

Exponential-Affine Models

(Started: April 12, 2012; Revised: December 2, 2013)

We look at models in which logarithms of bond prices are linear functions of a state variable. This convenient form makes it relatively easy to explore their properties. An example is the behavior of term premiums, whose apparent variation and predictability seems like a feature of the data we'd like to understand better.

1 Vasicek model

Model. A more common version of the Vasicek model than our ARMA(1,1) pricing kernel consists of the two equations

$$\begin{aligned}\log m_{t+1} &= \delta - x_t + \lambda w_{t+1} \\ x_{t+1} &= \varphi x_t + \sigma w_{t+1},\end{aligned}$$

with $\{w_t\} \sim \text{NID}(0, 1)$. If you substitute for x_t , you can see that this is the ARMA(1,1) in different form. Here we've given the state variable x_t a mean of zero, something we'll come back to. We'll see that this model, like the others in this module, delivers bond prices whose logs are linear in the state variable x_t .

Recursions. We build up bond prices one at a time from the pricing relation

$$q_t^{n+1} = E_t(m_{t+1} q_{t+1}^n),$$

starting with $q_t^0 = 1$ (a dollar now is worth a dollar). Let's get warmed up with the short rate f_t^0 , corresponding to $n = 0$:

$$\begin{aligned}q_t^1 &= E_t m_{t+1} \\ -f_t^0 &= \log q_t^1 = \log E_t m_{t+1}.\end{aligned}$$

We'll solve this with the usual "mean plus variance over two" formula for lognormal random variables. As of date t , the conditional distribution of $\log m_{t+1}$ is normal with mean and variance

$$\begin{aligned}E_t m_{t+1} &= \delta - x_t \\ \text{Var}_t m_{t+1} &= \lambda^2.\end{aligned}$$

The short rate is therefore $f_t^0 = x_t - (\delta + \lambda^2/2)$. The short rate is x_t plus a constant, so it inherits the variance and autocorrelation of x_t .

For bonds of higher maturity, we guess prices of "exponential-affine" form

$$\log q_t^n = A_n + B_n x_t$$

for some coefficients $\{A_n, B_n\}$ yet to be determined. (And no, A_n isn't the cumulative sum as it was in our last set of notes, it's a stand-alone coefficient.) Think about this for some particular maturity n . The pricing relation tells us we need to evaluate the conditional mean of the product $m_{t+1}q_{t+1}^n$. We do this the usual way: take logs, then use the properties of lognormal random variables. Here we have

$$\log m_{t+1} + \log q_{t+1}^{n+1} = (\delta + A_n) + (\varphi B_n - 1)x_t + (\lambda + B_n\sigma)w_{t+1}.$$

The conditional mean and variance are

$$\begin{aligned} E_t(\log m_{t+1} + \log q_{t+1}^n) &= (\delta + A_n) + (\varphi B_n - 1)x_t \\ \text{Var}_t(\log m_{t+1} + \log q_{t+1}^{n+1}) &= (\lambda + B_n\sigma)^2. \end{aligned}$$

Using (again!) the “mean and variance over two” formula, we have

$$\begin{aligned} \log q_t^{n+1} &= (\delta + A_n) + (\varphi B_n - 1)x_t + (\lambda + B_n\sigma)^2/2 \\ &= A_{n+1} + B_{n+1}x_t. \end{aligned}$$

Lining up terms, we have

$$\begin{aligned} A_{n+1} &= A_n + \delta + (\lambda + B_n\sigma)^2/2 \\ B_{n+1} &= \varphi B_n - 1. \end{aligned}$$

Given values for the parameters $(\delta, \varphi, \sigma, \lambda)$ we compute these one at a time, starting with $A_0 = B_0 = 0$. It's easy to do this in (say) Matlab, or even a spreadsheet program.

