Notes on the Yield Curve

April, 2016

Q-Group Conference

Washington, DC

Ian Martin, LSE

Steve Ross, MIT

The Term Structure

- The term structure has fascinated because bond prices embed market forecasts
- Specifically, we use futures as forecasts of future spot rates
- Historical theories of the risk premium, e.g., expectations hypotheses, preferred habitat, risk models, general equilibrium
- Modern derivatives based theories, CIR, HJM – continuous time
- This talk is based on a simpler model of the term structure anayzed in a paper by Ian Martin and myself (Notes on the Yield Curve (2016))
- The model was introduced in the Recovery Theorem [Ross (2015)] and also in a working paper, The Long Bond, by Ian Martin

The Model Framework

- The framework is a straightforward discrete time, discrete state model

- The states evolve by an irreducible Markoff process with a one period transition probability matrix:

$$\Pi = [\pi_{ij}],$$

  where π_{ij} = prob of a transition from state i to state j; i,j = 1,…,m, and note that

$$\Pi e = e, \ e \equiv \langle 1,\ldots,1 \rangle$$

- The market is assumed to be complete for AD securities and the AD prices are given by the AD tableau:

$$A = [a_{ij}],$$

  where a_{ij} = the current state i price of $1 paid if state j occurs next period

- Notice that we are not differentiating states by path dependence – nevertheless summary statistics such as wealth could be surrogates for the path and part of the state

The Model Framework

- Following Ross (2015) we will assume for the moment that we can decompose A as

$$A = \delta D \Pi D^{-1}$$

where δ is a discount factor and D is a diagonal matrix with positive entries.

- As motivation for this decomposition, with a representative agent with utility function u(i) in state i and a discount factor of δ, the first order conditions at an optimum are

$$u'(i) a_{ij} = \delta \pi_{ij} u'(j), \quad D = diagonal\ with\ elements\ [1/u'(i)]$$

- Result 1: For an arbitrary A the above decomposition exists and is unique

- Proof: From the Perron Frobenius theorem there is a unique eigenvector, v, and a real dominant eigenvalue, φ, such that Av = φv. Letting D be diagonal with the elements of v on the diagonal, then

$$\Pi = (1/\varphi) D^{-1} AD$$

is a stochastic matrix. From the uniqueness of v and φ it follows that D is unique

□

The Maintained Hypothesis

- As shown in detail in Ross [2015] this means that we can recover a probability distribution from observed state prices

- Hypothesis 1:

    - The probability distribution recovered in the above fashion is the true objective or natural probability distribution

- It is worth noting that this is a major point of disagreement between us and Hansen, et.al. [2013]. They choose to generalize the dynamic pricing kernel with a multiplicative martingale process

- This is, of course, an empirical issue and while any assumptions can be generalized that doesn’t make it either ‘correct’ to do so or useful – it depends on the data

The Yield Curve

- If the continuously compounded t period yield is given by $y_t(i)$, then the bond price at time t is

$$B_t(i) = e^{-y_t(i)t}$$

- By the law of iterated expectations

$$B_t(i) = \sum_j A^t(i,j)$$

- Hence,

$$y_t(i) = -(1/t)\log \sum_j A^t(i,j)$$

- And the conditional short rate is

$$r_f(i) = y_1(i) = -\log \sum_j A(i,j)$$

The Yield Curve

- An important construct is the long bond which is the above as $t \to \infty$

$$y_{\infty}(i) = -\lim_{t\to\infty} (1/t) \log \sum_{j} A^{t}(i,j)$$

- The return on a t period bond is given by

$$R_{t}(i,j) = B_{t-1}(j) / B_{t}(i)$$

and using lower case for logs,

$$r_{t}(i,j) = \log R_{t}(i,j)$$

and

$$r_{\infty}(i,j) = \lim_{t\to\infty} r_{t}(i,j) \;, \quad R_{\infty}(i,j) = \lim_{t\to\infty} R_{t}(i,j)$$

The Yield Curve

- Conditional expected bond returns are given by:

$$R_t(i) = \sum_j \pi(i,j) R_t(i,j)$$

and, for example, the unconditional expected log return on the long bond is given by

$$r_\infty(i) = \sum_{i,j} \pi(i) r_\infty(i,j)$$

where π_i is the stationary distribution for the Markoff process as given by the left eigenvector of Π, $\pi\Pi = \pi$.

