中級・応用計量経済学 宿題 2 解答例（修正版）

吉村 有博*

2013 年 1 月 28 日

1. 本問の解答例では、追加的な仮定として z と x は完全相関ではないとする。つまり、$|Corr(z,x)| \neq 1$ とする。

(i) Hausman 検定の帰無仮説は、$H_0 : E(ux) = 0$、対立仮説は $H_1 : E(ux) \neq 0$.

(ii) データベクトルを講義ノートのように定義すれば、

$$\hat{\beta}_{OLS} = (X'X)^{-1}X'Y,$$
$$\hat{\beta}_{2SLS} = (Z'X)^{-1}Z'Y.$$

(iii) 帰無仮説の下では $E(ux) = 0$、つまり x は外生変数なので、OLS 推定量は一致性・漸近正規性を持つ。均一分散の仮定より、OLS 推定量の漸近分散は講義ノート 2 の (17) 式から

$$V_{OLS} = \frac{\sigma_u^2}{E(x^2)} = \frac{\sigma_u^2}{\sigma_x^2}.$$

一方、帰無仮説の下で、2SLS 推定量は同様に一致性・漸近正規性を持つため[*1]、漸近分散は講義ノート 3 の (18) 式と、繰り返し期待値の法則から、

$$\begin{aligned} V_{2SLS} &= \frac{Var(z_i u_i)}{(Cov(z_i, x_i))^2} \\ &= \frac{\sigma_u^2 E(z^2)}{(E(zx))^2} = \frac{\sigma_u^2 \sigma_z^2}{(\sigma_{xz})^2}. \end{aligned}$$

[*2]

(iv) 真のパラメータを足し引きして、

$$\begin{aligned} \hat{\beta}_{2SLS} - \hat{\beta}_{OLS} &= (\hat{\beta}_{2SLS} - \beta) - (\hat{\beta}_{OLS} - \beta) \\ &= (Z'X)^{-1}Z'u - (X'X)^{-1}X'u. \end{aligned}$$

---

* TA: 経済学研究科博士後期課程 2 年

*1 適切な操作変数があるような状況では、2SLS 推定量は対立仮説の下でも一致性・漸近正規性を持つことに注意。

*2 補足：ここで、二つの推定量の漸近分散を簡単に比較できる。コーシー＝シュワルツの不等式から

$$(E(zx))^2 \leq E(z^2)E(x^2)$$

であるので、$V_{2SLS} \geq V_{OLS}$ である。ただし、等式が成り立つのは z と x が（確率 1 で）線形関係にある時その時のみである。つまり、z と x が完全相関でさえなければ、$V_{2SLS} > V_{OLS}$ が成り立ち、つまり OLS の方が漸近分散が厳密に小さいことが分かる。このように、2SLS 推定量は x が内生的か外生的かに依らずに一致推定できるというロバストな推定量ではあるが、反面、OLS 推定量よりも分散が大きくなってしまう側面がある。従って、できれば OLS 推定量が使いたいので、それを確認したいというのがこの Hausman 検定のモチベーションである。

(v)　中心極限定理が働く形式に書き直せばよい。帰無仮説の下で、

$$\begin{aligned}
\sqrt{n}(\hat{\beta}_{2SLS}-\hat{\beta}_{OLS}) &= \sqrt{n}(\hat{\beta}_{2SLS}-\beta)-\sqrt{n}(\hat{\beta}_{OLS}-\beta) \qquad (1)\\
&= \frac{\frac{1}{\sqrt{n}}\sum_i z_i u_i}{\frac{1}{n}\sum_i z_i x_i}-\frac{\frac{1}{\sqrt{n}}\sum_i x_i u_i}{\frac{1}{n}\sum_i x_i^2}\\
&= \frac{\frac{1}{\sqrt{n}}\sum_i z_i u_i}{E(zx)}-\frac{\frac{1}{\sqrt{n}}\sum_i x_i u_i}{E(x^2)}+(small)\\
&= \frac{1}{E(zx)E(x^2)}\frac{1}{\sqrt{n}}\sum_i \{E(x^2)z_i - E(zx)x_i\}u_i+(small).
\end{aligned}$$

ただし、ここで二つの推定量の分母をまとめるために、連続写像定理より分母の標本平均を期待値で置き換えて、置き換えによる残った誤差をオーダーの小さな項（small）としている[*3]。これより、帰無仮説の下では第一項の和の中は iid で、期待値ゼロ。その分散を V_1 と書けば、中心極限定理とスラツキーの補題より、

$$\begin{aligned}
\sqrt{n}(\hat{\beta}_{2SLS}-\hat{\beta}_{OLS}) &\to^d \frac{1}{E(zx)E(x^2)}N(0,V_1)\\
&\sim N\left(0,\frac{1}{E(zx)^2E(x^2)^2}V_1\right).
\end{aligned}$$

