var(R_t^{e1})} \\
&= \frac{\frac{1}{N}var(R^e)}{var(R^e)} = \frac{1}{N}
\end{aligned}$$

You can see that the coefficient of one comes from having the same thing on the left and right hand sides, and the tiny R^2 is the sign that you're explaining x with x plus noise.

Since the betas are driven by the same portfolio on left and right phenomenon, the hml betas are *not* larger where expected returns are larger. The hml's rise in a step function from left to right, not continuously within groupings as the expected returns do.

(c) Explain the pattern in α_i relative to $E(R^{ei})$.

Answer: The hml betas are step functions, while the average returns increase linearly, so the alphas rise within groups. What we do not see is the FF pattern that betas rise with expected returns even within groups.

(d) Now, what part of the FF Table 1 results might be spurious? What parts of FF's results do not look like these results? How might FF have run things a bit differently to clear up the small problems you do see here?

Answer: There is a spurious beta that comes from running the same thing on left and right. FF's very large R^2 and betas that vary within groups do not come from this result. FF could have omitted portfolio i from the market, hml, and smb for each regression. When you do this, you get exactly the same result, which is why they did not bother in the final version.

In sum, one of FF's central findings is *comovement* – high R2, the fact that portfolios with similar expected returns *move together.* That is the central ingredient missing in my simulation.

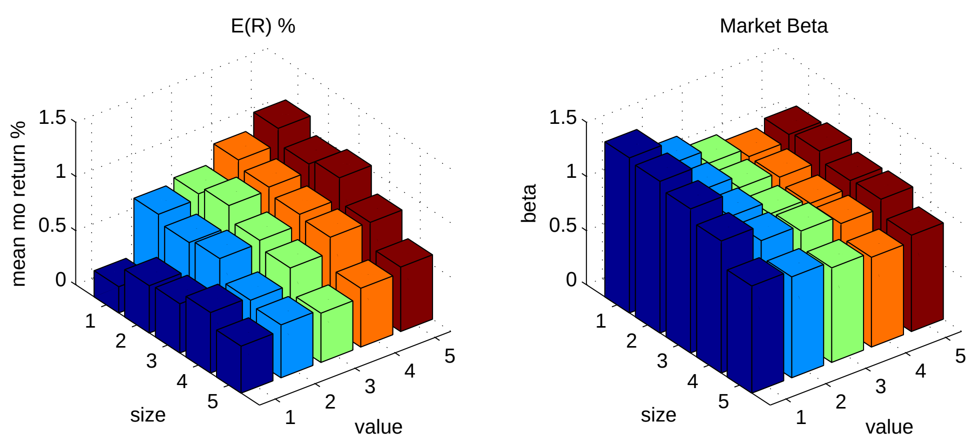
(e) The b and h in the simulation will drop to zero. The b, h, s in FF are essentially unaffected. The already small R^2 in the simulation drops to zero. The R^2 in FF is unaffected. How would the alphas be affected?

2. My results with commentary. First, the CAPM in tabular format. You can see that means rise to the northeast as for FF. Capm betas *do* vary. They are not all one. The size pattern is ok, they rise to the north. Alas the rise to the north is more pronounced in the growth portfolios where the returns do not rise to the north than it is for the value portfolios where they do. The value pattern is wrong, the betas rise to the northwest not the northeast. Thus, the spread in alphas across portfolios is larger than the spread of mean returns across portfolios. Alphas are composed of mean returns going one way and betas going the opposite way. The CAPM R2 are in the 60–80% range which is typical for large portfolios. There are lots of alpha t stats above 2

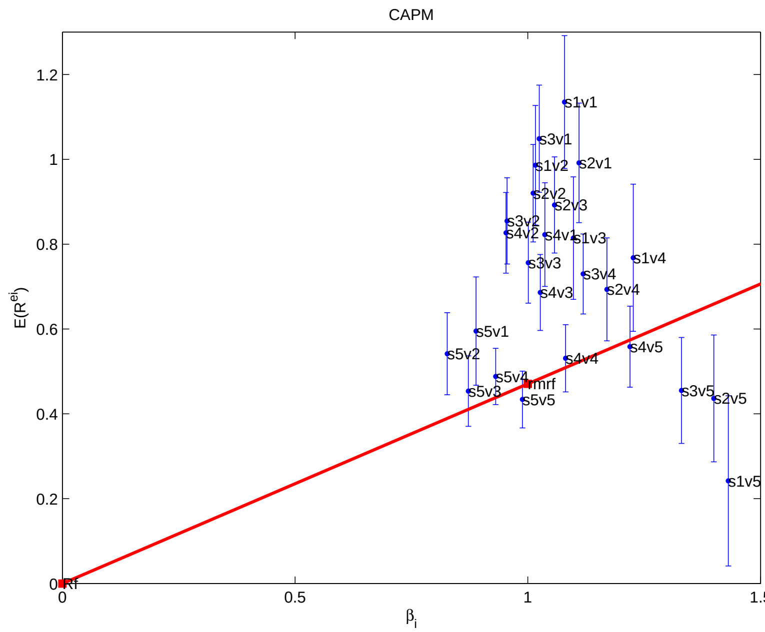
```
Data sample
     196301.00     201308.00

Time series regression results
As in FF all results are in boxes with size and book to market
mean return
          0.27          0.79          0.83          1.01          1.16
          0.46          0.71          0.91          0.94          1.02
          0.48          0.77          0.77          0.87          1.08
          0.59          0.56          0.70          0.85          0.86
          0.45          0.50          0.47          0.56          0.63
CAPM betas
          1.43          1.23          1.10          1.02          1.08
          1.40          1.17          1.06          1.01          1.11
          1.33          1.12          1.00          0.96          1.03
          1.22          1.08          1.03          0.95          1.04
```

