6.5 STATA コード

ここでは『21 世紀出生児縦断調査』を用いてダイナミックパネル推定を行ってみよう。これまで北村 (2007、2008, 2009) で行ってきた出生児の身体成長パターンの測定を行ってきた。これまでの研究ではダイナミックパネル推定ではなく、身長・体重の対数を出生からの経過日数とその二乗、および子育費用によって説明する固定効果モデルを採用してきた。

ここではこれまでのモデルに一次のラグ項を入れている。例えば体重については次のようなモデルを考えている。

$$\begin{aligned}\ln bdywht_{it} &= \alpha + \gamma \ln bdwht_{it-1} + \beta \ln survivalday_{it} + \delta \ln survaivalday_{it}^2 \\ &+ \zeta \ln \cos tofchildcare_{it} + u_{it}\end{aligned}$$

ここで $\ln bdywht =$ 体重の対数値、$\ln survivalday =$ 生存日数の対数値、$\ln \cos tofchildcare =$ 子育費用の対数値。身長の場合は $\ln bdywht$ に代えて $\ln bdyhgt$ を用いる。

推計に使った STATA コードは次のようになる。

```
/*dynamic panel analysis*/
tsset idno rirekino
/**weight**/
/*logarithm*/
reg lnbdywht lnbdywht_1 survivalday survival2 lncostofchildcare  /*OLS*/
xtreg lnbdywht lnbdywht_1 survivalday survival2 lncostofchildcare, mle /*Maximum
Likelihood*/

xtabond lnbdywht survivalday survival2 lncostofchildcare, lags(1) pre( lnbdyhgt_1)
artests(2) vce(gmm) /*Arellano and Bond GMM*/
estat sargan

xtabond lnbdywht survivalday survival2 lncostofchildcare, lags(1) pre( lnbdyhgt_1)
artests(2) vce(robust) /*Arellano and Bond GMM*/
estat abond
/**hight**/
/*logarithm*/
reg lnbdyhgt lnbdyhgt_1 survivalday survival2 lncostofchildcare /*OLS*/
xtreg lnbdyhgt lnbdyhgt_1 survivalday survival2 lncostofchildcare, mle /*Maximum
Likelihood*/
xtabond lnbdyhgt survivalday survival2 lncostofchildcare, lags(1) pre(lnbdywht_1)
artests(2) vce(gmm) /*Arellano and Bond GMM*/
```

```
estat sargan
xtabond lnbdyhgt survivalday survival2 lncostofchildcare, lags(1) pre(lnbdywht_1)
artests(2) vce(robust) /*Arellano and Bond GMM*/
estat abond
```

Dependent Variable: weight	OLS		MLE		GMM	
	Coef.	t-statistics	Coef.	z-statistics	Coef.	z-statistics
llnbdywht L1.					-0.033	-7.02
lnbdywht_1	0.246	231.58	0.088	133.71	0.261	20.06
survivalday	0.000	-56.75	0.000	114.60	0.000	100.14
survival2	1.24E-07	122.16	7.21E-09	11.78	-6.44E-09	-9.52
lncostofchildcare	0.013	35.14	0.003	12.73	0.000	-0.66
_cons	2.122	1,582.70	2.093	2,377.90	1.200	27.04
/sigma_u			0.098			
/sigma_e			0.057			
rho			0.750			
Diagnostic Test						
Number of observation	179,479		179,479		116,725	
Number of groups	43,169		43,169		36,854	
R-squared	0.841					
Adj R-squared	0.841					
Root MSE	0.107					
LR chi2(3)			450,209.24			
Sargan test					chi2(28) = 1584.489 Prob > chi2 = 0.0000	
Wald test					Chi2(4) = 314446.08 Prob > chi2 = 0.0000	
Arellano-Bond test for residual AR(1)					z = -24.467 Prob > z = 0.0000	
Arellano-Bond test for residual AR(2)					z = 8.675 Prob > z = 0.0000	
Instruments for differenced equation						
GMM-type:					L(2/.).lnbdywht L(1/.).lnbdyhght_1	
Standard:					D.survivalday D. survival2	

図表 6.1　ダイナミックパネル推定（体重）

図表 6.1 はプーリング OLS 推定、最尤法（M L E）推定, Arellano-Bond GMM 推定の結果を報告してある。GMM 推定については Sargan の過剰識別制約検定を行ったが、これが有効であるという帰無仮説は棄却された。すなわち、このモデルでは操作変数が直行条件を満たしていないか、外生変数が本当は内生変数であるか、モデルの重要な説明変数が落ちているかなどの問題があることがわかる。また Arellano-Bond の系列相関検定でも、系列相関が AR(1) でも AR(2) でも残っていることから、操作変数の直行条件が満たされていない可能性が示唆されている。

ここでは体重のラグ項のパラメータ γ がどのような値をとるかに関心がある。パラメータの順序は次のようになる。

$$MLE(0.088) < OLS(0.246) < GMM(0.261)$$

GMM 推定は先に述べたように問題が多いが、推定値は OLS とほぼ等しい。

それに対して図表 6.2 では身長を推定した。GMM 推定の統計的問題点は全く同じであるが、γ の値は次のような順序になる。

Dependent Variable: hight	OLS		MLE		GMM	
	Coef.	t-statistics	Coef.	z-statistics	Coef.	z-statistics
llnbdywht L1.					0.085	13.35
lnbdywht_1	0.213	196.5	0.079	103.98	−0.008	−3.50
survivalday	5.09E−05	32.81	0.000	206.8	0.000	155.71
survival2	2.19E−08	55.72	−1.75E−08	−65.38	−2.13E−08	−64.56
lncostofchildcare	0.003	22.37	0.000	−4.49	−0.001	−5.62
_cons	3.510	959.77	3.960	1553.86	3.937	180.96
/sigma_u			0.033			
/sigma_e			0.023			
rho			0.674			
Diagnostic Test						
Number of observation	170,124		170124		113742	
Number of groups			42528		35748	
R−squared	0.925					
Adj R−squared	0.925					
Root MSE	0.038					
LR chi2(3)			524,682.810			
Sargan test					chi2(28) = 993.153 Prob > chi2 = 0.0000	
Wald test					Chi2(4) = 442673.89 Prob > chi2 = 0.0000	
Arellano−Bond test for residual AR(1)					z = −37.335 Prob > z = 0.0000	
Arellano−Bond test for residual AR(2)					z = 4.3317 Prob > z = 0.0000	
Instruments for differenced equation						
GMM−type:					L(2/.).lnbdywht L(1/.).lnbdyhght_1	
Standard:					D.survivalday D. survival2	

図表 6.2　ダイナミックパネル推定（身長）

$$GMM(-0.008) < MLE(0.079) < OLS(0.213)$$

この場合は GMM 推定の値は最低になり、OLS 推定が最大値を取っている。この二つの推定では MLE 推定が 0.08 程度、OLS 推定が 0.2 程度と安定している。どちらかを選ぶとすれば、経験的に考えて、OLS 推定がここでは選ばれていいのではないだろうか。

しかし、我々の判断としてはここで用いた $\ln\cos tofchildcare$ 変数が内生変数である可能性が高いと思う、そこで次章ではこの問題について取り組みたい。

参考文献

阿部修人・小黒曜子（2004）「社長交代と外部出身取締役−− Semi Parametric 推計による分析」、『経済研究』、55(1)、pp.72-84.

