Problem set 2: Empirical Methods for Applied Microeconomics

General instructions. Please work in a group no larger than 3. When you write up your results, please let me know who is in your group. (Only turn in 1 completed homework.). Present your answers in a concise way (typed is highly preferred). Please include relevant Stata output and well-commented do files and ado files for all the exercises (or equivalent in the package of your choice.) Please do NOT include lots of undigested log files. Problem set is due on March 12.

Put the do files in an appendix and make clear reference to the regression output and/or figures.

Problem 1

IV in real data.

Download the data from Project Star, the Tennessee class size experiment from the 1980s from the Stock and Watson book http://fmwww.bc.edu/ec-p/data/stockwatson, star.doc and webstar.dta. For more about this, see Alan Krueger, 1999, “Experimental Estimates of Education Production Functions”, QJE, 114(2), 497-532. http://qje.oxfordjournals.org/content/114/497.short.

We are going to explore some aspects of IV. But first, we need to create some new variables. First, for the purposes of this assignment, we will ignore the assignment to a class with an aid (drop cltypek=3).

Then, create “percentiles” from the regular classes for math and reading, and then merge these onto the regular data, and then create imputed classes. We'll copy this code from the krueger.do program in the MHE archive.

```
/* create Krueger scaled scores */
/* reading score */
clear
use webstar
keep if cltypek > 1/* regular classes */
keep if treadssk ~= .
sort treadssk
gen pread0 = 100*_n/_N
egen pread = mean(pread0), by(treadssk)/* percentile score in
reg. classes */
keep treadssk pread
sort tread
keep if tread ~= tread[_n-1]
save tempr, replace
/* math score */
use webstar
keep if cltypek > 1/* regular classes */
keep if tmathssk ~= .
sort tmathssk
gen pmath0 = 100*_n/_N
egen pmath = mean(pmath0), by(tmathssk)
keep tmathssk pmath
sort tmath
keep if tmath ~= tmath[_n-1]
save tempm, replace
/* merge percentile scores back on */
```

```
use webstar
keep if stark == 1
sort treadssk
merge treadssk using tempr
ipolate pread treadssk, gen(pr) epolate
drop _merge
sort tmathssk
merge tmathssk using tempm
ipolate pmath tmathssk, gen(pm) epolate
replace pm = 0 if pm < 0
drop _merge
egen pscore = rowmean(pr pm)
/* make class ids */
egen classid1 = group(schidkn cltypek)
egen cs1 = count(classid1), by(classid1)
egen classid2 = group(classid1 totexpk hdegk cladk) if cltypek==1 & cs >= 20
egen classid3 = group(classid1 totexpk hdegk cladk) if cltypek>1 & cs >= 30
gen temp = classid1*100
egen classid = rowtotal(temp classid2 classid3)
egen cs = count(classid), by(classid)
gen female = ssex == 2
gen nwhite = srace >= 2 & srace <= 6 if srace ~= .
keep if cs <= 27 & pscore ~= .
```

We will focus on the children's kindergarten assignment and the average of their math and reading percentile scores. For the purposes of the first part of this question, unless I specify otherwise, ignore the fact that students are not randomly drawn from the population, but that whole classrooms are in the data.

(i) Did the random assignment work? Do a joint test that the variables nonwhite, female, number of kids in class (cs), and free lunch status (sesk=1), and total teacher experience (totexpk) are the same across the treatment and control groups. Krueger suggests that random assignment was only good within school. Are there any significant differences controlling for school dummies? (The id for school is schidkn.)

What if you allow for arbitrary correlation within school also?

Does this make you worry about the experiment (1 paragraph max).

(ii) Use the regular class room sample. What is the effect of class size on the average percentile score here (pscore)? Is this large? Meaningful?

(iii) Check that there is a first stage. First do this without any Xs in the model. Is the first stage strong using the rule of thumb number? Other values? What if you control for nonwhite, free and reduced price lunch, gender, and school dummies?

(iv) What does the reduced form look like without controls? With controls? Are smaller classes good for students?

(v) What is the Wald estimate?

(vi) Do two stage least squares by hand with the controls. Check you get the same thing with canned 2SLS. (Extra credit, do your own adjustment to the errors.)

(vii) What if you fail to include the Xs in the first stage? Why might this not change the estimates here?

(viii) Now we will worry about the SEs. First, get robust SEs for the results above in part (vi). Does this matter for your conclusions about the effect of a small class on scores?

(ix) Now you remember that there is a problem with the SEs because everyone in the same classroom has the same value for the class size variable and might share other shocks. You cluster the SEs at the classroom level (allow for arbitary correlation within classroom). Does this matter? Does this change your conclusion in (viii)?

