Columbia Business School

Changes in the Treasury Yield Curve: Data Analysis, Models, and Applications

The *yield curve* is the graph of bond yields versus maturities. This note investigates *changes* in the Treasury yield curve and its implications for bond portfolio management. The concept of duration is based on a *parallel shift* of the yield curve, i.e., yields changing by an equal amount for all maturities. A natural question to ask is, "Does the yield curve shift in a parallel manner?" The optimization of bond portfolios using linear programming requires the user to specify future yield curve scenarios. What yield change scenarios and probabilities are reasonable? How has the yield curve changed historically?

Figure 1 shows the Treasury yield curve on three dates in 1988. Casual examination of the figure indicates that the change from October 31, 1988 to November 30, 1988 is roughly parallel, i.e., the yield changes are approximately equal for all maturities. By contrast, the change from November 30, 1988 to December 30, 1988 is not at all parallel. The change is characterized by short-term yields rising significantly and long-term yields dropping slightly. The yield curve changes from upward sloping to almost flat in shape.

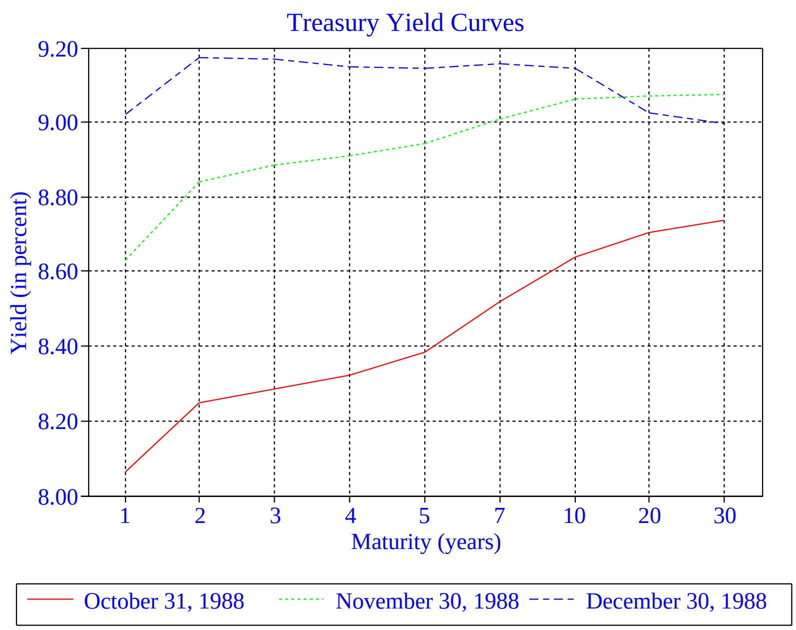

Figure 1.

This note investigates the structure of changes in the yield curve and its implications. First, a preliminary data analysis is carried out, and a model for yield curve changes is proposed. This model was originally suggested by Garbade (1986). Then the results of the analysis are applied to portfolio and hedging problems with Treasury bonds.

1. Historical Yield Curve Changes: A Preliminary Data Analysis

In this section, monthly bond yields from January, 1985 through December, 1988 are examined to determine what fraction of the yield curve changes could be classified as *parallel* and what fraction could be classified as *other* changes.

The yield curve at any point in time can be summarized by a list of yields of bonds of different maturities. For example, the yield curve on October 31, 1988 can be summarized by the yields

$$y = (y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8, y_9) = (8.06, 8.25, 8.29, 8.32, 8.38, 8.52, 8.64, 8.70, 8.74),$$

at corresponding maturities 1, 2, 3, 4, 5, 7, 10, 20, and 30 years, respectively. The yields are expressed in percent. The entire yield *curve* is summarized by nine discrete points. Additional points could be used, but nine is sufficient to illustrate the ideas. On November 30, 1988, the yield curve changed to

$$y' = (y'_1, y'_2, y'_3, y'_4, y'_5, y'_7, y'_7, y'_8, y'_9) = (8.63, 8.84, 8.89, 8.91, 8.94, 9.01, 9.06, 9.07, 9.08).$$

The vector of yield curve changes is

$$\Delta y = (y'_1 - y_1, y'_2 - y_2, \dots, y'_9 - y_9) = (0.57, 0.59, 0.60, 0.59, 0.56, 0.49, 0.42, 0.37, 0.34).$$

The average yield change, denoted $\overline{\Delta y}$, is

$$\overline{\Delta y} = \frac{1}{9} \sum_{i=1}^{9} \Delta y_i = 0.5033.$$

On average, the yield curve increased by 50 basis points in the month of November, 1988.

One can think of the yield curve change from y to y' as happening in two hypothetical steps. The first change to the curve is an average change that affects all maturities equally, i.e., a *parallel shift*, and the second change is the remainder. Algebraically this can be written as $\Delta y = \overline{\Delta y} + \epsilon$, where ϵ represents the vector of remaining changes:

$$\begin{aligned}
\Delta y &= (0.57,\ 0.59,\ 0.60,\ 0.59,\ 0.56,\ 0.49,\ 0.42,\ 0.36,\ 0.34) \\
\overline{\Delta y} &= (0.50,\ 0.50,\ 0.50,\ 0.50,\ 0.50,\ 0.50,\ 0.50,\ 0.50,\ 0.50) \\
\epsilon &= (0.07,\ 0.09,\ 0.10,\ 0.08,\ 0.05,\ -.01,\ -.08,\ -.14,\ -.16)
\end{aligned}$$

(Depending on the context, $\overline{\Delta y}$ is used to represent a single number, the average change, or a vector of numbers all equal to the average change.) Notice that the parallel part of the change represents most of the total change in the yield curve. This observation can be quantified by comparing the sum of squares of the components for each vector:

$$\sum_{i=1}^{9} \Delta y_i^2 = 2.36, \qquad \sum_{i=1}^{9} \overline{\Delta y}_i^2 = 2.28, \qquad \sum_{i=1}^{9} \epsilon_i^2 = 0.08.$$

The total change has a sum of squares of 2.36, the average change has a sum of squares of 2.28, and the remainder has a sum of squares of 0.08. It can be shown that the sum of squares created in this way will always add up, i.e., 2.36 = 2.28 + 0.08. This is a "Pythagorean theorem" that appears frequently in regression analysis and in the analysis of variance. Of the total yield curve shift, 96.4% (= 2.28/2.36) can be attributed to a parallel shift factor and 3.6% (= 0.08/2.36) can be attributed to other factors.

The situation is quite different between November, 1988 and December, 1988. Repeating the previous analysis gives the following results. The yield curve on December 30, 1988 is given by

$$y'' = (9.02, 9.17, 9.17, 9.15, 9.15, 9.16, 9.14, 9.03, 9.00).$$

The vector of changes from November, 1988 to December, 1988, the average change, and remaining changes are given by

$$\begin{aligned}
\Delta y &= (0.39, 0.33, 0.28, 0.24, 0.21, 0.15, 0.08, -.04, -.08)\\
\overline{\Delta y} &= (0.17, 0.17, 0.17, 0.17, 0.17, 0.17, 0.17, 0.17, 0.17)\\
\epsilon &= (0.22, 0.16, 0.11, 0.07, 0.03, -.03, -.09, -.22, -.25)
\end{aligned}$$

The vector Δy shows that the yields of the short-term maturity bonds have increased, but the yields of the long-term maturity bonds have decreased. After accounting for the average change, $\overline{\Delta y}$, the remaining changes are still very significant. That is, the yield curve change from November to December is not primarily due to a parallel shift factor.

The previous results can be quantified using the sum of squares criterion. The total change has a sum of squares of 0.48, the average change has a sum of squares of 0.27, and the remainder has a sum of squares of 0.21. Of the total yield curve shift, only 56.1% (= 0.27/0.48) can be attributed to a parallel shift factor while 43.9% (= 0.21/0.48) can be attributed to other factors.

