

計量経済 I : 宿題 6

村澤 康友

提出期限 : 2020 年 7 月 14 日 (My KONAN で提出)

注意 : すべての質問に解答しなければ提出とは認めない。授業の HP の解答例を正確に再現すること (乱数は除く)。グループで取り組んでよいが、個別に提出すること。解答例をコピーしたり、他人の名前で提出した場合は、提出点を 0 点とし、再提出も認めない。すべての結果をワープロ文書に貼り付け、pdf ファイルに変換して提出すること。

1. (教科書 p. 209, 実証分析問題 8-A) データセット「8_income.dta」を gretl に読み込み、以下の分析を行いなさい。
 - (a) 教科書 p. 198 の「年収 (対数値)」を「修学年数」で説明する単回帰モデルの推定結果を再現しなさい。
 - (b) 教科書 p. 198 の「本人の修学年数」を「父親の修学年数」で説明する単回帰モデルの推定結果を再現しなさい。
 - (c) gretl で 2SLS を実行する手順は以下の通り。
 - i. メニューから「モデル」→「操作変数法」→「2 段階最小二乗法」を選択。
 - ii. 「従属変数」を選択。
 - iii. 「説明変数 (回帰変数)」を選択。
 - iv. 「操作変数」を選択。
 - v. 「OK」をクリック。教科書 p. 199 の「年収 (対数値)」を「修学年数」で説明する線形モデルの 2SLS による推定結果を再現しなさい。
2. (教科書 p. 209, 実証分析問題 8-B) 前問と同じデータを用いて、以下の分析を行いなさい。
 - (a) 教科書 p. 207 の 2SLS によるミンサー方程式の推定結果を再現しなさい。
 - (b) 前問の分析に「母親の修学年数」を IV に加えて 2SLS でミンサー方程式を推定しなさい。
 - (c) さらに「生まれ月」を IV に加えて 2SLS でミンサー方程式を推定しなさい。
3. (教科書 p. 236, 実証分析問題 9-A) データセット「9_1_cig_xt.dta」を gretl に読み込み、以下の分析を行いなさい。
 - (a) 教科書 p. 225 の「喫煙本数」を「生活の満足度」で説明する単回帰モデルの 2007 年と 2009 年の推定結果を再現しなさい。
 - (b) 教科書 p. 226 の「喫煙本数」の差分を「生活の満足度」の差分で説明する単回帰モデルの推定結果を再現しなさい。※変数の変換方法は資料「gretl 入門」を参照。
 - (c) 前問の分析に所得を共変量として加えた重回帰モデルを推定しなさい。
4. (教科書 p. 236, 実証分析問題 9-B) データセット「9_2_life_xt.dta」を gretl に読み込み、以下の

分析を行いなさい。

- (a) 教科書 p. 228 の「生活の満足度」を「怪我・病気ダミー」と「年収」で説明する重回帰モデルの 2009 年の推定結果を再現しなさい。
- (b) 教科書 p. 229 の「生活の満足度」の差分を「怪我・病気ダミー」と「年収」の差分で説明する重回帰モデルの推定結果を再現しなさい。

解答例

1. (a) 単回帰

モデル 1: 最小二乗法 (OLS), 観測: 1-734

従属変数: lincome

| | 係数 | Std. Error | t-ratio | p 値 |
|--------------------|-----------|--------------------|---------|----------|
| const | 5.38769 | 0.0870176 | 61.91 | 0.0000 |
| yeduc | 0.0553906 | 0.00609099 | 9.094 | 0.0000 |
| Mean dependent var | 6.170857 | S.D. dependent var | | 0.356020 |
| Sum squared resid | 83.47680 | S.E. of regression | | 0.337697 |
| R^2 | 0.101508 | Adjusted R^2 | | 0.100280 |
| $F(1, 732)$ | 82.69835 | P-value(F) | | 8.86e-19 |
| Log-likelihood | -243.6648 | Akaike criterion | | 491.3296 |
| Schwarz criterion | 500.5267 | Hannan-Quinn | | 494.8770 |

(b) 2SLS の第 1 段階

モデル 2: 最小二乗法 (OLS), 観測: 1-734

従属変数: yeduc

| | 係数 | Std. Error | t-ratio | p 値 |
|--------------------|-----------|--------------------|---------|----------|
| const | 10.5220 | 0.350154 | 30.05 | 0.0000 |
| payeduc | 0.295540 | 0.0280256 | 10.55 | 0.0000 |
| Mean dependent var | 14.13896 | S.D. dependent var | | 2.047800 |
| Sum squared resid | 2668.439 | S.E. of regression | | 1.909295 |
| R^2 | 0.131883 | Adjusted R^2 | | 0.130697 |
| $F(1, 732)$ | 111.2046 | P-value(F) | | 2.66e-24 |
| Log-likelihood | -1515.202 | Akaike criterion | | 3034.405 |
| Schwarz criterion | 3043.602 | Hannan-Quinn | | 3037.952 |

