Problem Set 4 Solutions

Part I readings. Give one-sentence answers.

1. Novy-Marx, The Profitability Premium. Preview: We see that gross profitability forecasts returns, a lot; its power is not explained by hml slopes (in fact, hml goes the wrong way.) A "profitability factor" digests a long list of anomalies, just as FF found for HML.

   (a) $p_t - d_t = E_t \sum \rho^{j-1} (\Delta d_{t+j} - r_{t+j})$. Does this sort of identity suggest that variables, such as current gross profitability, which have strong power to predict future dividends (or earnings; "good company" variables) should forecast returns better alone, or better in combination with book/market or similar valuation ratios? (If you run a regression $R_{t+1} = a + bx_t + \varepsilon_{t+1}$ and a regression $R_{t+1} = a + bx_t + cy_t + \varepsilon_{t+1}$, sometimes including y_t improves x's forecast power (raises b), and sometimes including y_t lowers $x's$ forecast power (lowers b). Which is it in this case?)

   A: That's the whole idea really. Novy-Marx is enchanted with the power of profitability by itself, but both profitability and B/M are mixing two things seen best when they are included together. Higher Δd by itself can mean just higher p. Higher Δd *controlling for p must* mean higher returns. Abstract.

   (b) p.5. Fama and French found that profitability did not help to forecast returns. In a one-sentence nutshell (and inspired by the identity), why does Novy-Marx's measure work so much better?. (Start with, why is it different? Note, this "true economic profitability" stuff seems farfetched to me. There is something deeper at work.)

   A: Gross profitability is a better forecast of future earnings. For example, if you're investing a lot in R&D, that's a great sign of future earnings, but bad for earnings today.

   (c) Table 1 Panel A. Are returns here regressed on contemporaneous values of the right hand variable, $R_t = a + bx_t + \varepsilon_t$, or lagged, $R_t = a + bx_{t-1} + \varepsilon_t$? What's the punchline of Panel A?

   A:$R_t = a + bx_{t-1} + \varepsilon_t$?, the punchline is that gross profitability helps to forecast returns, controlling for B/M.

   (d) What danger does Novy-Marx worry about with Fama-MacBeth regressions, which motivates looking at portfolios? (A little graph might help.)

   A: p.9 "because they weight each observation equally, put tremendous weight on the nano- and micro-cap stocks, which make up roughly two-thirds of the market by name but less than 6% of the market by capitalization. The Fama-MacBeth regressions are also sensitive to outliers, and impose a potentially misspecified parametric relation between the variables, making the economic significance of the results difficult to judge. "

   (e) Table 2. More profitable firms will have higher prices, so just sorting on profitability can isolate low-return growth firms. If you sort just based on gross profitability, without controlling for B/M, do expected returns still rise with profitability, or do they go the other way? Do HML loadings explain the variation in expected returns?

   A: Even not controlling for B/M, higher profitability means higher expected returns. HML loadings go the wrong way.

   (f) Table 4. Does the power of profit - to asset sorts evaporate in the biggest firms, as "dissecting anomalies" suggested? (Where do we look?)

   A: Bottom row, top left box panel A. It's weaker (0.30-0.55) than the top row (0.40-1.07) but still there.

(g) Table 6 (Important!) Is the profit effect in average returns strong while controlling for book/market? Is the value effect strong in average returns controlling for profits? Do hml betas (and alphas) explain the value dimension of average returns, controlling for profits? Do hml betas (and alphas) or other Fama French betas explain the profit dimension of returns controlling for book to market?

A: Top left block, yes and yes. Each effect is stronger controlling for the others, as you would expect. β_{hml} column (far right) and row (lower block, panel A) yes for value, no for profitability where hml betas are all the same.

(h) Table 10 (Important). Like Fama and French, Novy-Marx wants to know if a factor formed on his anomaly, PMU, can explain a zoo of other anomalies. How many anomalies in Table 10 are left unexplained by the FF+momentum model, that Novy-Marx model captures? Hint: What's the central column of Table 10?) (For counting purposes, we can use t statistics greater or less than two, despite my sermons about t statistics)

A: I.e. how many α_{ff4} are above 2 with α less than 2. I count 11 cases, but the principle and getting you to look at the table is the main thing.

2. Fama and French Five Factor model. Reading notes. Fama and French spend a lot of time on showing you various different ways of doing things to show how it comes out the same. Don't get lost in the variations, first get the basic idea and figure out their point with one way of doing things. In particular, get the general idea of "factor definitions" section III but don't get too deep in the different versions. Section V sems the most important to me. I focused on the 32 portfolios and 2x2x2x2 factors.

