Predicting the Direction of Swap Spreads

Paul Teetor
March, 2007

Abstract

A model is developed for predicting the direction of 10-year swap spreads from related financial time series, such as Treasury bond prices, short-term interest rates, stock indexes, and swap spreads of other maturities. The model is shown to be statistically significant, and is shown to be (theoretically) profitable using an in-sample test.

Table of Contents

Introduction

The goal of this project is to predict the future direction of swaps spreads.

The definition of *swap spread* (or simply *spread*) can get quite technical, but for our immediate purposes, the reader can think of swap spreads as *a measure of risk in the banking system* – not necessarily catastrophic risk such as a total system failure, but rather the day-to-day business risk caused by shifting interest rates and liquidity in the money markets.

There is an active market for swap spreads, letting traders hedge their risk and speculate on the direction of spreads. The market is enormous. The notional trading in swap spreads is on the order of billions of dollars per day, and successful spread traders can generate substantial profits.

Generally, predicting the future of financial prices is quite difficult. A large body of research based on the *Efficient Market Hypothesis* predicates that such predictions are not merely difficult, but actually impossible. See [Brealey 1983] and [Malkiel 1985] for an introduction to that research.

I, too, am normally pessimistic regarding market forecasting, but the case of swap spreads holds some interesting possibilities.

First, the spread is range-bound. It resembles a mean-reverting variable rather than a trending variable.

Second, the spread is well-correlated with other financial data. It seems feasible that we could develop a linear regression of its value.

Third, forecasting the spread's direction would be sufficient. Forecasting the magnitude is more difficult, but not necessarily required for practical applications.

Based on these considerations, I decided to explore the possibility of predicting the *direction* of swap spreads using logistic regression.

Conventions

Unless otherwise stated, all tests are performed at the 95% confidence level ($\alpha = 0.05$).

Most, but not all, statistical studies were conducted using R, a language and environment for statistical computing [Venables 2002]. Since the R system is interactive, I distinguish user commands from system output by showing the user's input in **`bold`**.

Background

What is a “swap spread”?

Technically, a *swap spread* is the difference between two interest rates. For this project, however, I will use a simplified definition based on futures traded on the Chicago Board of Trade (CBOT).

The *5-year swap spread* is calculated as

$$SS_5 = T_5 - S_5$$

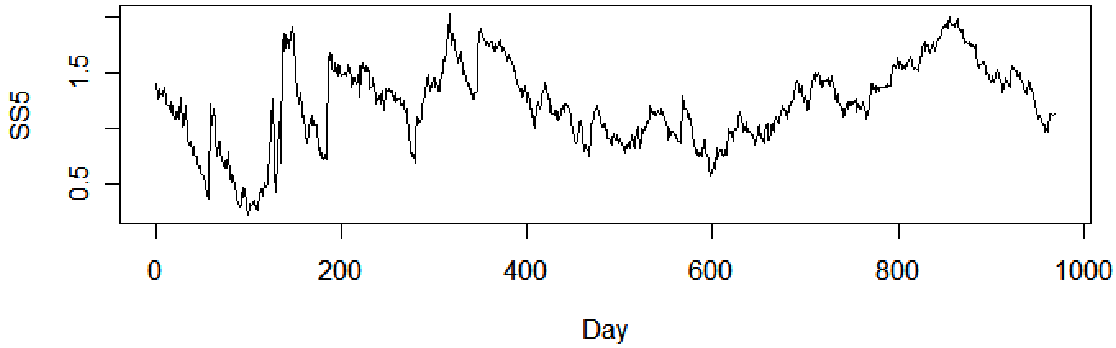
where T_5 is the price of the 5-year Treasury notes futures, and S_5 is the price of the 5-year interest rate swap futures.

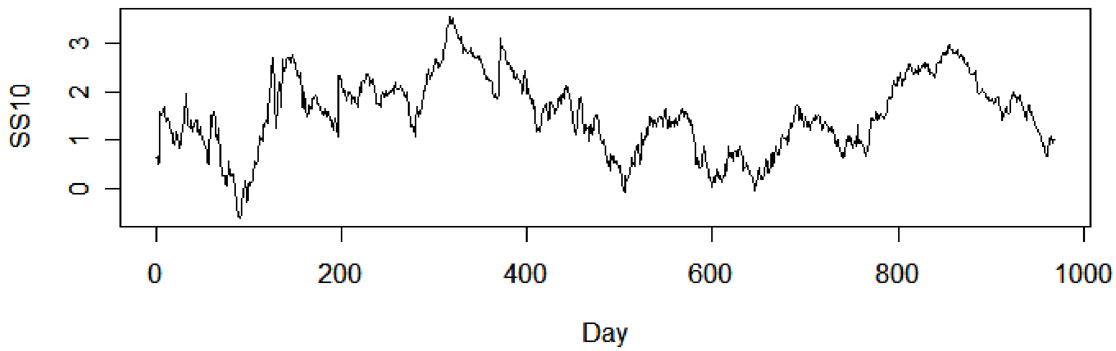
Likewise, the *10-year swap spread* is calculated as

$$SS_{10} = T_{10} - S_{10}$$

where and T_{10} is the price of the 10-year Treasury note futures, and S_{10} is the price of the 10-year interest rate swap futures.

Both swap spreads are always positive since Treasury prices are always larger than swap prices. Furthermore, the spreads oscillate within a relatively narrow range. Here, for example, are graphs of the 5-year and 10-spreads collected for this study. There are about 250 trading days in one calendar year, so these graphs represent about 4 years of data.





For this report, I will focus on the 10-year spreads, and henceforth "the spread" is understood to be the 10-year swap spread. The report logic is applicable to 5-year spread, too, but the constraint of time prevented me from investigating their models.

Influential Factors

Based on discussions with several traders, I knew the following are influential factors in pricing 10-year swap spreads.

- The general level of interest rates, especially rates on Treasury bonds.
- The level of short-term interest rates, such as 90-day money market rates.
- The general level of US stock prices.
- The prices of bank stocks in particular.
- The levels of other swap spreads, such as the 5-year spread.

These factors guided my selection of data for explanatory variables.

Research Hypotheses and Planned Tests

There are two hypotheses in this report.

> Hypothesis #1: We can model the fair value of swap spreads based on a linear regression of influential market data.

If the hypothesis is correct and we can model the spread's fair value, then we can calculate the *market mispricing* as the difference between the modeled value and the actual price. Those mispricings (residuals) are the basis of the second hypothesis.

> Hypothesis #2: Using the mispricing data and simple technical indicators as explanatory variables, we can predict the direction of swap spreads using a logistic regression.

A *technical indicator* is a simple statistic calculated from recent market history. There are many, many technical indicators; see [Colby 1988], for example. A large area of market research called *technical analysis* is devoted to characterizing and predicting market behavior based on those indicators. Some technical analysis is solid, respectable work; some is little more than codified superstition.

I am familiar with the successful work reported in [Harland 2000] which used a class of technical indicators called *momentum indicators* to predict prices of Treasury bond futures. Based on that work, I decided to include simple momentum indicators in the logistic regression to evaluate their effectiveness in predicting the swap spread's direction.

Planned Tests

Hypothesis #1 -- that we can predict fair value -- is not directly testable *per se*. There is no objective standard of "fair value", so we have nothing against which to compare our predictions. We are limited to testing the hypothesis using the usual goodness of fit criteria

We will *indirectly* test the first hypothesis, however, when we build the logistic regression to predict spread direction from the mispricing data: If the fair value model is meaningful, then the mispricings will be meaningful, too, and the logistic regression will identify them as a significant predictor.

Hypothesis #2 – predicting direction from mispricings – presents more opportunities. First, we can test its goodness of fit. Second, we can use its output in a simulated trading process to evaluate its practical significance as a viable market predictor.

Data Collected

All data for this study are time series data. Most of the relevant time series have long histories, but the price history of the two CBOT interest rate swap futures is limited. Those futures began trading in 2001 and 2002, so this study can only span the most recent 5 years of market history.

The final data set contained 968 observations, one for every trading day from 2/4/2003 to 12/14/2006.

Basic Time Series Data

I obtained these data from Commodity Systems Inc. of Boca Raton, FL (http://www.csidata.com).

- TSY5 – Price history for the CBOT 5-year Treasury note futures.
- TSY10 – Price history for the CBOT 10-year Treasury note futures.
- SWP5 – Price history for the CBOT 5-year interest rate swap futures.
- SWP10 – Price history for the CBOT 10-year interest rate swap futures.
- LIBOR – Price history for the London Inter-Bank Offering Rate (LIBOR), a widely used measure of short-term interest rates for 90-day, dollar-denominated deposits in international banks.

