SYR-MBA FIN 654 Financial Analytics Project 2 (PRJ04)

Khan and Khalifa

February 19, 2017

Project 2 (PRJ04)

This project uses recent market data to support a decision regarding the provision of captive financing to customers. The focus is on building and interpreting a forward curve, testing more than one model, plotting results, and applying the results to a a collateral management scenario.

Problem

As we researched how to provide captive financing and insurance for our customers, we found that we needed to understand the relationships among lending rates and various terms and conditions of typical equipment financing contracts. Now we will extend that analysis to the long term provision of financing terms for customers. This will involve building a monitoring and management system of revolving credit for customer accounts.

We will focus on one question:

> *What is the influence of the term structure on the management of collateral to support long term financial arrangements with customers?*

Note: "A captive finance company is a subsidiary whose purpose is to provide financing to customers buying the parent company's product." [Wikipedia]

Data

The data set `termstrc.csv` contains data from the Wall Street Journal's market data site, which we will use to get some high level insights. The cross-sectional data covers the maturity date and price of STRIPs principal-only prices, otherwise known as zero coupon bond prices.

Variable	Description	Units of Measure
PRICE	market value of zero coupon bond	percent of face value
MATURITY	date at which zero coupon bond expires	date month/day/year

Source: Wall Street Journal marketdata.

Data Description

STRIPs (Separate Trading of Registered Interest and Principal) are zero-coupon bonds created from coupon bonds, based on the timeline for the cash flows for the coupon payment and principal payment. Each coupon payment and the principal are traded as separate securities with different maturities. STRIPs offer no interest payment; payment is only at maturity. Such bonds are attractive to investors who see current yields high enough to lock in, because they believe interest rates will decline in years to come.

Work Flow

Data Exploration

```r
# Use `read.csv` to read the data into `R`, with the working
# directory set appropriately
x.data <- read.csv("data/termstr20170127.csv")
# Use `na.omit()` to clean the data
x.data <- na.omit(x.data)
# Inspect some of the data at the top and bottom
head(x.data, n = 3)
```

```
##    PRICE MATURITY
## 1 99.991       18
## 2 99.795       46
## 3 99.769       77
```

```r
tail(x.data, n = 3)
```

```
##      PRICE MATURITY
## 189 39.621    10699
## 190 39.415    10791
## 191 39.302    10883
```

```r
# Check the columns, sort the data and express Maturity in years
# (divide by 365)
names(x.data)
```

```
## [1] "PRICE"    "MATURITY"
```

```r
x.data <- x.data[order(x.data$MATURITY), ]
x.data$Tm <- x.data$MATURITY/365  # express MATURITY in years
# Cut out the 2nd column (MATURITY)
x.data <- x.data[, -2]
names(x.data)
```

```
## [1] "PRICE" "Tm"
```

```r
# Inspect some of the data and the structure
head(x.data, n = 3)
```

```
##    PRICE         Tm
## 1 99.991 0.04931507
## 2 99.795 0.12602740
## 3 99.769 0.21095890
```

```r
tail(x.data, n = 3)
```

```
##      PRICE       Tm
## 189 39.621 29.31233
## 190 39.415 29.56438
## 191 39.302 29.81644
```

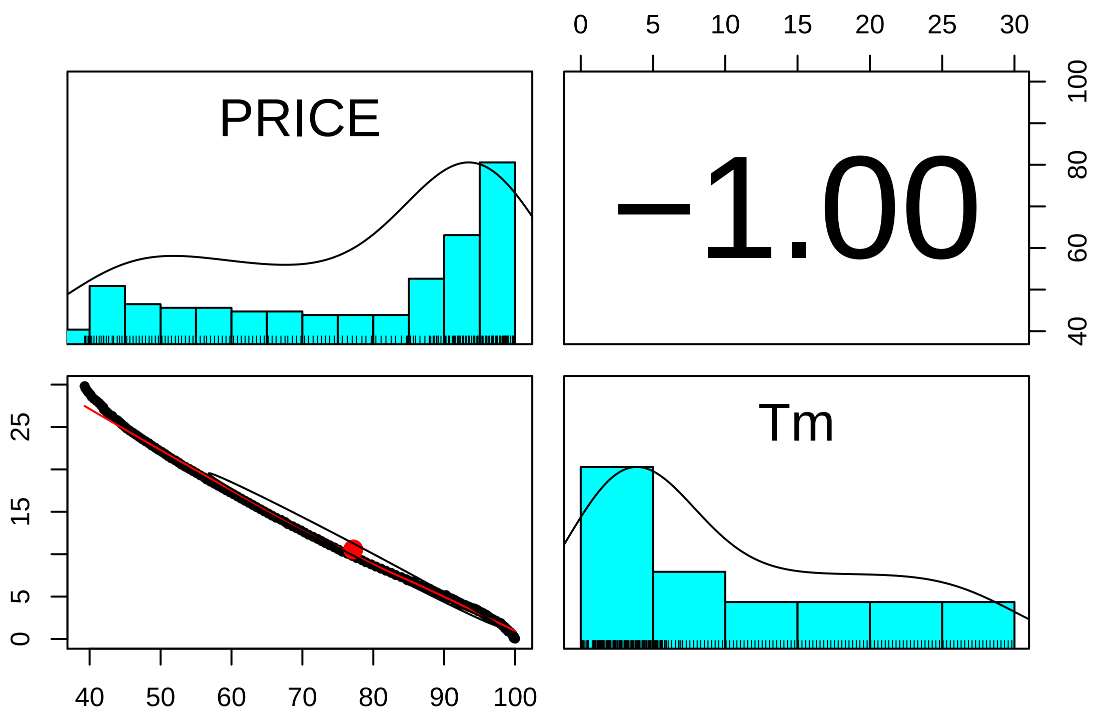
```r
str(x.data)
```

```
## 'data.frame':    191 obs. of  2 variables:
##  $ PRICE: num  100 99.8 99.8 99.8 99.7 ...
##  $ Tm   : num  0.0493 0.126 0.211 0.2932 0.3781 ...
```

```r
# Show summary statistics
summary(x.data)
```

