

Miért pont úgy kombinálja kétfokozatú legkisebb négyzetek módszere (2SLS) az instrumentumokat, ahogy?

Kézirat

A Hunyadi László 60. születésnapjára készülő könyvbe

Kézdi Gábor

2004. július

A Budapesti Corvinus Egyetem rövid életű Ökonometria Csoportjának vezetőjeként Hunyadi Laci fontosnak tartotta, hogy a tanításon túl is foglalkozzunk ökonometria problémákkal, és ezekről beszéljünk is egymás között. A cél az volt, hogy egymást jobban megismertessük érdekes alkalmazásokkal, bonyolultabb modellekkel, és az egyébként általunk viszonylag gyakran használt módszerek mélyebb hátterével. Ez a műhelymunka valóban elindult, bár a hétköznapi viharai és az Ökonometria Csoportnak és oktatóinak a jövője körüli bizonytalanság miatt nem futott fel teljesen. De nem adtuk fel a terveinket. Ez az írás is ezt akarja bizonyítani: egy olyan kérdést boncolgat, amelyet Laci többször felvetett, ám rendszeresen sosem beszéltünk végig.

A kérdés a következő. A kétfokozatú legkisebb négyzetek módszere több instrumentum esetén az endogén magyarázó változóknak az egzogén magyarázó változókra (instrumentumokra) való lineáris projekcióját használja a becsléshez. Ezáltal a túl sok instrumentumból azok egyféle lineáris kombinációjával hoz létre éppen elegendő számú instrumentumot. De vajon mi a megfontolás pont e lineáris kombináció mögött? Optimális megoldás-e ez, és ha igen, milyen értelemben, és milyen feltételek mellett? És végül, elképzelhető-e olyan szituáció, amikor van a 2SLS-nél jobb megoldás?

Tekintsünk egy független azonos eloszású (iid) mintát és rajta egy kétváltozós lineáris modellt, ahol a egyetlen magyarázó változó van, amely endogén:

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

$$E(u_i) = 0$$

$$\text{Cov}(u_i, x_i) \neq 0$$

Az elemzés során végig feltesszük, hogy a modell korrektül specifikált, vagyis a lineáris függvényforma a megfelelő, és x hatása y -ra minden egyes i esetén β_1 .

x endogenitása miatt β_1 OLS becslése inkonzisztens:

$$\begin{aligned}\hat{\beta}_{1_OLS} &\equiv \frac{\sum (y_i - \bar{y})(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \\ p \lim \hat{\beta}_{1_OLS} &= \frac{\text{Cov}(y_i, x_i)}{V(x_i)} = \frac{\text{Cov}(\beta_0 + \beta_1 x_i + u_i, x_i)}{V(x_i)} = \\ &= \frac{\beta_1 V(x_i)}{V(x_i)} + \frac{\text{Cov}(u_i, x_i)}{V(x_i)} \neq \beta_1\end{aligned}$$

Ha találunk megfelelő instrumentumot, β_1 konzisztensen becsülhető. Megfelelő (érvényes) instrumentum korrelálatlan a nem megfigyelt komponenssel és korrelált az endogén magyarázó változóval:

$$Cov(z_i, u_i) = 0$$

$$Cov(z_i, x_i) \neq 0$$

Ezeket az instrumentum momentumfeltételeinek nevezzük. Az első momentumfeltétel azt köti ki, hogy egy érvényes instrumentum nem korrelálhat a nem megfigyelhető heterogenitással; a második azt, hogy a megfigyelhető magyarázó változóval viszont korrelálnia kell. E két feltétel eredményeként az instrumentum közvetlenül nem, a megfigyelt magyarázó változón keresztül viszont hat az eredményváltozóra. Az instrumentális identifikáció, és az arra épülő instrumentális becslőfüggvény ezt használja ki: z és y megfigyelt együttmozgása két hatás eredője: z hatása x -re, és x hatása y -ra (z közvetlenül nem hat y -ra). Ha a megfigyelt z és y együttmozgásból "kiszűrjük" z hatását x -re, megkapjuk x hatását y -ra, vagyis β_1 -et.