- Recalling that $Av = \varphi v$, for any asset with return $R(i,j)$,

$$E\left(\varphi \frac{v(j)}{v(i)}\right) R(i,j)\, i = \sum_j \pi(i,j) \left(\varphi \frac{v(j)}{v(i)}\right) R(i,j) = \sum_j a_{ij} R(i,j) = 1,$$

verifying that $\varphi \frac{v(j)}{v(i)}$ is the stochastic discount factor and, equivalently, that $v(j)$ is the pricing kernel with a discount factor of φ

The Long Bond as the SDF

- Lemma: $A^T/\varphi^T \to L \equiv [v w'] = [v(i)w(j)]$
- Result 2: The return on the long bond is proportional to the SDF*
- Proof:

$$R_\infty(i,j) = \lim_{T\to\infty} \frac{\sum_k A^{T-1}(j,k)}{\sum_k A^T(i,k)} = (1/\varphi) \lim_{T\to\infty} \frac{\sum_k A^{T-1}(j,k)/\varphi^{T-1}}{\sum_k A^T(i,k)/\varphi^T} = (1/\varphi) \frac{v(j)\sum_k w(k)}{v(i)\sum_k w(k)} = 1/\varphi \; v(j)/v(i)$$

□

*This was first shown by Kazemi [1992] in a continuous time diffusion model

To put it another way, the SDF is equivalent to the unobservable theoretical construct of the long bond

The Long Bond as the SDF

- Now we link the eigenvalue φ to the yield curve
- Result 3: The long rate and the expected log return,

$$y_{\infty}(i) = r_{\infty} = -\log\varphi$$

- Proof: Note first,

$$y_{\infty}(i) = -\lim_{T\to\infty} \frac{1}{T}\log\sum_{j} A^{T}(i,j) = -\log\varphi - \lim_{T\to\infty}\frac{1}{T}\log\sum_{j}\left(A/\varphi\right)^{T}(i,j) = -\log\varphi$$

and

$$r_{\infty}(i,j) = -\log\varphi + \log\left(v(j)/v(i)\right)$$

hence

$$r_{\infty} = -\log\varphi + \sum_{i,j}\pi(i)\pi(i,j)\log v(j)/v(i) = -\log\varphi + \sum_{j}\pi(j)\log v(j) - \sum_{i}\pi(i)\log v(i) = -\log\varphi$$

□

The Long Bond as the SDF

- Notice that it might seem paradoxical to have the long yield approach a constant while the returns are equivalent to the variable SDF

- The resolution of this seeming paradox is that the return on a T-period zero coupon bond is

$$r_T(i,j) = y_T(i) - (T-1)[y_{T-1}(j) - y_T(j)]$$

hence while the conditional yield in state j approaches the current yield in state i, the duration grows with T and offsets the convergence in yields producing a variable return

- Summarizing what we've found, the long bond reveals the pricing kernel and the long yield is the log of the eigenvalue which is the time discount factor of the pseudo representative agent

Speed of Convergence

- Define
  $Q \equiv log[\max_{k} v(k) / \min_{k} v(k)] = log \max_{k} v(k) - log \min_{k} v(k) > 0$
- Hence,
  $$y_{\infty} - Q \leq r_{\infty}(i,j) = y_{\infty} + log v(j) - log v(i) \leq y_{\infty} + Q$$

- Now we link the eigenvalue φ to the yield curve

- Result 4: The difference between the T-period yield and the yield on the long bond is uniformly bounded across the states,

$$|y_{T}(i) - y_{\infty}| \leq Q/T$$

- from which it follows that the difference in yields of any two bonds of maturities T_1 and T_2 is also bounded,

$$|y_{T_1} - y_{T_2}| \leq Q[1/T_1 + 1/T_2]$$

Speed of Convergence

- Proof: Since for all i,

$$B_T(i) = \sum_j A^T(i,j) = \varphi^T v(i) \sum_j (1/v(j))\, \pi_{ij}$$

we have

$$\varphi^T e^{-Q} \le B_T(i) \le \varphi^T e^{Q}$$

hence

$$y_\infty - Q/T \le y_T(i) \le y_\infty + Q/T$$

which verifies both results.

