%let wrds = wrds.wharton.upenn.edu 4016;

options comamid=TCP remote=wrds;

signon user='rcurtis' password='XXXXX';

\* Remember to change your LIBNAME and file directory to your personal settings;

libname DGM 'C:\PHD\Teaching\2015 SAS Camp\Replications\DGM';

options errors=**3** noovp; options nocenter ps=max ls=**120**; options mprint source nodate symbolgen macrogen;

options msglevel=i;

\* Remember to change your MACROS file directory to your personal settings;

%include "C:\PHD\SAS Camp\macros.sas";

ods html close;

ods preferences;

ods html newfile=proc;

\*Pull down IBES variables and calculate\*;

rsubmit;

**data** ibes; set ibes.statsumu\_epsus;

if year(fpedats) ge **1986** and year(fpedats) le **1999**;

if fpi = '1';

if cusip = '' then delete;

keep cusip ticker oftic cname statpers fiscalp fpi numest meanest fpedats;

**run**;

**proc** **download** data=ibes;

**run**;

endrsubmit;

\*pull down actual reported eps values from unadjusted file\*;

rsubmit;

**data** ibes1; set ibes.actu\_epsus;

if pdicity = 'ANN';

if cusip = '' then delete;

if value=**.** then delete;

if year(pends) ge **1986** and year(pends) le **1999**;

keep cusip ticker oftic cname pdicity pends value anndats;

**run**;

**proc** **download** data=ibes1;

**run**;

endrsubmit;

**data** DGM.ibes\_dgm; set ibes; **run**; **quit**;

**data** DGM.ibes1\_dgm; set ibes1; **run**; **quit**;

\*alternate way of retrieving data;

rsubmit;

**proc** **sql**;

create table test as select distinct

a.cusip, a.ticker, a.ofitc, a.fpedats, b.value

from ibes.statsumu\_epsus as a, ibes.actu\_epsus as b

where a.oftic=b.oftic and a.fpedats=b.pends and a.fpi='1' and b.periodicty='ANN'

and year(a.fpedats)>=**1986** and year(a.fpedats)<=**1999** and a.cusip ne '' and b.cusip ne ''; **run**; **quit**;

**proc** **download** data=test; **run**; **quit**;

endrsubmit;

\*merge actual values with consensus analyst forecasts.

Reduces count to 672,922 [7/31/15]\*;

**proc** **sql**;

create table ibes2 as

select a.\*, b.value, b.anndats

from DGM.ibes\_dgm as a, DGM.ibes1\_dgm as b

where a.cusip = b.cusip and a.fpedats = b.pends and a.cusip ne '';

**quit**;

\*deleting forecasts after announcement reduces count to 671,270 [7/31/15];

**data** ibes2b; set ibes2;

miss\_forecast = meanest - value; \*calculate miss\_forecast variable\*;

miss\_forecast = round(miss\_forecast,**.01**); \*round miss forecast to the nearest cent\*;

difference = statpers - anndats; \*create variable to see how long from time analysts issued forecasts and earnings were announced\*;

if difference > **0** then delete; \*remove forecasts that happened after earnings were announced\*;

**run**;

\*Only keep the last forecast given before the earnings announcement date.

Reduces count to 64,484 [7/31/15]\*;

**proc** **sql**;

create table ibes3 as select distinct \*

from ibes2b

group by oftic, fpedats

having statpers = max(statpers);

**quit**;

\*Bring down compustat quarterly data to calculate variables\*;

rsubmit;

**data** compq;

set comp.fundq;

if indfmt='INDL';

if datafmt='STD';

if popsrc='D';

if consol='C';

if gvkey = **.** then delete;

if cusip = '' then delete;

if fyearq ge **1985** and fyearq le **1999**;

keep gvkey datadate fyearq fqtr cusip tic conm txtq piq;

**run**;

**proc** **download** data=compq;

**run**;

endrsubmit;

**data** DGM.compq\_gm; set compq; **run**; **quit**;

\*\*\*Remove all qtr 4 observations from dataset;

\*reduces observations from 611,683 to 463,182 [7/31/15];

**data** q3;

set DGM.compq\_gm;

if fqtr = **4** then delete;

**run**; **quit**;

\*\*\*sum up taxes and income over the first three quarters to develop an ETR for Q1-3;

**proc** **sql**;

create table EtrQ3 as select

a.\*, sum(a.txtq) as txtQ3, sum(a.piq) as piQ3, sum(a.fqtr) as accQ3

from q3 as a

group by gvkey, fyearq

order by gvkey, fyearq;

**quit**;

\*\*\*delete quarters 1 and 2 as well as observations that do not have data for all three quarters, create ETRQ3.

