Kamakura Corporation
2222 Kalakaua Avenue
Suite 1400
Honolulu, Hawaii 96815, USA
phone 1 808 791 9888
facsimile 1 808 791 9898
www.kamakuraco.com

Advanced Financial Risk Management:

Tools and Techniques for Integrated Credit Risk, Market Risk, Liquidity Risk and Interest Rate Risk Management

April 2012

Donald R. van Deventer
Kenji Imai
Mark Mesler

Part 2: Risk Management Techniques for Interest Rate Analytics

Chapter 5:
Yield curve smoothing

In this chapter, we have two objectives. The first objective is to supply realistic zero coupon bond prices, like those in Chapter 4, to a sophisticated user for analysis. The second objective for this chapter is to illustrate how to build a yield curve which meets the criteria for "best" that are defined by the user. Given a specific set of criteria, what yield curve results? By the criteria defined, how do we measure "best" and compare the best yield curve we derive with alternatives? There is a most important set of comparisons we need to make—among N different sets of criteria for "best yield curve," which set of criteria is in fact "best," and what does "best" mean in this context? The vast majority of literature on yield curve smoothing omits the definition of "best" and the related derivation of what yield curve smoothing technique is consistent with the selected definition of best.

In this chapter we will report on the strengths and merits of this list of potential yield curve smoothing techniques:

Example A	Yield step function
Example B	Linear yield function
Example C	Linear forward rate function
Example D	Quadratic yield function
Example E	Quadratic forward rate function
Example F	Cubic yield function

Example G Cubic forward rate function

Example H Quartic forward rate function

These techniques are reviewed in great detail with worked examples in a series of blogs on the Kamakura Corporation web site www.kamakuraco.com:

Basic Building Blocks of Yield Curve Smoothing, Part 1, November 2, 2009.

Basic Building Blocks of Yield Curve Smoothing, Part 2: A Menu of Alternatives, November 17, 2009.

Basic Building Blocks of Yield Curve Smoothing, Part 3: Stepwise Constant Yields and Forwards versus Nelson-Siegel," November 18, 2009.

Basic Building Blocks of Yield Curve Smoothing, Part 4: Linear Yields and Forwards versus Nelson-Siegel, November 20, 2009.

Basic Building Blocks of Yield Curve Smoothing, Part 5: Linear Forward Rates and Related Yields versus Nelson-Siegel, November 30, 2009.

Basic Building Blocks of Yield Curve Smoothing, Part 6: Quadratic Yield Splines and Related Forwards versus Nelson-Siegel, December 3, 2009.

Basic Building Blocks of Yield Curve Smoothing, Part 7: Quadratic Forward Rate Splines and Related Yields versus Nelson-Siegel," December 8, 2009.

Basic Building Blocks of Yield Curve Smoothing, Part 8: Cubic Yield Splines and Related Forwards versus Nelson-Siegel, December 10, 2009.

Basic Building Blocks of Yield Curve Smoothing, Part 9: Cubic Forward Rate Splines and Related Yields versus Nelson-Siegel, December 30, 2009.

Basic Building Blocks of Yield Curve Smoothing, Part 10: Maximum Smoothness Forward Rates and Related Yields versus Nelson-Siegel, January 5, 2010.

Basic Building Blocks of Yield Curve Smoothing, Part 11: The Shimko Test for Measuring Accuracy of Smoothing Techniques," January 13, 2010.

Basic Building Blocks of Yield Curve Smoothing, Part 12: Smoothing with Bond Prices as Inputs," January 20, 2010.

In this chapter, space permits us to focus only on selected techniques, examples A, D, F and H. These examples include a yield step function, quadratic yield splines, cubic yield splines, and quartic forward rate (maximum smoothness) splines. Readers who are interested in a more detailed treatment of these topics are referred to the original blogs above.

For each of the techniques, we compare them with the well-known Nelson-Siegel (1987) yield curve approximation technique to show the stark contrasts between the results given the definition of "best" and what is produced by Nelson Siegel.

In this section, we talk about how the definition of "best" yield curve or forward rate curve and the constraints one imposes on the resulting yield curve implies the mathematical function that is "best." This is the right way to approach smoothing. The wrong way to approach smoothing is the exact opposite: choose a mathematical function from the infinite number of functions one could draw and argue qualitatively why your choice is the "right one." We discuss the definition of "best" yield curve and the constraints commonly placed on the smoothing process.

The list of potential yield curve smoothing methodologies given above is not arbitrary. Instead, the mathematical yield curve or forward rate smoothing formulation can be derived as "best" after going through the following steps to define what "best" means and what constraints should be imposed upon the yield curve function(s) in deriving the "best" curve subject to those constraints:

Step 1: Should the smoothed curves fit the observable data exactly?

1a. Yes

1b. No

Except in the case where the data is known from the outset to be bad, there is no reason for "yes" not to be the standard choice here. If one does chose yes, the Nelson-Siegel approach to yield curve smoothing will be eliminated as a candidate because it is not flexible enough to fit 5 or more data points perfectly except by accident. We discuss the problems of the commonly used but flawed Nelson-Siegel approach at the end of this chapter.

Step 2: Select the element of the yield curve and related curves for analysis

2a. Zero coupon yields

2b. Forward rates

2c. Continuous credit spreads

2d. Forward continuous credit spreads

Smoothing choices 2a and 2b, smoothing with respect to either yields or forward rates, should only be selected when smoothing a curve with no credit risk. If the issuer of the securities being analyzed has a non-zero default probability, smoothing should only be done with respect to credit spreads (the risky zero coupon yield minus the risk-free zero coupon yield of the same maturity) or forward credit spreads (the risky forward curve less the risk-free forward curve). See Chapter 17 on why this distinction is important.

Step 3: Define "best curve" in explicit mathematical terms

3a. Maximum smoothness

3b. Minimum length of curve

3c. Hybrid approach

The real objective of yield curve smoothing, broadly defined, is to estimate those points on the curve that are not visible in the marketplace with the maximum amount of realism. "Realism" in the original Adams and van Deventer (1994, summarized later in this chapter) paper on maximum smoothness forward rates was defined as "maximum smoothness" for two reasons. First, it was a lack of smoothness in the forward rate curves produced by other techniques that was the criterion for their rejection by many analysts. Second, "maximum smoothness" was viewed as consistent with the feeling that economic conditions rarely change in a non-smooth way and, even when they do (i.e. if the Central Bank manages short term rates in such a way that they jump in 0.125%

increments or multiples of that amount) the uncertainty about the *timing* of the jump makes future rate expectations smooth. The maximum smoothness yield curve or forward rate curve g(s) (where s is the years to maturity for the continuous yield or forward) is the one that minimizes this function Z:

$$Z = \int_{t_1}^{t_2} [g''(s)]^2 ds$$

There is another criterion that can be used for "best" curve, the curve that is the shortest in length. This again can be argued to be most realistic in the sense that one would not expect big swings in the yield, even if the swings are smooth. The following article on www.wikipedia.com explains how to calculate the length of a curve given the mathematical function that produced the curve:

http://en.wikipedia.org/wiki/Arc_length

The length s of a yield curve or forward rate curve described by function f(x) between maturities a and b is

$$s = \int_a^b \sqrt{1 + [f'(x)]^2}\, dx.$$

where f'(x) is the first derivative of the yield curve or forward rate curve. Some analysts have suggested using "tension splines" that take a hybrid approach to these two criteria for best. Leif Andersen proposes such an approach in the *Review of Derivatives Research* in 2007.

Once the mathematical criterion for "best" is selected, we move on to specifications on curve fitting that will add to the realism of the fitted curves.

Step 4: Is the curve constrained to be continuous?

4a. Yes

4b. No

Why is this question being asked? It can be shown that the smoothest way to connect N points on the yield curve is **NOT** to fit a single function, like the Nelson-Siegel function below, to the data.

This was proven in an appendix contributed by Oldrich Vasicek in the Adams and van Deventer paper, corrected in van Deventer and Imai Financial Risk Analytics, 1996, and reproduced in Appendix A. The nature of the functional form that is consistent with the criterion for best and the constraints imposed can be derived using the calculus of variations, as Vasicek does in his proof. Any other functional form will not be "best" by the method selected by the analyst.

Given this, we now need to decide whether or not we should constrain the yield curve or forward rate curve to "connect" in a continuous fashion where each of the line segments meet. It is not obvious that "yes" is the only answer to this question. As an example, Robert Jarrow and Yildiray Yildirim (2003) use a four-step piece-wise constant function for forward rates to model Treasury Inflation Protected Securities. Such a forward rate curve would not be continuous but it is both perfectly smooth (except where the four segments join) and shorter in length than any non-constant function. Most analysts, however, would answer "yes" in this step.

Step 5: Is the curve differentiable?

5a. Yes

5b. No

If one answers "yes" to this question, the first derivative of the two line segments that join at "knot point" N have to be equal at that point in time. If one answers "no," that allows the first derivatives to be different and the resulting "curve" will have kinks or elbows at each knot point, like you would get with linear yield curve smoothing—where each segment is linear and continuous over the full length of the curve but not differentiable at the points where the linear segments join. Most analysts would answer "yes" to this question.

Step 6: Is the curve twice differentiable?

6a. Yes

6b. No

This constraint means that the second derivatives of the line segments that join at maturity N have to be equal. If the answer to this question is "yes," the segments will be neither linear nor quadratic.

Step 7: Is the curve thrice differentiable?

7a. Yes

7b. No

One can still further constrain the curve so that, even for the third derivative, the line segments paste together smoothly.

Step 8: At the spot date, time 0, is the curve constrained?

8a. Yes, the first derivative of the curve is set to zero or a non-zero value x.

8b. Yes, the second derivative of the curve is set to zero or a non-zero value y.

8c. No

This constraint is typically imposed in order to allow an explicit solution for the parameters of the curve for each line segment. See Appendix A as an example of how this assumption is used. The choice of non-zero values in 8a or 8b was suggested by Tibor Janosi and Robert Jarrow in a 2002 paper.

Step 9: At the longest maturity for which the curve is derived, time T, is the curve constrained?

9a. Yes, the first derivative of the curve is set to zero or a non-zero value j at time T.

9b. Yes, the second derivative of the curve is set to zero or a non-zero value k at time T.

9c. No

Step 9 is taken for two reasons. First, it is often necessary for an explicit, unique set of parameters to be derived for each segment as in step 8. Second and more importantly, it may be important for realistic curve fitting. If the longest maturity is 100 years, for instance, most analyst would expect the yield curve to be flat (choice 9a=0), rather than rising or falling rapidly, at the 100 year point. If the longest maturity is only 2 years, this constraint would not be appropriate and one would probably make choice 9b. Again, the suggestion of non-zero values in 9a and 9b is useful in this case.

Once all of these choices have been made, both the functional form of the line segments and the parameters that are consistent with the data can be explicitly derived. The resulting forward rate curve or yield curve that is produced by this method has these attributes:

- Given the constraints imposed on the curve and the raw data, the curve is the "best" than can be draw consistent with the analyst's definition of "best"
- The data will be fit perfectly
- All constraints will be adhered to exactly

Example A: Stepwise Constant Yields and Forwards versus Nelson-Siegel

In this section, we choose a common definition of "best" yield curve or forward rate curve and a simple set of constraints. We derive from our definition of "best" and the related constraints the fact that the "best" yield curve in this case is stepwise constant yields and forward rates. We then show that even this simplest of specifications is more accurate and "better" by our definition that the popular but flawed Nelson-Siegel approach.

For each of our smoothing techniques to be comparable, we will use the same set of data throughout the series. We will limit the number of data points to six points and smooth the five intervals between them. In order to best demonstrate how the definition of best affects the quality of the answer, we choose a simple set of sample data with two "humps" in the yield data.

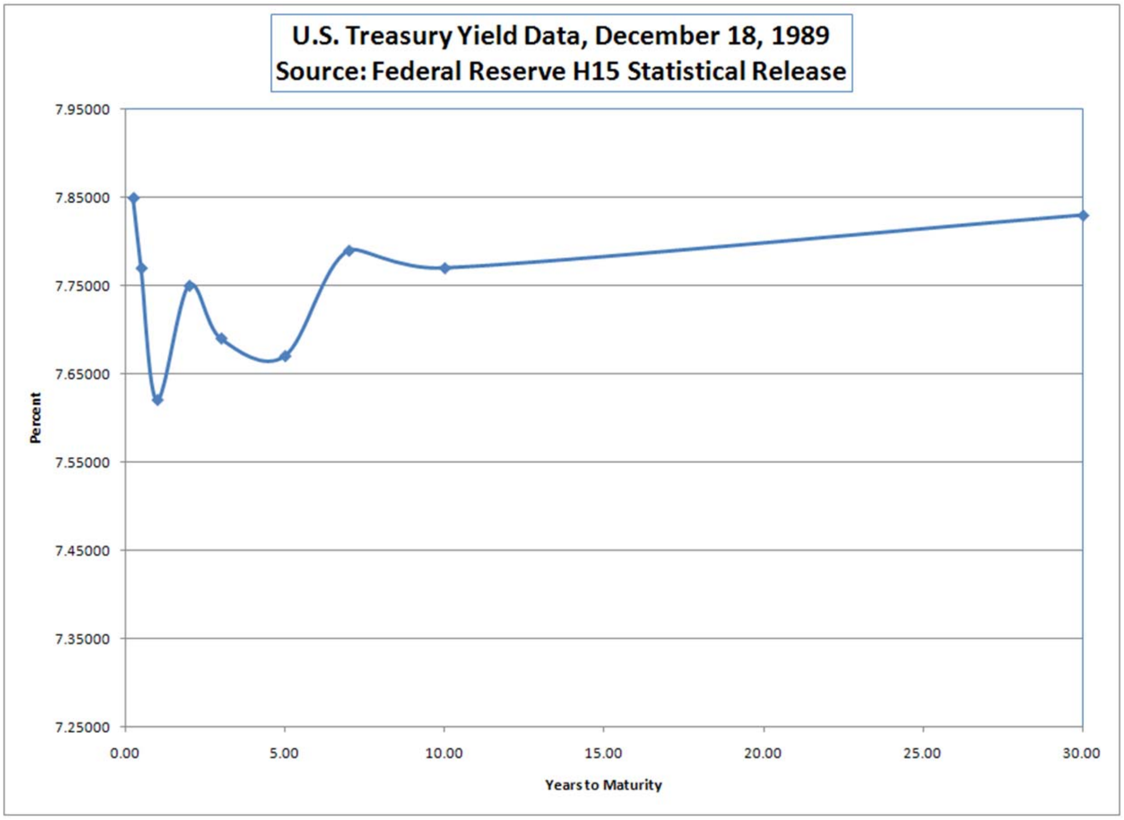
Figure 1

Actual Input Data		
Maturity in Years	Continuously Compounded Zero Coupon Yield	Zero Coupon Bond Price
0.000	4.000%	1.000000
0.250	4.750%	0.988195
1.000	4.500%	0.955997
3.000	5.500%	0.847894
5.000	5.250%	0.769126
10.000	6.500%	0.522046

Choosing a pattern of yields that falls on a perfectly straight line, for example, would lead a naïve analyst to the logical but naïve conclusion that a straight line is always the most accurate and "best" way to fit a yield curve. Such a conclusion would be nonsense. Similarly, restricting oneself to a yield curve functional form that allows only one "hump" would also lead one to simplistic conclusions about what is "best."