This procedure was repeated for yield curve changes for each month from January, 1985 to December, 1988. Combining the results for all months gives:

94.8%	average percent of total change due to parallel shift factor
5.2%	average percent of total change due to all other factors

The results show that the vast majority (94.8%) of yield curve changes can be explained by parallel shifts in the yield curve. Only a small percent (5.2%) of the changes are explained by other factors.

2. Historical Yield Curve Changes: A Principal Components Analysis

In the previous section, the relative importance of a parallel shift factor on changes in the Treasury yield curve was investigated. In this section, the results are extended using *principal components analysis*. Using this technique, the factors that affect the yield curve are the *outputs* of the analysis, not the *inputs*. It is not assumed that there is a parallel shift factor that influences changes in the yield curve. In addition to identifying a primary factor, principal components analysis identifies other factors that affect the yield curve, and quantifies the relative importance of each factor. In the next section, the results are used to create risk measures for bonds.

In principal components analysis the structure of the covariance matrix (or the correlation matrix) is analyzed. Since we are interested in changes in the Treasury yield curve, the variables used are changes in yields for various maturities, and the correlation matrix is determined for those variables.

The analysis in this note is performed using data from the yield curve of on-the-run Treasury securities. It would be theoretically more appealing to perform the analysis on the spot yield curve. However, the results would not be significantly different than the ones given here. Also, the analysis is performed on monthly data. Similar results hold for daily or weekly data.

For all consecutive months of data, define the vector of monthly yield changes by

$$\Delta y = (y_1' - y_1, y_2' - y_2, \ldots, y_9' - y_9) = (\Delta y_1, \Delta y_2, \ldots, \Delta y_9),$$

as was done in the previous section. For any yield change variable, Δy_i, its mean and standard deviation can be calculated. For any pair of variables, Δy_i and Δy_j, the correlation can be computed. A MATLAB program to do these calculations is given in Appendix B in Figure 9. The results are shown in Table 1.

Table 1. Correlation Matrix of Monthly Yield Changes using data from January, 1985 through December, 1988

	Δy_1	Δy_2	Δy_3	Δy_4	Δy_5	Δy_6	Δy_7	Δy_8	Δy_9
Δy_1	1.000	0.960	0.948	0.935	0.921	0.887	0.852	0.820	0.789
Δy_2	0.960	1.000	0.992	0.977	0.968	0.956	0.931	0.887	0.880
Δy_3	0.948	0.992	1.000	0.996	0.991	0.978	0.955	0.922	0.897
Δy_4	0.935	0.977	0.996	1.000	0.998	0.985	0.966	0.940	0.903
Δy_5	0.921	0.968	0.991	0.998	1.000	0.992	0.976	0.953	0.916
Δy_6	0.887	0.956	0.978	0.985	0.992	1.000	0.995	0.969	0.949
Δy_7	0.852	0.931	0.955	0.966	0.976	0.995	1.000	0.984	0.968
Δy_8	0.820	0.887	0.922	0.940	0.953	0.969	0.984	1.000	0.969
Δy_9	0.789	0.880	0.897	0.903	0.916	0.949	0.968	0.969	1.000
Std Dev (bp)	39.49	41.67	43.41	44.39	44.43	43.56	43.23	42.85	39.96
Mean (bp)	−.16	−1.63	−2.52	−3.11	−3.52	−4.02	−4.38	−4.65	−4.67
Maturity (yr)	1	2	3	4	5	7	10	20	30

In Table 1, the means and standard deviations of monthly yield changes are given in basis points. The standard deviations of the yield changes were approximately equal across maturities, ranging from a low of 39 basis points to a high of 44 basis points. The standard deviations are a measure of the monthly yield *volatility* of the Treasury bond market. The average monthly yield changes over the time period were close to zero. In fact, a standard statistical test shows that the mean yield changes are not significantly different than zero. The correlations of yield changes are highest for bonds that are close in maturity. Similarly, the correlations decline as the maturity difference in the bonds increases. This implies that bond hedges become less reliable as the maturity difference between the bonds increases.

The structure of the correlation matrix can be better understood using principal components analysis. Mathematical details are given in Appendix A. Additional reading about principal components analysis can be found in Johnson and Wichern (1982), Maxwell (1977), and many other books on multivariate statistical analysis. Principal components analysis is used heavily in many fields, including marketing, psychology, and other social sciences. Principal components analysis can also be used to identify factors in the APT (arbitrage pricing theory) model.

Principal components analysis identifies factors that can be used to model the original variables. The *factor loadings*, f, are easily computed using MATLAB, SAS/IML, GAUSS, APL, or other matrix languages. A MATLAB program to do this is listed in Figure 10 in Appendix B. The results are summarized in Table 2. The first component, f_1, has approximately equal loadings. It can be interpreted as a *parallel shift* component. Note that f_1 does not represent an exact parallel shift in the yield curve. The second component, f_2, has loadings that increase approximately linearly. It can be interpreted as a *slope change* component. The third component, f_3, has loadings that decrease and then increase. It can be interpreted as a *curvature change* component. The first three components from Table 2 are plotted in Figure 2.

Table 2. Principal Components (in basis points)

Effect on	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9
Δy_1	36.47	−13.65	4.67	4.40	1.31	−0.30	−.02	0.01	0.00
Δy_2	40.62	−7.75	2.52	−4.08	−.70	1.76	0.09	−.03	0.00
Δy_3	42.99	−5.37	−1.15	−1.97	−1.49	.17	−.09	0.06	0.00
Δy_4	44.09	−3.48	−3.45	0.35	−1.21	−1.10	−.09	−.02	0.00
Δy_5	44.21	−1.33	−3.90	0.64	−.14	−1.35	0.05	−.04	0.00
Δy_6	43.32	2.92	−2.34	−1.29	2.12	−0.35	0.20	0.03	0.00
Δy_7	42.58	6.87	−1.11	−.66	2.29	1.24	−.20	−.01	0.00
Δy_8	41.31	10.07	−.10	4.79	−1.61	1.57	0.08	0.00	0.00
Δy_9	37.69	11.17	6.75	−1.79	−.50	−1.66	−.02	0.00	0.00
% explained	95.22	3.50	0.67	0.42	0.11	0.08	0.00	0.00	0.00
Cumulative %	95.22	98.72	99.38	99.81	99.92	100.0	100.0	100.0	100.0

The row "% explained" in Table 2 shows how much of the total variation in yield curve changes is explained by each of the factors. The first factor explains 95.2% of the changes in the yield curve. This is almost the same result that was obtained in the previous section. The difference is due to the fact that f_1 is does not represent an exactly parallel shift. The first factor does a slightly better job of explaining shifts in the yield curve compared to an exactly parallel shift. The method of principal components finds a first factor that *best* explains changes in the yield curve. The second factor accounts for 3.5% of the changes in the yield curve. The third factor accounts for 0.7% of the changes in the yield curve. Together, these three factors account for 99.4% of changes in the yield curve. For the rest of this note, the remaining factors will be ignored, because they have a relatively insignificant effect. (While the analysis that follows only uses the first three factors, the fourth or any additional factors could be included in the analysis in a similar manner.)

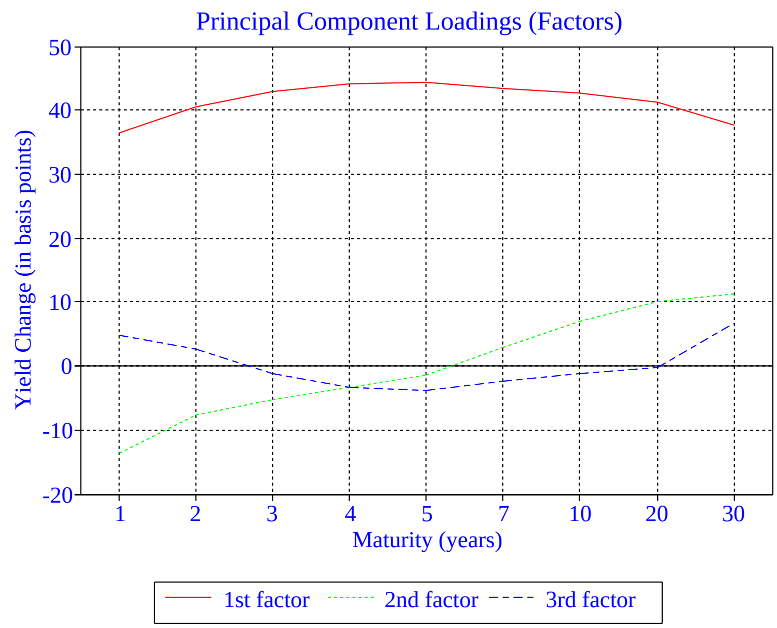

Figure 2.