Reduces sample to 152,467 [7/31/15];

**data** EtrQ3a;

set EtrQ3;

if fqtr = **1** then delete;

if fqtr = **2** then delete;

if accQ3 ne **6** then delete;

EtrQ3 = txtQ3 / piQ3;

**run**; **quit**;

\*\*\*Calculate ETR for Q4\*\*\*\*;

**proc** **sql**;

create table EtrQ4 as select

a.\*, sum(a.txtq) as txtQ4, sum(a.piq) as piQ4, sum(a.fqtr) as accQ4

from DGM.compq\_gm as a

group by gvkey, fyearq

order by gvkey, fyearq;

**quit**;

\*retaining Q4 reduces sample to 148,383 [7/31/15];

**data** EtrQ4a;

set EtrQ4;

if fqtr = **1** then delete;

if fqtr = **2** then delete;

if fqtr = **3** then delete;

if accQ4 ne **10** then delete;

EtrQ4 = txtQ4 / piQ4;

**run**; **quit**;

\*merge q3 ETR into Q4 Etr. Result is 148,383 observations [7/31/15];

**proc** **sql**;

create table Etr as select

a.\*, b.EtrQ3, b.piQ3

from EtrQ4a as a, EtrQ3a as b

where (a.gvkey = b.gvkey) and (a.fyearq = b.fyearq);

**quit**;

\*Create dependent variable Etr4\_Etr3 which meaures the difference between

the final year end ETR and the ETR after the first three quarters. Reduces count to 129,764 [7/31/15];

**data** Etr2;

set Etr;

Etr4\_Etr3 = EtrQ4 - EtrQ3;

if Etr4\_Etr3 = **.** then delete;

**run**; **quit**;

\*Bring down compustat annual data to calculate variables.

Retrieves 161,703 observations [7/31/15]\*;

rsubmit;

**data** comp;

set comp.funda;

if indfmt='INDL';

if datafmt='STD';

if popsrc='D';

if consol='C';

if gvkey = **.** then delete;

if cusip = '' then delete;

if fyr le **5** then year = (fyear+**1**);

if fyr ge **6** then year = fyear;

if year ge **1985** and year le **1999**;

keep gvkey datadate fyear fyr year tick cusip conm PI CSHPRI AT TXDI TXDFED TXDS TXDFO TXP TXR ACT CHE LCT DLC DP;

**run**;

**proc** **download** data=comp;

**run**;

endrsubmit;

**data** DGM.comp\_dgm; set comp; **run**; **quit**;

\*reduces count to 140,473 [7/31/15];

**data** comp1;

set DGM.comp\_dgm;

tax\_owed = (TXP - TXR) / PI; /\*DGM use tax return data but use this in sensitivity analyses as a measure that can be constructed from WRDS\*/

accruals = (IB - OANCF) / PI;

STR = **.35**;

if year = **1986** then STR = **.46**;

if year = **1987** then STR = **.40**;

if year ge **1988** and year le **1992** then STR = **.34**;

lag\_gvkey = lag(gvkey);

lag\_fyear = lag(fyear);

lag\_ACT = lag(act);

lag\_CHE = lag(CHE);

lag\_LCT = lag(lct);

lag\_DLC = lag(DLC);

lag\_TXP = lag(TXP);

if lag\_gvkey ne gvkey then delete;

if (lag\_fyear + **1**) ne fyear then delete;

CH\_CA = ACT - lag\_ACT;

CH\_Cash = CHE - lag\_che;

CH\_CL = LCT - lag\_LCT;

CH\_STD = DLC - lag\_DLC;

CH\_TP = TXP - lag\_TXP;

if TXDI = **.** then TXDI = TXDFED + TXDS + TXDFO; \*not sure if I should do this or not;

if accruals = **.** then accruals = (CH\_CA - CH\_Cash - CH\_CL + CH\_STD + CH\_TP - DP) / PI;

Deferred\_tax = TXDI / PI;

CUSIP8=SUBSTR(CUSIP,**1**,**8**); \*Create 8 digit CUSIP;

**run**; **quit**;

\*alternative way to retrieve code;

rsubmit;

**proc** **sql**;

create table sql\_way as select distinct

a.gvkey, a.datadate, a.fyear, a.fyr, a.tick, a.cusip, a.conm, a.PI, a.CSHPRI, a.AT, a.TXDI, a.TXDFED, a.TXDS, a.TXDFO, a.TXP, a.TXR, a.ACT, a.CHE, a.LCT, a.DLC, a.DP,

b.act as lag\_act, b.che as lag\_che, b.lct as lag\_lct, b.dlc as lag\_dlc, b.txp as lag\_txp

from comp.funda as a, comp.funda as b

where a.gvkey=b.gvkey and a.fyear=b.fyear+**1** and a.gvkey ne **.** and a.cusip ne '' and

and a.indfmt='INDL' and a.datafmt='STD' and a.popsrc='D' and a.consol='C'

and b.indfmt='INDL' and b.datafmt='STD' and b.popsrc='D' and b.consol='C'; **run**; **quit**;

**proc** **download** data=sql\_way;

endrsubmit;

\*Merge quarterly data of compustat in with annual compustat data.

Generates 117,469 observations [7/31/15]\*;

**proc** **sql**;

create table merge as select

a.\*, b.EtrQ3, b.Etr4\_Etr3, b.EtrQ4, b.piQ4, b.piQ3

from comp1 as a, etr2 as b

where (a.gvkey = b.gvkey) and (a.datadate = b.datadate);

**quit**;

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\* Above completed by Mark on Day 4 \*

\* Spencer will complete the rest on Day 5 \*