□

Speed of Convergence

- The next result provides a bound on the return of a sufficiently long dated assets with payoffs at the horizon (see Hansen and Scheinkman [2009] for a similar result)

- Result 5: The return on a long dated asset paying x(j) in state j at time T satisfies

$$R_T(i,j) = R_\infty(i,j) + O(\epsilon^T)$$

where ϵ satisfies $\psi/\varphi < \epsilon < 1$, ψ is the absolute value of the second largest eigenvalue of A, and ϵ can be chosen to be arbitrarily close to ψ/φ

- Proof: Recalling the definition of L,

$$\max_{i,j} |A^T(i,j) - L(i,j)| = \max_{i,j} |A^T(i,j) - v(i)w(j)| \leq C\delta^T$$

where C is a constant and $\psi < \delta < \varphi$ and can be chosen arbitrarily close to ψ

The value of the asset in state i:

$$p_T(i) = A^T x = \varphi^T (Lx)_i + O(\delta^T) = v(i)K\varphi^T + O(\delta^T),$$

$$K \equiv \sum_j w(j)x(j)$$

Speed of Convergence

It follows that the asset's return

$$R(p_T)(i,j) = \frac{p_{T-1}(j)}{p_T(i)} = \frac{K v(i) \varphi^{T-1} + O(\delta^{T-1})}{K v(i) \varphi^{T} + O(\delta^{T})}$$

$$\approx \frac{v(j)}{\varphi v(i)} = R_\infty(i,j)$$

□

- The decomposition in Result 5 provides a neat interpretation of pricing as the product of a time discount factor, φ^T, the economy wide kernel, v(i), which captures equilibrium risk aversion, and K which captures asset specific information.

- Notice, too, that

$$v(i)K = v(i) \sum_j w(j) x(j) = \sum_j \pi(j) \frac{v(i)}{v(j)} x(j),$$

the undiscounted, stationary value of the payoff

Traps

- Another way to analyze the speed of convergence is to think about for how long rates could remain or be 'trapped' or stuck in an abnormally high or low regime

- Roughly, we define a trap as a set of states, S, from which it is difficult to exit

- Definition 1: The market has an ϵ-trap if there is a set of states S , such that

$$Prob\,exit\ S\,in\ S \leq \epsilon\,and\ Prob(enter\ S \mid outside\ S) \leq \epsilon$$

- The trap is stronger the closer ϵ can be set to zero

- The next result directly links the strength of the trap, ϵ, to observables

Traps

- Result 6: Let x be any random function of the states, and define the variance of changes in x and the variance of the level of x, respectively as

$$\sigma^2(\Delta x) \equiv \sum_j \pi(i)\pi(i,j)[x(i)-x(j)]^2 \quad and \quad \sigma^2(x) \equiv \sum_i \pi(i)x(i)^2 - \left(\sum_i \pi(i)x(i)\right)^2$$

then the market has an ϵ-trap with

$$\epsilon = \sigma(\Delta x)/\sigma(x)$$

Proof: Uses the Cheeger inequality for directed graphs see Martin and Ross [2015]

□

- This is a very powerful and general result - for example, suppose x is the 10 year constant maturity Treasury yield, y. From April, 1953 to May, 2015 the monthly $\sigma(y)$ = 2.61% and $\sigma(\Delta(y))$=0.25%, then ϵ = 0.096. More dramatically, if we set x equal to a moving average of yields, we can drive ϵ down to 0.033 and the expected time spent in the trap is then 30 months and the probability of spending more than 5 years in a trap is more than 13%.

- But, it isn't obviously constructive – what is the trapped set?

Valuing Long Dated Assets

- Result 7: The long bond is growth optimal, i.e., in every state the long bond has the highest expected log return
- Proof: Since $\varphi v(j)/v(i)$ is the stochastic discount factor , Result 2 verifies that the reciprocal of the return on the long bond

$$1/R_{\infty}(i,j) = \varphi v(j)/v(i)$$

is the SDF, hence for any asset return, R

$$E[R(i,j)/R_{\infty}(i,j)] = 1 \text{ ,}$$

which implies that

$$E[\log R(i,j)] - E[\log R_{\infty}(i,j)] = E[\log(R(i,j)]/R_{\infty}(i,j))] \leq \log(E(R(i,j)/R_{\infty}(i,j))) = \log 1 = 0$$

□

Valuing Long Dated Assets

- Notice that in a complete market the SDF is unique, so the inverse of the long bond return is the unique SDF

- Another approach is to observe that an asset's price is the expected discounted value of its payoffs discounted at the return on the long bond, hence the cheapest way to achieve a long run payoff is an investment in the long bond

- Alternative to the long bond, other long dated assets can replicate the pricing kernel

- Let p be the price of an asset with a constant payoff, x:

$$p = Ax + A^2 x + \cdots = A^* x ,$$

where

$$A^* = A + A^2 + \cdots = A[I + A + A^2 + \cdots] = A[I - A]^{-1}$$

converges since $\varphi < 1$.

