

## Problem Set 6

1. These are “intuition” questions. Just give a sentence or two answer to each one. Don’t spend too much time on these.
  - (a) Consider the following tasks. In each case, state which of the following factors you would use: Fama and French’s *rmrf*, *hml*; a tradeable S&P500 index (which does not capture the whole market) and a tradeable value ETF (exchange-traded fund); a good measure of consumption growth; industry portfolios. State what regression you would run, what are the crucially important numbers you’d look at, and what else you might do with the regression. (Don’t worry about the zoo of additional factors that you would use in real life for any of these situations.) State whether the  $R^2$ ,  $\alpha$ , or other statistics are important.
    - i. A colleague finds that by sorting stocks based on transactions volume, the low-turnover stocks get much greater average returns than the high-turnover stocks. Is this pattern reflected by known risk factors?
    - ii. A colleague bugs you that Fama and French just explain size and book/market stock returns by HML and SMB returns. But, to him, the HML premium itself is sign of “irrationality.”
    - iii. You’re starting a hedge fund. Your clients demand that you have no market or value exposure and want as little volatility as possible.
    - iv. You want to evaluate an active manager – you suspect his returns are just due to market, value-growth exposure. Yet, you think that “value” is a not a “rational” risk factor.
    - v. You are running an event study, and you want to isolate the component of a company’s return on a given day that corresponds to firm-specific news.
  - (b) A buddy runs the CAPM regression.

$$R_{t+1}^i - R_t^f = \alpha_i + \beta_{i,R^m} [R_{t+1}^m - R_t^f] + \varepsilon_{t+1}^i$$

- i. He finds that the  $t$  statistics on  $\beta$  are below one. He concludes that the CAPM is a bad model. Can you offer any hope?
  - ii. He also finds that the  $R^2$  is only 20% in this time series regression. In regression class, you looked for good  $R^2$ . Does low  $R^2$  mean that the CAPM is not a good model?
  - iii. What effect does low  $R^2$  (Large  $\sigma^2(\varepsilon^i)$ ) have on standard errors  $\sigma(\hat{\beta}_{i,R^m})$  and  $\sigma(\hat{\alpha}_i)$ ?
  - iv. Another buddy comes along and says “Of course you’re getting small  $R^2$  in your time-series CAPM regressions. Everybody knows that stock prices are like random walks. You can’t predict returns.” Buddy 3 was paying attention in the first lecture, and says “No, returns are forecastable from d/p ratios. The  $R^2$  is fine.” Who is right?
- (c) Suppose you know (you do) that the CAPM is false, in that HML and SMB are important factors. Are there situations in which you would expect that the CAPM would still perform well and be an appropriate model to use? (Hint: The CAPM can be an APT too. Think of some assets that should really involve only market factors.)
  - (d) Looking across portfolios, you see that the regression errors are correlated with each other,  $\varepsilon_t^i$  is correlated with  $\varepsilon_t^j$ . Does this fact mean that the CAPM is not a good model? Do we need errors  $\varepsilon_t^i$  and  $\varepsilon_t^j$  to be uncorrelated with each other?

- (e) Noticing d, you try including an industry portfolio in the right hand side; for each stock you include the return on an industry index,

$$R_{t+1}^i - R_t^f = \alpha_i + \beta_{i,R^m} \left[ R_{t+1}^m - R_t^f \right] + \beta_{i,I} R_{t+1}^{Industry_i} + \varepsilon_{t+1}^i$$

You find that  $\beta_{i,f}$  is significant, and now the errors  $\varepsilon$  are much less correlated. Does this mean the CAPM is wrong? If you conclude the CAPM is right, does that mean you should always leave out the industry index? Are there reasons to leave it in (possibly for different uses of the model)?

- (f) If the CAPM were really right, how should the FF3F exercise have come out differently? Specifically, is it possible that you see the same  $\alpha_i$ , the same  $b_i, s_i, h_i$  as Fama-French do in

$$R_t^{ei} = \alpha_i + b_i r_{mrf_t} + s_i smb_t + h_i hml_t + \varepsilon_{it},$$

and that b,s,h are all strongly significant, but the CAPM is still right? What needs to be different for the CAPM to still survive? (You can assume that smb and hml are uncorrelated with r\_{mrf} if this helps you. What is the crucial piece of the FF3F model that I left out in my description?)

- (g) A buddy comes by with the following table of mean returns. The portfolios are made up of stocks sorted on his favorite trading rule, the “ratio-to-moving average” indicator. (Made up strategy and numbers)

low	2	3	4	high
10%	8%	5%	3%	1%

If we admit the phenomenon is real, i.e. not a statistical fluke or the result of selection and survival bias, does this imply the CAPM is dead and we need to include a “ratio-to-moving average” factor? Does it imply the 3 factor model is dead in the same way?