We recently reviewed 4,435 days of "on the run" U.S. Treasury yields reported by the Federal Reserve in its H15 statistical release at maturities (in years) of 0.25, 0.5, 1, 2, 3, 5, 7,10 and 30 years. In comparing same-day yield differences between maturities (i.e. the difference between yields at 0.25 and 0.5 years, 0.5 and 1 year, etc.), there were 813 days where there were at least two decreases in yields as maturities lengthened. For example, using the smoothing embedded in common spreadsheet software, here is the yield curve for December 18, 1989:

Figure 2



Therefore, it is very important that our sample data have a similarly realistic complexity to it or we will avoid challenging the various yield curve smoothing techniques sufficiently to see the difference in accuracy among them. It is important to note that the "yields" in these pictures are the simple "yields to maturity" on instruments which predominately are coupon bearing instruments. These yields should NEVER be smoothed directly as part of a risk management process. The graphs above are included only as an introduction to how complex yield curve shapes can be. When we want to do accurate smoothing, we are very careful to do it in the following way:

a. We work with raw bond net present values (price plus accrued interest), not "yields," because the standard yield to maturity calculation is riddled with inaccurate and inconsistent assumptions about interest rates over the life of the underlying bond.

b. We only apply smoothing to prices in the risk free bond and money markets. If the underlying issuer of fixed income securities has a non-zero default probability, we apply our smoothing to credit spreads or forward credit spreads in such a way that bond mispricing is minimized. We never smooth yields or prices of a risky issuer directly, because this ignores the underlying risk-free yield curve.

We will illustrate these points later in our series on the basic building blocks of yield curve smoothing. We now start with Example A, the simplest approach to yield curve smoothing.

From this point on in this chapter, unless otherwise noted, "yields" are always meant to be continuously compounded zero coupon bond yields and "forwards" are the continuous forward rates that are consistent with the yield curve. The first step in exploring a yield curve smoothing technique is to define our criterion for best and to specify what constraints we impose on the "best" technique to fit our desired trade-off between simplicity and realism. We answer the nine questions posed earlier in this chapter.

Step 1: Should the smoothed curves fit the observable data exactly?

1a. Yes. With only six data points at six different maturities, it would be a poor exercise in smoothing if we could not fit this data exactly. We note later that the flawed Nelson-Siegel function is unable to fit this data and fails our first test.

Step 2: Select the element of the yield curve and related curves for analysis

2a. Zero coupon yields is our choice. We find in the end that 2a and 2b are equivalent given our other answers to the following questions. If we were dealing with a credit-risky issuer of securities, we would have chosen 2c or 2d, but we have assumed our sample data is free of credit risk.

Step 3: Define "best curve" in explicit mathematical terms

3b. Minimum length of curve. This is the easiest definition of "best" to start with. We'll try it and show its implications. We now move on to specifications on curve fitting that represent our desired trade-off between realism and ease of calculation.

Step 4: Is the curve constrained to be continuous?

4b. No. By choosing no, we are allowing discontinuities in the yield curve.

Step 5: Is the curve differentiable?

5b. No. Since the answer we have chosen above, 4b, does not require the curve to be continuous, it will not be differentiable at every point along its length.

Step 6: Is the curve twice differentiable?

6b. No. For the same reason, the curve will not be twice differentiable at some points on the full length of the curve.

Step 7: Is the curve thrice differentiable?

7b. No. Again, the reason is due to our choice of 4b.

Step 8: At the spot date, time 0, is the curve constrained?

8c. No. For simplicity, we answer No to this question. We relax this assumption in later posts in this series.

Step 9: At the longest maturity for which the curve is derived, time T, is the curve constrained?

9c. No. Again, we choose No for simplicity and relax this assumption later in the blog.

Now that all of these choices have been made, both the functional form of the line segments and the parameters that are consistent with the data can be explicitly derived from our sample data.

Deriving the Form of the Yield Curve Implied by Example A

The key question in the list of 9 questions above is question 4. By our choice of 4b, we allow the "pieces" of the yield curve to be discontinuous. By virtue of our choices in questions 5-9, these yield curve pieces are also not subject to any constraints. All we have to do to get the "best" yield curve is to apply our criterion for "best"—the curve that produces the yield curve with shortest length.

The length of a straight line between two points on the yield curve has a length that is known, thanks to Pythagoras, who rarely gets the credit he deserves in the finance literature. The length of a straight segment of the yield curve that goes from maturity t_1 and yield y_1 to maturity t_2 and yield y_2 is the square root of (the square of [t_2-t_1] plus the square of [y_2-y_1]. As we noted above, the general formula for the length of any segment of the yield curve between maturity a and maturity b is given by this formula, which is the function one derives as the difference between t_2 and t_1 becomes infinitely small:

$$s = \int_a^b \sqrt{1 + [f'(x)]^2}\, dx.$$

where f'(x) is the first derivative of the line segment at each maturity point, say x. If the line segment happens to be straight, the segment can be described as

y=f(t)=mt+k

and the first derivative f'(t) is of course m. How can we make this line segment as short as possible? By making f'(t)=m as small as possible in absolute value, that is m=0. Very simply, we have DERIVED the fact that the yield curve segments which are "best" (which have the shortest length) are flat line segments. We are allowed to use flat line segments to fit the yield curve because our answer 4b does not require the segments to join each other in continuous fashion. The functional form of the "best" yield curve given our definition and constraints can be derived more elegantly using the calculus of variations as Oldrich Vasicek did in the proof of the maximum smoothness forward rate approach in Adams and van Deventer (1994), reproduced in Appendix A of this chapter.

Given that we have 6 data points, there are five intervals between points. We have taken advantage of our answer in 4b to treat the "sixth interval" from 0 years to maturity to 0 years to maturity as a separate segment. Given our original data, we know by inspection that our "sixth segment" has as a value the given value 4.000% for a maturity of zero.

In the rest of this book, we will repeatedly use these four relationships between zero coupon bond yields y, zero coupon bond prices p, and continuous forward rates f. The relationship between forward rates and zero coupon bond yields is given by the fourth relationship:

$$y(t) = \frac{-1}{t}\ln[p(t)]$$

$$p(t) = \exp[-\int_0^t f(s)ds]$$

$$p(t) = \exp[-y(t)t]$$

$$f(t) = y(t) + ty'(t)$$

Since the first derivative of the yield "curve" in each case is zero in this Example A, the forward rates are identical to the zero yields.

How did we do in terms of minimizing the length of the yield "curve" over its 10 year span? We know the length of a flat line segment from t_1 to t_2 is just t_2-t_1, so the total length of our discontinuous yield curve is

Length=(0-0)+(0.25-0)+(0.5-0.25)+(1-0.5)+(3-1)+(5-3)+(7-5)+(10-7)=10.000

We now compare our results for our "best" Example A yield curve and constraints to the popular but flawed Nelson-Siegel approach.

Fitting the Nelson-Siegel Approach to Sample Data

We now want to compare Example A, the model we derived from our definition of "best" and related constraints, to the Nelson-Siegel approach. The following table emphasizes the stark differences between even the basic Example A model and Nelson-Siegel:

Figure 3

	Yield Curve Smoothing Methodology	
		Example A
Name	Nelson-Siegel	Yield step function
1. Fit observed data exactly?	No	Yes
2. Element of analysis (yields or forwards)	Yields	Yields
3. Max smoothness or minimum length?	Neither	Minimum length
4. Continuous?	Yes	No
5. Differentiable?	Yes	No
6. Twice differentiable?	Yes	No
7. Thrice differentiable?	Yes	No
8. Time zero constraint?	No	No
9. Longest maturity time T constraint?	No	No

As the table notes, we know even before we start this exercise that the Nelson-Siegel function will not fit the actual market data we've assumed because there are more data points than there are Nelson-Siegel parameters AND the functional form of Nelson-Siegel is not flexible enough to handle the kind of actual U.S. Treasury data we reviewed earlier in this blog. The other thing that is important to note is that Nelson-Siegel will NEVER be superior to a function that has the same constraints and is derived from either (a) the minimum length criterion for best or (b) the maximum smoothness criterion.

We have explicitly chosen to answer “no” to questions 4-9 in Example A, while the answers for Nelson-Siegel are “yes.” This is not a virtue of Nelson-Siegel; it is just a difference in modeling assumptions. In subsequent installments of this blog, we will impose the constraints in questions 4-9 and we will find the results are still superior in every case to Nelson-Siegel.

In fitting Nelson-Siegel to our actual data, we have to make a choice of the function we are optimizing:

a. Minimize the sum of squared errors versus the actual zero coupon yields
b. Minimize the sum of squared errors versus the actual zero coupon bond prices

If we were using coupon bond price data, we would always optimize on the sum of squared pricing errors versus true net present value (price plus accrued interest) because legacy yield quotations are distorted by inaccurate embedded forward rate assumptions. In this case, however, all of the assumed inputs are on a zero coupon basis, and we have another issue to deal with. The zero coupon bond price at a maturity=0 is 1 for all yield values, so using the zero price at a maturity of zero for optimization is problematic. This means that we need to optimize in such a way that we minimize the sum of squared errors versus the zero coupon yields at all of the input maturities, including the zero point. We need to make 2 other adjustments before we can do this optimization using common spreadsheet software:

- We optimize versus the sum of squared errors in yields times 1 million in order to minimize the effect of rounding error and tolerance settings embedded in the optimization routine
- We note that, at maturity zero, the Nelson-Siegel function “blows up” because of a division by zero. Since y(0) is equal to the forward rate function at time zero f(0), we make that substitution to avoid dividing by zero.

That is, at the zero maturity point, instead of using the Nelson-Siegel yield function

$$y(t) = \alpha + (\beta + \gamma)\frac{\delta}{t}\left(1 - \exp\left(\frac{-t}{\delta}\right)\right) - \gamma\exp\left(\frac{-t}{\delta}\right)$$

we use the forward rate function:

$$f(t) = \alpha + \exp\left[\frac{-t}{\delta}\right]\left[\beta + \frac{\gamma}{\delta}t\right]$$

We now chose the values of alpha, beta, delta and gamma that minimize the sum of squared errors (times 1 million) in the actual and fitted zero yields. The results of that optimization are summarized in this table:

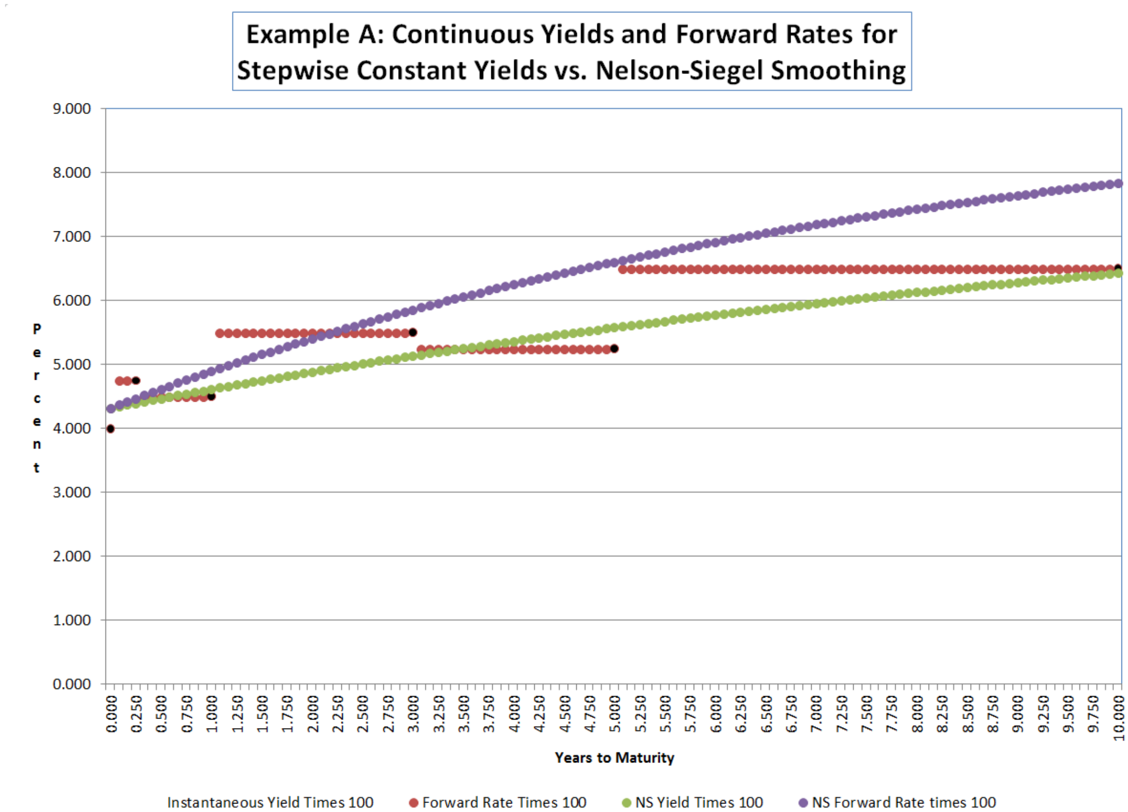
Figure 4

Base Data for All Examples			Nelson-Siegel Results			
Maturity	Instantaneous Yield	Zero Price	NS Yield	NS Zero Price	NS Squared Price Error	NS Squared Yield Error
0	4.000%	1.000000	4.323%	1.000000	na	0.0000104
0.25	4.750%	0.988195	4.399%	0.989062	0.0000008	0.0000123
1	4.500%	0.955997	4.618%	0.954872	0.0000013	0.0000014
3	5.500%	0.847894	5.140%	0.857110	0.0000849	0.0000130
5	5.250%	0.769126	5.584%	0.756384	0.0001624	0.0000112
10	6.500%	0.522046	6.436%	0.525396	0.0000112	0.0000004
Sum of Squared Errors times 1 Million					260.5679209	48.70568
Nelson Siegel Segment Length (After Optimization)				10.23137293		
Nelson Siegel Parameters (After Optimization)						
alpha	0.0926906					
beta	-0.0494584					
gamma	0.0000039					
delta	8.0575982					

After two successive iterations, we reach the best fitting parameter values for alpha, beta, delta and gamma. To those who have not used the Nelson-Siegel function before, the appallingly bad fit might come as a surprise. Even after the optimization, the errors in fitting zero yields are 32 basis points at the zero maturity, 36 basis points at 0.25, almost 12 basis points at 1 year, 36 basis points at 3 years, 33 basis points at five years, and 6 basis points at 10 years. The Nelson-Siegel formulation fails to meet the necessary condition for the consideration of a yield curve technique in this series: the technique MUST fit the observable data. Given this, why do people use the Nelson-Siegel technique? The only people who would use Nelson-Siegel are people who are willing to assume that the model is true and the market data is false. That's not a group likely to have a long career in risk management.

The graph below shows that our naïve model Example A, which has step-wise constant (and identical) forward rates and yields, fits the observable yields perfectly. The observable yields are plotted as black dots. The red dots represent the step-wise constant forwards and yields of Example A. The lower smooth line is the Nelson-Siegel zero coupon yield curve and the higher smooth line is Nelson-Siegel forward rate curve.

Figure 5



Now we pose a different question: given that we have defined the "best" yield curve as the one with the shortest length, how does the length of the Nelson-Siegel curve compare with Example A's length of 10 units?

There are two ways to answer this question. First, we could evaluate this integral using the first derivative of the Nelson-Siegel yield formula to evaluate the length of the curve, substituting y' for f' below:

$$s = \int_a^b \sqrt{1 + [f'(x)]^2}\, dx.$$

The second alternative is suggested by the link on length of an arc (given above): we can approximate the Nelson-Siegel length calculation by using a series of straight line segments and the insights of Pythagoras to evaluate length numerically. When we do this at monthly time intervals, the first 12 months' contributions to length are as follows:

	Zero Yield Curve	
Maturity	NS Yield Times 100	NS Segment Length
0.000	4.323222	0.000000
0.083	4.348709	0.087134
0.167	4.374023	0.087093
0.250	4.399164	0.087043
0.333	4.424133	0.086994
0.417	4.448931	0.086945
0.500	4.473558	0.086896
0.583	4.498018	0.086849
0.667	4.522311	0.086802
0.750	4.546437	0.086756
0.833	4.570399	0.086710
0.917	4.594198	0.086665
1.000	4.617835	0.086621

We sum up the lengths of each segment, 120 months in total, to get a total line length of 10.2314, compared to a length of 10.0000 for Example A's derived "best" curve, a step function of forward rates and yields. Note that the length of the curve, when the segments are not flat, depends on whether yields are displayed in percent (4.00) or decimal (0.04). To make the differences in length more obvious, our length calculations are based on yields in percent.

