세미나 자료집 13-S42

STATA를 이용한 패널데이터 분석방법

일시 | 2013년 10월 16일(수) 오전 10:00~12:00

장소 | 한국청소년정책연구원 10층 회의실

주최 | 한국청소년정책연구원

한국청소년정책연구원
National Youth Policy Institute

세 부 일 정

- STATA를 이용한 패널데이터 분석방법 -

시 간	내 용	비 고
10:00 ~ 10:10	개회 및 발표자 소개	양계민 연구위원
10:10 ~ 10:20	환영사	이재연 원장
10:20 ~ 11:40	콜로키움 발표	민인식(경희대 경제학과)
11:40 ~ 12:00	질의응답	참석자 전원
12:00	폐회	양계민 연구위원

STATA를 이용한 패널데이터 분석방법

민 인 식

STATA를 이용한
아동-청소년 패널(CYPS) 데이터 활용

발표자: 민 인 식 (경희대학교 경제학과)

2013년 10월 16일

【 목 차 】

I. STATA에서 패널데이터 구축

Panel data: repeated measurements at different time points on the same subject (Cross-sectional time series data)

Longitudinal data (종단데이터): 같은 subject에 대해서 시간에 따라 반복적으로 관찰한 데이터이다. 주로 health 나 biology에서 쓰는 표현이다.

Multilevel data: 하위 레벨의 관측치가 상위 레벨에 nested 되어 있는 구조의 데이터이다. 패널데이터나 종단데이터 역시 멀티레벨 데이터의 일종이다.

1. 패널데이터 분석의 장・단점

1) 추정량의 효율성
More observations : Pooling of cross-sectional and time-series data
More degree of freedom : stems from more observations
Reducted multicollinearity
==> 추정치의 효율성(efficiency)를 높이게 된다. 따라서 유의성이 높아질 가능성이 많다.

2) 다양한 통계적 분석이 가능
Wider range of problems : dynamics of change (노동시장에서의 실업-재취업의 문제)
Causality discussion : Time structure facilitates discussion
==> 개인이나 가구의 행동이나 policy changes에 대한 새로운 가설을 설정하여 test 할 수 있다.

3) 패널데이터 단점
패널데이터 구축에 시간과 비용이 많이 든다
패널에서 빠져나가는 attrition의 문제가 발생한다.
패널 그룹 간 상관관계가 존재할 수 있다. (같은 학교 내의 학생들을 표본추출하는 경우)

2. 패널데이터 만들기

1) 패널데이터 구축 과정
CYPS 데이터의 경우 초등학교 1학년, 초등학교 4학년, 중학교 1학년의 3개 코호트(cohort)를 3년씩 조사하였기 때문에 각 코호트 별로 $n = 2300$ 과 $T = 3$인 패널데이터 구조가 된다.

표 1. CYPS 데이터 구조

코호트	조사년도	표본 수	변수 수
초등학교 1학년	1차년도 (2010)	2342명	762개
	2차년도 (2011)	2264명	709개
	3차년도 (2012)	2200명	748개
초등학교 4학년	1차년도 (2010)	2378명	772개
	2차년도 (2011)	2264명	557개
	3차년도 (2012)	2219명	661개
중학교 1학년	1차년도 (2010)	2451명	843개
	2차년도 (2011)	2280명	717개
	3차년도 (2012)	2259명	795개

Objective: 중학교 1학년 ~ 3학년 동안 조사된 중학생에 대해 분석하기 위해서는 학생변수와 가구특성변수를 포함하여 패널데이터로 완성한다.

[1단계] 중학교 1학년 코호트의 3개 조사데이터에서 연구자 원하는 변수(가령, 성적변수, 학생특성변수 그리고 가구특성 변수)만을 남긴 후 새로 저장한다. ==> 3개의 데이터 파일이 남게 된다.

[2단계] 1단계 데이터에서 반드시 학생 id변수와 조사시점(년도)에 대한 변수가 존재해야 한다. 즉 group변수와 time 변수를 각 데이터가 포함해야 한다.

[3단계] 3개의 데이터를 하나로 append한다. 주의해야 할 점은 각 데이터에서 같은 contents의 변수는 같은 변수이름(variable name)을 가지고 있어야 한다.

[4단계] 3단계 작업 후 만들어진 데이터 학생id와 time 변수에 대해서 sort하면 Stata에서 통계분석에 활용할 수 있는 패널데이터로 완성된다.

