Szigma, XXI. 1989/90. 1. sz.

CHENG HSIAO — GRANT TAYLOR

Néhány megjegyzés a mérési hiba és az identifikáció kérdéséhez panelmodellekben*

1. Bevezetés

A gazdasági mennyiségeket sokszor hibával mérjük. Továbbá, az alkalmazáskor sok esetben nincs mért megfelelője az elméleti változóknak. Az ezeket formalizáló hiba-a-változókban (errors-in-variables) modellekben a szokásos eljárások egy adott paraméter identifikációjához többlet információt igényelnek, vagy kiegészítő adatok (ismételt mintavétel és/vagy instrumentális változók), vagy további feltevések formájában. (pl. AIGNER et al. (1984)). A panelmodellek esetén azonban a hiba-a-változókban modellek különböző típusai identifikálhatók és becsülhetők külső információ nélkül is (pl. ASHENFELTER, DEATON és SOLON (1984), GRILICHES és HAUSMAN, (1986)).

Vegyük például a következő egy egyenletes modellt:

$$y_{it} = \alpha_i + \beta x_{it} + u_{it}, \qquad i = 1, \ldots, N \quad t = 1, \ldots, T, \tag{1}$$

ahol az egyedhatásokat α_i képviseli, ami lehet korrelált x_{it}-vel, vagy sem (lásd MUNDLAK (1978)), de az u_{it} hibatagokról feltételezzük, hogy függetlenek, azonos eloszlásúak és függetlenek α_i -től és x_{it}-től.

Tegyük fel, hogy x_{it} valódi értéke közvetlenül nem figyelhető meg. Megfigyelhető azonban x_{it} hibával mért mértéke:

$$z_{it} = x_{it} + \varepsilon_{it}, \tag{2}$$

ahol ε_{it} a mérési hiba. Feltételezzük továbbá, hogy ε_{it} független x_{it} -től és független, azonos eloszlású minden i -re és t -re nulla várható értékkel és σ_ε^2 szórással.

Behelyettesítve z_{it} -t x_{it} helyére azt kapjuk, hogy:

$$y_{it} = \alpha_i + \beta z_{it} + \eta_{it}, \tag{3}$$

ahol

$$\eta_{it} = u_{it} - \beta \varepsilon_{it}. \tag{4}$$

Világos, hogy a (3)-ra a legkisebb négyzetek alkalmazása nem vezet konzisztens esztimátorhoz z_{it} és y_{it}, valamint az x_i és y_i korreláltsága miatt.

---

* Cheng HSIAO és Grant TAYLOR: „Some Remarks on Measurement Errors and the Identification of Panel Data Models". *(Fordította:* NEMÉNYI Judit)

A panelmodellek egyik gyakran emlegetett kedvező tulajdonsága, hogy lehetőséget biztosít a nem megfigyelhető egyedhatások kimutatására, amellett, hogy — mint a hiba-a-változókban modellek — identifikálhatók és becsülhetők külső instrumentális változók felhasználása nélkül. Például a differencia módszerrel a specifikációból kiszűrhető a nem megfigyelhető idővariáns hatás és csökkenthető a becsült paraméter hibája.

Az időbeli első defferenciát véve a specifikációból kiküszöbölhető α_i, és a következőt kapjuk:

$$y_{it} - y_{i,t-1} = \beta(z_{it} - z_{i,t-}+[(u_{it} - \beta\varepsilon_{it}) - (u_{i,t-1} - \beta\varepsilon_{i,t-1})],$$
$$i = 1, \dots, N \; t = 2, \dots, T. \tag{5}$$

Ha ε_{it} nulla, akkor β legkisebb négyzetek becslése (OLS) konzisztens. Ha ε_{it} nullától különböző, akkor az (5) legkisebb négyzetek becslése a következő értékhez konvergál:

$$\operatorname*{plim}_{N\to\infty} \hat{\beta}_d = \beta \left(1 - \frac{2\sigma_\varepsilon^2}{\operatorname{Var}(z_{it} - z_{i,t-1})}\right). \tag{6}$$

A panelmodelleknél olyan adattranszformációkra nyílik lehetőség, amelyekből különböző, a becsült paraméter torzítottságát csökkentő, változások származnak, s így felhasználhatók a mérési hibák feltárására és a „valódi” paraméterek meghatározására. Tehát, ha x_{it} szeriálisan korrelált, világos, hogy $z_{i,t-2}$, vagy $(z_{i,t-2} - z_{i,t-3})$ felhasználható $(z_{it} - z_{i,t-1})$ instrumentális változójaként. Sőt, mivel T feltehetően véges, a kapott IV (instrumentális változó) esztimátor konzisztens, miközben N tart a végtelenbe.

