```
* Computation of Variables of Tax Aggressiveness;
proc sort data=tax.aggressive; by gvkey fyear;run;
data tax; set tax.aggressive;by gvkey fyear;
if 4900<=sich<=4949 then delete;
if 6000<=sich<=6999 then delete;
if sich=. then delete;if OANCF=. then delete;
industry=100*floor(sich/100);
if fyr<=5 then calyear=fyear+1;else if fyr>=6 then calyear=fyear;
if tlcf=. then tlcf=0;if (pi-spi)>0 then singleCashETR=TXPD/(PI-SPI);

lag_at=lag(at);lagtlcf=lag(tlcf);lagsale=lag(sale);lagAR=lag(rect);
if gvkey ne lag(gvkey) or fyear ne lag(fyear)+1 then
do;lag_at=.;lagtlcf=.;lagAR=.;lagsale=.;end;

if OANCF<0 then limitedCFO=1;else limitedCFO=0;

TXFED1=sum(TXT,-1*TXDI,-1*TXS, -1*TXO);
if TXFED^=. then PERMDIFF = ((PI-(sum(TXFED,TXFO)/0.35))-(TXDI
/0.35))/lag_AT;
else PERMDIFF = ((PI-(sum(TXFED1,TXFO)/0.35))-(TXDI /0.35))/lag_AT;
tempdiff=(TXDI /0.35)/lag_AT;
lag_PERmDIFF=lag(permdiff);if gvkey ne lag(gvkey) or fyear ne lag(fyear)+1
then do;lag_PERMDIFF=.;end;

            AT_Inv=1/lag_AT;
            if INTAN^=. then INTANG=INTAN/lag_AT;
            else if INTAN=. then INTANG=0;
            else if INTAN=.c then INTANG=GDWL/lag_AT;
            If ESUB^=. then UNCON=ESUB/lag_AT;
            else UNCON=0;
            If MII^=. then MI=MII/lag_AT;
            else MI=0;
            if txs ^=. then CSTE=TXS/lag_AT;
            else CSTE=0;
            chg_NOL=(TLCF-lagTLCF)/lag_AT;
            chg_SALE=(SALE-lagSALE)/lag_AT;
            chg_RECT=(RECT-lagAR)/lag_AT;

            if TXFED^=. then BTD=(PI-(sum(TXFED,TXFO)/0.35-(TLCF-
            lagTLCF)))/lag_AT;
            else BTD=(PI-((TXT-TXDI-TXS-TXO)/0.35-(TLCF-lagTLCF)))/lag_AT;
            Size=log(AT);
            if CSHO*PRCC_F>0 then MV=CSHO*PRCC_F;
            if ceq ne . then BM=ceq/MV;
            LEV=DLTT/AT;
            LEVERAGE=sum(DLTT,DLC)/AT;
            Firm_Lev=DLTT/MV;
            Firm_Leverage=sum(DLTT,DLC)/MV;
            ROA=PI/AT;ROA1=PI/lag_AT;
            if XRD^=. then RD=XRD/lag_AT;
            else RD=0;
            RD1=XRD/lag_AT;
            if PIFO^=. then for_income=1;
            else for_income=0;
            if PIFO^=. then FI=PIFO/lag_AT;
            else FI=0;
            if TLCF>0 then NOL=1;
```

```
            else NOL=0;
            if txtubend^=. then UTB=(txtubend/lag_AT)*100;
            else if txtubend=. and fyear<2007 then utb=.;else if txtubend=.