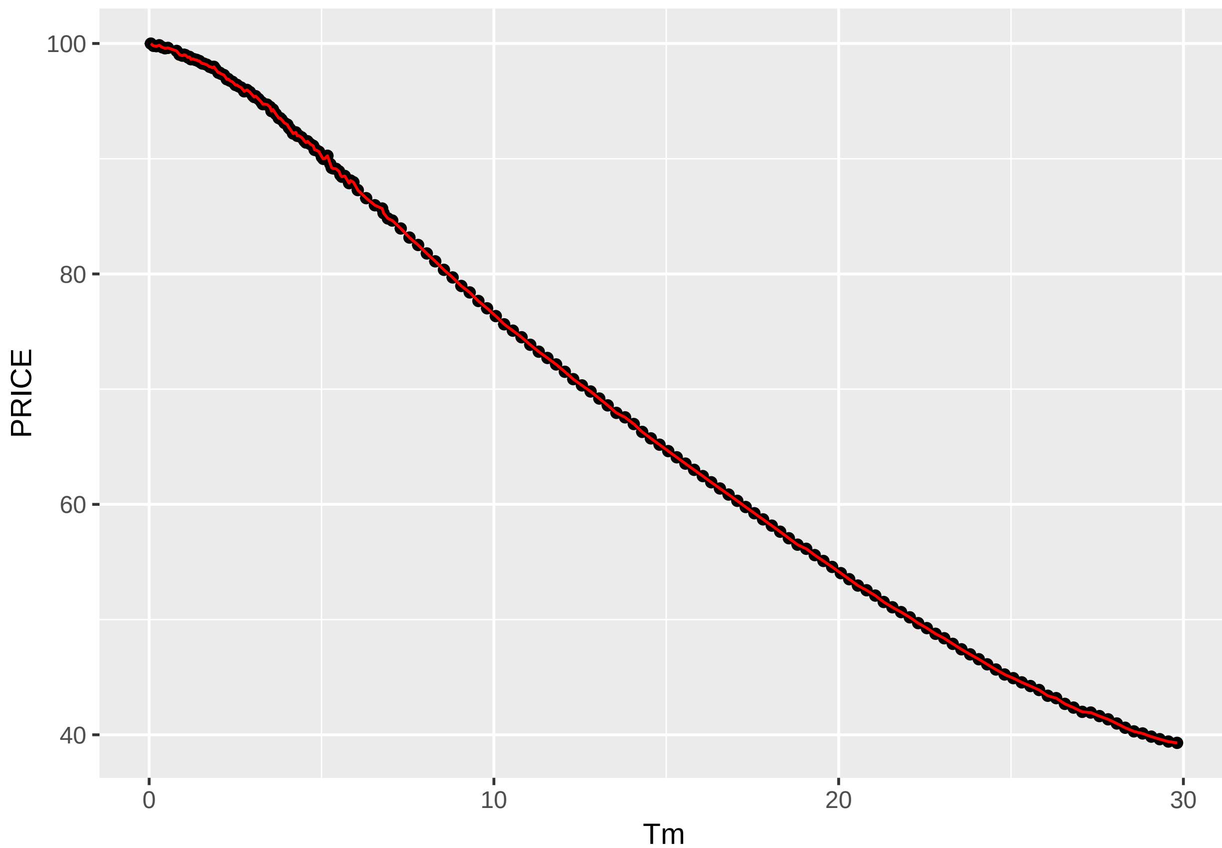
```
##      PRICE             Tm
##  Min.   :39.30   Min.   : 0.04931
##  1st Qu.:58.43   1st Qu.: 3.06849
##  Median :85.96   Median : 6.54795
##  Mean   :77.17   Mean   :10.55737
##  3rd Qu.:95.39   3rd Qu.:17.93425
##  Max.   :99.99   Max.   :29.81644
```

```r
# Load the psych library and produce a scatterplot matrix
require(psych)
pairs.panels(x.data)
```



Diagonals show histogram and trends for the variables: PRICE and Tm. Upper right of diagonal indicates that PRICE and Tm have perfect inverse correlation. Lower left of diagonal is the scatter plots (partial correlation) between PRICE and Tm.

```r
#' Plot the graph of PRICE against Tm
#' plot(x.data$Tm, x.data$PRICE, main = 'STRIPs', xlab = 'Tm', ylab = 'Price')
#' Or, use ggplot2()
require(ggplot2)
ggplot(data = x.data, aes(x = Tm, y = PRICE)) + geom_point() + geom_line(colour = "red")
```



The graph above confirms how bond prices decline over time as they reach maturity.

Data Analysis

The amount and pricing of collateral is often impacted by the term structure (yield curve) of interest rates. Modeling bond prices helps in understanding the underlying dynamics of the term structure, and manage cash volatility. In any such model, we formulate forward rates of return (future yields) on bonds of various maturities, and average these rates across the time to maturity in a yield calculation. Having sampled the STRIPs data and expressed the maturity in years, we are ready to estimate the empirical forward curve and interpret our findings.

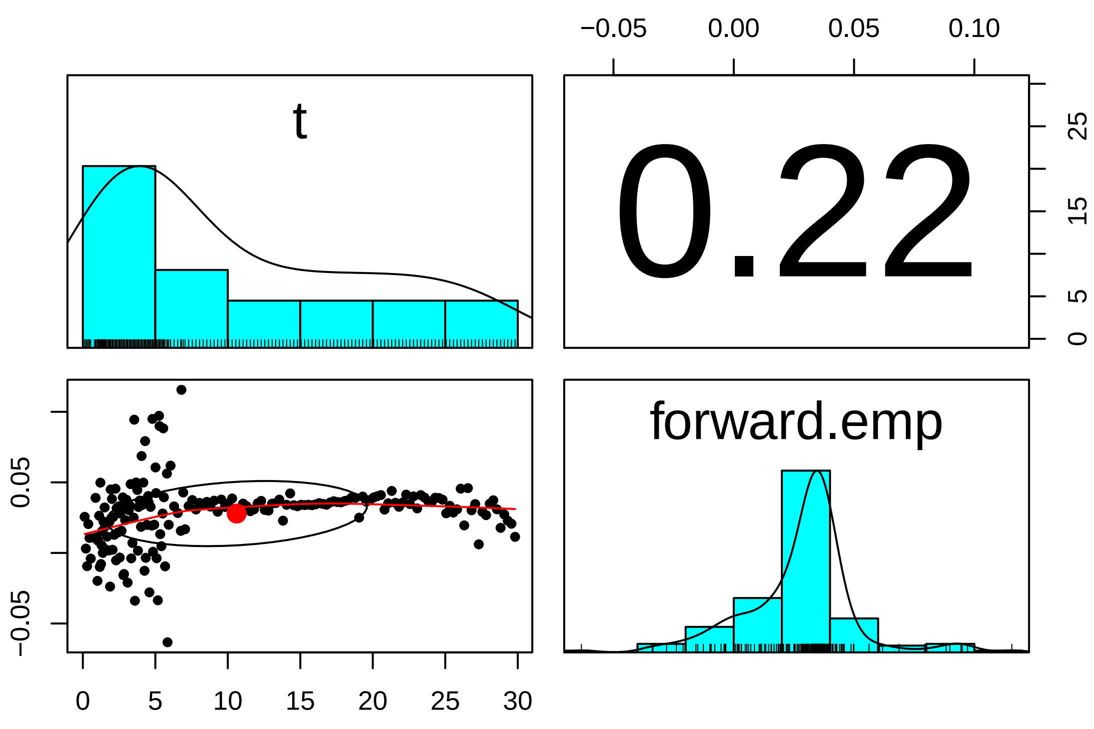
We can estimate the rate r(t) from the bond prices using nonlinear regression. The empirical forward-rate estimates give a general impression of the forward-rate curve and are useful for comparing with estimates from parametric models. These estimates are based on the formula

$$-\Delta \frac{log\{P(T_i)\}}{\Delta T_i} = -\frac{log\{P(T_i)\} - log\{P(T_{i-1})\}}{T_i - T_{i-1}}$$

```
t <- x.data$Tm[-1]  # seq(0,30,length = 100)
forward.emp <- -diff(log(x.data$PRICE))/diff(x.data$Tm)
forward.emp.df <- data.frame(t = t, forward.emp = forward.emp)
str(forward.emp.df)
```

```
## 'data.frame':    190 obs. of  2 variables:
##  $ t          : num  0.126 0.211 0.293 0.378 0.46 ...
##  $ forward.emp: num  0.02558 0.00307 -0.00939 0.02042 0.01062 ...
```