(x) Now instead you decide to take classroom means and get the IV estimates using them (a la Donald and Lang). Does this change your conclusion?

(xi) Extra credit. Use a block bootstrap where you resample entire classrooms. Does this change your conclusion?

Problem 2

IV simulation.

Start with 1000 observations of a standard normal u.

Create e, $v1$, and $v2$. Let e and $v1$ both be u + 0.5 times another standard normal.

Let $v2$ be u plus a standard normal.

Create 41 instruments (all standard normals), $z0, z1, ..., z40$.

Create $x1 = 0.2 \cdot z0 + v1$ and $x2 = 0.2 \cdot z0 + v2$.

Create $y1 = x1 + e$ and $y2 = x2 + e$.

(i) Which x is more highly correlated with the error term, $x1$ or $x2$?

(ii) Run OLS with $y1$ on $x1$, and then separately with $y2$ on $x2$.

(iii) Run 2SLS with the 1 good instrument for both $y1$ and $y2$ (separately). Do you get substantively different answers than in (ii)?

(iv) Run 2SLS with the 1 good and 40 bad instruments for both $y1$ and $y2$ (separately). What are the first stage Fs? Which estimate is further from the true value? Is this what you expected?

(v) Run LIML for both. Does LIML help?

Problem 3

Quantile regression.

Here we will use some data from a paper by Angrist, Kevin Lang, and Phil Oreopoulous on another experiment. Here is the link to the paper:

http://homes.chass.utoronto.ca/ oreo/research/compositions/Incentives_and_Services_For_College_Achievement.p

The data are at

http://econ-www.mit.edu/faculty/angrist/data1/data,

click on Angrist, Lang, and Oreopoulos (2009).

Load in the data. We will use the following sample.

```
(control==1 | sfsp==1) & noshow==0 & (GPA_year1!=. &
grade_20059_fall!=.)
```

We will in some specifications use some of the following controls.

```
List of controls:
female english hsgroup1 hsgroup2  numcourses6 numcourses5 lastminusof
lastminocc momedlehs momedsomcol dadedlehs dadedsomcol


Definitions:
hsgroup1 = hsgroup==1
hsgroup2 = hsgroup==2
hsgroup3 = hsgroup==3
gen numcourses6 = numcourses_==6
gen numcourses5 = numcourses_==5
gen numcoursesle4 = numcourses_<=4
gen lastminusof = lastmin==1 | lastmin==2
```

```
gen lastminocc = lastmin==3
gen lastminrar = lastmin==4 | lastmin==5
gen momedlehs = mom_edn<=2
gen momedsomcol = mom_edn>=3 & mom_edn<=5
gen momedgecol = mom_edn>=6 & mom_edn<=9
gen dadedlehs = dad_edn<=2
gen dadedsomcol = dad_edn>=3 & dad_edn<=5
gen dadedgecol = dad_edn>=6 & dad_edn<=9
```

(i) Start by regressing GPA in year 1 on the SFSP dummy and the controls above in our preferred sample (control group or SFSP group, no noshows).

(ii) Now calculate quantile regression estimates of the effects of the program on GPA in year 1, at the first decile, the median, the 9th decile, using the same sample and controlling for the Xs.

(iii) Run the same quantile regression without any other Xs but the treatment dummy in the specification. Why did the coefficients change from (ii)?

(iv) Now we will calculate the same results by hand. Figure out the 10th, 50th, and 90th percentile of the GPA year 1 distribution for the SFSP group and for the control group. Use these to get estimates of the QTE at these percentiles. Do you get the same estimates for the QTE as in (iii)?

(v) Now we will bootstrap the calculation with replacement. So, for 999 replicates, you will draw the data with replacement, calculate the relevant percentiles of the treatment and control groups, and then for the each percentile, get the QTE within the bootstrap replicates, sort it, and get the 90% CI for each percentile (pointwise) as the 50th smallest bootstrap estimate and 950th largest.

(vi) Now we will do the Abadie, Angrist, and Imbens IV QTE. Download ivqte from Blaise Melly's webpage http://www.econ.brown.edu/fac/Blaise_Melly/code_ivqte.html. You may also need to install "moremata" and "kdens".

First estimate 2SLS (instrument for sfsp_p with sfsp), with all the controls in.

Then estimate the AAI IV QTE using the ivqte command. (See right below, for command with ";" as the delimiter and for the 0.1 decile.

```
ivqte GPA_year1 (sfsp_p=sfsp) if (control==1 | sfsp==1) &
noshow==0 ,  q(.1) variance dummy(female hsgroup1 hsgroup2
numcourses6 numcourses5  lastminusof lastminocc momedlehs momedsomcol
dadedlehs dadedsomcol)  aai;
```

Do you think effects of the program are constant across the distribution?