Case Study: Factor Modeling

Dr. Kempthorne

October 28, 2013

Contents

1 Load R Libraries and Federal Reserve Data

The R script ("fm_casestudy_7_0.r") collects daily US Treasury yield data from
FRED, the Federal Reserve Economic Database, and stores them in the R
workspace "casestudy_7.RData".

The following commands re-load the data. The series are constant-maturity
yields on US Treasury bills, notes and bonds.

```
> source("fm_casestudy_0_InstallOrLoadLibraries.r")
> # load the R workspace created by the script file
> #    fm_casestudy_7_0.r
>
> dbnames0<-load(file="fm_casestudy_7_0.RData")
>  print(dbnames0)

[1] "fred.data00" "date.start"  "date.end"

>
>

> library ("graphics")
> library("quantmod")
> fred.data00.0<-na.omit(fred.data00)
> head(fred.data00.0)

           DGS3MO DGS6MO DGS1 DGS2 DGS3 DGS5 DGS7 DGS10 DGS20
2000-01-03   5.48   5.81 6.09 6.38 6.42 6.50 6.65  6.58  6.94
2000-01-04   5.43   5.75 6.00 6.30 6.34 6.40 6.56  6.49  6.84
2000-01-05   5.44   5.74 6.05 6.38 6.43 6.51 6.68  6.62  6.95
2000-01-06   5.41   5.69 6.03 6.35 6.39 6.46 6.63  6.57  6.86
2000-01-07   5.38   5.66 6.00 6.31 6.35 6.42 6.58  6.52  6.82
2000-01-10   5.42   5.64 6.07 6.38 6.42 6.49 6.62  6.57  6.86

> tail(fred.data00.0)

           DGS3MO DGS6MO DGS1 DGS2 DGS3 DGS5 DGS7 DGS10 DGS20
2013-05-23   0.05   0.08 0.12 0.26 0.42 0.91 1.40  2.02  2.82
2013-05-24   0.04   0.07 0.12 0.26 0.41 0.90 1.39  2.01  2.80
2013-05-28   0.05   0.09 0.13 0.29 0.49 1.02 1.53  2.15  2.95
2013-05-29   0.05   0.08 0.14 0.30 0.49 1.02 1.51  2.13  2.91
2013-05-30   0.04   0.07 0.13 0.31 0.49 1.01 1.51  2.13  2.92
2013-05-31   0.04   0.07 0.14 0.30 0.52 1.05 1.55  2.16  2.95

> fred.data00.diff=zoo(diff((coredata(fred.data00.0))), order.by=time(fred.data00.0)[-1])
> par(mfcol=c(1,1))
> ts.plot(as.ts(fred.data00.0),col=rainbow(NCOL(fred.data00.0)),bg="black",
+         main="FRED Rates Data: 2001-2013\n(Period 1:2001-2005)")
> legend(x=2880,y=7,
```

```
+         legend=dimnames(fred.data00.0)[[2]],
+         lty=rep(1,times=NCOL(fred.data00.0)),
+         col=rainbow(NCOL(fred.data00.0)),
+         cex=0.70)
> for (year0 in c(1999:2011)){
+   # year0<-2000
+   abline(v=sum(time(fred.data00.0)< as.Date(paste(as.character(year0),"-12-31",sep=""))))
+ }
> for (year0 in c(2000:2005)){
+   # year0<-2000
+   abline(v=sum(time(fred.data00.0)< as.Date(paste(as.character(year0),"-12-31",sep=""))),
+         col=3,lwd=2)
+ }
```
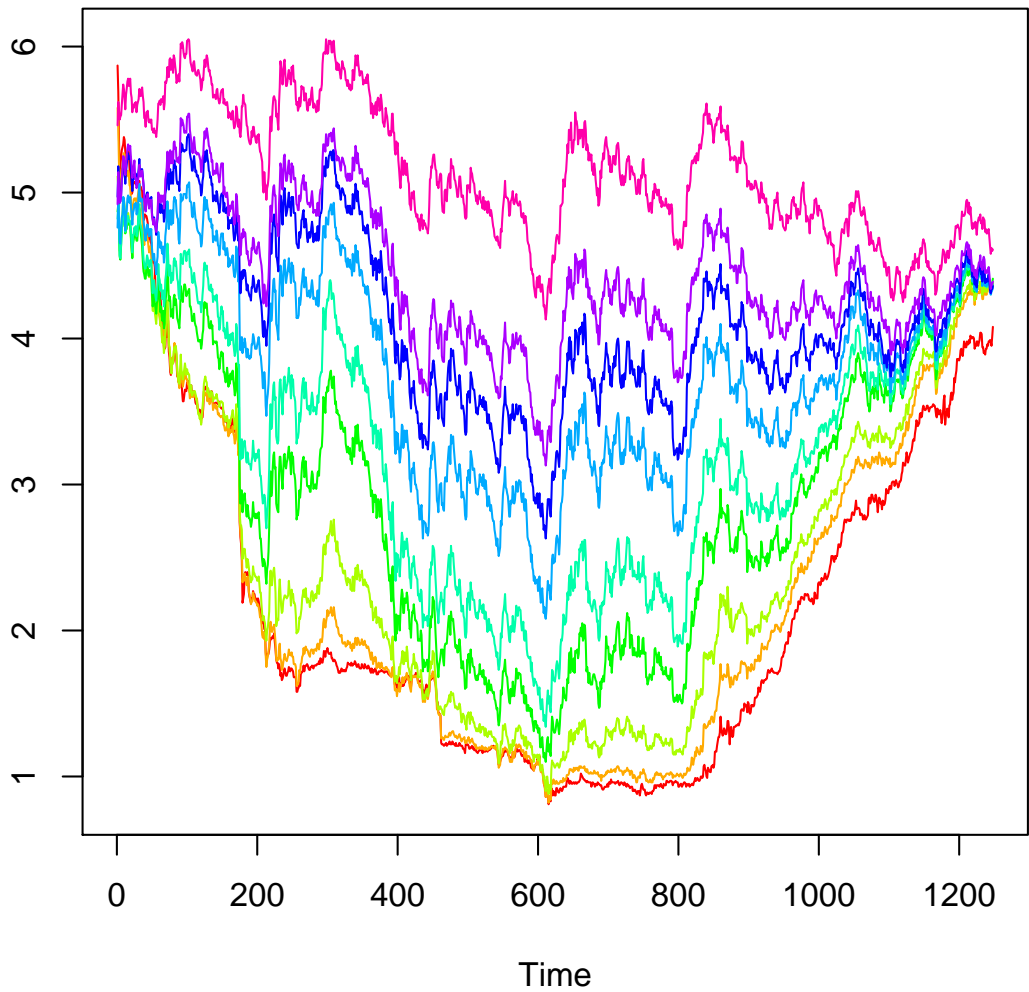
FRED Rates Data: 2001–2013
(Period 1:2001–2005)
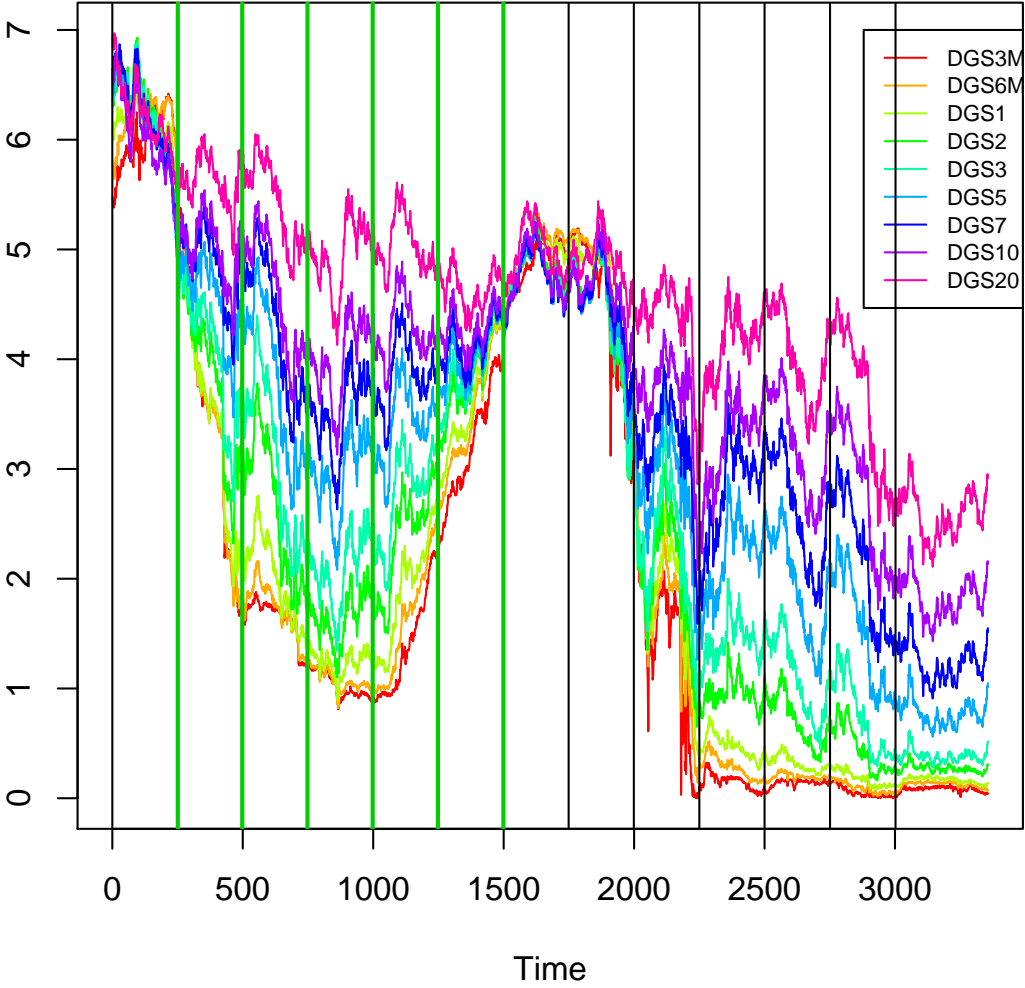
2 US Treasury Yields: 2001-2005

```
> fred.data00.diff.period1<-window(fred.data00.diff, start=as.Date("2001-01-01"),
+                                   end=as.Date("2005-12-31"))
> period1.name0<-"(2001-2005)"
> ts.plot(as.ts(fred.data00.0["2001::2005-12"]),col=rainbow(NCOL(fred.data00.0)),bg="black",
+         main=paste("FRED Data:  Rates", period1.name0,sep="\n"))
> #
>
```

FRED Data: Rates
(2001–2005)