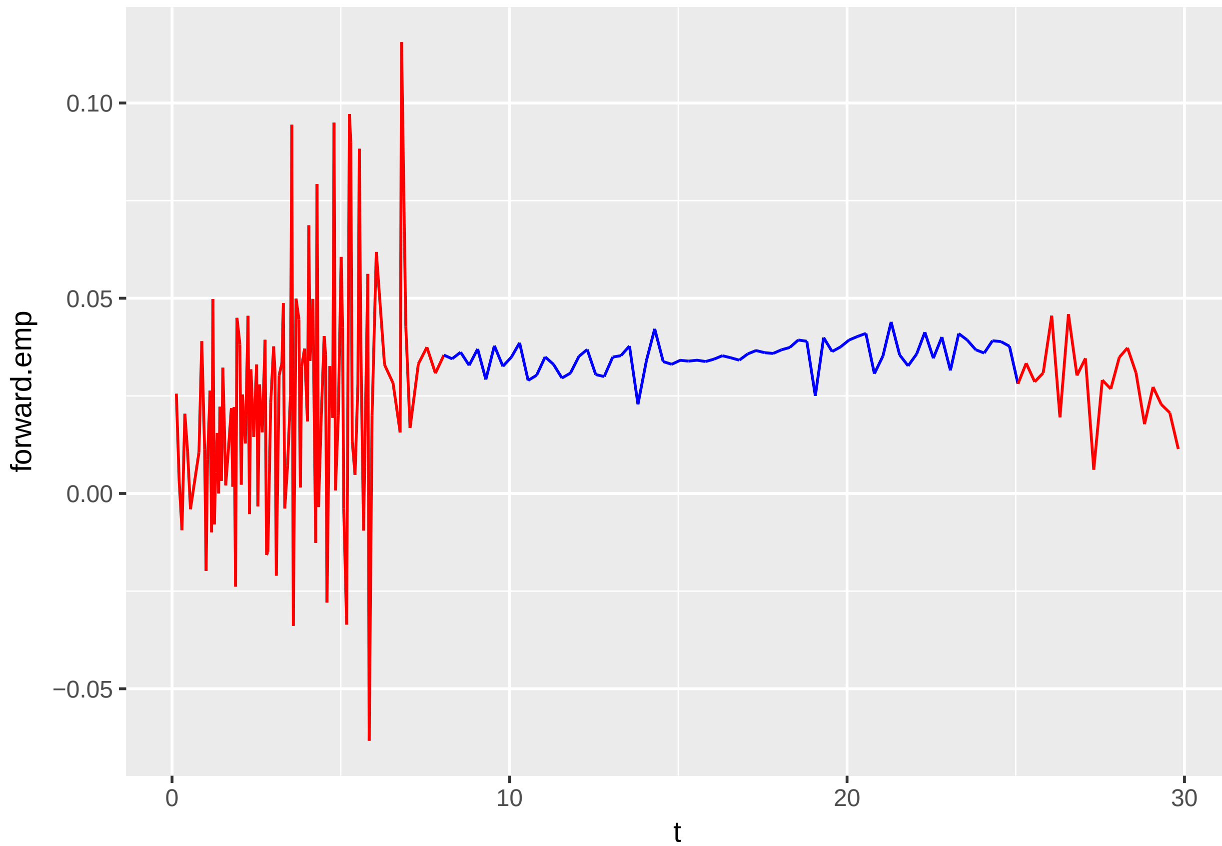
```
pairs.panels(forward.emp.df)
```



We see significant clustering and volatility across maturities less than eight years. More stability is found from years eight through twenty-five. This suggests at least two knots, one at maturity of eight years and the other at twenty-five years. Effectively there appear to be at least three distributions of forward rates conditional on maturities in this data.

We further use a line graph and colors to highlight the possible three or more distributions of forward rates.

```
col.emp <- ifelse(forward.emp.df$t < 8, "red", ifelse(forward.emp.df$t >
    25, "red", "blue"))
forward.emp.df <- data.frame(t = t, forward.emp = forward.emp, col.emp = col.emp)
ggplot(forward.emp.df, aes(x = t, y = forward.emp)) + geom_line(colour = col.emp,
    group = 1) + scale_colour_identity()
```



We now look at the parametric models. We again re-iterate the placement of at least two knots in the parametric models, one at maturity of eight years and the other at twenty-five years.

First, we fit a cubic spline to the zero price and maturity combinations.

```
fit.spline <- nls(PRICE ~ 100 * exp(-theta_0 * Tm - (theta_1 * Tm^2)/2 -
    (theta_2 * Tm^3)/3 - (Tm < 8) * (theta_3 * (Tm - 8)^3)/3 - (Tm >
    25) * (theta_4 * (Tm - 25)^3)/3), data = x.data, start = list(theta_0 = 0.01,
    theta_1 = 0.0024, theta_2 = -1e-04, theta_3 = 1e-05, theta_4 = 1e-04))
```

Then, we fit 5%, 50% and 95% envelopes.

```
require(quantreg)
taus <- c(0.05, 0.5, 0.95)
fit.spline.q.05 <- nlrq(PRICE ~ 100 * exp(-theta_0 * Tm - (theta_1 *
    Tm^2)/2 - (theta_2 * Tm^3)/3 - (Tm < 8) * (theta_3 * (Tm - 8)^3)/3 -
    (Tm > 25) * (theta_4 * (Tm - 25)^3)/3), data = x.data, tau = taus[1],
    start = list(theta_0 = 0.01, theta_1 = 0.0024, theta_2 = -1e-04,
```

```
        theta_3 = 1e-05, theta_4 = 1e-04))
fit.spline.q.50 <- nlrq(PRICE ~ 100 * exp(-theta_0 * Tm - (theta_1 *
    Tm^2)/2 - (theta_2 * Tm^3)/3 - (Tm < 8) * (theta_3 * (Tm - 8)^3)/3 -
    (Tm > 25) * (theta_4 * (Tm - 25)^3)/3), data = x.data, tau = taus[2],
    start = list(theta_0 = 0.01, theta_1 = 0.0024, theta_2 = -1e-04,
        theta_3 = 1e-05, theta_4 = 1e-04))
fit.spline.q.95 <- nlrq(PRICE ~ 100 * exp(-theta_0 * Tm - (theta_1 *
    Tm^2)/2 - (theta_2 * Tm^3)/3 - (Tm < 8) * (theta_3 * (Tm - 8)^3)/3 -
    (Tm > 25) * (theta_4 * (Tm - 25)^3)/3), data = x.data, tau = taus[3],
    start = list(theta_0 = 0.01, theta_1 = 0.0024, theta_2 = -1e-04,
        theta_3 = 1e-05, theta_4 = 1e-04))
```

Finally, we fit a quadratic polynomial, for comparison.

```
fit.quad <- nls(PRICE ~ 100 * exp(-theta_0 * Tm - (theta_1 * Tm^2)/2 -
    (theta_2 * Tm^3)/3), data = x.data, start = list(theta_0 = 0.01,
    theta_1 = 0.0024, theta_2 = -1e-04))
```

What anomalies appear based on these procedures? We have noted the irregularities in the first 8 years and then some after 25 years. We now look at the standarad deviations for the models, and the p-values associated with the coefficients.

```
knitr::kable(summary(fit.spline)$coefficients, digits = 4)
```

	Estimate	Std. Error	t value	Pr(>|t|)
theta_0	0.0144	3e-04	54.7512	0.0000
theta_1	0.0028	1e-04	39.7473	0.0000
theta_2	-0.0001	0e+00	-26.5398	0.0000
theta_3	0.0001	0e+00	8.7560	0.0000
theta_4	0.0006	3e-04	2.2262	0.0272

```
(sigma.spline <- (summary(fit.spline)$sigma)^0.5)
```

```
## [1] 0.6778881
```

```
knitr::kable(summary(fit.spline.q.05)$coefficients, digits = 4)
```

	Value	Std. Error	t value	Pr(>|t|)
theta_0	0.0156	0.0013	12.2286	0.0000
theta_1	0.0028	0.0004	7.7567	0.0000
theta_2	-0.0001	0.0000	-5.9081	0.0000
theta_3	0.0000	0.0000	0.5370	0.5919
theta_4	0.0005	0.0006	0.8415	0.4011

```
knitr::kable(summary(fit.spline.q.50)$coefficients, digits = 4)
```

	Value	Std. Error	t value	Pr(>|t|)
theta_0	0.0148	4e-04	39.1211	0.0000
theta_1	0.0026	1e-04	26.4951	0.0000
theta_2	-0.0001	0e+00	-17.4343	0.0000
theta_3	0.0001	0e+00	5.2501	0.0000
theta_4	0.0000	3e-04	0.0854	0.9321

```r
knitr::kable(summary(fit.spline.q.95)$coefficients, digits = 4)
```

	Value	Std. Error	t value	Pr(>|t|)
theta_0	0.0122	7e-04	18.7522	0.0000
theta_1	0.0030	1e-04	20.5864	0.0000
theta_2	-0.0001	0e+00	-13.3028	0.0000
theta_3	0.0001	0e+00	3.1460	0.0019
theta_4	0.0000	4e-04	0.0061	0.9952

```r
knitr::kable(summary(fit.quad)$coefficients, digits = 4)
```

	Estimate	Std. Error	t value	Pr(>|t|)
theta_0	0.0137	3e-04	51.5601	0
theta_1	0.0029	1e-04	43.4853	0
theta_2	-0.0001	0e+00	-30.7040	0

```r
(sigma.quad <- (summary(fit.quad)$sigma)^0.5)
```

```
## [1] 0.7453241
```

The standard deviation is lower for the cubic spline than for the quadratic polynomial. Based on the p-values, theta_4 coefficients are not significant for any of the models. Additionally, theta_3 coefficient is not significant for the quadratic polynomial, as expected.

