Programming used for "Anonymous Equity Research"

Travis Dyer and Eunjee Kim

Contents

SAS Code

Macros

```
/*******************************************************
This macro displays all duplicate observations in a dataset based on specified unique keys.
Output datasets are named 'duplicates' and 'nonduplicates.'

dset = dataset
keys = unique ids or combination of ids

E.g. %dupl(dset=data,keys=gvkey datadate);

*******************************************************/
%macro dupl(dset=,keys=);

data _null_;
	keys="&keys.";
	call symput('sql',tranwrd(keys,' ',', '));
	call symput('ord',tranwrd(keys,' ',', b.'));
run;

%let num=%eval(%sysfunc(countc(&keys,' '))+1);
%let crit= ;
%let str= ;

%do a=1 %to &num;
	%let key&a.=%scan(&keys,%eval(&a.),' ');
	%if &a>1 %then %let str=%str( & );
	%let crit= &crit.&str.a.&&key&a..%str(=)b.&&key&a..;
%end;

data dl;
	set &dset (keep=&keys);
	occur=1;
run;

proc sql;
create table dlx as select distinct
&sql, sum(occur) as repeat
	from dl
	group by &sql
	having repeat=1;

create table nonduplicates as select
b.*
from dlx as a inner join &dset as b
```

```
            on &crit.
            order by b.&ord.;
quit;

proc sql;
create table dlx as
select distinct &sql, sum(occur) as repeat
            from dl
            group by &sql
            having repeat>1;

create table duplicates as
select b.* from dlx as a inner join &dset as b
            on &crit.
            order by b.&ord.;
quit;


option nonotes;
proc datasets lib=work nolist;
            delete dl dlx;
run;
quit;
option notes;

%mend dupl;


/*************************************************************************
This macro calculates the exact date of X trading days pre or post a given date.

dset=                   Dataset containing the event date of interest.
t =                     Event date of interest.
prior =                 Indicator of whether the trading days are before or after the event date.
trading_days =          Number of trading days apart from the given event date.
outdset =               Output dataset contains a new variable with the exact date of X trading
                        days pre or post a given date. The naming convention for the new date
                        provided is date_(trading_days)_[p|f]

e.g.: %trading_days(dset=eventdata,t=eventdate,Prior=Y ,trading_days=5,outdset=eventdata2);

new variable name will be: date_5_p

NOTE: This macro considers weekends and listed NYSE and NASDAQ holidays
*************************************************************************/
```

```
%macro trading_days(dset=, t=, prior=,trading_days=, outdset=);
data &outdset; set &dset;
%let i=1;
%do %until (&i > &trading_days);
	%if %upcase(&prior.)=Y %then %do;
	date__&i = intnx('weekday',&t,-&i);
	%end;
	%if %upcase(&prior.)^=Y %then %do;
	date__&i = intnx('weekday',&t,&i);
	%end;
	%let i=%eval(&i+1);
%end;

array vars date__:;
	do over vars ;
	if vars ne HOLIDAY("NEWYEAR",year(vars)) and vars ne
HOLIDAY("MLK",year(vars)) and vars ne HOLIDAY("USPRESIDENTS",year(vars)) and vars
ne HOLIDAY("USPRESIDENTS",year(vars)) and vars ne HOLIDAY("EASTER",year(vars))-2
and  /*good friday*/ vars ne HOLIDAY("MEMORIAL",year(vars)) and vars ne
HOLIDAY("USINDEPENDENCE",year(vars)) and vars ne HOLIDAY("LABOR",year(vars))
and vars ne HOLIDAY("THANKSGIVING",year(vars)) and vars ne
HOLIDAY("CHRISTMAS",year(vars))  then vars = 0; else vars = 1; end;

%if %upcase(&prior.)=Y %then %do;
date__ = -&trading_days;
%end;
%if %upcase(&prior.)^=Y %then %do;
date__ = &trading_days;
%end;

%let i=1;
%do %until (&i > &trading_days);
	%if %upcase(&prior.)=Y %then %do;
	date__ = date__ - date__&i;
	%end;
	%if %upcase(&prior.)^=Y %then %do;
	date__ = date__ + date__&i;
	%end;
	%let i=%eval(&i+1);
%end;

%if %upcase(&prior.)=Y %then %do;
date_&trading_days._p = intnx('weekday',&t,date__);
%end;
%if %upcase(&prior.)^=Y %then %do;
date_&trading_days._f = intnx('weekday',&t,date__);
```

```
%end;

drop date__:;
format date_&trading_days: YYMMDDN8.;
run;
%mend;

/*****************************************************************************/
1) data: input dataset containing variables that will be win/trunc.
2) out: output dataset (leave blank to overwrite data)
3) BYVAR: variable(s) used to form groups (type 'none' for total sample
4) VARS: variable(s) that will be winsorized/truncated
5) TYPE: = W to winsorize and = T (or anything else) to truncate
6) PCTL = percentile points (in ascending order) to truncate/winsorize values.  Default is 1st and
99th percentiles.

DESCRIPTION: This macro is capable of both truncating and winsorizing one or multiple
variables.  Truncated values are replaced with a missing observation rather than deleting the
observation.  This gives the user more control over the resulting dataset.
                                                                                          */
E.g., %WT(data = mydata, out = mydata2, byvar = year, vars = assets earnings, type = W, pctl =
0 98)

==> Winsorizes by year at 98% and puts resulting dataset into mydata2 */
/*****************************************************************************/

%macro WT(data=_last_, out=, byvar=none, vars=, type = W, pctl = 1 99, drop= N);

	%if &out = %then %let out = &data;

	%let varLow=;
	%let varHigh=;
	%let xn=1;

	%do %until (%scan(&vars,&xn)= );
	%let token = %scan(&vars,&xn);
	%let varLow = &varLow &token.Low;
	%let varHigh = &varHigh &token.High;
	%let xn = %EVAL(&xn + 1);
	%end;

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

	data xtemp;
		set &data;
```

```
%let dropvar = ;
%if &byvar = none %then %do;
        data xtemp;
set xtemp;
xbyvar = 1;

%let byvar = xbyvar;
%let dropvar = xbyvar;
%end;

proc sort data = xtemp;
        by &byvar;

/*compute percentage cutoff values*/
proc univariate data = xtemp noprint;
by &byvar;
var &vars;
output out = xtemp_pctl PCTLPTS = &pctl PCTLPRE = &vars PCTLNAME = Low
High;

data &out;
merge xtemp xtemp_pctl; /*merge percentage cutoff values into main dataset*/
by &byvar;
array trimvars{&xn} &vars;
array trimvarl{&xn} &varLow;
array trimvarh{&xn} &varHigh;

do xi = 1 to dim(trimvars);
/*winsorize variables*/
%if &type = W %then %do;
        if trimvars{xi} ne . then do;
                if (trimvars{xi} < trimvarl{xi}) then trimvars{xi} = trimvarl{xi};
                if (trimvars{xi} > trimvarh{xi}) then trimvars{xi} = trimvarh{xi};
        end;
%end;
/*truncate variables*/
%else %do;
        if trimvars{xi} ne . then do;
                if (trimvars{xi} < trimvarl{xi}) then trimvars{xi} = .T;
                if (trimvars{xi} > trimvarh{xi}) then trimvars{xi} = .T;
        end;
%end;

                %if &drop = Y %then %do;
                  if trimvars{xi} = .T then delete;
                %end;
```

```
		end;
	drop &varLow &varHigh &dropvar xi;

	/*delete temporary datasets created during macro execution*/
	proc datasets library=work nolist;
		delete xtemp xtemp_pctl; quit; run;

%mend WT;
```