2.1 Means/Volatilities/Correlations of Yield Changes

```
> yieldchanges.period1<-fred.data00.diff.period1
> # Means and St Deviations of the yield changes
> yieldchanges.mean<-apply(yieldchanges.period1,2,mean)
> yieldchanges.vol<-sqrt(apply(yieldchanges.period1,2,var))
> print( round (data.frame(
```

```
+   daily.mean=yieldchanges.mean,
+   daily.vol=yieldchanges.vol), digits=4))

        daily.mean daily.vol
DGS3MO    -0.0015    0.0384
DGS6MO    -0.0011    0.0384
DGS1      -0.0008    0.0486
DGS2      -0.0006    0.0663
DGS3      -0.0006    0.0698
DGS5      -0.0005    0.0692
DGS7      -0.0006    0.0666
DGS10     -0.0006    0.0627
DGS20     -0.0008    0.0559


> # On average, all the yields declined from 2001-2005.
> # The yield with the highest standard deviation is the 3-year.
>
> # Correlations between yield changes
> yieldchanges.cor<-cor(fred.data00.diff.period1)
> print(round(yieldchanges.cor, digits=3))

        DGS3MO DGS6MO  DGS1  DGS2  DGS3  DGS5  DGS7 DGS10 DGS20
DGS3MO   1.000  0.813 0.638 0.479 0.443 0.395 0.341 0.312 0.236
DGS6MO   0.813  1.000 0.864 0.698 0.654 0.595 0.542 0.502 0.414
DGS1     0.638  0.864 1.000 0.876 0.835 0.781 0.735 0.697 0.612
DGS2     0.479  0.698 0.876 1.000 0.972 0.920 0.883 0.843 0.762
DGS3     0.443  0.654 0.835 0.972 1.000 0.956 0.923 0.889 0.813
DGS5     0.395  0.595 0.781 0.920 0.956 1.000 0.974 0.953 0.894
DGS7     0.341  0.542 0.735 0.883 0.923 0.974 1.000 0.979 0.942
DGS10    0.312  0.502 0.697 0.843 0.889 0.953 0.979 1.000 0.960
DGS20    0.236  0.414 0.612 0.762 0.813 0.894 0.942 0.960 1.000
```
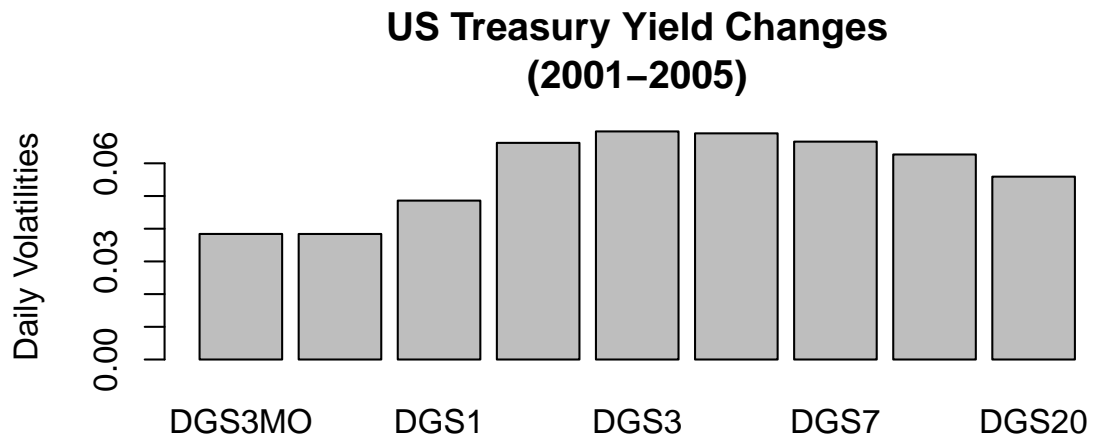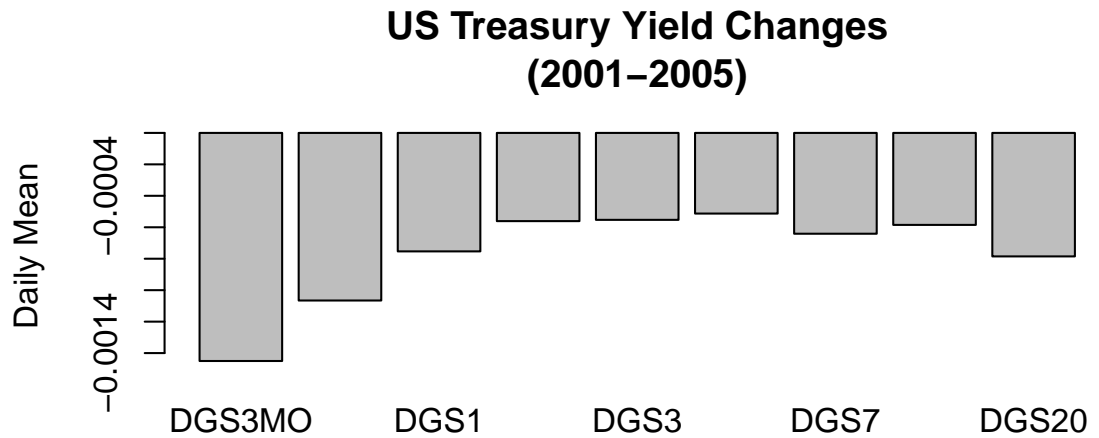
The closer the tenors of the yields, the higher the correlations.

```
> par(mfcol=c(2,1))
> barplot(yieldchanges.mean,
+          main=paste("US Treasury Yield Changes",period1.name0,sep="\n"),
+          ylab="Daily Mean")
> barplot(yieldchanges.vol,
+          main=paste("US Treasury Yield Changes",period1.name0,sep="\n"),
+          ylab="Daily Volatilities")
```
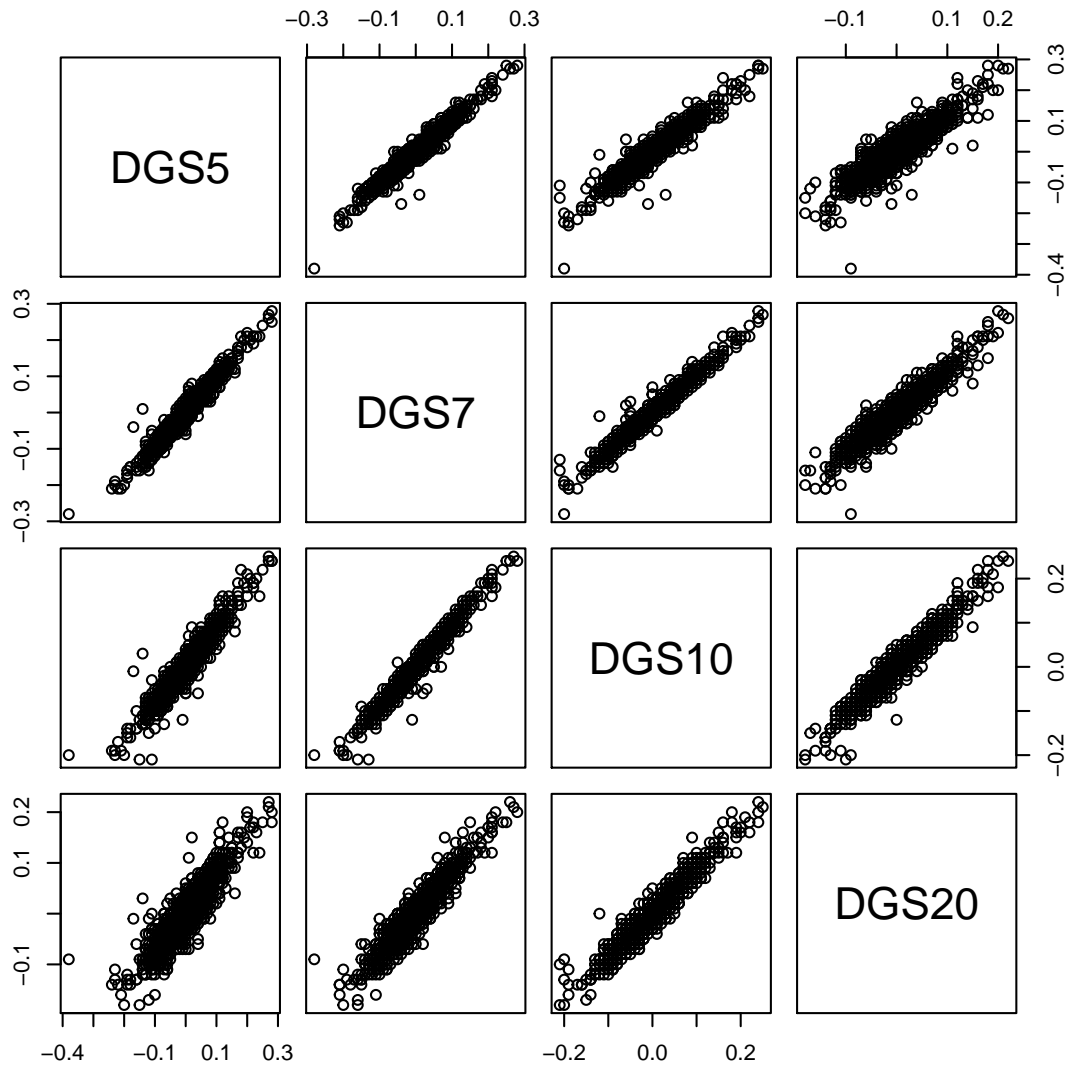
US Treasury Yield Changes
(2001−2005)

US Treasury Yield Changes
(2001−2005)

```
> pairs(yieldchanges.period1[,1:5])
```
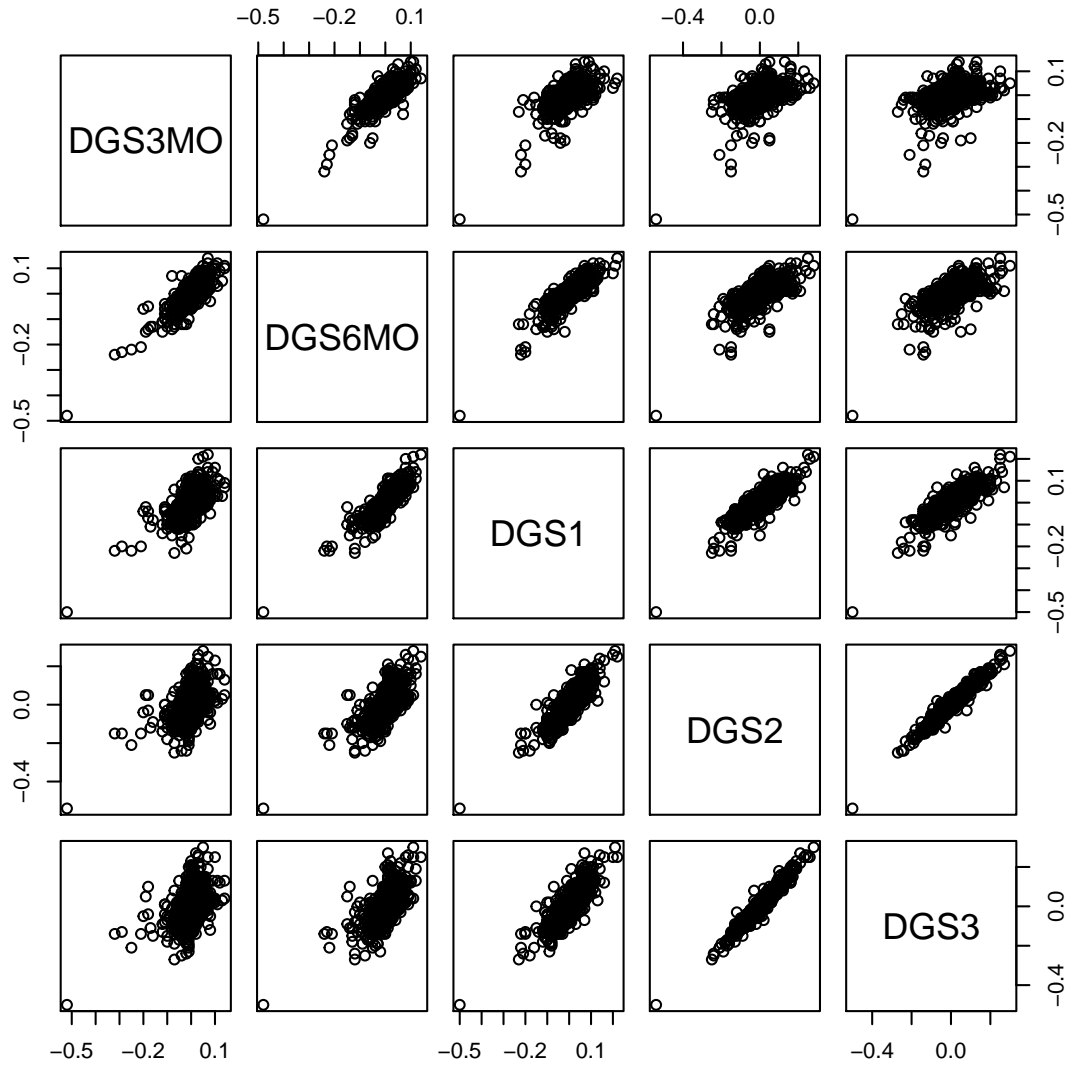
```
> pairs(yieldchanges.period1[,6:9])
```

2.2 Principal Components Analysis

```
> options(width=80)
> obj.princomp0<-princomp(coredata(fred.data00.diff.period1))
> print(summary(obj.princomp0))

Importance of components:
                            Comp.1     Comp.2     Comp.3     Comp.4      Comp.5
Standard deviation       0.1618033 0.05323436 0.02988340 0.01918045 0.013461432
Proportion of Variance   0.8494434 0.09194833 0.02897475 0.01193650 0.005879525
Cumulative Proportion    0.8494434 0.94139169 0.97036644 0.98230294 0.988182464
                              Comp.6      Comp.7      Comp.8      Comp.9
Standard deviation        0.011196400 0.009568498 0.009009979 0.008131889
Proportion of Variance    0.004067397 0.002970621 0.002633948 0.002145570
Cumulative Proportion     0.992249861 0.995220482 0.997854430 1.000000000


> #Note:
> #    The first principal component variable explains 84.9% of the total variability
> #    The second principal component variable explains 9.2% of the total variability
> #    The first three principal component variables explain 97.0% of the total variability
>
> par(mfcol=c(2,1))
> # The screeplot is a barplot of the variances of the Principal Component Variables
> # It provides a graphical display of the summary.
> screeplot(obj.princomp0,main=paste(
+    "Scree Plot PCA US Treasury Yield Changes", period1.name0,sep="\n") )
> # The plot below the scree plot compares the standard deviations (daily) of the
> # yield changes for US Treasury yields to those for the principal component variables.
> #
> # The principal components variables are translations and orthogonal rotations of the
> # original variables.  The first principal component variable has the highest variance
> # and the last has the smallest.
>
> yieldchanges.pcvars.vol<-sqrt(apply(obj.princomp0$scores,2,var))
> #
> ts.plot(cbind(
+    as.matrix(yieldchanges.vol),
+    as.matrix(yieldchanges.pcvars.vol)), type="b", col=c(3,4),
+         xlab="Variables: US Treasury (Green) and  PC Vars (Blue)",
+         main="Volatility of Yield Changes")
```
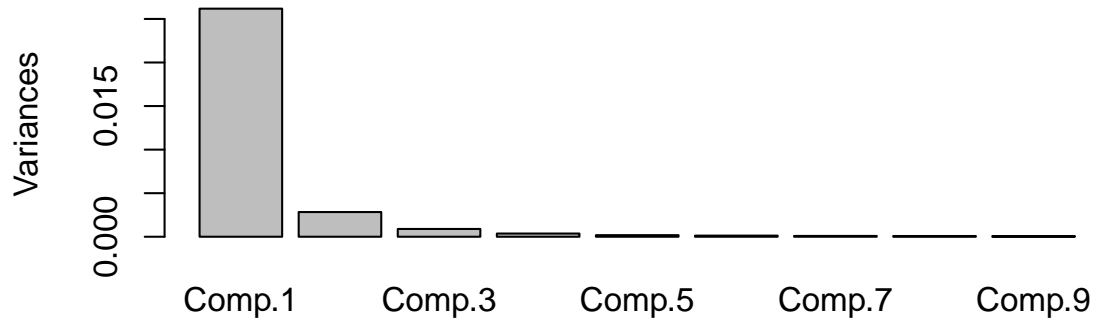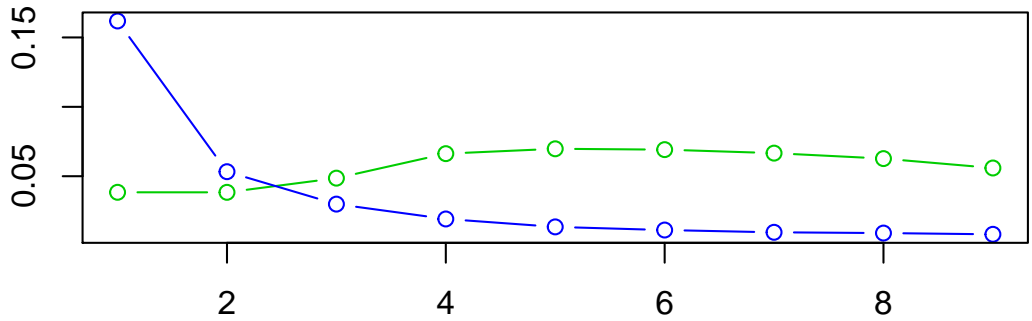
Scree Plot PCA US Treasury Yield Changes
(2001–2005)

