

Creating LaTeX and HTML documents from within Stata using texdoc and webdoc

Example 2

Ben Jann
University of Bern, ben.jann@soz.unibe.ch

Nordic and Baltic Stata Users Group meeting
Oslo, September 13, 2016

Contents

1	The texdoc source file	2
2	The resulting \LaTeX source file	6
3	The resulting PDF	8

1 The texdoc source file

— *crosswise09.texdoc* —

```
texdoc init crosswise09, replace logdir logall

/**
\documentclass{article}
\usepackage{stata}
\usepackage{graphicx, hyperref}

\title{My Analysis of the Crosswise09 Data}
\author{Ben Jann}
\date{\today}

\begin{document}

\maketitle

\tableofcontents

\section{Settings}

***/

about
version 14.1
clear all
set linesize 100
set type double
set more off
// required user packages: fre, rrreg, rrlogit, estout, coefplot

/**

\section{Number of Observations}

***/

use crosswise09.dta
fre uni
fre version
fre f7
tab uni version, chi2 exact
gen byte touse = (f7!=1) | inlist(b1,1,2) | inlist(b2,1,2) ///
                | inlist(f11,1,2) | inlist(f12,1,2)

fre touse
fre version if touse
tab uni version if touse, chi2 exact

/**
```

```
\section{Item-Nonresponse}
```

```
***/
```

```
gen byte partial    = b1==1 if inlist(b1,1,2) & version==1 & touse
gen byte severe     = b2==1 if inlist(b2,1,2) & version==1 & touse
replace partial    = f11==1 if inlist(f11,1,2) & version==2 & touse
replace severe     = f12==1 if inlist(f12,1,2) & version==2 & touse
gen byte missing = (partial>=.)
tab version missing if touse, missing row
drop missing
gen byte missing = (severe>=.)
tab version missing if touse, missing row
drop missing
gen byte missing = (partial>=.) | (severe>=.)
tab version missing if touse, missing row
drop missing
count if (partial<.) & (severe>=.) & touse
count if (partial>=.) & (severe<.) & touse
```

```
/***
```

```
\section{Table 1: Descriptives}
```

```
***/
```

```
su f2 if touse
egen byte agecat = cut(f2) if touse, at(18,23,28,`r(max)')
tab agecat version if touse, chi2 exact col
tab f1 version if touse, chi2 exact col
tab f3 version if touse, chi2 exact col
```

```
/***
```

```
\section{Table 2: Prevalence estimates}
```

```
***/
```

```
gen byte crosswise = version==1 if touse
gen pyes           = cond(crosswise, 0.25, 1) if touse
```

```
// - direct questioning (DQ)
```

```
reg partial if crosswise==0
reg severe  if crosswise==0
```

```
// - crosswise model (CM)
```

```
rrreg partial if crosswise, pw(pyes)
rrreg severe  if crosswise, pw(pyes)
```

```

// - difference between DQ and CM

rrreg partial crosswise, pw(pyes) robust hc2
rrreg severe crosswise, pw(pyes) robust hc2

/**

\section{Table 3: Regression estimates}

***/

gen byte female      = f1==2 if inlist(f1,1,2)
gen byte internet    = f8_4==1
gen byte students    = f8_5==1
gen byte papers3or4 = f7==3 if f7<.
gen byte papers5     = f7==4 if f7<.
gen byte zurich      = uni==1
gen byte munich      = uni==2

su partial crosswise zurich munich female ///
  papers3or4 papers5 internet students if partial<.
su partial crosswise zurich munich female ///
  papers3or4 papers5 internet students if partial<. & crosswise==1

eststo reg1: rrreg partial ///
  zurich munich ///
  female papers3or4 papers5 internet students ///
  if crosswise, pw(pyes) robust
test zurich = munich

eststo logit1: rrlogit partial ///
  zurich munich ///
  female papers3or4 papers5 internet students ///
  if crosswise, pw(pyes) robust
test zurich = munich

eststo reg2: rrreg partial crosswise ///
  zurich munich ///
  female papers3or4 papers5 internet students ///
  , pw(pyes) robust
test zurich = munich

eststo logit2: rrlogit partial crosswise ///
  zurich munich ///
  female papers3or4 papers5 internet students ///
  , pw(pyes) nolog robust
test zurich = munich

esttab reg1 reg2 logit1 logit2, order(crosswise) ///
  compress mtitle nonumber star(+ 0.1 * 0.05 ** 0.01 *** 0.001)