(a) p.3 FF say "higher expected future earnings imply a higher expected return." But in a present value model, higher expected earnings can coexist with a constant expected return – it just generates a higher stock price. How can FF say this?

A: This paragraph fixes M_t i.e the price. If the price doesn't change, then higher earnings must have come with higher expected returns. The point of the question is to review that FF are making partial statements holding things – like M – constant. Also p.7 " The valuation model does not predict that B/M, OP, and Inv effects show up in average returns without the appropriate controls."

(b) What are RMW and CMA?

A: p.4 the difference between returns on diversified portfolios of stock with robust and weak profitability and...low and high investment .. which we call conservative and aggressive.

(c) How do FF define and measure "profitability" and "investment?" Are these in dollars or ratios somehow?

A: p.6. "profitability (measured with accounting data for the fiscal year ending in t-1) is annual revenues minus cost of goods sold, interest expense, and selling, general, and administrative expenses, all divided by book equity at the end of fiscal year t-1." p.7; " Inv is the growth of total assets for the fiscal year ending in t-1 divided by total assets at the end of t-1."

(d) p.8 "we would like to sort jointly on size, B/M, OP, and Inv". How did "dissecting anomalies" solve this problem? (Optional) Why don't they use the same approach here?

A: by running cross sectional regressions, in which case multiple regression slopes tell you the power of each variable. Why do they not use that approach here? Because in a second stage they want to see if portfolio expected returns line up with factor betas.

(e) Table 1. For given size (ME), do expected returns rise for more profitable firms? Or does the "good company/good stock" fallacy apply? Is there a difference between small and large firms?

A: Table 1, middle. Higher profits mean higher expected returns (moving to the right), a bit more in small firms (30 bp not 20 bp).

(f) For given size, do firms that invest more have greater returns?

A: Table 1, bottom, It's the other way around. Though most of the effect seems concentrated in a few places.

(g) Table 2. Is there a value effect in returns, holding profits constant? Is there a profits effect on returns, holding value constant?

A: Top panel, expected returns rise as we go to the right and down. p.8

(h) Table 4 C. (p.13-14) Are the factor returns pretty much uncorrelated?

A: FF say so on p.13-14, but I see a strong 0.6-0.7 correlation in each row between HML and one of RMW and CMA. So the ability to drop HML later doesn't seem so surprising to me.

Note: As suggested in Problem set 4, FF Table 5 focus on average absolute alphas and average R^2. Ignore GRS statistics. To keep our job manageable, focus on the 2x2x2x2 factors and 32 portfolios (Table 5 page 2 right.)

(i) (Table 5 F) In portfolios formed on size, OP and inv, can we get by with just HML and one of RMW and CMA, or does it look like we really need both RMW and CMA factors to explain those sorts?

A: As I read it, to get 0.1 average alpha, we need both RMW and CMA. Table 5 Panel F.

(j) Note: As suggested in Problem set 4, the way to see if you can drop a factor for pricing (not for R^2) is to run it on the other factors and see if there is an intercept. Reading Table 6 2x2x2x2, are there any of the 5 factors that we can drop for our description of expected returns?

A: Look at the intercepts, both their size (%/mo) and t statistics. It looks like we can drop HML! That's because, as we saw in T4, HML is quite correlated with RMW and CMA. Why does problem set 4 drop size and FF don't? I think it is because they have RMW and CMA in the regression

Note: See the last paragraph for FF's bottom line. FF don't say anything about economics. If I were writing the article I would say that RMW captures earnings, and CMA captures cost of capital. Given those, B/M itself is redundant.

3. Frazzini and Pedersen, Betting against Beta. Reading note: Skip the theory, let's get to the facts starting section III. (I think the theory is a bit silly anyway. Let them buy options. Let firms leverage. But you can't argue with facts.) Let's also focus on US equities. Focus on Table III and IV, and discussion on p. 21

(a) In Table III, the average returns are the same across portfolios. How can there be a puzzle?

A: Puzzles are joint puzzles of expected returns and betas. Beta without expected return is just as much a puzzle as expected return without beta!

(b) If the average returns are the same in row 1, how can the BAB strategy have a high average return in the last column. $E(R^{eA} - R^{eB}) = E(R^{eA}) - E(R^{eB})$, no?

