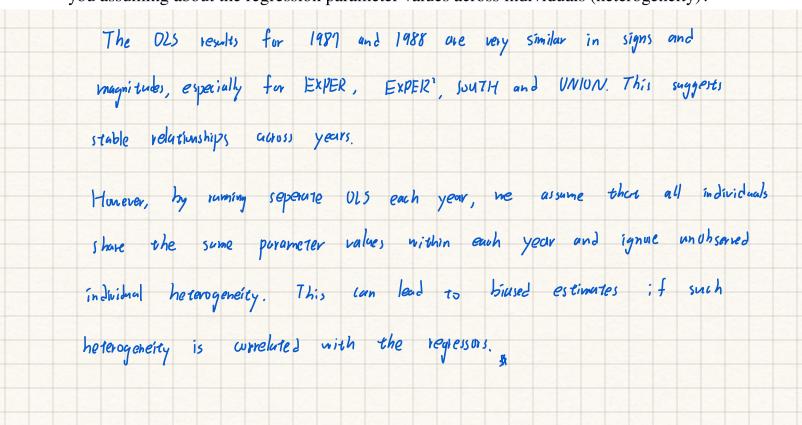
15.6 Using the NLS panel data on N = 716 young women, we consider only years 1987 and 1988. We are interested in the relationship between ln(WAGE) and experience, its square, and indicator variables for living in the south and union membership. Some estimation results are in Table 15.10.

	(1) OLS 1987	(2) OLS 1988	(3) FE	(4) FE Robust	(5) RE
C	0.9348	0.8993	1.5468	1.5468	1.1497
	(0.2010)	(0.2407)	(0.2522)	(0.2688)	(0.1597)
EXPER	0.1270	0.1265	0.0575	0.0575	0.0986
	(0.0295)	(0.0323)	(0.0330)	(0.0328)	(0.0220)
$EXPER^2$	-0.0033	-0.0031	-0.0012	-0.0012	-0.0023
	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0007)
SOUTH	-0.2128	-0.2384	-0.3261	-0.3261	-0.2326
	(0.0338)	(0.0344)	(0.1258)	(0.2495)	(0.0317)
UNION	0.1445	0.1102	0.0822	0.0822	0.1027
	(0.0382)	(0.0387)	(0.0312)	(0.0367)	(0.0245)
N	716	716	1432	1432	1432

(standard errors in parentheses)

a. The OLS estimates of the ln(WAGE) model for each of the years 1987 and 1988 are reported in columns (1) and (2). How do the results compare? For these individual year estimations, what are you assuming about the regression parameter values across individuals (heterogeneity)?



b. The ln(WAGE) equation specified as a panel data regression model is

$$\ln(WAGE_{it}) = \beta_1 + \beta_2 EXPER_{it} + \beta_3 EXPER_{it}^2 + \beta_4 SOUTH_{it}$$
$$+ \beta_5 UNION_{it} + (u_i + e_{it})$$
(XR15.6)

Explain any differences in assumptions between this model and the models in part (a).

	Th	e	Joa	ine	1	mod	el		allu	ws	f	iv	1	ndiı	i dua	1-5	וי שלן	lic	eff	ects	Ni	,		
(0)	TUV	ing		unc	obs <i>o</i>	ned		he	ano	gen	eity	•												
U)	LS	íh		Ja1 .	t l	n)	ig	7M C	5	th	ese	e	ffe	cts	,	Ca S Su	ming	1	all	ind	livi d	nals	are	
ide	ntí	cul	ap	art		fro	5	oh	serv	الح.	vo	ırja)	bles.											

c. Column (3) contains the estimated fixed effects model specified in part (b). Compare these estimates with the OLS estimates. Which coefficients, apart from the intercepts, show the most difference?

d. The F-statistic for the null hypothesis that there are no individual differences, equation (15.20), is 11.68. What are the degrees of freedom of the F-distribution if the null hypothesis (15.19) is true? What is the 1% level of significance critical value for the test? What do you conclude about the null hypothesis.

Numerator
$$df = N-1 = 116-1 = 115$$

| Denominator $df = NT-N-K_5 = 1432-116-5 = 111$

| 1% critical value for $F(n|s, n|1) \approx 1.48$

| Since $F = 11.48 > 1.48$, we reject the null

| There are significant individual effects \Rightarrow fixed effects model is preferred.

e. Column (4) contains the fixed effects estimates with cluster-robust standard errors. In the context of this sample, explain the different assumptions you are making when you estimate with and without cluster-robust standard errors. Compare the standard errors with those in column (3). Which ones are substantially different? Are the robust ones larger or smaller?

(Instar - rubust SEs (column 4) only for serial correlation within Individuals, unlike regular SEs (column 3), which assume i.i.d. errus. Robust SEs are larger, especially for SOUTH (U.1258->U.7495), meaning inference may change.