**Problem set 10**

Part I. Questions

1. The log price of one, two and three year bonds is $p_0^{(1)} = -0.05$, $p_0^{(2)} = -0.15, p_0^{(3)} = -0.30$

   (a) Find today's yields and forward rates
   (b) Plot the expected evolution of these bonds' prices over time, according to the expectations hypothesis.
   (c) Plot the expected evolution of these bonds' prices for the first year, according to the Fama Bliss regressions, specializing all the coefficients to 1 and 0 as appropriate.
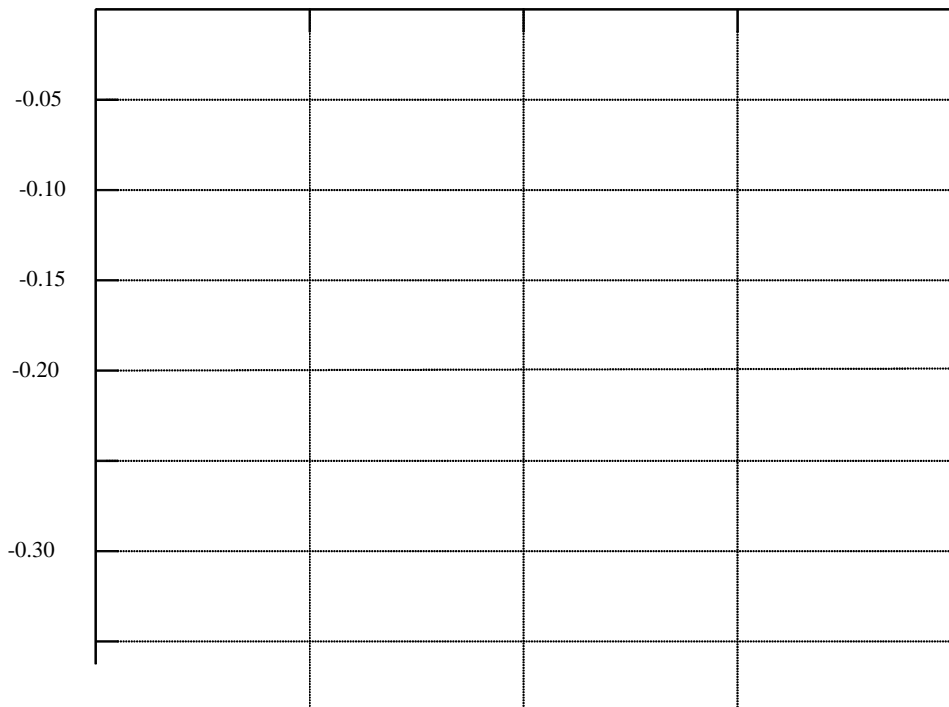
2. Forward curve

   (a) If the expectations hypothesis is true, what is the expected value at time t of future *forward rates* $E_t(f_{t+k}^{(n)})$ for $1 < k < n$? Hint: $E_t(E_{t+1}(x)) = E_t(x)$.
   (b) Suppose today's forward curve is (1%, 2%, 2.5%, 3%, 3.25%) at maturities 1,2,3,4,5 years. Plot the expected forward curves in the next 5 years (forward rate on y axis, maturity on x axis) as predicted by the expectations hypothesis. Visually, how are these curves connected?
   (c) Now, let's generalize a little bit. Suppose the forward curve is upward sloping $f_t^{(n)} > f_t^{(n-1)} > y_t^{(1)}$ for every n. Under the expectations hypothesis, we know this means that we expect short rates to rise in the future, $E_t\left(y_{t+n}^{(1)}\right) > y_t^{(1)}$ and $E_t\left(y_{t+1}^{(1)}\right) > y_t^{(1)}$ in particular. But what about other rates? Does the expectations hypothesis imply that longer-maturity forward rates are expected to rise, decline, or stay the same for the next year? I.e. if $f_t^{(n+1)} > f_t^{(n)}$ is $E_t\left(f_{t+1}^{(n)}\right)$ greater, less than or equal to $f_t^{(n)}$?

