Notes on the Yield Curve

February 8, 2016

Introduction

The literature on the term structure of interest rates is one of the oldest and earliest analytic areas of economics and finance. No doubt one of the attractions of the area has been the potential to use forward rates as predictors of future interest rates. But, it has long been known that while forward rates are informative, they also embody risk-premiums and are not unbiased predictors.

It has also long been accepted that fixed income yields are stationary. Much effort has been expended at the statistical examination of this issue, but, as Jon Ingersoll once remarked during a seminar twenty five years ago in which the speaker was employing a turgid and labored spectral analysis to show that interest rates were non stationary, 'Interest rates were about 5% 5000 years ago in Babylonian times and they're still about 5% today – seems pretty stationary to me.'[1] In keeping with this observation, this paper builds a simple stationary model of the fixed income market and uses it both to provide simplified proofs of some old results and to obtain some new ones. With complete markets, we are able to use the Recovery Theorem (Ross (2015)) to derive the unobserved probability distribution of fixed income returns from the observed pricing of bonds. We also verify the result that the stochastic discount factor – also called the pricing kernel - is the return on the long bond, and we are able to analyze the speed of convergence of observed bond pricing to that of the unobserved long bond. In addition, we obtain a number of new results on the general character of yield curves. This is accomplished in a discrete time, non-parametric setting. A distinctive and appealing feature of the framework is that it allows for the possibility of cycles and traps, both of which blur the distinction between stationary and non-stationary models. This stands in contrast to much of the modern theory of the term structure which is typically couched in a continuous time setting and often within a general equilibrium setting (see, for example, the CIR model, Cox, Ingersoll and Ross [1985]). We will briefly consider the differences between these approaches and ours.

Section I presents our framework. The basic model is one where states of nature evolve by a finite state Markov chain. This is about as simple a model as one could use that still allows for the rich variety of phenomena we observe in the term structure of interest rates. It is the same model used in the Recovery Theorem so, not surprisingly, we are be able to show that from state space pricing we can derive the natural motion of interest rates. Section II proves a variety of results on the term structure and, in particular, it highlights the relation between the pricing kernel, the eigenvector of the transition matrix, and the 'long' bond, i.e., the limiting behavior of the discount bonds as the maturity goes to infinity. Section III examines the speed of this convergence of yields and forward rates to the long bond and introduces the notion of a trap which captures the intuition that the economy may stay in a particular state or a set of states for long periods of time. Section IV shows that the long bond is the growth optimal asset and also

[1] John Cochrane and others have made the same point.

briefly extends the analysis to long-dated coupon-paying assets.   Section V derives the result that in the framework of the model, the asymptotic yield, i.e., the yield on the long bond, unconditionally dominates the yield on all other bonds. It also shows that some conjectured results on the monotonicity of the convergence of yields as the maturity increases will not be correct. Section VI concludes the paper and briefly highlights some differences between our model and some of the canonical models in the literature.

Section I. The Framework

We assume that there are m states of nature and that the physical economy transitions from one state to another according to an irreducible Markoff transition process, where

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

denotes the matrix of transition probabilities. Letting e be the vector of ones we have

$$\Pi e = e .$$

The matrix of Arrow-Debreu prices for a security paying one unit in state j if the economy is currently in state i are given by

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

Following Ross (2015) assume for the moment that the pricing matrix can be decomposed as

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

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

To motivate these assumptions, in a setting with assets priced by a representative expected utility maximizing agent, with utility function u(i) in state i, and a time discount factor, δ, the first order conditions for an optimum are

$$u'(i)a_{ij} = \delta\pi_{ij}u'(j).$$

Equation (1) is the matrix version of this equation with the elements of D as the inverses, v(i) = 1/u′(i), of the marginal utilities,

$$D = \begin{bmatrix} v(1) & 0 & 0 \\ 0 & v(i) & 0 \\ 0 & 0 & v(m) \end{bmatrix},$$

While this example illustrates how the decomposition of (1) will arise, as the next result shows, such a decomposition always exists and is unique whether or not there is a utility maximizing representative agent.

Result 1. For an arbitrary irreducible (or more strongly, strictly positive) matrix A of Arrow-Debreu prices, there exists a unique decomposition of the form of (1) where D is determined up to a scalar multiple.

