Graphs of marginal effect tables from regressions

Modesto Escobar (modesto@usal.es) Cristina Calvo (cristinacalvolopez@usal.es)

University of Salamanca

16th ESA Conference

Porto, 30th August 2024

Ayuda PDC2022-133355-100 financiada por:

The aims of this presentation are:

- To recall the way of presenting several **regressions** at the same time.
- To insist on the meaning and importance of **margins** and **marginal effects** in regression.
- To criticize the wide use of the **base categories** of qualitative independent variables in regression.
- To propose the presentation of margins and marginal effects of several regressions in **parallel**.
- To extend this presentation format to logistic and **multilogistic** models.
- To provide a **visual representation** of several regressions in unison.

Definition: The marginal effects of a generalized regression model refer to the change in the dependent variable that implies a one-unit change in each of the independent variables, keeping all other variables constant.

In other words, they show the degree to which the explanatory variables affect the response (dependent variable).

In the case of being in a multiple linear regression:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$$

Where:

- Y is the dependent or response variable.
- X_1 and X_2 are predictors or independent variables.
- β_1 and β_2 are the regression coefficiente, while β_0 is the intercept or constant.
- ϵ is the error or residual term.

The marginal effect of X_1 on Y is in this case β_1, which means that for every one unit increase in X_1, Y will change by β_1 units, on the assumption that X_2 remains constant.

Marginal effects

Estimation

The following formula is used to estimate marginal effects:

$$\hat{p} = 1/w_{\cdot} \sum_{j=1}^{N} \delta_j(S_p) w_j h(z_j, \hat{\theta})$$

where $\delta_j(S_p)$ are the components of the subpopulation S_p to accomplish the expected prediction, $w_{\cdot}$ es $\sum_{j=1}^{N} \delta_j(S_p) w_j$, and $h(z_j, \hat{\theta})$ is

- $\frac{\partial f(z_j, \hat{\theta})}{\partial x}$ for continuous predictors.
- $f(z_j, \hat{\theta}|A = k) - f(z_j, \hat{\theta}|A = base)$ in the case of a discrete variable A for each k of its K categories or values.

where $f(z_j, \hat{\theta})$ is the function of regression of a general linear model.

Data

Social satisfaction and vote in GB

Data from the ESS-round 11 (2023/24).
(Random sample of 1,684 citizens in Great Britain)

Table: Variables and categories

Variables	Categories
Satisfaction	[0-10] (6 variables: Life, Education, Health, Democracy, Economy, Government)
Vote in the last elections	(Conservative, Labor, Liberal Democrat, Scotish National Party, Others)
Gender	(Male, Female)
Age	(Continuous)
Ideology	(Left (0-4), Center (5), Right (6-10))
Religion	(Religious, Non religious)

Data distribution

Frequencies, percentages, means and standard deviations

Table: Descriptive statistics

	N/Mean	(%)/SD
Life satisfaction	7.0	2.2
Education satisfaction	5.4	2.2
Health satisfaction	4.8	2.6
Democracy satisfaction	4.3	2.6
Economy satisfaction	3.3	2.2
Government satisfaction	3.0	2.4
Party voted		
Conservative	380	(36%)
Labour	416	(40%)
Liberal Democrat	113	(11%)
Scottish National P.	53	(5%)
Others	81	(8%)
Gender		
Male	824	(49%)
Female	860	(51%)
Age	53.6	19.0
Ideology		
Left	532	(38%)
Center	516	(37%)
Right	363	(26%)
Religion		
Religious	709	(42%)
Non religious	964	(58%)

Regression

Government satisfaction on gender, age, ideology and religion

Table: Regression of government satisfaction (0/10)

```
. regress government_s i.gender age i.ideo i.religion
```

Source	SS	df	MS		
Model	1213.67037	5	242.734075	Number of obs	= 1,363
Residual	6925.11759	1,357	5.10325541	F(5, 1357)	= 47.56
				Prob > F	= 0.0000
				R-squared	= 0.1491
				Adj R-squared	= 0.1460
Total	8138.78797	1,362	5.97561525	Root MSE	= 2.259

government_s	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
gender						
Female	.0099489	.1231876	0.08	0.936	-.2317099	.2516076
age	-.001627	.0033898	-0.48	0.631	-.0082769	.0050229
ideo						
Center	1.248514	.1430336	8.73	0.000	.9679235	1.529105
Right	2.09363	.1615159	12.96	0.000	1.776782	2.410478
religion						
Non religious	-.6847867	.126781	-5.40	0.000	-.9334947	-.4360787
_cons	2.474867	.234688	10.55	0.000	2.014476	2.935257

Margins on factor variables

Margins of Government satisfaction on gender, ideology and religion

Table: Margins of government satisfaction

```
. margins gender ideo religion
Predictive margins                                  Number of obs = 1,363
Model VCE: OLS
Expression: Linear prediction, predict()
```

	Margin	Delta-method std. err.	t	P>|t|	[95% conf.	interval]
gender						
Male	2.982732	.0852646	34.98	0.000	2.815467	3.149997
Female	2.992681	.0884059	33.85	0.000	2.819254	3.166108
ideo						
Left	1.990669	.1010192	19.71	0.000	1.792499	2.18884
Center	3.239184	.1013302	31.97	0.000	3.040403	3.437965
Right	4.084299	.1229369	33.22	0.000	3.843132	4.325466
religion						
Religious	3.387447	.096053	35.27	0.000	3.199018	3.575875
Non religious	2.70266	.0807815	33.46	0.000	2.54419	2.861131

Marginal efects

Marginal effects of government satisfaction on gender, ideology and religion

Table: Average marginal effects of government satisfaction

```
. margins, dydx(*)
Average marginal effects                        Number of obs = 1,363
Model VCE: OLS
Expression: Linear prediction, predict()
dy/dx wrt:  2.gender age 2.ideo 3.ideo 2.religion
```

	dy/dx	Delta-method std. err.	t	P>|t|	[95% conf.	interval]
gender						
Female	.0099489	.1231876	0.08	0.936	-.2317099	.2516076
age	-.001627	.0033898	-0.48	0.631	-.0082769	.0050229
ideo						
Center	1.248514	.1430336	8.73	0.000	.9679235	1.529105
Right	2.09363	.1615159	12.96	0.000	1.776782	2.410478
religion						
Non religious	-.6847867	.126781	-5.40	0.000	-.9334947	-.4360787

Note: dy/dx for factor levels is the discrete change from the base level.

