

### Problem Set 3

Due by Sat 12:00, week 3

**Part I. Reading questions:** These refer to the reading assignment in the syllabus. Please hand in *short* answers. Where appropriate, cite the text or tables for your answers. Example:

In the introduction, do Fama and French claim that their model explains all the puzzles thrown at it?

A: no, p. 56 bottom, "Equation (1), however, cannot explain the continuation of short-term returns." i.e. momentum.

These are not deep questions, they're just "did you get the point of the reading" questions.

Fama and French, "Multifactor explanations:"

1. In Table I, which kinds of stocks have higher average returns?
2. Would a stock with strong earnings growth be a "growth" stock by FF's definition? Would a stock with a small number of employees be "small"? (Hint: This is a trick question. How do FF define growth and value?)
3. Does the spread in average returns in Table 1A present a puzzle, by itself? (Hint: why might you not just go buy small value stocks based on the evidence of this table?)
4. How are FF's "SMB" and "HML" factors constructed? (one sentence)
5. Which gets better returns going forward, stocks that had great past growth in sales over the last 5 years, or stocks that had poor past growth in sales?
6. Which results show the "long-term reversal" effect in average returns best? Which show the "momentum" effect best? (Table and line)
7. It looks like we should all buy value, but we can't all buy value, someone has to hold the growth stocks. If we all try to buy value, the value effect will disappear because we drive up the prices. How to Fama and French address this conundrum? (hint, p. 76, 77)

FF "Dissecting Anomalies"

1. How do FF "Dissecting Anomalies" define "Microcap" and "small" stocks? What percentage of stocks are "Micro"? What fraction of market value do "micro" stocks comprise? How can the percentile breakpoint that defines "tiny" be different from the fraction of tiny stocks in the sample?
2. Are the average returns in Table II raw, excess, or adjusted somehow? Do they represent returns, or alphas, or something else?
3. Which anomalies produce strong average returns for all size groups in Table II? What are the most important numbers in Table II that document your answer?
4. Explain what the first two rows of MC and B/M columns mean in Table IV.
5. "The novel evidence is that the market cap (MC) result draws [size effect] much of its power from microcaps." (p. 1667) What numbers in Table IV are behind this conclusion?

Cochrane, “Discount rates” p. 1053-1064, and 1098-1099

1. Figure 6 says expected returns are higher for value portfolios. Does the paper say this observation constitutes the value puzzle?
2. What central feature of Figure 6 captures FF’s “explanation” of the value puzzle?
3. On p. 1060 I say "Covariance is in a sense Fama and French’s central result." What table or set of numbers in Fama and French convey this result?
4. What kind of regression does “Discount rates” suggest to provide the same information as FF’s Table 1A, in the same way we forecast returns last week? Write a regression equation that implements the “discount rates” idea.
5. Does cay help to forecast one year stock returns? How about long run stock returns?
6. In the final column of Figure 5, which components of the present value identity also change so that cay can help to forecast one-year returns without changing the forecast of the long-run dividend yield?

## Part II

How much of the variation of prices and returns that we have seen over time is due to shocks to news about cashflows vs. shocks to news about returns? We’re going to run some regressions and make some plots to answer this question. Use the same regressions you ran last week

$$r_{t+1} = a_r + b_r dp_t + \varepsilon_{t+1}^r \quad (1)$$

$$\Delta d_{t+1} = a_d + b_d dp_t + \varepsilon_{t+1}^d \quad (2)$$

$$dp_{t+1} = a_{dp} + \phi dp_t + \varepsilon_{t+1}^{dp} \quad (3)$$

to answer the following questions. Along the way you get some practice in inferring long-run features of the data from these sorts of regressions. (You can either use the real dividend growth data or dividends inferred from the identity. The solutions will use real data.)

1. We can use the approximate price-dividend identity

$$p_t - d_t = E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} - E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}$$

to decompose the price-dividend ratio *at each date* into the part due to expected future dividends  $E_t \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j}$  and the part due to expected future returns  $E_t \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}$ . Use the VAR (1)-(3) to calculate  $E_t \left( \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} \right)$ ,  $E_t \left( \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} \right)$  at each date Plot  $p_t - d_t$  over time along with the two terms on the right hand side of this equation.

