## UBC, ECONOMICS 627
## 2019 MIDTERM EXAMINATION

You have 80 minutes. Answer all questions.

1. **(30 points)** Consider the linear IV model

$$Y = X\beta + U,$$
$$X = Z\Pi + V,$$

where $Y$ is the $n$-vector of observations on some dependent variable, $X$ is the $n \times k$ matrix of observations on potentially endogenous regressors, and $Z$ is the $n \times l$ matrix of observations on IVs. Suppose that there are no common variables between $X$ and $Z$, and that the usual exogeneity conditions hold:

$$EZ_iU_i = 0 \text{ and } EZ_iV_i' = 0,$$

where $Z_i'$ is the $i$-th row of the matrix $Z$ (the $i$-th observation), and $U_i$ and $V_i$ are defined similarly with $Z$ replaced by $U$ and $V$ respectively. Assume also that data are iid. Write

$$Y = X\beta + V\lambda + \epsilon, \text{ where}$$
$$\lambda = \left(EV_iV_i'\right)^{-1} EV_iU_i,$$
$$\epsilon = U - V\lambda.$$

(a) Let $\hat{V}$ denote the residuals from the OLS regression of $X$ against $Z$. Let $\hat{\lambda}$ denote the estimated coefficient on $\hat{V}$ from the OLS regression of $Y$ against $X$ and $\hat{V}$:

$$Y = X\hat{\beta} + \hat{V}\hat{\lambda} + \hat{\epsilon},$$

Show that

$$\hat{\lambda} = \left(V'M_ZM_XM_ZV\right)^{-1} V'M_ZM_X\left(V\lambda + \epsilon\right),$$

where $M_X = I - P_X$, $P_X = X(X'X)^{-1}X'$, and $M_Z$ is defined similarly with $X$ replaced by $Z$.

solution:

$$X = \begin{pmatrix} x_1'_{1\times k} \\ \vdots \\ x_n'_{1\times k} \end{pmatrix} \qquad Z = \begin{pmatrix} z_1'_{1\times l} \\ \vdots \\ z_n'_{1\times l} \end{pmatrix}$$

$$\Pi = \begin{pmatrix} \underset{l\times 1}{\pi_1} & \cdots & \underset{l\times 1}{\pi_k} \end{pmatrix} \qquad V = \begin{pmatrix} v_1'_{1\times k} \\ \vdots \\ v_n'_{1\times k} \end{pmatrix}$$

$$\Rightarrow \underset{1\times k}{x_i'} = \underset{1\times l}{z_i'} \underset{l\times k}{\Pi} + \underset{1\times k}{v_i'}$$

$$z_i x_i' = z_i z_i' \Pi + z_i v_i'$$

$$\frac{1}{n}\sum^{n} z_i x_i' = \frac{1}{n}\sum z_i z_i' \hat{\Pi} + \underbrace{\frac{1}{n}\sum_{i=1}^{\hat{n}} z_i v_i'}_{\text{sample moment condition}}$$

$$\hat{\Pi} = \left(\frac{1}{n}\sum z_i z_i'\right)^{-1}\left(\frac{1}{n}\sum z_i x_i'\right)$$

$$\Rightarrow \boxed{\hat{\Pi} = (Z'Z)^{-1}(Z'X)}$$

$$\hat{V} = X - Z\hat{\Pi} = X - \underset{n\times l}{Z}(\underset{l\times l}{Z'Z})^{-1}\underset{l\times n}{Z'}X = \underset{n\times n}{M_Z}\underset{n\times k}{X}$$

$$\hat{\lambda} = (\hat{V}'M_X\hat{V})^{-1}\hat{V}M_XY$$

$$= (X'M_Z M_X M_Z X)^{-1} X'M_Z M_X Y$$

$$= (X'M_Z M_X M_Z X)^{-1} X'M_Z M_X (X\beta + V\lambda + \varepsilon)$$

(*) $M_X X = 0$ by construction.

$$= (X'M_Z M_X M_Z X)^{-1} X'' M_Z M_X (V\lambda + \varepsilon)$$

(*) $M_Z X = M_Z Z\Pi + M_Z V = M_Z V$.

$$= (V'M_Z M_X M_Z V)^{-1} V'M_Z M_X (V\lambda + \varepsilon).$$

(*) Key idea: $Z$ and $V$ are asymptotically uncorrelated so the only correlation between $X$ and $V$ should be driven by $V$ inside of $X$! ($X = Z\Pi + V$)

(b) Find the probability limit of $\hat{\lambda}$. Specify additional assumptions you have to make.

