# How many factors are really important in explaining the fluctuations of the yield curve?

Paolo Pasquariello[1]
Ph-D candidate
New York University
Stern School of Business

This draft: May 1998

[1] Please address any comment at ppasquar@stern.nyu.edu

A yield curve can be described as a sequence of yields on debt instruments with specified maturities, or as a set of discrete yields plotted versus maturity in a graph. Even if in most of the cases the points are connected to form a piecewise linear but smooth graph, the real data is always a finite set of discrete points. A simple observation of the chart below shows that the shape as well as the level of the yield curve for Treasury securities changes over time. As a consequence, interest-rate risk cannot be summarized just in terms of parallel shifts of the curve.
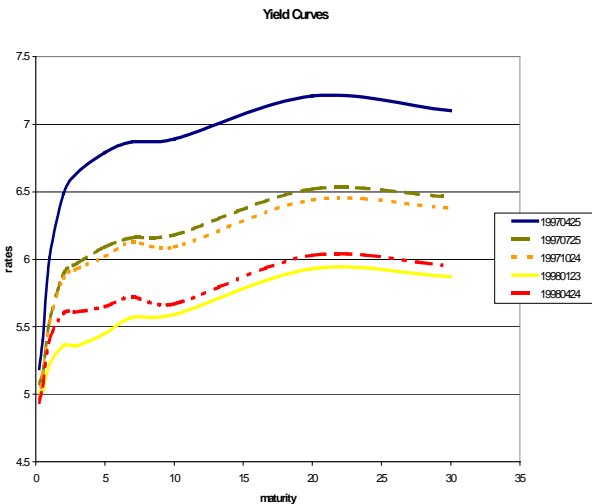
Yield Curves

Changes in yields on securities of different maturities can be separated into independent modes of fluctuations. In broad terms, shifts in the curve can be characterized as changes in level, slope and curvature. Higher order effects eventually contribute to the shape of the term structure of interest rates. However, statistical analysis shows that the first three components explain more than 90% of the volatility of each of the points on the curve[1]. In most of the financial and economic applications related to the analysis of the yield curve, representing the curve itself and its modes of fluctuations as continuous functions is a desirable task to implement. This paper examines the problem of representing a yield curve and its modes of

fluctuations with *polynomial* functions of maturity. This will imply the attempt of expressing the yield on a default-free debt instrument as a power function of its maturity and, more particularly in this paper, of obtaining numerical estimates of the coefficients of the chosen power function. In the next paragraph we state some general assumptions on the structure of the curve and its variation through time. We then analyze and implement a 10-steps method for the estimation of the coefficients of the selected polynomial.

Basic assumptions of the model[2]

Let R(t,y) denote the yield at time t of a debt instrument with a term to maturity of y years[3]. R(t,y) is a *continuous* function of y for fixed t, where y runs continuously from zero to some upper limit such as 30 years. We assume the yield curve at time t can be expressed as:

$$R(t, y) = R_0(y) + \sum_{j=1}^{J} w_j(t) \cdot f_j(y) \qquad [1]$$

Here, R0(y) and f1(y)………fJ(y) are stationary functions of maturity. Equation [1] simply states that the curve at time t can be constructed as a "baseline" function R0(y) plus a linear combination of *shift functions* f1(y)……fJ(y). Wj(y) is the weight on the jth shift function at time t. Hence, in this model the yield curve represented in [1] varies over time, in shape as well as level, as a result of temporal variation in the Wj(t) weights. We assume throughout this paper that the weights are normally distributed random walks with zero drift, unity variance per week and statistically independent of each other. If we let t0, t1, t2……be a sequence of times one week apart, we can rewrite equation [1] as:

$$R(tk, y) = R_0(y) + \sum_{j=1}^{J} w_j(tk) \cdot f_j(y) \qquad [2]$$

---

[1] For a complete analysis of this and most of the theoretical issues of this paper we refer to Kenneth Garbade, "Fixed Income Analytics", The MIT Press, Cambridge, MA, 1996, especially in chapters 16 to 20.

[3] The instrument can be a Treasury bill, a zero-coupon bond, a spot claim or a coupon-bearing security like a Treasury Note or Bond. For the purpose of this exercise, we will consider Treasury Bills, Notes and Bonds.