A: It's a zero-beta BAB factor. Notice the zero beta in the ex ante beta row. Intuitively, they buy more of the low beta stocks and short fewer of the high beta stocks. The beta of wR^{ea} is $w\beta_A$ (beta is linear). So if we weight a portfolio by the inverse of beta, $w = a/\beta_A$ then the beta of R^{eA}/β^A is one. So the portfolio is $R^{eBAB} = R^{eA}/\beta^A - R^{eB}/\beta^B$. If $E(R^{eA}) = E(R^{eB}) = E(R^e)$ then $E\left(R^{eBAB}\right) = E(R^e)(1/\beta^A - 1/\beta^B)$, and $\beta\left(R^{eBAB}\right) = 0$. Any difference in beta with no difference in mean return generates a BAB portfolio mean. (Note all these returns are excess returns, so I don't have to have weights summing to one.)

(c) How do we know that the betas will persist going forward? I.e. if we form a portfolio of betting against beta, how do we know that the differences in beta don't disappear going forward and we lose all our money? (Hint: In table III why are there two rows of beta?)

A: The portfolios are evaluated using post-ranking betas. They are formed using pre-ranking betas, but you can form portfolios any way you want. The two rows of beta show the differences between pre and post ranking betas.

Part II Computer

I start with mean returns (FF Table 1A). Where FF found 0.31 to 1.18 (first row) I get 0.33 to 1.15. Last row, 0.37 to 0.71 is now 0.34 to 0.69. The data changes do matter, but the pattern is the same

I start with the CAPM again in this sample (I didn't ask you to do that.) The results are quite similar to what you found last week in the 1963-today sample. Capm betas rise as you go from big to small, but decline from growth to value so the CAPM is a disaster.

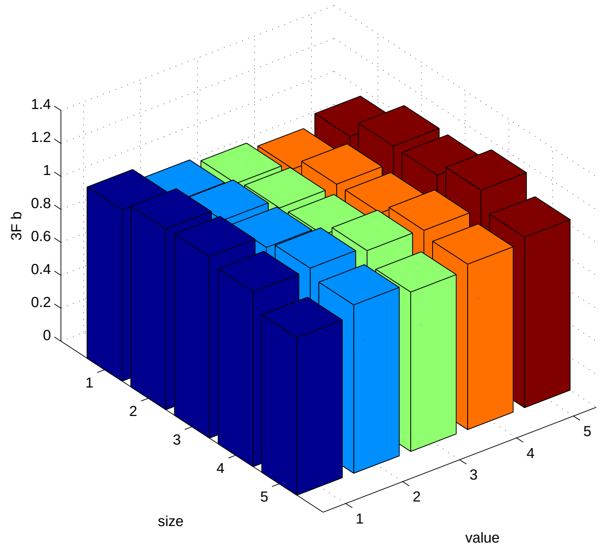
In the FF model, as in FF the market b are all about one. I get the same result as FF out to the second decimal, but differences there. Note how these *multiple regression* betas are different from the *single regression* CAPM betas. The market, hml, and smb are somewhat correlated, so multiple regressions assign some of what seemed to be movement with the market to movement with hml.

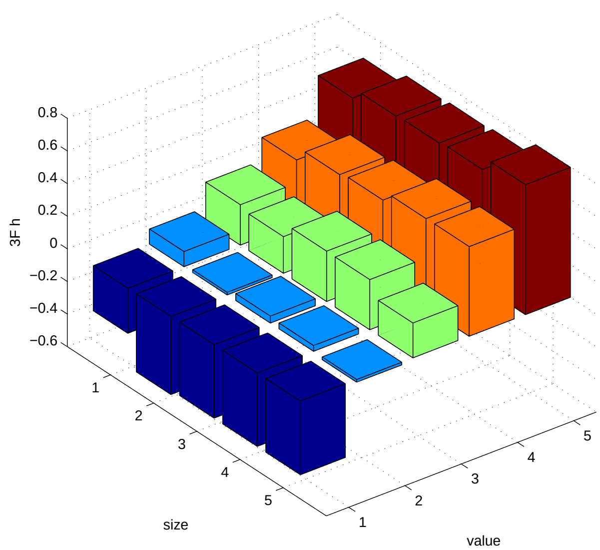
The h coefficients rise as we go to the right and the s coefficients rise as we go up. The alphas are about as in FF. Small growth stocks *underperform* dramatically. Note that this underperformace is not so much bad mean returns – they are the same as other mean returns. It comes from the combination 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.