```

```
/***/  
  
\section{Graph: Regression estimates}  
  
***/  
  
coefplot reg2 reg1 || logit2 logit1, drop(_cons) xline(0) ///  
    bylabels(LPM Logit) byopts(xrescale legend(off))  
texdoc graph  
  
/***/  
  
\end{document}  
  
— end of file —
```

2 The resulting L^AT_EX source file

Applying

```
. texdoc do crosswise09.texdoc
```

generates to the following L^AT_EX file.

— *crosswise09.tex* —

```
\documentclass{article}
\usepackage{stata}
\usepackage{graphicx, hyperref}

\title{My Analysis of the Crosswise09 Data}
\author{Ben Jann}
\date{\today}

\begin{document}

\maketitle

\tableofcontents

\section{Settings}

\begin{stlog}\input{crosswise09/1.log.tex}\end{stlog}

\section{Number of Observations}

\begin{stlog}\input{crosswise09/2.log.tex}\end{stlog}

\section{Item-Nonresponse}

\begin{stlog}\input{crosswise09/3.log.tex}\end{stlog}

\section{Table 1: Descriptives}

\begin{stlog}\input{crosswise09/4.log.tex}\end{stlog}

\section{Table 2: Prevalence estimates}

\begin{stlog}\input{crosswise09/5.log.tex}\end{stlog}

\section{Table 3: Regression estimates}

\begin{stlog}\input{crosswise09/6.log.tex}\end{stlog}

\section{Graph: Regression estimates}

\begin{stlog}\input{crosswise09/7.log.tex}\end{stlog}
\begin{center}
\includegraphics{crosswise09/7}
```

`\end{center}`

`\end{document}`

— end of file —

3 The resulting PDF

The following pages display the resulting PDF after compiling the \LaTeX source file.

My Analysis of the Crosswise09 Data

Ben Jann

September 13, 2016

Contents

1	Settings	1
2	Number of Observations	2
3	Item-Nonresponse	3
4	Table 1: Descriptives	5
5	Table 2: Prevalence estimates	6
6	Table 3: Regression estimates	8
7	Graph: Regression estimates	12

1 Settings

```
. about
Stata/MP 14.2 for Mac (64-bit Intel)
Revision 06 Sep 2016
Copyright 1985-2015 StataCorp LP
Total physical memory:      8.01 GB
30-user 2-core Stata network perpetual license:
  Serial number: 501406208640
  Licensed to: Ben Jann
                University of Bern

. version 14.1
. clear all
. set linesize 100
. set type double
. set more off
. // required user packages: fre, rrreg, rrlogit, estout, coefplot
```

2 Number of Observations

```
. use crosswise09.dta
```

```
. fre uni
```

```
uni — location of data collection
```

		Freq.	Percent	Valid	Cum.
Valid	1 ETH Zurich	111	23.42	23.42	23.42
	2 LMU Munich	90	18.99	18.99	42.41
	3 University Leipzig	273	57.59	57.59	100.00
	Total	474	100.00	100.00	

```
. fre version
```

```
version — experimental condition
```

		Freq.	Percent	Valid	Cum.
Valid	1 crosswise	358	75.53	75.53	75.53
	2 direct	116	24.47	24.47	100.00
	Total	474	100.00	100.00	

```
. fre f7
```

```
f7 — number of papers
```