The change in yield from time t(k-1) to time t(k) at a maturity of y years, i.e. at a particular point across the curve is:

$$DR(tk, y) = R(y, tk) - R(y, t(k-1)) \qquad [3]$$

Using equation [2] and [3], as well as denoting:

$$ej(tk) = wj(tk) - wj(t(k-1)) \qquad [4]$$

we can express [3] as:

$$DR(y, tk) = \sum_{j=1}^{J} f_j(y) \cdot ej(tk) \qquad [5]$$

Thus, this function describes the shift in the yield curve from time t(k-1) to t(k) as a linear combination of the shifts functions f1(y)….fJ(y). Based on the assumptions regarding the behavior of the weights wj(t), we can now claim that the ej(tk) terms are i.i. normally distributed random variables with zero drift and unit variance. Hence, the adopted model relies on the assumption that the shift functions represent normalized and statistically independent modes of fluctuation in the yield curve.

The estimation problem

The problem of interest in this paper is the estimation of the Ro(y) and f1(y)…..fJ(y) functions and of the weights w1(t)….wJ(t) at any point in time.

We assume that the baseline function R0(y) and the fj(y) shift functions are five-degree polynomials in y, i.e. that:

$$R_0(y) = \sum_{i=1}^{I} \beta_{i0} y^{i-1} \qquad [6]$$

$$f_j(y) = \sum_{i=1}^{I} \beta_{ij} y^{i-1} \qquad j = 1,....,J \quad [7]$$

where I is equal to 6, β_{i0} are the coefficients of the baseline polynomial and β_{ij} are the coefficients of the shift polynomials. We will also assume without loss of generality that J=I, i.e. we will consider six modes of fluctuations of the curve through time.

As a consequence of equations [6] and [7], we can restate the initial problem as: How can we estimate the β_{i0} and β_{ij} coefficients and the wj(t) weights at any point in time t?

Replacing the shift functions in [5] with their polynomial counterparts, we obtain:

$$DR(y, tk) = \sum_{i=1}^{6} y^{i-1} \left| \sum_{j=1}^{6} \beta_{ij} \cdot ej(tk) \right| \qquad [8]$$

We will now focus on a set of n=10 > I specific and distinct (but arbitrary) maturities:

Y1(t) = 0.25 years   Y2(t) = 0.5 years
Y3(t) = 1 year   Y4(t) = 2 years
Y5(t) = 3 years   Y6(t) = 5 years
Y7(t) = 7 years   Y8(t) = 10 years
Y9(t) = 20 years   Y10(t) = 30 years

and on a specific time horizon of one year (53 weeks) from April 25th 1997 to April 24th 1998[4]. We then define the vectors:

$$DR(tk) = \begin{vmatrix} DR(y1, tk) \\ . \\ . \\ . \\ DR(y10, tk) \end{vmatrix} \qquad e(tk) = \begin{bmatrix} e1(tk) \\ . \\ . \\ . \\ e6(tk) \end{bmatrix}$$

and the matrices:

$$Y = \begin{vmatrix} 1 & y1 & y1^2 & . & y1^5 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & y10 & y10^2 & . & y10^5 \end{vmatrix}$$

and:

$$B = \begin{vmatrix} \beta_{11} & . & . & . & \beta_{1J} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ \beta_{I1} & . & . & . & \beta_{IJ} \end{vmatrix}$$

As a result, we can write equation [8] in matrix form as:

$$DR(tk) = Y \cdot B \cdot e(tk) \qquad [9]$$

Equation [9] says that the vector of one-week changes DR(tk) is a linear function of the random vector e(tk), normally distributed with a mean-vector of zero and a covariance matrix equal to the JxJ identity matrix.

Let Ω(DR) denote the 10x10 covariance matrix of DR(tk). Since Ω(DR) is a symmetric matrix, we can represent it, through the well-known

---

[4] The original Garbade contribution correctly relies on theoretical-spot yields as baseline of analysis. For the purposes of this paper, benchmark-Treasury non-par bond yields were used, instead.

procedure of "diagonalization", in the following form:

$$\Omega(DR) = V(DR) \cdot D(DR) \cdot V(DR)' \qquad [\ 10\ ]$$

where the columns of V(DR) are the eigenvectors of Ω(DR) and D(DR) is a diagonal matrix whose entries are the eigenvalues of Ω(DR). Note that e(tk) is a J-dimensional vector, DR(tk) is a 10-dimensional vector and 10 > J: this will imply that Ω(DR) has only J positive eigenvalues. Hence, without loss of generality, we can assume the following partitions for V(DR) and D(DR):

$$V(DR) = [V'(DR) \quad V''(DR)]$$

$$D(DR) = \begin{vmatrix} D'(DR) & 0 \\ 0 & 0 \end{vmatrix}$$

where there are J column vectors in V'(DR) and where D'(DR) is a JxJ matrix with all the positive eigenvalues on its diagonal. It follows that:

$$\Omega(DR) = V'(DR) \cdot D'(DR) \cdot V'(DR)' \qquad [\ 11\ ]$$

We consider now a generic J-normally distributed random vector e with a mean vector of zero and a covariance matrix equal to the identity matrix. Further, we define the 10-dimensional vector:

$$DR = V'(DR) \cdot (D'(DR)^\wedge 1/2) \cdot e \qquad [\ 12\ ]$$

where D'(DR)^1/2 is a JxJ diagonal matrix whose diagonal entries are the square roots of the corresponding entries on the diagonal of D'(DR). It is then easy to prove that:

$$COV(DR) = V'(DR) \cdot D'(DR) \cdot V'(DR)' = \Omega(DR)$$

Thus, DR in the left side equation [12] and DR(tk) in the left side of equation [9] both are normally distributed with a mean vector zero and the same covariance matrix Ω(DR). Moreover, the "residual" vectors on the right side of both equations are normally distributed with mean zero and covariance matrix equal to the JxJ identity matrix. Thus, we can identify the matrices on the right side of both equations:

$$Y \cdot B = V'(DR) \cdot (D'(DR)^\wedge 1/2) \qquad [\ 13\ ]$$

This implies that we can solve the equation [13] for the unknown matrix B:

$$B = (Y'Y)^{-1} \cdot Y'V'(DR) \cdot (D'(DR))^{1/2} \qquad [\ 14\ ]$$

Equation [14] permits to reduce the problem of estimating the coefficients in B to the problem of estimating the covariance matrix Ω(DR).

To do so, we need to refer again to equation [2]. Replacing the baseline function and the shift functions with their polynomial counterparts, equation [2] becomes:

$$R(y, tk) = \sum_{i=1}^{6} ai(tk) y^{i-1} \qquad [\ 15\ ]$$

$$ai(tk) = \beta_{i0} + \sum_{j=1}^{6} \beta_{ij} w_j(tk) \qquad [\ 16\ ]$$

If we further define:

$$Dai(tk) = ai(tk) - ai(t(k-1)) \qquad [\ 17\ ]$$

$$Dai(tk) = \sum_{j=1}^{6} \beta_{ij} ej(tk) \quad i = 1,....,6 \qquad [\ 18\ ]$$

using the definition of ei(tk), then we can specify the following I-dimensional vector:

$$Da(tk) = \begin{vmatrix} Da1(tk) \\ . \\ . \\ . \\ DaI(tk) \end{vmatrix}$$

This permits us to represent equation [18] in matrix form as

$$Da(tk) = B \cdot e(tk) \qquad [\ 19\ ]$$

Equation [19] says that the vector Da(tk) of composite coefficients changes is a linear function of the normally distributed random vector e(tk). Thus, Da(tk) is normally distributed, as well. The expected value of Da(tk) is again the zero vector. We define Ω(DA) as the covariance matrix of the vector Da(tk). Why is Ω(DA) interesting for our purposes? Because Ω(DA) is related to Ω(DR). Indeed, looking at equations [19] and [9], it is clear that:

$$DR(tk) = Y \cdot B \cdot e(tk) = Y \cdot Da(tk) \qquad [\ 20\ ]$$

3

Thus, it is easy to verify that:

$$\Omega(DR) = COV(DR(tk)) = E[DR(tk) \cdot DR(tk)'] =$$
$$= E[Y \cdot Da(tk) \cdot Da(tk)' \cdot Y'] =$$
$$= Y \cdot \Omega(DA) \cdot Y' \qquad [\ 21\ ]$$

Equation [21] finally provides us a method for computing $\Omega(DR)$ through the estimation of $\Omega(DA)$ and the use of the known matrix Y.

The estimation procedure

The estimation procedure implemented in this paper consists of 10 steps relying on the main results of the Feasible GLS estimation approach.