→ 위 작업을 위한 Stata coding 작업은 연구자가 직접 해야 하는 부분이다.

표 2. 중학교 1학년 ~ 3학년 패널데이터 예

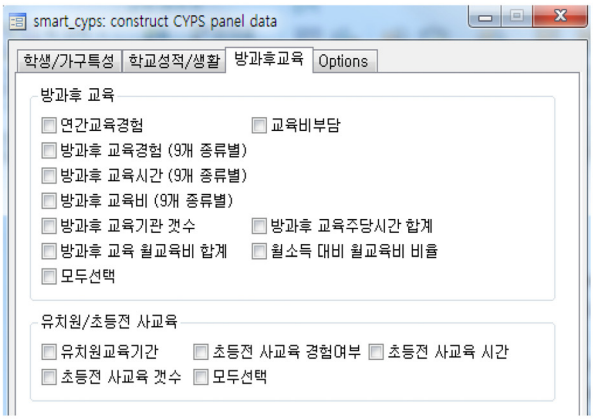
	cohort	id	id1	wave	sclid	region1	region2	sgrade	gender
1	ms1	14201	714201	1	234	10	1009	1	1
2	ms1	14201	714201	2	234	10	1009	2	1
3	ms1	14201	714201	3	234	10	1009	3	1
4	ms1	14203	714203	1	234	10	1009	1	0
5	ms1	14203	714203	2	234	10	1009	2	0
6	ms1	14203	714203	3	.	.	.	.	.
7	ms1	14204	714204	1	234	10	1009	1	1
8	ms1	14204	714204	2	234	10	1009	2	1
9	ms1	14204	714204	3	234	10	1009	3	1
10	ms1	14205	714205	1	234	10	1009	1	0
11	ms1	14205	714205	2	234	10	1009	2	0
12	ms1	14205	714205	3	.	.	.	.	.
13	ms1	14206	714206	1	234	10	1009	1	1
14	ms1	14206	714206	2	234	10	1009	2	1
15	ms1	14206	714206	3	234	10	1009	3	1
16	ms1	14207	714207	1	234	10	1009	1	1
17	ms1	14207	714207	2	234	10	1009	2	1
18	ms1	14207	714207	3	234	10	1009	3	1

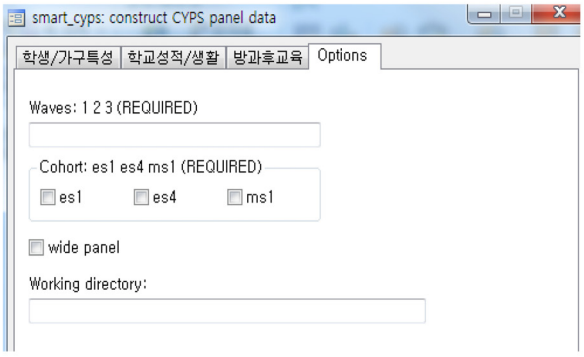
2) CYPS 패널데이터 구축 시 문제점

1) 통계패키지에 익숙치 않은 연구자들에게는 년도별 데이터를 연결하여 패널데이터로 만드는 작업이 여전히 어려운 문제이다.
2) 통계패키지에 익숙한 연구자들도 같은 데이터를 이용하여 새로운 연구를 할 때마다 패널데이터를 만들어야 하는 반복작업을 해야 한다.
3) raw data에는 많은 변수들이 있지만 대부분의 연구에서 자주 사용하는 변수들이 정해져 있고 그 숫자는 50여개 미만이다. 실제 연구에서 필요한 변수만을 선택해야 하는 작업을 다시 해야 한다.
4) 특정 연구자가 패널데이터 작성을 위한 programming code를 다른 연구자가 사용하기 어려운 경우가 많다.
5) raw data를 이용하여 패널데이터를 만들었을지라도 기존 변수이름을 연구자가 다시 바꾸는 작업을 하게 되는 경우가 많다.
6) Stata에서는 결측치가 “.”으로 표시되어야 하기 때문에 결측치 코딩을 새로 해야 하는 번거로움이 있다.
7) 단순히 각 년도 데이터를 연결해서 패널데이터를 만들었다고 해서 데이터 작업이 완료된 것이 아니다. 기존 변수를 이용하여 새로운 변수를 만들어야 하는 경우 복잡한 코딩작업을 거쳐야 한다. 가령, 방과 후 교육 개수, 방과 후 교육 총 시간, 방과 후 교육 총 교육비 등

3) smart_cyps 모듈을 이용한 패널데이터 구축
강사가 직접 작성한 모듈을 이용하면 CYPS 데이터를 쉽게 패널데이터로 전환할 수 있다.