Firm-quarter-level observations

```
*get quarterly data for firms;
rsubmit;
data compq; set comp.fundq
(keep = gvkey datadate rdq indfmt datafmt popsrc consol curcdq fyr fqtr fyearq tic ATq NIQ
IBQ SALEQ CEQQ XRDQ CSHOQ PRCCQ INTANQ dlttq dlcq);
where (indfmt='INDL') and (datafmt='STD') and (popsrc='D') and (consol='C') and
(curcdq='USD') and (atq >= 0) and (year(datadate)> 1990);
drop indfmt datafmt popsrc consol curcdq;
MVE = CSHOQ*PRCCQ;
MTB = MVE/CEQQ;
ROA = NIQ/ATQ;
Leverage = (dlttq+dlcq)/atq;
if missing(MVE) then delete;
if missing(MTB) then delete;
if missing(ROA) then delete;
if missing(leverage) then delete;
if missing(tic) then delete;
run;

*remove duplicates and constrain sample to 2006+;
%dupl(dset=compq,keys=gvkey datadate);
data nonduplicates1; set nonduplicates; if year(rdq) < 2006 then delete; run;
%dupl(dset=nonduplicates1,keys=gvkey rdq);

****** Merge on Permno	***;
proc sql; create table data2 as select distinct
	a.*, b.lpermno as permno
	from nonduplicates as a left join crsp.ccmxpf_linktable
	(where=(linktype in ('LU', 'LC', 'LD', 'LF', 'LN', 'LO', 'LS', 'LX'))) as b
	on (a.gvkey = b.gvkey)
	and (b.linkdt <= a.datadate or b.linkdt = .B)
	and (a.datadate <= b.linkenddt or b.linkenddt = .E);
quit;
```

```
*getting rid of duplicates;
%dupl(dset=data2,keys=gvkey datadate);
data tempx1; set duplicates; permno = .; run;
data tempx2; set nonduplicates tempx1; run;
proc sort data = tempx2 nodupkey; by gvkey datadate; run;
%dupl(dset=tempx2,keys=permno datadate);
data tempx3; set duplicates; permno = .; run;
data data3; set nonduplicates tempx3; run;

*remove firms without permno;
data data3; set data3; if missing(permno) then delete; run;

*get prior period bid-ask spread, turnover, and returns;
data dsf1; set crsp.dsf
(keep = permno date ret prc vol ASKHI BIDLO Ask Bid shrout);
where year(date) >= 1994;
prc = abs(prc);
if ask < bid then ask = .;
if ask-bid > 10 then ask = .;
run;

proc sql; create table illiqtemp as select
	a.permno, a.rdq, abs(b.ret)/(abs(b.prc) * abs(b.vol)) as illiq_ratio,
	(b. ASK - b. BID)/((b. ASK + b. BID)/2) as spread2_ratio,
	b. ret, b. vol/(b. shrout*1000) as dailyturn, b. date
	from data3 as a left join dsf1 as b
	on a.permno eq b.permno
	and INTNX( 'day', a.rdq, -90, 'S' ) <= b.date < a.rdq;
quit;

proc sql; create table illiqtemp1 as select
	a.*, b.vwretd, (a.ret - b.vwretd) as ret_mrktadj
	from illiqtemp as a left join crsp.dsi as b
	on a.date eq b.date;
quit;

proc sql; create table illiqtemp2 as select distinct
	permno, rdq, mean(illiq_ratio) as ILLIQ_mean,
	mean(spread2_ratio) as spread2_mean,
	mean(dailyturn) as TURN_mean, std(ret) as ret_vol,
	count(illiq_ratio) as count1, count(spread2_ratio) as count2a,
	count(ret) as count3, count(dailyturn) as count4,
	(exp(sum(log(1+ret_mrktadj))) - 1) as CAR
	from illiqtemp1 group by permno, rdq;
quit;
```

```
*remove values with less than 55 observations;
data illiqtemp3; set illiqtemp2;
if count1 < 55 then ILLIQ_mean = .;
if count2a < 55 then spread2_mean = .;
if count3 < 55 then ret_vol = .;
if count3 < 55 then CAR = .;
if count4 < 55 then TURN_mean = .;
run;

proc sql; create table compq2 as select distinct
	a. *, b. ILLIQ_mean, b. spread2_mean, b. ret_vol, b. turn_mean, b.CAR
	from data3 as a left join illiqtemp3 as b
	on a. permno = b. permno
	and a. rdq = b. rdq;
quit;

proc download data= compq2 out = dk.compq_20190426; run;
proc download data=nonduplicates out = dk.nonduplicates; run;
endrsubmit;


***get anonymous and non-anonymous count for period;
/*dk.matched_article3 <- Seeking Alpha article data w/ author categorization*/
data matched_article3; set dk.matched_article3;

data anonymous non_anonymous; set matched_article3;
if anonymous = 1 then output anonymous;
if anonymous = 0 then output non_anonymous;
run;

proc sql; create table temp1 as select
	a.gvkey, a.datadate, b.date
	from dk.compq_20190426 a left join anonymous b
	on a.gvkey eq b.gvkey
	and a.datadate <= b.date <= INTNX( 'day', a.datadate, +90, 'S' );
quit;

proc sql; create table temp2 as select
	gvkey, datadate, count(date) as anon_count
	from temp1 group by gvkey, datadate
	order by gvkey, datadate;
quit;

proc sql; create table tempor1 as select
	a.gvkey, a.datadate, b.date
	from dk.nonduplicates a left join non_anonymous b
	on a.gvkey eq b.gvkey
```

```
        and a.datadate <= b.date <= INTNX( 'day', a.datadate, +90, 'S' );
quit;

proc sql; create table tempor2 as select
        gvkey, datadate, count(date) as non_anon_count
        from tempor1 group by gvkey, datadate
        order by gvkey, datadate;
quit;

proc sql; create table quarter as select
        a.*, b.anon_count
        from dk.compq_20190426 a left join temp2 b
        on a.gvkey eq b.gvkey
        and a.datadate=b.datadate;
quit;

proc sql; create table quarter1 as select
        a.*, b.non_anon_count
        from quarter a left join tempor2 b
        on a.gvkey eq b.gvkey
        and a.datadate=b.datadate;
quit;

proc means data=quarter1 n nmiss; run;

data quarter1; set quarter1; id=_n_; PS= MVE/saleq; run;

*****merge IBES data;
data ibes; set dk.ibes_sum; run; /*IBES>Summary History>Summary statistics*/
data ibes; set ibes; if FPI=6; run; /*quarterly estimates*/

proc sort data=ibes; by TICKER ANNDATS_ACT descending STATPERS; run;

data ibes; set ibes;
if STATPERS> ANNDATS_ACT then delete;
if ANNDATS_ACT=. then delete;
run;

proc sort data=ibes nodupkey; by TICKER ANNDATS_ACT; run;

data ibes; set ibes; id=_n_; run;

/*add permno to I/B/E/S dataset using iclink file*/
proc sql; create table ibes as select
        a.*, b.ticker, b.permno, c.permno
```

```
        from ibes as a, dk.iclink as b, quarter1 as c /*iclink: https://wrds-
        www.wharton.upenn.edu/pages/support/research-wrds/macros/wrds-macro-iclink/ */
        where a.ticker=b.ticker and b.permno=c.permno;
quit;

proc sort data=ibes nodupkey; by id; run;

/*get a lagged stock price from quarter1 file*/
data ibes; set ibes; l_datadate=INTNX( 'month', FPEDATS, -2)-1; run;

proc sql; create table ibes as select
        a.*, b.datadate, b.prccq, b.permno
        from ibes as a left join quarter1 as b
        on a.permno=b.permno
        and a.l_datadate=b.datadate;
quit;

proc sort data=ibes nodupkey; by id; run;
data ibes; set ibes; if prccq=. then delete; run;
data ibes; set ibes; earnsurp=(MEDEST-ACTUAL)/prccq; run;

proc sql; create table quarter1 as select
        a.*, b.permno, b.earnsurp, b.STDEV, b.FPEDATS
        from quarter1 as a left join ibes as b
        on a.permno=b.permno
        and a.datadate=b.FPEDATS;
quit;

proc sort data=quarter1 nodupkey; by id; run;
proc means data=quarter1 n nmiss; run;

/**Analyst following**/
data ibes_detail; set dk.ibesdetail_09292019; run;
data ibes_detail; set ibes_detail; id=_n_; run;

/*add permno to I/B/E/S dataset using iclink file*/
proc sql; create table ibes_detail as select
        a.*, b.ticker, b.permno
        from ibes_detail as a, dk.iclink as b
        where a.ticker=b.ticker;
quit;

proc sort data=ibes_detail nodupkey; by id; run;

proc sql; create table quarter1 as select
        a.*, b.ANNDATS, b.permno, b.ANALYS
```

```
        from quarter1 as a left join ibes_detail as b
        on a.permno=b.permno
        and a.datadate-365<b.ANNDATS<a.datadate;
quit;

proc sort data=quarter1 nodupkey; by id ANALYS; run;

proc sql; create table quarter1 as select
        *, count(*) as ana_follow
        from quarter1 group by id;
quit;

proc sort data=quarter1 nodupkey; by id; run;

data quarter1; set quarter1;
if ANALYS=. then ana_follow=0;
drop FPEDATS ANNDATS ANALYS;
run;

*****merge on short interest during the quarter (comp);
data shortint; set dk.comp_shortint; run; /*<- Compustat North America Fundamentals quarterly
(/wrdslin/comp/sasdata/d_na/fundq)*/
data shortint; set shortint;
shortdate=INTNX( 'month', DATADATE, 1)-1; format shortdate yymmddn.;
run;

proc sql; create table quarter2 as select
        a.*, b.gvkey, b.SHORTINT, b.shortdate
        from quarter1 as a left join shortint as b
        on a.gvkey=b.gvkey
        and a.datadate=b.shortdate;
quit;

proc sort data=quarter2 nodupkey; by id; run;
data quarter2; set quarter2; short=SHORTINT/(cshoq*1000000); run;

*****merge on # of employees (comp.funda);
data employee; set dk.compa_employee; run; /*Compustat North America Fundamentals
Annual*/

proc sql; create table quarter3 as select
        a.*, b.gvkey, b.datadate as emp_datadate, b.EMP, b.sic, b.cik, b.CUSIP
        from quarter2 as a left join employee as b
        on a.gvkey=b.gvkey;
quit;
```

```
data quarter3; set quarter3;
if datadate<emp_datadate then delete;
run;

proc sort data=quarter3; by id descending emp_datadate; run;

proc sort data=quarter3 nodupkey; by id; run;

data quarter3; set quarter3; drop emp_datadate shortdate; run;

*****merge on institutional ownership (Bushee's measure);
proc import out= work.bushee datafile= "F:\Main_External\1-Databases\Institutional
Ownership\Bushee Data\Bushee Data 20181005.csv" dbms=csv replace; getnames=YES;
datarow=2; run;

*duplicates come from Brian Bushee's database -- these all have identical perm_TQD_class
variables -- so I am deleting duplicates here;
proc sort data = bushee nodupkey; by mgrno year; run;

rsubmit;
proc upload data = bushee; run;

* Note shares outstanding in this database is messed up - pull from CRSP;
data ii; set tfn.s34 (keep = cusip mgrno fdate rdate shares); where year(rdate) ge 1993;
year=year(rdate); run;

%dupl(dset=ii,keys=mgrno cusip rdate);
* select the latest filing date to remove duplicates;
proc sort data = duplicates; by  mgrno cusip rdate decending fdate; run;

data fix; set duplicates; by mgrno cusip rdate; if first.rdate then output fix; run;
data ii2; set nonduplicates fix; run;

proc sql; create table ii3 as select
        a.*, b. Perm_TQD_Class
        from ii2 as a left join bushee as b
        on a.mgrno eq b.mgrno
        and a. year eq b. year;
quit;

proc sql; create table iisum as select
        cusip, rdate, sum(shares) as INSTshares, count(shares) as instcount
        from ii3 group by cusip, rdate
        order by cusip, rdate;
quit;
```

```
proc sql; create table iisum2 as select
        cusip, rdate, Perm_TQD_Class, sum(shares) as instshares, count(shares) as instcount
        from ii3 group by cusip, rdate, Perm_TQD_Class
        order by cusip, rdate, Perm_TQD_Class;
quit;

proc sql; create table iisum3 as select
        a.*, b. instshares as dedshares, b. instcount as dedcount
        from iisum as a left join iisum2 as b
        on a.cusip eq b.cusip
        and a. rdate eq b. rdate
        and b. Perm_TQD_Class = 'DED';
quit;

proc sql; create table iisum4 as select
        a.*, b. instshares as qixshares, b. instcount as qixcount
        from iisum3 as a left join iisum2 as b
        on a.cusip eq b.cusip
        and a. rdate eq b. rdate
        and b. Perm_TQD_Class = 'QIX';
quit;

proc sql; create table iisum5a as select
        a.*, b. instshares as trashares, b. instcount as tracount
        from iisum4 as a left join iisum2 as b
        on a.cusip eq b.cusip
        and a. rdate eq b. rdate
        and b. Perm_TQD_Class = 'TRA';
quit;

data iisum5; set iisum5a; if missing(cusip) then delete; run;

proc sql; create table iisum6 as select
        a.*, b.permno
        from iisum5 as a left join crsp.stocknames (where=(ncusip ne ''))  as b
        on substr(a.cusip,1,8) eq substr(b.ncusip,1,8)
        and b. namedt <= a. rdate <= b. nameenddt;
quit;

data iisum6; set iisum6; if missing(permno) then delete; run;

* Get shares outstanding from CRSP | calculate ratios;
proc sql; create table iisum7 as select
        a.*, (b.shrout*1000) as shroutraw, a.instshares/(b.shrout*1000) as instown,
        a.dedshares/(b.shrout*1000) as dedown, a.qixshares/(b.shrout*1000) as qixown,
        a.trashares/(b.shrout*1000) as traown
```

```
        from iisum6 as a left join crsp.msf as b
        on a.permno eq b.permno
        and month(a.rdate) eq month(b.date)
        and year(a.rdate) eq year(b.date)
        order by permno, rdate;
quit;

proc download data = dk.Iisum7_20181005; run;
endrsubmit;

data inst; set dk.Iisum7_20181005; ncusip=substr(cusip, 1, 6); run;
data quarter3; set quarter3; cusip6=substr(cusip, 1, 6); run;

proc sql; create table quarter4 as select
        a.*, b.ncusip, b.instown, b.dedown, b.qixown, b.traown, b.rdate
        from quarter3 as a left join inst as b
        on a.cusip6=b.ncusip
        and a.datadate=b.rdate;
quit;

proc sort data=quarter4; by id; run;
data quarter4; set quarter4; drop cusip ncusip rdate; run;

proc means data=quarter4 n nmiss; run;

*****merge on # of Insider trading (Thomson Reuters);
data work; set dk.table1;run; /* Thomson Reuters > Insider Data > Table 1*//*Sample horizon:
2006-2018*/

data work; set work; if TRANCODE="P" or TRANCODE="S"; run; /*1. open market
transactions*/

data work; set work;
drop ADDRESS2 CITY ZIPCODE PHONE AMEND ADDRESS1 STATE COUNTRY
CUSIP2 ROLECODE2 ROLECODE3 ROLECODE4;
run;

data work; set work;
if cusip6="" then delete; if CLEANSE="R" or CLEANSE="H" or CLEANSE="R";
run; /*2. clean data: Heron and Lie 2007 JFE*/

proc sort data=work; by CUSIP6 TRANDATE PERSONID; run;
/*3. leave only directors and officers (Ali Wei Zhou 2011 JAR)*/

data work1; set work;
```

```
if ROLECODE1="CB" or ROLECODE1="D" or ROLECODE1="DO" or ROLECODE1="H"
or ROLECODE1="OD" or ROLECODE1="VC" or ROLECODE1="AV" or
ROLECODE1="CEO" or ROLECODE1="CFO" or ROLECODE1="CI" or
ROLECODE1="CO" or ROLECODE1="CT" or ROLECODE1="EVP" or ROLECODE1="O"
or ROLECODE1="OB" or ROLECODE1="OP" or ROLECODE1="OS" or
ROLECODE1="OT" or ROLECODE1="OX" or ROLECODE1="P" or ROLECODE1="S" or
ROLECODE1="SVP" or ROLECODE1="VP"; id=_n_;
run;

/*Insider trading*/
proc sql; create table quarter4 as select
        a.*, b.cusip6, b.TRANDATE, b.id as insiderid, b.SHARES_ADJ
        from quarter4 as a left join work1 as b
        on a.cusip6=b.cusip6
        and a.datadate-92<=b.TRANDATE<=datadate;
quit;

proc sort data=quarter4 nodupkey; by id insiderid; run;

proc sql; create table quarter4 as select
        *,count(*) as no_insider, sum(SHARES_ADJ) as sum_SHARES_ADJ
        from quarter4 group by id;
quit;

data quarter4; set quarter4;
if insiderid=. then no_insider=0;
if insiderid=. then sum_SHARES_ADJ=0;
run;

data quarter4; set quarter4; insider_amt=abs(sum_SHARES_ADJ/(CSHOq*1000000))*100;
drop TRANDATE insiderid SHARES_ADJ;
run;

proc sort data=quarter4 nodupkey; by id; run;

/*Insider ownership*/
proc sql; create table quarter4_1 as select
        a.*, b.cusip6, b.TRANDATE, b.PERSONID, b.SHARESHELD
        from quarter4 as a left join work1 as b
        on a.cusip6=b.cusip6
        and a.datadate-1095<=b.TRANDATE<=datadate;
quit;

proc sort data=quarter4_1 nodupkey; by id PERSONID descending TRANDATE; run;
proc sort data=quarter4_1 nodupkey; by id PERSONID; run;
```

```
data quarter4_1; set quarter4_1; if SHARESHELD=. then SHARESHELD=0; run;

proc sql; create table quarter4_1 as select
          *, sum(SHARESHELD) as insider_shares
          from quarter4_1 group by id;
quit;

proc sort data=quarter4_1 nodupkey; by id; run;
data quarter4_1; set quarter4_1;
insider_own=(insider_shares/(CSHOq*1000000));
run;

proc means data=quarter4_1 n nmiss; run;


*****cleaning the sample;
data quarter4_1; set quarter4_1; by gvkey; if first.gvkey then firmid+1; run; /*firm-id*/

data quarter4_1; set quarter4_1;
if ("28Feb2014"d < date) then post=1;
if ("28Feb2014"d >= date) then post=0;
abs_car = abs(car);
run;

proc means data=quarter4_1 n nmiss; run;

data quarter5; set quarter4_1;
if xrdq=. then xrdq=0;
if saleq=. then delete;
if ret_vol=. then delete;
if spread2_mean=. then delete;
if ILLIQ_mean=. then delete;
if TURN_mean=. then delete;
if emp=. then delete;
if short=. then delete;
if instown=. then instown=0; if instown>1 then instown=1;
if dedown=. then dedown=0; if dedown>1 then dedown=1;
if qixown=. then qixown=0; if qixown>1 then qixown=1;
if traown=. then traown=0; if traown>1 then traown=1;
if year(datadate)<2007 then delete;
if non_anon_count=. then non_anon_count=0;
diff=(anon_count)/(anon_count+non_anon_count);
if ps=. then delete;
run;

proc means data=quarter5 n nmiss; run;
```

```
data quarter5; set quarter5; retail=1-instown-insider_own; run;

*Merge on RavenPack data (Dow Jones News article) – dk.rp_count_deter contains the number
of Dow Jones news articles on RavenPack with a relevance score of 95 or higher, excluding SA
coverage, during the firm-quarter;
data news; set dk.rp_count_deter; run;

proc sql; create table quarter6 as select
        a.*, b.gvkey, b.datadate, b.num_RP_dj
        from quarter5 as a left join news as b
        on a.gvkey=b.gvkey
        and a.datadate=b.datadate;
quit;

*winsorize;
%WT(data = quarter6, out = quarter6, byvar = none, vars = quarter5, out = quarter6, byvar =
none, vars = anon_count non_anon_count diff ana_follow insider_own retail ROA ATQ MVE
MTB PS AGE spread2_mean ret_vol turn_mean car short instown dedown qixown traown
no_insider insider_amt emp xrdq num_RP_dj, type = W, pctl = 1 99)

/**** Table 1, Panel B ****/
proc means data = quarter6 n mean std q1 median q3 min max;
var anon_count non_anon_count num_RP_dj MVE ana_follow ROA car retail insider_own short
turn  mean spread2_mean ret_vol MTB xrdq;
run;

data dk.determinant_0723; set quarter6; run; /*188,850*/
```