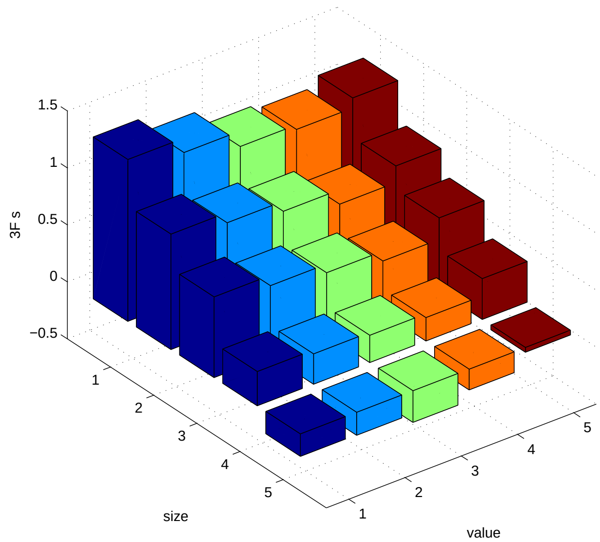
```
Data sample
     196301.00     199312.00

Time series regression results
As in FF all results are in boxes with size and book to market
mean return
          0.33          0.71          0.76          0.97          1.15
          0.39          0.69          0.90          0.98          1.11
          0.44          0.75          0.71          0.91          1.05
          0.47          0.42          0.67          0.84          0.95
          0.34          0.38          0.39          0.56          0.69
CAPM betas
          1.42          1.24          1.15          1.07          1.10
          1.43          1.23          1.10          1.04          1.12
          1.36          1.16          1.03          0.97          1.07
          1.22          1.13          1.04          0.97          1.08
          1.00          0.98          0.86          0.85          0.86
CAPM alphas
         -0.30          0.16          0.26          0.50          0.66
         -0.23          0.15          0.42          0.52          0.62
         -0.15          0.24          0.26          0.49          0.58
         -0.07         -0.07          0.22          0.41          0.48
```

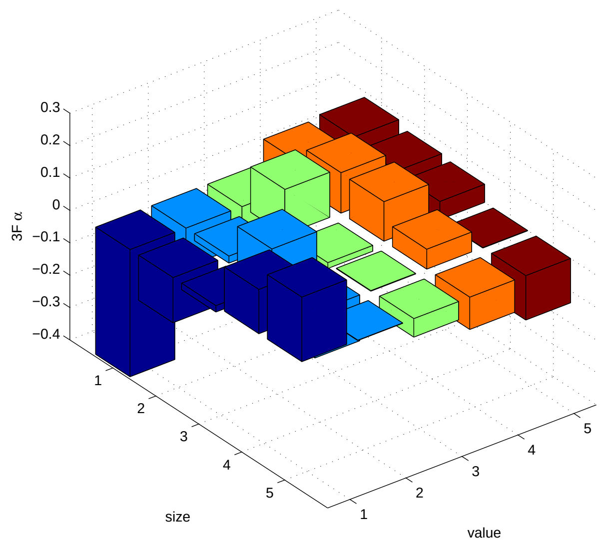
-0.10	-0.05	0.02	0.19	0.32
T on CAPM alphas				
-1.33	0.84	1.47	2.86	3.36
-1.37	1.06	3.11	4.02	3.81
-1.16	2.21	2.34	4.44	3.85
-0.71	-0.85	2.37	3.86	3.38
-1.07	-0.64	0.17	1.84	2.13
CAPM R2				
0.68	0.69	0.70	0.67	0.62
0.79	0.80	0.78	0.77	0.72
0.85	0.86	0.83	0.81	0.73
0.89	0.90	0.87	0.81	0.76
0.87	0.91	0.81	0.79	0.65
3F model b				
1.04	0.96	0.93	0.89	0.95
1.10	1.02	0.96	0.97	1.07
1.10	1.02	0.97	0.97	1.06
1.06	1.07	1.05	1.04	1.15
0.96	1.02	0.97	1.00	1.03
3F model h				
-0.28	0.09	0.25	0.39	0.64
-0.48	0.02	0.22	0.47	0.70
-0.45	0.04	0.31	0.49	0.71
-0.45	0.03	0.31	0.55	0.72
-0.45	-0.02	0.21	0.55	0.80
3F model s				
1.42	1.29	1.15	1.11	1.20
1.01	0.92	0.83	0.71	0.85
0.71	0.62	0.54	0.45	0.64
0.30	0.27	0.24	0.20	0.36
-0.20	-0.20	-0.28	-0.18	-0.04
3F model alphas				
-0.39	-0.12	-0.09	0.07	0.08
-0.14	-0.02	0.14	0.12	0.06
-0.02	0.11	-0.02	0.12	0.06
0.14	-0.14	-0.00	0.06	-0.00
0.20	-0.00	-0.06	-0.10	-0.14
T on 3F alphas				
-3.70	-1.55	-1.49	1.20	1.24
-1.71	-0.32	2.21	2.05	0.97
-0.27	1.53	-0.25	1.87	0.69
1.89	-1.76	-0.02	0.78	-0.01
2.99	-0.02	-0.65	-1.39	-1.22
3F R2				
0.93	0.95	0.96	0.96	0.96
0.96	0.96	0.96	0.95	0.96
0.96	0.94	0.93	0.94	0.92
0.95	0.92	0.91	0.90	0.89
0.93	0.92	0.85	0.90	0.81

