Strategies and tricks for teaching and researching with Stata

2017 Spanish Stata Users Group meeting

Andrés Santana

José Rama

Universidad Autónoma de Madrid

ADO to open DBs

- Stata12
  - **ssc install "use12", all**
  - **use12 "Data1.dta"**
- Stata13
  - **ssc install "use13", all**
  - **use13 "Data1.dta"**
- SPSS
  - **ssc install "usespss", all**
  - **usespss "Data1.sav"**

+ on other formats

- Other ways to find Radyakin's ADOs
  - https://ideas.repec.org/c/boc/bocode/s456957.html (usespss)
  - http://radyakin.org/transfer/use13/use13.htm (use13)
  - http://radyakin.org/transfer/use14/ (use14, not written yet)

- Save in SPSS
  - **ssc install "savespss" // findit savespss**
  - http://www.radyakin.org/transfer/savespss/savespss.htm

Codes & Labels?

- Traditional
  - **<u>ta</u>bulate P7r** // labels
  - **<u>ta</u>bulate P7r, nolabel** // codes
- How to see both?
  - **numlabel, <u>a</u>dd**
  - **<u>ta</u>bulate P7r**
- Back to default
  - **numabel, <u>r</u>emove**

```
Situación |
polít(1-5) |      Freq.      Percent        Cum.
-----------+-----------------------------------
1. Muy buena |        7         0.11         0.11
   2. Buena |        240         3.89         4.00
  3. Regular |      2,103        34.08        38.08
    4. Mala |      2,395        38.81        76.89
5. Muy mala |      1,426        23.11       100.00
-----------+-----------------------------------
     Total |      6,171       100.00
```

+ on numlabel

- **browse**

	fampol2	orgterri	orguesp
1	2. Socialista	2. menos autonomía	3. Bastante
2	4. Feminista	3. status quo	3. Bastante
3	.b. [NC]	3. status quo	3. Bastante
4	4. Feminista	3. status quo	1. Nada
5	6. Socialdemócrata	3. status quo	4. Mucho
6	.b. [NC]	4. más autonomía	3. Bastante
7	8. Liberal	5. posibilidad secesión	2. Poco
8	3. Ecologista	4. más autonomía	2. Poco
9	1. Comunista	4. más autonomía	1. Nada
10	2. Socialista	4. más autonomía	1. Nada
11	.b. [NC]	4. más autonomía	1. Nada
12	8. Liberal	3. status quo	2. Poco
13	.a. [NS]	3. status quo	3. Bastante
14	6. Socialdemócrata	4. más autonomía	2. Poco
15	3. Ecologista	4. más autonomía	1. Nada

Graphing Qualies (Ws)

- Traditional (cases)
  - **tabulate P7r**
  - **generate uniweigh = 1/r(N)**
  - **graph bar (count) uniweigh, over(P7r)**
- Traditional (probabilities)
  - tabulate P7r
  - generate uniweigh = 1/r(N)
  - **graph bar (sum) uniweigh, over(P7r)**
- Traditional (percentages)
  - tabulate P7r
  - **generate perweigh = 100/r(N)**
  - **graph bar (sum) perweigh, over(P7r)**

Just install 'catplot'…

- **ssc install "catplot"**
- https://ideas.repec.org/c/boc/bocode/s431501.html

… and enjoy!

- **catplot P7r** // cases
- **catplot P7r, fraction** // probabilities
- **catplot P7r, percent** // percentages

+ on catplot

- catplot P7r
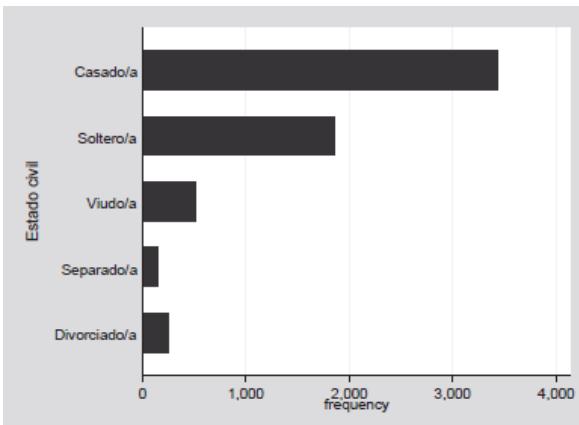
- catplot P7r, recast(bar)
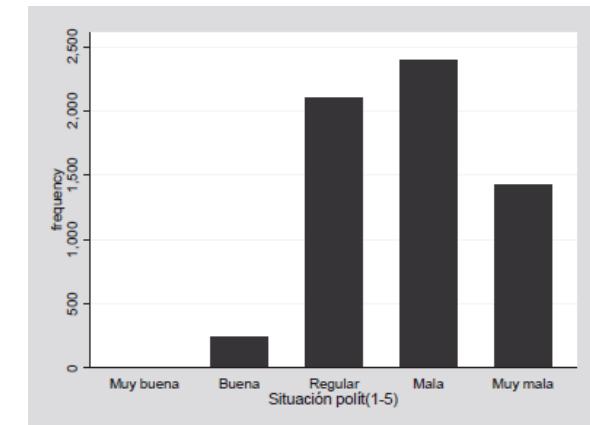