I obtained these data from the Yahoo! Finance web site (http://finance.yahoo.com).

- SPX – Price history for the Standard & Poor's 500, a widely diversified index of US stock prices.
- BIX – Price history for the Standard & Poor's' Bank Stock Index, a narrow index of US bank stock prices.

Computed Time Series Data

From the basic time series, I applied our working definition of swap spread to compute these time series.

- SS5 = TSY5 – SWP5
- SS10 = TSY10 – SWP10

Computed Categorical Data: Momentum and Acceleration

The study used *momentum and acceleration indicators* as explanatory variables for the logistic regression. There are categorical variables derived from numerical calculations.

Numerically, momentum is simply the change over a fixed look-back period.

- mom3 = SS10[0] – SS10[3]
- mom5 = SS10[0] – SS10[5]
- mom8 = SS10[0] – SS10[8]
- mom13 = SS10[0] – SS10[13]
- mom21 = SS10[0] – SS10[21]

where SS10[0] is today's swap spread, SS10[3] is the swap spread from 3 days ago, etc.

Numerically, acceleration is the rate of change of momentum; in this case, the 1-day change.

- acc3 = mom3[0] – mom3[1]
- acc5 = mom5[0] – mom5[1]
- acc8 = mom8[0] – mom8[1]
- acc13 = mom13[0] – mom13[1]
- acc21 = mom21[0] – mom21[1]

All momentum and acceleration numerical values were converted to categorical values (“Pos” or “Neg”) according to their sign: Positive values were converted to “Pos”, and negative values were converted to “Neg”.

Computed Categorical Data: Forward Direction

The response variables for the logistic regressions were categorical variables which indicated the future change of the spread, 10 days forward. There were two binary-valued variables, one to indicate a rising spread (buying opportunity), and one to indicate a falling spread (selling opportunity).

The calculation of the response variables proceeded in several steps. First, I formed these two time series.

1. $\Delta SS10_i$ = Change in SS10, 10 days into the future
2. Z_i = normalized $\Delta SS10 = (\Delta SS10_i - \text{mean}(\Delta SS10)) / \text{sd}(\Delta SS10)$.

Then the daily categorical values were calculated as

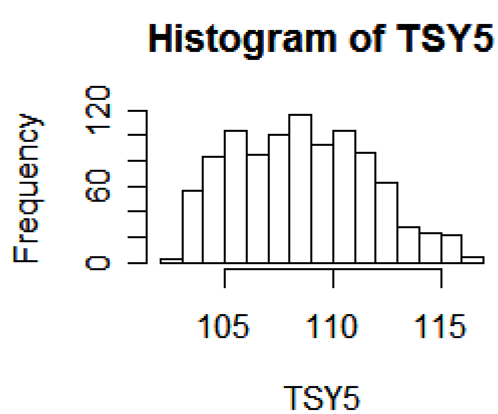
3. fwd10.buy_i = If $Z_i > 0.25$, then *true*; otherwise *false*.
4. fwd10.sell_i = If $Z_i < -0.25$, then *true*; otherwise *false*.

I ignored days with Z values between -0.25 and 0.25, reasoning that such small changes in price were too “noisy” to provide useful information.

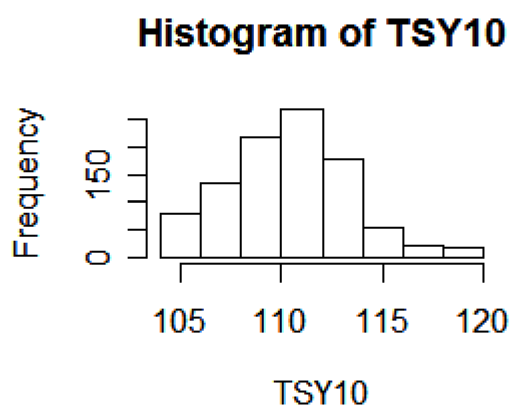
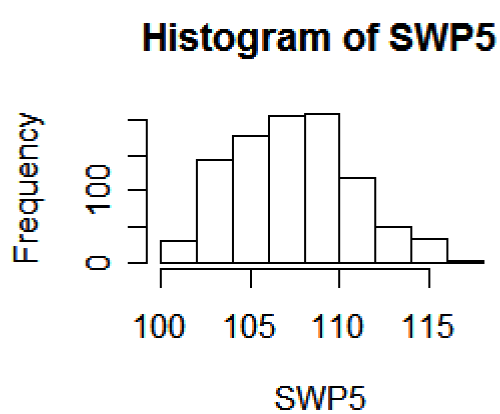
Sample Statistics: Quantitative Variables

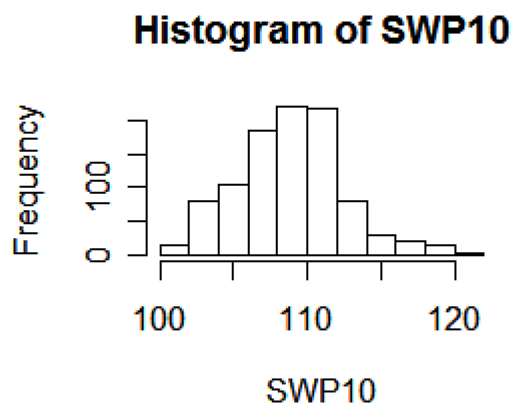
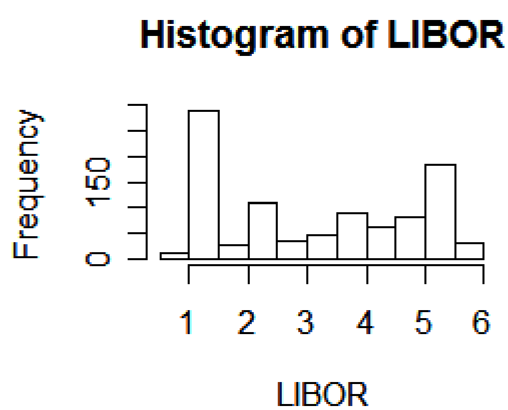
These are the descriptive statistics for the study’s quantitative variables.

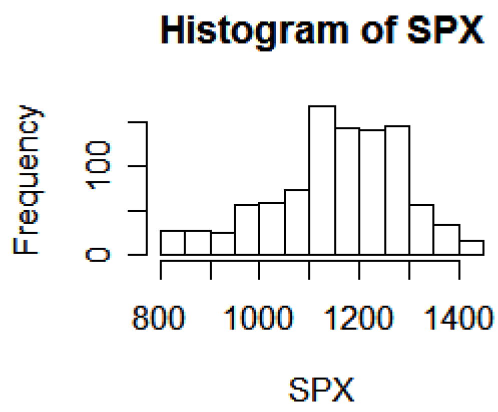
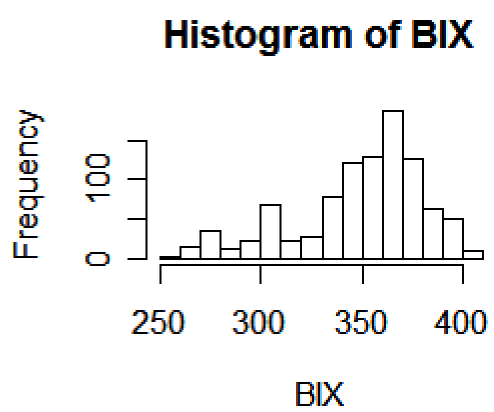
Variable	Summary, Standard Deviation, and Histogram

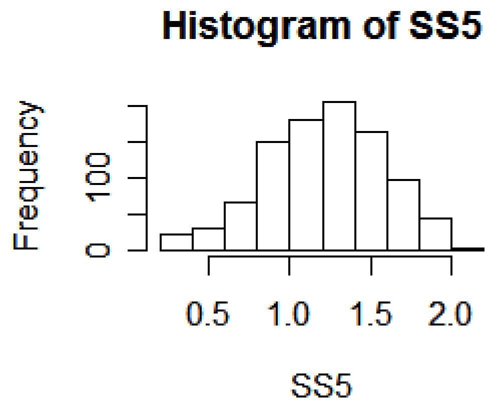
TSY5	`Min. 1st Qu.  Median  Mean 3rd Qu.  Max.`
	`102.9  106.0  108.5  108.6  110.8  116.6`
	