<u>solution:</u>

$$\hat{\lambda} = (\underbrace{V'M_Z M_X M_Z V}_{n})^{-1} \underbrace{V'M_Z M_X (V\lambda + \varepsilon)}_{n}$$

(*) our hope is that WLLN can help us simplify this!

- $$\frac{V'M_Z M_X M_Z V}{n} = \frac{1}{n}V'(I - P_Z)(I - P_X)(I - P_Z)V$$

$$= \frac{1}{n}V'(I - P_X - P_Z + P_Z P_X - P_Z + P_X P_Z - P_Z P_X P_Z)V$$

(*) Notice that $\frac{Z'V}{n} \to E z_i V_i' = 0$ so the terms that involve $P_Z V$ or $(P_Z V)'$ will be $o_p(1)$.

$$= \frac{1}{n}V'(I - P_X)V + o_p(1)$$

$$= \frac{V'V}{n} - \frac{V'X}{n}\left(\frac{X'X}{n}\right)^{-1}\frac{X'V}{n} + o_p(1)$$

(*) $\frac{V'X}{n} = \frac{V'Z\Pi}{n} + \frac{V'V}{n} = o_p(1) + \frac{V'V}{n}$ because $E Z_i V_i' = 0$.

$$= \frac{V'V}{n} - \frac{V'V}{n}\left(\frac{X'X}{n}\right)^{-1}\frac{V'V}{n} + o_p(1)$$

$$= E V_i V_i' - E V_i V_i' (E X_i X_i')^{-1} E V_i V_i' + o_p(1)$$

- $$\frac{V' M_Z M_X V}{n} = \frac{1}{n} V' (I - P_Z)(I - P_X) V$$
$$= \frac{1}{n} V' (I - P_X - P_Z + P_Z P_X) V$$
$$= \frac{1}{n} V'V - \frac{1}{n} V' P_X V - \frac{1}{n} V' \underbrace{P_Z V}_{\substack{= o_p(1) \\ \text{by construction}}} + \frac{1}{n} \underbrace{V' P_Z}_{\substack{(P_Z V)' = o_p(1) \text{ by} \\ \text{construction!}}} P_X V$$
$$= E V_i V_i' - E V_i V_i' (E X_i X_i')^{-1} E V_i V_i' + o_p(1)$$

- $$\frac{V' M_Z M_X \varepsilon}{n} = \frac{1}{n} V' (I - P_Z)(I - P_X) \varepsilon$$
$$= \frac{1}{n} V' (I - P_X - P_Z + P_Z P_X) \varepsilon$$
$$= \underbrace{\frac{1}{n} V'\varepsilon}_{\substack{E V_i \varepsilon_i = 0 \\ \text{so this is} \\ o_p(1)}} - \underbrace{\frac{1}{n} V' P_X \varepsilon}_{\substack{\underbrace{\frac{V'V}{n}}\left(\underbrace{\frac{X'X}{n}}\right)^{-1}\underbrace{\frac{X'\varepsilon}{n}} + o_p(1) \\ = o_p(1) \text{ because} \\ E X_i \varepsilon_i = 0}} - \frac{1}{n} V' \underbrace{P_Z \varepsilon}_{= o_p(1)} + \frac{1}{n} V' \underbrace{P_Z P_X \varepsilon}_{= o_p(1)}$$
$$\ldots \text{ similar arguments}$$
$$= o_p(1).$$

Therefore,

$$\hat{\lambda} = \left(E V_i V_i' - E V_i V_i' (E X_i X_i')^{-1} E V_i V_i' + o_p(1)\right)^{-1} \times$$
$$\left[E V_i V_i' - E V_i V_i' (E X_i X_i')^{-1} E V_i V_i' + o_p(1)\right] \lambda + o_p(1)$$
$$= \lambda + o_p(1) = (E V_i V_i')^{-1} E V_i u_i + o_p(1)$$

(c) Using the result in (b), explain how the estimator $\hat{\lambda}$ can be used for testing whether the regressors $X$ are exogenous. (This test is known as the Durbin-Wu-Hausman test.)

solution:

$\lambda = 0 \Leftrightarrow E V_i u_i = 0$ and since $\hat{\lambda} = \lambda + o_p(1)$, we can use $\hat{\lambda} = 0$ to test for exogeneity of $X$.