The components from Table 2 can be used to decompose changes in the yield curve as follows. The change in yield of the i^{th} bond can be written as

$$\Delta y_i = f_{i1}z_1 + f_{i2}z_2 + f_{i3}z_3 + \cdots + f_{i9}z_9, \tag{1}$$

where f_{ij} is the i^{th} loading of factor j and z_j is the magnitude of factor j. Since the last 6 factors have a relatively insignificant effect, equation (1) can be rewritten as

$$\Delta y_i = f_{i1}z_1 + f_{i2}z_2 + f_{i3}z_3 + \epsilon, \tag{2}$$

where the error term, ϵ, is very small, on average.

Equation (2), with the factors from Table 2, can be illustrated as follows. Recall that the vector of yield curve changes from October, 1988 to November, 1988 is given by

$$\Delta y = (0.57, 0.59, 0.60, 0.59, 0.56, 0.49, 0.42, 0.36, 0.34).$$

The total change can be decomposed into four parts. The first three parts correspond to effects from the first three factors. The last part is the remainder. Algebraically this can be written as $\Delta y = f_1 z_1 + f_2 z_2 + f_3 z_3 + \epsilon$:

$$\begin{aligned}
\Delta y \quad &= (0.57,\ 0.59,\ 0.60,\ 0.59,\ 0.56,\ 0.49,\ 0.42,\ 0.36,\ 0.34) \\
1.212 f_1 &= (0.44,\ 0.49,\ 0.52,\ 0.53,\ 0.54,\ 0.52,\ 0.52,\ 0.50,\ 0.46) \\
-1.156 f_2 &= (0.16,\ 0.09,\ 0.06,\ 0.04,\ 0.02,\ -.03,\ -.08,\ -.12,\ -.13) \\
-0.069 f_3 &= (0.00,\ 0.00,\ 0.00,\ 0.00,\ 0.00,\ 0.00,\ 0.00,\ 0.00,\ 0.00) \\
\epsilon \quad &= (-.02,\ 0.01,\ 0.01,\ 0.01,\ 0.00,\ 0.00,\ -.01,\ -.02,\ 0.02)
\end{aligned}$$

The values z_j were determined by solving the system of equations (1).[1] The result was $z_1 = 1.212$, $z_2 = -1.156$, and $z_3 = -0.069$. The relative importance of the factors can be seen by computing the sum of squares of each vector. The sum of squares of the original change vector is $\sum_{i=1}^{9} \Delta y_i^2 = 2.36$. The sum of squares of the next four vectors are 2.28, 0.08, 0.00, and 0.00, respectively. In other words, out of the total change, 96.7% (= 2.28/2.36) is due to f_1, 3.2% (= 0.08/2.36) is due to f_2, and a negligible amount is due to the remaining factors. The results are directly comparable to those obtained in the previous section.

Similarly, the change in the yield curve from November, 1988 to December, 1988 can be explained as follows.

$$\begin{aligned}
\Delta y \quad &= (0.39,\ 0.33,\ 0.28,\ 0.24,\ 0.21,\ 0.15,\ 0.08,\ -.04,\ -.08) \\
0.416 f_1 &= (0.15,\ 0.17,\ 0.18,\ 0.18,\ 0.18,\ 0.18,\ 0.18,\ 0.17,\ 0.16) \\
-1.908 f_2 &= (0.26,\ 0.15,\ 0.10,\ 0.07,\ 0.03,\ -.06,\ -.13,\ -.19,\ -.21) \\
-0.259 f_3 &= (-.01,\ -.01,\ 0.00,\ 0.01,\ 0.01,\ 0.01,\ 0.00,\ 0.00,\ -.02) \\
\epsilon \quad &= (-.01,\ 0.02,\ 0.00,\ -.02,\ -.01,\ 0.02,\ 0.03,\ -.02,\ -.01)
\end{aligned}$$

The values z_j were determined by solving the system of equations (1). The result was $z_1 = 0.416$, $z_2 = -1.908$, and $z_3 = -0.259$. The sum of squares of the original change vector is $\sum_{i=1}^{9} \Delta y_i^2 = 0.48$. The sum of squares of the next four vectors are 0.27, 0.21, 0.00, and 0.00, respectively. In other words, out of the total change, 56.0% (= 0.27/0.48) is due to f_1, 43.3% (= 0.20/0.48) is due to f_2, and a negligible amount is due to the remaining factors.

Equation (2) shows that change in the yield of the i^{th} bond depends on three known factors, f_{i1}, f_{i2}, and f_{i3}, and on three magnitudes z_1, z_2, and z_3:

$$\Delta y_i = f_{i1} z_1 + f_{i2} z_2 + f_{i3} z_3 + \epsilon. \tag{2}$$

[1] Note that the yields and yield changes are expressed in percent, and the principal component loadings are expressed in basis points in Table 2. This means that there is a factor of 100 to be accounted for. For example, the first component of $1.212 f_1$ is 0.44, not 44.0.

If changes to the yield curve are generated by equation (2), where each z_j is a random variable with mean zero and standard deviation one, and z_j and z_k are uncorrelated for $j \neq k$, then the standard deviations and correlations of yield changes (Δy_i) will be the same as in Table 1. If fact, the method of principal components *solves* for the factors f so that this result holds.

Using equation (2), the variance of Δy_i, denoted $\sigma^2(\Delta y_i)$, is given by:

$$\begin{aligned}\sigma^2(\Delta y_i) &\approx f_{i1}^2\sigma^2(z_1) + f_{i2}^2\sigma^2(z_2) + f_{i3}^2\sigma^2(z_3)\\ &\qquad + 2f_{i1}f_{i2}\sigma(z_1,z_2) + 2f_{i1}f_{i3}\sigma(z_1,z_3) + 2f_{i2}f_{i3}\sigma(z_2,z_3)\\ &= f_{i1}^2 + f_{i2}^2 + f_{i3}^2 \end{aligned} \tag{3}$$

The first equality holds approximately because the error term, ϵ, is small. The last equality uses $\sigma^2(z_j) = 1$ (i.e., the standard deviations of the z_j's are one) and $\sigma(z_j, z_k) = 0$ (i.e., z_j and z_k are uncorrelated for $j \neq k$). If all nine factors were included, the result would hold exactly. For example, for a 1-year maturity bill, equation (3) gives

$$\sigma^2(\Delta y_1) \approx 36.47^2 + (-13.65)^2 + 4.67^2 = 1538.19.$$

Taking square roots gives

$$\sigma(\Delta y_1) \approx \sqrt{1538.19} = 39.22.$$

This is quite close to the result in Table 1 that shows the standard deviation of yield changes of a 1-year maturity bill to be 39.49 basis points. In a similar manner, the standard deviations and correlations of the remaining variables generated by equation (2) can be shown to agree with the numbers in Table 1.

The results of this section show that changes in the yield curve are primarily generated by three factors. The three factors correspond to changes in the *level*, *slope*, and *curvature* of the yield curve. Of the three factors, the first factor "explains" 95% of the changes in the yield curve. In addition, the factors are uncorrelated. Roughly speaking, this means that an increase in the level of the yield curve (i.e., a positive realization of z_1) gives no information about the change in the slope of the yield curve.

