

The economic relation between the climate change and the production elasticity of cereals about each countries' macro data

January 20, 2009

Abstract

The production function of cereals regressed by cross-section macro data of each countries and their climate data is useful. That regression implied that the amount of cereal production is higher in the areas where climate is more stable, moister and more clement. If the climatic variation like abnormal weather becomes intense by global warming, it negatively affects the cereal production. The cereal production of the temperature doesn't always deteriorate if the cereal of production is not chosen in the global warming.

Keyword: Macroeconomics; Macro data; Cross section data; Microeconomics; Applied Econometrics; Agricultural economics; Geography; Production; Grain; Cereal; Climate; マクロ経済学のミクロ的基礎; 応用計量経済学; 農業経済学; 地理学; 穀物; 気候; 生産

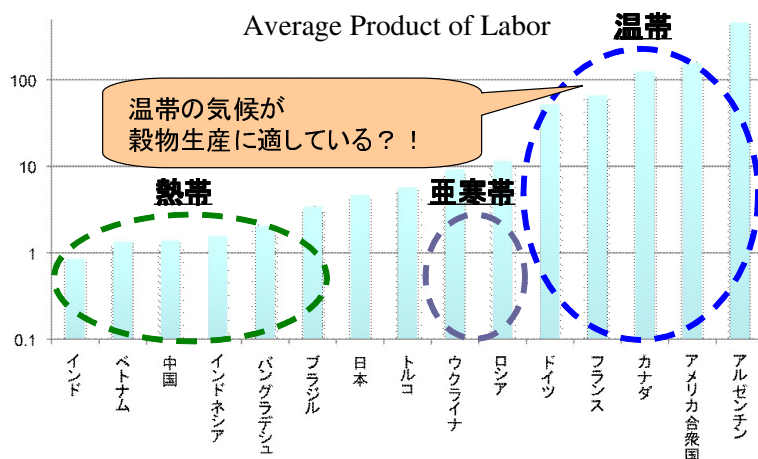
目次

1.動機と資料	3
2.仮説	5
3.先行研究	6
4.分析の方法	7
4.1.方針	
4.2.モデル式の導出	
4.3.気候因子の選択	
4.4.regression	
4.5.data	
4.6.適値と乖離	
4.7.結果と改善	
4.8 本研究と先行研究の比較	
4.9 仮説の判断	
5.分析の結果	23
5.1.OLS の Assumption	
5.2.方針と手順	
5.3.OLS の output	
5.4.homoskedasticity	
5.5.GLS の Assumption	
5.6.GLS の output	
5.7.他の Assumption の検証	
6.考察と結論	31
7.この研究の問題点	33
参考文献	34
R 言語のプログラム	36
Table1, 2, 3	

1. 動機と資料

●資料：総務省統計局(2006)および ILO(2006)（編集）

・・・一人当たりの穀物生産（ton/人）を比較（※縦軸は対数目盛）

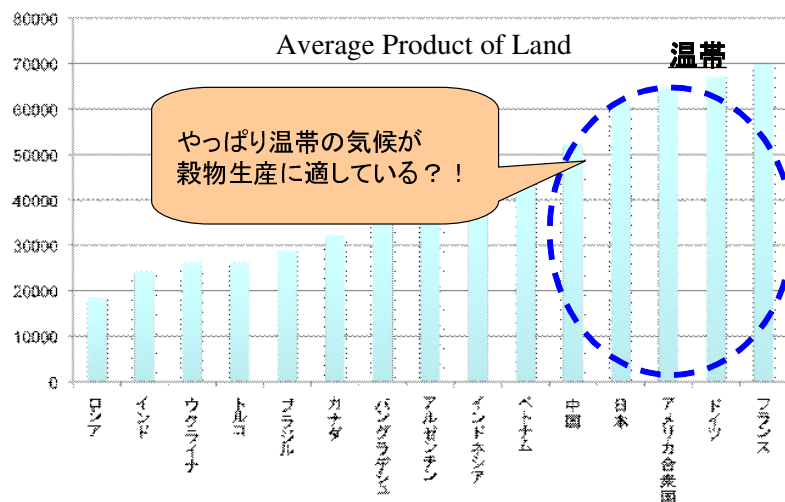


穀物は、米、小麦、大麦、らい麦、えん麦、とうもろこし、いも、ばれいしょ、大豆、かんしょ、落花生を採用。サンプル(国)は総務省統計局に掲載されている全ての国(15カ国)を掲載。総務省で得られた穀物生産量を、ILOで得られた農業・林業人口で割った。

(出所) 日本国総務省統計局 <http://www.stat.go.jp/>, ILO (国際労働機関) <http://laborsta.ilo.org/>

●資料：FAO(2005)（編集）

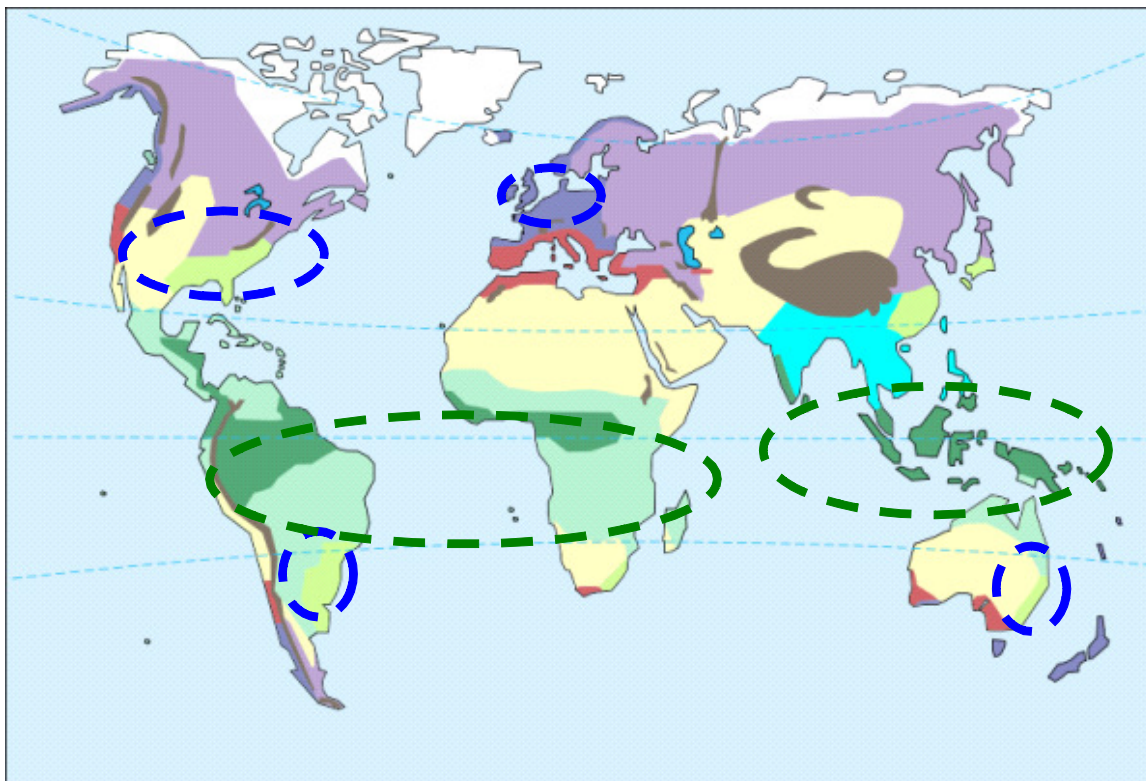
・・・一定面積当たりの穀物生産（hectogram/hectare）を比較



穀物は、大麦、そば、カナリア種子、フォニオ、トウモロコシ、ミレー、えん麦、キヌア、米、もみ、ライ麦、モロコシ、トリティカーレ、小麦、穀物を混ぜたもの、不特定な穀物を採用。サンプルは、APL の15カ国にそろえた。

(出所) FAO (国連食糧農業機関) <http://faostat.fao.org/>

資料：Köppen(1923)の区分によれば，気候は温度と降水量，その推移で決定される。



(出所) 気候区分の図は http://ja.wikipedia.org/wiki/%E7%94%BB%E5%83%8F:Climates_in_the_world.svg

