Replication code Serial Acquirers to get the serial acquirer classification (1/2)

Macias, A. J., P. R. Rau, and A. Stouraitis. 2023. Solving serial acquirer puzzles. Review of Corporate Finance Studies. Forthcoming.

From main text:

3.1. Data

We obtain our sample of acquisitions of U.S. targets (public, private, and subsidiary firms) announced by U.S. public acquirers during 1989-2018 from the Thomson One's Securities Data Company (SDC) database. We require that the bidder seeks to acquire more than 50% ownership of the target, and that the Center for Research in Security Prices (CRSP) and COMPUSTAT provide information for the acquirer. We obtain stock return and accounting data for the universe of U.S. publicly listed firms from CRSP and COMPUSTAT as of the quarter prior to the announcement date. To alleviate truncation concerns at the start and end of our sampling window, and to avoid concerns about the coverage of SDC in the early 1980s (Netter, Stegemoller, and Wintoki 2011), we require that (i) the acquirer has not conducted any acquisition in the 4 years prior to the start of our sample period (1985-1988); and (ii) that the acquirer conducts its first acquisition at the latest by the year 2016. Our initial sample consists of 55,482 mergers and acquisitions conducted by 8,640 unique acquirers. Our sample increases to 27,813 unique firms (239,120 firm-year observations) when we include firms that do not conduct any acquisitions within our sample period.

3.2. Classifying serial acquirers

We use cluster analysis to classify acquirers into one of four distinct types that we denote as loners, occasional acquirers, sprinters, and marathoners. Our classification is based on three common-sense dimensions. We perform the cluster analysis (and reclassify acquirers) annually, using data for the past 5 years. Based on three dimensions (total number of acquisitions over the past 5 years, number of acquisition blocks and average intensity), we conduct a *k*-median cluster analysis to classify acquirers into different categories, within 5-year rolling windows (starting with 1989-1993 and ending with 2014-2018) as depicted in Figure 1.

0. Preliminary code to set up the dataset that contains the serial acquirer classification used in the rest of the analysis.

Since our data comes from proprietary databases (e.g., Compustat, CRSP, SDC) we will focus on submitting the replication code assuming that any reader with access to such databases would be able to create the original merged dataset from such sources instead of us uploading such data that our licenses do not allow us to upload open to the public for people without such licenses. The paper describes the commonly used filters and merging methods used to generate the dataset from such proprietary databases. We include selected extracts of the original whole dataset where convenient to exemplify how the code would work with the whole dataset.

In general, we handle and clean the data in SAS and run the econometric analysis in Stata. Depending on the type of descriptive statistics, we use either SAS or Stata. In several tables we build the tables in Excel with the appropriate format using the output from the SAS or Stata. The analysis is based on two levels, namely, deal and firm-year level. To avoid sharing an individual dataset for each table, the shared datasets might already contain some of the defined variables, hence dropping some of them to be used in the provided shared sub-sample would be needed to see how such variable is used. We left in the dataset several variables used in the robustness and ancillary analyses. The main variables of interest are defined in the appendix or its labels are self-explanatory based on the analysis and corresponding code.

(We acknowledge that more experienced coders would have found more efficient ways to get the serial acquirer classification. However, the longer code enables for easier checking for errors and for obtaining intermediate variables or analysis]

For this preliminary code we split the process in four main parts:

0.1 SAS code to get 5- year rolling window sub-samples for then sending unto Stata for the cluster analysis with the goal to get the ex ante serial acquirer classification in each 5-year rolling window sub-sample.

0.2 Ex ante Serial Acquirer Classification based on 5-year rolling windows.

0.3 Stata code to get 5- year rolling window sub-samples with cluster analysis with the goal to get the ex ante serial acquirer classification in each 5-year rolling window sub-sample.

0.4 SAS code that uses the serial acquirer classification in each 5-year rolling window sub-sample generated in Stata. This code continues to prepare the sample to start with the actual analysis reported in the Tables.

0.4.A First for LARGE deals, and

0.4.B for All deals without size filter

0.1 SAS code to get 5- year rolling window sub-samples for then sending unto Stata for the cluster analysis with the goal to get the ex ante serial acquirer classification in each 5-year rolling window sub-sample.

Starting sample: merged and clean data set

/* SAS code for data on whole sample */

```
/* First, we present the required steps needed for the serial acquirer classification and then we start
presenting the code for the replication of the tables in a subsequent file*/

/* SAS */
/* Get information at the acquirer level [bdrcusip6] for the whole sample 1990-2019 */
/* Note how the original whole dataset is called "sdc1980Mar2019APASPPaft1989", yet I build a sub-sample to avoid issues of making public
whole dataset for SDC, Compustat and CRSP datasets. The subsequent code could simply change the name of the whole data set and instead use
the sub-sample created for a sub-set of years */
data sdc1980Mar2019APASPPaft1989; set serialac.sdc1980Mar2019APASPPaft1989; run; *deal level;
data sdc19952001APASPPinitialDlLvl; set sdc1980Mar2019APASPPaft1989; * Initial sample;
            if annyear => 1995 and annyear <= 2001; run;
data serialac.sdc19952001APASPPinitialDlLvl; set sdc19952001APASPPinitialDlLvl; run; * Initial shared Sub-
sample 1995-2001 deal level;


data bdrcusip6anndate1deal; set sdc1980Mar2019APASPPaft1989; keep bdrcusip6anndate  bdrcusip6 anndate; run;
      proc sort data=bdrcusip6anndate1deal ;by bdrcusip6anndate; run;
      proc sort data=bdrcusip6anndate1deal nodupkey; by bdrcusip6anndate; run;
      proc sort data=bdrcusip6anndate1deal ; by bdrcusip6 anndate; run;
data bdrcusip6anndate1dealA; set bdrcusip6anndate1deal;
      by bdrcusip6;
            format laganndatepriordeal   yymmdd9.;
            lagbdrcusip6 =  lag(bdrcusip6);
            laganndatepriordeal =  lag(anndate);
            if first.bdrcusip6 then firstbdrcusip6=1;
            if first.bdrcusip6 then laganndatepriordeal=. ;
            diffdayspriordeal75p = anndate - laganndatepriordeal ; *in days;
            if first.bdrcusip6 then lagbdrcusip6=. ;
```

```
            if first.bdrcusip6 then diffdayspriordeal75p = .;
            if first.bdrcusip6 then timesincelastdeal=. ;
            timesincelastdeal = diffdayspriordeal75p;
                  run;
      proc sort data=sdc1980Mar2019APASPPaft1989 ; by bdrcusip6 anndate; run;
      proc sort data=bdrcusip6anndate1dealA ; by bdrcusip6 anndate; run;
data sdc1980Mar2019APASPaft1989A0; merge sdc1980Mar2019APASPPaft1989(in=a)  bdrcusip6anndate1dealA;
      if a; by bdrcusip6 anndate; run;

*vars to get: mnadealindx mnadealinblockindx75p  mnablockindx75p ;
proc sort data=sdc1980Mar2019APASPaft1989A0; by bdrcusip6 anndate; run;
data sdc1980Mar2019APASPaft1989A; set sdc1980Mar2019APASPaft1989A0;
      by bdrcusip6;
            format laganndatepriordeal   yymmdd9.;
            lagbdrcusip6 =  lag(bdrcusip6);
            if first.bdrcusip6 then firstbdrcusip6=1;
            if first.bdrcusip6 then mnablockindx75p=1;
            if first.bdrcusip6 then mnadealinblockindx75p=1;
            mnadealindx +1;
            mnadealinblockindx75p +1;
            if first.bdrcusip6 then mnadealinblockindx75p=1;
            lagmnablockindx75p =  lag(mnablockindx75p);
            laganndatepriordeal =  lag(anndate);
            if laganndatepriordeal =  anndate and lagbdrcusip6 = bdrcusip6 then sameanndate =1;
            windowdaysthreshold75p= 407   ;  * 75th percentile (407 days = 1.15 yrs = 13.3 months) Defined
above ;
            if first.bdrcusip6 then mnablockindx75p=1;
            if diffdayspriordeal75p > windowdaysthreshold75p then mnablockindx75p + 1; *using threshold in
days;
            if firstbdrcusip6 ne 1 and diffdayspriordeal75p > windowdaysthreshold75p then
startnewblockbdrcusip6=1; *using threshold in days;
            if lagmnablockindx75p ne mnablockindx75p and firstbdrcusip6 ne 1 then mnadealinblockindx75p =1;
            if first.bdrcusip6 then mnablockindx75p=1;
            bdrcusip6mnablockindx75p= bdrcusip6||mnablockindx75p;
            keep  dealnosdc  bdrcusip6 lagbdrcusip6 anndate  mnadealindx mnablockindx75p lagmnablockindx75p
                  mnadealinblockindx75p diffdayspriordeal75p      startnewblockbdrcusip6 sameanndate
                  bdrcusip6anndate bdrcusip6mnablockindx75p totdeals  diffdayspriordeal75p
laganndatepriordeal ;
      run;
/* proc means data=sdc1980Mar2019APASPaft1989A mean n min p25 median p75 p90 p95 max; var
diffdayspriordeal75p mnadealinblockindx75p totdeals; run; */
proc sort data=sdc1980Mar2019APASPaft1989A0; by dealnosdc; run;
```

```
proc sort data=sdc1980Mar2019APASPaft1989A; by dealnosdc; run;
data sdc1980Mar2019APASPaft1989A1; merge sdc1980Mar2019APASPaft1989A0(in=a)  sdc1980Mar2019APASPaft1989A;