		Freq.	Percent	Valid	Cum.
Valid	1 none	65	13.71	13.74	13.74
	2 one or two	178	37.55	37.63	51.37
	3 three or four	111	23.42	23.47	74.84
	4 five or more	119	25.11	25.16	100.00
	Total	473	99.79	100.00	
Missing	.	1	0.21		
Total		474	100.00		

```
. tab uni version, chi2 exact
```

```
Enumerating sample-space combinations:
```

```
stage 3: enumerations = 1
```

```
stage 2: enumerations = 2
```

```
stage 1: enumerations = 0
```

location of data collection	experimental condition		Total
	crosswise	direct	
ETH Zurich	85	26	111
LMU Munich	68	22	90
University Leipzig	205	68	273
Total	358	116	474

```
Pearson chi2(2) = 0.0942 Pr = 0.954
```

```
Fisher's exact = 0.967
```

```
. gen byte touse = (f7!=1) | inlist(b1,1,2) | inlist(b2,1,2) ///  
> | inlist(f11,1,2) | inlist(f12,1,2)
```

```
. fre touse
```

```
touse
```

		Freq.	Percent	Valid	Cum.
Valid	0	64	13.50	13.50	13.50
	1	410	86.50	86.50	100.00
	Total	474	100.00	100.00	

```
. fre version if touse
version — experimental condition
```

		Freq.	Percent	Valid	Cum.
Valid	1 crosswise	313	76.34	76.34	76.34
	2 direct	97	23.66	23.66	100.00
	Total	410	100.00	100.00	

```
. tab uni version if touse, chi2 exact
```

Enumerating sample-space combinations:

stage 3: enumerations = 1

stage 2: enumerations = 1

stage 1: enumerations = 0

location of data collection	experimental condition		Total
	crosswise	direct	
ETH Zurich	55	16	71
LMU Munich	66	21	87
University Leipzig	192	60	252
Total	313	97	410

Pearson chi2(2) = 0.0639 Pr = 0.969
Fisher's exact = 0.972

3 Item-Nonresponse

```
. gen byte partial = b1==1 if inlist(b1,1,2) & version==1 & touse
(164 missing values generated)
. gen byte severe = b2==1 if inlist(b2,1,2) & version==1 & touse
(164 missing values generated)
. replace partial = f11==1 if inlist(f11,1,2) & version==2 & touse
(96 real changes made)
. replace severe = f12==1 if inlist(f12,1,2) & version==2 & touse
(96 real changes made)
. gen byte missing = (partial>=.)
. tab version missing if touse, missing row
```

Key
<i>frequency</i>
<i>row percentage</i>

```
experimental |
al | missing
```

condition	0	1	Total
crosswise	310 99.04	3 0.96	313 100.00
direct	96 98.97	1 1.03	97 100.00
Total	406 99.02	4 0.98	410 100.00

```
. drop missing
. gen byte missing = (severe>=.)
. tab version missing if touse, missing row
```

Key
frequency row percentage

experimental condition	missing		Total
	0	1	
crosswise	310 99.04	3 0.96	313 100.00
direct	96 98.97	1 1.03	97 100.00
Total	406 99.02	4 0.98	410 100.00

```
. drop missing
. gen byte missing = (partial>=.) | (severe>=.)
. tab version missing if touse, missing row
```

Key
frequency row percentage

experimental condition	missing		Total
	0	1	
crosswise	309 98.72	4 1.28	313 100.00
direct	96 98.97	1 1.03	97 100.00
Total	405 98.78	5 1.22	410 100.00

```
. drop missing
. count if (partial<.) & (severe>=.) & touse
1
```

```
. count if (partial>=.) & (severe<.) & touse
1
```

4 Table 1: Descriptives

```
. su f2 if touse
```

Variable	Obs	Mean	Std. Dev.	Min	Max
f2	407	23.24324	3.799442	18	60

```
. egen byte agecat = cut(f2) if touse, at(18,23,28,`r(max)`)
(68 missing values generated)
```

```
. tab agecat version if touse, chi2 exact col
```

