Business 35905 **John H. Cochrane**

Problem Set 5

1. Today's yield curve is $y_t^{(1)} = 0.05; f_t^{(2)} = 0.10; f_t^{(3)} = 0.15; f_t^{(4)} = 0.15$.Make a graph of the expected bond price through time, with bond price on the vertical axis and time on the horizontal axis. Describe the expected path

   (a) According to the expectations hypothesis

   (b) According to Fama Bliss, assuming all of the one-year regression coefficients are 0 or 1 as appropriate. (Ignore the multi-year regressions and set all constants to zero througout the problem) .

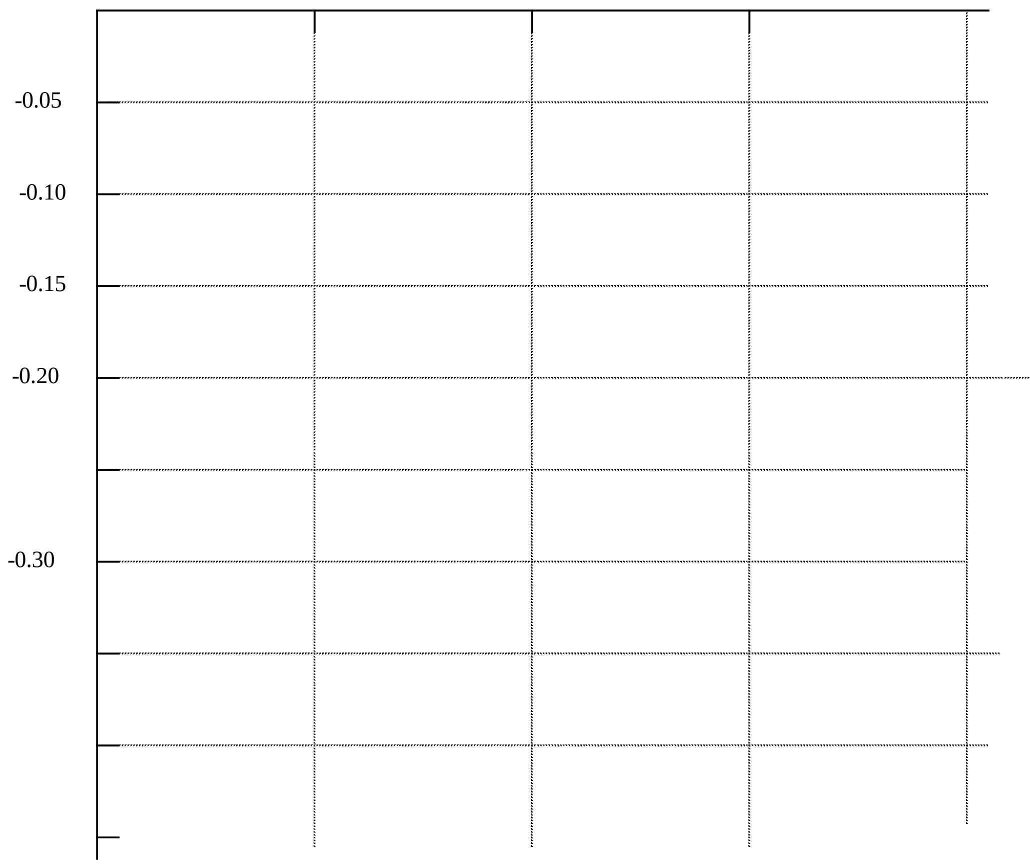

Figure 1:

   (c) Fama and Bliss find that the short rate is essentially a random walk. (Again simplify coefficients to 0 and 1 and forget about constants.) What do they say about longer rates; what $E_t f_{t+1}^{(n)}$ is implied by their regressions? (Hint, the picture will help. Then try to prove it.)

   (d) By contrast, what is $E_t f_{t+1}^{(n)}$ under the expectations hypothesis? We know that "if the yield (forward) curve is upward sloping, then we expect short term rates $y_{t+n-1}^{(1)}$ to rise in the future", i.e. $E_t y_{t+n-1}^{(1)} - y_t^{(1)} = f_t^{(n)} - y_t^{(1)}$. When do we expect *long term* rates to rise in the future? What is the signal that $E_t f_{t+1}^{(n)} > f_t^{(n)}$?

(e) Is your graph consistent with Fama and Bliss' long-run ($y^{(1)}_{t+1} - y^{(1)}_t$ on $f^{(n)}_t - y^{(1)}_t$) regressions? If not, how might you modify the setup (Just give a sentence and the general idea, you don't have to work out a lot of algebra).

2. In this problem we'll explore bond returns following up on Cochrane Piazzesi etc. First, get the Fama-Bliss discount bond price data from the class website. (It's from CRSP (via wrds is the easy way. Ask wrds for "quarterly updates" to get the most recent data.)) As in CP, start in 1964 (the bond data aren't very reliable before that). We'll use the full sample so we can see how CP and FB are doing.