Principal components analysis offers a general procedure for reducing the complexity of a large data set. When applied to the Treasury yield curve, it shows that the majority of the fluctuations in the yield curve can be described by only three or four important factors.

In the next section it is shown how equation (2) can be used to develop simple risk measures for a bond portfolio. These risk measures can be used for problems in hedging and general portfolio management. A later section shows how equation (2) can be used to simulate yield curve changes. It can also be used to generate scenarios for optimizing bond porfolios.

3. Application to Risk Measures for Bonds and Bond Portfolios

Let P_i and y_i denote the prices and yields of bonds maturing at times t_i, for $i = 1, \ldots, 9$, where $t_1 = 1$, $t_2 = 2$, $t_3 = 3$, $t_4 = 4$, $t_5 = 5$, $t_6 = 7$, $t_7 = 10$, $t_8 = 20$, $t_9 = 30$ and the times are expressed in years. Let $(dP_i/dy_i)/P_i$ be the return sensitivity (i.e., the negative of modified duration) of bond i, for $i = 1, \ldots, 9$. Then the return of bond i for a change in yield of Δy_i can be approximated by

$$R_i = \frac{\Delta P_i}{P_i} \approx \frac{dP_i}{dy_i}\frac{1}{P_i}\Delta y_i.$$

The change in the yield of bond i, Δy_i, can be modeled using equation (2) from the previous section. This gives a return of bond i of

$$R_i \approx \frac{dP_i}{dy_i}\frac{1}{P_i}(f_{i1}z_1 + f_{i2}z_2 + f_{i3}z_3). \tag{4}$$

Define the risk measures β_{ij} by $\beta_{i1} = (dP_i/dy_i)(1/P_i)f_{i1}$, $\beta_{i2} = (dP_i/dy_i)(1/P_i)f_{i2}$, and $\beta_{i3} = (dP_i/dy_i)(1/P_i)f_{i3}$. Thus β_{ij} denotes the return risk of bond i with respect to factor j. The risk parameter β_{i1} indicates the sensitivity of the return of bond i to factor 1, i.e., the return sensitivity to changes in the *level* of the yield curve. Similarly, the risk parameter β_{i2} indicates the return sensitivity to factor 2, i.e., to changes in the *slope* of the yield curve, and β_{i3} indicates the sensitivity to factor 3, i.e., to changes in the *curvature* of the yield curve.

Substituting the definition of β_{ij} into equation (4), the return of bond i can be written as

$$R_i \approx \beta_{i1}z_1 + \beta_{i2}z_2 + \beta_{i3}z_3. \tag{5}$$

Equation (5) says that the return of bond i depends on the risk parameters, β_{ij}, and the random magnitudes z_j, for $j = 1, 2, 3$.

Following the same reasoning that led to equation (3), the variance of the return of bond i can be expressed as a function of its risk parameters as follows:

$$\sigma^2(R_i) \approx \beta_{i1}^2 + \beta_{i2}^2 + \beta_{i3}^2. \tag{6}$$

To illustrate the previous ideas, Table 3 gives the risk measures for nine hypothetical bonds on October 31, 1988.

Table 3. Bond Characteristics and Risk Parameters

Bond	Yield (%)	Coupon (%)	Matur	$-\frac{dP}{dy}\frac{1}{P}$	β_1	β_2	β_3
1	8.06	0.000	1	0.9613	−0.3506	0.1313	−0.0449
2	8.25	8.250	2	1.8096	−0.7350	0.1402	−0.0456
3	8.29	8.750	3	2.5971	−1.1164	0.1395	0.0299
4	8.32	8.750	4	3.3260	−1.4663	0.1157	0.1146
5	8.38	9.000	5	3.9799	−1.7596	0.0529	0.1554
6	8.52	8.625	7	5.1816	−2.2448	−0.1513	0.1214
7	8.64	9.250	10	6.5198	−2.7760	−0.4478	0.0723
8	8.70	9.250	20	9.2943	−3.8393	−0.9357	0.0093
9	8.74	9.125	30	10.5036	−3.9586	−1.1735	−0.7086

Equation (5) and the results in Table 3 can be illustrated as follows. Suppose the magnitude of the first factor, z_1, is 1, and the magnitudes of the other two factors are zero. According to Table 2, this means that the yield curve will increase in a nearly parallel manner, by about 40 basis points. For the 1-year maturity bill, this will cause a return −0.35%.[2] For the 30-year bond, this will cause a return of −3.96%. The 30-year bond has more than 10 times the return risk to changes in the first factor compared to the 1-year bill. Now suppose that $z_1 = 1$, $z_2 = 1$, and $z_3 = 0$. Then the 1-year bond will have a return of −0.22% (= −0.35% + 0.13%) and the 30-year bond will have a return of −5.13% (= −3.96% − 1.17%). The change in slope further depresses the return of the 30-year bond, but has an offsetting effect on the 1-year bill.

The return risk of a bond can be succinctly summarized by its triple of betas, $(\beta_1, \beta_2, \beta_3)$. From equation (6), the standard deviation of the return of the bond is determined by the triple of betas. Each β_j indicates the contribution to the return risk of the bond due to changes in factor j of the yield curve.

Risk Measures for a Portfolio of Bonds

The previous results are easily extended to risk measures for a portfolio of bonds. Suppose an investor owns a quantity Q_i of bond i that is priced at P_i. The initial value of the portfolio is $W = \sum_{i=1}^n P_i Q_i$. Let $x_i = P_i Q_i / W$ denote the fraction of the value of the portfolio that is invested in bond i. Then the return of the portfolio, R_P, is given by

$$\begin{aligned}
R_P &= \frac{\Delta W}{W} = \frac{1}{W}\sum_{i=1}^n Q_i \Delta P_i \\
&= \sum_{i=1}^n \frac{Q_i P_i}{W}\frac{\Delta P_i}{P_i} = \sum_{i=1}^n x_i \frac{\Delta P_i}{P_i} \\
&\approx \sum_{i=1}^n x_i(\beta_{i1} z_1 + \beta_{i2} z_2 + \beta_{i3} z_3) \qquad \text{(from equation (5))} \\
&= (\sum_{i=1}^n x_i \beta_{i1}) z_1 + (\sum_{i=1}^n x_i \beta_{i2}) z_2 + (\sum_{i=1}^n x_i \beta_{i3}) z_3 \\
&= \beta_1^P z_1 + \beta_2^P z_2 + \beta_3^P z_3. \qquad (7)
\end{aligned}$$

The last equality defines the portfolio risk parameters $\beta_j^P = \sum_{i=1}^n x_i \beta_{ij}$, for $j = 1, 2, 3$. The main result is that the risk parameters of the portfolio are a linear combination of the risk parameters of the bonds in the portfolio. The risk of the portfolio of bonds can be summarized by the triple of parameters β_1^P, β_2^P, and β_3^P.

Following the same reasoning that led to equation (3), the variance in the return of the portfolio can be written as

$$\sigma^2(R_P) = \sum_{j=1}^9 (\beta_j^P)^2 \approx (\beta_1^P)^2 + (\beta_2^P)^2 + (\beta_3^P)^2. \qquad (8)$$

[2] Since the yield changes were expressed in percent, the β's given in Table 2 express the return in percent as well.

A portfolio is hedged, or protected, against changes in the yield curve if $R_P = 0$. Since the average yield changes are approximately zero, $R_P = 0$ when $\sigma^2(R_P) = 0$. From equation (8), this happens if $\beta_j^P = 0$ for all j. Also, $\sigma^2(R_P)$ is very nearly zero if $\beta_j^P = 0$ for $j = 1, 2, 3$.

For example, suppose the available bonds are those in Table 3 and suppose an investor wants to hedge \$1,000,000 face amount of bond 4 using bonds 1, 5, and 9. What quantities should be used?