We now build the forward curves and examine the structures.

```r
coef.spline <- summary(fit.spline)$coef[, 1]
forward.spline <- coef.spline[1] + coef.spline[2] * t + coef.spline[3] *
    t^2 + (t < 8) * coef.spline[4] * (t - 8)^2 + (t > 25) * coef.spline[5] *
    (t - 25)^2
coef.spline <- summary(fit.spline.q.05)$coef[, 1]
forward.spline.q.05 <- coef.spline[1] + coef.spline[2] * t + coef.spline[3] *
    t^2 + (t < 8) * coef.spline[4] * (t - 8)^2 + (t > 25) * coef.spline[5] *
    (t - 25)^2
coef.spline <- summary(fit.spline.q.50)$coef[, 1]
forward.spline.q.50 <- coef.spline[1] + coef.spline[2] * t + coef.spline[3] *
    t^2 + (t < 8) * coef.spline[4] * (t - 8)^2 + (t > 25) * coef.spline[5] *
    (t - 25)^2
coef.spline <- summary(fit.spline.q.95)$coef[, 1]
forward.spline.q.95 <- coef.spline[1] + coef.spline[2] * t + coef.spline[3] *
    t^2 + (t < 8) * coef.spline[4] * (t - 8)^2 + (t > 25) * coef.spline[5] *
    (t - 25)^2
coef.quad <- summary(fit.quad)$coef[, 1]
forward.quad <- coef.spline[1] + coef.spline[2] * t + coef.spline[3] *
    t^2
str(forward.spline)
```

```
##  num [1:190] 0.0189 0.0191 0.0192 0.0193 0.0195 ...
```

```r
str(forward.spline.q.50)
```

```
##  num [1:190] 0.0206 0.0207 0.0208 0.0209 0.021 ...
```

```r
str(forward.quad)
```

```
##  num [1:190] 0.0126 0.0129 0.0131 0.0134 0.0136 ...
```

Interpretation and Results

The parameters (coefficients) may be interpreted thus:

- theta_0 is independent of maturity and thus represents the long-run average forward rate.
- theta_1 helps to measure the average sensitivity of forward rates to a change in maturity.
- theta_2 helps to measure the maturity risk of the forward curve for the instrument.
- theta_3 helps to compensate for the directional change in the forward curve.
- theta_4 helps to further qualify the change in the forward curve, when appicable.

We pull the coefficients from a `summary()` of the `fit.spline` object for the plot, and then compare the different models.

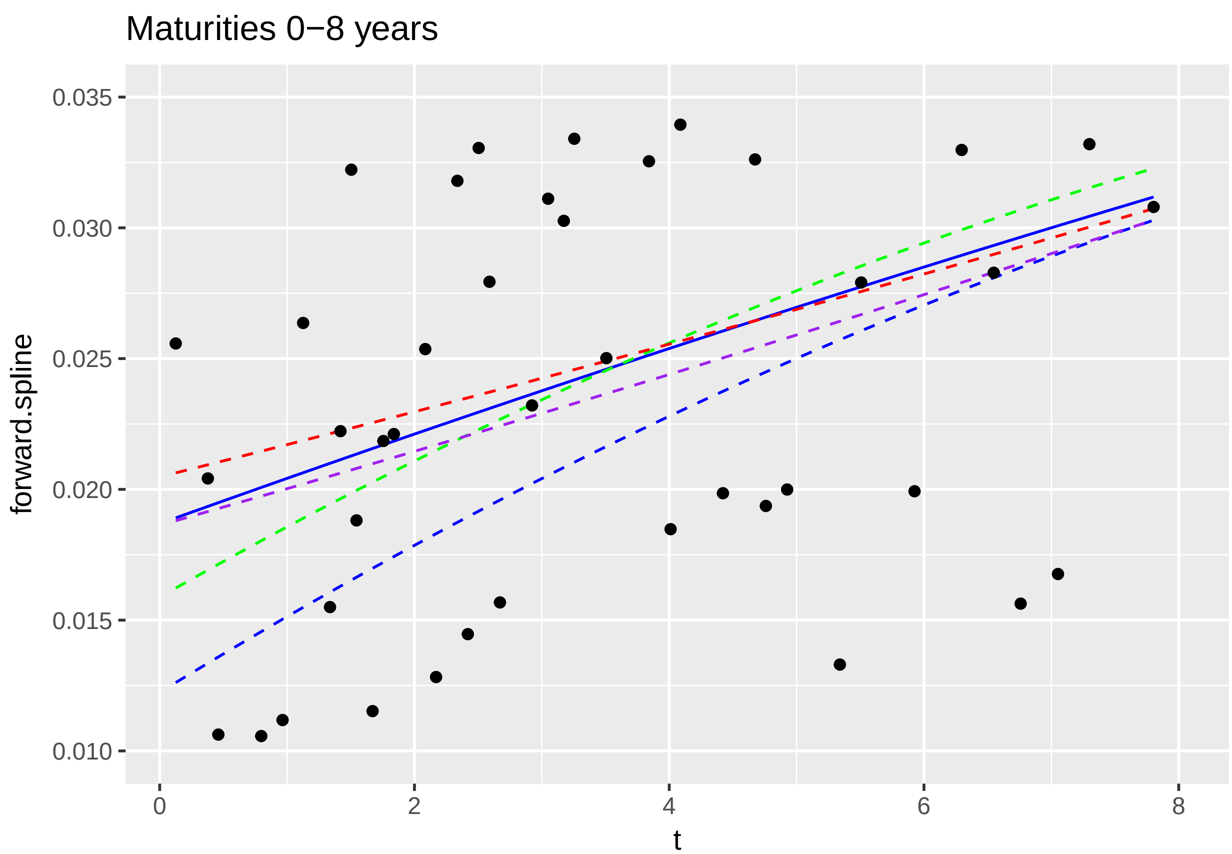
We do that by first preparing the data frame to include all the spline and polynomial models, and examine the structure.

We then zoom into the curves to see the action.