- catplot P7r, recast(dot)
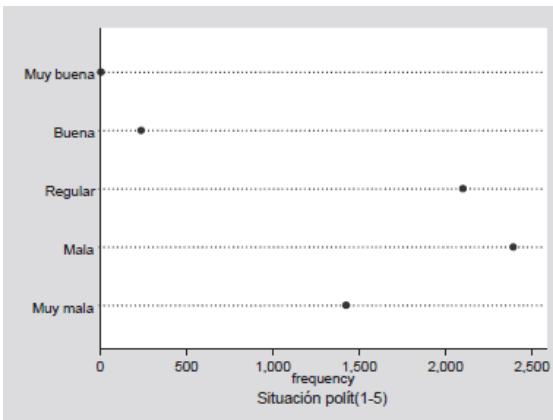
Just forget about
- **graph hbar**
- **graph bar**
- **graph dot**

… & for several Ws

– **catplot voto2015d sexo, percent(sexo) blabel(bar)**
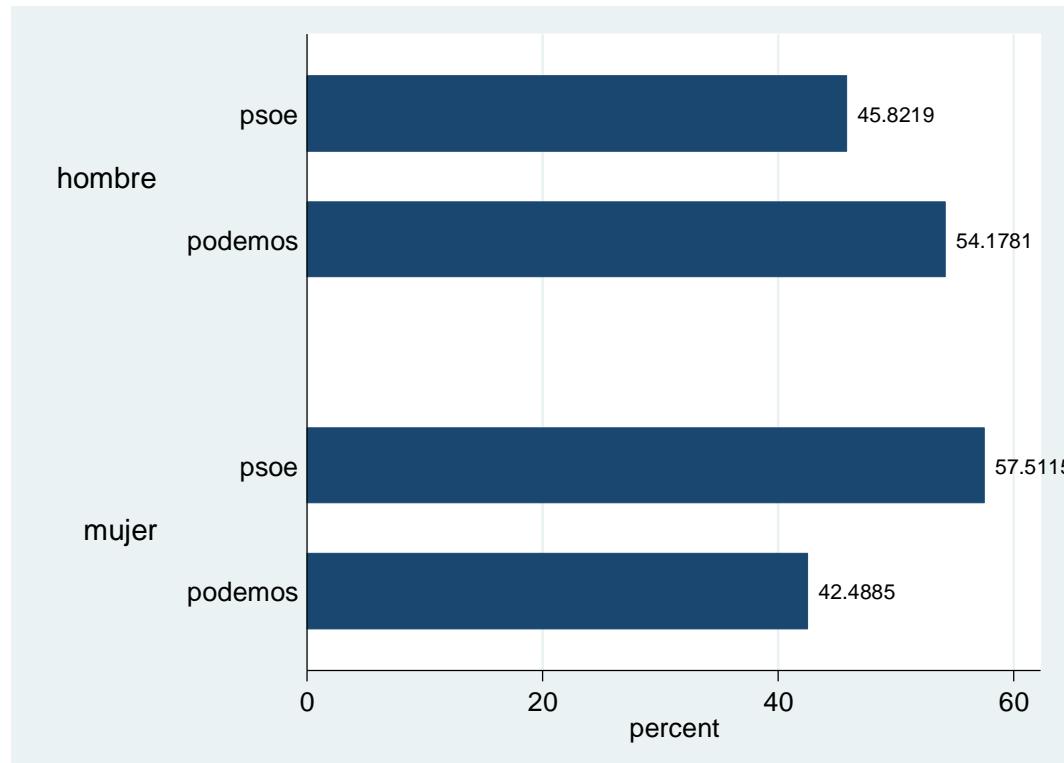
V vs W
compare means