Article-level observations

```
*For each article in "matched_article3" (Seeking Alpha article data w/ author categorization ) get the returns t+0, t+2 around the article date;
data article; set dk.matched_article3; run;

%trading_days(dset=article, t=date, Prior=N,trading_days=2,outdset=article1);
%trading_days(dset=article1, t=date_2_f, Prior=N,trading_days=15,outdset=article1a);
%trading_days(dset=article1a, t=date_2_f, Prior=N,trading_days=30,outdset=article2);

rsubmit;
proc upload data = article2; run;

data dsf1; set crsp.dsf
(keep = permno cusip date ret prc vol ASKHI BIDLO Ask Bid shrout);
where year(date) >= 1994;
prc = abs(prc);
if ask < bid then ask = .;
if ask-bid > 10 then ask = .;
run;

/* Get returns over the 3 day window [0,2]*/
proc sql; create table temp as select
        a.id, a.permno, b.cusip, a.date, b. ret, b. date as m_date
        from article2 as a left join dsf1 as b
        on a.permno eq b.permno
        and a.date <= b.date <= a.date_2_f;
quit;

proc sql; create table temp1 as select
        a.*, b.vwretd, (a.ret - b.vwretd) as ret_mrktadj
        from illiqtemp as a left join crsp.dsi as b
        on a.m_date eq b.date;
quit;

proc sql; create table temp2 as select distinct
        id, permno, cusip, date, count(ret) as count3,(exp(sum(log(1+ret_mrktadj))) - 1) as CAR
        from temp1 group by id, permno, date;
quit;

data temp3; set temp2; if count3 < 3 then CAR = .; run; *remove values with less than 3 observations;

proc sql; create table article3 as select distinct
        a. *, b.CAR
        from article2 as a left join temp3 as b
        on a. id = b. id;
```

```
quit;


/* Get returns over the following 15 days [3,17]*/
proc sql; create table tempA as select
        a.id, a.permno, b.cusip, a.date, b.ret, b.date as m_date
        from article2 as a left join dsf1 as b
        on a.permno eq b.permno
        and a.date_2_f < b.date <= a.date_15_f;
quit;

proc sql; create table tempA1 as select
        a.*, b.vwretd, (a.ret - b.vwretd) as ret_mrktadj
        from tempA as a left join crsp.dsi as b
        on a.m_date eq b.date;
quit;

proc sql; create table tempA2 as select distinct
        id, permno, cusip, date, count(ret) as count3, (exp(sum(log(1+ret_mrktadj))) - 1) as CAR
        from tempA1 group by id, permno, date;
quit;

data tempA3; set tempA2; if count3 < 9 then CAR = .; run; *remove values with less than 9
observations;

proc sql; create table article6 as select distinct
        a. *, b.CAR as CAR_15
        from article5 as a left join tempA3 as b
        on a. id = b. id;
quit;

/*Get returns over the following 30 days [3,32]*/
proc sql; create table tempB as select
        a.id, a.permno, b.cusip, a.date, b. ret, b. date as m_date
        from article2 as a left join dsf1 as b
        on a.permno eq b.permno
        and a.date_2_f < b.date <= a.date_30_f;
quit;

proc sql; create table tempB1 as select
        a.*, b.vwretd, (a.ret - b.vwretd) as ret_mrktadj
        from tempBas a left join crsp.dsi as b
        on a.m_date eq b.date;
quit;

proc sql; create table tempB2 as select distinct
```

```
        id, permno, cusip, date, count(ret) as count3, (exp(sum(log(1+ret_mrktadj))) - 1) as CAR
        from tempB1 group by id, permno, date;
quit;

data tempB3; set tempB2; if count3 < 25 then CAR = .; run; *remove values with less than 25
observations;

proc sql; create table article7 as select distinct
        a. *, b.CAR as CAR_30
        from article6 as a left join tempB3 as b
        on a. id = b. id;
quit;

*merge on firm attributes;
data compq; set comp.fundq
(keep = gvkey cusip datadate rdq indfmt datafmt popsrc consol curcdq fyr fqtr fyearq tic CSHOq
ATq NIQ IBQ SALEQ CEQQ XRDQ CSHOQ PRCCQ INTANQ dlttq dlcq);
where (indfmt='INDL') and (datafmt='STD') and (popsrc='D') and (consol='C') and
(curcdq='USD') and (atq >= 0) and (year(datadate)> 1990);
drop indfmt datafmt popsrc consol curcdq;
MVE = CSHOQ*PRCCQ;
MTB = MVE/CEQQ;
ROA = NIQ/ATQ;
Leverage = (dlttq+dlcq)/atq;
run;

%dupl(dset=compq,keys=gvkey datadate);
data nodup; set nonduplicates; run;
%dupl(dset=nodup,keys=gvkey rdq);

proc sql; create table article8 as select distinct
        a. *, b.ATq, b.MVE, b.MTB, b.ROA, b.Leverage, b.rdq, b.cusip, b.CSHOq, b.prccq
        from article7 as a left join nonduplicates as b
        on a. gvkey = b. gvkey
        and INTNX( 'day', a.date, -100, 'S' )<= b.rdq <= a.date;
quit;

proc sort data = article8; by id date decending rdq; run;
proc sort data = article8 nodupkey; by id date; run;

*get prior market attributes;
%trading_days(dset=article8, t=date, Prior=Y,trading_days=64,outdset=article8a);

proc sql; create table temp as select
        a.id, a.permno, a.date,(b. ASK - b. BID)/((b. ASK + b. BID)/2) as spread2_ratio, b. ret, b.
        vol/(b. shrout*1000) as dailyturn, b. date as m_date
```

```
        from article8a as a left join dsf1 as b
        on a.permno eq b.permno
        and a.date_64_p <= b.date < a.date;
quit;

proc sql; create table temp1 as select
        a.*, b.vwretd, (a.ret - b.vwretd) as ret_mrktadj
        from temp as a left join crsp.dsi as b
        on a.m_date eq b.date;
quit;

proc sql; create table temp2 as select distinct
        id, permno, date, mean(spread2_ratio) as spread2_mean, mean(dailyturn) as
        TURN_mean, std(ret) as ret_vol, count(spread2_ratio) as count2a, count(ret) as count3,
        count(dailyturn) as count4, (exp(sum(log(1+ret_mrktadj))) - 1) as CAR
        from temp1 group by id, permno, date;
quit;

proc sql; create table article9 as select distinct
        a. *, b. spread2_mean as p_spread2_mean,
        b. ret_vol as p_ret_vol, b. turn_mean as p_turn_mean, b.CAR as p_CAR
        from article8a as a left join temp2 as b
        on a. id = b. id;
quit;

proc download data = article9; run;
endrsubmit;

*get date of last 8-k — dk.edgar_filings_1993to2018 contains the list of SEC filings on EDGAR;
data filing; set dk.edgar_filings_1993to2018; if formtype="8-K"; if year(date)>2004; run;
rsubmit; proc download data = comp.names; run; endrsubmit;
data names; set names; cik1 = cik*1;run;

proc sql; create table filing1 as select distinct
        a. *, b. gvkey
        from filing as a inner join names as b
        on a. cik = b. cik1
        and year1 <= year(a.date) <= year2;
quit;

proc sql; create table article9 as select distinct
        a. *, b. date as date_8k
        from article9 as a left join filing1 as b
        on a. gvkey = b. gvkey
        and INTNX( 'day', a.date, -100, 'S' )<= b.date<= a.date;
quit;
```

```
proc sort data = article9; by id date decending date_8k; run;
proc sort data = article9 nodupkey; by id date; run;

data article9; set article9;
if ("28Feb2014"d < date) then post=1; if ("28Feb2014"d >= date) then post=0;
run;


*merge on article attributes (See Perl code in this file for details on how text attributes were calculated);
proc import datafile = 'J:\JGSM PHD Home\ek594\Travis\SA_text_attributes.csv' out =
work.Attributes dbms = CSV replace; run;

proc sql; create table attributes1 as select distinct
        a.*, b.*
        from outofhere as a left join Attributes as b
        on (a.id= b.id);
quit;

data attributes3; set attributes1;
ln_words = log(num_words);
numnum_rate = num_of_num/num_words;
tone = (positive_words-negative_words)/num_words;
modal_rate = ((strongmodal_words-weakmodal_words)/num_words)*100;
financial_rate = (financial_words/num_words)*100;
forward_rate = (future_count/num_words)*100;
run;

/*both anonymous & non-anonymous articles exists on the same date --> delete*/
proc sql; create table test as select distinct
        gvkey, date, sum(anonymous) as count_anonymous, count(anonymous) as total
        from attributes3 group by gvkey, date;
quit;

data test1; set test;
count_nonanonymous = total - count_anonymous;
if count_anonymous & count_nonanonymous > 0 then delete;
run;

proc sql; create table attributes4 as select
        a.*
        from attributes3 as a inner join test1 as b
        on a.gvkey eq b.gvkey
        and a.date = b.date;
quit;
```

```
proc means data=attributes4 n nmiss; run;

data attributes4; set attributes4;
if p_ret_vol=. then delete;
if leverage=. then delete;
if mtb=. then delete;
if num_words=. then delete;
if roa=. then delete;
if mve=. then delete;
if car=. then delete;
if CAR_30=. then delete;
if p_CAR=. then delete;
abs_car = abs(car);
rdq_diff = date - rdq;
file_diff = date - date_8k;
if file_diff=. then file_diff = 100;
if rdq_diff < 7 then recent_report = 1; else recent_report=0;
if file_diff < 7 then recent_filing = 1; else recent_filing=0;
run;

data acc; set attributes4;
rename date=p_date;
if tone>0 AND CAR_15>0 then ACC15=1;
if tone>0 AND CAR_15<=0 then ACC15=0;
if tone<=0 AND CAR_15>0 then ACC15=0;
if tone<=0 AND CAR_15<=0 then ACC15=1;
if ACC15=1 then ACC15_2=abs(tone)*abs(CAR_15);
if ACC15=0 then ACC15_2=abs(tone)*abs(CAR_15)*-1; /*15-day accuracy*/
run;

proc sql; create table attributes4_1 as select
        a.*, b.author1, b.p_date, b.acc15_2, abs(b.car) as impact
        from attributes4 as a
        left join acc as b
        on a.author1=b.author1
        and a.date-365<=b.p_date<date-30;
quit;

proc sort data=attributes4_1; by id p_date; run;

proc sql; create table attributes4_1 as select
        *, mean(acc15_2) as m_ACC15
        from attributes4_1 group by id;
quit;
```

```
proc sort data=attributes4_1 nodupkey; by id; run;

data attributes4_2; set attributes4_1; drop p_date acc15_2; run;

/**Analyst following**/
data ibes_detail; set dk.ibesdetail_09292019; run; /*I/B/E/S>Detail History (Detail file with
history)*/
data ibes_detail; set ibes_detail; id=_n_; run;

/*add permno to I/B/E/S dataset using iclink file*/
proc sql; create table ibes_detail as select
        a.*, b.ticker, b.permno
        from ibes_detail as a, dk.iclink as b
        where a.ticker=b.ticker;
quit;

proc sort data=ibes_detail nodupkey; by id; run;

proc sql; create table attributes5 as select
        a.*, b.ANNDATS, b.permno, b.ANALYS
        from attributes4_2 as a left join ibes_detail as b
        on a.permno=b.permno
        and a.date-365<b.ANNDATS<a.date;
quit;

proc sort data=attributes5 nodupkey; by id ANALYS; run;

proc sql; create table attributes5 as select
        *, count(*) as ana_follow
        from attributes5 group by id;
quit;

proc sort data=attributes5 nodupkey; by id; run;

data attributes5; set attributes5;
if ANALYS=. then ana_follow=0;
drop FPEDATS ANNDATS ANALYS;
run;

*Merge on RavenPack data (Dow Jones News article) – dk.rp_count_ret contains The number of
Dow Jones news articles on RavenPack with a relevance score of 95 or higher, excluding SA
coverage, over event days [-6, 0] relative to the date of article posting;
data news; set store.rp_count_ret; run;

proc sql; create table attributes5_1 as select
        a.*, b.id, b.num_RP, b.num_rp_dj
```

```
        from attributes5 as a left join news as b
        on a.id=b.id;
quit;


*****merge on institutional ownership – See above for calculation of dk.Iisum7_20181005;
data inst; set dk.Iisum7_20181005; ncusip=substr(cusip, 1, 6); run;
data attributes5_1; set attributes5_1; cusip6=substr(cusip, 1, 6); run;

proc sql; create table attributes5_2 as select
        a.*, b.ncusip, b.instown, b.rdate
        from attributes5_1 as a left join inst as b
        on a.cusip6=b.ncusip
        and a.date-180<b.rdate<a.date;
quit;

proc sort data=attributes5_2; by id descending rdate; run;
proc sort data=attributes5_2 nodupkey; by id; run;
data attributes5_2; set attributes5_2; drop cusip ncusip rdate; run;


*****merge on # of Insider trading;
data work; set dk.table1;run; /* Thomson Reuters > Insider Data > Table 1*/
data work; set work;
if TRANCODE="P" or TRANCODE="S";
run; /*open market transaction*/

data work; set work;
drop ADDRESS2 CITY ZIPCODE PHONE AMEND ADDRESS1 STATE COUNTRY
CUSIP2 ROLECODE2 ROLECODE3 ROLECODE4;
run;

data work; set work;
if cusip6="" then delete;
if CLEANSE="R" or CLEANSE="H" or CLEANSE="R";
run; /* clean data: Heron and Lie 2007 JFE*/

proc sort data=work; by CUSIP6 TRANDATE PERSONID; run; /*leave only directors and
officers (Ali Wei Zhou 2011 JAR)*/

data work1; set work;
if ROLECODE1="CB" or ROLECODE1="D" or ROLECODE1="DO" or ROLECODE1="H"
or ROLECODE1="OD" or ROLECODE1="VC" or ROLECODE1="AV" or
ROLECODE1="CEO" or ROLECODE1="CFO" or ROLECODE1="CI" or
ROLECODE1="CO" or ROLECODE1="CT" or ROLECODE1="EVP" or ROLECODE1="O"
or ROLECODE1="OB" or ROLECODE1="OP" or ROLECODE1="OS" or
```

```
ROLECODE1="OT" or ROLECODE1="OX" or ROLECODE1="P" or ROLECODE1="S" or
ROLECODE1="SVP" or ROLECODE1="VP"; id=_n_;
run;

data temp; set attributes5_2; keep cusip6 date id; run;

/*Insider ownership*/
proc sql; create table temp1 as select
        a.*, b.cusip6, b.TRANDATE, b.PERSONID, b.SHARESHELD
        from temp as a left join work1 as b
        on a.cusip6=b.cusip6
        and a.date-1095<=b.TRANDATE<=date;
quit;

proc sort data=temp1 nodupkey; by id PERSONID descending TRANDATE; run;
proc sort data=temp1 nodupkey; by id PERSONID; run;

data temp1; set temp1; if SHARESHELD=. then SHARESHELD=0; run;

proc sql; create table temp1 as select
        *, sum(SHARESHELD) as insider_shares
        from temp1 group by id;
quit;

proc sort data=temp1 nodupkey; by id; run;

proc sql; create table attributes5_2 as select
        a.*, b.id, b.insider_shares
        from attributes5_2 as a left join temp1 as b
        on a.id=b.id;
quit;

data attributes5_3; set attributes5_2; insider_own=(insider_shares/(CSHOq*1000000));
if instown>1 then instown=1;
if instown<0 then instown=0;
retail=1-insider_own-instown;
run;

data attributes5_4; set attributes5_3;
if m_acc15=. then delete;
if p_spread2_mean=. then delete;
run;

*import in Seeking Alpha comment data – dk.comments_data_like contains the number of
comments and likes that each article on SA received;
proc sql; create table article_id as select
```

```
        article_id, count(article_id) as num_comments, sum(_likes_count) as likes_count
        from dk.comments_data_like group by article_id
        order by article_id;
quit;

proc sql; create table article_id_alt1 as select distinct
        a.*, b.likes_count, b.num_comments
        from store.matched_article3 as a left join article_id as b
        on a.id = b.article_id;
quit;

data article_id_alt2; set article_id_alt1;
if missing(likes_count) then likes_count = 0;
if missing(num_comments) then num_comments = 0;
run;

*merge to main dataset;
data copy; set attributes5_4; run;

proc sql; create table temp as select distinct
        a.id, a.date, b.id as prior_id, b.date as prior_date
        from attributes5_4 as a left join copy as b
        on a.author1 = b.author1
        and INTNX('YEAR', a.date, -1, 'S') <= b.date <= a.date
        and a.id ne b.id;
quit;

proc sql; create table temp1 as select distinct
        a.*, b.num_comments, b.likes_count from temp as a left join article_id_alt2 as b
        on a.prior_id = b.id;
quit;

proc sql; create table temp2 as select distinct
        id, date, mean(num_comments) as ave_comment_counts, mean(likes_count) as ave_likes
        from temp1 group by id, date
        order by id, date;
quit;

data temp2a; set temp2;
if missing(ave_comment_counts) then ave_comment_counts = 0;
if missing(ave_likes) then ave_likes = 0;
run;

proc sql; create table attributes5_4 as select distinct
        a.*, b.ave_comment_counts, b.ave_likes
        from attributes5_4 as a left join temp2a as b
```

```
        on a.id = b.id;
quit;

%WT(data = attributes5_4, out = attributes5_5, byvar = none,
vars = ave_comment_counts ave_likes
anonymous abs_car post mve ana_follow roa p_car retail insider_own mtb p_turn_mean
p_spread2_mean p_ret_vol
recent_report recent_filing num_RP_dj
forward_rate ln_words modal_rate fog financial_rate tone m_ACC15, type = W, pctl = 1 99)

proc rank data=attributes5_5 groups=10 out=attributes5_5; var ave_comment_counts; ranks
d_histcomments; run;
proc rank data=attributes5_5 groups=10 out=attributes5_5; var ave_likes; ranks
d_commentlikes; run;
proc rank data=attributes5_5 groups=10 out=attributes5_5; var m_acc15; ranks d_m_acc15;
run;

data attributes5_5; set attributes5_5;
reputation=(d_m_acc15+d_histcomments+d_commentlikes);
run;

proc rank data=attributes5_5 groups=10 out=attributes5_5; var reputation; ranks d_reputation;
run;

data attributes5_5; set attributes5_5;
d_reputation=d_reputation/9;
d_histcomments=d_histcomments/9;
d_commentlikes=d_commentlikes/9;
d_m_acc15=d_m_acc15/9;
run;

data dk.return_08062020; set attributes5_5; run;
data attributes5_5; set dk.return_08062020; run;

/**merge with shortint**/
data shortint; set dk.comp_shortint; run; /*<- Compustat North America Fundamentals quarterly
(/wrdslin/comp/sasdata/d_na/fundq)*/
data shortint; set shortint;
shortdate=INTNX( 'month', DATADATE, 1)-1;
format shortdate yymmddn.;
run;

proc sql; create table attributes6 as select
        a.*, b.gvkey, b.SHORTINT, b.shortdate
        from attributes5_5 as a left join shortint as b
        on a.gvkey=b.gvkey and a.date-92<=b.shortdate<a.date;
```

```
quit;

proc sort data=attributes6 nodupkey; by id descending shortdate; run;
proc sort data=attributes6 nodupkey; by id; run;

data attributes6; set attributes6;
short=SHORTINT/(cshoq*1000000);
if short=. then short=0;
run;

/**add R&D**/
proc sql; create table attributes6 as select distinct
          a. *, b.xrdq
          from attributes6 as a left join nonduplicates as b
          on a. gvkey = b. gvkey
          and INTNX( 'day', a.date, -100, 'S' )<= b.rdq <= a.date;
quit;

proc sort data = attributes6; by id date decending rdq; run;
proc sort data = attributes6 nodupkey; by id date; run;

data attributes6; set attributes6; if xrdq=. then xrdq=0; run;

%WT(data = attributes6, out = attributes6, byvar = none, vars = short xrdq, type = W, pctl = 1
99)

/*Table 1 Panel A*/
proc means data = attributes6 n mean std q1 median q3 min max;
var anonymous abs_car post recent_report recent_filing num_RP_dj
mve ana_follow roa p_car retail insider_own short p_turn_mean p_spread2_mean p_ret_vol mtb
xrdq /*determinants*/
m_acc15 ave_comment_counts ave_likes /*reputation*/
forward_rate ln_words modal_rate fog financial_rate tone /*text attributes*/; run;

/*Table 1 Panel C*/
proc corr data=attributes6 pearson;
var anonymous abs_car mve roa mtb p_ret_vol recent_report recent_filing num_RP_dj
d_m_acc15 d_histcomments d_commentlikes d_reputation; run;
proc corr data=attributes6 spearman;
var anonymous abs_car mve roa mtb p_ret_vol recent_report recent_filing num_RP_dj
d_m_acc15 d_histcomments d_commentlikes d_reputation; run;

data dk.return_09252020; set attributes6; run;
```