	`SD: 3.088915`
	
	

TSY10	`Min. 1st Qu.  Median  Mean 3rd Qu.  Max.`

	104.0 108.3 110.6 110.4 112.2 119.8 SD: 3.010735 
SWP5	Min. 1st Qu. Median Mean 3rd Qu. Max. 100.9 104.7 107.3 107.4 109.6 116.3 SD: 3.301308 
SWP10	Min. 1st Qu. Median Mean 3rd Qu. Max. 101.2 106.7 109.1 108.9 111.0 120.3 SD: 3.493300

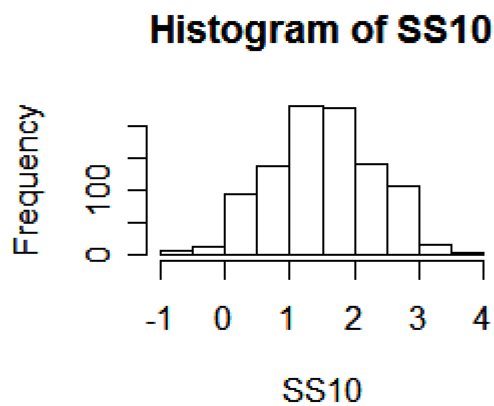
	
LIBOR	Min. 1st Qu. Median Mean 3rd Qu. Max. 0.910 1.325 3.180 3.167 4.811 5.690 SD: 1.633334 
SPX	Min. 1st Qu. Median Mean 3rd Qu. Max. 800.7 1094.0 1176.0 1156.0 1256.0 1425.0 SD: 130.1263

	
BIX	Min. 1st Qu. Median Mean 3rd Qu. Max. 257.6 334.7 355.7 349.4 370.4 406.4 SD: 31.42592 
SS5	Min. 1st Qu. Median Mean 3rd Qu. Max. 0.2188 0.9688 1.2190 1.2190 1.4690 2.0310 SD: 0.365106

<table>
<tr><td></td><td>
</td></tr>
<tr><td>SS10</td><td>

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-0.6094  1.0000  1.5310  1.5430  2.0780  3.5620

 SD: 0.7863546
```


</td></tr>
</table>

This is the correlation matrix for the quantitative variables.

	SS10	TSY10	TSY5	LIBOR	SPX	BIX	SS5
SS10	1.0000	-0.532	-0.273	-0.0264	0.0485	0.119	0.800
TSY10	-0.5323	1.000	0.932	-0.7409	-0.6809	-0.598	-0.700
TSY5	-0.2726	0.932	1.000	-0.9236	-0.8395	-0.744	-0.543
LIBOR	-0.0264	-0.741	-0.924	1.0000	0.8732	0.768	0.318
SPX	0.0485	-0.681	-0.840	0.8732	1.0000	0.944	0.320
BIX	0.1190	-0.598	-0.744	0.7679	0.9437	1.000	0.332
SS5	0.8003	-0.700	-0.543	0.3176	0.3201	0.332	1.000

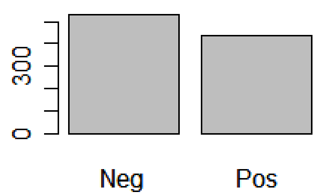
Sample Statistics: Categorical Variables

Many traders believe that the usefulness of a momentum indicator depends upon its look-back period; that a 3-day momentum indicator has a different meaning than a 13-day momentum indicator, for example. This belief was the motivation for including momentum and acceleration indicators with various look-back periods.

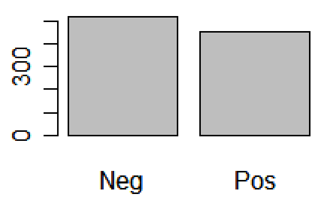
This belief was not born out by statistical testing. For the 2x2 tables formed by pairing different momentum indicators against each other, the resulting χ^2 tests revealed a very strong statistical dependency between every pair of momentum indicators: the *p*-values were essentially zero. Ergo, the different indicators are largely redundant[1].

Because of the large number of momentum and acceleration variables, and because they all have very similar distributions, only two momentum and two acceleration variables are shown here.

***mom5** – 5-day momentum indicator*

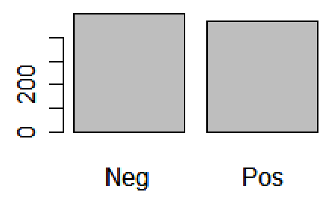


***mom8** – 8-day momentum indicator*

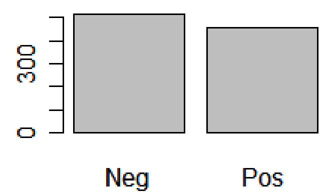


[1] The degree of redundancy between momentum indicators was further evidenced during the logistic regression, discussed later. The forward-selection process showed that one momentum indicator was useful, but an additional indicator added no value.

acc5 *– 5-day acceleration indicator*

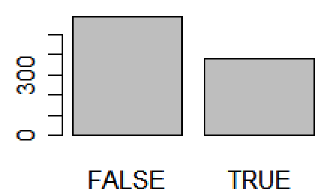


acc8 *– 8-day acceleration indicator*

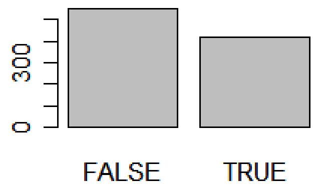


Here are the bar charts for the response variables, both for *buy* and *sell* opportunities.

fwd10.buy *– Response variable for "buy" model*



fwd10.sell *– Response variable for "sell" model*



Analysis

Linear Regression Model

I modeled the spread's fair value as a linear regression on the predictor variables.

$$E(SS10) = \alpha + \beta_1 \cdot TSY10 + \beta_2 \cdot TSY5 + \beta_3 \cdot SS5 + \beta_4 \cdot LIBOR + \beta_5 \cdot BIX + \beta_6 \cdot SPX$$

The regression produced this model.

```
Call:
lm(formula = SS10 ~ TSY10 + TSY5 + SS5 + LIBOR + BIX + SPX, data = export)

Residuals:
     Min       1Q   Median       3Q      Max
-1.30790 -0.16043  0.04775  0.20770  0.83401