그림 1. smart_cyps 대화창 screen shot





○ 장점

1) 패널데이터를 만드는데 많은 시간을 save할 수 있다.
2) dialog box를 이용해서 명령문과 옵션을 작성하기 때문에 Stata 초보자도 쉽게 사용할 수 있다.
3) 결측치 및 기본적으로 필요한 변수에 대한 recoding 작업이 되어 있는 상태로 패널데이터를 만들어 준다.
4) 서로 다른 연구자들이 smart_cyps 모듈을 이용해서 데이터를 만들면 같은 변수이름을 가지고 있기 때문에 서로 의사소통이 편리하다.

5) 매년 조사결과 제공 시 같은 내용의 변수에 대해서 같은 변수 이름을 지정해야 주는 불편함을 피할 수 있다.

○ 단점

1) dialog box에 없는 변수를 추가하고자 할 때 연구자가 smart_cyps 관련 코딩파일을 수정하는 게 불가능하다. 즉 dialog box의 유지보수는 원저자에 의해서만 가능하다.
2) 연구자의 연구주제에 따라 분석에 필요한 변수가 dialog box에 포함되지 않을 수 있다. 즉 smart_cyps 모듈에는 topic-specific 변수보다는 관련 연구자에게 꼭 사용하는 필수적인 변수들이 포함되게 될 것이다.
3) 연구자의 data management 능력을 저하시킬 가능성이 있다.

II. 패널데이터 구조 및 기초통계량

1. wide/long type panel

long-type panel

	cohort	id	wave	height	weight
1	ms1	14201	1	.	.
2	ms1	14201	2	168	73
3	ms1	14201	3	172	64
4	ms1	14203	1	.	.
5	ms1	14203	2	161	65
6	ms1	14203	3	.	.
7	ms1	14204	1	.	.
8	ms1	14204	2	172	88
9	ms1	14204	3	176	88
10	ms1	14205	1	.	.
11	ms1	14205	2	162	60
12	ms1	14205	3	.	.
13	ms1	14206	1	.	.
14	ms1	14206	2	179	84
15	ms1	14206	3	181	96

wide-type panel

	id	height1	height2	height3
1	14201	.	168	172
2	14203	.	161	.
3	14204	.	172	176
4	14205	.	162	.
5	14206	.	179	181
6	14207	.	163	166
7	14208	.	178	180
8	14209	.	173	177
9	14210	.	161	162
10	14211	.	175	178
11	14212	.	148	153
12	14213	.	168	173
13	14214	.	160	161

2. short panel and long panel

short 패널: many subjects (students or households) and short time periods

: T is fixed and $n \to \infty$

long 패널 : $T \to \infty$ and n is small

- dynamic model 설정이 가능하다.
- 패널 그룹 내에서 오차항의 자기상관 가능성이 크다.
- 패널 그룹효과(group effect)를 그룹더미로 설정할 수 있다 (고정효과 모형)

3. 균형패널과 불균형 패널

balanced panel: 각 그룹의 데이터 포괄기간이 서로 동일하다.

	state_code	year	pop	area
1	1	1970	.1369841	.871691
2	1	1971	.6184582	.871691
3	1	1972	.4241557	.871691
4	1	1973	.2648021	.871691
5	2	1970	.6432207	.4611429
6	2	1971	.0610638	.4611429
7	2	1972	.8983462	.4611429
8	2	1973	.9477426	.4611429

unbalanced panel: 각 그룹의 데이터 포괄기간이 서로 다르다.

	state_code	year	pop	area
1	1	1970	.1369841	.871691
2	1	1971	.6184582	.871691
3	1	1972	.4241557	.871691
4	1	1973	.2648021	.871691
5	2	1970	.6432207	.4611429
6	2	1971	.0610638	.4611429
7	3	1970	.5578017	.4216726
8	3	1971	.5552388	.4216726
9	3	1972	.5219247	.4216726
10	3	1973	.2769154	.4216726