      if a; by dealnosdc; run;
data serialac.sdc1980Mar2019APASPaft1989A1; set sdc1980Mar2019APASPaft1989A1; run;

*Estimate index within blockindex taking into consideration deals within same date;
data sdc1980Mar2019APASPaft1989A1; set serialac.sdc1980Mar2019APASPaft1989A1; run;
data sdc1980Mar2019anndate ; set sdc1980Mar2019APASPaft1989A1;
                  keep bdrcusip6 bdrcusip6anndate bdrcusip6mnablockindx75p  sameanndate; run;
proc sort data=sdc1980Mar2019anndate nodupkey; by bdrcusip6anndate; run;
proc sort data=sdc1980Mar2019anndate ; by bdrcusip6mnablockindx75p; run;
data bdrcrspcomp19802017anndateblckA; set sdc1980Mar2019anndate;
      by bdrcusip6mnablockindx75p;
      mnadealinblockindx75psametime +1;
      if first.bdrcusip6mnablockindx75p then mnadealinblockindx75psametime = 1;
run;
*get index within block regardless of same time dealnosdcs;
data bdrcrspcomp19802017blck ; set sdc1980Mar2019APASPaft1989A1;
                  keep bdrcusip6 bdrcusip6anndate bdrcusip6mnablockindx75p sameanndate; run;
proc sort data=bdrcrspcomp19802017blck ; by bdrcusip6 anndate; run;
proc sort data=bdrcrspcomp19802017blck ; by bdrcusip6mnablockindx75p; run;
data bdrcrspcomp19802017blckA; set bdrcrspcomp19802017blck;
      by bdrcusip6mnablockindx75p;
      mnadealinblockindx75pall +1;
      if first.bdrcusip6mnablockindx75p then mnadealinblockindx75pall = 1;
run;
*Estimate index for cusip6 taking into consideration deals within same date;
data bdrcrspcomp19802017anndate ; set sdc1980Mar2019APASPaft1989A1;
                  keep bdrcusip6 bdrcusip6anndate  sameanndate; run;
proc sort data=bdrcrspcomp19802017anndate ; by bdrcusip6anndate descending sameanndate; run;
proc sort data=bdrcrspcomp19802017anndate nodupkey; by bdrcusip6anndate ; run;
proc sort data=bdrcrspcomp19802017anndate ; by bdrcusip6 ; run;
data bdrcrspcomp19802017anndateA; set bdrcrspcomp19802017anndate;
      by bdrcusip6;
      lagbdrcusip6 =  lag(bdrcusip6);
      mnadealindxsametime +1;
      if first.bdrcusip6 then mnadealindxsametime = 1; run;
*Merge back this refined index taking into consideration same day dealnosdcs;
*with all information;
data bdrcrspcomp19802017indexes ; set sdc1980Mar2019APASPaft1989A1;
```

```
                  keep  anndate bdrcusip6  bdrcusip6anndate bdrcusip6mnablockindx75p mnadealindx
mnablockindx75p mnadealinblockindx75p sameanndate laganndatepriordeal; run; * 54490 observations and 9
variables.;
proc sort data= bdrcrspcomp19802017indexes; by bdrcusip6anndate  descending sameanndate; run;
*with mnadealinblockindx75psametime  index within block adj for same time;
proc sort data= bdrcrspcomp19802017anndateblckA; by bdrcusip6anndate; run;
*with mnadealinblockindx75psametime  index WITHOUT block adj for same time;
proc sort data= bdrcrspcomp19802017blckA; by bdrcusip6anndate; run;
*with mnadealinblockindx75psametime  index WITHOUT block adj for same time;
proc sort data= bdrcrspcomp19802017anndateA; by bdrcusip6anndate; run;
data bdrcrspcomp19802017anndateindx; merge bdrcrspcomp19802017anndate(in=a) bdrcrspcomp19802017blckA(in=b)
            bdrcrspcomp19802017anndateblckA(in=c) bdrcrspcomp19802017indexes(in=d);
      by bdrcusip6anndate; run;
*get mnadealindxsametime then allocating minimum bdrcusip6mnablockindx75p to ALL the deals taking place in
the same day;
      *first total number of deals and blocks, second at bdrcusip6 level,  and then at bdrcusip6block;
*1-first total number of deals and blocks;
proc means data=bdrcrspcomp19802017anndateindx noprint;
      by bdrcusip6 ;
      var anndate;
      output out= maxnadealindx     (drop=_FREQ_ _TYPE_)
            n(anndate )                          = dealsforbdrcusip6
            max(mnablockindx75p)    = totalmnablockindx75p;  run;
      *dealsforbdrcusip6 totalmnablockindx75p;
*Merge back;
proc sort data=bdrcrspcomp19802017anndateindx;  by bdrcusip6 ; run;
proc sort data=maxnadealindx;       by bdrcusip6 ; run;
data bdrcrspcomp19802017anndateindxA; merge bdrcrspcomp19802017anndateindx maxnadealindx;
      by bdrcusip6 ; run;

*2-at bdrcusip6 level   > get n of deals in each day and get min index [overall and within block] taking
into consideration same date];
proc sort data=bdrcrspcomp19802017anndateindxA nodupkey; by dealnosdc; run; *to ensure no false duplicates;
proc sort data=bdrcrspcomp19802017anndateindxA; by bdrcusip6 anndate; run;
proc means data=bdrcrspcomp19802017anndateindxA noprint;
      by bdrcusip6 anndate;
      var anndate;
      output out= minmnadealindx (drop=_FREQ_ _TYPE_)
            n(anndate) = dealsinday
            min(mnadealindx mnadealinblockindx75pall )= minmnadealindx minmnadealinblockindx75pall;  run;
 *dealsinday minmnadealindx minmnadealinblockindx75pall;   *!!! to be deleted aferwards: mnadealindx
mnadealinblockindx75pall after  not needed in estimations;
```

```
*Merge back;
proc sort data=bdrcrspcomp19802017anndateindxA;       by bdrcusip6 anndate; run;
proc sort data=minmnadealindx;                              by bdrcusip6 anndate; run;
data bdrcrspcomp19802017anndateindxB; merge bdrcrspcomp19802017anndateindxA minmnadealindx;
      by bdrcusip6 anndate; run;

*3-at bdrcusip6block level;
proc sort data= bdrcrspcomp19802017anndateindxB; by bdrcusip6 mnablockindx75p; run;
proc means data=bdrcrspcomp19802017anndateindxB noprint;
      by bdrcusip6 mnablockindx75p;
      var mnadealinblockindx75pall;
      output out= maxbdrcusip6mnablockindx75p  (drop=_FREQ_ _TYPE_)
            max(mnadealinblockindx75pall)= maxmnadealinblockindx75pall;  run; *21180 observations and 5
variables;  *CORRECT FOR WITHIN BLOCK;
            *maxmnadealinblockindx75pall;
*Merge back;
proc sort data=bdrcrspcomp19802017anndateindxB; by bdrcusip6 mnablockindx75p; run;
proc sort data=maxbdrcusip6mnablockindx75p;                                 by bdrcusip6 mnablockindx75p; run;
data bdrcrspcomp19802017anndateindxC; merge bdrcrspcomp19802017anndateindxB maxbdrcusip6mnablockindx75p;
      by bdrcusip6 mnablockindx75p;
      if dealsinday > 1 then concurrentdealsinday = 1; else concurrentdealsinday = 0;
      ain = minmnadealinblockindx75pall; *main var of interest!!! at block level;
      drop   mnadealindx mnadealinblockindx75p mnadealinblockindx75pall; * These indexes not yet corrected
for same date.
            The corrected ones that we end up using are the ones estimated above using the appropriate same
day approach;
run;
*AIN Main variables of interest;
proc means data=bdrcrspcomp19802017anndateindxC n mean min median max ;
      var minmnadealinblockindx75pall  ain  minmnadealindx mnablockindx75p dealsforbdrcusip6
concurrentdealsinday maxmnadealinblockindx75pall dealsinday  ; run;
 data bdrcrspcomp19802017ain401C0 ; set bdrcrspcomp19802017anndateindxC;
      if minmnadealinblockindx75pall = . Then minmnadealinblockindx75pall = 1;
      if ain = . Then ain = 1;
      if minmnadealindx = . Then minmnadealindx = 1;
      if mnablockindx75p = . Then mnablockindx75p = 1;
      if dealsforbdrcusip6 = . Then dealsforbdrcusip6 = 1;
      if maxmnadealinblockindx75pall = . Then maxmnadealinblockindx75pall = 1;
      if concurrentdealsinday = . Then concurrentdealsinday = 0;
      *!!! rename vars for better understanding of names;
            ainoverall        = minmnadealindx;
            totdealsoverall   = dealsforbdrcusip6;
```

```
            ainblocklevel     = ain;
            totdealsinblock = maxmnadealinblockindx75pall;
run; *21 variables   1989-2019 ;
* I estimate Intensity later at the 5-yr rolling window level;
data bdrcrspcomp19802017ain401C; set bdrcrspcomp19802017ain401C0;
 if sameanndate =1 then delete;
      if dealsinday > 1 then concurrentdeals=1; else concurrentdeals=0;
run; *20 variables.;
/*proc means data=bdrcrspcomp19802017ain401C n mean min median max ;
      var ainoverall  totdealsoverall ainblocklevel  totdealsinblock minmnadealinblockindx75pall  ain
minmnadealindx
mnablockindx75p dealsforbdrcusip6 concurrentdealsinday maxmnadealinblockindx75pall dealsinday    ; run; */
data serialac.bdrcrspcomp19802017ain401C; set bdrcrspcomp19802017ain401C; run; *21 variables;
*merge back;
data sdc1980Mar2019APASPaft1989A1; set serialac.sdc1980Mar2019APASPaft1989A1; run;
data bdrcrspcomp19802017ain401C;                set serialac.bdrcrspcomp19802017ain401C; run;
proc sort data=sdc1980Mar2019APASPaft1989A1;    by bdrcusip6 anndate; run;
proc sort data=bdrcrspcomp19802017ain401C;      by bdrcusip6 anndate; run;
data sdc1980Mar2019APASPaft1989A2; merge sdc1980Mar2019APASPaft1989A1(in=a) bdrcrspcomp19802017ain401C;
      if a; by bdrcusip6 anndate;
      annmonth = month(anndate);
run;
/* proc means data=sdc1980Mar2019APASPaft1989A1 mean; var timesincelastdeal  diffdayspriordeal75p; run;*/
*Imp File on MnA Deals;
OPTIONS nofmterr;
data serialac.sdc1980Mar2019APASPaft1989A2; set  sdc1980Mar2019APASPaft1989A2; run;

/* Key Vars at the whole overall sample 1990-2019 level
ainoverall              used the min index dealno adj for concurrent deals: TO BE USED AS INDX y
      Taking in consideration ALL BLOCKs, hence not = to AIN
totdealsoverall         Total number of deals for each acquiree in all blocks
ainblocklevel           used the min index dealno adj for concurrent deals: TO BE USED AS INDX y
       Adjusted for correct WITHIN BLOCK
totdealsinblock         Total number of deals in the block
concurrentdeals         indicator for more than one deal in same day
mnablockindx75p         index for block index
firstbdrcusip6          Indicator for First Deal by bdrcusip6 (it is also indicated in case there are
concurrent deals in such first deal)
*/

/*******************************/
```

