Talk Less, Learn More: Strategic Disclosure in Response to Managerial Learning from the Options Market

Co-authored by Yangyang Chen, Jeffrey Ng, and Xin Yang

STATA and SAS Code

```
**********************************
***              Options trading            ***
**********************************
*** CRSP stock price
use cusip permno date prc using "dsf_19252017.dta", clear
replace prc=abs(prc)
rename cusip cusip8
sort permno date
duplicates drop
duplicates list permno date
save stkprice, replace


*** CRSP stock volume
use cusip permno date prc vol using "dsf_19252017.dta", clear
duplicates drop
/*Volume is set to -99 if the value is missing.*/
drop if vol==. | vol==-99

rename cusip cusip8
gen year=year(date)
bys cusip8 year: egen stkvol=sum(vol)
gen value=abs(prc)*vol
bys cusip8 year: egen dstkvol=sum(value)
```

```
keep cusip8 permno year stkvol dstkvol
duplicates drop
duplicates list cusip8 year
save stkvol, replace

use "stocknames.dta", clear
sort ncusip namedt
save "stocknames.dta", replace



*** Option data
** 1996-2009
foreach j in 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008a 2008b
2008c 2008d 2009a 2009b 2009c 2009d {

use date cp_flag strike_price best_bid best_offer volume open_interest cusip using `j'.dta, clear
rename cusip ncusip

* generate new normal date
tostring date, replace
gen year=substr(date,1,4)
gen month=substr(date,5,2)
gen day=substr(date,7,2)
drop date
destring year month day, replace
gen date=mdy(month,day,year)
format date %td

* Link ncusip to permno
sort ncusip date
joinby ncusip using "stocknames.dta"
```

```
drop permco ticker comnam hexcd exchcd shrcd shrcls st_date end_date namedum
keep if date>=namedt & date<=nameenddt
drop namedt nameenddt cusip siccd

gen xprc=strike_price/1000
gen optprc=(best_bid+best_offer)/2

gen qtr=1
replace qtr=2 if month==4|month==5|month==6
replace qtr=3 if month==7|month==8|month==9
replace qtr=4 if month==10|month==11|month==12

merge m:1 permno date using stkprice
keep if _m==3
drop _m

gen value=optprc*volume*100
sort ncusip year qtr
bys ncusip year qtr: egen optvol=sum(volume)
bys ncusip year qtr: egen doptvol=sum(value)

* Moneyness
gen temp1=abs(ln(prc/xprc))
bys ncusip year qtr: egen temp2=sum(volume)
bys ncusip year qtr: egen money=mean(temp1)
bys ncusip year qtr: egen vwmoney=sum(temp1*volume/temp2)
drop temp*

* Open interest
bys ncusip year qtr: egen open=mean(open_interest)
```

```
keep ncusip permno year qtr optvol doptvol money vwmoney open
duplicates drop

sort ncusip year qtr
save `j'_Clean.dta, replace
}


** 2010-2016
foreach j in 2010a 2010b 2010c 2010d 2011a 2011b 2011c 2011d 2012a 2012b 2012c 2012d ///
2013a 2013b 2013c 2013d 2014a 2014b 2014c 2014d 2015a 2015b 2015c 2015d 2016a 2016b
2016c 2016d {

use date cp_flag strike_price best_bid best_offer volume open_interest cusip using `j'.dta, clear
rename cusip ncusip

gen year=year(date)
gen month=month(date)

* Link ncusip to permno
sort ncusip date
joinby ncusip using "stocknames.dta"
drop permco ticker comnam hexcd exchcd shrcd shrcls st_date end_date namedum
keep if date>=namedt & date<=nameenddt
drop namedt nameenddt cusip siccd

gen xprc=strike_price/1000
gen optprc=(best_bid+best_offer)/2

gen qtr=1
replace qtr=2 if month==4|month==5|month==6
```

```
replace qtr=3 if month==7|month==8|month==9
replace qtr=4 if month==10|month==11|month==12

merge m:1 permno date using stkprice
keep if _m==3
drop _m

gen value=optprc*volume*100
sort ncusip year qtr
bys ncusip year qtr: egen optvol=sum(volume)
bys ncusip year qtr: egen doptvol=sum(value)

gen temp1=abs(ln(prc/xprc))
bys ncusip year qtr: egen temp2=sum(volume)
bys ncusip year qtr: egen money=mean(temp1)
bys ncusip year qtr: egen vwmoney=sum(temp1*volume/temp2)
drop temp*

bys ncusip year qtr: egen open=mean(open_interest)

keep ncusip permno year qtr optvol doptvol money vwmoney open
duplicates drop

sort ncusip year qtr
save `j'_Clean.dta,replace
}


*** Merge data
use 1996_Clean, clear
```

```
foreach j in 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008a 2008b 2008c
2008d 2009a 2009b 2009c 2009d ///
2010a 2010b 2010c 2010d 2011a 2011b 2011c 2011d 2012a 2012b 2012c 2012d 2013a 2013b
2013c 2013d ///
2014a 2014b 2014c 2014d 2015a 2015b 2015c 2015d 2016a 2016b 2016c 2016d {
append using `j'_Clean.dta
}
sort ncusip year qtr
save OptionTrading.dta, replace



*** Convert quarterly data to annual data
use OptionTrading, clear
sort ncusip year qtr

foreach j in optvol doptvol {
by ncusip year: egen N`j'=sum(`j')
drop `j'
rename N`j' `j'
}

foreach j in money vwmoney open{
by ncusip year: egen N`j'=mean(`j')
drop `j'
rename N`j' `j'
}

drop qtr
duplicates drop
sort ncusip year
duplicates drop permno year, force
```

```
save OptionTrading_Annual.dta, replace


*** Relative trading activity in options and stock
use OptionTrading_Annual, clear
merge 1:1 permno year using stkvol
keep if _m==3
drop _m
gen d_os=ln(doptvol/dstkvol)
gen s_os=ln(optvol/stkvol)
keep permno ncusip year d_os s_os
save OS, replace



**********************************************
***                Management forecast                ***
**********************************************
* Use SAS code to link IBES ticker with gvkey and permno
* set libname;
libname linkwrds 'D:\Data\link\LinkingsuitByWrds';
libname link 'D:\Data\link';
libname linkm 'D:\Data\link\ibes_ticker_permno';
libname comp 'D:\Data\COMP\compa';
libname crsp 'D:\Data\CRSP';
libname ibes 'D:\Data\IBES';

* import mf data;
proc import out=mf
        datafile='D:\Data\IBES\mf19922017.dta'
        dbms=dta replace;
```

```
run;

* lnk ibes ticker with permno;
* lnk table:
* 1) WRDS Linking Suit;
* 2) IBES_Ticker_permno - a file that matches IBES-Ticker with CRSP Permno available here:
http://www.bhwang.com/research.html (includes data until the end of 2017)
* 3) iclink;

* Check for iclink;
* keep unique permno for each ticker;
proc sort data=link.iclink; by permno score; run;
data iclink; set link.iclink; by permno score; if first.permno;run;

* keep unique ticker for each permno;
proc sort data=iclink; by ticker score; run;
data iclink; set iclink; by ticker score; if first.ticker; run;


* Link permno using the three link table;
proc sql;
        create table gdetlink1
        as select a.*,
                b.permno,
                c.permno as cpermno,
                d.permno as dpermno
        from mf as a
        left join linkwrds.ibes_tkr_permno (where=(score=1 or 2)) as b
        on a.ticker=b.ticker
        and ((b.sdate<=a.anndats<=b.edate) or (b.sdate<=a.anndats and b.edate=.) or (b.sdate=.
