CHAPTER 19

YIELD CURVE MATHEMATICS

§1. The Term Structure of Rates

The *term structure of interest rates* refers to the familiar fact that borrowing money "costs" different amounts depending upon how long one wishes to borrow it for. By "costs" we mean the rate of interest one must pay to borrow. This information can be observed in the financial marketplace. There are several equivalent ways to quote the term structure. We discuss them here.

For simplicity, we think of entering into lending and borrowing agreements as risk-free in the sense that any money promised to be paid in the future will, with certainty, be paid. We assume there are no transaction costs or other hidden expenses associated with becoming either a lender or a borrower. We further assume that anyone may become a lender or borrower of any amount at any time.

The Discount Function. Perhaps the most direct way to express the term structure is through the *discount function.* Suppose that, for all $t \geq 0$, the bond market tells us that one can borrow $d(t)$ dollars today with the agreement to satisfy the loan with a single payment of \$1 at time t. Here t measures time in years and $t = 0$ corresponds to today. Then $d(t)$ is the *present value* of \$1 to be paid at time t in the future: $d(t)$ dollars received today is worth exactly the same as \$1 to be received at time t. Clearly, then, $d(0) = 1$ and $d(t)$ is decreasing in t.

With complete knowledge of the discount function, we can value any sequence of cash flow as follows. Suppose a security promises to pay the sequence of cash flow C_i, $i = 1, 2, 3 \ldots$, where the i^{th} cash flow is to be paid at time $t_i > 0$. Then the value today of that security is

$$\textit{value today} \;=\; \sum_i d(t_i) \cdot C_i. \tag{1}$$

We may think of this security as a *portfolio* comprising, for each i, C_i units of a security that pays \$1 at time t_i. Formula (1) uses the basic assumption of *linearity*: the value of a portfolio is the sum of the values of its constituent securities, and the value of n units of a security of unit value V is nV.

The Zero Coupon Yield Curve. We convert this information to a rate of interest as follows. At an annual *discount rate* of r with the convention of *continuous compounding*, \$1 paid at time t has a present value (PV) of

$$\textit{PV with continuous compounding} \;=\; \$1 \cdot e^{-rt}.$$

With the convention of *periodic compounding*, the present value is given by

$$\textit{PV with periodic compounding} \;=\; \$1 \cdot \left(1 + \frac{r}{n}\right)^{-nt}.$$

Here n is the *compounding frequency*; for example, $n = 2$ corresponds to semiannual compounding; $n = 4$ corresponds to quarterly; etc. The term "continuous compounding" comes from the fact that it represents the limiting value of periodic compounding as we increase the compounding frequency:

$$\begin{aligned} \lim_{n\to\infty} \left(1 + \frac{r}{n}\right)^{-nt} &= \left[\lim_{m\to\infty} \left(1 + \frac{1}{m}\right)^{m}\right]^{-rt} \quad \text{(substitute } m = n/r\text{)} \\ &= e^{-rt}. \end{aligned}$$

The *zero coupon yield curve*, also called *the spot curve*, presents the information contained in the discount function as the associated discount rates. Using $z(t)$, $t > 0$, to denote this curve, with the convention of continuous compounding it is related to $d(t)$ as follows:

$$d(t) = e^{-z(t)t}, \quad \text{or} \quad z(t) = -\frac{1}{t}\log d(t).$$

The same information can be presented with the convention of periodic compounding. For example, with semiannual compounding, the relationship between $d(t)$ and $z(t)$ would be

$$d(t) = \left(1 + \frac{z(t)}{2}\right)^{-2t}, \quad \text{or} \quad z(t) = 2\big(d(t)^{-1/(2t)} - 1\big).$$

It is apparent that $d(t)$ and $z(t)$ are just different ways to present the same information: if you are told $d(t)$, you can compute $z(t)$, and vice-versa.

The Par Yield Curve. A *bond*, for our present purposes, is a security whose promised cash flow depends on four features: the annual *coupon rate* (c), the *yearly payment frequency* (k), the bond's *maturity* (T), and the bond's *face amount* (F). The bond makes k interest payments per year in the amount of $F \cdot \frac{c}{k}$ until time T, at which time it makes a final interest payment (same amount) plus a return of *principal* in the amount of F.