Proof: The Perron-Frobenius theorem states that A has a unique (up to scale) positive eigenvector, v, and real eigenvalue, φ, satisfying Av = φv, where φ and v are strictly positive and where φ is the dominant eigenvalue. Letting D be the diagonal matrix with v on the diagonal, and setting v = De,

$$\Pi \equiv \left(\frac{1}{\varphi}\right) D^{-1} A D,$$

is a stochastic matrix, i.e., it is non-negative and Πe = e. Let C be a different positive diagonal decomposition of A as in (1). It follows that Ce is an eigenvector of A and since A has a unique positive eigenvector, it follows that Ce is proportional to De and, therefore, up to a scale, C = D with φ as the unique eigenvalue. It is immediate that Π is also unique.

□

Result 1 is a generalization of the Ross Recovery Theorem (Ross [2015]}; if Π is the true probability transition matrix and the stochastic pricing kernel is (in the terminology of Ross [2015]) transition independent then Result 1 assures that Π can be uniquely recovered from A. But the decomposition is well defined and unique whether or not there is a representative agent and there is no need to assume that there is a representative agent. Rather, this result can be thought of as defining the marginal utilities and the time discount factor of a pseudo-representative agent. We will take the following hypothesis as a given.

Hypothesis 1. The probability distribution recovered in the above fashion is the true objective or natural probability distribution.

Whether this hypothesis holds is fundamentally an empirical question and is not an issue to be determined on a priori theoretical grounds; it can only be resolved by comparing the recovered distribution, Π, with observed realizations in the market. In the remainder of the paper we will assume that Hypothesis 1 holds. This is similar to the expectations hypothesis which asserts that pricing is risk neutral. This has provided a useful way to think about interest rates and related phenomena and is the basis for a substantial literature. It is worth noting that Hansen, et. al. [2013], take issue with this point of view and argue that Hypothesis 1 does not hold, but we will briefly allude to some disagreements with their assessment of the current state of empirical work on this matter.

Section II. The Yield Curve

Let $B_t(i)$ denote the current – time 0 – price of a pure discount (zero coupon) bond paying 1 in t periods if the current state is i. The continuously compounded t period yield, y_t (i) and the simple yield Y_t (i) are given by

$$B_t(i) = e^{-y_t(i)t} = \frac{1}{[1 + Y_t(i)]^t}.$$

By the law of iterated expectations, we have that

$$B_t(i) = \sum_j A^t(i,j) , \qquad (2)$$

In other words, the price of a t period bond is the sum of the elements in the i^{th} row of A^t. Another way to see this is to observe that element (i,j) of A^t is the current state i price of a pure contingent security that pays 1 in t periods if the state is then j.

It follows that

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

and the conditional short rate is

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

We will define the conditional long rate as the limiting yield as the horizon goes to infinity,

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

The realized return on a t period bond from a transition from state i to state j in a single period is

$$R_t(i,j) \equiv \frac{B_{t-1}(j)}{B_t(i)} ,$$

Using lower case letters to indicate logs,

$$r_t(i,j) = \log R_t(i,j)$$

$$and\ r_\infty(i,j) \equiv \lim_{t\to\infty} r_t(i,j), \quad R_\infty(i) \equiv \lim_{t\to\infty} R_t(i,j) .$$

We obtain unconditional expected returns by summing over the transition probabilities, for example,

$$\bar{R}_\infty(i) = \sum_j \pi(i,j) R_\infty(i,j)\ ,$$

where the stationary distribution for the Markoff process, π, is given by the left eigenvector of Π,

$$\pi' \Pi = \pi'.$$

For example, the unconditional expected return on the long bond is given by,

$$\bar{r}_\infty = \sum_{i,j} \pi(i)\pi(i,j) r_\infty(i,j).$$

Now we'll take a closer look at the vector, v, of Result 1. Letting R(i,j) be the return on an arbitrary asset, the conditional expectation of R(i,j) weighted by v(j)/v(i) gives

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

verifying that $\phi v(j)/v(i)$ is the stochastic discount factor in the current state i, and, equivalently, that v(j) is the pricing kernel with a discount factor of φ.