- A* inherits the same dominant eigenvector, v, as A and its associated maximum eigenvalue is $\varphi/(1-\varphi)$

Valuing Long Dated Assets

- Result 8: The perpetual v-asset which pays v(i) in every period is the unique, infinitely lived limited liability asset with a constant dividend yield. Its yield is (1/φ) -1 and its returns replicate the returns on the long bond. Additionally no asset can have a uniformly higher or lower dividend yield.

- Proof: Denote the price of the perpetual v-asset as p_v and its dividend yield is

$$v/p_v = v/A^* v = 1-\varphi/\varphi$$

and its returns are

$$R_{ij}^v = p_v(j)+v(j)/p_v(i) = (\varphi/1-\varphi)v(j)+v(i)/(\varphi/1-\varphi)v(i) = (1/\varphi)(v(j)/v(i))$$

which is the return on the long bond. Uniqueness follows from uniqueness of the positive eigenvector. From Theorem 8.1.26 of Horn and Johnson [1990],the return on any perpetual asset with dividends of x is $(A_i^* x)/x_j$ which is not uniformly greater or less than the maximum eigenvalue of A*

□

Some General Properties of the Yield Curve

- Result 9: On average, i.e., unconditionally, the forward curve is below the long run yield

- Proof: By definition forward rates satisfy

$$f_1 + \cdots + f_T = -\log B_T$$

hence

$$f_T = -\log B_T + \log B_{T-1}$$

Notice that

$$f_1(i) = r_f(i) = y_1(i) = -\log \sum_j A(i,j)$$

and the T period forward rate,

$$f_T = -\log B_T + \log B_{T-1} = -\log[\varphi^T D \Pi^T D^{-1} e + \log[\varphi^{T-1} D \Pi^{T-1} D^{-1} e = -\log \varphi + \log[\Pi^{T-1} s] - \log[\Pi^T s]$$

where

$$s = D^{-1} e$$

Some General Properties of the Yield Curve

Applying Jensen's inequality row by row of the Markoff matrix,

$$E[f_T] = -\log \varphi + E[\log[\Pi^{T-1} s]] - E[\log[\Pi^{T} s]] \;=\; -\log \varphi + E[\log[\Pi^{T-1} s]] - E[\log[\Pi \Pi^{T-1} s]]$$

$$= -\log \varphi + E[\log[\Pi^{T-1} s]] - E[\log[E \Pi^{T-1} s_i]]$$

$$< -\log \varphi + E[\log[\Pi^{T-1} s]] - E[E[\log[\Pi^{T-1} s_i]]]$$

$$= -\log \varphi + E[\log[\Pi^{T-1} s]] - E[\log[\Pi^{T-1} s]] = -\log \varphi$$

□

- This verifies that, unconditionally, the forward curve approaches the long run yield from below – of course conditionally anything is possible

- Result 10: On average – unconditionally - the yield curve is below the long run yield

- Proof: Follows immediately from Result 9 since

$$y_T = -(1/T) \log B_T = (1/T) \sum_{1}^{T} f_T$$

□

Some General Properties of the Yield Curve

- Is the convergence of the yield to the long run yield as maturity increases monotone?
- This is a delicate question and the answer is that it is not generally the case
- The answer turns on the interplay amongst the eigenvectors and the eigenvalues below the dominant one
- Generally there could be m-1 cycles with local peaks and troughs

Conclusion

- There is a lot more to do and to be learned
- How does this model fit with existing models?
- For example, the CIR model is also stationary and Markoff but the long run yield in that model doesn't dominate the yield curve and, unconditionally, it can be humped
- Within the model itself there is an enormous amount to do:
  - We would like a simple approximate expression for the forward risk premium
  - We need to better understand how fast is convergence
- Critically for some current debates, how close are T period bonds to the unobservable long bond? Is the 30 year yield a good surrogate?
- No!