PERL Code (Text Attributes)

```
#!usr/bin/perl
#use strict;
use warnings;
use File::Slurp;
use Text::CSV;
use Lingua::StopWords qw( getStopWords );
use Lingua::Stem::En;
use Lingua::EN::Fathom;
use Lingua::EN::Sentence qw( get_sentences add_acronyms );

#####################################################################################
#	Code to get Textual attributes from Seeking Alpha Articles;
#####################################################################################
#Specify the directory containing the files that you want to read;

my $direct="G:\\SA_files";

#Specify the directory where you want to place the csv file;
my $outputfile= "C:\\Users\\tad99\\Dropbox\\1-DK\\programs\\SA_text_attributes.csv"; #csv
output file

my $slash='\\';

#####################################################################################
###			Acronyms added to help identify end of sentences.
#####################################################################################

add_acronyms('tech','bldg', 'mfg', 'asst', 'acct', 'llc', 'lp', 'llp', 'lllp', 'pllc', 'dba', 'prtnr', 'prtnrshp',
'exp', 'adr','mgmt');

#####################################################################################
###			create array of words lists;
#####################################################################################

my @positive =
("able","abundance","abundant","acclaimed","accomplish","accomplished","accomplishes","acc
omplishing","accomplishment","accomplishments","achieve","achieved","achievement","achiev
ements","achieves","achieving","adequately","advancement","advancements","advances","advan
cing","advantage","advantaged","advantageous","advantageously","advantages","alliance","allia
nces","assure","assured","assures","assuring","attain","attained","attaining","attainment","attain
ments","attains","attractive","attractiveness","beautiful","beautifully","beneficial","beneficially",
"benefit","benefited","benefiting","benefitted","benefitting","best","better","bolstered","bolsterin
g","bolsters","boom","booming","boost","boosted","breakthrough","breakthroughs","brilliant","c
haritable","collaborate","collaborated","collaborates","collaborating","collaboration","collaborati
ons","collaborative","collaborator","collaborators","compliment","complimentary","compliment
```

ed","complimenting","compliments","conclusive","conclusively","conducive","confident","constructive","constructively","courteous","creative","creatively","creativeness","creativity","delight","delighted","delightful","delightfully","delighting","delights","dependability","dependable","desirable","desired","despite","destined","diligent","diligently","distinction","distinctions","distinctive","distinctively","distinctiveness","dream","easier","easily","easy","effective","efficiencies","efficiency","efficient","efficiently","empower","empowered","empowering","empowers","enable","enabled","enables","enabling","encouraged","encouragement","encourages","encouraging","enhance","enhanced","enhancement","enhancements","enhances","enhancing","enjoy","enjoyable","enjoyably","enjoyed","enjoying","enjoyment","enjoys","enthusiasm","enthusiastic","enthusiastically","excellence","excellent","excelling","excels","exceptional","exceptionally","excited","excitement","exciting","exclusive","exclusively","exclusiveness","exclusives","exclusivity","exemplary","fantastic","favorable","favorably","favored","favoring","favorite","favorites","friendly","gain","gained","gaining","gains","good","great","greater","greatest","greatly","greatness","happiest","happily","happiness","happy","highest","honor","honorable","honored","honoring","honors","ideal","impress","impressed","impresses","impressing","impressive","impressively","improve","improved","improvement","improvements","improves","improving","incredible","incredibly","influential","informative","ingenuity","innovate","innovated","innovates","innovating","innovation","innovations","innovative","innovativeness","innovator","innovators","insightful","inspiration","inspirational","integrity","invent","invented","inventing","invention","inventions","inventive","inventiveness","inventor","inventors","leadership","leading","loyal","lucrative","meritorious","opportunities","opportunity","optimistic","outperform","outperformed","outperforming","outperforms","perfect","perfected","perfectly","perfects","pleasant","pleasantly","pleased","pleasure","plentiful","popular","popularity","positive","positively","preeminence","preeminent","premier","premiere","prestige","prestigious","proactive","proactively","proficiency","proficient","proficiently","profitability","profitable","profitably","progress","progressed","progresses","progressing","prospered","prospering","prosperity","prosperous","prospers","rebound","rebounded","rebounding","receptive","regain","regained","regaining","resolve","revolutionize","revolutionized","revolutionizes","revolutionizing","reward","rewarded","rewarding","rewards","satisfaction","satisfactorily","satisfactory","satisfied","satisfies","satisfy","satisfying","smooth","smoothing","smoothly","smooths","solves","solving","spectacular","spectacularly","stability","stabilization","stabilizations","stabilize","stabilized","stabilizes","stabilizing","stable","strength","strengthen","strengthened","strengthening","strengthens","strengths","strong","stronger","strongest","succeed","succeeded","succeeding","succeeds","success","successes","successful","successfully","superior","surpass","surpassed","surpasses","surpassing","transparency","tremendous","tremendously","unmatched","unparalleled","unsurpassed","upturn","upturns","valuable","versatile","versatility","vibrancy","vibrant","win","winner","winners","winning","worthy");
my @negative =
("abandon","abandoned","abandoning","abandonment","abandonments","abandons","abdicated","abdicates","abdicating","abdication","abdications","aberrant","aberration","aberrational","aberrations","abetting","abnormal","abnormalities","abnormality","abnormally","abolish","abolished","abolishes","abolishing","abrogate","abrogated","abrogates","abrogating","abrogation","abrogations","abrupt","abruptly","abruptness","absence","absences","absenteeism","abuse","abused","abuses","abusing","abusive","abusively","abusiveness","accident","accidental","accidentally","accidents","accusation","accusations","accuse","accused","accuses","accusing","acquiesce","acquiesced","acquiesces","acquiescing","acquit","acquits","acquittal","acquittals","acquitted","acquitting","adulterate","adulterated","adulterating","adulteration","adulterations","adversarial","advers

aries","adversary","adverse","adversely","adversities","adversity","aftermath","aftermaths","agai
nst","aggravate","aggravated","aggravates","aggravating","aggravation","aggravations","alerted"
,"alerting","alienate","alienated","alienates","alienating","alienation","alienations","allegation","
allegations","allege","alleged","allegedly","alleges","alleging","annoy","annoyance","annoyance
s","annoyed","annoying","annoys","annul","annulled","annulling","annulment","annulments","a
nnuls","anomalies","anomalous","anomalously","anomaly","anticompetitive","antitrust","argue",
"argued","arguing","argument","argumentative","arguments","arrearage","arrearages","arrears","
arrest","arrested","arrests","artificially","assault","assaulted","assaulting","assaults","assertions",
"attrition","aversely","backdating","bad","bail","bailout","balk","balked","bankrupt","bankruptci
es","bankruptcy","bankrupted","bankrupting","bankrupts","bans","barred","barrier","barriers","b
ottleneck","bottlenecks","boycott","boycotted","boycotting","boycotts","breach","breached","bre
aches","breaching","break","breakage","breakages","breakdown","breakdowns","breaking","bre
aks","bribe","bribed","briberies","bribery","bribes","bribing","bridge","broken","burden","burde
ned","burdening","burdens","burdensome","burned","calamities","calamitous","calamity","cance
l","canceled","canceling","cancellation","cancellations","cancelled","cancelling","cancels","carel
ess","carelessly","carelessness","catastrophe","catastrophes","catastrophic","catastrophically","c
aution","cautionary","cautioned","cautioning","cautions","cease","ceased","ceases","ceasing","c
ensure","censured","censures","censuring","challenge","challenged","challenges","challenging",
"chargeoffs","circumvent","circumvented","circumventing","circumvention","circumventions","
circumvents","claiming","claims","clawback","closed","closeout","closeouts","closing","closing
s","closure","closures","coerce","coerced","coerces","coercing","coercion","coercive","collapse"
,"collapsed","collapses","collapsing","collision","collisions","collude","colluded","colludes","col
luding","collusion","collusions","collusive","complain","complained","complaining","complains
","complaint","complaints","complicate","complicated","complicates","complicating","complica
tion","complications","compulsion","concealed","concealing","concede","conceded","concedes",
"conceding","concern","concerned","concerns","conciliating","conciliation","conciliations","con
demn","condemnation","condemnations","condemned","condemning","condemns","condone","c
ondoned","confess","confessed","confesses","confessing","confession","confine","confined","co
nfinement","confinements","confines","confining","confiscate","confiscated","confiscates","conf
iscating","confiscation","confiscations","conflict","conflicted","conflicting","conflicts","confront
","confrontation","confrontational","confrontations","confronted","confronting","confronts","con
fuse","confused","confuses","confusing","confusingly","confusion","conspiracies","conspiracy",
"conspirator","conspiratorial","conspirators","conspire","conspired","conspires","conspiring","co
ntempt","contend","contended","contending","contends","contention","contentions","contentious
","contentiously","contested","contesting","contraction","contractions","contradict","contradicte
d","contradicting","contradiction","contradictions","contradictory","contradicts","contrary","cont
roversial","controversies","controversy","convict","convicted","convicting","conviction","convic
tions","corrected","correcting","correction","corrections","corrects","corrupt","corrupted","corru
pting","corruption","corruptions","corruptly","corruptness","costly","counterclaim","counterclai
med","counterclaiming","counterclaims","counterfeit","counterfeited","counterfeiter","counterfei
ters","counterfeiting","counterfeits","countermeasure","countermeasures","crime","crimes","cri
minal","criminally","criminals","crises","crisis","critical","critically","criticism","criticisms","cri
ticize","criticized","criticizes","criticizing","crucial","crucially","culpability","culpable","culpabl
y","cumbersome","curtail","curtailed","curtailing","curtailment","curtailments","curtails","cut","
cutback","cutbacks","cyberattack","cyberattacks","cyberbullying","cybercrime","cybercrimes","
cybercriminal","cybercriminals","damage","damaged","damages","damaging","dampen","dampe