We will assume that bonds make semiannual payments ($k = 2$) and have a face amount $F = \$100$. We fix Δt to be 0.5 years and put $t_i = i\Delta t = 0.5i$, for

$i = 0, 1, 2, \ldots$. For example, a bond with a 5% annual coupon rate maturing in two years will make four payments: \$2.5, \$2.5, \$2.5, and \$102.5, at times $t_1 = .5$, $t_2 = 1.0$, $t_3 = 1.5$, and $t_4 = 2.0$ years, respectively.

A *par bond* is a bond whose coupon rate is such that the value of the bond is equal to its face amount. Using (1), we see that the coupon rate for a par bond maturing at time t_n, call it $c(t_n)$, is therefore given by:

$$100 \;=\; \sum_{i=1}^{n} 0.5c(t_n) \cdot 100 \cdot d(t_i) \;+\; 100 \cdot d(t_n),$$

producing

$$c(t_n) \;=\; \frac{1 - d(t_n)}{0.5 \sum_{i=1}^{n} d(t_i)}.$$

Re-writing this as

$$c(t_n) \;=\; \frac{1 - d(t_n)}{0.5 \sum_{i=1}^{n-1} d(t_i) + 0.5d(t_n)},$$

we may solve for $d(t_n)$ producing

$$d(t_n) \;=\; \frac{1 - 0.5c(t_n) \sum_{i=1}^{n-1} d(t_i)}{1 + 0.5c(t_n)}. \tag{2}$$

This looks problematic, since the discount function appears on both sides of (2). But (2) expresses $d(t_n)$ in terms of the par curve and the discount function valued at times *earlier* than time t_n. So we may first use (2) to compute $d(t_1)$ (where the summation in the numerator is vacuous), then $d(t_2)$, then $d(t_3)$, etc. This process is sometimes called *bootstrapping*.

Single-Period Forward Rates. Yet another way to report the term structure is through the *single-period forward rates* $f(t_i)$. With Δt still $\frac{1}{2}$ year, today's single-period forward rate at time t_i is the interest rate that you would agree to *today* for loaning *in the future* for the $\frac{1}{2}$ year period from time t_i to time t_{i+1}. To be clear, the rate is agreed to today, but the transaction (the $\frac{1}{2}$ year loan) takes place in the future at time t_i.

To calculate this, compute the amount A such that if one agrees today to borrow \$1 at time t_i and repay amount \$$A$ at time t_{i+1}, the transaction is valued at \$0. Using (1), we wish that $0 = 1 \cdot d(t_i) - A \cdot d(t_{i+1})$, so

$$A \;=\; \frac{d(t_i)}{d(t_{i+1})}. \tag{3}$$

Agreeing now (at time 0) to borrow \$1 at time t_i and repay it one period (one half year) later with amount \$$A$ is a *fair* transaction — i.e. one that is valued at \$0. The notion of single-period forward rate converts this to an interest rate (with the convention of continuous compounding) as follows:

$$1 \;=\; Ae^{-0.5f(t_i)} \quad \text{or} \quad f(t_i) = \frac{\log A}{0.5} \;=\; 2\log(d(t_i)/d(t_{i+1})). \tag{4}$$

The analog of this reflecting a discrete compounding period of length 0.5 (that is, compounding two times per year) is:

$$1 \;=\; A(1+0.5f(t))^{-1} \quad \text{or} \quad f(t) = \frac{A-1}{0.5} \;=\; 2((d(t_i)/d(t_{i+1}))-1)$$

With continuous compounding, from (3) and the left side of (4) we see that $d(t_i) = d(t_{i-1})e^{-0.5f(t_{i-1})}$. Repeated application of this (using that $d(t_0) = 1$) gives

$$\begin{aligned} d(t_n) \;&=\; e^{-0.5f(0)} \cdot e^{-0.5f(t_1)} \cdot e^{-0.5f(t_2)} \cdots e^{-0.5f(t_{n-1})} \\ &=\; e^{-\sum_{i=0}^{n-1} 0.5f(t_i)}; \end{aligned} \tag{5}$$

so the discount function, for the times t_n, can be reconstituted from the single-period forward rates.