(Notes: The formulas all apply to demeaned variables, i.e.  $x_t - E(x_t)$ . We’re interested in how things vary over time, not the levels. I took out all the means. The answer looks too simple to be true; each term will be some constant times  $p_t - d_t$  itself. That’s ok. If you forecast  $y_{t+1} = a + bx_t + \varepsilon_{t+1}$  then the forecast  $E_t y_{t+1}$  just follows  $x_t$ . If we use multiple forecasters on the right hand side, then the terms will not be perfectly correlated, but I wanted to keep it simple. )

2. But people use more than dp to forecast. As an opposite extreme, calculate the perfect-foresight decomposition. (This is a modern version of a Shiller graph):

$$p_t - d_t = \sum_{j=1}^k \rho^{j-1} \Delta d_{t+j} - \sum_{j=1}^k \rho^{j-1} r_{t+j} + \rho^k (p_t - d_t)$$

Use all the available data, i.e. at each date  $t$  up to the end of the sample  $T$ , i.e. plot each term of

$$p_t - d_t = \sum_{j=1}^{T-t} \rho^{j-1} \Delta d_{t+j} - \sum_{j=1}^{T-t} \rho^{j-1} r_{t+j} + \rho^{(T-t)} (p_T - d_T)$$

To start, plot the left hand side and the sum of the three terms on the right hand side. This lets you see how well the approximation works. Then, plot  $p_t - d_t$  with the return term and final term, and  $p_t - d_t$  with the dividend term and final term. (I found the graph with all three terms confusing.)

So, if people could forecast dividends and returns so well as to know the actual future, to what extent do expected future dividends vs. expected future returns account for variation in dividend yields?

3. Now, let's think about the astonishing 20% per year volatility of stock *returns*. Why are returns so volatile? Start with the identity

$$p_t - d_t = \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} - \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j}.$$

Now, take  $E_{t+1} - E_t$  of both sides, or equivalently think about the residual after running both sides on  $p_t - d_t$ . ( $\varepsilon_{t+1}^r = E_{t+1} - E_t(r_{t+1})$ .) That gives us

$$0 = (E_{t+1} - E_t) \left( \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} - \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} \right)$$

and solving, we get the unexpected return decomposition

$$\varepsilon_{t+1}^r = r_{t+1} - E_t r_{t+1} = (E_{t+1} - E_t) \left( \Delta d_{t+1} + \sum_{j=1}^{\infty} \rho^j \Delta d_{t+1+j} - \sum_{j=1}^{\infty} \rho^j r_{t+1+j} \right) \quad (4)$$

Return shocks must come from unexpected dividends, from news about future dividend growth, and from news about future returns. Positive news about future returns, holding future dividends fixed, means bad news about  $p_{t+1}$  and thus a lower return.

- (a) Use the regressions (1)-(3) and the approximate return identity to calculate the three components of the unexpected return  $\varepsilon_{t+1}^r$ , one due to current dividends, one due to future dividends and one due to future returns. (Hint: you end up with a formula of the form

$$(E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j \Delta d_{t+1+j} = \frac{??}{1 - \rho\phi} \varepsilon_{t+1}^{dp}.$$

This one is a constant times the dividend yield *shock* just as the last one was a constant times the dividend yield itself.

- (b) Plot the unexpected return  $\varepsilon_{t+1}^r$  and its three components defined by (4). (These won't all fit on the same graph, so split them up artistically. It's a good idea to start by making a plot to verify that the identity (4) works!)
- (c) How important are shocks to current dividends, shocks to future dividends and shocks to future returns in understanding the variation of returns? You should see a result quite different from the price-dividend ratio variance calculation. Why?

4. To quantify the graphs question of part b, we can do the same decomposition of variance for unexpected returns that we did for price-dividend ratios. From (4) you can multiply both sides by  $\varepsilon_{t+1}^r$  and take the mean, resulting in

$$\sigma^2(\varepsilon_{t+1}^r) = \text{cov}(\varepsilon_{t+1}^r, \varepsilon_{t+1}^d) + \text{cov}\left(\varepsilon_{t+1}^r, (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j \Delta d_{t+1+j}\right) - \text{cov}\left(\varepsilon_{t+1}^r, (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j r_{t+1+j}\right)$$

