/\*Import data\*/

**proc** **import** datafile = "\\tsclient\rosehuang\Downloads\group project data.csv"

out= p1data dbms=csv replace;

getnames= yes;

**run**;

**proc** **transreg** data=p1data ss2 short;

model identity(crime)=pspline(nw/ nknots =**1** degree = **1**);

**run**;

/\*Create new variables

add police average\*/

**data** p1data;

set p1data;

police = po1-po2;

ur = u2/u1;

sop = so\*prob;

sowealth = so\*wealth;

sour = so\*ur;

sopop = so\*pop;

sopolice = so\*police;

soed = so\*ed;

wealthed = wealth\*ed;

probed = prob \*ed;

probwealth = prob\*wealth;

pola = (po1+po2)/**2**;

if nw >**10** then splinenw = (nw-**10**);

if nw <**10** then splinenw = **0**;

if lf >**0.575** then splinelf = (lf-**0.575**);

if lf< **0.575** then splinelf = **0**;

if time > **25** then splinetime = (time-**25**);

if time < **25** then splinetime = **0**;

lcrime = log(crime);

/\*Figure out a by eyeballing where the cutoff is\*/

**run**;

**proc** **contents** data=p1data;

**run**;

**proc** **print** data=p1data;**run**;

/\*Univariate analysis\*/

**proc** **univariate** data=p1data;

var m so ed po1 po2 lf mpf pop nw u1 u2 wealth ineq prob time crime;

histogram;

**run**;

/\*For table 1\*/

**proc** **freq** data=p1data;

table so;

**run**;

**proc** **means** data=p1data;

var m so ed po1 po2 lf mpf pop nw u1 u2 wealth ineq prob time crime;

**run**;

/\*Bivariate analysis\*/

**proc** **reg** data=p1data;

model crime = m;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = so;

**run**;

**quit**;

**proc** **glm** data=p1data;

class so;

model crime = so;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = ed;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = po1;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = po2;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = lf;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = mpf;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = pop;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = nw;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = u1;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = u2;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = wealth;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = ineq;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = prob;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = time;

**run**;

**quit**;

/\*Create new variable for difference in police protection from 1960 to 1959\*/

**data** p1data;

set p1data;

police = po1-po2;

uemprt = u2/u1;

**run**;

**proc** **reg** data=p1data;

model crime = police;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = uemprt;

**run**;

**quit**;

**proc** **univariate** data=p1data;

var ed wealth uemprt police;

histogram;

**run**;

**proc** **reg** data=p1data;

model crime = prob wealth ur pop police ed;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = prob so sop;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = prob wealth so sowealth;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = prob ur so sour;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = prob pop so sopop;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = prob police so sopolice;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = prob ed so soed;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = prob time;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = prob wealth ed wealthed;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = prob ed probed;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = prob wealth probwealth;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = prob ur pop police ed;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = prob ur police ed;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = prob ur police;

**run**;

**quit**;

**proc** **reg** data=p1data;

model crime = prob police ed;

**run**;

**quit**;

**proc** **corr** data=p1data;

var ed wealth;

**run**;

**proc** **corr** data=p1data;

var ur wealth;

**run**;

**proc** **corr** data=p1data;

var m mpf;

**run**;

**proc** **reg** plots=diagnostics(stats=(default aic));

model crime = time prob/clb;

**run**;

**quit**;

**proc** **reg** plots=diagnostics(stats=(default aic));

model crime = time prob m so ed lf police po1 po2 mpf pop nw u1 u2 ur wealth ineq/selection=stepwise include = **2** clb;

**run**;

**quit**;

**proc** **reg** plots=diagnostics(stats=(default aic));

model crime = time prob m so ed pola po1 po2 mpf pop nw splinenw lf splinelf splinetime u1 u2 ur wealth ineq/selection=stepwise include = **2** clb;

**run**;

**quit**;

**proc** **reg** plots=diagnostics(stats=(default aic));

model crime = time prob so m ed pola po1 po2 mpf pop nw splinenw lf splinelf splinetime u1 u2 ur wealth ineq/selection=stepwise include = **3**;

**run**;

**quit**;

**proc** **reg** plots=diagnostics(stats=(default aic));

model lcrime = time prob m so ed pola po1 po2 mpf pop nw splinenw lf splinelf splinetime u1 u2 ur wealth ineq/selection=stepwise include = **2** clb;

**run**;

**quit**;

**proc** **corr** data=p1data;

var wealth ineq;

**run**;

**proc** **reg** plots=diagnostics(stats=(default aic));

model crime = time prob m ed po1 mpf lf splinelf ur ineq/clb;

**run**;

**quit**;

**proc** **reg** plots=diagnostics(stats=(default aic));

model lcrime = time prob m ed po1 ur wealth ineq/ clb;

**run**;

**quit**;