KTH Engineering Sciences
Harald Lang 28/2– 07

Övning 34 och Frisch-Waugh-Lovells sats

Övning 34

Låt oss först tänka intuitivt. Om jag skattar ekvationen med OLS, så vet jag att jag kan tolka resultatet som en prediktions-ekvation. Det är då inte samma sak som den struktur-ekvation jag egentligen vill skatta. (Den ekvation jag vill skatta är skall uppenbarligen tolkas som "hur påverkar x y?".)

Om jag nu tänker på den givna strukturekvationen, och så vill jag göra en prediktion av y för givet x. Om då x_1 är mycket stor, så är min prediktion att även residualen är stor, dvs. att y är större än $x'\beta$. Min prediktion bör alltså vara $x'\hat{\beta}$ där $\hat{\beta}_1 > \beta_1$. Så OLS väntar jag mig ger ett för stort värde.

Nu gör vi en mer matematiskt formell analys. Det visar sig då att vi bara kan dra en säker slutsats om x_1 inte är korrelerad med de övriga x-variablerna.

$$y = x_1\beta_1 + x_2\beta_2 + \cdots + e$$

Tag nu kovariansen av bägge led med x_1. Om nu x_1 inte är korrelerad med de övriga x-variablerna får vi

$$\mathrm{Cov}(x_1, y) = \mathrm{Var}(x_1)\beta_1 + \mathrm{Cov}(x_1, e) \tag{1}$$

där alltså $\mathrm{Cov}(x_1, e) > 0$ enligt förutsättningarna i uppgiften. Men OLS-skattningen $\hat{\beta}$ förutsätter ju att x_i är okorrelerad med residualen, så där har vi

$$y = x_1\hat{\beta}_1 + x_2\hat{\beta}_2 + \cdots + \varepsilon$$

med $\mathrm{Cov}(x_1, \varepsilon) = 0$. Alltså

$$\mathrm{Cov}(x_1, y) = \mathrm{Var}(x_1)\hat{\beta}_1 \tag{2}$$

Sammanställer vi nu (1) och (2) får vi

$$\mathrm{Var}(x_1)\hat{\beta}_1 = \mathrm{Var}(x_1)\beta_1 + \mathrm{Cov}(x_1, e)$$

dvs.

$$\hat{\beta}_1 > \beta_1$$

Om nu x_1 är korrelerad med övriga x-variabler så påverkas *alla* parameterskattningarna, inte bara β_1, så kalkylen ovan håller inte. Men rimligtvis står sig resultatet att OLS-skattningen blir för stor om x_1 inte är alltför starkt korrelerad med de övriga x-variablerna. Korrelationen får ju inte heller vara alltför stark, eftersom vi då har multikollinearitet.

Frisch-Waugh-Lovells sats

Jag vill göra en mer intuitiv (enligt min subjektiva mening) härledning än den i Hansen.

$$y = x'\beta + +z'\alpha + e \tag{1}$$

där x och z alltså är förklaringsvariabler (här är alltså z bara en grupp av förklaringsvariabler, inte några instrumentvariabler), och vi vill skatta (1) med OLS. Gör vi det får vi sambanden (med våra vanliga matrisbeteckningar)

$$Y = X\hat{\beta} + Z\hat{\alpha} + \hat{e} \tag{2}$$

Nu är vi egentligen bara intresserade av $\hat{\beta}$, inte av $\hat{\alpha}$, så vi skulle vilja bli av med z-variablerna. Om dessa vore helt okorrelerade med x-variablerna kunde vi bara ta bort dem ut ekvationen – det har vi visat i en övning tidigare. Men nu förutsätter vi inte att z-na är okorrelerade med x-na. Då gör vi så här: vi kör en regeression (OLS) av y på z:

$$Y = Z\hat{\gamma} + \hat{e}_2 \tag{3}$$

och en regression av varje x_j-variabel på z:

$$X_j = Z\hat{\delta}_j + \hat{\varepsilon}_j \tag{4}$$

Satsen säger nu att vi får koefficienterna $\hat{\beta}$ i (2) genom att köra en regression av residualerna i (3) på residualerna i (4):

$$\hat{e}_2 = \hat{\varepsilon}\hat{\beta} + \hat{e} \tag{5}$$

och att residualerna i denna OLS-regression är desamma som i (2).

Nu bevisar vi att det är så. Notera att eftersom (2), (3) och (4) är skattade med OLS, så gäller att

$$X'\hat{e} = 0,\ Z'\hat{e} = 0,\ Z'\hat{e}_2 = 0 \ \text{ och } \ Z'\hat{\varepsilon}_j = 0 \tag{6}$$

Vi härleder nu (5). Vi utnyttjar i tur och ordning (3), (2) och (4):

$$\begin{aligned}
\hat{e}_2 &= Y - Z\hat{\gamma} = X\hat{\beta} + Z\hat{\alpha} + \hat{e} - Z\hat{\gamma} \\
&= \sum_j \big(Z\hat{\delta}_j + \hat{\varepsilon}_j\big)\hat{\beta}_j + Z\hat{\alpha} + \hat{e} - Z\hat{\gamma} \\
&= \sum_j \hat{\varepsilon}_j\hat{\beta}_j + \hat{e} + Z\hat{a} \qquad (7)\\
\text{där } \ \hat{a} &= \Big(\sum_j \hat{\delta}_j\hat{\beta}_j + \hat{\alpha} - \hat{\gamma}\Big)
\end{aligned}$$

Nu visar jag att $Z\hat{a} = 0$. Multiplicera nämligen likheten (7) med Z' och utnyttja (6):

$$0 = Z'Z\hat{a}$$

vilket också medför att

$$\big(Z\hat{a}\big)'\big(Z\hat{a}\big) = 0$$

Men $Z\hat{a}$ är en kolonn-vektor, och här står att dess längd i kvadrat är noll. Alltså är den noll-vektorn, dvs. $Z\hat{a} = 0$. Ekvation (7) blir alltså

$$\hat{e}_2 = \sum_j \hat{\varepsilon}_j \hat{\beta}_j + \hat{e}$$

och nu är saken klar, om vi kan visa att detta verkligen är OLS-skattningen av $\hat{e}_2$ på $\hat{\varepsilon}$. För att bevisa det är allt vi behöver göra är att visa att $\hat{\varepsilon}_j' \hat{e} = 0$, ty det är ju dessa likheter som definierar OLS! Men

$$\begin{aligned} \hat{\varepsilon}_j' \hat{e} &= \Big(X_j - Z\hat{\delta}\Big)' \hat{e} \\ &= X_j' \hat{e} - \hat{\delta}' Z' \hat{e} = 0 \end{aligned}$$

den sista likheten p.g.a. (6). Därmed är beviset klart.

Observera att jag inte menar att denna sats är en del av teorin i kursen. Jag tar upp den som en övning i OLS. Men faktum är att satsen ibland används, särskilt för paneldata (som vi inte berört i kursen).