```
/*
This do-file creates the data set and runs the analysis in "Discrimination
Begins in the Womb: Evidence of Sex-Selective Prenatal Investments"
using the Individual Recode section of the India NHFS.  The datasets

Leah K. Nelson
Last edited: August 2010
*/

#delimit;
clear;
set mem 800m;
set matsize 800;


/* ********1992******** */
/* Collect data from 1992 survey round */

/* Wealth data from the HH Recode */
#delimit;
use IAWI22FL.dta, clear;
split whhid, gen(hhid);
sort hhid1 hhid2 hhid3;
save India1992_wealth.dta, replace;

/* Individual recode */
#delimit;
use IAIR23FL.DTA, clear;

/* Merge in wealth data */
split caseid, gen(hhid);
sort hhid1 hhid2 hhid3;
merge hhid1 hhid2 hhid3 using India1992_wealth.dta;
drop if _merge==2;
drop _merge;
ren wlthindf wealth;
ren wlthind5 wealth_q;

/* Information on the first 2 births of each woman. */

/* Replace missing values with . */
#delimit;
foreach num of numlist 1 2 3 4 5 6 { ;
replace m2a_`num'=. if m2a_`num'==9;
replace m2b_`num'=. if m2b_`num'==9;
replace m2d_`num'=. if m2d_`num'==9;
replace m2e_`num'=. if m2e_`num'==9;
replace m2f_`num'=. if m2f_`num'==9;
replace m2g_`num'=. if m2g_`num'==9;
replace m2i_`num'=. if m2i_`num'==9;
replace m2j_`num'=. if m2j_`num'==9;
replace m2k_`num'=. if m2k_`num'==9;

/* Variable: whether or not received any prenatal care */
```

```
egen prenatcare`num' = rowmax(m2a_`num' m2b_`num' m2d_`num' m2e_`num' m2f_`num'
m2g_`num' m2i_`num' m2j_`num' m2k_`num');
} ;

/* Other info */
#delimit;
foreach num of numlist 1 2 3 4 5 6 {;
gen birthyear`num'=b2_0`num' + 1900;
gen teta`num'=1 if  m1_`num'>=1 & m1_`num' <8 & m1_`num'~=.;
replace teta`num'=0 if m1_`num'==0;
gen numteta`num'= m1_`num';
replace numteta`num'=. if m1_`num'==8;
gen prenatvisits`num'= m14_`num';
replace prenatvisits`num'=. if m14_`num'==98;
gen male`num'=1 if b4_0`num'==1;
replace male`num'=0 if b4_0`num'==2;
gen birthassistance`num' = 1 if m3n_`num'==0;
replace birthassistance`num' = 0 if m3n_`num'==1;
gen nonhomedelivery`num'=1 if m15_`num'>=21 & m15_`num'~=. & m15_`num'~=96;
replace nonhomedelivery`num'=0 if m15_`num'<=13 & m15_`num'~=. & m15_`num'~=96;
gen iron`num'=1 if s413_`num'==1 ;
replace iron`num'=0 if s413_`num'==0;
gen firstcheck`num'= m13_`num';
};

gen age_mom=v012;
gen educ = v106;
gen survey=1992;
ren v005 iweight;
gen urban=1 if v025==1;
replace urban=0 if v025==2;
ren v130 religion;
ren caseid case_id;
gen double caseid = 2+ v003*10+1000*v002+10000000*v001+10000000000*v024;
ren v024 state;

/* Mark obs from northern states */
gen north=1 if (state==19| state==7 | state==8 | state==20 | state==30 | state==24);

#delimit;
keep birthyear* teta* male* prenatcare* age_mom caseid case_id
survey urban state b4_* bord* educ north caseid firstch* religion
iweight birthassistance* prenatvisits* numteta* iron* nonhome*  wealth* ;

save India1992_new.dta, replace;

/* *******1998******* */
/* Collect data from 1998 survey round */

#delimit;
use IAWI41FL.dta, clear;
split whhid, gen(hhid);
sort hhid1 hhid2 hhid3;
save India1998_wealth.dta, replace;
```

```
/* Individual recode */
#delimit;
use IAIR42FL.dta, clear;

/* Merge in wealth data */
split caseid, gen(hhid);
sort hhid1 hhid2 hhid3;
merge hhid1 hhid2 hhid3 using India1998_wealth.dta;
drop if _merge==2;
drop _merge;
ren wlthindf wealth;
ren wlthind5 wealth_q;

/* Information on the first 2 births of each woman. */

#delimit;
foreach num of numlist 1 2{;
gen birthyear`num'=b2_0`num' + 1900;
gen teta`num'=1 if  m1_`num'>=1 & m1_`num' <8 & m1_`num'~=.;
replace teta`num'=0 if m1_`num'==0;
gen numteta`num'= m1_`num';
replace numteta`num'=. if m1_`num'==8;
