PROPOSED SOLUTIONS
HOMEWORK 8
GRA6039 ECONOMETRICS WITH PROGRAMMING
AUTUMN 2020

EMIL A. STOLTENBERG

Solutions to Ex. 1. Consider the model

$$Y_i = x_i^{\rm t}\beta + \varepsilon_i = \beta_0 + \beta_1 x_{i,1} + \cdots + \beta_{p-1} x_{i,p-1} + \varepsilon_i, \quad \text{for } i = 1,\ldots,n,$$

where $\varepsilon_1,\ldots,\varepsilon_n$ are i.i.d. random variables with $\mathrm{E}\,\varepsilon_1 = 0$ and $\mathrm{Var}(\varepsilon_1) = \sigma^2$, the covariates are fixed numbers (not rv's), and $n > p$. The least squares estimator is the minimiser of

$$g(\beta) = g(\beta_0,\ldots,\beta_{p-1}) = \sum_{i=1}^n (Y_i - x_i^{\rm t}\beta)^2.$$

(a) Differentiate with respect to β_0, β_1 and so on, and set the partial derivatives equal to zero,

$$\begin{aligned}
\frac{\partial}{\partial\beta_0} g(\beta) &= -2\sum_{i=1}^n (Y_i - x_i^{\rm t}\beta) = 0,\\
\frac{\partial}{\partial\beta_1} g(\beta) &= -2\sum_{i=1}^n x_{i,1}(Y_i - x_i^{\rm t}\beta) = 0,\\
&\;\;\vdots\\
\frac{\partial}{\partial\beta_1} g(\beta) &= -2\sum_{i=1}^n x_{i,p-1}(Y_i - x_i^{\rm t}\beta) = 0.
\end{aligned}$$

We see that

$$\begin{aligned}
\frac{\partial}{\partial\beta} g(\beta) &= \begin{pmatrix} \frac{\partial}{\partial\beta_0} g(\beta) \\ \frac{\partial}{\partial\beta_1} g(\beta) \\ \vdots \\ \frac{\partial}{\partial\beta_1} g(\beta) \end{pmatrix} = -2 \begin{pmatrix} \sum_{i=1}^n (Y_i - x_i^{\rm t}\beta) \\ \sum_{i=1}^n x_{i,1}(Y_i - x_i^{\rm t}\beta) \\ \vdots \\ \sum_{i=1}^n x_{i,p-1}(Y_i - x_i^{\rm t}\beta) \end{pmatrix} \\
&= -2\sum_{i=1}^n \begin{pmatrix} 1 \\ x_{i,1} \\ \vdots \\ x_{i,p-1} \end{pmatrix} (Y_i - x_i^{\rm t}\beta) = -2\sum_{i=1}^n x_i(Y_i - x_i^{\rm t}\beta) = -2X^{\rm t}(Y - X\beta) = 0
\end{aligned}$$

Date: November 4, 2020.

where the 0 on the right is a $p\times 1$ vector of zeros. We also use that $\sum_{i=1}^n x_iY_i = X^{\rm t}Y$ and $\sum_{i=1}^n x_ix_i^{\rm t} = X^{\rm t}X$.

(b) We assume that $X^{\rm t}X$ is invertible. From the expression above we have

$$X^{\rm t}Y = X^{\rm t}X\beta.$$

Multiply by $(X^{\rm t}X)^{-1}$ on both sides and we get the estimator

$$\widehat{\beta} = (X^{\rm t}X)^{-1}X^{\rm t}Y.$$

Using the representation $Y = X\beta + \varepsilon$ of the regression model, that X is a matrix of fixed numbers, and that $\mathrm{E}\,\varepsilon = 0$,

$$\mathrm{E}\,[\widehat{\beta}] = (X^{\rm t}X)^{-1}X^{\rm t}\mathrm{E}\,[Y] = (X^{\rm t}X)^{-1}X^{\rm t}(X\beta + \mathrm{E}\,[\varepsilon]) = (X^{\rm t}X)^{-1}X^{\rm t}X\beta = \beta.$$

because $(X^{\rm t}X)^{-1}X^{\rm t}X = I_p$ (the identity matrix). For the variance, use that $\mathrm{Var}(AY) = A\mathrm{Var}(Y)A^{\rm t}$, and that $\mathrm{Var}(Y) = \sigma^2 I_n$,

