Solutions to Advanced Econometrics Homework 3
Class Size and Scholastic Achievement

Ashwin Iyenggar (1521001)

ashwin.iyenggar15@iimb.ernet.in

December 6, 2016

Question (a)

Assume that you knew nothing about Maimonides Rule. Estimate equation (3) in the paper using Ordinary Least Squares (OLS). The equation posits a relationship between average test scores (class-level) and class size. Ideally, you might be interested in estimating equation (2) that posits a relationship between class size and individual test scores. However, no such data is available to estimate equation 2. Comment on the assumptions you need to make in order to get to equation (3) from equation (2). Use the same specification as that used in Table II.

Since we are regressing log wages on years of education, this means that $\%\Delta(wi) = (100 * b) * \Delta(Si)$. Therefore the economic meaning of the slope coefficient b is that a unit change in number of years of education, leads to a 100*b% increase in wages. That would have ideally been the case had Si been an exogenous regressor of ln wi. We discuss why that may not quite be the case in the following question. Table **??** nevertheless presents the regression results from this regression under model (1) where yeduc shows a positive and significant effect on the logarithm of hourly wages

```
reg lhwage yeduc
outreg2 using 'imagepath'Ia.tex, ctitle(baseline) tex(pretty frag) dec(4)
    replace
```

Question (b)

A worry in interpreting coefficients estimated using OLS is that there might be factors not included in the regression that are correlated with both class size and test scores. Invoking Maimonides rule you may be able to use school enrollment as an instrumental variable for class size. If class size is causally linked to test scores, then enrollment must be strongly related with test scores. Run a reduced form regression with test scores as the dependent variable and enrollment as one of the independent variables. Basically, populate Table III. Briefly discuss what you learn from this exercise.

Since we are regressing log wages on years of education, this means that $\%\Delta(wi) = (100 * b) * \Delta(Si)$. Therefore the economic meaning of the slope coefficient b is that a unit change in number of years of education, leads to a 100*b% increase in wages. That would have ideally been the case had Si been an exogenous regressor of ln wi. We discuss why that may not quite be the case in the following question. Table **??** nevertheless presents the regression results from this regression under model (1) where yeduc shows a positive and significant effect on the logarithm of hourly wages

```
reg lhwage yeduc
outreg2 using 'imagepath'Ia.tex, ctitle(baseline) tex(pretty frag) dec(4)
    replace
```

Question (c)

Use enrollment >=40 as an IV to re-estimate equation (3) in the paper. Use a similar specification as that used in Table IV and Table V in the paper. Using your estimates, populate Tables IV and V. How do your results compare with that in part (a)?

Since we are regressing log wages on years of education, this means that $\%\Delta(wi) = (100 * b) * \Delta(Si)$. Therefore the economic meaning of the slope coefficient b is that a unit change in number of years of education, leads to a 100*b% increase in wages. That would have ideally been the case had Si been an exogenous regressor of ln wi. We discuss why that may not quite be the case in the following question. Table **??** nevertheless presents the regression results from this regression under model (1) where yeduc shows a positive and significant effect on the logarithm of hourly wages

```
reg lhwage yeduc
outreg2 using 'imagepath'Ia.tex, ctitle(baseline) tex(pretty frag) dec(4)
    replace
```

Question (d)

A concern with using the IV approach to estimate the causal effect of class size on scholastic achievement is that we may not be able to generalize the effects to those schools not affected by the IV. Explain why.

Since we are regressing log wages on years of education, this means that $\%\Delta(wi) = (100 * b) * \Delta(Si)$. Therefore the economic meaning of the slope coefficient b is that a unit change in number of years of education, leads to a 100*b% increase in wages. That would have ideally been the case had Si been an exogenous regressor of ln wi. We discuss why that may not quite be the case in the following question. Table **??** nevertheless presents the regression results from this regression under model (1) where `yeduc` shows a positive and significant effect on the logarithm of hourly wages

```
reg lhwage yeduc
outreg2 using 'imagepath'Ia.tex, ctitle(baseline) tex(pretty frag) dec(4)
    replace
```