gen prenatcare`num'=1 if  m2n_`num'==0;
replace prenatcare`num'=0 if m2n_`num'==1;
gen prenatvisits`num'= m14_`num';
replace prenatvisits`num'=. if m14_`num'==98;
gen male`num'=1 if b4_0`num'==1;
replace male`num'=0 if b4_0`num'==2;
gen nonhomedelivery`num'=1 if m15_`num'>=21 & m15_`num'~=. & m15_`num'~=96;
replace nonhomedelivery`num'=0 if m15_`num'<=13 & m15_`num'~=. & m15_`num'~=96;
gen iron`num'=1 if s415_`num'==1 ;
replace iron`num'=0 if s415_`num'==0;
gen ultra`num' = 1 if s411i_`num'==1;
replace ultra`num'= 0 if s411i_`num'==0;
gen firstcheck`num'= m13_`num';
gen blind`num' = 1 if s414a_`num'==1;
replace blind`num'= 0 if s414a_`num'==0;
gen blurr`num' = 1 if s414b_`num'==1;
replace blurr`num'= 0 if s414b_`num'==0;
gen conv`num' = 1 if s414c_`num'==1;
replace conv`num'= 0 if s414c_`num'==0;
gen swell`num' = 1 if s414d_`num'==1;
replace swell`num'= 0 if s414d_`num'==0;
gen fatigue`num' = 1 if s414e_`num'==1;
replace fatigue`num'= 0 if s414e_`num'==0;
gen anemia`num' = 1 if s414f_`num'==1;
replace anemia`num'= 0 if s414f_`num'==0;
gen bleed`num' = 1 if s414g_`num'==1;
replace bleed`num'= 0 if s414g_`num'==0;
/* Whether or not mother has had any pregnancy complications */
egen comp`num' = rowmax(blind`num' blurr`num' conv`num' swell`num' fatigue`num' anemia`num'
bleed`num');
replace comp`num'= 0 if blind`num'==0 & conv`num'==0 & swell`num'==0 & fatigue`num'==0 &
```

```
anemia`num'==0 & bleed`num'==0;
gen birthweight`num'=m19_`num' if m19_`num'<9995;
gen childage`num'=b8_0`num'
gen breastfed`num'=1 if m4_`num'<=35 | m4_`num'==95 | m4_`num'==96;
replace breastfed`num'=0 if m4_`num'==94;
replace breastfed`num'=. if m4_`num'==97|m4_`num'==98|m4_`num'==.;
gen immuniz`num'=(h10_`num'==1);
replace immuniz`num'=. if h10_`num'==8;
};

/* Other Information */
#delimit;
gen age_mom=v012;
gen educ = v106;
egen everabort1 = rowmax(s222b s222c);
gen everabort = (everabort1>0);
replace everabort = . if everabort1==.;
gen abort_induced = (s222c>0);
replace abort_induced = . if s222c==.;
egen livebirths = rownonmiss(b4*);
egen numpregs = rowtotal(livebirths s222a s222b s222c);
egen numabort = rowtotal(s222b s222c);
gen abortratio = numabort/numpregs;
gen indratio = s222c/numpregs;
replace indratio = . if indratio<0;
gen lang=slang;
gen caste=v131;
gen husbanded=v701;
gen agefirstmarriage=v511;
gen survey=1998;
ren v005 iweight;
gen readsnews = (v109==1);
replace readsnews = . if v109==.;
gen urban=1 if v025==1;
replace urban=0 if v025==2;
ren v130 religion;
ren v024 state;

ren caseid case_id;
gen double caseid = 2+ v003*10+1000*v002+10000000*v001+10000000000*state;

/* Mark obs from northern states */
gen north=1 if (state==19| state==7 | state==8 | state==20 | state==30 | state==24);

#delimit;
keep birthyear* teta* male* prenatcare* age_mom ultra* case_id childage*
survey urban state b4_* bord* educ north caseid firstch* religion
blind* blurr* conv* swell* fatigue* anemia* bleed* comp* iweight
prenatvisits* numteta* iron* nonhome* wealth*
birthweight* breastfed* immuniz* readsnews abort* everabort
numabort numpregs abortratio indratio lang caste husbanded agefirstmarriage;

save India1998.dta, replace;
```

```
/* *******2005-6******* */
/* Collect data from 2005-6 survey round */

/* Wealth data from the HH Recode */
#delimit;
use IAHR51FL.dta, clear;
keep hhid hv270 hv271;
ren hv270 wealth_q;
ren hv271 wealth;
split hhid;
sort hhid1 hhid2 hhid3;
save India2005wealth.dta, replace;

#delimit;
use IAIR51FL.dta, clear;

gen husbanded=v701;
drop d1* v2* v3* v4* v5* v6* v7* v8* ml* hw*;

/* Merge in wealth data */
#delimit;
split caseid, gen(hhid);
sort hhid1 hhid2 hhid3;