Step 1

For each date t(k) in a sequence t0……..t52 from April 25th 1997 to April 24th 1998 we need to estimate the values of the composite coefficients a1(tk)……a6(tk) of equation [15]:

$$R(y) = \sum_{i=1}^{6} ai\, y^{i-1} + e \qquad [\ 21\ ]$$

where e is a random error term. We define R as a vector of all observed yields from 0.25 years to 30 years. R is a 10x1 vector. ε is a vector or error terms. We finally define the following regressors-matrix, using scaled regressors:

$$Z = \begin{bmatrix} 1 & \dfrac{y1}{30} & \left|\dfrac{y1}{30}\right|^2 & . & . & \left(\dfrac{y1}{30}\right)^5 \\ . & . & . & & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . & . \\ 1 & \dfrac{y10}{30} & \left(\dfrac{y10}{30}\right)^2 & . & . & \left(\dfrac{y10}{30}\right)^5 \end{bmatrix}$$

Z is a (6x10) matrix in this application. Using equation [21], we can write:

$$R = Z \cdot \Upsilon + \varepsilon \qquad [\ 22\ ]$$

where Υ is the following vector of scaled coefficients:

$$\gamma = \begin{vmatrix} \gamma_1 = a1 \\ \gamma_2 = a2 \cdot 30 \\ . \\ . \\ \gamma_6 = a6 \cdot 30^5 \end{vmatrix}$$

We assume that ε is normally distributed with a mean vector equal to zero and a covariance matrix given by $\sigma 2F$, where σ is an unknown scalar and F is a still unknown positive definite matrix.

We will assume that F is of the form:

$$F = \begin{vmatrix} y1^{-2\alpha} & \rho_{12} & \rho_{13} & \rho_{14} & . & . & \rho_{110} \\ \rho_{21} & y2^{-2\alpha} & . & . & . & . & \rho_{210} \\ \rho_{31} & \rho_{32} & . & . & . & . & . \\ \rho_{41} & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & \rho_{910} \\ \rho_{101} & \rho_{102} & . & . & . & . & y30^{-2\alpha} \end{vmatrix}$$

The selected form for F is based on the observation that interest rate-change volatility is generally lower for lower maturities and that the correlation among disturbances arises from correlation in the yield-estimation errors. While we assume that diagonal elements in the matrix vary through time, we maintain that our "rhos" will be constant across the sample, in line with what would be reasonable to assume for the yield estimation errors. It is common practice to assume that this correlation declines when the absolute difference in bond maturities increases. We have still to estimate "alpha" and "rho". To do so, we observe that:

$$VAR(\varepsilon_i) = \sigma^2 y_i^{-2\alpha}$$

$$LN(VAR(\varepsilon_i)) = LN(\sigma^2) + \alpha \cdot \left(-2\, LN(y_i)\right)$$

Thus, if we find a consistent estimator for ε, we can use it to regress its natural log on –2LN(yi). As a result we will obtain a consistent estimator for "alpha"[5].

The residuals in a LSE regression of R on Z will be consistent estimators for the true disturbances. Thus, we will regress:

LN(e) = LN(R – Z[((Z'Z)^(-1))Z'R])

on –2LN(Y) to estimate "alpha" for each of the sample weeks.

---

[5] It is worth noting that the constant term of the proposed regression won't be a consistent estimator for LN(σ2).

4

The following chart shows the pattern for "alpha" across the sample.

Alpha
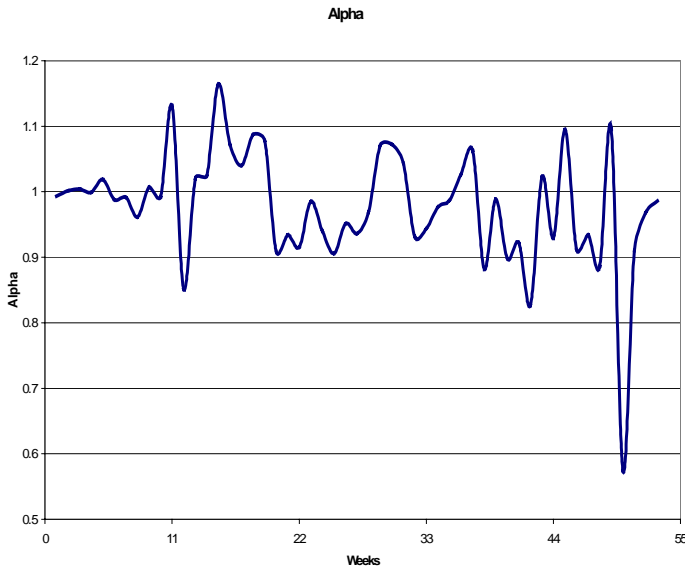
Step 2

A problem remains unsolved: "rhos" are still unknown. Even in this case, we will use a very "natural" approach: our estimator for "rho(ij)" will be the observed correlation among the estimated residuals ei and ej.