Az instrumentális változó (IV) becslőfüggvény:

$$\hat{\beta}_{1-IV} \equiv \frac{\sum (y_i - \bar{y})(z_i - \bar{z})}{\sum (x_i - \bar{x})(z_i - \bar{z})}$$

$\hat{\beta}_{1-IV}$ konzisztens becslőfüggvénye β_1 -nek:

$$\begin{aligned} p \lim \hat{\beta}_{1-IV} &= \frac{Cov(y_i, z_i)}{Cov(x_i, z_i)} = \frac{Cov(\beta_0 + \beta_1 x_i + u_i, z_i)}{Cov(x_i, z_i)} = \\ &= \frac{\beta_1 Cov(x_i, z_i)}{Cov(x_i, z_i)} + \frac{Cov(u_i, z_i)}{Cov(x_i, z_i)} = \beta_1 \end{aligned}$$

Ha olyan szerencsés helyzetben vagyunk, hogy nemcsak egy, hanem több érvényes instrumentumunk is van, a bőség zavara vet fel egy újabb problémát. Ha a modell korrekten specifikált, bármelyik felhasználásával konzisztensen becsülhetjük β_1 -et. Kombinálásuk azonban hatásosabb becslőfüggvényhez vezethet: több instrumentum több információt tartalmazhat, mint egy. Ha mindegyik instrumentum korrelálatlan u -val, úgy bármilyen lineáris kombinációjuk is korrelálatlan:

$$\begin{aligned} Cov(z_{1i}, u_i) &= Cov(z_{2i}, u_i) = \dots = Cov(z_{Li}, u_i) = 0 \\ \Rightarrow Cov\left[\left(\sum_{l=1}^L \lambda_l z_{li}\right), u_i\right] &= \sum_{l=1}^L \lambda_l Cov(z_{li}, u_i) = 0 \end{aligned}$$

A kétfokozatú legkisebb négyzetek módszere (2SLS) a következő kombinációt alkalmazza:

$$x_i^* = \hat{\gamma}_0 + \hat{\gamma}_1 z_{1i} + \hat{\gamma}_2 z_{2i} + \dots + \hat{\gamma}_L z_{Li},$$

ahol a $\hat{\gamma}$ paraméterek az OLS becslések abban a lineáris regresszióban, melynek eredményváltozója az x , magyarázó változói pedig a z -k. A 2SLS ekkor

$$\hat{\beta}_{1_2SLS} \equiv \frac{\sum (y_i - \bar{y})(x_i^* - \bar{x}^*)}{\sum (x_i^* - \bar{x}^*)^2}.$$

A kérdés az, hogy van-e olyan lineáris kombináció, amelynek kisebb (aszimptotikus) varianciája van, mint a többinek, és ha igen, ezt hogyan határozhatjuk meg - és hogy vajon a 2SLS ilyen becslőfüggvény-e.

A Momentumok Általánosított Módszere (GMM) keretében választ kaphatunk ezekre a kérdésekre. A kiindulópontot az érvényes instrumentum momentumfeltételei jelentik. Legyen $\mathbf{x}_i = (1 \ x_i)'$ a modell magyarázó változóinak (esetünkben a konstans és az egyetlen x) a vektora; $\boldsymbol{\beta} = (\beta_0 \ \beta_1)'$ a becülendő paraméterek vektora, és $\mathbf{z}_i = (1 \ z_{1i} \ \dots \ z_{Li})'$ az instrumentumok vektora. \mathbf{x}_i és $\boldsymbol{\beta}$ 2×1 -es, \mathbf{z}_i pedig $(L+1) \times 1$ -es oszlopvektor (L instrumentum és a konstans).