   (a) Form log prices, log yields, log forward rates, one year log returns and one year excess log returns. (Hint: You're using monthly data so $r^{(n)}_{t+12} = p^{(n-1)}_{t+12} - p^{(n)}_t$. The first row of returns from jan 1964 to jan 1965 is, in percent units, 3.6962 3.9504 3.9390 4.0114.) Plot log yields, log forward rates and excess log returns to make sure you are transforming the data correctly.

   (b) Make an eigenvalue factor decomposition for i) yields ($Q\Lambda Q'$ =eig(cov(y)) ii) Δyields iii) forward rates iv) excess returns. (See note below) In each case,

      i. Make a table of the variance of the 5 (or 4 for excess returns) factors as a fraction of overall variance ($100 \times \lambda_i / \sum \lambda_i$), and the standard deviation of the factors in percent units ($100 \times \lambda_i^{0.5}$). This table (and the above graphs) should convince you that the first two or three factors capture almost all the variance of all these quantities

      ii. Make a plot of the loadings, the columns of Q. Make two plots: 1) $Q(:, i) * \lambda_i$, which expresses the loadings on unit variance shocks 2) Q which shows the loadings on shocks with variance λ_i. Together, you should see that in all these cases the main factor has a "level" pattern, there is a second smaller "slope" factor, a third very small "curvature" factor. The additional factors are tiny, and don't have particularly interpretable loadings – their loadings represent idiosyncratic movements of the individual bonds. (Note: if there is a "level" shock to yields, what should happen to returns of different maturities? Yes, this is the same "level" shock even though it has a different shape.)

      iii. Make a plot of the fitted value of yields, forwards, and excess returns, using only the first three principal components, together with the actual values (as presented in class, but extending to forwards and excess returns). (My plot focuses on 1998 - present for clarity, but play with the visual to tell a story.) Do all three series work as well as yields? Can you explain any differences?

      iv. Save the level slope and curvature factors in forward rates – you'll need them later.

   (c) Now let's reproduce and update Fama-Bliss

      i. Run the four Fama-Bliss one-year return regressions

$$rx^{(n)}_{t+1} = a_n + b_n(f^{(n)}_t - y^{(1)}_t) + \varepsilon^{(n)}_{t+1}$$

      present the coefficients, t statistics, and R^2. Is this still working as FB said it would? (Note: the data are overlapping observations of annual returns. so the ε_t are correlated. Use a Hansen-Hodrick correction – see the GMM chapter in *Asset Pricing.* You may use my olsgmm. Do you have to correct R^2?)

      ii. We showed in class that the coefficient b_2 in $rx^{(2)}_{t+1} = a_2 + b_2(f^{(2)}_t - y^{(1)}_t) + \varepsilon^{(2)}_{t+1}$ is equal to one minus the coefficient b^*_2 in $y^{(1)}_{t+1} - y^{(1)}_t = a^*_2 + b^*_2(f^{(2)}_t - y^{(1)}_t) + \varepsilon^{*(2)}_{t+1}$. Find the "complementary regressions" to the remaining excess return regressions, run them, and confirm that the coefficients add up to one.

      iii. Verify FB's long-run interest rate change regressions,

$$y^{(1)}_{t+n-1} - y^{(1)}_t = a_n + b_n\left(f^{(n)}_t - y^{(1)}_t\right) + \varepsilon^{(n)}_{t+n-1}$$

Does the expectations hypothesis still seem to work in the long run?

(d) Now, we'll do "What CP should have done" (or what they would have done, if they had figured this out in time!) The essence of the CP observation is that a single factor dominates variation over time in *expected* excess returns. So, let's write this as a single factor model, just as you did above with yields, etc.

i. Start by forming an unrestricted forecast of excess returns using all forward rates, i.e. run

$$\begin{aligned} rx_{t+1} &= \alpha + \beta f_t + \varepsilon_{t+1} \\ 4 \times 1 &= 4 \times 1 + (4 \times 5)(5 \times 1) \end{aligned}$$

$$f_t = \left[\begin{array}{cccc} y_t^{(1)} & f_t^{(2)} & ... & f_t^{(5)} \end{array} \right]'$$

Print and plot the coefficients and R^2. You'll see some differences relative to CP. If you want to see CP results, stop the data in 2006. (Note: matlab lets you run several regressions with one command. Thus, if rx is $T \times 4$ and f is $T \times 5$, then [ones(T,1) f]\rx performs all the regressions at once.)

ii. Now we have an unconstrained model,

$$E_t(rx_{t+1}) = \alpha + \beta f_t$$

a) Plot these expected returns over time. Do they look like they have a factor structure? Find the eigenvalue decomposition of the covariance matrix of *expected* returns, $Q\Lambda Q' = eig(cov(E_t\,(rx_{t+1}))$.