The suggested graphical exposition:

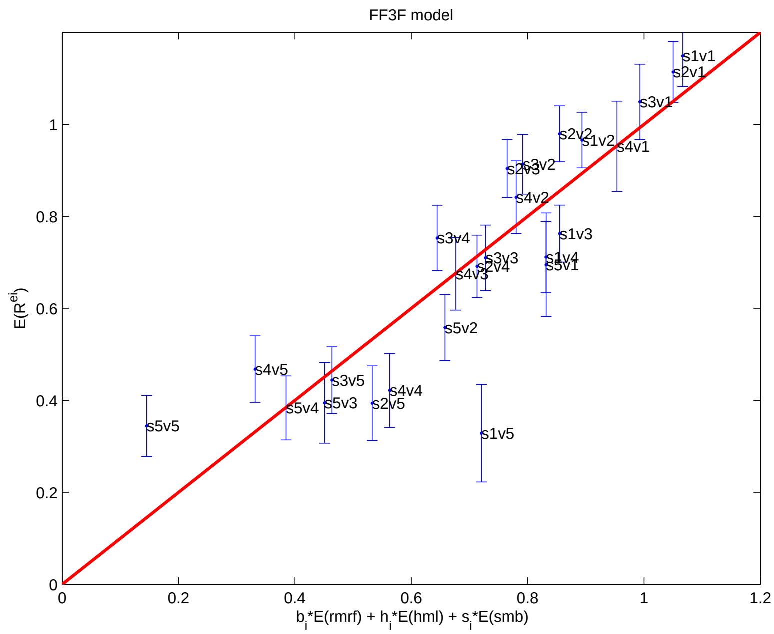


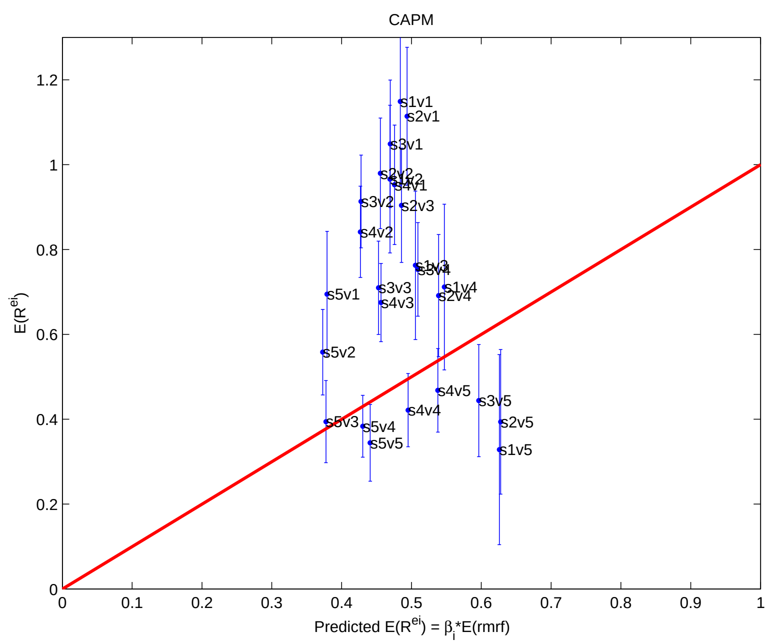






Here's the actual vs. predicted plot, and the CAPM plot repeated in this data sample for comparison





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.