漸近分散 V は、均一分散の仮定と繰り返し期待値の法則より、

$$\begin{aligned}
V &= \frac{1}{E(zx)^2E(x^2)^2}V_1\\
&= \frac{1}{E(zx)^2E(x^2)^2}E((E(x^2)z_i - E(zx)x_i)^2u_i^2)\\
&= \frac{\sigma_u^2}{E(zx)^2E(x^2)^2}\left(E(z^2)E(x^2)^2 - E(x^2)E(zx)^2\right)\\
&= \frac{\sigma_u^2E(z^2)}{E(zx)^2}-\frac{\sigma_u^2}{E(x^2)}\\
&= V_{2SLS}-V_{OLS}.
\end{aligned}$$

つまり、推定量の差の漸近分散が、それぞれの漸近分散の差で書ける。

(vi)　(1) 式の Hausman 検定統計量は（漸近的に）正規分布の二乗（一般には二次形式）なのでカイ二乗分布する (講義ノート付録の 4 ページ参照)。カイ二乗分布の自由度については、漸近分散がスカラーで、かつ z と x は完全相関でない仮定によりゼロに退化しないので、$rank((V_{2SLS}-V_{OLS})^{-1})=1$ より、自由度は 1.

(vii)　帰無仮説が間違っていた場合、つまり対立仮説の下では、OLS 推定量は漸近バイアスを持つ。つまり、

$$\hat{\beta}_{OLS}-\beta \to^p \frac{E(xu)}{E(x^2)}$$

であるので、バイアスの符号に応じて、

$$\sqrt{n}(\hat{\beta}_{OLS}-\beta) \to^p \begin{cases} +\infty & \text{if } E(xu)>0\\ -\infty & \text{if } E(xu)<0 \end{cases}$$

[*3] この置き換えは、帰無仮説の下では正当化される。

となる。一方、対立仮説の下でも 2SLS 推定量は平均 0 を中心とした正規分布に分布収束する。従って以上の結果から、対立仮説の下では、(v) の (1) 式は $+\infty$ か $-\infty$ に発散するので、Hausman 検定統計量は $+\infty$ に発散する。*4

2. (i) 講義ノート 4 の (17) 式より、固定効果推定値は、$\hat{\beta}_{FE} = 1.97$.
   (ii) (21)、(22) 式より、漸近分散は、$\hat{V}_{FE} = 0.68$.
   (iii) t 値は、$\frac{\hat{\beta}_{FE}-0}{\sqrt{\frac{\hat{V}_{FE}}{n}}} = 23.95 > 1.96$ より、有意水準 5 %で帰無仮説を棄却する。
3. (i) 最小二乗法による (a, b) の推定値は、$(\hat{a}_{OLS}, \hat{b}_{OLS}) = (-0.06, 1.19)$.
   (ii) 講義ノート 2 の (16) 式の不均一分散ロバスト分散推定量を用いると、傾き b の分散推定値は $\hat{V} = 1.07$ で、t 値は $\frac{\hat{b}_{OLS}-4}{\sqrt{\frac{\hat{V}}{n}}} = -27.11 < -1.96$ より、有意水準 5 %で帰無仮説を棄却する。
   (iii) プロビット回帰による推定値は、$(\hat{\alpha}_{probit}, \hat{\beta}_{probit}) = (-1.97, 4.20)$.
   (iv) 傾き β の漸近分散推定値は、$\hat{V}_{probit} = 48.73$. 従って、t 値は $\frac{\hat{\beta}_{probit}-4}{\sqrt{\frac{\hat{V}_{probit}}{n}}} = 0.29 < 1.96$ より、帰無仮説を受容（採択）する。*5

*4 この性質により、Hausman 検定の（漸近的な）検出力が保障される。

*5 補足：本問は、モデルの想定を間違えると検定結果が大きく変わってしまうことを確認する問題である。真のモデルは誤差項が正規分布するような probit モデルなので、(iii) の推定法は正しい推定を行っており、検定結果も信頼できる。一方、(i) の最小二乗法は、仮に真のモデルが線形確率モデルならば正しい推定・検定が行える（講義ノート 5、1 ページ参照）が、本問ではそれは間違った想定にあたるため、検定結果は信頼できない。実際に、(ii) において傾きが 4 であるという仮説が棄却されてしまったが、このデータは実は真の傾きが 4 として生成されたものであるため、間違った推論を導いてしまうことが確認できる。