With continuous compounding we also have $d(t_n) = e^{-z(t_n)t_n}$ so equating exponents yields that $z(t_n)t_n = \sum_{i=0}^{n-1} 0.5f(t_i)$, or, since $t_n = 0.5n$,

$$z(t_n) \;=\; \frac{1}{n}\sum_{i=0}^{n-1} f(t_i). \tag{6}$$

That is, the zero coupon rate to time t_n is the arithmetic average of the first n single-period forward rates. While this fact is exact with continuous compounding, it is only approximately true for periodic compounding.

Exercise 1. Show that with periodic compounding $1 + z(t_n)$ is the geometric mean of the numbers $1 + f(t_i)\Delta t$, $0 \le i < n$.

Continuous-Time Limit. The calculations above assumed discretized time with $\Delta t = 0.5$ years. One may envision a discount function $d(t)$ available for a continuous time variable t with the corresponding zero-coupon curve $z(t)$ and *instantaneous* forward rates $f(t)$. This useful mathematical construct is not observable in the marketplace. With this notion, (5) and (6) become

$$d(t) \;=\; e^{-\int_0^t f(s)\,ds}, \quad \text{and} \quad z(t) \;=\; \frac{1}{t}\int_0^t f(s)\,ds. \tag{7}$$

Differentiation then yields

$$z'(t) \;=\; \frac{1}{t}f(t) - \frac{1}{t^2}\int_0^t f(s)\,ds \;=\; \frac{1}{t}f(t) - \frac{1}{t}z(t),$$

and solving for $f(t)$ produces

$$f(t) \;=\; z(t) + tz'(t). \tag{8}$$

From (8) we see that the instantaneous forward rate t years forward lies above the t year zero coupon rate if and only if the the zero coupon curve is positively sloped at maturity t. Also, locations where $z'(t)$ is discontinuous will create discontinuities in $f(t)$.

Additionally, if we assume that the forward rates $f(t)$ are bounded, then the spot curve is also bounded (the same bound will do, see (7), right side). It follows that

$$z'(t) \;=\; \frac{f(t) - z(t)}{t} \;\to\; 0, \quad \text{as } t \text{ gets large.}$$

In other words, the spot curve should flatten with maturity. This, in fact, is observed in the real world. The assumption that $f(t)$ is bounded is not crazy; after all the Fed controls short-term rates.

§2. Four Examples

Presently we will look at several par yield curves of varying shapes and their resulting spot curve and single-period forward rates. But first a few words about some software that we will use repeatedly to generate par curves of various shapes. Shown in Figure 1 are what we'll call the *Nelson-Siegel* "shift", "twist" and "hump" functions (my terminology). They are given by

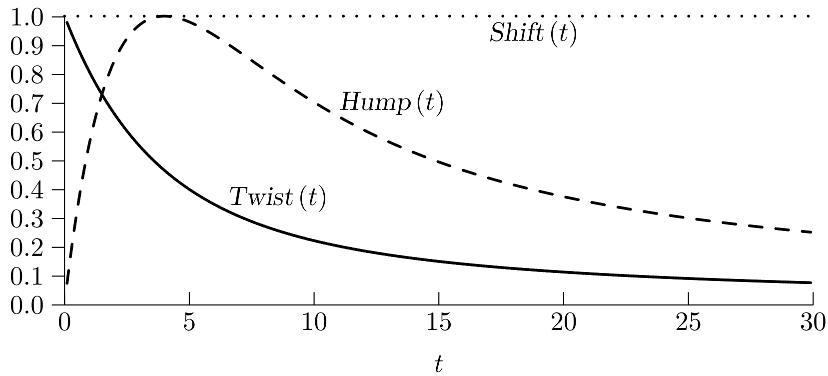
$$\begin{aligned}
\mathit{Shift}\,(t) \;&=\; 1,\\
\mathit{Twist}\,(t) \;&=\; \frac{1 - e^{-0.44832t}}{0.44832t}, \quad \text{and}\\
\mathit{Hump}\,(t) \;&=\; 3.3509 \cdot \left[\frac{1 - e^{-0.44832t}}{0.44832t} - e^{-0.44832t}\right].
\end{aligned}$$

The constant 0.44832 scales time so that the maximum value of $\mathit{Hump}\,(t)$ occurs at time $t = 4$ years; the constant 3.3509 scales *Hump* so that its maximum value is 1. [The Nelson-Siegel functions include a parameter, τ, specifying the the time at which the hump is maximal. Here, for convenience, we always take $\tau = 4$ years.] The par yield curves that we will consider are of the form

$$c(t) \;=\; A \cdot \mathit{Shift}\,(t) + B \cdot \mathit{Twist}\,(t) + C \cdot \mathit{Hump}\,(t), \tag{9}$$

for numbers A, B, and C.