merge hhid1 hhid2 hhid3 using India2005wealth.dta;
drop if _merge==2;
drop _merge;

/* Information on the first 2 births of the woman. */
#delimit;
foreach num of numlist 1 2{;
    gen birthyear`num'=b2_0`num';
    gen teta`num'=1 if  m1_`num'>=1 & m1_`num' <8 & m1_`num'~=.;
    replace teta`num'=0 if m1_`num'==0;
    gen numteta`num'= m1_`num';
    replace numteta`num'=. if m1_`num'==8;
    gen prenatcare`num'=1 if  m2n_`num'==0;
    replace prenatcare`num'=0 if m2n_`num'==1;
    gen prenatvisits`num'= m14_`num';
    replace prenatvisits`num'=. if m14_`num'==98;
    gen male`num'=1 if b4_0`num'==1;
    replace male`num'=0 if b4_0`num'==2;
    gen nonhomedelivery`num'=1 if m15_`num'>=21 & m15_`num'~=. & m15_`num'~=96;
    replace nonhomedelivery`num'=0 if m15_`num'<=12 & m15_`num'~=. & m15_`num'~=96;
    gen iron`num'=1 if m45_`num'==1 ;
    replace iron`num'=0 if m45_`num'==0;
    gen daysiron`num'= m46_`num';
    replace daysiron`num'=. if m46_`num'==998;
    replace daysiron`num'=0 if m45_`num'==0;
    gen firstcheck`num'= m13_`num';
    replace firstcheck`num'= . if m13_`num'==98;
    gen blind`num' = 1 if m48_`num'==1;
    replace blind`num'= 0 if m48_`num'==0;
    gen blurr`num' = 1 if m47_`num'==1;
```

```
    replace blurr`num'= 0 if m47_`num'==0;
    gen conv`num' = 1 if s427_`num'==1;
    replace conv`num'= 0 if s427_`num'==0;
    gen swell`num' = 1 if s428_`num'==1;
    replace swell`num'= 0 if s428_`num'==0;
    gen fatigue`num' = 1 if s429_`num'==1;
    replace fatigue`num'= 0 if s429_`num'==0;
    gen bleed`num' = 1 if s430_`num'==1;
    replace bleed`num'= 0 if s430_`num'==0;
    egen comp`num' = rowmax(blind`num' blurr`num' conv`num' swell`num' fatigue`num' bleed`num');
    replace comp`num'= 0 if blind`num'==0 & conv`num'==0 & swell`num'==0 & fatigue`num'==0 &
    bleed`num'==0;
    gen birthweight`num'=m19_`num' if m19_`num'<9995;
    gen childage`num'=b8_0`num';
    gen breastfed`num'=1 if m4_`num'<=60 | m4_`num'==95 ;
    replace breastfed`num'=0 if m4_`num'==94;
    replace breastfed`num'=. if m4_`num'==98|m4_`num'==.;
    gen immuniz`num'=(h10_`num'==1);
    replace immuniz`num'=. if h10_`num'==8;
};

/* Ultrasound data */

/* First, collect only births and terminations from calendar.
Note that we only need to preserve order, not actual dates,
to match birth history and ultrasound data.  The calendar
is recorded as a long string so the following code extracts
only the information on births and terminations from that
string. */

#delimit;
gen test1= subinstr(vcal_1,"P","P ",.);
gen test2= subinstr(test1,"0","",.);
gen test3= subinstr(test2,"1","",.);
gen test4= subinstr(test3,"2","",.);
gen test5= subinstr(test4,"3","",.);
gen test6= subinstr(test5,"4","",.);
gen test7= subinstr(test6,"5","",.);
gen test8= subinstr(test7,"6","",.);
gen test9= subinstr(test8,"7","",.);
gen test10= subinstr(test9,"8","",.);
gen test11= subinstr(test10,"9","",.);
gen test12= subinstr(test11,"W","",.);
gen test13= subinstr(test12,"N","",.);
gen test14= subinstr(test13,"C","",.);
gen test15= subinstr(test14,"F","",.);
gen test16= subinstr(test15,"?","",.);
gen test17= subinstr(test16,"B","B ",.);
gen test18= subinstr(test17,"T","T ",.);
split test18, gen(discard);
gen discardthis=1 if discard1=="P";
gen test19a = subinstr(test18,"P","",.);
gen test19b = trim(test19a);
gen test19 = itrim(test19b);
```

```
split test19, gen(preghist);

#delimit;
foreach x in 1 2 3 4 5 6 { ;
gen pregstat`x' = 1 if preghist`x'=="T"|preghist`x'=="B";
} ;
/* Number of past pregnancies */
egen pastpregs=rownonmiss(pregstat*);

/* Next, collect ultrasound data (also entered as a long string).