Volatility of Yield Changes

Variables: US Treasury (Green) and PC Vars (Blue)

```
> # The loadings component specifies the principal components variables:
>
> print(round(obj.princomp0$loadings[,1:9],digits=4))

        Comp.1  Comp.2  Comp.3  Comp.4  Comp.5  Comp.6  Comp.7  Comp.8  Comp.9
DGS3MO 0.1088 -0.5225  0.5434 -0.5408  0.2107  0.2865 -0.0041 -0.0093 -0.0291
DGS6MO 0.1581 -0.4817  0.2508  0.2555 -0.2581 -0.7343  0.0634  0.0473  0.0524
DGS1   0.2531 -0.4107 -0.0514  0.6687 -0.0306  0.5244 -0.1674 -0.0914 -0.0712
DGS2   0.3905 -0.2071 -0.4750 -0.0673  0.4529 -0.0503  0.5029  0.2849  0.1809
DGS3   0.4197 -0.0873 -0.4022 -0.2965  0.0657 -0.2045 -0.5551 -0.4197 -0.1900
DGS5   0.4206  0.1082 -0.0311 -0.2101 -0.5932  0.1841 -0.1817  0.4776  0.3455
DGS7   0.4009  0.2289  0.1475 -0.0204 -0.2167  0.0356  0.3865  0.0047 -0.7530
DGS10  0.3697  0.2837  0.2710  0.0666 -0.0173  0.0351  0.2898 -0.6191  0.4859
DGS20  0.3101  0.3621  0.3944  0.2316  0.5286 -0.1515 -0.3749  0.3472 -0.0033
```

The first principal component variable is a weighted average of all the yield changes, giving highest weight to the 5-year tenor.
The second principal component measures the spread between long-tenor and short-tenor yields.
The third principal component variable measures the curvature in the yield curve.
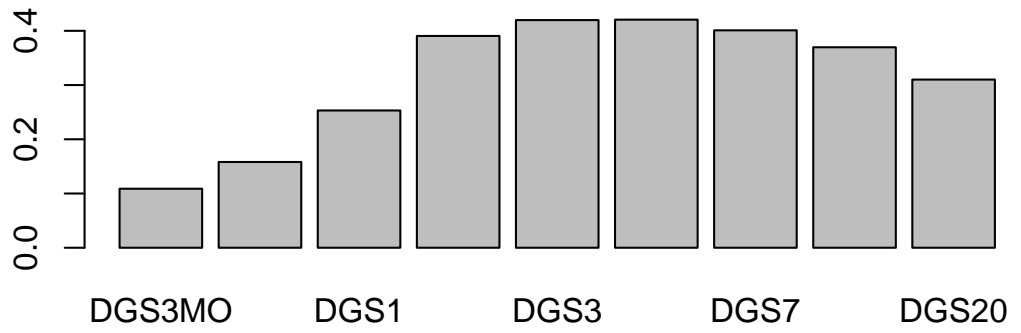
The last (9th) principal component variable measures the 'hedge' of the 7-year yield with the 5-year and 10-year yields.
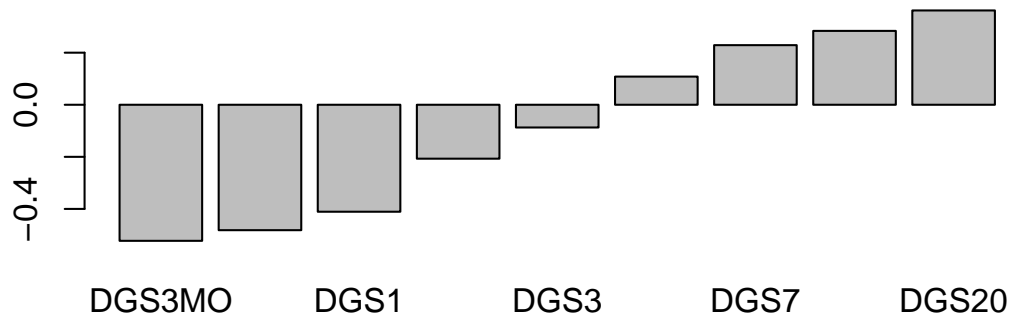
```
> par(mfcol=c(2,1))
> barplot(obj.princomp0$loadings[,1],
+        main="US Treasury Yield Changes: PCA\n Loadings: PC1")
> barplot(obj.princomp0$loadings[,2],
+        main="Loadings: PC2")
>
```

US Treasury Yield Changes: PCA
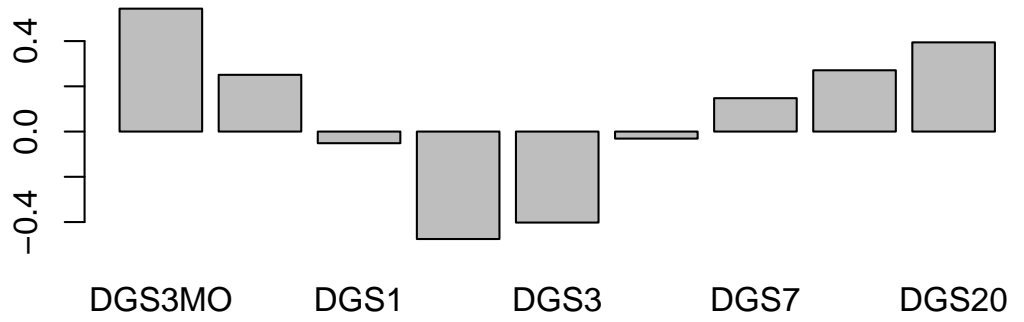**Loadings: PC1**

Loadings: PC2

```
> par(mfcol=c(2,1))
> barplot(obj.princomp0$loadings[,3],
+          main="US Treasury Yield Changes: PCA\nLoadings: PC3")
> barplot(obj.princomp0$loadings[,9],
+          main="Loadings: PC9")
> # PC9 has the smallest variability
> # It is dominated by the matching of 7-Year yield changes by
> # weighted average of 5-Year and 10-Year
>
```
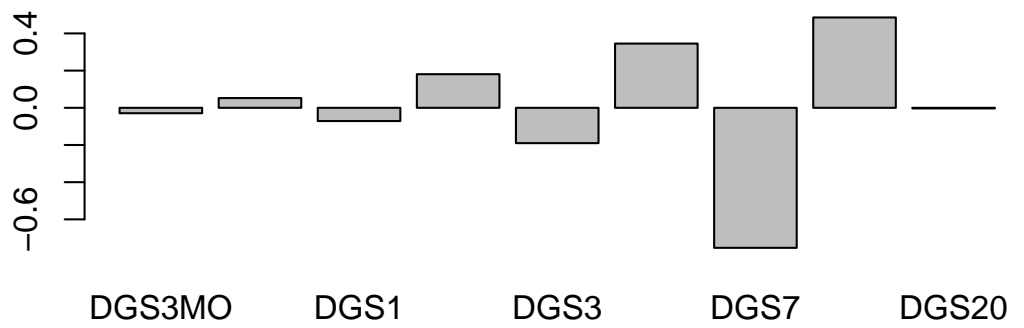
US Treasury Yield Changes: PCA
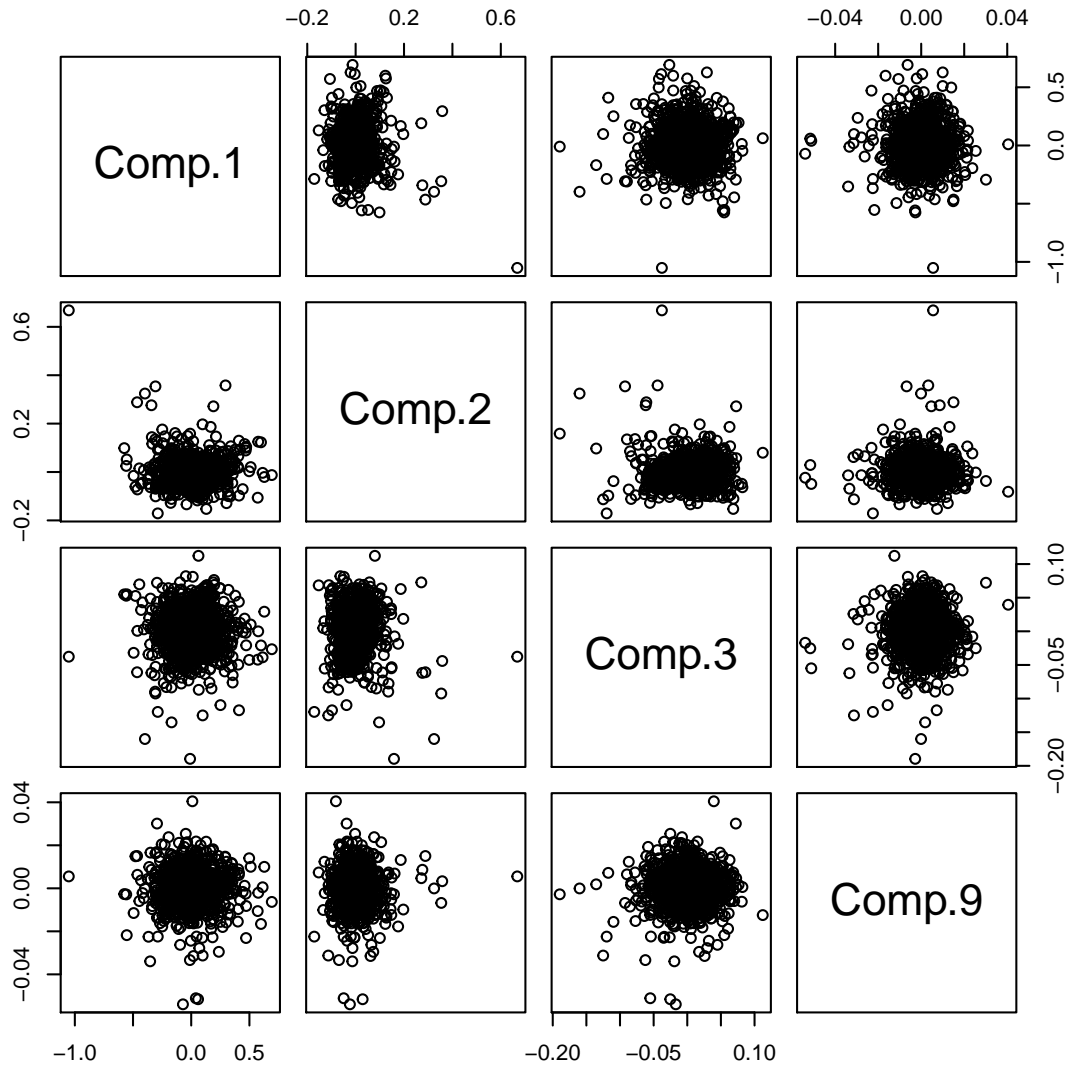## Loadings: PC3