and a.anndats<=b.edate))
```

```
        left join iclink (where=(score=0 or 1)) as c
        on a.ticker=c.ticker
        left join linkm.ibes_ticker_permno (where=(dup1=0 and dup2=0 and
matching_inconsistencies="")) as d
        on a.ticker=d.ticker
        and (d.startdate<=a.anndats<=d.enddate)
        order by a.ticker, prd_yr, prd_mon, anndats,
                pdicity, measure, val_1, val_2;
quit;


* check dup;
proc sort data=gdetlink1
        nouniquekey out=check;
        by ticker prd_yr prd_mon anndats
                pdicity measure val_1 val_2;
run;


data gdetlink1 (drop=cpermno dpermno);
        set gdetlink1;
        if permno=. and dpermno ^=.
        then permno=dpermno;
        if permno=. and dpermno=. and cpermno ^=.
        then permno=cpermno;
run;


proc sort
        data=gdetlink1 nodupkey;
        by ticker permno prd_yr prd_mon anndats
```

```
                pdicity measure val_1 val_2;
run;



* lnk mf to gvkey;
* lnk table:
* 1) cibeslnk;
* 2) comp.security;
proc import out=security
        datafile='D:\Data\COMP\gvkey_ibtic_link_sas.dta'
        dbms=dta replace;
run;

proc sql;
        create table gdetlink2 as
        select a.*, b.gvkey, c.gvkey as dgvkey
        from gdetlink1 as a
        left join link.cibeslnk as b
        on a.ticker=b.ticker
        and ((b.fdate<=a.anndats<=b.ldate) or (b.fdate<=a.anndats and b.ldate=.E)
        or (b.fdate=.B and a.anndats <=b.ldate))
        left join security as c
        on a.ticker=c.ibtic
        order by ticker, prd_yr, prd_mon, anndats,
                pdicity, measure, val_1, val_2;
quit;

proc sort data=gdetlink2
        nouniquekey out=check;
        by ticker gvkey prd_yr prd_mon anndats
```

```
                pdicity measure val_1 val_2;
run;

data gdetlink2 (drop=dgvkey);
   set gdetlink2;
        if gvkey=. and dgvkey ^=.
   then gvkey=dgvkey;
run;



proc sort
        data=gdetlink2 nodupkey;
        by ticker gvkey prd_yr prd_mon anndats
                pdicity measure val_1 val_2;
run;

proc sort
        data=gdetlink2 nodupkey;
        by ticker prd_yr prd_mon anndats
                pdicity measure val_1 val_2;
run;



* more gvkey from ccm;
proc sql;
        create table gdetlink3 as
        select a.*, b.gvkey as cgvkey
        from gdetlink2 as a
        left join crsp.ccmxpf_lnkhist
        (where=(linkprim in ('P', 'C') and linktype in ('LU', 'LC'))) as b
        on a.permno=b.lpermno
```

```
        and ((b.linkdt<=a.anndats<=b.linkenddt)
        or (b.linkdt<=a.anndats and b.linkenddt=.E) or (b.linkdt=.B and a.anndats<=b.linkenddt))
        order by ticker, prd_yr, prd_mon, anndats,
                pdicity, measure, val_1, val_2;
quit;

data gdetlink3 (drop=cgvkey);
        set gdetlink3;
        if gvkey=. and cgvkey ^=.
        then gvkey=cgvkey;
run;

proc sort
        data=gdetlink3 nodupkey;
        by ticker permno gvkey
                prd_yr prd_mon anndats
                pdicity measure val_1 val_2;
run;

*more permno from ccm;
proc sql;
        create table gdetlink4 as
        select a.*, b.lpermno
        from gdetlink3 as a
        left join crsp.ccmxpf_lnkhist
        (where=(linkprim in ('P', 'C') and linktype in ('LU', 'LC'))) as b
        on a.gvkey=b.gvkey
        and ((b.linkdt<=a.anndats<=b.linkenddt)
        or (b.linkdt<=a.anndats and b.linkenddt=.E) or (b.linkdt=.B and a.anndats<=b.linkenddt))
        order by ticker, prd_yr, prd_mon, anndats,
                pdicity, measure, val_1, val_2;
```

```
quit;

data gdetlink4 (drop=lpermno);
        set gdetlink4;
        if permno=. and lpermno ^=.
        then permno=lpermno;
run;

proc sort
        data=gdetlink4 nodupkey;
        by ticker permno gvkey
                prd_yr prd_mon anndats
                pdicity measure val_1 val_2;
run;


data gdet_lnk_new; set gdetlink4;run;

data gdet_lnk_new;
        set gdet_lnk_new;
        prd_datadate=intnx('month',mdy(prd_mon,1,prd_yr),0,'e');
        format prd_datadate MMDDYY10.;
        label prd_datadate="mf period end date";
run;

proc sort
        data=gdet_lnk_new nodupkey;
        by ticker prd_yr prd_mon anndats
                pdicity measure val_1 val_2;
run;
```

```
proc export data=gdet_lnk_new
        outfile='D:\Final version\gdet_lnk_new.dta'
        dbms=dta replace;
run;

* Frequency of management earnings forecasts
use "gdet_lnk_new.dta", clear
destring gvkey, replace
keep if pdicity=="QTR" | pdicity=="ANN"
keep if measure=="EPS"
drop if gvkey==. | permno==.

