Empirical Asset Pricing:
Problem Set 6

March 18, 2024

In this problem set, we first use the methodology that we have seen so far in equity markets to understand which shocks are priced in bond markets. This will help us to specify an affine term structure model that we will subsequently estimate.

During the lectures we have seen the following facts

1. Bond risk premia increase in maturity.
2. Bond returns are predictable by a single factor, the Cochrane and Piazzesi (CP) factor.
3. Two or three factors capture most of the variation in bond yields, but the CP factor is not well explained by the three factors.

We explore an affine term structure model with three factors: the CP factor and first two principal components.

1. Data preparation:

   (a) Download yield data from 1971.11 to 2024.03 from here. Select the first 10 zero-coupon yields (maturities from 1 year to 10 years). Keep the month-end values.

   (b) Construct *annual* forward rates as

   $$f_t(n) = ny_t(n) - (n-1)y_t(n-1).$$

   (c) Compute *annual* excess (log) returns on each bond, at a monthly frequency,

   $$rx_{t+1}(n) = p_{t+1}(n-1) - p_t(n) - y_t(1),$$

   where $p_t(n) = \ln P_t(n)$ denotes the log bond price.

(d) Construct the CP factor as discussed in the notes or CP's original paper using the first 5 forward rates to predict returns, but using the average excess returns of 2- to 10-year bonds. Note that we use annual returns, but monthly overlapping data.

(e) Regress (contemporaneously) each of the forward rates on the CP factor (including an intercept) and collect the residuals

$$f_t(n) = c_n + d_n CP_t + f_t^{\perp}(n).$$

(f) Construct the first two principal components based on the panel of $f_t^{\perp}(n)$. Call these factors $PC_t = (PC_{1t}, PC_{2t})'$, where PC_{it} corresponds to the $i-$th principal component.

(g) Remove the sample mean from CP_t and the principal components, PC_t.

(h) Report

  i. The R-squared of the CP regression.

  ii. Report the R-squared separately when estimating the model using data until 1999.12 and thereafter. Comment on the differences.

2. We now have 3 pricing factors, $X_t = (CP_t, PC_t')'$. We can write an affine model as

$$\begin{aligned}
X_{t+1} &= \Gamma X_t + \epsilon_{t+1}, \epsilon_t \sim N(0, \Sigma), \\
y_t(1) &= \delta_0 + \delta_1 X_t, \\
\lambda_t &= \lambda_0 + \lambda_1 X_t, \\
m_{t+1} &= -y_t(1) - \frac{1}{2}\lambda_t' \Sigma \lambda_t - \lambda_t' \epsilon_{t+1},
\end{aligned}$$

where $\lambda_0 \in R^{3\times 1}$ and $\lambda_1 \in R^{3\times 3}$. Note that X_t has mean zero as we de-meaned the factors.

The parameters we need to estimate are $(\Gamma, \Sigma, \delta_0, \delta_1, \lambda_0, \lambda_1)$.

Without further restrictions, we have to estimate 12 risk prices (3 in λ_0 and 9 in λ_1)! Let us use the facts we know and establish new facts to simplify the model (a lot).

First, we know that a single factor, the CP factor, explains variation in bond risk premia. Variation in bond risk premia is captured by λ_1. Therefore, λ_1 should only depend on CP_t, which means that we start from

$$\lambda_1 = \begin{bmatrix} \lambda_{1(1,1)} & 0 & 0 \\ \lambda_{1(2,1)} & 0 & 0 \\ \lambda_{1(3,1)} & 0 & 0 \end{bmatrix}.$$

We now have “only” 6 risk prices left. In the remainder of the problem set we estimate the parameters of the term structure model and simplify the risk prices even further.

3. Estimate δ_0 and δ_1 by regressing the short rate on the factors,

$$y_t(1) = \delta_0 + \delta_1' X_t + u_t.$$

Report the estimates for δ_0 and δ_1 as well as the R-squared of the regression. What should the R-squared be if the affine model is correct? Comment on the results.

4. Next, we estimate the VAR,

$$X_{t+1} = \Gamma X_t + \ \epsilon_{t+1},$$

where we omit the intercept as we demeaned the factors. Note that we lag the factors by 12 months and we again use monthly overlapping data. Report the estimates for Γ and Σ, the covariance matrix of the residuals.

5. We now have 3 potentially priced shocks, ϵ_t. Let's use two-pass regressions to figure out which shock is most important in pricing the cross-section of bond returns. Use the following steps for *for each factor separately*

Step 1: Regress excess returns, $rx_t(n)$, for each maturity on ϵ_{it}. Collect the betas.

Step 2: Run a cross-sectional regressions of $E_T(rx_t(n))+\frac{1}{2}Var_T(rx_{t+1}(n))$ on the betas. Note that we get the Jensen term because we are using geometric (=log) returns.

Report a table with

(a) Rows: The bonds' maturities.

(b) Columns: The alpha for each of the 3 pricing models.

(c) The final row should report the mean absolute pricing error for each model.

The level factor should stand out as pricing the cross-section best.

It suggests another major simplification of our model. *Shocks to the level factor are priced and the variation in price of risk are driven by the CP factor.*

This implies a pricing model of the form

$$\lambda_t = \begin{bmatrix} 0 \\ \lambda_{0(2)} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ \lambda_{1(2,1)} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} X_t.$$

So we are down from 12 parameters to just 2! To complete the estimation, we need to estimate $\lambda_{0(2)}$ and $\lambda_{1(2,1)}$.

6. The affine model implies

$$p_t(n) = A(n) + B(n)'X_t.$$

Derive expressions for $A(n)$ and $B(n)$ in terms of $(\delta_0, \delta_1, \mu^\star, \Gamma^\star, \Sigma)$, where

$$\begin{aligned} \Gamma^\star &= \Gamma - \Sigma\lambda_1, \\ \mu^\star &= -\Sigma\lambda_0. \end{aligned}$$

$\Gamma^\star$ and $\mu^\star$ are the $\mathbb{Q}$-parameters of the VAR model.

Note that given the simplification in the risk prices before,

$$-\Sigma\lambda_0 = -\Sigma_{(:,2)}\lambda_{0(2)},$$

where $\Sigma_{(:,2)}$ denotes the second column of Σ.

7. In the affine model, we have

$$E_t(rx_{t+1}(n)) + \frac{1}{2}B(n-1)'\Sigma B(n-1) = B(n-1)'\Sigma\lambda_0 + B(n-1)'\Sigma\lambda_1 X_t.$$

The second term on left-hand side is the standard Jensen correction when we use geometric returns, $\frac{1}{2}Var_t(rx_{t+1}(n))$.

$B(n)$ only depends on λ_1. Use the 10-year bond and estimate using annual returns and monthly overlapping data

$$rx_{t+1}(10) = a + \theta CP_t + u_{t+1}.$$

Hence, CP_t is lagged by 12 months and $rx_{t+1}(10)$ is the annual return on a 10-year bond. Report θ, the predictive coefficient.

Use $\lambda_{1(2,1)}$ to match the predictive coefficient.

8. Lastly, the unconditional risk premium is given by

$$E(rx_{t+1}(n)) + \frac{1}{2}B(n-1)'\Sigma B(n-1) = B(n-1)'\Sigma\lambda_0.$$

   Choose $\lambda_{0(2)}$ to match the risk premium on the 10-year bond and report the estimate.

Done!