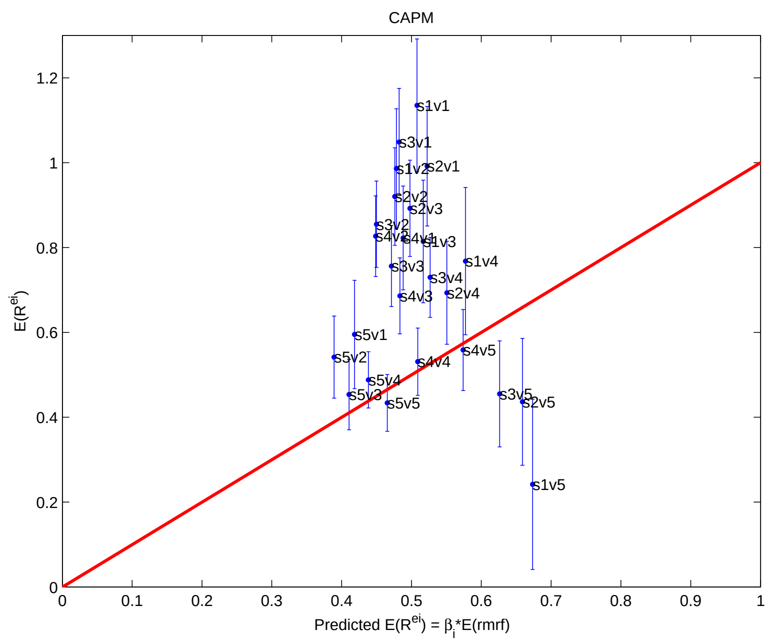
	0.99	0.93	0.87	0.83	0.89
CAPM alphas					
	-0.43	0.18	0.29	0.51	0.63
	-0.22	0.14	0.39	0.44	0.48
	-0.17	0.22	0.28	0.40	0.57
	-0.01	0.03	0.20	0.39	0.35
	-0.04	0.04	0.05	0.15	0.19
T on CAPM alphas					
	-2.18	1.07	2.06	3.66	4.09
	-1.51	1.17	3.53	3.91	3.45
	-1.39	2.37	3.02	4.03	4.60
	-0.10	0.35	2.29	4.11	2.90
	-0.59	0.68	0.55	1.61	1.54
CAPM R2					
	0.64	0.63	0.67	0.64	0.62
	0.75	0.76	0.75	0.73	0.68
	0.80	0.83	0.79	0.75	0.70
	0.85	0.87	0.82	0.78	0.71
	0.88	0.87	0.79	0.72	0.62

The same thing in pictures: You see FF's pattern of mean returns – increasing to small and value, except in the growth category. You can see that mean returns are not related to betas. If anything the growth companies have higher betas and lower returns. Notice these betas are larger than FF's, and the betas you will find below. Single regression betas are different from multiple regression betas because hml and smb are not exactly uncorrelated with $rmrf$.



I also made expected return vs. beta and expected return vs. beta*market premium plots, the latter can be compared with the 3 factor model: (I added standard error bars. These are the standard errors of alpha. This was not required on the problem set.)



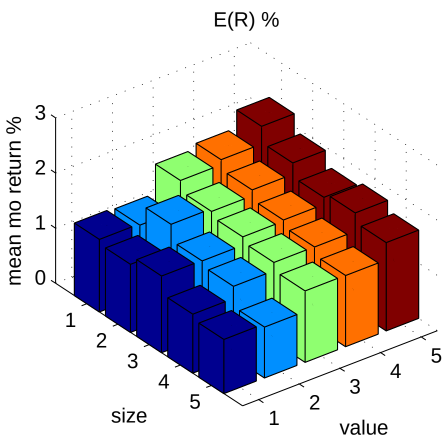


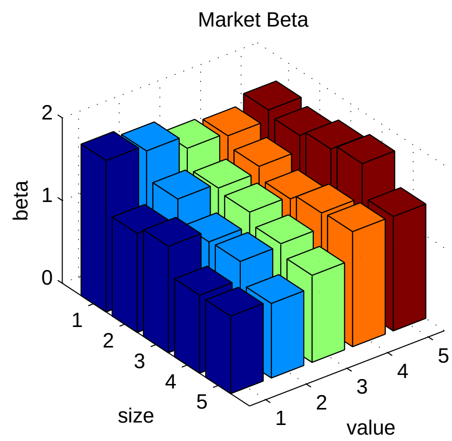
"s1" means small and "s5" means big. "v1" means value and "v5" means growth. These graphs show that there is sort of a blob with some hope for the CAPM, except for the size effect among deep growth firms. As you go from s1v5 (large growth) to s5v5 (small growth) betas rise but average excess returns decline, and these points are way off the line. In the bar graphs, you could see mean returns declining as you go "back." But the blob is way too much of a blob for a successful theory.