2)

`Data sample`

199401.00 201311.00

Time series regression results
As in FF all results are in boxes with size and book to market

mean return

0.24	0.95	0.99	1.13	1.21
0.62	0.79	0.97	0.92	0.94
0.59	0.84	0.92	0.86	1.16
0.81	0.81	0.79	0.90	0.75
0.65	0.72	0.64	0.59	0.57

3F model b

1.14	0.96	0.87	0.85	0.98
1.12	0.96	0.93	0.95	1.08
1.07	1.02	0.98	0.99	1.02
1.05	1.04	1.07	0.99	1.09
0.96	0.94	0.98	0.95	1.08

3F model h

-0.36	-0.01	0.29	0.49	0.71
-0.30	0.23	0.54	0.67	0.91
-0.41	0.29	0.54	0.72	0.81
-0.36	0.38	0.57	0.62	0.85
-0.29	0.22	0.39	0.64	0.71

3F model s

1.31	1.30	1.05	0.97	0.99
0.98	0.87	0.77	0.77	0.91
0.77	0.48	0.37	0.37	0.45
0.50	0.22	0.15	0.25	0.14
-0.26	-0.21	-0.15	-0.22	-0.15

3F model alphas

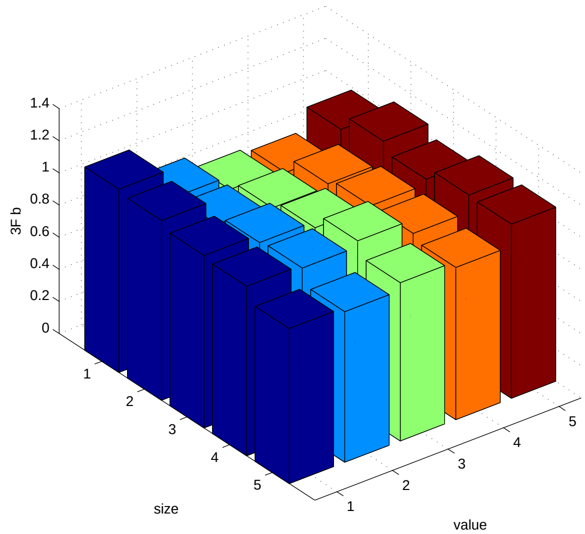
-0.65	0.09	0.16	0.29	0.23
-0.20	-0.04	0.11	0.02	-0.13
-0.14	0.04	0.12	0.01	0.25
0.14	0.04	-0.04	0.10	-0.14
0.19	0.14	-0.02	-0.09	-0.23

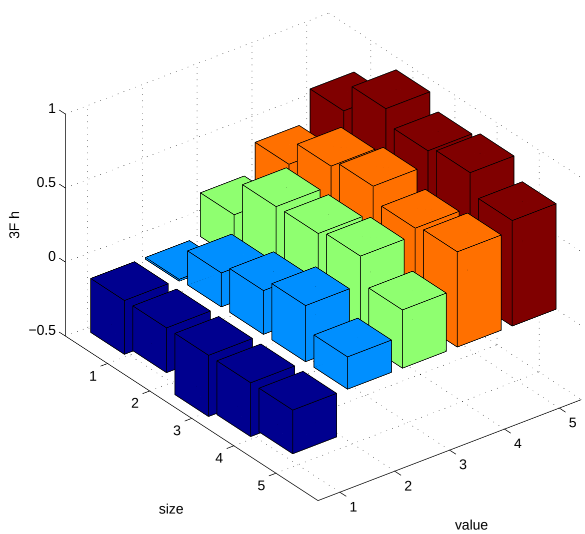
T on 3F alphas

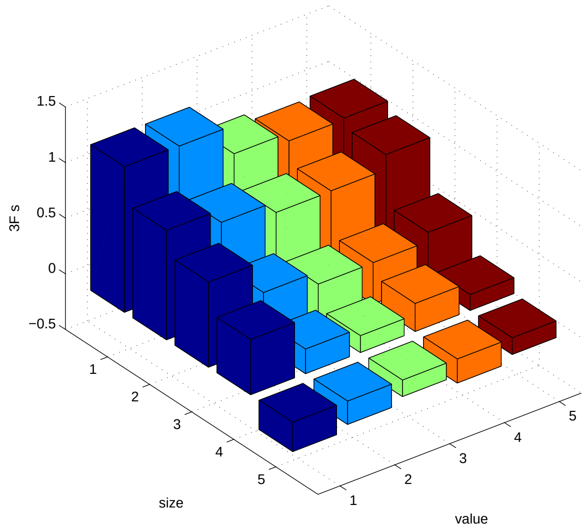
-3.67	0.70	1.65	2.87	2.25
-1.79	-0.39	1.19	0.20	-1.35
-1.26	0.30	0.96	0.05	1.86
1.31	0.33	-0.27	0.85	-1.07
3.12	1.55	-0.23	-0.88	-1.42