Step 3

Step 1 and 2 permitted us to obtain an estimate of the matrix F(tk) for each of the 53 weeks in our sample. These estimates will permit us to provide values for the unknown coefficients in the model. As we assumed that the disturbances are normally distributed, a Feasible GLS estimator for the vector of coefficients Υ(tk) for each of the sample weeks will have the same properties of a Maximum Likelihood Estimator. The proposed estimator is the following:

$$\gamma^{FGLS} = \left(Z' \cdot \hat{F}^{-1} \cdot Z \right)^{-1} \cdot Z' \cdot \hat{F}^{-1} \cdot R$$

We obtain series of 53 estimator vectors for each of the sample weeks. From the definition of Υ(tk), we derive the corresponding series of 53 vectors of estimators for a(tk), i.e. our estimates for ai(tk)....a6(tk), for tk = 0,......,52.

Step 4

We now use the estimated values of the composite coefficients ai(tk) to estimate the covariance matrix of weekly changes in the composite coefficients. In other terms, we can now provide an estimate for the matrix Ω(DA). Each of the entries in Ω(DA) will be replaced by the observed COV(Dai, Daj).

Step 5

For the sequence of maturities of our sample, y1,..., y10, we construct the matrix Y defined in the previous paragraph. Y, Y' and the estimated Ω(DA) will permit us to estimate Ω(DR) through equation [21].

Step 6

We compute the eigenvalues and eigenvectors of the estimated Ω(DR), in order to identify the matrices V(DR) and D(DR), as defined in equation [10].

Step 7

We now extract V'(DR) and D'(DR) through the partition suggested at the bottom of equation [10].

Step 8

We can finally compute the coefficients β_{ij} applying the formula in equation [14] to derive the estimate for the matrix B.

Step 9

It is now time to examine the problem of estimating the β_{i0} coefficients of the baseline polynomial function R0(y) defined in equation [6]. This step is quite short, because these coefficients cannot simply be estimated. Any polynomial representation of the yield curve is consistent with any polynomial representation of the baseline function. Next step will help in clarifying this point.

Step 10

Given the basic representation of the yield curve in equation [2], we now want to examine the problem of computing the fj(y) shift functions and the wj(tk) weights.

As from equation [16] it is clear the nature of ai(tk) as a composite indicator, we can define a(tk) as a Ix1 vector containing the estimated values for a1(tk).....a6(tk) in the week tk.

fj(tk) are estimated simply applying equation [7] and using our estimate for B.

Then define:

$$\beta_0 = \begin{vmatrix} \beta_{10} \\ . \\ . \\ . \\ \beta_{I0} \end{vmatrix} \qquad w(tk) = \begin{bmatrix} w_1(tk) \\ . \\ . \\ . \\ w_J(tk) \end{bmatrix}$$

It is clear that, using the matrix B, we can represent equation [16] in the following matrix form:

$$a(tk) = \beta 0 + B \cdot w(tk) \qquad [\ 23\ ]$$

We already estimated both a(tk) and B. Thus, we can solve equation [23]. As J=I, we are sure that B is invertible. We finally derive the following equation for w(tk):

$$w(tk) = B^{\wedge}(-1) \cdot [a(tk) - \beta 0] \qquad [\ 24\ ]$$

Equation [24] suggests an interesting interpretation that helps us in clarifying step 9. We just said, in equation [1], that the yield at t(k) can be represented as a baseline function plus a linear combination of shift functions. We also said that the yield curve R(t,y) and the R0(y) baseline functions are polynomials of degree I-1. Then, the difference between those two functions has to be a polynomial of degree I-1. This difference polynomial can be expressed exactly as a linear combination of an arbitrary set of I linearly independent polynomials of degree I-1. If the matrix B is nonsingular, then the shift functions are independent polynomials. Thus R(t,y) – R0(y) can be expressed exactly as a linear combination of the shift functions. This argument is valid for each chosen R0(y). Thus, we can calculate a vector w(tk) of weights for any baseline function or for any coefficient vector *β0*. Hence, we can comfortably specify exogenously the benchmark against which we want to measure the fluctuations in the yield curve. We will define the baseline function as the actual yield curve at t0 = April 25th 1997, so that R0(y) = R(y,t0) or *βi0* = ai(t0) for I=1,…,6. The resulting weights will reflect the movement of the yield curve relative to the curve which prevailed in week t0.

