```
dm 'log;clear';
dm 'output;clear';
dm 'odsresults;clear';
proc datasets library = work kill; quit;


* SAS 1: Connect WRDS from your computer **********************;

%let wrds = wrds.wharton.upenn.edu 4016;
options comamid = TCP;
signoff;
signon wrds username = ****  password = ****;
*signon username = _prompt_;

dm 'log;clear';
dm 'output;clear';
dm 'odsresults;clear';
proc datasets library = work kill; quit;

* define the folder where my final data will be stored;
%let my_directory = D:\SAS;
libname my_lib 'D:\SAS';


rsubmit inheritlib = (my_lib = my_lib);
```

```
* ************************************************************************************
* ******************** W R D S   R E S E A R C H   M A C R O S **********************
* ************************************************************************************
* WRDS Macro: CRSPMERGE
* Summary   : Merges CRSP Stocks and Events Data
* Date      : April 14, 2009
* Author    : Rabih Moussawi and Luis Palacios, WRDS
* Variables : - S: Monthly/Daily, defaults to Monthly, but s=d for CRSP Daily data
*             - START, END: Start and End Dates. Example Date Format: 01JAN2000
*             - SFVARS: Stock File Vars to extract. e.g. PRC VOL SHROUT
*             - SEVARS: Event File Vars to extract. e.g. TICKER NCUSIP SHRCD EXCHCD
*                  warning: including DIVAMT may result in multiple obs per period
*             - FILTERS: Additional screens using variables in SFVARS or SEVARS
*                        (default no filters)
*             - OUTSET: Output Dataset Name (default names crsp_m or crsp_d)
* ************************************************************************************;

%MACRO CRSPMERGE (S =, OUTSET = crsp_&s.,
START=, END=,
SFVARS =,
SEVARS =,
FILTERS =);

/* Check Series: Daily or Monthly and define datasets - Default is Monthly */
%if %sysfunc(lowcase(&s)) = d %then %let s = d; %else %if %sysfunc(lowcase(&s)) ne d %then %let s = m;

%let sf = crsp.&s.sf;
%let se = crsp.&s.seall;
%let senames = crsp.&s.senames;

%put #### START. Merging CRSP Stock File (&s.sf) and Event File (&s.se);

*options nonotes;
%let sdate = %sysfunc(putn("&start"d,5.));
%let edate = %sysfunc(putn("&end"d,5.));

%let sevars = %sysfunc(compbl(&sevars));
%let sevars = %sysfunc(lowcase(&sevars));
%let nsevars = %sysfunc(countw(&sevars));

/* create lag event variable names to be used in the RETAIN statement */
%let sevars_l = lag_%sysfunc(tranwrd(&sevars,%str( ),%str( lag_)));
```

```
%if %length(&filters) > 2 %then %let filters = and &filters;
%else %let filters = %str( );

/* Get stock data */

proc sql;
     create table __sfdata
    as select *
    from &sf (keep = permno date &sfvars)
    where date between &sdate and &edate and permno in
    (select distinct permno
     from &senames( where = (&edate>=NAMEDT and &sdate<=NAMEENDT) keep = permno namedt nameendt )
     )
    order by permno, date;
quit;

/* Get event data */

proc sql;
     create table __sedata
     as select a.*
     from &se (keep= permno date &sevars) as a,
    (select distinct permno,
           case when min(namedt)<&sdate then max(namedt*(namedt<&sdate)) else min(namedt) end as minnamedt
           from &senames( where = (&edate>=NAMEDT and &sdate<=NAMEENDT) keep = permno namedt nameendt )
           group by permno) as b
     where a.date>=b.minnamedt and a.date<=&edate and a.permno = b.permno
     order by a.permno, a.date;
quit;

/* Merge stock and event data */

/* create lag event variable names to be used in the RETAIN statement */

%let eventvars = ticker comnam ncusip shrout siccd exchcd shrcls shrcd shrflg trtscd nmsind mmcnt nsdinx;

