8. Presenting Panel Data Models

The key question now is, "Which information do we have to report? And how?" Some studies report parameter estimates and their statistical significances only; Others include standard errors but exclude goodness-of-fit measures; And oftentimes researchers fail to interpret the results substantively for readers. This section discusses general guidelines for presenting panel data models. However, specific pieces of information to be presented and their styles depend on research questions and purpose of studies.

8.1 Presenting All Possible Models? No!

Some studies present all possible models including the pooled OLS, fixed effect model, random effect models, and two-way effect model. Is this practice reasonable? No. Strictly speaking, if one model is "right," the other models are "wrong." It must be absurd to present wrong models together unless comparison of models is the goal of the study. If a fixed effect turns out insignificant, why are you trying to present the "wrong" model? In short, you just need to report a "right" model or your final model only.

8.2 Which Information Should Be Reported?

You MUST report goodness-of-fit measures, parameter estimates with their standard errors, and test results (See Table 8.1).

8.2.1 Goodness-of-fit Measures

Goodness-of-fit examines the extant that the model fits data. In case of poor goodness-of-fit, you need to try other model. The essential goodness-of-fit measures that you need to report are,

- F-test (or likelihood rato test) to test the model and its significance (p-value).
- Sum of squared errors (residual), degrees of freedom for errors, and N (*nT*).
- R^2 in OLS and fixed effect models.
- Theta θ and variance components $\hat{\sigma}_u$ estimated in a random effect model.

Keep in mind that some estimation methods report incorrect statistics and standard errors. For example, `.xtreg` returns incorrect R^2 in a fixed effect model because the command fits the "within" estimator (running OLS on transformed data with the intercept suppressed). Both between R^2 and overall R^2 displayed in Stata output are almost meaningless. In order to get correct R^2 for a fixed effect model, use LSDV1 or `.areg`. Use macro variables, if needed, to obtain various goodness-of-fit measures that are not displayed in Stata output.

8.2.2 Parameter Estimates of Regressors

Obviously, you must report parameter estimates and their standard errors.
Fortunately, `.regress` and `.xtreg` produce correct parameter estimates and their adjusted standard errors. But the "within" estimation itself produces incorrect standard errors due to incorrect (larger) degrees of freedom (see Table 3.2).

8.2.3 Parameter Estimates of Dummy Variables

In a fixed effect model, a question is if individual intercepts need to be reported. In general, parameter estimates of regressors are of primary interest in most cases and accordingly individual intercepts are not needed. However, you have to report them if audience wants to know or individual effects are of main research interest. The combination of LSDV1 or LSDV2 will give you easy solutions for this case (see 4.2).[20] Do not forget that LSDV1, LSDV2, and LSDV3 have different meanings of dummy parameters and that null hypotheses of t-test differ from one another (see Table 4.3).

8.2.4 Test Results

Finally, you should report if fixed and/or random effect exists because panel data modeling is to examine fixed and/or random effects. Report and interpret the results of F-test for a fixed effect model and/or Breusch-Pagan LM test for a random effect model. When both fixed and random effects are statistically significant, you need to conduct a Hausman test and report its result.[21] If you doubt constant slopes across group and/or time, conduct a Chow test to examine the poolability of data.

Table 8.1 Examples of Presenting Analysis Results

	Pooled OLS	Fixed Effect Model	Random Effect Model
Ouput index	.8827**	.9193**	.9067**
	(.0133)	(.0299)	(.0256)
Fuel price	.4540**	.4175**	.4228**
	(.0203)	(.0152)	(.0140)
Loading factor	-1.6275**	-1.0704**	-1.0645**
	(.3453)	(.2017)	(.2001)
Intercept (baseline)	9.5169**	9.7930**	9.6279**
	(.2292)	(.2637)	(.2102)
Airline 1 (dummy)		-.0871	
		(.0842)	
Airline 2 (dummy)		-.1283	
		(.0757)	
Airline 3 (dummy)		-.2960**	
		(.0500)	
Airline 4 (dummy)		.0975**	
		(.0330)	
Airline 5 (dummy)		-.0630**	
		(.0239)	
F-test (model)	2419.34**	3935.79**	19642.72**
DF	86	81	81
R^2	.9883	.9974	
SSE (SRMSE)	1.3354	.2926	.3116
SEE or $\hat{\sigma}_v$	.1246	.0601	.0602
$\hat{\sigma}_u$			.1249
θ			.8767
Effect Test		57.7319**	334.8496**
N	90	90	90

Source: http://pages.stern.nyu.edu/~wgreene/Text/tables/tablelist5.htm
* Standard errors in parenthesis; ** Statistical significance: * <.05, ** <.01

[20] In some uncommon cases, you need to report variance components estimated in a random effect model. Unlike the SAS MIXED procedure, `.xtreg` does not report these statistics.