For each pregnancy, a Y or N or ? is recorded. */

#delimit;
gen test20 = trim(vcal_6);
ren test20 test21;
gen test22 = itrim(test21);
gen test23 = subinstr(test22,"Y","Y ",.);
gen test24= subinstr(test23,"N","N ",.);
gen test25 = subinstr(test24,"?","? ",.);
split test25, gen(ult);

#delimit;
foreach x in 1 2 3 4 5 6 { ;
gen ultstat`x' = 1 if ult`x'~="";
} ;
egen pastults=rownonmiss(ultstat*);

/* Get rid of all ultrasounds not corresponding to termination or birth */
#delimit;
replace ult1 = "" if preghist1==";
*Get rid of all ultrasounds corresponding to current pregnancy;
replace ult1 = "" if discardthis==1 & pastpregs<pastults;

/*Next, only collect ultrasound info for births, not terminations (only
for most recent 2 births)*/

#delimit;
foreach x in 1 2 3 4 5 6 { ;
replace ult`x' = "" if preghist`x'=="T";
} ;
foreach x in 1 2 3 4 5 6 { ;
replace preghist`x' = "" if preghist`x'=="T";
} ;

gen birthhist = preghist1+preghist2+preghist3+preghist4+preghist5+preghist6;
gen ulthist= ult1+ult2+ult3+ult4+ult5+ult6;
gen test26 = subinstr(ulthist,"Y","Y ",.);
gen test27 = subinstr(test26,"N","N ",.);
gen test28 = subinstr(test27,"?","? ",.);
split test28, gen(ultr);

gen ultra1=1 if ultr1=="Y";
replace ultra1=0 if ultr1=="N";
gen ultra2=1 if ultr2=="Y";
```

```
replace ultra2=0 if ultr2=="N";

/* Other info */
#delimit;
gen age_mom=v012;
ren v130 religion;
gen survey=2005;
ren v005 iweight;
gen urban=1 if v025==1;
replace urban=0 if v025==2;
/* Note that the state codes are different from the previous rounds.
To fix this, I created a crosswalk between new and old codes by hand
in Excel. */
sort v024;
merge v024 using oldstatecode.dta;
drop _merge;
ren oldcode state;
ren v106 educ;
gen lang = slang;
replace lang = 8 if slang==9;
replace lang = 9 if slang==10;
replace lang = 10 if slang==11;
replace lang = 11 if slang==12;
replace lang = 12 if slang==13;
replace lang = 13 if slang==14;
replace lang = 14 if slang==8;
gen caste=s46;
gen agefirstmarriage=s310c;
gen readsnews = (v157==2| v157==3);
replace readsnews = . if v157==.;
gen watchestv = (v159==2| v159==3);
replace watchestv = . if v159==.;
ren caseid case_id;
gen double caseid = 3+v003*10+1000*v002+100000*v001+10000000000*state;

gen north=1 if (state==19| state==7 | state==8 | state==20 | state==30 | state==24);

#delimit;
keep birthyear* teta* male* prenatcare* age_mom educ survey b4_* bord* north caseid ultra*
firstch* blind* blurr* conv* swell* fatigue* bleed* comp*  wealth* childage* religion
birthweight* iweight state urban prenatvisits* numteta* iron* daysiron* nonhome* case_id
breastfed* immuniz* watchestv readsnews lang caste husbanded agefirstmarriage;

save India2005.dta, replace;


/* Append all survey years, create important variables */
#delimit;
use India2005.dta, clear;
append using India1992_new.dta;
append using India1998.dta;

/* Drop birthyears that are before 4 years previous to survey round and that
occur after the survey year. */
```

```
drop if birthyear1<1989 & survey==1992;
drop if birthyear1<1995 & survey==1998;
drop if birthyear1<2001 & survey==2004;
drop if birthyear1>1992& survey==1992;
drop if birthyear1>1999& survey==1998;
drop if birthyear1>2006& survey==2004;

/* Sex-ratio - note that this is sex ratio before last birth */
foreach var of varlist b4* { ;
    replace `var'=0 if `var'==2;
} ;
egen sexratio=rmean(b4_02-b4_20);

/* Gender of the firstborn */
foreach x of varlist b4_01-b4_20{;
tostring `x', gen(s`x');
};
gen test=  sb4_20+ sb4_19+ sb4_18+ sb4_17+ sb4_16+ sb4_15+ sb4_14+ sb4_13+ sb4_12+ sb4_11+
sb4_10+ sb4_09+ sb4_08+ sb4_07+ sb4_06+ sb4_05+ sb4_04+ sb4_03+ sb4_02+ sb4_01;
gen test2 = subinstr(test,"."," ",.);
gen test3=itrim(test2);
gen test4=subinstr(test3,"1","1 ",.);
gen test5=subinstr(test4,"0","0 ",.);
split test5, gen(firstchild);
gen firstmale=1 if firstchild1=="1"; /* Designates whether firstborn is male */
replace firstmale=0 if firstchild1=="0";

/* Outcomes of interest */
gen ante = 1 if prenatvisits1>1 & prenatvisits1~=.;
replace ante = 0 if prenatvisits1<2 & prenatvisits1~=.;

gen tets = 1 if numteta1>1 & numteta1~=.;
replace tets = 0 if numteta1<2 & numteta1~=.;

/* Timing of prenatal care */
#delimit;
gen earlycheck_1=1 if firstcheck1<=3 & prenatcare1==1;
replace earlycheck_1=0 if prenatcare1==0|(firstcheck1>=4 & prenatcare1==1);
gen earlycheck_2=1 if firstcheck1<=4 & prenatcare1==1;
replace earlycheck_2=0 if prenatcare1==0|(firstcheck1>=5 & prenatcare1==1);

gen latecheck_1=1 if firstcheck1>=4 & prenatcare1==1;
replace latecheck_1=0 if firstcheck1<=3 | prenatcare1==0;
gen latecheck_2=1 if firstcheck1>=5 & prenatcare1==1;