知久馬　（1995）『生存時間解析　SAS による生物統計』、東大出版会

北村行伸 (2005) 『パネルデータ分析』、岩波書店

北村行伸 (2009) 『ミクロ計量経済学入門』、日本評論社

北村行伸（2007）「21 世紀成年者縦断調査に基づく子供の成長パターンの測定」、厚生労働科学研究費補助金・統計情報高度利用総合研究事業平成 18 年度報告書『パネル調査（縦断調査）に関する総合的分析システムの開発研究』、2007 年 3 月、pp.101-123.

北村行伸 (2008) 「21 世紀出生児縦断調査に基づく子供の成長パターンの測定（II)」、厚生労働科学研究費補助金・統計情報総合研究事業　平成 19 年度　総括研究報告書『パネル調査（縦断調査）に関する総合的分析システムの開発研究』、2008 年 3 月、pp.71-89.

北村行伸 (2009) 「21 世紀出生児縦断調査に基づく子供の成長パターンの測定（III)」、厚生労働科学研究費補助金・統計情報総合研究事業　平成 20 年度　総括研究報告書『パネル調査（縦断調査）に関する統合的高度統計分析システムの開発研究』、2009 年 3 月、pp.127-151.

中村　剛（2001）『Cox 比例ハザードモデル』、医学統計学シリーズ、朝倉書店

山口一男（2001-2)「イベントヒストリー分析 (1)-(15)」、『統計』2001 年 9 月号－ 2002 年 11 月号

Anderson, T.W. and Hsiao, C. (1981) "Estimation of Dynamic Models with Error Components," *Journal of the American Statistical Association*, 76, pp.598-606.

Anderson, T.W. and Hsiao, C. (1982) "Formulation and Estimation of Dynamic Models Using Panel Data," *Journal of Econometrics*, 18, pp.47-82.

Arellano, M. (1989) "A Note on the Anderson-Hsiao Estimator for Panel Data," *Economics Letters*, 31, pp.337-341.

Arellano, M. (2003) *Panel Data Econometrics*, Oxford University Press.

Arellano, M. and Bond, S. (1991) "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58, pp.277-297.

Balestra, P. and Nerlove, M. (1966) "Pooling Cross-Section and time-Series Data in the Estimation of a Dynamic Model: the Demand for Natural Gas," *Econometrica*, 34, pp.585-612.

Baltagi, B.H. (2001) *Econometric Analysis of Panel Data*, 2nd ed, John Wiley & Sons.

Baltagi, B.H. and Li, Q. (1991a) "A Transformation that Will Circumvent the Problem of Autocorrelation in an Error Component Model," *Journal of Econometrics*, 48, pp.385-393.

Blundell, R. and Bond, S. (1998) "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87, pp.115-143.

Chamberlain, G. (1982) "Multivariate Regression Models for Panel Data," *Journal of Econometrics*, 18, pp.5-46.

Chamberlain, G. (1984) "Panel Data, in Z. Griliches and M. Intrilligator (eds.), *Handbook of Econometrics*, North-Holland, Vol.2., Chapter 22, pp.1247-1318.

Fisher, R.A.(1973a) *Statistical Methods for Research Workers*, 14th ed, Hafner Publishing.

Fujiki, H., Hsiao, C. and Shen, Y. (2002) "Is There a Stable Money Demand Function under the Low Interest Rate Policy? A Panel Data Analysis," *Monetary and Economic Studies*, 20(2), pp.1-23.

Holtz-Eakin, D.(1988) "Testing for Individual Effects in Autoregressive Models", *Journal of Econometrics*, 39, pp.297-307.

Holtz-Eakin, D., Newey, W. and Rosen, H.S. (1988) "Estimating Vector Autoregressions with Panel Data, " *Econometrica*, 56, pp.1371-1395.

Hsiao, C. (2003) *Analysis of Panel Data 2nd ed.*, Cambridge University Press.

Hsiao, C. and Pesaran, M.H. (2004) "Random Coefficient Panel Data Models," mimeo.

Lillard, L.A. and Weiss, Y. (1979) "Components of Variation in Panel Earnings Data: American Scientists 1960-70," *Econometrica*, 47(2), pp.437-454.

Lillard, L.A. and Willis, R.J.(1978) "Dynamic Aspects of Earning Mobility," *Econometrica*, 46(5), pp.985-1012.

Maddala, G.S. (1971a) "The Use of Variance Components Models in Pooling Cross-Section and Time-Series Data," *Econometrica* 39, pp.341-358.

Maddala, G.S. (1971b) "The Likelihood Approach to Pooling Cross-Section and time Series Data," *Econometrica*, 39, pp.939-53.

Maddala, G.S. (2001) *Introduction to Econometrics*, 3rd ed, John Weily & Sons.

Maddala, G.S., Li, H., Trost, R.P. and Joutz, F. (1997) "Estimation of Short-Run and Long-Run Elasticities of Energy Demand from Panel Data Using Shrinkage Estimators," *Journal of Business and Econmic Statistics*, 15, pp.90-100.

Nelson, C.R and Startz, R. (1990a) "The Distribution of the Instrumental Variables Estimator and Its t-Ratio When the Instrument is a Poor One," *Journal of Business*, 63(1), pp.X125-S140.

Nelson, C.R and Startz, R. (1990b) "Some Further Results on the Exact Small Sample Properties of the Instrumental Variable Estimator," *Econometrica* 58 (4), pp.967-976.

Nerlove, M. (2000) "Growth Rate Convergence, Fact or Artifact? An Essay onPanel Data

Econometrics, in Krishnakumar, J. and Ronchetti, E. (eds) *Panel Data Econometrics: Future Directions*, North Holland, Chapter 1, pp.3-33.

Nerlove, M. (2002) *Essays in Panel Data Econometrics*, The Cambridge University Press.

Nickell, S. (1981) "Biases in Dynamic Models with Fixed Effects, *Econometrica*, 49, pp.1417-1426.

Ridder, G. and Wansbeek, T. (1990) "Dynamic Models for Panel Data," in Von der Plog, R (ed.) *Advanced Lectures in Quantitative Economics*, Academic Press, pp.557-582.

Sevestre, P. and Trognon, A. (1996) "Dynamic Linear Models," in Mátyás, L. and Sevestre, P. (eds.), *The Econometrics of Panel Data*, Chapter 7, pp.120-144, Kluwer Academic Publishers.

Smith, R.P. and Fuertes, A-M. (2004) "Panel Time-Series," mimeo.

Staiger, D. and Stock, J. (1997) "Instrumental Variables Estimation with Weak Instruments," *Econometrica*, 65, pp.557-586.

Trognon, A. (1978) "Miscellaneous Asymptotic Properties of Ordinary Least Square and Maximum Likelihood Methods in Dynamic Error Components Models," *Annales de L'INSEE*, 30-31, pp.631-357.

Vella, F. and Verbeek, M. (1998) "Whose Wages Do Unions Raise? A Dynamic Model of Unionism and Wage Rate Determination for Young Men," *Journal of Applied Econometrics*, 13, pp.163-183.

Wooldridge, J.M. (2003) *Introductory Econometrics: A Modern Approach*, 2nd edition, South-Western, Thomson Learning.