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 12.735824   1.875479   6.791 1.95e-11 ***
TSY10       -0.268633   0.020827 -12.898  < 2e-16 ***
TSY5         0.151297   0.033472   4.520 6.95e-06 ***
SS5          1.088649   0.044823  24.288  < 2e-16 ***
LIBOR       -0.200270   0.033151  -6.041 2.18e-09 ***
BIX          0.011778   0.001089  10.813  < 2e-16 ***
SPX         -0.002391   0.000342  -6.993 5.03e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3177 on 961 degrees of freedom
Multiple R-Squared: 0.8377,     Adjusted R-squared: 0.8367
F-statistic: 826.9 on 6 and 961 DF,  p-value: < 2.2e-16
```

The fitted regression equation is

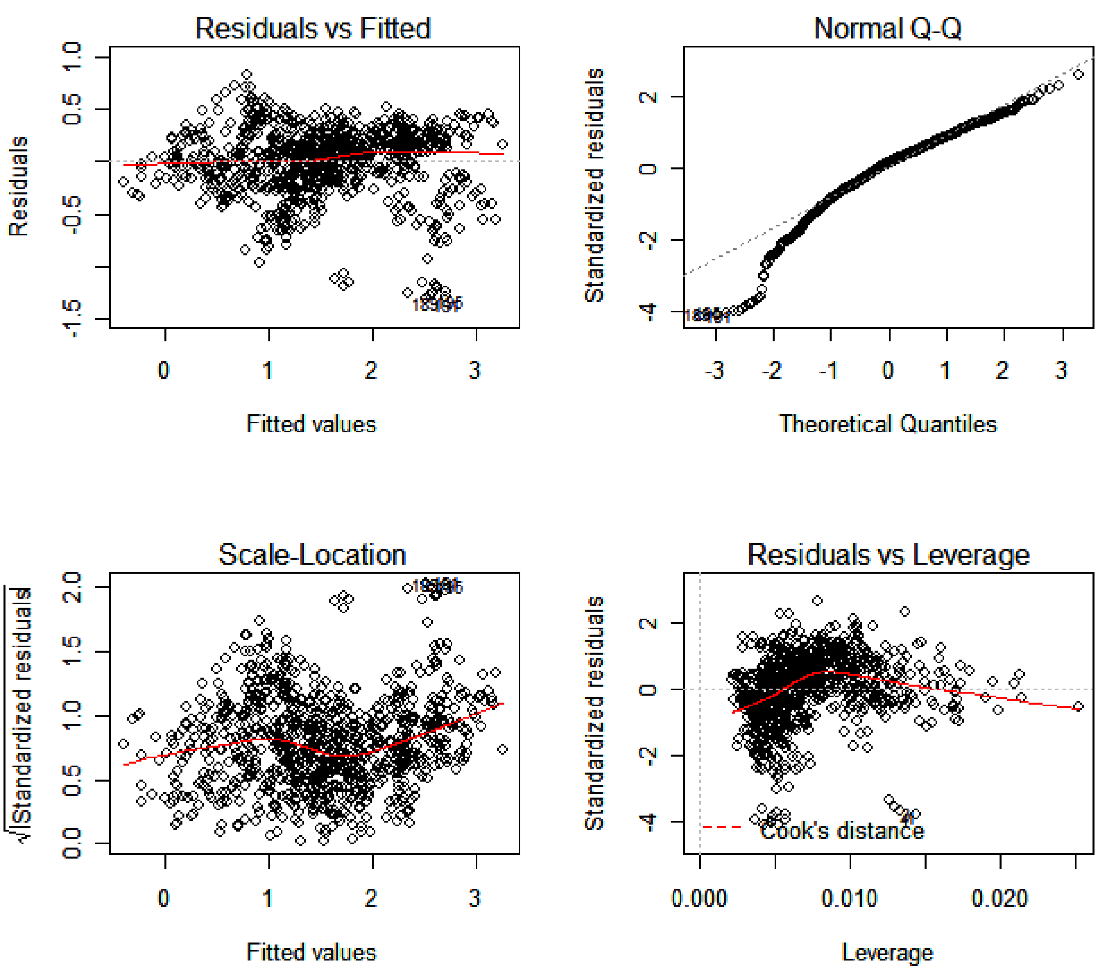
$$E(SS10) = 12.7 - 0.269 \cdot TSY10 + 0.151 \cdot TSY5 + 1.09 \cdot SS5 - 0.2 \cdot LIBOR + 0.0118 \cdot BIX - 0.00239 \cdot SPX$$

Evaluation of the Linear Regression

The model of fair value is statistically significant: All parameters have *p*-values which are essentially zero, indicating that every parameter coefficient is non-zero and, hence, every parameter is significant.

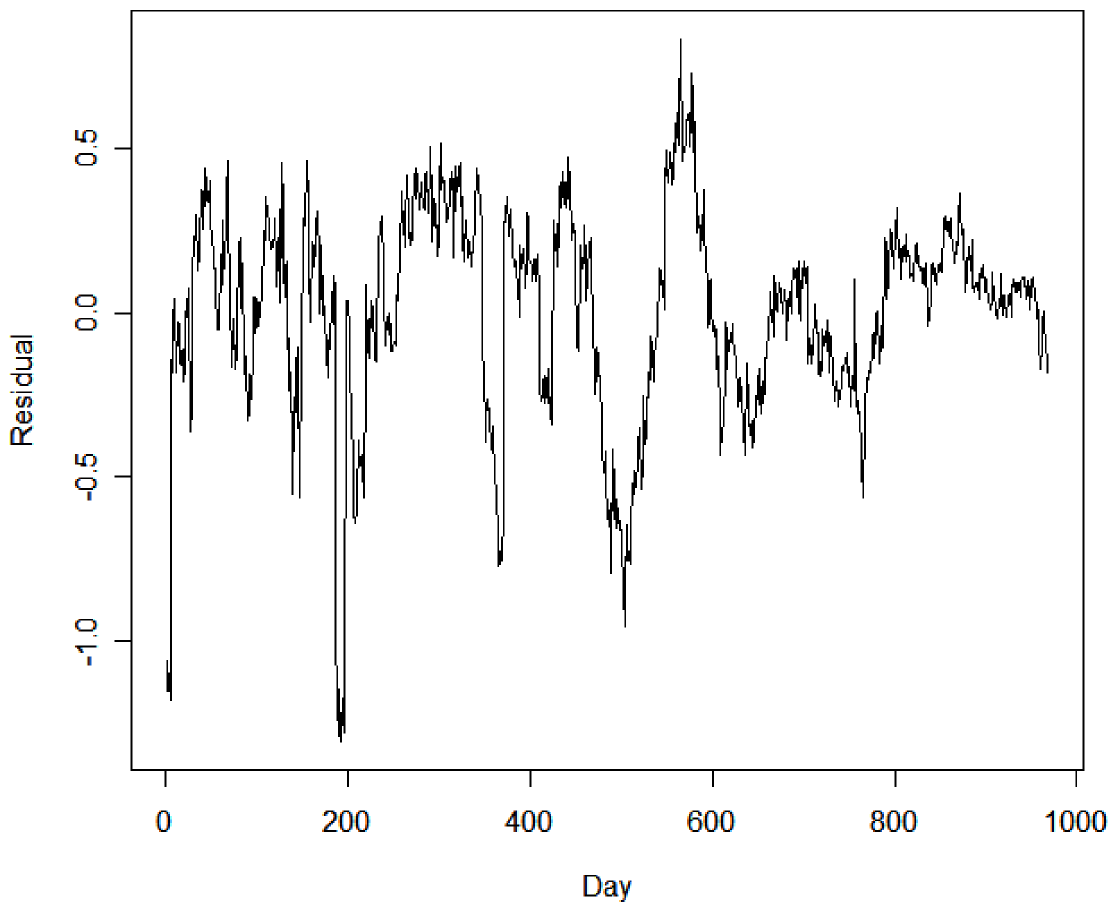
The R^2 is 0.8377, indicating that the model explains over 83% of the variation in swap spreads, which is very good.

The standard diagnostic plots of the residuals do not reveal any serious problems with the fit.



One data point appears to have high leverage, but its residual is nearly zero and, hence, the problem is not significant.

The reader should be aware of something, however. The diagnostic plot of residuals appears benign, but when the residuals are re-plotted as a time series, we get another perspective.



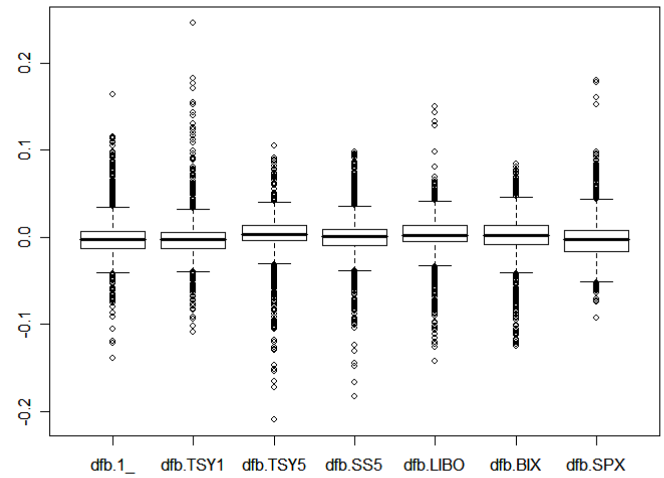
When viewed as a time series, the residuals do not appear random. Rather, they show a serial dependence. Is this a problem? *For our immediate purposes, it is not.*

In pure time series modeling, we might be disappointed that the residuals seem to contain useful information, and we might enhance our model to exploit that information. But here, our intention is the opposite: We will use the residuals as one input to the logistic regression of the spread direction. We want the residuals to contain useful information regarding mispricings.

As a final evaluation of the model, we can check the influence measures. Given the sample size, checking individual observations would be tedious. Instead, we start with a broad check of the DFFIT and DFBETA values, then isolate problematic observations if necessary. The following R commands will calculate the influence measures of the sample, and then produce box plots of the DFBETA values (which in this case are the first 7 columns of the influence measures.)

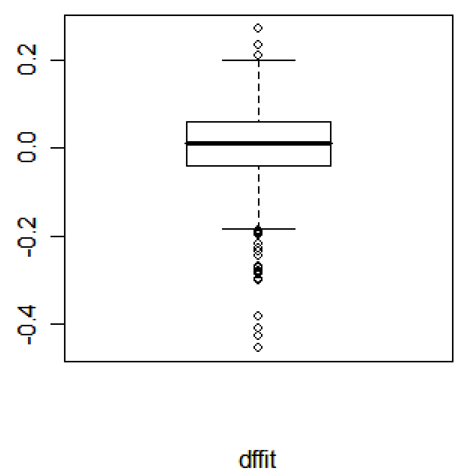
```
> infl = influence.measures(models$ss10.spot.full)
> boxplot(as.data.frame(infl$infmat[,1:7]))
```

This is the resulting plot.



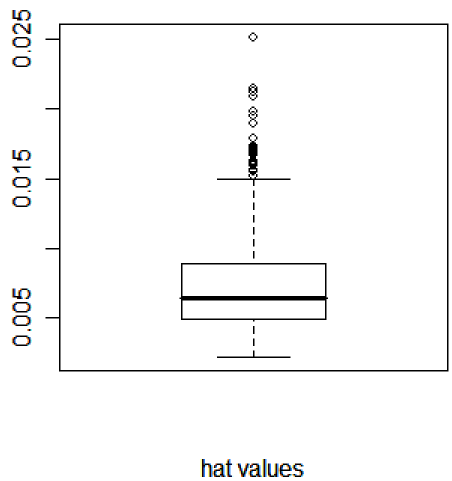
Clearly, none of the DFBETA values exceed our guideline of ±1.0.

Similarly, we can create a box plot of the DFFIT values.



Again we see that no value exceeds our guideline of ±1.0.

The plot of "hat" values does reveal one observation above our guideline of 3n/p = 3×7/968 = 0.02169.



That observation, however, is the extreme leverage observation identified in the *Residuals vs. Leverage* plot, above. As noted, that observation is not a problem due to its low residual.

Reducing the Linear Regression Model

By "piling on" all the explanatory variables into the initial regression, we may have included extraneous variables which provide no explanatory power beyond a reduced model. We can test this conjecture using the `drop1` function of R, which performs sequential deletions of single terms from the model, performing a χ^2 calculation to test the hypothesis that the full model and the reduced model have equivalent explanatory power. (The `drop1` function is essentially one step of a backward selection procedure.)