The idea is to solve for Q_1, Q_5, and Q_9 so that the risk parameters of the resulting portfolio is zero. Using $Q_4 = 1$ and the parameters from Table 3, the equations for the portfolio betas are:

$$\begin{aligned}
\beta_1^P &= -0.3506Q_1 - 1.7596Q_5 - 3.9586Q_9 - 1.4663 = 0 \\
\beta_2^P &= 0.1313Q_1 + 0.0529Q_5 - 1.1735Q_9 + 0.1157 = 0 \\
\beta_3^P &= -0.0449Q_1 + 0.1554Q_5 - 0.7086Q_9 + 0.1146 = 0
\end{aligned}$$

Setting the portfolio betas to zero and solving gives $Q_1 = -0.421$, $Q_5 = -0.786$, and $Q_9 = 0.016$. That is, the investor should short \$421,000 face amount of bond 1, short \$786,000 face amount of bond 5, and purchase \$16,000 face amount of bond 9. Since changes in the yield curve can be summarized by three independent factors, it takes three bonds to (almost completely) hedge any given position.

Using equation (8), the variance of return is

$$\sigma^2(R_P) = \sum_{j=1}^{9} (\beta_j^P)^2 = \sum_{j=4}^{9} (\beta_j^P)^2,$$

where the last equality follows because $\beta_j^P = 0$ for $j = 1, 2$, and 3. Computing β_j^P for $j = 4, \ldots, 9$, the previous equation gives $\sigma(R_P) = 0.05\%$. The standard deviation of the return of the portfolio is very nearly zero.

Now suppose that only bonds 1 and 9 are available to hedge the position in bond 4. Then the best that the investor can hope to do is to solve for Q_1 and Q_9 so that $\beta_1^P = 0$ and $\beta_2^P = 0$. The equations to solve are:

$$\begin{aligned}
\beta_1^P &= -0.3506Q_1 - 3.9586Q_9 - 1.4663 = 0 \\
\beta_2^P &= 0.1313Q_1 - 1.1735Q_9 + 0.1157 = 0
\end{aligned}$$

The solution is $Q_1 = -2.340$ and $Q_9 = 0.163$. In this case, $\beta_3^P = 0.335$ and the standard deviation of the return of the portfolio is $\sigma(R_P) = 0.346\%$. If fewer than three bonds are used to hedge, substantial return risk can remain.

The β_j^P risk measures can also be used for speculating. Suppose that an investor does not have a view about changes in the level or curvature of the yield curve, but does expect that the yield curve will be steeper in the future. The investor could form a portfolio with $\beta_1^P = 0$, $\beta_2^P > 0$, and $\beta_3^P = 0$. Since a steepening of the yield curve corresponds to $z_2 > 0$, a portfolio with $\beta_2^P > 0$ will experience a positive return. The return of the portfolio will be insensitive to changes in the level and the curvature of the yield curve because $\beta_1^P = 0$ and $\beta_3^P = 0$.

In *Bond Analytics*, the quadratic program $(QPBH)$ was used to solve the bond hedging problem. In this section, the quadratic program has been simplified to solving a linear system of equations. The solution of $(QPBH)$ and the linear system of equations will be *nearly* identical because the yield change factors are consistent with the standard deviations and correlations of bond yield changes. A simpler system of linear equations suffices to solve the quadratic programming problem because the "work" was already done in computing the principal components.

The three parameters β_1^P, β_2^P, and β_1^P are simple and useful measures of the return risk of a bond portfolio. Together they do a much better job of measuring risk than the concept of duration alone. While the concept of duration relies heavily on the assumption of parallel shifts in the yield curve, the β_j^P measures of risk explicitly take into account non-parallel movements of the yield curve.

The analysis done here focused on the yield curve defined by the most actively traded Treasury bonds, i.e., the on-the-run issues. To the extent that the yields of other bonds move independently of the yields of on-the-run Treasury bonds, other bonds will have additional risk factors. For example, off-the-run Treasuries, agency, corporate, and municipal bonds have a "spread risk" associated with the change in the difference between their yields and the yields of comparable on-the-run Treasury issues. The analysis this section can be extended to include risk measures for bonds other than on-the-run Treasury securities. (This is another idea for a project for the course.)

4. Application to Bond Pricing

This section illustrates how principal components analysis can be used to price bonds. Suppose that on October 31, 1988, an investor holds a portfolio of the nine bonds given in Table 3. The investor wishes to determine the value of the portfolio at the end of the next month. However, the database service can only supply prices and yields for bonds 1, 4, 7, and 9. The prices and yields of the other bonds are unavailable. How can the investor estimate the prices and yields of the remaining bonds?

From equation (1), the change in yield of the i^{th} bond can be written as

$$\Delta y_i = f_{i1}z_1 + f_{i2}z_2 + f_{i3}z_3 + \cdots + f_{i9}z_9, \tag{1}$$

where f_{ij} is the i^{th} component of factor j and z_j is the magnitude of factor j. The investor has information about the four yield changes Δy_1, Δy_4, Δy_7 and Δy_9. Since the principal component loadings f_{ij} are known, the investor can use equation (1) to solve for four of the z's: z_1, z_2, z_3, and z_4.

To give a numerical example, suppose that $\Delta y_1 = -0.1417$, $\Delta y_4 = -0.3444$, $\Delta y_7 = -0.3819$, and $\Delta y_9 = -0.3300$. Then using the first four factors f_j and the rows corresponding to bonds 1, 4, 7, and 9 the system of equations to solve is:

$$\begin{aligned}
0.3647z_1 - 0.1365z_2 + 0.0467z_3 + 0.0440z_4 &= -0.1417\\
0.4409z_1 - 0.0348z_2 - 0.0345z_3 + 0.0035z_4 &= -0.3444\\
0.4258z_1 + 0.0687z_2 - 0.0111z_3 - 0.0066z_4 &= -0.3819\\
0.3769z_1 + 0.1117z_2 + 0.0675z_3 - 0.0179z_4 &= -0.3300
\end{aligned}$$

The f_{ij} are from Table 2. The solution is $z_1 = -0.7827$, $z_2 = -0.5117$, $z_3 = 0.6034$, and $z_4 = 1.0383$. Setting $z_5 = \cdots = z_9 = 0$, equation (1) can be used to predict the remaining values of Δy_i. The predicted changes in yield are $\Delta y_2' = -0.3055$, $\Delta y_3' = -0.3364$, $\Delta y_5' = -0.3561$, $\Delta y_6' = -0.3816$, and $\Delta y_8' = -0.3257$. The predicted changes in yield can be used to predict the prices of the bonds using the price-yield equation.

5. Application to Scenario Generation

Equation (2) showed how the yield change of a bond could be decomposed using three known factors and three independent magnitudes. Denote the yield change of bond i by Δy_i, the three known factors by f_{i1}, f_{i2}, and f_{i3}, and the three random magnitudes by z_1, z_2, and z_3. Equation (2) relates these quantities by

$$\Delta y_i = f_{i1}z_1 + f_{i2}z_2 + f_{i3}z_3 + \epsilon. \tag{2}$$

Suppose that changes to the yield curve are generated by equation (2), where each z_j is a random variable with mean zero and standard deviation one, and z_j and z_k are uncorrelated for $j \neq k$. In short, suppose that the covariance matrix of the z's is the identity matrix, i.e., $\mathrm{Cov}(z) = I$. Then the standard deviations and correlations of yield changes (Δy_i) will be the same as in Table 1.

Thus, equation (2) offers a method for *simulating* yield curve changes that are consistent with a covariance matrix of historical yield changes. The procedure is to generate random values for z_1, z_2, and z_3 such that each z_j has mean zero (by assumption future yield curve changes have a mean of zero) and standard deviation one. Also, z_j and z_k should be uncorrelated for $j \neq k$. For example, each z_j could be generated from a standard normal distribution. The reason for this choice of distribution is given in the next section. This procedure could be repeated over and over to generate as many yield curve scenarios as desired.