duplicates tag ticker anndats val_1 val_2 prd_datadate, gen(dup)
sort ticker anndats val_1 val_2 prd_datadate
bro if dup!=0
drop if pdicity=="QTR" & dup!=0
drop dup

duplicates list gvkey anndats val_1 val_2 prd_datadate
duplicates drop gvkey anndats val_1 val_2 prd_datadate, force
duplicates list permno anndats val_1 val_2 prd_datadate

* Drop pre-announcement
gen dif_daya=prd_datadate-anndats
count if dif_daya<0
drop if dif_daya<0

* freq of forecast
destring gvkey, replace
bys gvkey annyear: egen freq_all=count(anndats)
```

```
* width
* mean diff of all forecast
gen range=4 if inlist(range_desc,"02","14")
replace range=3 if inlist(range_desc,"01","06","08")
gen diff=val_2-val_1 if range==3
replace diff=0 if range==4
bys gvkey annyear: egen m_width=mean(diff)

* keep related variables
keep gvkey annyear freq_all m_width
duplicates drop
sort gvkey annyear
replace annyear=annyear-1
local mf_a freq_all m_width
foreach v of local mf_a {
        rename `v' `v'_f1
}
save mf_all, replace


* the stock price at the end of the year
use "msf_19252017.dta", clear
keep permno date prc altprc
gen year=year(date)
gen absprc=abs(prc)
gen absaltprc=abs(altprc)
gsort permno year -date
bys permno year: keep if _n==1
keep permno year absprc absaltprc
save prc, replace
```

```
**********************************************
***                Other control variables                ***
**********************************************
*** SEO
import excel using "SEOall.xls", sheet("SEO4") firstrow clear case(lower)
rename cusip cnum
drop if cnum==""
gen year=year(issuedate)
sort cnum year
bys cnum year: keep if _n==1
gen seo=1
keep cnum year seo
save seo, replace


*** Institutional ownership
* CRSP monthly data
use "msf_19252017.dta", clear
gen year=year(date)
gen month=month(date)
keep cusip ncusip date year month prc shrout permno
duplicates drop
drop cusip date
rename ncusip cusip8
drop if cusip8==""
duplicates list permno year month
duplicates list cusip8 year month
save csrp_msf_prcshr, replace

* import data
```

```
use "tfn_S34_19802017.dta", clear
drop if cusip==""
gen year=year(rdate)
gen month=month(rdate)
gen cnum=substr(cusip,1,6)
rename cusip cusip8

* drop duplicates
duplicates tag mgrno cusip8 year month, gen(dup)
bro if dup!=0
sort mgrno cusip8 year month
/*for dup!=0, keep obs with less missing variables*/
egen miss=rowmiss(fdate-dup)
sort mgrno cusip8 year month miss
bys mgrno cusip8 year month: keep if _n==1
drop dup miss

* keep related variables
keep rdate mgrno typecode cusip8 shares shrout1 prc change year month cnum
duplicates drop
rename prc prc1
rename shrout1 shrout_cda

merge m:1 cusip8 year month using csrp_msf_prcshr
drop if _merge==2
drop _m

* total IH
sort cusip8 year month
bys cusip8 year month: egen ihtt=sum(shares)
```

```
* gen mean value by cusip8 year
sort cusip8 year month
bys cusip8 year: egen mihtt=mean(ihtt)
bys cusip8 year: egen mshrout_cda=mean(shrout_cda)
drop if mshrout_cda==.

replace mihtt=0 if mihtt==.
gen pmihtt=mihtt/(mshrout_cda*1000000)
replace pmihtt==1 if pmihtt>1

* drop duplicates
keep cusip8 permno year pmihtt
duplicates drop
duplicates list cusip8 year
duplicates tag cusip8 year, gen(dup)
drop if permno==. & dup!=0
drop dup
duplicates list cusip8 year
save IH_Yr_Average.dta, replace

* link to gvkey
use IH_Yr_Average.dta, clear
keep if permno==.
merge 1:1 cusip8 year using "gvkey_cusip.dta"
keep if _m==3
drop _m
destring gvkey, replace
tempfile temp1
save "`temp1'", replace

* Link permno with gvkey
```

```
use "ccmxpf_lnkhist.dta" if inlist(linktype, "LC","LU")==1, clear
destring gvkey, replace
rename lpermno permno
joinby permno using IH_Yr_Average.dta
keep if year(linkdt)<=year & (year(linkenddt)>=year | linkenddt==.)
duplicates tag permno year, gen(duptag)
drop if duptag > 0 & linkprim != "P"
duplicates tag gvkey year, gen(dups)
drop if dups > 0 & linkprim != "P"
drop duptag dups
duplicates drop permno year, force
duplicates drop gvkey year, force
drop linkprim liid linktype linkdt linkenddt lpermco
tempfile temp2
save "`temp2'", replace

* Append
use "`temp1'", clear
append using "`temp2'"

duplicates tag gvkey year, gen(dup)
bro if dup!=0
sort gvkey year
drop if permno==. & dup!=0
drop dup
duplicates list gvkey year
drop cusip cnum
save IH_Yr_Average_gvkey.dta, replace


*** Stock volatility
```

```
use "dsf_19252017.dta", clear
gen year=year(date)
gen month=month(date)
gen day=day(date)
keep permno year month day ret
drop if ret==.