```
> drop1(models$ss10.spot.full, test="Chisq")
Single term deletions

Model:
SS10 ~ TSY10 + TSY5 + SS5 + LIBOR + BIX + SPX
       Df Sum of Sq      RSS      AIC   Pr(Chi)
<none>                 97.02 -2212.69
TSY10   1     16.80   113.82 -2060.14 < 2.2e-16 ***
TSY5    1      2.06    99.08 -2194.33 6.401e-06 ***
SS5     1     59.55   156.58 -1751.40 < 2.2e-16 ***
LIBOR   1      3.68   100.71 -2178.61 1.893e-09 ***
BIX     1     11.81   108.83 -2103.54 < 2.2e-16 ***
SPX     1      4.94   101.96 -2166.64 4.162e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From this output, we see that all *p*-value of every model was essentially zero, and we reject the hypothesis that any reduced model had equivalent explanatory power. We conclude that the full model cannot be reduced without damaging its value.

Logistic Regression Model: Minimal Model

I performed two logistic regressions, one to predict the probability of a rising market (the "buy" model), and the other to predict the probability of a falling market (the "sell" model). I started with these minimal models

$$\log(\text{odds}(\text{Up})) = \alpha_{up} + \text{beta}_{up} \times \text{RESID}$$

$$\log(\text{odds}(\text{Dn})) = \alpha_{dn} + \text{beta}_{dn} \times \text{RESID}$$

where *RESID* is the time series of residuals from the fair value model (above). The logistic regression produced these fits of the minimal models.

Buy model:

```
Call:
glm(formula = object$ss10.fwd10.buy ~ resids$ss10.spot.full,
    family = binomial(), data = export)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.5594  -0.9710  -0.8479   1.2875   1.7530

Coefficients:
                      Estimate Std. Error z value Pr(>|z|)
(Intercept)           -0.44550    0.06736  -6.613 3.76e-11 ***
resids$ss10.spot.full -1.39703    0.22187  -6.297 3.04e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1297.8  on 967  degrees of freedom
Residual deviance: 1254.8  on 966  degrees of freedom
AIC: 1258.8

Number of Fisher Scoring iterations: 4
```

Sell model:

```
Call:
glm(formula = object$ss10.fwd10.sell ~ resids$ss10.spot.full,
    family = binomial(), data = export)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.4882  -1.0801  -0.8337   1.2132   1.8840

Coefficients:
                      Estimate Std. Error z value Pr(>|z|)
(Intercept)           -0.28378    0.06648  -4.268 1.97e-05 ***
resids$ss10.spot.full  1.39232    0.23198   6.002 1.95e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)
```

```
    Null deviance: 1325.0  on 967  degrees of freedom
Residual deviance: 1284.8  on 966  degrees of freedom
AIC: 1288.8

Number of Fisher Scoring iterations: 4
```

The fitted equations are

$$\log(\text{odds}(\text{Up})) = -0.446 - 1.40 \times \text{RESID} \qquad (\textit{Buy model})$$

$$\log(\text{odds}(\text{Up})) = -0.284 + 1.39 \times \text{RESID} \qquad (\textit{Sell model})$$

We want to know if the RESID variable is significant; that is, do the residuals from the linear regression really contain information useful to the logistic regression?

For the *buy* model, the log-likelihood test for the significance of β has a *p*-value of $\chi^2(42.95, 1)$, which is essentially zero. For the *sell* model, the *p*-value is $\chi^2(40.19, 1)$, which is also essentially zero. So for the both models, yes, the RESID data is significant.

Logistic Regression Model: Forward Selection

As stated, the initial logistic regression was a minimal model. We also want to incorporate a simple momentum indicator and determine if that improves the model.

The R system includes the `add1` function which performs one step of the forward-selection process, letting us test new explanatory variables for inclusion in the model. We can use `add1` to test the inclusion of momentum indicators in both models, *buy* and *sell*.

Buy model:

```
Single term additions

Model:
object$ss10.fwd10.buy ~ resids$ss10.spot.full
           Df Deviance     AIC     LRT  Pr(Chi)
<none>         1254.81 1258.81
SS10.mom3   1  1250.83 1256.83    3.98 0.046041 *
SS10.mom5   1  1244.61 1250.61   10.20 0.001406 **
SS10.mom8   1  1247.37 1253.37    7.44 0.006378 **
SS10.mom13  1  1253.40 1259.40    1.42 0.234008
SS10.mom21  1  1253.29 1259.29    1.52 0.217429
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Sell model:

```
Single term additions

Model:
object$ss10.fwd10.sell ~ resids$ss10.spot.full
           Df Deviance     AIC     LRT  Pr(Chi)
<none>         1284.77 1288.77
```

```
SS10.mom3   1  1280.35 1286.35    4.42 0.035590 *
SS10.mom5   1  1276.86 1282.86    7.91 0.004916 **
SS10.mom8   1  1275.13 1281.13    9.63 0.001911 **
SS10.mom13  1  1281.36 1287.36    3.41 0.064818 .
SS10.mom21  1  1284.54 1290.54    0.23 0.631080
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

For both models, the *mom5* and *mom8* indicators were significant additions, as indicated by the low *p*-values to test the equivalence of the minimal models *versus* the expanded models. It appears to me that the *mom5* indicator has a slight advantage, so I incorporated it into the logistic regressions.

After incorporating *mom5*, however, further expansion of the model yielded no improvements. Here is the `add1` output for inclusion of an additional indicator.

Buy model:

```
Single term additions

Model:
object$ss10.fwd10.buy ~ (resids$ss10.spot.full + SS10.mom5)
           Df Deviance     AIC     LRT Pr(Chi)
<none>         1244.61 1250.61
SS10.mom3   1  1244.59 1252.59    0.02  0.8880
SS10.mom8   1  1243.56 1251.56    1.05  0.3051
SS10.mom13  1  1244.58 1252.58    0.03  0.8545
SS10.mom21  1  1244.52 1252.52    0.09  0.7625
```

Sell model:

```
Single term additions

Model:
object$ss10.fwd10.sell ~ (resids$ss10.spot.full + SS10.mom5)
           Df Deviance     AIC     LRT Pr(Chi)
<none>         1276.86 1282.86
SS10.mom3   1  1276.55 1284.55    0.30 0.58200
SS10.mom8   1  1273.73 1281.73    3.13 0.07705 .
SS10.mom13  1  1276.34 1284.34    0.52 0.47052
SS10.mom21  1  1276.71 1284.71    0.15 0.70169
```

The *p*-values are for the comparison of the *mom5*-only model against the model with both *mom5* and another momentum indicator. They are uniformly large; in fact, all but one are remarkably large. We conclude that inclusion of an addition momentum indicator – *any* additional momentum indicator – is unproductive.