- **tabulate** particpd, summarize(ideo)

```
              |     Summary of Ideología (1-10)
     abstenc  |        Mean    Std. Dev.         Freq.
--------------+-----------------------------------------
        Voto  |   4.6196494    2.0075426          4906
   Abstenció  |   4.7852761    1.6598279           489
--------------+-----------------------------------------
       Total  |   4.6346617    1.9789829          5395
```

Equality of means easier output

- **pwmean** particip, summarize(ideo) cimeans


```
Pairwise comparisons of means with equal variances
over          : participd
---------------------------------------------------------------------
        ideo |        Mean    Std. Err.     [95% Conf. Interval]
-------------+-------------------------------------------------------
    participd |
        Voto |    4.619649    .0282484     4.564271      4.675028
  Abstención |    4.785276    .0894752     4.609869      4.960684
---------------------------------------------------------------------
```

+ on pwmean
(for several categories)

– **pwmean ideo, over(voto2015) cimeans sort cformat(%6.2f)**

```
Pairwise comparisons of means with equal variances
------------------------------------------------------------
     ideo |    Mean   Std. Err.   [95% Conf. Interval]
----------+-------------------------------------------------
voto2015  |
       iu |    2.87      0.10       2.68       3.06
  podemos |    3.09      0.04       3.00       3.17
    otros |    3.70      0.06       3.57       3.82
     psoe |    3.71      0.04       3.63       3.80
   absten |    4.79      0.06       4.66       4.91
 protesta |    4.85      0.15       4.56       5.14
       cs |    5.50      0.06       5.39       5.61
       pp |    6.97      0.04       6.89       7.05
------------------------------------------------------------
```

& more…
(differences)

– **pwmean** ido, over(ecivil) sort cformat(%6.2f)

```
Pairwise comparisons of means with equal variances
----------------------------------------------------------------------------
               ideo |  Contrast   Std. Err.   [95% Conf. Interval]
--------------------+-------------------------------------------------------
             ecivil |
 Separado/a vs Viudo/a |    -0.92        0.19        -1.30       -0.54
Divorciado/a vs Viudo/a |    -0.76        0.16        -1.08       -0.44
Separado/a vs Casado/a |    -0.61        0.17        -0.95       -0.28
 Soltero/a vs Casado/a |    -0.60        0.06        -0.72       -0.48
Divorciado/a vs Casado/a |   -0.46        0.14        -0.72       -0.19
Separado/a vs Soltero/a |    -0.02        0.17        -0.36        0.33
Divorciado/a vs Soltero/a |    0.14        0.14        -0.13        0.42
Divorciado/a vs Separado/a |  0.16        0.21        -0.26        0.57
 Viudo/a vs Casado/a |     0.30        0.10         0.10        0.51
 Viudo/a vs Soltero/a |    0.90        0.11         0.69        1.12
----------------------------------------------------------------------------
```

… & some graphs

- graph bar ideo, over(voto2015) blabel(bar)
- graph hbar edad, over(misa) blabel(bar)
- graph hbar edad, over(ecivild) over(sexo) blabel(bar)
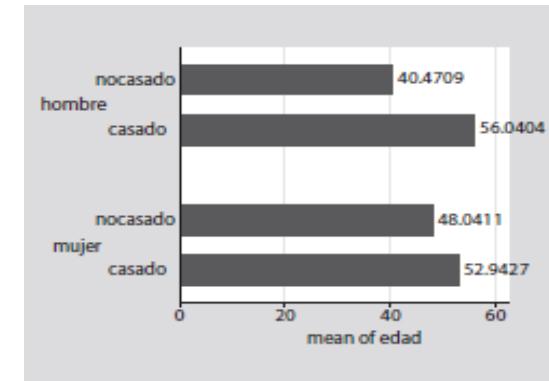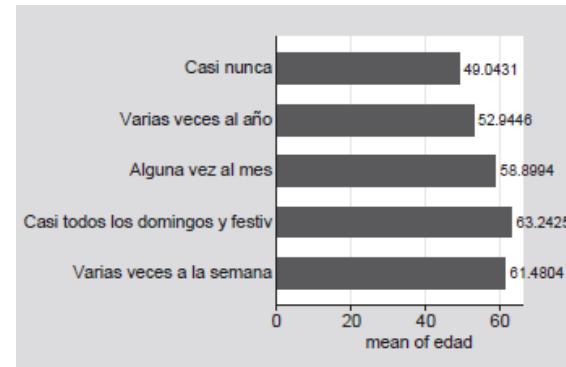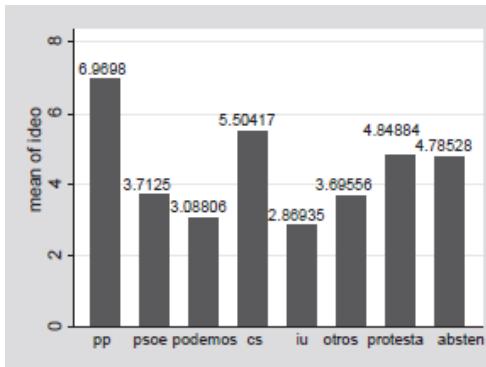
Correlations with pvalues

