Paper 249-2008

Exploring Treasuries with SAS®

Taras Zlupko, CRSP®, University of Chicago, Chicago, IL

ABSTRACT

Treasuries represent an important part of the financial securities universe and are indispensable in academic and industry research and in day-to-day financial decision making. Empirical Treasury data is therefore used in a variety of ways impacting research results and real-life actions. Completeness and integrity of data, convenience of access and opportunity to deploy relevant analytics are some of the vital components when working with Treasury securities catering to reliability and timeliness of achieving relevant goals. These components come together when using the US Treasury database available from the Center for Research in Security Prices (CRSP) in the SAS environment.

Overview of the US Treasury database available form CRSP is provided in the paper. Particulars and usage of some CRSP calculated data items are highlighted in the context of accessing and processing the data with SAS. Application of some SAS tools when working with Treasury securities are exemplified and discussed.

CRSP US TREASURY DATABASE

CRSP US Treasury database contains descriptive information, market and derived data including issue identification, characteristics, data ranges, prices, yields, durations, returns, accrued interest and other data items on direct, marketable Treasury issues. Given historical depth of database coverage, data was collected from different sources and included manual data entry for early periods when electronic sources were not available.

The US Treasury data from CRSP is organized in two databases – Daily and Monthly.

Some major distinctions between the Daily and Monthly Treasury databases are as follows:

Data frequency:	Daily or Monthly.
Historical coverage:	Daily data begins in June 1961, Monthly – in December 1925.
Data items:	Rates of Certificate of Deposit, Commercial Paper, and Federal Funds as well as when-issued prices are included only in the Daily database. Data on the Public debt of the United States is provided in a separate data set.
Derived files:	Fama files are available only in the Monthly database.

Master file and cross-sectional file approach is represented in both Daily and Monthly databases. The data is sorted by issue then by quotation date in the master file and by quotation date and by issue in the cross-sectional file.

Both daily and monthly versions of CRSP US Treasury database include supplemental Fixed Term Indices File. The file is designed to provide readily available values for constructing yield curves and contains annualized yields of selected Treasuries with maturities from one to thirty years. Some additional data items such as CRSP issue identification number, number of years left to maturity, holding period return, accrued interest, duration, bid and ask prices are also included in the file. Daily series begin in June 1961 reflective of the daily database start date. Monthly series for different maturities start on different dates from January 1941 through January 1942 due to the sparsity of Treasuries satisfying selection criteria set forth by the methodology in the early period covered by the database.

Fama files represent unique component of CRSP Monthly US Treasury database derived from underlying market data. These files have been made available by Professor Eugene Fama. Four groups of files are available in the database as follows:

Treasury Bill Term Structure Files	– bid, ask and bid-ask average prices and corresponding yields, forward rates and holding period returns of six and twelve month Treasuries;
Maturity Portfolios Returns Files	– returns of equal-weighted portfolios of Treasuries within either six or twelve month maturity intervals.

Fama-Bliss Discount Bonds Files – prices and yields of artificial discounted Treasuries constructed according to methodology developed by Fama and Bliss (1987) with data going back to 1952;

Risk Free Rates File – yields of one and three month Treasuries. Three values are included for each maturity based on bid, ask and bid-ask average price.

CRSP Monthly US Treasury Database Guide provides more information about particulars of the Fama files. Both Daily and Monthly US Treasury Database Guides are available online from CRSP and should be consulted when working with relevant database.

ACCESS TO TREASURIES

Treasury data from CRSP is available in SAS format. Therefore access to data is straightforward in the SAS environment and requires only setting up relevant libraries.

Standard directories are created when CRSP databases are installed. These directories differ by the date of database cut. The following statements set up libraries of daily and monthly Treasuries ending December 2007 called bdz and bmz respectively:

```
LIBNAME bdz ' C:\crspdata\BDZ200712\sas\';
LIBNAME bmz ' C:\crspdata\BMZ200712\sas\';
```