In words, the *variance* of returns must come from its covariance with dividends, or covariance with news about future dividends and returns. This looks like a bear, but in the last problem you reduced (4) to an equation of the form

$$\varepsilon_{t+1}^r = \varepsilon_{t+1}^d + (?d)\varepsilon_{t+1}^{dp} + (?r)\varepsilon_{t+1}^{dp}$$

so this variance decomposition is easy, just

$$\sigma^2(\varepsilon_{t+1}^r) = \text{cov}(\varepsilon_{t+1}^r, \varepsilon_{t+1}^d) + (?d)\text{cov}(\varepsilon_{t+1}^r, \varepsilon_{t+1}^{dp}) + (?r)\text{cov}(\varepsilon_{t+1}^r, \varepsilon_{t+1}^{dp})$$

Compute the four terms of this equation to numerically decompose the variance of unexpected returns. Express the answer as 1)  $100 \times \sigma^2(r)$  ( $0.20^2 = 0.04$  so multiplying by 100 makes it prettier) and 2) as percentages of  $\sigma^2(\varepsilon_{t+1}^r)$ , i.e.  $100 \times \text{cov}(\varepsilon_{t+1}^r, \varepsilon_{t+1}^d) / \sigma^2(\varepsilon_{t+1}^r)$ . Which components are most important?

### Part III Fama and French

We're going to replicate and extend Fama and French's basic results, using earlier data and extending the data to now. Get the 25 Fama French portfolios and the factors from the class website<sup>1</sup>. Start in 1963Q1 as FF did, but use all the data to the present. Subtract the risk free rate  $r_f$  from the 25 test assets to make them excess returns. The factors  $rmr$   $hml$   $smb$  are already excess returns. You do not have to form portfolios – Fama and French have done that for you.

- FF didn't show us that the CAPM does not work on their portfolios. Evaluate the CAPM on these 25 portfolios, as follows.

(a) Run time series regressions of  $R_t^{ei} = \alpha_i + \beta_i(rmr_t) + \varepsilon_t^i$   $t = 1, 2, \dots, T$  for each  $i$ .

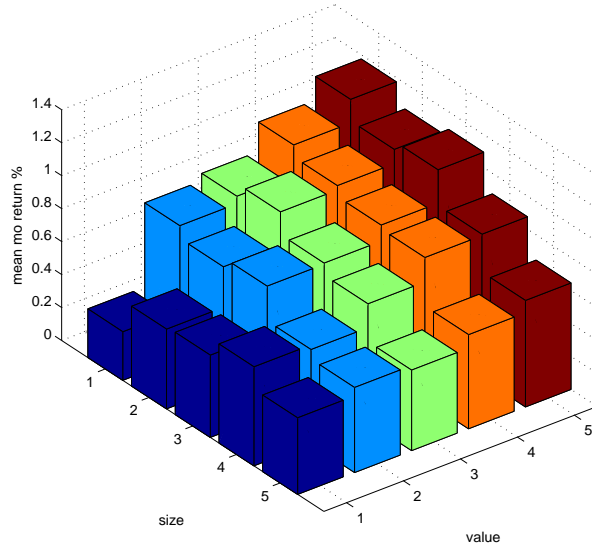
- Tabulate  $E(R^{ei})$ ,  $\hat{\alpha}$ ,  $\hat{\beta}$ ,  $t(\hat{\alpha})$ . Make tables in Fama-French format, i.e.

		Mean Return (% per month)		
		Low B/M	...	High B/M
		growth	...	value
small	.	.	.	.
	.	.	.	.
	.	.	.	.
	.	.	.	.
big	.	.	.	.
	.	.	.	.

- To see the patterns better, also make plots of expected excess returns, alphas, and betas in the 25 portfolios, i.e.

---

<sup>1</sup>If you ever want these data in the future, you can get them directly at [http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data\\_library.html](http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html)



and also

- iii. Make a plot of average excess returns  $E(R^{ei})$  (y axis) against betas (x axis), or average excess returns (y) against predicted values  $\beta_i E(R^{em})$  (x).
- (b) Does the CAPM work at all on the FF 25 portfolios? Look at the *statistical* significance – are the ts on the alphas greater than 2? Look at the *economic* significance – is there a pattern in average excess returns, and is that pattern mirrored in the betas so that  $E(R^e) = \beta\lambda$  has any hope? Does it work decently for some portfolios, but have trouble on a few, or is it pretty much not working anywhere?
- (In part we’re doing this to fill in the absence of a simple CAPM from Fama and French’s paper. You should always start by showing the last model does not work on your portfolios. FF did it in an earlier paper.)
- (c) Repeat the CAPM evaluation , using data from 193201 to 196212. How does the CAPM work in the earlier sample? (You don’t have to present everything, just the graph or table that best captures the similarities or differences you see in the earlier data.)