data &outset. (keep=permno date &sfvars &sevars);
     merge __sedata (in=eventdata) __sfdata (in=stockdata);
     by permno date;
     retain &sevars_l;

     %do i = 1 %to &nsevars;
          %let var = %scan(&sevars,&i,%str( ));
```

```
            %let var_l = %scan(&sevars_l,&i,%str( ));
            %if %sysfunc(index(&eventvars,&var))>0 %then %do;
            if eventdata or first.permno then &var_l = &var. ;
            else if not eventdata then &var = &var_l. ;
            %end;
      %end;

      if eventdata and not stockdata then delete;
      drop &sevars_l ;
run;

/* Some companies have many distribution on the same date (e.g. a stock and cash dist)  */
/* Records will identical except for different DISTCD and DISTAMT */

proc sort data = &outset. noduplicates;
where 1 &filters;
      by permno date;
run;

/* House Cleaning */

proc sql;
      drop table __sedata, __sfdata;
quit;

*options notes;
%put #### DONE . Dataset &outset. Created! ;

%MEND CRSPMERGE;

* ************************************************************************
* *************  Material Copyright Wharton Research Data Services  ******************
* ***************************** All Rights Reserved ********************************
* ************************************************************************;
```

```
* SAS 2: Define libraries and macro variables ****************************************;

* sample period;
%let comp_start_date = 01JAN1961; * we get the two years of accounting data before 1963;
%let start_date = 01JAN1961;
%let end_date = 31DEC2012;
```

```
*SAS 3: Construct BE data ************************************;

data compustat (drop = at indfmt datafmt popsrc consol);
     set compa.funda (keep = gvkey datadate at pstkl txdb
                itcb pstkrv seq pstk indfmt datafmt popsrc consol);
     where indfmt='INDL' and datafmt='STD' and popsrc='D' and consol='C'
          and datadate >="&comp_start_date"d and datadate <="&end_date"d;
run;

* Add permno and permco to BE data using the link-used table;
* The nobs might increase because a firm can be matched to multiple permno's;
proc sql;
     create table compustat_permno
     as select distinct a.*, b.upermno as permno, b.upermco as permco
     from compustat as a
     left join my_lib.ccmxpf_lnkused
          ( keep = ugvkey upermno upermco ulinkdt ulinkenddt usedflag ulinktype
          where = (usedflag = 1 and ulinktype in ("LU","LC")) ) as b
     on a.gvkey = b.ugvkey
     and (b.ulinkdt <= a.datadate or missing(b.ulinkdt) = 1)
     and (a.datadate <= b.ulinkenddt or missing(b.ulinkenddt) = 1)
     order by a.datadate, a.gvkey;
quit;
proc sort data = compustat_permno; by gvkey datadate; run;
```

```
* Calculate BE;
data BE;
      set compustat_permno (where = (missing(permno) = 0));

      year = year(datadate);

      if missing(ITCB) then ITCB = 0; * investment tax credit;

      BVPS = PSTKRV; * preferred stock - redemption value;
      if missing(BVPS) then BVPS = PSTKL; * preferred stock - liquidating value;
      if missing(BVPS) then BVPS = PSTK; * preferred stock- par value;
      if missing(BVPS) then BVPS = 0;

      BE = SEQ + TXDB + ITCB - BVPS; * If SEQ or TXDB is missing, BE, too, will be missing;
      if BE<=0 then BE = .; * If BE<0, the value of BE is taken to be missing;

      label datadate = "Fiscal Year End Date";
      keep gvkey datadate year BE permno permco;
run;

* In some cases, firms change the month in which their fiscal year ends,
* resulting in two entries in the Compustat database for the same calendar year y.
