Business 35150 **John H. Cochrane**

Problem set 9

Part I questions

1. Interest rates. What is a reasonable value for
   (a) the log price of a one year pure discount bond, in the conventions of the readings $p_t^{(1)}$?
   (b) the log yield of a one year pure discount bond, $y_t^{(1)}$?
   (c) the actual price $P_t^{(1)}$?
   (d) the actual yield $Y_t^{(1)}$?
   In each case, choose between 95.00, 0.95, 0.095, 5, 1.05, 0.05, -5, -1.05, -0.05.

The next set of problems give you some practice with the definitions, and derive some important relationships between prices, forward rates, yields, etc.

2. Show that the price is the discounted value of a dollar, using forward rates to discount. Do this in logs. You're looking for $p_t^{(n)} = -\left(y_t^{(1)} + f_t^{(2)} + ... + f_t^{(n)}\right)$. Hint, start with $-p_t^{(n)} = \left(0 - p_t^{(1)}\right) + (p_t^{(1)} - p_t^{(2)}) + ...$(this is an important point – you can always use forward rates to discount certain payments)

3. Therefore, show that the yield = average of intervening forward rates (and express that form)

4. And conclude that if forward rates equal expected future spot rates, $f_t^{(n)} = E_t y_{t+n-1}^{(1)}$ the expectations hypothesis yields = average of expected future spot rates holds

5. Now, differencing in time, show that the current price is the negative of the sum of bond returns all the way to maturity. Start with $-p_t^{(n)} = \left(0 - p_{t+n-1}^{(1)}\right) + \left(p_{t+n-1}^{(1)} - p_{t+n-2}^{(2)}\right) + ...$ and substitute to get the relation between $p_t^{(n)}$ and lots of $r_{t+k}^{(j)}$ getting all the j and k right. Does the fact that you have a variable known at time t on the left and lots of variables known only in the future on the right bother you? What does that tell you about bond returns?

6. Show that the two year forward spread mechanically equals the rise in the one year spot rate plus the excess return on a two year bond

$$f_t^{(2)} - y_t^{(1)} = \left(y_{t+1}^{(1)} - y_t^{(1)}\right) + \left(r_{t+1}^{(2)} - y_t^{(1)}\right)$$

   (hint, just start with the definitions)

7. *Do this one, its important* Therefore, show that if you run regressions of one year interest rate changes and returns on forward-spot spreads, the coefficients and errors add up *just as we found in week 2 with long-run return regressions on dp.* In equations, if you run

$$\begin{aligned} \left(y_{t+1}^{(1)} - y_t^{(1)}\right) &= a_y + b_y\left(f_t^{(2)} - y_t^{(1)}\right) + \varepsilon_{t+1}^y \\ \left(r_{t+1}^{(2)} - y_t^{(1)}\right) &= rx_{t+1}^{(2)} = a_r + b_r\left(f_t^{(2)} - y_t^{(1)}\right) + \varepsilon_{t+1}^r \end{aligned}$$

   Find the relationship between a_y, a_r, b_y,b_r and between ε^y and ε^r. How is this like our dividend yield regressions?

Part II computer

The bootstrap and monte carlo are very useful techniques. To explore them, get the data from problem set 1.

1. Start by running forecasting regressions of log returns and log dp on dp

$$r_{t+1} = a_r + b_r dp_t + \varepsilon^r_{t+1} \quad (1)$$
$$dp_{t+1} = a_{dp} + b_{dp} dp_t + \varepsilon^{dp}_{t+1} \quad (2)$$

   Use the 1947-end of sample period. Tabulate the slope coefficients b_r, b_{dp}, standard errors, and the correlation of the errors. That's important – the errors are quite correlated.

2. Now bootstrap. Resample the *errors* ε^r_t , ε^{dp}_t from the regression residuals, with replacement, keeping the errors at each date together to preserve their correlation. Then generate new data by feeding it through the system (1)-(2). Rerun the regression in each artificial data set. Do this many times (in a loop) and keep $\hat{b}_r$ and $\hat{b}_{dp}$ from each trial.

   Why? In any sampling experiment we have to decide what is correlated and what is not correlated. If we resampled the actual variables dp_t and r_t, we would lose the correlation over time that is so crucial to the whole business. By sampling the errors and reconstructing the variables, we preserve the correlation over time of the variables. We also sample the errors together, so as to preserve the correlation between the errors ε^r_t and ε^{dp}_t. We *are* making assumptions here – we are assuming that the errors t and $t-1$ are not correlated. We are also forming a null hypothesis around our sample estimates b_r and b_{dp}.