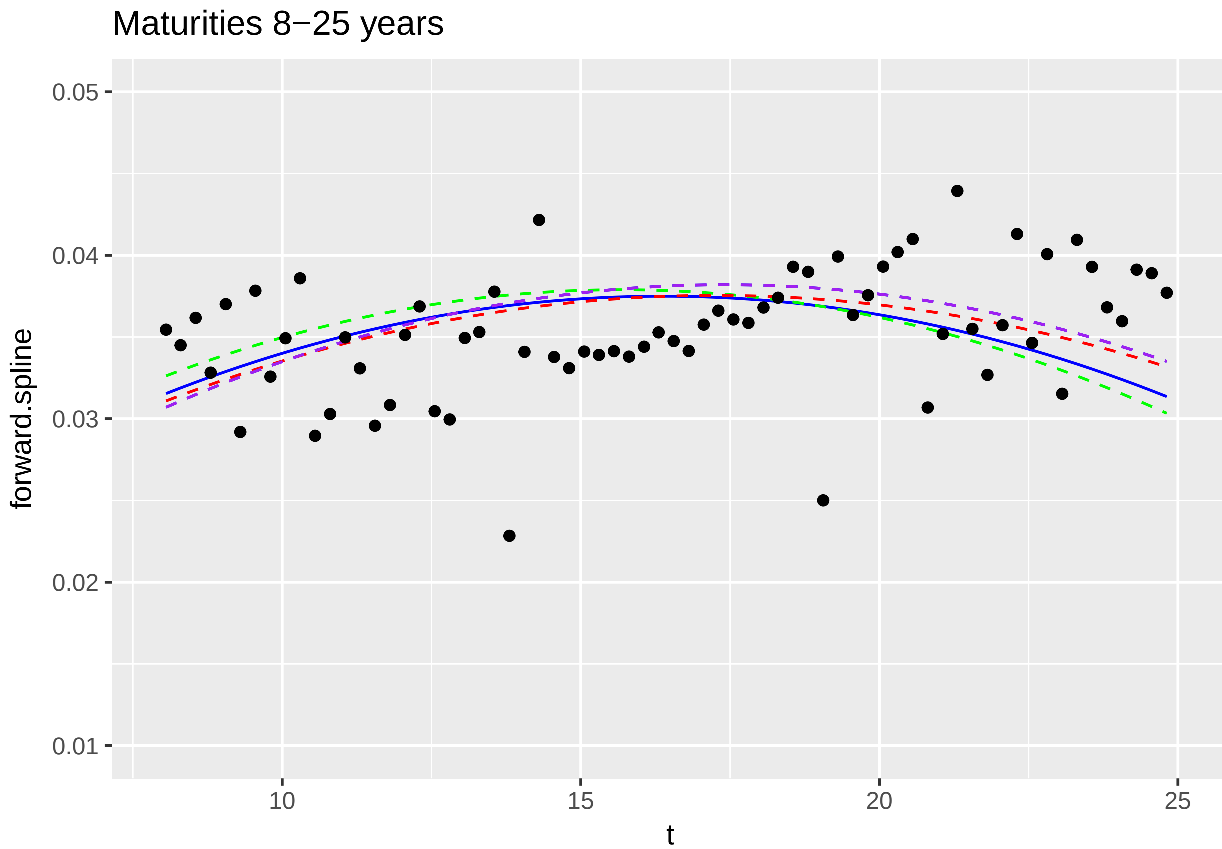
```r
x.spl <- data.frame(t = t, forward.spline = forward.spline, forward.spline.q.05 = forward.spline.q.05,
    forward.spline.q.50 = forward.spline.q.50, forward.spline.q.95 = forward.spline.q.95,
    forward.quad = forward.quad, forward.emp = forward.emp)
str(x.spl)
```

```
## 'data.frame':    190 obs. of  7 variables:
##  $ t                  : num  0.126 0.211 0.293 0.378 0.46 ...
##  $ forward.spline     : num  0.0189 0.0191 0.0192 0.0193 0.0195 ...
##  $ forward.spline.q.05: num  0.0162 0.0165 0.0167 0.0169 0.0171 ...
##  $ forward.spline.q.50: num  0.0206 0.0207 0.0208 0.0209 0.021 ...
##  $ forward.spline.q.95: num  0.0188 0.0189 0.019 0.0192 0.0193 ...
##  $ forward.quad       : num  0.0126 0.0129 0.0131 0.0134 0.0136 ...
##  $ forward.emp        : num  0.02558 0.00307 -0.00939 0.02042 0.01062 ...
```

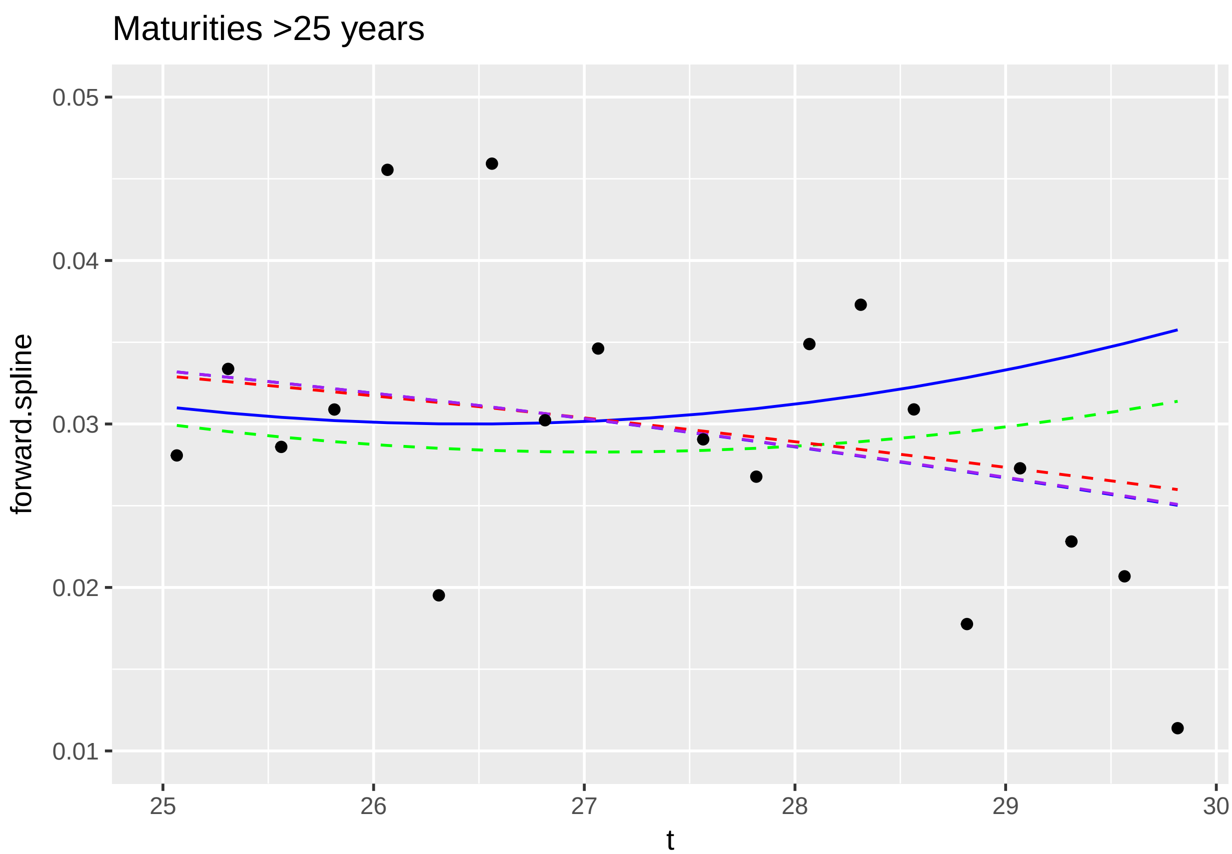
```r
# 0-8 years maturity
ggplot(data = x.spl, aes(x = t, y = forward.spline)) + geom_line(colour = "blue") +
    geom_line(aes(x = t, y = forward.quad), colour = "blue", linetype = "dashed") +
    geom_line(aes(x = t, y = forward.spline.q.05), colour = "green",
        linetype = "dashed") + geom_line(aes(x = t, y = forward.spline.q.50),
    colour = "red", linetype = "dashed") + geom_line(aes(x = t, y = forward.spline.q.95),
    colour = "purple", linetype = "dashed") + geom_point(aes(x = t,
    y = forward.emp)) + ylim(0.01, 0.035) + xlim(min(t), 8) + ggtitle("Maturities 0-8 years")
```




```r
# 8-25 years maturity
ggplot(data = x.spl, aes(x = t, y = forward.spline)) + geom_line(colour = "blue") +
    geom_line(aes(x = t, y = forward.quad), colour = "blue", linetype = "dashed") +
    geom_line(aes(x = t, y = forward.spline.q.05), colour = "green",
        linetype = "dashed") + geom_line(aes(x = t, y = forward.spline.q.50),
    colour = "red", linetype = "dashed") + geom_line(aes(x = t, y = forward.spline.q.95),
    colour = "purple", linetype = "dashed") + geom_point(aes(x = t,
    y = forward.emp)) + ylim(0.01, 0.05) + xlim(8, 25) + ggtitle("Maturities 8-25 years")
```




```
# > 25 years maturity
ggplot(data = x.spl, aes(x = t, y = forward.spline)) + geom_line(colour = "blue") +
    geom_line(aes(x = t, y = forward.quad), colour = "blue", linetype = "dashed") +
    geom_line(aes(x = t, y = forward.spline.q.05), colour = "green",
        linetype = "dashed") + geom_line(aes(x = t, y = forward.spline.q.50),
    colour = "red", linetype = "dashed") + geom_line(aes(x = t, y = forward.spline.q.95),
    colour = "purple", linetype = "dashed") + geom_point(aes(x = t,
    y = forward.emp)) + ylim(0.01, 0.05) + xlim(25, max(t)) + ggtitle("Maturities >25 years")
```

