1 Fixed Effects and Random Effects Estimation

1.1 Fixed Effects Introduction

- Fixed effects model:

$$y_{it} = \alpha + \beta x_{it} + f_i + \epsilon_{it}$$

$$E\left(\epsilon_{it}|x_{it}, f_i\right) = 0$$

Suppose we just run:

$$y_{it} = \alpha + \beta x_{it} + \epsilon_{it}$$

Then we get:

$$\hat{\beta} = \beta + \frac{cov\left(f_i, x_{it}\right)}{var\left(x_{it}\right)}$$

- Two estimation methods: (1.) Fixed effects estimation, (2.) First differencing

$$y_{it} = \alpha + \beta x_{it} + f_i + \epsilon_{it} \tag{1}$$

$$y_{it} - y_{it-1} = \beta\left(x_{it} - x_{it-1}\right) + \epsilon_{it} - \epsilon_{it-1} \tag{2}$$

If T=2, the estimators are the same though the standard errors for (2.) need to be corrected if you run OLS.

- What are the differences: standard errors.

For example. Suppose

$$\epsilon_{it} = \rho\epsilon_{it-1} + \mu_{it}$$

Then $var\left(\epsilon_{it}\right) = \rho var\left(\epsilon_{it}\right) + var\left(\mu_{it}\right)$

Thus:

$$\left(1-\rho\right) var\left(\epsilon_{it}\right) = var\left(\mu_{it}\right) \implies var\left(\epsilon_{it}\right) = \frac{var\left(\mu_{it}\right)}{1-\rho}$$

If $\rho = 1$ (i.e. a random walk), then the F.E. estimator should not be used as inference in theory is not possible (infinite standard errors). However, the first difference estimator works fine. In general, the first difference estimator is better when you gain more inference from comparing one observation to another close in time due to high serial correlation in the errors. Otherwise, fixed effects works better.

- What about fixed effects in nonlinear models? Often not identified. Fixed effects probit does not work. Incidental parameters problem (the effect is not "fixed"). It turns out the logit models work with fixed effects.

- Also, note that if you have a balanced panel (number of observations per "effect") is constant. Then let T be the number of observations per "effect"and K the

  number of effects. If N (number of observations $=$ KT) goes to infinity holding T constant, $p\lim \hat{\beta} = \beta$; however the probability limit of the f_i are not. If N goes to infinity keeping the ratio $\frac{K}{T}$ constant, then all parameters are identified. In other words, more than one way to do assymptotics. You have to be careful about what assumptions you make if you want to identify the fixed effects themselves and use assmyptotic standard errors.

1.2 Random Effects

- What if $cov\left(f_i, \epsilon_{it}\right) = 0$? Then we have the random effects model. In this case, OLS is consistent. There is also the random effects model which imposes $cov\left(f_i, \epsilon_{it}\right) = 0$. It is usually estimated using MLE but can also be estimated using FGLS. With MLE, estimate with VCV matrix

$$\begin{pmatrix} M & 0 & 0 & 0 & 0 \\ 0 & M & 0 & 0 & 0 \\ 0 & 0 & M & 0 & 0 \\ 0 & 0 & 0 & M & 0 \\ 0 & 0 & 0 & 0 & M \\ \\ \\ \\ \end{pmatrix}$$

where $M =$

$$\begin{pmatrix} \sigma_\epsilon^2 + \sigma_f^2 & \sigma_\epsilon^2 & \sigma_\epsilon^2 & \sigma_\epsilon^2 & \sigma_\epsilon^2 \\ \sigma_\epsilon^2 & \sigma_\epsilon^2 + \sigma_f^2 & \sigma_\epsilon^2 & \sigma_\epsilon^2 & \sigma_\epsilon^2 \\ \sigma_\epsilon^2 & \sigma_\epsilon^2 & \sigma_\epsilon^2 + \sigma_f^2 & \sigma_\epsilon^2 & \sigma_\epsilon^2 \\ \sigma_\epsilon^2 & \sigma_\epsilon^2 & \sigma_\epsilon^2 & \sigma_\epsilon^2 + \sigma_f^2 & \sigma_\epsilon^2 \\ \sigma_\epsilon^2 & \sigma_\epsilon^2 & \sigma_\epsilon^2 & \sigma_\epsilon^2 & \sigma_\epsilon^2 + \sigma_f^2 \end{pmatrix}$$

Benefits:

1. FGLS estimator: more flexible if you dont know error structure.

2. RE estimator more robust (less noisy) if you do know error structure.

1.3 Reconsideration of Fixed Effects

- Some problems with FE:

  1. Suppose β is heterogeneous β_{Δ} and $\beta_{\sim\Delta}$ if X changes or not; then FE only captures β_{Δ}.

  2. In the presence of measurement error, fixed effects estimation can increase coefficient attenuation.

- Attenuation:

True:

$$y_i = \beta x_i + \epsilon_i$$

Observed

$$\bar{x}_i = x_i + \delta_i$$

$$\hat{\beta} = \frac{\sum_{t=1}^{t} x_i y_i}{\sum_{t=1}^{t} x_i^2} =$$

$$\frac{\sum_{t=1}^{t} \bar{x}_i \left(\beta x_i + \epsilon_i\right)}{\sum_{t=1}^{t} \bar{x}_i^2} = \frac{\sum_{t=1}^{t} \left(\beta x_i^2 + \beta \delta_i x_i + x_i \epsilon_i + \delta_i \epsilon_i\right)}{\sum_{t=1}^{t} \left(x_i^2 + 2 x_i \delta_i + \delta_i^2\right)}$$

$$E\hat{\beta} = \frac{\beta \sum_{t=1}^{t} x_i^2}{\sum_{t=1}^{t} \left(x_i^2 + \delta_i^2\right)} = \frac{\beta \sigma_x^2}{\sigma_x^2 + \sigma_\delta^2} = \beta \left(1 - \frac{\sigma_\delta^2}{\sigma_x^2 + \sigma_\delta^2}\right) < \beta$$

- $\frac{\sigma_\delta^2}{\sigma_x^2 + \sigma_\delta^2}$ is called the reliability ratio

- Measurement Error Tradeoff

Suppose T=2;

$$y_{1i} = \alpha Z_i + \beta X_{1i} + \mu_i + \epsilon_{1i}$$

$$y_{2i} = \alpha Z_i + \beta X_{2i} + \mu_i + \epsilon_{2i}$$

$$\mu_i = \gamma X_{1i} + \gamma X_{2i} + \delta Z_i + \omega_i$$

Then $\hat{\beta}_{FE} = \hat{\beta}_{FD}$ comes from the regression $y_{1i} - y_{2i} = \beta\left(X_{1i} - X_{2i}\right) + \epsilon_{1i} - \epsilon_{2i}$

It can be shown that

$$\hat{\beta}_{FE} = \beta\left(1 - \frac{\sigma_\delta^2}{\left[\sigma_x^2 + \sigma_\delta^2\right](1 - \rho_X)}\right)$$

where ρ_X is the correlation coefficient of X within the "fixed effect" group: $\frac{cov(X_{1i}X_{2i})}{\sigma_X^2}$

- So there is a tradeoff: bias from exclusion of the fixed effect versus bias due to exacerbation of the attenuation in the presence of measurement error with highly correlated X's.

- Intuition: if the X's are highly correlated, then when using fixed effects, most of the variation left is measurement error.

TABLE 1

DESCRIPTIVE STATISTICS

	Overall Contested Elections (1)	Subsample of Repeat Meetings: First Meeting (2)	Subsample of Repeat Meetings: Later Meetings (3)
	A. Statistics		
Democratic percentage of vote	54.1	55.1	55.5
	(18.0)	(16.6)	(16.4)
Incumbent's percentage of vote	66.8	64.2	63.4
	(10.1)	(10.8)	(10.8)
Success rate for incumbents seeking reelection	94.8	94.1	89.5
	B. Breakdown by Status of Incumbent		
Democratic	52.8	55.5	62.6
Republican	36.2	34.8	37.4
Open seat	10.9	9.7	...
	C. Campaign Spending per Candidate (Thousands of 1990 Dollars)		
Incumbents	293	266	343
Challengers	136	134	173
Open seat	409	275	...
Observations	2,781	299	334

NOTE.—Numbers in parentheses are standard deviations. Col. 1, except for spending data, is drawn from the data set used in Levitt and Wolfram (1994). See n. 8 for further information. Spending data in col. 1 are unweighted averages of real spending for all major party candidates in general elections between 1972 and 1990. They are based on Common Cause (1974, 1976), Sorauf (1988, table 6-1), and multiple editions of the Federal Election Commission *Reports on Financial Activity*.

ple; in the subsample, this margin is reduced by about three percentage points. The percentage of beaten incumbents is higher in the subsample, especially when the opponents have met previously. The increased rate of challenger success in repeat bids is attributable to the fact that politicians appear to behave strategically (Jacobson 1989); repeat challenges are far more likely in those years in which national political conditions favor the challenger's party. For instance, in the aftermath of Watergate in 1974, 19 Democrats who had previously run for office chose to challenge again, compared to only three Republicans. Similarly, in the Reagan landslide of 1980, repeat Republican challengers outnumbered repeat Democratic challengers almost three to one. When national political conditions are controlled for in the regression analysis of the following section, the differences between first meetings and repeats disappear.