A similar check of the acceleration indicators, described above, revealed that they contained no useable explanatory power. Not a single acceleration indicator improved the model quality, according to the χ^2 test of the `add1` function.

The final regression models are

Buy model:

```
Call:
glm(formula = object$ss10.fwd10.buy ~ (resids$ss10.spot.full +
    SS10.mom5), family = binomial(), data = export)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.7062  -0.9950  -0.8071   1.2691   1.8057

Coefficients:
                      Estimate Std. Error z value Pr(>|z|)
(Intercept)           -0.64974    0.09443  -6.881 5.95e-12 ***
resids$ss10.spot.full -1.49669    0.22531  -6.643 3.08e-11 ***
SS10.mom5Pos           0.43651    0.13722   3.181  0.00147 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Equation: log(odds(Up)) = -0.640 – 1.50×RESID + 0.437×MOM5

Sell model:

```
Call:
glm(formula = object$ss10.fwd10.sell ~ (resids$ss10.spot.full +
    SS10.mom5), family = binomial(), data = export)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.4552  -1.0714  -0.8269   1.1885   2.0046

Coefficients:
                      Estimate Std. Error z value Pr(>|z|)
(Intercept)           -0.11389    0.08963  -1.271  0.20387
resids$ss10.spot.full  1.46583    0.23360   6.275  3.5e-10 ***
SS10.mom5Pos          -0.37728    0.13469  -2.801  0.00509 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Equation: log(odds(Down)) = -0.114 – 1.47×RESID - 0.377×MOM5

Simulated Trading

So far, this entire study has been "just statistics". The models pass their statistical tests, but are they useful?

A basic test of usefulness is the application of the models to simulate trading using the in-sample data. We would prefer out-of-sample data, of course. But if the models fail on in-sample data, there is no hope, so we can start there.

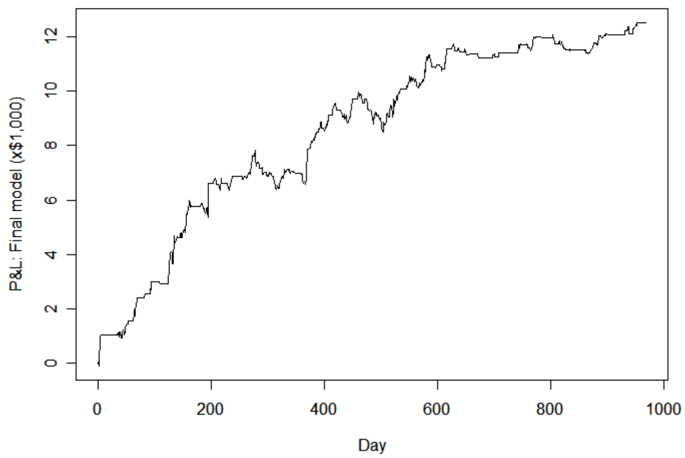
To use the logistic regressions for trading, we follow these rules, using the probabilities implied from the calculated log odds for each day.

Buy model: If prob(Up) > 0.5, buy today' market.

Sell model: If prob(Down) > 0.5, sell today's market.

Together, these rules produced a gross profit of $12,529.70 over the test period, trading one spread. That is approximately $3,200 *per annum*. (These are gross figures because they do not include transaction costs such as broker's fees.)

This is the graph of the cumulative profit and loss from the simulation.



Evaluation of Simulated Trading

To evaluate the simulated results, I started with the day-by-day output of the simulation, which is a time series of tuples, one tuple for each simulated day.

<*Decision*, *Outcome*>

Outcome is the market's subsequent direction, either "Up" or "Dn", after the day. The data set contained 451 observations. (Some days have no associated decision, either buy or sell, and so generated no tuple.)

From this data I formed the 2x2 contingency table expressing the relationship between decisions and outcomes.

Frequency Expected Cell Chi-Square Percent Row Pct Col Pct

Table of Signal by Outcome

Signal	Outcome		Total
	"Dn"	"Up"	
"Buy"	71 94.958 6.0446 15.74 44.10 26.69	90 66.042 8.6911 19.96 55.90 48.65	161 35.70
"Sell"	195 171.04 3.3558 43.24 67.24 73.31	95 118.96 4.8251 21.06 32.76 51.35	290 64.30
Total	266 58.98	185 41.02	451 100.00

Fortunately, the categorical variables, *Decision* and *Outcome*, are ordinal, not merely nominal, and we can apply the Cochran-Armitage trend test. A SAS script yielded these test results (see Appendix for SAS code).

Cochran-Armitage Trend Test	
Statistic (Z)	4.7871
One-sided Pr > Z	<.0001
Two-sided Pr > |Z|	<.0001

The *p*-value is essentially zero, and we can reject the hypothesis that *Outcome* is independent of *Decision*. The decisions are statistically significant.

Summary and Conclusions

- Using a linear regression, we can create a model for fair value of a swap spreads predicted from Treasury bond prices, short-term interest rates, the general level of stocks prices, an index of bank stock prices, and other swap spreads.

- Using a logistic regression, we can exploit the residuals from the linear regression to predict the direction of swap spreads.
- We can improve the logistic regression by incorporating a momentum indicator as an explanatory variable. One indicator is sufficient, however; additional indicators do not enhance the model's predictive power.
- Simulated trading using in-sample data verified the usefulness of the logistic regression: it did produce usable trading decisions, and the decisions were statistically significant.

Limitations

This study has several limitation, including:

- The study uses less than 5 years of data. I would prefer 10 years or more, but the CBOT swaps futures have been trading for only slightly longer than 5 years.
- I used a non-standard definition of "swap spread". While it is a practical alternative to the common definition, professional traders might not recognize it as equivalent.
- The models are relatively simple. In particular, the logistic model contains only two explanatory variables, and no interaction terms. Some preliminary exploration revealed that interaction terms might be useful.
- I explored only two classes of technical indicators for inclusion in the logistic regression. There are many other candidate indicators.
- The simulated trading did not include frictional costs, such as broker's fees and bid/ask spreads.
- The simulated trading used in-sample data.

Recommendations

To pursue this study further, I recommend these steps.

- Perform the simulated trading on out-of-sample data, verifying that the models have abstracted useful information from the original sample.
- Estimate the impact of frictional costs on trading.
- Explore the inclusion of additional or alternative technical indicators in the logistic regression.
- Explore the effective of interaction terms in the logistic regression.
- Some of my explorations into correlations indicated that controlling for short-term interest rates (LIBOR) revealed statistical interactions between other variables. This could be a very fruitful area for research.

References

[Brealey 1983] *An Introduction to Risk and Return from Common Stocks (2nd ed.)*, Richard A. Brealey, MIT Press, 1983 (ISBN 0-262-52116-4)

[Colby 1988] *The Encyclopedia of Technical Market Indicators*, Robert W. Colby and Thomas A. Meyers. Business One Irwin, 1988 (ISBN 1-556-23049-4)

[Harland 2000] "Using Nonlinear Neurogenetic Models with Profit Related Objective Functions to Trade the US T-bond Future," by Zac Harland. Reprinted in *Computational Finance 1999*, Abu-Mostafa, et al (ed.). MIT Press, 2000

[Malkiel 1985] *A Random Walk Down Wall Street (4th ed.)*, Burton G. Malkiel. W.W. Norton and Co., 1985 (ISBN 0-393-95460-9)

[Venables 2002] *An Introduction to R*, W. N. Venables, D. M. Smith, et al. Network Theory Limited. Also available at http://cran.r-project.org/doc/manuals/R-intro.html (ISBN 0-954-16174-2)

Appendix: Catalog of Variables

This is a catalog of the variables which I collected prior to beginning my statistical exploration. Note that not all variables were actually used in the final model.

Raw Data

Price, rate, and index data is all quantitative. These are the explanatory variables for the multivariate linear regression models.