Next week, we’ll replicate the FF 3 factor model, but that’s enough for now.

## Programming hints

I left the data files with their original names as on Ken French's website so you will recognize them if you go there in the future. The original files have more stuff in them, and I extracted the parts you need.

```
st = (you find and fill in starting observation number)
x = load('25_Portfolios_5x5_monthly.txt');
x = x(st:end,:); % start after NAs end
r = x(:,2:end); % get rid of date column
T1 = size(x,1);
x = load('F-F_Research_Data_Factors_monthly.txt');
x = x(st:T1,:); % note factors often have more data, so stop when the returns stop
rmrf = x(:,2);
smb = x(:,3);
hml = x(:,4);
rf = x(:,5);
disp('Data sample'); % always a good idea to print out the data sample so you are sure it's right
disp([x(1,1) x(end,1)]);
T = size(rmrf,1); % define T and N for later use.
N = size(r,2);
rx = r-rf*ones(1,N); % make T x N matrix of excess returns
```

The biggest conceptual hurdle in doing everything with matrices is understanding rows vs. columns. In thinking about returns as random variables, we usually line them up in a column, i.e.

$$R \equiv \begin{bmatrix} R^1 \\ R^2 \\ \vdots \\ R^{25} \end{bmatrix}$$

Thus, we think of means, standard deviations, etc. as column matrices. However, when we use data on returns, we usually put time on the rows and returns on the columns.

$$\mathbf{R} = \begin{bmatrix} R_{1947}^1 & R_{1947}^2 & \dots & R_{1947}^{25} \\ R_{1948}^1 & R_{1948}^2 & \dots & R_{1948}^{25} \\ \dots & \dots & \dots & \dots \\ R_{2008}^1 & R_{2008}^2 & \dots & R_{2008}^{25} \end{bmatrix}.$$

In programming up formulas, you have to remember which way is which.

A lovely feature of matlab is that you can run 25 regressions at the same time. If  $f$  is a  $T \times K$  matrix of factors (e.g.,  $T \times 3$  for `rmrf`, `hml`, `smb`) and  $rx$  is a  $T \times N$  matrix of excess returns (e.g. the  $T \times 25$  matrix of `ff` portfolio returns) then you can run all the regressions at one time with

```
X = [ones(T,1) rmrf hml smb];
b = X\rx;
e = rx-X*b;
sigma = cov(e);
```

This produces all 25 alphas, all 75 betas and the variance covariance matrix of the residuals  $\Sigma$  (covariance across portfolios) in one fell swoop. (You might want to pull the alphas and betas out and make column vectors with

```
alpha = b(1,:)';
beta = b(2,:)';
h = b(3,:)';
s = b(4,:)'; )
```

The only statistics you need this week are the standard errors of the alphas and betas. You can use a regression function, but you can also just program up the formulas: Recall from stats that if  $y$  is  $T \times 1$  and  $X$  is  $T \times (K + 1)$  (including a constant) and the regression is

$$y = Xb + \varepsilon$$

the formula is

$$\begin{aligned} \text{var}(\hat{b}) &= (X'X)^{-1}s^2; \\ s^2 &= \frac{1}{T-K-1} \sum_t \varepsilon_t^2 \end{aligned}$$

The standard errors are the square roots of the diagonals of the matrix  $\text{var}(\hat{b})$ . We have 25 of these regressions going on side by side. The  $(X'X)^{-1}$  matrix is the same for each of the 25 regressions. Thus, you can do everything as follows

```
X = [ones(T,1) rmrh hml smb];
b = X\rx;
u = rx-X*b;
s2 = (T-1)/(T-K-1)*var(u)'; % produces a vector, variance of the 25 errors
% NOTE var uses 1/T-1, I corrected to 1/T-K-1

mx = inv(X'*X);
dmx = diag(mx); % we're interested in standard errors,
% the diagonals of the covariance matrix of bs

sigma = (s2*dmx(1)).^0.5; % standard errors of alpha, beta
sigb = (s2*dmx(2:end)').^0.5; % s2 is a column vector of 25. dmx' is a
% row vector corresponding to factors.
% this produces a matrix the same size as
% the b coefficients.

alpha = b(1,:)';
beta = b(2:end,:)'; % is now 25 x 3 with portfolios up and down
R2 = 1-s2./(std(rx).^2)';
```

Warning: don't forget to use  $x.^0.5$  and not  $x^0.5$  if you want to take the square root of the elements of a matrix.