***0.2 Ex ante Serial Acquirer Classification based on 5-year rolling windows ;**

```
/* Ex ante Serial Acquirer Classification based on 5-year rolling windows*/
/* SAS
Long Macro used to estimate the required variables for each 5-year rolling window
Main Vars at bdr level for each rolling window:
            bdr5totdeals bdr5totdealsge1  bdr5totalnblocks  bdr5totalnblocksge1          bdr5compactnessrw
bdr5compactnessrwge1
            bdr5mntotdealsinblck  bdr5mntotdealsinblckge1   bdr5mxainblocklevel     bdr5mxainblocklevelge1
      bdr5mnconcurrentdeals   bdr5mnconcurrentdealsge1
            pprtn5ge1totdealsrw  pprtn5ge1totblocksrw
*at bidder level
Macro gets the information for ALL the deals and also for ONLY large deals within each 5-year rolling
window sub-sample
*/

OPTIONS nofmterr;
data sdc1980mar2019flxrwb; set serialac.sdc1980mar2019flxrwb; run; * Initial sample deal level;
data sdc19952001forrllngwndw; set sdc1980mar2019flxrwb; * Reduced sample;
             if annyear => 1995 and annyear <= 2001; run;
data serialac.sdc19952001forrllngwndw; set sdc19952001forrllngwndw; run; * Initial  Sub-sample 1995-2001;
OPTIONS nofmterr;
data sdc19952001forrllngwndw; set serialac.sdc19952001forrllngwndw; run; * Initial  Sub-sample 1995-2001
deal level;

*Clean of duplicated observations and deals out of sample;
data aaafew; set sdc19891996flxrwb;
             if annyear => 1989 and annyear <= 2018;
                  bdrcusip6anndate= bdrcusip6||anndate;
                  keep annyear anndate bdrcusip6 dealnosdc bdrcusip6anndate acqname trgname drelsizege1mm;
      run;
            proc sort data = aaafew nodupkey ; by bdrcusip6 dealnosdc; run;
            proc sort data = aaafew  ; by bdrcusip6 anndate; run;
data aaafew0; set aaafew; run;
* START MACRO  to get 5-year rolling window subsamples;
%%macro CreateRWSamplesallge1mm(firstyrrw, firstyrindex, lastyrindex, addyr);
      %do indexiter=&firstyrindex %to &lastyrindex;
      %let yrendrw_     = &indexiter ;                      * yrendrw_  yr in which Rolling window ends;
      %let yrstrtrw_    = %eval(&indexiter - &addyr); * yrstrtrw_ yr in which Rolling window starts;
      %let yrm3_        = &yrstrtrw_;                                 * yrstrtrw_ yr in which Rolling
window starts, req 3 clean yrs NOT req anymore after Jul2021;
/*BASED ON ACQ-date observations, instead of unique deal observations */
                  data aaafew; set aaafew0;
```

```
                        if annyear => &yrm3_ and annyear <= &yrendrw_;
                      run;
                proc sql;
                            create table aaafew2 as
                            select distinct  *
                            , min(annyear) as firstannyear
                            , count(annyear) as totdeals
                            from aaafew
                            group by  bdrcusip6
                            order by bdrcusip6 ;
                      quit;
      * I need to get correct laganndate to get correct  diffdayspriordeal75p from laganndatepriordeal but
using prior day> Done below;
                      data aaafew2blck1; set aaafew2; keep bdrcusip6anndate  bdrcusip6 anndate; run;
                            proc sort data=aaafew2blck1 ;by bdrcusip6anndate; run;
                            proc sort data=aaafew2blck1 nodupkey; by bdrcusip6anndate; run;*To delete
concurrent deals to get first deals in a given date;
                            proc sort data=aaafew2blck1 ; by bdrcusip6 anndate; run;
                      data aaafew2blck2; set aaafew2blck1;
                            by bdrcusip6;
                                  format laganndatepriordeal   yymmdd9.;
                                  lagbdrcusip6 =  lag(bdrcusip6);
                                  laganndatepriordeal =  lag(anndate);  *to be corrected later on taking
into consideration same-day deals;
                                  if first.bdrcusip6 then firstbdrcusip6=1;
                                  if first.bdrcusip6 then laganndatepriordeal=. ;
                                  diffdayspriordeal75p = anndate - laganndatepriordeal ; *in days; *to be
corrected later on taking into consideration same-day deals;
                                  if first.bdrcusip6 then lagbdrcusip6=. ;
                                  if first.bdrcusip6 then diffdayspriordeal75p = .;
                                  if first.bdrcusip6 then timesincelastdeal=. ;
                                  timesincelastdeal = diffdayspriordeal75p;
                                        run;
                            proc sort data=aaafew2 ; by bdrcusip6 anndate; run;
                            proc sort data=aaafew2blck2 ; by bdrcusip6 anndate; run;
                      data aaafew3; merge aaafew2(in=a)  aaafew2blck2;
                            if a; by bdrcusip6 anndate; run;
                *vars to get: mnadealindx mnadealinblockindx75p  mnablockindx75p ;
                      proc sort data=aaafew3; by bdrcusip6 anndate; run;
                      data aaafew4; set aaafew3;
                            by bdrcusip6;
                                  format laganndatepriordeal   yymmdd9.;
```

```
                                    lagbdrcusip6 =  lag(bdrcusip6);
                                    if first.bdrcusip6 then firstbdrcusip6=1;
                                    if first.bdrcusip6 then mnablockindx75p=1;
                                    if first.bdrcusip6 then mnadealinblockindx75p=1;
                                    mnadealindx +1;
                                    mnadealinblockindx75p +1;
                                    if first.bdrcusip6 then mnadealinblockindx75p=1;
                                    lagmnablockindx75p =  lag(mnablockindx75p);
                                    laganndatepriordeal =  lag(anndate);
                                    if laganndatepriordeal =  anndate and lagbdrcusip6 = bdrcusip6 then
sameanndate =1;
                                    windowdaysthreshold75p= 407   ;  * 75th percentile (407 days = 1.15 yrs
= 13.3 months) Defined above ;
                                    if first.bdrcusip6 then mnablockindx75p=1;
                                    if diffdayspriordeal75p > windowdaysthreshold75p then mnablockindx75p +
1; *using threshold in days;
                                    if firstbdrcusip6 ne 1 and diffdayspriordeal75p >
windowdaysthreshold75p then startnewblockbdrcusip6=1; *using threshold in days;
                                    if lagmnablockindx75p ne mnablockindx75p and firstbdrcusip6 ne 1 then
mnadealinblockindx75p =1;
                                    if first.bdrcusip6 then mnablockindx75p=1;
                                    bdrcusip6mnablockindx75p= bdrcusip6||mnablockindx75p;
                                    keep  dealnosdc  bdrcusip6 lagbdrcusip6 anndate  mnadealindx
mnablockindx75p lagmnablockindx75p
                                          mnadealinblockindx75p diffdayspriordeal75p
      startnewblockbdrcusip6 sameanndate
                                          bdrcusip6anndate bdrcusip6mnablockindx75p totdeals
diffdayspriordeal75p laganndatepriordeal ;
                              run;
                        proc sort data=aaafew4; by dealnosdc; run;
                        proc sort data=aaafew3; by dealnosdc; run;
                        data aaafew5; merge aaafew3(in=a)  aaafew4;
                              if a; by dealnosdc; run;
                        data aaafew5anndateblck1 ; set aaafew5;
                                          keep bdrcusip6 bdrcusip6anndate bdrcusip6mnablockindx75p
sameanndate; run;
                        proc sort data=aaafew5anndateblck1 nodupkey; by bdrcusip6anndate; run;
                        proc sort data=aaafew5anndateblck1 ; by bdrcusip6mnablockindx75p; run;
                        data aaafew5anndateblck2; set aaafew5anndateblck1;
                              by bdrcusip6mnablockindx75p;
                              mnadealinblockindx75psametime +1;
                              if first.bdrcusip6mnablockindx75p then mnadealinblockindx75psametime = 1;
```

```
                        run;
                        *get index witin block regardless of same time dealnosdcs;
                        data aaafew5blck3 ; set aaafew5;
                                        keep bdrcusip6 bdrcusip6anndate bdrcusip6mnablockindx75p
sameanndate anndate; run;
                        proc sort data=aaafew5blck3 ; by bdrcusip6 anndate; run;
                        proc sort data=aaafew5blck3 ; by bdrcusip6mnablockindx75p; run;
                        data aaafew5blck4; set aaafew5blck3;
                              by bdrcusip6mnablockindx75p;
                              mnadealinblockindx75pall +1;
                              if first.bdrcusip6mnablockindx75p then mnadealinblockindx75pall = 1;
                        run;
                        proc sort data=aaafew5blck4 ; by bdrcusip6 anndate; run; * 2 variables;
                        *Estimate index for cusip6 taking into consideration deals within same date;
                        data aaafew5anndate ; set aaafew5;
                                        keep bdrcusip6 bdrcusip6anndate  sameanndate; run;
                        proc sort data=aaafew5anndate ; by bdrcusip6anndate descending sameanndate; run;
                        proc sort data=aaafew5anndate nodupkey; by bdrcusip6anndate ; run; * 3 variables;
                        proc sort data=aaafew5anndate ; by bdrcusip6 ; run;
                        data aaafew5anndate1; set aaafew5anndate;
                              by bdrcusip6;
                              lagbdrcusip6 =  lag(bdrcusip6);
                              mnadealindxsametime +1;
                              if first.bdrcusip6 then mnadealindxsametime = 1;
                        run;*5 variables;

                        *with all information;
                        data aaafew5indexes ; set aaafew5;
                                        keep  anndate bdrcusip6  bdrcusip6anndate
bdrcusip6mnablockindx75p mnadealindx  mnablockindx75p mnadealinblockindx75p sameanndate
laganndatepriordeal; run; * 54490 observations and 9 variables.;
                        proc sort data= aaafew5indexes; by bdrcusip6anndate  descending sameanndate; run;
9 variables.;
                        *with mnadealinblockindx75psametime  index within block adj for same time;
                        proc sort data= aaafew5anndateblck2; by bdrcusip6anndate; run;* 4 variables;
                        *with mnadealinblockindx75psametime  index WITHOUT block adj for same time;
                        proc sort data= aaafew5blck4; by bdrcusip6anndate; run;*  5 variables.;
                        *with mnadealinblockindx75psametime  index WITHOUT block adj for same time;
                        proc sort data= aaafew5anndate1; by bdrcusip6anndate; run;* 5 variables;

                        data aaafew5anndateindx; merge aaafew5anndate1(in=a) aaafew5blck4(in=b)
                                    aaafew5anndateblck2(in=c) aaafew5indexes(in=d);
```

```
                        by bdrcusip6anndate; run;*10 variables;

                    *get mnadealindxsametime then allocating minimum bdrcusip6mnablockindx75p to ALL
the deals taking place in the same day;
                        *first total number of deals and blocks, second at bdrcusip6 level,  and then
at bdrcusip6block;
                    *1-first total number of deals and blocks;
                    proc means data=aaafew5anndateindx noprint;
                        by bdrcusip6 ;
                        var anndate;
                        output out= maxnadealindx     (drop=_FREQ_ _TYPE_)
                            n(anndate )                          = dealsforbdrcusip6
                            max(mnablockindx75p)    = totalmnablockindx75p;  run; *9500
observations and 5 variables;
                        *dealsforbdrcusip6 totalmnablockindx75p;
                    *Merge back;
                    proc sort data=aaafew5anndateindx;  by bdrcusip6 ; run;
                    proc sort data=maxnadealindx;       by bdrcusip6 ; run;
                    data aaafew5anndateindxA; merge aaafew5anndateindx maxnadealindx;
                        by bdrcusip6 ; run; *deal level 13 variables;

                    *2-at bdrcusip6 level   > get n of deals in each day and get min index [overall and
within block] taking into consideration same date];
                    proc sort data=aaafew5anndateindxA; by bdrcusip6 anndate; run;
                    proc means data=aaafew5anndateindxA noprint;
                        by bdrcusip6 anndate;
                        var anndate;
                        output out= minmnadealindx (drop=_FREQ_ _TYPE_)
                            n(anndate) = dealsinday
                            min(mnadealindx mnadealinblockindx75p laganndatepriordeal)=
minmnadealindx minmnadealinblockindx75pall adjlaganndatepriordeal;  run;
                         *dealsinday minmnadealindx minmnadealinblockindx75pall;   *!!! to be deleted
aferwards: mnadealindx mnadealinblockindx75pall;
                    *Merge back;
                    proc sort data=aaafew5anndateindxA;       by bdrcusip6 anndate; run;
                    proc sort data=minmnadealindx;                                  by bdrcusip6 anndate;
run;
                    data aaafew5anndateindxB0; merge aaafew5anndateindxA minmnadealindx;
                        by bdrcusip6 anndate;
                    run;
                     data aaafew5anndateindxB; set aaafew5anndateindxB0;
                     by bdrcusip6;
```

```
                                    *correcting variables now taking into consideration same-day deals;
                                    diffdayspriordeal75p = anndate - adjlaganndatepriordeal ; *in days;
*in days; *to be corrected later on taking into consideration same-day deals;
                                    timesincelastdeal = diffdayspriordeal75p; *in days; *to be corrected
later on taking into consideration same-day deals;
                                    if minmnadealindx = 1 then lagbdrcusip6=. ;
                                    if minmnadealindx = 1 then diffdayspriordeal75p = .;
                                    if minmnadealindx = 1 then timesincelastdeal=. ;
                                    if minmnadealindx = 1 then adjlaganndatepriordeal=. ;
                                    if minmnadealindx = 1 then lagmnablockindx75p=. ;
                        run;*deal level 6 variables;

                        *3-at bdrcusip6block level;  * note, it does not matter if I use
mnadealinblockindx75pall or mnadealinblockindx75p when estimating the max;
                        proc sort data= aaafew5anndateindxB; by bdrcusip6 mnablockindx75p; run;
                        proc means data=aaafew5anndateindxB noprint;
                              by bdrcusip6 mnablockindx75p;
                              var mnadealinblockindx75p;
                              output out= maxbdrcusip6mnablockindx75p  (drop=_FREQ_ _TYPE_)
                                    max(mnadealinblockindx75pall)= maxmnadealinblockindx75pall;  run; *5
variables;  *This is the CORRECT FOR WITHIN BLOCK;
                                    *maxmnadealinblockindx75pall;
                        *Merge back;
                        proc sort data=aaafew5anndateindxB; by bdrcusip6 mnablockindx75p; run;
                        proc sort data=maxbdrcusip6mnablockindx75p;                                 by
bdrcusip6 mnablockindx75p; run;
                        *I detect a serial acquirer which in his first ever acquisition it acquires 17
different targets, then the bidder continues acquiring up to 56 targets in the sample bdrcusip6: 023586.
day of first 17 acq: 23Dec1993;
                        data aaafew5anndateindxC; merge aaafew5anndateindxB maxbdrcusip6mnablockindx75p;
                              by bdrcusip6 mnablockindx75p;
                              if dealsinday > 1 then concurrentdealsinday = 1; else concurrentdealsinday =
0;
                              ain = minmnadealinblockindx75pall; *main var of interest!!! at block level;
                              drop   mnadealindx mnadealinblockindx75p mnadealinblockindx75pall; * indexes
not yer corrected for same date, yet used above to estimate the corrected ones;
                        run; *deal level and 21 variables;
                         data aaafew5anndateindxainC0 ; set aaafew5anndateindxC;
                        *create vars with value=1 if missing value. Recovered 200 more observations :-)  ;
                                    *minmnadealinblockindx75pall  ain minmnadealindx mnablockindx75p
dealsforbdrcusip6 concurrentdealsinday maxmnadealinblockindx75pall      ;
                         *sum maxmnadealinblockindx75pall maxminmnadealindx;
```

```
                        if minmnadealinblockindx75pall = . Then minmnadealinblockindx75pall = 1;
