

第 12 章 時系列とクロスセクションの統合

1. パネルデータとは

パネルデータは、時系列データとクロスセクションデータを統合したものである。たとえば、実質 GDP と高齢化比率が時系列データとして存在し、それが各時点で各国ごとにあるようなデータである。

1970、1980、1990 年の高齢化比率が、その後の実質 GDP 成長率とどの程度関係があるのかを調べてみよう。クロスセクションデータとして、カナダ、米国、英国、フランス、イタリア、日本を考える。

高齢化比率と実質経済成長率の関係

	カナダ			米国			英国		
	1970	1980	1990	1970	1980	1990	1970	1980	1990
65歳以上人口の全人口に占める比率(10年前)(%)	7.90	9.40	11.27	9.84	11.19	12.39	12.94	15.07	15.72
過去10年間の実質GDP平均成長率(年平均)(%)	4.26	2.77	2.84	3.26	3.17	3.26	1.94	2.67	2.30
	フランス			イタリア			日本		
	1970	1980	1990	1970	1980	1990	1970	1980	1990
65歳以上人口の全人口に占める比率(10年前)(%)	12.87	13.97	13.99	10.89	13.15	15.32	7.07	9.04	11.99
過去10年間の実質GDP平均成長率(年平均)(%)	3.31	2.47	1.87	3.61	2.26	1.57	4.46	4.09	1.41

(出所)世界銀行『世界開発指標2003』

2. パネルデータの取り込み方

まずワークファイルを作成する。今回の例のように、データが 5 年後ごとで連続していない場合は、Frequency を Integer date とする。1 が 1970、2 が 1980、3 が 1990 とみなし、1 から 3 までとする。

EViews では、パネルデータを、データ名とクロスセクションを表す記号を組み合わせで作成する。たとえば、カナダ (C A) の実質 GDP 成長率 (G R O W T H) のデータは次のように表示される。

G R O W T H C A

同様に、日本 (J P) の高齢化比率 (O L D) を表すデータは、「OLDJPN」、米国 (U S A) は「OLDUSA」という名前で登録される。

パネルデータの取り込み方には 2 種類ある。

2.1 最初にパネルデータの形にして変数を移す方法

パネルデータを分析する際に使うのは、[Pool]というオブジェクトである。

[Objects] [New Object...]で[Pool]を選ぶ。

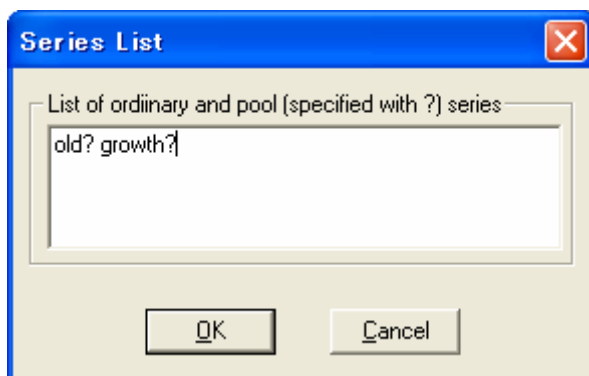
Cross Section Identifiers: (Enter identifiers below this line)

というメッセージが含まれた、テキストの貼り付けが可能な画面が出てくる。ここには、クロスセクションを識別する記号を入れる。国別データであれば国の記号、産業別データであれば産業別の記号、個票データであれば、個人を識別する記号などである。

たとえば、カナダ、フランス、イギリス、イタリア、日本、アメリカの国別データを入力する場合は、以下の記号を入力する。個々のデータの区切りはスペースでよい。エクセルなどに書いておいてコピーして張り付ければ、繰り返して使う場合には便利である。

CAN FRA GBR ITA JPN USA

[View] [Spread Sheet(Stacked Data...)]を選ぶと、次のようなメニューが出る。



ここに変数名を書き入れていく。パネルデータの場合は変数名の最後に「？」を加える。

国別にデータがないもの(たとえば原油価格など)がある場合は、「？」を除いて入力する。この違いは重要である。

[Edit+/-]を1回クリックしてデータの編集が可能な状態にして、エクセルなどからデータを貼り付ける。

列には成長率、高齢化比率などの系列のデータが並び、行は、国ごとに時系列データが並ぶ。一括して張り付ける場合は、あらかじめこの形式にエクセルのデータを並べ替えると便利である。