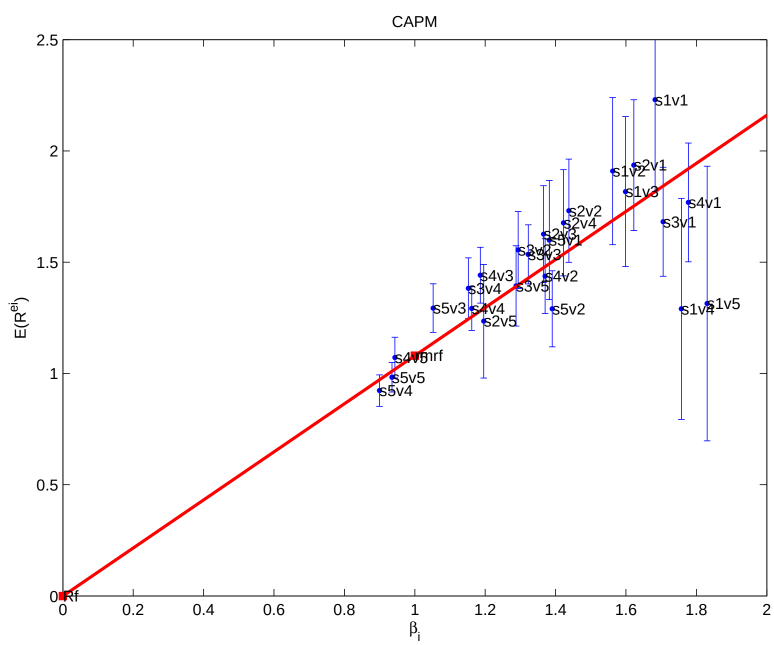
Value isn't a puzzle of average returns, it's a puzzle of lack of betas. It's perfectly natural that "value" stocks have high average returns, because it's perfectly natural that they are extra risky. The puzzle is that the beta is not here. That's particularly obvious in the lower plot. The puzzle is that when the market declines, value stocks *don't* get hammered. To "explain" the value puzzle, you have to explain this lack of beta, not the unusual expected returns. All the behavioral stories are looking in the wrong place.

c) Here are the numbers in case you want to compare. The graphs are much more insightful

```
Data sample
     193201.00     196212.00

Time series regression results
As in FF all results are in boxes with size and book to market
mean return
          1.31          1.29          1.82          1.91          2.23
          1.23          1.68          1.63          1.73          1.94
          1.39          1.38          1.54          1.55          1.68
          1.07          1.29          1.44          1.44          1.77
          0.98          0.92          1.29          1.29          1.60
CAPM betas
          1.83          1.76          1.60          1.56          1.68
          1.20          1.42          1.37          1.44          1.62
          1.29          1.15          1.32          1.29          1.71
          0.94          1.16          1.19          1.37          1.78
          0.94          0.90          1.05          1.39          1.38
CAPM alphas
         -0.66         -0.61          0.09          0.22          0.41
         -0.06          0.14          0.15          0.18          0.18
          0.00          0.14          0.11          0.16         -0.16
          0.05          0.04          0.16         -0.04         -0.15
         -0.03         -0.05          0.16         -0.21          0.11
T on CAPM alphas
         -1.07         -1.23          0.27          0.67          1.00
         -0.22          0.59          0.69          0.77          0.62
          0.01          1.01          0.80          0.91         -0.66
          0.57          0.37          1.28         -0.25         -0.57
         -0.41         -0.69          1.44         -1.24          0.40
CAPM R2
          0.48          0.57          0.71          0.70          0.64
          0.70          0.79          0.81          0.80          0.76
          0.84          0.88          0.91          0.86          0.84
          0.92          0.94          0.91          0.88          0.83
          0.95          0.95          0.91          0.88          0.74
```