$$\begin{aligned}\mathrm{Var}(\widehat{\beta}) &= (X^{\rm t}X)^{-1}X^{\rm t}\mathrm{Var}(Y)((X^{\rm t}X)^{-1}X^{\rm t})^{\rm t} = \sigma^2(X^{\rm t}X)^{-1}X^{\rm t}((X^{\rm t}X)^{-1}X^{\rm t})^{\rm t}\\ &= \sigma^2(X^{\rm t}X)^{-1}X^{\rm t}X(X^{\rm t}X)^{-1} = \sigma^2(X^{\rm t}X)^{-1},\end{aligned}$$

where we use that $((X^{\rm t}X)^{-1}X^{\rm t})^{\rm t} = X(X^{\rm t}X)^{-1}$, this is because $(AB)^{\rm t} = B^{\rm t}A^{\rm t}$, and $(A^{\rm t})^{-1} = (A^{-1})^{\rm t}$. To see that latter, $A^{-1}A = I$ by definition, and $I^{\rm t} = I$, that is, the transpose of the identity matrix is equal to the identity matrix,

$$(A^{-1}A)^{\rm t} = A^{\rm t}(A^{-1})^{\rm t} = I,$$

so $(A^{-1})^{\rm t} = (A^{\rm t})^{-1}$.

(c) (You should know the conclusion of this exercise, that the estimator is unbiased, but don't worry about the derivation) An unbiased estimator of σ^2 is

$$\widehat{\sigma}_n^2 = \frac{1}{n-p}\sum_{i=1}^n (Y_i - x_i^{\rm t}\widehat{\beta}_n)^2.$$

Write

$$\begin{aligned}(n-p)\widehat{\sigma}_n^2 &= \sum_{i=1}^n (Y_i - x_i^{\rm t}\widehat{\beta}_n)^2 = \sum_{i=1}^n (Y_i - x_i^{\rm t}\widehat{\beta}_n)(Y_i - x_i^{\rm t}\widehat{\beta}_n)\\ &= \sum_{i=1}^n (Y_i - x_i^{\rm t}\widehat{\beta}_n)Y_i = \sum_{i=1}^n Y_i^2 - \sum_{i=1}^n x_i^{\rm t}Y_i\widehat{\beta}_n = \sum_{i=1}^n Y_i^2 - \widehat{\beta}_n^{\rm t}X^{\rm t}X\widehat{\beta}_n.\end{aligned}$$

Then take the expectation,

$$\begin{aligned}
\mathrm{E}\,[(n-p)\widehat{\sigma}_n^2] &= \sum_{i=1}^n \mathrm{E}[\,Y_i^2] - \mathrm{E}\,[\widehat{\beta}_n^{\mathrm{t}} X^{\mathrm{t}} X \widehat{\beta}_n] \\
&= \sum_{i=1}^n (x_i^{\mathrm{t}}\beta)^2 + n\sigma^2 - \mathrm{E}\,\sum_{i=1}^n \widehat{\beta}_n^{\mathrm{t}} x_i x_i^{\mathrm{t}} \widehat{\beta}_n \\
&= \sum_{i=1}^n \beta^{\mathrm{t}} x_i x_i^{\mathrm{t}} \beta + n\sigma^2 - \mathrm{E}\,\sum_{i=1}^n \widehat{\beta}_n^{\mathrm{t}} x_i x_i^{\mathrm{t}} \widehat{\beta}_n \\
&= n\sigma^2 - \mathrm{E}\,\sum_{i=1}^n (\widehat{\beta}_n - \beta)^{\mathrm{t}} x_i x_i^{\mathrm{t}} (\widehat{\beta}_n - \beta) \\
&= n\sigma^2 - \sum_{i=1}^n \mathrm{E}\,(x_i^{\mathrm{t}}(\widehat{\beta}_n - \beta))^2 \\
&= n\sigma^2 - \sum_{i=1}^n \mathrm{Var}\,(x_i^{\mathrm{t}}(\widehat{\beta}_n - \beta)) = n\sigma^2 - \sum_{i=1}^n x_i^{\mathrm{t}} \mathrm{Var}\,(\widehat{\beta}_n - \beta) x_i \\
&= n\sigma^2 - \sigma^2 \sum_{i=1}^n x_i^{\mathrm{t}} (X^{\mathrm{t}} X)^{-1} x_i \\
&= n\sigma^2 - \sigma^2 \mathrm{tr}(X(X^{\mathrm{t}}X)^{-1} X^{\mathrm{t}}) = \sigma^2(n-p).
\end{aligned}$$

This explains why one divides by $n-p$ when estimating the variance σ^2.