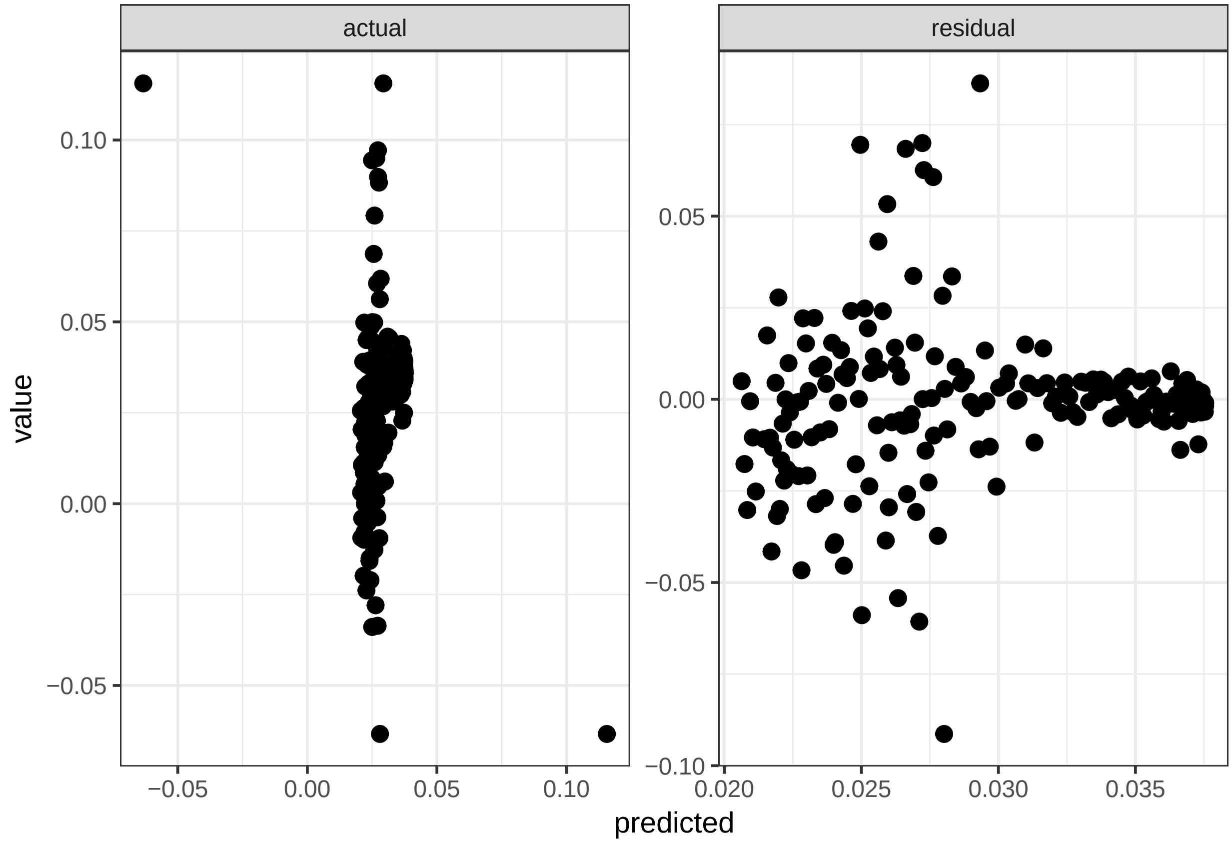


The cubic polynomial and quadratic spline models follow the empirical forward rates much more closely than the quadratic polynomial model. While they are all similar for the middle maturities 8-25 years. Though the quadratic polynomial appears to converge with the other models for the maturities 0-8 years, the polynomials are definitely not a good fit based on maturities greater than 25 years. These would come into play when considering [longer] term loans. The goal would be to negotiate the longer term rates down to where they are today.

We further compare the residuals to predicted for the median (50%) and empirical models.

```
require(reshape2)
actual <- forward.emp
predicted <- forward.spline.q.50
residual <- actual - predicted
results <- data.frame(actual = actual, predicted = predicted, residual = residual)
# Compute the range (minum and maximum) for the actual and
# predicted values
min_xy <- min(min(results$actual), min(results$predicted))
max_xy <- max(max(results$actual), max(results$predicted))
# Compute the maximum across the actual and predicted values
plot.melt <- melt(results, id.vars = "predicted")
# Use melt() from reshape2 to build the data frame with predicted
# as id and values of variables
plot.data <- rbind(plot.melt, data.frame(predicted = c(min_xy, max_xy),
    variable = c("actual", "actual"), value = c(max_xy, min_xy)))
# Graph the plots side by side for the values of predicted and
# residuals
p <- ggplot(plot.data, aes(x = predicted, y = value)) + geom_point(size = 2.5) +
```

```r
    theme_bw()
p <- p + facet_wrap(~variable, scales = "free")
p
```



There are some large "leakages" (anomalies) in the corners, otherwise the values are pretty tight. The bulge in the middle for the actuals is likely from maturities 0-8 years The bulge to the right in the residuals are also probably from maturities 0-8 years. They represent the short term. But some of the residuals (7 bonds) add to a lot of uncertainty.

Scenario

Scenario Description

1. Suppose we just bought a 10 year maturity zero-coupon bond to satisfy collateral requirements for workers' compensation in the (great) State of New York.
2. The forward rate can been estimated using a spline, a pure quadratic, and across quantiles. We believe that the 50%tile of zero-coupon bond prices is the appropriate threshold for estimation.

$$\int r_{0.50}(t)dt = \theta_{0,0.50}t + \theta_{1,0.50}\frac{t^2}{2} + \theta_{2,0.50}\frac{t^3}{3} + \theta_{3,0.50}\frac{(t-8)^3_-}{3} + \theta_{4,0.50}\frac{(t-25)^3_+}{3}$$

   We will estimate this curve (today's curve) and explain in detail the choice of a particular model.
3. In 6 months there is a very high likelihood we will need to exit all business in New York State, have no employees that can claim workers' compensation. If we do so, we will sell the 10 year maturity

zero-coupon bond. We project that the forward curve will be at the 75%tile of the current data, so that

$$\int r_{0.75}(t)dt = \theta_{0,0.75}t + \theta_{1,0.75}\frac{t^2}{2} + \theta_{2,0.75}\frac{t^3}{3} + \theta_{3,0.75}\frac{(t-8)^3_-}{3} + \theta_{4,0.75}\frac{(t-25)^3_+}{3}$$

We will estimate this curve (6 months out) and explain in detail the choice of a particular model.

4. How much would we gain or lose on this transaction at our exit?

Scenario Workflow

Let's recall the following: i) The forward rate is the rate of change of the yield-to-maturity ii) This means we integrate (i.e., take the cumulative sum of) forward rates to get the yield iii) The cumulative sum would then be some maturity times the components of the yield curve adjusted for the slope of the forward curve (the terms in Tm). iv) This adjustment is just one-half (1/2) of the slope term.