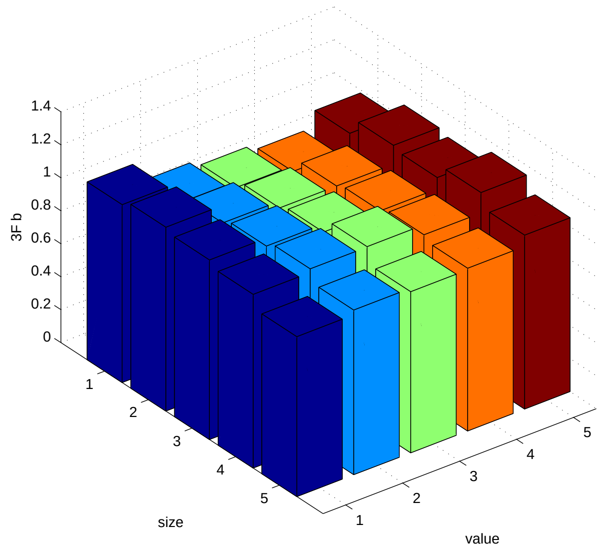


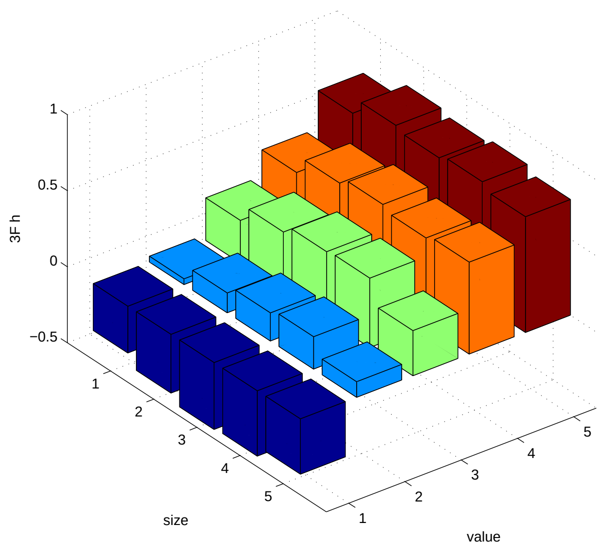
The CAPM works quite well in the earlier sample. What changed? The pattern of expected returns is the same. In fact, the small growth anomaly is "better" in the earlier sample too, in that small stocks earn more than large stocks even in the extreme growth bucket. *The betas changed.* Once again, the value puzzle is all about the betas, not really the average return! The big anomalies are still the small growth portfolios, s1v4 and s1v5, the back row on the left. They don't have great returns, but they have huge betas. Here the puzzle is "too much" beta.

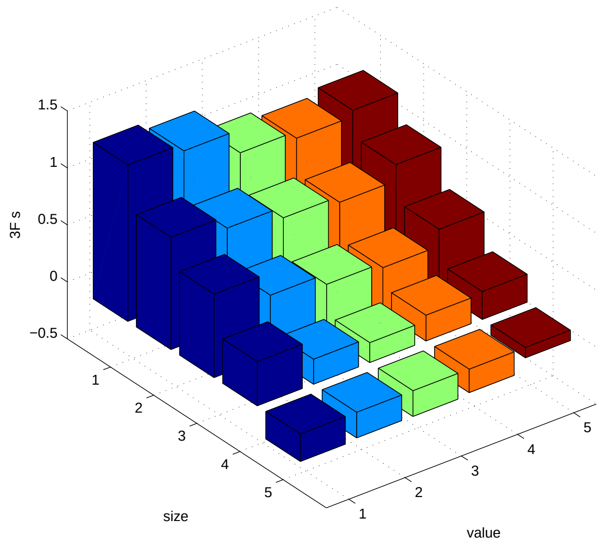
Now for the FF model. My results are consistent with FF in this larger sample. The CAPM betas are all about one. Note how these *multiple regression* betas are different from the *single regression* betas above. The market, hml, and smb are somewhat correlated, so multiple regressions assign some of what seemed to be movement with the market to movement with hml. The h coefficients rise as we go to the right and the s coefficients rise as we go up. The alphas are about as in FF, except the small growth alpha is much worse. Small growth stocks *underperform* dramatically. Note that this underperformace is not so much bad mean returns – they are the same as other mean returns. It comes from the combination of mean returns and betas. To take advantage of it, you don't short small growth stocks, you have to short small growth stocks *and* invest in hml. The 3F R2 are all above 90%, leading me to label the model more APT, as we will see.

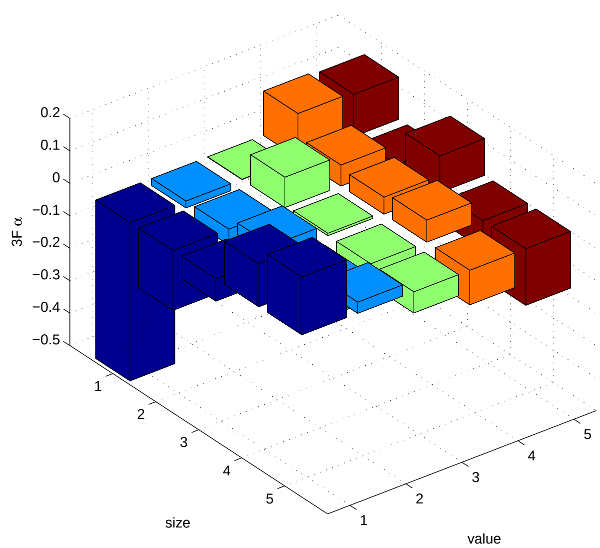
```
     196301.00     201209.00

Time series regression results
As in FF all results are in boxes with size and book to market
mean return
          0.24          0.77          0.81          0.99          1.13
          0.44          0.69          0.89          0.92          0.99
          0.46          0.73          0.76          0.85          1.05
          0.56          0.53          0.69          0.83          0.82
          0.43          0.49          0.45          0.54          0.60
3F model b
          1.08          0.96          0.92          0.88          0.98
          1.12          1.01          0.97          0.97          1.08
          1.09          1.04          0.99          0.99          1.06
          1.05          1.08          1.08          1.02          1.14
          0.97          1.00          0.98          0.99          1.05
3F model h
         -0.31          0.04          0.28          0.45          0.70
         -0.39          0.13          0.39          0.57          0.81
         -0.44          0.18          0.44          0.61          0.78
         -0.43          0.21          0.46          0.58          0.81
         -0.36          0.10          0.30          0.60          0.76
3F model s
          1.37          1.31          1.10          1.04          1.09
          0.98          0.87          0.78          0.72          0.86
          0.73          0.53          0.44          0.40          0.54
          0.39          0.22          0.17          0.22          0.24
         -0.24         -0.22         -0.23         -0.21         -0.09
3F model alphas
         -0.49         -0.02          0.00          0.14          0.13
         -0.18         -0.05          0.09          0.06         -0.05
         -0.07          0.04          0.01          0.05          0.11
          0.13         -0.11         -0.04          0.07         -0.09
          0.18          0.03         -0.07         -0.11         -0.17
T on 3F alphas
         -5.04         -0.30          0.01          2.45          2.17
         -2.74         -0.81          1.63          1.18         -0.78
         -1.11          0.52          0.11          0.80          1.50
          2.11         -1.54         -0.57          0.99         -1.06
          3.66          0.58         -0.93         -1.76         -1.86
3F R2
          0.92          0.94          0.95          0.95          0.95
          0.95          0.94          0.94          0.94          0.95
          0.95          0.91          0.90          0.90          0.89
```