(d) Suppose instead that Fama and Bliss's one-year horizon regressions are right, and set all their regression coefficients to one to make it simple, i.e.

$$rx_{t+1}^{(n)} = 0 + 1\left(f_t^{(n)} - y_t^{(1)}\right) + \varepsilon_{t+1}^{(n)}.$$

$$y_{t+1}^{(1)} - y_t^{(1)} = 0 + 0\left(f_t^{(n)} - y_t^{(1)}\right) + \delta_{t+1}^{(1)}.$$

Find the prediction for the forward curve at $t+1$, $E_t\left(f_{t+1}^{(n)}\right)$ given information at time t in this case. (Hint: Use the definition of returns to see what $E_t\left(rx_{t+1}^{(n)}\right) = 0 + 1\left(f_t^{(n)} - y_t^{(1)}\right)$ implies for $E_t\left(f_{t+1}^{(n)}\right)$.) Express your answer as we did in a and c, i.e. $E_t\left(f_{t+1}^{(n)}\right) = ...$ and find the relation between $E_t\left(f_{t+1}^{(n)}\right) - f_t^{(n)}$ and the slope of the forward curve.

(Note: I do not ask you to extend Fama-Bliss forecasts past a year. If you do, you will find that you cannot really use 0 and 1. The coefficients must be strictly less, like 0.9 and 0.1. If not, you will find that the expected one year rate never rises, where FB's long run regressions say it does eventually rise. I'll explain in the solutions, for now the main point is, we are only looking one year ahead with the 0 and 1 approximation.)

3. This problem is designed to help you think about where the data violate the expectations hypothesis. Recall the standard pattern of interest rates over the business cycle. What's the problem? To investigate, I made up an interest rate pattern that looks like a typical cycle. Let y1 = [6 6 6 6 4 2 1 1 1 2 4 6 6 6]; represent a time-series of the one-year rate $y_t^{(1)}$ as we go through a recession.
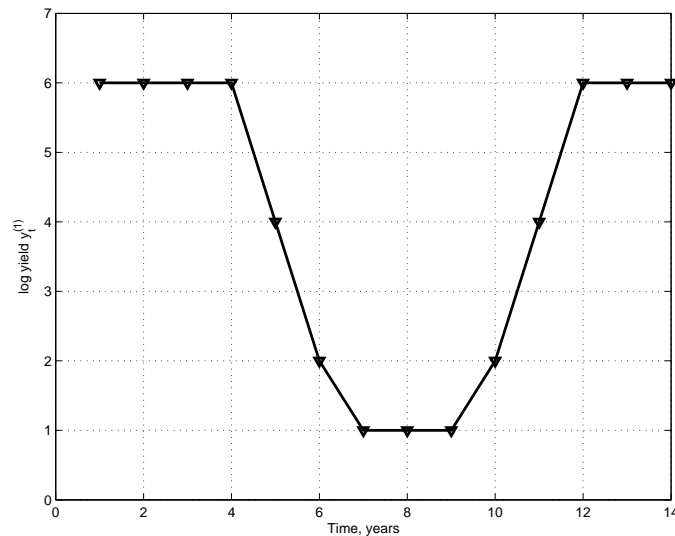

(a) Calculate the time series of the two and three year forward rates $f_t^{(2)}$ and $f_t^{(3)}$ at each date, assuming the pure expectations hypothesis and assuming everybody knows for sure that the given pattern of one year rates will happen. Make a plot of $y_t^{(1)}, f_t^{(2)}, f_t^{(3)}$. How is this plot different from the path of interest rates in a typical recession?

(b) To further explore, I made up a time series of "typical" forward rates in a recession,
    f2 = [6.2 6.2 6 5.75 4.5 3 2.5 2.5 2.5 3.5 5 6.2 6.2 6.2]';
    f3 = [6.4 6.4 6 5.5 5 4 3.5 3.5 3.5 4.5 5.5 6.4 6.4 6.4]';