Key
<i>frequency</i>
<i>column percentage</i>

Enumerating sample-space combinations:

stage 3: enumerations = 1

stage 2: enumerations = 6

stage 1: enumerations = 0

agecat	experimental condition		Total
	crosswise	direct	
18	154 49.52	47 49.47	201 49.51
23	126 40.51	42 44.21	168 41.38
28	31 9.97	6 6.32	37 9.11
Total	311 100.00	95 100.00	406 100.00

Pearson chi2(2) = 1.3053 Pr = 0.521

Fisher's exact = 0.557

```
. tab f1 version if touse, chi2 exact col
```

Key
<i>frequency</i>
<i>column percentage</i>

sex	experimental condition		Total
	crosswise	direct	
male	154 49.68	42 43.30	196 48.16
female	156	55	211

	50.32	56.70	51.84
Total	310	97	407
	100.00	100.00	100.00

Pearson chi2(1) = 1.2040 Pr = 0.273
 Fisher's exact = 0.296
 1-sided Fisher's exact = 0.163
 . tab f3 version if touse, chi2 exact col

Key
<i>frequency</i>
<i>column percentage</i>

nationality	experimental condition		Total
	crosswise	direct	
german or swiss	288 92.90	93 95.88	381 93.61
other	22 7.10	4 4.12	26 6.39
Total	310 100.00	97 100.00	407 100.00

Pearson chi2(1) = 1.0920 Pr = 0.296
 Fisher's exact = 0.351
 1-sided Fisher's exact = 0.214

5 Table 2: Prevalence estimates

```
. gen byte crosswise = version==1 if touse
(64 missing values generated)
. gen pyes = cond(crosswise, 0.25, 1) if touse
(64 missing values generated)
```

```
.
. // - direct questioning (DQ)
.
. reg partial if crosswise==0
```

Source	SS	df	MS	Number of obs	=	96
Model	0	0	.	F(0, 95)	=	0.00
Residual	6.48958333	95	.068311404	Prob > F	=	.
Total	6.48958333	95	.068311404	R-squared	=	0.0000
				Adj R-squared	=	0.0000
				Root MSE	=	.26136

partial	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
_cons	.0729167	.0266754	2.73	0.007	.0199593 .125874

```
. reg severe if crosswise==0
```

Source	SS	df	MS	Number of obs	=	96
				F(0, 95)	=	0.00

Model	0	0	.	Prob > F	=	.
Residual	.989583333	95	.010416667	R-squared	=	0.0000
Total	.989583333	95	.010416667	Adj R-squared	=	0.0000
				Root MSE	=	.10206

```

. // - crosswise model (CM)
.
. rrreg partial if crosswise, pw(pyese)

```

```

Randomized response regression          Number of obs   =      310
                                         F(  0,    309)  =      0.00
                                         Prob > F        =      .
                                         R-squared      =     0.0000
                                         Adj R-squared  =     0.0000
                                         Root MSE      =     0.9623

```

partial	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
_cons	.2225806	.0546551	4.07	0.000	.1150374 .3301239

```

Pr(non-negated question) = pyese
Pr(surrogate "yes")      = 0
Pr(surrogate "no")      = 0
. rrreg severe if crosswise, pw(pyese)

```

```

Randomized response regression          Number of obs   =      310
                                         F(  0,    309)  =      0.00
                                         Prob > F        =      .
                                         R-squared      =     0.0000
                                         Adj R-squared  =     0.0000
                                         Root MSE      =     0.8766

```

severe	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
_cons	.016129	.0497849	0.32	0.746	-.0818313 .1140894

```

Pr(non-negated question) = pyese
Pr(surrogate "yes")      = 0
Pr(surrogate "no")      = 0

```

```

. // - difference between DQ and CM
.
. rrreg partial crosswise, pw(pyese) robust hc2

```

```

Randomized response regression          Number of obs   =      406
                                         F(  1,    404)  =      6.05
                                         Prob > F        =     0.0143
                                         R-squared      =     0.0056
                                         Adj R-squared  =     0.0031
                                         Root MSE      =     0.8511

```

partial	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
crosswise	.149664	.0608271	2.46	0.014	.0300868 .2692411