第 7 章

同時方程式パネルデータ分析

7.1 はじめに

多くの経済変数は相互依存関係にあり、完全に外生的に決まっている変数はむしろ珍しいといってもいいかもしれない。もちろん、これは人間の経済活動自体が相互依存関係の中で決まっているという事実を勘案すれば当然のことであろう[*1]。

しかし、現実問題として、我々は実証研究で、同じ変数をある時は被説明（内生）変数に使い、ある時は説明（外生）変数に使ってきた。例えば、企業収益は企業活動の目的であると考えれば、これは被説明（内生）変数として扱うべきであるが、投資関数の推計においては企業収益を説明（外生）変数に使うことがある。同様に、賃金も被説明（内生）変数として使ったり、労働供給関数では説明（外生）変数として使うことがある。

変数が内生的であるという議論をする場合には、暗黙の内に同時方程式あるいは経済システムを考えていることが多い。経済活動の相互依存関係を捉えるという意味で一般均衡モデルを用いることが重要であるということは、多くの経済学者に認められてきたことであるし、経済政策分析においては益々その重要性が増してきているとも言える。実証研究においても、内生変数の取り扱いに注意して、それがもたらすバイアスを適切に取り除く必要があるということは広く認識されてきている。本章でもその手法について解説することを主眼としている。

同時に 1970 年代後半以後、大型計量経済モデルに対する批判が相次ぎ、同時方程式の研究自体も少なくなってきた。同時方程式パネルデータ分析も Baltagi (1981b)、Hausman and Taylor (1981) などの 1980 年代の研究から大きな進展がなく、Hsiao(2003) においても、旧版である Hisao(1986) から唯一変更がなかったのが同時方程式の第 5 章である。

本章でも同時方程式の相互依存関係を本格的に取り込んだ完全情報 3 段階最小二乗法（3SLS）や完全情報最小距離推定法などについては、簡単な解説にとどめ、より実用的な 2 段階最小二乗法 (2SLS) による単一方程式に議論を絞っている。また、同時方程式といっても全ての変数が相互依存関係にあるのではなく、一部の変数が影響を与えあっており、他の変数は外生変数と扱っても問題がない場合が往々にしてある。この特殊ケースとして三角配列システムがあるが、これについて

[*1] 本章は北村（2005）『パネルデータ分析』（岩波書店）の第 6 章に基づいている。詳細については北村（2005）第 6 章を参照されたい。実証分析は新たに『21 世紀出生児縦断調査』を用いて行っている。

はここでは論じない[*2]。

7.2 同時方程式パネルデータ分析の考え方

同時方程式パネルデータとは、次のような構造をもっている。

$$
\text{主体 } i \begin{cases} \boxed{1} & \boxed{Y_{1i} = \beta_{1i} X_{1i} + v_{1it}} \\ 2 & \vdots \\ \vdots & \vdots \\ g & \vdots \\ \vdots & \vdots \\ G & Y_{Gi} = \beta_{Gi} X_{Gi} + v_{Git} \end{cases}
$$

$$
\text{主体 } j \begin{cases} \boxed{1} & \boxed{Y_{1j} = \beta_{1j} X_{1j} + v_{1jt}} \\ 2 & \vdots \\ \vdots & \vdots \\ g & \vdots \\ \vdots & \vdots \\ G & Y_{Gj} = \beta_{Gj} X_{Gj} + v_{Gjt} \end{cases}
$$

例えば $\hat{\beta}_1$ を推計したい場合、

1. 各主体はクロスセクションで Y_{1i} を選択すると同時に $Y_{gi}\ (g = 1, \ldots, G)$ を選択している。
2. これを t 時点毎に行っている。
3. 一般の同時方程式問題では、各 i について各時点共通の β を推計している。
4. この β_g が異なる主体 i, j で共通しているとすると、誤差項の構造に違いに配慮した推計をする必要が出てくる (固定効果 A_i を用いる)。
5. 一般のパネル推計では時系列方向の平均 $\bar{Y} = \beta \bar{X}$ を用いたが、同時方程式パネル推計では、個別主体の同時方程式制約下での β の推定量と主体間のパネル推定量の差を最小にするような推定が有効かつ一致するように求めることが問題となる。

一般に、同時方程式体系を考えたとしても、パネルデータ分析の枠組みでは、その中から、関心の高い一本の方程式を選んで、それについて内生性の問題に配慮した推計方法を考えることが多

[*2] 三角配列システムを推計する方法は Chamberlain(1977) によって操作変数法が、Chamberlain and Griliches (1975) によって最尤法が提案されている。最尤法は有効推定ではあるが計算が難しい。操作変数法は有効ではないが、計算が簡単であり、一致推定は得られる。詳細については Hsiao(2003, chapter 5), pp.129-136 を参照されたい。

い[*3]。

$$\underset{NT\times 1}{\mathbf{y}_g} = \mathbf{Y}_g\gamma_g + \mathbf{X}_g\mathbf{B}_g + \mathbf{v}_g \tag{7.1}$$
$$= \mathbf{w}_g\theta_g + \mathbf{v}_g \quad g = 1,\ldots,G$$

しかし、この方法は基本的には制限情報推定法 (limited-information estimation method) に従っており、同時方程式体系全体から得られる情報を完全に反映した推計 (完全情報推定法：full-information estimation method) を行えば、より有効な推定量を得られる。

ここで

$$\mathbf{y} = (\mathbf{y}\prime,\ldots,\mathbf{y}'_G)', \quad \mathbf{v} = (\mathbf{v}'_1,\ldots,\mathbf{v}'_G)'$$
$$\mathbf{w} = \begin{bmatrix} \mathbf{w}_1 & \mathbf{o} & \cdots & \mathbf{o} \\ \mathbf{o} & \mathbf{w}_2 & & \mathbf{o} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{o} & \mathbf{o} & \cdots & \mathbf{w}_G \end{bmatrix}, \quad \theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_G \end{bmatrix}$$

G 本の構造方程式体系は次のように表わすことができる。

$$\mathbf{y} = \mathbf{w}\theta + \mathbf{v} \tag{7.2}$$

(7.2) 式を 3 段階最小 2 乗法（3SLS）で推計して有効なのは誤差項（$v_{1_{it}}, v_{2_{it}}, \ldots v_{G_{it}}$）が i と t に関して iid の場合のみである。不均一分散や系列相関がある場合には、完全情報最小距離推定 (the full information minimum-distance estimator) か一般化 3 段階最小 2 乗法（G3SLS) を用いる必要がある。

完全情報最小距離推定は次のように定式化される。T は固定、N は無限であるとすると、T 期の推計式に対して個人 i の G 本の構造方程式を考える。

$$\begin{aligned} \mathbf{y}_{1_i} &= \mathbf{w}_{1_i}\theta_1 + \mathbf{v}_{1i} \\ \mathbf{y}_{2_i} &= \mathbf{w}_{2_i}\theta_2 + \mathbf{v}_{2i} \\ &\vdots \\ \mathbf{y}_{G_i} &= \mathbf{w}_{G_i}\theta_G + \mathbf{v}_{G_i} \quad i = 1,\ldots N \end{aligned} \tag{7.3}$$

最小距離推定値 θ は次の目的 (距離) 関数を最小化するように $\hat{\theta}$ を選ぶことによって得られる。

$$\left[\widehat{\tilde{\pi}} - \tilde{\mathbf{f}}(\theta)\right] \widehat{\tilde{\Omega}}^{-1} \left[\widehat{\tilde{\pi}} - \tilde{\mathbf{f}}(\theta)\right]$$