The next result links the vector v to the long end of the yield curve. Let L be the matrix whose (i,j)-th element is v(i)w(j) where w' is the left eigenvector of A and where w and v are normalized so that w'v = 1. Notice that row i of L is v(i)w'. From (1) it is easy to see that the left eigenvector of A, w, such that w'A = φw', is $w = D^{-1}\pi$ with entries, $\pi_i/v(i)$. An important limiting result (see, for example, Theorem 8.5.1 of Horn and Johnson [1990]) as $t \to \infty$ is that

$$\frac{A^t}{\varphi^t} \to L \ \equiv [vw'] = [v(i)w(j)] \quad ,$$

where w is normalized so that w'v = 1.

Result 2. Let the v asset be an asset that pays off v(j) next period. Returns on the v-asset are identical to the returns on the long bond.

Proof: The price of the v asset is given by

$$P_v(i) = \sum_j A(i,j) v(j) = \varphi v(i)\ ,$$

So the return $R_v(i,j) = (1/\phi)v(j)/v(i)$. The return on the long bond is

$$
\begin{aligned}
R_{\infty}(i,j) &= \lim_{t\to\infty}\frac{\sum_k A^{t-1}(j,k)}{\sum_k A^t(i,k)} \\
&= \frac{1}{\varphi}\lim_{t\to\infty}\frac{\sum_k A^{t-1}(j,k)\,/\varphi^{t-1}}{\sum_k A^t(i,k)\,/\varphi^t} = \frac{1}{\varphi}\frac{v(j)\sum_k w(k)}{v(i)\sum_k w(k} = \frac{1}{\varphi}\frac{v(j)}{v(i)}\ . \qquad (3)
\end{aligned}
$$

□

Empirically, then, if we could observe the return on the long bond, Result 2 would allow us to observe the kernel as well.[2] Below we will see whether it is possible to approximate this return from the returns on sufficiently long dated bonds. We now link the eigenvalue, φ, to the long end of the yield curve.

Result 3. The long rate and the expected log return on the long bond both equal -log ϕ independent of the current state i.

$$
y_{\infty}(i) = \bar{r}_{\infty} = -\log\varphi\ .
$$

Proof: Note first that

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

where the summation goes to a constant matrix as shown in the proof of Result 2.

From (3) we have

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

hence,

$$
\begin{aligned}
\bar{r}_{\infty} = -\log\varphi + \sum_{i,j}\pi(i)\pi(i,j)\log\frac{v(j)}{v(i)} &= -\log\varphi + \sum_j \pi(j)\log v(j) - \sum_i \pi(i)\log v(i) \\
= -\log\varphi\ .
\end{aligned}
$$

□

It might at first seem contradictory to have the long rate converge to a constant while the returns on the long bonds are equivalent to the variable pricing kernel. Note, though, that the return on a T-period bond is

[2] Kazemi [1992] was the first to observe the connection between returns on the long bond and the pricing kernel in a continuous-time equilibrium diffusion model. Some other interpretations in Sections III and IV link the eigenvector, v, to the return on any long dated asset and to the unique infinitely-lived asset with a constant dividend yield.

$$r_t(\mathrm{i,j}) = y_t(i) - (t-1)[y_{t-1}(j) - y_t(i)] \ .$$

Hence, while Result 3 shows that the conditional yield in state j approaches the current yield in state i, the duration, t, grows as well and when the two effects are multiplied the result is that long-dated bonds have volatile returns.

Summarizing the findings of this section, returns on the long bond reveal the eigenvector, v, i.e., the pricing kernel, and the long yield is the log of the eigenvalue which is the time discount factor of the pseudo-representative agent.

Section III. Speed of Convergence

The results of Section II verified that the long end of the yield curve appropriately defined converges to the unknown pricing kernel, v(i). In this section we will explore the speed of this convergence. This is crucial for determining the empirical strength of these findings. For example, we would like to know how big t has to be to assure us that the yield on a t-period bond, $y_t(I)$, is within some $\varepsilon > 0$ of the infinite yield, - log ϕ. To address this issue we introduce the metric

$$Q \ \equiv \log\left[\frac{max_k v(k)}{min_k v(k)}\right] = \log max_k v(k) - \log min_k v(k) > 0 \ .$$

We can think of Q as measuring the extent to which risk aversion matters for pricing. If pricing is risk neutral then the pricing kernel, v = e and Q = 0. Put another way, if Q = 0 then e is an eigenvector of A which implies that the row sums of A equal the same discount factor, ϕ, and interest rates are constant. This means that the yield curve is flat so that the pricing of (fixed income) securities is risk neutral and risk neutral and natural probabilities coincide. From Result 2 we can see that Q is also a measure of the dispersion of long bond returns.