[21] The null hypothesis is that group/time specific effects are not correlated with any regressors. Either "A random effect model is better than the fixed effect model" or "A random effect model is consistent" is NOT a correct null hypothesis. If the null hypothesis is rejected, a random effect model violates a key OLS assumption 2 and ends up with biased and inconsistent estimates; however, a fixed effect model still remains unbiased and consistent.

8.3 Interpreting Results Substantively

If your model fits the data well and individual regressors turn out statistically significant, you have to interpret parameter estimators in a "sensical" way. You may not simply report signs and magnitude of coefficients. Do not simply say, for example, an independent variable is "significant," "negatively (or positively) related to…", or "insignificantly related..."

A standard form of interpretation is, "For one unit increase in an IV, DV is expected to increase by OO units, holding all other variables constant." You may omit the *ceteris paribus* assumption (holding all other variables constant). However, try to make interpretation more sense to audience who does not know much about econometrics. See 4.1 and 4.2 for examples of substantive interpretation.

Provide statistical significance in a table and the p-value in parenthesis at the end of the interpretation sentence.

8.4 Presenting Results Professionally

Many studies often present results in tables but some of them fail to construct professional tables. Common bad table examples include 1) large and various fonts, 2) too small and/or too large numbers, 3) colorful and stylish border lines, 4) badly aligned numbers, and 5) non-systematic order. The following is the list of checkpoints to be considered when constructing a professional table (see Table 8.1 for an example).

- Title should describe the contents of a table appropriately. Provide unit of measurement (e.g., Million Dollars) and period (e.g., Year 2010) if needed.
- Organize a table systemically and compactly.
- Provide parameter estimates and their standard errors in parenthesis.
- Do not use variable names used in computer software as labels. Use loading factor instead of `load`.
- Use 10 point *Times New Roman* for labels and 10 point *Courier New* for numbers. Do not use stylish fonts (e.g., *Cooperplate*) and too big or too small size.
- Rescale numbers appropriately in order to avoid such numbers as "0.00004455" or "75,845,341,697,785."
- Report up to three or four digits below the decimal point. Do not round numbers arbitrarily.
- Do not use stylish border lines (e.g., their colors, thickness, and type of lines).
- Minimize use of vertical and horizontal lines. Use no vertical line in general.
- Align numbers to the right and consider the location of decimal point carefully.
- Use "Standardized coefficients," if needed, rather than "Beta," "β," or "beta coefficients." Nobody knows the true value of β.
- Provide the source of data, if applicable, at the bottom of the table.
- Indicate statistical significance as *<.05, **<.01.

8.5 Common Mistakes and Awkward Expressions

It is not difficult to find awkward expressions even in academic papers. Consider following suggestions for common mistakes in presentation.

8.5.1 Statistical Significance

Do not say, "significant level," "at 5% level," or "at the level of significance α=5%," and the like. These expressions should be "significance level," "at the .05 level," and "at the .05 (significance) level," respectively. Use a specific significance level (e.g., "at the .01 significance level") rather than "at the conventional level."

8.5.2 Hypothesis

A hypothesis is a conjecture about the unknown (e.g, α, β, δ, and σ). Therefore, "$b_1 = 0$" is not a valid hypothesis, but "$\beta_1 = 0$" is. Because the b_1 is already known (estimated from the sample), you do not need to test $b_1 = 0$.

8.5.3 Parameter Estimates

Say, "parameter estimates of β_1" or "the coefficient of an independent variable 1" instead of "The coefficient of β_1." Also say, "standardized coefficients" instead of "Beta," β, or "beta coefficient."

8.5.4 P-values

Do not say, "The p-value is significant." A p-value itself is neither significant nor insignificant. You may say, "The p-value is small enough to reject H_0" or "A small p-value suggests rejection of H_0."

8.5.5 Reject or Do Not Reject the Null Hypothesis

Say, "reject" or "do not reject" the null hypothesis rather than "accept (or confirm)" the null hypothesis. Also say "reject the H_0 at the .01 level" instead of "I do not believe that the H_0 is true" or "The test provides decisive evidence that the H_0 is wrong" (no one knows if a H_0 is really true or wrong). Always be simple and clear.