duplicates drop
bys permno year: egen stkvol=sd(ret)
replace stkvol=stkvol*100
keep permno year stkvol
duplicates drop
save stkvol, replace


*** Stock Turnover
use "msf_19252017.dta", clear
gen year=year(date)
gen month=month(date)
keep if year>=1980
bys permno year month: gen turnover=(vol*100)/(shrout*1000)
bys permno year: egen stk_turnover=mean(turnover)
keep permno year stk_turnover
duplicates drop
duplicates list permno year
save stk_turnover, replace


*** Analyst coverage
****link gvkey with ibes ticker;
libname comp 'D:\Data\COMP\compa';
libname linkwrds 'D:\Data\link\LinkingsuitByWrds';
```

```
libname link 'D:\Data\link';
libname ibes 'D:\Data\IBES\Analyst';
libname crsp 'D:\Data\CRSP';

data comp1;
  set comp.funda;
  where fyear>=1986 and indfmt='INDL' and datafmt='STD' and popsrc='D' and consol='C';
  keep gvkey datadate fyear;
run;

proc sort data=comp1 out=comp1 nodupkey;
   by gvkey fyear descending datadate;
run;

* cibeslnk;
data cibeslnk;
  set link.cibeslnk;
  if ldate =.E then ldate ='31DEC2017'd ;
run;

proc import out=security
            datafile='D:\Data\COMP\compa\gvkey_ibtic_link_sas.dta'
            dbms=dta replace;
run;

*to add the IBES ticker;
proc sql;
      create table comp2
      as select a.*,b.ticker, c.ibtic
      from comp1 a
      left join cibeslnk b
```

```
      on a.gvkey= b.gvkey and b.fdate <= a.datadate <=b.ldate
            left join security c
            on a.gvkey=c.gvkey;
quit;

data comp2 (drop=ibtic);
    set comp2;
          if ticker="" and ibtic^="" then ticker=ibtic;
run;

proc sort data=comp2 out=comp3 nodupkey;
 where not missing(ticker);
 by gvkey datadate;
run;



* merge with statsumu_epsus;
data comp4;
set comp3;
IBES_date = intnx('month',datadate, 0,'end');
format datadate YYMMDD10. IBES_date YYMMDD10.;
run;



proc sql;
      create table comp5
      as select a.*, b.*
      from comp4 a left join ibes.statsumu_epsus b
      on a.ticker = b.ticker and a.datadate = b.fpedats and b.statpers <= a.IBES_date  and
b.statpers >= a.IBES_date - 90 and
            b.FPI = '1' and b.measure ='EPS';
```

```
quit;

proc sort data=comp5 out=comp6 nodupkey;
by gvkey datadate descending statpers;
run;


*right before the year end;
data comp7;
set comp6;
where not missing(numest);
by gvkey datadate descending statpers;
if first.datadate = 1;
num_anal = numest;
for_std = stdev;
for_mean = meanest;
dispersion = for_std/abs(for_mean);
format IBES_date yymmdd10. statpers yymmdd10. fpedats yymmdd10.;
run;


data comp7;
   set comp7 (rename=(fyear=year));
   where not missing(gvkey);
run;

data comp7(drop=gvkey1);
   set comp7(rename=(gvkey=gvkey1));
   gvkey=gvkey1*1;
run;
```

```
proc sort data=comp7 nodupkey;
by gvkey year;
run;

proc export data=comp7
        outfile="D:\Data\MyData\Analyst\Analstsum_coverage"
                dbms=stata replace;
run;



*** Basic data from Compustat
use gvkey cik tic cusip datadate fyear indfmt datafmt popsrc consol at sic sich dltt dlc re oibdp
dp ///
emp lt act lct invt rect oiadp ni csho prcc_f ceq xido sstk dltis ib sale ///
dvp dvc capx fopt ibc dpc txdc esubc sppiv fopo sstk prstkc che txdb xrd revt ///
using "D:\Data\COMP\compa\comp_funda_1950_2017.dta" ///
if      (indfmt=="INDL") & (datafmt=="STD") & (popsrc=="D") & (consol=="C") &
(gvkey != "")   & (fyear != .), clear

drop indfmt datafmt popsrc consol
drop if at==.
rename fyear year
destring gvkey, replace

* drop duplicates
gsort gvkey year -datadate
bys gvkey year: keep if _n==1
keep if (year>=1996) & (year<=2017)

gen profitmarg=oibdp/sale
gen roa=ib/at
```

```
gen lev=(dltt+dlc)/at

local temp act lct re oibdp dp
foreach v of local temp {
        replace `v'=0 if `v'==.
}
gen z_score=1.2*((act-lct)/at)+1.4*(re/at)+3.3*((oibdp-
dp)/at)+0.6*(prcc_f*csho/lt)+0.999*(sale/at)

merge 1:1 gvkey year using "IH_Yr_Average_gvkey.dta"
drop if _m==2
drop _m

merge 1:1 gvkey year using "Analstsum_coverage", keepusing(num_anal for_mean for_std
dispersion)
drop if _m==2
drop _m
save all, replace



* Link with permno
use "ccmxpf_lnkhist.dta" if inlist(linktype, "LC","LU")==1, clear
destring gvkey, replace
rename lpermno permno
joinby gvkey using all
keep if year(linkdt)<=year & (year(linkenddt)>=year | linkenddt==.)
duplicates tag permno year, gen(duptag)
drop if duptag > 0 & linkprim != "P"
duplicates tag gvkey year, gen(dups)
drop if dups > 0 & linkprim != "P"
drop duptag dups
```

```
duplicates list gvkey year
duplicates drop gvkey year, force
duplicates list permno year
duplicates drop permno year, force
drop linkprim liid linktype linkdt linkenddt lpermco
save all, replace


use all, clear
* Option trading volume
merge 1:1 permno year using "OptionTrading_Annual.dta"
drop if _m==2
drop _m

* stock volatility
merge 1:1 permno year using "stkvol.dta"
drop if _m==2
drop _m

* get ncusip
cap drop ncusip
sort permno year
joinby permno using "stocknames.dta"
drop permco ticker comnam hexcd exchcd shrcd shrcls st_date end_date namedum
keep if datadate>=namedt & datadate<=nameenddt
drop namedt nameenddt

* SEO
gen cnum=substr(ncusip,1,6)
merge m:1 cnum year using seo
drop if _m==2
```

```
drop _m
replace seo=0 if seo==.

replace pmihtt=0 if pmihtt==.
replace num_anal=0 if num_anal==.
gen lnanalst=ln(1+num_anal)

* replace missing sich from comp with siccd from crsp.stocknames
duplicates list gvkey year
order sich, after(sic)
order siccd, after(sic)
count if sich==.
replace siccd=. if siccd==0
replace sich=siccd if sich==. & siccd!=.