Konzisztens esztimátort eredményező lehetséges megoldás, ha egy adott modell különböző adattranszformációknak kitett változóinak a torzítását hasonlítjuk össze (GRILICHES és HAUSMAN (1986)). Például, ha kovariancia-transzformációt alkalmazunk a meg nem figyelhető egyedhatás kiszűrésére, akkor

$$(y_{it} - \bar{y}_i) = \beta(z_{it} - \bar{z}_i) + [(u_{it} - \bar{u}_i) - \beta(\varepsilon_{it} - \bar{\varepsilon}_i)], \tag{7}$$

ahol

$$\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}, \; \bar{z}_i = \frac{1}{T}\sum_{t=1}^{T} z_{it}, \; \bar{u} = \frac{1}{T}\sum_{t=1}^{T} u_{it}, \; \bar{\varepsilon}_i = \frac{1}{T}\sum_{t=1}^{T} \varepsilon_{it}.$$

A (7) legkisebb négyzetek becslése a következő értékhez konvergál:

$$\operatorname*{plim}_{N\to\infty} \hat{\beta}_w = \beta \left[1 - \frac{T-1}{T} \frac{\sigma_\varepsilon^2}{\operatorname{Var}(z_{it} - \bar{z}_i)}\right]. \tag{8}$$

Ekkor a β és a σ_ε^2 konzisztens esztimátora (6) és (8) alapján:

$$\hat{\beta} = \lambda_1 \hat{\beta}_w + (1 - \lambda_1)\hat{\beta}_d, \tag{9}$$

$$\hat{\sigma}_\varepsilon^2 = \frac{\hat{\beta} - \hat{\beta}_d}{\hat{\beta}} \cdot \frac{\operatorname{Var}(z_{it} - z_{i,t-1})}{2}, \tag{10}$$

ahol

$$\lambda_1 = \left[1 - \frac{T-1}{2T} \cdot \frac{\mathrm{Var}(z_{it} - z_{i,t-1})}{\mathrm{Var}(z_{it} - \bar{z}_i)}\right]^{-1}. \tag{11}$$

Egy másik alternatíva, NICKELL (1985) javaslatára: konzisztens esztimátorhoz juthatunk a within (csoporton belüli) kereten belül maradva. Az általánosság megsértése nélkül, feltéve, hogy z_{it} a 0. időpontra létezik, $\bar{z}_{i,-1}$ -et a következőképpen írjuk fel:

$$\bar{z}_{i,-1} = \frac{1}{T}\sum_{t=0}^{T-1} z_{it}. \tag{12}$$

Ekkor $(z_{i,t-1} - z_{i,-1})$ -et instrumentális változóként használjuk (7) -hez. Amennyiben az eredményül kapott becslést β_{wiv} -val jelöljük, azt kapjuk, hogy:

$$\begin{aligned}\operatorname*{plim}_{N\to\infty} \hat{\beta}_{wiv} &= \operatorname*{plim}_{N\to\infty} \frac{\sum_{i=1}^{N}(y_{it} - \bar{y}_i)(z_{i,t-1} - \bar{z}_i)}{\sum_{i=1}^{N}\sum_{t=1}^{T}(z_{it} - \bar{z}_i)(z_{i,t-1} - \bar{z}_{i,-1})} = \\ &= \beta\left[1 + \frac{T-1}{T^2} \cdot \frac{\sigma_\epsilon^2}{\mathrm{Cov}[(z_{i,t-1} - \bar{z}_{i,-1})(z_{it} - z_i)]}\right]\end{aligned} \tag{13}$$

(13) és (8) felhasználásával meghatározható β és $\hat{\beta}_2$ konzisztens esztimátora az alábbiak szerint:

$$\hat{\beta}_2 = \lambda_2\hat{\beta}_w + (1 - \lambda_2)\hat{\beta}_{wiv}, \tag{14}$$

ahol

$$\lambda_2 = \left[1 + \frac{T\mathrm{Cov}[(z_{i,t-1} - \bar{z}_{i,-1})(z_{it} - \bar{z}_i)]}{\mathrm{Var}(z_{it} - \bar{z}_i)}\right]^{-1}. \tag{15}$$