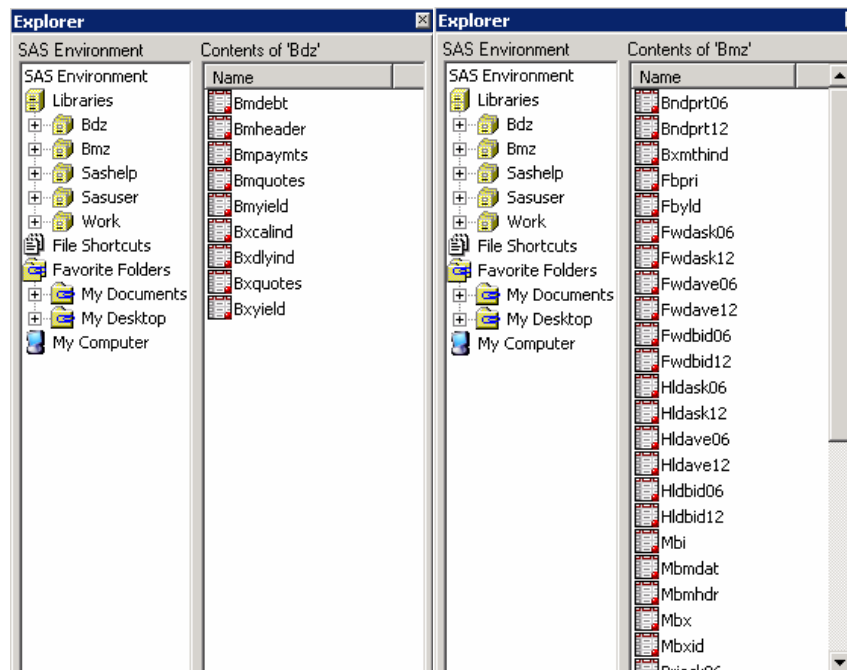Daily and monthly data layout in SAS Explorer window is shown in Figure 1 below.

Figure 1. Daily (left) and Monthly CRSP US Treasury databases in SAS Explorer.

As seen in Figure 1, monthly database contains more data sets than daily database. This is primarily due to the presence of derived Fama files mentioned earlier. There are twenty nine Fama files in total: Twenty four Treasury Bill Term Structure files, two Maturity Portfolios Returns files and two Fama-Bliss Discount Bonds files and the Risk Free Rates file. CRSP documentation referenced at the end of the paper provides detailed discussion of these files.

Although daily database consists of fewer data sets, covers smaller number of securities and shorter historic coverage than monthly database, daily database is more than ten times larger than monthly database which is due to higher frequency of data. There are about 116,400 observations and around 5,850 distinct Treasury securities included in CRSP monthly database and about 1,930,300 observations and 4,000

distinct Treasuries included in the daily database as of 2007 year-end. This is on average about twenty observations per security in monthly database and close to five hundred observations per security in daily database.

Identification and selection of Treasury securities in CRSP database can be done using CRSPID and/or CUSIP. CRSPID is the unique issue identification number assigned by CRSP to each Treasury security in the database. This is consistent with practice employed in other CRSP databases such as Stock, Indices and Mutual Funds which have their own identifiers. CRSPID incorporates maturity date, type of issue and uniqueness number. Another identifier included in the database is CUSIP. As mentioned in CRSP documentation, the Committee on Uniform Security Identification Procedures began assigning CUSIP identification numbers in 1968. The earliest observation of Treasury security with CUSIP assigned available in the database is dated February 15, 1969.

WORKING WITH TREASURIES

One of the benefits of CRSP US Treasury database is availability of supplemental files. These files are helpful in accomplishing some popular tasks related to Treasury securities analysis and are convenient to use in the SAS environment. For example, the Fixed Term Indices File can be used to create yield curves without going into the process of issue selection and yield calculation.

Yield curve is one of the most actively referenced indicators of the Treasuries market. Annualized yields to maturity for one-, two-, five-, seven-, ten-, twenty- and thirty-year Treasuries are pre-calculated and pre-selected by CRSP and provided in the Fixed Term Indices File. The file is available in both daily and monthly CRSP database.