From equation (3) we have that

$$\min_{i,j} r_\infty(i,j) = \ y_\infty - Q \ and \ \max_{i,j} r_\infty(i,j) = \ y_\infty + Q \ .$$

This allows us to prove the following convergence result.

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

$$|y_t(i) - \ y_\infty| \ \le \frac{Q}{t}.$$

From this it follows that the difference in the yields of any two bonds of maturities t_1 and t_2 are also bounded across the states,

$$\left|y_{t_1}(i) - y_{t_2}(i)\right| \le Q[\frac{1}{t_1} + \frac{1}{t_2}] .$$

Proof: Since for all i

$$B_t(i) = \sum_j A^t(i,j) = \varphi^t v_i \sum_j (\frac{1}{v_j})\pi_{ij} ,$$

we have

$$\varphi^t e^{-Q} \le B_t(i) \le \varphi^t e^{Q} \ \ for\ all\ i,$$

Hence,

$$y_\infty - \frac{Q}{t} \le y_t(i) \le y_\infty + \frac{Q}{t} ,$$

which verifies both results.

□

The next result provides a similar bound for the return on long dated discount bonds and, indeed, on all sufficiently long dated assets with payoffs at the horizon. This result is similar to that of Hansen and Scheinkman [2009].

Result 5. The return on a long dated asset paying x(j) in state j at time T (and zero for t < T) approaches the return on the long bond as T → ∞. More precisely,

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

where $\psi/\phi < \varepsilon < 1$, ψ is the absolute value of the second largest eigenvalue of A, and ε can be chosen arbitrarily close to ψ/ϕ. Thus the ratio of the next largest eigenvalue to the largest, ϕ, determines the rate at which the return on a T period bond is close to the return on the long bond.

Proof: Recalling the definition of the matrix L used in Result 2, the largest difference between the elements of L and A^T,

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

where C is a constant and δ satisfies $\psi < \delta < \phi$ and can be chosen to be arbitrarily close to ψ (see Horn and Johnson, Theorem 8.5.1). The value of the asset in state i, $p_T(i)$, is given by:

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

where

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

It follows that the return on the asset,

$$R(p_T)(i,j) = \frac{p_{T-1}(j)}{p_T(i)} = \frac{Kv(j)\varphi^{T-1} + O(\delta^{T-1})}{Kv(i)\varphi^T + O(\delta^T)} \approx \frac{v(j)}{\varphi v(i)} = R_\infty(i,j).$$

□

The decomposition in the proof of Result 5 allows us to interpret pricing as the product of a time discount factor, ϕ^T, the economy wide kernel, v(i), which captures risk aversion, and a term specific to the asset, K, plus other effects which are of negligible order. Notice that

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

is the expectation of the risk adjusted undiscounted value of the payoff where the expectation is, appropriately in the limit, taken with respect to the stationary distribution, π. Another way to think about the speed of convergence is in terms of how long rates could remain, or are 'trapped', e.g., stuck in an abnormally high or low regime. This is particularly interesting given the interest in models that features probabilistic shifts of regime. We will define a trap as a set of states that is difficult to exit from once entered. Formally, for any set S of the states, let

$$P(in\ S) = \textstyle\sum_{i\in S}\pi(i)\ \ and\ P(outside\ S) = \sum_{i\notin S}\pi(i).$$

Since π is the stationary distribution, these are the fraction of time the economy spends in or outside of S respectively. Additionally, we will be interested in the fraction of time spent entering S or exiting from S once it is in S,

$$P(exit\ S) = \textstyle\sum_{i\in S, j\notin S}\pi(i)\pi(i,j)\ \ and\ P(enter\ S) = \sum_{i\notin S, j\in S}\pi(i)\pi(i,j).$$

Notice that these are the stationary probabilities that the economy exits or enters S in the next time period.

Lastly, we will define the conditional probabilities of exiting S given that we start in S and of entering S given that we start outside of S,

$$P(exit\ S|start\ in\ S) = \frac{P(exit\ S)}{P(in\ S)}\ and\ P(enter\ S|start\ outside\ S) = \frac{P(enter\ S)}{P(outside\ S)}\,.$$

We could just define a trap such that the probability of exit is appropriately small, but that could occur trivially for one of two reasons. First, if S is a relatively small the economy will rarely be in S. We exclude this case by requiring that the probability of exiting S is small, conditional on starting in S. Second, the probability of exiting S could be small simply because S is very large.

