Creating LaTeX documents from within Stata

Ben Jann

ETH Zürich

German Stata Users Group Meeting

Bonn, June 26, 2009
Introduction

- texdoc is a new command to create \LaTeX\ documents from within Stata.
- texdoc is especially convenient to create \LaTeX\ documents that contain Stata output.
- texdoc is like weaving, but all Stata.
- I use it for teaching, e.g. to create solutions for class assignments.
- I also use it for Stata Journal articles.
Syntax and Usage

- Create a \LaTeX\ document (interactive mode)

  \texttt{texdoc\_init\ docname [\ , \_replace ]}
  \texttt{tex\ line 1}
  \texttt{tex\ line 2}
  \texttt{\ldots}
  \texttt{texdoc\_close}
Syntax and Usage

- Include Stata output in \LaTeX\ document

  \begin{verbatim}
  texdoc _init docname [ , replace ... ]
  ...
  
  texdoc _stlog [name]
  ... commands ...
  texdoc _stlog _close
  ...
  texdoc _close
  \end{verbatim}
Within `texdoc stlog`, type

```
texdoc _stlog oom command
```

... to suppress output (and print `"(output omitted)"`).

Furthermore, within `texdoc stlog` type

```
texdoc _stlog cnp
```

... to continue output on next page (and print `"(continued on next page)"`).
Syntax and Usage

- Non-interactive mode: Process a do-file containing `texdoc` commands.

```stata
texdoc do filename [, init(docname) close replace ]
```

- In non-interactive mode you can use the `/*tex ... tex*/` comment structure to include blocks of \LaTeX{} code.

- `init()` and `close` can also be specified within the do-file using `texdoc init` and `texdoc close`.

- Get rid of all \LaTeX{} and `texdoc` commands:

```stata
texdoc strip oldfile newfile [, replace ]
```
Examples
Create homework assignment (interactive mode)

. texdoc init assignment
(texdoc output file is assignment.tex)
. tex \documentclass[12pt]{article}
. tex
. tex \begin{document}
. tex
. tex \section*{Assignment A}
. tex
. tex \subsection*{Exercise 1}
. tex
. tex Open auto.dta and describe the data.
. tex
. tex \subsection*{Exercise 2}
. tex
. tex Run some regressions.
. tex
. tex \subsection*{Exercise 3}
. tex
. tex Draw a scatter plot.
. tex
. tex \subsection*{Exercise 4}
. tex
. tex Draw a histogram.
. tex
. tex \end{document}
. texdoc close (texdoc output written to assignment.tex)
Solutions to assignment (non-interactive mode)

```
. type solutions.do
* Solutions to Assignment A
*/tex
\documentclass[12pt]{article}
\usepackage{stata, graphicx}
\begin{document}
\section*{Assignment A}
\subsection*{Exercise 1}
Open auto.dta and describe the data.
\begin{verbatim}
texdoc stlog
sysuse auto(1978 Automobile Data)
summarize
\end{verbatim}

\begin{verbatim}
<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>make</td>
<td>74</td>
<td>6165.257</td>
<td>2949.496</td>
<td>3291</td>
<td>15906</td>
</tr>
<tr>
<td>price</td>
<td>74</td>
<td>21.2973</td>
<td>5.785503</td>
<td>12</td>
<td>41</td>
</tr>
<tr>
<td>mpg</td>
<td>74</td>
<td>3.405797</td>
<td>.9899323</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>trunk</td>
<td>74</td>
<td>13.75676</td>
<td>4.277404</td>
<td>5</td>
<td>23</td>
</tr>
<tr>
<td>weight</td>
<td>74</td>
<td>187.9324</td>
<td>22.26634</td>
<td>142</td>
<td>233</td>
</tr>
<tr>
<td>length</td>
<td>74</td>
<td>39.64865</td>
<td>4.399354</td>
<td>31</td>
<td>51</td>
</tr>
<tr>
<td>disp</td>
<td>74</td>
<td>197.2973</td>
<td>91.83722</td>
<td>79</td>
<td>425</td>
</tr>
<tr>
<td>gear_ratio</td>
<td>74</td>
<td>3.014865</td>
<td>.4562871</td>
<td>2.19</td>
<td>3.89</td>
</tr>
<tr>
<td>foreign</td>
<td>74</td>
<td>.2972973</td>
<td>.4601885</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
\end{verbatim}
As we can see, the mean price is 6165.
\begin{verbatim}
\end{verbatim}
\subsection*{Exercise 2}
Run some regressions.
\begin{verbatim}
texdoc stlog
regress price weight mpg
\end{verbatim}
\begin{verbatim}
| Source | SS | df | MS | Number of obs = 74
|--------|----|----|----|----|
| Model  | 186321280 | 2 | 93160639.9 | Prob > F = 0.0000
| Residual | 448744116 | 71 | 6320339.67 | R-squared = 0.2934
| Total   | 635065396 | 73 | 8699525.97 | Adj R-squared = 0.2735
\end{verbatim}
\begin{verbatim}
| price  | Coef. | Std. Err. | t | P>|t| |
|--------|-------|-----------|---|-----|
| weight | 1.746559 | .6413538 | 2.72 | 0.008 |
| mpg    | -49.51222 | 86.15604 | -0.57 | 0.567 |
| _cons  | 1946.069 | 3597.05 | 0.54 | 0.590 |
\end{verbatim}
\begin{verbatim}
\end{verbatim}
\subsection*{Exercise 3}
Draw a scatter plot.
\begin{verbatim}
texdoc stlog
scatter price mpg
\end{verbatim}
\begin{verbatim}
\end{verbatim}
\subsection*{Exercise 4}
Draw a histogram.
\begin{verbatim}
texdoc stlog
hist price(bin=8, start=3291, width=1576.875)
\end{verbatim}
\begin{verbatim}
\end{verbatim}
\end{document}
```
Issues

- texdoc do always runs everything, that is it
  - cannot process \LaTeX{} without running Stata commands,
  - cannot run Stata commands without processing \LaTeX{}.

- An option to copy the pieces of Stata output directly into the \LaTeX{} document instead of using external log files would be nice.

- texdoc stlog relies on sjlog, which has some limitations:
  - linesize is fixed
  - closes the (unnamed) default log

- texdoc do does not exit the do-file on exit

- Overall, texdoc is only for small documents. I would not use it to produce a whole book or so. Also, texdoc is suited primarily for documents where Stata plays an important role. If the document is mostly \LaTeX{} then an alternative approach should probably be followed . . .
Alternative approach

- Make \LaTeX{} the default and tag Stata commands.
- That is, define a \LaTeX{} document containing blocks of Stata code such as

\begin{verbatim}
\begin{stata}
... commands ...
\end{stata}
\end{verbatim}

\begin{verbatim}
\begin{stlog}
... commands ...
\end{stlog}
\end{verbatim}

and then process the file e.g. as follows

\texttt{dotex filename}

Implementation would not be much more complicated than the implementation of \texttt{texdoc}. 

Thanks for listening!