save all, replace



*** Other Control variables
libname comp "D:\Data\COMP\compa";
libname link 'D:\Data\link';
libname crsp 'D:\Data\CRSP';

* Basic financial data from Compustat;
%let fyear1= 1984; %let fyear2= 2017;
%let vars = gvkey datadate fyear sich cusip cik
            prcc_f csho AJEX at ceq ib OANCF sale
            INDFMT DATAFMT POPSRC CONSOL;

data controls;
  set comp.funda (keep= &vars);
  where fyear between &fyear1 and &fyear2;
```

```
   if indfmt='INDL' and datafmt='STD' and popsrc='D' and consol='C';
   drop indfmt datafmt popsrc consol;
   endfyr= datadate; format endfyr date9.;
   begfyr= intnx('month',endfyr,-11,'beg'); format begfyr date9.;
run;
proc sort; by gvkey endfyr; run;


* One year lag for at, ib, and sale;
proc sql;
   create table controls as
   select a.*,
   b.at as lat1,
   b.ib as lib1,
   b.sale as lsale1
from controls a left join controls b
on a.gvkey=b.gvkey and a.fyear=b.fyear+1;
quit;
proc sort nodupkey; by gvkey datadate; run;


* Firm size;
* Book-to-market;
* Loss;
* Sale growth;
data controls;
  set controls;
  mv=prcc_f*csho;
  logmv=log(mv);
  logat=log(at);
  if ceq>0 and mv>0 then bm=ceq/mv; else bm=.;
```

```
 if not missing(ib) then do;
  loss=0;
  if ib<0 then loss=1;
 end;
 if lat1>0 then sales_growth=(sale-lsale1)/lat1; else sales_growth=.;
run;


* Link gvkey with permno;
proc sql;
        create table controls as
        select a.*, b.lpermno as permno
        from controls as a
        left join crsp.ccmxpf_lnkhist
        (where=(linkprim in ('P', 'C') and linktype in ('LU', 'LC'))) as b
        on a.gvkey=b.gvkey
        and ((b.linkdt<=a.endfyr<=b.linkenddt)
        or (b.linkdt<=a.endfyr and b.linkenddt=.E) or (b.linkdt=.B and a.endfyr=<b.linkenddt));
quit;

data controls;
 set controls;
 if not missing(permno);
run;


* Earnings volatility;
proc sql undo_policy = none;
 create table controls
 as select distinct a.*,count(b.roa) as numroa
 from controls a left join controls b
```

```
  on a.gvkey=b.gvkey
  and 0<a.fyear-b.fyear<=10
  and not missing(b.roa)
  group by a.gvkey, a.fyear
  order by a.gvkey, a.fyear;
quit;
proc sort nodupkey; by gvkey fyear;run;

proc sql undo_policy = none;
  create table controls
  as select distinct a.*,std(b.roa) as earnvol
  from controls a left join controls b
  on a.gvkey=b.gvkey
  and 0<a.fyear-b.fyear<=10
  and a.numroa>=5 and not missing(b.roa)
  group by a.gvkey, a.fyear
  order by a.gvkey, a.fyear;
quit;
proc sort nodupkey; by gvkey fyear;run;


* Add CRSP siccd code;
proc sql;
      create table controls
      as select a.*,b.siccd
      from controls a  left join crsp.stocknames b
      on a.permno = b.permno and a.datadate >= b.namedt and a.datadate <= b.nameenddt;
quit;
proc sort nodupkey; by gvkey datadate;run;

* Fill the missing sich with siccd;
```

```
data controls; set controls;
  if siccd=0 then siccd=.;
  if sich=. & siccd ^=. then sich=siccd;
run;

*Fill the missing sich;
data sich;
        set controls;
        if not missing(sich);
        keep gvkey sich;
run;
proc sort nodupkey; by gvkey; run;
proc sql;
        create table controls as
        select a.*,b.sich as sica
        from controls a left join sich b
        on a.gvkey=b.gvkey;
quit;

data controls;
 set controls;
 if missing(sich) then sich=sica;
 run;


* FPS: equals 1 if the firm is in the biotech, computer, electronics, or retail industry, and 0
otherwise;
data controls;
 set controls;
  fpslitig=0;
  if not missing(sich) then do;
```

```
  if sich>=2833 and sich<=2836 then fpslitig=1;
  else if sich>=8731 and sich<=8734 then fpslitig=1;
  else if sich>=3570 and sich<=3577 then fpslitig=1;
  else if sich>=7370 and sich<=7374 then fpslitig=1;
  else if sich>=3600 and sich<=3674 then fpslitig=1;
  else if sich>=5200 and sich<=5961 then fpslitig=1;
 end;
run;


* CRSP monthly data;
data mret2;
  set controls;
  keep permno gvkey datadate begfyr endfyr;
run;
proc sql;
  create table mret2 as select *
  from mret2 as a, crsp.msf0918 as b
  where a.permno = b.permno
  and intck('month',a.begfyr,b.date) between 0 and 11
  and not missing(a.permno);
quit;
proc sort nodupkey; by gvkey datadate date; run;


* Indices - Year-end Cap. Deciles with Monthly Returns - NYSE/AMEX/NASDAQ;
proc sql;
  create table mret2 as select a.*, b.decret as sret, b.capn
  from mret2 as a, crsp.ermport1 as b
  where a.permno = b.permno and a.date=b.date;
quit;
```

```
proc sort nodupkey; by gvkey datadate date; run;

* Stock - Market Indexes Monthly NYSE/AMEX/NASDAQ/ARCA;
proc sql;
  create table mret2 as select a.*,b.vwretd
  from mret2 as a, crsp.msi as b
  where a.date=b.date;
quit;
proc sort nodupkey; by gvkey datadate date; run;

* delist return;
data delist; set crsp.dsedelist;
        keep permno dlstdt dlstcd dlpdt dlret;
        format dlstdt dlpdt date9.;
run;
proc sort; by permno dlstdt; run;

proc sql;
        create table mret2
        as select a.*, b.dlstdt, b.dlret, b.dlstcd
        from mret2 as a left join delist as b
        on a.permno = b.permno
        and a.begfyr <= b.dlstdt <= a.endfyr;
quit;


* Return adjusted by dlret;
data mret2; set mret2;
        ret1 = ret;
        if year(dlstdt) = year(date) and month(dlstdt) = month(date) and not missing(dlret)
                then ret1 = dlret;
```

```
        if year(dlstdt) = year(date) and month(dlstdt) = month(date) and missing(dlret)
                then ret1 = -0.999999;
        if not missing(ret1) and ret1 <= -1 then ret1 = -0.999999;
        drop dlret ret;
run;


* Stock return;
data mret2;
  set mret2 (rename=(ret1=ret));
  where not missing(ret);
  logret=log(1+ret);
  logsret=log(1+sret);
  logvwretd=log(1+vwretd);
  if missing(logret) then logsret=0;
  if missing(logret) then logvwretd=0;
keep gvkey permno datadate date gvkey ret sret
logret logsret logvwretd vol shrout;
run;


* Buy-and-hold size-adjusted return;
* Market-adjusted 12-month stock return;
* Turnover;
data mret3; set _null_;
proc sql undo_policy = none;
  create table mret2 as select * from mret2
    group by gvkey, datadate
    having count(*)=12