replace latecheck_2=0 if firstcheck1<=4 | prenatcare1==0;

#delimit;
gen latecheck_3 = 1 if firstcheck1>4 & prenatcare1==1;
replace latecheck_3 = 0 if firstcheck1<=4 & prenatcare1==1 & prenatvisits1<2;
replace latecheck_3 = 1 if firstcheck1<=4 & prenatcare1==1 & prenatvisits1>1;
replace latecheck_3=0 if prenatcare1==0;

#delimit;
```

```
gen latecheck_4 = 1 if firstcheck1>3 & prenatcare1==1;
replace latecheck_4 = 0 if firstcheck1<=3 & prenatcare1==1 & prenatvisits1<2;
replace latecheck_4 = 1 if firstcheck1<=3 & prenatcare1==1 & prenatvisits1>1;
replace latecheck_4=0 if prenatcare1==0;

/* Rainfall information */
sort state birthyear1;
merge state birthyear1 using rain.dta,
keep(rain rain_bad);
drop if _merge==2;
drop _merge;

save India_1992-2006.dta, replace;


/* Panel Data Set
This code simply reshapes the above data set so that the result is a panel
that has the information for the most recent 2 births of each woman.  */

#delimit;
use India_1992-2006.dta", clear;

#delimit;
egen sexratio2 = rowmean(b4_03-b4_20);

gen ante2 = 1 if prenatvisits2>1 & prenatvisits2~=.;
replace ante2 = 0 if prenatvisits2<2 & prenatvisits2~=.;

gen tets2 = 1 if numteta2>1 & numteta2~=.;
replace tets2 = 0 if numteta2<2 & numteta2~=.;

ren bord_01 bord1;
ren bord_02 bord2;
ren sexratio sexratio1;
ren ante ante1;
ren tets tets1;

/* Drop duplicate mother observations (while in wide format) */
sort caseid;
by caseid: gen dup = cond(_N==1,1,_n);
drop if dup>1;

/* Reshape to long format */
reshape long teta numteta iron prenatcare prenatvisits daysiron
nonhomedelivery  male birthyear bord sexratio ultra ante tets
blurr blind conv swell fatigue bleed comp anemia firstcheck childage,
i(caseid) j(num);

tsset caseid num;

/* Mark observations that have at least 2 observations for births */
#delimit;
sort caseid;
by caseid: egen ptet = count(teta);
```

```
by caseid: egen pntet = count(numteta);
by caseid: egen piro = count(iron);
by caseid: egen pdiro = count(daysiron);
by caseid: egen pcar = count(prenatcare);
by caseid: egen pvis = count(prenatvisits);
by caseid: egen phom = count(nonhomedeliv);

#delimit;
xi i.birthyear, prefix(B) noomit;
xi i.state, prefix(S) noomit;
xi i.survey, prefix(s) noomit;

save India_panel.dta, replace;




* ******************************************************************
/* REGRESSIONS */
* ******************************************************************


/* Figure 1: Gender differentials in prenatal care by birth parity */

#delimit;
use India_1992-2006.dta, clear;

#delimit;
xi: reg prenatcare1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban
[pw=iweight] if survey>=1998 & sexratio==1 & bord_01==2, robust;
xi: reg prenatcare1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban
[pw=iweight] if survey>=1998 & sexratio==0 & bord_01==2, robust;
xi: reg prenatcare1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban
[aw=iweight] if survey>=1998 & sexratio==1 & bord_01==3, robust;
xi: reg prenatcare1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban
[aw=iweight] if survey>=1998 & sexratio==0 & bord_01==3, robust;
xi: reg prenatcare1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban
[aw=iweight] if survey>=1998 & sexratio==1 & bord_01==4, robust;
xi: reg prenatcare1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban
[aw=iweight] if survey>=1998 & sexratio==0 & bord_01==4, robust;
xi: reg prenatcare1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban
[pw=iweight] if survey>=1998 & sexratio==1 & bord_01==5, robust;
xi: reg prenatcare1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban
[pw=iweight] if survey>=1998 & sexratio==0 & bord_01==5, robust;



/* Table 1 - Prenatal Investments and Gender of the Child */

#delimit;
use India_1992-2006.dta, clear;

#delimit;
```

```
/* First we need a variable that ensures that we have the same sample
across all specs */
xi: reg ante male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01 sexratio
[pw=iweight] if survey>=1998, robust;
predict wholesample if e(sample);

#delimit;
* most basic regression - no controls;
xi: reg ante male1  [pw=iweight] if wholesample~=., robust;
* adding in geographic and survey variables;
xi: reg ante male1 i.state i.survey urban [pw=iweight] if wholesample~=., robust;
* adding in child level variables;
xi: reg ante male1 i.state i.survey urban i.birthyear1 bord_01 sexratio [pw=iweight] if
wholesample~=., robust;
* adding in mother level variables;
xi: reg ante male1 i.state i.survey urban i.birthyear1 bord_01 sexratio age_mom educ [pw=iweight
] if wholesample~=., robust;
* adding in household level variables;