Loadings: PC9

The principal component variables are uncorrelated. This is evident from the correlation matrix of the variables and the pairs plot:

```
> print(round(cor(obj.princomp0$scores)))
```

```
       Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8 Comp.9
Comp.1      1      0      0      0      0      0      0      0      0
Comp.2      0      1      0      0      0      0      0      0      0
Comp.3      0      0      1      0      0      0      0      0      0
Comp.4      0      0      0      1      0      0      0      0      0
Comp.5      0      0      0      0      1      0      0      0      0
Comp.6      0      0      0      0      0      1      0      0      0
Comp.7      0      0      0      0      0      0      1      0      0
Comp.8      0      0      0      0      0      0      0      1      0
Comp.9      0      0      0      0      0      0      0      0      1
```
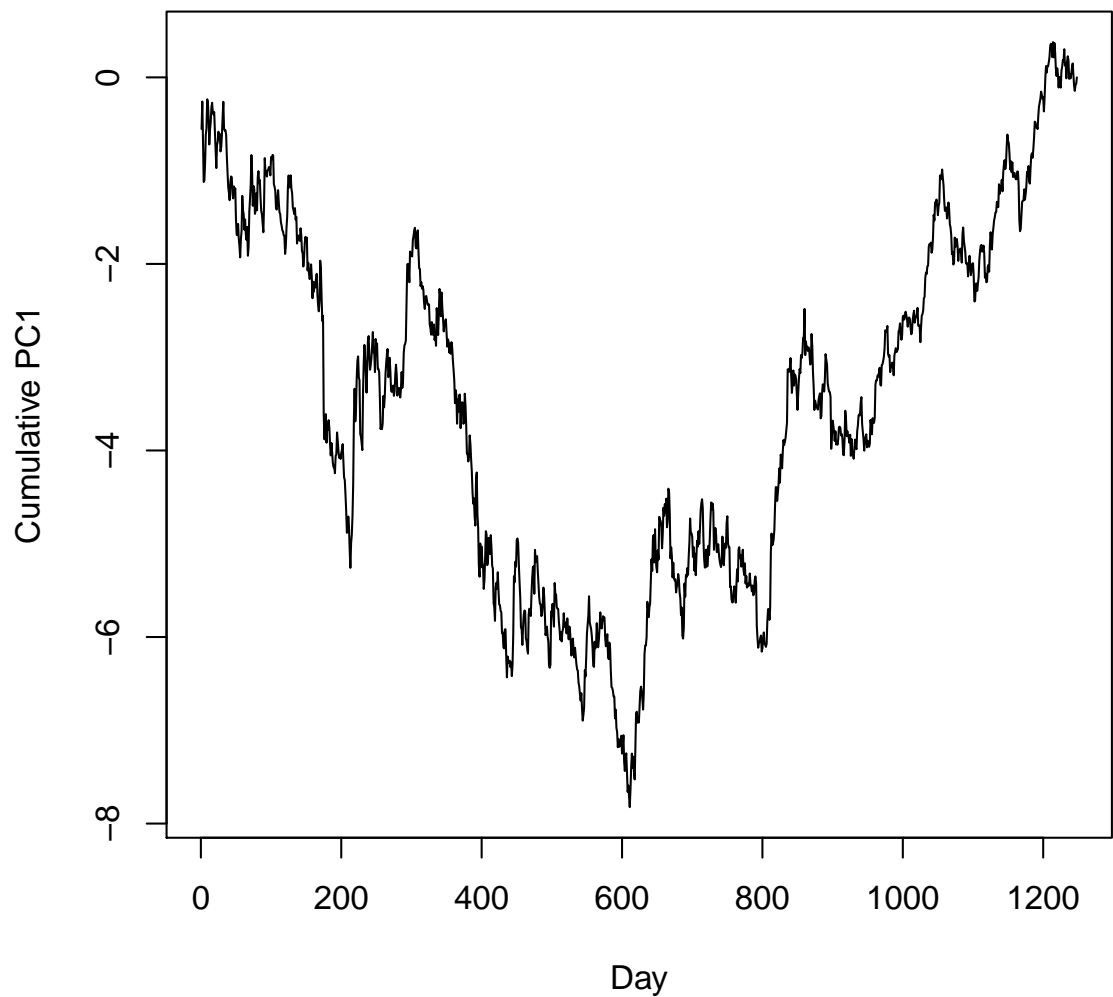
```
> pairs(obj.princomp0$scores[,c(1:3,9)])
```

Plotting the cumulative scores of the principal components variables displays the evolution of these factors characterizing the variability in the treasury yield data.
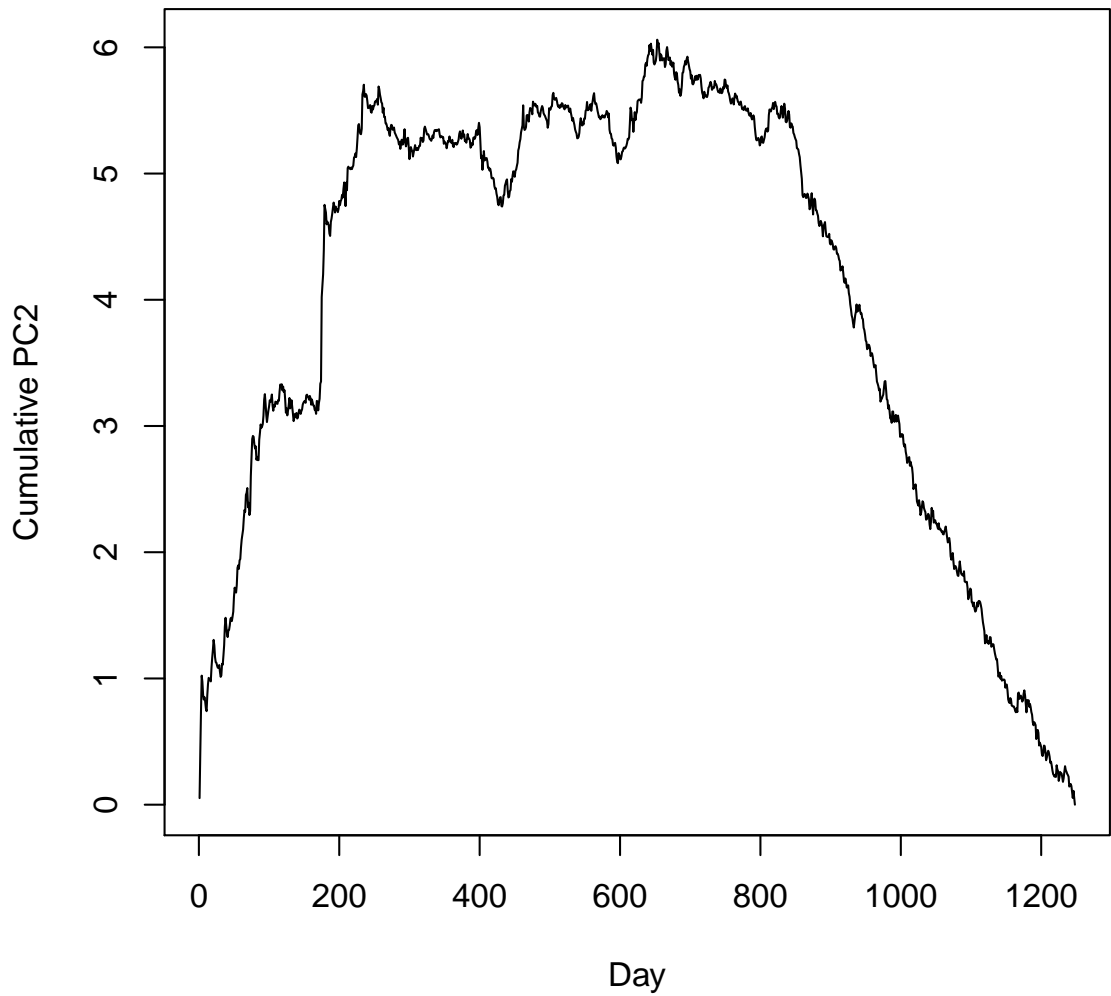
```
> par(mfcol=c(1,1))
> ts.plot(cumsum(obj.princomp0$scores[,1]),
+          main="Cumulative PC Variables\nPC1 (Level-Shifts)",
+          ylab="Cumulative PC1",xlab="Day")
```

Cumulative PC Variables
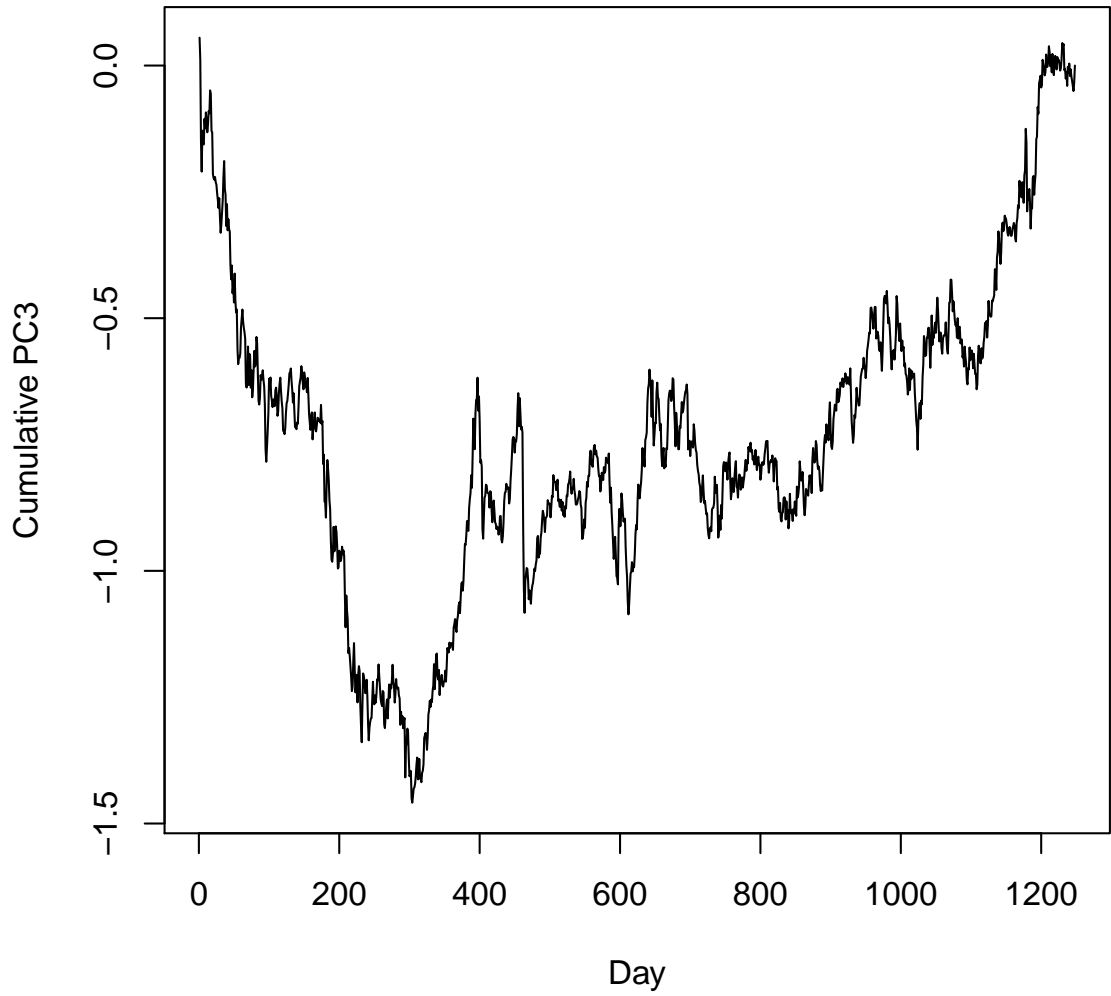## PC1 (Level−Shifts)

```
> ts.plot(cumsum(obj.princomp0$scores[,2]),
+         main="Cumulative PC Variables\nPC2 (Spread Between Long and Short Maturities)",
+         ylab="Cumulative PC2",xlab="Day")
```

Cumulative PC Variables
PC2 (Spread Between Long and Short Maturities)

```
> ts.plot(cumsum(obj.princomp0$scores[,3]),
+         main="Cumulative PC Variables\nPC3 (Curvature of Term Structure)",
+         ylab="Cumulative PC3",xlab="Day")
```

Cumulative PC Variables
PC3 (Curvature of Term Structure)

```
> ts.plot(cumsum(obj.princomp0$scores[,9]),
+         main="Cumulative PC Variables\nPC9 (7 Year vs 5 and 10 Year)",
+         ylab="Cumulative PC9",
+         xlab="Day")
> ########################
>
```

Cumulative PC Variables
PC9 (7 Year vs 5 and 10 Year)

2.3 Factor Analysis

The r function *factanal()* performs the maximum-likelihood factor analysis.
The output includes the results of the likeihood ratio test for the sufficiency of
the given number of factors.

```
> # The computations lead to rejecting the hypothesis that 4 factors is sufficient.
> print(factanal(yieldchanges.period1,factors=4))

Call:
factanal(x = yieldchanges.period1, factors = 4)

Uniquenesses:
DGS3MO DGS6MO   DGS1   DGS2   DGS3   DGS5   DGS7  DGS10  DGS20
 0.316  0.005  0.106  0.005  0.026  0.012  0.016  0.017  0.041

Loadings:
        Factor1 Factor2 Factor3 Factor4
DGS3MO   0.113   0.815
DGS6MO   0.265   0.947   0.168
DGS1     0.489   0.705   0.395
DGS2     0.673   0.445   0.586
DGS3     0.751   0.392   0.501
DGS5     0.867   0.330   0.330   0.140
DGS7     0.922   0.272   0.242
DGS10    0.948   0.235   0.171
DGS20    0.959   0.154


               Factor1 Factor2 Factor3 Factor4
SS loadings      4.758   2.671   0.988   0.039
Proportion Var   0.529   0.297   0.110   0.004
Cumulative Var   0.529   0.825   0.935   0.940


Test of the hypothesis that 4 factors are sufficient.
The chi square statistic is 26.76 on 6 degrees of freedom.
The p-value is 0.00016

> # The computations also lead to rejecting the hypothesis that 5 factors is sufficient.
> print(factanal(yieldchanges.period1,factors=5))

Call:
factanal(x = yieldchanges.period1, factors = 5)

Uniquenesses:
DGS3MO DGS6MO   DGS1   DGS2   DGS3   DGS5   DGS7  DGS10  DGS20
 0.231  0.062  0.005  0.005  0.026  0.010  0.016  0.016  0.042
```

```
Loadings:
       Factor1 Factor2 Factor3 Factor4 Factor5
DGS3MO  0.124   0.864
DGS6MO  0.278   0.895   0.206   0.130
DGS1    0.479   0.667   0.396   0.405
DGS2    0.680   0.420   0.589
DGS3    0.758   0.370   0.505
DGS5    0.871   0.309   0.333           0.145
DGS7    0.925   0.251   0.242
DGS10   0.950   0.217   0.167
DGS20   0.959   0.136


               Factor1 Factor2 Factor3 Factor4 Factor5
SS loadings      4.797   2.530   1.008   0.217   0.034
Proportion Var   0.533   0.281   0.112   0.024   0.004
Cumulative Var   0.533   0.814   0.926   0.950   0.954


Test of the hypothesis that 5 factors are sufficient.
The chi square statistic is 4.33 on 1 degree of freedom.
The p-value is 0.0374
```

A six-factor model cannot be fit to the data. For this period, the structure of the yield changes is such that a reduced-dimension model for the correlation structure is rejected by the likelihood ratio test.