Data

log yield $y_t^{(1)}$ and fwd. rates

Time, years
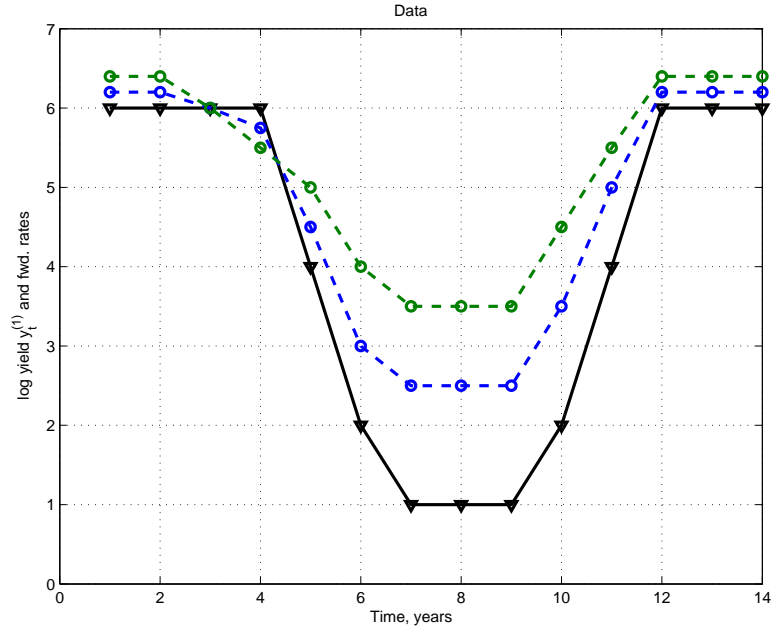
I put in a little forward curve inversion before the recession, and then the usual pattern of an upward sloping forward curve on the way out of the recession. Using these forward rates, calculate excess returns to 2, and 3 period bonds $rx_{t+1}^{(2)}$ and $rx_{t+1}^{(3)}$ through this episode, and put them on your plot.

Interpretation: If people know what the path of interest rates is going to be, then these are expected returns. Then the plot shows you the pattern of expected returns - risk premiums - that generates the forward rates in the first place. Remember "recessions" are the part where interest rates are falling and zero; "expansions" are the part where interest rates are rising. You should see a risk premium in "recessions."' If people don't know the path of interest rates, these are still the ex-post excess returns that people earn – high going in to recessions, low on the way out. The only issue is how much of them is expected– a premium – and how much is unexpected.

(c) Now, let's calculate the other side of the story. For each $y_t^{(1)}$, suppose the expectations hypothesis holds, and calculate and graph the forecast of $E_t y_{t+j}^{(1)}$. I did the first few to get you going. Fill in the rest

4

Expected path from forwards
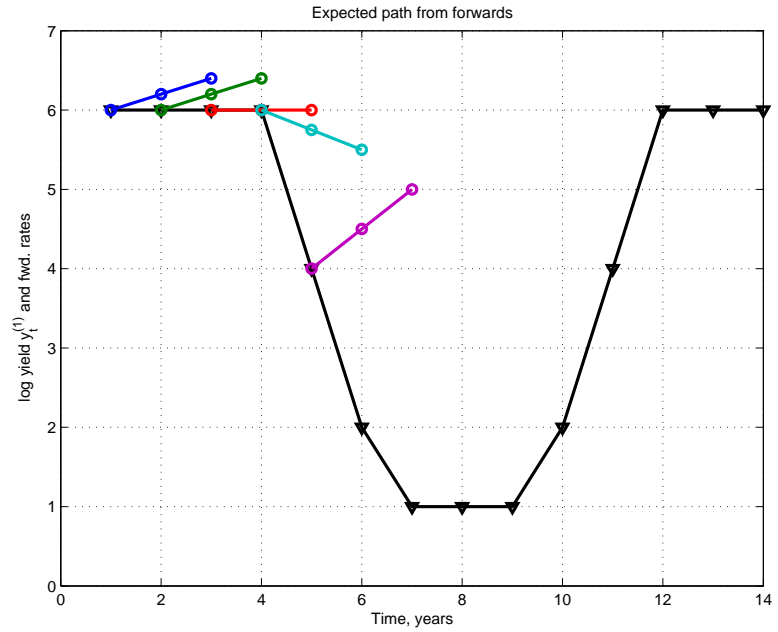
Interpretation. If the expectations hypothesis is correct (or under "risk neutral measure" if you know what that is) you're finding the pattern of expected future spot rates that makes sense of each day's forward rates. Then the puzzle is, though, that people's expectations are so wrong – that they expect interest rates to start rising right away for so long while interest rates keep declining and stay low.

Part II. Computer

1 Reproduce the regression coefficients in the Fama-Bliss regressions, using the updated data through 20131231 on the class website. Fill in the dots.

| | $rx_{t+1}^{(n)} =$ | | | $y_{t+n-1}^{(1)} - y_t^{(1)} =$ | | |
| | $a + b\left(f_t^{(n)} - y_t^{(1)}\right) + \varepsilon_{t+1}$ | | | $a + b\left(f_t^{(n)} - y_t^{(1)}\right) + \varepsilon_{t+n-1}$ | | |
n	a	b	R^2	a	b	R^2
2	0.15	0.83	0.12	-0.15	0.17	0.01
3	0.04	1.15	0.14	-0.57	0.51	0.05
4	.	.	.	.	.	.
5	.	.	.	.	.	.
	forecasting one year *returns* on n-year bonds			forecasting one year *rates* n years from now		

Use the sample starting in 1964:01 up to the most recent date.