- **pwcorr** espvi gsal gini exp npar inm**, sig star(0.05)**

	espvida	gsalud	gini	exports	npartidos	inmigran
espvida	1.0000					
gsalud	**0.5041***	1.0000				
	0.0000					
gini	-0.4971*	-0.2969*	1.0000			
	0.0000	0.0007				
exports	**0.1721***	**0.0313**	-0.0550	1.0000		
	0.0306	**0.6918**	0.5472			
npartidos	0.3078*	0.2656*	-0.2437*	-0.0860	1.0000	
	0.0003	0.0017	0.0111	0.3385		
inmigran	0.3189*	0.1806*	-0.2748*	0.4775*	**0.1101**	1.0000
	0.0004	0.0474	0.0057	0.0000	**0.2706**	

+ on pwcorr

— **pwcorr espvi gsal gini exp npa inm, print(0.01)**

```
             |   espvida    gsalud      gini   exports nparti~s inmigran
-------------+------------------------------------------------------------
     espvida |    1.0000
      gsalud |    0.5041    1.0000
        gini |   -0.4971   -0.2969    1.0000
     exports |    0.1721                        1.0000
   npartidos |    0.3078    0.2656   -0.2437              1.0000
    inmigran |    0.3189    0.1806   -0.2748    0.4775             1.0000
```

— **pwcorr gini pib frac_rel, sig casewise**

```
             |      gini       pib  frac_rel
-------------+------------------------------
        gini |    1.0000
         pib |   -0.3724    1.0000
             |    0.0000
    frac_rel |    0.0905    0.0157    1.0000
             |    0.3157    0.8617
```

& more…

– **pwcorr** espvi gsal gini**, obs star(0.01)**

```
            |   espvida    gsalud      gini
------------+---------------------------------
    espvida |   1.0000
            |      174
     gsalud |   0.5041*   1.0000
            |      174        189
       gini |  -0.4971*  -0.2969*   1.0000
            |      126        126       126
```

Scatter plot (combine legend)

- **global** frac12 "frac_eth frac_rel"
- **global** frac13 "frac_eth frac_rel"
- **qui gr tw scatter $frac12** || **lfit $frac12**, name(A1)
- **qui gr tw scatter $frac13** || **lfit $frac13**, name(A2)

- **gr combine A1 A2**
- **findit** grc1leg
- grc1leg A1 A2
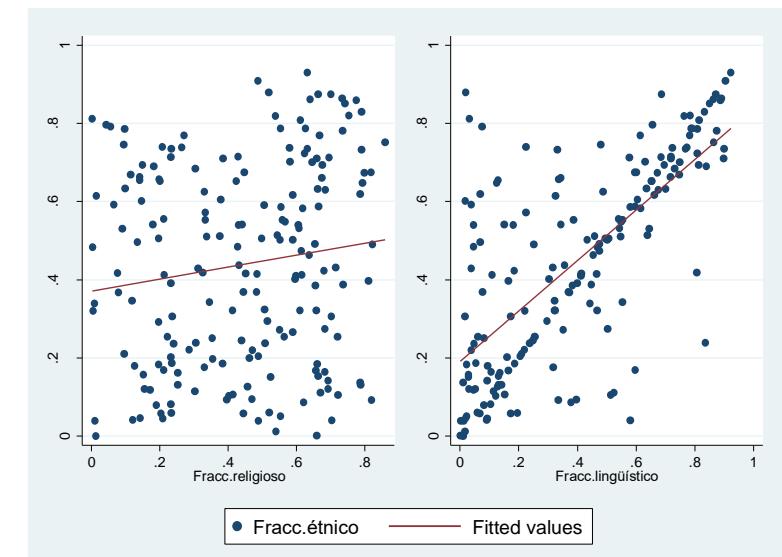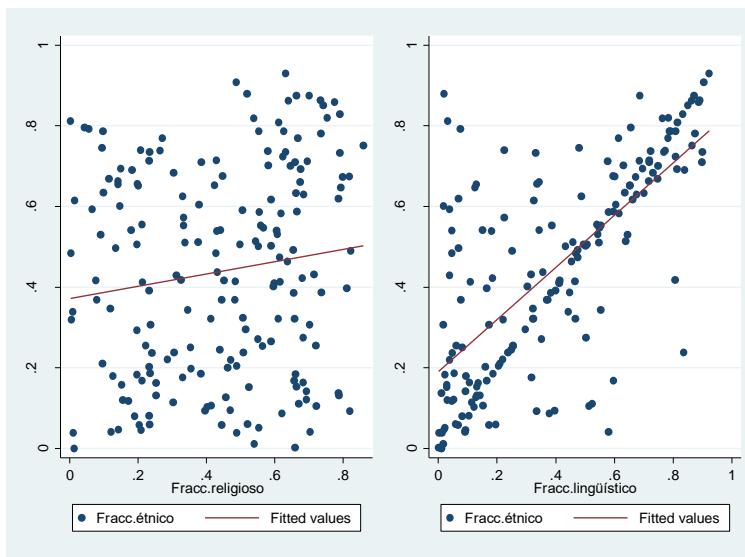
Macros & Models