In some applications it is desirable to generate as few scenarios as possible that are still consistent with historical yield curve changes. The problem is to choose scenarios (z_1, z_2, z_3) so that $\mathrm{Cov}(z) = I$. One easy way to do this is illustrated in Table 4.

Table 4. Yield Curve Scenario Generation

	z_1	z_2	z_3	Probability
Scenario 1	1.732	0	0	1/6
2	−1.732	0	0	1/6
3	0	1.732	0	1/6
4	0	−1.732	0	1/6
5	0	0	1.732	1/6
6	0	0	−1.732	1/6

Table 4 shows how to generate $n = 6$ scenarios that are consistent with a covariance matrix of historical yield curve changes. To prove this consistency, it is necessary to show that $\mathrm{Cov}(z) = I$. Since the expected value of each z_j is zero, the variance of each z_j is

$$\mathrm{Var}(z_j) = E(z_j^2) = 1/6[1.732^2 + (-1.732)^2 + 0 + 0 + 0 + 0] = 1.$$

Since $n = 6$ scenarios are generated, each nonzero value of z_j is chosen to be $\pm\sqrt{n/2}$ precisely so that $\mathrm{Var}(z_j) = 1$. A variance of z_j of one implies $\sigma(z_j) = 1$. Similarly, it is easy to show that $\mathrm{Cov}(z_j, z_k) = 0$ for $j \neq k$.

Example. Using the factors from Table 2 and the z-values from Table 4, six yield curve change scenarios are generated and shown in Table 5. Each scenario occurs with equal probability 1/6. The final yield curves are calculated by adding the initial yield curve to the yield curve changes. Since the factors from Table 2 were generated from monthly data, the yield curve scenarios reflect possible yield curves one month in the future.

Table 5. Yield curve change scenarios (in basis points)

Scenario	Bond 1	2	3	4	5	6	7	8	9
1	63.2	70.4	74.5	76.4	76.6	75.0	73.8	71.6	65.3
2	−63.2	−70.4	−74.5	−76.4	−76.6	−75.0	−73.8	−71.6	−65.3
3	23.6	13.4	9.3	6.0	2.3	−5.1	−11.9	−17.4	−19.3
4	−23.6	−13.4	−9.3	−6.0	−2.3	5.1	11.9	17.4	19.3
5	8.1	4.4	−2.0	−6.0	−6.8	−4.1	−1.9	−0.2	11.7
6	−8.1	−4.4	2.0	6.0	6.8	4.1	1.9	0.2	−11.7

Obviously six scenarios cannot adequately represent the full distribution of yield curves one month in the future. Nevertheless, they represent a range of outcomes that are consistent with the covariance matrix of historical yield curve changes. The scenarios given are more representative of possible scenarios than are simple parallel yield curve shifts.

If a more detailed range of scenarios is needed, equation (2) can be used to simulate a larger range of future scenarios. In this case, more information is needed on the distribution of the z-values. That is the subject of the next section.

The simulation procedure described in this section is used in *Structured Bond Portfolios*. This simulation procedure was used before in the section "Model Risk" in *Portfolio Optimization*. In that application, stock returns were simulated. The simulated stock returns were generated to be consistent with a given vector of mean returns and a given covariance matrix. Equation (2) was also used to do that simulation (with a slight adjustment for the nonzero mean returns).

6. Analysis of the z-values

In the previous section it was shown how to generate yield curve scenarios consistent with a covariance matrix of historical yield curve changes. However, there are many distributions of z-values that satisfy $\mathrm{Cov}(z) = I$. Are the z-values normally distributed? lognormal? other? In order to identify a reasonable distribution, the historical z-values can be examined.

Once the yield curve change factors have been identified, the z-values of the historical changes can be found by solving equation (1), which is repeated next.

$$\Delta y_i = f_{i1}z_1 + f_{i2}z_2 + f_{i3}z_3 + \cdots + f_{i9}z_9. \tag{1}$$

Using equation (1), each historical yield curve change can be associated with a vector of z-values. The z_1-value summarizes the level of the yield curve at any given time. Similarly, the z_2-value summarizes the slope and the z_3-value summarizes the curvature of the yield curve at any given time.

The z-values obtained from solving equation (1) can be analyzed to better understand their distribution. A total of 48 months of data from January, 1985 through December, 1988 was analyzed. The same approach could be used to analyze weekly or daily yield curve changes. The MATLAB program PRINCOMP.M that does the calculations of this section is listed in Figure 10 in Appendix B.

Figures 3 to 5 give frequency histograms of z_1, z_2 and z_3, respectively. The frequency histograms appear to be consistent with *standard normal distributions* of z-values. However, much more data would be needed to rigorously test this hypothesis.

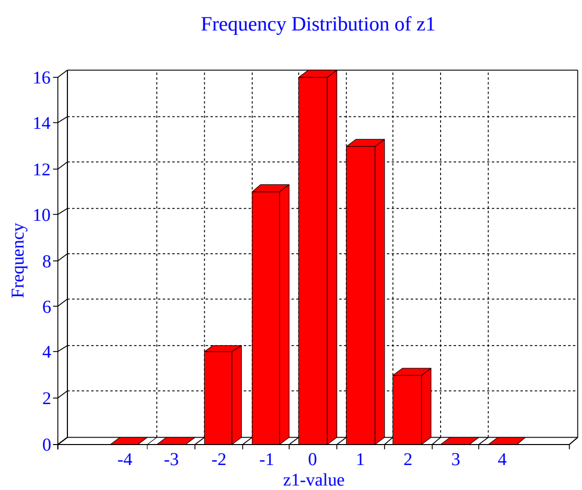

Figure 3.

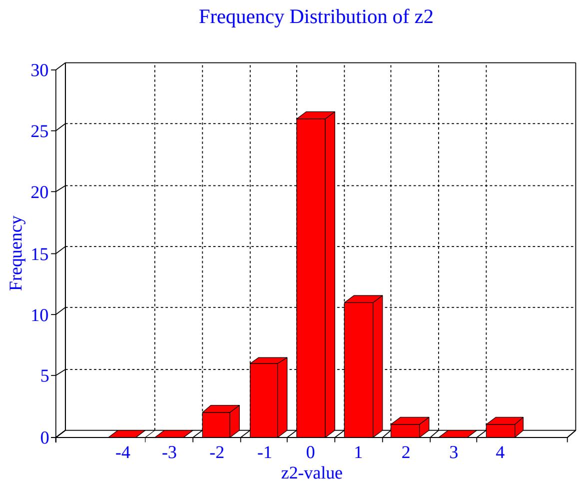

Figure 4.

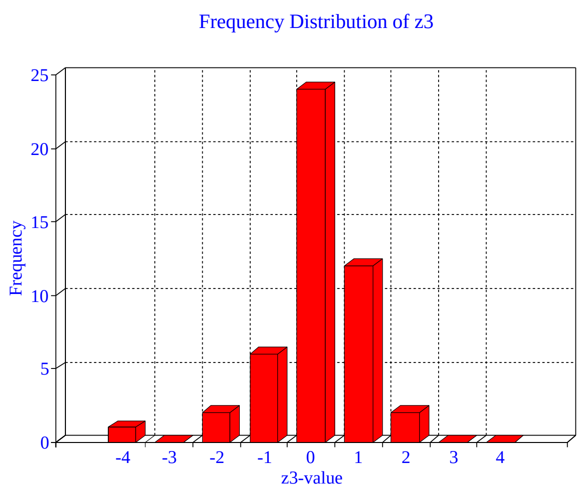

Figure 5.

The z-values can also be cumulated to examine their behavior over time. The time series of z-values can be used to determine if the yield curve is *mean-reverting*. For example, if the yield curve reaches a high level, is it more likely to decline than increase? This behavior would be indicated by z_1-values that don't stray too far from zero. Inefficiencies in the bond market could be observed in the time series behavior of the z-values.

