Tables en R con Stargazer

14 de octubre de 2021

INSTITUTO TECNOLOGICO AUTONOMO DE MEXICO (ITAM)

Seminario de Investigación Económica

Instructor: Horacio Larreguy

Asistente: Manuel Quintero

1. Introducción

Los siguientes enlaces son a páginas que incluyen distintos ejemplos de formatos de tablas en Stargazer y que argumentos incluir para distintas modificaciones:

1. https://www.jakeruss.com/cheatsheets/stargazer/.

2. https://cran.r-project.org/web/packages/stargazer/vignettes/stargazer. pdf.

2. Tabla formato presentable

```r
aux <- c("y1","y2","y3","y4","y5")
resultados_data <- datos[aux]

dep_var <- c("No education",
             "\\shortstack{Primary \\\\ education}",
             "\\shortstack{Technical or \\\\ Secondary \\\\ education}",
             "\\shortstack{College \\\\ education \\\\ or equivalent}",
             "\\shortstack{Higher \\\\ education}")

tabla1 <- stargazer(reg1, reg2, reg3, reg4, reg5,
                    header=FALSE,
                    font.size="scriptsize",
                    dep.var.caption = "",
                    dep.var.labels.include = FALSE,
                    table.placement = "H",
                    omit = c("Constant","lagged_var"),
                    column.labels = dep_var,
                    covariate.labels = c("Mother education", "Father education", "
                        Total population", "Social gap index"),
                    omit.stat=c("f", "ser","adj.rsq"),
                    add.lines = list(c("Outcome mean", round(res_mean, 3)),
                                     c("Outcome std. dev.", round(res_sd, 3))),
                    title = "Analysis of the different educational levels",
                    type = "latex")


note.latex <- "\\multicolumn{6}{l} {\\parbox[t]{13cm}{ \\textit{Notes:}
```

```
We report estimates from OLS regressions, including municipal and time fixed effects.
All specifications further include the lagged dependent variable as a control.
Standard errors clustered by municipality are in parentheses.
* denote p$<$0.1, ** denote p$<$0.05, y *** denote p$<$0.01.}} \\\\"
tabla1[grepl("Note", tabla1)] <- note.latex


cat(tabla1, file = 'tabla1.tex')
```

Tabla 1: Analysis of the different educational levels

	No education	Primary education	Technical or Secondary education	College education or equivalent	Higher education
	(1)	(2)	(3)	(4)	(5)
Mother education	1.082**	0.948***	1.658***	1.000***	1.663***
	(0.393)	(0.043)	(0.263)	(0.000)	(0.255)
Father education	5.939***	0.508***	2.008***	1.000***	2.308***
	(0.143)	(0.007)	(0.092)	(0.000)	(0.091)
Total population		0.008	0.152*		0.146*
		(0.010)	(0.071)		(0.072)
Social gap index				−0.000	0.133
				(0.000)	(0.132)
Outcome mean	12.969	3.58	5.982	5.358	9.708
Outcome std. dev.	4.136	1.531	2.029	0.989	2.357
Observations	100	100	100	100	100
R^2	0.994	1.000	0.991	1.000	0.994

Notes: We report estimates from OLS regressions, including municipal and time fixed effects. All specifications further include the lagged dependent variable as a control. Standard errors clustered by municipality are in parentheses. * denote p<0.1, ** denote p<0.05, y *** denote p<0.01.

3.   Nombre único para todas las columnas

```
tabla2 <- stargazer(reg1, reg2, reg3, reg4, reg5,
                    header=FALSE,
                    font.size="scriptsize",
                    dep.var.caption = "Variables de educación",
                    dep.var.labels.include = FALSE,
                    table.placement = "H",
                    omit = c("Constant","lagged_var"),
                    column.labels = "",
                    covariate.labels = c("Mother education", "Father education", "
                        Total population", "Social gap index"),
                    omit.stat=c("f", "ser","adj.rsq"),
                    add.lines = list(c("Outcome mean", round(res_mean, 3)),
                                     c("Outcome std. dev.", round(res_sd, 3))),
                    title = "Analysis of the different educational levels",
                    type = "latex")


note.latex <- "\\multicolumn{6}{l} {\\parbox[t]{11cm}{ \\textit{Notes:}
We report estimates from OLS regressions, including municipal and time fixed effects.
All specifications further include the lagged dependent variable as a control.
Standard errors clustered by municipality are in parentheses.
* denote p$<$0.1, ** denote p$<$0.05, y *** denote p$<$0.01.}} \\\\"
tabla2[grepl("Note", tabla2)] <- note.latex

cat(tabla2, file = 'tabla2.tex')
```