– global soc "V1 V2 V3"

– global eco "V4 V5"

– global pol "V6 V7 V8"

– global full "$soc $eco $pol" // macro of macros

– ssc install estout

– eststo m1, ti(M1): qui reg Y $soc

– eststo m2, ti(M2): qui reg Y $soc $eco

– eststo m4, ti(M4): qui reg Y $full

Graphing alternative models w coefplot

- eststo Modelo1: qui logit cs $soc
- eststo Modelo2: qui logit cs $eco
- ssc install coefplot
- coefplot Modelo1 || Modelo2, xline(0) drop(_cons)
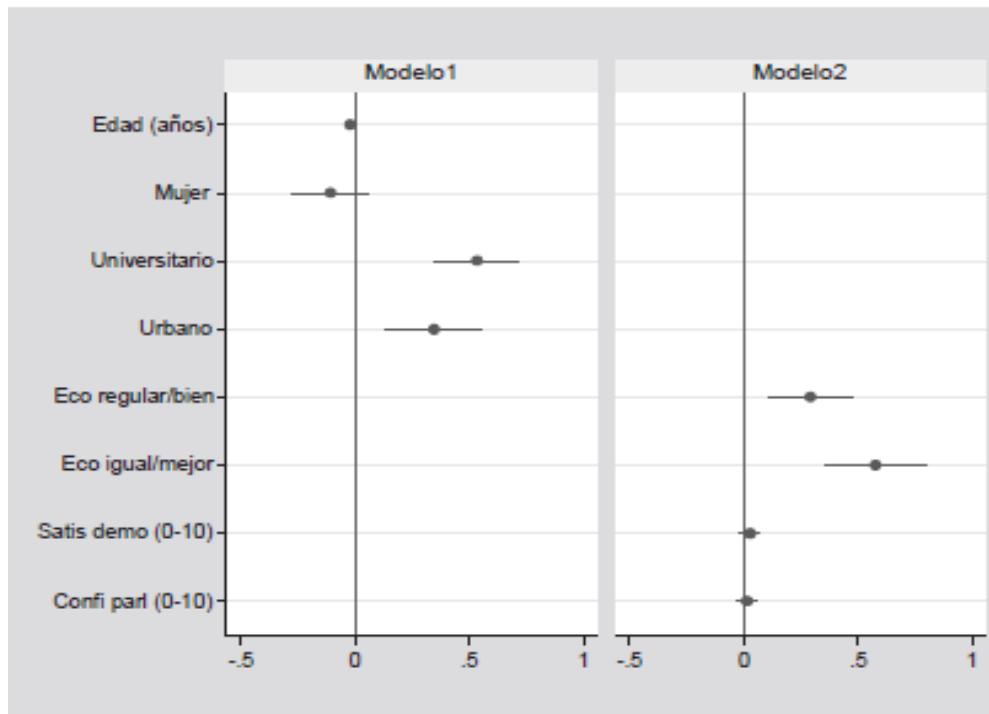
+ on coefplot