Az eddigiekben bemutatott különböző becslések mind azt a tulajdonságot használják ki, hogy a mérési hibák és a magyarázó változók időbeli és egyedek szerinti szerkezete eltérő. Az ilyen konzisztens becsléshez vezető adattranszformációs eljárások azonban sokszor *ad hoc* jellegűek és nem feltétlenül hatékonyak. Például ezeknek a módszereknek a működéséhez szükséges, hogy a magyarázó változók szeriálisan korreláltak legyenek. Amennyiben x_{it} szeriálisan korrelálatlan, akkor (6) és (8) ugyanahhoz a $\beta\left[1 - \frac{\sigma_\epsilon^2}{\sigma_\epsilon^2 + \mathrm{Var}(x_{it})}\right]$-hoz konvergál, és (14) nem korlátos.

A cikkben megmutatjuk, hogy a likelihood megközelítés egységes keretet biztosít a hiba-a-változókban típusú feltételek identifikációjához. Ezen túlmenően, a likelihood megközelítés alkalmazásával kimutatható, hogy az általános felfogással ellentétben a paneladatok felhasználása lehetővé teszi a hiba-a-változókban modellek identifikációját és becslését anélkül, hogy követelmény lenne a magyarázó változók szeriális korreláltsága.

2. Strukturális hiba-a-változókban modellek

Legyen az S struktúra az y véletlen változó, $F(y)$ valószínűségi eloszlásfüggvényének teljes valószínűségi specifikációja. Modellnek nevezzük az összes *a priori* szóbajöhető struktúra J halmazát. Az identifikáció a struktúrák elbírálásából áll, adott J modell és y megfigyelések mellett (pl. HSIAO (1983)). A legtöbb alkalmazáskor y -ról feltételezik, hogy $F(y|\tilde{\gamma}^0)$ ismert, feltételes valószínűségi eloszlásfüggvényű folyamat generálja, de az $(m \times 1)$ méretű $\tilde{\gamma}^0$ ismeretlen. Ezáltal a struktúra megkülönböztetés problémája a paraméterpontok közötti választás problémájára redukálódik. Ebben a vizsgálati keretben vegyük a következőket:

1. definíció: Legyen Ω az $\mathbb{R}^m$ konvex, kompakt altere. Ekkor $\tilde{\gamma}^0 \in \Omega$ -ra nézve az $F(y|\tilde{\gamma}^0)$ struktúra identifikált, ha nincs más $\tilde{\gamma}^1 \in \Omega$, amire igaz, hogy $F(y|\tilde{\gamma}^0) = F(y|\gamma^1)$ minden y-ra.

Az *1. definíció* alapján a struktúra lokális identifikációjának általános feltétele, melyet ROTEMBERG (1971) dolgozott ki, a következő:

1. tétel: Legyen $\tilde{\gamma}^0 \in \Omega$ az információs mátrix egy reguláris pontja, azaz az információs mátrixnak konstans a rangja $\tilde{\gamma}$-ra $\tilde{\gamma}^0$ nyitott környezetében. Ekkor $\tilde{\gamma}^0$ lokálisan identifikálható akkor és csak akkor, ha az információs mátrix $\tilde{\gamma}^0$-nál nem szinguláris.

Az 1. tétel felhasználásával származtathatók a panelmodellek identifikálhatóságának a feltételei. Például, tegyük fel, hogy y és x valódi értékének együttes eloszlását a $\tilde{\Theta}^0(p \times 1)$ méretű vektor jellemzi. Tételezzük fel, hogy x nem megfigyelhető, de a hibával mért megfelelője igen:

$$\tilde{z} = \tilde{x} + \tilde{\epsilon} \tag{16}$$

Jelölje az $\tilde{y}$ és a $\tilde{z}$ együttes likelihoodját $f(\tilde{y}, \tilde{z}; \tilde{\gamma}^0)$, ahol $\gamma^0 = (\tilde{\Theta}^0, \tilde{\delta}^0)((p+s) \times 1)$ méretű paramétervektor. Ilyenkor szokásos eljárás az $f(\tilde{y}|\tilde{z}_i; \tilde{\theta}^0\tilde{\delta}^0)$ együttes likelihood függvényt átírni az $f(\tilde{y}|\tilde{z}; \tilde{\theta}^0, \tilde{\delta}^0)$ feltételes likelihood és az $f(\tilde{z}; \tilde{\delta}^0)$ marginális likelihood szorzatára. Ekkor a következő adódik:

$$\begin{aligned} \log f(\tilde{y}, \tilde{z}, \tilde{\theta}^0, \tilde{\delta}^0) &= \log f(\tilde{y}|\tilde{z}; \tilde{\theta}^0, \tilde{\delta}^0) + \log f(\tilde{z}; \tilde{\delta}^0) = \\ &= \log \int f(\tilde{y}|\tilde{x}; \tilde{\theta}^0) f(\tilde{x}|\tilde{z}; \tilde{\delta}^0) d\tilde{x} + \log f(\tilde{z}; \tilde{\delta}^0), \end{aligned} \tag{17}$$

ahonnan a $(\tilde{\theta}^0, \tilde{\delta}^0)$ információs mátrixa az alábbi formában írható fel:

$$I = I_{\tilde{y}|\tilde{z}}(\tilde{\theta}^0, \delta^0) + I_z(\tilde{\delta}^0). \tag{18}$$

Elemi sor és oszlop műveletekkel $I_{y|z}$ és I_z sokszor kifejezhető az alábbi formában:

$$\begin{pmatrix} \tilde{A} & \tilde{0} \\ \tilde{0} & \tilde{0} \end{pmatrix}, \tag{19}$$

és

$$\begin{pmatrix} \tilde{0} & \tilde{0} \\ \tilde{0} & \tilde{D} \end{pmatrix}, \tag{20}$$

ahol $\tilde{A}$ és $\tilde{D}(p \times p)$ és $(s \times s)$ méretű mátrixok. Az $\tilde{A}$ mátrix rangja p, ha θ^0 identifikálható és $\bar{x}$ megfigyelhető. Ebből következik, hogy a mérési hibának kitett modell akkor identifikálható, ha $\tilde{D}$ rangja s. Az ismételt megfigyelésekkel egyenértékű paneladatok elérhetősége meglehetősen általános, a magyarázó változókra és a mérési hibákra vonatkozó strukturális feltételek mellett biztosítja a modell identifikálhatóságát.

Illusztrációképpen vegyük az (1) és az (2) modellt. Azonban ahelyett, hogy GRILICHES és HAUSMAN (1986), valamint NICKELL (1985) nyomán indulófeltételünk az x_{it}-k szeriális korreláltsága lenne, feltételezzük, hogy az x_{it}-k i szerint független eloszlásúak, t szerint pedig független és normális eloszlásúak μ_i várható értékkel és $\sigma_{x_i}^2$ szórással. Az egyszerűség kedvéért azt is feltesszük, hogy ε_{it} és u_{it} függetlenek és normális eloszlásúak. Ekkor y_{it} és z_{it} együttes likelihood függvénye az alábbiak szerint írható fel:

$$
\begin{aligned}
L &= \prod_{i=1}^{N} f(y_{i1}, \ldots, y_{iT}, z_{i1}, \ldots, z_{iT}) = \\
&= \prod_{i=1}^{N} f(y_{i1}, \ldots, y_{iT} | z_{i1}, \ldots, z_{iT}) f(z_{i1}, \ldots, z_{iT}) = \\
&= \prod_{i=1}^{N} (2\pi\sigma_{y_i}^2)^{\frac{T}{2}} \\
&\exp\left\{ -\frac{1}{2\sigma_{y_i}^2} \sum_{t=1}^{T} \left[y_{it} - \alpha_i - \sigma_\varepsilon^2 \left(\sigma_{x_i}^2 + \sigma_\varepsilon\right)^{-1} \mu_i - \sigma_{x_i}^2 \left(\sigma_{x_i}^2 + \sigma_\varepsilon^2\right)^{-1} \beta z_{it} \right]^2 \right\} = \\
&= \prod_{i=1}^{N} \left[2\pi \left(\sigma_{x_i}^2 + \sigma_\varepsilon^2\right) \right]^{\frac{T}{2}} \exp\left\{ -\frac{1}{2(\sigma_{x_i}^2 + \sigma_\varepsilon^2)} \sum_{t=1}^{T} (z_{it} - \mu_i)^2 \right\},
\end{aligned}
\tag{21}
$$

ahol

$$
\sigma_{y_i}^2 = \sigma_u^2 + \beta^2 \sigma_{x_i}^2 \sigma_\varepsilon^2 (\sigma_{x_i}^2 + \sigma_\varepsilon^2)^{-1}. \tag{22}
$$