You only need to reproduce the coefficients and R^2–you do not have to produce standard errors. These are overlapping monthly observations of annual returns, so, like the d/p regressions, the errors are correlated over time due to the overlap. In my answers, I correct for this overlap using the formula in *Asset Pricing* p.210 and the program olsgmm. (If you want to do this, use 12 lags and no weighting in the left hand panel and 12*n lags and no weighting in the right panel, but it's optional. As with returns on d/p, you can also do a rough job of standard errors by rerunning the regression using nonoverlapping data. We dealt with this issue running returns on d/p, so I will save your programming time for other tasks this week.)

The point of the problem, of course, is to verify that you understand the regression. The most common source of trouble is not implementing the definitions of forward rate and return correctly. The first log return on a two year bond is the Jan 1965 log price of a *one* year bond minus the Jan 1964 price of a *two* year bond. The second log return on a two year bond is the Feb 1965 log price of a one year bond minus the Feb 1964 log price of a two year bond. To make sure you're on the right track, here are the first few data points.

```
first prices
    19640831  -0.03813  -0.07468  -0.11706  -0.15964  -0.20099
    19640930  -0.03897  -0.07594  -0.11806  -0.15992  -0.19966
    19641030  -0.03896  -0.07735  -0.11931  -0.16016  -0.20049
first forwards
    19640831   0.03655   0.04238   0.04258   0.04135
    19640930   0.03697   0.04212   0.04186   0.03974
    19641030   0.03839   0.04197   0.04084   0.04033
first r
    19650831   0.03296   0.03533   0.03528   0.03450
    19650930   0.03248   0.03326   0.03190   0.02884
    19651029   0.03358   0.03397   0.03135   0.02642
first rx
    19650831  -0.00517  -0.00280  -0.00285  -0.00363
    19650930  -0.00649  -0.00571  -0.00707  -0.01013
    19651029  -0.00538  -0.00499  -0.00761  -0.01254
```

3. a) Replicate the basic Cochrane-Piazzesi regression. Run each bond excess return on all the forward rates, i.e.

$$rx_{t+1}^{(n)} = a^{(n)} + \beta_1^{(n)} y_t^{(1)} + \beta_2^{(n)} f_t^{(2)} + ... + \beta_5^{(n)} f_t^{(5)} + \varepsilon_t^{(n)} \tag{1}$$

Use data 1964-end for everything in this problem. Don't worry about standard errors. Make a plot of $\beta_i^{(n)}$ vs. i for each maturity n, as in the paper. Do the coefficients all have the same pattern across

maturity? Is it still tent-shaped, telling to look at the middle of the forward curve vs. the ends? Or is it slope-shaped, as Fama and Bliss suggest, telling you to look at the slope of the curve? Do you see the improved R^2 over the FB regressions you ran above?

b) Plot the fitted values of these regressions over time. There will be four lines,

$$E_t\left(rx_{t+1}^{(n)}\right) = a^{(n)} + \beta_1^{(n)}y_t^{(1)} + \beta_2^{(n)}f_t^{(2)} + \dots + \beta_5^{(n)}f_t^{(5)}; \; n = 2, 3, 4, 5.$$

This is a plot, at each time t of the model's expectation for returns of each bond over the next year. Do these lines seem to move together, or do they seem to wander around independently of each other? I.e., When $E_t(rx_{t+1}^{(2)})$ goes up, does $E_t\left(rx_{t+1}^{(3)}\right)$ seem to go up at the same time, though possibly more? Or do expected returns of bonds of different maturities seem to go their own ways? (A "one factor" model does not mean that they all move *by the same amount*. It means that they all move at the same time, with longer-maturities always moving by more than shorter maturities.)