Tabla 2: Analysis of the different educational levels

	Variables de educación				
	(1)	(2)	(3)	(4)	(5)
Mother education	1.082**	0.948***	1.658***	1.000***	1.663***
	(0.393)	(0.043)	(0.263)	(0.000)	(0.255)
Father education	5.939***	0.508***	2.008***	1.000***	2.308***
	(0.143)	(0.007)	(0.092)	(0.000)	(0.091)
Total population		0.008	0.152*		0.146*
		(0.010)	(0.071)		(0.072)
Social gap index				−0.000	0.133
				(0.000)	(0.132)
Outcome mean	12.969	3.58	5.982	5.358	9.708
Outcome std. dev.	4.136	1.531	2.029	0.989	2.357
Observations	100	100	100	100	100
R^2	0.994	1.000	0.991	1.000	0.994

Notes: We report estimates from OLS regressions, including municipal and time fixed effects. All specifications further include the lagged dependent variable as a control. Standard errors clustered by municipality are in parentheses. * denote $p<0.1$, ** denote $p<0.05$, y *** denote $p<0.01$.

4.   Ajuste de fuente y ajuste horizontal

Note que el ajuste horizontal lo hicimos manualmente, directo en Overleaf.

```
dep_var <- c("\\shortstack{No \\\\ education  \\\\ Mex}",
             "\\shortstack{Primary \\\\ education \\\\  Mex}",
             "\\shortstack{Technical or \\\\ Secondary \\\\ education \\\\  Mex}",
             "\\shortstack{College \\\\ education \\\\ or equivalent \\\\  Mex}",
             "\\shortstack{Higher \\\\ education \\\\  Mex}",
             "\\shortstack{No  \\\\  education  \\\\ USA}",
             "\\shortstack{Primary \\\\ education  \\\\ USA}",
             "\\shortstack{Technical or \\\\ Secondary \\\\ education  \\\\ USA}",
             "\\shortstack{College \\\\ education \\\\ or equivalent  \\\\ USA}",
             "\\shortstack{Higher \\\\ education  \\\\ USA}")

tabla3 <- stargazer(reg1, reg2, reg3, reg4, reg5, reg6, reg7, reg8, reg9, reg10,
                    header=FALSE,
                    font.size="tiny", # so that it fits horizontally
                    dep.var.caption = "",
                    dep.var.labels.include = FALSE,
                    table.placement = "H",
                    omit = c("Constant","lagged_var"),
                    column.labels = dep_var,
                    covariate.labels = c("Mother education", "Father education", "
                        Total population", "Social gap index"),
                    omit.stat=c("f", "ser","adj.rsq"),
                    add.lines = list(c("Outcome mean", round(res_mean, 3)),
                                     c("Outcome std. dev.", round(res_sd, 3))),
                    title = "Analysis of the different educational levels",
                    type = "latex")


note.latex <- "\\multicolumn{11}{l} {\\parbox[t]{20cm}{ \\textit{Notes:}
We report estimates from OLS regressions, including municipal and time fixed effects.
All specifications further include the lagged dependent variable as a control.
Standard errors clustered by municipality are in parentheses.
Columns 6 to 10 include the US unemployment index as a control.
* denote p$<$0.1, ** denote p$<$0.05, y *** denote p$<$0.01.}} \\\\"
tabla3[grepl("Note", tabla3)] <- note.latex

cat(tabla3, file = 'tabla3.tex')
```