3   US Treasury Yields: 2009-2012

```
> fred.data00.diff.period2<-window(fred.data00.diff, start=as.Date("2009-01-01"),
+                                   end=as.Date("2013-12-31"))
> period2.name0<-"(2009-2013)"
> ts.plot(as.ts(fred.data00.0["2009::2013"]),col=rainbow(NCOL(fred.data00.0)),bg="black",
+         main=paste("FRED Data:  Rates", period2.name0,sep="\n"))
> #
>
```

FRED Data: Rates
(2009−2013)

Time
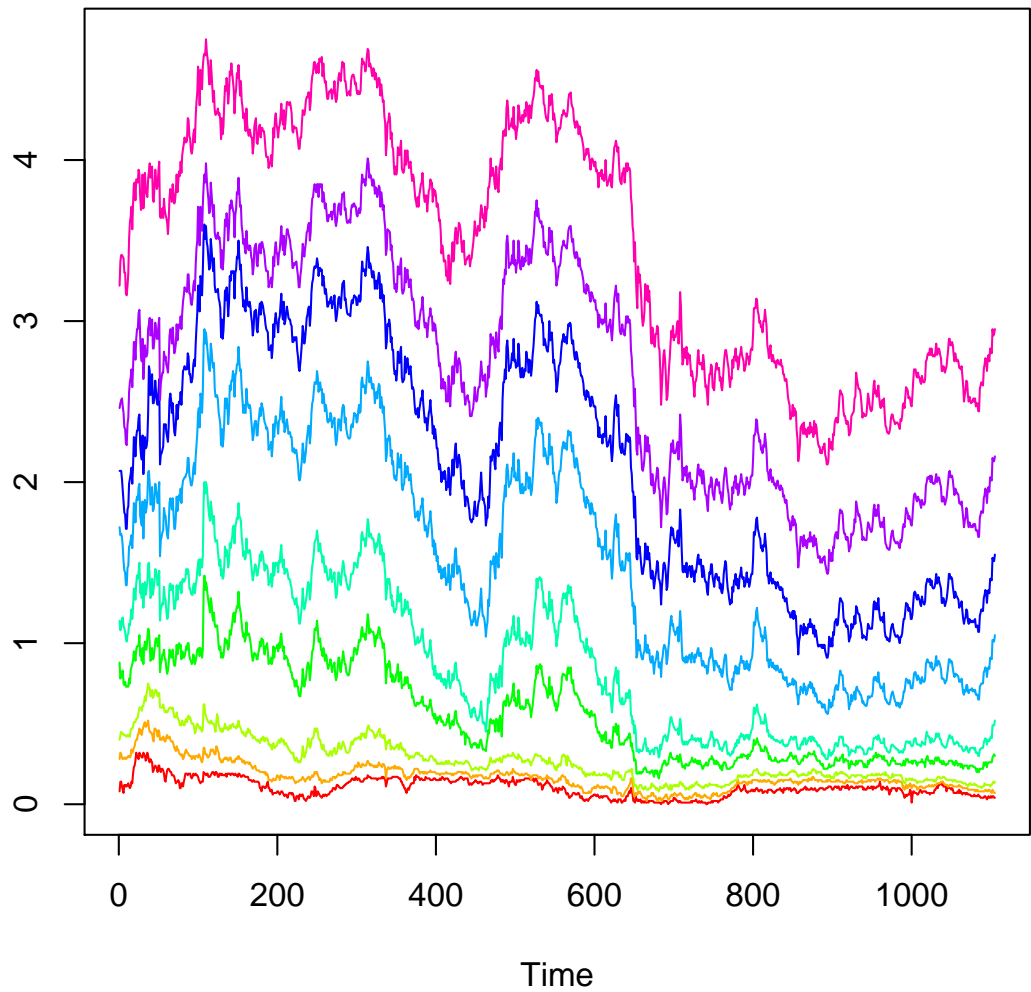
3.1   Means/Volatilities/Correlations of Yield Changes

```
> yieldchanges.period2<-fred.data00.diff.period2
> # Means and St Deviations of the yield changes
> yieldchanges.mean<-apply(yieldchanges.period2,2,mean)
> yieldchanges.vol<-sqrt(apply(yieldchanges.period2,2,var))
> print( round (data.frame(
```

```
+    daily.mean=yieldchanges.mean,
+    daily.vol=yieldchanges.vol), digits=4))

       daily.mean daily.vol
DGS3MO     -1e-04    0.0125
DGS6MO     -2e-04    0.0116
DGS1       -2e-04    0.0154
DGS2       -4e-04    0.0352
DGS3       -4e-04    0.0448
DGS5       -5e-04    0.0597
DGS7       -3e-04    0.0659
DGS10      -1e-04    0.0653
DGS20      -1e-04    0.0658

> # On average, all the yields declined since 2009.
> # The yields with the highest standard deviations are the 7, 10, and 20 year.
>
> # Correlations between yield changes
> yieldchanges.cor<-cor(fred.data00.diff.period2)
> print(round(yieldchanges.cor, digits=3))

        DGS3MO DGS6MO  DGS1  DGS2  DGS3  DGS5  DGS7 DGS10 DGS20
DGS3MO   1.000  0.460 0.313 0.168 0.148 0.119 0.092 0.101 0.091
DGS6MO   0.460  1.000 0.510 0.384 0.354 0.316 0.269 0.258 0.238
DGS1     0.313  0.510 1.000 0.674 0.619 0.555 0.487 0.446 0.385
DGS2     0.168  0.384 0.674 1.000 0.904 0.829 0.746 0.688 0.596
DGS3     0.148  0.354 0.619 0.904 1.000 0.936 0.868 0.815 0.720
DGS5     0.119  0.316 0.555 0.829 0.936 1.000 0.962 0.929 0.847
DGS7     0.092  0.269 0.487 0.746 0.868 0.962 1.000 0.967 0.907
DGS10    0.101  0.258 0.446 0.688 0.815 0.929 0.967 1.000 0.957
DGS20    0.091  0.238 0.385 0.596 0.720 0.847 0.907 0.957 1.000
```
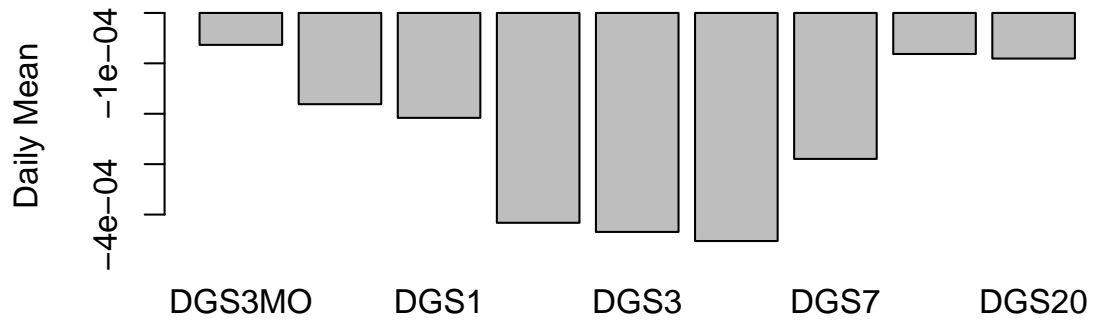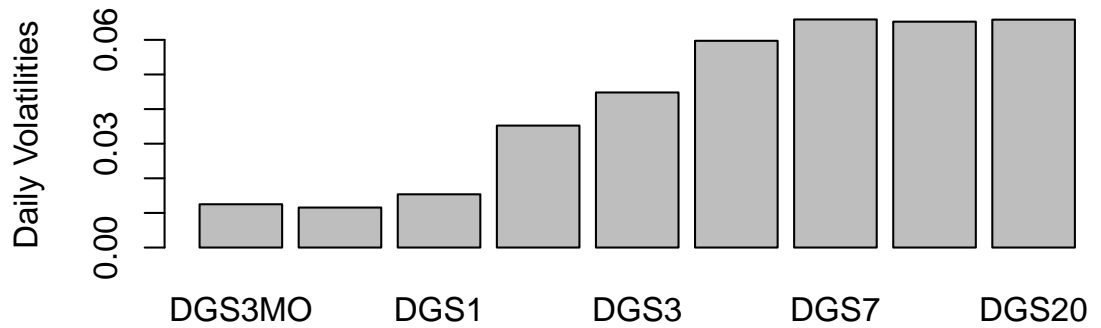
The closer the tenors of the yields, the higher the correlations.

```
> par(mfcol=c(2,1))
> barplot(yieldchanges.mean,
+         main=paste("US Treasury Yield Changes",period2.name0,sep="\n"),
+         ylab="Daily Mean")
> barplot(yieldchanges.vol,
+         main=paste("US Treasury Yield Changes",period2.name0,sep="\n"),
+         ylab="Daily Volatilities")
```