We present selected charts for the shift functions and the weights related to the three most important modes of fluctuations of the yield curve, level, slope and curvature.
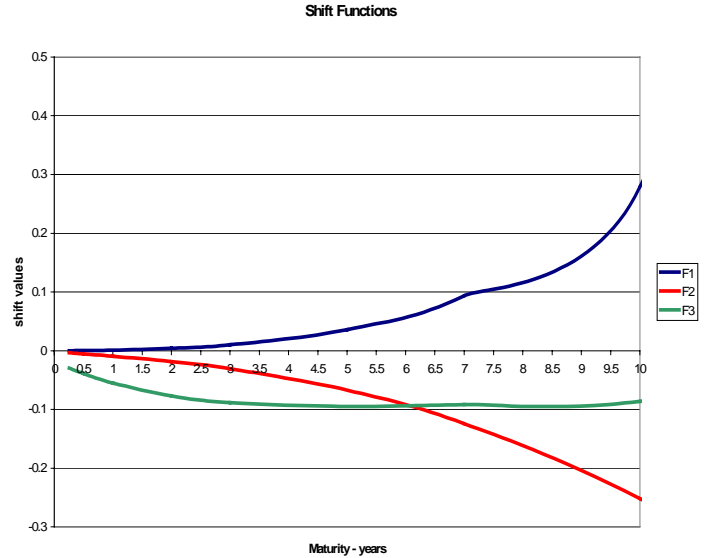
Shift Functions

The twists in the shift functions explain the twists observed in the chart at the beginning of the paper for the term-structure of interest rates in the U.S. F1 changes the level of the yield curve, F2 changes its slope, while F3 determines the curvature (convexity or concavity).

Weights

Short-term interest rates move up (down) and longer-term interest rates move down (up), i.e. the spread goes up (down) as w2(t) increases (decreases). Short and long term interest rates

move down (up) and yields in the middle move up (down), as w3(t) increases (decreases).

The three remaining shift functions, f4, f5 and f6, cause more complex twisting of the yield curve, but the magnitude of their effect on the term structure is almost negligible.

An estimated value of σ for each week of our sample interval has been computed using the unbiased estimator:

$$\hat{\sigma} = \sqrt{\frac{\hat{\varepsilon}' \cdot \hat{F}^{-1} \cdot \hat{\varepsilon}}{n - I}} \qquad [\,25\,]$$

$$\hat{\varepsilon} = R - Z \cdot \gamma^{GLS} \qquad [\,26\,]$$

The following chart plots the resulting estimation over the sample interval:

Standard Deviation (in bps)
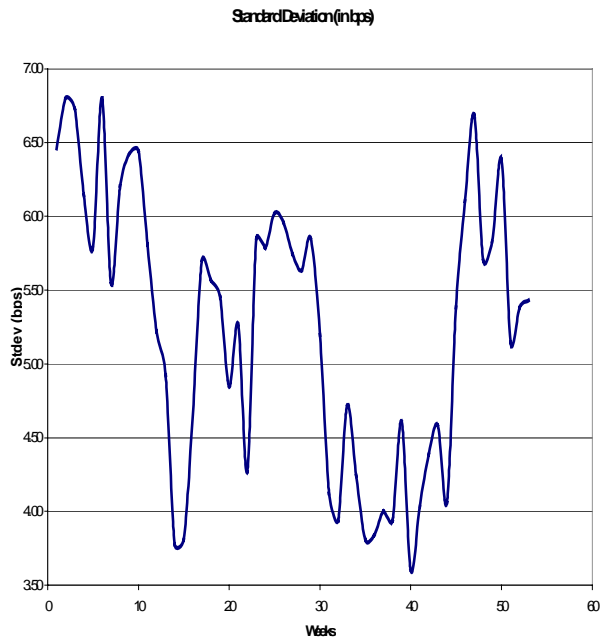
In general, lower values of I, the number of coefficients in the selected polynomial, will shift up the graph and higher values of I will lower it. This reflects the greater relative representation errors associated with polynomial of lower degrees.

Conclusions

The model examined in this paper is useful when an analyst needs to represent a yield curve and its modes of fluctuation with continuous functions rather than with tabulation of the behavior of yields at a finite number of discrete maturities. This paper suggests a 10-step FGLS approach to the estimation of the parameters of a polynomial representation of the yield curve.

Is the proposed model really effective? The shift function F1, positively sloped, rather than parallel to the x-axis, as theoretically expected, casts some doubts regarding some of the initial assumptions. More work is needed to improve some of the assumptions of the model and to reach a better fit.

Bibliography

Kenneth D. Garbade, *Fixed Income Analytics*, MIT Press, Cambridge, MA, 1996.

William H. Greene, *Econometric Analysis*, Prentice Hall, NJ, 1997.