Később használandó referenciaként jegyezzük meg, hogy $\boldsymbol{\beta}$ 2SLS becslőfüggvénye e jelölések alapján a következő:

$$\hat{\boldsymbol{\beta}}_{2SLS} = \left(n^{-1} \sum \mathbf{x}_i^* \mathbf{x}_i^{*'} \right)^{-1} \left(n^{-1} \sum \mathbf{x}_i^* y_i \right).$$

A továbbiakban végig feltesszük, hogy $E(\mathbf{z}_i \mathbf{x}_i')$ rangja 2 (ez a megfelelője az érvényes instrumentumok második momentumfeltételének, vagyis az x -szel való korreláltságnak). A kiinduló momentumfeltételeket ekkor a következő egyenletrendszer foglalja össze:

$$E(\mathbf{z}_i u_i) = E[\mathbf{z}_i (y_i - \mathbf{x}_i' \boldsymbol{\beta})] = \mathbf{0}.$$

$\boldsymbol{\beta}$ GMM becslőfüggvénye az analógia elvén alapul: a várható értéket annak mintabeli megfelelőjével, a mintaátlaggal helyettesíti:

$$\frac{1}{n} \sum \mathbf{z}_i u_i = \frac{1}{n} \sum \mathbf{z}_i (y_i - \mathbf{x}_i' \boldsymbol{\beta}) = S_{zy} - S_{zx} \boldsymbol{\beta} \rightarrow E(\mathbf{z}_i u_i) \text{ in prob.}$$

$$S_{zy} \equiv \frac{1}{n} \sum_i \mathbf{z}_i y_i$$

$$S_{zx} \equiv \frac{1}{n} \sum_i \mathbf{z}_i \mathbf{x}_i'$$

Mint hogy véges mintában nulla valószínűséggel lesz mindegyik mintaátlag pontosan nulla még ha a várható értékek mind nullák is, a GMM azt a paramétert keresi, amely mellett a momentumfeltételből képzett kvadratikus forma a legközelebb van nullához:

$$\begin{aligned} \hat{\boldsymbol{\beta}}_{GMM} &\equiv \arg \min_{\boldsymbol{\beta}} (S_{zy} - S_{zx} \boldsymbol{\beta})' \mathbf{A}^{-1} (S_{zy} - S_{zx} \boldsymbol{\beta}) \\ &= (S_{zx}' \mathbf{A}^{-1} S_{zx})^{-1} (S_{zx}' \mathbf{A}^{-1} S_{zy}) \end{aligned}$$

ahol a második egyenlőségénél az optimum elsőrendű feltételét írtuk le (kihasználjuk hogy a kvadratikus forma konvex, sőt 1 valószínűséggel szigorúan konvex, így az elsőrendű feltétel elégséges). S_{zy} a mintabeli négyzetösszeg (osztva a mintaelemszámmal), amely egy $(L+1) \times 1$ dimenziójú oszlopvektor; S_{zx} pedig ezzel analóg mátrix, dimenziója $(L+1) \times 2$. \mathbf{A} bármilyen olyan $(L+1) \times (L+1)$ dimenziójú valószínűségi mátrix lehet, amely valószínűségben valamilyen pozitív definit mátrixhoz konvergál:

$$p \lim \mathbf{A} = \Psi \quad \text{poz.def.}$$

A GMM becslőfüggvény konzisztenciáját igen egyszerű belátni:

$$\begin{aligned} \hat{\beta}_{GMM} &= (S_{zx}' \mathbf{A}^{-1} S_{zx})^{-1} (S_{zx}' \mathbf{A}^{-1} S_{zy}) \\ &= (S_{zx}' \mathbf{A}^{-1} S_{zx})^{-1} (S_{zx}' \mathbf{A}^{-1} [S_{zx} \beta + S_{zu}]) \\ &= \beta + (S_{zx}' \mathbf{A}^{-1} S_{zx})^{-1} (S_{zx}' \mathbf{A}^{-1} S_{zu}) \end{aligned}$$

$$S_{zu} \equiv \frac{1}{n} \sum_i \mathbf{z}_i u_i$$

$$p \lim \hat{\beta}_{GMM} = \beta + [E(zx)' \Psi^{-1} E(zx)]^{-1} [E(zx)' \Psi^{-1} E(zu)] = \beta$$