   Tabulate the mean and standard deviation of the $\hat{b}_r$ and $\hat{b}_{dp}$ across your trials. Compare the means to the regression values Are the estimates biased? (Is the mean of the bootstrapped estimates close to the assumed values b_r and b_{dp}?) Did the OLS standard error formulas understate or overstate the sampling uncertainty – standard deviation of $\hat{b}$ across trials – revealed in the bootstrap?

3. Make a histogram of your $\hat{b}_r$ $\hat{b}_{dp}$ estimates, and include the "true" (assumed null hypothesis) values from the original regression. Are the distributions visibly different from a normal? (Note: who cares? A: Hypothesis tests care. If you computed 5% values, they would be based on normal distributions. "Fat tails" of the $\hat{b}_r$ $\hat{b}_{dp}$ distributions mean these tests are wrong.)

4. Now, you will see that the distributions of $\hat{b}_r, \hat{b}_{dp}$ are not normal. Does this come because the errors ε were not normal? Repeat, but this time in each simulation choose $\left\{\varepsilon^r_t, \varepsilon^{dp}_t\right\}$ to be normally distributed with mean zero and covariance matrix equal to the covariance matrix of the residuals. (See programming hints.) Simulate again, and tabulate as above. Do you get distributions of $\hat{b}_r$ and $\hat{b}_{dp}$ that are less biased or more normally distributed?

5. Now, these simulations are centered around the regression values $b_r \approx 0.12$ and $b_{dp} \approx 0.94$. A "hypothesis test" asks the question, suppose the truth were $b_r = 0$, $b_{dp} = 0.94$, how often would we see a $\hat{b}_r$ equal to or higher than the $b_r \approx 0.12$ that we see in our data? Let's find out.

   Redo your bootstrap, resampling the errors $\left\{\varepsilon^r_t, \varepsilon^{dp}_t\right\}$ as before, but this time form your simulated data r, dp from a vector autoregression with $b_r = 0$ and $b_{dp} = 0.94$. Run the regression in simulated data, and tabulate again the mean and standard deviation of $\hat{b}_r, \hat{b}_{dp}$. This time, also tabulate the fraction of your simulations that produce a sample $\hat{b}_r$ greater or equal to the value $\hat{b}_r \approx 0.12$ that we observed. This is the "hypothesis test" and the number is supposed to be less than 5%. Is it?

6. To gain insight, repeat your plots of the distribution of $\hat{b}_r$ and $\hat{b}_{dp}$, including the sample values $b_r \approx 0.12$ and $b_{dp} \approx 0.94$ we found before.

7. You will conclude that it's hard to reject $b_r = 0$. But also make a scatterplot of $\hat{b}_r$ *vs* $\hat{b}_{dp}$, and include our sample values $b_r \approx 0.12$, $b_{dp} \approx 0.94$. Are samples in which $\hat{b}_r$ is falsely high also

samples in which $\hat{b}_{dp}$ is off in one way or another? Is it likely that we see *both* high b_r *and* high b_{dp}?

Note: The last point is important in the return predictability papers. A bunch of papers noticed the upward bias in b_r and claimed predictability is all spurious. Other papers, (including my "dog that did not bark") noticed the joint behavior of b_r and b_{dp}. All the samples with spuriously high b_r have spuriously low b_{dp} since the true b_{dp} can't be greater than one, and our b_{dp} is so large, this means that our world can't come from one with much lower b_r than what we actually see.

Programming hints:

```
selectors = floor(rand(T,1)*(T)+1)
```

produces a vector T long of random integers, each between 1 and T. Then

```
error_trial = err(selectors,:)
```

takes a T x 2 vector of errors (return, dp) and selects from them randomly with replacement.

```
hist(betas,n)
```

produces histograms, with n bins

```
plot(betas(:,1),betas(:,2),'.')
```

produces a scatterplot of the first simulated beta against the second one

```
sigma = cov(err);
D = chol(sigma);
error_trial = randn(T,2)*D;
```

If err is Tx2, sigma is 2x2 covariance matrix. this cool technique produces a Tx2 draw of random normals with the same covariance matrix. Try

```
cov(error_trial)
```

and you should see that up to sampling error it is the same as the original sigma.