* In such cases, data from the latest in the given calendar year y are used.;
proc sort data = BE; by gvkey permno year datadate; run;
data BE;
      set BE;
      by gvkey permno year datadate;
      if last.year;
run;
proc sort data = BE nodupkey; by gvkey permno year datadate; run;
```

```
* SAS 4: Merge CRSP stock and event file and add risk-free rate ******************;

%let filter=%str(shrcd in (10,11) and exchcd in (1,31,2,32,3,33));

%crspmerge(s = m, outset = CRSP,
start = &start_date, end = &end_date,
sfvars = permco ret vol shrout prc altprc,
sevars = siccd shrcd exchcd dlstcd dlret,
filters=&filter);

proc sql;
     create table CRSP_M
     as select distinct a.*, b.rf
     from CRSP as a
     left join ff.factors_monthly as b
     on year(a.date) = year(b.date) and month(a.date) = month(b.date);
quit;
```

```
* SAS 5: Construct ME and return data ************************************;

* Calculate excess return adjusted for delising;
data CRSP_M2;
      set CRSP_M;

      year = year(date);

      * calculate market capitalization;
      if abs(altprc)>0 and shrout>0 then Meq = abs(altprc)*shrout/1000;

      * if dlret is missing, follow Shumway (1997) to determine dlret;
      if missing(dlstcd) = 0 and missing(dlret) =1 then do;
            if dlstcd in (500, 520, 574, 580, 584) or (dlstcd>=551 and dlstcd<=573)
                  then dlret = -0.3;
            else dlret = -1;
      end;

      * calculate return adjusted for delisting;
      if missing(dlstcd) = 0 then do;
            if missing(ret) = 0 then retadj = (1+ret)*(1+dlret)-1;
            else retadj = dlret;
      end;
      else retadj = ret;

      eretadj = retadj - rf;
run;
proc sort data = CRSP_M2; by date permco Meq; run;
```

```
* There are cases when the same firm (permco) has two or more securities (permno)
at the same date.
* We aggregate all ME for a given permco and date,
* and assign this aggregated ME to the permno with the largest ME;
data CRSP_M3;
      set CRSP_M2;
      by date permco Meq;
      retain ME;

      if first.permco and last.permco then do;
            ME = Meq;
            output;
      end;
      else do;
            if first.permco then ME = Meq;
            else ME = sum(Meq, ME);
            If last.permco then output;
      end;
run;

proc sort data = crsp_m3 nodupkey; by permno date; run;
```

```
* SAS 6: Merge BE and ME with Return Data ************************************;

* Calculate BM from the previous year and June ME from this year for each permno;
data ME_Jun;
      set CRSP_M3 (where = (month(date) = 6 & missing(ME) = 0));
      t = year(date);
      ME_Jun = ME;
      keep permno t ME_Jun;
run;
proc sort data = ME_Jun; by permno t; run;
data ME_last_Dec;
      set CRSP_M3 (where = (month(date) = 12 & missing(ME) = 0));
      t = year(date)+1;
      ME_last_Dec = ME;
      keep permno t ME_last_Dec;
run;
proc sort data = ME_last_Dec; by permno t; run;
data BE_last_year;
      set BE (where = (missing(BE) = 0));
      t = year+1;
      BE_last_year = BE;
      keep permno t BE_last_year;
run;
proc sort data = BE_last_year; by permno t; run;

data ME_BM;
      merge ME_Jun (in = a) BE_last_year (in = b) ME_last_Dec (in = c);
      by permno t;
      if a & b & c;

      BM = BE_last_year/ME_last_Dec;
      keep permno t ME_Jun BM;
run;
```

```
* Match each permno's monthly return to the corresponding BM and ME;
data ret;
      set CRSP_M3;
      if month(date)>6 then t = year(date);
      else t = year(date)-1;
run;
proc sort data = ret; by permno t date; run;
data ret_ME_BM;
      merge ret (in = a) ME_BM (in = b);
      by permno t;
      if a;
run;

* Also add the mktcap and stock price from the previous month;
proc sort data = ret_ME_BM; by permno date; run;
data ret_ME_BM;
      set ret_ME_BM;

      altprc_lag1 = lag1(altprc);
      ME_lag1 = lag1(ME);

      permno_lag1 = lag1(permno);
      date_lag1 = lag1(date);

      if (permno NE permno_lag1) or (intck('month',date_lag1,date)>1) then do;
            altprc_lag1 = .;
            ME_lag1 = .;
            end;
run;

* Exclude observations with missing values;
data assignment1_data;
      retain permno date year exchcd siccd retadj eretadj altprc_lag1 ME_lag1 ME_Jun BM;
      set ret_ME_BM;
      if nmiss(retadj, ME_lag1, ME_Jun, BM) = 0;
      keep permno date year exchcd siccd retadj eretadj altprc_lag1 ME_lag1 ME_Jun BM;
run;
```

```
* Download data from the WRDS server to my PC;
proc download data = compustat_permno out = my_lib.compustat_permno; run;
proc download data = CRSP_M out = my_lib.CRSP_M; run;
proc download data = assignment1_data out = my_lib.assignment1_data; run;

endrsubmit;
```