0.94	0.89	0.88	0.89	0.87
0.94	0.90	0.86	0.89	0.80

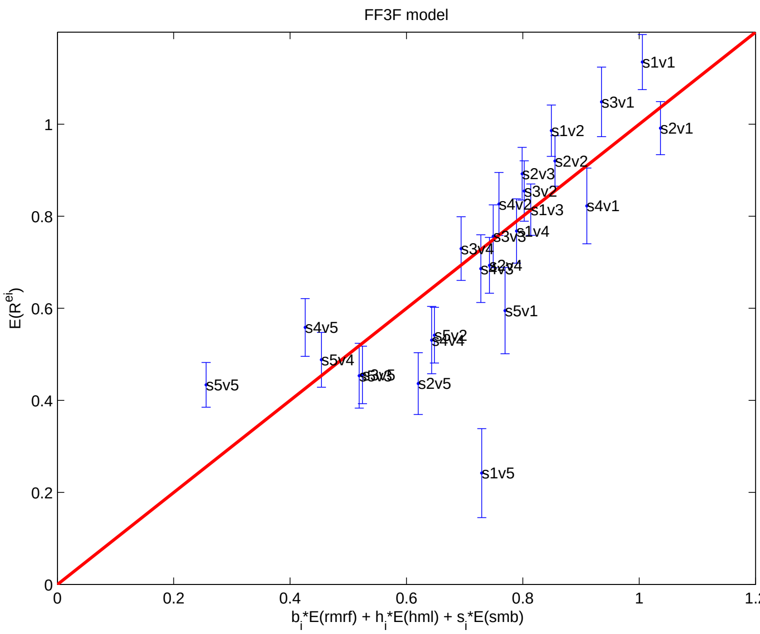


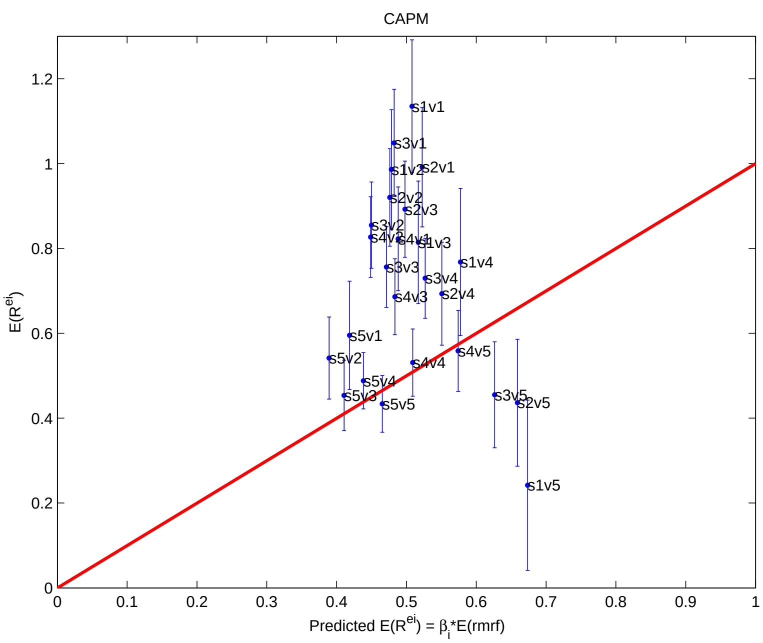






Here's the actual vs. predicted plot, and the CAPM plot repeated for comparison

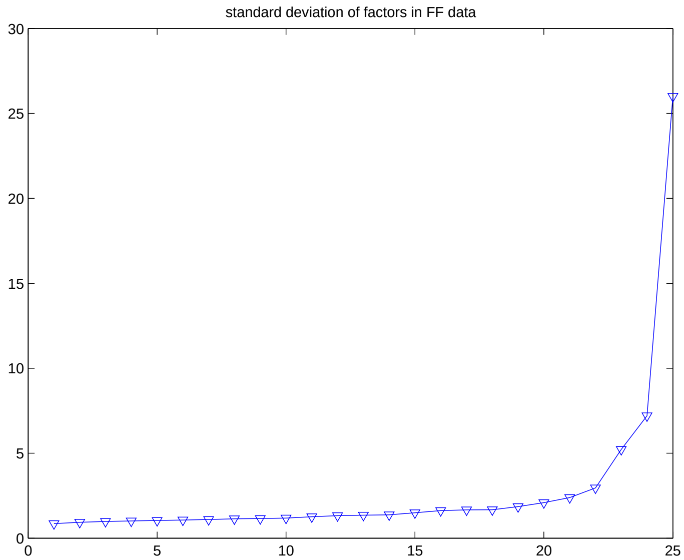




How are we doing better? The "predicted" has a much bigger range, which is good. Most of the portfolios like pretty close to the line, and much closer than in the CAPM case. Sometimes models "do better" because the standard errors get bigger. This one, the standard errors actually get smaller, but the points move a lot closer to the line. Thee model is still not doing well on s2v5 and s1v5 – the small growth firms. In fact, it isn't improving at all there. So the big improvement is really in the better spread along the line for all the other portfolios.