Cumulative values for z_1, z_2, and z_3 are given in Figure 6. Cumulative values for z_4, z_5, and z_6 are given in Figure 7. The graph of z_1 in Figure 6 shows that the level of the yield curve declined from January, 1985 through July, 1986. Then the level rose until December, 1988, but not back to its original level.

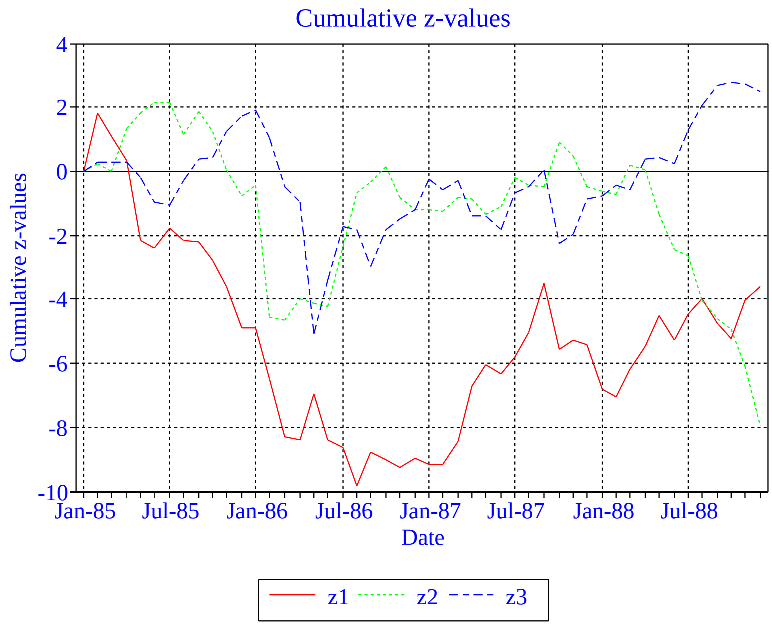

Figure 6.

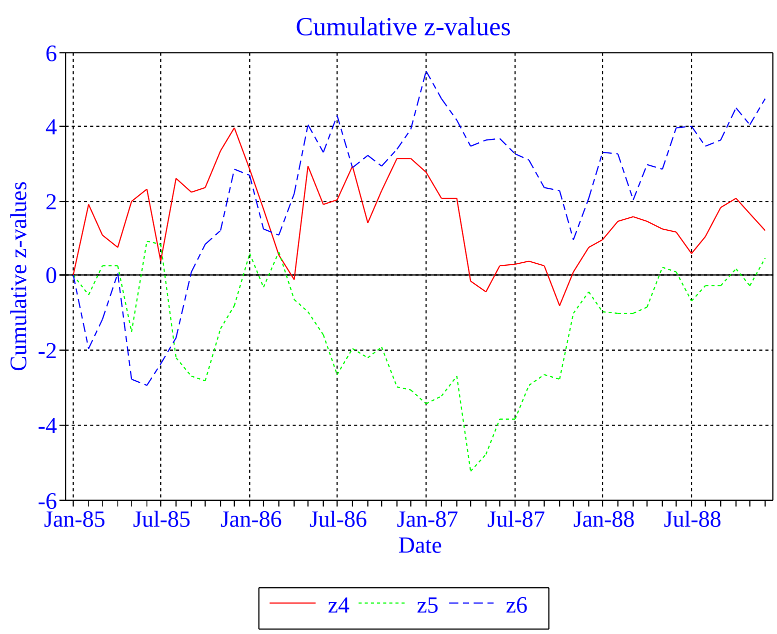

Figure 7.

7. Homework Problems

1. Compute 95% confidence intervals for each of the nine mean yield changes given in Table 1. For each confidence interval, specify whether or not zero is contained in the interval. Which of the mean yield changes are significantly different than zero?

2. *(Risk Measures for Bonds and Bond Portfolios)* Suppose you want to bet on a steepening of the yield curve while remaining hedged against parallel shifts and changes in curvature. In other words, you want to design a portfolio with maximal β_2^P but with $\beta_1^P = \beta_3^P = 0$. Using the β_{ij}s in Table 3, choose portfolio weights $x_1, \ldots, x_9$ (summing to 1) to solve this problem. Specifically, address the following:

   (a) Can 1^{st}-factor risk be hedged out if all x_i are constrained to be nonnegative? Answer this question in terms of the β_{i1}s in the table and explain why your answer makes sense financially.

   (b) Solve the portfolio optimization problem described above. Specifically, choose portfolio weights $x_1, \ldots, x_9$, summing to 1, that maximize β_2^P while setting β_1^P and β_3^P equal to zero. Limit short sales by imposing the constraint $x_i \geq -2$, $i = 1, \ldots, 9$. (This is a linear program.)

3. (*Principal components analysis of daily yield curve changes*) The file BDAILY.PRN contains daily bond yield and return information from December 12, 1988 through December 7, 1989. The file BDAILY.PRN is in the same format as BMONTHLY.PRN which is described in lines [5] – [9] of Figure 10. Determine the principal component loadings (as in Table 2) corresponding to daily yield changes. Reproduce the graphs in Figures 2 through 7 for daily data. Are there any significant differences between the daily and monthly results? Perform a statistical goodness-of-fit test for the hypothesis that z_1 is normally distributed. Repeat the test for z_2 and z_3.

4. (*Security return simulation and efficient frontier estimation*) Use the analysis in the section "Application to Scenario Generation" to simulate 36 months of security returns for the true parameters given in Figure 13 of *Portfolio Optimization*. Use the simulated returns to estimate the mean security returns and their covariances. Compare the true efficient frontier (using the true parameters in Figure 13) with the efficient frontier based on your estimated parameters. How different are the results? How much closer is the estimated efficient frontier to the true frontier if 72 months of simulated data are used?

8. References

[1] Kenneth Garbade, “Modes of fluctuation in bond yields — an analysis of principal components,” Bankers Trust, June, 1986, No. 20, reprinted in Garbade, K., *Fixed Income Analytics*, MIT Press, Cambridge, Mass, 1996.

[2] Johnson and Wichern, *Applied Multivariate Statistical Analysis*, Prentice-Hall, Englewood Cliffs, New Jersey, 1982.

[3] Maxwell, *Multivariate Analysis in Behavioural Research*, Wiley, New York, 1977.

9. Appendix A: Mathematics of Principal Components Analysis

This appendix gives a brief description of the mathematics involved in computing the principal component loadings given in Table 2. A MATLAB program to carry out the calculations is given in Appendix B.

Let Δy denote the vector of yield changes and let $C = \text{Cov}(\Delta y)$ denote the covariance matrix of the yield changes. The covariance matrix C can be decomposed as

$$C = VDV^T,$$

where V is the matrix whose columns are eigenvectors of C, D is the diagonal matrix of eigenvalues of C, and V^T is the transpose of V. The eigenvalue matrix is normalized so that the length of each eigenvector is one, i.e., $VV^T = I$. Furthermore, assume that the columns of V are ordered from the largest eigenvalue to the smallest.

Let $D^{1/2}$ denote the diagonal matrix whose elements are the square-roots of the eigenvalues. Let $D^{-1/2}$ denote the diagonal matrix whose elements are the reciprocals of the square-roots of the eigenvalues. Define the linear transformation from the vector z to Δy by

$$\Delta y = VD^{1/2}z = Fz,$$

where $F = VD^{1/2}$. Now if the covariance matrix of the z's is the identity matrix (i.e., z_j has standard deviation 1 for all j and z_j and z_k are uncorrelated for all $j \neq k$) then

$$\text{Cov}(Fz) = F\text{Cov}(z)F^T = (VD^{1/2})\, I\, (VD^{1/2})^T = VDV^T = C.$$

That is, if yield curve changes Δy are generated by Fz where $\text{Cov}(z) = I$, then the resulting covariance matrix of yield changes will be consistent with the original covariance matrix C.