A (21) információs mátrixa az alábbi

$$
I = \begin{pmatrix} W & P \\ P' & S \end{pmatrix}, \tag{23}
$$

ahol W egy $(3N \times 3N)$ méretű blokkdiagonális mátrix, amelynek diagonálisa:

$$
W_i = E \begin{pmatrix} \frac{\partial^2 \log L}{\partial \alpha_i^2} & \frac{\partial^2 \log L}{\partial \alpha_i \partial \mu_i} & \frac{\partial^2 \log L}{\partial \alpha_i \partial \sigma_{x_i}^2} \\ & \frac{\partial^2 \log L}{\partial \mu_i^2} & \frac{\partial^2 \log L}{\partial \mu_i \partial \sigma_{x_i}^2} \\ & & \frac{\partial^2 \log L}{\partial (\sigma_{x_i}^2)^2} \end{pmatrix} = \tag{24}
$$

$$
= - \begin{pmatrix} \frac{T}{\sigma_{y_i}^2} & \frac{\sigma_\varepsilon^2 T}{\sigma_{y_i}^2 \sigma_{x_i}^2} & \frac{T}{\sigma_{y_i}^2 \sigma_{x_i}^2} [\beta \bar{z}_i - \sigma_{z_i}^{-2} a_i] \\ & \frac{\sigma_\varepsilon^4 T}{\sigma_{y_i}^2 \sigma_{x_i}^4} + \frac{T}{\sigma_{x_i}^2} & \frac{T \sigma_\varepsilon^2}{\sigma_{y_i}^2 \sigma_{x_i}^4} [\beta_t \bar{z}_i - \sigma_{z_i}^{-2} a_i] \\ & & b_i \end{pmatrix},
$$

P $(3N \times 3)$ méretű mátrix, amelynek i-edik blokkja:

$$P_i = E\begin{pmatrix} \frac{\partial^2 \log L}{\partial\alpha_i\partial\beta} & \frac{\partial^2 \log L}{\partial\alpha_i\partial\sigma_\varepsilon^2} & \frac{\partial^2 \log L}{\partial\alpha_i\partial\sigma_u^2} \\ \frac{\partial^2 \log L}{\partial\mu_i\partial\beta} & \frac{\partial^2 \log L}{\partial\mu_i\partial\sigma_\varepsilon^2} & \frac{\partial^2 \log L}{\partial\mu_i\partial\sigma_u^2} \\ \frac{\partial^2 \log L}{\partial\sigma_{z_i}^2\partial\beta} & \frac{\partial^2 \log L}{\partial\sigma_{z_i}^2\partial\sigma_\varepsilon^2} & \frac{\partial^2 \log L}{\partial\sigma_{z_i}^2\partial\sigma_u^2} \end{pmatrix} = \tag{25}$$

$$= -\begin{pmatrix} \frac{\sigma_{z_i}^2 T\bar{Z}_i}{\sigma_{y_i}^2\sigma_{z_i}^2} & \frac{T^2}{\sigma_{y_i}}\sigma_{z_i}^2[\mu_i - \sigma_{z_i}^{-2}a_i] & 0 \\ \frac{\sigma_{z_i}^2\sigma_\varepsilon^2 T\bar{Z}_i}{\sigma_{y_i}^2\sigma_{z_i}^2} & \frac{T\sigma_\varepsilon^2}{\sigma_{y_i}^2\sigma_{z_i}^2}[\beta\bar{z}_i - \sigma_{z_i}^{-2}a_i] & 0 \\ c_i & d_i & \frac{T\beta^2\sigma_\varepsilon^4}{2\sigma_{y_i}^4\sigma_{z_i}^4} \end{pmatrix},$$