To exclude this case we will require that the probability of entering S is small conditional on starting outside S.

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

$$P(exit\ S|in\ S) \leq \varepsilon\ and\ P(enter\ S|outside\ S) \leq \varepsilon\ .$$

The closer to zero ε can be chosen, the stronger is the property of having an ε-trap. Thus ε can be thought of as indexing the propensity of the economy to experience traps. If ε = 1, then this is no restriction at all since every set of states satisfies the definition of a 1-trap. On the other hand, ε = 0 is a black hole: once S is entered it cannot be exited, however if you don't start in S you also cannot enter it so S is an isolated set of states. We've ruled out this possibility by assuming irreducibility.

The next result justifies these considerations by directly linking ε to the observable data. To do so we will employ the Cheeger inequality for directed graphs (proved by Chung [2005]) which, we believe, is a new and potentially powerful tool for finance and economics.

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

$$\sigma^2(\Delta x) \equiv \sum_{i,j} \pi(i)\pi(i,j)[x(i) - x(j)]^2\ and\ \sigma^2(x) \equiv \sum_i \pi(i)x(i)^2 - \left(\sum_i \pi(i)x(i)\right)^2.$$

Then the economy has an ε-trap with

$$\varepsilon = \frac{\sigma(\Delta x)}{\sigma(x)}.$$

Proof: By the Cheeger inequality for directed graphs,

$$inf_S \frac{P(exit\ S)}{\min\{P(in\ S), 1 - P(\ in\ S)\}} \leq \sqrt{2\lambda}\,,$$

where λ is the second smallest eigenvalue of the Laplacian, L, defined by

$$L\ \equiv I - \frac{1}{2}\left[\Phi^{\frac{1}{2}}\Pi\Phi^{-\frac{1}{2}} +\ \Phi^{\frac{1}{2}}\Pi'\Phi^{-\frac{1}{2}}\right],$$

and, where, Φ is a diagonal matrix with the entries of π on the diagonal. By Corollary 4.2 of Chung [2005], λ satisfies

$$\lambda =\ inf_x sup_c \frac{\sum_{i,j} \pi(i)\pi(i,j)[x(i) - x(j)]^2}{2\sum_j \pi(j)(x(j) - c)^2}.$$

The inner supremum is attained by setting

$$c = \sum_k \pi(k)x(k)\ ,$$

so that we can write

$$\lambda = \ inf_x \ \frac{\sigma^2(\Delta x)}{2\sigma^2(x)}.$$

By the Cheeger inequality, then, there is a set of states S such that for all x,

$$\frac{P(exit\ S)}{\min\{P(in\ S), 1 - P(\ in\ S)\}} \le \frac{\sigma(\Delta x)}{\sigma(x)},$$

which verifies the result.

□

As an example, we can apply this result with x equal to the 10 year constant maturity Treasury yield. In monthly data over the period from April, 1953 to May, 2015, $\sigma(y_{10})$ = 2.61% and $\sigma(\Delta y_{10})$ = 0.25%, and from the above result we can set ε = 0.096.

Result 6 is very general and can be applied using any function x on the state space. For example, we can slightly improve on the above example by choosing x = max{ y_{10}, 10.00%} which results in an ε-trap with $\varepsilon = \sigma(\Delta x)/\sigma(x) = 0.0893$. Intuitively, by capping x at 10% we remove the period when interest rates were in the teens at the end of the 1970's and the beginning of the 1980's. Another interesting example is to explore the implications of setting x equal to a moving average of yields. This and similar choices for x are fertile grounds for future research.

Section IV. Valuing Long Dated Assets

Since returns on the long bond are equivalent to the pricing kernel, it is not surprising that the yield on the long bond is central to the valuation of long dated assets. We begin with a well-known result on the optimality of the long bond.

Result 7. The long bond is growth optimal, i.e., in every state the expected log return on the long bond exceeds that of any other asset.

