

Template using MarkDoc and L^AT_EX within Stata .do File

Matthew C. Ingram¹

Nakissa Jahanbani¹

César Rentería²

¹ Department of Political Science

² Department of Public Affairs & Policy
Rockefeller College of Public Affairs & Policy

September 7, 2017

Abstract

Summary. Type your summary or abstract here, if desired. If there is none or if this part is not relevant, then omit.

1 Purpose of this document

This document offers you a template to improve your workflow in Stata using L^AT_EX and the MarkDoc package.

2 Template in greater detail

This is the template for all assignments you produce for RPAD/RPOS 517. Please use this template for each assignment, and please submit all assignments as PDF documents.

There are three versions of the template:

- 1 one for Stata using MarkDoc and L^AT_EX (.do file)
- 2 one for R using Sweave and L^AT_EX (.Rnw)
- 3 one for R using knitr and markdown (.Rmd file)

This template is the first one: Stata with MarkDoc and L^AT_EX.

Each template offers a combination of three main tools, including a statistical package (Stata or R), a text processing and presentation system or language (L^AT_EX, markdown, and/or HTML), and a system for weaving together code from the statistical package and the text and formatting from the text processing and presentation system (MarkDoc, Sweave, or knitr).

What you choose will likely depend on your needs, and this may vary from one project to another. For instance, HTML may be a better choice for producing web pages and other web content, and markdown may be a better choice for producing short reports with little formatting. It may also be the case that one collaboration you have uses one set of tools and another team you work with uses a different set of tools.

In general, however, I recommend learning and using L^AT_EX; this is a powerful document preparation and typesetting system that will meet all of your needs.

3 Getting started

At this point, I assume the following:

- you have access to Stata

- for student and educational pricing, see [here](#)
- I do not recommend Small Stata; you will quickly run into limitations
- you have downloaded and installed a .tex distribution and .tex editor see [Notes on TeX](#))

If all of the above is correct, open Stata.

This .do file using MarkDoc and L^AT_EX can then be easily opened and edited in Stata.

MarkDoc enhances your workflow in Stata by allowing you to integrate your Stata syntax with document preparation and typesetting tools to control the appearance of your content. Specifically, MarkDoc is designed to use three tools (markdown, HTML, and L^AT_EX).

For more on MarkDoc, see package documentation [here](#).

Once you have opened this .do file in Stata, you can run this .do file by:

- (a) placing your cursor at start of document and clicking the “Execute (do)” button at top-right above;
- (b) highlighting the parts you want to run and hitting CTRL-D (CMD-D on Mac), or
- (c) run the whole file by simply hitting CTRL-D (CMD-D on Mac).

MarkDoc converts your smcl logfile (.smcl) to a L^AT_EX file (.tex), which you will then need to compile. For compiling your .tex file and any related figures, tables, or other output, you will need a .tex distribution and a .tex editor. There are several distribution and editors available, but I recommend the MiKTeX distribution and TeXstudio editor for Windows, and the MacTeX distribution and TeXshop editor for Mac. Running your .do-file in Stata will generate a .tex file, and you will then open and compile this file in your .tex editor (e.g., TeXstudio), which will output a finished .pdf document (though other formats are available). To read a basic introduction to L^AT_EX, including suggestions for .tex editors, see [Notes on L^AT_EX](#).

When you run this do-file, a .tex file will be generated that includes both text content as well as any embedded Stata code and the output of any Stata code blocks within the document. You can embed a chunk of Stata code by placing it outside the forward slash and three asterisks and three asterisks and a forward slash as illustrated below with the “verbatim” command from L^AT_EX.

4 Quick intro to L^AT_EX

A few useful L^AT_EX commands for writing are:

5 This is heading 1

This is a heading 1 without numbering

5.1 This is heading 2

This is heading 2 without numbering

This is a text paragraph which requires no syntax.

This is a list of items:

- item 1
- item 2

This is a line introducing Equation 1 and Equation 2 below.

$$\alpha = \sqrt{\beta} \tag{1}$$

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon \tag{2}$$

In the blocks of Stata code below, we illustrate several typical steps in the course of writing a paper, including:

- 1 loading some data,
- 2 describing the data, including a table of summary statistics,
- 3 graphing the data,
- 4 running a simple regression model,
- 5 graphing the results.

6 Data Analysis Example

This document provides an example about how to use MarkDoc to add text to your logfile. To do so, I will use the winter olympics dataset for Michael Bailey’s statistics text, Real Stats (Oxford, 2016).

Data are in a folder called “data”, and all figures are exported to and then imported from a folder called “figures”.

7 Loading Data

In this section, we load our data.

All Stata commands have been suppressed with ‘/**/’ syntax at start of each line of code.

All unwanted results have also been suppressed by using the command ‘quietly’.

8 Descriptive Statistics

Here we import the table of summary statistics.

Now lets generate and add a figure that visualizing one of these variables.

Table 1: Summary Statistics

(1)					
	count	mean	sd	min	max
id	1122	58.87611	33.98619	1	118
year	1122	1996.52	10.6987	1980	2014
host	1121	.0089206	.0940687	0	1
country	0
temp	1100	53.94518	19.33287	3.9	90.5
precipitation	1100	56.036	45.02825	.3	280
elevation	1100	3017.599	2194.859	43	8850
gold	1122	.587344	1.926375	0	14
silver	1122	.583779	1.837421	0	16
bronze	1121	.5807315	1.695831	0	13
population	1113	51.92986	179.0457	0	1350.695
gdp	966	1.169092	1.806789	.011	14.523
participate	1122	.600713	.4899703	0	1
medals	1122	1.751337	5.170563	0	37
athletes	1122	18.16934	36.8805	0	230
time	1122	5.423351	2.882459	1	10
<i>N</i>	1122				

We can write some text and reference the figure by noting that Figure 1 shows a histogram of our data. **NOTE:** you may need to compile your document twice in order for the figure reference to be correct. \LaTeX compiles in layers, so if you do not see a recent change reflected in your compiled document, try compiling again.