青は土地生産性が高い地域 緑は土地生産性が低い地域

土地生産性が高い地域は

・温帯（一部，亜寒帯）・降水量が多すぎない・気温の変化が小さい
のではないかと？

2.仮説

仮説 1

→温度や降水量，その標準偏差は，穀物の生産額の説明変数として有効である．

仮説 2

→気候変動（温度や降水量の標準偏差）は小さい方（つまり気候が安定している方）が，穀物の生産により寄与する．

仮説 3

→気温よりも降水量の方が，穀物の生産により寄与する．

3. 先行研究

Furuya & Koyama (2005)

$$\ln YH_t = a + b_1 T + b_2 \ln PRC_t + b_3 \ln TMP_t$$

YH_t = Yield (APT)

TMP_t = Temperature

PRC_t = Precipitation

T_t = Time trend,

$$\ln YH_t = a + b_1 T + b_2 \ln PRC_t + b_3 \ln TMP_t$$

	Furuya & Koyama (2005)
Data	Micro Time Series Data
Regressand	Yield (APT)
Regressor	Temp., Prep., Time trend, Subsidy
Estimation	OLS, AR(1), AR(2), AR(3), AR(8)
# years	40
# Countries	14
# Cereals	5
Climate	Partial
備考	①それぞれの穀物、それぞれの国ごとに回帰分析を実施している。 ②温度・降水量データは、穀物ごとに地域を変え、また一部の時期のみ採用。 ③穀物・国ごとの実態に着目

4.分析の方法

4.1. 方針

重回帰分析を用いて、気候の指標に対する生産の弾力性を求める。

※弾力性

→ある指標が1%増加すれば、生産は何%増加するか？

弾力性を表す paramator を比較することで、仮説を検証

4.2. モデル式の導出

Assumption & Notation

The cereal is produced from the three inputs.

- (1) The production function is f .
- (2) The inputs are Capital ($\equiv K$), Land ($\equiv T$) and Labor ($\equiv L$),.
(•• The Agriculture sector of the specific factors model)
- (3) The output (cereal) is Quantity ($\equiv Q$).
- (4) The cereal market is perfect competitive, so the price of output (cereal) P is constant.
- (5) $K, T, L, P, Q \in R^+$

The production function is $Q = f(K, T, L) \quad \dots(3.1)$

Production function の型の選定

Production function の実証でよく使用されるものとして, CES 型, Translog 型, Cobb-Douglas 型の 3 つ挙げられる.

まず, Cobb-Douglas 型 (Cobb and Douglas[1928]) よりも複雑なものは, 以下の 2 つである.

- CES 型 (Arrow, Chenery, Minhas and Solow[1961])

生産要素の代替弾力性を一定にしたものである. CRS (1 次同時) であることが前提であるが, 今回, そのような仮定はおかないので, 使用できない.

- Translog 型 (Christensen, Jorgenson and Lau[1970])

Cobb-Douglas 型を通常 2 次の項までテイラー展開したものであり, regressor の数が膨大になる. 通常 K 個の regressor で済むところが, constant を除いて

${}_{K+1}H_2 - 1 = {}_{K+2}C_2 - 1 = \frac{(K+2)(K+1)}{2} - 1$ 個の regressor で分析しなくてはならない.

よって, Excel では regressor は 16 個までしか使えないので, この制限を越える可能性が大きくなる. 現に, 本論文では生産を説明する要素に 8 個考えるので, Translog 型では regressor が 44 個 (定数項を除く) となり, Excel が使用できない. また, White test を行うと更に増えてしまう.

以上より, 今回は単に Cobb-Douglas 型を使用することにする.

f is Cobb-Douglas function: $f(K, T, L) = AK^\alpha T^\beta L^\gamma$

The production function is

$$Q = f(K, T, L) \quad \dots(3.1)$$

f is Cobb-Douglas function:

$$f(K, T, L) = AK^\alpha T^\beta L^\gamma$$

α, β, γ

Substitute (3.1) for (3.2),

$$Q = AK^\alpha T^\beta L^\gamma$$

The both sides $\times P$,

$$PQ = PAK^\alpha T^\beta L^\gamma$$

$$PQ = PAK^\alpha T^\beta L^\gamma$$

Take natural log for the both sides,

$$\begin{aligned} \ln PQ &= \ln PAK^\alpha T^\beta L^\gamma \\ &= \ln PA + \ln K^\alpha + \ln T^\beta + \ln L^\gamma \\ &= \ln PA + \alpha \ln K + \beta \ln T + \gamma \ln L \end{aligned}$$

($\ln PA, \alpha, \beta, \gamma$ are constant.)

$$\ln PQ = \ln PA + \alpha \ln K + \beta \ln T + \gamma \ln L$$

Add other factors and rewrite the constant parameters,

$$\begin{aligned} \ln PQ = & \beta_1 + \alpha \ln K + \beta \ln T + \gamma \ln L \\ & + \beta_6 \ln x_6 + \beta_7 \ln x_7 + \dots \end{aligned}$$

(気候因子)

x_6 … 月降水量の標準偏差 x_8 … 年降水量

x_7 … 月平均気温の標準偏差 x_9 … 年平均気温

4.3. 気候因子の選択

(気候因子) X_8 ……年降水量 X_9 ……年平均気温

<気候因子の選択に対する検討1>

=「年降水量」「年平均気温」だと、値が高すぎたり、または低すぎたりして、生産が悪化することを考慮していないのではないか？適値を計算し、それとの乖離をregressorにすべきではないだろうか？=

○値が、ある穀物の適値と比べて仮に高すぎたり低すぎたりしていても、その値に適する穀物を生産するだけなので、「穀物」全体を対象にしている本研究では乖離を計算する必要性は高くない？

○穀物(というか植物)に厳密な適切な年間の値というものは存在しない。

→今回は、乖離を計算せず値をそのままregressorにした場合と、適値からの乖離をregressorにした場合の2つのregressionを行う。

(気候因子) X_6 ……月降水量の標準偏差 X_7 ……月平均気温の標準偏差

<気候因子の選択に対する検討2>

=気候変動の指標として、分散ではなく標準偏差を採用した理由=

・分散は本来の指標を2乗して計算されるので、sample間の差が著しく拡大してしまうのを防ぐため

・一般に変動の指標としては、分散より標準偏差が採用されているから
(例)ファイナンスのリスクの指標は標準偏差

・分散を採用しようが、標準偏差を採用しようが、Cobb-Douglas functionに natural logを取る場合には、係数の有意性やモデルの当てはまり方が変わることはない(Lemma 1 を参照)

Lemma 1

Cobb-Douglas function に natural log を取る場合には, regressor に, 分散を採用しようが, 標準偏差を採用しようが, 係数の有意性やモデルの当てはまり方が変わることはない.

(Proof)

Cobb-Douglas function

$$y = Ax^\alpha z^\beta \quad (x, y, z, \alpha, \beta, A \in R)$$

を regression することを考える.

これに natural log を取ると,

$$\ln y = \ln A + \alpha \ln x + \beta \ln z \text{ となる.}$$

ここで, $z = \sigma$ (標準偏差) とすると,

$$\ln y = \ln A + \alpha \ln x + \beta \ln \sigma \text{ となる.} \quad \cdots(4.3.1)$$

次に, $z = \sigma^2$ (分散) とすると,

$$\ln y = \ln A + \alpha \ln x + \beta \ln \sigma^2$$

$$= \ln A + \alpha \ln x + 2\beta \ln \sigma \text{ となり, } \gamma = 2\beta \text{ とすれば,}$$

$$\ln y = \ln A + \alpha \ln x + \gamma \ln \sigma \text{ となる.} \quad \cdots(4.3.2)$$

ここで, (4.3.1)と(4.3.2)を比較すれば, 一部の係数以外, variable も含めて式の形状は全く同じである. よって, 係数の値が $\gamma = 2\beta$ の関係になる以外は, regression の結果は全く同じになる.