ned","danger","dangerous","dangerously","dangers","deadlock","deadlocked","deadlocking","de adlocks","deadweight","deadweights","debarment","debarments","debarred","deceased","deceit" ,"deceitful","deceitfulness","deceive","deceived","deceives","deceiving","deception","deceptions ","deceptive","deceptively","decline","declined","declines","declining","deface","defaced","defa cement","defamation","defamations","defamatory","defame","defamed","defames","defaming"," default","defaulted","defaulting","defaults","defeat","defeated","defeating","defeats","defect","d efective","defects","defend","defendant","defendants","defended","defending","defends","defens ive","defer","deficiencies","deficiency","deficient","deficit","deficits","defraud","defrauded","de frauding","defrauds","defunct","degradation","degradations","degrade","degraded","degrades"," degrading","delay","delayed","delaying","delays","deleterious","deliberate","deliberated","delib erately","delinquencies","delinquency","delinquent","delinquently","delinquents","delist","delist ed","delisting","delists","demise","demised","demises","demising","demolish","demolished","de molishes","demolishing","demolition","demolitions","demote","demoted","demotes","demoting" ,"demotion","demotions","denial","denials","denied","denies","denigrate","denigrated","denigrat es","denigrating","denigration","deny","denying","deplete","depleted","depletes","depleting","de pletion","depletions","deprecation","depress","depressed","depresses","depressing","deprivation ","deprive","deprived","deprives","depriving","derelict","dereliction","derogatory","destabilizati on","destabilize","destabilized","destabilizing","destroy","destroyed","destroying","destroys","d estruction","destructive","detain","detained","detention","detentions","deter","deteriorate","deter iorated","deteriorates","deteriorating","deterioration","deteriorations","deterred","deterrence","d eterrences","deterrent","deterrents","deterring","deters","detract","detracted","detracting","detri ment","detrimental","detrimentally","detriments","devalue","devalued","devalues","devaluing"," devastate","devastated","devastating","devastation","deviate","deviated","deviates","deviating"," deviation","deviations","devolve","devolved","devolves","devolving","difficult","difficulties","d ifficultly","difficulty","diminish","diminished","diminishes","diminishing","diminution","disadv antage","disadvantaged","disadvantageous","disadvantages","disaffiliation","disagree","disagree able","disagreed","disagreeing","disagreement","disagreements","disagrees","disallow","disallo wance","disallowances","disallowed","disallowing","disallows","disappear","disappearance","di sappearances","disappeared","disappearing","disappears","disappoint","disappointed","disappoin ting","disappointingly","disappointment","disappointments","disappoints","disapproval","disapp rovals","disapprove","disapproved","disapproves","disapproving","disassociates","disassociating ","disassociation","disassociations","disaster","disasters","disastrous","disastrously","disavow"," disavowal","disavowed","disavowing","disavows","disciplinary","disclaim","disclaimed","discla imer","disclaimers","disclaiming","disclaims","disclose","disclosed","discloses","disclosing","di scontinuance","discontinuances","discontinuation","discontinuations","discontinue","discontinue d","discontinues","discontinuing","discourage","discouraged","discourages","discouraging","dis credit","discredited","discrediting","discredits","discrepancies","discrepancy","disfavor","disfav ored","disfavoring","disfavors","disgorge","disgorged","disgorgement","disgorgements","disgor ges","disgorging","disgrace","disgraceful","disgracefully","dishonest","dishonestly","dishonesty ","dishonor","dishonorable","dishonorably","dishonored","dishonoring","dishonors","disincentiv es","disinterested","disinterestedly","disinterestedness","disloyal","disloyally","disloyalty","dism al","dismally","dismiss","dismissal","dismissals","dismissed","dismisses","dismissing","disorder ly","disparage","disparaged","disparagement","disparagements","disparages","disparaging","dis paragingly","disparities","disparity","displace","displaced","displacement","displacements","dis places","displacing","dispose","dispossess","dispossessed","dispossesses","dispossessing","dispr oportion","disproportional","disproportionate","disproportionately","dispute","disputed","dispute s","disputing","disqualification","disqualifications","disqualified","disqualifies","disqualify","dis qualifying","disregard","disregarded","disregarding","disregards","disreputable","disrepute","dis rupt","disrupted","disrupting","disruption","disruptions","disruptive","disrupts","dissatisfaction", "dissatisfied","dissent","dissented","dissenter","dissenters","dissenting","dissents","dissident","di ssidents","dissolution","dissolutions","distort","distorted","distorting","distortion","distortions"," distorts","distract","distracted","distracting","distraction","distractions","distracts","distress","dis tressed","disturb","disturbance","disturbances","disturbed","disturbing","disturbs","diversion","d ivert","diverted","diverting","diverts","divest","divested","divesting","divestiture","divestitures", "divestment","divestments","divests","divorce","divorced","divulge","divulged","divulges","div ulging","doubt","doubted","doubtful","doubts","downgrade","downgraded","downgrades","dow ngrading","downsize","downsized","downsizes","downsizing","downsizings","downtime","dow ntimes","downturn","downturns","downward","downwards","drag","drastic","drastically","draw back","drawbacks","dropped","drought","droughts","duress","dysfunction","dysfunctional","dys functions","easing","egregious","egregiously","embargo","embargoed","embargoes","embargoin g","embarrass","embarrassed","embarrasses","embarrassing","embarrassment","embarrassments ","embezzle","embezzled","embezzlement","embezzlements","embezzler","embezzles","embezz ling","encroach","encroached","encroaches","encroaching","encroachment","encroachments","e ncumber","encumbered","encumbering","encumbers","encumbrance","encumbrances","endange r","endangered","endangering","endangerment","endangers","enjoin","enjoined","enjoining","en joins","erode","eroded","erodes","eroding","erosion","erratic","erratically","erred","erring","erro neous","erroneously","error","errors","errs","escalate","escalated","escalates","escalating","evad e","evaded","evades","evading","evasion","evasions","evasive","evict","evicted","evicting","evi ction","evictions","evicts","exacerbate","exacerbated","exacerbates","exacerbating","exacerbatio n","exacerbations","exaggerate","exaggerated","exaggerates","exaggerating","exaggeration","ex cessive","excessively","exculpate","exculpated","exculpates","exculpating","exculpation","excul pations","exculpatory","exonerate","exonerated","exonerates","exonerating","exoneration","exo nerations","exploit","exploitation","exploitations","exploitative","exploited","exploiting","exploi ts","expose","exposed","exposes","exposing","expropriate","expropriated","expropriates","expro priating","expropriation","expropriations","expulsion","expulsions","extenuating","fail","failed", "failing","failings","fails","failure","failures","fallout","falsely","falsification","falsifications","fa lsified","falsifies","falsify","falsifying","falsity","fatalities","fatality","fatally","fault","faulted","f aults","faulty","fear","fears","felonies","felonious","felony","fictitious","fined","fines","fired","fi ring","flaw","flawed","flaws","forbid","forbidden","forbidding","forbids","force","forced","forci ng","foreclose","foreclosed","forecloses","foreclosing","foreclosure","foreclosures","forego","fo regoes","foregone","forestall","forestalled","forestalling","forestalls","forfeit","forfeited","forfeit ing","forfeits","forfeiture","forfeitures","forgers","forgery","fraud","frauds","fraudulence","fraud ulent","fraudulently","frivolous","frivolously","frustrate","frustrated","frustrates","frustrating","f rustratingly","frustration","frustrations","fugitive","fugitives","gratuitous","gratuitously","grieva nce","grievances","grossly","groundless","guilty","halt","halted","hamper","hampered","hamperi ng","hampers","harass","harassed","harassing","harassment","hardship","hardships","harm","har med","harmful","harmfully","harming","harms","harsh","harsher","harshest","harshly","harshne ss","hazard","hazardous","hazards","hinder","hindered","hindering","hinders","hindrance","hind rances","hostile","hostility","hurt","hurting","idle","idled","idling","ignore","ignored","ignores", "ignoring","ill","illegal","illegalities","illegality","illegally","illegible","illicit","illicitly","illiquid ","illiquidity","imbalance","imbalances","immature","immoral","impair","impaired","impairing" ,"impairment","impairments","impairs","impasse","impasses","impede","impeded","impedes","i

mpediment","impediments","impeding","impending","imperative","imperfection","imperfection
s","imperil","impermissible","implicate","implicated","implicates","implicating","impossibility",
"impossible","impound","impounded","impounding","impounds","impracticable","impractical","
impracticalities","impracticality","imprisonment","improper","improperly","improprieties","imp
ropriety","imprudent","imprudently","inability","inaccessible","inaccuracies","inaccuracy","inac
curate","inaccurately","inaction","inactions","inactivate","inactivated","inactivates","inactivating
","inactivation","inactivations","inactivity","inadequacies","inadequacy","inadequate","inadequa
tely","inadvertent","inadvertently","inadvisability","inadvisable","inappropriate","inappropriatel
y","inattention","incapable","incapacitated","incapacity","incarcerate","incarcerated","incarcerat
es","incarcerating","incarceration","incarcerations","incidence","incidences","incident","incident
s","incompatibilities","incompatibility","incompatible","incompetence","incompetency","incom
petent","incompetently","incompetents","incomplete","incompletely","incompleteness","inconcl
usive","inconsistencies","inconsistency","inconsistent","inconsistently","inconvenience","inconv
eniences","inconvenient","incorrect","incorrectly","incorrectness","indecency","indecent","indef
easible","indefeasibly","indict","indictable","indicted","indicting","indictment","indictments","in
effective","ineffectively","ineffectiveness","inefficiencies","inefficiency","inefficient","inefficie
ntly","ineligibility","ineligible","inequitable","inequitably","inequities","inequity","inevitable","i
nexperience","inexperienced","inferior","inflicted","infraction","infractions","infringe","infringe
d","infringement","infringements","infringes","infringing","inhibited","inimical","injunction","in
junctions","injure","injured","injures","injuries","injuring","injurious","injury","inordinate","inor
dinately","inquiry","insecure","insensitive","insolvencies","insolvency","insolvent","instability",
"insubordination","insufficiency","insufficient","insufficiently","insurrection","insurrections","in
tentional","interfere","interfered","interference","interferences","interferes","interfering","interm
ittent","intermittently","interrupt","interrupted","interrupting","interruption","interruptions","inte
rrupts","intimidation","intrusion","invalid","invalidate","invalidated","invalidates","invalidating"
,"invalidation","invalidity","investigate","investigated","investigates","investigating","investigati
on","investigations","involuntarily","involuntary","irreconcilable","irreconcilably","irrecoverabl
e","irrecoverably","irregular","irregularities","irregularity","irregularly","irreparable","irreparabl
y","irreversible","jeopardize","jeopardized","justifiable","kickback","kickbacks","knowingly","l
ack","lacked","lacking","lackluster","lacks","lag","lagged","lagging","lags","lapse","lapsed","lap
ses","lapsing","late","laundering","layoff","layoffs","lie","limitation","limitations","lingering","li
quidate","liquidated","liquidates","liquidating","liquidation","liquidations","liquidator","liquidat
ors","litigant","litigants","litigate","litigated","litigates","litigating","litigation","litigations","loc
kout","lockouts","lose","loses","losing","loss","losses","lost","lying","malfeasance","malfunctio
n","malfunctioned","malfunctioning","malfunctions","malice","malicious","maliciously","malpr
actice","manipulate","manipulated","manipulates","manipulating","manipulation","manipulation
s","manipulative","markdown","markdowns","misapplication","misapplications","misapplied","
misapplies","misapply","misapplying","misappropriate","misappropriated","misappropriates","m
isappropriating","misappropriation","misappropriations","misbranded","miscalculate","miscalcul
ated","miscalculates","miscalculating","miscalculation","miscalculations","mischaracterization",
"mischief","misclassification","misclassifications","misclassified","misclassify","miscommunica
tion","misconduct","misdated","misdemeanor","misdemeanors","misdirected","mishandle","mis
handled","mishandles","mishandling","misinform","misinformation","misinformed","misinformi
ng","misinforms","misinterpret","misinterpretation","misinterpretations","misinterpreted","misin
terpreting","misinterprets","misjudge","misjudged","misjudges","misjudging","misjudgment","m
isjudgments","mislabel","mislabeled","mislabeling","mislabelled","mislabels","mislead","mislea

ding","misleadingly","misleads","misled","mismanage","mismanaged","mismanagement","mism
anages","mismanaging","mismatch","mismatched","mismatches","mismatching","misplaced","
misprice","mispricing","mispricings","misrepresent","misrepresentation","misrepresentations","
misrepresented","misrepresenting","misrepresents","miss","missed","misses","misstate","misstat
ed","misstatement","misstatements","misstates","misstating","misstep","missteps","mistake","mi
staken","mistakenly","mistakes","mistaking","mistrial","mistrials","misunderstand","misunderst
anding","misunderstandings","misunderstood","misuse","misused","misuses","misusing","mono
polistic","monopolists","monopolization","monopolize","monopolized","monopolizes","monopo
lizing","monopoly","moratoria","moratorium","moratoriums","mothballed","mothballing","nega
tive","negatively","negatives","neglect","neglected","neglectful","neglecting","neglects","neglig
ence","negligences","negligent","negligently","nonattainment","noncompetitive","noncomplianc
e","noncompliances","noncompliant","noncomplying","nonconforming","nonconformities","non
conformity","nondisclosure","nonfunctional","nonpayment","nonpayments","nonperformance","
nonperformances","nonperforming","nonproducing","nonproductive","nonrecoverable","nonrene
wal","nuisance","nuisances","nullification","nullifications","nullified","nullifies","nullify","nulli
fying","objected","objecting","objection","objectionable","objectionably","objections","obscene"
,"obscenity","obsolescence","obsolete","obstacle","obstacles","obstruct","obstructed","obstructin
g","obstruction","obstructions","offence","offences","offend","offended","offender","offenders",
"offending","offends","omission","omissions","omit","omits","omitted","omitting","onerous","o
pportunistic","opportunistically","oppose","opposed","opposes","opposing","opposition","opposi
tions","outage","outages","outdated","outmoded","overage","overages","overbuild","overbuildin
g","overbuilds","overbuilt","overburden","overburdened","overburdening","overcapacities","ove
rcapacity","overcharge","overcharged","overcharges","overcharging","overcome","overcomes","
overcoming","overdue","overestimate","overestimated","overestimates","overestimating","overe
stimation","overestimations","overload","overloaded","overloading","overloads","overlook","ov
erlooked","overlooking","overlooks","overpaid","overpayment","overpayments","overproduced"
,"overproduces","overproducing","overproduction","overrun","overrunning","overruns","oversha
dow","overshadowed","overshadowing","overshadows","overstate","overstated","overstatement"
,"overstatements","overstates","overstating","oversupplied","oversupplies","oversupply","oversu
pplying","overtly","overturn","overturned","overturning","overturns","overvalue","overvalued","
overvaluing","panic","panics","penalize","penalized","penalizes","penalizing","penalties","penal
ty","peril","perils","perjury","perpetrate","perpetrated","perpetrates","perpetrating","perpetration
","persist","persisted","persistence","persistent","persistently","persisting","persists","pervasive",
"pervasively","pervasiveness","petty","picket","picketed","picketing","plaintiff","plaintiffs","ple
a","plead","pleaded","pleading","pleadings","pleads","pleas","pled","poor","poorly","poses","po
sing","postpone","postponed","postponement","postponements","postpones","postponing","preci
pitated","precipitous","precipitously","preclude","precluded","precludes","precluding","predator
y","prejudice","prejudiced","prejudices","prejudicial","prejudicing","premature","prematurely","
pressing","pretrial","preventing","prevention","prevents","problem","problematic","problematica
l","problems","prolong","prolongation","prolongations","prolonged","prolonging","prolongs","p
rone","prosecute","prosecuted","prosecutes","prosecuting","prosecution","prosecutions","protest
","protested","protester","protesters","protesting","protestor","protestors","protests","protracted",
"protraction","provoke","provoked","provokes","provoking","punished","punishes","punishing",
"punishment","punishments","punitive","purport","purported","purportedly","purporting","purpo
rts","question","questionable","questionably","questioned","questioning","questions","quit","quit
ting","racketeer","racketeering","rationalization","rationalizations","rationalize","rationalized","r