Proof: Since

$$\varphi \frac{v(j)}{v(i)}$$

is the stochastic discount factor, Result 2 verifies that the reciprocal of the return on the long bond

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

is a stochastic discount factor. Hence, for any asset return R, we have

$$E\left[\frac{R(i,j)}{R_\infty(i,j)}\right] = 1 \,,$$

which implies that

$$E[\log(\mathrm{R(i,j)})] - E[\log(R_\infty(i,j)] = E[log\left(\frac{R(i,j)}{R_\infty(i,j)}\right)] \le log\left(E\left(\frac{R(i,j)}{R_\infty(i,j)}\right)\right) = \log(1) = 0 \,.$$

□

Notice that since in a complete market the stochastic discount factor is unique, the inverse of the return on the long bond is the unique stochastic discount factor. Another route to Result 6 is to note that any asset's price is the expected discounted value of its payoffs where the payoffs are discounted at the rate of return on the long bond. This implies that the cheapest way to achieve a long run payoff is an investment in the long bond.

The growth optimality of the long bond has been noted before this (see, particularly, Bansal and Lehmann [1997]). Alvarez and Jermann [2005] use this feature to build a case that the stochastic discount factor cannot be stationary. Their empirical analysis relies on a reprise of the equity risk premium puzzle by noting that the log return on the overall market is too much greater than that on, say, the thirty year bond to be consistent with a stationary model where the long bond must be log optimal. Given the previous work on the speed of convergence, however, it's far from clear that T = 30 years is close to T = ∞. Whether or not the long bond prices all assets including equities, though, there is nothing that prevents it from being the projection of the economy wide stochastic discount factor on the fixed income market. In that sense, since there is broad agreement that yields are stationary, we are comfortable in thinking of it as the appropriate discount factor for fixed income assets.

While the returns on the long bond replicate the pricing kernel, there are other assets that can also serve as the pricing operator. The next result shows that an infinitely lived asset with a constant dividend yield is also a surrogate for the stochastic discount factor. To see this, consider an infinitely lived asset that pays off, x(i), at each time depending on the state at that time. If x(i) is constant then the asset is simply a consol. The value, p_∞, of the asset is

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

where

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

converges since, by assumption $\varphi < 1$. Thus $p_\infty = A^*x$. Notice that A^* inherits the same dominant eigenvector as A, namely v, and has the associated maximum eigenvalue is $\frac{\varphi}{1-\varphi}$. We can now prove the following result.

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 dividend yield is $1/\varphi - 1$ and its returns perfectly replicate the returns on the one period v-asset and on the long bond. Additionally, no asset can have a uniformly higher or lower dividend yield than the perpetual v-asset.

Proof: From the above discussion, the dividend yield of the perpetual v-asset is

$$\frac{v}{p_\infty} = \frac{v}{A^*v} = \frac{1-\varphi}{\varphi}.$$

Its returns are

$$R_{ij}^{\infty} = \frac{p_\infty(j)+v_j}{p_\infty(i)} = \frac{\left(\frac{\varphi}{1-\varphi}\right)v_j + v_j}{\left(\frac{\varphi}{1-\varphi}\right)v_i} = \frac{1}{\varphi}\left(\frac{v_j}{v_i}\right),$$

which are the returns on the v-asset and the long bond. Uniqueness follows immediately from being the unique positive eigenvector of A^*. The return, R_{ij}, on any perpetual asset with dividends of x is $(A_i^*x)/x_i$ and from Theorem 8.1.26 of Horn and Johnson [1990] this is not uniformly greater or less than the maximum eigenvalue of A^*.

□

Section V. Some General Properties of the Yield Curve

The above results allow us to uncover some new results on pricing in the fixed income markets.

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

Proof: The forward curve is defined by the difference between a T-1 period bond and a T period bond,

$$f_1 + \cdots + f_T = -logB_T,$$

hence, the forward rate from T-1 to T is given by

$$f_T = -logB_T + logB_{T-1} \ .$$

where f and B are understood to be m dimensional vectors with the states suppressed.

Notice that

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

The T period forward rate,

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

where $s \equiv D^{-1}e$.