_cons	.0729167	.0266017	2.74	0.006	.0206216	.1252117
-------	----------	----------	------	-------	----------	----------

```
Pr(non-negated question) = pyes
Pr(surrogate "yes")      = 0
Pr(surrogate "no")      = 0
```

```
. rrreg severe crosswise, pw(pyese) robust hc2
```

```
Randomized response regression      Number of obs   =      406
                                     F( 1, 404)      =      0.01
                                     Prob > F         =      0.9107
                                     R-squared        =      0.0000
                                     Adj R-squared     =     -0.0025
                                     Root MSE        =      0.7682
```

severe	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
crosswise	.0057124	.0508988	0.11	0.911	-.0943471 .1057719
_cons	.0104167	.0103879	1.00	0.317	-.0100044 .0308377

```
Pr(non-negated question) = pyes
Pr(surrogate "yes")      = 0
Pr(surrogate "no")      = 0
```

6 Table 3: Regression estimates

```
. gen byte female      = f1==2 if inlist(f1,1,2)
(3 missing values generated)
. gen byte internet    = f8_4==1
. gen byte students    = f8_5==1
. gen byte papers3or4  = f7==3 if f7<.
(1 missing value generated)
. gen byte papers5     = f7==4 if f7<.
(1 missing value generated)
. gen byte zurich      = uni==1
. gen byte munich      = uni==2
.
. su partial crosswise zurich munich female ///
>   papers3or4 papers5 internet students if partial<.
Variable | Obs      Mean      Std. Dev.      Min      Max
-----+-----+-----+-----+-----+-----
partial  | 406      .5049261      .5005926      0         1
crosswise| 406      .7635468      .4254279      0         1
zurich   | 406      .1724138      .3782058      0         1
munich   | 406      .2093596      .4073535      0         1
female   | 403      .5186104      .5002746      0         1
-----+-----+-----+-----+-----+-----
papers3or4| 405      .2691358      .4440592      0         1
papers5   | 405      .2888889      .4538068      0         1
internet  | 406      .8940887      .3081038      0         1
students  | 406      .2413793      .4284478      0         1
. su partial crosswise zurich munich female ///
>   papers3or4 papers5 internet students if partial<. & crosswise==1
```


Variable	Obs	Mean	Std. Dev.	Min	Max
partial	310	.6387097	.4811511	0	1
crosswise	310	1	0	1	1
zurich	310	.1741935	.3798891	0	1
munich	310	.2096774	.4077365	0	1
female	307	.504886	.5007924	0	1
papers3or4	309	.2944984	.4565563	0	1
papers5	309	.2912621	.455081	0	1
internet	310	.883871	.3208976	0	1
students	310	.2290323	.4208894	0	1

```
.
. eststo reg1: rrreg partial ///
>   zurich munich ///
>   female papers3or4 papers5 internet students ///
>   if crosswise, pw(pyes) robust
```

```
Randomized response regression          Number of obs   =       306
                                         F( 7, 298)       =       1.23
                                         Prob > F         =       0.2841
                                         R-squared       =       0.0258
                                         Adj R-squared   =       0.0029
                                         Root MSE       =       0.9599
```

partial	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
zurich	.1297747	.1679073	0.77	0.440	-.2006596	.4602089
munich	-.2022313	.1352413	-1.50	0.136	-.4683803	.0639178
female	.0313942	.1168534	0.27	0.788	-.1985681	.2613565
papers3or4	-.1129604	.1359154	-0.83	0.407	-.380436	.1545152
papers5	.0879665	.1432817	0.61	0.540	-.1940055	.3699386
internet	.1726924	.1687123	1.02	0.307	-.1593261	.5047109
students	.1897438	.1378786	1.38	0.170	-.0815952	.4610829
_cons	.0344005	.1662536	0.21	0.836	-.2927793	.3615803

```
Pr(non-negated question) = pyes
Pr(surrogate "yes")      = 0
Pr(surrogate "no")      = 0
```

```
. test zurich = munich
( 1)  zurich - munich = 0
      F( 1, 298) = 2.97
      Prob > F = 0.0860
```

```
.
. eststo logit1: rrlogit partial ///
>   zurich munich ///
>   female papers3or4 papers5 internet students ///
>   if crosswise, pw(pyes) robust
```