In this section, we have accomplished three things:

a. We have shown that the functional form used for yield curve fitting can and should be derived from a mathematical definition of "best" rather than being asserted for qualitative reasons
b. We have shown that the Nelson-Siegel yield curve fitting technique fails to fit yield data with characteristics often found in the U.S. Treasury bond market
c. We have shown that the Nelson-Siegel technique is inferior to the step-wise constant forward rates and yields that were derived from Example A's specifications: that "best" means the shortest length and that a continuous yield function is not required, consistent with the Jarrow and Yildirim paper above

Example D: Quadratic Yield Splines and Related Forward Rates

In this section, we turn to Example D, in which we seek to create smoother yield and forward rate functions by requiring that the yield curve be continuous and where the first derivatives of the curve segments we fit be equal at the knot points where the segments are joined together. We use a quadratic spline of yields to achieve this objective, and we optimize to produce the "maximum tension/minimum length" yields and forwards consistent with the quadratic splines. Finally, we compare the results to the popular but flawed Nelson-Siegel approach and gain still more insights on how to further improve the realism of our smoothing techniques.

In this section, we make two more modifications in our answers to the nine questions above and derive, not assert, the best yield curve consistent with the definition of "best" given the constraints we impose.

Step 1: Should the smoothed curves fit the observable data exactly?

1a. Yes. Our answer is unchanged. With only six data points at six different maturities, a technique which cannot fit this simple data (like Nelson-Siegel cannot) is too simplistic for practical use.

Step 2: Select the element of the yield curve and related curves for analysis

2a. Zero coupon yields is our choice for Example D. We continue to observe that we would never choose 2a or 2b to smooth a curve where the underlying securities issuer is subject to default risk. In that case, we would make the choices in either 2c or 2d. We do that later in this series.

Step 3: Define "best curve" in explicit mathematical terms

3b. Minimum length of curve. We continue with this criterion for best.

Step 4: Is the curve constrained to be continuous?

4b. Yes. We now insist on continuous yields and see what this implies for forward rates.

Step 5: Is the curve differentiable?

5a. Yes. This is another major change in Example D. We seek to create smoother yields and forward rates by requiring that the first derivatives of two curve segments be equal at the knot point where they meet.

Step 6: Is the curve twice differentiable?

6b. No. The curve will not be twice differentiable at some points on the full length of the curve given this answer, and we evaluate whether this assumption should be changed given the results we find.

Step 7: Is the curve thrice differentiable?

7b. No. The reason is due to our choice of 6b.

Step 8: At the spot date, time 0, is the curve constrained?

8c. No. For simplicity, we again answer No to this question. We wait until later in this chapter to explore this option.

Step 9: At the longest maturity for which the curve is derived, time T, is the curve constrained?

9a. Yes. This is the third major change in Example D. As we explain below, our other constraints and our definition of "best" put us in a position where the parameters of the best yield curve are not unique without one more constraint. We impose this constraint to obtain unique coefficients and then optimize the parameter used in this constraint to achieve "the best" yield curve.

Now that all of these choices have been made, both the functional form of the line segments for forward rates and the parameters that are consistent with the data can be explicitly derived from our sample data.

Deriving the Form of the Yield Curve Implied by Example D

Our data set has observable yield data at maturities of 0, 0.25 years, 1, 3, 5 and 10 years. For each segment of the yield curve, we have two constraints per segment (that the segment produces the market yields at the left side and right side of the segment). In addition to this, our imposition of the requirement that the full curve be differentiable means that there are other constraints. We have 6 knot points in total and 4 interior knot points (0.25, 1, 3, and 5) where we require the adjacent yield curve segments to produce the same first derivative of yields at a given knot point. That means we have 14 constraints (not including the constraint in step 9), but a linear yield spline would only have 2 parameters per segment and 10 parameters in total.

That means that we need to step up to a functional form for each segment that has more parameters. The following two links to Wikipedia articles nicely summarize the progression from linear to quadratic to cubic splines:

http://en.wikipedia.org/wiki/Spline_interpolation

http://en.wikipedia.org/wiki/Spline_%28mathematics%29

Our objective in these blogs is not to reproduce the intellectual history of splines, because it's too voluminous, not to mention too hard. Indeed, a google search yesterday on "quadratic splines" produced applications to automotive design (where much of the original work on splines was done), rocket science (really), engineering, computer graphics, and astronomy. Instead, we want to focus on the unique aspect of using splines and other smoothing techniques in finance, where both yields and related forwards are important. The considerations for "best" in finance are dramatically different than they would for the designer of the hood for a new model of the Maserati.

In the current Example D, we are still defining the "best" yield curve as the one with the shortest possible length or maximum "tension." When we turn to cubic splines, we will discuss the proof that cubic splines produce the "smoothest" curve. In this example, we will numerically force the quadratic splines we derive to be the shortest possible quadratic splines that are consistent with the constraints we're imposing in our continuing search for greater financial realism.

We can derive the shortest possible quadratic spline implementation in two ways. First, we could evaluate the function s given in Step 3, since for the first time the full yield curve will be continuously differentiable, even at the knot points. Alternatively, we can again measure the length the curve by approximating the curve with a series of line segments, thanks to Pythagoras, whom we invoked in Example A.

We chose the latter approach for consistency with our earlier examples. We break the 10 year yield curve into 120 monthly segments to numerically measure length. We now derive the set of five quadratic spline coefficients that are consistent with the constraints we have imposed.

Each yield curve segment has the quadratic form

$$y_i(t) = a_i + b_{i1}t + b_{i2}t^2$$

The subscript i refers to the segment number. The segment from 0 to 0.25 years is segment 1, the segment from 0.25 years to 1 year is segment 2, and so on. As mentioned above, we have 10 constraints that require the left hand side and the right hand of each segment to equal the actual yield at that maturity, which we call y*. For the jth line segment, we have two constraints, one at the left hand side of the segment, where the years to maturity are t_j, and one at the right hand side of the segment t_{j+1}.

$$y^*(t_j) = a_j + b_{j1}t_j + b_{j2}t_j^2$$
$$y^*(t_{j+1}) = a_j + b_{j1}t_{j+1} + b_{j2}t_{j+1}^2$$

In addition we have four constraints at the interior knot points. At each of these knot points, the first derivatives of the two segments that join at that point must be equal:

$$b_{j1} + 2b_{j2}t_{j+1} = b_{j+1,1} + 2b_{j+1,2}t_{j+1}$$

When we solve for the coefficients, we will rearrange these four constraints in this manner:

$$b_{j1} + 2b_{j2}t_{j+1} - b_{j+1,1} - 2b_{j+1,2}t_{j+1} = 0$$

Our last constraint is imposed at the right hand side of the yield curve, time T=10 years. We will discuss alternative specifications later, but for now we assume that we want to constrain the first derivative of the yield curve at T=10 to take a specific value x:

$$b_{j1} + 2b_{j2}T = x$$

In the equation above, j=5, the fifth line segment, and T=10. The most commonly used value for x is zero, the constraint that the yield curve be flat where T=10. Our initial implementation will use x=0.

In matrix form, our constraints look like this:

Figure 6

Coefficient Matrix for Quadratic Yield Curve Line Segments

Coefficient of what parameter?

Equation Number	a1	b11	b12	a2	b21	b22	a3	b31	b32	a4	b41	b42	a5	b51	b52	Coefficient Vector		y Vector
1	1	0.000	0.000	0	0	0	0	0	0	0	0	0	0	0	0	a1		4.000%
2	1	0.25	0.0625	0	0	0	0	0	0	0	0	0	0	0	0	b11		4.750%
3	0	0	0	1	0.250	0.063	0	0	0	0	0	0	0	0	0	b12		4.750%
4	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	a2		4.500%
5	0	0	0	0	0	0	1	1.000	1.000	0	0	0	0	0	0	b21		4.500%
6	0	0	0	0	0	0	1	3	9	0	0	0	0	0	0	b22		5.500%
7	0	0	0	0	0	0	0	0	0	1	3.000	9.000	0	0	0	a3		5.500%
8	0	0	0	0	0	0	0	0	0	1	5	25	0	0	0	b31	=	5.250%
9	0	0	0	0	0	0	0	0	0	0	0	0	1	5.000	25.000	b32		5.250%
10	0	0	0	0	0	0	0	0	0	0	0	0	1	10	100	a4		6.500%
11	0	1	0.5	0	-1	-0.5	0	0	0	0	0	0	0	0	0.000	b41		0.000%
12	0	0	0	0	1	2	0	-1	-2	0	0	0	0	0	0.000	b42		0.000%
13	0	0	0	0	0	0	0	1	6	0	-1	-6	0	0	0.000	a5		0.000%
14	0	0	0	0	0	0	0	0	0	0	1	10	0	-1	-10	b51		0.000%
15	0	0	0	0	0	0	0	0	0	0	0	0	0	1	20.000	b52		0.000%

Note that it is the last element of the "y Vector" matrix where we have set the constraint that the first derivative of the yield curve be zero at 10 years. When we invert the coefficient matrix, we get the following result:

Figure 7

Inverted Coefficient Matrix

Row	a1	b11	b12	a2	b21	b22	a3	b31	b32	a4	b41	b42	a5	b51	b52
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	-8	8	2.666667	-2.66667	-1	1	1	-1	-0.4	0.4	-1	1	-1	1	-1
3	16	-16	-10.6667	10.66667	4	-4	-4	4	1.6	-1.6	4	-4	4	-4	4
4	0	0	1.777778	-0.77778	-0.33333	0.333333	0.333333	-0.33333	-0.13333	0.133333	0	0.333333	-0.33333	0.333333	-0.33333
5	0	0	-3.55556	3.555556	1.666667	-1.66667	-1.66667	1.666667	0.666667	-0.66667	0	-1.66667	1.666667	-1.66667	1.666667
6	0	0	1.777778	-1.77778	-1.33333	1.333333	1.333333	-1.33333	-0.53333	0.533333	0	1.333333	-1.33333	1.333333	-1.33333
7	0	0	0	0	2.25	-1.25	-1.5	1.5	0.6	-0.6	0	0	1.5	-1.5	1.5
8	0	0	0	0	-1.5	1.5	2	-2	-0.8	0.8	0	0	-2	2	-2
9	0	0	0	0	0.25	-0.25	-0.5	0.5	0.2	-0.2	0	0	0.5	-0.5	0.5
10	0	0	0	0	0	0	6.25	-5.25	-3	3	0	0	0	7.5	-7.5
11	0	0	0	0	0	0	-2.5	2.5	1.6	-1.6	0	0	0	-4	4
12	0	0	0	0	0	0	0.25	-0.25	-0.2	0.2	0	0	0	0.5	-0.5
13	0	0	0	0	0	0	0	0	4	-3	0	0	0	0	10
14	0	0	0	0	0	0	0	0	-0.8	0.8	0	0	0	0	-3
15	0	0	0	0	0	0	0	0	0.04	-0.04	0	0	0	0	0.2

We solve for the following coefficient values:

Coefficient Vector		Values
a1		0.04
b11		0.08416667
b12		-0.21666667
a2		0.05527778
b21		-0.03805556
b22		0.02777778
a3	=	0.02125
b31		0.03
b32		-0.00625
a4		0.105625
b41		-0.02625
b42		0.003125
a5		0.015
b51		0.01
b52		-0.0005

These coefficients give us the five quadratic functions that make up the yield curve. We use the fourth relationship below to derive the forward rate curve:

$$y(t) = \frac{-1}{t}\ln[p(t)]$$

$$p(t) = \exp[-\int_0^t f(s)ds]$$

$$p(t) = \exp[-y(t)t]$$

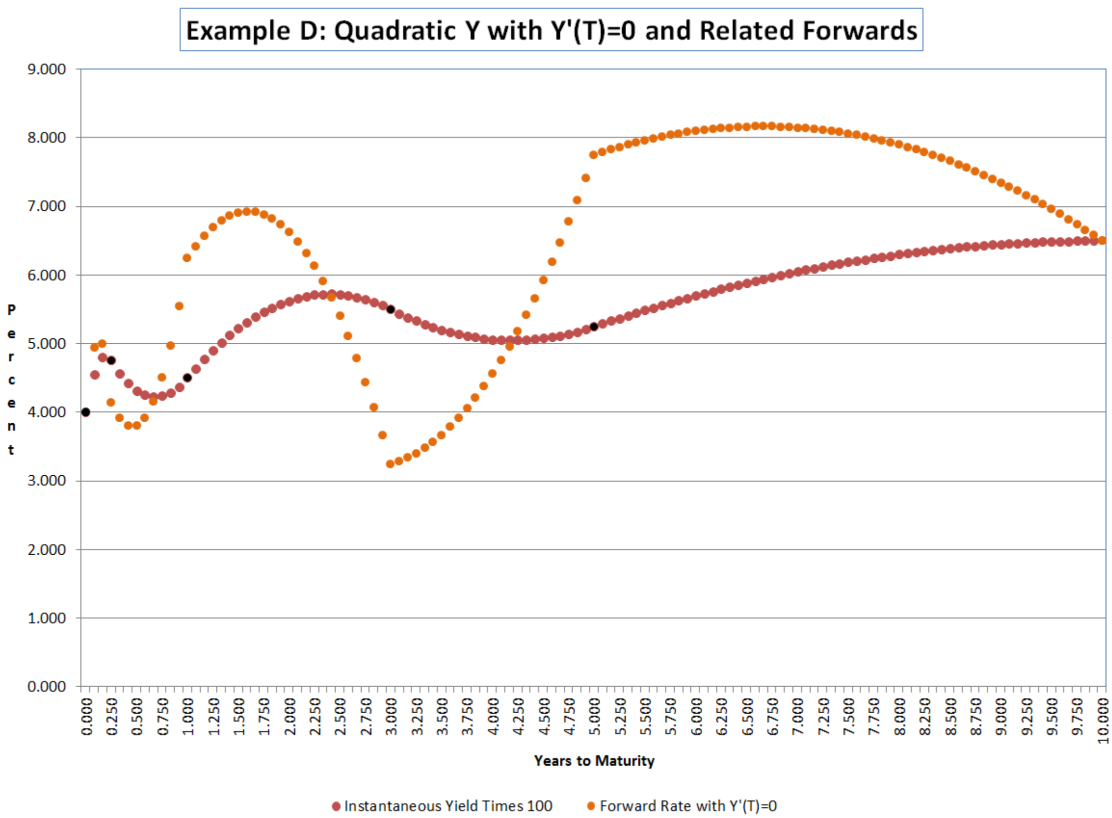
$$f(t) = y(t) + ty'(t)$$

The forward rate function for any segment j is

$$f_j(t) = a_j + 2b_{j1}t + 3b_{j2}t^2$$

We then plot the yield and forward rate curves to see whether or not our definition of "best yield curve" and the related constraints we imposed produced a realistic yield curve and forward rate curve:

Figure 8



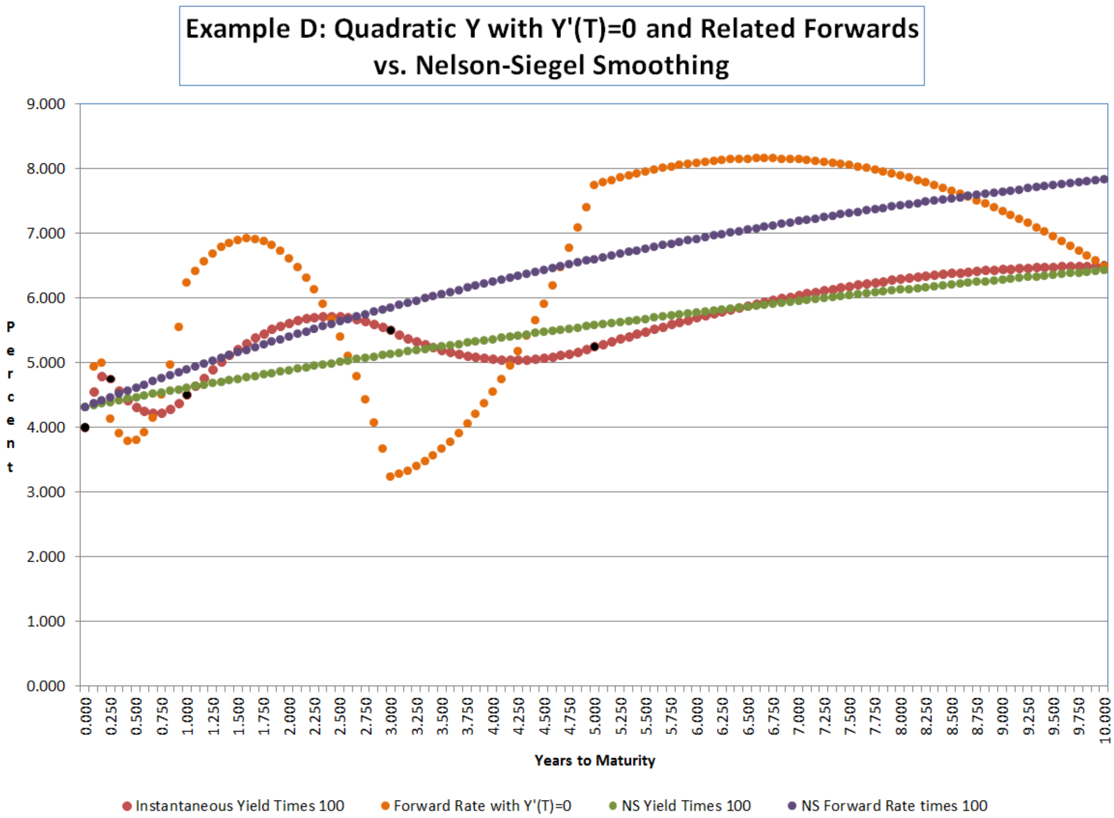
The answer for Example D so far is mixed. The yield curve itself fits the data perfectly, our standard "necessary" condition. The yield curve is also reasonably smooth and looks quite plausible given where the observable yields (black dots) are. The bad news, as usual, relates to forward rates. As in prior examples, it is much easier to spot flaws in the assumptions when looking at forward rates than it is when looking at yields alone. This is one of the reasons why "best practice" in the financial application of splines is so different than it is in computer graphics or automotive design, where there is no equivalent of the forward rate curve.

The lack of realism in the forward rates can be summarized as follows:

- The forward curve segments do not have first derivatives that are equal at the knot points where the forward segments meet; the forward curve has serious "kinks" in it at the knot point
- The movements of the forward curve are "not smooth" by any definition

It's possible that our 15th constraint, $y'(10)=0$, should be changed to some other value $y'(10)=x$. We explore that below. Before doing so, however, we overlay the yield and forward curves on the best fitting (but not perfectly fitting) Nelson-Siegel curve:

Figure 9



The graph makes two key points. First of all, the Nelson-Siegel curves are simply wrong and do not pass through the actual data points (in black). We know that already. Why, then, does anyone use Nelson-Siegel? For some, the smoothness of the forward curves is so seductive that the fact that valuation using the curve is wrong seems not to concern these analysts.

We clearly need to do better than a quadratic yield spline to get the smooth forward rates that we want. Let's first ask the question "Can we improve the quadratic spline approach by optimizing the value of x in the 15th constraint, $y'(10)=x$?" This insight comes from a working paper by Tibor Janosi and Robert A. Jarrow in 2002, who noted that this constraint and others like it could have important implications for the "best" curve. Our first implementation was $x=0$, a flat yield curve on the long end of the curve. We now look at alternative values for x.

There are two criteria for "best" that we could use given our statement above that the "best" curve is the one that is the shortest, the one with maximum tension. We could choose the value of x that produces the shortest yield curve or the value of x that produces the shortest forward rate curve. We do both.

If we use common spreadsheet software and iterate x so that the length (calculated numerically at 120 monthly intervals) of the yield curve is minimized, we get this set of coefficients:

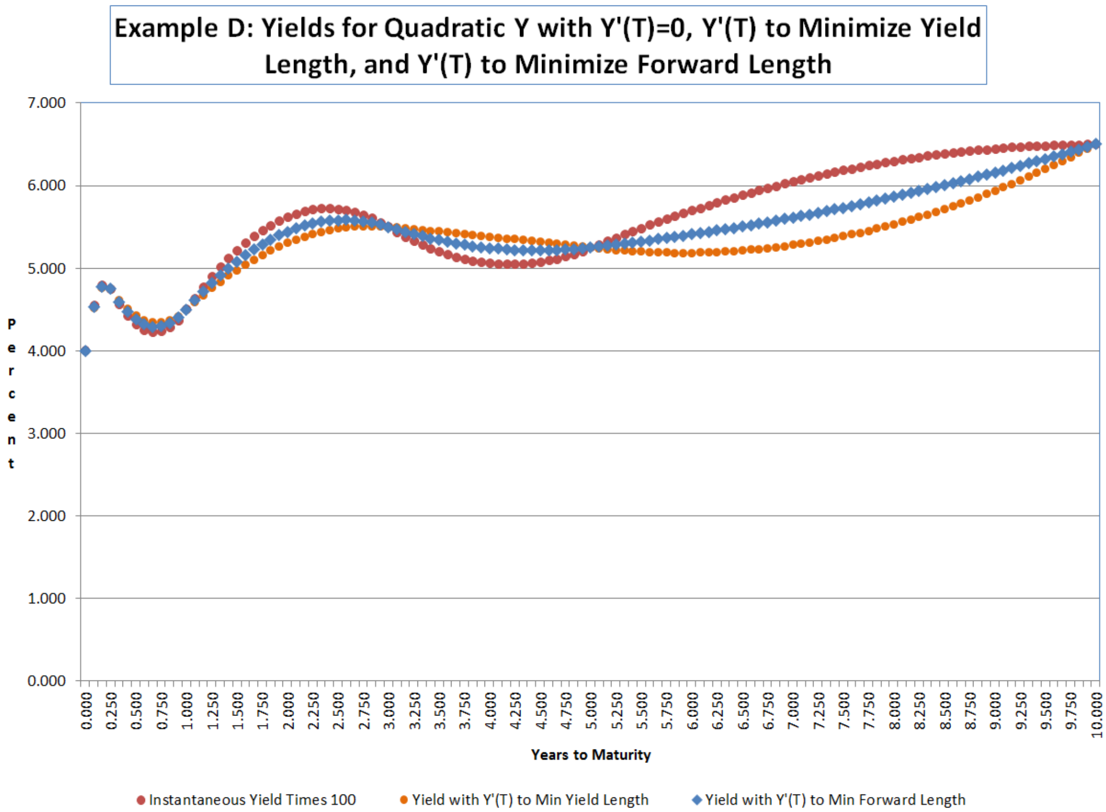
Coefficient Vector		Values
a1		0.04
b11		0.07776339
b12		-0.19105355
a2		0.05314335
b21		-0.02738342
b22		0.01924007
a3	=	0.03085492
b31		0.01719344
b32		-0.00304836
a4		0.05760041
b41		-0.00063689
b42		-7.6639E-05
a5		0.07903279
b51		-0.00920984
b52		0.00078066

If we iterate x so that the length of the forward rate curve is minimized, we get still another set of coefficients:

Coefficient Vector		Values
a1		0.04
b11		0.080529
b12		-0.20211601
a2		0.05406522
b21		-0.03199278
b22		0.02292756
a3	=	0.0267065
b31		0.02272467
b32		-0.00443117
a4		0.07834252
b41		-0.01169935
b42		0.00130617
a5		0.05137664
b51		-0.00091299
b52		0.00022753

The value of x that produces the shortest yield curve is 0.00640. The value of x that produces the shortest forward rate curve is 0.00364. We can compare these three variations on the yield curve produced by the quadratic spline approach:

Figure 10



Our "base case" is the one which is highest on the right hand side of the graph, and the flatness of the yield curve on the right hand side of the curve is readily apparent. Relaxing that constraint to produce the shortest yield curve produces the curve which is lowest on the right hand side of the graph, where the yield curve is rising sharply at the 10 year point. The shortest forward rate curve produces the yield curve in the middle. When we look at forward rates associated with these yield curves, we get an improvement. Nonetheless, it is clear that we can remove the kinks in the forward rate curves like those seen above and generate more realistic movements. For that reason we move on to the next example.

Example F: Cubic Yield Splines and Related Forwards

In this section, we make an important change in our definition of the "best yield curve." In this post we change the definition of "best" to be the yield curve with "maximum smoothness," which we define mathematically. Given the other constraints we impose for realism, we find that this definition of "best" implies that yield curve should be formed from a series of cubic polynomials. We then turn to Example F, in which we apply cubic splines to yields and derive the related forward rates.

We make a number of changes in our answers to the nine questions we have posed in each example so far.

Step 1: Should the smoothed curves fit the observable data exactly?

1a. Yes. Our answer is unchanged.

Step 2: Select the element of the yield curve and related curves for analysis

2a. Zero coupon yields is our choice for Example F. We continue to observe that we would never choose 2a or 2b to smooth a curve where the underlying securities issuer is subject to default risk. In that case, we would make the choices in either 2c or 2d. We do that later in this book.

Step 3: Define "best curve" in explicit mathematical terms

3a. Maximum smoothness. For the first time, we choose maximum smoothness as an intuitive criterion for best, since in past installments we have critiqued the results from some assumptions because of the lack of smoothness in either yields, forward rates, or both. Smoothness is defined as the variable z such that

$$z = \int_{0}^{T} g''(s)^2 ds$$

In this case, the function g(s) is defined as the yield curve. In order to evaluation z over the full maturity spectrum of the yield curve, the yield segments must be twice differentiable at each point, including the knot points. We want to minimize z over the full length of the yield curve. Our answers in Steps 4-6 make the analytical valuation of s possible just as when we evaluated length in Example D.

Step 4: Is the curve constrained to be continuous?

4b. Yes. We insist on continuous yields and see what this implies for forward rates.

Step 5: Is the curve differentiable?

5a. Yes. This is a major change imposed in Example D. We seek to take the spikes out of yields and forward rates by requiring that the first derivatives of two yield curve segments be equal at the knot point where they meet. This constraint also means that a linear "curve" isn't sufficiently rich to satisfy the constraints we've imposed.

Step 6: Is the curve twice differentiable?

6a. Yes. For the first time, we insist that the yield curve be twice differentiable everywhere along the curve, including the knot points. This would allow us to evaluate smoothness analytically if we wished to do so, and it insures that the yield curve segments will join in a visibly smooth way at the knot points.

Step 7: Is the curve thrice differentiable?

7b. No. We do not need to specify this constraint as yet.

Step 8: At the spot date, time 0, is the curve constrained?

8b. Yes. For the first time, we find it necessary to constrain the yield curve at its left hand side in order to derive a unique set of coefficients for each segment of the yield curve.

Step 9: At the longest maturity for which the curve is derived, time T, is the curve constrained?

9a. Yes. As in question 8, our other constraints and our definition of "best" put us in a position where the parameters of the best yield curve are not unique without one more constraint. We impose this constraint to obtain unique coefficients and then optimize the parameter used in this constraint to achieve "the best" yield curve. We explore choice 9b later in this chapter. The constraint here is imposed on the yield curve, not the forward rate curve.

Now that all of these choices have been made, both the functional form of the line segments for yields and the parameters that are consistent with the data can be explicitly derived from our sample data.

Deriving the Form of the Yield Curve Implied by Example F Assumptions

Our data set has observable yield data at maturities of 0, 0.25 years, 1, 3, 5 and 10 years. We have 6 knot points in total and 4 interior knot points (0.25, 1, 3, and 5) where we require (a) the adjacent yield curve segments to produce the same yield values, (b) the same first derivatives of yields, and (c) the same second derivatives of yields.

That means that we need to step up to a functional form for each yield curve segment that has more parameters than the quadratic segments we used in Examples D and E. The following two links to Wikipedia articles nicely summarize the progression from linear to quadratic to cubic splines, as we mentioned in Example D:

http://en.wikipedia.org/wiki/Spline_interpolation

http://en.wikipedia.org/wiki/Spline_%28mathematics%29

As it says in the first link, "Amongst all twice continuously differentiable functions, clamped and natural cubic splines yield the least oscillation about the function [g] which is interpolated." From our definition of "best," maximum smoothness, there is no question which mathematical function is the best: it is a spline of cubic polynomials. This can be proven using the calculus of variations, as was done in the case of maximum smoothness forward rates in the proof contributed by Oldrich Vasicek, reproduced in Appendix A of this chapter. As is the case throughout the series, we are not *asserting* that a functional form is best like advocates of the Nelson-Siegel approach. Instead, we are defining "best" mathematically, imposing constraints that are essential to realism in financial applications and DERIVING the fact that, in Example F, it is the cubic spline that is "best."

As in prior examples where we measured "tension" or length, we can measure the smoothness of a yield curve in two ways. First, we could explicitly evaluate the integral that defines the smoothness statistic z since our constraints for Example F will result in a yield curve that is twice differentiable over its full length. The other alternative is to use a discrete approximation to evaluation of the integral. This is attractive because it allows us to calculate smoothness for the other examples in this series for which the relevant functions are not twice differentiable. For that reason, we use a discrete measure of smoothness at 1/12 year maturity intervals over the 120 months of the yield curve we derive.

For any function x, the first difference between points i and i-1 is

$$x(i) - x(i-1)$$

The first derivative can be approximated by

$$\frac{x(i)-x(i-1)}{\Delta i}$$

The second derivative is

$$\frac{x'(i)-x'(i-1)}{\Delta i}=\frac{x(i)-2x(i-1)+x(i-2)}{(\Delta i)^2}$$

To evaluate smoothness numerically, we calculate the second derivative at 1/12 year intervals, square it, and sum over the full length of the yield curve. With this background out of the way, we now derive the maximum smoothness yield curve.

Each yield curve segment has the cubic form

$$y_i(t)=a_i+b_{i1}t+b_{i2}t^2+b_{i3}t^3$$

The subscript i refers to the segment number. The segment from 0 to 0.25 years is segment 1, the segment from 0.25 years to 1 year is segment 2, and so on. The first 10 constraints require the yield curve to be equal to the observable values of y at either the left or right hand side of that segment of the yield curve, so for each of the 5 segments there are two constraints like this where y* denotes observable data:

$$y^*(t_j)=y(t_j)=a_j+b_{j1}t_j+b_{j2}t_j^2+b_{j3}t_j^3$$

In addition we have four constraints that require the first derivatives of the yield curve segments to be equal at the four interior knot points:

$$y'_j(t_{j+1})=b_{j1}+2b_{j2}t_{j+1}+3b_{j3}t_{j+1}^2=y_{j+1}'(t_{j+1})=b_{j+1,1}+2b_{j+1,2}t_{j+1}+3b_{j+1,3}t_{j+1}^2$$

at the interior knot points. We rearrange these four constraints so that there is zero on the right hand side of the equation:

$$b_{j1}+2b_{j2}t_{j+1}+3b_{j3}t_{j+1}^2-b_{j+1,1}-2b_{j+1,2}t_{j+1}-3b_{j+1,3}t_{j+1}^2=0$$

At each of these knot points, the second derivatives must also be equal:

$$2b_{j2}+6b_{j3}t_{j+1}=2b_{j+1,2}+6b_{j+1,3}t_{j+1}$$

When we solve for the coefficients, we will rearrange these four constraints in this manner:

$$2b_{j2}+6b_{j3}t_{j+1}-2b_{j+1,2}-6b_{j+1,3}t_{j+1}=0$$

So far, we have 18 constraints to solve for 5x4=20 coefficients. Our 19th constraint was imposed in our answer to question 8, where we required that the second derivative of the yield curve was zero when maturity t_j is zero or some other value x_1:

$$2b_{j2} + 6b_{j3}t_j = x_1$$

Our last constraint is imposed by our answer to question 9 at the right hand side of the forward rate curve, time T=10 years. We will discuss the alternative specifications later in the chapter, but for now assume we want to constrain the first derivative of the yield curve at t_{j+1}=10 to take a specific value x_2, a value commonly selected to be zero:

$$y'_j(t_{j+1}) = b_{j1} + 2b_{j2}t_{j+1} + 3b_{j3}t_{j+1}^2 = x_2$$

Our initial implementation will use x_1 and x_2=0.