The screenshot shows the EViews software window titled 'Pool: UNTITLED Workfile: UNTITLED'. The main data table is as follows:

obs	OLD?	GROWTH?			
CAN-2001	7.900000	4.260000			
CAN-2002	9.400000	2.770000			
CAN-2003	11.300000	2.840000			
FRA-2001	12.900000	3.300000			
FRA-2002	14.000000	2.500000			
FRA-2003	14.000000	1.900000			
GBR-2001	12.900000	1.900000			
GBR-2002	15.100000	2.700000			
GBR-2003	15.700000	2.300000			
ITA-2001	10.900000	3.600000			
ITA-2002	13.100000	2.300000			
ITA-2003	15.300000	1.600000			
JPN-2001	7.100000	4.500000			
JPN-2002	9.000000	4.100000			
JPN-2003	12.000000	1.400000			
USA-2001	9.800000	3.300000			
USA-2002	11.200000	3.200000			
USA-2003	12.400000	3.300000			

2.2 データをワークファイルに移した後、パネルデータにする方法

まず個別のデータを入力し、最後にパネルデータにする方法もある。最初に、OLD JAPAN, OLD USAなどのデータを作っておき、それを通常の変数と同じようにEViewsに取り込んだ後、上記パネルデータを作成する操作（個別の識別記号の入力と、系列名の入力）を行えば、自動的にパネルデータとデータが入力される。

3 . パネルデータの推計

データが入力されれば、推計できる環境になる。

[Estimate]を押すと、次のメニューになる。

Pool Estimation

Specification Options

Dependent variable
growth?

Estimation method
Fixed and Random Effects
Cross-section: None
Period: None
Weights: No weights

Estimation settings
Method: LS - Least Squares (and AR)
Sample: 1 3 Balance Sample

Regressors and AR() terms
Common coefficients:
c old?
Cross-section specific coefficients:
Period specific coefficients:

OK キャンセル

<メニューの説明>

Dependent variable 被説明変数

Pegressors and AR()terms 説明変数

Estimation method 推計法

3 . 1 推計法によって説明変数を入力する場所が変わる

説明変数は、時系列、クロスセクション共通の係数とするもの、クロスセクション別に係数を推計するもの、時系列別に係数を推計する場合の3種類がある。

Common Coefficients に C (定数項) と old?を入力した場合

Dependent Variable: GROWTH?
Method: Pooled Least Squares
Date: 09/17/05 Time: 12:25
Sample: 1 3
Included observations: 3
Cross-sections included: 6
Total pool (balanced) observations: 18

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.106886	0.698077	8.748156	0.0000
OLD?	-0.271747	0.057501	-4.725956	0.0002

R-squared	0.582623	Mean dependent var	2.876111
Adjusted R-squared	0.556537	S.D. dependent var	0.900321
S.E. of regression	0.599551	Akaike info criterion	1.919168
Sum squared resid	5.751382	Schwarz criterion	2.018098
Log likelihood	-15.27251	F-statistic	22.33466
Durbin-Watson stat	2.038150	Prob(F-statistic)	0.000228

Dependent Variable: GROWTH? Path = m:#views DB = none WF = untitled

Cross-section specific coefficients に C (定数項) old?を入力した場合

Dependent Variable: GROWTH?
Method: Pooled Least Squares
Date: 09/17/05 Time: 12:32
Sample: 1 3
Included observations: 3
Cross-sections included: 6
Total pool (balanced) observations: 18

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CAN--C	7.082537	1.864673	3.798273	0.0090
FRA--C	16.20000	7.084582	2.286656	0.0622
GBR--C	-0.649080	3.269679	-0.198515	0.8492
ITA--C	8.454545	1.982093	4.265465	0.0053
JPN--C	9.484844	1.279016	7.415735	0.0003
USA--C	3.288583	2.834347	1.160261	0.2900
CAN--OLDCAN	-0.397819	0.193545	-2.055432	0.0856
FRA--OLDFRA	-1.000000	0.519276	-1.925758	0.1024
GBR--OLDGBR	0.202454	0.223701	0.905022	0.4003
ITA--OLDITA	-0.454545	0.149902	-3.032283	0.0230
JPN--OLDJPN	-0.656745	0.133489	-4.919828	0.0027
USA--OLDUSA	-0.001969	0.253431	-0.007767	0.9941