Az utolsó sorban azt használjuk ki, hogy $p \lim S_{zu} = E(\mathbf{z}_i u_i) = \mathbf{0}$, hogy $\hat{\beta}_{GMM}$ ennek folytonos függvénye és így alkalmazható a Slutsky-tétel¹, valamint hogy az \mathbf{A} mátrix valószínűségi határának (Ψ) létezik az inverze.²

Az \mathbf{A} mátrixtól függően végtelen sok GMM becslőfüggvény létezik, és mindegyik konzisztens. Ez annak az újrafogalmazása, hogy érvényes instrumentumok bármilyen lineáris kombinációjával készíthető konzisztens becslőfüggvény: a kombinációhoz használt súlyok mátrixa nem más, mint $\mathbf{A}^{-1/2}$.

A GMM becslőfüggvény szimptotikusan normális:

$$\sqrt{n} (\hat{\beta}_{GMM} - \beta) \xrightarrow{D} N(0, \mathbf{\Lambda})$$

$$\mathbf{\Lambda} \equiv \mathbf{\Lambda} \mathbf{\Omega} \mathbf{\Lambda}'$$

$$\mathbf{\Lambda} \equiv [E(\mathbf{z}_i \mathbf{x}_i)' \Psi^{-1} E(\mathbf{z}_i \mathbf{x}_i)]^{-1} E(\mathbf{z}_i \mathbf{x}_i)' \Psi^{-1}$$

$$\mathbf{\Omega} \equiv V(\mathbf{z}_i u_i) = E(\mathbf{z}_i u_i^2 \mathbf{z}_i') = E(u_i^2 \mathbf{z}_i \mathbf{z}_i')$$

A legkisebb aszimptotikus varianciát az a GMM becslőfüggvény adja, amelyben

$$\mathbf{A} = \hat{\mathbf{\Omega}}$$

$$p \lim \hat{\mathbf{\Omega}} = \mathbf{\Omega} = E(u_i^2 \mathbf{z}_i \mathbf{z}_i')$$

¹ A Slutsky-tétel azt mondja ki, hogy ha $p \lim \xi = \mu$, akkor bármely f folytonos függvényre

$$p \lim f(\xi) = f(\mu).$$

² Elvileg nem kell, hogy az inverz véges mintában is létezzon, csak olyankor általánosított inverzt (Moore-Penrose) kell használni. A lényeg megértéséhez ettől a finomságtól nyugodtan eltekinthetünk, és feltehetjük, hogy \mathbf{A} maga is pozitív definit.

Ezt a becslőfüggvényt Optimális GMM-nek (OGMM) nevezik:

$$\hat{\beta}_{OGMM} = \left(S_{zx}' \hat{\Omega}^{-1} S_{zx} \right)^{-1} \left(S_{zx}' \hat{\Omega}^{-1} S_{zy} \right)$$

Kicsit pongyolán fogalmazva, az optimális súlymátrix az instrumentumok (z) és a nem megfigyelt heterogenitás (u) szorzata “szórásának” (a covarianciamátrix $\frac{1}{2}$ hatványának) az inverze. Az intuíció gyakorlatilag ugyanaz, mint az általánosított legkisebb négyzetek módszerénél (GLS): a legjobb lineáris kombináció az, ahol az egyes instrumentumok annál kisebb súlyt kapnak, minél zajosabbak (minél inkább szóródnak véges mintában a 0 momentumfeltétel körül). A minimális aszimptotikus variancia bizonyítást itt nem vezetjük le;³ lényegében ugyanarra a kaptafára megy, mint a minimális variancájú becslőfüggvények bizonyításai általában (Gauss-Markov tétel).