(c) 2SLS

モデル 1: 二段階最小二乗法 (2SLS), 観測: 1-734

従属変数: lincome

内生変数 (instrumented): yeduc

操作変数: const payeduc

| | 係数 | 標準誤差 | t 値 | p 値 | |
|--------------------|-----------|--------------------|----------|-----------|-----|
| const | 5.75290 | 0.240370 | 23.93 | 5.86e-094 | *** |
| yeduc | 0.0295608 | 0.0169771 | 1.741 | 0.0821 | * |
| Mean dependent var | 6.170857 | S.D. dependent var | 0.356020 | | |
| Sum squared resid | 85.52760 | S.E. of regression | 0.341820 | | |
| R-squared | 0.101508 | Adjusted R-squared | 0.100280 | | |
| F(1, 732) | 3.031835 | P-value(F) | 0.082066 | | |
| Log-likelihood | -5192.300 | Akaike criterion | 10388.60 | | |
| Schwarz criterion | 10397.80 | Hannan-Quinn | 10392.15 | | |

ハウスマン (Hausman) 検定 -

帰無仮説: OLS 推定値は一致性を持つ

漸近的検定統計量: カイ二乗 (1) = 2.74972

なお、p 値 (p-value) = 0.0972716

弱操作変数 (weak instrument) の検定 -

第 1 段階の F 統計量 (1, 732) = 111.205

名目 5% の有意水準で検定を行う場合の望ましい TSLS 最大サイズに対する臨界値:

| | | | | |
|-------|-------|------|------|------|
| size | 10% | 15% | 20% | 25% |
| value | 16.38 | 8.96 | 6.66 | 5.53 |

2. (a) IV : 就業可能年数, 就業可能年数の 2 乗, 父親の修学年数, 兄弟姉妹数

モデル 1: 二段階最小二乗法 (2SLS), 観測: 1-734

従属変数: `lincome`

内生変数 (instrumented): `yeduc`

操作変数: `const exper exper2 payeduc sibs`

| | 係数 | 標準誤差 | t 値 | p 値 | |
|--------------------|-------------|--------------------|----------|-----------|-----|
| const | 4.52414 | 0.328680 | 13.76 | 1.71e-038 | *** |
| yeduc | 0.0699093 | 0.0217875 | 3.209 | 0.0014 | *** |
| exper | 0.0609592 | 0.0160773 | 3.792 | 0.0002 | *** |
| exper2 | -0.00104174 | 0.000610360 | -1.707 | 0.0883 | * |
| Mean dependent var | 6.170857 | S.D. dependent var | 0.356020 | | |
| Sum squared resid | 70.30899 | S.E. of regression | 0.310344 | | |
| R-squared | 0.246920 | Adjusted R-squared | 0.243825 | | |
| F(3, 730) | 23.08824 | P-value(F) | 2.75e-14 | | |

ハウスマン (Hausman) 検定 -

帰無仮説: OLS 推定値は一致性を持つ

漸近的検定統計量: カイ二乗 (1) = 0.477582

なお、p 値 (p-value) = 0.48952

Sargan の過剰識別検定 -

帰無仮説: 全ての操作変数は有効 (valid) である

検定統計量: LM = 0.403198

なお、p 値 (p-value) = $P(\text{カイ二乗}(1) > 0.403198) = 0.525442$

弱操作変数 (weak instrument) の検定 -

第 1 段階の F 統計量 (2, 729) = 32.831

名目 5% の有意水準で検定を行う場合の望ましい TSLS 最大サイズに対する臨界値:

| size | 10% | 15% | 20% | 25% |
|-------|-------|-------|------|------|
| value | 19.93 | 11.59 | 8.75 | 7.25 |

(b) IV : 就業可能年数, 就業可能年数の 2 乗, 父親の修学年数, 兄弟姉妹数, 母親の修学年数

モデル 2: 二段階最小二乗法 (2SLS), 観測: 1-734

従属変数: lincome

内生変数 (instrumented): yeduc

操作変数: const exper exper2 payeduc sibs moyeduc

| | 係数 | 標準誤差 | t 値 | p 値 | |
|--------------------|-------------|--------------------|----------|-----------|-----|
| const | 4.54345 | 0.320583 | 14.17 | 1.87e-040 | *** |
| yeduc | 0.0685564 | 0.0211846 | 3.236 | 0.0013 | *** |
| exper | 0.0612705 | 0.0160450 | 3.819 | 0.0001 | *** |
| exper2 | -0.00106162 | 0.000606077 | -1.752 | 0.0803 | * |
| Mean dependent var | 6.170857 | S.D. dependent var | 0.356020 | | |
| Sum squared resid | 70.40867 | S.E. of regression | 0.310564 | | |
| R-squared | 0.246519 | Adjusted R-squared | 0.243423 | | |
| F(3, 730) | 23.11941 | P-value(F) | 2.64e-14 | | |