This looks a lot to me like FF's results, so there isn't much change in the longer sample. Mostly the point was to force you to really understand what FF did.

3. Here are the standard deviations of the factors. There's a huge first factor, 3 smaller factors, and then a tail. A picture like this usually means take the first 4 factors and ignore the rest.



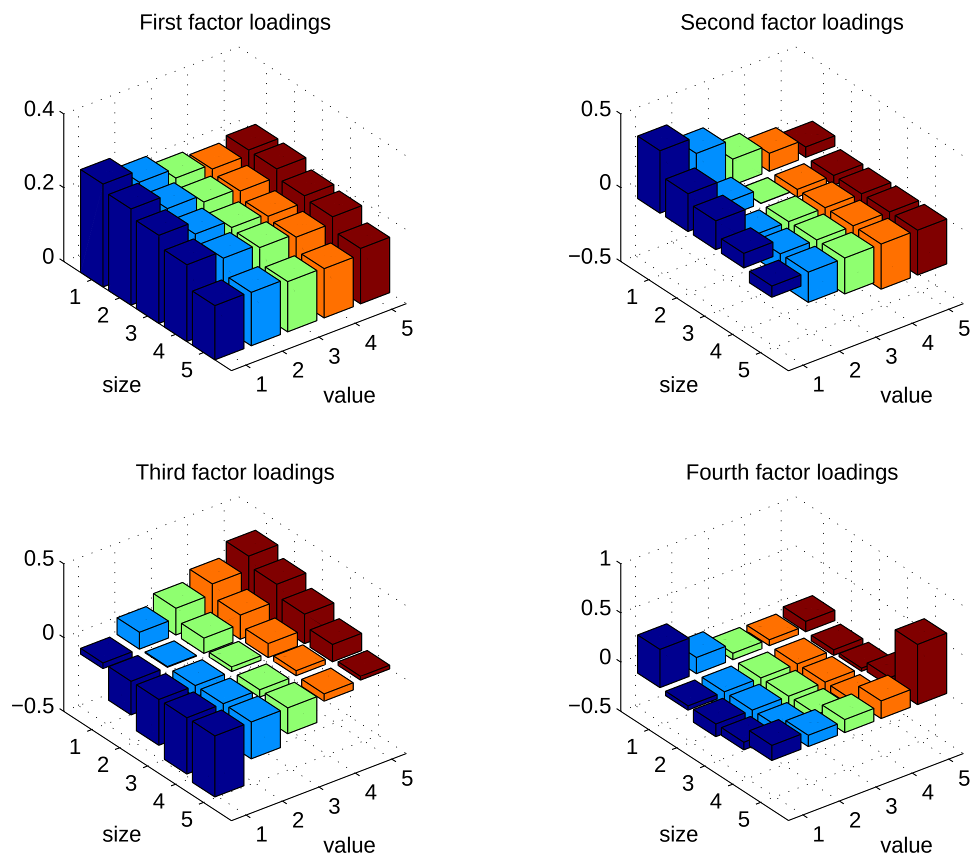
Next, the loadings. See the plots below.

The first factor is a "level" factor, moving everyone in the same direction. It's not flat – it moves small firms more than big firms, so it contains a bit of size as well as rmrf. The loadings are more like the single-regression loadings on the market that we saw when evaluating the CAPM on these portfolios. With interpretation 2 – loadings as the method for constructing the factor portfolio – we see that the first factor puts more weight on small firms. (Even equal weights would make the first factor the equally weighted market, not the value weighted market. It's quite interesting that the first factor is not the value-weighted market, which would have a hugely larger loading on the big portfolios.)

The second and third factors combine size and value. The second one is a bit more size-oriented. It moves small firms up and large firms down, but it also has a bit of value. The third factor has a bit more value – it moves the growth stocks down and the value stocks up – but it also has a size tilt. Clearly, between the two factors we have size and value factors. Using interpretation 2, we form these factor pretty much by combining small - big and the third as value - growth.

You're seeing one subtle issue here in factor models. The factors are separately "identified" by our attempt to order the variances. Since the variances of factors 2,3,4 are pretty much the same, the program has a hard time telling them apart.

The fourth factor moves small growth and large value up, and everyone else slightly down. This is really cool – small growth and large value are the Fama French puzzles in terms of alpha and look, they are also the next factor in the covariance matrix! Once again, high returns correspond to comovement, not to arbitrage.



This is nice, but again, it's all about covariances. How do mean returns line up with these covariances? The betas are the same as the columns of Q. Here are the alphas. I start with mean returns, and verify the usual pattern, higher mean with small size and especially value. The 1 factor model leaves substantial alphas. You have to watch the vertical axis here, as it changes. The alphas also have a pattern, they are larger for value stocks. There is a size pattern too – alphas are negative for small growth but positive for small value.

The 3 factor model does just about as well as the FF3 factor model. Most of the alphas are small. Here is a table. Most of these are smaller than the FF3F counterparts. There is a huge negative alpha for small growth and large value, as with FF3F.