Kérdés marad az, hogy pontosan mi is $\hat{\Omega}$, amelyről eddig annyit tudunk, hogy konzisztensnek kell lennie Ω -ra. Egyszerű választ ad erre az analógia elve:

$$\hat{\Omega} = \frac{1}{n} \sum \hat{u}_i^2 \mathbf{z}_i \mathbf{z}_i'$$

ami kétlépcsős OGMM becslési eljárást jelent: első lépcsőben megfelelően kell \hat{u} -kat becsülnünk $\hat{\Omega}$ -hoz, majd második lépcsőben e (Ω -ra konzisztensen) becsült $\hat{\Omega}$ felhasználásával kapjuk meg az OGMM becslőfüggvényt. Ez megint a GLS módszerrel analóg, illetve annak megvalósítható változatával (FGLS). Minden ilyen becslőfüggvény konzisztens lesz Ω -ra, ha olyan $\hat{u}_i = y_i - \mathbf{x}_i' \hat{\beta}$ változók szerepelnek benne, ahol $\hat{\beta}$ konzisztens becslőfüggvénye β -nak. Tudjuk, hogy a rendelkezésre álló z változók bármilyen lineáris kombinációjával konzisztens instrumentális becslőfüggvény készíthető, vagy másképpen fogalmazva, bármely pozitív definit \mathbf{A} mátrix konzisztens GMM becslőfüggvényhez vezet. Így bármelyiket használhatjuk az első lépcsőben, de praktikus szempontból az legegyszerűbb, ha az első lépcsőben $\mathbf{A}=\mathbf{I}$ (az $(L+1) \times (L+1)$ dimenziójú egységmátrix).

A kérdés az, hogy van-e ehhez az optimális GMM-hez bármi köze a 2SLS-nek. A válasz: igen, bizonyos feltételek mellett. Tegyük fel, hogy u homoszkedasztikus z -re kondicionálva, vagyis $V(u_i | \mathbf{z}_i) = V(u_i)$, és ezért

$$E(u_i^2 \mathbf{z}_i \mathbf{z}_i') = E(u_i^2) E(\mathbf{z}_i \mathbf{z}_i') \equiv \sigma^2 E(\mathbf{z}_i \mathbf{z}_i').$$

Ebben az esetben Ω konzisztens becslőfüggvénye a következő:

$$\hat{\Omega}_{hom} = \frac{1}{n} \sum \hat{u}_i^2 \frac{1}{n} \sum \mathbf{z}_i \mathbf{z}_i' \equiv \hat{\sigma}^2 S_{zz}$$

Elvileg ez is kétlépcsős eljárást tenne szükségessé akárcsak az OGMM általános esetében, ám egy szerencsés “véletlen” ettől megóv minket: minthogy a reziduális variancia (illetve annak

³ Formális bizonyítást lásd például Wooldridge (2002), 8.3. fejezet.

reciproka) a becslőfüggvénynek mind a “nevezőjében”, mind a “számlálójában” szerepel, egyszerűen kiesik:

$$\begin{aligned}\hat{\beta}_{OGMM_hom} &= (S_{zx}' \hat{\Omega}_{hom}^{-1} S_{zx})^{-1} (S_{zx}' \hat{\Omega}_{hom}^{-1} S_{zy}) = \left[S_{zx}' (\hat{\sigma}^2 S_{zz})^{-1} S_{zx} \right]^{-1} \left[S_{zx}' (\hat{\sigma}^2 S_{zz})^{-1} S_{zy} \right] \\ &= (S_{zx}' S_{zz}^{-1} S_{zx})^{-1} (S_{zx}' S_{zz}^{-1} S_{zy})\end{aligned}$$