Applying Jensen's inequality to each row of the state dependent matrix,

$$\begin{aligned} E[f_T] = -log\varphi + E[\log[\Pi^{T-1}s]] - \mathrm{E}[\log[\Pi^T s]] \\ = -log\varphi + E[\log[\Pi^{T-1}s]] - \mathrm{E}[\log[\Pi\Pi^{T-1}s]] \\ = -log\varphi + E[\log[\Pi^{T-1}s]] - \mathrm{E}[\log[E[\Pi^{T-1}s|i]]] \\ < -log\ \varphi + E[\log[\Pi^{T-1}s]] - \mathrm{E}[\mathrm{E}[\log[\Pi^{T-1}s|i]]] \\ = -log\varphi + E[\log[\Pi^{T-1}s]] - \mathrm{E}[\log[\Pi^{T-1}s]] = -log\varphi, \end{aligned}$$

This verifies that the forward curve unconditionally approaches the yield on the long bond from below. Notice, of course, that this is an unconditional result and that this result will be reversed when conditioning on states with a relatively high current forward rate.

□

An alternative and shorter proof of Result 9 follows by noting that

$$f_T(i) = r_T(i,j) + \log B_{T-1}(i) - \log B_{T-1}(j) \ .$$

Using Results 3 and 7, together with the fact that

$$\sum_{i,j} \pi\,(i)\pi(i,j)[\log B_{T-1}(i) - \log B_{T-1}(j)] = 0,$$

we have the result:

$$\sum_i \pi(i) f_T(i) = \sum_{i,j} \pi(i)\pi(i,j) f_T(i) = \sum_{i,j} \pi\,(i)\pi(i,j) r_T(i,j) < \bar{r}_\infty = y_\infty.$$

This makes it clear that Result 9 is a direct consequence of our previous results.

Result 10. On average, i.e., unconditionally, the yield curve is below the long run yield.

Proof: Follows immediately from Result 9 since

$$y_T = -\left(\frac{1}{T}\right) log[B_T] = \left(\frac{1}{T}\right) \sum_{1}^{T} f_t \quad .$$

□

Notice that the approach to the long run is not guaranteed to be monotone in either Result 9 or 10. What is assured is that the long run forward rate and yield is above the average rates for any maturity. The question of monotonicity is delicate and turns on the interplay amongst the eigenvectors and the eigenvalues below the dominant one.

To see why this is so, assume, for example, that Π has distinct eigenvalues and consider the forward curve as a function of the horizon, T. In this case there is a linearly independent set of right eigenvectors and left eigenvectors and if v^j denotes the right eigenvector and x^i the left eigenvector associated with two distinct eigenvalues then v^j and x^i are orthogonal. From the spectral decomposition of Π,

$$f_T = -log\varphi + \log\left[\left[e\pi' + \sum_{1}^{m-1} \gamma_j^{T-1} v^j x^j\right] s\right] - \log\left[\left[e\pi' + \sum_{1}^{m-1} \gamma_j^{T} v^j x^j\right] s\right] ,$$

and the behavior of the expected forward curve as a function of T is governed by the properties of

$$\pi' \log\left[[e\pi' + \sum_{1}^{m-1} \gamma_j^T v^j x^j] s\right] \equiv E[\log[e(\pi' s) + \epsilon_T]] ,$$

where

$$\epsilon_T \equiv \sum_{1}^{m-1} \gamma_j^T v^j x^j s \quad .$$

Notice that

$$E[\epsilon_T] = \sum_{1}^{m-1} \gamma_j^T \pi' v^j (x^j s) = 0 \quad .$$

The properties of ε_T are dependent on how the eigenvectors are interrelated and how those relations change with T. For example, suppose that m = 3 and thus

$$\epsilon_T \equiv \sum_{1}^{m-1} \gamma_j^T v^j x^j s = (x^1 s)\,\gamma_1^T v^1 + (x^2 s)\gamma_2^T v^2 \ .$$

Since by the Perron-Frobenius Theorem the eigenvalues are below 1 in magnitude, it is tempting to conjecture that this term becomes less volatile as T is increased. However, if v^1 and v^2 are negatively correlated and the ratio of their associated eigenvalues is sufficiently close to one, increasing T can make this term more volatile. In this case, as T increases the forward rates could move away from the long run yield, -log ϕ, before they converge to it. In general, the forward curve could have m peaks and troughs before it asymptotes to the long run yield, –log ϕ.

It is important to observe that Results 9 and 10 do not necessarily hold for some typical affine models. For example, in the CIR model, if the volatility of interest rates is sufficiently large, then the yield curve will be downward sloping with maturity and the long yield does not majorize unconditional forward rates and finite maturity yields .