3 factor alphas

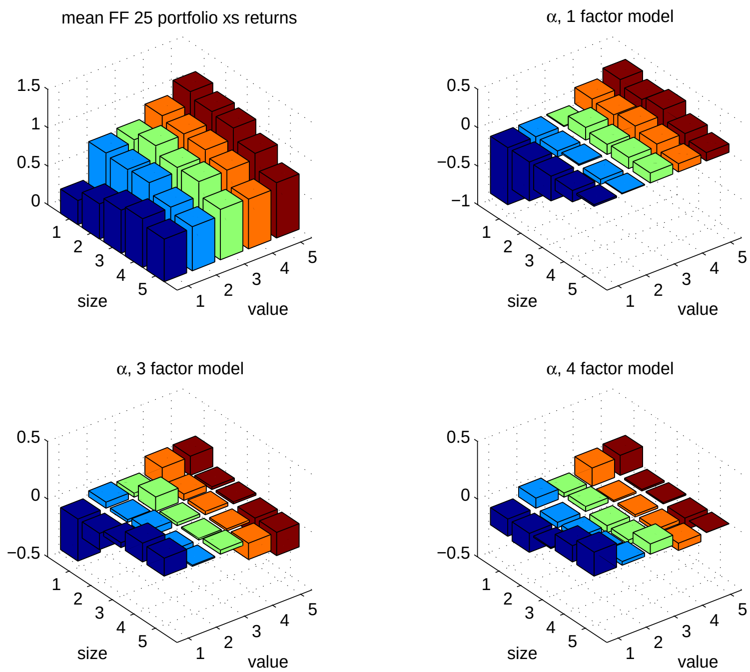
-0.37	0.05	0.03	0.16	0.14
-0.09	-0.02	0.10	0.06	-0.04
0.03	0.06	-0.01	0.03	0.09
0.22	-0.11	-0.06	0.06	-0.11
0.19	0.02	-0.06	-0.12	-0.16

4 factor alphas

-0.27	0.09	0.05	0.17	0.17
-0.10	-0.07	0.04	0.01	-0.05
-0.01	-0.01	-0.07	-0.01	0.08
0.20	-0.15	-0.10	0.04	-0.09
0.23	0.05	-0.03	-0.07	-0.01

3 factor r2

	0.96	0.97	0.97	0.96	0.96
	0.97	0.95	0.95	0.94	0.94
	0.97	0.94	0.93	0.94	0.92
	0.96	0.94	0.94	0.92	0.90
	0.88	0.90	0.85	0.88	0.78
4 factor R2					
	0.98	0.97	0.97	0.96	0.96
	0.97	0.96	0.96	0.95	0.95
	0.97	0.96	0.95	0.94	0.92
	0.96	0.95	0.94	0.92	0.91
	0.89	0.91	0.86	0.90	0.91



The 4th factor model loads on the large value and small growth, interestingly enough. Then, when we add it to the mix, it reduces the large value-small growth puzzle. (In the same way that a momentum factor eliminates the momentum puzzle.) Even the 4 factor model doesn't completely get rid of the small growth alpha. Interesting. However, we're now fishing deep in the small eigenvalues. The 25 factor alpha does the trick for sure, and going past 4 factors and 98% R2 is very very dangerous in terms of fishing, measurement errors, etc.

Again, there is no guarantee that covariances will explain alphas. That's a result, not a mathematical certainty. If it were not true of course there would be high Sharpe ratio opportunities. Thus it's wonderful to see each factor in turn beat down the alpha puzzles of the previous factor model.

Disclaimer: Of course we should be cautious in the use of too many factors. They may not be stable out of sample. Also, the size factor had questionable economics, the value factor only had FF's speculations about human capital in depressed industries, and momentum has no economics. I have no hint of the economics behind a small growth - large value factor. Thus, you should probably view it as the momentum factor, an ad hoc device that may be useful for performance evaluation, but still on shaky ground for fundamental asset pricing.

Bottom line: In facto *The R^2 is perhaps the biggest finding of FF Table 1!* My simulation with uncorrelated assets had no R^2, and R^2 in Table 1 and the factor model is large. *There are really only 3 factors in the covariance matrix of the 25 portfolios.* Second, the APT works pretty well. Or, remarkably badly, actually. Given the huge R^2 the Sharpe ratio for shorting small growth stocks is awfully large. Too bad it's really hard to short small growth stocks.

It's interesting that FF chose to make ad-hoc factors that are awfully close to what comes out of factor analysis. Why didn't they do what we just did, and remove the question "where did you come up with those factors?" I think the answer is, transparency. Now that we know and accept this style of modeling it makes perfect sense to use the factor analysis. But in writing the first paper, adding a black box would not have helped their cause. Watching how FF write their papers and the choices they make in their empirical work is very instructive.

4.

Here are my results.

```
mean and t stat of
             rmrf,        hml,         smb
mean         0.49         0.39         0.25
sharpe       0.11         0.14         0.08
t            2.70         3.36         2.01
```

The small stock premium has been roaring back btw. A few years ago it was quite small. Still, on a univariate basis, we have a just barely significant t stat, though I would judge a fairly economically signifiant mean and Sharpe ratio. It's clearly the weakest of the three factors, and our candidate for dropping if we want to drop one