Közben szép csendesen elérkeztünk a gondolatmenet végéhez. Ez a becslőfüggvény ugyanis nem más mint a 2SLS:

$$\begin{aligned}\hat{\beta}_{2SLS} &= (n^{-1} \sum \mathbf{x}_i^* \mathbf{x}_i^{*'})^{-1} (n^{-1} \sum \mathbf{x}_i^* y_i) = \left[n^{-1} \sum (\hat{\gamma}' \mathbf{z}_i) (\hat{\gamma}' \mathbf{z}_i)' \right]^{-1} \left[n^{-1} \sum (\hat{\gamma}' \mathbf{z}_i) y_i \right] \\ &= \left[n^{-1} \sum \hat{\gamma}' \mathbf{z}_i \mathbf{z}_i' \hat{\gamma} \right]^{-1} \left[n^{-1} \sum \hat{\gamma}' \mathbf{z}_i y_i \right] = \left[n^{-1} \sum (S_{zz}^{-1} S_{zx})' \mathbf{z}_i \mathbf{z}_i' (S_{zz}^{-1} S_{zx}) \right]^{-1} \left[n^{-1} \sum (S_{zz}^{-1} S_{zx})' \mathbf{z}_i y_i \right] \\ &= \left[n^{-1} \sum S_{zx}' S_{zz}^{-1} \mathbf{z}_i \mathbf{z}_i' S_{zz}^{-1} S_{zx} \right]^{-1} \left[n^{-1} \sum S_{zx}' S_{zz}^{-1} \mathbf{z}_i y_i \right] \\ &= \left[S_{zx}' S_{zz}^{-1} (n^{-1} \sum \mathbf{z}_i \mathbf{z}_i') S_{zz}^{-1} S_{zx} \right]^{-1} \left[S_{zx}' S_{zz}^{-1} (n^{-1} \sum \mathbf{z}_i y_i) \right] = (S_{zx}' S_{zz}^{-1} S_{zz} S_{zz}^{-1} S_{zx})^{-1} (S_{zx}' S_{zz}^{-1} S_{zy}) \\ &= (S_{zx}' S_{zz}^{-1} S_{zx})^{-1} (S_{zx}' S_{zz}^{-1} S_{zy}) = \hat{\beta}_{OGMM_hom}\end{aligned}$$

ahol a második sorban kihasználtuk, hogy definíció szerint $\hat{\gamma} = \hat{\gamma}_{OLS} = S_{zz}^{-1} S_{zx}$.⁴

A 2SLS tehát pontosan megegyezik az Optimális GMM-mel ha a nem megfigyelt heterogenitás az instrumentumokra kondicionáltan homoszkedasztikus. Vagyis ebben az esetben 2SLS az a lineáris kombinációja az instrumentumoknak, amelyika legkisebb aszimptotikus varianciájú becslést biztosítja. Magyarul: a legjobb. Amennyiben a nem megfigyelt heterogenitás heteroszkedasztikus, akkor viszont nem az.

De vajon praktikus szempontból lényeges-e az OGMM és a 2SLS közötti különbség heteroszkedasztikus esetben? Ezt a kérdés két dolog is motiválja. Egyrészt az OGMM bonyolultabb eljárást igényel, ezért ha praktikus szempontból nem nagy az előnye, kár vele bajlódni. Másrészt ráadásul az OGMM kétlépcsős eljárása nemcsak bonyolultabb, de véges mintában bizonytalanabb, sőt torz is lehet (lásd pl. Podivinsky, 1999).

A kérdés vizsgálatához egy egyszerű Monte Carlo szimulációt végeztünk el. A szimuláció során két különböző adatgeneráló folyamatot (DGP), egy homoszkedasztikus (DGP₁) és egy heteroszkedasztikus (DGP₂) vizsgáltunk. Mindkét folyamatban egy endogén magyarázó változó (x) és 2 érvényes instrumentum volt (z_1 és z_2); a két instrumentum közül z_1 jobban korrelált x -szel (vagyis erősebb), z_2 kevésbé (gyengébb). DGP₂-ben az instrumentumok négyzetei korreláltak a nem megfigyelt heterogenitás (u) négyzetével, így u feltételesen heteroszkedasztikus volt ($\text{Corr}(u^2, z_1^2) = \text{Corr}(u^2, z_2^2) = 0.25$). A heteroszkedaszticitás mértéke közepesen erősnek mondható.