xi: reg ante male1 i.state i.survey urban i.birthyear1 bord_01 sexratio age_mom educ i.wealth_q
[pw=iweight] if wholesample~=., robust;
xi: reg ante male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01 sexratio
[pw=iweight] if survey>=1998 & north==1, robust;
xi: reg ante male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01 sexratio
[pw=iweight] if survey>=1998 & sexratio<0.5 & sexratio~=., robust;
xi: logit ante male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998 & prenatcare1==1 & ultra1==1, robust;
xi: logit ante male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998 & prenatcare1==1 & ultra1==0, robust;



/* Table 2 - Sex-Selective Prenatal Investments in India:
Births to Women with and without Ultrasounds (Logit specification) */

#delimit;
use "${dir}Data\RegData\India_1992-2006.dta", clear;

/* POOLED DATA (1998-9, 2005-6), with prenatal care restriction */

/* Full sample */
#delimit;
xi: logit ante male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998 & prenatcare1==1 & ultra1==1, robust;
xi: logit ante male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998 & prenatcare1==1 & ultra1==0, robust;
xi: logit teta1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998 & prenatcare1==1 & ultra1==1, robust;
xi: logit teta1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998 & prenatcare1==1 & ultra1==0, robust;

/* North sample */
#delimit;
xi: logit ante male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998 & prenatcare1==1 & north==1 & ultra1==1, robust;
```

```
xi: logit ante male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998 & prenatcare1==1 & north==1 & ultra1==0, robust;
xi: logit teta1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998 & prenatcare1==1 & north==1 & ultra1==1, robust;
xi: logit teta1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998 & prenatcare1==1 & north==1 & ultra1==0, robust;

/* Majority female sample */
#delimit;
xi: logit ante male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998 & prenatcare1==1 & sexratio<0.5 & ultra1==1, robust;
xi: logit ante male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998 & prenatcare1==1 & sexratio<0.5 & ultra1==0, robust;
xi: logit teta1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998 & prenatcare1==1 & sexratio<0.5 & ultra1==1, robust;
xi: logit teta1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998 & prenatcare1==1 & sexratio<0.5 & ultra1==0, robust;


/* ******2005-6 data only****** */

/* Full sample */
#delimit;
xi: logit ante male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey==2005 & ultra1==1, robust;
xi: logit ante male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey==2005 & ultra1==0, robust;
xi: logit teta1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey==2005 & ultra1==1, robust;
xi: logit teta1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey==2005 & ultra1==0, robust;

/* North sample */
#delimit;
xi: logit ante male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey==2005 & north==1 & ultra1==1, robust;
xi: logit ante male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey==2005 & north==1 & ultra1==0, robust;
xi: logit teta1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey==2005 & north==1 & ultra1==1, robust;
xi: logit teta1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey==2005 & north==1 & ultra1==0, robust;

/* Majority female sample */
#delimit;
xi: logit ante male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey==2005 & sexratio<0.5 & ultra1==1, robust;
xi: logit ante male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey==2005 & sexratio<0.5 & ultra1==0, robust;
xi: logit teta1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey==2005 & sexratio<0.5 & ultra1==1, robust;
xi: logit teta1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey==2005 & sexratio<0.5 & ultra1==0, robust;
```

```
/* Table 3 - Accounting for Stopping Rules and
Reverse Causality due to Sex Selective Abortions */

#delimit;
use "${dir}Data\RegData\India_1992-2006.dta", clear;

* Age Restrictions
#delimit;
xi: reg ante male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01 sexratio
[pw=iweight] if survey>=1998 & childage1<=1 & childage1~=. , robust;
xi: reg ante male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01 sexratio
[pw=iweight] if survey>=1998 & childage1<1 & childage1~=. , robust;
* Additional care
xi: reg ante male1 i.state i.survey i.birthyear1 age_mom educ urban bord_01 sexratio i.wealth_q
[pw=iweight] if survey>=1998 & prenatcare1==1, robust;