After some simple additional transformations of the file using SAS, yields of different maturity Treasuries available in the Fixed Term Indices File are plotted to form the yield curves. Figure 2 shows yields of Treasuries of different maturity on days when highest positive and negative yield spreads between thirty-year and one-year Treasury issues were recorded according to data in the monthly database. Yield spread is the difference in yields between Treasuries of different maturity.
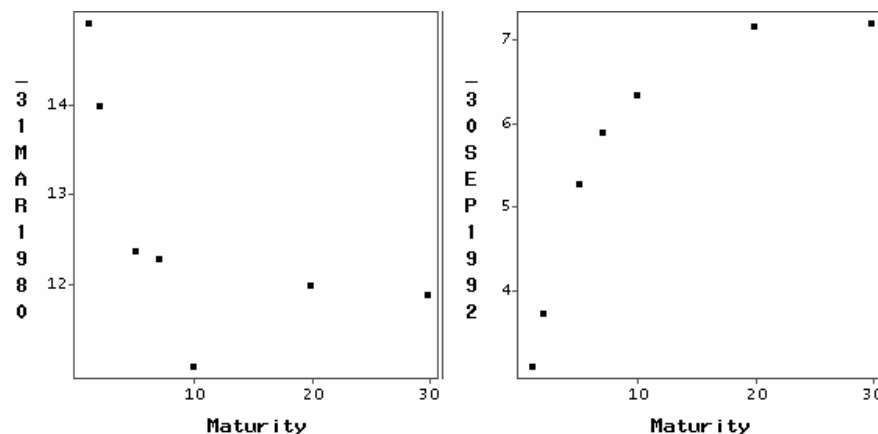

Figure 2. Yields on days of extreme negative (left) and positive yield spreads between
one- and thirty-year US Treasuries.

Highest positive yield spread between thirty-year and one-year Treasuries was recorded at the end of September 1992. Steep implied yield curve is evident on that day from Figure 2. Extreme negative yield spread occurred on March 30, 1980. Yields on that day imply the inverted yield curve. The yield of ten-year Treasury appears to be outside of the path of the implied inverted yield curve and may be an outlier requiring further attention. Also, note that the overall level of yields is quite different on two days.

The Fixed Term Indices File does not include short-term Treasury Bills. These can be added to the picture from the Risk Free Rates File which is readily available in monthly database. Yield based on the bid/ask average price needs to be selected from the Risk Free Rates File and date range restricted to that of the rest of the data. Of course, securities for the yield curve can be selected straight from the market-reported data available in the database. This is unavoidable if daily yield curves including Treasury Bills to be constructed since there is no Risk Free Rates File in CRSP daily Treasury database. The user may also

want to consider preferences between active on-the-run and off-the-run issues when selecting Treasuries for yield curve.

Yield curve represents a cross-section of yields on any given date. The time-series of yields provides the insight of changing yield levels over time. Figure 3 plots annualized yields of one- and thirty-year Treasuries from CRSP Fixed Term Indices File indicating the points of highest and lowest yield spread identified earlier and used in Figure 2.

Dotted line in Figure 3 represents yields of thirty-year Treasuries. Periods of inverted yield curve exist when solid line representing one-year Treasuries is crossing above the dotted line. Highest Treasury yields visible in Figure 3 are during early 1980th. Largest negative yield spread between thirty- and one-year Treasuries also occurred during that period and is identified by 'x' in Figure 3 as is the largest positive yield spread.
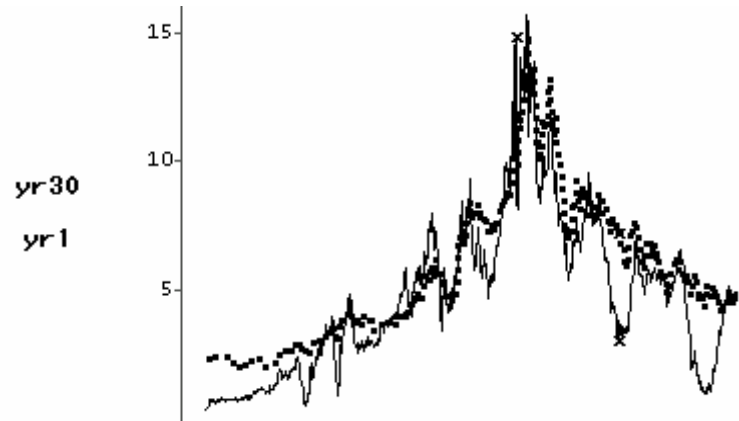

Figure 3. Annualized yield to maturity of one- and thirty-year Treasuries from
CRSP Fixed Term Indices File, 1942-2007.