9 Regression

Let's end this template with a regression analysis.

Regression was saved as a clean table in \LaTeX file to folder “output”, and we can now embed that clean,

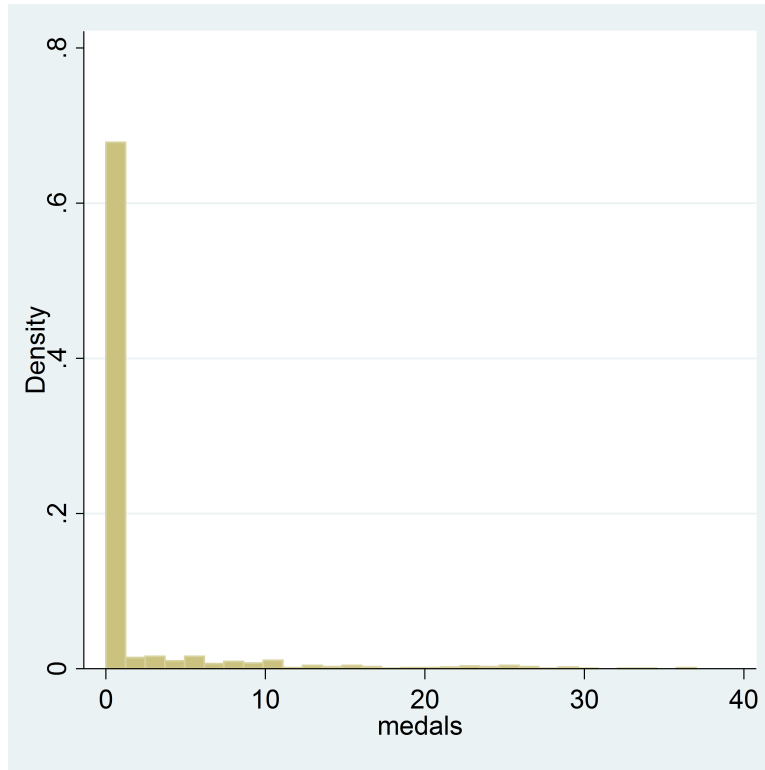


Figure 1: This graph shows the histogram of the price variable

formatted table directly into the document we are creating.

This is table with *outreg* command.

	medals
athletes	0.113 (0.003)**
gdp	0.172 (0.054)**
temp	0.011 (0.005)*
_cons	-1.225 (0.347)**
R-squared	0.71
Adjusted R2	0.71
RMSE	2.82
<i>N</i>	956
* $p < 0.05$; ** $p < 0.01$	

This is table with *esttab* command.

	(1)	(2)
	medals	medals
athletes	0.113*** (0.00266)	0.113*** (0.00274)
gdp	0.123* (0.0566)	0.172** (0.0540)
temp		0.0114* (0.00543)
_cons	-0.477*** (0.119)	-1.225*** (0.347)
<i>N</i>	966	956

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

And we can also generate a graph that reports a visualization of the coefficients from the regression analysis. Graphs like this can be an extremely useful and more intuitive way of communicating your results.

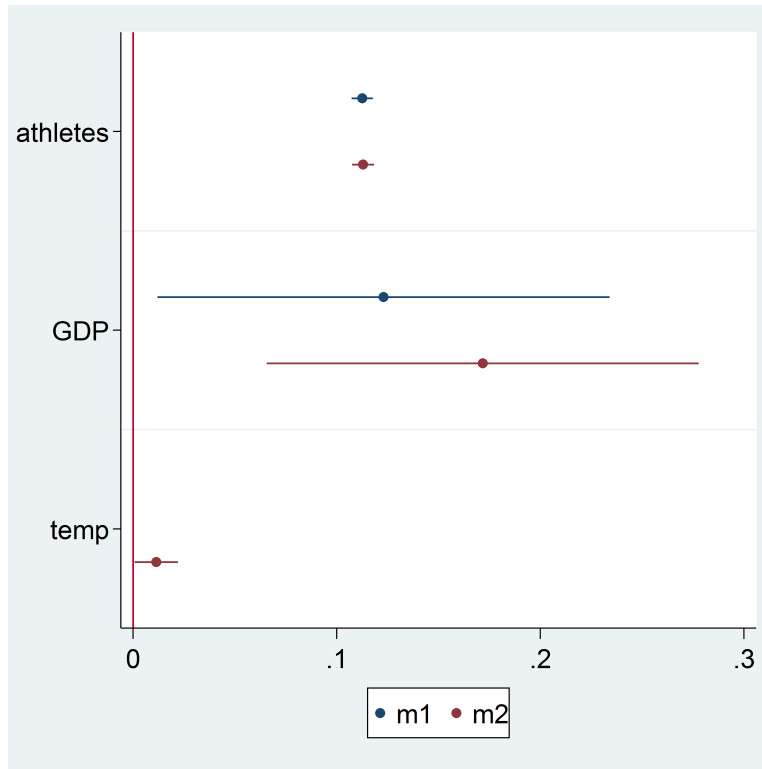


Figure 2: This graph shows the coefficients for both regressions.

Then we can write some text and reference the figure by noting that Figure 2 reports the coefficients. **NOTE:** Again, you may need to compile your document twice in order for the figure reference to be correct. \LaTeX compiles in layers, so if you do not see a recent change reflected in your compiled document, try compiling again.

10 Conclusion

Finally, we would write our conclusion here.