⁴ $\hat{\gamma}$ az az $(L+1) \times 2$ dimenziójú OLS paramétervektor, amelynek első oszlopában \mathbf{x} első elemének – a konstansnak –, második oszlopában \mathbf{x} második elemének – x -nek – a \mathbf{z} vektoron futtatott regressziós paramétereit becsüljük. Minthogy a konstans nem szóródik, annak paraméterei mind nullák, így $\hat{\gamma}$ első oszlopa is nullvektor.

A Monte Carlo szimulációban 50 ezerszer generáltunk mintát az adott DGP alapján, és ezeken a mintákon egyenként megbecsültük β_1 OLS, IV₁ (IV csak z_1 -gyel), IV₂ (IV csak z_2 -vel), 2SLS és OGMM becsléseit. Az 50 ezer ismétlés után megvizsgáltuk a különböző módokon becsült β_1 -ek átlagos relatív eltérését a valóságtól (Rel.Bias vagyis relatív torzítás), a szóródását (Std vagyis szórás), valamint a torzítás és a szóródás együttes hatásaként adódó teljes eltérésnégyzetet (RMSE, root mean squared error). A Monte Carlo szimulációt elvégeztük 100, 1000 és 10 000 elemű mintákra is.

Az alábbi táblázat foglalja össze szimulációk eredményeit. A DGP-k pontos leírását a Függelék tartalmazza.

Táblázat: A Monte Carlo szimulációk eredményei

	DGP1 (homoszkedasztikus)					DGP2 (heteroszkedasztikus)				
	OLS	IV(z1)	IV(z2)	2SLS	OGMM	OLS	IV(z1)	IV(z2)	2SLS	OGMM
n=100										
Rel.bias	0.463	-0.084	-0.175	0.010	0.011	0.463	-0.044	1.334	0.014	0.012
Std	0.089	27.442	77.039	0.529	0.534	0.092	9.928	482.118	0.660	0.660
RMSE	0.472	27.442	77.039	0.529	0.534	0.472	9.928	482.119	0.660	0.660
n=1000										
Rel.bias	0.464	-0.008	-0.094	0.000	0.000	0.464	-0.006	-0.105	0.001	0.001
Std	0.028	0.131	3.324	0.119	0.119	0.029	0.167	7.247	0.153	0.153
RMSE	0.465	0.131	3.325	0.119	0.119	0.464	0.167	7.247	0.153	0.153
n=10 000										
Rel.bias	0.464	-0.001	-0.004	0.000	0.000	0.464	-0.001	-0.005	0.000	0.000
Std	0.009	0.040	0.102	0.037	0.037	0.009	0.052	0.132	0.048	0.048
RMSE	0.464	0.040	0.102	0.037	0.037	0.464	0.052	0.132	0.048	0.048

A szimulációk rendben kimutatják az ismert eredményeket: az OLS torz, és bár szórása minden mintanagyság mellett kisebb mint bármi más becslőfüggvényé, végeredményben nagyon mellélő. Az egyetlen instrumentumot használó instrumentális becslőfüggvények kis mintában torzok, ám nagy mintában ez eltűnik (konzisztencia). Az erősebb IV-t (z_1) használó becslőfüggvénynek a kismintás torzítása és a szórása is kisebb, mint a gyengébb instrumentumot (z_2) használóé. Az instrumentumokat kombináló becslőfüggvényekben (2SLS, OGMM) gyakorlatilag eltűnik a kismintás torzítás, és a szórás is mindig jóval kisebb, mint az egy instrumentumot használó IV-k esetén. A heteroszkedasztikus DGP esetében bizonytalanabbak becslések: a torzítások és a szórások is általában nagyobbak.