R-squared	0.905289	Mean dependent var	2.876111
Adjusted R-squared	0.731654	S.D. dependent var	0.900321
S.E. of regression	0.466386	Akaike info criterion	1.547114
Sum squared resid	1.305094	Schwarz criterion	2.140695
Log likelihood	-1.924026	F-statistic	5.213722
Durbin-Watson stat	4.009185	Prob(F-statistic)	0.027336

Path = m:#views DB = none WF = untitled

Period specific coefficients に C (定数項) old?を入力した場合 (教科書 p204)

Dependent Variable: GROWTH?
Method: Pooled Least Squares
Date: 09/17/05 Time: 12:31
Sample: 1 3
Included observations: 3
Cross-sections included: 6
Total pool (balanced) observations: 18

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C--1	6.727464	1.192024	5.643733	0.0001
C--2	5.066206	1.359473	3.726595	0.0029
C--3	4.183156	2.072954	2.017968	0.0665
OLD?--1	-0.317151	0.113613	-2.791506	0.0163
OLD?--2	-0.178652	0.111596	-1.600884	0.1354
OLD?--3	-0.145712	0.152957	-0.952633	0.3596

R-squared	0.661487	Mean dependent var	2.876111
Adjusted R-squared	0.520440	S.D. dependent var	0.900321
S.E. of regression	0.623475	Akaike info criterion	2.154186
Sum squared resid	4.664649	Schwarz criterion	2.450975
Log likelihood	-13.38766	F-statistic	4.689834
Durbin-Watson stat	2.249813	Prob(F-statistic)	0.013212

3.2 固定効果と変量効果による推計

パネルデータを推計する場合は次のように考える。

$$\text{growth}_{it} = c + \beta \text{old}_{it} + \alpha_i + \alpha_t + e_{it}$$

α_i は国ごとに変わる要因、 α_t は期間ごとに変わる要因である。 α_i や α_t の想定の方の違いで固定効果モデルと変量効果モデルという2種類の推計法が考えられる。

クロスセクションの変動要因の場合で考えると、固定効果モデルでは、クロスセクション個別の効果を各国固有の要因によって決まると考え、各国別にダミー変数を使った場合と同じである。Regressor and AR()の Common Coefficients に old?、Cross-section specific coefficients に C (定数項) を入力した場合と同じ結果になることを確認しよう。

変量効果モデルでは、個別の効果を誤差項と同じようなランダムなものとして推計する。個別効果の平均はゼロとなり、正負均等に分布する。

Estimate method の coross section を Fixed にすると各国ごとにダミー変数が付く固定効果、Random にすると各国の効果が変量効果で推計される。(変量効果の場合、options の weghinin option の Random effect method を wansbeek-kaptein にしないとうまくいかない)

期間ごとに個別効果を考える場合は、Esmate method の period の方を変更する。

3.3 ハウスマン検定

変量効果と固定効果のどちらを選ぶかを検定するのがハウスマン検定だ。固定効果と変量効果の推計値には次のような特徴がある。

ハウスマン検定は、固定効果、変量効果の両者を推計し、係数が近ければ変量効果を選び、離れていれば固定効果を選ぶというものである。

	固定効果	変量効果
推計法	ダミー変数での推計	一般化最小二乗法
使用する場合	誤差と説明変数に相関がない場合	誤差と説明変数に相関がある場合
推定量	不偏性を持つ (推計値が真の期待値と一致)	有効性を持つ (誤差の分散が最小)
真のモデルが固定効果の場合	正しい推定	不偏性なし
真のモデルが変量効果の場合	不偏性あり	正しい推定

EViews での操作

まず変量効果で推計した後、

[view] fixed/random effect testing correlated random effects Hauseman Tests

を選択する。

検定等計量(Chi-Sq. Statistic)が計算される。P 値が小さければ変量効果モデルを選び、大きければ固定効果モデルを選ぶ。

