

```

*****;
rsubmit;
*download Compustat variables;
data Comp00;
set comp.funda;
where indfmt = 'INDL' and datafmt = 'STD' and consol = 'C' and popsrc = 'D';
keep gvkey cik comm datadate fyear at sale pi pifo pidom spi tii;
run;
*download and clean segment variables;
PROC SQL;
CREATE TABLE geo
AS SELECT gvkey, datadate, geotp, sales
FROM comp.WRDS_SEGMERGED
WHERE stype='GEOSEG'
Group BY gvkey, datadate having srcdate = min(srcdate)
order by gvkey, datadate;
QUIT;
proc sql;
create table geol
as select distinct gvkey, datadate, geotp, sum(sales) as total_sale
from geo
group by gvkey, datadate, geotp;
quit;
proc transpose data=geol out=geo2
(rename=(_1=obsolete _2=dsale _3=fsale));
by gvkey datadate;
id geotp;
var total_sale;
run;
*download and merge company identifiers;
data company;
set comp.company;
keep gvkey fic sic;
run;
proc sql;
create table Comp01
as select a.*, b.dsale, b.fsale, c.fic, c.sic
from Comp00 as a left join geo2 as b
on a.gvkey = b.gvkey and a.datadate=b.datadate
left join Company c
on a.gvkey = c.gvkey
order by a.gvkey, a.fyear;
quit;

*clean data;
data comp01;
set comp01;
if pifo=. then pifo=pi-pidom;
if pidom=. then pidom=pi-pifo;
if fsale=. then fsale=sale-dsale;
if dsale=. then dsale=sale-fsale;
sic2=input(substr(sic,1,2),12.);
ciknumber=cik+1-1;
format ciknumber Z10.;
informat ciknumber BEST32.;
run;

```

```

*Create lags;
PROC SQL;
CREATE TABLE comp02 AS
SELECT a.*, b.at as lag_at, b.pifo as lag_pifo, b.pidom as lag_pidom, b.dsale
as lag_dsale, b.fsale as lag_fsale
FROM Comp01 as a left join Comp01 as b
on a.gvkey = b.gvkey and a.fyear = b.fyear+1;
QUIT;
PROC download data=Comp02 out=fx.DM1;
run;
endrsubmit;

*Sample selection;
data DM2;
set fx.DM1;
if 1997<=fyear<=2017;
if 0.99*sale<=dsale+fsale<=1.01*sale;
if 0.99*abs(pi)<=abs(pidom+pifo)<=1.01*abs(pi);
if 6000<=sic<=6999 then delete; *Financial;
if 4900<=sic<=4999 then delete; *Utilities;
if find(conm,"TRUST") then delete;
if find(conm,"LP") then delete;
if find(conm,"PARTNERS") then delete;
if fic='USA';
if spi=. then spi=0;
if tii=. then tii=0;
if abs(spi)>0.1*sale then delete;
if abs(tii)>0.1*sale then delete;
if at<1 then delete;
if dsale<1 then delete;
if fsale<1 then delete;
run;

*First differencing key variables;
data DM3(keep=gvkey ciknumber datadate fyear sic2 dPIFO dPIDOM dSALEFO
dSALEDOM);
set DM2;
dPIFO=(pifo-lag_pifo)/lag_at;
dPIDOM=(pidom-lag_pidom)/lag_at;
dSALEFO=(fsale-lag_fsale)/lag_at;
dSALEDOM=(dsale-lag_dsale)/lag_at;
if not missing(dPIFO+dPIDOM+dSALEFO+dSALEDOM);
run;

*Merge with tax incentive data;
proc sql;
create table DM4
as select a.*, b.avgftr
from DM3 as a left join fx.avgftr as b
on a.gvkey = b.gvkey and a.fyear=b.fyear;
quit;

*Add controls;
rsubmit;
data control;
set comp.funda;
where indfmt = 'INDL' and datafmt = 'STD' and consol = 'C' and popsrc = 'D';

```

```

ROS=pi/sale;
if xrd=. then xrd=0;
if xad=. then xad=0;
if intan=. then intan=0;
if ch=. then ch=0;
if dt=. then dt=0;
rd_at=xrd/at; if rd_at=. then rd_at=0;
ads_at=xad/at; if ads_at=. then ads_at=0;
intan_at=intan/at; if intan_at=. then intan_at=0;
cash_at=ch/at; if cash_at=. then cash_at=0;
debt_at=dt/at; if debt_at=. then debt_at=0;
size=log(1+at);
keep gvkey datadate fyear ros rd_at ads_at intan_at cash_at debt_at size;
run;
PROC download data=control out=fx.DM_control;
run;
endrsubmit;

*Merge above datasets;
proc sql;
create table DM5
as select a.*, b.*
from DM4 as a left join fx.DM_control as b
on a.gvkey = b.gvkey and a.fyear=b.fyear;
quit;

*Run Main Program to obtain z7;

*Merge with FX risk data;
proc sql;
create table DM6 as
select distinct a.*, b.datadate as matched_date 'matched_date', b.score*-1 as
score
from DM5 as a, z7 as b
where a.ciknumber=b.ciknumber and -12<=intck('month',
intnx('year',b.datadate,+1), a.datadate)<=12
order by a.gvkey, a.fyear;
quit;

*Clean data;
data DM7;
set DM6;
mo=intck('month',matched_date,datadate);
if not missing(avgftr);
run;
proc sort data=DM7; by gvkey fyear mo; run;
proc sort data=DM7 nodupkey; by gvkey fyear dPIFO; run;

*Winsorize continuous variables here and then export;

```