The columns of F are the vectors of *principal component loadings,* or *factors*, and the vector z is the vector of *principal components.*

Some intuition behind principal components analysis can be gained by considering a small example. Suppose that there are only two bonds. Denote their yield changes by $(\Delta y_1, \Delta y_2)$. Each yield change vector can be plotted as a point in two-dimensional space as illustrated by the shaded region in Figure 8. If all of the changes were exactly parallel, then the scatter plot would lie on the line $\Delta y_1 = \Delta y_2$.

Principal components analysis solves for a vector f_1 of length one that explains as much of the variability of Δy_1 and Δy_2 as possible. In the example in Figure 8, f_1 lies on the line $\Delta y_1 = \Delta y_2$. Next, principal components analysis solves for the vector f_2 of length one that is perpendicular to f_1 and explains as much of the remaining variability as possible.

The situation is harder to visualize in higher dimensions, but the procedure is the same. It turns out that the sequence of vectors, i.e., factors, that explain the maximum amount of remaining variability are the eigenvectors of the covariance matrix of yield changes. For a more complete discussion of principal components analysis, see Johnson and Wichern (1982) or Maxwell (1977).

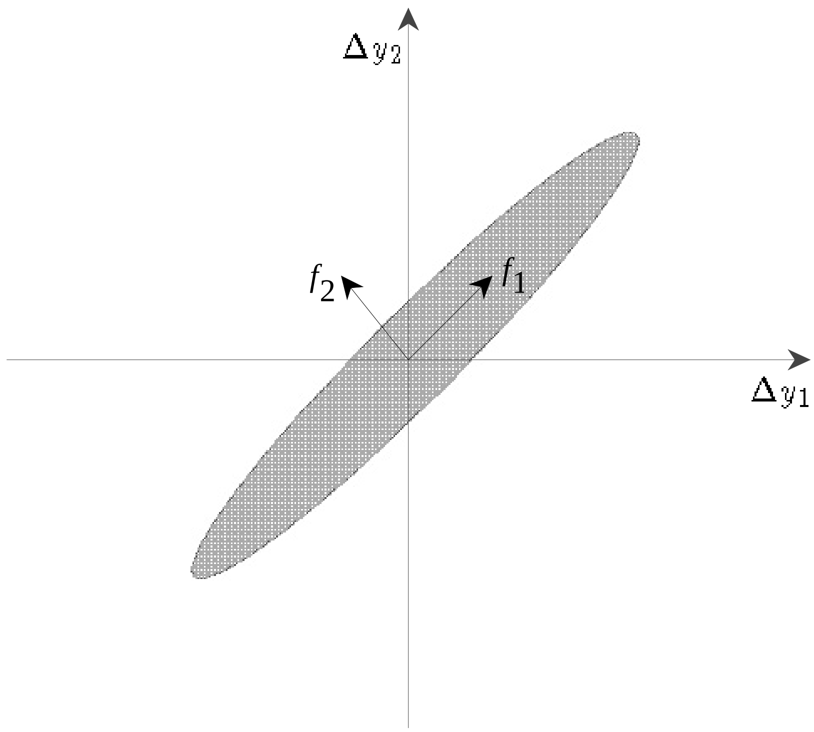

Figure 8. Illustration of Principal Components Analysis

10. Appendix B: MATLAB Programs

This appendix gives MATLAB programs that can be used to do the computations in this note. Figure 9 contains a listing of the MATLAB program BONDYLD.M that is used to compute the bond yield change statistics given in Table 1. Note that this is essentially the same program that was appeared in the Appendix to *Bond Analytics*. The main difference is that is was run on monthly data rather than daily data.

PRINCOMP.M is a MATLAB program that performs the computations described in Appendix A and used throughout this note. A listing of the program is given in Figure 10.

The first lines of PRINCOMP.M are similar to the program BONDYLD.M. Line [12] is used to read the data into the matrix `bmonthly`. Lines [11] and [13] are used to extract only the yield information. Line [14] creates a matrix of yield differences, and line [15] computes the covariance matrix of yield changes.

The main work of the program is done in line [16] where the eigenvalues and eigenvectors are computed. Lines [17] – [19] are used to sort the results in order of decreasing eigenvalues. The principal component loadings, i.e., the f matrix, is computed in line [20].

Lines [23] – [27] display the results, including the explanatory power of each factor. Line [30] is used to compute the z-values from equation (1). Histograms of z-values are plotted in lines [32] – [37]. Lines [39] – [49] are used to plot cumulative z-values.

```
% bondyld.m
% program to compute bond yield change statistics

% the data in the file bmonthly.prn is in the format:
%  (1) date, (2) y-1yr, (3) r-1yr, (4) y-2yr, (5) r-2yr, (6) y-3yr,
%  (7) r-3yr, (8) y-4yr, (9) r-4yr, (10) y-5yr, (11) r-5yr,
%  (12) y-7yr, (13) r-7yr, (14) y-10yr, (15) r-10yr, (16) y-20yr,
%  (17) r-20yr, (18) y-30yr, (19) r-30yr

yld_col   = [ 2  4  6  8 10 12 14 16 18 ];

load        bmonthly.prn;
yld_mat   = bmonthly(:,yld_col);
yld_chng  = diff(yld_mat);

disp('means'),   disp(mean(yld_chng));
disp('std dev'), disp(std(yld_chng));  pause;
disp('correlation matrix'), disp(corrcoeff(yld_chng));

% end of bondyld.m
```

Figure 9. MATLAB program to compute bond yield change statistics

```
% princomp.m
% program to compute principal component loadings for
% bond yield change data

% the data in the file bmonthly.prn is in the format:
%  (1) date, (2) y-1yr, (3) r-1yr, (4) y-2yr, (5) r-2yr, (6) y-3yr,
%  (7) r-3yr, (8) y-4yr, (9) r-4yr, (10) y-5yr, (11) r-5yr,
%  (12) y-7yr, (13) r-7yr, (14) y-10yr, (15) r-10yr, (16) y-20yr,
%  (17) r-20yr, (18) y-30yr, (19) r-30yr

yld_col  = [ 2  4  6  8 10 12 14 16 18 ];
load       bmonthly.prn;
yld_mat  = bmonthly(:,yld_col); % only keep yield columns
yld_chng = diff(yld_mat);       % compute matrix of yield changes
yld_cov  = cov(yld_chng);       % compute covariance of yield changes
[v, d]   = eig(yld_cov);        % compute eigenvalues and eigenvectors
[tmp, k] = sort(-diag(d));      % sort v and d into decreasing order
v        = v(:,k);
d        = d(k,k);
f        = v*sqrt(d);           % compute principal component loadings
expl     = 100*diag(d)/sum(diag(d));  % compute percent explained

% display results
disp('principal components (in basis points)');
disp(100*f); pause;
disp('percent explained'), disp(expl');
disp('cumulative percent explained'), disp(cumsum(expl'));

% compute and plot the distribution of z-values
z = f \ yld_chng';  zbin = [-3 : 3];

hist(z(1,:),zbin); title('Frequency Distribution of z1');
grid; xlabel('z1-value'); ylabel('Frequency'); pause;
hist(z(2,:),zbin); title('Frequency Distribution of z2');
grid; xlabel('z2-value'); ylabel('Frequency'); pause;
hist(z(3,:),zbin); title('Frequency Distribution of z3');
grid; xlabel('z3-value'); ylabel('Frequency'); pause;

% plot cumulative z-values
[nr, nc] = size(z);
z = [zeros(nr,1)  z]; z = cumsum(z');

plot(0:nc,z(:,1), 0:nc,z(:,2), 0:nc,z(:,3));
title('Cumulative z1, z2, and z3-values'); xlabel('Time'); grid;
ylabel('Cumulative z-values'); pause;

plot(0:nc,z(:,4), 0:nc,z(:,5), 0:nc,z(:,6));
title('Cumulative z4, z5, and z6-values'); xlabel('Time'); grid;
ylabel('Cumulative z-values'); pause;

% end of princomp.m
```

Figure 10. MATLAB program to compute principal component loadings