Figure 1. The Nelson-Siegel shift, twist, and hump functions.



The program `YieldMath.cpp` implements this, generating a par curve via (9) and the corresponding spot and forward rate curves. The program asks the user for *short rate*, *long rate*, and *hump*, all to be input as percents. In (9), C is the user-specified value of *hump*. The coefficients A and B are then solved for so that the 6-month par yield, $c(0.5)$, as computed in (9), is *short rate* and the 30-year par yield, $c(30.0)$, as computed in (9), is *long rate*. The par curve is then computed using (9) for times $t_n = n \cdot 0.5$, $1 \leq n \leq 60$. (Throughout this text bonds pay interest semiannually and our interest rate models will extend for 30 years.)

Exercise 2. Calculate A, B, and C if the *short rate* is 2%, the *long rate* is 5%, and the *hump* is 1%. For this yield curve, what is the 10-year par yield?

Remark. Nelson and Siegel use this methodology to parametrically model the *spot* curve, rather than the par curve.

Figure 2 illustrates this for two upward sloping par curves (this is typical). In the first case, the short- and long-rates are 1% and 4%, respectively, with no hump ($C = 0$). The second case shows the same short- and long-rates with a hump of 3%. When there is no hump, the par curve (solid line) is monotonically increasing. Note that the spot curve (dashes) lies above the par curve, but almost exactly coincides with the par curve for short maturities. To understand this heuristically, think of a par bond as a portfolio of zero coupon bonds — one for each coupon payment and a final, and much larger one, for the final coupon and principal payment. The yield on the par bond should be a weighted average of the yields on the underlying zeros. Most of the weight should be placed on the yield of the zero corresponding to the par bond's payment at maturity. This, after all, is by far the largest payment and, at least for relatively short par bonds, contributes the majority of the bond's value. A positively sloped spot yield curve is therefore associated with a positively sloped par curve. The par curve should lie below the spot curve, however, because the t-year par bond yield is a blend of the t-year spot yield and, to a lesser extent, (smaller) spot yields of shorter maturity. For very short par bonds, the final payment contributes the overwhelming majority of the bond's value. Such a bond's yield should therefore roughly correspond to the yield on a zero of the same maturity. Note also that the single-period forward rates (the $\bullet$s), lie above the spot curve. This is because $z'(t) > 0$ — recall formula (8).

For the humped curve, the situation is more complicated. Note that sometime after the hump's peak at 4 years, the spot curve passes through the par curve. This portion of the curve is inverted (see discussion below). Note also that the forward rates pass through the spot curve just as the spot curve peaks, i.e., where $z'(t) = 0$.

Figure 3 shows the situation for two inverted par curves. In both cases, the short- and long-rates are 7% and 4%, respectively. The first case has no hump ($C = 0$); the second case has a hump of -3%. The first case is a mirror image of the positively sloped par curve with no hump. Here, the spot curve is inverted but lies below the par curve. This is because the t-year par yield is a blend of the t-year spot yield and (bigger) spot yields of shorter maturity. Note that the single-period forward rates lie below the spot curve as here $z'(t) < 0$.

The reader should be sure to understand the relationship between the three curves in the case where the hump is -3%.

Figure 2. Two upward sloping par yield curves.