ationalizes","rationalizing","reassessment","reassessments","reassign","reassigned","reassigning
","reassignment","reassignments","reassigns","recall","recalled","recalling","recalls","recession"
,"recessionary","recessions","reckless","recklessly","recklessness","redact","redacted","redacting
","redaction","redactions","redefault","redefaulted","redefaults","redress","redressed","redresses
","redressing","refusal","refusals","refuse","refused","refuses","refusing","reject","rejected","reje
cting","rejection","rejections","rejects","relinquish","relinquished","relinquishes","relinquishing"
,"relinquishment","relinquishments","reluctance","reluctant","renegotiate","renegotiated","reneg
otiates","renegotiating","renegotiation","renegotiations","renounce","renounced","renouncement
","renouncements","renounces","renouncing","reparation","reparations","repossessed","reposses
ses","repossessing","repossession","repossessions","repudiate","repudiated","repudiates","repudi
ating","repudiation","repudiations","resign","resignation","resignations","resigned","resigning","
resigns","restate","restated","restatement","restatements","restates","restating","restructure","rest
ructured","restructures","restructuring","restructurings","retaliate","retaliated","retaliates","retali
ating","retaliation","retaliations","retaliatory","retribution","retributions","revocation","revocatio
ns","revoke","revoked","revokes","revoking","ridicule","ridiculed","ridicules","ridiculing","riski
er","riskiest","risky","sabotage","sacrifice","sacrificed","sacrifices","sacrificial","sacrificing","sc
andalous","scandals","scrutinize","scrutinized","scrutinizes","scrutinizing","scrutiny","secrecy","
seize","seized","seizes","seizing","sentenced","sentencing","serious","seriously","seriousness","s
etback","setbacks","sever","severe","severed","severely","severities","severity","sharply","shock
ed","shortage","shortages","shortfall","shortfalls","shrinkage","shrinkages","shut","shutdown","s
hutdowns","shuts","shutting","slander","slandered","slanderous","slanders","slippage","slippage
s","slow","slowdown","slowdowns","slowed","slower","slowest","slowing","slowly","slowness"
,"sluggish","sluggishly","sluggishness","solvencies","solvency","spam","spammers","spamming
","staggering","stagnant","stagnate","stagnated","stagnates","stagnating","stagnation","standstill"
,"standstills","stolen","stoppage","stoppages","stopped","stopping","stops","strain","strained","st
raining","strains","stress","stressed","stresses","stressful","stressing","stringent","subjected","sub
jecting","subjection","subpoena","subpoenaed","subpoenas","substandard","sue","sued","sues","
suffer","suffered","suffering","suffers","suing","summoned","summoning","summons","summo
nses","susceptibility","susceptible","suspect","suspected","suspects","suspend","suspended","sus
pending","suspends","suspension","suspensions","suspicion","suspicions","suspicious","suspicio
usly","taint","tainted","tainting","taints","tampered","tense","terminate","terminated","terminates
","terminating","termination","terminations","testify","testifying","threat","threaten","threatened
","threatening","threatens","threats","tightening","tolerate","tolerated","tolerates","tolerating","to
leration","tortuous","tortuously","tragedies","tragedy","tragic","tragically","traumatic","trouble",
"troubled","troubles","turbulence","turmoil","unable","unacceptable","unacceptably","unaccount
ed","unannounced","unanticipated","unapproved","unattractive","unauthorized","unavailability",
"unavailable","unavoidable","unavoidably","unaware","uncollectable","uncollected","uncollecti
bility","uncollectible","uncollectibles","uncompetitive","uncompleted","unconscionable","uncon
scionably","uncontrollable","uncontrollably","uncontrolled","uncorrected","uncover","uncovered
","uncovering","uncovers","undeliverable","undelivered","undercapitalized","undercut","underc
uts","undercutting","underestimate","underestimated","underestimates","underestimating","unde
restimation","underfunded","underinsured","undermine","undermined","undermines","undermin
ing","underpaid","underpayment","underpayments","underpays","underperform","underperform
ance","underperformed","underperforming","underperforms","underproduced","underproduction
","underreporting","understate","understated","understatement","understatements","understates",
"understating","underutilization","underutilized","undesirable","undesired","undetected","undete

rmined","undisclosed","undocumented","undue","unduly","uneconomic","uneconomical","unec
onomically","unemployed","unemployment","unethical","unethically","unexcused","unexpected
","unexpectedly","unfair","unfairly","unfavorability","unfavorable","unfavorably","unfavourable
","unfeasible","unfit","unfitness","unforeseeable","unforeseen","unforseen","unfortunate","unfor
tunately","unfounded","unfriendly","unfulfilled","unfunded","uninsured","unintended","unintent
ional","unintentionally","unjust","unjustifiable","unjustifiably","unjustified","unjustly","unknow
ing","unknowingly","unlawful","unlawfully","unlicensed","unliquidated","unmarketable","unme
rchantable","unmeritorious","unnecessarily","unnecessary","unneeded","unobtainable","unoccup
ied","unpaid","unperformed","unplanned","unpopular","unpredictability","unpredictable","unpre
dictably","unpredicted","unproductive","unprofitability","unprofitable","unqualified","unrealistic
","unreasonable","unreasonableness","unreasonably","unreceptive","unrecoverable","unrecovere
d","unreimbursed","unreliable","unremedied","unreported","unresolved","unrest","unsafe","unsa
lable","unsaleable","unsatisfactory","unsatisfied","unsavory","unscheduled","unsellable","unsold
","unsound","unstabilized","unstable","unsubstantiated","unsuccessful","unsuccessfully","unsuit
ability","unsuitable","unsuitably","unsuited","unsure","unsuspected","unsuspecting","unsustaina
ble","untenable","untimely","untrusted","untruth","untruthful","untruthfully","untruthfulness","u
ntruths","unusable","unwanted","unwarranted","unwelcome","unwilling","unwillingness","upset
","urgency","urgent","usurious","usurp","usurped","usurping","usurps","usury","vandalism","ver
dict","verdicts","vetoed","victims","violate","violated","violates","violating","violation","violati
ons","violative","violator","violators","violence","violent","violently","vitiate","vitiated","vitiate
s","vitiating","vitiation","voided","voiding","volatile","volatility","vulnerabilities","vulnerability
","vulnerable","vulnerably","warn","warned","warning","warnings","warns","wasted","wasteful"
,"wasting","weak","weaken","weakened","weakening","weakens","weaker","weakest","weakly",
"weakness","weaknesses","willfully","worries","worry","worrying","worse","worsen","worsene
d","worsening","worsens","worst","worthless","writedown","writedowns","writeoff","writeoffs",
"wrong","wrongdoing","wrongdoings","wrongful","wrongfully","wrongly","false");

my @strongmodal =
("always","best","clearly","definitely","definitively","highest","lowest","must","never","strongly
","unambiguously","uncompromising","undisputed","undoubtedly","unequivocal","unequivocall
y","unparalleled","unsurpassed","will");
my @middlemodal =
("can","frequently","generally","likely","often","ought","probable","probably","rarely","regularl
y","should","tends","usually","would");
my @weakmodal =
("almost","apparently","appeared","appearing","appears","conceivable","could","depend","depe
nded","depending","depends","may","maybe","might","nearly","occasionally","perhaps","possib
le","possibly","seldom","seldomly","sometimes","somewhat","suggest","suggests","uncertain","
uncertainly");

my @financial = ("10-c","10-k","10-q","12b-1 fees","13-d","14a","14d-9","20-f","360 degree
feedback","401(k) plan","5/20 rule","52 week high low","529 plan","6-k","8-k","abandonment
value","abs","accidental death benefit","accounts payable","accounts
receivable","accretion","accrual accounting","accrued expenses","accrued
income","accumulation","achromatic lens","acid-test ratio","acquisition","actual cash
value","actuaries","adr","advance","adverse selection","advertising","advertorial","affirmative

action","affordable housing","agent","aggressive allocation funds","agreement of sale","aicte","algorithm","all ordinaries","all-or-none orders","alpha","american depositary receipt","amex","amortization","amortized loan","analysis","analyst","angel investor","annotations","annual financial statement","annual percentage rate","annuity","appraisal value","appropriation bill","apr","arbitrage","ashes series","asian games","assessed value","asset","assignor","associate","at-the-money","audit","aurora","auto insurance","auto it tool","autocratic leadership","average collection period","back office","backwardation","bailout","balance of payment","balance sheet","balanced funds","balloon payment","bancassurance","bargaining power of buyers","bargaining power of suppliers","barley","bars","base rate","basel accords","basel iii","basis point","bcci","bea","behavioral finance","behavioural competency","ben bernanke","benchmark","beneficiary","beta","bid","big board","bill of lading","bill of sale","black-scholes","block deal","blue chip","blue ocean strategy","board of directors","board of governors","bollinger bands","bond","bonus share","book value","bounce rate","boutique","bps","brand tribe","break the buck","breakaway gap","breakeven","break-even pricing","breakout","brexit","bridge loan","broker","bubble","budgetary deficit","buffer overflow","bulge bracket","bull","bureau of labor statistics","butterfly","buy and hold strategy","buyback","buyout","buy-side","bv / share","cac","calabash tool","call","canadian deposit insurance corporation","capacity cost","capex","capital","capitulation","capm","carpet area","casa","cash & carry trade","cash & cash equivalents","cash accounting","cash cows","cash flow","catch up effect","cdo","cds","cede or ceding company","celss","central bank","central plan","ceo","certified financial planner","cfa","cfo","chana","chandrayaan 1","channel management","chapter 11","chartists","chicago board of trade","chicago mercantile exchange","chief executive officer","chief financial officer","chromosphere","ciphertext","circuit breaker","claim amount","classified board","clean price","clearing price","c-level","clickjacking","close","cme group","cnbc","cogs","collateral","college savings plan","commencement certificate","commercial bank","commitment fee","commodity","common area maintenance","common shareholders","common shares","competency mapping","completion certificate","completion guarantee","compound interest","compounding","concealment","conference board, the","confirmation bias","consolidated fund","consolidation","conspicuous consumption","consumer surplus","contagion","contango","contingency fund","contingent claims","contra fund","contractionary policy","contrarians","coo","corona","corporation tax","correlated","cost and freight","cost leadership","cost of carry","cost of goods sold","countervailing duties","country funds","coupon payments","coupon rate","covenant","coverdell education savings account","credit quality","cross elasticity of demand","crossing systems","cross-site request forgery","cross-site scripting","crowding out effect","crown corporations","crude palm oil","cryogenic engine","cryptography","currency appreciation","currency deposit ratio","current liabilities","current ratio","custodian","customs duty","dark matter","data mining","dating back","davis cup","dax","day order","dcf","dead cat bounce","deadweight loss","dealer","debenture","debugging","decimalization","decision tree model","decline","decryption","default","deferred income tax","defined-benefit plans","defined-contribution plans","deflation","delisting","delta","demand sensitivity","demands for grants","denial-of-service attack","depreciation","depression","digital fingerprinting","dilution","direct tax","direct to home","dirty price","discount rate","discounting","distribution","distributive