- (h) In regression class, you were often advised to test for nonlinearities. For example, you should try regressions like

$$R_{t+1}^i - R_t^f = \alpha_i + \beta_{i,R^m} \left[ R_{t+1}^m - R_t^f \right] + \beta_{i,R^{m2}} \left[ R_{t+1}^m - R_t^f \right]^2 + \dots + \varepsilon_{t+1}^i$$

You do, you find the squared terms are important, and the alphas change a lot. Does this mean the CAPM is wrong, either as an equilibrium model or as an APT, and you should include nonlinear terms? (Hint: If you find some alphas, what portfolio would you use to take advantage of them? Does that make any sense? )

- (i) You run a time-series regression of the consumption-based model,

$$R_{t+1}^{ei} = \alpha_i + \beta_{i,\Delta c} \Delta c_t + \varepsilon_{t+1}^i$$

where  $\Delta c_t$  is aggregate consumption growth. You find really big alphas. Could this model be right anyway? (Hint: what portfolio would you form to take advantage of the alpha?) .

- (j) You find that the cumulative return on hml  $\sum_{j=0}^k hml_{t-j}$  does a good job of forecasting the market return,  $R_{t+1} = a + b \left( \sum_{j=0}^k hml_{t-j} \right) + \varepsilon_{t+1}$ .

- i. How might this observation be interesting?
- ii. Why was it sensible to cumulate rather than just run  $R_{t+1} = a + b \times hml_t + \varepsilon_{t+1}$ ?

- (k) The expected return- beta model is  $E(R^{ei}) = \beta_i \lambda$ . This is true by construction – pick any  $\lambda$ , then just define  $\beta_i = E(R^{ei})/\lambda$ . What went wrong with this logic – how does the model have any testable content?

2. We derived the continuous-time capm,

$$(dR_t^i - r^f dt) = \beta_{i,W} (dR_t^W - r^f dt) + \sigma^i dz_t^i$$

i.e., the intercept is zero. Assume that returns are iid (which the CAPM assumes) lognormal diffusions, and a constant risk-free rate.

(a) Should you

i. Run the CAPM in logs, and look for  $\alpha^i = 0$

$$\log R_t^i - \log R^f = \alpha^i + \beta_{i,W} (\log R_t^W - \log R^f) + \varepsilon_t^i$$

ii. Run the CAPM in levels, and look for  $\alpha^i = 0$ . After all, portfolios are about arithmetic, not log returns.

$$R_t^i - R^f = \alpha^i + \beta_{i,W} (R_t^W - R^f) + \varepsilon_t^i$$

iii. Run the CAPM in logs, but look for an adjusted  $\alpha^i$

iv. Do something else?

(b) Evaluate how far off the wrong answers are for  $\sigma_\varepsilon = 100\%$  and  $\sigma_\varepsilon = 20\%$ , at both monthly and annual horizons. (i.e. give the “alpha” you might falsely measure, in units of monthly and then annual percent return. Yes, very small growth stocks have  $\sigma = 100\%$ . For monthly data, you can refer to Fama and French Table 1 for a sense of what is “big” and “small.” ) Use  $\beta_{i,W} = 1$  for this evaluation.

(**Hint:** think of value processes  $dV_t^W/V_t^W = dR_t^W$  and  $dV_t^i/V_t^i = dR_t^i$ ; You’re looking for a relationship between  $R_T^i = V_T^i/V_0^i$  and  $R_T^W$  that you can estimate by regression, and which looks as much like the CAPM as possible. State the right way to do it. You will don’t have to analyze deeply what happens with the wrong ways of doing it. One of them leads to an easy alpha calculation, for the other(s) just indicate qualitatively what goes wrong.)

3. Chapter 9, problem 2. The problem is a bit cryptic, so here is a restatement.

We tried to develop bounds on  $p(x)$  knowing the price  $p(f)$ , using only the law of one price. It failed. We showed that for given  $x$ , (or  $\varepsilon$ ), no matter how small  $\varepsilon$ , the solutions to

$$\min_{\{m\}} \text{ and } \max_{\{m\}} p(x) = E(mx) \text{ s.t. } p(f) = E(mf)$$

are infinite in both directions. Graphically, as the  $m$  sweeps out the entire plane through  $f^*$  in figure 9.1, we generate all possible prices for  $x$  and  $\varepsilon$ .

However, as suggested on p. 181 (“a natural first idea..”) it seems that by restricting  $m > 0$ , we restrict how far the discount factor can go on this plane, so we will produce upper and lower bounds for the price of  $x$  and of  $\varepsilon$  – just what we’re looking for!