Not only yields but also returns on Treasuries are available from CRSP. Treasury return is defined by CRSP as price change plus amount of interest paid divided by last price where price includes accrued interest. As pointed out in Fisher and Lorie [1977], Treasury returns "computed in this manner are the precise analogs of returns including cash dividends" – total returns of stocks. Returns together with yields are included in both Fixed Term Indices File and the Risk Free Rates File. Therefore it is straightforward to compare returns for different benchmark securities. Figure 4 and Figure 5 plot returns of one- and thirty-year Treasuries which are available in the Fixed Term Indices File.
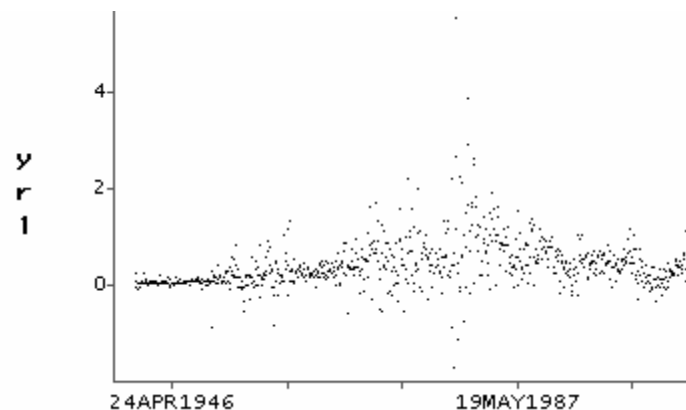

Figure 4. Returns of one-year Treasuries from CRSP Fixed Term Indices File, 1942-2007.
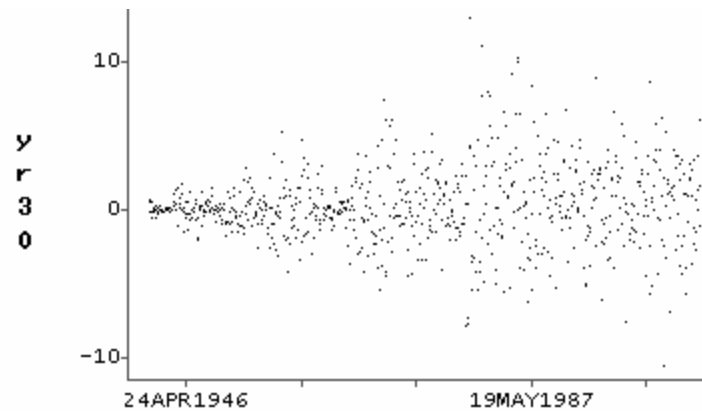
Figure 5. Returns of thirty-year Treasuries from CRSP Fixed Term Indices File, 1942-2007.

As seen in Figures 4 and 5, returns of thirty-year Treasuries are more dispersed than returns of one-year Treasuries. They are also spread over larger interval of values – from about minus ten to ten, whereas one-year Treasuries are largely contained within the minus two to four range. It is also visible in Figure 5 that dispersion is increasing in the early part of the period under consideration.

Spreads between bid and ask prices of Treasury securities and their trends are of interest when discussing the Treasuries market. Bid/ask spread is the difference between lowest ask and highest bid prices at any given time. The commonly used measure of the bid/ask spread is the quoted spread. It is the ratio of difference between ask and bid prices to ask price expressed in percentage:

$$sprdq = 100 * (ask - bid)/ask$$

Both daily and monthly CRSP databases contain bid and ask prices of individual Treasuries. Below is example of code computing annual average bid/ask quoted spread using monthly data. The code checks for missing data and bid/ask consistency before proceeding with calculations. Pric1r and pric2r are bid and ask prices respectively. The data is from the cross-sectional MBX file available in monthly database. There is also master MBMDAT file which includes prices in the database. The primary difference between the two data sets is that MBX is sorted by quote date and MBMDAT is sorted by issue. Using MBX eliminates the need in sorting by date before running *proc means*.

```
/*compute individual bid/ask spreads*/

data spreads; set bmz.mbx;
    where pric2r>0 and pric1r>0 and pric2r^=pric1r;
        sprdq=100*(pric2r-pric1r)/pric2r;
            year=year(qdate);
run;

/*compute annual means of spreads*/

proc means data=spreads noprint;
    by year;
        var sprdq;
            output out=annual_mean(drop=_type_)
                    mean=spread;
run;
```