bargaining","divergence","diversification","diversified conglomerate","dividend","djia","dns spoofing","dodd-frank","dog","dollar cost averaging","domestic institutional investo","do-not-reduce order","doosra","doppler effect","dow jones industrial average","dow jones transportation average","downgrade","dry docking","dsn","duckworth-lewis system","due date rate","due diligence","earnings","ease of doing business","ebit","ebt","economies of scale","education cess","education ira","effective annual rate","el classico","e-learning","electromagnetic waves","electronics communications networks","embedded value","emi","e-mini contracts","emotional intelligence","employer matching","endorsements","enterprise value","entrances","eps","equity","escalation clause","esop","etf","eurex","euro","e-wallets","excess grants","excessive euphoria","exchange rate","exchange traded fund","exchange-traded fund","exclusions","execution","exercise date","exercise price","exit multiple","expenditure budget","expense ratio","expiration date","expiry date","extraordinary items","extrinsic rewards","face value","fair trade price","fallen angel","fannie mae","fcf","fdic","federal deposit insurance corporation","federal funds rate","federal reserve","fiat money","fibonacci","fiduciary responsibility","fiis","fill-or-kill order","film insurance","finance bill","financial advisor","financial crisis","financial engineering","financial futures","financial institutions group","financial statements","fire and special perils policy","first class life","fiscal deficit","fiscal year","fitch","fixed claim","fixed costs","fixed income","fixed-term employment","flash crash","focus group","football field graph","forbes investors advisory institute","forbes special situation survey","force india","forex","forward contract","forward price","fpo","fractional banking","freddie mac","freehold property","front office","front running","frontage","ftse 100","ftse100","fund of funds","funds for future appropriation","future delivery","futures contract","gaap","gagan","gamma","gantt chart","garp","gasoline","gdp","general insurance","generic strategies","geometric average return","geospace","geo-synchronous satellite","gic","gifting","glass-steagall act","global funds","gnp","gold fund","good-till-canceled","goodwill","gordon growth method","government sponsored enterprise","graph theory","greater fool theory","gross budgetary support","gross domestic product","gross domestic saving","gross national product","gross profit","group policy","growth at a reasonable price","growth funds","gslv","guaranteed interest plans","guaranteed survival benefit","gur","hacking","haircut","hang seng","hawthorne effect","head and shoulders","heating oil","hedge","hedging","heliopause","herding","high density polyethylene","historical price","historical value","home loan","homosphere","housing starts","https","hubble space telescope","ifrs","i-league","immediate-or-cancel order","impact cost","impulsive buying","in the money","income funds","income statement","income trust","indemnity","independent directors","index","indifference curve","individual retirement accounts","industrial production","inferior goods","inflation","information ratio","infrastructure investment trus","inside information","insiders","insolvent","insurability","insurable interest","insurance advisor","intelsat","interest expense","interest income","interest payment","interest rate","internalization","international funds","international monetary fund","interval schemes","in-the-money","intrinsic value","introduction stage","inventory","investing","investment bank","investment horizon","investment newsletters","investor sentiment","invisible hand","invoice finance","involvement","ipo","irnss","iron condor","irr","isro","iss","jobs growth relief reconciliation act of 2003","joint venture","kansas city board of trade","key employee or keyman","key performance indicator","key result areas","kpi","lagging indicators","langrangian point","lapsed policy","large and midcap funds","large cap","latent bug","law of demand","law of supply","lbo","leading indicators","leaps","lender of last resort","letter of intent","leverage","liability","libor","life assured","liffe","limit","line of credit","liquid","load","loan-to-value","london interbank offer rate","long","loss adjustment expense","loss aversion","loss carryforward","loss leaders","lot size","lse","lumpsum","m&a","macro","maintenance requirement","management buy out","management by walking around","management fee","managing director","margin","mark up","markdown","market","markup","matrix organization","maturity","maven","mba","mbo","mcc","mean","megafund","melted menthol flakes","memory corruption","merger","mibor","mid cap","middle office","mid-year discount","minority interest","misrepresentation","mitigation","mobile automation","modigliani and modigliani","monetary policy","money back plans","money supply","monopoly","monte carlo simulation","moody's","moral hazard","mortgage","moving average","multibagger","multicap funds","mutation","mutual fund","myers-briggs type indicator","naics","naked","nasdaq","natural gas","nav","nda","nebula","negotiable certificates of deposit","net income","net national income","net present value","net property plant and equipment","net worth","new business profit","new business strain","new fund offer","new york mercantile exchange","nikkei","nim","no claim bonus","nol","nominee","non disclosure agreement","noncyclical","non-plan expenditure","north american industry classification system","note","number theory","nvocc","nyse","occupancy certificate","odd lots","odd-lot order","oecd","off price retailers","offer for sale","office of the superintendent of financial institutions","oligopoly","onechicago","opportunity cost","option","otc","out of the money","out-of-the-money","outside owner","outsiders","over-the-counter","p/b","p/e","paid up policy","paper","par","path traversal","patwari","payback period","payment for order flow","payments banks","payout ratios","peer appraisal","penetration","pension","percentage lease","perfect competition","performance planning","perpetuity","personal selling","perturbation","phillips curve","pink sheet","pink slip","pitch book","pivot table","plant, property & equipment","plowback rate","poison pill","polyvinylchloride","ponzi scheme","portfolio","position","poverty trap","pp&e","ppe","precedent transactions","precious metals","pre-emptive pricing","preferences","preferencing","preferred shares","prepaid expenses","prepaid income","prepaid tuition plan","price ceiling","price deflator","price floor","price mechanism","price-book ratio","pricing strategies","primary deficit","prime rate","principal","priority in claims","producer surplus","professional money-management companies","profit and loss","profitability ratio","promotions","property tax","proportional tax","proposal form","proprietary desk","prospect theory","prospectus","prototype model","proxy statement","pseudocode","pslv","public","pure research shops","put","qip","qqq","quant","quasar","quick ratio","quota sampling","r&d","radio astronomy","rail transport","rally","random sampling","ranji trophy","ranorex tool","rapid application development","rate of return","rating","rational behaviour","rationing","reaction wheels","real business cycle theory","real economic growth rate","real estate","realized gain","realized loss","rebranding","receipts budget","receivables","recession","record date","red chilli","reference price","referring sites","reg fd","registered retirement savings plan","regressive tax","regulation fair disclosure","reinstatement","reinsurance","reit","rent seeking","repo","requirement traceability matri","research & development","reserve ratio","residual claim","resistance level","resistance line","restructuring","retail banking","retainer","retracement","return grade","return on investment","revenue","reversing trade","revival period","revolver","riders","rights","risk","roa","roe","roi","roth individual retirement funds","r-squared","run rate","runner","russell 2000","russell 3000","s&p 500","s&p500","s-1","s-4","safety reserve","sales","santosh trophy","sarbanes-oxley act","satisficing","scalability","scalar chain","scalping","scheme category","scheme class","scientific revolution","search costs","seasonal adjustment","sec","segmentation","seigniorage","selenium web driver","selling group","selling, general & administrative costs","sell-side","seniority","sensitivity table","series exams","service tax","settlement date","sg&a","shanghai composite","share pledge","shareholder value","shares outstanding","sharpe ratio","short","sikuli tool","simple interest","sinking fund","site/project plans","small cap","soapui","socially responsible funds","soft currency","soft loans","software engineering","software maintenance","solar panel","sortino ratio","sox","special investment vehicle","special purpose vehicle","specialists","specialty products","speculation","speculative motive","speculators","spiral model","spot price","spread","spyware","sql injection","squaring off","staggered board","stimulus package","stochastic oscillator","stock","stop loss","stop order","straddle","stratified sampling","strike price","style box","subprime","subsidy","subvention","succession planning","super area","super montage","superday","supplementary grants","support level","support line","surcharge","surrender value","survey of current business","swap","swot","systematic investment plan","systematic transfer plan","systematic withdrawal plan","systems design","t20","takeover","tarp","tax deferred","tax efficient","tax exempt","tax planning funds","tax sheltered","taxation","taxes","tax-exempt funds","tax-free savings account","t-bill","tender period","term deposit","terminal multiple","terminal value","terrestrial planet","testing","theory x & theory y","theta","threat of new entrants","threat of substitutes","tick","tigrs","time value","times interest earned","to-i","tombstone","total permanent disability","tour de france","trade union","trader","trading","traditional insurance plans","traffic source","tranche","transaction multiple","transport planning","treasury bills","treasury investment growth receipts","trend","treynor ratio","trojan","true cost economics","trustee","tsx","tulip mania","tulipmania","turmeric","under writer","underwriter","underwriting","unemployment","union budget","union excise duty","upgrade","utm","valuation","valued added tax","van allen radiation belts","vanguard 500","var","vat","vega","velocity of circulation","vesting bonus","vice president","vix","volatility","volcker rule","volume","voluntary retirement scheme","wacc","wage curve","wage drift","walkthrough","warehouse clubs","warrant","wash sale","waterfall model","watir","wavelength","web server","whistleblower","wholesale banking","williams %r","wimbledon","windfall gains","world bank","yield","zenith","zero bug bounce","zero gravity","zero-sum investment");

my @wordarray = ("anticipate", "anticipates", "assume", "assumes", "believe", "believes", "commit", "commits", "expect", "expects", "foresee", "foresees", "intend", "intends", "seek", "seeks", "will");
my @bigramarray = ("corporation aims", "corporation estimates", "corporation forecasts", "corporation hopes", "corporation plans", "corporation projects", "corporation targets", "management aims", "management estimates", "management forecasts", "management hopes", "management plans", "management projects", "management targets", "currently aim", "currently aimed", "currently aiming", "currently aims", "currently anticipated", "currently anticipating", "currently assumed", "currently assuming", "currently believed", "currently believing", "currently committed", "currently committing", "currently estimate", "currently estimated",

"currently estimates", "currently estimating", "currently expected", "currently expecting", "currently forecast", "currently forecasted", "currently forecasting", "currently forecasts", "currently foreseeing", "currently foreseen", "currently hope", "currently hoped", "currently hopes", "currently hoping", "currently intended", "currently intending", "currently plan", "currently planed", "currently planning", "currently plans", "currently project", "currently projected", "currently projecting", "currently projects", "currently seeking", "currently sought", "currently target", "currently targeted", "currently targeting", "currently targets", "currently willing", "normally aim", "normally aims", "normally estimate", "normally estimates", "normally forecast", "normally forecasts", "normally hope", "normally hopes", "normally plan", "normally plans", "normally project", "normally projects", "normally target", "normally targets", "company aims", "company estimates", "company forecasts", "company hopes", "company plans", "company projects", "company targets", "still aim", "still aimed", "still aiming", "still aims", "still anticipated", "still anticipating", "still assumed", "still assuming", "still believed", "still believing", "still committed", "still committing", "still estimate", "still estimated", "still estimates", "still estimating", "still expected", "still expecting", "still forecast", "still forecasted", "still forecasting", "still forecasts", "still foreseeing", "still foreseen", "still hope", "still hoped", "still hopes", "still hoping", "still intended", "still intending", "still plan", "still planed", "still planning", "still plans", "still project", "still projected", "still projecting", "still projects", "still seeking", "still sought", "still target", "still targeted", "still targeting", "still targets", "still willing", "also aim", "also aims", "also estimate", "also estimates", "also forecast", "also forecasts", "also hope", "also hopes", "also plan", "also plans", "also project", "also projects", "also target", "also targets", "does not aim", "does not estimate", "does not forecast", "does not hope", "does not plan", "does not project", "does not target", "firm aims", "firm estimates", "firm forecasts", "firm hopes", "firm plans", "firm projects", "firm targets", "and aim", "and aims", "and estimate", "and estimates", "and forecast", "and forecasts", "and hope", "and hopes", "and plan", "and plans", "and project", "and projects", "and target", "and targets", "are aimed", "are aiming", "are anticipated", "are anticipating", "are assumed", "are assuming", "are believed", "are believing", "are committed", "are committing", "are estimated", "are estimating", "are expected", "are expecting", "are forecasted", "are forecasting", "are foreseeing", "are foreseen", "are hoped", "are hoping", "are intended", "are intending", "are planed", "are planning", "are projected", "are projecting", "are seeking", "are sought", "are targeted", "are targeting", "are willing", "but aim", "but aims", "but estimate", "but estimates", "but forecast", "but forecasts", "but hope", "but hopes", "but plan", "but plans", "but project", "but projects", "but target", "but targets", "not aimed", "not aiming", "not anticipated", "not anticipating", "not assumed", "not assuming", "not believed", "not believing", "not committed", "not committing", "not estimated", "not estimating", "not expected", "not expecting", "not forecasted", "not forecasting", "not foreseeing", "not foreseen", "not hoped", "not hoping", "not intended", "not intending", "not planed", "not planning", "not projected", "not projecting", "not seeking", "not sought", "not targeted", "not targeting", "not willing", "now aim", "now aimed", "now aiming", "now aims", "now anticipated", "now anticipating", "now assumed", "now assuming", "now believed", "now believing", "now committed", "now committing", "now estimate", "now estimated", "now estimates", "now estimating", "now expected", "now expecting", "now forecast", "now forecasted", "now forecasting", "now forecasts", "now foreseeing", "now foreseen", "now hope", "now hoped", "now hopes", "now hoping", "now intended", "now intending", "now plan", "now planed", "now planning", "now plans", "now project", "now projected", "now projecting", "now projects", "now seeking", "now sought", "now target", "now targeted", "now targeting", "now

```
targets", "now willing", "do not aim", "do not estimate", "do not forecast", "do not hope", "do not
plan", "do not project", "do not target", "is aimed", "is aiming", "is anticipated", "is anticipating",
"is assumed", "is assuming", "is believed", "is believing", "is committed", "is committing", "is
estimated", "is estimating", "is expected", "is expecting", "is forecasted", "is forecasting", "is
foreseeing", "is foreseen", "is hoped", "is hoping", "is intended", "is intending", "is planed", "is
planning", "is projected", "is projecting", "is seeking", "is sought", "is targeted", "is targeting",
"is willing", "we aim", "we estimate", "we forecast", "we hope", "we plan", "we project", "we
target");


#####################################################################################
#####################################################################################
#####           Create the CSV file
#####################################################################################
#####################################################################################

#create the csv file you will write output to using Text::CSV
my @rows;
my $csv = Text::CSV->new ( { binary => 1, eol => "\n", auto_diag => 1 } ) #choose some
recommended Text::CSV settings
            or die "Cannot use CSV: ".Text::CSV->error_diag ();
open(my $mycsv, ">", $outputfile) or die "$outputfile: $!";#open csv file you will write to

#name the columns of the csv you will generate
my
@columns=('id','fog','num_words','positive_words','negative_words','strongmodal_words','middl
emodal_words','weakmodal_words','financial_words','future_count');

$csv->print ($mycsv, [@columns]);


#####################################################################################
#       identify files to open;
#####################################################################################
open(my $fh, '<:encoding(UTF-8)', 'C:\Users\tad99\Dropbox\1-DK\programs\okids.csv') or die
"Could not open file $!";


#####################################################################################
#       Separate array into parts;
#####################################################################################
my $txt = ".txt";

while (my $id = <$fh>) {
        chomp $id;
        my $file = $id.$txt;
```

```
        next unless (-e $direct.$slash.$file);
        my $firstfile = read_file($direct.$slash.$file);


####################################################################################
#       Only forms with more than 100 characters;
####################################################################################
        my $firstfile1 = $firstfile;
        my $firstfile2 = lc($firstfile);
        my $firstfile3 = lc($firstfile);

        $firstfile1 =~ s/[[:punct:]]//g;  #removes punctuation
        $firstfile1 =~ s/\s*//g;              #removes spaces

        my $characters = length($firstfile1);
        if($characters < 100){next};             # skip files with less than 100 characters.


####################################################################################
####################################################################################
####################################################################################
####################################################################################
####################################################################################
####################################################################################

        my $id=$file;
        $id =~ s/\.txt//ig;
        my @array = split('\.',$firstfile2);
        my @array2 = split('\.',$firstfile2);
        my $num_sentences1=0;

    my $sentences=get_sentences($firstfile2);   ## Get the sentences.
    foreach my $sentence (@$sentences) {
                $num_sentences1 ++;                           #count number of real sentences
        }


####################################################################################
#       Calculate text characteristics
####################################################################################

        my $text = new Lingua::EN::Fathom;
        my $accumulate = 0;
        $text->analyse_block($firstfile2,$accumulate);
```

```
	my $num_words		= $text->num_words;#only includes real word, not numbers
	my $percent_complex_words = $text->percent_complex_words;
	my %text_hash = $text->unique_words;

	my $fog = 0.4*(($num_words/$num_sentences1)+$percent_complex_words);

	my $positive_words = 0;
	my $negative_words = 0;
	my $strongmodal_words = 0;
	my $middlemodal_words = 0;
	my $weakmodal_words = 0;
	my $financial_words = 0;
	my $future_count = 0;


#get word counts;
	foreach my $word (@positive){if (exists($text_hash{$word})){$positive_words = 
$positive_words + $text_hash{$word};}}
	foreach my $word (@negative){if (exists($text_hash{$word})){$negative_words = 
$negative_words + $text_hash{$word};}}

	foreach my $word (@strongmodal){if (exists($text_hash{$word})){$strongmodal_words 
= $strongmodal_words + $text_hash{$word};}}
	foreach my $word (@middlemodal){if 
(exists($text_hash{$word})){$middlemodal_words = $middlemodal_words + 
$text_hash{$word};}}
	foreach my $word (@weakmodal){if (exists($text_hash{$word})){$weakmodal_words = 
$weakmodal_words + $text_hash{$word};}}

	foreach my $word (@financial){ my $word1 = lc($word); my $count = () = $firstfile2 =~ 
/$word1/g; $financial_words = $financial_words + $count;}

	$ngrams = Lingua::EN::Bigram->new;
	$ngrams->text($firstfile2);
	$word_count = $ngrams->word_count;
	$count = $ngrams->bigram_count;


	foreach my $word(@wordarray){
		if (defined $$word_count{$word}) {
			$future_count = $future_count + $$word_count{$word};
		}
	}
	foreach my $word2(@bigramarray){
		if (defined $$count{$word2}) {
			$future_count = $future_count + $$count{$word2};
```

```
		}
	}

###########################################################################
#	Save to CSV
###########################################################################

my @my_array= ($id, $fog, $num_words, $positive_words, $negative_words,
$strongmodal_words, $middlemodal_words,
$weakmodal_words,$financial_words,$future_count);

$csv->print ($mycsv, [@my_array]);#or die?

}

close $mycsv or die "$!";	#close csv file;
```