- Date – Date of observation
- SS10 – Swap spread price, 10-year maturity
- SS5 – Swap spread price, 5-year maturity
- TSY10 – Treasury bond price, 10-year maturity
- TSY5 – Treasury bond price, 5-year maturity
- SWP10 – Swap price, 10-year maturity
- SWP5 – Swap price, 5-year maturity
- LIBOR – Short-term interest rate
- BIX – Index of bank stock prices
- SPX – S&P 500 index of stock prices

Technical details, of interest only to dedicated finance geeks:

1. The time series derived from futures prices (TSY5, TSY10, SWP5, SWP10) were constructed by Commodity Systems Inc. (CSI) using their proprietary *Perpetual Contract* algorithm.
2. I computed the LIBOR time series from prices of the near-by CME Eurodollar futures contract as $L = 100 - ED$, where ED is the futures price. Strictly speaking, then, this is a forward rate, not the spot rate. (Using the CME 1-month Libor contract would have yielded a closer approximation to the spot rate.) Using spot rates is problematic, however, due to timing issues between the time at which LIBOR is fixed in London, and the time at which the US futures market closes. Using the Eurodollar-implied rate provides a timelier estimate of the short-term rate.

Derived Values: Technical Indicators

Momentum and acceleration indicators are binary (positive/negative). These are explanatory variables for the logistic models.

- SS10.mom3 – SS10, 3-day momentum indicator
- SS10.mom5 – SS10, 5-day momentum indicator
- SS10.mom8 – SS10, 8-day momentum indicator
- SS10.mom13 – SS10, 13-day momentum indicator
- SS10.mom21 – SS10, 21-day momentum indicator
- SS5.mom3 – SS5, 3-day momentum indicator
- SS5.mom5 – SS5, 5-day momentum indicator

- SS5.mom8 – SS5, 8-day momentum indicator
- SS5.mom13 – SS5, 13-day momentum indicator
- SS5.mom21 – SS5, 21-day momentum indicator
- SS10.acc3 – SS10, 3-day acceleration indicator
- SS10.acc5 – SS10, 5-day acceleration indicator
- SS10.acc8 – SS10, 8-day acceleration indicator
- SS10.acc13 – SS10, 13-day acceleration indicator
- SS10.acc21 – SS10, 21-day acceleration indicator
- SS5.acc3 – SS5, 3-day acceleration indicator
- SS5.acc5 – SS5, 5-day acceleration indicator
- SS5.acc8 – SS5, 8-day acceleration indicator
- SS5.acc13 – SS5, 13-day acceleration indicator
- SS5.acc21 – SS5, 21-day acceleration indicator

Derived Values: Forward Directions

These are binary valued: "Up" or "Down". They are the response values for the logistic regression. Our goal is to predict these values.

- SS10.fwd1 – SS10, one day forward
- SS10.fwd5 – SS10, 5 days forward
- SS10.fwd10 – SS10, 10 days forward
- SS5.fwd1 – SS5, one day forward
- SS5.fwd5 – SS5, 5 days foward
- SS5.fwd10 – SS5, 10 days forward
- SWP10.fwd5 – SWP10, 5 days forward
- SWP10.fwd10 – SWP10, 10 days forward
- SWP5.fwd5 – SWP5, 5 days forward
- SWP5.fwd10 – SWP5, 10 days forward

Appendix: R Code

Although the body of this report, above, shows several R commands used to perform statistical calculations, three R scripts were used first to load the data into R, create the models, and perform some model evaluation.

load.export.R

The raw data was held in a CSV file called `export.csv`. This script loaded the data, converted some columns into factors (R's term for a categorical variables), and pre-computed some additional columns used during the regressions and the evaluation.

```
#
#  Load export.csv into R
#
```

```
export = read.csv("export.csv", header=T)

# Convert momentum and acceleration numbers into factors
#
export$SS10.mom3 = as.factor(ifelse(export$SS10.mom3 > 0, "Pos", "Neg"))
export$SS10.mom5 = as.factor(ifelse(export$SS10.mom5 > 0, "Pos", "Neg"))
export$SS10.mom8 = as.factor(ifelse(export$SS10.mom8 > 0, "Pos", "Neg"))
export$SS10.mom13 = as.factor(ifelse(export$SS10.mom13 > 0, "Pos", "Neg"))
export$SS10.mom21 = as.factor(ifelse(export$SS10.mom21 > 0, "Pos", "Neg"))

export$SS5.mom3 = as.factor(ifelse(export$SS5.mom3 > 0, "Pos", "Neg"))
export$SS5.mom5 = as.factor(ifelse(export$SS5.mom5 > 0, "Pos", "Neg"))
export$SS5.mom8 = as.factor(ifelse(export$SS5.mom8 > 0, "Pos", "Neg"))
export$SS5.mom13 = as.factor(ifelse(export$SS5.mom13 > 0, "Pos", "Neg"))
export$SS5.mom21 = as.factor(ifelse(export$SS5.mom21 > 0, "Pos", "Neg"))

export$SS10.acc3 = as.factor(ifelse(export$SS10.acc3 > 0, "Pos", "Neg"))
export$SS10.acc5 = as.factor(ifelse(export$SS10.acc5 > 0, "Pos", "Neg"))
export$SS10.acc8 = as.factor(ifelse(export$SS10.acc8 > 0, "Pos", "Neg"))
export$SS10.acc13 = as.factor(ifelse(export$SS10.acc13 > 0, "Pos", "Neg"))
export$SS10.acc21 = as.factor(ifelse(export$SS10.acc21 > 0, "Pos", "Neg"))

export$SS5.acc3 = as.factor(ifelse(export$SS5.acc3 > 0, "Pos", "Neg"))
export$SS5.acc5 = as.factor(ifelse(export$SS5.acc5 > 0, "Pos", "Neg"))
export$SS5.acc8 = as.factor(ifelse(export$SS5.acc8 > 0, "Pos", "Neg"))
export$SS5.acc13 = as.factor(ifelse(export$SS5.acc13 > 0, "Pos", "Neg"))
export$SS5.acc21 = as.factor(ifelse(export$SS5.acc21 > 0, "Pos", "Neg"))

# Calculate forward deltas
#
export$SS10.delta10 = export$SS10.fwd10 - export$SS10
export$SS10.delta5  = export$SS10.fwd5  - export$SS10
export$SS10.delta1  = export$SS10.fwd1  - export$SS10

export$SS5.delta10 = export$SS5.fwd10 - export$SS5
export$SS5.delta5  = export$SS5.fwd5  - export$SS5
export$SS5.delta1  = export$SS5.fwd1  - export$SS5

export$SWP10.delta10 = export$SWP10.fwd10 - export$SWP10
export$SWP10.delta5  = export$SWP10.fwd5  - export$SWP10

export$SWP5.delta10 = export$SWP5.fwd10 - export$SWP5
export$SWP5.delta5  = export$SWP5.fwd5  - export$SWP5

# Define objective functions
#
MIN_Z = 0.25

z.score <- function(x) (x - mean(x)) / sd(x, na.rm=T)

object = data.frame(
   ss10.fwd10.buy   = z.score(export$SS10.delta10)  >   MIN_Z,
   ss10.fwd10.sell  = z.score(export$SS10.delta10)  < -(MIN_Z),
   ss5.fwd10.buy    = z.score(export$SS5.delta10)   >   MIN_Z,
   ss5.fwd10.sell   = z.score(export$SS5.delta10)   < -(MIN_Z),
   swp10.fwd10.buy  = z.score(export$SWP10.delta10) >   MIN_Z,
   swp10.fwd10.sell = z.score(export$SWP10.delta10) < -(MIN_Z),
   swp5.fwd10.buy   = z.score(export$SWP5.delta10)  >   MIN_Z,
   swp5.fwd10.sell  = z.score(export$SWP5.delta10)  < -(MIN_Z)
```

```
    )
```

make.models.R

This script performed both the linear regression and the logistic regression, storing the resulting models in a list called *models*.

```
#
#  Create the models for the spot data
#
models = list()

# Full models for swap spreads
#
models$ss10.spot.full = lm(SS10 ~ TSY10 + TSY5 + SS5 + LIBOR + BIX + SPX,
export)

#
#  Calculate residuals for the spot model
#
resids = list()

resids$ss10.spot.full = residuals(models$ss10.spot.full)

#
#  Create models for the forward trading objectives
#

# Minimal models for swap spreads
#
models$ss10.fwd10.buy.min  = glm(object$ss10.fwd10.buy ~
resids$ss10.spot.full,
                                    family=binomial(),
                                    data=export)
models$ss10.fwd10.sell.min = glm(object$ss10.fwd10.sell ~
resids$ss10.spot.full,
                                    family=binomial(),
                                    data=export)

# Final models
#
models$ss10.fwd10.buy.red = glm(object$ss10.fwd10.buy
                              ~ (resids$ss10.spot.full + SS10.mom5),
                              family=binomial(), data=export)

models$ss10.fwd10.sell.red = glm(object$ss10.fwd10.sell
                              ~ (resids$ss10.spot.full + SS10.mom5),
                              family=binomial(), data=export)
```