TABLE 2

CROSS-SECTIONAL ESTIMATES FOR SUBSAMPLE OF REPEAT CHALLENGERS: LINEAR MODEL

Variable	First Meeting (1)	Later Meetings (2)	Later Meetings (3)
Challenger spending	−2.7	−2.7	−.8
	(.3)	(.3)	(.2)
Incumbent spending	−.2	−.2	.1
	(.3)	(.2)	(.1)
Competitor party strength:			
Vote share (−1)	.54		.78
	(.07)		(.04)
Vote share in last election with different challenger		.21	
		(.05)	
Constant	34.0	48.0	14.2
	(5.2)	(7.8)	(3.2)
Adjusted R^2	.56	.55	.76

NOTE.—Dependent variable is incumbent's share of the two-party vote. Standard errors are in parentheses. Spending variables are in terms of $100,000 of 1990 dollars. Year dummies (not shown) were included in all regressions.

candidate 2.7 percent of the vote, whereas marginal spending by the incumbent has essentially no impact on the election outcome. There does not appear to be a systematic difference in the effects of campaign spending between the first time two candidates meet and subsequent elections. It is important to note that the results in columns 1 and 2 are indistinguishable from the results of previous studies using cross-sectional data, collected in table 3. As a consequence, any differences between the results obtained in applying the panel data model of the following section and the results of past cross-sectional analyses must be attributed to the methodological approach, not the sample being analyzed.

Column 3 of table 2 provides an informal test of the standard cross-sectional approach. If challenger quality were adequately controlled for using a cross-sectional approach, the results in columns 1–3 should be similar. Note, however, that in column 3, where a better control for challenger quality is available, the impact of challenger spending shrinks to less than one-third of the previous estimates. The proportion of the variance explained by the model rises substantially as well. The results of column 3 suggest that failure to adequately control for challenger quality in previous cross-sectional analyses has led to an upward bias in the effects of challenger spending. The estimates obtained in the following section further reinforce that conclusion.

TABLE 4

RESULTS OF THE REGRESSION MODEL

Variable	Linear Spending (1)	Square Root Spending (2)	Log Spending (3)
Challenger spending	.30	.13	1.04
	(.19)	(.06)	(.50)
Incumbent spending	.09	.06	.61
	(.13)	(.06)	(.75)
Open-seat spending	.17	.09	.67
	(.44)	(.13)	(1.40)
Incumbency	3.2	3.7	3.5
	(.8)	(1.1)	(1.7)
Scandal dummy	4.8	4.8	5.0
	(1.4)	(1.4)	(1.4)
1990	...	...	...
1988	.6	.7	.6
	(.7)	(.7)	(.7)
1986	1.6	1.7	1.7
	(.9)	(.9)	(.9)
1984	−1.4	−1.3	−1.4
	(1.1)	(1.1)	(1.1)
1982	2.0	2.2	2.1
	(1.3)	(1.3)	(1.3)
1980	...	...	...
1978	2.3	2.3	2.5
	(.7)	(.7)	(.7)
1976	4.0	4.0	4.2
	(.9)	(.9)	(.9)
1974	5.0	4.9	5.1
	(1.1)	(1.1)	(1.1)
1972	−2.0	−1.9	−1.7
	(1.4)	(1.4)	(1.4)
Adjusted R^2	.24	.24	.24
F-test*	.85	1.64	1.36

NOTE.—The dependent variable is the Democratic percentage of the two-party vote. White heteroskedasticity-consistent standard errors are in parentheses. Spending variables are in terms of $100,000 of 1990 dollars. All variables except for year dummies are multiplied by the incumbency indicator variable (see Sec. III for further explanation). Year dummies for the 1970s are relative to 1980; year dummies for 1980s are relative to 1990. Adjusted R^2 value refers to the percentage of variance explained after the fixed-effects transformation. In col. 3, candidates spending less than $1,000 are treated as though they spent $1,000. Degrees of freedom are equal to 320 in all regressions.

* F-test of spending coefficients equal to zero.

roots, and natural logs, respectively.[13] Heteroskedasticity-consistent standard errors (see White 1980) are in parentheses. The adjusted R^2 values are similar across regressions, as are the values and stan-

[13] The model was also estimated using various permutations of the ratio of campaign spending between the incumbent and the challenger. The results of those regressions were completely consistent with the results presented below and are available from the author on request.