                        if ain = . Then ain = 1;
                        if minmnadealindx = . Then minmnadealindx = 1;
                        if mnablockindx75p = . Then mnablockindx75p = 1;
                        if dealsforbdrcusip6 = . Then dealsforbdrcusip6 = 1;
                        if maxmnadealinblockindx75pall = . Then maxmnadealinblockindx75pall = 1;
                        if concurrentdealsinday = . Then concurrentdealsinday = 0;
                        *!!! rename vars for better understanding of names;
                              ainoverall        = minmnadealindx;
                              totdealsoverall   = dealsforbdrcusip6;
                              ainblocklevel     = ain;
                              totdealsinblock = maxmnadealinblockindx75pall; *need to estimate number
below after adj fo same-day deals: totcum_suminblock;
                  run; *deal level obs with 21 variables   1989-2019 ;
                  * I estimate Intensity later below, around lines 2,200 following SAS code:
20161018_Serial Acquirers Update.sas around lines 4,700 ;
                  data aaafew5anndateindxainC; set aaafew5anndateindxainC0;
                   if sameanndate =1 then delete;
                        if dealsinday > 1 then concurrentdeals=1; else concurrentdeals=0;
                  run;
                  proc sort data=aaafew5;       by bdrcusip6 anndate; run;
                  proc sort data=aaafew5anndateindxainC;    by bdrcusip6 anndate; run;
                  data aaafew5statsA1; merge aaafew5(in=a) aaafew5anndateindxainC;
                        if a; by bdrcusip6 anndate;
                  run;
                  proc sort data=aaafew5statsA1;      by bdrcusip6mnablockindx75p; run;
                  *estimate total deals in each block;
                  *Note that ainblocklevel might be lower if there are more than 1 deal in same day;
                  proc means data=aaafew5statsA1 noprint;
                        by bdrcusip6mnablockindx75p ;
                        var anndate;
                        output out= maxnadealindxinblock    (drop=_FREQ_ _TYPE_)
                              n(anndate )             =     totcum_suminblock
                              min(anndate)            =     firstanndatebdrblockrw
                              max(anndate)            =     lastanndatebdrblockrw;
                  ;  run;
                  proc sort data=aaafew5statsA1;      by bdrcusip6 anndate; run;
                        data aaafew5statsA2; merge aaafew5statsA1(in=a) maxnadealindxinblock;
                        if a; by bdrcusip6mnablockindx75p;
                  run;
                  proc sort data=aaafew5statsA2;      by bdrcusip6 anndate; run;
                  data aaafew5A2cleanstartA; set aaafew5statsA2; run;
```

```
                                        data  aaafew5A2bybdrblocktot; set  aaafew5A2cleanstartA;
                                              format firstanndatebdrblockrw mmddyy9.;format
lastanndatebdrblockrw mmddyy9.;
                                              totdealsinblck = totcum_suminblock;
                                              bdrblockduration = lastanndatebdrblockrw-
firstanndatebdrblockrw+1;
                                              bdrblockintensity = totcum_suminblock/bdrblockduration;
                                              if bdrblockduration = 1 and totcum_suminblock = 1 then
bdrblockintensityadj = 0; else      bdrblockintensityadj = bdrblockintensity;
                                              drop totcum_suminblock;
                                        run;
                                        *get max for each bdrcusip6;
                                        proc sql;
                                          create table aaafew5A2bybdrblock as
                                          select distinct
                                               bdrcusip6,
                                                                    max(bdrblockduration)               as
      maxbdrblockduration,
                                                                    mean(bdrblockduration)              as
      meanbdrblockduration,
                                                                    median(bdrblockduration)            as
      mdnbdrblockduration,
                                                                    min(bdrblockduration)               as
      minbdrblockduration,
                                                         max(bdrblockintensityadj)          as
maxbdrblockintensityadj,
                                                         mean(bdrblockintensityadj)         as
mnbdrblockintensityadj
                                          from   aaafew5A2bybdrblocktot
                                          group by  bdrcusip6;
                                        quit; * 3 columns;
                                        *merge max based on "by mna19851996bybdrblocktot" to the bdruirer
level;
                                        proc sql;
                                          create table aaafew5A2bybdrvarsforclstA as
                                          select *
                                          from   aaafew5A2bybdrblocktot, aaafew5A2bybdrblock
                                             where
aaafew5A2bybdrblocktot.bdrcusip6=aaafew5A2bybdrblock.bdrcusip6
                                          order by aaafew5A2bybdrblocktot.bdrcusip6;
                                        quit;
                                        data aaafew5A2bybdrvarsforclstB; set aaafew5A2bybdrvarsforclstA;
```

```
                                                compactnessrw = maxbdrblockintensityadj;
                                                compactnessmnrw = mnbdrblockintensityadj;
                                                totalnblocks = totalmnablockindx75p;
                                          run;
                  proc sort data=aaafew5A2bybdrvarsforclstB;      by bdrcusip6 anndate; run;
                  proc means data=aaafew5A2bybdrvarsforclstB n mean min p1 p5 p10 p25 median p75 p90 p95
p99 max; var totdeals totdealsinblck totalmnablockindx75p  compactnessrw ; run;
                  data selvarsbybdrall8918; set aaafew5A2bybdrvarsforclstB;
                              keep acqname trgname drelsizege1mm bdrcusip6 anndate annyear dealnosdc
totdeals totdealsinblck totalnblocks compactnessrw ainoverall ainblocklevel concurrentdeals;
      run;

/* Now after having estimated the statistics using the whole sample, the code below is for the macros that
I will use for the rolling windows,
                  for which I only need much fewer variables:
                  1- total number of deals in sample     overall: totdeals   and in block:
totdealsincrrntblock,
                  2- total number of blocks   totalmnablockindx75p, and
                  3- compactness (intensity)  compactnessrw  (using max of all compactness for, if ,
various blocks)

                  other useful variables adjusted for deals in same day:
                  ainoverall  AND ainblocklevel */

/* Also get total deals, total blocks and intensity as of prior year
And also get stats now ONLY for drelsizege1mm = 1 SEE Code below which emulates code above but now with
smaller sample for economically significant deals (drelsizege1mm=1)
*/
/*****************************************/
/*  ESTIMATE SERIAL ACQUIRER STATISTICS, adjusted for concurrent deals in same date, FOR SIGNIFICANTLY
LARGE SIZE DEALS (large)
      Main Vars to be estimated for drelsizege1mm=1 deals:
                  acqname trgname drelsizege1mm bdrcusip6 anndate annyear dealnosdc
                  totdealsge1       totdealsinblckge1 totalnblocksge1   compactnessrwge1  ainoverallge1
      ainblocklevelge1  concurrentdealsge1
                              */
                  /*    data sdc1980mar2019flxrwb; set serialac.sdc1980mar2019flxrwb; run;*/
                        *Clean of duplicated observations and deals out of sample;
                        data aaafew; set aaafew0;
                         if annyear => &yrm3_ and annyear <= &yrendrw_;
                              if drelsizege1mm = 1;
                        run;
```

```
proc sort data = aaafew nodupkey ; by bdrcusip6 dealnosdc; run; *20283 observations
and 693 variables;

proc sql;
      create table aaafew2 as
      select distinct  *
      , min(annyear) as firstannyear
      , count(annyear) as totdeals
      from aaafew
      group by  bdrcusip6
      order by bdrcusip6 ;
quit;

* I get correct laganndate to get correct  diffdayspriordeal75p from
laganndatepriordeal but using prior day in code below after required adjustment;
data aaafew2blck1; set aaafew2; keep bdrcusip6anndate  bdrcusip6 anndate; run;
      proc sort data=aaafew2blck1 ;by bdrcusip6anndate; run;
      proc sort data=aaafew2blck1 nodupkey; by bdrcusip6anndate; run; *to delete
concurrent deals to get first deals in a given date;
      proc sort data=aaafew2blck1 ; by bdrcusip6 anndate; run;
data aaafew2blck2; set aaafew2blck1;
      by bdrcusip6;
            format laganndatepriordeal   yymmdd9.;
            lagbdrcusip6 =  lag(bdrcusip6);
            laganndatepriordeal =  lag(anndate);  *to be corrected later on taking
into consideration same-day deals;
            if first.bdrcusip6 then firstbdrcusip6=1;
            if first.bdrcusip6 then laganndatepriordeal=. ;
            diffdayspriordeal75p = anndate - laganndatepriordeal ; *in days; *to be
corrected later on taking into consideration same-day deals;
            if first.bdrcusip6 then lagbdrcusip6=. ;
            if first.bdrcusip6 then diffdayspriordeal75p = .;
            if first.bdrcusip6 then timesincelastdeal=. ;
            timesincelastdeal = diffdayspriordeal75p;
                  run;
      proc sort data=aaafew2 ; by bdrcusip6 anndate; run;
      proc sort data=aaafew2blck2 ; by bdrcusip6 anndate; run;
data aaafew3; merge aaafew2(in=a)  aaafew2blck2;
      if a; by bdrcusip6 anndate; run;*12 variables.;
*vars to get: mnadealindx mnadealinblockindx75p  mnablockindx75p ;
proc sort data=aaafew3; by bdrcusip6 anndate; run;
data aaafew4; set aaafew3;
```

```
                              by bdrcusip6;
                                    format laganndatepriordeal   yymmdd9.;
                                    lagbdrcusip6 =  lag(bdrcusip6);
                                    if first.bdrcusip6 then firstbdrcusip6=1;
                                    if first.bdrcusip6 then mnablockindx75p=1;
                                    if first.bdrcusip6 then mnadealinblockindx75p=1;
                                    mnadealindx +1;
                                    mnadealinblockindx75p +1;
                                    if first.bdrcusip6 then mnadealindx=1;
                                    if first.bdrcusip6 then mnadealinblockindx75p=1;
                                    lagmnablockindx75p =  lag(mnablockindx75p);
                                    laganndatepriordeal =  lag(anndate);
                                    if laganndatepriordeal =  anndate and lagbdrcusip6 = bdrcusip6 then
sameanndate =1;
                                    /*windowdaysthreshold75p=     407   ;  * FOR ALL DEALS: 75th percentile
(407 days = 1.15 yrs = 13.3 months) Defined above ; */
                     windowdaysthreshold75p= 826   ;  * FOR Large DEALS: 75th percentile (826 days = 2.25 yrs
= 27.53 months) Defined above ;
                                    if first.bdrcusip6 then mnablockindx75p=1;
                                    if diffdayspriordeal75p > windowdaysthreshold75p then mnablockindx75p +
1; *using threshold in days;
                                    if firstbdrcusip6 ne 1 and diffdayspriordeal75p >
windowdaysthreshold75p then startnewblockbdrcusip6=1; *using threshold in days;
                                    if lagmnablockindx75p ne mnablockindx75p and firstbdrcusip6 ne 1 then
mnadealinblockindx75p =1;
                                    if first.bdrcusip6 then mnablockindx75p=1;
                                    bdrcusip6mnablockindx75p= bdrcusip6||mnablockindx75p;
                                    keep  dealnosdc  bdrcusip6 lagbdrcusip6 anndate  mnadealindx
mnablockindx75p lagmnablockindx75p
                                          mnadealinblockindx75p diffdayspriordeal75p
      startnewblockbdrcusip6 sameanndate
                                          bdrcusip6anndate bdrcusip6mnablockindx75p totdeals
diffdayspriordeal75p laganndatepriordeal ;
                               run;
                         proc sort data=aaafew4; by dealnosdc; run;
                         proc sort data=aaafew3; by dealnosdc; run;
                         data aaafew5; merge aaafew3(in=a)  aaafew4;
                               if a; by dealnosdc; run;
                         data aaafew5anndateblck1 ; set aaafew5;
                                          keep bdrcusip6 bdrcusip6anndate bdrcusip6mnablockindx75p
sameanndate; run;
                         proc sort data=aaafew5anndateblck1 nodupkey; by bdrcusip6anndate; run;
```

```
                        proc sort data=aaafew5anndateblck1 ; by bdrcusip6mnablockindx75p; run;
                        data aaafew5anndateblck2; set aaafew5anndateblck1;
                              by bdrcusip6mnablockindx75p;
                              mnadealinblockindx75psametime +1;
                              if first.bdrcusip6mnablockindx75p then mnadealinblockindx75psametime = 1;
                        run;
                        *get index within block regardless of same time dealnosdcs;
                        data aaafew5blck3 ; set aaafew5;
                                          keep bdrcusip6 bdrcusip6anndate bdrcusip6mnablockindx75p
sameanndate anndate; run;
                        proc sort data=aaafew5blck3 ; by bdrcusip6 anndate; run;
                        proc sort data=aaafew5blck3 ; by bdrcusip6mnablockindx75p; run; *2 variables;
                        data aaafew5blck4; set aaafew5blck3;
                              by bdrcusip6mnablockindx75p;
                              mnadealinblockindx75pall +1;
                              if first.bdrcusip6mnablockindx75p then mnadealinblockindx75pall = 1;
                        run;
                        proc sort data=aaafew5blck4 ; by bdrcusip6 anndate; run;
                        *Estimate index for cusip6 taking into consideration deals within same date;
                        data aaafew5anndate ; set aaafew5;
                                          keep bdrcusip6 bdrcusip6anndate  sameanndate; run;
                        proc sort data=aaafew5anndate ; by bdrcusip6anndate descending sameanndate; run;
                        proc sort data=aaafew5anndate nodupkey; by bdrcusip6anndate ; run;
                        proc sort data=aaafew5anndate ; by bdrcusip6 ; run;
                        data aaafew5anndate1; set aaafew5anndate;
                              by bdrcusip6;
                              lagbdrcusip6 =  lag(bdrcusip6);
                              mnadealindxsametime +1;
                              if first.bdrcusip6 then mnadealindxsametime = 1;
                        run;* 5 variables;
                        *with all information;
                        data aaafew5indexes ; set aaafew5;
                                          keep  anndate bdrcusip6  bdrcusip6anndate
bdrcusip6mnablockindx75p mnadealindx  mnablockindx75p mnadealinblockindx75p sameanndate
laganndatepriordeal; run;
                        proc sort data= aaafew5indexes; by bdrcusip6anndate  descending sameanndate; run;
                        *with mnadealinblockindx75psametime  index within block adj for same time;
                        proc sort data= aaafew5anndateblck2; by bdrcusip6anndate; run;
                        *with mnadealinblockindx75psametime  index WITHOUT block adj for same time;
                        proc sort data= aaafew5blck4; by bdrcusip6anndate; run;
                        *with mnadealinblockindx75psametime  index WITHOUT block adj for same time;
                        proc sort data= aaafew5anndate1; by bdrcusip6anndate; run;
```

```
                        data aaafew5anndateindx; merge aaafew5anndate1(in=a) aaafew5blck4(in=b)
                                    aaafew5anndateblck2(in=c) aaafew5indexes(in=d);
                              by bdrcusip6anndate; run;
                        *get mnadealindxsametime then allocating minimum bdrcusip6mnablockindx75p to ALL
the deals taking place in the same day;
                              *first total number of deals and blocks, second at bdrcusip6 level,  and then
at bdrcusip6block;
                        *1-first total number of deals and blocks;
                        proc means data=aaafew5anndateindx noprint;
                              by bdrcusip6 ;