US Treasury Yield Changes
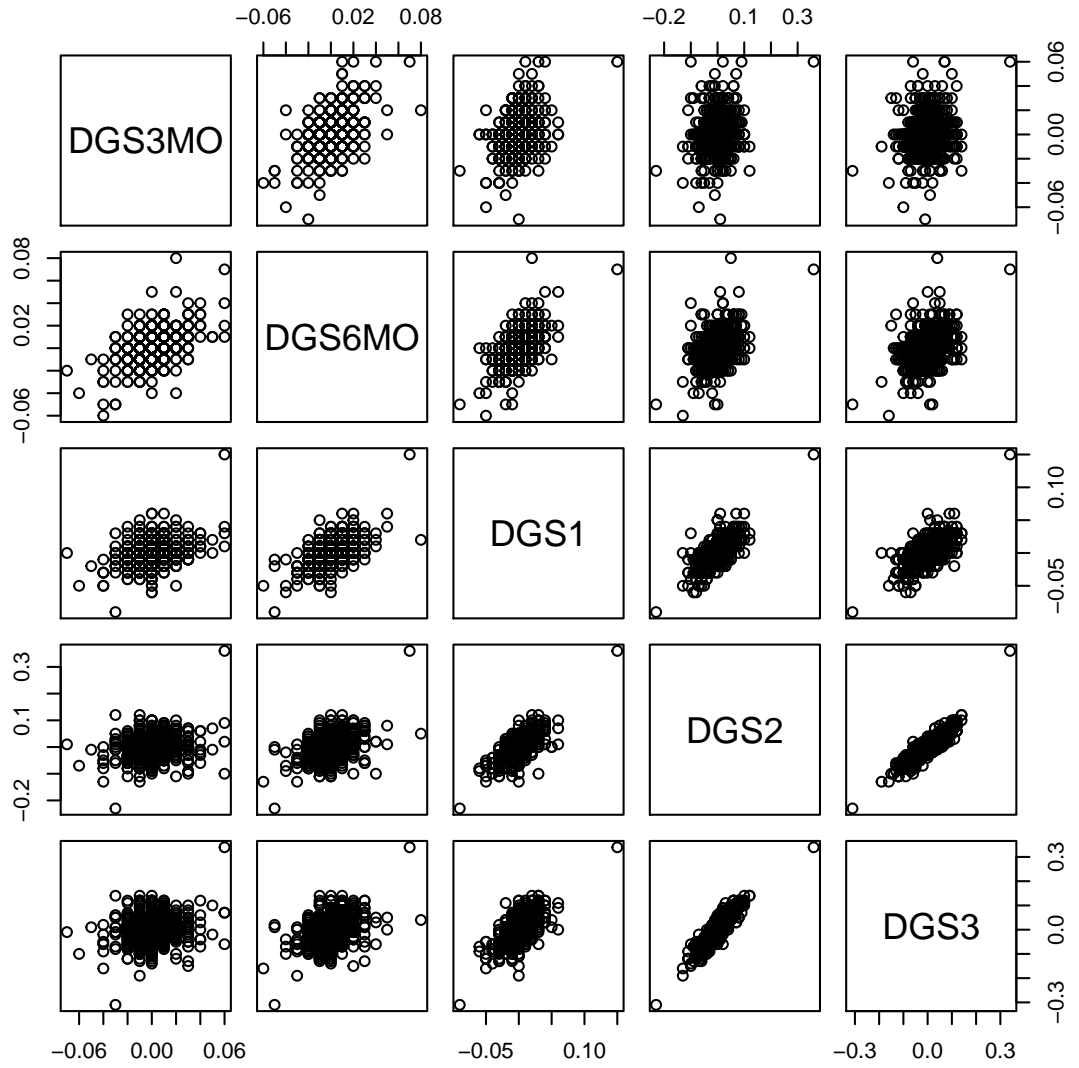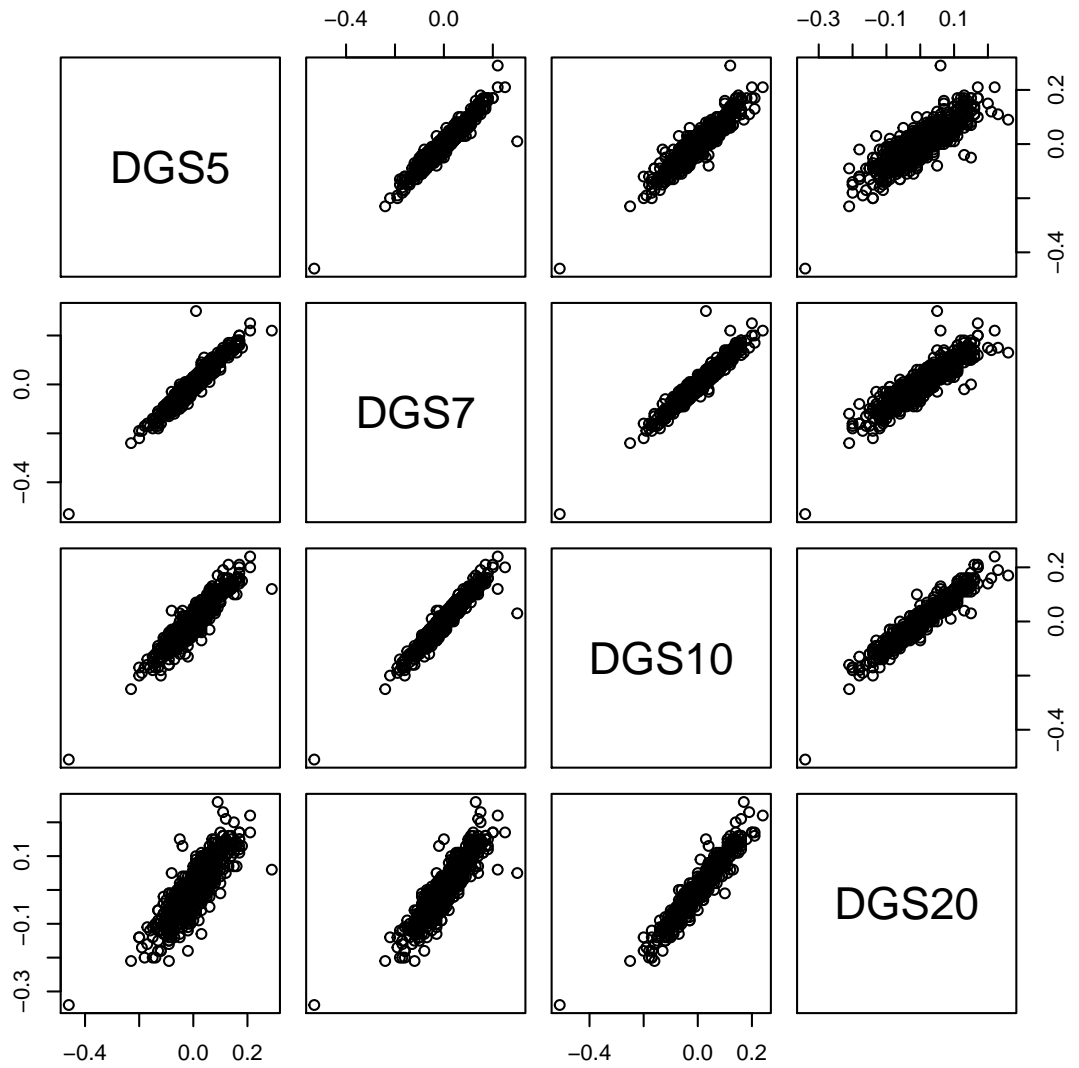## (2009–2013)

US Treasury Yield Changes
(2009–2013)

The yields on average fell during the period from 2009-2013.

The volatilty of the yield changes is highest for tenors greater than 5 years (7, 10, and 20 years).

```
> pairs(yieldchanges.period2[,1:5])
```

```
> pairs(yieldchanges.period2[,6:9])
```

3.2 Principal Components Analysis

```
> options(width=80)
> obj.princomp0<-princomp(coredata(fred.data00.diff.period2))
> print(summary(obj.princomp0))

Importance of components:
                           Comp.1     Comp.2     Comp.3     Comp.4      Comp.5
Standard deviation      0.1339363 0.03706452 0.01778919 0.01391865 0.011029487
Proportion of Variance  0.8856111 0.06782084 0.01562280 0.00956403 0.006005609
Cumulative Proportion   0.8856111 0.95343190 0.96905471 0.97861873 0.984624344
                            Comp.6      Comp.7      Comp.8      Comp.9
Standard deviation      0.009714252 0.009053910 0.008359889 0.008075984
Proportion of Variance  0.004658704 0.004046865 0.003450225 0.003219862
Cumulative Proportion   0.989283047 0.993329912 0.996780138 1.000000000


> #Note:
> #   The first principal component variable explains 88.6% of the total variability
> #   The second principal component variable explains 6.8% of the total variability
> #   The first three principal component variables explain 96.9% of the total variability
> #
> #   When compared to the period 2001-2005, the first principal component
> #   variable explains modestly more of the variability 88.6% vs 84.9%
> par(mfcol=c(2,1))
> # The screeplot is a barplot of the variances of the Principal Component Variables
> # It provides a graphical display of the summary.
> screeplot(obj.princomp0,main=paste(
+    "Scree Plot PCA US Treasury Yield Changes", period2.name0,sep="\n") )
> # The plot below the scree plot compares the standard deviations (daily) of the
> # yield changes for US Treasury yields to those for the principal component variables.
> #
> # The principal components variables are translations and orthogonal rotations of the
> # original variables.  The first principal component variable has the highest variance
> # and the last has the smallest.
>
> yieldchanges.pcvars.vol<-sqrt(apply(obj.princomp0$scores,2,var))
> #
> ts.plot(cbind(
+   as.matrix(yieldchanges.vol),
+   as.matrix(yieldchanges.pcvars.vol)), type="b", col=c(3,4),
+         xlab="Variables: US Treasury (Green) and  PC Vars (Blue)",
+         main="Volatility of Yield Changes")
```
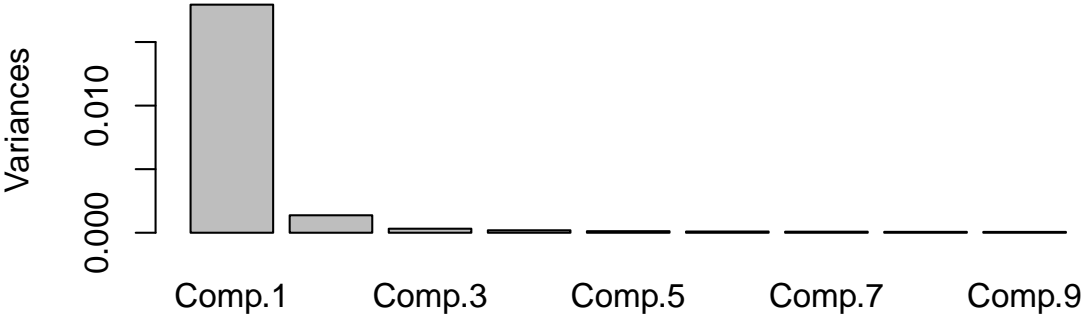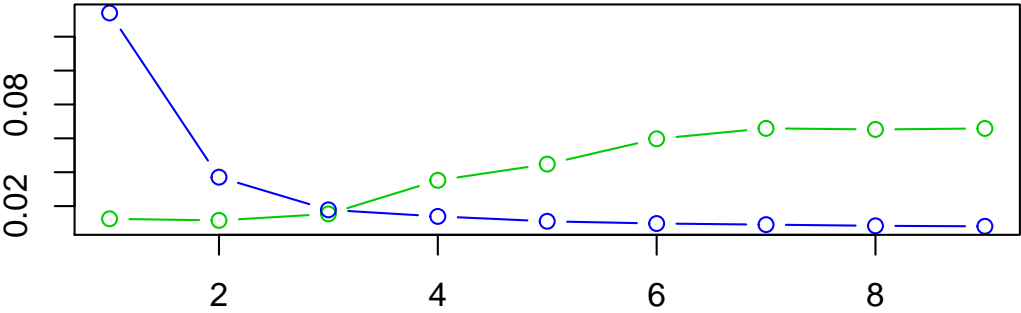
Scree Plot PCA US Treasury Yield Changes (2009–2013)

Volatility of Yield Changes

Variables: US Treasury (Green) and PC Vars (Blue)

```
> # The loadings component specifies the principal components variables:
>
> print(round(obj.princomp0$loadings[,1:9],digits=4))

          Comp.1  Comp.2  Comp.3  Comp.4  Comp.5  Comp.6  Comp.7  Comp.8  Comp.9
DGS3MO  -0.0109 -0.0453  0.3093  0.6247  0.2362  0.5371  0.1003 -0.1827 -0.3524
DGS6MO  -0.0260 -0.0775  0.2938  0.4393  0.0478 -0.1151  0.0822  0.4843  0.6762
DGS1    -0.0598 -0.1828  0.3444  0.3222 -0.3242 -0.6821 -0.1406 -0.3046 -0.2401
DGS2    -0.2059 -0.5100  0.4134 -0.3489 -0.4829  0.3778 -0.0515  0.1593 -0.0288
DGS3    -0.2992 -0.4904  0.0520 -0.1910  0.5624 -0.1072  0.1746 -0.4448  0.2732
DGS5    -0.4340 -0.2815 -0.2780  0.1083  0.2730 -0.1627 -0.2200  0.5624 -0.4216
DGS7    -0.4848  0.0125 -0.4418  0.2468 -0.4281  0.0542  0.5455 -0.1341  0.0836
DGS10   -0.4782  0.2616 -0.0577  0.1008 -0.0742  0.1890 -0.7076 -0.2611  0.2815
DGS20   -0.4601  0.5568  0.5001 -0.2681  0.1550 -0.1169  0.2890  0.1167 -0.1464
```

The first principal component variable is a weighted average of all the yield changes, giving greater weights to the longer tenors. In contrast to the 2001-2005 period, the short-tenor yields have low weights. This is due to the low-magnitude yields at the short tenors which consequently have very low variabilitity (and thus little contriution to the first principal component variable). highest weight to the 5-year tenor.

The second principal component measures the spread between long-tenor and medium-tenor yields.

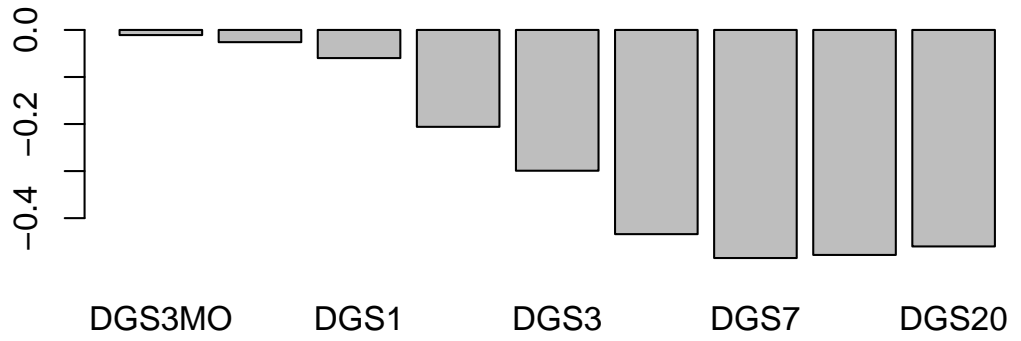The third principal component variable measures the curvature in the yield curve.