We then calculate the six month (0.5 year maturity) return.

$$R_{6m} = \frac{P(T-0.5) - P(T)}{P(T)} = \frac{P(T-0.5)}{P(T)} - 1$$

Use appropriate taus - we really need 50th and 75th only.

```
require(quantreg)
taus <- c(0.25, 0.5, 0.75)
fit.spline.q.25 <- nlrq(PRICE ~ 100 * exp(-theta_0 * Tm - (theta_1 *
    Tm^2)/2 - (theta_2 * Tm^3)/3 - (Tm < 8) * (theta_3 * (Tm - 8)^3)/3 -
    (Tm > 25) * (theta_4 * (Tm - 25)^3)/3), data = x.data, tau = taus[1],
    start = list(theta_0 = 0.01, theta_1 = 0.0024, theta_2 = -1e-04,
        theta_3 = 1e-05, theta_4 = 1e-04))
fit.spline.q.50 <- nlrq(PRICE ~ 100 * exp(-theta_0 * Tm - (theta_1 *
    Tm^2)/2 - (theta_2 * Tm^3)/3 - (Tm < 8) * (theta_3 * (Tm - 8)^3)/3 -
    (Tm > 25) * (theta_4 * (Tm - 25)^3)/3), data = x.data, tau = taus[2],
    start = list(theta_0 = 0.01, theta_1 = 0.0024, theta_2 = -1e-04,
        theta_3 = 1e-05, theta_4 = 1e-04))
fit.spline.q.75 <- nlrq(PRICE ~ 100 * exp(-theta_0 * Tm - (theta_1 *
    Tm^2)/2 - (theta_2 * Tm^3)/3 - (Tm < 8) * (theta_3 * (Tm - 8)^3)/3 -
    (Tm > 25) * (theta_4 * (Tm - 25)^3)/3), data = x.data, tau = taus[3],
    start = list(theta_0 = 0.01, theta_1 = 0.0024, theta_2 = -1e-04,
        theta_3 = 1e-05, theta_4 = 1e-04))
```

Compare that with the quad model (optional).

```
knitr::kable(summary(fit.spline.q.25)$coefficients, digits = 4)
```

	Value	Std. Error	t value	Pr(>|t|)
theta_0	0.0155	5e-04	29.2905	0.0000
theta_1	0.0027	2e-04	15.8831	0.0000
theta_2	-0.0001	0e+00	-10.8555	0.0000
theta_3	0.0000	0e+00	2.8243	0.0053
theta_4	0.0004	3e-04	1.2313	0.2198

```
(sigma.spline <- (summary(fit.spline)$sigma)^0.5)
```

```
## [1] 0.6778881
```

```r
knitr::kable(summary(fit.spline.q.50)$coefficients, digits = 4)
```

	Value	Std. Error	t value	Pr(>|t|)
theta_0	0.0148	4e-04	35.8996	0.0000
theta_1	0.0026	1e-04	23.7817	0.0000
theta_2	-0.0001	0e+00	-15.4801	0.0000
theta_3	0.0001	0e+00	4.2669	0.0000
theta_4	0.0000	3e-04	0.0827	0.9342

```r
(sigma.spline <- (summary(fit.spline)$sigma)^0.5)
```

```
## [1] 0.6778881
```

```r
knitr::kable(summary(fit.spline.q.75)$coefficients, digits = 4)
```

	Value	Std. Error	t value	Pr(>|t|)
theta_0	0.0139	4e-04	37.4669	0.0000
theta_1	0.0027	1e-04	34.2952	0.0000
theta_2	-0.0001	0e+00	-22.9120	0.0000
theta_3	0.0001	0e+00	8.0406	0.0000
theta_4	0.0001	2e-04	0.3184	0.7505

```r
(sigma.spline <- (summary(fit.spline)$sigma)^0.5)
```

```
## [1] 0.6778881
```

```r
knitr::kable(summary(fit.quad)$coefficients, digits = 4)
```

	Estimate	Std. Error	t value	Pr(>|t|)
theta_0	0.0137	3e-04	51.5601	0
theta_1	0.0029	1e-04	43.4853	0
theta_2	-0.0001	0e+00	-30.7040	0

```r
(sigma.quad <- (summary(fit.quad)$sigma)^0.5)
```

```
## [1] 0.7453241
```

Derive coefficients and build forward curves.

```r
(coeff.25 <- summary(fit.spline.q.25)$coef[, 1])
```

```
##      theta_0      theta_1      theta_2      theta_3      theta_4
## 1.547564e-02 2.690580e-03 -8.228395e-05 4.759137e-05 4.245202e-04
```

```r
forward.spline.q.25 <- coeff.25[1] + coeff.25[2] * t + coeff.25[3] *
    t^2 + (t < 8) * coeff.25[4] * (t - 8)^2 + (t > 25) * coeff.25[5] *
    (t - 25)^2
(coeff.50 <- summary(fit.spline.q.50)$coef[, 1])
```

```
##      theta_0      theta_1      theta_2      theta_3      theta_4
## 1.482673e-02 2.633964e-03 -7.634591e-05 8.828272e-05 2.148477e-05
```

```
forward.spline.q.50 <- coeff.50[1] + coeff.50[2] * t + coeff.50[3] *
    t^2 + (t < 8) * coeff.50[4] * (t - 8)^2 + (t > 25) * coeff.50[5] *
    (t - 25)^2
(coeff.75 <- summary(fit.spline.q.75)$coef[, 1])
```

```
##       theta_0       theta_1       theta_2       theta_3       theta_4
##  1.392043e-02  2.727890e-03 -7.835498e-05  1.167364e-04  7.047712e-05
```

```
forward.spline.q.75 <- coeff.75[1] + coeff.75[2] * t + coeff.75[3] *
    t^2 + (t < 8) * coeff.75[4] * (t - 8)^2 + (t > 25) * coeff.75[5] *
    (t - 25)^2
```

Take a few points on the 50%tile model, plot the graph and validate.

```
coeff.now <- as.vector(coeff.50)
t.now <- 7
(r07.50 <- coeff.now[1] + coeff.now[2] * (t.now) + coeff.now[3] *
    (t.now^2) + (t.now < 8) * coeff.now[4] * (t.now - 8)^3/t.now +
    (t.now > 25) * coeff.now[5] * (t.now - 25)^3/t.now)
```

```
## [1] 0.02951092
```

```
t.now <- 14
(r14.50 <- coeff.now[1] + coeff.now[2] * (t.now) + coeff.now[3] *
    (t.now^2) + (t.now < 8) * coeff.now[4] * (t.now - 8)^3/t.now +
    (t.now > 25) * coeff.now[5] * (t.now - 25)^3/t.now)
```

```
## [1] 0.03673844
```