                              var anndate;
                              output out= maxnadealindx     (drop=_FREQ_ _TYPE_)
                                    n(anndate )                         = dealsforbdrcusip6
                                    max(mnablockindx75p)    = totalmnablockindx75p;  run;
                              *dealsforbdrcusip6 totalmnablockindx75p;
                        *Merge back;
                        proc sort data=aaafew5anndateindx;  by bdrcusip6 ; run;
                        proc sort data=maxnadealindx;       by bdrcusip6 ; run;
                        data aaafew5anndateindxA; merge aaafew5anndateindx maxnadealindx;
                              by bdrcusip6 ; run;*13 variables;
                        *2-at bdrcusip6 level   > get n of deals in each day and get min index [overall and
within block] taking into consideration same date];
                        proc sort data=aaafew5anndateindxA; by bdrcusip6 anndate; run;
                        proc means data=aaafew5anndateindxA noprint;
                              by bdrcusip6 anndate;
                              var anndate;
                              output out= minmnadealindx (drop=_FREQ_ _TYPE_)
                                    n(anndate) = dealsinday
                                    min(mnadealindx mnadealinblockindx75p laganndatepriordeal)=
minmnadealindx minmnadealinblockindx75pall adjlaganndatepriordeal;  run;
                                     *dealsinday minmnadealindx minmnadealinblockindx75pall;   *!!! to be deleted
afterwards: mnadealindx mnadealinblockindx75pall;
                        *Merge back;
                        proc sort data=aaafew5anndateindxA;       by bdrcusip6 anndate; run;
                        proc sort data=minmnadealindx;                                  by bdrcusip6 anndate;
run;
                        data aaafew5anndateindxB0; merge aaafew5anndateindxA minmnadealindx;
                              by bdrcusip6 anndate;                       run;
                         data aaafew5anndateindxB; set aaafew5anndateindxB0;
                         by bdrcusip6;
                                    *correcting variables now taking into consideration same-day deals;
```

```
                                    diffdayspriordeal75p = anndate - adjlaganndatepriordeal ; *in days;
*in days; *to be corrected later on taking into consideration same-day deals;
                                    timesincelastdeal = diffdayspriordeal75p; *in days; *to be corrected
later on taking into consideration same-day deals;
                                    if minmnadealindx = 1 then lagbdrcusip6=. ;
                                    if minmnadealindx = 1 then diffdayspriordeal75p = .;
                                    if minmnadealindx = 1 then timesincelastdeal=. ;
                                    if minmnadealindx = 1 then adjlaganndatepriordeal=. ;
                                    if minmnadealindx = 1 then lagmnablockindx75p=. ;
                        run;* 6 variables;
                        *3-at bdrcusip6block level;  * note, it does not matter if I use
mnadealinblockindx75pall or mnadealinblockindx75p when estimating the max;
                        proc sort data= aaafew5anndateindxB; by bdrcusip6 mnablockindx75p; run;
                        proc means data=aaafew5anndateindxB noprint;
                              by bdrcusip6 mnablockindx75p;
                              var mnadealinblockindx75p;
                              output out= maxbdrcusip6mnablockindx75p  (drop=_FREQ_ _TYPE_)
                                    max(mnadealinblockindx75pall)= maxmnadealinblockindx75pall;  run;
*CORRECTED now FOR WITHIN BLOCK;
                    *Merge back;
                        proc sort data=aaafew5anndateindxB; by bdrcusip6 mnablockindx75p; run;
                        proc sort data=maxbdrcusip6mnablockindx75p;                                  by
bdrcusip6 mnablockindx75p; run;
                        data aaafew5anndateindxC; merge aaafew5anndateindxB maxbdrcusip6mnablockindx75p;
                              by bdrcusip6 mnablockindx75p;
                              if dealsinday > 1 then concurrentdealsinday = 1; else concurrentdealsinday =
0;
                              ain = minmnadealinblockindx75pall; *main var of interest!!! at block level;
                              drop   mnadealindx mnadealinblockindx75p mnadealinblockindx75pall; * indexes
not yet corrected for same date, yet used above to estimate the corrected ones;
                        run;
                         data aaafew5anndateindxainC0 ; set aaafew5anndateindxC;
                              if minmnadealinblockindx75pall = . Then minmnadealinblockindx75pall = 1;
                              if ain = . Then ain = 1;
                              if minmnadealindx = . Then minmnadealindx = 1;
                              if mnablockindx75p = . Then mnablockindx75p = 1;
                              if dealsforbdrcusip6 = . Then dealsforbdrcusip6 = 1;
                              if maxmnadealinblockindx75pall = . Then maxmnadealinblockindx75pall = 1;
                              if concurrentdealsinday = . Then concurrentdealsinday = 0;
                              *!!! rename vars for better understanding of names;
                                    ainoverall        = minmnadealindx;
                                    totdealsoverall   = dealsforbdrcusip6;
```

```
                                    ainblocklevel     = ain;
                                    totdealsinblock = maxmnadealinblockindx75pall; *need to estimate number
below after adj for same-day deals: totcum_suminblock;                          run;
                          data aaafew5anndateindxainC; set aaafew5anndateindxainC0;
                           if sameanndate =1 then delete;
                                if dealsinday > 1 then concurrentdeals=1; else concurrentdeals=0; run;
                          proc sort data=aaafew5;       by bdrcusip6 anndate; run;
                          proc sort data=aaafew5anndateindxainC;    by bdrcusip6 anndate; run;
                          data aaafew5statsA1; merge aaafew5(in=a) aaafew5anndateindxainC;
                                if a; by bdrcusip6 anndate;
                          run;
                          proc sort data=aaafew5statsA1;      by bdrcusip6mnablockindx75p; run;
                          *estimate total deals in each block;
                          *Note that ainblocklevel might be lower if there are more than 1 deal in same day;
                          proc means data=aaafew5statsA1 noprint;
                                by bdrcusip6mnablockindx75p ;
                                var anndate;
                                output out= maxnadealindxinblock    (drop=_FREQ_ _TYPE_)
                                      n(anndate )             =     totcum_suminblock
                                      min(anndate)            =     firstanndatebdrblockrw
                                      max(anndate)            =     lastanndatebdrblockrw;
                    ;  run;
                          proc sort data=aaafew5statsA1;      by bdrcusip6 anndate; run;
                          data aaafew5statsA2; merge aaafew5statsA1(in=a) maxnadealindxinblock;
                                if a; by bdrcusip6mnablockindx75p;
                          run;
                          proc sort data=aaafew5statsA2;      by bdrcusip6 anndate; run;
                    data aaafew5A2cleanstartA; set aaafew5statsA2; run;
                                        data  aaafew5A2bybdrblocktot; set  aaafew5A2cleanstartA;
                                              format firstanndatebdrblockrw mmddyy9.;format
lastanndatebdrblockrw mmddyy9.;
                                              totdealsinblck = totcum_suminblock;
                                              bdrblockduration = lastanndatebdrblockrw-
firstanndatebdrblockrw+1;
                                              bdrblockintensity = totcum_suminblock/bdrblockduration;
                                              if bdrblockduration = 1 and totcum_suminblock = 1 then
bdrblockintensityadj = 0; else      bdrblockintensityadj = bdrblockintensity;
                                              drop totcum_suminblock;
                                        run;
                                        *get max for each bdrcusip6;
                                        proc sql;
                                          create table aaafew5A2bybdrblock as
```

```
                                        select distinct
                                             bdrcusip6,
                                                                max(bdrblockduration)               as
      maxbdrblockduration,
                                                                mean(bdrblockduration)              as
      meanbdrblockduration,
                                                                median(bdrblockduration)            as
      mdnbdrblockduration,
                                                                min(bdrblockduration)               as
      minbdrblockduration,
                                                    max(bdrblockintensityadj)           as
maxbdrblockintensityadj,
                                                    mean(bdrblockintensityadj)          as
mnbdrblockintensityadj
                                        from   aaafew5A2bybdrblocktot
                                        group by  bdrcusip6;
                                      quit;
                                      *merge max based on "by mna19851996bybdrblocktot" to the
bdracquirer level;
                                      proc sql;
                                        create table aaafew5A2bybdrvarsforclstA as
                                        select *
                                        from   aaafew5A2bybdrblocktot, aaafew5A2bybdrblock
                                           where
aaafew5A2bybdrblocktot.bdrcusip6=aaafew5A2bybdrblock.bdrcusip6
                                        order by aaafew5A2bybdrblocktot.bdrcusip6;
                                      quit;
                              data aaafew5A2bybdrvarsforclstB; set aaafew5A2bybdrvarsforclstA;
                                           compactnessrw = maxbdrblockintensityadj;
                                           compactnessmnrw = mnbdrblockintensityadj;
                                           totalnblocks = totalmnablockindx75p;
                                      run;
                            proc sort data=aaafew5A2bybdrvarsforclstB;      by bdrcusip6 anndate; run;
                            data aaafew5A2bybdrvarsforclstBge1; set aaafew5A2bybdrvarsforclstB;
                            data selvarsbybdrgsz8918; set aaafew5A2bybdrvarsforclstBge1;
                                  totdealsge1             = totdeals;
                                  totdealsinblckge1       = totdealsinblck;
                                  totalnblocksge1         = totalnblocks ;
                                  compactnessrwge1  = compactnessrw ;
                                  ainoverallge1           = ainoverall;
                                  ainblocklevelge1  = ainblocklevel;
                                  concurrentdealsge1      = concurrentdeals;
```

```
                              keep acqname trgname drelsizege1mm bdrcusip6 anndate annyear dealnosdc
                              totdealsge1       totdealsinblckge1 totalnblocksge1   compactnessrwge1
      ainoverallge1     ainblocklevelge1  concurrentdealsge1;
                  run;

/* main     vars for ALL deals and for only large deals
data selvarsbybdrall8918
      totdeals          totdealsinblck          totalnblocks      compactnessrw           ainoverall
      ainblocklevel           concurrentdeals
data selvarsbybdragsz8918
      totdealsge1       totdealsinblckge1 totalnblocksge1   compactnessrwge1  ainoverallge1
      ainblocklevelge1  concurrentdealsge1
                              */
*merge all dals and ge1mmusd deals statistics for bdrcusip6 and anndate;
      proc sort data=selvarsbybdrall8918;  by bdrcusip6 dealnosdc; run;
      proc sort data=selvarsbybdrgsz8918; by bdrcusip6 dealnosdc; run;
data selvarsallgszrw&yrstrtrw_&yrendrw_; merge selvarsbybdrall8918(in=a) selvarsbybdrgsz8918 (in=b);
      if a; by bdrcusip6 dealnosdc; run;
      proc sort data=selvarsbybdrall8918;  by bdrcusip6 anndate; run;
*get data at the bidder level for each window to use in cluster analysis in Stata;
proc means data=selvarsallgszrw&yrstrtrw_&yrendrw_ noprint;
      by bdrcusip6 ;
      var anndate;
      output out= sbdrallgsz5yrrw&yrstrtrw_&yrendrw_  (drop=_FREQ_ _TYPE_)
            max(totdeals)                 =     bdr5totdeals
            max(totdealsge1)        =     bdr5totdealsge1
            max(totalnblocks)       =     bdr5totalnblocks
            max(totalnblocksge1)    =     bdr5totalnblocksge1
            max(compactnessrw)            =     bdr5compactnessrw
            max(compactnessrwge1)   =     bdr5compactnessrwge1
            mean(totdealsinblck)    =     bdr5mntotdealsinblck
            mean(totdealsinblckge1) =     bdr5mntotdealsinblckge1
            max(ainblocklevel)            =     bdr5mxainblocklevel
            max(ainblocklevelge1)   =     bdr5mxainblocklevelge1
            mean(concurrentdeals)   =     bdr5mnconcurrentdeals
            mean(concurrentdealsge1)=     bdr5mnconcurrentdealsge1
                  ;  run;
data sbdrallgsz5yrrw&yrstrtrw_&yrendrw_; set sbdrallgsz5yrrw&yrstrtrw_&yrendrw_;
            pprtn5ge1totdealsrw=bdr5totdealsge1/bdr5totdeals;
            pprtn5ge1totblocksrw=bdr5totalnblocksge1/bdr5totalnblocks;
/*Main Vars at bdr level for each rolling window:
```

```
            bdr5totdeals bdr5totdealsge1  bdr5totalnblocks  bdr5totalnblocksge1           bdr5compactnessrw
bdr5compactnessrwge1
            bdr5mntotdealsinblck  bdr5mntotdealsinblckge1   bdr5mxainblocklevel     bdr5mxainblocklevelge1
      bdr5mnconcurrentdeals   bdr5mnconcurrentdealsge1
            pprtn5ge1totdealsrw  pprtn5ge1totblocksrw */ *at bdr level;
data serialac.selvarsallgszrw&yrstrtrw_&yrendrw_; set selvarsallgszrw&yrstrtrw_&yrendrw_; run; *for all
yrs: 56087 observations and 21 variables.;
            proc means data=sbdrallgsz5yrrw&yrstrtrw_&yrendrw_ n mean min p25 median p75 p99 max sum;
            var bdr5totdeals bdr5totdealsge1    bdr5totalnblocks  bdr5totalnblocksge1
      bdr5compactnessrw bdr5compactnessrwge1
                  bdr5mntotdealsinblck  bdr5mntotdealsinblckge1   bdr5mxainblocklevel
      bdr5mxainblocklevelge1  bdr5mnconcurrentdeals   bdr5mnconcurrentdealsge1
                  pprtn5ge1totdealsrw  pprtn5ge1totblocksrw  ; run;
data serialac.sbdrallgsz5yrrw&yrstrtrw_&yrendrw_; set sbdrallgsz5yrrw&yrstrtrw_&yrendrw_; run; *for all
yrs: 56087 observations and 21 variables.;
          quit;
      %end;
%mend CreateRWSamplesallge1mm;

   *END of MACRO;


%CreateRWSamplesallge1mm(1989,1989, 1989, 0)
%CreateRWSamplesallge1mm(1989,1990, 1990, 1)
%CreateRWSamplesallge1mm(1989,1991, 1991, 2)
%CreateRWSamplesallge1mm(1989,1992, 1992, 3)
%CreateRWSamplesallge1mm(1989,1993, 1993, 4)
%CreateRWSamplesallge1mm(1990,1994, 1994, 4)
%CreateRWSamplesallge1mm(1991,1995, 1995, 4)
%CreateRWSamplesallge1mm(1992,1996, 1996, 4)
%CreateRWSamplesallge1mm(1993,1997, 1997, 4)
%CreateRWSamplesallge1mm(1994,1998, 1998, 4)
%CreateRWSamplesallge1mm(1995,1999, 1999, 4)
%CreateRWSamplesallge1mm(1996,2000, 2000, 4)
%CreateRWSamplesallge1mm(1997,2001, 2001, 4)
%CreateRWSamplesallge1mm(1998,2002, 2002, 4)
%CreateRWSamplesallge1mm(1999,2003, 2003, 4)
%CreateRWSamplesallge1mm(2000,2004, 2004, 4)
%CreateRWSamplesallge1mm(2001,2005, 2005, 4)
%CreateRWSamplesallge1mm(2002,2006, 2006, 4)
%CreateRWSamplesallge1mm(2003,2007, 2007, 4)
%CreateRWSamplesallge1mm(2004,2008, 2008, 4)
```

```
%CreateRWSamplesallge1mm(2005,2009, 2009, 4)
%CreateRWSamplesallge1mm(2006,2010, 2010, 4)
%CreateRWSamplesallge1mm(2007,2011, 2011, 4)
%CreateRWSamplesallge1mm(2008,2012, 2012, 4)
%CreateRWSamplesallge1mm(2009,2013, 2013, 4)
%CreateRWSamplesallge1mm(2010,2014, 2014, 4)
%CreateRWSamplesallge1mm(2011,2015, 2015, 4)
%CreateRWSamplesallge1mm(2012,2016, 2016, 4)
%CreateRWSamplesallge1mm(2013,2017, 2017, 4)
%CreateRWSamplesallge1mm(2014,2018, 2018, 4)

/* Start using the generated 5-year rolling window sub-samples in Stata to get the serial acquirer
classifications
using cluster analysis in Stata code: 20220708_SerialAcq_ClstrAnlyssRllngWndw13ge1mmusdrw
And then, bring the results back to SAS
Main var of interest: srw5acqtype or srw7acqtype13gsz, depending on the analysis (e.g. 5 yrs, 7 yrs, gz for
large size deals, etc.
Note that for efficiency sometimes I adjust the name in SAS or in Stata according to the double-checked
subsample used in the analysis of interest */
```