ここで $\widehat{\tilde{\pi}}$ は $\mathbf{y}_i$ の無制約の最小 2 乗法推定量、$\widehat{\tilde{\mathbf{\Omega}}}$ は $\widehat{\tilde{\pi}}$ と個人 i における $\tilde{\pi}$ の差の分散共分散行列の一致推定量 ($\tilde{\mathbf{\Omega}} = \sqrt{N}(\widehat{\tilde{\pi}} - \tilde{\pi})$)。もし $\mathbf{v}_i$ が iid であれば $\sqrt{N}(\hat{\theta} - \theta)$ は漸近的に正規分布となる。

G3SLS は次のように求められる。

$$\hat{\theta}_{G3SLS} = (S_{wx}\hat{\Psi}^{-1}S'_{wx})^{-1}(S_{wx}\hat{\Psi}^{-1}S_{xy}) \tag{7.4}$$

[*3] 以下の議論は Hsiao (2003,pp.124-6) を参照している。

ここで

$$S_{wx} = \begin{bmatrix} \tilde{S}_{w1x} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \tilde{S}_{w2x} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \tilde{S}_{wGx} \end{bmatrix},$$

$$\tilde{S}_{w_g\mathbf{x}} = \left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{w}_{gi1}\mathbf{x}'_{i1}, \frac{1}{N}\sum_{i=1}^{N}\mathbf{w}_{gi2}\mathbf{x}'_{i2}, \ldots \frac{1}{N}\sum_{i=1}^{N}\mathbf{w}_{giT}\mathbf{x}'_{it}\right],$$

$$S_{xy} = \begin{bmatrix} \mathbf{S}_{xy1} \\ \mathbf{S}_{xy2} \\ \vdots \\ \mathbf{S}_{xy\alpha} \end{bmatrix},$$

$$\underset{TK\times 1}{S_{xy_g}} = \begin{bmatrix} \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_{i1}y_{gi1} \\ \vdots \\ \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_{iT}y_{git} \end{bmatrix}$$

G3SLS は漸近的に最小距離推定に一致する。G3SLS も 3SLS も一致推定量ではあるが、誤差項の分散共分散に誤差構成要素が入っていると有効推定ではなくなる。

7.3 単一方程式推定

前節では同時方程式パネルデータ分析の考え方を示し、同時方程式体系全体の情報を有効に使うことで、はじめて有効一致推定を得ることができると論じた。しかし、現実的に複数の連立方程式に複雑な誤差構成要素を取り込んで計算することは極めてやっかいなことであり、現実のパネルデータにおける同時方程式の推計は次のように簡便化した2段階で行っている。

1. 観察不可能な（latent variable）効果を一階の階差を取って消去する。
2. 内生変数に対して操作変数を見つけて 2SLS 推計する。この場合、操作変数は時間とともに変化する変数を用いる。

次のようなモデルを考えよう。

$$y_{it1} = \alpha_1 y_{it2} + z_{it1}\beta_1 + \alpha_{i1} + v_{it1} \tag{7.5}$$
$$y_{it2} = \alpha_2 y_{it1} + z_{it2}\beta_2 + \alpha_{i2} + v_{it2} \tag{7.6}$$

ここで z_{it1}, z_{it2} は外生変数、一般モデルでは固定効果 α_{i1} と α_{i2} は全ての説明変数と相関している。誤差項 v_{it1} と v_{it2} は z とは無相関である。y_{it2} は v_{it1} と、y_{it1} は v_{it2} と相関している。

(7.5) 式を推計する場合、$\alpha_{i1} + v_{it1}$ は全ての説明変数と相関しているので OLS 推計は不適切である。そこで α_{i1} を階差を取って消去し、プーリング 2SLS で推計する。

$$\triangle y_{it1} = \alpha_1 \triangle y_{it2} + \triangle z_{it1}\beta_1 + \triangle v_{it1} \tag{7.7}$$

この場合、誤差項 $\triangle v_{it1}$ は $\triangle z_{it1}$ とは無相関となる。

しかし $\triangle y_{it2}$ と $\triangle v_{it1}$ は相関している可能性があり、$\triangle y_{it2}$ に対して操作変数をあてがう必要がある。一般には z_{it2} に含まれていて z_{it1} に含まれていない変数であり、かつ時間とともに変化する変数を用いる。

このようにして 2 段階最小二乗法あるいは操作変数法によってパラメータ β_1 をなるべくバイアスを少なくするように推定するというのが定石である。ここでの特徴は操作変数を外から探してくるのではなく、すでに同時方程式体系の中に含まれている外生変数を操作変数として用いるということである。問題は誤差項が一つの確率変数で表わされているのではなく、誤差構成要素が複数あり、それぞれの誤差分布に配慮しなければならないということである。

この問題をより厳密にするために、次のような同時方程式モデルを考えよう[*4]。

$$y_1 = Z_1\delta_1 + u_1 \tag{7.8}$$

ここで $Z_1 = [Y_1, X_1]$, $\delta_1' = (\gamma_1', \beta_1')$

Y_1 は要素 g_1 の内生変数、X_1 は要素 k_1 の外生変数、$X = [X_1, X_2]$ は同時方程式体系共通の外生変数である。この方程式は X_2 の式から除外されている外生変数の数 k_2 が $g_1 - 1$ と同数かそれより大きいときに識別できる[*5]。

誤差構成要素を次のように仮定する。

$$u_1 = Z_\mu \mu_1 + v_1 \tag{7.9}$$

ここで $Z_\mu = (I_N \otimes \iota_T)$, $\mu_1' = (\mu_{11}, \ldots, \mu_{N1})$ と $v_1' = (v_{111}, \ldots, v_{NT1})$ は平均ゼロの確率変数である。

$$E\begin{pmatrix}\mu_1\\ v_1\end{pmatrix}(\mu_1', v_1') = \begin{bmatrix} \sigma^2_{\mu_{11}} I_N & 0 \\ 0 & \sigma^2_{v_{11}} I_{NT} \end{bmatrix} \tag{7.10}$$

(7.8) は次のように変換できる。$Q = I_{NT} - P$, $P = I_N \otimes \bar{J}_T$,

$$Qy_1 = QZ_1\delta_1 + Qu_1 \tag{7.11}$$

$\tilde{y}_1 = Qy_1$, $\tilde{Z}_1 = QZ_1$ として (7.11) 式を 2SLS 推計する。その際 $\tilde{X} = QX$ を操作変数として用いる。

[*4] 以下の議論は Baltagi(2001、pp.111-15) に依拠している。

[*5] 識別のための必要条件は、モデル全体に含まれる外生（先決）変数 K から当該方程式に含まれる外生（先決）変数の数 k の差が、当該方程式に含まれる内生変数の数 g との間に次のような関係を持つことである。$K - k \geqq g - 1$

Within2SLS 推計は

$$\tilde{\delta}_{1with2SLS} = (\tilde{Z}_1' P_{\tilde{X}} \tilde{Z}_1)^{-1} \tilde{Z}_1' P_{\tilde{X}} \tilde{y}_1 \tag{7.12}$$
$$var(\tilde{\delta}_{1with2SLS}) = \sigma^2_{v_{11}} (\tilde{Z}_1' P_{\tilde{X}} \tilde{Z}_1)^{-1}$$