b) Plot Q and $Q\Lambda^{1/2}$, and c) tabulate $\lambda_i / \sum \lambda_i$ to get a sense of the shapes and names of the factors, and how many are important. Conclude that there is a *single* dominant factor that explains the variation over time of expected excess returns of all maturities, i.e. that expected excess returns of each maturity move together. Does your dominant column of Q have a tent shape? If not, why not?

iii. Now, let's keep only the biggest factor. That means we keep the factor (a scalar)

$$x_t = Q(:, end)' E_t(rx_{t+1});$$

and then we model expected returns across maturities as functions of the factor by

$$Erx_t^* = Q(:, end) x_t$$

(check that $Q(:, end)$ corresponds to the largest eigenvalue) just as you would have done in keeping a single factor model of yields. To see how this idea works, add Erx_t^* below your plot of the unconstrained estimate $E_t(rx_{t+1})$ above, to see if they look the same, as we did for yields

(e) We saw above that you span all the economically interesting movement in yields, forward rates, and so forth with level, slope and curve factors. Surely we can understand this new factor x as a linear combination of those, right? To see if this is true, let's run regressions using the yield factors,

$$rx_{t+1}^{(n)} = a_n + b_n x_t + c_n (\text{level}_t) + d_n\,(\text{slope}_t) + e_n\,(\text{curve}_t) + \varepsilon_{t+1}^{(n)}.$$

If, as the first section surely suggests, the information in x_t is spanned by the remaining facotrs, then we should see b_n small and statistically insignificant. Is it?

(Moral of the story: factors that explain nearly all the variance of yields and forward rates do not necessarily capture all the variance of *expected returns.* It is a very common practice in finance and macro to *first* reduce a large number of variables to a few common factors, and *then* use the common factors to forecast. For example, Jim Stock and Mark Watson, "Forecasting Using Principal Components from a Large Number of Predictors," Journal of the American Statistical Association, 97 (December 2002), 1167–1179 and a large number of subsequent applied papers. This problem is a counterexample. The linear combination that is good for forecasting is not well spanned by the linear combinations of yields that account for variance. It's almost guaranteed with near unit root series – spreads will be good for forecasting, but levels will account for variance.

If you're bored, try it with forward rate factors. You'll see the factor model does a much better job. If you're patient, wait for the solutions. Also if you wonder what happened to R^2 try the sample up to 2006.)

3. Let's modify the basic Vasicek term structure model, and see if we can account for Fama-Bliss regressions. The basic model has a constant market price of risk. We need to have a time-varying price of risk. The obvious way to do that is just to make the price of risk depend on the single factor. So, let's pursue the obvious extension, in which rather than just λ we have a time-varying $\lambda_t = \lambda_1 x_t$,

$$\begin{aligned} x_{t+1} - \delta &= \rho(x_t - \delta) + \varepsilon_{t+1} \\ \log m_{t+1} &= -x_t - \frac{1}{2}(\lambda_1 x_t)^2 \sigma_\varepsilon^2 - (\lambda_1 x_t)\varepsilon_{t+1} \end{aligned}$$

(In real life you'd allow a constant too, $\lambda_t = \lambda_0 + \lambda_1 x_t$, but I'm simplifying the algebra a bit.)

(a) Find $p_t^{(1)}, p_t^{(2)}$, hence $y_t^{(1)}, f_t^{(2)}$, in this model. Hint: they are still linear functions (stuff)+ (stuff) x_t! You have to use $Ee^x = e^{Ex+\frac{1}{2}\sigma^2 x}$, exactly as we did in lecture. Follow the steps as in lecture!

(b) Find $E_t rx_{t+1}^{(2)} = E_t p_{t+1}^{(1)} - p_t^{(2)} + p_t^{(1)}$. Again it's a function $E_t rx_{t+1}^{(2)} =$ (stuff) + (stuff) x_t.

(c) Find the predicted value of the Fama-Bliss slope coefficients, i.e. find β in $E_t rx_{t+1}^{(2)} = \alpha + \beta(f_t^{(2)} - y_t^{(1)})$. All you're doing here is substituting out the previous results. You had

$$E_t rx_{t+1}^{(2)} = a + bx_t$$

and

$$(f_t^{(2)} - y_t^{(1)}) = c + dx_t$$

so you're just getting rid of x_t on the right hand side in favor of $f_t^{(2)} - y_t^{(1)}$.

$$E_t rx_{t+1}^{(2)} = (\text{mess}) + \frac{b}{d}(f_t^{(2)} - y_t^{(1)})$$

Forget the mess in the constant, we're only interested in the slope coefficient, b/d.

(d) Can we find λ_1 so that this model captures the Fama-Bliss slope coefficient that b/d is approximately 1? (λ_1 describes how much the market price of interest rate risk varies over time. Note ρ is pinned down by the slope of the yield curve so you don't really get to play with it. It is a number less than one, closer to one the flatter the yield curve is. You may comment how ρ closer to one helps the cause of a Fama Bliss slope coefficient of one however.)