    order by gvkey, datadate, date;
```

```
   create table mret3 as select distinct
     gvkey, datadate,
     exp(sum(logret))-1 as ret12,
          exp(sum(logsret))-1 as sret12,
          exp(sum(logvwretd))-1 as mret12,
          sum(vol/shrout) as mturnover12
     from mret2
   group by gvkey, datadate
   order by gvkey, datadate, date;
quit;
proc sort nodupkey data=mret3; by gvkey datadate;
data mret3;
set mret3;
saret12=ret12-sret12;
maret12=ret12-mret12;
run;


* Return Standard deviation;
* Return skewness;
data mret_otherst;
  set controls;
  keep permno gvkey datadate begfyr;
run;

proc sql;
  create table mret_otherst as select *
  from mret_otherst as a, crsp.msf0918 as b
  where a.permno = b.permno and intck('month',a.begfyr,b.date) between 0 and 11
  and not missing(a.permno);
quit;
```

```
proc sort nodupkey; by gvkey datadate date; run;

proc sql undo_policy = none;
   create table mret_otherst as select * from mret_otherst
   group by gvkey, datadate
   having count(*)=12
   order by gvkey, datadate, date;
quit;

proc means data=mret_otherst noprint;
by gvkey datadate;
var ret;
output out=mret_otherst1 std=mstdret12 skew=mskewret12;
run;


*merge Compustat data with Crsp data;
proc sql;
create table controls as
select a.*,
  b.saret12 as csaret12,
  b.maret12 as cmaret12,
  b.ret12 as cret12,
  b.sret12 as csret12,
  b.mret12 as cmret12,
  b.mturnover12 as cmturnover12,
  c.mstdret12 as cmstdret12,
  c.mskewret12 as cmskewret12
from controls a left join mret3 b
on a.gvkey=b.gvkey and a.datadate=b.datadate
left join mret_otherst1 as c
```

```
on a.gvkey=c.gvkey and a.datadate=c.datadate;
quit;
proc sort nodupkey; by gvkey datadate; run;


*Class action litigation risk;
data controls;
 set controls;

kslagc_litigation=-7.883
+0.566*fpslitig
+0.518*logat
+0.982*sales_growth
+0.379*cmaret12
-0.108*cmskewret12
+25.635*cmstdret12
-(0.00007/1000)*cmturnover12;

data controls;
  set controls;
  gkslagc_litigation=pdf('normal',kslagc_litigation,0,1);
 run;

data mfcontro; set controls; run;

* Export to stata file;
proc export data=mfcontro
          outfile='D:\Data\mfcontro.dta'
          dbms=dta replace;
run;
```

```
use "D:\Data\mfcontro.dta", clear
keep gvkey fyear logmv bm loss earnvol csaret12 gkslagc_litigation
destring gvkey, replace
duplicates list gvkey fyear
rename fyear year
save contro2, replace


*** Financial statement comparability
use "verdi_2011_jn_benefitsfinancial_data_2013.dta", clear
destring gvkey, replace
gsort gvkey year -datadate
bys gvkey year: keep if _n==1
save fin_compa, replace


*** Cash flow volatility
use gvkey cik cusip datadate fyear indfmt datafmt popsrc consol at ib oancf sale revt ///
using "comp_funda_1950_2018.dta" ///
if        (indfmt=="INDL") & (datafmt=="STD") & (popsrc=="D") & (consol=="C") &
(gvkey != "")   & (fyear != .), clear

gsort gvkey fyear -datadate
bys gvkey fyear: keep if _n==1

destring gvkey, replace
xtset gvkey fyear
gen meancfo4 = (l1.oancf + l2.oancf + l3.oancf + l4.oancf)/4
gen cfovol4 = (((l1.oancf-meancfo4)^2 + (l2.oancf-meancfo4)^2 + (l3.oancf-meancfo4)^2 +
(l4.oancf-meancfo4)^2)/4)^.5
```

```
replace cfovol4=cfovol4/at

keep gvkey fyear cfovol4
rename fyear year
duplicates list gvkey year
save "Volatility_a.dta", replace


*** Managerial insider option trading
use "TFN_Table2_19962018.dta", clear
drop if cleanse=="S" | cleanse=="A"
keep if formtype=="4"
keep if inlist(rolecode1,"CEO","CFO","CB","CO","D","O","P","VP") |
inlist(rolecode2,"CEO","CFO","CB","CO","D","O","P","VP") | ///
inlist(rolecode3,"CEO","CFO","CB","CO","D","O","P","VP") |
inlist(rolecode4,"CEO","CFO","CB","CO","D","O","P","VP")
keep if trancode=="P" | trancode=="S"
tab derivative
keep if derivative=="OPTNS" | derivative=="CALL" | derivative=="PUT"

egen cusip=concat(cusip6 cusip2)
drop if cusip==""
gen year=year(trandate)
tab year
keep if inrange(year,1996,2017)

gen de_mam_mktvau=shares*xprice
keep cusip year trandate de_mam_mktvau

* link with permno
joinby cusip using "stocknames.dta"
```

```
drop ticker comnam hexcd exchcd shrcd shrcls st_date end_date namedum ncusip siccd permco
keep if trandate>=namedt & trandate<=nameenddt
drop namedt nameenddt

* total insider dollar option trading volume
bys permno year: egen de_mam_mktvau_y=sum(de_mam_mktvau)

keep permno year de_mam_mktvau_y
duplicates drop
sort permno year
save de_mam_mktvau_y, replace



*** one-year forward investment
use gvkey datadate fyear indfmt datafmt popsrc consol at capx xrd ppent ///
using "comp_funda_1950_2018.dta" ///
if        (indfmt=="INDL") & (datafmt=="STD") & (popsrc=="D") & (consol=="C") &
(gvkey != "")   & (fyear != .), clear

drop indfmt datafmt popsrc consol
drop if at==.
rename fyear year
destring gvkey, replace
keep if (year>=1996) & (year<=2017)
tab year

xtset gvkey year
gen INV3=capx/l.ppent

replace year=year-1
rename INV3 INV3_f1
```

```
keep gvkey year INV3_f1
duplicates list gvkey year
save INV, replace



*** The current year MEF
use "gdet_lnk_new.dta", clear
destring gvkey, replace
keep if pdicity=="QTR" | pdicity=="ANN"
keep if measure=="EPS"

drop if gvkey==. | permno==.

duplicates tag ticker anndats val_1 val_2 prd_datadate, gen(dup)
sort ticker anndats val_1 val_2 prd_datadate
bro if dup!=0
drop if pdicity=="QTR" & dup!=0
drop dup

duplicates list gvkey anndats val_1 val_2 prd_datadate
duplicates drop gvkey anndats val_1 val_2 prd_datadate, force
duplicates list permno anndats val_1 val_2 prd_datadate

* Drop pre-announcement
gen dif_daya=prd_datadate-anndats
count if dif_daya<0
drop if dif_daya<0

* freq of forecast
destring gvkey, replace
bys gvkey annyear: egen freq_all=count(anndats)
```

```
* keep related variables
keep gvkey annyear freq_all
duplicates drop
sort gvkey annyear
rename annyear year
save freq_all, replace


*** Annual 8k
use "8k_98_119.dta", clear
keep cik fdate nitem
keep if cik!=""
destring cik, replace
gen year=year(fdate)

gen vol=1 if (year<2004) & (nitem=="5" | nitem=="9" | nitem=="12")
replace vol=1 if (year>=2004) & (nitem=="8.01" | nitem=="7.01" | nitem=="2.02")
replace vol=0 if vol==.