4. 패턴 및 기초통계량 계산

1) xtdes : 패널데이터 description

```
. tsset id wave
       panel variable:  id (unbalanced)
        time variable:  wave, 1 to 3, but with gaps
                delta:  1 unit

. xtdes

      id:  14201, 14203, ..., 167042                         n =       2351
    wave:  1, 2, ..., 3                                      T =          3
           Delta(wave) = 1 unit
           Span(wave)  = 3 periods
           (id*wave uniquely identifies each observation)

Distribution of T_i:   min      5%     25%       50%       75%     95%     max
                         1       2       3         3         3       3       3

     Freq.  Percent    Cum.  Pattern
 ---------------------------------------
      2226    94.68   94.68  111
        50     2.13   96.81  11.
        42     1.79   98.60  1..
        33     1.40  100.00  1.1
 ---------------------------------------
      2351   100.00          XXX
```

[해석]

2) 기초통계량

```
. xtsum height weight gpa_avg
```

Variable		Mean	Std. Dev.	Min	Max	Observations
height	overall	164.7108	7.823539	110	190.4	N = 4528
	between		7.550862	125	188.7	n = 2308
	within		2.043132	137.2108	192.2108	T-bar = 1.96187
weight	overall	55.3088	10.7381	30	120	N = 4510
	between		10.29707	30.5	100	n = 2303
	within		3.017098	20.3088	90.3088	T-bar = 1.95832
gpa_avg	overall	3.081514	1.300807	1	5	N = 4490
	between		1.237853	1	5	n = 2306
	within		.4098745	1.248181	4.914848	T-bar = 1.94709

[해석]

3) xttab

```
. xttab coedu
```

	Overall		Between		Within
coedu	Freq.	Percent	Freq.	Percent	Percent
1	474	10.45	243	10.52	98.56
2	567	12.50	291	12.60	98.80
3	3494	77.05	1788	77.44	99.66
Total	4535	100.00	2322	100.56	99.44
			(n = 2309)		

[해석]

4) xttrans 명령어

```
. xttrans gpa_math
```

수학성적	수학성적 1	2	3	4	5	Total
1	75.07	8.50	10.48	4.96	0.99	100.00
2	45.32	14.78	20.20	16.26	3.45	100.00
3	23.64	15.76	27.72	19.57	13.32	100.00
4	13.74	5.92	19.67	36.97	23.70	100.00
5	2.39	1.52	6.94	23.43	65.73	100.00
Total	36.02	8.33	15.37	18.70	21.57	100.00

[해석]

III. 패널데이터를 이용한 회귀분석: 연속형 종속변수

1. 데이터 구성 및 변수 설명

종속변수: 국어+수학+영어 평균 성적(5점 만점 척도)

데이터 구성: 중학교 1학년 코호트의 3년간 패널데이터를 사용함. 성적변수는 2차년도와 3차년도에만 조사되었다.

```
. tsset id wave
        panel variable:  id (unbalanced)
         time variable:  wave, 1 to 3, but with gaps
                 delta:  1 unit
```

: unbalanced and time gaps

```
. tab wave
```

¼-°£ÀÌÂ÷¼ö	Freq.	Percent	Cum.
1	2,351	34.14	34.14
2	2,277	33.06	67.20
3	2,259	32.80	100.00
Total	6,887	100.00	

```
. xtsum gpa_avg attitude study_time3
```

Variable		Mean	Std. Dev.	Min	Max	Observations
gpa_avg	overall	3.081125	1.300924	1	5	N = 4491
	between		1.237737	1	5	n = 2306
	within		.4098439	1.247792	4.914459	T-bar = 1.94753
attitude	overall	2.735923	.5176787	1	4	N = 6887
	between		.433953	1.1	4	n = 2351
	within		.287237	1.469256	4.069256	T-bar = 2.92939
study_~3	overall	3.494427	2.248819	0	13.5	N = 6887
	between		1.854575	0	11.41667	n = 2351
	within		1.305861	-2.144462	10.49443	T-bar = 2.92939

2. Pooling regression : no error-component model

Pooled OLS 추정

$$GPA_{it} = \alpha + \beta_1 Attitude_{it} + \beta_2 StudyTime_{it} + \epsilon_{it},$$

$$i = 1,2,\cdots,n \text{ 및 } t = 1,2,\cdots,T_i \qquad \text{(식 1)}$$

[단점] 패널그룹 간 이분산성, 자기상관 문제, contemporaneous correlation을 고려하지 않는다. 그룹 heterogeneity를 고려하지 않는다. 설명변수 내생성(endogeneity)이 없다고 가정한다.

```
. reg gpa_avg attitude study_time3
```

