Informal Rules on Figures and Tables

Josh Clinton
Vanderbilt University

How you present your work means at least as much as what the results show. If the results are hard to read or interpret, the importance of your findings may be lost. If the results appear to be quickly written (or shoddily presented), I will draw inferences about the quality of the work that generated the estimates. If you do not care enough to spend the time to make the tables and figures attractive and compelling, I will not spend the time to pay much attention to them. Because tables and figures are the heart of your paper and contain the core of any empirical contribution this is not the outcome you should desire.

Here are some rules of thumb. Not everyone will agree with these, but these are just my thoughts on the issue.

- Do not include Tables or Figures you do not talk about: Figures and tables take up a lot of room and if you do not care about them enough to talk about them in the paper you should certainly not include them.

- Variables should be Interpretable by your Mom: Calling a variable "underscore cons" (or other STATA-speak) is horrific. Calling a variable reppct96 is bad. Calling it *reppct96* is better because you are at least acknowledging that the word is not a word. Ideal would be *1996 % Republican Two-Party Presidential Vote* (or *96 Rep. Two-Pty Pres. Vote %.*

- Two digits post decimal point: Unless you have reason to do so, you should never report results to three-decimal points. Do you really think your results are that precise? Long digits make for visually messy tables.

- Common length: As much as possible, report the same number of digits – e.g., "4.00" or "0.93." This provides for visually clean tables.

- Logged Variables: Should be *log(votes)* or *logged votes*, not *logvotes*.

- 0.00000*: A personal peeve – what does this mean? Either the coefficient is 0 (in which case it cannot be significant), or it is not in which case you have provided no information other than the fact it is small. Either: rescale the variable, or break the above rule and include enough digits to get a non-zero.

- Labels: Every Table or Figure should be clearly labelled. What Model is being run? Using which data? I should be able to understand what your paper is doing by looking only at your figures and tables.

- Table Consistency: Your tables should look the same – it you have a line separating the table from the text it should have such a line on the top and the bottom for every table. Do not distract the reader by changing table formats unless necessary. Formatting should not detract from my ability to read the paper.

- Centering: Tables should be centered in the text. Words should not wrap around tables and figures.

- LINES: Lines in a table are useful for separating rows and columns by focusing the eye and how readers read the column. Do you want people to read across? down? Having both row and column borders is confusing – what are you telling your readers about what they should focus on?

- STATA $\neq$ FORMATTING: Just because STATA reports something does not mean it should find its way into a table. Use your expertise to determine what you think readers need to know.

- EXCEL $\neq$ FORMATTING: See above. Show evidence that you thought about the presentation of the data beyond cutting and pasting.

- COLUMN REDUNDANCY: Don't have columns that contain the same information – Coefficient and Standard Error, "Z-Score" "$P > Z$" and "95 % Confidence Intervals" all contain the same information. Depending on how many models are in a Table, I may or may not include confidence intervals. I would never include "Z-Score" or "P>Z."

- OBSERVATIONS IN DESCRIPTIVE STATISTICS: Unless this changes across variables, report this in the caption, do not waste a column for this in the table.

- COLUMN LABELS IN A TABLE: Capitalize and use periods to denote abbreviations.

- REPORT MEANINGFUL FIT STATISTICS: Model F-stats (or Wald χ^2) are silly. Of course you can jointly reject the hypothesis that all coefficients are zero. R^2 has a meaning statistically. "Adjusted R^2 does not. Report the former. (The one exception is *maybe* "psuedo-R^2".)

- VARIABLES IN TEXT: Referring to a variable in the text should be italicized to distinguish it from the concept. Vote is a concept, *Vote* is a variable.

- FIGURE BACKGROUND: The background to your figures should be white – the same color as the paper. Do not have a "grey" border around the figures.

- FIGURE AXES LABELS: Should be capitalized and meaningful.

- FIGURE AXES: Graphs that are side-by-side should have the scale on the y-axis (unless they are histograms or density plots and that would be misleading). Axis should be similarly labeled is much as possible in terms of range and the increments of tick-marks.

- FIGURE LINES: If it is a line, you should plot a line, not a bunch of closely connected points. If you are overlaying the line over data, make sure the line is thick enough to see through the points.