ハウスマン (Hausman) 検定 -

帰無仮説: OLS 推定値は一致性を持つ

漸近的検定統計量: カイ二乗 (1) = 0.609293

なお, p 値 (p-value) = 0.435054

Sargan の過剰識別検定 -

帰無仮説: 全ての操作変数は有効 (valid) である

検定統計量: LM = 0.471859

なお, p 値 (p-value) = P(カイ二乗 (2) > 0.471859) = 0.789836

弱操作変数 (weak instrument) の検定 -

第 1 段階の F 統計量 (3, 728) = 23.2762

Critical values for TSLS bias relative to OLS:

| bias | 5% | 10% | 20% | 30% |
|-------|-------|------|------|------|
| value | 13.91 | 9.08 | 6.46 | 5.39 |

Relative bias is probably less than 5%

名目 5% の有意水準で検定を行う場合の望ましい TSLS 最大サイズに対する臨界値:

| size | 10% | 15% | 20% | 25% |
|-------|-------|-------|------|------|
| value | 22.30 | 12.83 | 9.54 | 7.80 |

(c) IV : 就業可能年数, 就業可能年数の 2 乗, 父親の修学年数, 兄弟姉妹数, 母親の修学年数, 生まれ月

モデル 3: 二段階最小二乗法 (2SLS), 観測: 1-734

従属変数: `lincome`

内生変数 (instrumented): `yeduc`

操作変数: `const exper exper2 payeduc sibs moyeduc mbirth`

| | 係数 | 標準誤差 | t 値 | p 値 | |
|--------------------|-------------|--------------------|----------|-----------|-----|
| const | 4.51230 | 0.319307 | 14.13 | 2.95e-040 | *** |
| yeduc | 0.0707394 | 0.0210926 | 3.354 | 0.0008 | *** |
| exper | 0.0607683 | 0.0160224 | 3.793 | 0.0002 | *** |
| exper2 | -0.00102955 | 0.000604888 | -1.702 | 0.0892 | * |
| Mean dependent var | 6.170857 | S.D. dependent var | 0.356020 | | |
| Sum squared resid | 70.25229 | S.E. of regression | 0.310219 | | |
| R-squared | 0.247140 | Adjusted R-squared | 0.244046 | | |
| F(3, 730) | 23.42145 | P-value(F) | 1.76e-14 | | |

ハウスマン (Hausman) 検定 -

帰無仮説: OLS 推定値は一致性を持つ

漸近的検定統計量: カイ二乗 (1) = 0.454561

なお, p 値 (p-value) = 0.500177

Sargan の過剰識別検定 -

帰無仮説: 全ての操作変数は有効 (valid) である

検定統計量: LM = 2.12962

なお, p 値 (p-value) = $P(\text{カイ二乗}(3) > 2.12962) = 0.545943$

弱操作変数 (weak instrument) の検定 -

第 1 段階の F 統計量 (4, 727) = 17.5575

Critical values for TSLS bias relative to OLS:

| bias | 5% | 10% | 20% | 30% |
|-------|-------|-------|------|------|
| value | 16.85 | 10.27 | 6.71 | 5.34 |

Relative bias is probably less than 5%

名目 5% の有意水準で検定を行う場合の望ましい TSLS 最大サイズに対する臨界値:

| size | 10% | 15% | 20% | 25% |
|-------|-------|-------|-------|------|
| value | 24.58 | 13.96 | 10.26 | 8.31 |

3. (a) 2007 年

モデル 1: 最小二乗法 (OLS), 観測: 1-3022
 従属変数: ncig

| | 係数 | Std. Error | t-ratio | p 値 |
|--------------------|-----------|--------------------|----------|--------|
| const | 5.73678 | 0.281568 | 20.37 | 0.0000 |
| life | -0.685414 | 0.102435 | -6.691 | 0.0000 |
| Mean dependent var | 3.966992 | S.D. dependent var | 5.346335 | |
| Sum squared resid | 85088.68 | S.E. of regression | 5.308018 | |
| R^2 | 0.014609 | Adjusted R^2 | 0.014282 | |
| $F(1, 3020)$ | 44.77234 | P-value(F) | 2.63e-11 | |
| Log-likelihood | -9331.411 | Akaike criterion | 18666.82 | |
| Schwarz criterion | 18678.85 | Hannan-Quinn | 18671.15 | |