xi: reg ante male1 i.state i.survey i.birthyear1 age_mom educ urban bord_01 sexratio i.wealth_q
[pw=iweight] if survey>=1998 & latecheck_2==1, robust;
* FE
#delimit;
use "${dir}Data\RegData\India_panel.dta", clear;
xtreg ante male Bb* bord if survey==1998 & pcar==2 & twogend==0.5 & wealth_q~=. ,fe robust;


/* Table 4 - Accounting for Sex Selective Abortions */

#delimit;
use India_1992-2006.dta, clear;

#delimit;
xi: reg ante male1 i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio i.wealth_q abortratio [pw=iweight] if survey==1998, robust;
xi: reg ante male1 i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio i.wealth_q indratio [pw=iweight] if survey==1998, robust;
xi: reg ante male1 i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio i.wealth_q rain_bad rain [pw=iweight] if survey==1998, robust;
xi: reg ante male1 i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio i.wealth_q firstmale[pw=iweight] if survey==1998, robust;
xi: reg ante male1 i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio i.wealth_q abortratio indratio rain_bad rain firstmale [pw=iweight] if survey==1998,
robust;


/* Table 5 - Sex-Selective Prenatal Investments in India: Pre-Ultrasound Period
(births in 1992 and earlier) */

#delimit;
use India_1992-2006.dta, clear;

#delimit;
xi: reg ante male1 i.state  i.wealth_q i.survey i.birthyear1 age_mom educ urban bord_01 sexratio
 [pw=iweight] if birthyear1<1993, robust;
xi: reg prenatcare1 male1 i.state  i.wealth_q i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if birthyear1<1993, robust;
```

```
xi: reg prenatvisits1 male1 i.state  i.wealth_q i.survey i.birthyear1 age_mom educ urban bord_01
 sexratio [pw=iweight] if birthyear1<1993 & prenatvisits1<9, robust;
xi: reg teta1 male1 i.state  i.wealth_q i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if birthyear1<1993, robust;
xi: reg numteta1 male1 i.state  i.wealth_q i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if birthyear1<1993, robust;
xi: reg iron1 male1 i.state  i.wealth_q i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if birthyear1<1993, robust;
xi: reg nonhomedelivery1 male1 i.state  i.wealth_q i.survey i.birthyear1 age_mom educ urban
bord_01 sexratio [pw=iweight] if birthyear1<1993, robust;


/* Table 6 - Sex-Selective Prenatal Investments in India: Timing of Prenatal Care */

#delimit;
use India_1992-2006.dta, clear;

#delimit;
xi: reg earlycheck_2 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998, robust;
xi: reg latecheck_3 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998 , robust;


/* Table 7 - Gender and Pregnancy Complications in India */

#delimit;
use "${dir}Data\RegData\India_1992-2006.dta", clear;

#delimit;
xi: reg blind1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio if survey>=1998, robust;
xi: reg blurr1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio if survey>=1998, robust;
xi: reg conv1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01 sexratio
 if survey>=1998, robust;
xi: reg swell1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio if survey>=1998, robust;
xi: reg fatigue1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio if survey>=1998, robust;
xi: reg anemia1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio if survey>=1998, robust;
xi: reg bleed1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio if survey>=1998, robust;
xi: reg comp1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01 sexratio
 if survey>=1998, robust;


/* Table 8 - Sex-Selective Prenatal Care in Other Countries */

* Available upon request
```

```
/* Appendix Table 3 - Sex-Selective Prenatal Investments in India: Any Prenatal Care */

#delimit;
use India_1992-2006.dta, clear;

* first we need a variable that ensures that we have the same sample across all specs;
xi: reg prenatcare1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998, robust;
predict wholesample if e(sample);
* most basic regression - no controls;
xi: reg prenatcare1 male1  [pw=iweight] if wholesample~=., robust;
* adding in geographic and survey variables;
xi: reg prenatcare1 male1 i.state i.survey urban [pw=iweight] if wholesample~=., robust;
* adding in child level variables;
xi: reg prenatcare1 male1 i.state i.survey urban i.birthyear1 bord_01 sexratio [pw=iweight] if
wholesample~=., robust;
* adding in mother level variables;
xi: reg prenatcare1 male1 i.state i.survey urban i.birthyear1 bord_01 sexratio age_mom educ [pw=
iweight] if wholesample~=., robust;
* adding in household level variables;
xi: reg prenatcare1 male1 i.state i.survey urban i.birthyear1 bord_01 sexratio age_mom educ i.