és S az alábbi (3×3)-as mátrix

$$S = E\begin{pmatrix} \frac{\partial^2 \log L}{\partial\beta^2} & \frac{\partial^2 \log L}{\partial\beta\partial\sigma_\varepsilon^2} & \frac{\partial^2 \log L}{\partial\beta\partial\sigma_u^2} \\ & \frac{\partial^2 \log L}{\partial(\sigma_\varepsilon^2)^2} & \frac{\partial^2 \log L}{\partial\sigma_\varepsilon^2\partial\sigma_u^2} \\ & & \frac{\partial^2 \log L}{\partial(\sigma_u^2)^2} \end{pmatrix} = \tag{26}$$

$$= -\begin{pmatrix} \sum_{i=1}^{n}\sigma_{y_i}^{-2}\sigma_{z_i}^{-2}\sigma_{x_i}^2\sum_{t=1}^{T} z_{it}^2 & f & T\beta\sigma_\varepsilon^2\sum_{i=1}^{N}\sigma_{y_i}^{-4}\sigma_{z_i}^{-2}\sigma_{x_i}^2 \\ & g & h \\ & & \frac{T}{2}\sum_{i=1}^{N}\frac{1}{\sigma_{y_i}^4} \end{pmatrix},$$

ahol

$$\begin{aligned} \sigma_{z_i}^2 &= (\sigma_{x_i}^2 + \sigma_\varepsilon^2), \\ a_i &= (\sigma_\varepsilon^2\mu_i + \beta\sigma_{x_i}^2\bar{z}_i), \end{aligned}$$

továbbá

$$\begin{aligned} b_i &= (2\sigma_{y_i}^4)^{-1}T\beta^2\sigma_\varepsilon^2\sigma_{z_i}^{-2}[1 - \sigma_{z_i}^{-2}\sigma_{x_i}^2] + \\ &\quad + \sigma_{y_i}^{-2}\sigma_{z_i}^{-2}\sum_{t=1}^{T}[\sigma_{z_i}^{-2}(\sigma_\varepsilon^2\mu_i + \beta\sigma_{x_i}^2 z_{it}) - \beta z_{it}]^2 + T(2\sigma_{z_i}^2)^{-1} \\ c_i &= T\beta^3\sigma_\varepsilon^4\sigma_{x_i}^2\sigma_{y_i}^{-4}\sigma_{z_i}^{-4}[1 - \sigma_{z_i}^{-2}\sigma_{x_i}^2] + \\ &\quad + \sigma_{y_i}^{-2}\sigma_{z_i}^{-4}\sigma_{x_i}^2\sum_{t=1}^{T}\left[\left(\frac{\sigma_{x_i}^2}{\sigma_{z_i}^2} - 1\right)\beta z_{it} - \frac{\sigma_\varepsilon^2}{\sigma_{z_i}^2}\mu_i\right] z_{it} \\ d_i &= T\sigma_{y_i}^{-2}\beta^4\sigma_\varepsilon^2\sigma_{x_i}^2\sigma_{z_i}^{-4}[1 - \sigma_{z_i}^{-2}\sigma_{x_i}^2][1 - \sigma_{z_i}^{-2}] + \\ &\quad + \sigma_{y_i}^{-2}\sigma_{z_i}^{-8}\sum_{t=1}^{T}[(\sigma_\varepsilon^2\mu_i + \beta\sigma_{x_i}^2 z_{it}) - \sigma_{z_i}^2\mu_i][\sigma_\varepsilon^2\mu_i + \beta\sigma_{x_i}^2 z_{it}) - \sigma_{z_i}^2\beta z_{it}] + \frac{T}{2\sigma_{z_i}^4} \end{aligned}$$

$$f = T\sum_{i=1}^{N}\frac{1}{\sigma_{y_i}^4}[\beta^3\sigma_{x_i}^4\sigma_\varepsilon^2\sigma_{z_i}^{-4}(1-\sigma_{z_i}^{-2})]-$$

$$-\sum_{i=1}^{N}\frac{\sigma_{x_i}^2}{\sigma_{y_i}^2\sigma_{z_i}^4}\sum_{t=1}^{T} z_{it}[\mu_i - \sigma_{z_i}^{-2}(\sigma_\varepsilon^2\mu_i + \beta\sigma_{x_i}^2 z_{it})]$$

$$g = -\frac{T}{2}\sum_{i=1}^{N}\sigma_{y_i}^{-4}[\beta^2\sigma_{x_i}^2\sigma_{z_i}^{-2}(1-\sigma_{z_i}^{-2})]^2-$$

$$-\sum_{i=1}^{N}\frac{1}{\sigma_{y_i}^2}\sum_{t=1}^{T}[\sigma_{z_i}^{-2}\mu_i - \sigma_{z_i}^{-4}(\sigma_\varepsilon^2\mu_i + \beta\sigma_{x_i}^2 z_{it})]^2 + \frac{T}{2}\sum_{i=1}^{N}\sigma_{z_i}^{-4}$$