Since both models are stationary and since the CIR has a representative agent, although it is cast in continuous time, in some senses it is a specialization of our model. But, a principal difference is that the marginal utility of the representative agent is determined by a power utility function and is further restricted to a log utility function of the form, $e^{-\rho t}$ ln(c(t)). This implies that the pricing kernel is neither bounded away from zero nor unbounded in wealth and even though the interest rate is stationary the random evolution of wealth can permit the long run stationary yield to differ from the discount factor, ρ.

This suggests that the inconsistency in the results occurs because the kernel in the CIR model depends on aggregate wealth which, in turn, depends not just on the current state variable but on the integral of interest rates over the future path. When the expectation is taken in the proof of Result 9 the conditioning is over the prior path and this violates the Markov assumption of our model. Exploring the exact differences in the models is an important area for future research.

Section VI. Conclusion

The model we have developed provides a complete, rich, and empirically testable set of implications for the fixed income markets. We have shown that the long bond, whether or not it is observable, is a very useful theoretical construct that simplifies many existing results and points the way to some new ones. In particular, the dependence of the speed of convergence on the second eigenvalue of the pricing matrix allows us to ascertain how far the limiting asymptotic results will be from observed results for traded bonds. In addition, the introduction of the concept of traps and the nature of the convergence of long yields to the infinitely long yield should prove powerful for future research.

As we have seen, too, there are some important differences between the model of this paper and neoclassical factor models such as the CIR model. To begin with, while it is easily possible to impose a factor structure on the state space pricing matrix A and, therefore, on the probability matrix Π, we have made no use of such an assumption. Doing so would no doubt allow us to sharpen some of the results, for example, speed of convergence and the eigenvalue structure would become functions of the factor loadings. In a two factor world where state information is captured by a single interest rate level and a second factor that captures the slope of the term structure or, perhaps, the current level of volatility, we would expect the largest and the second largest eigenvalues to be functionally dependent on the loadings on the two factors. This would be quite interesting to explore.

We have also found results that are inconsistent with a standard model such as CIR For example, Results 9 and 10 tell us that the long yield majorizes the unconditional forward rate and the yield for all maturities. In the CIR model, though, this is not generally the case and, particularly, if the volatility of interest rates is sufficiently large, then the curve will be downward sloping with maturity (see CIR, 1985). Reconciling these results with the model we have developed should be a major item for future research.

In a different direction, it might be fruitful to consider state spaces that exhibit cycles, traps, hysteresis, non-reversible transitions, and so on. The theoretical literature (and, arguably, recent macroeconomic experience) has provided many examples in which such phenomena occur, and our framework is flexible enough to accommodate them.

References

Alvarez, F., and U. J. Jermann (2005), "Using Asset Prices to Measure the Persistence of the Marginal Utility of Wealth," Econometrica, 73:6:1977-2016.

Backus, D. K., A. W. Gregory, and S. E. Zin (1989), "Risk Premiums in the Term Structure: Evidence from Artificial Economies," Journal of Monetary Economics, 24:371-399.

Bansal, R., and B. N. Lehmann (1994), Addendum to "Growth-Optimal Portfolio Restrictions on Asset Pricing Models," working paper.

Bansal, R., and B. N. Lehmann (1997), "Growth-Optimal Portfolio Restrictions on Asset Pricing Models," Macroeconomic Dynamics, 1:333-354.

Breeden, D. T., and R. H. Litzenberger (1978), " Prices of State-Contingent Claims Implicit in Option Prices," Journal of Business, 51:4:621-651.

Cox, J.C., Ingersoll, J.E., and Ross, S.A., "A Theory of the Term Structure of Interest Rates," Econometrica 53, No. 2, March 1985, 385-407.

Dybvig, P. H., J. E. Ingersoll, and S. A. Ross (1996), "Long Forward and Zero-Coupon Rates Can Never Fall," Journal of Business, 69:1:1-25.

Hansen, L. P., and J. A. Scheinkman (2009), "Long-Term Risk: An Operator Approach," Econometrica, 77:1:177-234.

Horn, R. A., and C. R. Johnson (1990), Matrix Analysis, Cambridge University Press, Cambridge, UK.

Kazemi, H. B. (1992), "An Intemporal Model of Asset Prices in a Markov Economy with a Limiting Stationary Distribution," Review of Financial Studies, 5:1:85-104.

Ross, S. A. (2015), "The Recovery Theorem," Journal of Finance, Vol 70, 2, 2015, 615-648.