wealth_q [pw=iweight] if wholesample~=., robust;
xi: reg prenatcare1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998 & north==1, robust;
xi: reg prenatcare1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998 & sexratio<0.5 & sexratio~=., robust;
xi: reg prenatcare1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998 & childage1<=1 & childage1~=. , robust;
xi: reg prenatcare1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01
sexratio [pw=iweight] if survey>=1998 & childage1<1 & childage1~=. , robust;

* FE
use India_panel.dta, clear;

#delimit;
xtreg prenatcare male Bb* bord if survey==1998 & pcar==2 & twogend==0.5 & wealth_q~=. ,fe robust;



/* Appendix Table 5 - Sex-Selective Prenatal Investments in India: Tetanus shot during
pregnancy */

#delimit;
use India_1992-2006.dta, clear;

#delimit;
* first we need a variable that ensures that we have the same sample across all specs;
xi: reg teta1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01 sexratio
 [pw=iweight] if survey>=1998, robust;
predict wholesample if e(sample);

#delimit;
* most basic regression - no controls;
```

```
xi: reg teta1 male1  [pw=iweight] if wholesample~=., robust;
* adding in geographic and survey variables;
xi: reg teta1 male1 i.state i.survey urban [pw=iweight] if wholesample~=., robust;
* adding in child level variables;
xi: reg teta1 male1 i.state i.survey urban i.birthyear1 bord_01 sexratio [pw=iweight] if
wholesample~=., robust;
* adding in mother level variables;
xi: reg teta1 male1 i.state i.survey urban i.birthyear1 bord_01 sexratio age_mom educ [pw=
iweight] if wholesample~=., robust;
* adding in household level variables;
xi: reg teta1 male1 i.state i.survey urban i.birthyear1 bord_01 sexratio age_mom educ i.wealth_q
 [pw=iweight] if wholesample~=., robust;
xi: reg teta1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01 sexratio
 [pw=iweight] if survey>=1998 & north==1, robust;
xi: reg teta1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01 sexratio
 [pw=iweight] if survey>=1998 & sexratio<0.5 & sexratio~=., robust;
xi: reg teta1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01 sexratio
 [pw=iweight] if survey>=1998 & childage1<=1 & childage1~=. , robust;
xi: reg teta1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban bord_01 sexratio
 [pw=iweight] if survey>=1998 & childage1<1 & childage1~=. , robust;

* FE
#delimit;
use India_panel.dta, clear;

#delimit;
xtreg teta male Bb* bord if survey==1998 & pcar==2 & twogend==0.5 & wealth_q~=. ,fe robust;


/* Appendix Table 6 - Sex-Selective Prenatal Investments in India: Non-Home Delivery */

#delimit;
use India_1992-2006.dta, clear;

#delimit;
*first we need a variable that ensures that we have the same sample across all specs;
xi: reg nonhomedelivery1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban
bord_01 sexratio [pw=iweight] if survey>=1998, robust;
predict wholesample if e(sample);

#delimit;
* most basic regression - no controls;
xi: reg nonhomedelivery1 male1  [pw=iweight] if wholesample~=., robust;
* adding in geographic and survey variables;
xi: reg nonhomedelivery1 male1 i.state i.survey urban [pw=iweight] if wholesample~=., robust;
* adding in child level variables;
xi: reg nonhomedelivery1 male1 i.state i.survey urban i.birthyear1 bord_01 sexratio [pw=iweight]
 if wholesample~=., robust;
* adding in mother level variables;
xi: reg nonhomedelivery1 male1 i.state i.survey urban i.birthyear1 bord_01 sexratio age_mom educ
 [pw=iweight] if wholesample~=., robust;
* adding in household level variables;
xi: reg nonhomedelivery1 male1 i.state i.survey urban i.birthyear1 bord_01 sexratio age_mom educ
 i.wealth_q [pw=iweight] if wholesample~=., robust;
```

```
xi: reg nonhomedelivery1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban
bord_01 sexratio [pw=iweight] if survey>=1998 & north==1, robust;
xi: reg nonhomedelivery1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban
bord_01 sexratio [pw=iweight] if survey>=1998 & sexratio<0.5 & sexratio~=., robust;
xi: reg nonhomedelivery1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban
bord_01 sexratio [pw=iweight] if survey>=1998 & prenatcare1==1 & ultra1==1, robust;
xi: reg nonhomedelivery1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban
bord_01 sexratio [pw=iweight] if survey>=1998 & prenatcare1==1 & ultra1==0, robust;
xi: reg nonhomedelivery1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban
bord_01 sexratio [pw=iweight] if survey>=1998 & childage1<=1 & childage1~=. , robust;
xi: reg nonhomedelivery1 male1 i.wealth_q i.state i.survey i.birthyear1 age_mom educ urban
bord_01 sexratio [pw=iweight] if survey>=1998 & childage1<1 & childage1~=. , robust;
xi: reg nonhomedelivery1 male1 i.state i.survey i.birthyear1 age_mom educ urban bord_01 sexratio
 i.wealth_q [pw=iweight] if survey>=1998 & prenatcare1==1, robust;
xi: reg nonhomedelivery1 male1 i.state i.survey i.birthyear1 age_mom educ urban bord_01 sexratio
 i.wealth_q [pw=iweight] if survey>=1998 & latecheck_2==1, robust;


* FE
#delimit;
use India_panel.dta, clear;

#delimit;
xtreg nonhomedelivery male Bb* bord if survey==1998 & pcar==2 & twogend==0.5 & wealth_q~=. ,fe
robust;
```