Average annual bid/ask spreads computed above are represented by solid line in Figure 6 and are in percentage. Number of observations (end-month quotes) used to create the average annual bid/ask spreads is also shown in Figure 6.
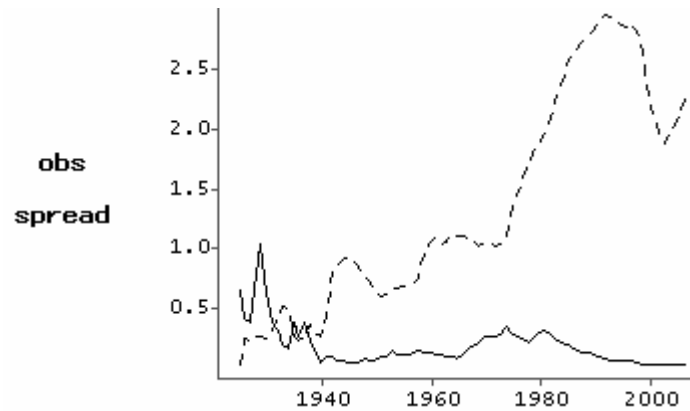
Figure 6. Average annual bid/ask spread (percentage) and quote count (thousands), 1925-2007.

Treasury spreads tend to fall over time which is evident in Figure 6. The number of tradable Treasury securities has grown over the period under consideration as seen from the count of quotes used to produce the average annual spreads figures.

In addition to market-reported data, CRSP database includes total and publicly held face value outstanding of each Treasury security. These figures are published by the US Treasury. Amounts outstanding are available in both daily and monthly database. Having the data for each security it is straightforward to find out the amounts outstanding of all Treasuries on any given date. This is illustrated below using the BMDEBT data set from daily database which is dedicated to amounts outstanding. Monthly database does not have separate data set containing this data. Instead, amounts outstanding are included in the cross-sectional MBX file. The BMDEBT data set has to be checked for possible missing values and sorted by date before *proc means* is used:

```
proc sort data=bdz.bmdebt out=bdz_bmdebt;
          by qdate;
run;

proc means data=bdz_bmdebt noprint; by qdate;
          var totout pubout;
              output out=debt
                  sum=total public;
                      format qdate date9.;
run;
```

Results of the above are shown in Figure 7. Solid upper line is representing face value outstanding of US Treasuries. Bottom line is the publicly held face value outstanding of Treasuries in CRSP database.
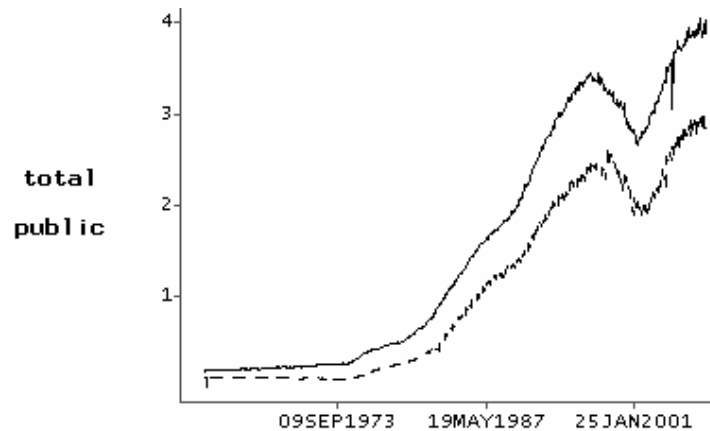

Figure 7. Total and publicly held face value outstanding of US Treasuries (trillions), 1961-2007.

6

Breakdown of total amount outstanding of US Treasuries shown in Figure 7 by maturity of securities at any given time is of further interest. It is possible to easily produce relevant statistics with SAS since CRSP issue identifier number (CRSPID) includes a character specifying type of Treasury. Figure 8 shows total face value outstanding of Treasury bills, notes and bonds over the period from 1961 to 2007.
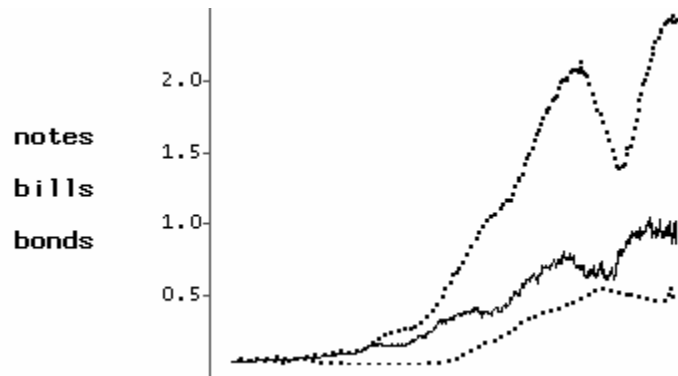

Figure 8. Total face value outstanding of bills, notes and bonds (trillions), 1961-2007.