Source	SS	df	MS	Number of obs =	4491
Model	1931.26537	2	965.632684	F(2, 4488) =	764.65
Residual	5667.62245	4488	1.26283923	Prob > F =	0.0000
				R-squared =	0.2542
				Adj R-squared =	0.2538
Total	7598.88781	4490	1.69240263	Root MSE =	1.1238

gpa_avg	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
attitude	.8439375	.0338619	24.92	0.000	.7775515	.9103235
study_time3	.1624756	.0081163	20.02	0.000	.1465636	.1783876
_cons	.2290902	.0890899	2.57	0.010	.05443	.4037503

[해석]

패널 GLS 추정

오차항에 대해서 패널 그룹 간 이분산성, 1계 자기상관 그리고 동시적 상관관계를 가정하고 추정할 수 있다. => 효율적인 추정량

[단점] samll n and large T인 경우에만 주로 사용한다. 따라서 패널그룹의 수가 많은 경우 추정이 어렵다. time gaps이 있는 경우 1계 자기상관을 가정하기 어렵다. 추정해야 할 모수가 많아질수록 자유도가 줄어든다.

```
. xtgls gpa_avg attitude study_time3, p(hetero)
matsize too small - should be at least 2306
r(908);
```

```
. xtgls gpa_avg attitude study_time3, corr(ar1)
(note: 121 observations dropped because only 1 obs in group)
matsize too small - should be at least 2185
r(908);
```

그룹의 수(n)을 줄인 후 만든 패널데이터를 이용하여 다시 추정함.

```
. xtgls gpa_avg attitude study_time3, p(hetero)

Cross-sectional time-series FGLS regression

Coefficients:  generalized least squares
Panels:        heteroskedastic
Correlation:   no autocorrelation

Estimated covariances      =       159          Number of obs      =       304
Estimated autocorrelations =         0          Number of groups   =       159
Estimated coefficients     =         3          Obs per group: min =         1
                                                               avg =   1.91195
                                                               max =         2
                                                Wald chi2(2)       =    927.86
                                                Prob > chi2        =    0.0000
```

gpa_avg	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
attitude	.8774098	.0510964	17.17	0.000	.7772627	.977557
study_time3	.1755272	.0085502	20.53	0.000	.1587692	.1922852
_cons	.1262547	.1546725	0.82	0.414	-.1768977	.4294072

[해석]

```
. xtgls gpa_avg attitude study_time3, corr(ar1) force
(note: 14 observations dropped because only 1 obs in group)

Cross-sectional time-series FGLS regression

Coefficients:  generalized least squares
Panels:        homoskedastic
Correlation:   common AR(1) coefficient for all panels  (-0.5891)

Estimated covariances      =         1          Number of obs      =       290
Estimated autocorrelations =         1          Number of groups   =       145
Estimated coefficients     =         3          Time periods       =         2
                                                Wald chi2(2)       =    189.12
                                                Prob > chi2        =    0.0000
```

gpa_avg	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
attitude	.8867394	.139239	6.37	0.000	.6138359	1.159643
study_time3	.1926456	.0275754	6.99	0.000	.1385989	.2466923
_cons	.0294114	.3529826	0.08	0.934	-.6624217	.7212445

[해석]

3. one-way error component models

$$GPA_{it} = \alpha + \beta_1 Attitude_{it} + \beta_2 Study\,Time_{it} + u_i + e_{it} \qquad \text{(식 2)}$$

오차항 ϵ_{it} 을 2개의 오차항으로 구분하였다.

u_i : time-invariant error term

e_{it}: time-varying error term

상수항 $(\alpha + u_i)$이 패널그룹에 따라서 서로 다르다고 가정한다. group heterogeneity를 고려한다. 그러나 모든 i에 대해서 β_1과 β_2는 서로 같다고 가정한다.

Fixed effects model (고정효과 모형)

오차항 u_i : Unobserved Heterogeneity를 고정효과 즉 추정해야 할 모수로 간주한다.

고정효과 모형 추정의 기본 아이디어

- u_i를 직접 추정한다 (LSDV 모형): incidental parameter problem
- u_i를 없애준 후 모형을 추정한다 (within transformation 모형, FD 모형)

고정효과 모형의 장점:

$cov(x_{it}, u_i) \neq 0$ (omitted variable bias)일지라도 여전히 일치추정량을 구할 수 있다.