(d)–(g) Implement the code below, and make sure you understand it.

```
cd("~/Dropbox/H2020/GRA6039_H20/");

hw8 = readtable("hw8.txt");
y = hw8.y; x1 = hw8.x1;x2 = hw8.x2;
n = length(y);
X = [1 + zeros(n,1),x1,x2]; % The design matrix
p = length(X(1,:)); % Useful later

betahat = inv(transpose(X)*X)*transpose(X)*y;

sigma2hat = sum((y - X*betahat).^2)/(n - p);

sebetahat = sqrt(diag(sigma2hat*inv(transpose(X)*X)));

tvals = betahat./sebetahat;

pvals = 2.*(1 - normcdf(abs(tvals)));
out= round([betahat,sebetahat,tvals,pvals],3);
out = array2table(out);
```

```
out.Properties.VariableNames = {'betahat' 'se' 'z' 'p-value'};
out

% Here is the linear regression function that
% is already in Matlab. Compare the output to that
% you get with the code above.
X2 = [x1,x2];
fitlm(X2,y)
```

Solutions to Ex. 2. Let Z_1 and Z_2 be independent standard normal random variables, and set $Z = (Z_1, Z_2)^{\mathrm{t}}$.

(a) By the result given in Eq. (1) at the start of the Homework set, we know that $AZ + \mu$ has a normal distribution, with expectation $\mathrm{E}[AZ + \mu] = A\mathrm{E}\,[Z] + \mu = \mu$, and variance

$$\mathrm{Var}(AZ + \mu) = AA^{\mathrm{t}} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2. \end{pmatrix}$$

In conclusion

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathrm{N}_2\Big(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2. \end{pmatrix} \Big).$$

(b) We have

$$\begin{aligned} X &= \mu_1 + \sigma_1 Z_1, \\ Y &= \mu_2 + \sigma_2\rho Z_1 + \sigma_2(1-\rho^2)^{1/2} Z_2. \end{aligned}$$

(b) We are to find the condition expectations,

$$\begin{aligned} \mathrm{E}\,[Y \mid X] &= \mathrm{E}\,[\mu_2 + \sigma_2\rho Z_1 + \sigma_2(1-\rho^2)^{1/2} Z_2 \mid X] \\ &= \mu_2 + \sigma_2\rho\mathrm{E}\,[Z_1 \mid X] = \mu_2 + \sigma_2\rho\mathrm{E}\,[\frac{X-\mu_1}{\sigma_1} \mid X] \\ &= \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(X - \mu_1). \end{aligned}$$

Here we use that since X and Z_2 are independent $\mathrm{E}\,[Z_2 \mid X] = \mathrm{E}\,[Z_2] = 0$, for the second equality; and, for the last equality, the rule that says $\mathrm{E}\,[h(X) \mid X] = h(X)$, for some function h. We can write

$$\mathrm{E}\,[Y \mid X] = \mathrm{E}\,[Y] + \frac{\mathrm{Cov}(Y,X)}{\mathrm{Var}(Y)}(X - \mathrm{E}\,[X]).$$

By the symmetry relation between X and Y,

$$\mathrm{E}\,[X \mid Y] = \mathrm{E}\,[X] + \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}(Y - \mathrm{E}\,[Y]) = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}(Y - \mu_2).$$