Estimation. Now where do the parameters come from? We'll use properties of forward rates, starting with the short rate f_t^0 . In principle they're interconnected, but here we can solve for them one at a time:

- Autocorrelation. Because f_t^0 is linear in x_t , and x_t is an AR(1), the autocorrelation of the short rate is φ . So we set φ equal to the observed autocorrelation of the short rate in the data. Using the table distributed in class, we have $\varphi = 0.959$.
- Variance. Similarly, the variance of the short rate is the variance of x , which is $\sigma^2/(1-\varphi^2)$. Using data from the same table gives us ??? Note that we have converted the numbers in the table to monthly numbers.
- Long forward rates. Next we have λ , which controls the pricing of interest rate risk and therefore the slope of the forward rate curve. Recall that forward rates are

$$f_t^n = \log(q_t^n/q_t^{n+1}) = (A_n - A_{n+1}) + (B_n - B_{n+1})x_t.$$

Since x_t has a zero mean, we have

$$E(f^n - f^0) = (A_n - A_{n+1}) - (A_0 - A_1).$$

There's nothing particularly intuitive about this formula, but if we choose values of φ , σ , and λ , we can compute the A 's and see how they line up with the evidence. If we're off, we vary λ until we get it.

[Comment: this works because the δ 's cancel out. Show this?]

- Mean short rate. This doesn't really matter, but the mean short rate is

$$E(f^0) = -(\delta + \lambda^2/2).$$

Given our value of λ , we choose δ to match the observed mean short rate.

If any of this seems mysterious, see the Matlab program.

An alternative parameterization (optional). Suppose we change the model (slightly) to

$$\begin{aligned}\log m_{t+1} &= -\lambda^2/2 - x_t + \lambda w_{t+1} \\ x_{t+1} &= (1 - \varphi)\delta + \varphi x_t + \sigma w_{t+1}.\end{aligned}$$

Show that the short rate is now exactly x_t . The model is equivalent to the previous one, we've simply moved the mean to a different part of the model.

[Start with this?]

2 Duration revisited

[next time]

3 The Cox-Ingersoll-Ross model

Model. Here's a popular variant of the Vasicek model in which risk and risk premiums move around. More on the latter shortly. The model is

$$\begin{aligned}\log m_{t+1} &= -(1 + \lambda^2/2)x_t + \lambda x_t^{1/2} w_{t+1} \\ x_{t+1} &= (1 - \varphi)\delta + \varphi x_t + x_t^{1/2} \sigma w_{t+1},\end{aligned}$$

with (as usual) independent standard normal innovations w_t . Here x_t plays two roles: It governs the conditional mean of the pricing kernel, as it did before, but it also governs the conditional variance.

The so-called “square-root” process for x_t guarantees that it remains positive — at least it does in continuous time. The idea is that the variance shrinks as x_t gets close to zero. As a result, the change in x_t near zero comes from the mean reversion term,

$$x_{t+1} - x_t \approx (1 - \varphi)\delta > 0.$$

As long as $\delta > 0$, that pushes x_t away from zero and keeps it positive.

Despite the square-root process, x_t is still an AR(1). Its mean is

$$E(x_t) = (1 - \varphi)\delta + \varphi E(x_{t-1}) \Rightarrow E(x) = \delta.$$

The variance is

$$\begin{aligned}\text{Var}(x) &= E(x_t - \delta)^2 = E\left[\varphi(x_{t-1} - \delta) + \sigma x_{t-1}^{1/2} w_t\right]^2 \\ &= \varphi^2 \text{Var}(x) + E(\sigma^2 x_{t-1}) = \varphi^2 \text{Var}(x) + \sigma^2 \delta,\end{aligned}$$

or

$$\text{Var}(x) = \sigma^2 \delta / (1 - \varphi^2).$$

What about the autocovariance? The first autocovariance is

$$\begin{aligned}\text{Cov}(x_t, x_{t-1}) &= E(x_t - \delta)(x_{t-1} - \delta) \\ &= E\left\{[\varphi(x_{t-1} - \delta) + \sigma x_{t-1}^{1/2} w_t](x_{t-1} - \delta)\right\} \\ &= \varphi \text{Var}(x).\end{aligned}$$

The first autocorrelation is therefore φ .