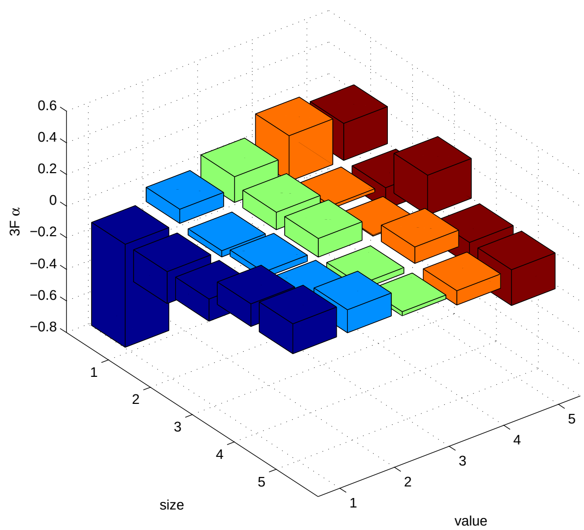
3F R2

0.90	0.93	0.94	0.92	0.93
0.95	0.93	0.93	0.93	0.94
0.94	0.87	0.86	0.86	0.86
0.93	0.87	0.85	0.86	0.87
0.96	0.89	0.87	0.89	0.80

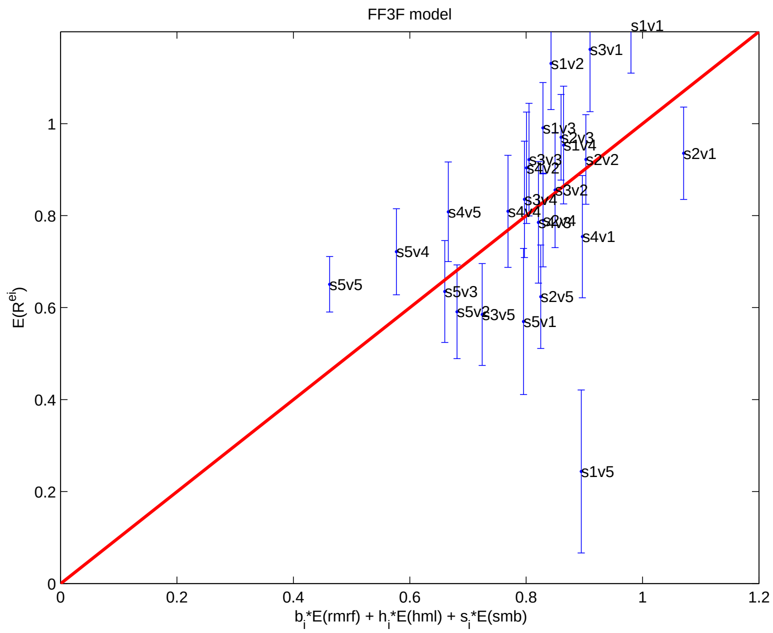

1.4
1.2
1
0.8
0.6
0.4
0.2
0
3F b
1
2
3
4
5
size
1
2
3
4
5
value


1
0.5
0
−0.5
3F h
1
2
3
4
5
size
1
2
3
4
5
value





Here's the actual vs. predicted plot



Overall, this is amazingly consistent performance out of sample. The t statistics are smaller, but that's because the sample is smaller. (One more warning against watching t statistics too closely!) The pattern of expected returns, alphas, b h, s, R^2 are all quite consistent. Interestingly the small growth puzzle got a lot worse.

3.

	rmrf	hml	smb
mean (pct/mo)	0.51	0.39	0.26
Sharpe	0.11	0.14	0.08
t	2.80	3.35	2.04
alpha and beta of on rmrf			
alpha		0.48	0.15
alpha sharpe		0.18	0.05
alpha t		4.34	1.23
beta		-0.19	0.21
beta t		-7.67	8.00

	rmse(a)	mean|a|	R2
FF3F model	0.136	0.099	0.913
rmrf hml	0.189	0.162	0.801
rmrf only	0.325	0.273	0.746
1 factor ff	0.136	0.099	0.399
weights			
	3.27	6.89	2.66
1 factor mv	0.000	0.000	0.100

```
weights
        -13.81          -2.01          -5.02          11.75           2.89
          5.27          -0.07           2.31          -9.33          -0.01
         -1.54           4.85          -0.67          -3.03           1.47
         10.38          -0.84          -0.81           4.65          -2.91
          6.51          -5.30           5.83          -3.39           0.38
```

a) The market premium of 0.5% per month is 6% per year. OK, about right. hml and smb have lower premiums – but everything here is a long-short portfolio, so the level of the premium isn't really that important.