(c) The conditional variance of Y given X is

$$\begin{aligned}\mathrm{Var}(Y \mid X) &= \mathrm{E}\,[Y^2 \mid X] - (\mathrm{E}\,[Y \mid X])^2 \\ &= \mu_2^2 + 2\mu_2\rho\frac{\sigma_2}{\sigma_1}(X-\mu_1) + \rho^2\frac{\sigma_2^2}{\sigma_1^2}(X-\mu_1)^2 + \sigma_2^2(1-\rho^2) - (\mathrm{E}\,[Y \mid X])^2 \\ &= \sigma_2^2(1-\rho^2),\end{aligned}$$

which you see by expanding the squares in $(\mathrm{E}\,[Y \mid X])^2 = \{\mu_2 + \rho\sigma_2/\sigma_1(X-\mu_1)\}^2$. By the symmetry relation between X and Y again,

$$\mathrm{Var}(X \mid Y) = \sigma_1^2(1-\rho^2).$$

Solutions to Ex. 3. Suppose that the true regression model is

$$Y_i = \beta_0 + \beta_1 X_{i,1} + \beta_1 X_{i,2} + \varepsilon_i, \quad \text{for } i = 1,\dots,n, \tag{1}$$

where $\varepsilon_1,\dots,\varepsilon_n$ are i.i.d. standard normals independent of $X_{1,1},\dots,X_{n,1}$ and $X_{1,2},\dots,X_{n,2}$, while $(X_{1,1},X_{1,2}),\dots,(X_{n,1},X_{n,2})$ is are i.i.d. bivariate normals with distribution

$$\begin{pmatrix} X_{i,1} \\ X_{i,2}\end{pmatrix} \sim \mathrm{N}_2\Big(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1 & \rho \\ \rho & 1\end{pmatrix}\Big), \quad \text{for } i = 1,\dots,n.$$

For some reason it is impossible to collect data on $X_{1,2},\dots,X_{n,2}$, so this variable is not in your dataset. You decide to estimate β_0 and β_1 by the estimators $\widetilde{\beta}_0$ and $\widetilde{\beta}_1$ defined as the minimisers of

$$g(\beta_0,\beta_1) = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_{i,1})^2. \tag{2}$$

Let $X_1^{(n)} = (X_{1,1},\dots,X_{n,1})$, and $X_2^{(n)} = (X_{1,2},\dots,X_{n,2})$, while $X^{(n)} = (X_1^{(n)}, X_2^{(n)})$. We'll use

$$\bar{X}_{n,1} = \frac{1}{n}\sum_{i=1}^n X_{i,1}, \quad \text{and} \quad \bar{X}_{n,2} = \frac{1}{n}\sum_{i=1}^n X_{i,2},$$

for the empirical means.

(a) We use the rule that say $\mathrm{E}\,[h(X) \mid X] = h(X)$. Here $Y_i = \beta_0 + \beta_1 X_{i,1} + \beta_1 X_{i,2} + \varepsilon_i$ is a function of $X_{i,1}, X_{i,2}$, while ε_i is independent of these two.

$$\begin{aligned}&\mathrm{E}\,[Y_i \mid (X_{i,1},X_{i,2})]\mathrm{E}\,[\beta_0 + \beta_1 X_{i,1} + \beta_1 X_{i,2} + \varepsilon_i \mid (X_{i,1},X_{i,2})] \\ &\qquad = \beta_0 + \beta_1 X_{i,1} + \beta_1 X_{i,2} + \mathrm{E}\,[\varepsilon_i \mid (X_{i,1},X_{i,2})] = \beta_0 + \beta_1 X_{i,1} + \beta_1 X_{i,2},\end{aligned}$$

because $\mathrm{E}\,[\varepsilon_i \mid (X_{i,1},X_{i,2})] = \mathrm{E}\,[\varepsilon_i] = 0$, by independence.

(b) Write

$$\frac{1}{n}\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})^2 = \frac{1}{n}\sum_{i=1}^n X_{i,1}^2 - \bar{X}_{n,1}^2.$$

Here $\bar{X}_{n,1} \to_p 0$ by the LLN, and since $g(x)x^2$ is a continuous function, $\bar{X}_{n,1}^2 \to_p 0$ by PLIM.1. The $X_{1,1}^2,\dots,X_{n,1}^2$ are i.i.d. with expectation 1 and variance 2 (read about the moments of the normal distribution on wikipedia). The LLN therefore gives $(1/n)\sum_{i=1}^n X_{i,1}^2 \to_p$