고정효과 모형의 단점:

설명변수 중 time-invariant 수의 추정계수를 얻을 수 없다.

finite sample에서는 자유도의 손실

Random effects 모형 (확률효과 모형)

group heterogeneity를 일종의 latent variable로 본다.

$$GPA_{it} = \alpha + \beta_1 Attitude_{it} + \beta_2 StudyTime_{it} + u_i + e_{it}$$

(관찰되지 않는 그룹의 특성) u_i을 random component(random variable)로 간주한다. u_i를 정규분포를 따르는 확률변수로 간주한다. 추정해야 할 모수는 u_i의 분산인 σ_u^2이다.

위 모형에서 오차항은 $(u_i + e_{it})$이 되고 이 오차항이 일반적인 가정을 만족하면 OLS 추정량은 좋은 추정량이 된다.

OLS 추정량의 문제점: $cov(u_i + e_{it}, u_i + e_{it-1}) \neq 0$

즉 오차항의 1계 자기상관 (first-order autocorrelation)이 존재하게 된다.

오차항의 1계 자기상관을 고려한 FGLS 추정 또는 (정규분포를 가정한) ML 추정

$$(y_{it} - \theta\bar{y}_i) = \alpha(1-\theta) + \beta(x_{it} - \theta\bar{x}_i) + u_i(1-\theta) + (e_{it} - \theta\bar{e}_i)$$

RE 추정량의 단점:

- $cov(x_{it}, u_i) = 0$ 가정이 성립하지 않는다면, inconsistent estimator가 된다.

RE 추정량의 장점:

- $cov(x_{it}, u_i) = 0$ 가정이 성립한다면, 자유도의 손실 없이 효율적인 추정량을 얻을 수 있다. FE와 BE보다 더 효율적인 추정량이 된다.
- time-invariant 설명변수의 추정계수를 구할 수 있다.

```
. xtreg gpa_avg attitude study_time3, fe

Fixed-effects (within) regression               Number of obs      =      4492
Group variable: id                              Number of groups   =      2307

R-sq:  within  = 0.0094                         Obs per group: min =         1
       between = 0.3048                                        avg =       1.9
       overall = 0.2304                                        max =         2

                                                F(2,2183)          =     10.41
corr(u_i, Xb)  = 0.4493                         Prob > F           =    0.0000
```

gpa_avg	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
attitude	.0660062	.0358466	1.84	0.066	-.0042908	.1363032
study_time3	.0320882	.008062	3.98	0.000	.0162781	.0478983
_cons	2.79325	.09969	28.02	0.000	2.597753	2.988747
sigma_u	1.1960393					
sigma_e	.58499737					
rho	.80695225	(fraction of variance due to u_i)				

F test that all u_i=0: F(2306, 2183) = 6.24 Prob > F = 0.0000

[해석]

```
. xtreg gpa_avg attitude study_time3, re theta

Random-effects GLS regression                   Number of obs      =      4492
Group variable: id                              Number of groups   =      2307

R-sq:  within  = 0.0084                         Obs per group: min =         1
       between = 0.3315                                        avg =       1.9
       overall = 0.2541                                        max =         2

                                                Wald chi2(2)       =    577.19
corr(u_i, X)   = 0 (assumed)                    Prob > chi2        =    0.0000

                    theta
    min        5%         median        95%        max
 0.4631     0.4631     0.5897     0.5897     0.5897
```

gpa_avg	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
attitude	.473969	.0301961	15.70	0.000	.4147857	.5331522
study_time3	.101311	.0069972	14.48	0.000	.0875967	.1150252
_cons	1.439895	.0840116	17.14	0.000	1.275235	1.604554
sigma_u	.91927694					
sigma_e	.58499737					
rho	.71176264	(fraction of variance due to u_i)				

[해석]

Q) rho 값의 의미는 무엇인가?

Q) theta 값의 의미는?

RE 모형: ML 추정량

xtreg lwage gender edu tenure metro, mle

Hausman 검정 : 고정효과 vs. 확률효과

```
. xtreg gpa_avg attitude study_time3, fe
. estimates store FE
. xtreg gpa_avg attitude study_time3, re
. estimates store RE

. hausman FE RE, sigmamore

                ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |       FE           RE         Difference          S.E.
-------------+----------------------------------------------------------------
    attitude |    .0660062      .473969       -.4079627        .0228399
 study_time3 |    .0320882      .101311       -.0692228        .0048525
------------------------------------------------------------------------------
                          b = consistent under Ho and Ha; obtained from xtreg
           B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(2) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =      466.50
                Prob>chi2 =      0.0000
```

[해석]

MEMO

MEMO

MEMO

MEMO