Sharpe ratios are a better measure. ($E\left[2\times(R^a-R^b)\right]=2\times E\left[(R^a-R^b)\right]$ but $E\left[2\times(R^a-R^b)\right]/\sigma\left[2\times(R^a-R^b)\right]=E\left[(R^a-R^b)\right]/\sigma\left[(R^a-R^b)\right]$. Sharpe ratios are invariant to leverage or position size.) All three portfolios have about the same Sharpe ratio. So far they look about equivalent.

b) However, smb has a positive beta and hml has a negative beta. This means that smb's alpha is about half its premium, while hml's alpha is more than its premium. We have already seen that small stocks are priced by betas, and that the value effect betas go the wrong way. Here we see it again.

In the next row, I show the alpha sharpe ratio, $\alpha/\sigma(\varepsilon)$. This is the Sharpe you get from $smb-\beta_s rmrf = \alpha_s+\varepsilon_s$. You see that the small alpha sharpe is worse even than the small sharpe itself.

The alpha t statistic is the statistical test for dropping a factor. It says that smb is not *statistically* helping to account for expected returns of other assets. Beware statistics though, because longer samples make the same number more "significant"

c) So let's see how the models do. The FF model has about 10-13 bp alphas and 0.91 R^2, bottom line evaluation of "average returns" and "returns" respectively. The two factor model, without smb, does result in a somewhat higher alpha – 16-19 bp. The R^2 drops substantially. Remember, $\alpha_s=0$ only tells us that the model of *average* returns isn't affected. "Unpriced" factors help R^2.

Dropping hml is a much worse idea. the alphas rise to 27-35 bp, and the R^2 drops to 0.75.

d) The one-factor version of FF gives *exactly the same alphas as the FF model!* That's a theorem. *You can always coalesce a multifactor model into an equivalent single factor model of average returns. The mean-variance efficient combination of multiple factors gives you that magic single factor.*

Also, the weights spread out between market, a strong value tilt and a smaller size tilt, are pretty reasonable.

But the R^2 drops to 0.4! (And all the t-statistics, not shown, are much bigger.) The single factor model is *not* just as good a model of return variance. So, why did FF not reexpress their model this way? Because with the R^2 at 0.4, it's a much worse model *of returns.* And the t stats and standard errors are needlessly large.

e) 0?? Really? Yes indeed. *The ex-post mean-variance efficient portfolio prices all assets perfectly in a single-beta model.* This is the "Roll theorem" after UCLA's Richard Roll who first proved it*. It's an important point – factor fishing needs some rules of the game or you end up with perfection, even in a single factor.

This isn't a practical factor model, however. The weights are all over the place – and very unstable across subsamples.

And $R^2=0.1$. This is a horrible model of return variance.

* If you like looking at proofs, here is one. R^e is a $N\times 1$ vector (i.e. $N=25$), $\Sigma=cov(R^e,R^{e\prime})$ is

a $N \times N$ matrix.

$$\begin{aligned}
w &= \Sigma^{-1} E\left(R^{e}\right) \\
f &= w' R^{e} = E(R^{e})' \Sigma^{-1} R^{e} \\
\beta &= \frac{cov(R^{e}, f)}{var(f)} = \frac{cov\left[R^{e}, R^{e\prime} \Sigma^{-1} E(R^{e})\right]}{cov(f,f) = cov\left[E(R^{e})' \Sigma^{-1} R^{e}, R^{e\prime} \Sigma^{-1} E(R^{e})\right]} \\
\beta &= \frac{\Sigma \Sigma^{-1} E(R^{e})}{E(R^{e})' \Sigma^{-1} \Sigma \Sigma^{-1} E(R^{e})} = \frac{E(R^{e})}{E(R^{e})' \Sigma^{-1} E(R^{e})} \\
E(f) &= E(R^{e})' \Sigma^{-1} E\left(R^{e}\right) \\
\beta &= \frac{E(R^{e})}{E(f)} \\
E(R^{e}) &= \beta E(f)
\end{aligned}$$