(気候因子) X_6 … 月降水量の標準偏差 X_8 … 年降水量

<気候因子の選択に対する検討3>

=日照時間はregressorに必要ではないか？=

- ・日照時間は、降水量と強い負の関係にあるはず。
→regressorは互いに、相関関係があってはならない(multi-collinearity: Assumption__を参照)ので、日照時間を入れてはならない。
- ・本研究は、気候が生産に及ぼす影響を分析しているが、Köppen(1923)の定義では、気候は気温と降水量のみで決まるので、入れなくて良い。
- ・そもそも日照時間のデータが見つからなかった…。

4.4. regression

- ・線形重回帰分析

<被説明変数: regressand>

穀物生産額の自然対数の値 ($\ln PQ$)

<説明変数: regressor>

x_i (気候因子, $i = 6, 7, 8, 9$) 及び,

土地面積 (T), 農業従事者数 (L), 農業資本 (K) の自然対数の値

前述の式 (モデル式) の変数の数値を, 世界の国と地域の統計から, sample を集めて分析

4.5. data

＝穀物に関するデータ＝

FAO（国連食糧農業機関） <http://faostat.fao.org/> にて、各国の穀物の Gross および Net の生産額 (PQ), 穀物の栽培面積 (T), 農業従事者数 (L), 農業資本 (K) のデータを採用

穀物の定義

国際連合食糧農業機関（FAO）において、Cereals として統計に参入されているものを、本論文における穀物と定義する。よって、以下の 15 種となる。

Table 穀物の定義

English	日本語
Barley	大麦
Buckwheat	そば
Canary seed	カナリア種子
Fonio	フォニオ
Maize	トウモロコシ
Millet	ミレー
Oats	燕麦（えん麦）
Quinoa	キヌア
Rice, paddy	ライス, もみ
Rye	ライ麦
Sorghum	モロコシ
Triticale	マカロニコムギ
Wheat	小麦
Mixed grain	穀物を混ぜたもの
Cereals, nes	不特定な穀物

4.6. 適値と乖離

適値の算出方法

本項では、各国の穀物生産において最適な年間降水量と年平均気温を算出した方法について説明する。

(1) それぞれの穀物の適値

様々な穀物生産に関する農業の書籍・論文を参照したが、多くの穀物で明確な適値は存在しない。そのような穀物においては、適するとされる気候のモデルとなる年間降水量・年平均気温を当てはめた。実際には複数の気候で適する場合が多いので、降水量は 250 の倍数、気温は 2.5 の整数倍から最適と判断したものを適温とした。但し、より良い適値が別の方法で得られた場合には、それを採用する。（「Table3: 穀物の最適年間降水量と最適年平均温度」を参照せよ。）

(2) 各国において生産している穀物の価格と気温のデータの取り扱い

各国の最適気温および最適降水量を算出するため、各穀物の 3 カ年（基本的には、1999 年から 2001 年）の価格と生産量を求めなければならない。本項ではその際、生じた 8 つの問題の取り扱いを説明する。

1. Belgium と Luxembourg について

2000 年と 2001 年のデータが分離されている。土地・資本などその他のデータは全て統合されているので、この価格と生産量のデータも統合しなければならない。両年については、まずはそれぞれの国のそれぞれの財で、 Σ （各年の価格×各年の生産量）を計算し、各穀物について、それぞれ両国のものを合計するという手法を採った。

2. Georgia の Buckwheat

1999年のデータ（価格、生産量ともに）が存在しないので、1999年のデータを全穀物、2002年のデータに入れ替えた。

3. Czech の Buckwheat

全ての年で価格のデータが存在しない。チェコと同じような経済水準・気候かつ地理的にも近いスロバキアとハンガリーと比較してみたが、スロバキアの方が他の穀物の価格動向が似ていた。よって、スロバキアのそれぞれの年のデータを採用した。

4. Estonia の Triticale

1999年と2000年のTriticaleのデータがないので、1999年と2000年のデータを全品目、2002年と2002年のデータに入れ替えた。

5. Bahamas, Liberia, Cape Verde, Djibouti, Solomon Islands

価格のデータが一切存在しない。しかし、穀物が1財しかなく、そもそも価格が要らないので、特に求めないことにする。

6. Viet Nam

価格のデータが一切存在しない。周辺の東南アジア諸国の価格データを見たところ、平均すると大体、穀物3つが同じような価格になることから、価格比を同じにした。

7. その他

以下の国（"→"の前の国）も価格のデータが一切存在しない。

Afghanistan → Tajikistan

Angola → Namibia

Central African Republic → Cameroon

Chad → Cameroon
Fiji → Philippines
Gabon → Congo
Guinea-Bissau → Guinea
Guyana → Suriname
Guatemala → Mexico
Korea, Democratic People's Republic of → China
Libyan Arab Jamahiriya → Algeria
Mauritania → Mali
Oman → Yemen
Papua New Guinea → Philippines
Senegal → Guinea
Sierra Leone → Guinea
Tanzania, United Republic of → Mozambique
Uzbekistan → Kyrgyzstan
Zambia → Zimbabwe

そこで、3カ年の Price のデータを一括して、他の国（"→"の後の国）に入れ換えた。その際の選定の基準は以下となる。

- * 同じような経済水準であること。
- * 経済交流が盛んであること。
- * 気候が似ていること。
- * 食文化が似ていること。
- * 貿易の観点から地形が似ていること（例：同じく内陸国である）。
- * 地理的に近いこと。
- * 入れ換えられる国が生産している穀物を、できるだけ多くの種類において生産していること。かつ、それらの種類のその国における share ができるだけ大きいこと。

以上の基準の上で、入れ換えた国で生産されていない穀物がある場合には、その穀物を data set から消去して考える。

8. Cereals, nes と mixed grain

適温、適降水量を求めるのは不可能なので無視する。

(3) 各国の適値の算出

(1) によって求められた穀物 i の最適な年間降水量と年平均気温をそれぞれ $R_i, T_i (\in R_+)$ とする。また、各国における穀物 i の j 年度における価格と生産量をそれぞれ、 $P_{ij}, Q_{ij} (\in R_+)$ とする。つまり、生産額は $P_{ij}Q_{ij}$ となるが、3年間（基本的には 1999 年から 2001 年まで）の平均値は、 $\frac{1}{3} \sum_{j=1}^3 P_{ij}Q_{ij}$ となり、穀物全体の

生産額の年平均は、 $\frac{1}{3} \sum_t \sum_{j=1}^3 P_{ij}Q_{ij}$ となる。これを用いた、穀物 i の生産額におけ

る国内の生産シェアは $\frac{\sum_{j=1}^3 P_{ij}Q_{ij}}{\sum_t \sum_{j=1}^3 P_{ij}Q_{ij}}$ となる。 $P_{ij}, Q_{ij} (\in R_+)$ より、これは 0 から 1

の間をとり、かつ、 $\sum_t \frac{\sum_{j=1}^3 P_{ij}Q_{ij}}{\sum_t \sum_{j=1}^3 P_{ij}Q_{ij}} = \frac{\sum_t \sum_{j=1}^3 P_{ij}Q_{ij}}{\sum_t \sum_{j=1}^3 P_{ij}Q_{ij}} = 1$ となる。このシェアを使って、

降水量や気温の加重平均を求めることによって、適値を算出すれば、最適年間降水量は、

$$\sum_t \frac{\sum_{j=1}^3 P_{ij}Q_{ij}}{\sum_t \sum_{j=1}^3 P_{ij}Q_{ij}} R_t = \frac{\sum_t \sum_{j=1}^3 R_t P_{ij}Q_{ij}}{\sum_t \sum_{j=1}^3 P_{ij}Q_{ij}} \text{ となる。同様に最適年平均気温は、 } \frac{\sum_t \sum_{j=1}^3 T_t P_{ij}Q_{ij}}{\sum_t \sum_{j=1}^3 P_{ij}Q_{ij}} \text{ と}$$

なる。

4.7. 結果と改善

線形重回帰分析で2つの改善方法

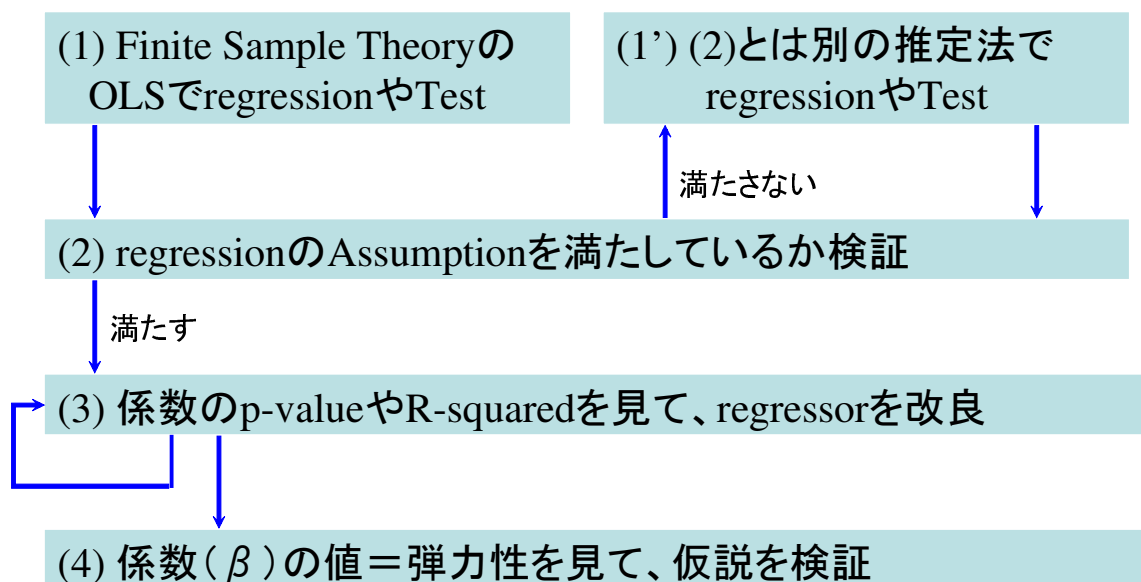
○まずは Finite-Sample Theory における Ordinary Least Squares = OLS（通常最小二乗法）で regression や Test を行う

