**data** ratio;

set Tmp1.Funda\_large (keep=FYEAR DATADATE GVKEY NI TEQ PI EBIT SALE AT PRCC\_F EPSPX COGS INVT SICH MKVALT BKVLPS DLTT DLC ACT LCT);

if fyear in (**2000**,**2008**,**2017**);

ROE=ni/teq;

taxburden=ni/pi;

interestburden=pi/ebit;

operatingmargin=ebit/sale;

TAT=sale/at;

EM=at/teq;

PM=ni/sale; ROA=ni/at; PE=prcc\_f/epspx; IT=cogs/invt;

year=year(datadate); Sic2=int(sich/**100**); MarketToBookValue=MKVALT/BKVLPS;

Leverage=(DLTT + DLC)/teq; CurrentRatio=act/lct; ATO=sale/at;

**run**;

**proc** **means** data=ratio nway n mean median min p1 p5 p95 p99 max;

var AT ROA ROE TAT PE PM IT EM Sic2 ATO CurrentRatio Leverage MarketToBookValue taxburden interestburden operatingmargin ;

output out= ratio1;

**run**;

%include "E:\Downloads\WINSOR.SAS";

%***winsor*** (dsetin=ratio, byvar=none, dsetout=final, vars=tat AT ROA ROE PE PM IT EM Sic2 ATO CurrentRatio Leverage MarketToBookValue taxburden interestburden operatingmargin, pctl=**1** **99**); **run**;

**proc** **means** data=final nway n mean median min p1 p5 p95 p99 max;

var AT ROA ROE TAT PE PM IT EM Sic2 ATO CurrentRatio Leverage MarketToBookValue taxburden interestburden operatingmargin ;

output out= ratio2;

**run**;

**proc** **sort** data=final noduplicate out=finalsorted;

by at;

**run**;

**proc** **rank** data=finalsorted groups=**3** ties=low

out=finalsortedranked ;

var at;

ranks at\_rank;

**run**;

**proc** **means** data=finalsortedranked nway n mean median;

class at\_rank;

VAR ROE ROA PM tat;

output out=C.final\_YL mean= ROE ROA PM tat median= ROE\_ ROA\_ PM\_ tat\_;

**run**;

**proc** **means** data=final nway n mean median;

class year;

VAR ROA PM TAT;

output out=final2 mean= ROA PM TAT median= ROA\_ PM\_ TAT\_;

**run**;

**PROC** **SGPLOT** DATA = final2 ;

SERIES X = year Y = roa / LEGENDLABEL ="ROA";

SERIES X = year Y = pm/ LEGENDLABEL = "PM";

SERIES X = year Y = tat / LEGENDLABEL ="TAT";

TITLE "Du Pont Analysis 3 factors";

**RUN**;

**data** sic2;

set final;

if sic2=**42**;

**run**;

**proc** **means** data=sic2 nway n mean median;

class YEAR;

VAR ROA PM TAT;

output out=sic242 mean= ROA PM TAT median= ROA\_ PM\_ TAT\_;

**run**;

**PROC** **SGPLOT** DATA = sic242 ;

SERIES X = year Y = roa / LEGENDLABEL ="ROA";

SERIES X = year Y = pm/ LEGENDLABEL = "PM";

SERIES X = year Y = tat / LEGENDLABEL ="TAT";

TITLE "Du Pont Analysis 3 factors-sic2";

**RUN**;

**data** add;

set final;

logsize=log(at);

**run**;

**proc** **surveyreg** data=add;

cluster SIC2;

model ROE= logsize MarketToBookValue leverage CurrentRatio ATO ROA ;

**quit**;