            and fyear>=2007 then utb=0;
            utb1=(txtubend/lag_AT)*100;
            if CAPX^=. then CAPEX=CAPX/AT;
            else CAPEX=0;
            CAPEX1=CAPX/lag_AT;
            Accruals=(IBC-(OANCF-XIDOC))/lag_AT;
            PPE=PPENT/lag_AT;

drop ib ibc ppent tlcf lagtlcf xrd csho btd prcc_f startdate n_periods exchg
OANCF txfed txfo txdi PIFO AQC SPPE intan txdfed txdfo txs mii esub XIDOC
lagAR ceq capx dpc lagsale sale TXTUBEND TXTUBBEGIN ni txo rect invt
GDWL;run;

*Best Employer;
proc sort data=google.bestemp;by ticker_symbol year ranking_state;run;
PROC FREQ data=google.bestemp noprint;tables ticker_symbol*year/out=bestemp
noprint;run;


*Voting %pass;
data voting;set comp.voting;where agendageneraldesc="Advisory Vote to Ratify"
and 2006<=year(meetingdate)<=2013; run;

*CASH/GAAP ETR;
Proc sql;create table ETR as select a.gvkey,a.fyear,b.fyear as
prior_yrs,b.txpd,b.pi,b.spi,b.ROA,b.ROA1,b.txt,b.singleCASHETR from tax as a,
tax as b where a.gvkey=b.gvkey and intnx('month',a.datadate, -48,
"end")<b.datadate<=a.datadate;quit;

proc sort data=ETR; by gvkey fyear prior_yrs; run;

proc sql;create table ETR as select gvkey,fyear,sum(txpd)/(sum(pi)-sum(SPI))
as cash_ETR,sum(pi) as sum_PI,std(ROA) as std_ROA,std(ROA1) as std_ROA1,
   std(singleCASHETR) as std_cashETR from ETR group by gvkey, fyear having
count(*)>=4;quit;run;

*Discretionary Perm Diff;
*Require 15 observations per industry-year;
data perm;set tax; if PERMDIFF=. or AT_Inv=. or INTANG=. or UNCON=. or MI=.
or CSTE=. or chg_NOL=. or lag_PERMDIFF=. then delete;run;
proc sql undo_policy = none;create table perm as select * from perm group by
industry,fyear having count(*)>=15 order by industry,fyear;quit;
proc sort data=perm; by industry fyear; run;
Proc reg data = perm  noprint ;
     by industry fyear;
     model PERMDIFF = AT_Inv INTANG UNCON MI CSTE chg_NOL lag_PERMDIFF;
     output out=dis_PERM r=DTAX;run;

*Discretionary Accruals;
* Require 15 observations per industry-year;
data acc;set tax;if Accruals=. or AT_Inv=. or chg_SALE=. or chg_RECT=. or
PPE=. then delete;run;
proc sql undo_policy = none;create table acc as select * from acc group by
industry, fyear having count(*)>=15 order by industry,fyear;quit;
```

```sas
* Calculate regression coefficients by industry and year;
proc sort data=acc; by industry fyear; run;
Proc reg data = acc  noprint ;
     by industry fyear;
     model Accruals = AT_Inv chg_SALE chg_RECT PPE ROA;
     output out=dis_ACC r=DACC;run;

*merge compustat vars;
proc sql;
create table tax as select a.*,b.tsscore as ts_lisowsky from tax as
a,tax.lisowsky as b where a.gvkey=put(b.gvkey, Z6.) and a.fyear=b.fyear;
create table tax as select a.*,b.count from tax as a left join bestemp as b
on a.tic=b.ticker_symbol and a.calyear=b.year;
create table tax as select a.*,b.frozen from tax as a left join
comp.pensionfreeze as b on a.gvkey=put(b.gvkey, Z6.) and a.fyear=b.year;
create table tax as select a.*,b.votedfor,b.voteresult from tax as a left
join voting as b on a.tic=b.ticker and a.fyear=year(b.meetingdate);
create table tax as select a.*,b.DACC from tax as a left join dis_Acc as b on
a.gvkey=b.gvkey and a.fyear=b.fyear;
create table tax as select
a.*,b.cash_ETR,b.std_cashETR,b.sum_PI,b.std_ROA,b.std_ROA1 from tax as a left
join ETR as b on a.gvkey=b.gvkey
and a.fyear=b.fyear;
create table tax as select a.*,b.Dtax from tax as a left join dis_perm as b
on a.gvkey=b.gvkey and a.fyear=b.fyear;
create table tax as select a.*,b.pbpro as plan_size,b.pplao as plan_asset
from tax as a left join tax.pension as b on a.gvkey=b.gvkey and
a.datadate=b.datadate;
quit;run;

proc sort noduprec;by _all_;run;
data taxclean;set tax;where 2006<=fyear<=2013;where firm_leverage>=0 and
bm>0;