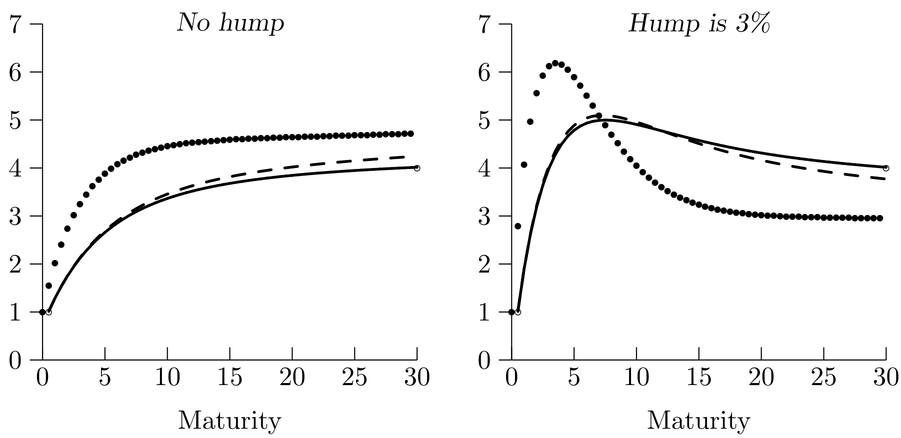
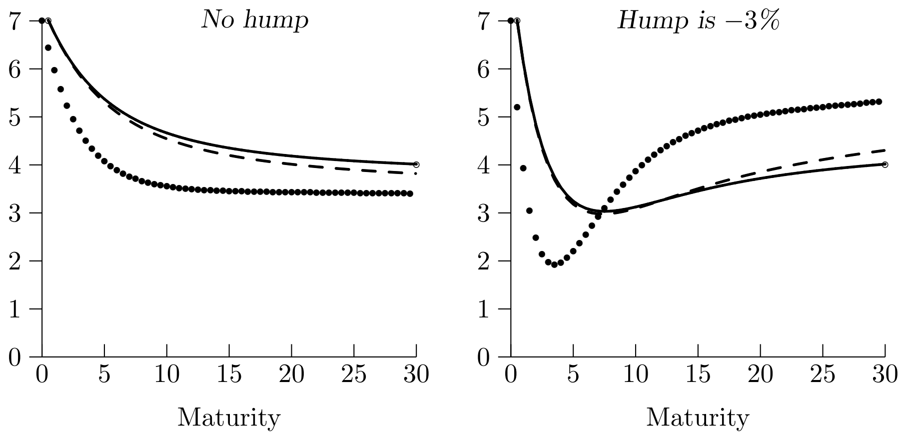


Figure 3. Two inverted par yield curves.



Exercise 3. With the par curve of Exercise 2, use the formulae developed in §1 to calculate the discount function out to 4 years: $d(0.5), d(1.0), \ldots, d(4.0)$. Calculate the corresponding zero-coupon yields and the single-period forward rates $f(0), f(0.5), \ldots, f(3.5)$ with the convention of continuous compounding.

§3. Yield Curve Interpolation

We have based our discussion on the notion of risk-free borrowing and lending: all obligations were assumed to be paid in full and on time with certainty. In practice, typical sources of this information are sovereign yield curves and the LIBOR swap curve. These yield curves serve as input to models used to price other debt-related securities. The choice of curve will depend on the application. For our purposes, the source of this information is irrelevant as the mathematics is the same regardless of the source.

In the US debt market, for example, the US Treasury often plays this role. Debt incurred by the Federal government is fully backed by the government's taxing authority and is generally regarded as risk-free (except perhaps by S&P). We mean risk-free in the sense that all promised coupon and principal payments will be made on time with virtual certainty. One source of information comes from the "on-the-run" Treasury bills, notes, and bonds. These are bonds of specific maturities that have been recently issued. They are very liquid and actively traded producing good prices. Additionally, the on-the-runs typically trade near par value because they are recently issued and when issued their coupon rates were selected to be priced near par.

The US Treasury gathers data from the major banks each day about the current prices of the on-the-runs (see `www.treasury.gov`). Figure 4 shows this data on September 30, 2013. This yield curve has the typical positive slope, although the short end of the curve has historically low rates.

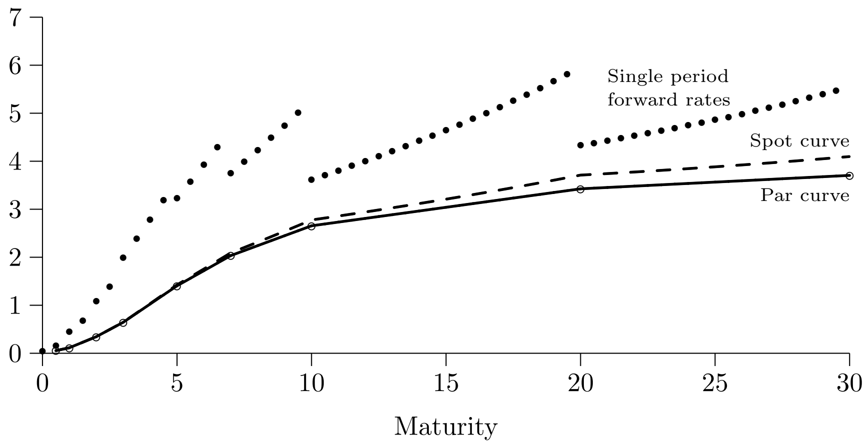
Figure 4. On-the-run Treasury Data on 9/30/2013.