Recursions. We follow the same logic as in the Vasicek model, starting with the short rate

$$-f_t^0 = \log E_t m_{t+1}.$$

The conditional distribution of $\log m_{t+1}$ is normal with mean and variance

$$\begin{aligned}E_t m_{t+1} &= -(1 + \lambda^2/2)x_t \\ \text{Var}_t m_{t+1} &= \lambda^2 x_t.\end{aligned}$$

The new ingredient is that the variance is now linear in x_t . The short rate is therefore $f_t^0 = x_t$. That's why we started with a somewhat odd coefficient of x_t in the pricing kernel — so this would work out.

For bonds of arbitrary maturity, we use the guess $\log q_t^n = A_n + B_n x_t$. That gives us

$$\log m_{t+1} + \log q_t^{n+1} = [A_n + (1 - \varphi)\delta] + [\varphi B_n - (1 + \lambda^2/2)]x_t + (\lambda + B_n \sigma)x_t^{1/2} w_{t+1}.$$

whose conditional mean and variance are

$$\begin{aligned}E_t (\log m_{t+1} + \log q_t^{n+1}) &= [A_n + B_n(1 - \varphi)\delta] + [\varphi B_n - (1 + \lambda^2/2)]x_t \\ \text{Var}_t (\log m_{t+1} + \log q_t^{n+1}) &= (\lambda + B_n \sigma)^2 x_t.\end{aligned}$$

The “mean and variance over two” formula gives us

$$\begin{aligned}\log q_t^{n+1} &= [A_n + B_n(1 - \varphi)\delta] + [\varphi B_n - (1 + \lambda^2/2)]x_t + (\lambda + B_n \sigma)^2/2 x_t \\ &= A_{n+1} + B_{n+1} x_t.\end{aligned}$$

Lining up terms, we have

$$\begin{aligned}A_{n+1} &= A_n + B_n(1 - \varphi)\delta \\ B_{n+1} &= \varphi B_n - (1 + \lambda^2/2) + (\lambda + B_n \sigma)^2/2.\end{aligned}$$

This is a mess, but it's easy to program up, which is what we'll do. We start, as always, with $A_0 = B_0 = 0$.

Estimation. We estimate parameters much as we did before:

- Autocorrelation. We set φ equal to the observed autocorrelation of the short rate in the data, just as we did before.
- Mean. We set δ equal to the mean short rate.
- Variance. Similarly, the variance of the short rate is the variance of x , which is $\sigma^2\delta/(1-\varphi^2)$. Given values for φ and δ , the variance of the short rate nails down σ .
- Long forward rates. Again, λ controls the mean forward rate curve. Recall that forward rates are

$$f_t^n = \log(q_t^n/q_t^{n+1}) = (A_n - A_{n+1}) + (B_n - B_{n+1})x_t.$$

Since x_t has a mean of δ , we have

$$E(f_t^n - f_t^0) = [(A_n - A_{n+1}) - (A_0 - A_1)] + [(B_n - B_{n+1}) - (B_0 - B_1)]\delta.$$

We simply vary λ until it gives us the right mean forward spread.

See the Matlab program.

4 Term premiums

Now let's think about *term premiums*: (risk) premiums on bonds of different terms or maturities. We connect term premiums reflected in bond returns and forward rates. We start with the latter; the others follow from their definitions. Then we describe how term premiums can be predicted.

Definitions. Let q_t^n be the price at date t of an n -period zero-coupon bond. Yields y and forward rates f are defined from prices by

$$-\log q_t^n = ny_t^n = f_t^0 + f_t^1 + \dots + f_t^{n-1}.$$

The definition of forward rates implies $f_t^n = \log(q_t^n/q_t^{n+1})$. One-period returns are $r_{t+1}^n = q_{t+1}^{n-1}/q_t^n$. The *short rate* is $\log r_{t+1}^1 = y_t^1 = f_t^0$.

We define (forward) *term premiums* tp by

$$f_t^n = E_t f_{t+n}^0 + tp_t^n. \tag{1}$$

We define *bond premiums* from (log) excess returns:

$$bp_t^n = E_t(\log r_{t+1}^n - \log r_{t+1}^1).$$

The question is how the two are related.