1. From PLIM.2 we can conclude that $(1/n)\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})^2 \to_p 1$.
(b) Write

$$\frac{1}{n}\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})(X_{i,2} - \bar{X}_{n,2}) = \frac{1}{n}\sum_{i=1}^n X_{i,1}X_{i,2} - \bar{X}_{n,1}\bar{X}_{n,2},$$

here $\bar{X}_{n,1} \to_p 0$ and $\bar{X}_{n,2} \to_p 0$, so by PLIM.2 $\bar{X}_{n,1}\bar{X}_{n,2} \to_p 0$. The $X_{1,1}X_{1,2}, \dots, X_{n,1}X_{n,2}$ are i.i.d. random variables with

$$\mathrm{E}\,[X_{1,1}X_{1,2}] = \mathrm{Cov}(X_{1,1}, X_{1,2}) + \mathrm{E}\,[X_{1,1}]\,\mathrm{E}\,[X_{1,2}] = \mathrm{Cov}(X_{1,1}, X_{1,2}) = \rho,$$

and finite variance (to find the variance of $X_{i,1}X_{i,2}$, use techniques from Ex. 2). The LLN then yields

$$\frac{1}{n}\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})(X_{i,2} - \bar{X}_{n,2}) \xrightarrow{p} \rho.$$

(c) The minimisers of $g(\beta_0, \beta_1)$ are the least squares estimators $\widetilde{\beta}_0$ and $\widetilde{\beta}_1$ where

$$\widetilde{\beta}_1 = \frac{\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})Y_i}{\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})^2}.$$

Since $Y_i = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \varepsilon_i$, we can write

$$\begin{aligned}\widetilde{\beta}_1 &= \frac{\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})Y_i}{\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})^2} = \frac{\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})(\beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \varepsilon_i)}{\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})^2} \\ &= \beta_1 + \beta_2 \frac{\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})(X_{i,2} - \bar{X}_{n,2})}{\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})^2} + \frac{\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})\varepsilon_i}{\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})^2}\end{aligned}$$

using that $\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1}) = 0$, and that $\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})X_{i,2} = \sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})(X_{i,2} - \bar{X}_{n,2})$. Using the rule $\mathrm{E}\,[AB \mid B] = B\mathrm{E}\,[A \mid B]$, and that $\mathrm{E}\,[A \mid B] = \mathrm{E}\,[A]$, when A and B are independent,

$$\begin{aligned}\mathrm{E}\,\Big[\frac{\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})\varepsilon_i}{\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})^2} \mid X^{(n)}\Big] &= \frac{\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})\mathrm{E}\,[\varepsilon_i \mid X^{(n)}]}{\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})^2} \\ &= \frac{\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})\mathrm{E}\,[\varepsilon_i]}{\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})^2} = 0.\end{aligned}$$

Then, using that $\mathrm{E}\,[h(X) \mid X] = h(X)$,

$$\mathrm{E}\,[\widetilde{\beta}_1 \mid X^{(n)}] = \beta_1 + \beta_2 \frac{\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})(X_{i,2} - \bar{X}_{n,2})}{\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})^2}. \tag{3}$$

(d) We can combine the results from (b) and (c), and use PLIM.2 to conclude that

$$\frac{\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})(X_{i,2} - \bar{X}_{n,2})}{\sum_{i=1}^n (X_{i,1} - \bar{X}_{n,1})^2} \xrightarrow{p} \rho.$$

But this means that

$$\mathrm{E}\,[\widetilde{\beta}_1 \mid X^{(n)}] \xrightarrow{p} \beta_1 + \rho\beta_2,$$

(e) We have seen that

$$\mathrm{E}\,[\widetilde{\beta}_1 \mid X^{(n)}] + \beta_1 + \mathrm{bias}_n$$

where the bias term is as given in (3). This bias is referred to as omitted variable bias. But if $\rho = 0$ or very close to 0 and n is big, then bias_n is going to be very close to zero with high probability. The moral is that omitting a variable is okay when that variable is uncorrelated with the included independent variable whose effect we are interested in accurately estimating.