Term	Yield (%)
6-Month	0.04
12-Month	0.10
2-Year	0.33
3-Year	0.63
5-Year	1.39
7-Year	2.02
10-Year	2.64
20-Year	3.41
30-Year	3.69

From Figure 4 we see that $c(.5) = .0004$, $c(1) = .0010$, $c(2) = .0033$, $c(3) = .0063$, $c(5) = .0139$, $c(7) = .0202$, $c(10) = .0264$, $c(20) = .0341$, and $c(30) = .0369$. But to calculate the discount function thereby enabling us to value a stream of cash flow we need to know $c(t)$ for *all* values of t from 0.5 to 30 in steps of 0.5 years.

Linear Interpolation. One approach to estimating par bond coupon rates for maturities not indicated by the on-the-runs is to interpolate. Figure 5, for example, shows the on-the-runs (as ∘s) together with a naive linear interpolation. In this scheme, for example, $c(25) = \frac{1}{2}c(20) + \frac{1}{2}c(30) = .0355$. With the value of $c(t)$ similarly established for all $t = 0.5, 1, 1.5, \ldots, 30$, it becomes possible compute the corresponding spot curve $z(t)$ for the same values of t and the single-period forward rates (here for $t = 0, 0.5, 1, 1.5, \ldots, 29.5$) via formulae developed in §1. These are also reported in Figure 5.

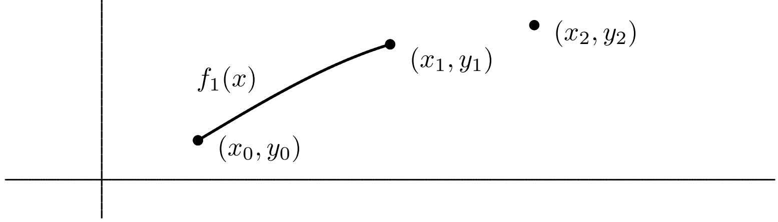
Figure 5. Linearly interpolated Treasury term structure on 9/30/2013.



Observe that the forward rates lie above the spot curve, which lies above the par curve, as expected in this positively sloped term structure. Perhaps the most glaring feature of Figure 5 is the several discontinuities in the forward rates $f(t)$ occurring at each on-the-run date. By virtue of our linear interpolation scheme, the slope of the par curve changes substantially at these dates. This, in turn, produces obvious changes in the slope of the spot curve. A look at (8) reveals that discontinuities in the slope of the spot curve produce discontinuities in the forward rates. In fact, since $z'(t)$ is multiplied by t in (8), the discontinuity in $f(t)$ caused by a given jump of $z'(t)$ gets magnified the bigger t is. While (8) is technically valid for instantaneous forward rates (rather than single-period rates with $\Delta t = 0.5$), it is the same phenomenon at work in this instance.

Cubic Spline Interpolation. Our goal here is to interpolate in such a fashion that $c'(t)$ experiences no discontinuities. A common approach to this is to use *cubic spline interpolation*. This approach actually assures that both $c'(t)$ and $c''(t)$ are continuous everywhere — at the on-the-run data points in particular. To see how this works, imagine we are trying to construct a function $f(x) = f_1(x)$ whose graph connects the two points (x_0, y_0) and (x_1, y_1) where $x_0 < x_1$ (see Figure 6).

Figure 6.