Within2SLS は次の式を GLS 推計することによっても導出可能である。

$$\tilde{X}'\tilde{y}_1 = \tilde{X}\tilde{Z}_1\delta_1 + \tilde{X}'\tilde{u}_1 \tag{7.13}$$

$\bar{y}_1 = Py_1, \bar{Z}_1 = PZ_1$ とおき $\bar{X} = PX$ を操作変数として (7.5) 式を 2SLS 推計すると Between2SLS 推計が得られる。

$$\hat{\delta}_{1btw2SLS} = (\bar{Z}_1' P_{\bar{x}} \bar{Z}_1)_1^{-1} \bar{Z}' P_{\bar{x}} \bar{y}_1 \tag{7.14}$$
$$var(\hat{\delta}_{1btw2SLS}) = \sigma^2_{1_{11}} (\bar{Z}_1' P_{\bar{x}} \bar{Z}_1)^{-1}$$
$$\sigma^2_{1_{11}} = T\sigma^2_{\mu_{11}} + \sigma^2_{v_{11}}$$

Between2SLS は GLS 推計としても導出可能である。

$$\bar{X}'\bar{y}_1 = \bar{X}\bar{Z}_1\delta_1 + \bar{X}'\bar{u}_1 \tag{7.15}$$

(7.13) と (7.15) を同時方程式として扱う。

$$\begin{pmatrix} \tilde{X}'\tilde{y}_1 \\ \bar{X}'\bar{y}_1 \end{pmatrix} = \begin{pmatrix} \tilde{X}'\tilde{Z}_1 \\ \bar{X}'\bar{Z}_1 \end{pmatrix} \delta_1 + \begin{pmatrix} \tilde{X}'\tilde{u}_1 \\ \bar{X}'\bar{u}_1 \end{pmatrix} \tag{7.16}$$

ここで

$$E\begin{pmatrix} \tilde{X}'\tilde{u}_1 \\ \bar{X}'\bar{u}_1 \end{pmatrix} = 0, \quad var\begin{pmatrix} \tilde{X}'\tilde{u}_1 \\ \bar{X}'\bar{u}_1 \end{pmatrix} = \begin{bmatrix} \sigma^2_{v_{11}}\tilde{X}'\tilde{X} & 0 \\ 0 & \sigma^2_{1_{11}}\bar{X}'\bar{X} \end{bmatrix}$$

(7.16) 式を GLS 推計すると the error component two-stage least squares (EC2SLS) を得る。

$$\hat{\delta}_{1,EC2SLS} = \left[\frac{\tilde{Z}_1' P_{\tilde{X}} \tilde{Z}_1}{\sigma^2_{v_{11}}} + \frac{\bar{Z}_1' P_{\bar{X}} \bar{Z}_1}{\sigma^2_{1_{11}}}\right]^{-1} \left[\frac{\tilde{Z}_1' P_{\tilde{X}} \tilde{y}_1}{\sigma^2_{v_{11}}} + \frac{\bar{Z}_1' P_{\bar{X}} \bar{y}_1}{\sigma^2_{1_{11}}}\right] \tag{7.17}$$

ここで

$$\hat{\delta}_{1EC2SLS} = w_1\hat{\delta}_{1with2SLS} + w_2\hat{\delta}_{1btw2SLS}$$
$$w_1 = \left[\frac{\tilde{Z}_1' P_{\tilde{X}} \tilde{Z}_1}{\sigma^2_{v_{11}}} + \frac{\bar{Z}_1' P_{\bar{X}} \bar{Z}_1}{\sigma^2_{1_{11}}}\right]^{-1} \left[\frac{\tilde{Z}_1' P_{\tilde{X}} \tilde{Z}_1}{\sigma^2_{v_{11}}}\right]$$
$$w_2 = \left[\frac{\tilde{Z}_1' P_{\tilde{X}} \tilde{Z}_1}{\sigma^2_{v_{11}}} + \frac{\bar{Z}_1' P_{\bar{X}} \bar{Z}_1}{\sigma^2_{1_{11}}}\right]^{-1} \left[\frac{\bar{Z}_1' P_{\bar{X}} \bar{Z}_1}{\sigma^2_{1_{11}}}\right]$$

$$\hat{\sigma}^2_{v_{11}} = (y_1 - Z_1\tilde{\delta}_{1with2SLS})'Q(y_1 - Z_1\tilde{\delta}_{1with2SLS})/N(T-1) \tag{7.18}$$
$$\hat{\sigma}^2_{1_{11}} = (y_1 - Z_1\tilde{\delta}_{1btw2SLS})'P(y_1 - Z_1\hat{\delta}_{1btw2SLS})/N \tag{7.19}$$
$$\hat{\sigma}^2_{\mu_{11}} = (\hat{\sigma}^2_{1_{11}} - \hat{\sigma}^2_{v_{11}})/T > 0$$

この結果は、第1章でも繰り返し論じたように、EC2SLS 推定はウイズイン推定とビトウィーン推定の加重平均になっている。

(7.8) 式に $\Omega_{11}^{-1/2}$ をかけて

$$y_1^* = Z_1^* \delta_1 + u_1^* \tag{7.20}$$

ここで $y_1^* = \Omega_{11}^{-1/2} y_1, \quad Z_1^* = \Omega_{11}^{-1/2} Z_1, \quad u_1^* = \Omega_{11}^{-1/2} u_1$

$$\begin{aligned} \Omega_{11}^{-1/2} &= (P/\sigma_{1_{11}}) + (Q/\sigma_{v_{11}}) \\ y_{1_{it}}^* &= (y_{1_{it}} - \theta_1 \bar{y}_{1i.})/\sigma_{v_{11}}, \quad \theta_1 = 1 - (\sigma_{v_{11}}/\sigma_{1_{11}}) \\ \bar{y}_{1i.} &= \sum_{t=1}^{T} y_{1it}/T \end{aligned} \tag{7.21}$$

操作変数 A を用いて (7.20) 式を 2SLS 推計すると

$$\hat{\delta}_{1,2SLS} = (Z_1^{*'} P_A Z_1^*)^{-1} Z_1^{*'} P_A y_1^* \tag{7.22}$$

ここで $P_A = A(A'A)^{-1}A'$
最適操作変数を次のように表す。

$$X^* = \Omega_{11}^{-1/2} X = \frac{QX}{\sigma_{v_{11}}} + \frac{PX}{\sigma_{1_{11}}} = \frac{\tilde{X}}{\sigma_{v_{11}}} + \frac{\bar{X}}{\sigma_{1_{11}}}$$

$A = X^*$ とすると G2SLS を得る。

$$\hat{\delta}_{1,G2SLS} = (Z_1^{*'} p_{X^*}^* Z_1)^{-1} Z_1^* P_{X^*} y_1^* \tag{7.23}$$

(7.20) 式に操作変数 $A = [QX, PX] = \left[\tilde{X}, \bar{X}\right]$ を用いて 2SLS 推計を行う。ここで QX は PX に直交しており、$P_A = P_{\tilde{X}} + P_{\bar{X}}$ である。

$$\begin{aligned} P_A Z_1^* &= (P_{\tilde{X}} + P_{\bar{X}}) \left[\Omega_{11}^{-1/2} Z_1\right] \\ &= (P_{\tilde{X}} + P_{\bar{X}}) \left[\frac{Q}{\sigma_{v_{11}}} + \frac{P}{\sigma_{\ell_{11}}}\right] Z_1 = \frac{P_{\tilde{X}} \tilde{Z}_1}{\sigma_{v_{11}}} + \frac{P_{\bar{X}} \bar{Z}_1}{\sigma_{1_{11}}} \end{aligned} \tag{7.24}$$