if count ne . then bestemp=1;else if count=. then bestemp=0; if
voteresult="Pass" then pass=1;else pass=0;
array a{3} frozen plan_size votedfor;
array b{3} missfrozen nopension missvote;
do i = 1 to 3;if a(i)=. then b(i)=1;else b(i)=0; end;
do i = 1 to 3;if a(i)=. then a(i)=0;end;
drop spi txpd txt singlecashETR tic tempdiff permdiff lag_permdiff count
txfed1 intang uncon mi cste chg_sale chg_rect i fyr chg_nol
accruals voteresult;run;

*CEO compensation;
data CEO;set comp.CEOCFOcomp;where 2006<=year<=2013 and debt_equity3>=0 and
insidedebt>=0 and CEOd=1 and vega_equity>0 and delta_equity>0;
keep gvkey year execid equity_den3 insidedebt debt_equity3 total_curr
vega_equityold delta_equityold delta_equity vega_equity ;run;

*Merge tax variables with CEO compensation;
proc sql;create table taxclean as select distinct a.*,b.execid,b.total_curr
as BS,b.insidedebt,(b.insidedebt/a.lag_at) as debt,
(b.equity_den3/a.lag_at) as equity,b.debt_equity3 as CEO_DE,b.vega_equity as
vega,b.delta_equity as delta from taxclean as a,CEO as b where
a.gvkey=put(b.gvkey, Z6.) and a.fyear=b.year;
```

```sas
create table taxclean as select a.*,b.age from taxclean as a left join
comp.exeage as b on a.gvkey=b.gvkey and put(a.execid,z5.)=b.execid and
a.fyear=b.year;
create table taxclean as select a.*,n(gvkey) as n_firms from taxclean as a
group by gvkey;quit;
proc sort noduprec; by _all_;run;
proc sort data=taxclean;by gvkey fyear;
data cleanf;set taxclean;by gvkey fyear;
        If (CEO_DE^=. and Firm_leverage>0) then
        Relative_DE=CEO_DE/Firm_leverage;
        if (relative_DE>1 or (debt>0 and firm_leverage=0))then highdebt=1;else
        if debt=0 then highdebt=0;else highdebt=0;
        lnCEO_DE=log(1+CEO_DE);lnrelative_DE=log(1+relative_DE);
        array f(3) BS vega delta;
        array g(3) lnBS lnvega lndelta;do i=1 to 3;g(i)=log(f(i));end;
        drop lag_at dltt dlc sich cik_num i pi sum_pi at_INV at mv n_firms PPE
firm_lev;run;

data all;set cleanf;if  ts_lisowsky=. or BS=. or debt=. or CEO_DE=.  or
lpermno=90655 or age=. or std_ROA=. then delete;run;

*SAS Macro for winsorizing;
%macro winsor(dsetin=, dsetout=, byvar=none, vars=, type=winsor, pctl=1 99);

%if &dsetout = %then %let dsetout = &dsetin;

%let varL=;
%let varH=;
%let xn=1;

%do %until ( %scan(&vars,&xn)= );
    %let token = %scan(&vars,&xn);
    %let varL = &varL &token.L;
    %let varH = &varH &token.H;
    %let xn=%EVAL(&xn + 1);
%end;

%let xn=%eval(&xn-1);

data xtemp;
    set &dsetin;
    run;

%if &byvar = none %then %do;

    data xtemp;
        set xtemp;
        xbyvar = 1;
        run;

    %let byvar = fyear;

%end;

proc sort data = xtemp;
    by &byvar;
    run;
```

```sas
proc univariate data = xtemp noprint;
    by &byvar;
    var &vars;
    output out = xtemp_pctl PCTLPTS = &pctl PCTLPRE = &vars PCTLNAME = L H;
    run;

data &dsetout;
    merge xtemp xtemp_pctl;
    by &byvar;
    array trimvars{&xn} &vars;
    array trimvarl{&xn} &varL;
    array trimvarh{&xn} &varH;

    do xi = 1 to dim(trimvars);

        %if &type = winsor %then %do;
            if not missing(trimvars{xi}) then do;
                if (trimvars{xi} < trimvarl{xi}) then trimvars{xi} =
trimvarl{xi};
                if (trimvars{xi} > trimvarh{xi}) then trimvars{xi} =
trimvarh{xi};
            end;
        %end;

        %else %do;
            if not missing(trimvars{xi}) then do;
                if (trimvars{xi} < trimvarl{xi}) then delete;
                if (trimvars{xi} > trimvarh{xi}) then delete;
            end;
        %end;

    end;
    drop &varL &varH xi;
    run;

%mend winsor;
%winsor(dsetin=all, dsetout=all, byvar=none, vars=std_ret lnBS debt lnCEO_DE
capex CAPEX1 lnrelative_DE std_ROA std_ROA1 bm lnvega
lndelta,type=winsor,pctl=1 99);
run;quit;
```