STATA Code

Firm-quarter-level observations (Table 2)

```
gen ln_anon_count=log(anon_count+1)
gen ln_non_anon_count=log(non_anon_count+1)
gen ln_mve=log(mve)

reghdfe ln_anon_count ln_non_anon_count num_rp_dj ln_mve ana_follow, absorb(fyearq
firmid) vce(cluster firmid) keepsingletons
outreg2 using results.xls, replace excel ctitle(anon_count) label bdec(3) tdec(2) rdec(3) adjr2
tstat addtext (Year FE, yes, Firm FE)

reghdfe ln_anon_count ln_non_anon_count num_rp_dj ln_mve ana_follow roa car,
absorb(fyearq firmid) vce(cluster firmid) keepsingletons
outreg2 using results.xls, append excel ctitle(anon_count) label bdec(3) tdec(2) rdec(3) adjr2
tstat addtext (Year FE, yes, Firm FE)

reghdfe ln_anon_count ln_non_anon_count num_rp_dj ln_mve ana_follow roa car retail
insider_own short, absorb(fyearq firmid) vce(cluster firmid) keepsingletons
outreg2 using results.xls, append excel ctitle(anon_count) label bdec(3) tdec(2) rdec(3) adjr2
tstat addtext (Year FE, yes, Firm FE)

reghdfe ln_anon_count ln_non_anon_count num_rp_dj ln_mve ana_follow roa car retail
insider_own short turn_mean spread2_mean ret_vol mtb xrdq, absorb(fyearq firmid) vce(cluster
firmid) keepsingletons
outreg2 using results.xls, append excel ctitle(anon_count) label bdec(3) tdec(2) rdec(3) adjr2
tstat addtext (Year FE, yes, Firm FE)
```

Article-level observations (Table 3-7)

```
egen authorid = group(author1)
egen new_ticker = group(ticker)
gen ln_mve=log(mve)


*Table 3*

reghdfe abs_car anonymous recent_report recent_filing num_rp_dj ln_mve ana_follow roa p_car
retail insider_own short mtb p_turn_mean p_spread2_mean p_ret_vol xrdq d_reputation,
a(constant) vce(cluster new_ticker) keepsingleton
outreg2 using results.xls, replace excel ctitle(abs_car) label bdec(3) tdec(2) rdec(3) adjr2 tstat
addtext (Year FE, no, Firm FE, no )
```

reghdfe abs_car anonymous recent_report recent_filing num_rp_dj ln_mve ana_follow roa p_car retail insider_own short mtb p_turn_mean p_spread2_mean p_ret_vol xrdq d_reputation, a(year) vce(cluster new_ticker) keepsingleton
outreg2 using results.xls, append excel ctitle(abs_car) label bdec(3) tdec(2) rdec(3) adjr2 tstat addtext (Year FE, yes, Firm FE, no)

reghdfe abs_car anonymous recent_report recent_filing num_rp_dj ln_mve ana_follow roa p_car retail insider_own short mtb p_turn_mean p_spread2_mean p_ret_vol xrdq d_reputation, a(new_ticker year) vce(cluster new_ticker) keepsingleton
outreg2 using results.xls, append excel ctitle(abs_car) label bdec(3) tdec(2) rdec(3) adjr2 tstat addtext (Year FE, yes, Firm FE, yes)

Table 4

reghdfe abs_car c.anonymous##c.post recent_report recent_filing num_rp_dj ln_mve ana_follow roa p_car retail insider_own short mtb p_turn_mean p_spread2_mean p_ret_vol xrdq d_reputation, a(constant) vce(cluster new_ticker) keepsingleton
outreg2 using results.xls, replace excel ctitle(abs_car) label bdec(3) tdec(2) rdec(3) adjr2 tstat addtext (Year FE, no, Firm FE, no, Author FE, no)

reghdfe abs_car c.anonymous##c.post recent_report recent_filing num_rp_dj ln_mve ana_follow roa p_car retail insider_own short mtb p_turn_mean p_spread2_mean p_ret_vol xrdq d_reputation, a(new_ticker) vce(cluster new_ticker) keepsingleton
outreg2 using results.xls, append excel ctitle(abs_car) label bdec(3) tdec(2) rdec(3) adjr2 tstat addtext (Year FE, no, Firm FE, yes)

reghdfe abs_car c.anonymous##c.post recent_report recent_filing num_rp_dj ln_mve ana_follow roa p_car retail insider_own short mtb p_turn_mean p_spread2_mean p_ret_vol xrdq d_reputation, a(new_ticker authorid) vce(cluster new_ticker) keepsingleton
outreg2 using results.xls, append excel ctitle(abs_car) label bdec(3) tdec(2) rdec(3) adjr2 tstat addtext (Year FE, no, Firm FE, yes, Author FE, yes)

Table 5

reghdfe abs_car c.anonymous##c.d_reputation recent_report recent_filing num_rp_dj ln_mve ana_follow roa p_car retail insider_own short mtb p_turn_mean p_spread2_mean p_ret_vol xrdq, a(new_ticker year) vce(cluster new_ticker) keepsingleton
outreg2 using results.xls, append excel ctitle(abs_car) label bdec(3) tdec(2) rdec(3) adjr2 tstat addtext (Year FE, yes, Firm FE, yes)

reghdfe abs_car c.anonymous##c.d_m_acc15 recent_report recent_filing num_rp_dj ln_mve ana_follow roa p_car retail insider_own short mtb p_turn_mean p_spread2_mean p_ret_vol xrdq, a(new_ticker year) vce(cluster new_ticker) keepsingleton outreg2 using results.xls, append excel ctitle(abs_car) label bdec(3) tdec(2) rdec(3) adjr2 tstat addtext (Year FE, yes, Firm FE, yes)

reghdfe abs_car c.anonymous##c.d_histcomments recent_report recent_filing num_rp_dj ln_mve ana_follow roa p_car retail insider_own short mtb p_turn_mean p_spread2_mean p_ret_vol xrdq, a(new_ticker year) vce(cluster new_ticker) keepsingleton
outreg2 using results.xls, append excel ctitle(abs_car) label bdec(3) tdec(2) rdec(3) adjr2 tstat addtext (Year FE, yes, Firm FE, yes)

reghdfe abs_car c.anonymous##c.d_commentlikes recent_report recent_filing num_rp_dj ln_mve ana_follow roa p_car retail insider_own short mtb p_turn_mean p_spread2_mean p_ret_vol xrdq, a(new_ticker year) vce(cluster new_ticker) keepsingleton
outreg2 using results.xls, append excel ctitle(abs_car) label bdec(3) tdec(2) rdec(3) adjr2 tstat addtext (Year FE, yes, Firm FE, yes)

Table 6 Panel A

reghdfe ln_words anonymous ln_mve roa mtb, absorb(new_ticker year) vce(cluster authorid) keepsingletons
outreg2 using results.xls, replace excel ctitle(Length) label bdec(3) tdec(2) rdec(3) tstat addtext (Year FE, yes, Firm FE, yes)

reghdfe fog anonymous ln_mve roa mtb, absorb(new_ticker year) vce(cluster authorid) keepsingletons
outreg2 using results.xls, append excel ctitle(Complexity) label bdec(3) tdec(2) rdec(3) tstat addtext (Year FE, yes, Firm FE, yes)

reghdfe modal_rate anonymous ln_mve roa mtb, absorb(new_ticker year) vce(cluster authorid) keepsingletons
outreg2 using results.xls, append excel ctitle(Modal) label bdec(3) tdec(2) rdec(3) tstat addtext (Year FE, yes, Firm FE, yes)

reghdfe forward_rate anonymous ln_mve roa mtb, absorb(new_ticker year) vce(cluster authorid) keepsingletons
outreg2 using results.xls, append excel ctitle(Forward-looking) label bdec(3) tdec(2) rdec(3) tstat addtext (Year FE, yes, Firm FE, yes)

reghdfe financial_rate anonymous ln_mve roa mtb, absorb(new_ticker year) vce(cluster authorid) keepsingletons
outreg2 using results.xls, append excel ctitle(Finance) label bdec(3) tdec(2) rdec(3) tstat addtext (Year FE, yes, Firm FE, yes)

reghdfe tone anonymous ln_mve roa mtb, absorb(new_ticker year) vce(cluster authorid) keepsingletons outreg2 using results.xls, append excel ctitle(Tone) label bdec(3) tdec(2) rdec(3) tstat addtext (Year FE, yes, Firm FE, yes)

Table 6 Panel B. (w/ text attributes)

reghdfe abs_car anonymous recent_report recent_filing num_rp_dj ln_mve ana_follow roa p_car retail insider_own short mtb p_turn_mean p_spread2_mean p_ret_vol xrdq d_reputation forward_rate ln_words modal_rate fog financial_rate tone, a(constant) vce(cluster new_ticker) keepsingleton
outreg2 using results.xls, replace excel ctitle(abs_car) label bdec(3) tdec(2) rdec(3) adjr2 tstat addtext (Year FE, no, Firm FE, no)

reghdfe abs_car anonymous recent_report recent_filing num_rp_dj ln_mve ana_follow roa p_car retail insider_own short mtb p_turn_mean p_spread2_mean p_ret_vol xrdq d_reputation forward_rate ln_words modal_rate fog financial_rate tone, a(year) vce(cluster new_ticker) keepsingleton
outreg2 using results.xls, append excel ctitle(abs_car) label bdec(3) tdec(2) rdec(3) adjr2 tstat addtext (Year FE, yes, Firm FE, no)

reghdfe abs_car anonymous recent_report recent_filing num_rp_dj ln_mve ana_follow roa p_car retail insider_own short mtb p_turn_mean p_spread2_mean p_ret_vol xrdq d_reputation forward_rate ln_words modal_rate fog financial_rate tone, a(new_ticker year) vce(cluster new_ticker) keepsingleton
outreg2 using results.xls, append excel ctitle(abs_car) label bdec(3) tdec(2) rdec(3) adjr2 tstat addtext (Year FE, yes, Firm FE, no)

Table 7 Panel A (w/o co-occurring information events)

preserve
drop if recent_report==1
drop if recent_filing==1
drop if num_rp_dj>0

reghdfe abs_car anonymous ln_mve ana_follow roa p_car retail insider_own short mtb p_turn_mean p_spread2_mean p_ret_vol xrdq d_reputation, a(constant) vce(cluster new_ticker) keepsingleton
outreg2 using results.xls, replace excel ctitle(abs_car) label bdec(3) tdec(2) rdec(3) adjr2 tstat addtext (Year FE, no, Firm FE, no)

reghdfe abs_car anonymous ln_mve ana_follow roa p_car retail insider_own short mtb p_turn_mean p_spread2_mean p_ret_vol xrdq d_reputation, a(year) vce(cluster new_ticker) keepsingleton outreg2 using results.xls, append excel ctitle(abs_car) label bdec(3) tdec(2) rdec(3) adjr2 tstat addtext (Year FE, yes, Firm FE, no)

reghdfe abs_car anonymous ln_mve ana_follow roa p_car retail insider_own short mtb p_turn_mean p_spread2_mean p_ret_vol xrdq d_reputation, a(new_ticker year) vce(cluster new_ticker) keepsingleton
outreg2 using results.xls, append excel ctitle(abs_car) label bdec(3) tdec(2) rdec(3) adjr2 tstat addtext (Year FE, yes, Firm FE, yes)
restore

Table 7 Panel B (w/Entropy Balancing)

ebalance anonymous num_rp_dj ln_mve ana_follow roa p_car retail insider_own short p_turn_mean p_spread2_mean p_ret_vol mtb xrdq, target(2) gen(weight)
reghdfe abs_car anonymous recent_report recent_filing num_rp_dj ln_mve ana_follow roa p_car retail insider_own short mtb p_turn_mean p_spread2_mean p_ret_vol xrdq d_reputation [aweight=weight], a(constant) vce(cluster new_ticker) keepsingleton
outreg2 using results.xls, replace excel ctitle(abs_car) label bdec(3) tdec(2) rdec(3) adjr2 tstat addtext (Year FE, no, Firm FE, no)

reghdfe abs_car anonymous recent_report recent_filing num_rp_dj ln_mve ana_follow roa p_car retail insider_own short mtb p_turn_mean p_spread2_mean p_ret_vol xrdq d_reputation [aweight=weight], a(year) vce(cluster new_ticker) keepsingleton
outreg2 using results.xls, append excel ctitle(abs_car) label bdec(3) tdec(2) rdec(3) adjr2 tstat addtext (Year FE, yes, Firm FE, no)

reghdfe abs_car anonymous recent_report recent_filing num_rp_dj ln_mve ana_follow roa p_car retail insider_own short mtb p_turn_mean p_spread2_mean p_ret_vol xrdq d_reputation [aweight=weight], a(new_ticker year) vce(cluster new_ticker) keepsingleton
outreg2 using results.xls, append excel ctitle(abs_car) label bdec(3) tdec(2) rdec(3) adjr2 tstat addtext (Year FE, yes, Firm FE, yes)