és végül

$$h = \frac{T\beta^2}{2}\sum_{i=1}^{N}\frac{\sigma_{x_i}^2}{\sigma_{y_i}^2\sigma_{z_i}^2}(1-\sigma_{z_i}^{-2}).$$

Könnyű belátni, hogy (23) elemi mátrixműveletekkel diagonális mátrix-szá transzformálható, melynek rangja $(3N+3)\times(3N+3)$. Paneladatok felhasználásával tehát (1) és (2) minden paramétere identifikálható akkor is, ha x_{it} szeriálisan korrelálatlan.

A fentiek általánosítása könnyen elvégezhető olyan lineáris modellekre is, amelyek több, hibával mért magyarázó változót tartalmaznak. Az eljárás hasonló akkor is, ha az x_{it}-k r-ed rendű autoagresszív, mozgó átlagolású, vagy bármely típusú autoregresszív (integrált) és mozgó átlagolású folyamatot követnek.

3. Következtetések

A paneladatok felhasználásának egyik előnye, hogy lehetővé teszik a nem megfigyelhető egyedi jellemzők feltárását. (pl. HSIAO (1985,86)). Azonban a mérési hiba paneladatoknál komoly zavart okozhat. Például az Institute for Social Research, a Michigan Egyetemen készített egy panelfelmérést egy olyan vállalat alkalmazottai körében, ahol már pontos adatokkal rendelkeztek a keresetekre és a foglalkoztatásra. DUNCAN és HILL (1985) beszámol arról, hogy az éves keresetekre és a ledolgozott órákra vonatkozó felmérés alapján kapott órabérek mérési hibája az összes szórás 0.7-szerese volt. Szerencsére a többi változót nagyobb megbízhatósággal mérték, de a bérváltozó kirívó megbízhatatlansága „vészjelző". Tekintettel a panelmódszerek nagyfokú érzékenységére a mérési hibákra, a kutatóknak különös gonddal kell eljárniuk adataik mérési hibájának felderítésekor. Ebben a cikkben bemutattuk, hogy a nem megfigyelhető egyedhatások és a magyarázó változók esetleges korrelációja miatti sok aggodalom ellenére — mely a panelmodelleknél gyakran előfordul — a likelihood megközelítéssel egységes keret biztosítható a mérési hibának kitett panelmodellek identifikációjához. Valójában, ez hatásos becslési módszerek alapjául is szolgál.

Hivatkozások:

AIGNER, D. J., C. HSIAO, A. KAPTEYN and T. WANSBEEK (1985): „Latent Variable Models in Econometrics", in *Handbook of Econometrics*, vol. II., ed. by Z. GRILICHES and M. D. IINTRILIGATOR, Amsterdam: North-Holland, 1322-1393.

ASHENFELTER, O., A. DEATON and G. SOLON (1984): „Does it Make Sense to Collect Panel Data in Developing Countries?" Mimeo.

DUNCAN, G. J., and D. H. HILL (1985): „An Investigation of the Extent and Consequences of Measurement Error in Labor-economic Survey Data", *Journal of Labor Economics*, 3., 508-532.

GRILICHES, Z. and J. A. HAUSMAN (1986): „Errors-in-Variables in Panel Data", *Journal of Econometrics*, 31., 93-118.

HSIAO, C. (1984): „Identification", in *Handbook of Econometrics*, vol. I., ed. by Z. GRILICHES and M. D. IINTRILIGATOR, Amsterdam: North-Holland, 1322-1393.

HSIAO, C. (1985): „Benefits and Limitations of Panel Data", *Econometrics Reviews*, 4., 121-174.

HSIAO, C. (1986): *Analysis of Panel Data*, Cambridge, Cambridge University Press.

MUNDLAK, Y. (1978): „On the Pooling of Time Series and Cross Section Data", *Econometrica*, 46., 69-85.

NICKELL, S. (1985): „Comment on Benefits and Limitations of Panel Data", *Econometric Reviews*, 4., 179-182.