0.3 Stata code to get 5- year rolling window sub-samples with cluster analysis with the goal to get the ex ante serial acquirer classification in each 5-year rolling window sub-sample.

/* UPDATED for 5 yrs Rolling window

Serial Acquirer classification, using 5-yr rolling window datasets generated in SAS

For RCFS

Set max for loners to 1, [1,1]

set max for occ to 4, [2,4]

set max for sprinters to 9 [5,7]

then, marthn [>10]

because of shorter rolling window of 5 years

Files at bidder level: sbdrallgsz5yrrw

Main vars of interest at deal-level bidder anndate (dealnosdc):

(For ALL-deals sample: totdeals totdealsinblck totalnblocks compactnessrw ainoverall ainblocklevel concurrentdeals)

For ge1MMusd-deals sample: totdealsge1 totdealsinblckge1 totalnblocksge1 compactnessrwge1 ainoverallge1 ainblocklevelge1 concurrentdealsge1

*Adj based on analysis and averages of whole set of rolling windows, to smooth clusters and reduce noise

Thresholds for ALL-deals sample

global trhshldocc 1

global trhshldspr 5

global trhshldmrth 7

Set max for lnr to 3,

Set max for occ to 3,

set max for sprinters to 7

because of shorter rolling window of 5 years and larger deals

Thresholds for Ge1MMusd-deals sample

Set max for loners to 1, [1,1]

set max for occ to 4, [2,4]

set max for sprinters to 9 [5,7]

then, marthn [>10]

because of shorter rolling window of 5 years

No Reqrmnt prior 3 yrs :

*samples for only large deal size

%CreateRWSamplesallge1mm(1989,1989, 1989, 0)

%CreateRWSamplesallge1mm(1989,1990, 1990, 1)

%CreateRWSamplesallge1mm(1989,1991, 1991, 2)

%CreateRWSamplesallge1mm(1989,1992, 1992, 3)

%CreateRWSamplesallge1mm(1989,1993, 1993, 4)

%CreateRWSamplesallge1mm(1990,1994, 1994, 4)

%CreateRWSamplesallge1mm(1991,1995, 1995, 4)

%CreateRWSamplesallge1mm(1992,1996, 1996, 4)

%CreateRWSamplesallge1mm(1993,1997, 1997, 4)

%CreateRWSamplesallge1mm(1994,1998, 1998, 4)

%CreateRWSamplesallge1mm(1995,1999, 1999, 4)

%CreateRWSamplesallge1mm(1996,2000, 2000, 4)

%CreateRWSamplesallge1mm(1997,2001, 2001, 4)

%CreateRWSamplesallge1mm(1998,2002, 2002, 4)

%CreateRWSamplesallge1mm(1999,2003, 2003, 4)

%CreateRWSamplesallge1mm(2000,2004, 2004, 4)

%CreateRWSamplesallge1mm(2001,2005, 2005, 4)

%CreateRWSamplesallge1mm(2002,2006, 2006, 4)

%CreateRWSamplesallge1mm(2003,2007, 2007, 4)

%CreateRWSamplesallge1mm(2004,2008, 2008, 4)

%CreateRWSamplesallge1mm(2005,2009, 2009, 4)

%CreateRWSamplesallge1mm(2006,2010, 2010, 4)

%CreateRWSamplesallge1mm(2007,2011, 2011, 4)

%CreateRWSamplesallge1mm(2008,2012, 2012, 4)

%CreateRWSamplesallge1mm(2009,2013, 2013, 4)

%CreateRWSamplesallge1mm(2010,2014, 2014, 4)

%CreateRWSamplesallge1mm(2011,2015, 2015, 4)

%CreateRWSamplesallge1mm(2012,2016, 2016, 4)

%CreateRWSamplesallge1mm(2013,2017, 2017, 4)

%CreateRWSamplesallge1mm(2014,2018, 2018, 4)

Files at bidder anndate (dealnosdc) level: selvarsallgszrw

Files at bidder level: sbdrallgsz5yrrw

Main vars of interest at deal-level bidder anndate (dealnosdc):

For ALL-deals sample: totdeals totdealsinblck totalnblocks compactnessrw ainoverall ainblocklevel concurrentdeals

For ge1MMusd-deals sample: totdealsge1 totdealsinblckge1 totalnblocksge1 compactnessrwge1 ainoverallge1 ainblocklevelge1 concurrentdealsge1

Main vars of interest at bidder level:

[bdr5totdealsge1 bdr5totalnblocksge1 bdr5compactnessrwge1]

bdr5totdeals bdr5totdealsge1 bdr5totalnblocks bdr5totalnblocksge1 bdr5compactnessrw bdr5compactnessrwge1

bdr5mntotdealsinblck bdr5mntotdealsinblckge1 bdr5mxainblocklevel bdr5mxainblocklevelge1 bdr5mnconcurrentdeals bdr5mnconcurrentdealsge1

pprtn5ge1totdealsrw pprtn5ge1totblocksrw

First I determine the clusters for gsz (ge1mmusd=1) sample

and then for ALL-feals sample (shown below)

I just need to change the vars for the clusters and then change names

```
*/      cd "D:\serialac\"
            set more off
            display _newline(100)
            clear
import sas   sbdrallgsz5yrrw19891989, case(lower) clear
import sas   sbdrallgsz5yrrw19891990, case(lower) clear
import sas   sbdrallgsz5yrrw19891991, case(lower) clear
import sas   sbdrallgsz5yrrw19891992, case(lower) clear
import sas   sbdrallgsz5yrrw19891993, case(lower) clear
import sas   sbdrallgsz5yrrw19901994, case(lower) clear
import sas   sbdrallgsz5yrrw19911995, case(lower) clear
```

```
import sas   sbdrallgsz5yrrw19921996, case(lower) clear
import sas   sbdrallgsz5yrrw19931997, case(lower) clear
import sas   sbdrallgsz5yrrw19941998, case(lower) clear
import sas   sbdrallgsz5yrrw19951999, case(lower) clear
import sas   sbdrallgsz5yrrw19962000, case(lower) clear
import sas   sbdrallgsz5yrrw19972001, case(lower) clear
import sas   sbdrallgsz5yrrw19982002, case(lower) clear
import sas   sbdrallgsz5yrrw19992003, case(lower) clear
import sas   sbdrallgsz5yrrw20002004, case(lower) clear
import sas   sbdrallgsz5yrrw20012005, case(lower) clear
import sas   sbdrallgsz5yrrw20022006, case(lower) clear
import sas   sbdrallgsz5yrrw20032007, case(lower) clear
import sas   sbdrallgsz5yrrw20042008, case(lower) clear
import sas   sbdrallgsz5yrrw20052009, case(lower) clear
import sas   sbdrallgsz5yrrw20062010, case(lower) clear
import sas   sbdrallgsz5yrrw20072011, case(lower) clear
import sas   sbdrallgsz5yrrw20082012, case(lower) clear
import sas   sbdrallgsz5yrrw20092013, case(lower) clear
import sas   sbdrallgsz5yrrw20102014, case(lower) clear
import sas   sbdrallgsz5yrrw20112015, case(lower) clear
import sas   sbdrallgsz5yrrw20122016, case(lower) clear
```

```
import sas   sbdrallgsz5yrrw20132017, case(lower) clear
import sas   sbdrallgsz5yrrw20142018, case(lower) clear
/* For easiness of running, copy from above the desired 5-yr RW */
          import sas  sbdrallgsz5yrrw19921996, case(lower) clear
                    *drop if bdrcusip6 == "023586"  /*17 deals in just one day*/
           renvars , lower
                              compress
                              set more off
          *                   des
          *          bdrtotdeals       bdr5totalnblocks          bdr5compactnessrw   /  *bdr5totdealsge1 bdr5totalnblocksge1 bdr5compactnessrwge1
          global vars       bdr5totdeals              bdr5totalnblocks                  bdr5compactnessrw
          sum $vars , det
          cluster kmed $vars , k(4) name(kmed4) measure(abs) start(prand(8723))
          cluster kmed $vars , k(5) name(kmed5) measure(abs) start(prand(8723))
          cluster kmed $vars , k(6) name(kmed6) measure(abs) start(prand(8723))
          cluster kmed $vars , k(7) name(kmed7) measure(abs) start(prand(8723))

          cluster stop      kmed4
          cluster stop      kmed5
          cluster stop      kmed6
          cluster stop      kmed7
```

```
table  kmed4 kmed5
table  kmed4 kmed6
table  kmed4 kmed7

tabstat   $vars    , by(kmed4) stat(   mean n min median max ) long columns(statistics)
tabstat   $vars    , by(kmed5) stat(   mean n min median max ) long columns(statistics)
tabstat   $vars    , by(kmed6) stat(   mean n min median max ) long columns(statistics)
tabstat   $vars    , by(kmed7) stat(   mean n min median max ) long columns(statistics)

        cluster kmed $vars , k(8) name(kmed8) measure(abs) start(prand(8723))
        cluster stop       kmed8
        tabstat   $vars    , by(kmed8) stat(   mean n min median max ) long columns(statistics)

/*****************************************************/
/********* FIRST RUN CODE ABOVE THEN THE According CODE BELOW FOR EACH ROLLING WINDOW
**************************************/
/*****************************************************/
*Adj based on analysis and averages of whole set of rolling windows, to smooth clusters and reduce noise
*ge1mmusd  thresholds:          trhshldocc 1 trhshldspr 4 trhshldmrth 7
*All deals thresholds:  trhshldocc 2 trhshldspr 6 trhshldmrth 9
```

* for shorter windows of 5 yrs I change threshold for Sprinters from 6 to 5, and from 9 to 8 for marathoners (i.e., max for occ is 5, max for sprinters is 8)

*ge1mmusd varsadd bdr5mntotdealsinblckge1 bdr5mnconcurrentdealsge1 pprtn5ge1totdealsrw

*All deals varsadd bdr5mntotdealsinblck bdr5mnconcurrentdeals pprtn5ge1totdealsrw

histogram bdr5totdeals if bdr5totdeals<10

global varsadd bdr5mntotdealsinblck bdr5mnconcurrentdeals pprtn5ge1totdealsrw

global trhshldocc 2

global trhshldspr 5

global trhshldmrth 8

sum $vars $varsadd

********** bdr5totdealsge1 totalnblocksge1 compactnessrwge1

* mna19891989bybdrvarsforclstflxrw

* modified for window 4 bins hierarchical for # of selected clusters according to criteria using overall sample

drop srw5acqtype

gen srw5acqtype = 1 if kmed5==4|kmed5==5

replace srw5acqtype = 2 if kmed5==1|kmed5==3|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))

replace srw5acqtype = 3 if (kmed5==99)|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))

replace srw5acqtype = 4 if kmed5==2|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))

tabstat $vars , by(srw5acqtype) stat(mean n min median max) long columns(statistics)