To convert 25 x 1 vectors into FF style boxes the reshape function is useful,

```
meanr = mean(rx)';
disp(reshape(meanr,5,5)');
```

The mean function creates a row vector of means; ' makes it a column vector. The final ' puts size in the columns and value across the rows.

I used the bar3 function to make the bar graphs,

```
figure;
```

```

bar3(reshape(meanr,5,5)');
ylabel('size');
xlabel('value');
zlabel('mean mo return %');
print -depsc2 ps3a.eps;

```

The print command makes files that are easy for me to paste in to my word processor (latex). If you use word,

```

print -dpng ps3a.png;
print -djpeg ps3a.jpg;

```

etc. might work better. Use

```
doc print
```

for a list. The matlab function

```
cov(x)
```

will find the N x N covariance matrix of a T x N matrix, and

```
corcoef(x)
```

makes correlation matrices.

You can add things to a plot with

```

figure
plot(..)
hold on;
plot(..)

```

This is a good time to start learning about functions in matlab. I wrote a program which I call `tsregress.m`, it takes any  $T \times N$  matrix of excess returns, any  $T \times K$  matrix of factors, and runs time-series regressions of the returns on the factors, returning alphas, betas, t statistics,  $R^2$  and so forth. Since we will do this again, all you have to do next time is load new data and use the same program. Of course, cut and paste works too.

If there are any holdouts trying to do this in xcel, here's what I know about how to do it. (Courtesy of Toby Moskowitz and Gene Fama). Organize data in a  $T \times N$  array,

$$\begin{array}{c}
| \\
| \\
\text{Time} \\
| \\
|
\end{array}
\begin{bmatrix}
- & \text{industry} & - \\
rx_1^1 & rx_1^2 & \cdots \\
\cdots & \cdots & \cdots \\
rx_t^1 & rx_t^2 & \cdots \\
\cdots & \cdots & \cdots \\
rx_T^1 & rx_T^2 & \cdots
\end{bmatrix}
\begin{array}{l}
\text{market} \\
rmrf_1 \\
\cdots \\
rmrf_t \\
\cdots \\
rmrf_T
\end{array}$$

Then store mean returns, alphas, betas, etc. in rows across the bottom.



When you have to run lots of regressions (the 10 time series regressions here, or the 25 regressions coming up when we redo the Fama French model in problem set 5), the functions in 2-5 can be very useful. Using them, you can set up one regression over a data range (e.g. first return on the market), then copy that cell over to adjoining cells and automatically have betas etc. calculated for the other 9 (or 24) regressions.

1) The easiest way to run a regression in Excel is by using Excel's regression function: click on Tools, then Data Analysis (if this option is not available, click on Add-Ins, and select the Analysis ToolPak), then Regression, and then input your X and Y variables as directed.

2) Additionally, you could use the "slope" function (e.g., type `"=SLOPE(known_y's, known_x's)"`), which will output only the slope in the linear regression line through data points in known\_y's and known\_x's (e.g., in the market model equation above).

3) For multivariate regressions, you can also try the "linest" function, which works as follows: highlight a large output region, then type (without de-highlighting the output region) `"=LINEST(known_y's, known_x's, TRUE/FALSE, TRUE/FALSE)"` and press CTRL-SHIFT-ENTER to enter the formula. Note that you can input more than one X-variable here, unlike with the "slope" function. Also, for the first TRUE/FALSE option, enter TRUE if you want the constant in the regression calculated normally, and FALSE if you want the constant forced to equal 0; for the second option, type TRUE if you want additional summary statistics in your output, and FALSE if you do not. The format of the output is as follows: the first line contains the regression coefficients (with the constant listed last), the second line contains the standard errors of the coefficients, the third line contains the R2 and the standard error of the dependent variable, the fourth line contains the F-stat and the degrees of freedom, and the fifth line contains the regression sum of squares and the residual sum of squares.

4) The "index" function can be helpful when running multivariate regressions as well, as it outputs only the regression coefficient for the particular independent variable that you specify. For example, if you type `"=INDEX(LINEST(known_y's, known_x's),1)"` and then enter, the function outputs the coefficient for the first independent variable in your "known\_x's" range.

5) You can make covariance matrices easily with tools, data analysis, covariance. (You have to have the add-in as above).