As shown in Figure 8, Treasury note was overall the most popular vehicle of debt issuance followed by bills and bonds. This is clearly visible during the most of the period under consideration.

Important supplementary data is provided in the CALENDAR data set, which is included in the daily CRSP Treasury database. This data set contains rates of certificates of deposit and commercial papers of different maturities as well as the federal funds effective rate. Figure 9 shows historical federal funds effective rate. As the rest of the data in the daily database, federal funds series begins in June 1961.
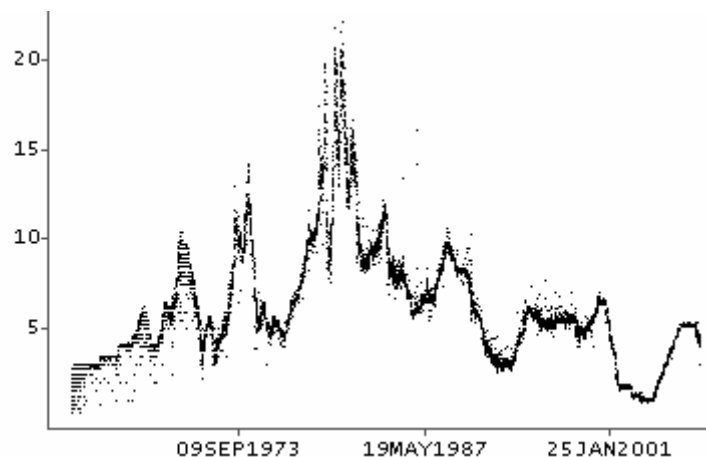

Figure 9. Federal funds effective rate, 1961-2007.

Supplemental data available in CRSP Treasury database can conveniently be used with the rest of Treasury data for further processing and analysis. For example, it may be of interest to look at average levels of and relations between the federal funds effective rate and Treasury securities of different maturity. This can be done by combining the federal funds effective rate data from the CALENDAR data set, yields of Treasuries with one or more years to maturity from the Fixed Term Indices File and yields of short term Treasuries from the Risk Free Rates File. Relevant statistics can be obtained by running *proc corr*.

Code below can be used to create data set containing data items of interest and calculate sixty-month rolling correlation, mean and standard deviation. Updated and simplified SAS macro developed in Zlupko (2007) is used to calculate rolling statistics.

7

```
/*sort and transpose Fixed Term Indices File*/

proc sort data=bmz.bxmthind(keep=termtype qdate ytm) out=rate;
            by qdate termtype;
proc transpose data=rate out=rate(drop=_name_ _label_);
            by qdate;
                id termtype;
run;

/*merge Calendar, Risk Free Rates and Fixed Term Indices files;
  keep data after June 1961 for completeness*/

data rate;
merge bdz.bxcalind(keep=qdate ffefrt)
        bmz.riskfree(keep=qdate ave_1 ave_3) rate(in=g);
            by qdate; if g=1;
                where qdate>='30Jun1961'd;
run;

/*calculate sixty-month rolling statistics for selected pair of
  data items*/

%let r1=ffefrt;
%let r2=ave_1;

%macro stats;
data stats;
%do i=1 %to 500;
        data temp; set rate; if %eval(&i)<=_n_<%eval(&i+60);
        proc corr data=temp outp=temp1 noprint;
                var &r1 &r2;
data stats; set stats temp1;
proc datasets; delete temp temp1;
run;
%end;
%mend;
%stats;
```