The answer follows from the definitions. The excess return on an $(n+1)$ -period bond is

$$\log r_{t+1}^n - \log r_{t+1}^1 = (f_t^1 - f_{t+1}^0) + \dots + (f_t^n - f_{t+1}^{n-1}).$$

The typical term is

$$f_t^j - f_{t+1}^{j-1} = (E_t f_{t+j}^0 - E_{t+1} f_{t+j}^0) + tp_t^j - tp_{t+1}^{j-1}.$$

That is, you get innovations in forward rates and changes in term premiums. Since the former have (conditional) mean zero by construction, only the latter show up in bond risk premiums:

$$\begin{aligned} bp_t^{n+1} &= E_t(tp_{t+1}^1 - tp_t^0) + E_t(tp_{t+1}^2 - tp_t^1) + \dots + E_t(tp_{t+1}^n - tp_t^{n-1}) \\ &= E_t(tp_{t+1}^1 - tp_t^1) + E_t(tp_{t+1}^2 - tp_t^2) + \dots + E_t(tp_{t+1}^n - tp_t^{n-1}) + tp_t^n. \end{aligned}$$

In short, bond premiums are functions of term premiums. In stationary ergodic settings, the unconditional means have a more simple form: $E(bp^{n+1}) = E(tp^n)$.

Evidence on term premiums. We consider three approaches to term premiums:

- Term premiums are zero: $tp_t^n = 0$. This is obviously false, because on average term premiums increase with maturity n .
- Term premiums are constant and depend only on maturity: $tp_t^n = tp^n$.
- Term premiums are arbitrary functions of the state. This allows almost anything we might consider, including exponential-affine models.

Let's see what they deliver.

Although it's clearly false, the first approach leads to some useful expressions for forward rates and bond returns. The first is that excess returns on bonds of all maturities are zero. We can look at forward rates or bond returns, but we see the same thing. We also have some strong implications for the dynamics of forward rates. If term premiums are zero, equation (1) becomes

$$f_t^n = E_t f_{t+n}^0.$$

That is: forward rates are expectations of future short rates. For this reason, the first approach is often referred to as the *expectations hypothesis*.

We're going to show that the expectations hypothesis implies that forward rates are martingales:

$$f_t^n = E_t f_{t+1}^{n-1}. \quad (2)$$

Why? It's a little subtle, but let's work through the logic. The key ingredient is the law of iterated expectations, an extension of what I learned as the double expectation theorem. If we have a stochastic process for a random variable x_t , then

$$E_t x_{t+n} = E_t (E_{t+1} x_{t+n}) = E_t (E_{t+1} [E_{t+2} x_{t+n}]) \dots$$

We've done stuff like this all along, computing things one period at a time. It works because as we move the expectation back in time, we base it on less information. This says we can

either do that directly (the left-hand side) or do it one period at a time (the right-hand side).

Now let's apply this to the expected future short rate. We'll do this for $n = 2$ to keep things simple. We have

$$f_t^2 = E_t f_{t+2}^0 = E_t(E_{t+1} f_{t+2}^0) = E_t f_{t+1}^1.$$

More generally, we get the martingale sequence $f_t^n, f_{t+1}^{n-1}, \dots, f_{t+n}^0$. Each is the expectation of the short rate at $t + n$, but they're expectations at different dates.

All of this is a long-winded way of setting up some regression evidence. Under the expectations hypothesis, forward rates are martingales. We can examine that idea by trying to see if we can come up with a better prediction than zero. We've already seen that a constant helps, since it adds a nonzero risk premium. But we can do better than that. If we estimate

$$E_t(f_{t+1}^{n-1} - f_t^0) = a_n + b_n(f_t^n - f_t^0) + \text{error}$$

for lots of maturities n , what do we see? If the expectations hypothesis is right, we should get $b_n = 1$. In fact we get $b^1 = 0.5$, $b^3 = 0.75$, etc. As n increases, the coefficients get closer to one, but at short maturities they're clearly not one.

So what does that tell us? We can predict changes in term premiums. It's not hard to show that this carries over to excess returns.

5 Predictability in affine models