The last (9th) principal component variable might be interpretable as the 'hedge' of the 6-month yield with the 3-month and 5-year yields.
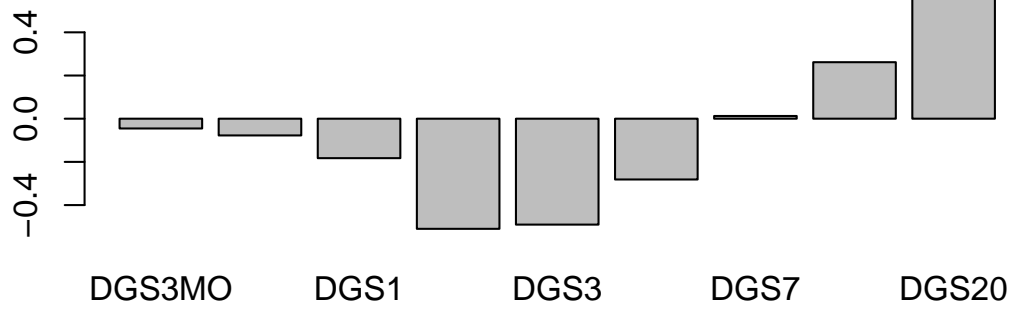
```
> par(mfcol=c(2,1))
> barplot(obj.princomp0$loadings[,1],
+         main="US Treasury Yield Changes: PCA\n Loadings: PC1")
> barplot(obj.princomp0$loadings[,2],
+         main="Loadings: PC2")
>
```

US Treasury Yield Changes: PCA
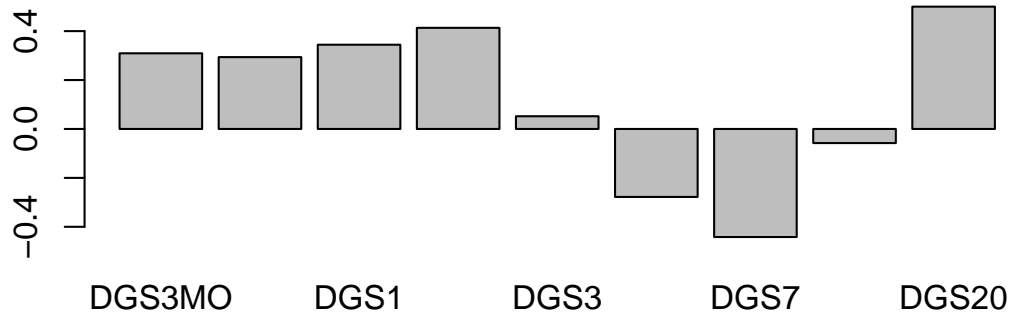**Loadings: PC1**

Loadings: PC2

```
> par(mfcol=c(2,1))
> barplot(obj.princomp0$loadings[,3],
+         main="US Treasury Yield Changes: PCA\nLoadings: PC3")
> barplot(obj.princomp0$loadings[,9],
+         main="Loadings: PC9")
> # PC9 has the smallest variability
> # It is dominated by the matching of 7-Year yield changes by
> # weighted average of 5-Year and 10-Year
>
```
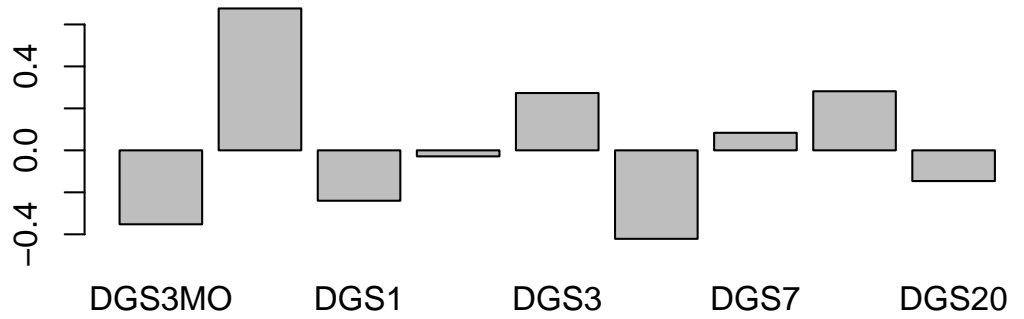
US Treasury Yield Changes: PCA
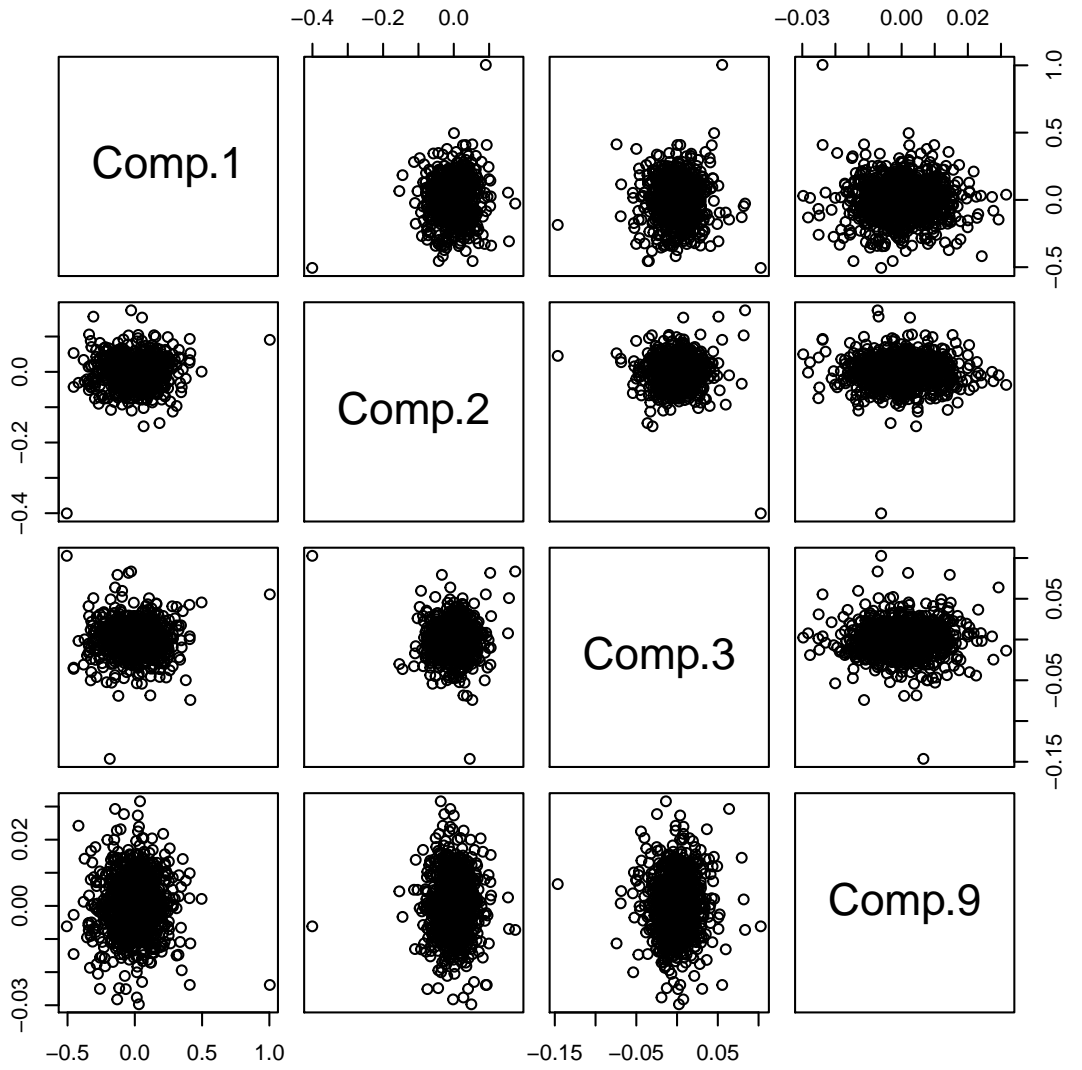## Loadings: PC3

Loadings: PC9

The principal component variables are uncorrelated. This is evident from the correlation matrix of the variables and the pairs plot:

```
> print(round(cor(obj.princomp0$scores)))
```

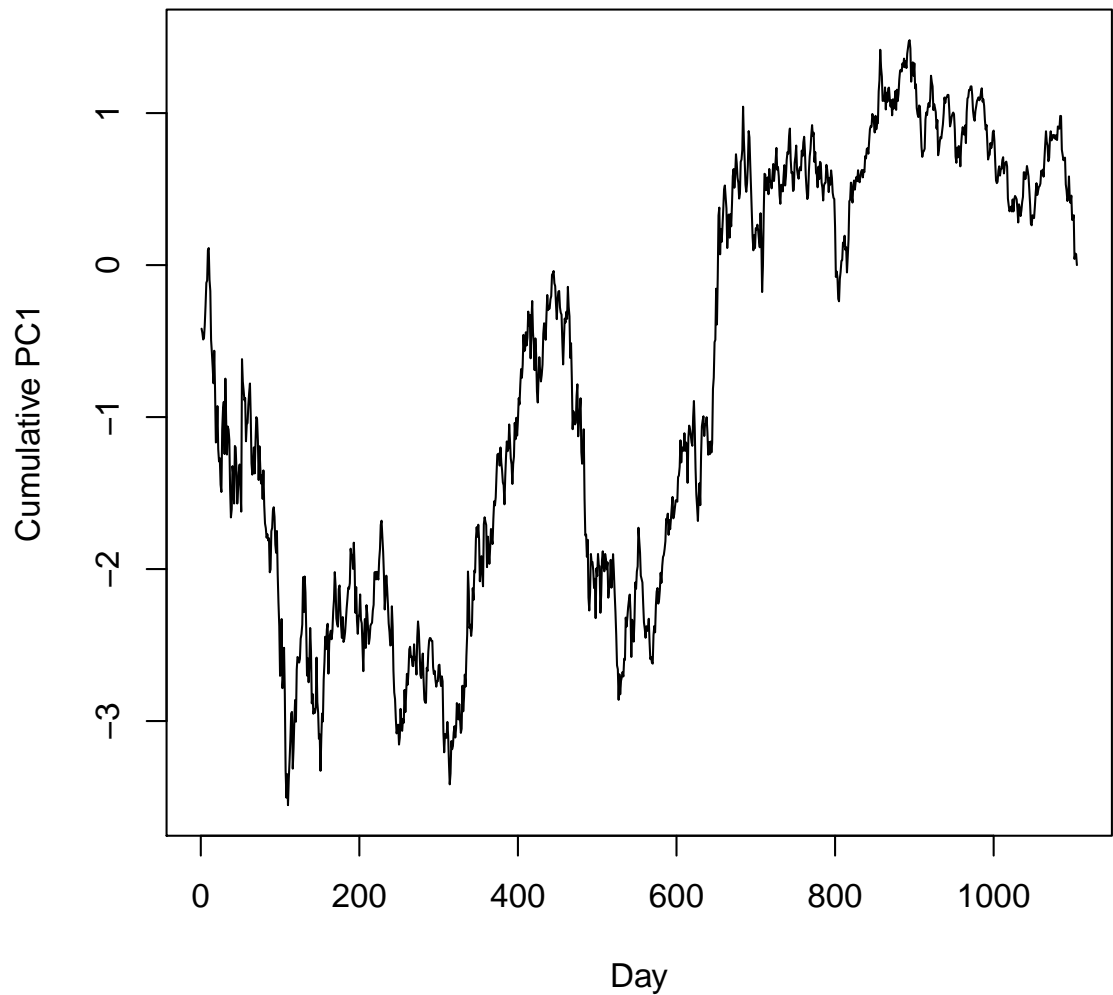|	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5	Comp.6	Comp.7	Comp.8	Comp.9
Comp.1	1	0	0	0	0	0	0	0	0
Comp.2	0	1	0	0	0	0	0	0	0
Comp.3	0	0	1	0	0	0	0	0	0
Comp.4	0	0	0	1	0	0	0	0	0
Comp.5	0	0	0	0	1	0	0	0	0
Comp.6	0	0	0	0	0	1	0	0	0
Comp.7	0	0	0	0	0	0	1	0	0
Comp.8	0	0	0	0	0	0	0	1	0
Comp.9	0	0	0	0	0	0	0	0	1

```
> pairs(obj.princomp0$scores[,c(1:3,9)])
```