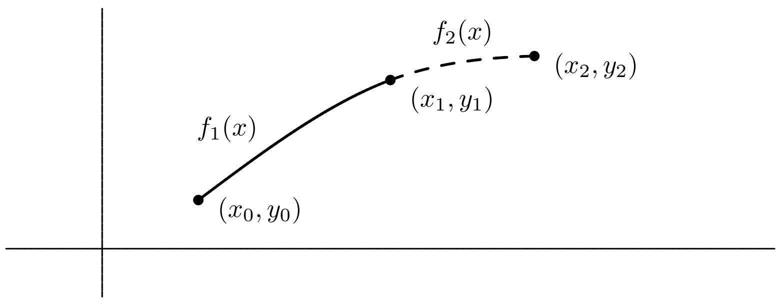


Suppose further that $f_1(x)$ is a cubic polynomial, so expanded about x_0 we have $f_1(x) = a_3(x - x_0)^3 + a_2(x - x_0)^2 + a_1(x - x_0) + a_0$, for some constants a_i. Additionally, say we wish that $f_1'(x_0) = m_0$ and $f_1''(x_0) = p_0$ for some pre-determined m_0 and p_0. The four conditions $f_1(x_0) = y_0$, $f_1(x_1) = y_1$, $f_1'(x_0) = m_0$, and $f_1''(x_0) = p_0$, are sufficient to fix the values of the four a_i. In particular, we find that:

$$\begin{aligned}
a_0 &= y_0 \\
a_1 &= m_0 \\
a_2 &= p_0/2 \quad \text{and} \\
a_3 &= \frac{y_1 - y_0}{(x_1 - x_0)^3} - \frac{m_0}{(x_1 - x_0)^2} - \frac{p_0}{2(x_1 - x_0)}.
\end{aligned}$$

Now imagine that we wish to extend $f(x)$ so that: (1) the graph goes through the point (x_2, y_2) where $x_1 < x_2$; (2) $f'(x)$ is continuous (in fact twice differentiable) at x_1; and (3) $f''(x)$ is continuous at x_1 (see Figure 7).

Figure 7.



To do this, simply construct a cubic polynomial $f_2(x)$ so that: $f_2(x_1) = y_1$, $f_2(x_2) = y_2$, $f_2'(x_1) = m_1$, and $f_2''(x_1) = p_1$, where $m_1 = f_1'(x_1)$ and $p_1 = f_1''(x_1)$. Since $f_1(x)$ has already been determined, m_1 and p_1 are readily computed and the four conditions just stated determine $f_2(x)$ in precisely the same manner that $f_1(x)$ was determined. Then put $f(x) = f_1(x)$ for $x \in [x_0, x_1]$ and $f(x) = f_2(x)$ for $x \in [x_1, x_2]$ (note that $f_1(x_1) = f_2(x_1)$).

Clearly this process can be continued indefinitely and the function $f(x)$ may be extended to pass through additional points (x_3, y_3), (x_4, y_4), etc., where $x_2 < x_3 < x_4 < \cdots < x_n$ in manner such that $f(x)$ is a piecewise cubic polynomial and $f'(x)$ and $f''(x)$ are continuous everywhere — at the x_i in particular. The resulting function is determined by the points $\{(x_i, y_i) : 0 \leq i \leq n\}$ and the initial choice of $m_0 = f'(x_0)$ and $p_0 = f''(x_0)$.

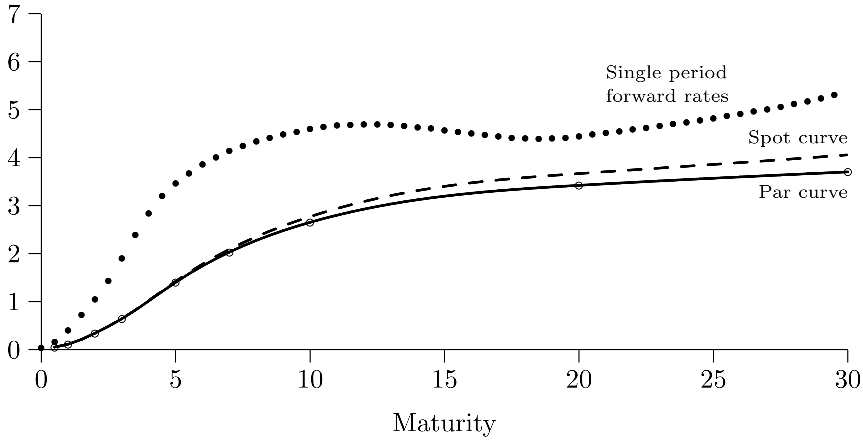
Let's see how this works for interpolating the term structure. Here, the points (x_i, y_i) are determined by the on-the-run data $(t, c(t))$ where $t = 0.5,\ 1, 2,\ 3,\ 5,\ 7,\ 10,\ 20,$ and 30. Any initial choice of m_0 and p_0 will produce a fit with continuous first and second derivatives. The actual selection of the initial m_0 and p_0 is more art than science: one seeks a fit that "looks reasonable." Remarkably, this happens when $p_0 = c''(.5) = 0$ and m_0 is chosen so that $c''(30) = 0$. The requirement that the second derivative is zero at both endpoints is called *natural boundary conditions*. We may solve for m_0 by some form of iteration — the graphs of the resulting c, z, and f are shown in Figure 8.