→しかし、OLS を行うには、いくつかの Assumption を満たしていることが前提
→regression の結果、もし Assumption を満たしていないなら、Large-Sample Theory を採用したり、別の推定法（GLS や GMM など）を使う必要がある。

○決定係数（R-squared）と、regressors の t-value を検証

・・・もし当てはまりが悪ければ、改善（指標の求め方がミスかも）または、外して再度回帰分析を行う
→繰り返して改善する。

• 分析のアルゴリズム



●係数 β は何を表すか??

$$\ln PQ = \beta_1 + \beta_2 \ln K_1 + \beta_3 \ln K_2 + \beta_4 \ln T + \beta_5 \ln L + \beta_6 \ln x_6 + \dots$$

これは1次式(linear)なので、全微分(明らかに可能)して、

$$d \ln PQ = \beta_2 d \ln K_1 + \beta_3 d \ln K_2 + \beta_4 d \ln T + \beta_5 d \ln L + \beta_6 d \ln x_6 + \dots$$

仮に x_6 以外は不変だとすると、それらの natural log は0

$$d \ln PQ = \beta_6 d \ln x_6 \Leftrightarrow \beta_6 = \frac{d \ln PQ}{d \ln x_6}$$

<復習> 対数の微分公式: $\frac{d}{dx} \ln x = \frac{1}{x} \Leftrightarrow d \ln x = \frac{dx}{x} \approx \frac{\Delta x}{x}$

$$\beta_6 = \frac{d \ln PQ}{d \ln x_6} \approx \frac{\Delta PQ / PQ}{\Delta x_6 / x_6}$$

これは、 x_6 の生産額に対する弾力性と等しい。

∴ 係数 β は、その説明変数の生産額に対する弾力性と等しい

4.8. 本研究と先行研究の比較

Table 4: Furuya & Koyama (2005) V.S. Fujishiro (2009)

Furuya & Koyama (2005): $\ln YH_t = a + b_1 T + b_2 \ln PRC_t + b_3 \ln TMP_t$

Fujishiro (2009): $\ln PQ = \beta_1 + \beta_2 \ln K_1 + \beta_3 \ln K_2 + \beta_4 \ln T + \beta_5 \ln L + \beta_6 \ln x_6 + \dots$

Furuya & Koyama (2005)	V.S.	Fujishiro & Kohno (2009)
Micro Time Series Data	Data	Macro Cross Section Data
Yield (APT)	Regressand	Net/Gross Price×Quantity
Temp., Prep., Time trend, Subsidy (only U.S. and EU)	Regressor	Capital, Land, Labor, Temp., Prep., these deviations
OLS, AR(1), AR(2), AR(3), AR(8)	Estimation	OLS, GLS
40	# years	3, 30
14	# Countries	126, 98
5	# Cereals	15
Partial	Climate	Annual
① それぞれの穀物，それぞれの国ごとに回帰分析を実施している。 ② 温度・降水量データは，穀物ごとに地域を変え，また一部の時期のみ採用。 ③ 穀物・国ごとの実態に着目	備考	① 穀物全体・世界全体で回帰分析を実施 ② 温度・降水量データは，適値からの乖離も考慮 ③ 気候（年間の温度・降水量）が生産関数に及ぼす影響に着目

4.9. 仮説の判断

•仮説 1

温度や降水量，その標準偏差は，穀物の生産額を説明する変数として有効である。

→有効であることは以下の2つが必要条件

回帰分析の結果，

(1)得られた **R-squared が 1 に近**ければ，モデル式全体はデーターに対して当てはまりがよい。

(2)それぞれの **regressor の t-value が有意**であれば，その変数は有効。

•仮説 2

気候変動（温度や降水量の標準偏差）は小さい方（つまり気候が安定している方）が，穀物の生産により寄与する。

→温度変動の弾力性 β_7 と降水量変動の弾力性 β_6 の正負によって判断

→**負なら，その分散が減少（変動が小さい）すれば，生産は増加すると確認できる。**

•仮説 3

気温よりも降水量の方が，穀物の生産により寄与する。

→気温と降水量の係数（弾力性）の大きさの比較によって判断

→**降水量の係数（弾力性）が（正の方向に）大きければ，降水量の方が生産に対して正のインパクトが大きいと確認できる。**

乖離の場合は負の方向に大きければ，同様にインパクトが大きいと考える。

5.分析の結果

5.1 . OLS (Ordinary Least Square)の Assumption

以上の仮定に基づいて, $\{\beta_1, \beta_2, \dots, \beta_K\}$ に相当するとされる OLS estimator が求まり, その estimator に対する Test (検定) が行われる.

Assumption in Finite-Sample Theory

以下, $y_i, x_{ik}, \varepsilon_i (i = 1, 2, \dots, n; k = 1, 2, \dots, K)$ をそれぞれ, regressand, regressor, error term とする (K は regressor の数, n は sample の数である).

(1) (linearity) $y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_K x_{iK} + \varepsilon_i$

(2) (strict exogeneity) $E(\varepsilon_i | x_1, x_2, \dots, x_K) = 0$

(3) (no multicollinearity) $\{x_1, x_2, \dots, x_K\}$ is linearly independent.

(4) (homoskedasticity) $E(\varepsilon_i^2 | x_1, x_2, \dots, x_K) = \sigma^2$

(5) (no correlation between observations)

$$E(\varepsilon_i \varepsilon_j | x_1, x_2, \dots, x_K) = 0 (i, j = 1, 2, \dots, n; i \neq j)$$

(6) (normality of the error term) $\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n | x_1, x_2, \dots, x_K\} \sim N(0, \sigma^2)$

Assumption of Large Sample Theory by OLS

Large Sample Theory では, "linearity", "strict exogeneity", "normality of the error term"が不要 (但し, ergodicity など別の Assumption が必要になる).

しかし, 後に述べる homoskedasticity を Test するための White Test では Large-Sample Theory に基づいて, OLS estimate の際に観測される deviation (残差) を求めている.

そのため, 本研究では linearity, exogeneity, normality が violate されているかどうかに関わらず, Finite-Sample Theory における t-Test や Wald-Test を実施している.

5.2. 方針と手順

.

(1) 今回は、乖離を計算せず値をそのまま regressor にした場合と、適値からの乖離を regressor にした場合の2つの regression を行う。

(2) regressand (PQ) として、Net production と Gross Production が考えられる。両方で行い、当てはまり方を比較する。

∴同一の sample, 同一の推定法 (Estimation) に対して、 $2 \times 2 = 4$ 通りの regression を行い、それぞれを比較する。

★R 言語で作成したプログラムで regression および Test

Step1: sample 未選択 (sample size:126)

←Combine データがない場合の自然対数の値を 0 にしたのがまずいかも

↓

Step2: Combine データがない Sample(後述)を全て除いた (sample size:98)

除かれた国：

Afghanistan, Angola, Barbados, Belarus, Bolivia, Cameroon, Cape Verde, Djibouti, Dominican Republic, Fiji, Gabon, Guinea-Bissau, Honduras, Jamaica, Korea, Democratic People's Republic of, Lao People's Democratic Rep., Liberia, Malawi, Mauritius, Mozambique, Namibia, Niger, Nigeria, Rwanda, Solomon Islands, Tanzania, United Republic of, Togo, Trinidad and Tobago

・・・Combine を使えない穀物の熱帯の国や発展途上国が多い。

←生産性が著しく悪化する??