Plotting the cumulative scores of the principal components variables displays the evolution of these factors characterizing the variability in the treasury yield data.
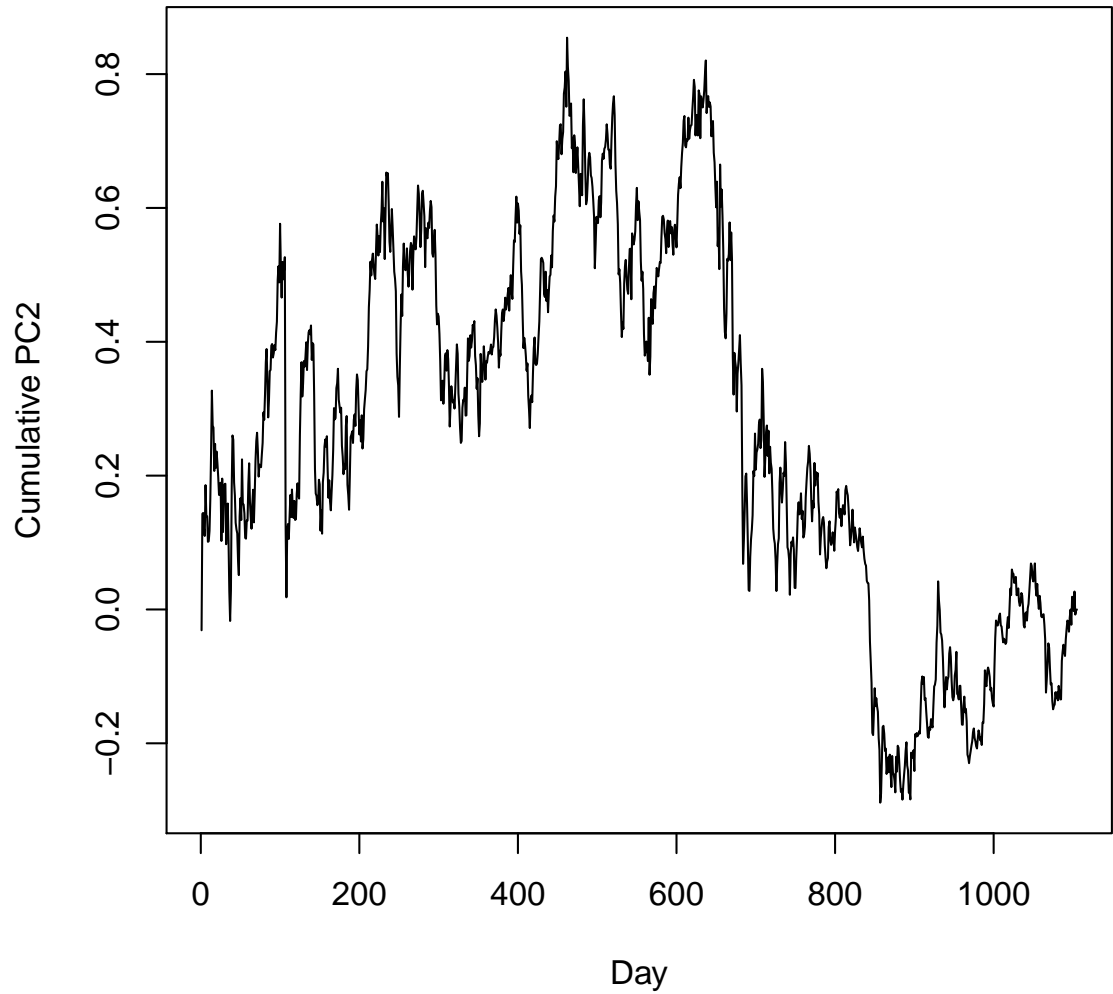
```
> par(mfcol=c(1,1))
> ts.plot(cumsum(obj.princomp0$scores[,1]),
+         main="Cumulative PC Variables\nPC1 (Level-Shifts)",
+         ylab="Cumulative PC1",xlab="Day")
```

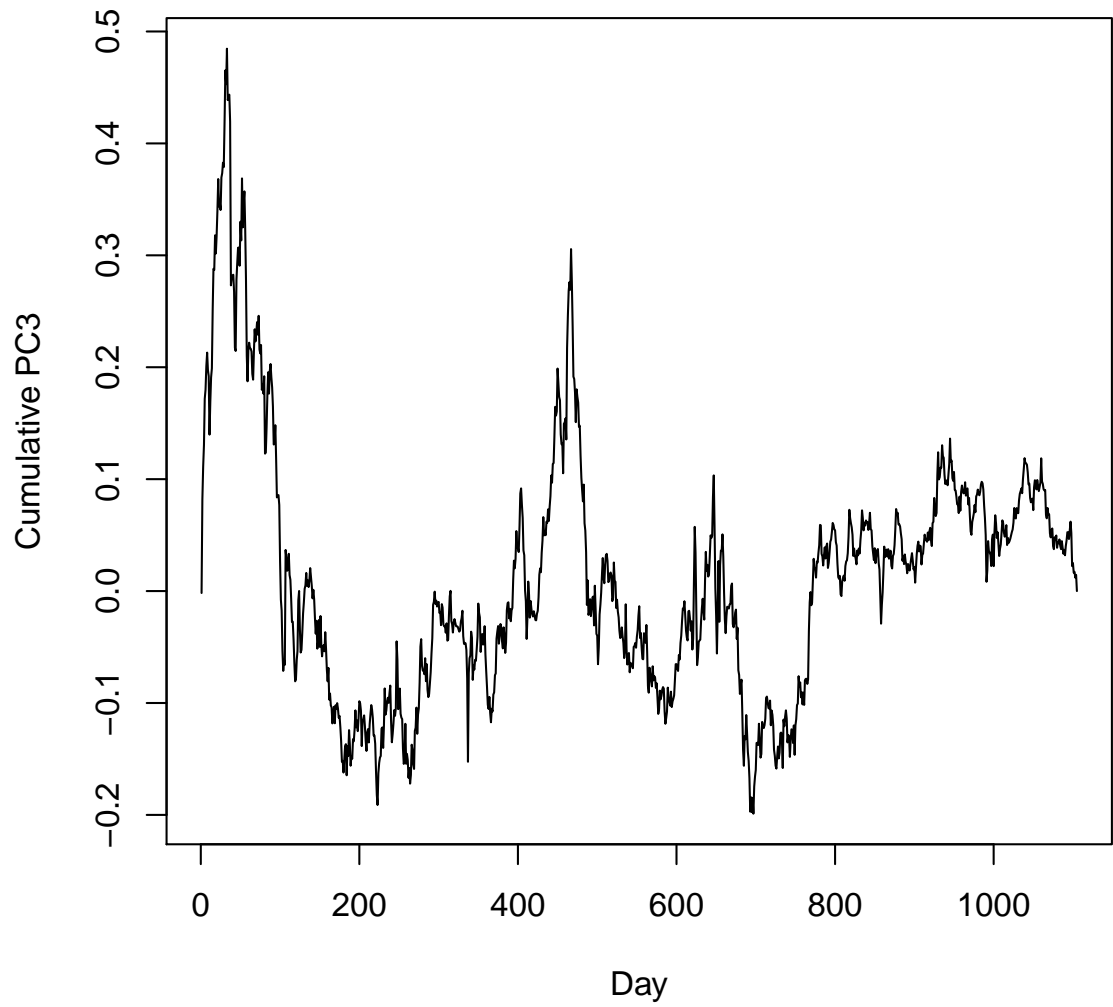## Cumulative PC Variables
PC1 (Level−Shifts)

```
> ts.plot(cumsum(obj.princomp0$scores[,2]),
+         main="Cumulative PC Variables\nPC2 (Spread Between Long and Short Maturities)",
+         ylab="Cumulative PC2",xlab="Day")
```

Cumulative PC Variables
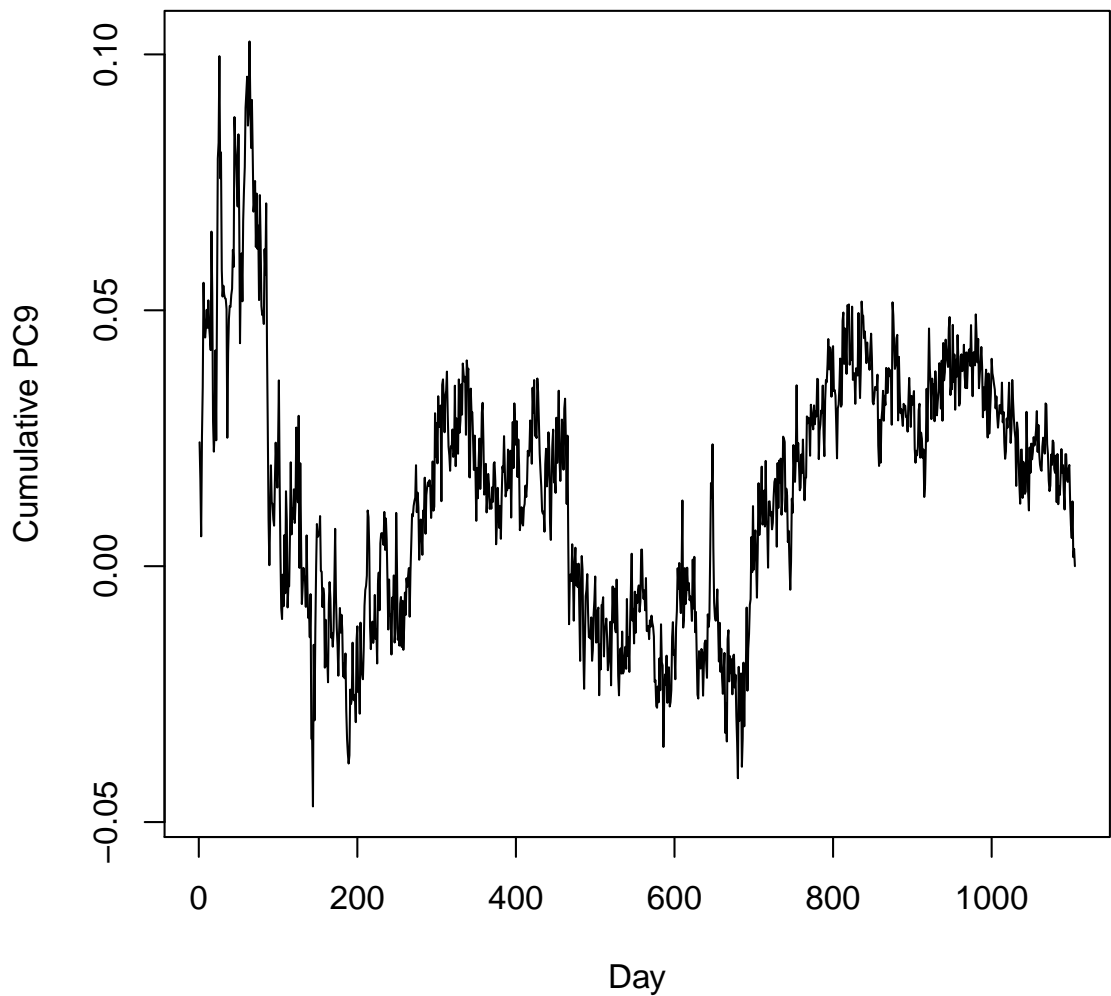## PC2 (Spread Between Long and Short Maturities)

```
> ts.plot(cumsum(obj.princomp0$scores[,3]),
+         main="Cumulative PC Variables\nPC3 (Curvature of Term Structure)",
+         ylab="Cumulative PC3",xlab="Day")
```

Cumulative PC Variables
PC3 (Curvature of Term Structure)

```
> ts.plot(cumsum(obj.princomp0$scores[,9]),
+         main="Cumulative PC Variables\nPC9 (7 Year vs 5 and 10 Year)",
+         ylab="Cumulative PC9",
+         xlab="Day")
> #########################
>
```

Cumulative PC Variables
PC9 (7 Year vs 5 and 10 Year)

3.3 Factor Analysis

The r function *factanal*() performs the maximum-likelihood factor analysis. The output includes the results of the likeihood ratio test for the sufficiency of the given number of factors.

```
> print(factanal(yieldchanges.period2,factors=3))

Call:
factanal(x = yieldchanges.period2, factors = 3)

Uniquenesses:
DGS3MO DGS6MO   DGS1   DGS2   DGS3   DGS5   DGS7  DGS10  DGS20
 0.677  0.487  0.386  0.140  0.044  0.023  0.033  0.006  0.062

Loadings:
       Factor1 Factor2 Factor3
DGS3MO                  0.567
DGS6MO 0.116   0.200   0.677
DGS1   0.232   0.496   0.561
DGS2   0.436   0.759   0.307
DGS3   0.589   0.752   0.211
DGS5   0.771   0.603   0.141
DGS7   0.866   0.456   0.101
DGS10  0.937   0.314   0.131
DGS20  0.941   0.175   0.144

              Factor1 Factor2 Factor3
SS loadings     3.712   2.128   1.302
Proportion Var  0.412   0.236   0.145
Cumulative Var  0.412   0.649   0.793

Test of the hypothesis that 3 factors are sufficient.
The chi square statistic is 231.03 on 12 degrees of freedom.
The p-value is 1.22e-42

> print(factanal(yieldchanges.period2,factors=4))

Call:
factanal(x = yieldchanges.period2, factors = 4)

Uniquenesses:
DGS3MO DGS6MO   DGS1   DGS2   DGS3   DGS5   DGS7  DGS10  DGS20
 0.651  0.360  0.412  0.022  0.061  0.005  0.033  0.005  0.062

Loadings:
       Factor1 Factor2 Factor3 Factor4
```

```
DGS3MO                         0.588
DGS6MO  0.119   0.200   0.765
DGS1    0.246   0.540   0.486
DGS2    0.435   0.861   0.211
DGS3    0.627   0.701   0.178   0.152
DGS5    0.803   0.537   0.140   0.203
DGS7    0.887   0.400   0.104
DGS10   0.947   0.291   0.116
DGS20   0.938   0.182   0.117  -0.104


            Factor1 Factor2 Factor3 Factor4
SS loadings   3.867   2.131   1.302   0.089
Proportion Var  0.430   0.237   0.145   0.010
Cumulative Var  0.430   0.666   0.811   0.821


Test of the hypothesis that 4 factors are sufficient.
The chi square statistic is 5.23 on 6 degrees of freedom.
The p-value is 0.515
```

For the period since 2009, the yield changes can be described by a 4-factor model.

The uniquenesses are the specific variances of the component yield changes; note the high the value for the 3-month yield changes (DGS3MO) as compared with the other tenors.

18.S096 Mathematical Applications in Financial Industry
Fall 2013