The reason that natural boundary conditions generally produce reasonable forward rates is that, as it turns out, this choice of m_0 and p_0 minimize the quantity

$$\textit{aggregate squared convexity} = \int_{0.5}^{30} c''(t)^2\, dt.$$

The motivation for minimizing this quantity is that large values of $|c''(t)|$ produce rapid changes in $c'(t)$ and hence rapid changes in $z'(t)$. Rapid changes in $z'(t)$ produce rapid changes in $f(t)$ via (8). It is these rapid swings in $f(t)$ that look unreasonable.

Figure 8. Treasury term structure on 9/30/2013 with cubic spline interpolation — natural boundary conditions.

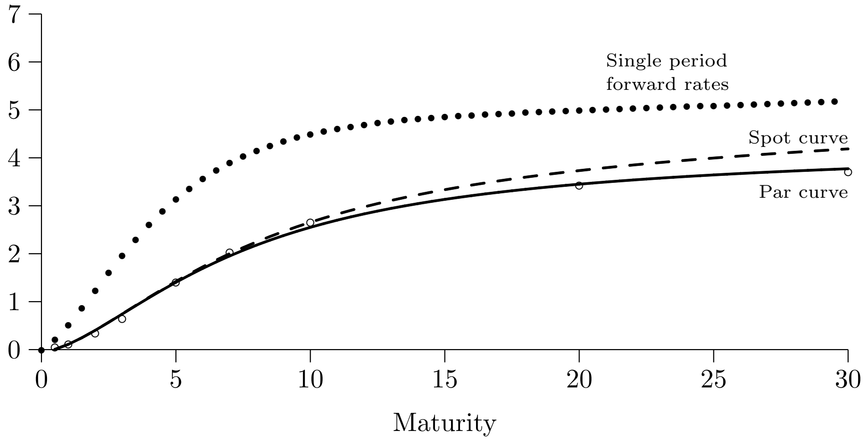


Parameterized Fits. Another approach is to assume that the data, here the par yield curve, follows some functional form described by a few parameters. The values of those parameters are then selected to produce a good overall fit of the observed data. We illustrate this general approach here by taking (9) as our functional form. The parameters A, B, C must be selected to explain the on-the-run yields as best as possible in some sense. Here, the "sense" we use is the minimization of the sum of the squared errors; we seek to minimize

$$SOS \;=\; \sum_{t\in T}\left[c(t) - (A\cdot \mathit{Shift}(t) + B\cdot \mathit{Twist}(t) + C\cdot \mathit{Hump}(t))\right]^2,$$

where the summation index set is $T = \{0.5, 1, 2, 3, 5, 7, 10, 20, 30\}$. This is natural here, because the determination of A, B, and C may be accomplished by simple regression techniques. The results of this process are shown in Figure 9. This approach produces a more stable variety of single-period forward rate curves. However, the fitted par curve does not precisely explain the observed data. For example, here the actual 30-year on-the-run par yield is 3.69% while the fitted 30-year yield is 3.76% — a difference of 7 basis points.

Figure 9. Treasury term structure on 9/30/2013 with best Nelson-Siegel fit.



Accompanying Code

The program `YieldMath.cpp` invites the user to create a par yield curve by specifying the curve's short rate, long rate, and hump, as defined in §2. It produces output files for viewing the resulting par and spot yield curves as well as the single-period forward rates. These may be viewed using the TeX program `Curves.tex`.

The programs `LinearFit.cpp`, `SplineFit.cpp` and `NelsonSiegelFit.cpp` implement these forms of interpolation for the Treasury yield curve on 9-30-2013. The resulting par, spot, and forward rate curves may be viewed with `Curves.tex`.