5.3. OLS の output

Table 1 および Table 2 も参照せよ.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Regressand	Net	Gross	Net	Gross	Net	Gross	Net	Gross
Sample size	126	126	126	126	98	98	98	98
R squared	0.962	0.964	0.959	0.961	0.948	0.951	0.945	0.949
Adjusted R squared	0.960	0.961	0.956	0.959	0.943	0.947	0.940	0.944
F-Test	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Wald statistic	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
const.	-3.462	-3.321	<u>-0.541</u>	<u>-0.499</u>	-3.211	-3.043	<u>-0.858</u>	<u>-0.811</u>
Capital (Combine)	0.058	0.057	0.066	0.065	0.080	0.077	0.093	0.090
Capital (Tractor)	0.097	0.096	0.120	0.119	0.108	0.109	0.121	0.121
Land	0.899	0.902	0.889	0.894	0.907	0.910	0.906	0.909
Labor	<u>0.039</u>	<u>0.034</u>	<u>0.017</u>	<u>0.012</u>	<u>0.018</u>	<u>0.013</u>	<u>-0.007</u>	<u>-0.011</u>
Prep.s.d.	-0.401	-0.386	<u>-0.020</u>	<u>-0.019</u>	-0.370	-0.351	<u>-0.059</u>	<u>-0.058</u>
Temp. s.d.	<u>-0.102</u>	<u>-0.090</u>	-0.226	-0.207	-0.160	-0.149	-0.285	-0.265
Annual prep.	0.451	0.439	—	—	0.348	0.330	—	—
Average temp.	<u>0.213 *</u>	<u>0.187 *</u>	—	—	0.259	0.233	—	—
Deviation of prep.	—	—	-0.074	-0.080	—	—	-0.026	-0.031
Deviation of temp.	—	—	<u>-0.093</u>	<u>-0.087</u>	—	—	-0.127	-0.119
White test	0.087	0.074	0.030	0.030	0.085	0.064	0.058	0.047

・式の当てはまりはかなり良い

(R-squared, Adjusted R-squared, F-Test, Wald-Test). ←Production に対する Land の貢献がかなり高いから？

・regressand において, Net Production より, Gross Production の方が当てはまりが良い. ←Gross の方が実際の値に近いから？

・Labor は有意でない. ←そもそも穀物の労働力ではないから？

・乖離を入れたとき, 降水量の標準偏差が有意でなくなる (R-squared も若干悪化する).

・Combine が存在しない国を除いた sample size 98 の方が, より有意になる. ←実際の値に近いから？

5.4. homoskedasticity

$$(4)(\text{homoskedasticity}) \quad E(\varepsilon_i^2 | x_1, x_2, \dots, x_K) = \sigma^2$$

White Test によって判断される。以下、この方法を述べる。
まず、 ε_i^2 を他の regressor とその全ての second moment ($x_i x_j (i, j = 1, 2, \dots, n)$) で regress する(つまり、このときの regressor の総数は $K+1$ H_2 となる)。その regression における R-squared を R^2 とすると、White test では、homoskedasticity が violate されればされるほど、 nR^2 は、degree of freedom (自由度) $K+1$ H_2 の χ^2 -distribution に従うとされる。つまり、null hypothesis が accept される方が良く、そのためには p-Value が大でなければならない。

ちなみに、 ε_i^2 は実際には観測不可能であるが、White Test では Large-Sample Theory に基づいて、OLS estimate の際に観測される deviation (残差) を使用している。そのため、本研究では(1) (linearity) (2) (strict exogeneity) (6) (normality of the error term) が violate されているかどうかに関わらず、Finite-Sample Theory における t-Test や F-Test を実施している。

→White Test を実施

confidence interval 95% で null を accept する。しかし confidence interval 90% で null を reject される場合がある。

→OLS で誤りと言えないが、疑義が生じている。

→GLS を行うべき。

(その他の Assumption は後で検証)

5.5. GLS (Generalized Least Square)の Assumption

GLS では、(1) OLS in Finite-Sample Theory における Assumption のうち、(1) (linearity) (2) (strict exogeneity) (3) (no multicollinearity)が成立していれば、estimator は有効 (正確には unbiased で efficient) となる。本研究では、Time Series Data を扱わないので、(5) (no correlation between observations)を前提にした Eicker-White の方法で GLS estimator を求めた。

ちなみに、 ε_i^2 は実際には観測不可能であるが、Eicker-White の方法では、一部、Large-Sample Theory に基づいて estimate を行っている。このため、White test の箇所でも述べたが、本研究では、(1) (linearity) (2) (strict exogeneity) (6) (normality of the error term)が violate されているかどうかに関わらず、Finite-Sample Theory における t-Test や F-Test を OLS と同様、GLS に対しても実施している。手順は OLS の場合と同様とする。

5.6. GLS の output

Table 1 および Table 2 も参照せよ.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Regressand	Net	Gross	Net	Gross	Net	Gross	Net	Gross
Sample size	126	126	126	126	98	98	98	98
R squared	0.962	0.964	0.959	0.961	0.948	0.951	0.945	0.948
Adjusted R squared	0.960	0.961	0.956	0.958	0.943	0.947	0.940	0.944
F-Test	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Wald statistic	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
const.	-3.677	-3.592	<u>-0.529</u>	<u>-0.407</u>	-3.149	-2.971	<u>-0.765</u>	<u>-0.639</u>
Capital (Combine)	0.058	0.059	0.067	0.067	0.074	0.076	0.083	0.084
Capital (Tractor)	0.104	0.103	0.121	0.122	0.107	0.106	0.129	0.128
Land	0.895	0.893	0.897	0.900	0.923	0.924	0.904	0.902
Labor	<u>0.032</u>	<u>0.033</u>	<u>0.001</u>	<u>0.003</u>	<u>0.006</u>	<u>-0.001</u>	<u>-0.010</u>	<u>-0.008</u>
Prep.s.d.	-0.415	-0.403	<u>-0.017</u>	<u>-0.034</u>	-0.381	-0.361	<u>-0.060</u>	<u>-0.081</u>
Temp. s.d.	<u>-0.101</u>	<u>-0.082</u>	-0.245	-0.234	-0.171	-0.165	-0.273	-0.266
Annual prep.	0.479	0.480	—	—	0.340	0.327	—	—
Average temp.	0.254	0.225	—	—	0.250	0.215	—	—
Deviation of prep.	—	—	<u>-0.070*</u>	-0.087	—	—	-0.031	-0.035
Deviation of temp.	—	—	<u>-0.106</u>	<u>-0.100</u>	—	—	-0.151	-0.143

Test や R-squared は OLS とほとんど同じ.

当てはまりはよいが、いくつかの係数で有意でない.

但し、OLS と比べて、有意になる係数（年平均温度）が増えた.

5.7. 他の Assumption の検証

2. Assumption に対する Test

$$(1)(\text{linearity}) \quad y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_K x_{iK} + \varepsilon_i$$

今回,

$$\ln PQ = \beta_1 + \beta_2 \ln K_1 + \beta_3 \ln K_2 + \beta_4 \ln T + \beta_5 \ln L + \beta_6 \ln x_6 + \cdots$$

とモデルを仮定しているので, OK.

$$(2)(\text{strict exogeneity}) \quad E(\varepsilon_i | x_1, x_2, \dots, x_K) = 0$$

$E(\varepsilon_i | x_1, x_2, \dots, x_K) = 0$ を null hypothesis とする t-Test を行う. これが accept される, つまり, p-Value が大であることが好ましい.

→accept されたので, 問題なし.

$$(3)(\text{no multicollinearity}) \quad \{x_1, x_2, \dots, x_K\} \text{ is linearly independent.}$$

ひとつの regressor を他の regressors で regression してみて, R-squared (以下, R^2 とする) が高ければ, regression 同士で相関が発生 (multicollinearity) していることになるとなる. この場合, estimator が正しい値を取らなかつたり, 有意になりにくくなつたりするので問題である. 一般には,

$$VIF \equiv \frac{1}{1 - R^2}$$

で定義される VIF (Variance Inflation Factor) が 10 を越えたら multicollinearity が発生していると考え. この VIF は, regressor それぞれにおいて存在するので, ひとつの regressors の set につき, その regressor の数の分だけ求めなければならぬ.

→結果, Step1 2 どちらにおいても, 10 を越える VIF は存在しなかつた.

∴ OK.