apply.models.R

This script performs some model evaluation by applying the logistic regressions to calculate log odds, converting log odds into probabilities and then into buy/sell signals, simulating the effects of those signals, calculating the resultant profit and loss (P&L), and plotting the P&L.

```
#
#  Apply trading models to data, giving P&L
#

# Several experiments showed that 0.5 here is a good
# balance of profit, drawdown, and sharpe ratio
#
PROB_THRESH = 0.5

gen.signal <- function(model) {
   pred = predict(model)
   odds = exp(pred)
   prob = odds / (1 + odds)
   sig = prob > PROB_THRESH

   return(sig)
}

calc.pl <- function(model, fwd) {
   pl = ifelse(gen.signal(model), fwd, 0)

   return(pl)
}

# Generate trade signals, buy and sell, for all models
#
signal = list()

signal$ss10.fwd10.buy.min  = gen.signal(models$ss10.fwd10.buy.min)
signal$ss10.fwd10.sell.min = gen.signal(models$ss10.fwd10.sell.min)

signal$ss10.fwd10.buy.red  = gen.signal(models$ss10.fwd10.buy.red)
signal$ss10.fwd10.sell.red = gen.signal(models$ss10.fwd10.sell.red)

# Generate actual outcomes
#
outcomes = list()

outcomes$ss10 = ifelse(export$SS10.delta5 > 0, "Up", "Dn")

# Generate 2x2 contingency tables for all models
#
tables = list()

make.table <- function(sig) {
   tbl = table(signal  = sig,
               outcome = outcomes$ss10,
               exclude = NA)
   return(tbl)
}

tables$ss10.fwd10.buy.min  = make.table(signal$ss10.fwd10.buy.min)
tables$ss10.fwd10.sell.min = make.table(signal$ss10.fwd10.sell.min)

tables$ss10.fwd10.buy.red  = make.table(signal$ss10.fwd10.buy.red)
tables$ss10.fwd10.sell.red = make.table(signal$ss10.fwd10.sell.red)
```

```
rm(make.table)

# Functions to calculate odds ratios
#
odds.ratio <- function(tbl) {
   return ((tbl[1,1] * tbl[2,2]) / (tbl[1,2] * tbl[2,1]))
}

odds.ratio.ase <- function(tbl) sqrt(1/tbl[1,1] + 1/tbl[1,2] + 1/tbl[2,1] +
1/tbl[2,2])

odds.ratio.ci <- function(tbl) {
   logOR = log(odds.ratio(tbl))
   ase = odds.ratio.ase(tbl)
   return (c(exp(logOR - 1.96*ase), exp(logOR + 1.96*ase)))
}

relative.risk <- function(tbl) {
   return (tbl[1,1] / (tbl[1,1] + tbl[1,2])) / (tbl[2,1] / (tbl[2,1] +
tbl[2,2]))
}

# Calculate P&L, buy and sell and combined, for all models
#
pl = list()

pl$ss10.fwd10.buy.min  = calc.pl(models$ss10.fwd10.buy.min,
export$SS10.delta1)
pl$ss10.fwd10.sell.min = calc.pl(models$ss10.fwd10.sell.min, -
(export$SS10.delta1))
pl$ss10.fwd10.min      = pl$ss10.fwd10.buy.min + pl$ss10.fwd10.sell.min

pl$ss10.fwd10.buy.red  = calc.pl(models$ss10.fwd10.buy.red,
export$SS10.delta1)
pl$ss10.fwd10.sell.red = calc.pl(models$ss10.fwd10.sell.red, -
(export$SS10.delta1))
pl$ss10.fwd10.red      = pl$ss10.fwd10.buy.red + pl$ss10.fwd10.sell.red

rm(calc.pl)

# Plot the cumulative P&L
#
par(mfrow=c(1,1))
plot(cumsum(pl$ss10.fwd10.red), typ='l', ylab='P&L: Final model (x$1,000)',
xlab='Day')

summary.pl <- function(v, buyTbl, sellTbl) {
   cat("\tP&L:",
          sum(v),
          "total =",
          sum(v)/(length(v)/250), "p.a.",
          "\n")
   cat("\tAnn Sharpe ratio:",
          250 * mean(v) / (sqrt(250) * sd(v)),
          "\n");
   cat("\tMax drawdown:", max(cummax(cumsum(v)) - cumsum(v)), "\n");
   cat("\tOdds ratios:", odds.ratio(buyTbl), "(buy)\n");
   cat("\tOdds ratio CIs:", odds.ratio.ci(buyTbl), "(buy)\n")
}
```

```
cat("Minimal model P&L:\n");
summary.pl(pl$ss10.fwd10.min, tables$ss10.fwd10.buy.min,
tables$ss10.fwd10.sell.min);

cat("Final model:\n");
summary.pl(pl$ss10.fwd10.red, tables$ss10.fwd10.buy.red);

rm(summary.pl)
```

Appendix: SAS Code and Output

This SAS script was used to test the 2x2 contingency table produced by simulated trading.

```
/*
 * MAT 442 Final Project: Signal analysis
 */


data trades;
   infile "trades.csv" dlm=",";
   input Signal $ Outcome $;
run;

ods html;

proc freq data=trades;
   title "Trade Analysis";
   tables Signal * Outcome / chisq expected cellchi2 trend measures ci;
run;

ods html close;
```

SAS output:

Trade Analysis

The FREQ Procedure

Frequency Expected Cell Chi-Square Percent Row Pct Col Pct

Table of Signal by Outcome			
Signal	**Outcome**		**Total**
	"Dn"	**"Up"**	
"Buy"	71	90	161

	94.958 6.0446 15.74 44.10 26.69	66.042 8.6911 19.96 55.90 48.65	35.70
"Sell"	195 171.04 3.3558 43.24 67.24 73.31	95 118.96 4.8251 21.06 32.76 51.35	290 64.30
Total	266 58.98	185 41.02	451 100.00

Statistics for Table of Signal by Outcome

Statistic	DF	Value	Prob
Chi-Square	1	22.9165	<.0001
Likelihood Ratio Chi-Square	1	22.8214	<.0001
Continuity Adj. Chi-Square	1	21.9700	<.0001
Mantel-Haenszel Chi-Square	1	22.8657	<.0001
Phi Coefficient		-0.2254	
Contingency Coefficient		0.2199	
Cramer's V		-0.2254	

Fisher's Exact Test	
Cell (1,1) Frequency (F)	71
Left-sided Pr <= F	1.452E-06
Right-sided Pr >= F	1.0000

Fisher's Exact Test	
Table Probability (P)	9.148E-07
Two-sided Pr <= P	2.363E-06

Statistic	Value	ASE	95% Confidence Limits	
Gamma	-0.4447	0.0811	-0.6036	-0.2859
Kendall's Tau-b	-0.2254	0.0466	-0.3168	-0.1340
Stuart's Tau-c	-0.2125	0.0443	-0.2994	-0.1256
Somers' D C|R	-0.2314	0.0479	-0.3252	-0.1376
Somers' D R|C	-0.2196	0.0457	-0.3091	-0.1301
Pearson Correlation	-0.2254	0.0466	-0.3168	-0.1340
Spearman Correlation	-0.2254	0.0466	-0.3168	-0.1340
Lambda Asymmetric C|R	0.1027	0.0650	0.0000	0.2300
Lambda Asymmetric R|C	0.0000	0.0000	0.0000	0.0000
Lambda Symmetric	0.0549	0.0355	0.0000	0.1244
Uncertainty Coefficient C|R	0.0374	0.0155	0.0069	0.0678
Uncertainty Coefficient R|C	0.0388	0.0161	0.0073	0.0704
Uncertainty Coefficient Symmetric	0.0381	0.0158	0.0071	0.0691

Estimates of the Relative Risk (Row1/Row2)			
Type of Study	Value	95% Confidence Limits	
Case-Control (Odds Ratio)	**0.3843**	0.2586	0.5711
Cohort (Col1 Risk)	**0.6558**	0.5415	0.7943
Cohort (Col2 Risk)	**1.7064**	1.3770	2.1147

Cochran-Armitage Trend Test	
Statistic (Z)	4.7871
One-sided Pr > Z	<.0001
Two-sided Pr > |Z|	<.0001

Sample Size = 451