Sixty-month rolling correlations and means of the federal funds effective rate and one-month and thirty-year Treasuries over the period from June 1961 to December 2007 produced by the code above are plotted in Figure 10 and Figure 11.
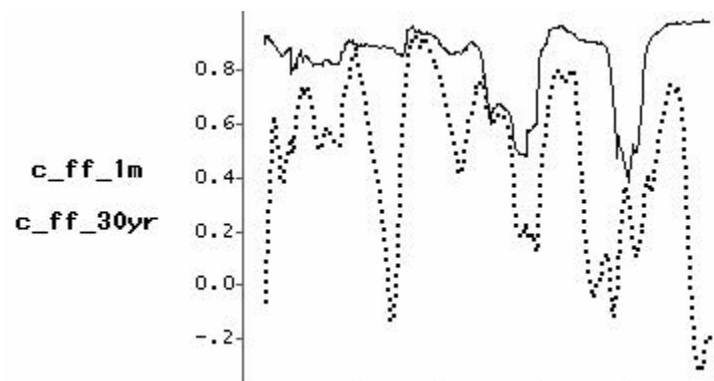

Figure 10. Sixty-month rolling correlations of the federal funds effective rate and one-month and thirty-year Treasury rates, 1961-2007.
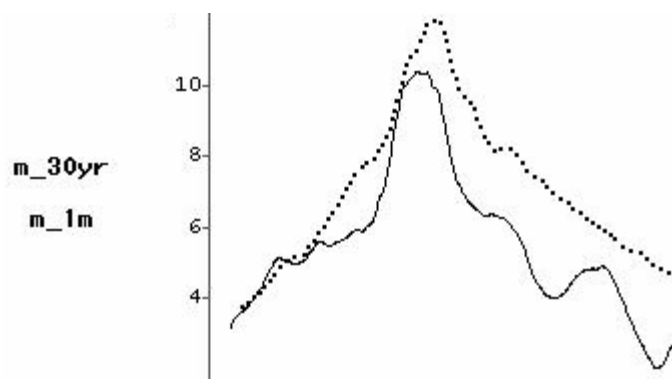
8

Figure 11. Sixty-month rolling means of one-month and thirty-year Treasury rates, 1961-2007.

Figure 10 shows that five-year rolling rates of one-month Treasury bills are overall more correlated with the federal funds effective rate than rates of thirty-year Treasury bonds. As seen in Figure 11, rolling average rates of short-term Treasuries are lower than the rates of long-term Treasuries.

CONCLUSIONS

Monthly and Daily US Treasury databases available from the Center for Research in Security Prices represent comprehensive universe of Treasuries and provide vital level of detail and history of this type of financial securities needed for research and decision-making purposes. There are also some other supplementary data items in the database. In addition to market-reported Treasury data CRSP supplies some relevant derived data items in the database. Values computed by CRSP make the use of the database still more convenient. Methodology of data collection and calculations are described in CRSP documentation and should be consulted before proceeding with data access and analysis.

Working with CRSP Treasury data is greatly facilitated by its availability in SAS format. Flexibility and efficiency of data analysis is exercised by possibility of starting the process from either market-reported values provided in the database or pre-computed derived files which contain some popular data items when applicable. Calculations can conveniently be implemented in SAS and may be performed when using derived values available from CRSP or when creating a new process if modification to methodology is required. SAS tools of data preparation and analysis create efficient environment when working with Treasury securities.

REFERENCES

Center for Research in Security Prices. 2007. *CRSP Daily Treasury US Database Guide. Version CA295.200701.* Chicago: The University of Chicago Graduate School of Business. http://www.crsp.chicagogsb.edu/support/documentation/pdfs/daily_treasury.pdf

Center for Research in Security Prices. 2007. *CRSP Monthly Treasury US Database Guide. Version CA295.200701.* Chicago: The University of Chicago Graduate School of Business. http://www.crsp.chicagogsb.edu/support/documentation/pdfs/monthly_treasury.pdf

Fama, E.F. and R.R.Bliss. 1987. "The Information in Long-Maturity Forward Rates". *American Economic Review* 77:680-692.

Fisher, L. and J.H.Lorie. 1977. A Half Century of Returns on Stocks and Bonds: Rates of Return on Investments in Common Stocks and on U.S. Treasury Securities, 1926-1976. Chicago: University of Chicago Graduate School of Business, 174p.

Zlupko, Taras. 2007. "CRSP Indices in SAS." *Proceedings of SAS Global Forum, Orlando, FL,* 167-2007.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author:
>Taras Zlupko
>CRSP, University of Chicago
>105 West Adams Street, Suite 1700
>Chicago, IL 60603
>Work Phone: (312)263-6654
>E-mail: Taras.Zlupko@crsp.ChicagoGSB.edu

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.