The point of this problem is to see visually if there is really only one "factor" in *expected* returns. Also, it helps you to understand the idea that *expected* returns vary over time. (It will help to plot this for a shorter time interval than the whole dataset, so you can distinguish lines more easily.) By plotting the expected returns, you may understand the "one factor" business more directly than with the tent plots.

c) Now use the CP two-step procedure to estimate the restricted one-factor model, i.e. run

$$\frac{1}{4}\sum_{n=2}^{n}rx_{t+1}^{(n)} = \gamma_0 + \gamma_1 y_t^{(1)} + \gamma_2 f_t^{(2)} + \dots + \gamma_5 f_t^{(5)} + \varepsilon_{t+1};$$

$$rx_{t+1}^{(n)} = b_n\left(\gamma_0 + \gamma_1 y_t^{(1)} + \gamma_2 f_t^{(2)} + \dots + \gamma_5 f_t^{(5)}\right) + \varepsilon_{t+1}$$

Report the γ and the b coefficients. Make a plot of the implied $\beta_i^{(n)} = b_n\gamma_i$ and compare them to the plot of the unrestricted $\beta_i^{(n)}$ you made above. Make a plot of this model's predictions for $E_t\left(rx_{t+1}^{(n)}\right)$ as a function of time, on the same scale as the part b plot. This model is a "single-factor model" by *construction*: The expected returns $E_t\left(rx_{t+1}^{(n)}\right)$ vary over time only as one common "factor" $\gamma'f$ varies over time. Can you see how each of the lines moves together with the longer-maturity lines moving more? How close does this plot seem to be to the plot of unrestricted regressions in b?

d) (Note: It may be easier to do the next problem first, to familiarize yourself with factor analysis, and then come back.) Now, here is what CP should have done, with the benefit of hindsight. (And we did do in our later paper when we figured it out.) If you think there is a single-factor model of expected returns, just factor-analyze the expected returns. Use the unrestricted regressions (1) to form a time series of $E_t(rx_{t+1}^{(n)})$ for each n. You already did this and plotted it in part b. Now, form the 4 x 4 covariance matrix of *expected* excess returns, $cov\left(E_t(rx_{t+1}^{(n)}), E_t(rx_{t+1}^{(m)})\right)$ and factor analyze it.

This takes longer to say than to do, once you get the idea that *expected* returns form a time-series just like actual returns. Call the 4×1 vector of fitted values of (1) $E_t r = \left[\; E_t\left(rx_{t+1}^{(2)}\right) \quad E_t\left(rx_{t+1}^{(3)}\right) \quad E_t\left(rx_{t+1}^{(4)}\right) \quad E_t\left(rx_{t+1}^{(5)}\right)\;\right]'$ and create a $T \times 4$ vector of expected returns $etrx = \left[\; E_1 r_2' \quad E_2 r_3' \quad E_3 r_4' \quad \dots \;\right]'$, then take [Q,L] = eig(cov(etrx)). Display the eigenvalues L, verify one of them is huge compared to the others. Use q= the column of Q corresponding to the largest eigenvalue to form the common factor $f_t = q'E_t r$, i.e. to form the $T \times 1$ data vector f = etrx*q. Then also use q to form the fitted values, $E_t r^{fit} = qf_t$ or the data vector etrfit = f*q'; (We usually put data on $N \times 1$ vectors into $T \times N$ matrices. You can do it the other way, but all my programs use that convention.)

Plot the fitted values of this single-factor model for expected returns, $E_t r^{fit}$ and compare your plot both to the actual expected returns (part b) and to the CP ad-hoc single factor model (part c).

4. Let's do an eigenvalue factor decomposition (principal components) for the Fama French 25

portfolios. This is what FF were basically doing. Load up the Fama French 25 portfolio and factors data, and subtract off Rf to make the portfolios into excess returns. (Use the postwar sample 1947 on. It's a little prettier.) Form the covariance matrix of the FF 25 portfolio excess returns, and take the eigenvalue decomposition. ([Q,L] = eig(cov(rx)))

a) Plot the standard deviation of the factors (square root of Λ diagonals), and look at numbers. Does it suggest you can stop after a few factors?

b) Now, let's look at the four factors with largest standard deviations. First we want to look at the corresponding columns of Q. I found the plots most revealing by using the same size-by-book/market display that I used to display the FF model before. For example, here's the first factor loadings. (bar3(reshape(Q(:,end),5,5)')). Note matlab puts the biggest eigenvalue last) Since the loadings are all positive, it is a sort of level factor, yet interestingly it weights small firms even more than large firms (rmrf is *value* weighted, so would put lots of weight on large. If the bars were level, this would be equal weighted. Since small stocks *vary* more than big stocks, this analysis is even more than equal weighted.)