Solutions to Ex. 4. Read carefully through the formulation of the model in this exercise. As we saw in Ex. 3 we saw that $\widetilde{\beta}_1$ is biased when there is an omitted independent variable that is correlated with the $X_{i,1}$. With IV-estimation we assume that there is a random variable Z_i that is correlated with $X_{i,1}$, and whose effect on Y_i is not direct, and goes through $X_{i,1}$. In our model

$$X_{i,1} = \gamma Z_i + \eta_i, \quad \text{for } i = 1, \dots, n,$$

where $\eta_1, \dots, \eta_n$ are independent of $Z_1, \dots, Z_n$. First we find the least-squares estimator $\widehat{\gamma}_n$ for γ, then replace $X_{i,1}$ by the predicted values $\widehat{\gamma}_n Z_i$ when estimating β_1. That is, find $\widehat{\beta}_0$ and $\widehat{\beta}_1$ that minimise

$$g_2(\beta_0, \beta_1) = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 \widehat{\gamma}_n Z_i)^2.$$

(b) The estimator $\widehat{\gamma}_n$ is

$$\widehat{\gamma}_n = \frac{\sum_{i=1}^n (Z_i - \bar{Z}_n) X_{i,1}}{\sum_{i=1}^n (Z_i - \bar{Z}_n)^2} = \frac{\sum_{i=1}^n (Z_i - \bar{Z}_n)(\gamma Z_i + \eta_i)}{\sum_{i=1}^n (Z_i - \bar{Z}_n)^2} = \gamma + \frac{\sum_{i=1}^n (Z_i - \bar{Z}_n)\eta_i}{\sum_{i=1}^n (Z_i - \bar{Z}_n)^2}.$$

Here $\sum_{i=1}^n (Z_i - \bar{Z}_n)\eta_i \to_p 0$, and $\sum_{i=1}^n (Z_i - \bar{Z}_n)^2 \to_p 1$, so by PLIM.2, $\widehat{\gamma}_n \to_p \gamma$.

(c) The instrumental variable estimator for β_1 is

$$\widehat{\beta}_{1,\text{iv}} = \frac{\sum_{i=1}^n (\widehat{\gamma}_n Z_i - \widehat{\gamma}_n \bar{Z}_n) Y_i}{\sum_{i=1}^n (\widehat{\gamma}_n Z_i - \widehat{\gamma}_n \bar{Z}_n)^2}.$$

(e) We can write

$$\begin{aligned}\widehat{\beta}_{1,\text{iv}} &= \frac{\sum_{i=1}^n (\widehat{\gamma}_n Z_i - \widehat{\gamma}_n \bar{Z}_n) Y_i}{\sum_{i=1}^n (\widehat{\gamma}_n Z_i - \widehat{\gamma}_n \bar{Z}_n)^2} = \frac{1}{\widehat{\gamma}_n} \frac{\sum_{i=1}^n (Z_i - \bar{Z}_n)(\beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \varepsilon_i)}{\sum_{i=1}^n (Z_i - \bar{Z}_n)^2} \\ &= \beta_1 \frac{\gamma}{\widehat{\gamma}_n} + \frac{\beta_1}{\widehat{\gamma}_n} \frac{\sum_{i=1}^n (Z_i - \bar{Z}_n)\{\beta_1 \eta_i + \beta_2 X_{i,2} + \varepsilon_i\}}{\sum_{i=1}^n (Z_i - \bar{Z}_n)^2}.\end{aligned}$$

We can work through this expression term by term. First, $\gamma/\widehat{\gamma}_n \to_p 1$ using PLIM.2. Second,

$$\frac{1}{n}\sum_{i=1}^n (Z_i - \bar{Z}_n)\eta_i \to_p 0, \quad \frac{1}{n}\sum_{i=1}^n (Z_i - \bar{Z}_n) X_{i,2} \to_p 0, \quad \frac{1}{n}\sum_{i=1}^n (Z_i - \bar{Z}_n)\varepsilon_i \to_p 0,$$

which you can argue for by using the LLN. Third, $\sum_{i=1}^n (Z_i - \bar{Z}_n)^2 \to_p 1$, and we use PLIM.2 to conclude that $\widehat{\beta}_{1,\text{iv}} \to_p \beta_1$.

Solutions to Ex. 5. I'll type this up soon.

Department of Economics, BI Norwegian Business School
Email address: emil.a.stoltenberg@bi.no