(5)(no correlation between observations)

$$E(\varepsilon_i \varepsilon_j | x_1, x_2, \dots, x_K) = 0 \quad (i, j = 1, 2, \dots, n; i \neq j)$$

本来は Durbin-Watson Test を実施する。しかし、Time Series Data ではなく、今回のように時期を固定した Cross section Data を用いた分析では、通常、この Assumption が violate されるとは考えないので、Test を行わない。

(6) (normality of the error term) $\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_K | x_1, x_2, \dots, x_K\} \sim N(0, \sigma^2)$

normality (正規性) の Test には、Shapiro-Wilk test を実施する。これは、normality に従うことを null hypothesis とした Test である。つまり、null hypothesis が accept される方が良く、そのためには p-Value が大でなければならない。

→Shapiro-Wilk test で null hypothesis が accept

→問題無し

6.考察と結論

=仮説 1 =

温度や降水量，その標準偏差は，穀物の生産額を説明する変数として有効

(1)R-squared がほとんど 0.95 を越えており，F-Test や Wald-Test も圧倒的に有意なので，式の当てはまり方は良い

(2)各係数において・・・

- ・ Combine, Tractor, 土地, 年間降水量は全ての場合で有意
- ・ 平均気温も適降水量からの乖離も，

Large Sample Theory まで拡大すれば有意.

- ・ Labor は全然ダメ

←やはり，農業従事労働者を穀物生産の労働者とするのは無理？

- ・ その他は，場合による.

=仮説 2 =

気候変動（温度や降水量の標準偏差）は小さい方（つまり気候が安定している方）が，穀物の生産により寄与する.

→温度変動の弾力性 β_7 と降水量変動の弾力性 β_6 の正負によって判断

→負なら，その分散が減少（変動が小さい）すれば，生産は増加すると確認できる.

- ・ 温度変動の弾力性 β_7 と降水量変動の弾力性 β_6 のどちらも負

→ 仮説は実証された！！

=仮説 3 =

気温よりも降水量の方が，穀物の生産により寄与する.

→気温と降水量の係数（弾力性）の大きさの比較によって判断

→降水量の係数（弾力性）が（正の方向に）大きければ，降水量の方が生産に

対して正のインパクトが大きいと確認できる.

→実際は, 降水量の係数の方が (正の方向に) 大

∴仮説は否定された.

→しかし, 乖離では, 温度の方が (負の方向に) 大

∴これは, 上の否定と矛盾

→本研究では, この仮説が正しいか間違っているか分からない.

=その他=

・ regressand において, Net Production より, Gross Production の方が当てはまりが良い. ←Gross の方が実際の値に近いから?

・ 乖離を入れたとき, 降水量の標準偏差が有意でなくなり, 逆に, 温度の標準偏差の有意さは向上する (R-squared は若干悪化する).

・ 降水量・温度とその乖離では, 前者の方が有意 (乖離を入れる意味はない??).

Combine が存在しない国を除いた sample size 98 の方が, より有意になる. ←実際の値に近いから

・ Combine よりも Tractor の方が寄与度が高い.

=Implication=

(1) 穀物生産額が高い条件は以下

気候変動の係数は負 = 気候が安定

温度や降水量の係数が正 = 温暖かつ湿潤

乖離の係数が負 = 温度や降水量が高すぎても低すぎてもダメ

→温暖がちょうど良い?

←実際の統計と整合性が高い.

(2) 地球温暖化により

異常気象など, 気候変動が激しくなれば穀物生産に悪影響.

温度の上昇しても, 穀物の種類を選ばなければ, 穀物生産が悪化することはない.

7.この研究の問題点

データーに対し，多少，手を加えている．

→実際，良い結果が出ていない．

適値の計算が厳密でない．

sample に入れている年が限られている（1999-2001）．

労働者のデーターがいかげん．

→実際，良い結果が出ていない．

References

Furuya, Jun and Osamu Koyama, "Impacts of Climate Change on World Agricultural Products Markets: Estimation of Macro Yield Functions," JARQ, 39, pp. 121-134, 2005

小西葉子, 『生産関数の特定化に関する統計的推測』. 経済科学, 51(2), 2003

小西葉子, 『わが国における地域別社会資本投資の生産性と効率性』. 経済科学, 52(3), pp. 75-86, 2004

小西葉子, 『生産関数の推定における人的資本の代理変数』. 経済科学, 50(4), pp. 83-95,, 2003

BOUBAKER Mohsen, YAMADA Toshiaki, "Screening Spring Wheat Genotypes (Triticum sp.) for Seedling Emergence under Optimal and Suboptimal Temperature Conditions," Japanese journal of breeding, Vol.41, No.2, pp. 381-387, 1991

MURAI HIROAKI, TAKAHASHI KIYOSHI, MASUI TOSHIHIKO, HARASAWA HIDEO, MATSUOKA YUZURU, "RESEARCH ON ASSESSMENT OF CLIMATE CHANGE IMPACTS ON GRAIN PRODUCTION TAKING ACCOUNT OF ADAPTATION," Environmental Systems Research, VOL.33, pp. 97-104, 2005

Hayashi, Fumio, "Econometrics," Princeton University Press, 2000

James D. Hamilton, "Time Series Analysis," Princeton University Press, 1994

E. Berndt, "The Practice of Econometrics," Addison-Wesley, 1991

W. Greene, "Econometric Analyses, fourth ed," Prentice-Hall, 1999

David A. Besanko, Ronald Braeutigam, "Microeconomics: An Integrated Approach,"
John Wiley & Sons Inc, 2001

山本拓, 『計量経済学』. 新世社, 1995

森棟公夫, 『統計学入門』. 新世社, 2000

田中勝人, 『基礎コース統計学』. 新世社, 1998

依田高典, 『トラスログ費用関数の推定』. 2001

Laurits R. Christensen, Dale W. Jorgenson, Lawrence J. Lau, "Transcendental
Logarithmic Production Frontiers Transcendental Logarithmic Production Frontiers,"
The Review of Economics and Statistics, Vol. 55, No. 1, pp. 28-45, 1973