— ssc install coefplot

— eststo m1: qui logit cs $soc

— eststo m2: qui logit cs $eco

— coefplot m1 || m2, xline(0) drop(_cons)

— eststo pp: qui logit pp $full

— eststo ps, ti(psoe): qui logit psoe $full

— coefplot pp ps po cs, xli(0) drop(_cons)
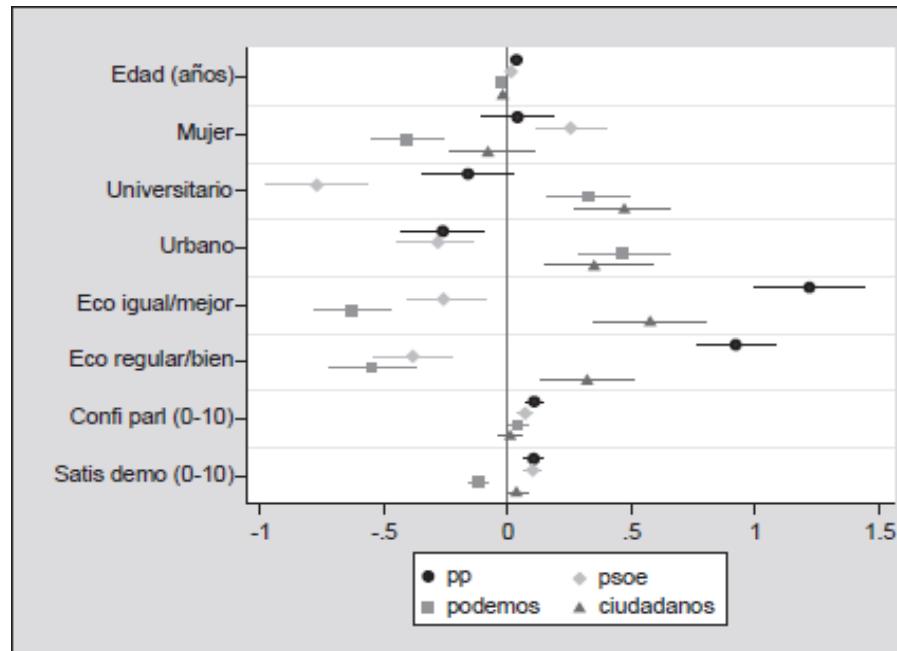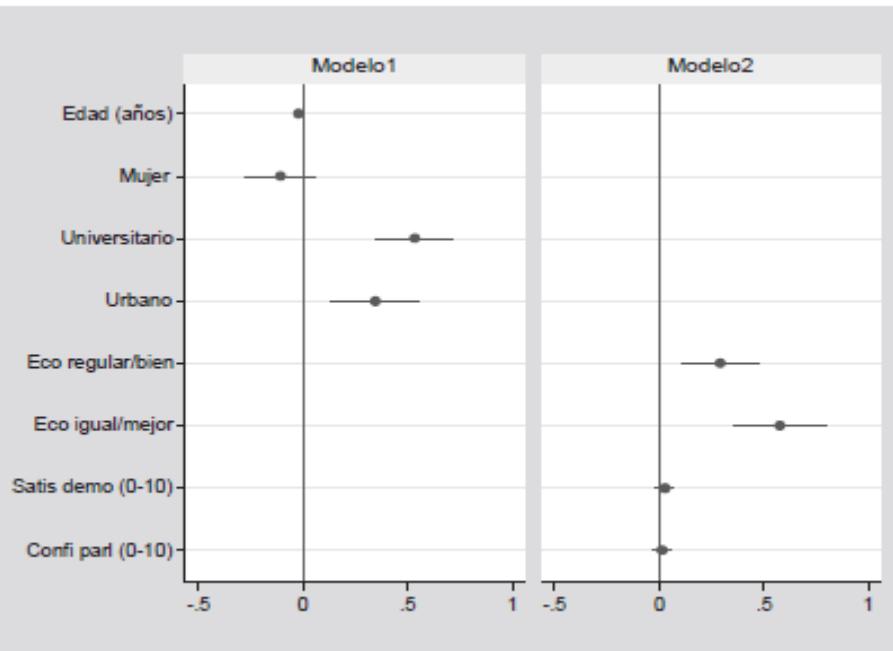
Summary

Other formats
Codes & labels
Graphing qualies with catplot
Comparing means w pwmean
Correlations w pwcorr
Combining scatterplots w grc1leg
Macros & Models
Graphing multiple models w coefplot

Thank U!

We hope that some of these tricks
are useful for You!