In matrix form, our constraints look like this:

Figure 11

Coefficient Matrix for Cubic Yield Curve Line Segments

Coefficient of what parameter?

Equation Number	a1	b11	b12	b13	a2	b21	b22	b23	a3	b31	b32	b33	a4	b41	b42	b43	a5	b51	b52	b53	Coefficient Vector		y Vector
1	1.00000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	a1		4.000%
2	1.00000	0.25000	0.06250	0.01563	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	b11		4.750%
3	0	0	0	0	1.00000	0.25000	0.06250	0.01563	0	0	0	0	0	0	0	0	0	0	0	0	b12		4.750%
4	0	0	0	0	1.00000	1.00000	1.00000	1.00000	0	0	0	0	0	0	0	0	0	0	0	0	b13		4.500%
5	0	0	0	0	0	0	0	0	1.00000	1.00000	1.00000	1.00000	0	0	0	0	0	0	0	0	a2		4.500%
6	0	0	0	0	0	0	0	0	1.00000	3.00000	9.00000	27.00000	0	0	0	0	0	0	0	0	b21		5.500%
7	0	0	0	0	0	0	0	0	0	0	0	0	1.00000	3.00000	9.00000	27.00000	0	0	0	0	b22		5.500%
8	0	0	0	0	0	0	0	0	0	0	0	0	1.00000	5.00000	25.00000	125.00000	0	0	0	0	b23	=	5.250%
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.00000	5.00000	25.00000	125.00000	a3		5.250%
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.00000	10.00000	100.00000	1000.00000	b31		6.500%
11	0	1.00000	0.50000	0.18750	0	-1.00000	-0.50000	-0.18750	0	0	0	0	0	0	0	0	0	0	0	0	b32		0.000%
12	0	0	0	0	0	1.00000	2.00000	3.00000	0	-1.00000	-2.00000	-3.00000	0	0	0	0	0	0	0	0	b33		0.000%
13	0	0	0	0	0	0	0	0	0	1.00000	6.00000	27.00000	0	-1.00000	-6.00000	-27.00000	0.00000	0.00000	0.00000	0.00000	a4		0.000%
14	0	0	0	0	0	0	0	0	0	0	0	0	0	1.00000	10.00000	75.00000	0.00000	-1.00000	-10.00000	-75.00000	b41		0.000%
15	0	0.00000	2.00000	1.50000	0.00000	0.00000	-2.00000	-1.50000	0	0	0	0	0	0	0	0	0	0	0	0	b42		0.000%
16	0	0	0	0	0	0	2.00000	6.00000	0	0	-2.00000	-6.00000	0	0	0	0	0	0	0	0	b43		0.000%
17	0	0	0	0	0	0	0	0	0	0	2.00000	18.00000	0	0	-2.00000	-18.00000	0	0	0	0	a5		0.000%
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.00000	30.00000	0	0	-2.00000	-30.00000	b51		0.000%
19	0	0	2.00000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	b52		0.000%
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.00000	20.00000	300.00000	b53		0.000%

Note that it is the last two elements of the "y Vector" matrix where we have set the constraints that the second derivative at time zero and the first derivative of the yield curve at 10 years be zero. We then invert the coefficient matrix and solve for the coefficients, which are shown here:

Coefficient		Values
a1		0.04
b11		0.03462
b12		-5.6E-19
b13		-0.07392
a2		0.038359
b21		0.054307
b22	=	-0.07875
b23		0.031082
a3		0.072955
b31		-0.04948
b32		0.025037
b33		-0.00351
a4		-0.06253
b41		0.086005
b42		-0.02012
b43		0.001505
a5		0.162438
b51		-0.04898
b52		0.006872
b53		-0.00029

These coefficients give us the five cubic functions that make up the yield curve. We use the fourth relationship below to derive the forward rate curve implied by the derived coefficients for the yield curve:

$$y(t) = \frac{-1}{t}\ln[p(t)]$$

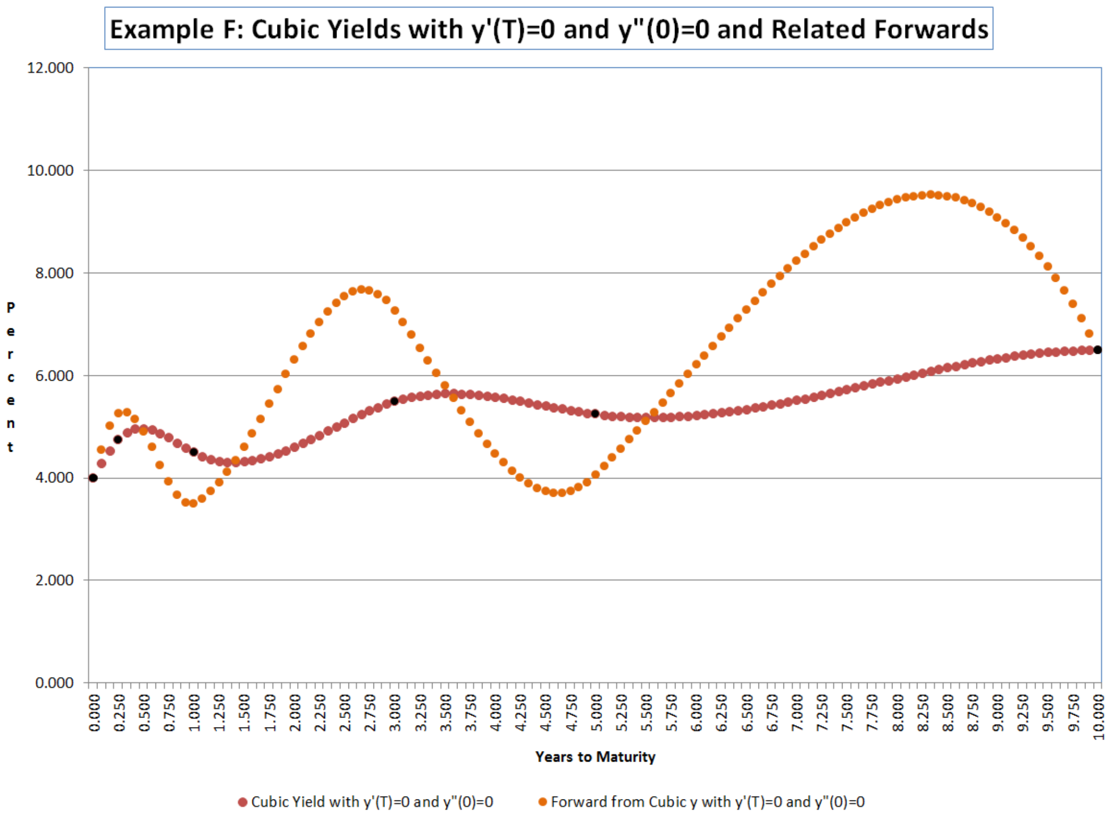
$$p(t) = \exp[-\int_0^t f(s)ds]$$

$$p(t) = \exp[-y(t)t]$$

$$f(t) = y(t) + ty'(t)$$

We can now plot the yield and forward rate curves to see whether or not our definition of “best yield curve” and the related constraints we imposed have produced a realistic yield curve and forward rate curve:

Figure 12



Here in Example F, the yield curve itself fits the data perfectly, our standard "necessary" condition. The yield curve is extremely smooth given the variations in it that are required by observable yields, noted as black dots. The forward rate curve is also reasonably smooth but it shows some more dramatic movements than expected. Is it possible to do even better by optimizing the discrete smoothness statistic by changing x_1 and x_2 in our 19th and 20th constraints? Yes. There are two criteria for "best" that we could use given our statement above that the "best" curve is the one that is the smoothest. We could choose the values of x_1 and x_2 that produce the smoothest yield curve or the values of x_1 and x_2 that produce the smoothest forward rate curve. We do both, as before.

If we use common spreadsheet software and iterate x_1 and x_2 so that the smoothness statistic z (calculated numerically at 120 monthly intervals) of the yield curve is minimized, we get this set of coefficients:

Coefficient Vector		Values
a1		0.04
b11		0.034517
b12		0.000644
b13		-0.07485
a2		0.038345
b21		0.054379
b22	=	-0.0788
b23		0.031078
a3		0.072892
b31		-0.04926
b32		0.024841
b33		-0.00347
a4		-0.05723
b41		0.080857
b42		-0.01853
b43		0.00135
a5		0.125941
b51		-0.02904
b52		0.003448
b53		-0.00012

If we iterate x_1 and x_2 so that the smoothness statistic z of the forward rate curve is minimized, we get still another set of coefficients:

Coefficient Vector		Values
a1		0.04
b11		0.041167
b12		-0.04209
b13		-0.0103
a2		0.039399
b21		0.048377
b22	=	-0.07093
b23		0.028158
a3		0.070823
b31		-0.04589
b32		0.023337
b33		-0.00327
a4		-0.05044
b41		0.075366
b42		-0.01708
b43		0.001226
a5		0.104982
b51		-0.01789
b52		0.001567
b53		-1.8E-05

The smoothest yield curve is produced by x_1=0.00129 and x_2=0.00533. The smoothest forward rate curve is produced by x_1=-0.08419 and x_2=0.00811. We will graph the differences that result when we review Example H, the maximum smoothness forward rate approach.

Example H: Maximum Smoothness Forward Rates and Related Yields

The authors wish to thank Keith Luna of Western Asset Management Company for helpful comments on this section of Chapter 5. Now, we present the final installment in yield curve and forward rate smoothing techniques before moving on to smoothing credit spreads in Chapter 17. We introduce the maximum smoothness forward rate technique introduced by Adams and van Deventer (1994) and corrected in van Deventer and Imai (1996), which we call Example H. We explain why a quartic function is needed to maximize smoothness of the forward rate function over the full length of the forward rate curve. Finally, we compare 23 different techniques for smoothing yields and forward rates that have been discussed in this series and show why the maximum smoothness forward rate approach is the best technique by multiple criteria.

Here are the answers to our normal 9 questions for Example H, quartic splines of forward rates:

Step 1: Should the smoothed curves fit the observable data exactly?

1a. Yes. Our answer is unchanged. Since the Nelson-Siegel specification cannot meet this minimum standard, as outlined in earlier blogs in this series, it is not included in the comparisons below.

Step 2: Select the element of the yield curve and related curves for analysis

2b. Forward rates is the choice for Example H. In addition to making this choice for empirically improving smoothness by changing parameter values, we add other "constraints" which improve our ability to maximize smoothness over the full length of the curve. We continue to observe that we would never choose 2a or 2b to smooth a curve where the underlying securities issuer is subject to default risk. In that case, we would make the choices in either 2c or 2d. We do that in Chapter 17.

Step 3: Define "best curve" in explicit mathematical terms

3a. Maximum smoothness. We unambiguously choose maximum smoothness as the primary criterion for "best." Within that class of functions that are maximum smoothness, we also take the opportunity to iterate parameter values to show which of the infinite number of "maximum smoothness" forward curves have "minimum length" and "maximum smoothness" for each set of constraints we impose. The two measures are very consistent.

As noted earlier in this chapter, we have critiqued the results from some yield curve smoothing techniques because of the lack of smoothness in either yields, forward rates, or both. Smoothness is defined as the variable z such that

$$z = \int_0^T g''(s)^2 ds$$

The smoothest possible function has the minimum z value. Since a straight line has a second derivative of zero, z is zero in that case and a straight line is perfectly smooth. In this case, the function g(s) is defined as the forward rate curve. In order to evaluate z over the full maturity spectrum of the forward curve, the forward rate segments must be at least twice differentiable at each point, including the knot points. We want to minimize z over the full length of the yield curve. Our answers in Steps 4-6 make the analytical valuation of z possible just as when we evaluated length in example D.

Recall that forward rates are related to zero coupon yields and bond prices by the same four relationships that we have been using throughout this book:

$$y(t) = \frac{-1}{t}\ln[p(t)]$$

$$p(t) = \exp[-\int_0^t f(s)ds]$$

$$p(t) = \exp[-y(t)t]$$

$$f(t) = y(t) + ty'(t)$$

Given that forward rates are in effect a “derivative” of zero yields and zero prices, Oldrich Vasicek uses the calculus of variations to prove that the smoothest forward rate function that can be derived is a thrice differentiable function, the quartic spline of forward rates. This proof, reproduced in Appendix A, was corrected in van Deventer and Imai (1996) from the original version in Adams and van Deventer (1994) thanks to helpful comments from Volf Frishling of the Commonwealth Bank of Australia and further improved by the insights of Kamakura’s Robert A. Jarrow. The choice of the quartic form is a ***derivation*** shown in the proof. Similarly, in order to achieve maximum smoothness, the same proof shows that the curves fitted together MUST be thrice differentiable over the entire length of the curve, which obviously includes the knot points. For that reason, we impose this constraint below.

Step 4: Is the curve constrained to be continuous?

4b. Yes. We insist on continuous forwards and see what this implies for yields.

Step 5: Is the curve differentiable?

5a. Yes. Again, this is the change first imposed in Example D. **Step 6: Is the curve twice differentiable?**

6a. Yes. We insist that the forward rate curve be twice differentiable everywhere along the curve, including the knot points. This would allow us to evaluate smoothness analytically if we wished to do so, and it insures that the forward rate curve segments will join in a visibly smooth way at the knot points.

Step 7: Is the curve thrice differentiable?

7b. Yes. For the first time, we impose this constraint because we have a mathematical proof that we will not obtain a unique maximum smoothness forward curve unless we do impose it.

Step 8: At the spot date, time 0, is the curve constrained?

8a and 8b. Both approaches will be used and compared. Like Example F, we find it necessary to constrain the forward rate curve at its left hand side and right hand side in order to derive a unique set of coefficients for each segment of the forward rate curve. We compare results using both answers 8a and 8b and reach conclusions about which is “best” for the sample data we are analyzing. The answers to question 8 can have a critical impact on the reasonableness of splines for forward rates in a financial context. In fact, in this example, we need to use 3 of the four constraints listed in questions 8 and 9.

Step 9: At the longest maturity for which the curve is derived, time T, is the curve constrained?

9a and 9b. Both approaches will be used and compared. For uniqueness of the parameters of the quartic forward rate segment coefficients, we again have to choose 3 of the four constraints in questions 8 and 9. We then optimize the parameters used in these constraints to achieve “the best” forward rate curve as suggested by Janosi and Jarrow (2002). The constraints here are again imposed on the forward rate curve, not the yield curve. We will use both “maximum smoothness” and “minimum length” as criterion for best, conditional on our choice of quartic forward curve segments and the other constraints we impose. We are reassured to see that the maximum smoothness forward rate technique produces results that are both the smoothest of any twice differentiable functions and the shortest length/maximum tension over the full length of the curve.

Deriving the Parameters of the Quartic Forward Rate Curves Implied by Example H Assumptions

Our data set has observable yield data at maturities of 0, 0.25 years, 1, 3, 5 and 10 years. We have 6 knot points in total and 4 interior knot points (0.25, 1, 3, and 5 years) where the curves that join must have equal first, second and third derivatives.

As in Example F in Part 8 of this series, we can measure the smoothness of a forward rate curve in two ways. First, we could explicitly evaluate the integral that defines the smoothness statistic z since our constraints for Example H will result in a forward rate curve that is twice differentiable over its full length. The other alternative is to use a discrete approximation to evaluation of the integral, like we did in Examples F and G. This is attractive because it allows us to calculate smoothness for the other examples in this series for which the relevant functions are not twice differentiable. For that reason, we use the same discrete measure of smoothness at 1/12 year maturity intervals over the 120 months of the forward curve and yield curves that we derive here.