Tabla 3: Analysis of the different educational levels

	No education Mex	Primary education Mex	Technical or Secondary education Mex	College education or equivalent Mex	Higher education Mex	No education USA	Primary education USA	Technical or Secondary education USA	College education or equivalent USA	Higher education USA
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Mother education	1.082**	0.948***	1.658***	1.000***	1.663***	1.082**	0.500***	1.327	1.019**	1.635***
	(0.393)	(0.043)	(0.263)	(0.000)	(0.255)	(0.393)	(0.000)	(0.966)	(0.425)	(0.263)
Father education	5.939***	0.508***	2.008***	1.000***	2.308***	5.939***	0.500***	4.250***	5.943***	3.015***
	(0.143)	(0.007)	(0.092)	(0.000)	(0.091)	(0.143)	(0.000)	(0.355)	(0.151)	(0.095)
Total population		0.008	0.152*		0.146*		0.000**	0.603*		0.153*
		(0.010)	(0.071)		(0.072)		(0.000)	(0.272)		(0.073)
Social gap index				−0.000	0.133				−0.189	0.130
				(0.000)	(0.132)				(0.218)	(0.136)
Outcome mean	12.969	3.58	5.982	5.358	9.708	12.845	3.034	13.088	12.557	11.033
Outcome std. dev.	4.136	1.531	2.029	0.989	2.357	3.95	1.025	6.316	4.444	2.682
Observations	100	100	100	100	100	100	100	100	100	100
R^2	0.994	1.000	0.991	1.000	0.994	0.993	1.000	0.985	0.994	0.995

Notes: We report estimates from OLS regressions, including municipal and time fixed effects. All specifications further include the lagged dependent variable as a control. Standard errors clustered by municipality are in parentheses. Columns 6 to 10 include the US unemployment index as a control. * denote $p<0.1$, ** denote $p<0.05$, y *** denote $p<0.01$.

5. Sidewaystable

```
tabla4 <- stargazer(reg1, reg2, reg3, reg4, reg5, reg6, reg7, reg8, reg9, reg10,
                    header=FALSE,
                    font.size="tiny", # so that it fits horizontally
                    dep.var.caption = "",
                    dep.var.labels.include = FALSE,
                    table.placement = "t", # change to top of the page
                    float.env = "sidewaystable", # environment to sidewysetable (add
                        rotating package to preamble in overleaf)
                    omit = c("Constant","lagged_var"),
                    column.labels = dep_var,
                    covariate.labels = c("Mother education", "Father education", "
                        Total population", "Social gap index"),
                    omit.stat=c("f", "ser","adj.rsq"),
                    add.lines = list(c("Outcome mean", round(res_mean, 3)),
                                    c("Outcome std. dev.", round(res_sd, 3))),
                    title = "Analysis of the different educational levels",
                    type = "latex")


note.latex <- "\\multicolumn{11}{l} {\\parbox[t]{23cm}{ \\textit{Notes:}
We report estimates from OLS regressions, including municipal and time fixed effects.
All specifications further include the lagged dependent variable as a control.
Standard errors clustered by municipality are in parentheses.
Columns 6 to 10 include the US unemployment index as a control.
* denote p$<$0.1, ** denote p$<$0.05, y *** denote p$<$0.01.}} \\\\"
tabla4[grepl("Note", tabla4)] <- note.latex

cat(tabla4, file = 'tabla4.tex')
```

Tabla 4: Analysis of the different educational levels

	No education	Primary education	Technical or Secondary education	College education or equivalent	Higher education	No education USA	Primary education USA	Technical or Secondary education USA	College education or equivalent USA	Higher education USA
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Mother education	1.082**	0.948***	1.658***	1.000***	1.663***	1.082**	0.500***	1.327	1.019**	1.635***
	(0.393)	(0.043)	(0.263)	(0.000)	(0.255)	(0.393)	(0.000)	(0.966)	(0.425)	(0.263)
Father education	5.939***	0.508***	2.008***	1.000***	2.308***	5.939***	0.500***	4.250***	5.943***	3.015***
	(0.143)	(0.007)	(0.092)	(0.000)	(0.091)	(0.143)	(0.000)	(0.355)	(0.151)	(0.095)
Total population		0.008	0.152*		0.146*		0.000**	0.603*		0.153*
		(0.010)	(0.071)		(0.072)		(0.000)	(0.272)		(0.073)
Social gap index				−0.000	0.133				−0.189	0.130
				(0.000)	(0.132)				(0.218)	(0.136)
Outcome mean	12.969	3.58	5.982	5.358	9.708	12.845	3.034	13.088	12.557	11.033
Outcome std. dev.	4.136	1.531	2.029	0.989	2.357	3.95	1.025	6.316	4.444	2.682
Observations	100	100	100	100	100	100	100	100	100	100
R^2	0.994	1.000	0.991	1.000	0.994	0.993	1.000	0.985	0.994	0.995

Notes: We report estimates from OLS regressions, including municipal and time fixed effects. All specifications further include the lagged dependent variable as a control. Standard errors clustered by municipality are in parentheses. Columns 6 to 10 include the US unemployment index as a control. * denote $p<0.1$, ** denote $p<0.05$, y *** denote $p<0.01$.