merge m:1 cik fdate using "cikgvkey.dta"
keep if _m==3
drop _m

bys gvkey year: egen countvol=sum(vol)
keep gvkey year countvol
duplicates drop
replace year=year-1
rename countvol countvol_f1
save 8k_ann, replace
```

```
*** Conference call
use "SHAREME_confcallanalytics02to15.dta", clear
gen year=year(rdq)
tab year
destring gvkey, replace
keep gvkey permno year rdq num_words_comp_qa
replace num_words_comp_qa=0 if num_words_comp_qa==.
gen lnqawords=ln(1+num_words_comp_qa)
bys gvkey year: egen m_qawords=mean(lnqawords)
keep gvkey year m_qawords
duplicates drop
replace year=year-1
rename m_qawords m_qawords_f1
save cofcall, replace


*** Good and bad news
use "mf_surprise.dta", clear
destring gvkey, replace
keep if pdicity=="QTR" | pdicity=="ANN"
keep if measure=="EPS"
drop if gvkey==. | permno==.

duplicates tag ticker ann_date val_1 val_2 prd_datadate, gen(dup)
sort ticker ann_date val_1 val_2 prd_datadate
bro if dup!=0
drop if pdicity=="QTR" & dup!=0
drop dup

duplicates list gvkey ann_date val_1 val_2 prd_datadate
```

```
duplicates drop gvkey ann_date val_1 val_2 prd_datadate, force
duplicates list permno ann_date val_1 val_2 prd_datadate

* Drop pre-announcement
gen dif_daya=prd_datadate-ann_date
count if dif_daya<0
drop if dif_daya<0

* freq of forecast
destring gvkey, replace
bys gvkey annyear: egen freq_all=count(ann_date)

*** Proportion of good and bad news
gen pc_md=(md_mfn_e1>0)
replace pc_md=. if md_mfn_e1==.
gen nc_md=(md_mfn_e1<0)
replace nc_md=. if md_mfn_e1==.
bys gvkey annyear: egen freq_pcmd_all=sum(pc_md)
bys gvkey annyear: egen freq_ncmd_all=sum(nc_md)
bys gvkey annyear: gen prop_pcmd_all=freq_pcmd_all/freq_all
bys gvkey annyear: gen prop_ncmd_all=freq_ncmd_all/freq_all

* keep related variables
keep gvkey annyear prop_pcmd_all prop_ncmd_all
duplicates drop
sort gvkey annyear
duplicates list gvkey annyear
rename annyear year
replace year=year-1
local var prop_pcmd_all prop_ncmd_all
foreach v of local var {
```

```
        rename `v' `v'_f1
}
save gb_all, replace


*** Firms with options trading
use all, clear
rename year annyear
merge 1:1 gvkey annyear using mf_all
drop if _m==2
drop _m
rename annyear year

* annual and quarterly forecast
gen forecast_all_f1=1 if freq_all_f1!=.
replace forecast_all_f1=0 if freq_all_f1==.
replace freq_all_f1=0 if freq_all_f1==.
gen lnfreq_all_f1=ln(1+freq_all_f1)

* merge other control variables
merge 1:1 gvkey year using contro2
drop if _m==2
drop _m

merge 1:1 permno year using stk_turnover
drop if _m==2
drop _m

merge 1:1 permno year using prc
drop if _m==2
drop _m
```

```
gen width_all_f1=m_width_f1/absprc
replace width_all_f1=width_all_f1*(-100)
gen dispersion1=for_std/absprc
gen lnoptvol=ln(1+doptvol/1000000)

* drop missing variables
global contro logmv bm roa pmihtt csaret12 earnvol loss gkslagc_litigation lnanalst lev z_score
seo stkvol stk_turnover
egen miss1=rowmiss($contro)
drop if miss1 !=0
drop miss1
drop if lnoptvol==.

* winsorize
winsor2 lnoptvol lnfreq_all_f1 width_all_f1 dispersion1 $contro, cuts(1 99) replace
winsor2 sale, cuts(1 99) replace

* Mid Z_score
bys year: quantiles z_score, gen(x_zscore) nq(5)
gen mid_zscore=1 if x_zscore==3
replace mid_zscore=0 if x_zscore!=3 & x_zscore!=.

gen Ind = floor(sic/100)
tab year, gen(d_year)

keep if (year>=1996) & (year<=2016)
rename (forecast_all_f1 lnfreq_all_f1) (DumMF FreqMF)
rename (logmv bm roa pmihtt csaret12 earnvol loss gkslagc_litigation lnanalst lev mid_zscore
seo stkvol stk_turnover) ///
(Size BM ROA IO Ret EarnVol Loss Litigation Analyst Lev Mid_Zscore Issue Stkvol Turnover)
```

```
rename lnoptvol LnOptvol

global contro Size Lev BM ROA Loss IO Analyst Ret EarnVol Stkvol Turnover Litigation
Mid_Zscore Issue
save regall, replace


**************************************
***        Regression for Main Tables          ***
**************************************
* Table 3
use regall, clear
reghdfe DumMF LnOptvol $contro d_year*, absorb(gvkey) cluster(gvkey)
reghdfe FreqMF LnOptvol $contro d_year*, absorb(gvkey) cluster(gvkey)
reghdfe DumMF LnOptvol $contro d_year*, absorb(Ind) cluster(gvkey)
reghdfe FreqMF LnOptvol $contro d_year*, absorb(Ind) cluster(gvkey)

* Panel C of Table 4
use "CDS_incep_all.dta", clear
gen cdsyear=year(date_incep)
gen month=month(date_incep)
keep gvkey cdsyear
save "CDSfirm.dta", replace

use regall, clear
merge m:1 gvkey using "CDSfirm.dta"
drop if _m==2
drop _m

keep if (year>=1997) & (year<=2015)   /*CDS data is from 1997 to 2015*/
sort gvkey year
```

```
gen CDSinitiation=(year>=cdsyear)
replace CDSinitiation=0 if cdsyear==.
gen CDStraded=1 if year>=cdsyear
sort gvkey CDStraded
bys gvkey: replace CDStraded=CDStraded[_n-1] if mi(CDStraded)
replace CDStraded=0 if cdsyear==.
replace CDStraded=0 if CDStraded==.