Ami fő kérdésünket, a 2SLS és az OGMM viszonyát illeti, az eredmények meglehetősen egyértelműek. Homoszkedasztikus DGP mellett a 2SLS kis mintában precízebb (hiszen kihasználja a homoszkedaszticitás feltevését, ami itt helyes), közepes és nagy mintában azonban a kettő teljesen azonos eredményt produkál. Heteroszkedasztikus DGP esetén a nagy mintás hasonlóság megmarad, de kis mintában sem jobb az OGMM. Az eredmények mögött valószínűleg az áll, hogy bár az OGMM gyorsabban konvergál a valós β -hoz (kisebb az aszimptotikus varianciája), ez az előny praktikus szempontból elenyésző. Ugyanakkor azonban az OGMM kétlépcsős eljárása plusz bizonytalanságot visz a becslésbe kis minta esetén, ezért kis mintában sincs meg az előnye a 2SLS-sel szemben.

Az analitikus és a szimulációs eredményeket a következőképpen foglalhatjuk össze. A 2SLS elvileg is a legjobb (legisebb aszimptotikus varianciát adó) módon kombinálja az instrumentumokat homoszkedasztikus esetben. Heteroszkedasztikus esetben elvileg van nála

jobb becslőfüggvény, praktikusban azonban ennek az elvileg jobb becslésnek több a hátránya, mint az előnye. A 2SLS megállja a helyét heteroszkedasztikus környezetben is, ezért használjuk csak bátran.

Hivatkozások:

Podivinsky, Jan M. (1999): Finite sample properties of GMM estimators and tests. In L. Mátyás (szerk): Generalized Method of Moments Estimation. Cambridge University Press.

Jeffrey M. Wooldridge (2002): Econometric analysis of cross section and panel data. MIT Press.

Függelék

A Monte Carlo szimulációkban használt adatgeneráló folyamatok (DGP-k) pontos leírása

DGP₁:

$$z_1 \sim \text{iidN}(0,0.5)$$

$$z_2 \sim \text{iidN}(0,0.5)$$

$$x = 0.5z_1 + 0.2z_2 + v$$

$$v \sim \text{iidN}(0,\sigma_v) \text{ úgy, hogy } \sigma_x=1$$

$$u \sim \text{N}(0,1) \text{ úgy, hogy } \text{Corr}(u,v)=0.5 \quad y = \beta_1 x + u$$

$$\beta_1 = 1$$

vagyis: endogén x , 2 érvényes instrumentum z_1 & z_2 , z_1 erősebb:

$$\text{Corr}(x,u) = 0.464$$

$$\text{Corr}(u,z_1) = \text{Corr}(u,z_2)=0.$$

$$\text{Corr}(z_1,x) \approx 0.25, \text{Corr}(z_2,x) \approx 0.10.$$

$$\text{Homoszkedaszticitás: } \text{Corr}(u^2, z_1^2) = \text{Corr}(u^2, z_2^2)=0.$$

DGP₂:

$$z_1 \sim \text{iidN}(0,0.5)$$

$$z_2 \sim \text{iidN}(0,0.5)$$

$$x = 0.5z_1 + 0.2z_2 + v$$

$$v \sim \text{iidN}(0,\sigma_v) \text{ úgy, hogy } \sigma_x=1$$

$$u \sim \text{N}(0,1) \text{ úgy, hogy } \text{Corr}(u,v)=0.5 \quad y = \beta_1 x + u$$

$$\beta_1 = 1$$

vagyis: endogén x , 2 érvényes instrumentum z_1 & z_2 , z_1 erősebb:

$$\text{Corr}(x,u) = 0.464$$

$$\text{Corr}(u,z_1) = \text{Corr}(u,z_2)=0.$$

$$\text{Corr}(z_1,x) \approx 0.25, \text{Corr}(z_2,x) \approx 0.10.$$

Heteroszkedaszticitás: $\text{Corr}(u^2, z_1^2) = \text{Corr}(u^2, z_2^2)=0.25$, amit egy autoregresszív kondicionális heteroszkedaszticitás (ARCH) modell generál:

$$u = \psi v + e$$

$$e = \varepsilon \times \sqrt{0.5 + z_1^2 + z_2^2}, \quad \varepsilon \sim iidN(0, \sigma_\varepsilon), \text{ hogy } \sigma_u = 1$$