小西葉子, 西山慶彦, 『生産関数のノンパラメトリック統計解析』. 統計関連学
会連合大会記録(日本統計学会第71回大会), 2003

日本国総務省統計局, <http://www.stat.go.jp/>, 2008

日本国気象庁, <http://www.jma.go.jp/jma/>

FAOSTAT, <http://faostat.fao.org/>, 2008

ILO, <http://www.ilo.org/>, 2008

帝国書院編集部編, 『新詳地理資料』. 帝国書院, 1997

帝国書院編集部編, 『新詳高等地図 初訂版』. 帝国書院, 1997

本研究の regression で使用した R 言語のプログラム

regressand が net production で、乖離を考慮しない sample size が 98 の場合

<input>

```
#normal_net_98
Y<-YN_98 #regressand
X<-X_98_normal #regressor

n<-nrow(Y) #sample size
K<-ncol(X) #number of regressor including const.
b<-solve(t(X)%*%X)%*%t(X)%*%Y #parameter of OLS
e<-Y-X%*%b #error term of OLS
sigma<-e%*%t(e) #making variance matrix of GLS st no-serial correlation
sigma[upper.tri(sigma)]<-0 #making variance matrix of GLS st no-serial correlation
sigma[lower.tri(sigma)]<-0 #making variance matrix of GLS st no-serial correlation
V<-sigma #variance matrix of GLS st no-serial correlation
C<-V^(-1/2) #C matrix of GLS st no-serial correlation
C[upper.tri(C)]<-0 #C matrix of GLS st no-serial correlation
C[lower.tri(C)]<-0 #C matrix of GLS st no-serial correlation
Y_G<-C%*%Y #regressand of GLS
X_G<-C%*%X #regressor of GLS
b_G<-solve(t(X_G)%*%X_G)%*%t(X_G)%*%Y_G #parameter of GLS
e_G<-Y-X%*%b_G #error term of GLS
R2<-1-(colSums(e*e))/(colSums((Y-colMeans(Y))^2)) #R_squared of OLS
A_R2<-R2-(K-1)*(1-R2)/(n-K) #adjusted_R_squared of OLS
R2_G<-1-(colSums(e_G*e_G))/(colSums((Y-colMeans(Y))^2)) #R_squared of GLS
A_R2_G<-R2_G-(K-1)*(1-R2_G)/(n-K) #adjusted_R_squared of GLS
XX<-solve(t(X)%*%X) #solve(t(X)%*%X)
XX[upper.tri(XX)]<-0 #making diagonal matrix
XX[lower.tri(XX)]<-0 #making diagonal matrix
XX<-XX%*%matrix(1,nrow=K,ncol=1) #change diagonal matrix into column vector
s2<-t(e)%*%e/(n-K) #variance of sample error
s2<-s2[1,1] #variance of sample error
t_V_f<-b/((s2*XX)^(1/2)) #t-Value in Finite-Sample Theory
p_V_t_f<-2*(1-pt(abs(t_V_f),n-K)) #p-Value of t-Test in Finite-Sample Theory
s2_G<-t(e_G)%*%e_G/(n-K) #variance of sample error of GLS
s2_G<-s2_G[1,1] #variance of sample error of GLS
t_V_f_G<-b_G/((s2_G*XX)^(1/2)) #t-Value of GLS in Finite-Sample Theory
p_V_t_f_G<-2*(1-pt(abs(t_V_f_G),n-K)) #p-Value of t-Test of GLS in Finite-Sample Theory
O_K_1<-matrix(0,nrow=K-1,ncol=1) #making linear hypotheses
I_K_1<-diag(K-1) #making linear hypotheses
R<-cbind(O_K_1,I_K_1) #making linear hypotheses
r<-O_K_1 #making linear hypotheses
E_Var_b<-s2*solve(t(X)%*%X) #natural estimator of Var(b|X) of OLS
F_r_f<-t(R%*%b-r)%*%solve(R%*%E_Var_b%*%t(R))%*%(R%*%b-r)/(K-1) # F-ratio of OLS in
Finite-Sample Theory
p_V_F_f<-1-pf(F_r_f,K-1,n-K) #p-Value of F-Test of OLS in Finite-Sample Theory
E_Var_b_G<-s2_G*solve(t(X)%*%X) #natural estimator of Var(b|X) of GLS
```

```

F_r_f_G<-t(R%*%b_G-r)%*%solve(R%*%E_Var_b_G%*%t(R))%*%(R%*%b_G-r)/(K-1) # F-ratio
of GLS in Finite-Sample Theory
p_V_F_f_G<-1-pf(F_r_f_G,K-1,n-K) #p-Value of F-Test of GLS in Finite-Sample Theory
B<-V #diagonal matrix of e2 of OLS
S<-(t(X)%*%B%*%X)/n #data matrix of S of OLS
SXX<-(t(X)%*%X)/n #data matrix of SXX
Avar_b<-solve(SXX)%*%S%*%solve(SXX) #asymptotic variance of OLS estimator
Avar_b[upper.tri(Avar_b)]<-0 #asymptotic variance of OLS estimator
Avar_b[lower.tri(Avar_b)]<-0 #asymptotic variance of OLS estimator
Avar_b<-Avar_b%*%matrix(1,nrow=K,ncol=1) #change diagonal matrix into column vector
SE_1<-(Avar_b/n)^(1/2) #standard error of OLS in Large-Sample Theory
t_V_1<-b/SE_1 #robust t-ratio of OLS = t-Value in Large-Sample Theory
p_V_t_1<-2*(1-pnorm(abs(t_V_1),0,1)) #p-Value of t-Test of OLS in Large-Sample Theory
B_G<-e_G%*%t(e_G) #making diagonal matrix of e2 of GLS
B_G[upper.tri(B_G)]<-0 #making diagonal matrix of e2 of GLS
B_G[lower.tri(B_G)]<-0 #making diagonal matrix of e2 of GLS
S_G<-(t(X)%*%B_G%*%X)/n #data matrix of S og GLS
Avar_b_G<-solve(SXX)%*%S_G%*%solve(SXX) #asymptotic variance of GLS estimator
Avar_b_G[upper.tri(Avar_b_G)]<-0 #asymptotic variance of GLS estimator
Avar_b_G[lower.tri(Avar_b_G)]<-0 #asymptotic variance of GLS estimator
Avar_b_G<-Avar_b_G%*%matrix(1,nrow=K,ncol=1) #change diagonal matrix into column vector
SE_1_G<-(Avar_b_G/n)^(1/2) #standard error of GLS in Large-Sample Theory
t_V_1_G<-b_G/SE_1_G #robust t-ratio of GLS = t-Value in Large-Sample Theory
p_V_t_1_G<-2*(1-pnorm(abs(t_V_1_G),0,1)) #p-Value of t-Test of GLS in Large-Sample Theory
Avar_b<-solve(SXX)%*%S%*%solve(SXX) #asymptotic variance of OLS estimator
Avar_b_G<-solve(SXX)%*%S_G%*%solve(SXX) #asymptotic variance of GLS estimator
W_s_1<-(t(R%*%b-r)%*%solve(R%*%Avar_b%*%t(R))%*%(R%*%b-r))*n # W-statistic of OLS in
Large-Sample Theory
p_V_W_1<-1-pchisq(W_s_1,K-1) #p-Value of W-stat of OLS in Large-Sample Theory
W_s_1_G<-(t(R%*%b_G-r)%*%solve(R%*%Avar_b_G%*%t(R))%*%(R%*%b_G-r))*n # W-statistic
of GLS in Large-Sample Theory
p_V_W_1_G<-1-pchisq(W_s_1_G,K-1) #p-Value of W-stat of GLS in Large-Sample Theory
i<-1 #regressand number in VIF
Y_V<-matrix(X[,c(i)],nrow=n,ncol=1) #regressand in VIF
X_V<-matrix(X[,-c(i)],nrow=n,ncol=K-1) #regressor in VIF
b_V<-solve(t(X_V)%*%X_V)%*%t(X_V)%*%Y_V #parameter of OLS in VIF
e_V<-Y_V-X_V%*%b_V #error term in VIF
R2_V<-1-(colSums(e_V*e_V))/(colSums((Y_V-colMeans(Y_V))^2)) #R_squared of OLS in VIF
VIF_1<-1/(1-R2_V) #VIF
i<-2 #regressand number in VIF
Y_V<-matrix(X[,c(i)],nrow=n,ncol=1) #regressand in VIF
X_V<-matrix(X[,-c(i)],nrow=n,ncol=K-1) #regressor in VIF
b_V<-solve(t(X_V)%*%X_V)%*%t(X_V)%*%Y_V #parameter of OLS in VIF
e_V<-Y_V-X_V%*%b_V #error term in VIF
R2_V<-1-(colSums(e_V*e_V))/(colSums((Y_V-colMeans(Y_V))^2)) #R_squared of OLS in VIF
VIF_2<-1/(1-R2_V) #VIF
i<-3 #regressand number in VIF
Y_V<-matrix(X[,c(i)],nrow=n,ncol=1) #regressand in VIF
X_V<-matrix(X[,-c(i)],nrow=n,ncol=K-1) #regressor in VIF
b_V<-solve(t(X_V)%*%X_V)%*%t(X_V)%*%Y_V #parameter of OLS in VIF
e_V<-Y_V-X_V%*%b_V #error term in VIF
R2_V<-1-(colSums(e_V*e_V))/(colSums((Y_V-colMeans(Y_V))^2)) #R_squared of OLS in VIF
VIF_3<-1/(1-R2_V) #VIF
i<-4 #regressand number in VIF
Y_V<-matrix(X[,c(i)],nrow=n,ncol=1) #regressand in VIF
X_V<-matrix(X[,-c(i)],nrow=n,ncol=K-1) #regressor in VIF
b_V<-solve(t(X_V)%*%X_V)%*%t(X_V)%*%Y_V #parameter of OLS in VIF