Fitting constant-only model:

```
Iteration 0: log pseudolikelihood = -212.10304
Iteration 1: log pseudolikelihood = -199.85397
Iteration 2: log pseudolikelihood = -199.85373
Iteration 3: log pseudolikelihood = -199.85373
```

Fitting full model:

```
Iteration 0: log pseudolikelihood = -199.85373
Iteration 1: log pseudolikelihood = -198.77043
Iteration 2: log pseudolikelihood = -197.73043
```

```
Iteration 3: log pseudolikelihood = -196.81341
Iteration 4: log pseudolikelihood = -196.68131
Iteration 5: log pseudolikelihood = -196.68051
Iteration 6: log pseudolikelihood = -196.68051
```

```
Randomized response logistic regression      Number of obs   =      306
                                             Nonzero outcomes =      196
P(non-negated question) = pyes             Zero outcomes   =      110
P(surrogate "yes")      = 0                 Wald chi2(7)    =      6.36
P(surrogate "no")       = 0                 Prob > chi2     =     0.4987
Log pseudolikelihood = -196.68051          Pseudo R2      =     0.0159
```

partial	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
zurich	.5216706	.9830921	0.53	0.596	-1.405155	2.448496
munich	-1.001754	.9174338	-1.09	0.275	-2.799891	.7963832
female	.2318339	.8098296	0.29	0.775	-1.355403	1.819071
papers3or4	-.6753206	.8620195	-0.78	0.433	-2.364848	1.014207
papers5	.2226108	.8264936	0.27	0.788	-1.397287	1.842509
internet	1.011847	1.588698	0.64	0.524	-2.101945	4.125639
students	.9554459	.7934316	1.20	0.229	-.5996515	2.510543
_cons	-2.303029	1.297016	-1.78	0.076	-4.845134	.2390766

```
. test zurich = munich
( 1) [partial]zurich - [partial]munich = 0
      chi2( 1) = 1.64
      Prob > chi2 = 0.2001
```

```
. eststo reg2: rrreg partial crosswise ///
> zurich munich ///
> female papers3or4 papers5 internet students ///
> , pw(pyes) robust
```

```
Randomized response regression      Number of obs   =      402
                                             F( 8, 393)     =      1.70
                                             Prob > F       =     0.0976
                                             R-squared     =     0.0268
                                             Adj R-squared =     0.0070
                                             Root MSE     =     0.8472
```

partial	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
crosswise	.1722132	.0640163	2.69	0.007	.046356	.2980703
zurich	.122656	.1280605	0.96	0.339	-.1291133	.3744253
munich	-.1326625	.1057303	-1.25	0.210	-.3405302	.0752052
female	.0409437	.0889149	0.46	0.645	-.1338647	.2157522
papers3or4	-.097614	.1074576	-0.91	0.364	-.3088777	.1136497
papers5	.0471034	.109144	0.43	0.666	-.1674756	.2616825
internet	.1475172	.1376815	1.07	0.285	-.1231672	.4182016
students	.1606791	.1024078	1.57	0.117	-.0406565	.3620148
_cons	-.1198158	.1376898	-0.87	0.385	-.3905164	.1508849

```
Pr(non-negated question) = pyes
Pr(surrogate "yes")      = 0
Pr(surrogate "no")       = 0
```

```
. test zurich = munich
( 1) zurich - munich = 0
```


	(0.21)	(-0.87)	(-1.78)	(-2.73)
N	306	402	306	402

t statistics in parentheses
+ p<0.1, * p<0.05, ** p<0.01, *** p<0.001

7 Graph: Regression estimates

```
. coefplot reg2 reg1 || logit2 logit1, drop(_cons) xline(0) ///
> bylabels(LPM Logit) byopts(xrescale legend(off))
```