table kmed4 kmed5

```
tabstat  $vars    $varsadd  , by(srw5acqtype) stat(  mean n min median max ) long columns(statistics)
saveold "D:\xold_mna19891989selvars5rw", version(12) replace

*****************************
* mna19891990selvars5rw
* modified for window  4 bins hierarchical  for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                  srw5acqtype = 1 if kmed7==1|kmed7==7
replace srw5acqtype = 2 if kmed7==4|kmed7==5|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed7==6|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed7==2|kmed7==3|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat  $vars   , by(srw5acqtype) stat(  mean n min median max ) long columns(statistics)
table  kmed4 kmed5

tabstat  $vars   $varsadd  , by(srw5acqtype) stat(  mean n min median max ) long columns(statistics)
saveold "D:\xold_mna19891990selvars5rw", version(12) replace

*****************************
* mna19891991selvars5rw
```

```
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                 srw5acqtype = 1 if kmed7==1|kmed7==7
replace srw5acqtype = 2 if kmed7==6|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed7==5|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed7==2|kmed7==3|kmed7==4|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat   $vars    , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed5

tabstat   $vars    $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna19891991selvars5rw", version(12) replace

*****************************
* mna19891992selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                 srw5acqtype = 1 if kmed7==1|kmed7==7
replace srw5acqtype = 2 if kmed7==6|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed7==4|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
```

```
replace srw5acqtype = 4 if kmed7==2|kmed7==3|kmed7==5|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat  $vars   , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed7

tabstat  $vars   $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna19891992selvars5rw", version(12) replace

*****************************
* mna19891993selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                 srw5acqtype = 1 if kmed7==7
replace srw5acqtype = 2 if kmed7==1|kmed7==5|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed7==4|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed7==2|kmed7==3|kmed7==6|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat  $vars   , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat  $vars   $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna19891993selvars5rw", version(12) replace
```

```
*****************************
* mna19901994selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                  srw5acqtype = 1 if kmed6==5
replace srw5acqtype = 2 if kmed6==2|kmed6==6|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed6==4|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed6==1|kmed6==3|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat   $vars    , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat   $vars    $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna19901994selvars5rw", version(12) replace

*****************************
* mna19911995selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
```

```
gen                 srw5acqtype = 1 if kmed5==1
replace srw5acqtype = 2 if kmed5==4|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed5==3|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed5==2|kmed5==5|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat   $vars    , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat   $vars    $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna19911995selvars5rw", version(12) replace

*****************************
* mna19891996selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                 srw5acqtype = 1 if kmed7==2
replace srw5acqtype = 2 if kmed7==1|kmed7==3|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed7==6|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed7==4|kmed7==5|kmed7==7|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat   $vars    , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6
```

```
tabstat  $vars   $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna19921996selvars5rw", version(12) replace

*****************************
* mna19891997selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                  srw5acqtype = 1 if kmed5==1
replace srw5acqtype = 2 if kmed5==2|kmed5==5|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed5==4|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed5==3|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat  $vars   , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat  $vars   $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna19931997selvars5rw", version(12) replace

*****************************
* mna19891998selvars5rw
```

```
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                 srw5acqtype = 1 if kmed6==2
replace srw5acqtype = 2 if kmed6==6|kmed6==5|kmed6==4|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed6==3|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed6==1|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat   $vars    , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat   $vars    $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna19941998selvars5rw", version(12) replace

*****************************
* mna19891999selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                 srw5acqtype = 1 if kmed4==1
replace srw5acqtype = 2 if kmed4==2|kmed4==4|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if (kmed4==99)|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
```

```
replace srw5acqtype = 4 if kmed4==3|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat   $vars    , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat   $vars    $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna19951999selvars5rw", version(12) replace

*****************************
* mna19892000selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                 srw5acqtype = 1 if kmed6==2
replace srw5acqtype = 2 if kmed6==3|kmed6==5|kmed6==6|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed6==4|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed6==1|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat   $vars    , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat   $vars    $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna19962000selvars5rw", version(12) replace
```

```
*****************************
* mna19892001selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                 srw5acqtype = 1 if kmed7==6
replace srw5acqtype = 2 if kmed7==3|kmed7==4|kmed7==5|kmed7==7|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed7==1|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed7==2|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat  $vars   , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat  $vars   $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna19972001selvars5rw", version(12) replace

*****************************
* mna19902002selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
```

```
gen                 srw5acqtype = 1 if kmed6==6
replace srw5acqtype = 2 if kmed6==1|kmed6==4|kmed6==5|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed6==3|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed6==2|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat   $vars    , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat   $vars    $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna19982002selvars5rw", version(12) replace

*****************************
* mna19912003selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                 srw5acqtype = 1 if kmed7==1
replace srw5acqtype = 2 if kmed7==3|kmed7==7|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed7==5|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed7==2|kmed7==4|kmed7==6|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat   $vars    , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6
```

```
tabstat  $vars   $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna19992003selvars5rw", version(12) replace

*****************************
* mna19922004selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                  srw5acqtype = 1 if kmed6==1
replace srw5acqtype = 2 if kmed6==6|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed6==4|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed6==2|kmed6==3|kmed6==5|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat  $vars   , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat  $vars   $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna20002004selvars5rw", version(12) replace

*****************************
* mna19932005selvars5rw
```

```
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                 srw5acqtype = 1 if kmed6==1
replace srw5acqtype = 2 if kmed6==6|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed6==4|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed6==2|kmed6==3|kmed6==5|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat   $vars    , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat   $vars    $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna20012005selvars5rw", version(12) replace

*****************************
* mna19942006selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                 srw5acqtype = 1 if kmed5==1
replace srw5acqtype = 2 if kmed5==5|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed5==4|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
```

```
replace srw5acqtype = 4 if kmed5==2|kmed5==3|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat   $vars    , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat   $vars    $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna20022006selvars5rw", version(12) replace

*****************************
* mna19952007selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                 srw5acqtype = 1 if kmed7==5
replace srw5acqtype = 2 if kmed7==1|kmed7==2|kmed7==6|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed7==3|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed7==4|kmed7==7|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat   $vars    , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat   $vars    $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna20032007selvars5rw", version(12) replace
```

```
*****************************
* mna19962008selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                 srw5acqtype = 1 if kmed7==1
replace srw5acqtype = 2 if kmed7==3|kmed7==6|kmed7==7|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed7==5|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed7==2|kmed7==4|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat   $vars    , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat   $vars    $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna20042008selvars5rw", version(12) replace

*****************************
* mna19972009selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
```

```
gen                 srw5acqtype = 1 if kmed7==1
replace srw5acqtype = 2 if kmed7==3|kmed7==6|kmed7==7|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed7==5|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed7==2|kmed7==4|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat   $vars    , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat   $vars    $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna20052009selvars5rw", version(12) replace

*****************************
* mna19982010selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                 srw5acqtype = 1 if kmed5==1
replace srw5acqtype = 2 if kmed5==3|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed5==4|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed5==2|kmed5==5|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat   $vars    , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6
```

```
tabstat  $vars   $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna20062010selvars5rw", version(12) replace

*****************************
* mna19992011selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                  srw5acqtype = 1 if kmed6==2
replace srw5acqtype = 2 if kmed6==1|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed6==5|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed6==3|kmed6==4|kmed6==6|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat  $vars   , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat  $vars   $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna20072011selvars5rw", version(12) replace

          *****************************
* mna20002012selvars5rw
```

```
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                 srw5acqtype = 1 if kmed5==1
replace srw5acqtype = 2 if kmed5==5|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed5==3|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed5==4|kmed5==2|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat   $vars    , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat   $vars    $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna20082012selvars5rw", version(12) replace

          ****************************
* mna20012013selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                 srw5acqtype = 1 if kmed6==1
replace srw5acqtype = 2 if kmed6==2|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed6==5|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
```

```
replace srw5acqtype = 4 if kmed6==3|kmed6==4|kmed6==6|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat  $vars   , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat  $vars   $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna20092013selvars5rw", version(12) replace

          ****************************
* mna20022014selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                 srw5acqtype = 1 if kmed5==1
replace srw5acqtype = 2 if kmed5==3|kmed5==4|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if (kmed5==99)|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed5==2|kmed5==5|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat  $vars   , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat  $vars   $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna20102014selvars5rw", version(12) replace
```

```
/* 14 Jul 2021   Aqui vopy. Faltan las sig de abajo y en XL crear esos campos y el resumen */


        ****************************

* mna20032015selvars5rw

* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample

*

drop srw5acqtype

gen                 srw5acqtype = 1 if kmed7==1

replace srw5acqtype = 2 if kmed7==5|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))

replace srw5acqtype = 3 if kmed7==6|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))

replace srw5acqtype = 4 if kmed7==2|kmed7==3|kmed7==4|kmed7==7|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))

tabstat  $vars   , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)

table  kmed4 kmed6


tabstat  $vars   $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)

saveold "D:\xold_mna20112015selvars5rw", version(12) replace


        ****************************

* mna20042016selvars5rw
```

```
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                 srw5acqtype = 1 if kmed7==1
replace srw5acqtype = 2 if kmed7==2|kmed7==3|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed7==7|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed7==4|kmed7==5|kmed7==6|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat   $vars    , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat   $vars    $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna20122016selvars5rw", version(12) replace

          ****************************
* mna20052017selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                 srw5acqtype = 1 if kmed7==1
replace srw5acqtype = 2 if kmed7==6|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed7==5|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
```

```
replace srw5acqtype = 4 if kmed7==2|kmed7==3|kmed7==4|kmed7==7|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat  $vars   , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat  $vars   $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna20132017selvars5rw", version(12) replace

          ****************************
* mna20062018selvars5rw
* modified for window  4 bins hierarchical   for # of selected clusters according to criteria using overall sample
*
drop srw5acqtype
gen                 srw5acqtype = 1 if kmed7==2
replace srw5acqtype = 2 if kmed7==1|kmed7==6|((bdr5totdeals>$trhshldocc)&!missing(bdr5totdeals))
replace srw5acqtype = 3 if kmed7==7|((bdr5totdeals>$trhshldspr)&!missing(bdr5totdeals))
replace srw5acqtype = 4 if kmed7==3|kmed7==4|kmed7==5|((bdr5totdeals>$trhshldmrth)&!missing(bdr5totdeals))
tabstat  $vars   , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
table  kmed4 kmed6

tabstat  $vars   $varsadd  , by(srw5acqtype) stat(   mean n min median max ) long columns(statistics)
saveold "D:\xold_mna20142018selvars5rw", version(12) replace
```

```
/*****************************************************************************************************/

/*APPEND ALL  see code in older cluster code 20210714_SerialAcq_ClstrAnlyssRllngWndw13flxrw*/

                        cd "D:\"

                        set more off

                                    display _newline(100)

                                    set more off

                                    clear

use xold_mna20072013bybdrvarsforclstflxrw,  replace

append using xold_mna19891995bybdrvarsforclstflxrw xold_mna19901996bybdrvarsforclstflxrw xold_mna19911997bybdrvarsforclstflxrw
xold_mna19921998bybdrvarsforclstflxrw xold_mna19921998bybdrvarsforclstflxrw xold_mna19931999bybdrvarsforclstflxrw
xold_mna19942000bybdrvarsforclstflxrw ///

xold_mna19952001bybdrvarsforclstflxrw xold_mna19962002bybdrvarsforclstflxrw xold_mna19972003bybdrvarsforclstflxrw
xold_mna19982004bybdrvarsforclstflxrw xold_mna19992005bybdrvarsforclstflxrw xold_mna20002006bybdrvarsforclstflxrw
xold_mna20012007bybdrvarsforclstflxrw ///

xold_mna20022008bybdrvarsforclstflxrw xold_mna20032009bybdrvarsforclstflxrw xold_mna20042010bybdrvarsforclstflxrw
xold_mna20052011bybdrvarsforclstflxrw xold_mna20062012bybdrvarsforclstflxrw xold_mna20072013bybdrvarsforclstflxrw
xold_mna20082014bybdrvarsforclstflxrw ///

xold_mna20092015bybdrvarsforclstflxrw xold_mna20102016bybdrvarsforclstflxrw xold_mna20112017bybdrvarsforclstflxrw
xold_mna20122018bybdrvarsforclstflxrw xold_mna20132019bybdrvarsforclstflxrw

saveold "D:\xold_mnaALL19892019bybdrvarsforclstflxrw", version(12) replace

/*************************************************************/

/*  Create routine for creating shorter Stata files only with key variables for later merging  */
```

```
cd "D:\"
set more off
        clear
use xold_mna19891989selvars5rw , replace
gen sracqtypgsz5rw = srw5acqtype
gen totdlsgsz5rw =bdr5totdeals
gen lstyrrw = 1989
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr8989", version(12) replace

use xold_mna19891990selvars5rw , replace
gen sracqtyp5rwyr8990 = srw5acqtype
gen  totdls5yr8990 =bdr5totdeals
gen lstyrrw = 1990
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr8990", version(12) replace

use xold_mna19891991selvars5rw , replace
gen sracqtyp5rwyr8991 = srw5acqtype
gen  totdls5yr8991 =bdr5totdeals
gen lstyrrw = 1991
```

```
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr8991", version(12) replace

use xold_mna19891992selvars5rw , replace
gen sracqtyp5rwyr8992 = srw5acqtype
gen  totdls5yr8992 =bdr5totdeals
gen lstyrrw = 1992
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr8992", version(12) replace

use xold_mna19891993selvars5rw , replace
gen sracqtyp5rwyr8993 = srw5acqtype
gen  totdls5yr8993 =bdr5totdeals
gen lstyrrw = 1993
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr8993", version(12) replace

use xold_mna19901994selvars5rw , replace
gen sracqtyp5rwyr9094 = srw5acqtype
gen  totdls5yr9094 =bdr5totdeals
gen lstyrrw = 1994
```

```
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr9094", version(12) replace

use xold_mna19911995selvars5rw , replace
gen sracqtyp5rwyr9195 = srw5acqtype
gen  totdls5yr9195 =bdr5totdeals
gen lstyrrw = 1995
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr9195", version(12) replace

use xold_mna19921996selvars5rw , replace
gen sracqtyp5rwyr9296 = srw5acqtype
gen  totdls5yr9296 =bdr5totdeals
gen lstyrrw = 1996
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr9296", version(12) replace

use xold_mna19931997selvars5rw , replace
gen sracqtyp5rwyr9397 = srw5acqtype
gen  totdls5yr9397 =bdr5totdeals
gen lstyrrw = 1997
```

```
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr9397", version(12) replace

use xold_mna19941998selvars5rw , replace
gen sracqtyp5rwyr9498 = srw5acqtype
gen  totdls5yr9498 =bdr5totdeals
gen lstyrrw = 1998
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr9498", version(12) replace

use xold_mna19951999selvars5rw , replace
gen sracqtyp5rwyr9599 = srw5acqtype
gen  totdls5yr9599 =bdr5totdeals
gen lstyrrw = 1999
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr9599", version(12) replace

use xold_mna19962000selvars5rw , replace
gen sracqtyp5rwyr9600 = srw5acqtype
gen  totdls5yr9600 =bdr5totdeals
gen lstyrrw = 2000
```

```
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr9600", version(12) replace

use xold_mna19972001selvars5rw , replace
gen sracqtyp5rwyr9701 = srw5acqtype
gen  totdls5yr9701 =bdr5totdeals
gen lstyrrw = 2001
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr9701", version(12) replace

use xold_mna19982002selvars5rw , replace
gen sracqtyp5rwyr9802 = srw5acqtype
gen  totdls5yr9802 =bdr5totdeals
gen lstyrrw = 2002
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr9802", version(12) replace

use xold_mna19992003selvars5rw , replace
gen sracqtyp5rwyr9903 = srw5acqtype
gen  totdls5yr9903 =bdr5totdeals
gen lstyrrw = 2003
```

```
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr9903", version(12) replace

use xold_mna20002004selvars5rw , replace
gen sracqtyp5rwyr0004 = srw5acqtype
gen  totdls5yr0004 =bdr5totdeals
gen lstyrrw = 2004
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr0004", version(12) replace

use xold_mna20012005selvars5rw , replace
gen sracqtyp5rwyr0105 = srw5acqtype
gen  totdls5yr0105 =bdr5totdeals
gen lstyrrw = 2005
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr0105", version(12) replace

use xold_mna20022006selvars5rw , replace
gen sracqtyp5rwyr0206 = srw5acqtype
gen  totdls5yr0206 =bdr5totdeals
gen lstyrrw = 2006
```

```
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr0206", version(12) replace

use xold_mna20032007selvars5rw , replace
gen sracqtyp5rwyr0307 = srw5acqtype
gen  totdls5yr0307 =bdr5totdeals
gen lstyrrw = 2007
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr0307", version(12) replace

use xold_mna20042008selvars5rw , replace
gen sracqtyp5rwyr0408 = srw5acqtype
gen  totdls5yr0408 =bdr5totdeals
gen lstyrrw = 2008
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr0408", version(12) replace

use xold_mna20052009selvars5rw , replace
gen sracqtyp5rwyr0509 = srw5acqtype
gen  totdls5yr0509 =bdr5totdeals
gen lstyrrw = 2009
```

```
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr0509", version(12) replace

use xold_mna20062010selvars5rw , replace
gen sracqtyp5rwyr0610 = srw5acqtype
gen  totdls5yr0610 =bdr5totdeals
gen lstyrrw = 2010
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr0610", version(12) replace

use xold_mna20072011selvars5rw , replace
gen sracqtyp5rwyr0711 = srw5acqtype
gen  totdls5yr0711 =bdr5totdeals
gen lstyrrw = 2011
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr0711", version(12) replace

use xold_mna20082012selvars5rw , replace
gen sracqtyp5rwyr0812 = srw5acqtype
gen  totdls5yr0812 =bdr5totdeals
gen lstyrrw = 2012
```

```
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr0812", version(12) replace

use xold_mna20092013selvars5rw , replace
gen sracqtyp5rwyr0913 = srw5acqtype
gen  totdls5yr0913 =bdr5totdeals
gen lstyrrw = 2013
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr0913", version(12) replace

use xold_mna20102014selvars5rw , replace
gen sracqtyp5rwyr1014 = srw5acqtype
gen  totdls5yr1014 =bdr5totdeals
gen lstyrrw = 2014
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr1014", version(12) replace

use xold_mna20112015selvars5rw , replace
gen sracqtyp5rwyr1115 = srw5acqtype
gen  totdls5yr1115 =bdr5totdeals
gen lstyrrw = 2015
```

```
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr1115", version(12) replace

use xold_mna20122016selvars5rw , replace
gen sracqtyp5rwyr1216 = srw5acqtype
gen  totdls5yr1216 =bdr5totdeals
gen lstyrrw = 2016
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr1216", version(12) replace

use xold_mna20132017selvars5rw , replace
gen sracqtyp5rwyr1317 = srw5acqtype
gen  totdls5yr1317 =bdr5totdeals
gen lstyrrw = 2017
keep bdrcusip6  srw5acqtype bdr5totdeals lstyrrw
saveold "D:\xold_sacqtpl5rwyr1317", version(12) replace

use xold_mna20142018selvars5rw , replace
gen sracqtyp5rwyr1418 = srw5acqtype
gen  totdls5yr1418 =bdr5totdeals
gen lstyrrw = 2018
```