```

```

e_V<-Y_V-X_V%*%b_V #error term in VIF
R2_V<-1-(colSums(e_V*e_V))/(colSums((Y_V-colMeans(Y_V))^2)) #R_squared of OLS in VIF
VIF_4<-1/(1-R2_V) #VIF
i<-5 #regressand number in VIF
Y_V<-matrix(X[,c(i)],nrow=n,ncol=1) #regressand in VIF
X_V<-matrix(X[,-c(i)],nrow=n,ncol=K-1) #regressor in VIF
b_V<-solve(t(X_V)%*%X_V)%*%t(X_V)%*%Y_V #parameter of OLS in VIF
e_V<-Y_V-X_V%*%b_V #error term in VIF
R2_V<-1-(colSums(e_V*e_V))/(colSums((Y_V-colMeans(Y_V))^2)) #R_squared of OLS in VIF
VIF_5<-1/(1-R2_V) #VIF
i<-6 #regressand number in VIF
Y_V<-matrix(X[,c(i)],nrow=n,ncol=1) #regressand in VIF
X_V<-matrix(X[,-c(i)],nrow=n,ncol=K-1) #regressor in VIF
b_V<-solve(t(X_V)%*%X_V)%*%t(X_V)%*%Y_V #parameter of OLS in VIF
e_V<-Y_V-X_V%*%b_V #error term in VIF
R2_V<-1-(colSums(e_V*e_V))/(colSums((Y_V-colMeans(Y_V))^2)) #R_squared of OLS in VIF
VIF_6<-1/(1-R2_V) #VIF
i<-7 #regressand number in VIF
Y_V<-matrix(X[,c(i)],nrow=n,ncol=1) #regressand in VIF
X_V<-matrix(X[,-c(i)],nrow=n,ncol=K-1) #regressor in VIF
b_V<-solve(t(X_V)%*%X_V)%*%t(X_V)%*%Y_V #parameter of OLS in VIF
e_V<-Y_V-X_V%*%b_V #error term in VIF
R2_V<-1-(colSums(e_V*e_V))/(colSums((Y_V-colMeans(Y_V))^2)) #R_squared of OLS in VIF
VIF_7<-1/(1-R2_V) #VIF
i<-8 #regressand number in VIF
Y_V<-matrix(X[,c(i)],nrow=n,ncol=1) #regressand in VIF
X_V<-matrix(X[,-c(i)],nrow=n,ncol=K-1) #regressor in VIF
b_V<-solve(t(X_V)%*%X_V)%*%t(X_V)%*%Y_V #parameter of OLS in VIF
e_V<-Y_V-X_V%*%b_V #error term in VIF
R2_V<-1-(colSums(e_V*e_V))/(colSums((Y_V-colMeans(Y_V))^2)) #R_squared of OLS in VIF
VIF_8<-1/(1-R2_V) #VIF
i<-9 #regressand number in VIF
Y_V<-matrix(X[,c(i)],nrow=n,ncol=1) #regressand in VIF
X_V<-matrix(X[,-c(i)],nrow=n,ncol=K-1) #regressor in VIF
b_V<-solve(t(X_V)%*%X_V)%*%t(X_V)%*%Y_V #parameter of OLS in VIF
e_V<-Y_V-X_V%*%b_V #error term in VIF
R2_V<-1-(colSums(e_V*e_V))/(colSums((Y_V-colMeans(Y_V))^2)) #R_squared of OLS in VIF
VIF_9<-1/(1-R2_V) #VIF
X1<-matrix(X[,c(1)],nrow=n,ncol=1) #regressor of White Test
X2<-matrix(X[,c(2)],nrow=n,ncol=1)
X3<-matrix(X[,c(3)],nrow=n,ncol=1)
X4<-matrix(X[,c(4)],nrow=n,ncol=1)
X5<-matrix(X[,c(5)],nrow=n,ncol=1)
X6<-matrix(X[,c(6)],nrow=n,ncol=1)
X7<-matrix(X[,c(7)],nrow=n,ncol=1)
X8<-matrix(X[,c(8)],nrow=n,ncol=1)
X9<-matrix(X[,c(9)],nrow=n,ncol=1)
X_W<-cbind(
X1*X1,X1*X2,X1*X3,X1*X4,X1*X5,X1*X6,X1*X7,X1*X8,X1*X9,
X2*X2,X2*X3,X2*X4,X2*X5,X2*X6,X2*X7,X2*X8,X2*X9,
X3*X3,X3*X4,X3*X5,X3*X6,X3*X7,X3*X8,X3*X9,
X4*X4,X4*X5,X4*X6,X4*X7,X4*X8,X4*X9,
X5*X5,X5*X6,X5*X7,X5*X8,X5*X9,
X6*X6,X6*X7,X6*X8,X6*X9,
X7*X7,X7*X8,X7*X9,
X8*X8,X8*X9,
X9*X9) #regressor of White Test

```

```

Y_W<-e #regressand of White Test
b_W<-solve(t(X_W)%*%X_W)%*%t(X_W)%*%Y_W #parameter of OLS in White Test
e_W<-Y_W-X_W%*%b_W #error term in White Test
R2_W<-1-(colSums(e_W*e_W))/(colSums((Y_W-colMeans(Y_W))^2)) #R_squared of OLS in
White Test
p_V_W<-1-pchisq(n*R2_W,(K+1)*K/2) #p-Value of White Test
p_V_E_e<-(1-pt(((n)*(1/2))*abs(colMeans(e)/s2),n-1))*2 #p-value of Expectation of error
p_V_E_e_G<-(1-pt(((n)*(1/2))*abs(colMeans(e_G)/s2_G),n-1))*2 #p-value of Expectation of error

title_98_normal_net<-data.frame(Regressand='Net Production',Regressor='with not Adjusted Temp.
and Prep.',Sample_size=n,row.names=c(['Regression_Title'])) #title
R2_of_98_normal_net<-data.frame(Normal=c(R2,R2_G),Adjusted=c(A_R2,A_R2_G),row.names=c(['
OLS_R_squared'],'[GLS_R_squared]')) #R_squared
F_of_98_normal_net<-data.frame(F_Test=c(p_V_F_f,p_V_F_f_G),Wald_Test=c(p_V_W_l,p_V_W_l_
G),row.names=c(['p-Value_OLS'],'[p-Value_GLS]')) #F test
OLS_estimator_98_normal_net<-data.frame(Parameter_of_OLS=b,Finite_p_Value=p_V_t_f,Large_p_
Value=p_V_t_l,row.names=c(['const.'],'[X2]','[X3]','[X4]','[X5]','[X6]','[X7]','[X8]','[X9]'))
#OLS_estimator
GLS_estimator_98_normal_net<-data.frame(Parameter_of_GLS=b_G,Finite_p_Value=p_V_t_f_G,Larg
e_p_Value=p_V_t_l_G,row.names=c(['const.'],'[X2]','[X3]','[X4]','[X5]','[X6]','[X7]','[X8]','[X9]'))
#GLS_estimator
Test_of_Assumption_98_normal_net<-data.frame(White_Test=p_V_W,Exogeneity_OLS=p_V_E_e,Ex
ogeneity_GLS=p_V_E_e_G,row.names=c(['p-Value'])) #Test_of_Assumption
VIF_98_normal_net<-data.frame(const.=VIF_1,X2=VIF_2,X3=VIF_3,X4=VIF_4,X5=VIF_5,X6=VIF
_6,X7=VIF_7,X8=VIF_8,X9=VIF_9,row.names=c(['VIF'])) #VIF
error_OLS_98_normal_net<-e #error term of OLS
error_GLS_98_normal_net<-e_G #error term of GLS

options(digits=3) #result

cat("\n");title_98_normal_net;cat("\n");R2_of_98_normal_net;F_of_98_normal_net;cat("\n");OLS_es
timator_98_normal_net;GLS_estimator_98_normal_net;cat("\n");Test_of_Assumption_98_normal_net
;VIF_98_normal_net;shapiro.test(error_OLS_98_normal_net);shapiro.test(error_GLS_98_normal_net)

```

< output >

	Regressand	Regressor	Sample_size
[Regression_Title]	Net Production	with not Adjusted Temp. and Prep.	98
	Normal	Adjusted	
[OLS_R_squared]	0.948	0.943	
[GLS_R_squared]	0.948	0.943	
	F_Test	Wald_Test	
[p-Value_OLS]	0	0	
[p-Value_GLS]	0	0	
	Parameter_of_OLS	Finite_p_Value	Large_p_Value
[const.]	-3.2105	0.000219	1.08e-05
[X2]	0.0797	0.025869	2.04e-02
[X3]	0.1084	0.009892	8.30e-03
[X4]	0.9072	0.000000	0.00e+00
[X5]	0.0181	0.739842	7.43e-01
[X6]	-0.3699	0.003858	1.46e-03
[X7]	-0.1603	0.081354	4.78e-02
[X8]	0.3475	0.007541	6.91e-03

	Parameter_of_GLS	Finite_p_Value	Large_p_Value
[X9]	0.2594	0.068189	8.30e-03
[const.]	-3.14899	0.000288	8.57e-06
[X2]	0.07371	0.039157	3.01e-02
[X3]	0.10682	0.011067	9.46e-03
[X4]	0.92261	0.000000	0.00e+00
[X5]	0.00632	0.907529	9.08e-01
[X6]	-0.38079	0.003008	1.14e-03
[X7]	-0.17098	0.063689	3.42e-02
[X8]	0.34031	0.008896	7.35e-03
[X9]	0.24954	0.079505	1.17e-02

	White_Test	Exogeneity_OLS	Exogeneity_GLS
[p-Value]	0.0854	1	0.55
[VIF]	0	2.75	3.44
		4.13	3.71
		5.33	2.36
		4.72	1.96

Shapiro-Wilk normality test

data: error_OLS_98_normal_net
W = 0.989, p-value = 0.5733

Shapiro-Wilk normality test

data: error_GLS_98_normal_net
W = 0.99, p-value = 0.666