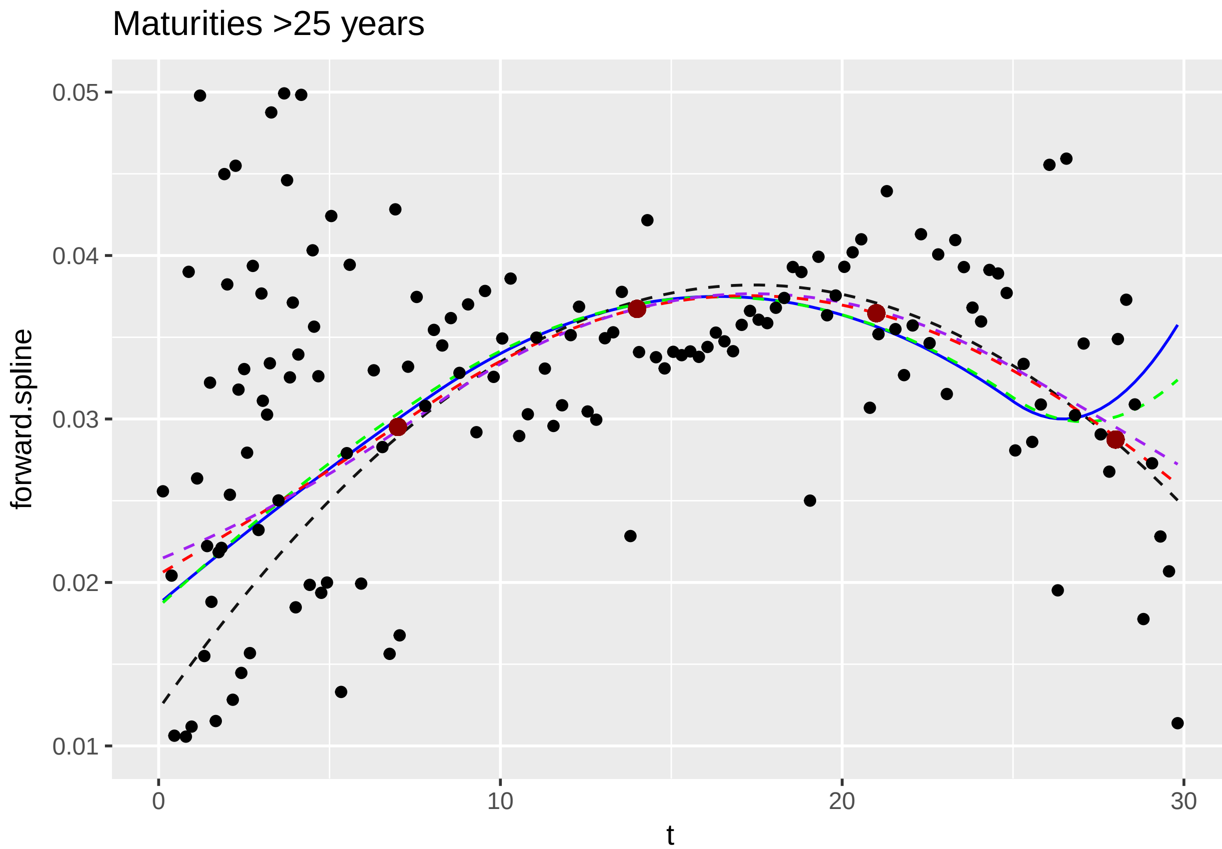
```
t.now <- 21
(r21.50 <- coeff.now[1] + coeff.now[2] * (t.now) + coeff.now[3] *
    (t.now^2) + (t.now < 8) * coeff.now[4] * (t.now - 8)^3/t.now +
    (t.now > 25) * coeff.now[5] * (t.now - 25)^3/t.now)
```

```
## [1] 0.03647144
```

```
t.now <- 28
(r28.50 <- coeff.now[1] + coeff.now[2] * (t.now) + coeff.now[3] *
    (t.now^2) + (t.now < 8) * coeff.now[4] * (t.now - 8)^3/t.now +
    (t.now > 25) * coeff.now[5] * (t.now - 25)^3/t.now)
```

```
## [1] 0.02874326
```

```
ggplot(data = x.spl, aes(x = t, y = forward.spline)) + geom_line(colour = "blue") +
    geom_line(aes(x = t, y = forward.quad), colour = "gray7", linetype = "dashed") +
    geom_line(aes(x = t, y = forward.spline.q.25), colour = "green",
        linetype = "dashed") + geom_line(aes(x = t, y = forward.spline.q.50),
    colour = "red", linetype = "dashed") + geom_line(aes(x = t, y = forward.spline.q.75),
    colour = "purple", linetype = "dashed") + geom_point(aes(x = t,
    y = forward.emp)) + geom_point(aes(x = 7, y = r07.50), colour = "darkred",
    size = 2.5) + geom_point(aes(x = 14, y = r14.50), colour = "darkred",
    size = 2.5) + geom_point(aes(x = 21, y = r21.50), colour = "darkred",
    size = 2.5) + geom_point(aes(x = 28, y = r28.50), colour = "darkred",
    size = 2.5) + ylim(0.01, 0.05) + xlim(min(t), max(t)) + ggtitle("Maturities >25 years")
```



Scenario Outcome

First, we set up today's yield curve and the curve 6 months out.

The yield curve for 50%tile is given by following equation.

$$y_{0.50}(T) = \theta_{0,0.50} + \theta_{1,0.50}\frac{T}{2} + \theta_{2,0.50}\frac{T^2}{3} + \theta_{3,0.50}\frac{(T-8)^3_-}{3T} + \theta_{4,0.50}\frac{(T-25)^3_+}{3T}$$

The yield curve for 75%tile is given by following equation.

$$y_{0.75}(T) = \theta_{0,0.75} + \theta_{1,0.75}\frac{T}{2} + \theta_{2,0.75}\frac{T^2}{3} + \theta_{3,0.75}\frac{(T-8)^3_-}{3T} + \theta_{4,0.75}\frac{(T-25)^3_+}{3T}$$

Then, using these yields we compute the bond prices (as a percentage of par) for today and for 6 months out as well.

$$p(T) = 100 * exp\{-T * y(T)\}$$

The return is the difference in the two relative to the first yield (multiplied by 2 when annualized) expressed as a percentage. Our exit transaction is long today's version of the bond and short the 6 month version.

```
require(formattable)

t.now <- 10
coeff.now <- as.vector(coeff.50)
(yield.now <- coeff.now[1] + coeff.now[2] * (t.now)/2 + coeff.now[3] *
    (t.now^2)/3 + (t.now < 8) * coeff.now[4] * (t.now - 8)^3/(3 *
    t.now) + (t.now > 25) * coeff.now[5] * (t.now - 25)^3/(3 * t.now))
```

```
## [1] 0.02545169
```

```r
(bond.price.now <- 100 * exp(-t.now * yield.now))
```

```
## [1] 77.52909
```

```r
t.6m <- 9.5
coeff.6m <- as.vector(coeff.75)
(yield.6m <- coeff.6m[1] + coeff.6m[2] * (t.6m)/2 + coeff.6m[3] *
    (t.6m^2)/3 + (t.6m < 8) * coeff.6m[4] * (t.6m - 8)^3/(3 * t.6m) +
    (t.6m > 25) * coeff.6m[5] * (t.6m - 25)^3/(3 * t.6m))
```

```
## [1] 0.02452073
```

```r
(bond.price.6m <- 100 * exp(-t.6m * yield.6m))
```

```
## [1] 79.21956
```

```r
percent(return <- bond.price.6m/bond.price.now - 1, 2)
```

```
## [1] 2.18%
```

```r
percent(a_return <- 2 * return, 2)  ## annualized return
```

```
## [1] 4.36%
```

Observations and Recommendations

Insights

We estimated forward-rate curves from zero coupon bond prices, from STRIPs data available on Wall Street Journal's market data site. After sorting the data and expressing maturity in years, we were able to plot a graph usng the empirical forward-rate estimates. We observed significant volatility for maturity 0-8 years and some volatility for maturity beyond 25 years. We used non-linear regression techniques for fitting models. Models for the forward curve included: quadratic polynomial, cubic polynomial, and quadratic polynomial spline with 2 knots. The p-values for theta values were low in general, except for theta_4. The small p-value of theta_3 is evidence that the spline model fits better than the quadratic polynomial.

We then looked at a scenario where a 10 year maturity zero-coupon bond is bought to satisfy collateral requirements for workers' compensation with a very high likelihood that we will need to exit all business in 6 months. We used the 50%tile of zero-coupon bond prices as the appropriate threshold for estimating returns on the bonds purchased. We used the 75%tile of the current data to estimate the returns should we sell the 10 year maturity zero-coupon bond in 6 months. Our computations resulted in a positive return.

We note how our modeling of public data related to bond price maturities has helped us draw the scenario and answer some critical business questions.

Recommendations

Based on the positive return that is projected for our exit transaction (buy 10 year maturity zero-coupon bonds and sell in 6 months), we recommend the purchase of the 10 year bonds as collateral for workers' compensation.

Sources

Wall Street Journal marketdata, Treasury STRIPs Bond Tutor, discussions with others, and varied sources.