ここで

$$\begin{aligned} Z_1^{*\prime} P_A Z_1^* &= \left(\frac{\tilde{Z}_1' P_{\tilde{X}} \tilde{Z}_1}{\sigma_{v_{11}}^2} + \frac{\bar{Z}_1' P_{\bar{X}} \bar{Z}_1}{\sigma_{1_{11}}^2} \right) \\ Z_1^{*\prime} P_A y_1^* &= \left(\frac{\tilde{Z}_1' P_{\tilde{X}} \tilde{y}_1}{\sigma_{v_{11}}^2} + \frac{\bar{Z}_1' P_{\bar{X}} \bar{y}_1}{\sigma_{1_{11}}^2} \right) \end{aligned}$$

(7.17) 式の $\hat{\delta}_{1,EC2SLS}$ は $A = \left[\tilde{X}, \bar{X}\right]$ の時 (7.23) 式と同値である。

すなわち、$EC2SLS$ は既存の 2SLS 推計法を用いて推計することが可能なのである。

第 1 ステップ：(7.8) 式 Within2SLS と (7.15) 式 Between2SLS を 2SLS で推計し、(7.12) と (7.14) 式を得る。

第 2 ステップ：$\hat{\sigma}^2_{v_{11}}$ と $\hat{\sigma}^2_{1_{11}}$ を (7.18)(7.19) によって推計し、(7.22) 式で用いる y_1^*, Z_1^*, X^* を得る。(7.8) 式を $\Omega_{11}^{-1/2}$ で変換して (7.20) 式を得る。

第 3 ステップ：操作変数 $A = X^*$ か $A = [QX, PX]$ を用いて (7.20) 式を 2SLS 推計すると、それぞれ (7.22) と (7.17) 式を得る。

7.4 内生性検定

内生性検定としては一般には、Wu-Hausman 検定として知られているもの（最小二乗法推定と操作変数法推定のパラメータをハウスマン検定する）や内生変数に関するモデルを最小二乗推定し、その式から得られた誤差をもともとの式に代入し、そのパラメータが 0 かどうかを t 検定するという方法が提案されている。

次のようなモデルで説明変数 y_2 の内生性の疑いがある時を考えよう[*6]。

$$y_1 = \beta_0 + \beta_1 y_2 + \beta_2 z_1 + \beta_3 z_2 + u_1$$

ここで z_1 と z_2 は外生変数であり、他に操作変数として z_3 と z_4 を考えることができる。この時、上の式を最小二乗法と操作変数法で推計し、パラメータが有意に違うかどうかを Hausman (1978) に従ってカイ二乗検定することによって、説明変数 y_2 が内生であるかどうかを確かめることができる[*7]。これは Durbin-Wu-Hausman (DWH) 検定として知られているが、基本的には Hausman 検定を敷衍したものである。すなわち、最小二乗法推定による推定パラメータ $\widehat{\beta}_e$ と操作変数法による推定パラメータ $\widehat{\beta}_c$ を用いて、次のようなカイ二乗統計量を計算する。

$$(\widehat{\beta}_c - \widehat{\beta}_e)'(var[\widehat{\beta}_c] - var[\widehat{\beta}_e])^{-1}(\widehat{\beta}_c - \widehat{\beta}_e) \backsim \chi(k_1)$$

ここで k_1 は内生性検定の対象となった内生変数の数である。この検定は内生性検定というよりも、最小二乗法と操作変数法を用いた場合に推定パラメータが有意に違うかどうかを検定したものである。

Wooldridge(2002, p.119) は次ような内生性検定法を紹介してる。ここで y_1 に関するモデルを考えよう。

$$y_1 = \beta_0 + \beta_1 y_2 + \beta_2 z_1 + \beta_3 z_2 + u_1$$

[*6] 以下の議論は北村 (2009)『ミクロ計量経済学入門』（日本評論社）第 5 章第 6 節を引用している。

[*7] 一般的には Durbin (1954)、Wu(1973)、Hausman(1978) によって形成された検定であり、Durbin-Wu-Hausman test として知られている。Bowden and Turkington (1984, pp.50-52) や Davidson and MacKinnon (2004, pp.338-340) を参照。

説明変数 y_2 が内生変数であると仮定して次のような式を推計する。

$$y_2 = \alpha_0 + \alpha_1 z_1 + \alpha_2 z_2 + \alpha_3 z_3 + \alpha_4 z_4 + v_2$$

操作変数の仮定により z_j は u_1 とは無相関であるので、v_2 が u_1 と無相関であれば、y_2 も u_1 とは無相関になる。ということは次式でパラメター $\delta_1 = 0$ が y_2 も u_1 とは無相関のための必要十分条件になる。

$$u_1 = \delta_1 v_2 + e_1$$

これを直接検定する方法はないので、y_2 式を最小二乗法で推計し、残差として $\hat{v}_2$ を計算し、これをもとの y_1 式に代入し最小二乗法で推計する。

$$y_1 = \beta_0 + \beta_1 y_2 + \beta_2 z_1 + \beta_3 z_2 + \delta_1 \hat{v}_2 + \varepsilon$$

t 検定で $\delta_1 = 0$ が棄却されれば、y_2 は内生変数であるということになる。

ここでは、従来の内生性の問題とは違って、固定効果が内生であり、説明変数と相関している場合を考える[*8]。はじめに全ての説明変数が固定効果に相関している場合を考え、次いで一部の説明変数のみが固定効果と相関している場合を考えよう。

$$y = \alpha\iota_{NT} + X\beta + Z_\mu \mu + v = Z\delta + Z_\mu \mu + v \tag{7.25}$$
$$\mu_i = \bar{X}_i' \pi + \varepsilon_i \tag{7.26}$$

ここで $\varepsilon_i \backsim iid(0, \sigma_\varepsilon^2)$, $\bar{X}_i'$ は $1 \times K$ ベクトル
(7.26) は次のように書き換えられる。

$$\mu_i = Z_\mu' X\pi / T + \varepsilon_i \tag{7.27}$$

ここで $\mu' = (\mu_1, \ldots, \mu_N)$, $Z_\mu = I_N \otimes \iota_T$, $\varepsilon_i' = (\varepsilon_1, \ldots, \varepsilon_N)$.
(7.27) を (7.25) に代入すると、

$$y = X\beta + PX_\pi + (Z_\mu \varepsilon + v) \tag{7.28}$$

ここで $P = I_N \otimes \bar{J}_T$, ε と v は無相関で、$(Z_\mu \mu + v)$ は平均ゼロで次のような分散共分散行列構造をもつ。

$$V = E(Z_\mu \varepsilon + v)(Z_\mu \varepsilon + v)' = \sigma_\varepsilon^2 (I_N \otimes J_T) + \sigma_v^2 I_{NT} \tag{7.29}$$

*8 以下は Baltagi (2001,pp.118-122) を引用している。これは賃金関数において、個人の固定効果である能力と学歴や職歴が相関している場合、生産関数において潜在変数である経営能力が労働や資本などの投入財と相関している場合など様々なケースで出てくる問題である。