Each forward rate curve segment has the quartic form

$$f_i(t) = c_i + d_{i1}t + d_{i2}t^2 + d_{i3}t^3 + d_{i4}t^4$$

The subscript i refers to the segment number. The segment from 0 to 0.25 years is segment 1, the segment from 0.25 years to 1 year is segment 2, and so on. The first constraint requires the first forward rate curve segment to be equal to the observable value of y at time zero since y(0)=f(0).

$$y^*(t_j) = f^*(t_j) = c_j + d_{j1}t_j + d_{j2}t_j^2 + d_{j3}t_j^3 + d_{j4}t_j^4$$

where for this first constraint $t_j = 0$. In addition we have four constraints that require the forward rate curves to be equal at the four interior knot points:

$$f_j(t_{j+1}) = c_j + d_{j1}t_{j+1} + d_{j2}t_{j+1}^2 + d_{j3}t_{j+1}^3 + d_{j4}t_{j+1}^4 = f_{j+1}(t_{j+1}) = c_{j+1} + d_{j+1,1}t_{j+1} + d_{j+1,2}t_{j+1}^2 + d_{j+1,3}t_{j+1}^3 + d_{j+1,4}t_{j+1}^4$$

at the interior knot points. We rearrange these four constraints, for j=1,4 like this:

$$c_j + d_{j1}t_{j+1} + d_{j2}t_{j+1}^2 + d_{j3}t_{j+1}^3 + d_{j4}t_{j+1}^4 - c_{j+1} - d_{j+1,1}t_{j+1} - d_{j+1,2}t_{j+1}^2 - d_{j+1,3}t_{j+1}^3 - d_{j+1,4}t_{j+1}^4 = 0$$

At each of these interior knot points, the first derivatives of the two forward rate segments that join at that point must be also be equal:

$$d_{j1} + 2d_{j2}t_{j+1} + 3d_{j3}t_{j+1}^2 + 4d_{j4}t_{j+1}^3 = d_{j+1,1} + 2d_{j+1,2}t_{j+1} + 3d_{j+1,3}t_{j+1}^2 + 4d_{j+1,4}t_{j+1}^3$$

When we solve for the coefficients, we will rearrange these four constraints in this manner:

$$d_{j1} + 2d_{j2}t_{j+1} + 3d_{j3}t_{j+1}^2 + 4d_{j4}t_{j+1}^3 - d_{j+1,1} - 2d_{j+1,2}t_{j+1} - 3d_{j+1,3}t_{j+1}^2 - 4d_{j+1}t_{j+1}^3 = 0$$

Next, we need to impose the constraint that the second derivatives of the joining forward rate curve segments are equal at each of the four interior knot points. This requires

$$2d_{j2} + 6d_{j3}t_{j+1} + 12d_{j4}t_{j+1}^2 = 2d_{j+1,2} + 6d_{j+1,3}t_{j+1} + 12d_{j+1,4}t_{j+1}^2$$

As usual we rearrange this as follows:

$$2d_{j2} + 6d_{j3}t_{j+1} + 12d_{j4}t_{j+1}^2 - 2d_{j+1,2} - 6d_{j+1,3}t_{j+1} - 12d_{j+1,4}t_{j+1}^2 = 0$$

Finally, the new constraint that the third derivatives be equal at each knot point requires that

$$6d_{j3} + 24d_{j4}t_{j+1} = 6d_{j+1,3} + 24d_{j+1,4}t_{j+1}$$

We arrange this constraint to read as follows:

$$6d_{j3} + 24d_{j4}t_{j+1} - 6d_{j+1,3} - 24d_{j+1,4}t_{j+1} = 0$$

So far, we have 1+4+4+4+4=17 constraints to solve for 5x5=25 coefficients. We have 5 more constraints that require that the forward curve segments produce observable zero coupon bond prices:

$$-\ln\left[\frac{p(t_k)}{p(t_j)}\right] = \left[\int_{t_j}^{t_k} f_j(s)ds\right] = c_j(t_k - t_s) + \frac{1}{2}d_{j1}(t_k^2 - t_j^2) + \frac{1}{3}d_{j2}(t_k^3 - t_j^3) + \frac{1}{4}d_{j3}(t_k^4 - t_j^4) + \frac{1}{5}d_{j4}(t_k^5 - t_j^5)$$

For our 23rd, 24th and 25th constraints, we can choose any 3 of the constraints in 8a, 8b, 9a, and 9b. We choose three combinations that are very common in financial applications where it is logical and realistic to expect the forward rate curve to be flat at the right hand side of the yield curve, f'(10)=0 in this example, where t_{j+1}=10.

$$f'_j(t_{j+1}) = c_{j1} + 2d_{j2}t_{j+1} + 3d_{j3}t_{j+1}^2 + 4d_{j4}t_{j+1}^3 = 0$$

Our 24th and 25th constraints come from setting the second derivatives at time t_{j+1}=0 to x_1 on the left hand side of the curve and x_2 on the right hand side of the curve where t_{j+1}=10:

$$f''(t_{j+1}) = 2d_{j2} + 6d_{j3}t_{j+1} + 12d_{j4}t_{j+1}^2 = x_k$$

Our initial implementation, which we label Example H-Qf1a (quadratic forwards 1a), will use x_1 and x_2=0. Our second implementation Example H-Qf1b uses the insights of Janosi and Jarrow (2002) to optimize x_1 and x_2 to minimize the length of the forward curve. The third implementation optimizes x_1 and x_2 to maximize the smoothness of the curve by minimizing the function z given above on a discrete 1/12 of a year basis.

In matrix form, with apologies to those readers with less than perfect vision, our constraints for example H-Qf1a look like this:

Figure 13

Coefficient Matrix for Quartic Forward Curve Line Segments

Coefficient of what parameter?

Equation Number	c1	d11	d12	d13	d14	c2	d21	d22	d23	d24	c3	d31	d32	d33	d34	c4	d41	d42	d43	d44	c5	d51	d52	d53	d54	Coefficient Vector		y Vector
1	1	0.000	0.000	0.000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	c1		4.000%
2	1	0.25	0.0625	0.01563	0.00391	-1	-0.25	-0.0625	-0.01563	-0.00391	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	d11		0
3	0	0	0	0	0	1	1	1	1	1	-1	-1	-1	-1	-1	0	0	0	0	0	0	0	0	0	0	d12		0
4	0	0	0	0	0	0	0	0	0	0	1	3	9	27	81	-1	-3	-9	-27	-81	0	0	0	0	0	d13		0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	5	25	125	625	-1	-5	-25	-125	-625	d14		0
6	0	1	0.5	0.1875	0.0625	0	-1	-0.5	-0.1875	-0.0625	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	c2		0
7	0	0	0	0	0	0	1	2	3	4	0	-1	-2	-3	-4	0	0	0	0	0	0	0	0	0	0	d21		0
8	0	0	0	0	0	0	0	0	0	0	0	1	6	27	108	0	-1	-6	-27	-108	0	0	0	0	0	d22	=	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	10	75	500	0	-1	-10	-75	-500	d23		0
10	0	0	2	1.5	0.75	0	0	-2	-1.5	-0.75	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	d24		0
11	0	0	0	0	0	0	0	2	6	12	0	0	-2	-6	-12	0	0	0	0	0	0	0	0	0	0	c3		0
12	0	0	0	0	0	0	0	0	0	0	0	0	2	18	108	0	0	-2	-18	-108	0	0	0	0	0	d31		0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	30	300	0	0	-2	-30	-300	d32		0
14	0	0	0	6	6	0	0	0	-6	-6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	d33		0
15	0	0	0	0	0	0	0	0	6	24	0	0	0	-6	-24	0	0	0	0	0	0	0	0	0	0	d34		0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	6	72	0	0	0	-6	-72	0	0	0	0	0	c4		0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	120	0	0	0	-6	-120	d41		0
18	0.25	0.03125	0.00521	0.00098	0.0002	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	d42		0.01188
19	0	0	0	0	0	0.75	0.46875	0.32813	0.24902	0.1998	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	d43		0.03313
20	0	0	0	0	0	0	0	0	0	0	2	4	8.66667	20	48.4	0	0	0	0	0	0	0	0	0	0	d44		0.12000
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	8	32.6667	136	576.4	0	0	0	0	0	c5		0.09750
22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	37.5	291.667	2343.75	19375	d51		0.38750
23	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	d52		0.00000
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	20	300	4000	d53		0.00000
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	60	1200	d54		0.00000

Note that it is the last three elements of the "y Vector" matrix where we have set the constraints that the second derivatives at time zero and 10 and the first derivative of the yield curve at 10 years be zero. Then invert the matrix and solve for these coefficients:

Coefficient Vector		Values
c1		0.0400000000
d11		0.0744582074
d12		0.0000000000
d13		-0.6332723810
d14		0.8530487205
c2		0.0363357405
d21	=	0.1330863589
d22		-0.3517689090
d23		0.3047780430
d24		-0.0850017035
c3		0.1272995768
d31		-0.2307689862
d32		0.1940141086
d33		-0.0590773021
d34		0.0059621327
c4		-0.5156595261
d41		0.6265098177
d42		-0.2346252934
d43		0.0361758984
d44		-0.0019756340
c5		0.8790434876
d51		-0.4892525933
d52		0.1001034299
d53		-0.0084545981
d54		0.0002558909

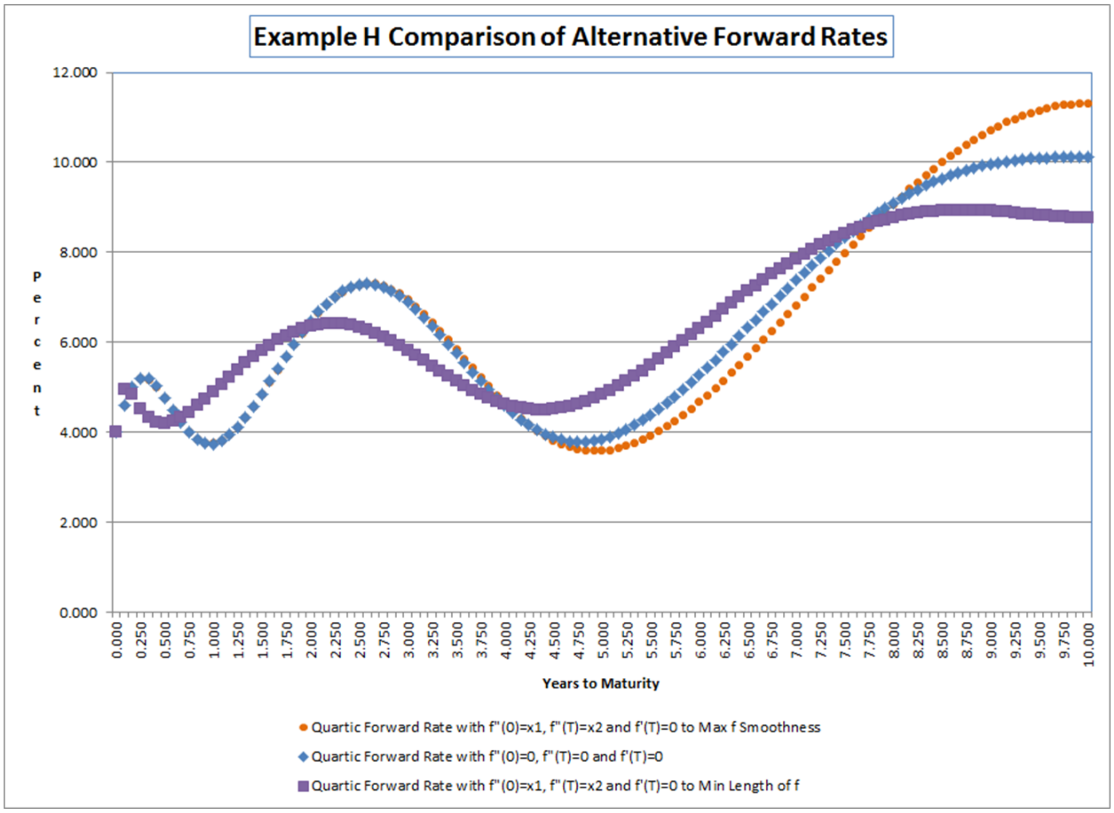
These coefficients give us the five quartic functions that make up the forward rate curve. We the second and third relationships listed above to solve for the derived coefficients for the zero coupon yield curve. The yield function for any segment j is

$$y_j(t) = \frac{1}{t}\left[y^*(t_j)t_j + c_j(t-t_j) + \frac{1}{2}d_{j1}(t^2-t_j^2) + \frac{1}{3}d_{j2}(t^3-t_j^3) + \frac{1}{4}d_{j3}(t^4-t_j^4) + \frac{1}{5}d_{j4}(t^5-t_j^5) \right]$$

We note that y* denotes the observable value of y at the left hand side of the line segment where the maturity is t_j. Within the segment, y is a quintic function of t, divided by t.

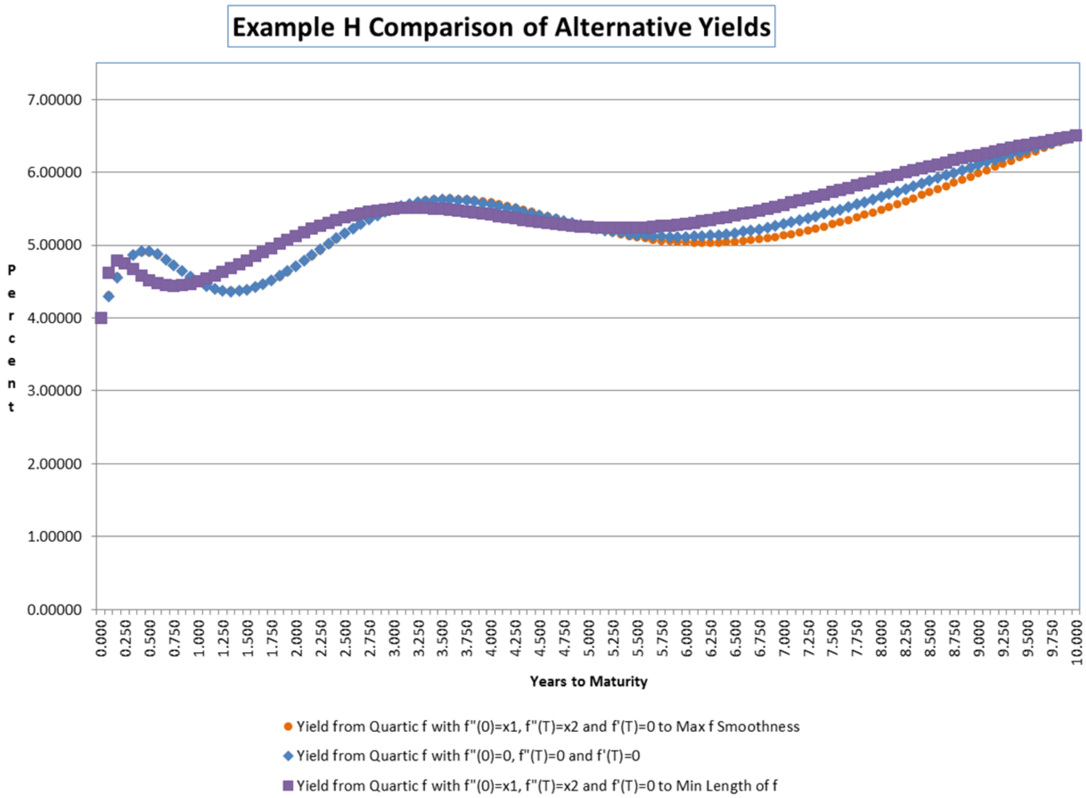
We can now plot the yield and forward rate curves to see the realism of our assumptions about maximum smoothness and the related constraints we have imposed. The results show something very interesting. When the infinite number of forward rate curves that are maximum smoothness subject to constraints 23, 24 and 25 are compared, it is seen that the real trade-off comes from sacrificing smoothness at the left hand side of the curve in order to get shorter length overall, with a smaller rise in the forward curve at the 10 year point. Example H-Qf1a, where we have set all three assumptions 23, 24 and 25 to zero falls between the two optimized results at the 10 year point:

Figure 14



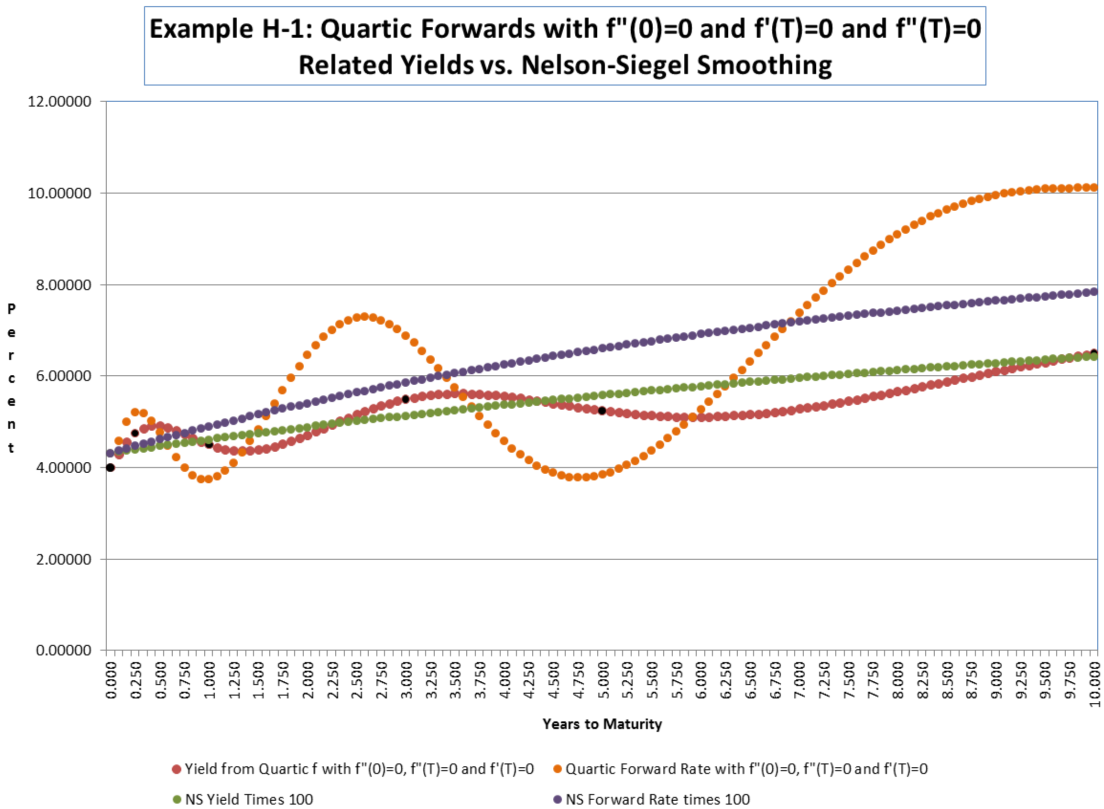
The results for all three examples are far superior to what we in earlier examples. When we compare the three yield curves from our three Example H alternatives, again the results are very reasonable:

Figure 15



Next we compare our base case, with assumptions 23, 24 and 25 all set to zero, with the arbitrary Nelson-Siegel formulation. The Nelson-Siegel functions are simply two lines drawn on a page with no validity because they don't match the observable data, which are the black dots on the yield curve derived from our base case:

Figure 16



We continue to be surprised that the Nelson-Siegel technique is used in academia given the fact that it is both inaccurate and more complex in calculation: the Nelson-Siegel formulation is incapable of fitting an arbitrary set of data and requires a non-linear optimization. The maximum smoothness forward rate approach, by contrast, fits the data perfectly in every case and requires (in our base case) just a matrix inversion.

For example H-Qf1b, when we use simple spreadsheet optimize to iterate on x_1 and x_2 to derive the maximum smoothness forward rate curve of shortest length, we get these coefficients:

Coefficient Vector		Values
c1		0.04
d11		0.232654282
d12		-1.80184349
d13		5.124447912
d14		-5.19776484
c2		0.060389164
d21	=	-0.09357235
d22		0.155516289
d23		-0.09517817
d24		0.021861242
c3		0.040261445
d31		-0.01306147
d32		0.034749971
d33		-0.01466729
d34		0.001733522
c4		-0.18940957
d41		0.293166555
d42		-0.11836404
d43		0.019358042
d44		-0.00110192
c5		0.653581615
d51		-0.3812264
d52		0.083953844
d53		-0.00761768
d54		0.000246863

When we iterate using the Example H-Qf1c objective, maximum smoothness, we get the following set of coefficients for our five quartic forward rate segments:

Coefficient Vector		Values
c1		0.04
d11		0.077396762
d12		-0.03383111
d13		-0.52363325
d14		0.736847176
c2		0.036799797
d21	=	0.128600013
d22		-0.34105061
d23		0.295618762
d24		-0.08240483
c3		0.124892139
d31		-0.22376935
d32		0.187503439
d33		-0.05675061
d34		0.005687509
c4		-0.4755168
d41		0.576775895
d42		-0.21276919
d43		0.032198867
d44		-0.00172495
c5		0.672902511
d51		-0.34195955
d52		0.062851448
d53		-0.00455055
d54		0.000112524

Comparing Yield Curve and Forward Rate Smoothing Techniques

As we have discussed throughout this blog, the standard process for evaluating yield curve and forward rate curves should involve the mathematical statement of the objective of smoothing, the imposition of constraints that are believed to generate realistic results, and then the testing of the results to establish whether they are reasonable or not. In the last section of this chapter, we discuss the "Shimko Test" used in Adams and van Deventer (1994) for measuring the accuracy of yield curve and forward rate smoothing techniques on large amounts of data. For expositional purposes in this section, we analyze only one case, our base case, and reach some tentative conclusions whose validity can be confirmed or denied on huge masses of data using the Shimko test.

Ranking 23 Smoothing Techniques by Smoothness of the Forward Rate Curve

The chart below ranks 23 smoothing techniques on their effectiveness in lowering the discrete smoothing statistic as much as possible as explained above:

	Ranking by Smoothness of Forward Rate Curve					
Rank	Example	Smoothing Technique	Length of Yield Curve Fitted	Length of Forward Curve Fitted	Smoothness of Yield Curve Fitted	Smoothness of Forward Curve Fitted
1	**Example H-Qf1c**	**Quartic Forward Rate with f"(0)=x1, f"(T)=x2 and f'(T)=0 to Max f Smoothness**	**11.58**	**21.01**	**505.76**	**4,287.35**
2	**Example H-Qf1a**	**Quartic Forward Rate with f"(0)=0, f"(T)=0 and f'(T)=0**	**11.51**	**19.91**	**495.93**	**4,309.97**
3	Example F	Cubic y, Max f Smoothness	11.75	24.94	510.10	4,564.23
4	Example F	Cubic y, Max y Smoothness	11.73	22.62	473.48	4,811.18
5	Example F	Cubic y, y'(T)=0 and y"(0)=0	11.64	23.14	475.92	5,038.87
6	Example G-4Cfb	Cubic Forward Rate with f'(0)=x1 and f'(T)=x2 to Max f Smoothness	14.15	48.89	687.21	5,931.25
7	Example G-3Cfb	Cubic Forward Rate with f"(0)=x1 and f'(T)=x2 to Max f Smoothness	14.15	48.89	687.21	5,931.25
8	Example G-3Cfe	Cubic Forward Rate with f"(0)=x1 and f'(T)=0 to Max f Smoothness	13.59	40.52	690.46	6,179.14
9	Example E	Quadratic f, f'(T)=0	11.07	15.28	1,061.01	7,767.53
10	Example E	Quadratic f, Min f Length	11.07	15.26	1,061.76	7,779.04
11	Example E	Quadratic f, Min y Length	11.07	15.28	1,062.55	7,791.63
12	Example G-3Cfa	Cubic Forward Rate with f"(0)=0 and f'(T)=0	15.36	52.23	597.41	8,118.72
13	Example G-4Cfc	Cubic Forward Rate with f'(0)=x1 and f'(T)=x2 to Min f Length	11.38	15.79	4,111.96	24,073.20
14	Example G-3Cfc	Cubic Forward Rate with f"(0)=x1 and f'(T)=x2 to Min f Length	11.38	15.79	4,111.96	24,073.23
15	Example G-3Cfd	Cubic Forward Rate with f"(0)=x1 and f'(T)=0 to Min f Length	11.40	15.87	4,200.81	24,615.22
16	Example G-4Cfa	Cubic Forward Rate with f'(0)=0 and f'(T)=0	19.83	78.05	2,128.84	25,398.14
17	**Example H-Qf1b**	**Quartic Forward Rate with f"(0)=x1, f"(T)=x2 and f'(T)=0 to Min Length of f**	**11.23**	**15.52**	**5,267.60**	**25,629.13**
18	Example C	Linear forwards	12.15	31.59	619.35	27,834.72
19	Example D	Quadratic y, Min y Length	11.46	18.70	3,346.49	35,901.19
20	Example D	Quadratic y, Min f Length	11.56	17.21	3,780.67	43,108.27
21	Example D	Quadratic y, y'(T)=0	12.00	20.48	4,409.57	56,443.05
22	Example A	Yield step function	13.12	13.12	123,120.00	123,120.00
23	Example B	Linear yields	10.99	16.87	1,776.50	319,346.00

As we would expect, the best result was the maximum smoothness forward rate approach where the optimization of the 24th and 25th constraints was done with respect to smoothness. The second best approach was maximum smoothness was the base case, Example H-Qf1a, where we set the derivatives with respect to constraints 23, 24, and 25. Technique 17 was also one of the maximum smoothness alternatives, where we intentionally sacrificed the smoothness of the curve (most notably on the short end of the curve) to minimize length. Linear yield smoothing produced the worst results when smoothness is the criterion for best.

Ranking 23 Smoothing Techniques by Length of the Forward Curve

Next we report on the discrete approximation for length of the forward curve if its length is the sole criterion. The results for 23 techniques are shown below:

Ranking by Length of Forward Rate Curve						
Rank	Example Description	Smoothing Technique	Length of Yield Curve Fitted	Length of Forward Curve Fitted	Smoothness of Yield Curve Fitted	Smoothness of Forward Curve Fitted
1	Example A	Yield step function	13.12	13.12	123,120.00	123,120.00
2	Example E	Quadratic f, Min f Length	11.07	15.26	1,061.76	7,779.04
3	Example E	Quadratic f, f'(T)=0	11.07	15.28	1,061.01	7,767.53
4	Example E	Quadratic f, Min y Length	11.07	15.28	1,062.55	7,791.63
5	**Example H-Qf1b**	**Quartic Forward Rate with f"(0)=x1, f"(T)=x2 and f'(T)=0 to Min Length of f**	**11.23**	**15.52**	**5,267.60**	**25,629.13**
6	Example G-3Cfc	Cubic Forward Rate with f"(0)=x1 and f'(T)=x2 to Min f Length	11.38	15.79	4,111.96	24,073.23
7	Example G-4Cfc	Cubic Forward Rate with f'(0)=x1 and f'(T)=x2 to Min f Length	11.38	15.79	4,111.96	24,073.20
8	Example G-3Cfd	Cubic Forward Rate with f"(0)=x1 and f'(T)=0 to Min f Length	11.40	15.87	4,200.81	24,615.22
9	Example B	Linear yields	10.99	16.87	1,776.50	319,346.00
10	Example D	Quadratic y, Min f Length	11.56	17.21	3,780.67	43,108.27
11	Example D	Quadratic y, Min y Length	11.46	18.70	3,346.49	35,901.19
12	**Example H-Qf1a**	**Quartic Forward Rate with f"(0)=0, f"(T)=0 and f'(T)=0**	**11.51**	**19.91**	**495.93**	**4,309.97**
13	Example D	Quadratic y, y'(T)=0	12.00	20.48	4,409.57	56,443.05
14	**Example H-Qf1c**	**Quartic Forward Rate with f"(0)=x1, f"(T)=x2 and f'(T)=0 to Max f Smoothness**	**11.58**	**21.01**	**505.76**	**4,287.35**
15	Example F	Cubic y, Max y Smoothness	11.73	22.62	473.48	4,811.18
16	Example F	Cubic y, y'(T)=0 and y"(0)=0	11.64	23.14	475.92	5,038.87
17	Example F	Cubic y, Max f Smoothness	11.75	24.94	510.10	4,564.23
18	Example C	Linear forwards	12.15	31.59	619.35	27,834.72
19	Example G-3Cfe	Cubic Forward Rate with f"(0)=x1 and f'(T)=0 to Max f Smoothness	13.59	40.52	690.46	6,179.14
20	Example G-3Cfb	Cubic Forward Rate with f"(0)=x1 and f'(T)=x2 to Max f Smoothness	14.15	48.89	687.21	5,931.25
21	Example G-4Cfb	Cubic Forward Rate with f'(0)=x1 and f'(T)=x2 to Max f Smoothness	14.15	48.89	687.21	5,931.25
22	Example G-3Cfa	Cubic Forward Rate with f"(0)=0 and f'(T)=0	15.36	52.23	597.41	8,118.72
23	Example G-4Cfa	Cubic Forward Rate with f'(0)=0 and f'(T)=0	19.83	78.05	2,128.84	25,398.14

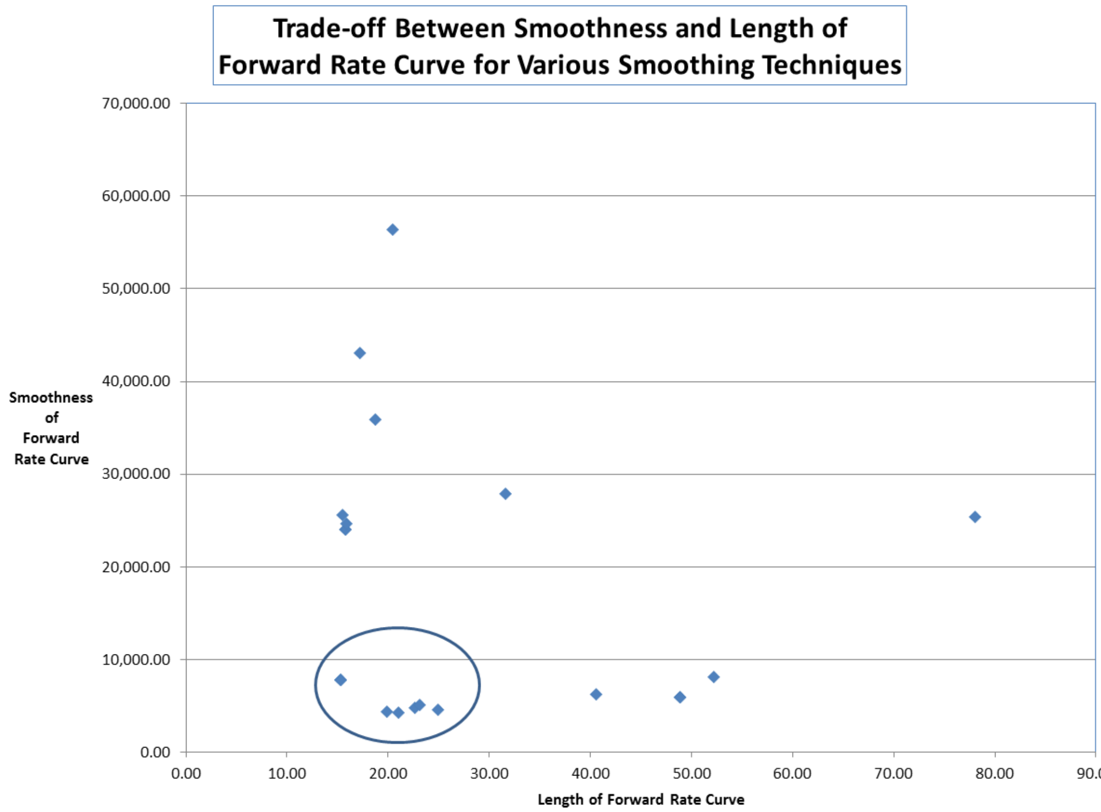
Not surprisingly, if one doesn't insist on continuity or "smoothness" at all, one gets a short forward curve. Ranks 1-4 are taken by a yield/forward step function and quadratic curve fitting. Among all of the techniques which are at least twice differentiable, the maximum smoothness forward rate technique, optimized to minimize length, was the winner. There is no need to sacrifice one attribute (length) for another (smoothness). In this example, the same functional form (quadratic forward rates) can produce the winner by either criterion.