reg DumMF LnOptvol $contro CDSinitiation CDStraded d_year*, r cluster(gvkey)
reg FreqMF LnOptvol $contro CDSinitiation CDStraded d_year*, r cluster(gvkey)
reghdfe DumMF LnOptvol $contro CDSinitiation CDStraded d_year*, absorb(Ind)
cluster(gvkey)
reghdfe FreqMF LnOptvol $contro CDSinitiation CDStraded d_year*, absorb(Ind)
cluster(gvkey)

* Table 5
use regall, clear
winsor2 money, cuts(1 99) replace
gen lnmoney=ln(money)

reghdfe LnOptvol lnmoney $contro d_year*, absorb(gvkey) cluster(gvkey)
reghdfe DumMF $contro d_year* (LnOptvol=lnmoney), absorb(gvkey) cluster(gvkey)
reghdfe FreqMF $contro d_year* (LnOptvol=lnmoney), absorb(gvkey) cluster(gvkey)
reghdfe LnOptvol lnmoney $contro d_year*, absorb(Ind) cluster(gvkey)
reghdfe DumMF $contro d_year* (LnOptvol=lnmoney), absorb(Ind) cluster(gvkey)
reghdfe FreqMF $contro d_year* (LnOptvol=lnmoney), absorb(Ind) cluster(gvkey)

* Table 6
use regall, clear
gen tobinq=((prcc_f*csho)+dltt)/at
gen cashflow2=(ib+dp)/at
```

```
merge 1:1 gvkey year using INV
drop if _m==2
drop _m

merge 1:1 gvkey year using freq_all
drop if _m==2
drop _m

replace freq_all=0 if freq_all==.
gen dum_all=(freq_all!=0)
gen lnfreq_all=ln(1+freq_all)

winsor2 INV3_f1 tobinq cashflow2 lnfreq_all, cuts(1 99) replace
replace INV3_f1=INV3_f1*100
gen lnoptvoltobinq=ln((1+doptvol/1000000)*tobinq)
winsor2 lnoptvoltobinq, cuts(1 99) replace
gen lnoptvoltobinqdum_all=lnoptvoltobinq*dum_all
gen lnoptvoltobinqlnfreq_all=lnoptvoltobinq*lnfreq_all
replace doptvol=doptvol/100000000
winsor2 doptvol, cuts(1 99) replace

reghdfe INV3_f1 lnoptvoltobinqdum_all lnoptvoltobinq dum_all doptvol tobinq cashflow2 Size
Ret d_year*, absorb(gvkey) cluster(gvkey)
reghdfe INV3_f1 lnoptvoltobinqdum_all lnoptvoltobinq dum_all doptvol tobinq cashflow2 Size
Ret d_year*, absorb(Ind) cluster(gvkey)
reghdfe INV3_f1 lnoptvoltobinqlnfreq_all lnoptvoltobinq lnfreq_all doptvol tobinq cashflow2
Size Ret d_year*, absorb(gvkey) cluster(gvkey)
reghdfe INV3_f1 lnoptvoltobinqlnfreq_all lnoptvoltobinq lnfreq_all doptvol tobinq cashflow2
Size Ret d_year*, absorb(Ind) cluster(gvkey)
```

```
* Table 7 and 8
use regall, clear

merge 1:1 permno year using de_mam_mktvau_y
drop if _m==2
drop _m

local var de_mam_mktvau_y
foreach v of local var {
        replace `v'=0 if `v'==.
        gen ln`v'=ln(1+`v')
}

merge 1:1 gvkey year using fin_compa
drop if _m==2
drop _m

merge 1:1 gvkey year using "Volatility_a.dta"
drop if _m==2
drop _m

hhi5 sale, by(Ind year)
winsor2 hhi_sale lnde_mam_mktvau_y dispersion m4_acctcomp cfovol4 profitmarg, cuts(1 99)
replace
save regall_c, replace

* By year median
local var m4_acctcomp dispersion hhi_sale cfovol4 profitmarg lnde_mam_mktvau_y
foreach v of local var {
   use regall_c, clear
```

```
    bys year: egen m_`v'=median(`v')
    gen h_`v'=1 if `v'>m_`v'
    replace h_`v'=0 if h_`v'==.
    replace h_`v'=. if `v'==.
    gen LnOptvolh_`v'=LnOptvol*h_`v'

    reghdfe DumMF LnOptvolh_`v' LnOptvol h_`v' $contro d_year*, absorb(gvkey)
cluster(gvkey)
    reghdfe FreqMF LnOptvolh_`v' LnOptvol h_`v' $contro d_year*, absorb(gvkey)
cluster(gvkey)
    reghdfe DumMF LnOptvolh_`v' LnOptvol h_`v' $contro d_year*, absorb(Ind) cluster(gvkey)
    reghdfe FreqMF LnOptvolh_`v' LnOptvol h_`v' $contro d_year*, absorb(Ind) cluster(gvkey)
}

use regall_c, clear
bys year: egen m_Litigation=median(Litigation)
gen h_Litigation=1 if Litigation>m_Litigation
replace h_Litigation=0 if h_Litigation==.
replace h_Litigation=. if Litigation==.
gen LnOptvolh_Litigation=LnOptvol*h_Litigation

reghdfe DumMF LnOptvolh_Litigation LnOptvol h_Litigation Size Lev BM ROA Loss IO
Analyst Ret EarnVol Mid_Zscore Issue Turnover Stkvol d_year*, absorb(gvkey) cluster(gvkey)
reghdfe FreqMF LnOptvolh_Litigation LnOptvol h_Litigation Size Lev BM ROA Loss IO
Analyst Ret EarnVol Mid_Zscore Issue Turnover Stkvol d_year*, absorb(gvkey) cluster(gvkey)
reghdfe DumMF LnOptvolh_Litigation LnOptvol h_Litigation Size Lev BM ROA Loss IO
Analyst Ret EarnVol Mid_Zscore Issue Turnover Stkvol d_year*, absorb(Ind) cluster(gvkey)
reghdfe FreqMF LnOptvolh_Litigation LnOptvol h_Litigation Size Lev BM ROA Loss IO
Analyst Ret EarnVol Mid_Zscore Issue Turnover Stkvol d_year*, absorb(Ind) cluster(gvkey)
```

```
* Table 9
use regall, clear
merge 1:1 permno year using OS
drop if _m==2
drop _m

winsor2 d_os s_os, cuts(1 99) replace
rename s_os ShOS
bys year: egen m_ShOS=median(ShOS)
gen h_ShOS=1 if ShOS>m_ShOS
replace h_ShOS=0 if h_ShOS==.
replace h_ShOS=. if ShOS==.
gen LnOptvolh_ShOS=LnOptvol*h_ShOS

reghdfe DumMF LnOptvolh_ShOS LnOptvol h_ShOS $contro d_year*, absorb(gvkey)
cluster(gvkey)
reghdfe FreqMF LnOptvolh_ShOS LnOptvol h_ShOS $contro d_year*, absorb(gvkey)
cluster(gvkey)
reghdfe DumMF LnOptvolh_ShOS LnOptvol h_ShOS $contro d_year*, absorb(Ind)
cluster(gvkey)
reghdfe FreqMF LnOptvolh_ShOS LnOptvol h_ShOS $contro d_year*, absorb(Ind)
cluster(gvkey)
```