The question, then, is whether this idea will work for continuously-distributed random variables such as stock returns, or whether the two-state diagram introduces other restrictions. Thus, the question is, can we derive the price bounds we’re looking for by

$$\min_{\{m\}} \text{ and } \max_{\{m\}} p(x) = E(mx) \text{ s.t. } p(f) = E(mf) \text{ and } m > 0$$

or

$$\min_{\{m\}} \text{ and } \max_{\{m\}} p(\varepsilon) = E(m\varepsilon) \text{ s.t. } p(f) = E(mf) \text{ and } m > 0$$

If we could do this, and presuming that the price bounds got tighter as  $\varepsilon$  got smaller in some sense (maybe  $\sup(\varepsilon)$  rather than  $E(\varepsilon^2)$ ), then we would have derived the APT as a limit, and we would have used the “arbitrage” that the title promises.

To answer this question, I found it useful (for once) not to construct  $m$  but instead to rely on theorems. We know that  $\exists m > 0$  such that  $p(x) = E(mx)$  and  $p(f) = E(mf)$  for any  $p(x)$ , so long as there are no arbitrage opportunities. So, we can see if there are upper and lower bounds on  $p(x)$  if we can find “arbitrage portfolios” of  $x$  and  $f$  that have non-negative payoffs, and insisting that those portfolios have non-negative prices.

With this insight, try to construct or show the absence of arbitrage portfolios, and hence upper and lower bounds in the following circumstances. Consider only a single  $f$ , and a risk free rate if you need it. Don’t worry about the difference between  $\geq$  and  $>$ .

- (a) A two-state example as in Figure 9.1. (Drawings will help here!)
  - (b)  $f$  and  $x$  are stock returns, which cannot be negative  $f \geq 0$  and  $x \geq 0$ . (Why?)
  - (c)  $f$  and  $x$  are excess returns, formed by borrowing at the risk free rate. These have  $f \geq -R^f$  and  $x \geq -R^f$  (Why?)
  - (d)  $f$  and  $x$  are excess returns, formed by selling one stock and buying another. These can range from  $-\infty$  to  $\infty$ .
  - (e) If you find that the two-state world of the figure was misleading, what did it implicitly assume that is not true of the real cases we have in mind?
4. Let’s try “what if Fama and French thought like Ross?” Get your 25 Fama French portfolio data and their factor data from the last problem set or Ken French’s website.
- (a) Start by replicating Fama and French. This week, you don’t have to do any statistics. Just
    - i. Make a table of mean excess returns on the 25 portfolios. Organize your table by size and value just like FF. (The reshape command is useful here.)
    - ii. Run each portfolio excess return on `rmrf`, `hml`, and `smb`. (You can do this in one command, `b = rhv\rets`. If `rets` is `T x 25`, this runs 25 regressions at once.)
    - iii. Make 5x5 tables of  $\alpha$   $b, h, s, R^2$
    - iv. Calculate  $\sqrt{\sum \alpha_i^2}/25$  as a simple measure of fit.
    - v. Make a plot of predicted  $b_i E(rmr f) + h_i E(hml) + s_i E(smb)$  vs. actual  $E(R^{ei})$  returns.
  - (b) Now, suppose they were thinking APT. Let’s start by factor-analyzing the covariance matrix of returns. You do this with `[Q L]=eig(cov(returns))`.  $Q$  is the matrix of eigenvectors and  $L$  the matrix of eigenvalues. Form factors from  $f = Q(:,x)'R_t$  i.e. *the factors are the columns of  $Q$  corresponding to the largest eigenvalues in turn*. Make 3 factors from the first three eigenvectors of  $Q$  (Note, matlab organizes eigenvalues from small to large, so these are the last three columns of  $Q$ )
    - i. Run each portfolio on the three eigenvalue factors
    - ii. Make 5x5 tables of  $\alpha$   $b_1, b_2, b_3, R^2$
    - iii. Calculate  $\sqrt{\sum \alpha_i^2}/25$  as a simple measure of fit.
    - iv. Make a plot of predicted  $b_i E(rmr f) + h_i E(hml) + s_i E(smb)$  vs. actual  $E(R^{ei})$  returns.
  - (c) Compare the APT model and the FF model.
    - i. Are the factors interpretable – do you see a pattern in the loadings  $Q(:,x)$  by which you form factors from returns?
    - ii. Do you see a relation between the  $b$  regression coefficients and the  $Q$  weights?
    - iii. Does the plot of predicted vs. actual returns look better or worse than FFs?
    - iv. Which gives smaller alphas?
    - v. How do the  $R^2$  compare?