(7.28) の GLS 推計は次のようになる。

$$\hat{\beta}_{GLS} = \tilde{\beta}_{with} = (X'QX)^{-1}X'Qy \tag{7.30}$$

$$\hat{\pi}_{GLS} = \hat{\beta}_{btw} - \tilde{\beta}_{with} = (X'PX)^{-1}X'Py - (X'QX)^{-1}X'Qy \tag{7.31}$$

$$var(\hat{\pi}_{GLS}) = var(\hat{\beta}_{btw}) + var(\tilde{\beta}_{with}) \tag{7.32}$$
$$= (T\sigma_\varepsilon^2 + \sigma_v^2)(X'PX)^{-1} + \sigma_v^2(X'QX)^{-1} \tag{7.33}$$

Mundlak(1978) が示したように、(7.25) 式の最良線形不偏推定量 (BLUE) は固定効果 (within) 推定である。ランダム効果推定は (7.26) 式を無視しておりバイアスが残る。(7.28) 式では全ての説明変数は固定効果に相関しているが、ランダム効果モデルでは説明変数と固定効果は無相関であることが想定されている。

Hausman and Taylor(1981) では、一部の説明変数のみが μ_i と相関しているというモデルを考えている。

$$y_{it} = X_{it}\beta + Z_i\gamma + \mu_i + v_{it} \tag{7.34}$$

Hausman and Taylor (1981) は $X = [X_1; X_2]$ と $Z = [Z_1; Z_2]$ を 2 分割した。すなわち、X_1 は $n \times k_1$, X_2 は $n \times k_2$, Z_1 は $n \times g_1$ Z_2 は $n \times g_2$, $n = NT$ に分割し、X_1 と Z_1 は外生変数、X_2 と Z_2 は内生変数で μ_i と相関し、v_{it} とは無相関であるとする。

ウィズイン誤差項を次のように求める。

$$\hat{d}_i = \bar{y}_i - \bar{X}_i\hat{\beta}_w \tag{7.35}$$

(7.35) の時間平均をとり、$\hat{d}_i$ を z_i に関して、操作変数 $A = [X_1, z_1]$ を用いた 2 SLS 推計を行う。

$$\hat{\gamma}_{2SLS} = (Z'P_AZ)^{-1}Z'P_A\hat{d} \tag{7.36}$$

ここで $P_A = A(A'A)^{-1}A'$。$Z'P_AZ$ は非特異 (non singular) 行列であり、次数条件 $k_1 \geqslant g_2 - 1$ を満たしている。

分散は次のように求められる。

$$\hat{\sigma}_v^2 = \tilde{y}'\bar{P}_{\tilde{X}}\tilde{y}/N(T-1) \tag{7.37}$$

ここで $\tilde{y} = Qy$, $\tilde{X} = QX$, $\bar{P}_A = 1 - P_A$

$$\sigma_1^2 = \frac{(y_{it} - X_{it}\tilde{\beta}_w - Z_i\hat{\gamma}_{2SLS})'P(y_{it} - X_{it}\tilde{\beta}_w - Z_i\hat{\gamma}_{2SLS})}{N} \tag{7.38}$$

(7.35) を次のように変換する。

$$\Omega^{-1/2}y_{it} = \Omega^{-1/2}X_{it}\beta + \Omega^{-1/2}Z_1\gamma + \Omega^{-1/2}u_{it} \tag{7.39}$$

Hausman and Taylor 推定は、(7.39) 式を $\mathrm{A}_{HT} = \left[\tilde{X}, \tilde{X}_1, Z_1\right]$ を操作変数とした 2SLS 推計であると理解できる。

1. $k_1 < g_2 - 1$ であれば過小識別、$\hat{\beta}_{HT} = \tilde{\beta}_{with}$ であり $\hat{\gamma}_{HT}$ は存在しない。
2. $k_1 = g_2 - 1$ であれば適正識別、$\hat{\beta}_{HT} = \hat{\beta}_{with}$, $\hat{\gamma}_{HT} = \hat{\gamma}_{2SLS}$ である。
3. $k_1 > g_2 - 1$ であれば過剰識別、(7.39) 式より得られた $\hat{\beta}_{HT}$ は $\hat{\beta}_{with}$ より有効である。

過剰識別テストは次の統計量によってテストできる。

$$\hat{m} = \hat{q}' \left[var(\tilde{\beta}_{with}) - var(\hat{\beta}_{HT}) \right]^{-1} \hat{q} \tag{7.40}$$

ここで $\hat{q} = \hat{\beta}_{HT} - \tilde{\beta}_{with}$, $\hat{\sigma}_v^2 \hat{m} \overset{H_0}{\to} X_\iota^2$, $\ell = \min[k_1 - q_2, NT - k]$ である。

7.5 不均一分散検定

不均一分散の問題は操作変数法にも残っており、検定を行い、不均一分散の問題を解決することが望ましい*9。また、パラメータの分散を不均一分散頑強標準誤差によって修正し、頑強 t 統計量を推定する必要がある。基本的な考え方は第 4 章で見た、Breusch and Pagan (1979) の検定と White (1980) の不均一誤差頑強推定を踏襲するものである。

しかし、Pagan and Hall (1983) が指摘したように、複数の内生変数をもつ操作変数法を考えるような一般的設定では、Breusch and Pagan (1979) の検定は、関心のある内生変数を含んだ式のみの不均一分散を検定しており、潜在的な他の連立方程式における不均一分散問題は無視している。Pagan and Hall (1983) および White (1982) は他の連立方程式に不均一分散問題が存在しているという設定でカイ二乗検定を提案している*10。

7.6 弱相関の操作変数の問題

実証研究上、適切な操作変数を見つけることは極めて難しいことが知られている*11。とりわけ z と x の相関が弱い場合には問題がある。変数 z と誤差項 u が相関している場合の操作変数法による推計値の確率極限は次のように表せる。

$$p\lim \hat{\beta}_1 = \beta_1 + \frac{Corr(z,u)}{Corr(z,x)} \cdot \frac{\sigma_u}{\sigma_x}$$

σ_u と σ_x は、u と x に関する標準偏差である。問題は例え、$Corr(z,u)$ が小さくても、$Corr(z,x)$ も小さければ操作変数による推計値 $\hat{\beta}_1$ は大幅な不一致推定となるということである。現実的に

*9 以下の議論は北村 (2009)『ミクロ計量経済学入門』(日本評論社) 第 5 章第 4 節を引用している。

*10 STATA では ivhettest というコマンドを使うことで、Pagan and Hall (1983)、White (1980)、Breusch and Pargan (1979)、Koenker (1981) などの一連の不均一分散検定を行うことができる。

*11 以下の議論は北村 (2009)『ミクロ計量経済学入門』(日本評論社) 第 5 章第 5 節を引用している。

考えて、操作変数を用いるよりも最小二乗法を用いた方が不一致性の程度が低くなることもあり得る。

これは**弱相関操作変数**（Weak Instrumental Variables）の問題として知られている。以下ではこの弱相関に関する主要な検定を紹介する。

Bound, Jaeger and Baker (1995) は内生変数を操作変数で回帰した第1段階の推定式の決定係数 (R^2) の計算において、いくつかの操作変数を落として推定することによって決定係数に変化があるかどうかを検定することを提案している。具体的には部分決定係数は $(RSS_{Z_2} - RSS_Z)/TSS$ と定義され、RSS_{Z_2} は操作変数 Z_2 だけを使って計算した誤差平方和であり、RSS_Z はすべての操作変数を用いた場合の誤差平方和を表す。この方法は弱相関の問題を検知するが、その結果として内生変数に対して操作変数が過小になる場合を排除できない。