Make and display the same plots for factors 2-4, and explain *how can you analogously interpret the remaining three of the first four factors?* (Yes 4 – let's go one past FF and see what happens.)

First factor loadings
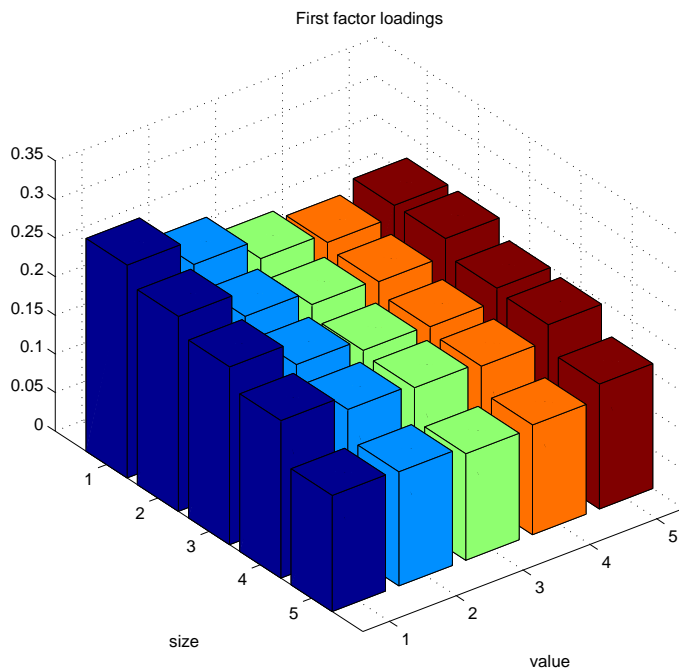
Note: in matching what comes out of this procedure with the Fama-French model, keep in mind that you can always use linear combinations of factors. Mathematically, if you have a factor model

$$R_t = q_1 x_t^{(1)} + q_2 x_t^{(2)} + q_3 x_t^{(3)}$$

then you could just as easily rewrite this as a factor model

$$\begin{aligned} R_t &= q_1 x_t^{(1)} + [q_2 + q_3] \left[x_t^{(2)} + x_t^{(3)} \right] + [q_2 - q_3] \left[x_t^{(2)} - x_t^{(3)} \right] \\ R_t &= q_1 x_t^{(1)} + \tilde{q}_2 \tilde{x}_t^{(2)} + \tilde{q}_3 \tilde{x}_t^{(3)} \end{aligned}$$

These factors will no longer have the special property that each in turn captures as much variance as possible, but once we've settled on how many factors we want, that's not so important anymore. The upshot is, this procedure may well produce factors that are *combinations* of hml and smb, especially since hml and smb have very similar variance.

c) Take the first four factors, construct a time series of the factors from the underlying 25 FF excess returns, $(x_t = q_i' y_t)$ and then run time series regressions of the 25 excess returns against these factors.

$$R_t^{ei} = \alpha_i + b_{1i} x_{1t} + b_{2i} x_{2t} + ... + \varepsilon_{it}$$

i.e. do just like Fama and French. Use i) the first three factors and then ii) the first four factors and, for comparison use iii) the Fama French factors.

You can use my tsregress2 or your program, so this should be a quick programming job.

```
factors = rx*Q(:,end-3:end); % produces T x 3 factors time series
[alpha, beta, R24, siga, sigb, chi2stat, chi2vals, chi2pv, N, Fstat, Fpvals, Fpv, TNK,u]...
= tsregress2(rx,factors);
```

You need only consider alphas, betas, and R^2.

i) Compare the regression coefficients b with the columns of Q. Do they have the expected relationship?

ii) How good are the 3 and 4 factor models as models of *variance?* Are the R^2 as high as FF found?

iii) How well do these models work as factor / factor pricing models? Are the alphas small? (There is no need to repeat GRS tests, etc. Just plot or make a table so you can see how big the alphas are. rmse alphas are a good summary.)

Note: It is *not* true that factor models will attack average returns or alphas one by one. The factor model finds in order the largest common movements in the *covariance* matrix of returns, looking for factors that produce high R^2. There is no mathematical reason that this should relate at all to the *mean* returns or alphas left over from the last factor model. In fact, you will see a pattern that each factor model attacks patterns seen in mean returns of the last factor model. This is *finance*, not math, it's a sign of the basic idea in finance working. The central idea in finance is that risk premia will attach only to common, undiversifiable movements in asset returns. And, lo, it does.