keep bdrcusip6 srw5acqtype bdr5totdeals lstyrrw

saveold "D:\xold_sacqtpl5rwyr1418", version(12) replace

0.4 SAS code that uses the serial acquirer classification in each 5-year rolling window sub-sample generated in Stata. This code continues to prepare the sample to start with the actual analysis reported in the Tables.

0.4.A First for LARGE deals (and below 0.4.B for All deals without size filter)

```
/*bring back from Stata the classif using 5 yr rolling windows
to merge use bdrcusip6 and prior year for the ex ante approach
[3 steps below for each rolling window: import, create var to merge, and sort)
                  ************/
/*First expanding 5-yr rolling windows gsz ge1mmusd sample  Main Vars: srw7acqtypegsz
      bdr5totdealsge1 */

proc import out= xold_sacqtplgsz5rwyr8989 datafile='d:\xold_sacqtplgsz5rwyr8989.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr8990 datafile='d:\xold_sacqtplgsz5rwyr8990.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr8991 datafile='d:\xold_sacqtplgsz5rwyr8991.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr8992 datafile='d:\xold_sacqtplgsz5rwyr8992.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr8993 datafile='d:\xold_sacqtplgsz5rwyr8993.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr9094 datafile='d:\xold_sacqtplgsz5rwyr9094.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr9195 datafile='d:\xold_sacqtplgsz5rwyr9195.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr9296 datafile='d:\xold_sacqtplgsz5rwyr9296.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr9397 datafile='d:\xold_sacqtplgsz5rwyr9397.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr9498 datafile='d:\xold_sacqtplgsz5rwyr9498.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr9599 datafile='d:\xold_sacqtplgsz5rwyr9599.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr9600 datafile='d:\xold_sacqtplgsz5rwyr9600.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr9701 datafile='d:\xold_sacqtplgsz5rwyr9701.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr9802 datafile='d:\xold_sacqtplgsz5rwyr9802.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr9903 datafile='d:\xold_sacqtplgsz5rwyr9903.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr0004 datafile='d:\xold_sacqtplgsz5rwyr0004.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr0105 datafile='d:\xold_sacqtplgsz5rwyr0105.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr0206 datafile='d:\xold_sacqtplgsz5rwyr0206.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr0307 datafile='d:\xold_sacqtplgsz5rwyr0307.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr0408 datafile='d:\xold_sacqtplgsz5rwyr0408.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr0509 datafile='d:\xold_sacqtplgsz5rwyr0509.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr0610 datafile='d:\xold_sacqtplgsz5rwyr0610.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr0711 datafile='d:\xold_sacqtplgsz5rwyr0711.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr0812 datafile='d:\xold_sacqtplgsz5rwyr0812.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr0913 datafile='d:\xold_sacqtplgsz5rwyr0913.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr1014 datafile='d:\xold_sacqtplgsz5rwyr1014.dta' dbms=DTA replace; run;
```

```
proc import out= xold_sacqtplgsz5rwyr1115 datafile='d:\xold_sacqtplgsz5rwyr1115.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr1216 datafile='d:\xold_sacqtplgsz5rwyr1216.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr1317 datafile='d:\xold_sacqtplgsz5rwyr1317.dta' dbms=DTA replace; run;
proc import out= xold_sacqtplgsz5rwyr1418 datafile='d:\xold_sacqtplgsz5rwyr1418.dta' dbms=DTA replace; run;


/* already added ins Stata var "lstyrrw" as of latest year in rolling window , yet add here cusipID plus yr
To modify code aboce: take code from above and replace as needed */

data xold_sacqtplgsz5rwyr8989 ; set xold_sacqtplgsz5rwyr8989; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr8990 ; set xold_sacqtplgsz5rwyr8990; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr8991 ; set xold_sacqtplgsz5rwyr8991; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr8992 ; set xold_sacqtplgsz5rwyr8992; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr8993 ; set xold_sacqtplgsz5rwyr8993; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr9094 ; set xold_sacqtplgsz5rwyr9094; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr9195 ; set xold_sacqtplgsz5rwyr9195; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr9296 ; set xold_sacqtplgsz5rwyr9296; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr9397 ; set xold_sacqtplgsz5rwyr9397; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr9498 ; set xold_sacqtplgsz5rwyr9498; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr9599 ; set xold_sacqtplgsz5rwyr9599; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr9600 ; set xold_sacqtplgsz5rwyr9600; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr9701 ; set xold_sacqtplgsz5rwyr9701; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr9802 ; set xold_sacqtplgsz5rwyr9802; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr9903 ; set xold_sacqtplgsz5rwyr9903; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr0004 ; set xold_sacqtplgsz5rwyr0004; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
```

```
data xold_sacqtplgsz5rwyr0105 ; set xold_sacqtplgsz5rwyr0105; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr0206 ; set xold_sacqtplgsz5rwyr0206; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr0307 ; set xold_sacqtplgsz5rwyr0307; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr0408 ; set xold_sacqtplgsz5rwyr0408; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr0509 ; set xold_sacqtplgsz5rwyr0509; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr0610 ; set xold_sacqtplgsz5rwyr0610; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr0711 ; set xold_sacqtplgsz5rwyr0711; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr0812 ; set xold_sacqtplgsz5rwyr0812; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr0913 ; set xold_sacqtplgsz5rwyr0913; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr1014 ; set xold_sacqtplgsz5rwyr1014; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr1115 ; set xold_sacqtplgsz5rwyr1115; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr1216 ; set xold_sacqtplgsz5rwyr1216; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr1317 ; set xold_sacqtplgsz5rwyr1317; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;
data xold_sacqtplgsz5rwyr1418 ; set xold_sacqtplgsz5rwyr1418; cusip6lstyrrw =  bdrcusip6||lstyrrw; drop
bdrcusip6 ; run;

proc sort data=xold_sacqtplgsz5rwyr8989 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr8990 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr8991 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr8992 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr8993 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr9094 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr9195 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr9296 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr9397 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr9498 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr9599 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr9600 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr9701 ; by  cusip6lstyrrw; run;
```

```
proc sort data=xold_sacqtplgsz5rwyr9802 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr9903 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr0004 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr0105 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr0206 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr0307 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr0408 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr0509 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr0610 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr0711 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr0812 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr0913 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr1014 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr1115 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr1216 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr1317 ; by  cusip6lstyrrw; run;
proc sort data=xold_sacqtplgsz5rwyr1418 ; by  cusip6lstyrrw; run;

/* end of bringing back from Stata to SAS 5 yr rw  gsz ge1mmusd sample with cluster classification
main vars srw7acqtypegsz bdr5totdealsge1   based on 5yr rw gsz ge1mmusd==1*/

/*  USE THIS SECOND MERGING PROCESS AT YR LEVEL which includes many variables already defined in Stata for
robustness to other rolling windows periods (i.., 7 and 13 years instead of 5 year rolling windows) */
proc import out= xold_serlacqtoaviddrunbyyr datafile='d:\xold_serlacqtoaviddrunbyyr.dta' dbms=DTA replace;
run;
data xold_serlacqtoaviddrunbyyrA; set xold_serlacqtoaviddrunbyyr;
      lstyrrw = calyear -1;
       cusip6lstyrrw =  cusip6||lstyrrw;
run; * firm-yr observations;

data xold_serlacqtoaviddrunbyyrB;
retain  cusip6lstyrrw   cusip6 lstyrrw calyear gvkey ngvkey; set xold_serlacqtoaviddrunbyyrA; run;

proc sort data=xold_serlacqtoaviddrunbyyrB nodupkey dupout=dups;
 by cusip6 calyear; run;

proc sort data=xold_serlacqtoaviddrunbyyrB nodupkey; by  cusip6 calyear; run;
proc sort data=xold_serlacqtoaviddrunbyyrB; by  cusip6lstyrrw; run;
data serlacqbygszflx5rw; merge xold_serlacqtoaviddrunbyyrB (in=a)
      xold_sacqtplgsz5rwyr8989 xold_sacqtplgsz5rwyr8990 xold_sacqtplgsz5rwyr8991 xold_sacqtplgsz5rwyr8992
xold_sacqtplgsz5rwyr8993 xold_sacqtplgsz5rwyr9094 xold_sacqtplgsz5rwyr9195
```

```
      xold_sacqtplgsz5rwyr9296 xold_sacqtplgsz5rwyr9397 xold_sacqtplgsz5rwyr9498 xold_sacqtplgsz5rwyr9599
xold_sacqtplgsz5rwyr9600 xold_sacqtplgsz5rwyr9701 xold_sacqtplgsz5rwyr9802
      xold_sacqtplgsz5rwyr9903 xold_sacqtplgsz5rwyr0004 xold_sacqtplgsz5rwyr0105 xold_sacqtplgsz5rwyr0206
xold_sacqtplgsz5rwyr0307 xold_sacqtplgsz5rwyr0408 xold_sacqtplgsz5rwyr0509
      xold_sacqtplgsz5rwyr0610 xold_sacqtplgsz5rwyr0711 xold_sacqtplgsz5rwyr0812 xold_sacqtplgsz5rwyr0913
xold_sacqtplgsz5rwyr1014 xold_sacqtplgsz5rwyr1115 xold_sacqtplgsz5rwyr1216
      xold_sacqtplgsz5rwyr1317 xold_sacqtplgsz5rwyr1418  ;
      if a ;            by  cusip6lstyrrw;
      if srw7acqtypegsz       = . then srw7acqtypegsz = 0; *initialize with zero if no information from the
rolling window 5 yrs gsz ge1mmusd sample;
      if srw7acqtypegsz       = . then srw7acqtypegsz = 0; *initialize with zero if no information from the
rolling window 5 yrs gsz ge1mmusd sample;
      if bdr5totdealsge1      = . then bdr5totdealsge1 = 0; *initialize with zero if no information from
the rolling window 5 yrs gsz ge1mmusd sample;
*Note that totdlsgsz5yr = 0 implies  NO gsz M&A transactions within the rolling window as of prior year;
run; * main vars for large size deals: srw7acqtypegsz bdr5totdealsge1  gsz geqmmus=1;

proc means data=serlacqbygszflx5rw n mean min p1 median p99 max ;
      var bdr5totdealsge1  ; class srw7acqtypegsz; run;
      ************** GOOD FILE WITH FlexRW classifPriorYR AT YEAR LEVEL for gsz )ge1mmusd=1) sample  and
whole sample ;
data serialac.serlacqbygszflx5rw; set serlacqbygszflx5rw; run; * firm-year observations with serial
acquirer classification;
/*Adj in Stata to take into consideration the prior deals within a calendar year at the deal level within
each calendar year, still keeping the ex-ante information expectation*/
*adj for flx rw 5 yrs all-deals;
global trhshldallocc 2
global trhshldallspr 5
global trhshldallmrth 8
replace srw5acqtype =4 if srw5acqtype==0 & ainoverall> $trhshldallmrth
replace srw5acqtype =3 if srw5acqtype==0 & ainoverall> $trhshldallspr
replace srw5acqtype =2 if srw5acqtype==0 & ainoverall> $trhshldallocc
replace srw5acqtype =1 if srw5acqtype==0 & ainoverall> 0
replace srw5acqtype =4 if srw5acqtype==1 & ainoverall> $trhshldallmrth
replace srw5acqtype =3 if srw5acqtype==1 & ainoverall> $trhshldallspr
replace srw5acqtype =2 if srw5acqtype==1 & ainoverall> $trhshldallocc
replace srw5acqtype =4 if srw5acqtype==2 & ainoverall> $trhshldallmrth
replace srw5acqtype =3 if srw5acqtype==2 & ainoverall> $trhshldallspr
replace srw5acqtype =4 if srw5acqtype==3 & ainoverall> $trhshldallmrth
```