2009 年

モデル 2: 最小二乗法 (OLS), 観測: 1-3022
 従属変数: ncig

| | 係数 | Std. Error | t-ratio | p 値 |
|--------------------|-----------|--------------------|----------|--------|
| const | 5.07249 | 0.272702 | 18.60 | 0.0000 |
| life | -0.507946 | 0.0946631 | -5.366 | 0.0000 |
| Mean dependent var | 3.691016 | S.D. dependent var | 4.964790 | |
| Sum squared resid | 73761.82 | S.E. of regression | 4.942109 | |
| R^2 | 0.009444 | Adjusted R^2 | 0.009116 | |
| $F(1, 3020)$ | 28.79208 | P-value(F) | 8.67e-08 | |
| Log-likelihood | -9115.560 | Akaike criterion | 18235.12 | |
| Schwarz criterion | 18247.15 | Hannan-Quinn | 18239.44 | |

(b) 差分の単回帰

モデル 3: Pooled OLS, 観測数: 3022
 クロスセクションユニット数: 3022
 時系列の長さ = 1
 従属変数: d.ncig

| | 係数 | Std. Error | t-ratio | p 値 |
|--------------------|-----------|--------------------|-----------|--------|
| d.life | -0.287266 | 0.109858 | -2.615 | 0.0090 |
| Mean dependent var | -0.275976 | S.D. dependent var | 5.322167 | |
| Sum squared resid | 85607.61 | S.E. of regression | 5.323299 | |
| Uncentered R^2 | 0.002258 | Centered R^2 | -0.000425 | |
| $F(1, 3021)$ | 6.837607 | P-value(F) | 0.008970 | |
| Log-likelihood | -9340.598 | Akaike criterion | 18683.20 | |
| Schwarz criterion | 18689.21 | Hannan-Quinn | 18685.36 | |

(c) 差分の重回帰

モデル 4: Pooled OLS, 観測数: 3022

クロスセクションユニット数: 3022

時系列の長さ = 1

従属変数: d_ncig

| | 係数 | Std. Error | t-ratio | p 値 |
|--------------------|------------|--------------------|----------|--------|
| d_life | -0.300654 | 0.109478 | -2.746 | 0.0061 |
| d_income | 0.00314364 | 0.000643058 | 4.889 | 0.0000 |
| Mean dependent var | -0.275976 | S.D. dependent var | 5.322167 | |
| Sum squared resid | 84935.49 | S.E. of regression | 5.303238 | |
| Uncentered R^2 | 0.010092 | Centered R^2 | 0.007429 | |
| $F(2, 3020)$ | 15.39386 | P-value(F) | 2.23e-07 | |
| Log-likelihood | -9328.688 | Akaike criterion | 18661.38 | |
| Schwarz criterion | 18673.40 | Hannan-Quinn | 18665.70 | |

4. (a) 2009 年

モデル 1: 最小二乗法 (OLS), 観測: 1-3020
従属変数: life

| | 係数 | Std. Error | t-ratio | p 値 |
|--------------------|-------------|--------------------|---------|----------|
| const | 2.67366 | 0.0305300 | 87.57 | 0.0000 |
| shock | -0.124873 | 0.0346006 | -3.609 | 0.0003 |
| income | 0.000282184 | 7.10062e-005 | 3.974 | 0.0001 |
| Mean dependent var | 2.703974 | S.D. dependent var | | 0.938548 |
| Sum squared resid | 2633.179 | S.E. of regression | | 0.934227 |
| R^2 | 0.009842 | Adjusted R^2 | | 0.009186 |
| $F(2, 3017)$ | 14.99441 | P-value(F) | | 3.31e-07 |
| Log-likelihood | -4078.226 | Akaike criterion | | 8162.452 |
| Schwarz criterion | 8180.491 | Hannan-Quinn | | 8168.939 |

(b) 差分

モデル 2: Pooled OLS, 観測数: 3020
クロスセクションユニット数: 3020
時系列の長さ = 1
従属変数: d.life

| | 係数 | Std. Error | t-ratio | p 値 |
|--------------------|-------------|--------------------|---------|----------|
| const | 0.215365 | 0.0313550 | 6.869 | 0.0000 |
| shock | -0.140117 | 0.0484445 | -2.892 | 0.0039 |
| d.income | 0.000223286 | 0.000161409 | 1.383 | 0.1667 |
| Mean dependent var | 0.164238 | S.D. dependent var | | 1.310712 |
| Sum squared resid | 5168.730 | S.E. of regression | | 1.308893 |
| R^2 | 0.003434 | Adjusted R^2 | | 0.002773 |
| $F(2, 3017)$ | 5.197346 | P-value(F) | | 0.005581 |
| Log-likelihood | -5096.623 | Akaike criterion | | 10199.25 |
| Schwarz criterion | 10217.29 | Hannan-Quinn | | 10205.73 |