Shea(1997) は操作変数間の相関を考慮した部分決定係数を提案している。すなわち、内生変数 i に関する部分決定係数は $R_p^2 = (v_{i,i,OLS})/(v_{i,i.IV})\{(1-R_{IV}^2)/(1-R_{OLS}^2)\}$ と表される。ここで $v_{i,i}$ は内生変数にかかる推定係数の漸近分散を表している[*12]。

Andreson(1984) とそれを敷衍した Hall, Rudebusch and Wilcox (1996) はより一般的なアプローチを提案している。ここでは内生変数 $\mathbf{X}$ と操作変数 $\mathbf{Z}$ の行列間の**正準相関**（canonical correlation）$Corr_i\ i = 1,2,..k$ を計算し、操作変数が有意であるということは、すべての相関が有意にゼロとは異なるはずであることを検定している。Anderson は最小の正準相関はゼロであるという帰無仮説を尤度比を用いて検定した。この統計量は自由度 $l-k+1$ のカイ二乗分布に従う。帰無仮説が棄却できなければ操作変数の弱相関問題だけではなく、識別に問題がある可能性を示唆することになる。

Hall and Peixe (2000) は正準相関を用いて操作変数の重複 (redundancy) を検定する方法を提案した。これは Anderson の検定に似ているが、重複していると疑われる操作変数を含んだ正準相関と含まない正準相関に関する尤度比検定であり、自由度が内生変数と重複している操作変数の積に等しくなるようなカイ二乗分布に従うことを示した[*13]。

操作変数の弱相関問題は、多くの無相関あるいは低相関の操作変数を用いると最小二乗法推定以上にバイアスをもたらすことが Hahn and Hausman (2002b) によって示されている。また、Staiger and Stock (1997) は弱相関問題は第1段階の推定において操作変数が有意であっても起こりうることを示している[*14]。

[*12] Shea(1997) の統計量は小さい程、操作変数の内生変数に関する説明力が低いことを意味しいてる。STATA のコマンド ivreg2 の中の first か ffirst というオプションを用いれば計算できる。

[*13] この検定は STATA では ivreg2 の中の redundant というオプションを用いればいい。また ivreg2 の中にでは Anderson and Rubin (1949)、Cragg and Donald (1993)、Stock and Wright (2000) らの検定が出来る。

[*14] Nelson and Sartz (1990)、Stock and Wright (2000)、Stock, Wright and Yogo (2002)、Hahn and Hausman (2002a, 2003)、Andrews and Stock (2005)、Chao and Swanson (2005)、Stock and Yogo (2005)、Hausman, Stock and Yogo (2005) なども参照。

7.7 STATA コード

ここでは『21 世紀出生児縦断調査』を用いて操作変数法パネル推定を行ってみよう。これまで北村 (2007、2008, 2009) で行ってきた出生児の身体成長パターンの測定を行ってきた。これまでの研究では操作変数パネル推定ではなく、身長・体重の対数を出生からの経過日数とその二乗、および子育費用によって説明する固定効果モデルを採用してきた。

しかし、第 6 章でモデルに一次のラグ項を入れた推定を行うと推定係数に内生性バイアスがあるように見受けられた。そこで、本章では子育費用の対数値を被説明変数とするモデルを考えてみたい。基本的な考え方は、子育費用は母親がフルタイムの職に就いていてそれなりの所得を得ているかどうかに依存するが、母親の就業は父親の所得とベビーシッターの利用可能性に依存しているというものである。

$$\begin{aligned} \ln\cos tofchildcare_{it} &= \alpha + \gamma fulltimework\_m_{it} + \beta workdummy\_m_{it} + u_{it} \\ workdummy\_m_{it} &= \delta \ln income\_f + \eta childcarewor\ker + e_{it} \end{aligned}$$

ここで $\ln\cos tofchildcare$ = 子育費用の対数値、$fulltimework\_m$ = 母親が週 40 時間以上の労働に従事している場合に 1 をとるダミー変数、$workdummy\_m$ = 母親の就業状況を表すダミー変数。フルタイム、パートタイム、自営業、内職などいずれかの職についていれば 1 をとる。$\ln income\_f$ は父親の所得の対数値、$childcarewor\ker$ は普段の保育者が保育ママやベビーシッターであれば 1 をとるダミー変数。

推計に使った STATA コードは次のようになる。

```
/*weak instruments tests*/
ivreg2 lncostofchildcare fulltimework_m (workdummy_m = lnincome_f childcareworker),
gmm2s orthog(lnincome_f)
ivhettest, all
/*Durbin-Wu-Hausman tests for endogeneity in Iv estimation*/
quietly ivreg2 lncostofchildcare fulltimework_m (workdummy_m = lnincome_f childcare-
worker), small
estimates store iv2
quietly regress lncostofchildcare fulltimework_m workdummy_m
hausman iv2, constant sigmamore
quietly ivreg2 lncostofchildcare fulltimework_m (workdummy_m = lnincome_f childcare-
worker), orthog(lnincome_f) small
ivendog
/*Panel Instrumental Variable Estimation*/
xtivreg lncostofchildcare fulltimework_m (workdummy_m = lnincome_f childcare-
```

```
worker), fe
est store fixed4
xtivreg lncostofchildcare fulltimework_m (workdummy_m = lnincome_f childcare-
worker), re
hausman fixed4
```

まず、図表 7.1 では 7.6 節で論じた操作変数の有意性検定を行った。ここではデータをプールして検定している。各検定の統計的意義や解釈については北村 (2009) を参照されたい。ここでは、Anderson Canonical Correlation 尤度比検定では帰無仮説が棄却できることで、操作変数が有意であることを示唆している。Cragg and Donald 検定でも弱相関問題が棄却されていること、そして、不均一分散検定に関しては、Pagan and Hall 検定、White/Koenker 検定、Breusch and Pagan 検定などがあり、不均一分散の存在が棄却できないことを示唆していることを報告しておきたい。

Dependent Variable: lncostofchildcare	2step GMM Estimated Coefficient	z-stat
workdummy_m	0.767	76.77
fulltimework_m	−0.008	−0.74
_cons	0.660	171.83
Number of observation	142886	
Centerd R2	−0.012	
Uncentered R2	0.549	
Root MSE	0.840	
Identification Tests		
Underidentification test (Anderson canonical Correlation LM statistic)	Chi2(2) = 3.6e+04 P-value = 0.0000	
Weak identification test (Cragg-Donald wald F statistic)	2.4e + 04	
Sargan statistic	Chi2(1) = 2188.256 P-value = 0.0000	
C Statistic	Chi2(1) = 2188.256 P-value = 0.0000	
IV heteroskedasticity tests		
Pagan-Hall general test statistic	Chi2(3) = 377.345 P-value = 0.0000	
Pagan-Hall test	Chi2(3) = 801.917 P-value = 0.0000	
White/Koenker nR2 test statistic	Chi2(3) = 365.547 P-value = 0.0000	
Breusch-Pagan/Godfrey/Cook-Weisberg	Chi2(3) = 805.520 P-value = 0.0000	
Instrumented:	Workdummy_m	
Included instruments:	lnincome_m fulltimework_m	
Excluded instruments:	lnincome_f childcareworker	

図表 7.1　操作変数の有意性検定