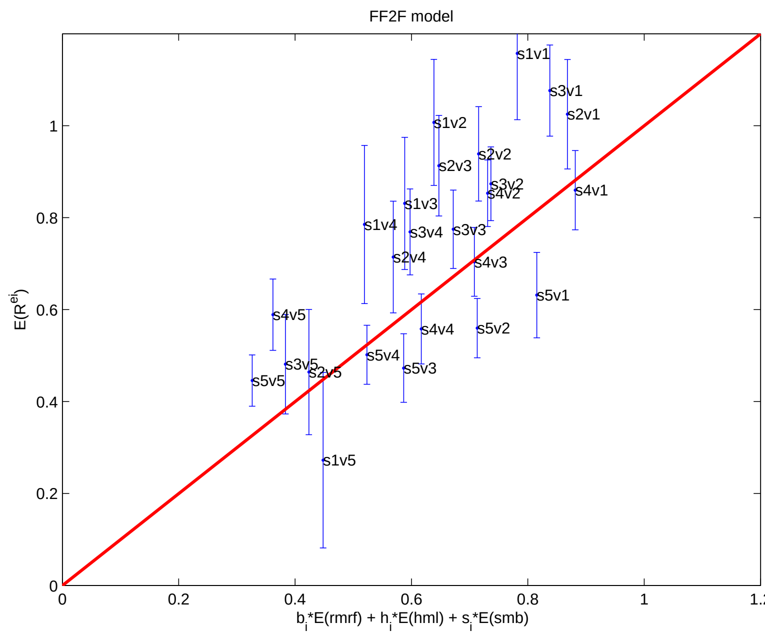
```
smb on rmrf, hml
alpha, beta, R2
             alpha    rmrf     hml      smb      R2
smb          0.23     0.18    -0.16              0.12
t            1.87     6.62    -3.70
smb          0.15     0.21                       0.10
t            1.23     7.99
hml          0.51    -0.16             -0.14     0.11
t            4.56    -6.23             -3.70
hml          0.48    -0.19                       0.09
t            4.33    -7.67
```

The first row is cool, and makes my major point. Thout $E(smb) > 0$ and is "significant", the smb alpha is not signifcant. Granted it's close – the premium only drops from 0.25 to 0.23 and the t stat from 2.01 to 1.87. So it's not overwhelming. But you at least see the methodological point that E(smb) is different from the question whethe rwe can drop smb. Interestingly, the smb alpha is seven smaller at 0.15 with even less t stat 1.23 relative to the CAPM only. Adding hml makes smb more necessary. The comparable exercise with hml shows it has huge alphas relative to the market or market and smb. We clearly need hml as a factor!

What does the "two factor" model look like? I appended 3 factor model results below to compare.

```
2F model b
         1.33          1.20          1.12          1.07          1.18
         1.30          1.17          1.11          1.10          1.24
         1.22          1.14          1.07          1.06          1.16
         1.13          1.12          1.11          1.06          1.19
         0.93          0.96          0.93          0.95          1.04
2F model h
        -0.52         -0.17          0.10          0.29          0.52
        -0.54         -0.01          0.27          0.45          0.67
        -0.55          0.10          0.37          0.55          0.69
        -0.49          0.18          0.42          0.54          0.77
        -0.33          0.14          0.33          0.63          0.78
2F model alphas
        -0.18          0.27          0.24          0.37          0.38
         0.04          0.15          0.27          0.22          0.16
         0.10          0.17          0.10          0.14          0.24
         0.23         -0.06         -0.00          0.12         -0.02
         0.12         -0.02         -0.11         -0.15         -0.18
t on 2F alphas
        -0.92          1.55          1.69          2.69          2.60
         0.30          1.20          2.43          2.17          1.32
         0.90          1.83          1.21          1.71          2.40
         2.93         -0.77         -0.06          1.69         -0.25
         2.14         -0.34         -1.53         -2.38         -1.98
2F R2
         0.67          0.64          0.67          0.66          0.68
         0.79          0.76          0.77          0.78          0.77
         0.85          0.83          0.83          0.85          0.81
         0.90          0.88          0.87          0.87          0.86
         0.92          0.88          0.83          0.87          0.80

2F model rmse alpha
         0.19

3F model rmse alpha
         0.14\bigskip
```

As you can see market betas are not all one – they increase from bottom to top. Variation in market betas now does a lot to explain the size effect. h coefficients seem largely unchanged increasing from left to right. The alphas are by and large a ltitle bigger, especially teh top two rows. Where most were below 10 bp now, many are greater. However, there is a theorem that tells us if smb alpha against rmrf smb is not significant, the rise in alphas is not signifcant either, given that we gain a degree of freedom. Interestingly the small growth alpha is no longer statistically or economically significant, though it still stands out from its peers. The t statisticss on alphas remain by and large mostly marginal. I don't present the GRS *the GRS statistic tests one model vs. perfection. You do not use GRS statsitsics to test one model vs. another.* And we're not really doing testing anyway. The rmse alphas do increase, and there is a test to say if this increase is statistically significant. It isn't. But the economic significance of 0.14 to 0.19 isn't trivial. The R^2 decline. It's not as much as the 0.6 R^2 of the CAPM but not as good as the 0.9+ R^2 of the three factor model. Well adding "unpriced" covariance factors does raise R^2. Here's the plot of actual vs. expected. The overall spread is bigger, but its interesting that small growth is no longer a special problem. Its mean return is the same, but because there is no huge size beta, it behaves more normally.



So, bottom line philosophical question: Should we include or leave out size? I think the answer is, it is so useful in explaining the *variance* of returns that one wants to include it for many practical purposes. Even if its true contribution is zero, its contribution in small samples may be large. Overall, this gives you a sense of how adding or subtracting a next to useless expected return factor with a big covariance presence affects the results.