If we look at both attributes, which techniques best balance the trade-offs between smoothness and length? We turn to that question next.

Trading Off Smoothness versus the Length of the Forward Rate Curve

In the graph below, we plot the 23 smoothing techniques on an XY graph where both length and smoothness of the forward curve is relevant. We truncated the graph to eliminate some extreme outliers.

Figure 17



A reasonable person would argue that the "best" techniques by some balance of smoothness and length would fall in the lower left hand corner of the graph. There are five techniques with a forward rate curve with smoothness below 7,000 and length below 30.00:

Best Techniques: Smoothness Under 7,000, Length Under 30, Ranked by Length of Forward Curve					
Example Description	**Smoothing Technique**	**Length of Yield Curve Fitted**	**Length of Forward Curve Fitted**	**Smoothness of Yield Curve Fitted**	**Smoothness of Forward Curve Fitted**
Example H-Qf1a	**Quartic Forward Rate with f"(0)=0, f"(T)=0 and f'(T)=0**	**11.51**	**19.91**	**495.93**	**4,309.97**
Example H-Qf1c	Quartic Forward Rate with f"(0)=x1, f"(T)=x2 and f'(T)=0 to Max f Smoothness	11.58	21.01	505.76	4,287.35
Example F	Cubic y, Max y Smoothness	11.73	22.62	473.48	4,811.18
Example F	Cubic y, y'(T)=0 and y"(0)=0	11.64	23.14	475.92	5,038.87
Example F	Cubic y, Max f Smoothness	11.75	24.94	510.10	4,564.23

It is interesting to see that the "best five" consist of 2 quartic forward rate smoothing approaches, ranked first and second by smoothness, and 3 cubic yield spline approaches. Of this "best 5" group, the quartic forward rate approach, optimized for smoothness, was the smoothest. The quartic spline approach, with all derivatives in constraints 23-25 set to zero, was the shortest. We now turn to a more comprehensive approach to performance measurement of smoothing techniques.

The Shimko Test for Measuring Accuracy of Smoothing Techniques

In 1994, Kenneth J. Adams and Donald R. van Deventer published "Fitting Yield Curves and Forward Rate Curves with Maximum Smoothness" (*Journal of Fixed Income*, June 1994). In 1993, our thoughtful friend David Shimko responded to an early draft of the Adams and van Deventer paper by saying "I don't care about a mathematical proof of 'best,' I want something that would have best estimated a data point that I intentionally leave out of the smoothing process-this to me is proof of which technique is most realistic." A statistician would add, "And I want something that is most realistic on a very large sample of data." We agreed that Dr. Shimko's suggestion was the ultimate proof of the accuracy and realism of any smoothing technique. The common academic practice of using one set of fake data and then judging which technique "looks good" or "looks bad" is as ridiculous as it is common. Therefore, we atone for the same sin, which we have used in the first 10 installments of this series, by explaining how to perform the Shimko test as in Adams and van Deventer [194].

The Shimko test works as follows. First we assemble a large data set, which in the Adams and van Deventer case was 660 days of swap data. Next, we select one of the maturities in that data set and leave it out of the smoothing process. For purposes of this example, say we leave out the 7 year maturity because that leaves a wide 5 year gap in the swap data to be filled by the smoothing technique. We smooth the 660 yield curves one by one. Using the smoothing results and the zero coupon bond yields associated with the 14 semiannual payment dates of the 7 year interest rate swap, we calculate the 7 year swap rate implied by the smoothing process. We have 660 observations of this estimated 7 year swap rate, and we compare it to the actual 7 year swap rates that we left out of the smoothing process. The "best" smoothing technique is the one which most accurately estimates the omitted data point over the full sample.

This test can be performed on any of the maturities that were inputs to the smoothing process, and we strongly recommend that all maturities are used, one at a time. Because this test suggested by David Shimko is a powerful test applicable to any contending smoothing techniques, we strongly recommend that no assertion of superior performance be made without applying the Shimko test on a large amount of real data. As an example of the volumes of data that can be analyzed, Dicker, Jarrow and van Deventer present a history of forward rate curves, zero coupon yield curves, and par coupon bond yield curves for every business day (more than 12,000 business days) from January 2, 1962 to August 22, 2011 in these volumes available on www.kamakuraco.com:

Dicker, Daniel T., Robert A. Jarrow and Donald R. van Deventer, "Inside the Kamakura Book of Yields: A Pictorial History of 50 Years of U.S. Treasury Forward Rates," Kamakura Corporation memorandum, September 13, 2011.

Dickler, Daniel T. and Donald R. van Deventer, "Inside the Kamakura Book of Yields: An Analysis of 50 Years of Daily U.S. Treasury Forward Rates," Kamakura blog, www.kamakuraco.com, September 14, 2011.

Dicker, Daniel T., Robert A. Jarrow and Donald R. van Deventer, "Inside the Kamakura Book of Yields: A Pictorial History of 50 Years of U.S. Treasury Zero Coupon Bond Yields," Kamakura Corporation memorandum, September 26, 2011.

Dickler, Daniel T. and Donald R. van Deventer, "Inside the Kamakura Book of Yields: An Analysis of 50 Years of Daily U.S. Treasury Zero Coupon Bond Yields," Kamakura blog, www.kamakuraco.com, September 26, 2011.

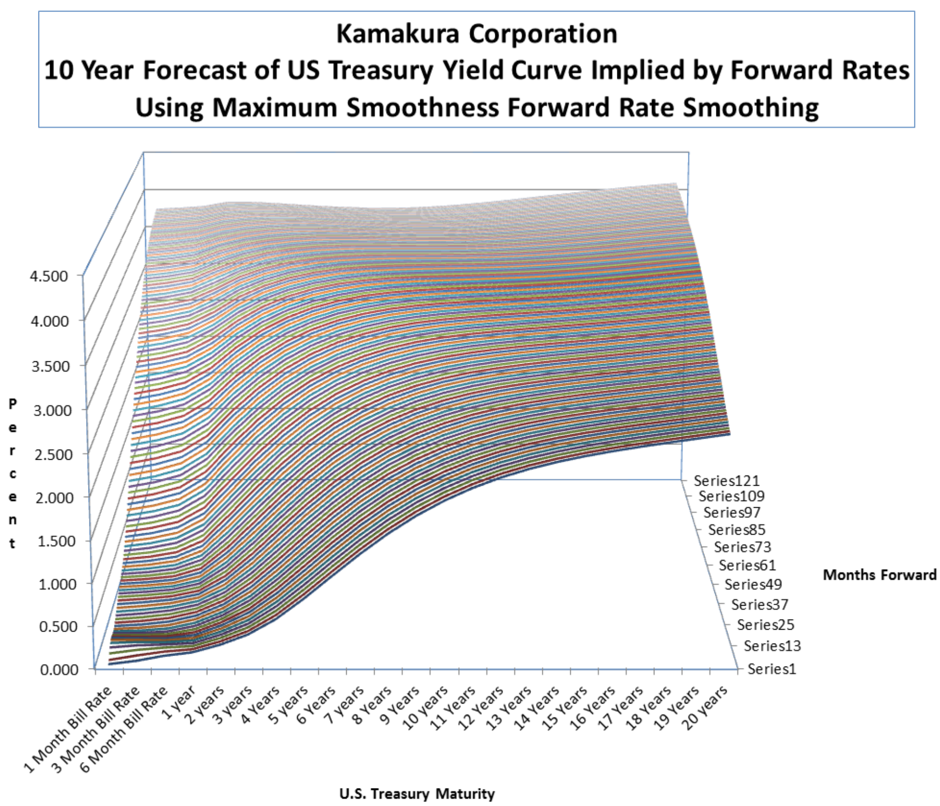
Dicker, Daniel T., Robert A. Jarrow and Donald R. van Deventer, "Inside the Kamakura Book of Yields: A Pictorial History of 50 Years of U.S. Treasury Par Coupon Bond Yields," Kamakura Corporation memorandum, October 5, 2011.

Dickler, Daniel T. and Donald R. van Deventer, "Inside the Kamakura Book of Yields: An Analysis of 50 Years of Daily U.S. Par Coupon Bond Yields," Kamakura blog, www.kamakuraco.com, October 6, 2011.

Smoothing Yield Curves Using Coupon Bearing Bond Prices as Inputs

For expositional purposes, we have assumed in this chapter that the raw inputs to this process are zero coupon bond yields. When, instead, the inputs are the prices and terms on coupon bearing bonds, the analysis changes in a minor way which we illustrate in Chapter 17. The initial zero coupon bond yields are guessed, and an iteration of zero coupon bond yields is performed that minimizes the sum of squared bond pricing errors. The Dickler, Jarrow and van Deventer analysis above was done in that manner using Kamakura Risk Manager version 8.0. Kamakura Corporation produces a weekly forecast of implied forward rates derived in this manner. The implied forward U.S. Treasury yields as of May 3, 2012 are shown in this graph:

Figure 18



We now turn to interest rate simulation using the smoothed yield curves that we have generated in this chapter.

Appendix A

Proof of the Maximum Smoothness Forward Rate Theorem[1]

Schwartz [1989] demonstrates that cubic splines produce the maximum smoothness discount functions or yield curves if the spline is applied to discount bond prices or yields respectively. In this appendix, we derive by a similar argument the functional form that produces the forward rate curve with maximum smoothness. Let f(t) be the current forward rate function, so that

$$P(t) = \mathbf{exp}\left(-\int_0^t f(s)ds \right)$$

(A1)

is the price of a discount bond maturing at time t. The maximum smoothness term structure is a function f with a continuous derivative that satisfies the optimization problem

$$\mathbf{min} \quad \int_0^T f''^2(s)ds$$

(A2)

subject to the constraints

$$\int_0^{t_i} f(s)ds = -\mathbf{log} \;\; P_i, \quad for\ i = 1,2,...,m.$$

(A3)

Here the $P_i = P(t_i)$, for i=1,2,...,m are given prices of discount bonds with maturities $0 < t_1 < t_2 <...< t_m < T$.

Integrating twice by parts we get the following identity:

$$\int_0^t f(s)ds = \frac{1}{2}\int_0^t (t-s)^2 f''(s)ds + tf(0) + \frac{1}{2}t^2 f'(0)$$

[1] This proof was kindly provided by Oldrich Vasicek. We also appreciate the comments of Volf Frishling,, who pointed out an error in the proof in Adams and van Deventer[1994]. Robert Jarrow also made important contributions to this proof.

(A4)

Put

$$g(t) = f''(t), \quad 0 \le t \le T$$

(A5)

and define the step function

$$\begin{aligned} u(t) &= 1 \quad for \;\; t \ge 0 \\ &= 0 \quad for \;\; t < 0. \end{aligned}$$

The optimization problem can then be written as

$$\mathbf{min} \int_0^T g^2(s)ds$$

(A6)

subject to

$$\frac{1}{2}\int_0^T (t_i - s)^2 u(t_i - s)g(s)ds = -\mathbf{log}\, P_i - t_i\, f(0) - \frac{1}{2} t_i^2\, f'(0)$$

(A7)

for i=1, 2, ..., m. Let λ_i 9 for i=1,2,..,m be the Lagrange multipliers corresponding to the constraints (A7). The objective then becomes

$$\begin{aligned} \mathbf{min} \quad Z[g] &= \int_0^T g^2(s)ds \\ &+ \sum_{i=1}^m \lambda_i \left(\frac{1}{2}\int_0^T (t_i - s)^2 u(t_i - s)g(s)ds + \mathbf{log}\, P_i + t_i\, f(0) + \frac{1}{2} t_i^2\, f'(0) \right) \end{aligned}$$

(A8)

According to the calculus of variations, if the function g is a solution to (A8), then

$$\frac{d}{d\varepsilon} Z[g + \varepsilon h] \quad {}_{\varepsilon=0} = 0$$

(A9)

for any function h(t) identically equal to w"(t) where w(t) is any twice differentiable function defined on [0,T] with w'(0)=w(0)=0[2]. We get

$$\frac{d}{d\varepsilon} Z[g+\varepsilon h] \quad _{\varepsilon=0} = 2\int_0^T \left[g(s) + \frac{1}{4}\sum_{i=1}^{m} \lambda_i (t_i - s)^2 u(t_i - s) \right] h(s)ds$$

In order that this integral is zero for any function h, we must have

$$g(t) + \frac{1}{4}\sum_{i=1}^{m} \lambda_i (t_i - t)^2 u(t_i - t) = 0$$

(A10)

for all t between 0 and T. This means that

$$g(t) = 12 e_i t^2 + 6 d_i t + 2 c_i \quad for \quad t_{i-1} < t \le t_i, \; i = 1,2,..,m+1,$$

(A11)

where

$$e_i = -\frac{1}{48}\sum_{j=i}^{m} \lambda_j$$
$$d_i = \frac{1}{12}\sum_{j=i}^{m} \lambda_j t_j$$
$$c_i = -\frac{1}{8}\sum_{j=i}^{m} \lambda_j t_j^2$$

(A12)

and we define $t_0 = 0$, $t_{m+1} = T$. Moreover (A10) implies that g and g' (and therefore f" and f''') are continuous. From (A4) we get

$$f(t) = e_i t^4 + d_i t^3 + c_i t^2 + b_i t + a_i, \quad t_{i-1} < t \le t_i, \; i = 1,2,..,m+1.$$

(A13)

Continuity of f, f', f" and f''' then implies that

---

2 We are grateful to Robert Jarrow for his assistance on this point.

$$e_i t_i^4 + d_i t^3 + c_i t_i^2 + b_i t_i + a_i = e_{i+1} t_i^4 + d_{i+1} t_i^3 + c_{i+1} t_i^2 + b_{i+1} t_i + a_{i+1}, \quad i = 1,2,..,m$$

(A14)

$$4 e_i t_i^3 + 3 d_i t_i^2 + 2 c_i t_i + b_i = 4 e_{i+1} t_i^3 + 3 d_{i+1} t_i^2 + 2 c_{i+1} t_i + b_{i+1}, \quad i = 1,2,..,m.$$

$$12 e_i t_i^2 + 6 d_i t_i + 2 c_i = 12 e_{i+1} t_i^2 + 6 d_{i+1} t_i + 2 c_{i+1}$$

$$24 e_i t_i + 6 d_i = 24 e_{i+1} t_i + 6 d_{i+1}$$

(A15)

The constraints (A3) become

$$\frac{1}{5} e_i (t_i^5 - t_{i-1}^5) + \frac{1}{4} d_i (t_i^4 - t_{i-1}^4) + \frac{1}{3} c_i (t_i^3 - t_{i-1}^3) + \frac{1}{2} b_i (t_i^2 - t_{i-1}^2) + a_i (t_i - t_{i-1}) = -\log\left[\frac{P_i}{P_{i-1}}\right],$$
$$i = 1,2,..,m$$

(A16)

where we define $P_0 = 1$. This proves the theorem.