Alternative marginal effects

Marginal effects of gender, ideology and religion **vs. mean**

Table: Marginal effects on government satisfaction **vs. mean**

```
. quietly: summarize government_s if e(sample)
. margins gender ideo religion, exp(predict(xb) - `r(mean)´)
Predictive margins                                      Number of obs = 1,363
Model VCE: OLS
Expression: predict(xb) - 2.987527512839325
```

	Margin	Delta-method std. err.	z	P>|z|	[95% conf.	interval]
gender						
Male	-.0047956	.0852647	-0.06	0.955	-.1719113	.1623201
Female	.0051533	.0884059	0.06	0.954	-.1681191	.1784257
ideo						
Left	-.9968583	.1010192	-9.87	0.000	-1.194852	-.7988643
Center	.2516561	.1013302	2.48	0.013	.0530525	.4502597
Right	1.096771	.1229369	8.92	0.000	.8558197	1.337723
religion						
Religious	.3999195	.096053	4.16	0.000	.211659	.5881799
Non religious	-.2848673	.0807815	-3.53	0.000	-.4431962	-.1265383

Multiple regression

Several response variables with the same predictors

Table: Regression of satisfaction on gender, age, religion and ideology

```
. quietly: mvreg *_s = i.gender age i.religion i.ideo
. quietly: etable, showeq column(index) showstars
. collect layout (colname#result[_r_b _r_se] result[N]) (coleq#stars)
```

	Life	Economy	Government	Democracy	Education	Health
Gender						
Female	0.055	-0.269 *	-0.047	-0.235	-0.376 **	-0.439 **
	(0.117)	(0.116)	(0.126)	(0.141)	(0.123)	(0.139)
Age	0.008 *	-0.000	-0.001	-0.002	-0.007 *	0.003
	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)
Religion						
Non religious	0.053	-0.350 **	-0.610 **	-0.455 **	-0.576 **	-0.457 **
	(0.120)	(0.119)	(0.129)	(0.145)	(0.126)	(0.142)
Ideology						
Center	0.227	0.773 **	1.303 **	0.579 **	-0.073	0.479 **
	(0.136)	(0.135)	(0.147)	(0.165)	(0.143)	(0.162)
Right	0.567 **	1.276 **	2.141 **	1.259 **	0.111	0.345
	(0.152)	(0.151)	(0.165)	(0.184)	(0.160)	(0.181)
Intercept	6.323 **	2.998 **	2.361 **	4.368 **	6.212 **	4.864 **
	(0.223)	(0.221)	(0.240)	(0.269)	(0.234)	(0.265)

'.ado' dime

New way of presenting results

A code is being developed to obtain the marginal effects with an elegant presentation format. It is in the form of ado:

dime (*Differenciated Marginal Effects*) has three possibilities at the moment.

- **mregress** for various multiple regressions separated by parentheses.
- **sureg** for seemingly unrelated regressions (Zellner model).
- **mlogit** for multinomial regressions.

Tidy marginal effects

Tidy marginal effects on gender, age, ideology and religion vs. mean

Table: Marginal effects of government satisfaction **vs. mean**

```
. dime regress (government_s i.gender age i.ideo i.religion)
Table of marginals and global mean differences
```

	Total	Government satisfaction			
Total	(1,363)	3.0			
Male	(706)	3.0	0.0	(0.09)	
Female	(657)	3.0	0.0	(0.09)	
Age	(1,363)	3.0	-0.0	(0.00)	
Left	(513)	2.0	-1.0	(0.10)	***
Center	(498)	3.2	0.3	(0.10)	**
Right	(352)	4.1	1.1	(0.12)	***
Religious	(567)	3.4	0.4	(0.10)	***
Non religious	(796)	2.7	-0.3	(0.08)	***
R2		0.15			

*** p<.001, ** p<.01, * p<.05

Multiple tidy marginal effects in sureg

Tidy marginal effects of gender, age, ideology and religion vs. mean

Table: Margins and marginal effects **vs. mean** of gov. satisfaction

```
. dime sureg (life_s i.gender age i.ideo i.religion) ///
> (democracy_s i.gender age i.ideo i.religion) ///
> (economy_s i.gender age i.ideo i.religion) ///
> (government_s i.gender age i.ideo i.religion)
Table of marginals and global mean differences
```

	Total	Life		Democracy		Economy		Government	
Total	(1,328)	7.0		4.4		3.3		3.0	
Male	(694)	7.0	-0.0	4.6	0.1	3.4	0.1	3.0	0.0
Female	(634)	7.0	0.0	4.3	-0.1	3.1	-0.1	2.9	-0.0
Age	(1,328)	7.0	0.0 *	4.4	-0.0	3.3	-0.0	3.0	-0.0
Left	(497)	6.8	-0.2 **	3.9	-0.5 ***	2.6	-0.6 ***	1.9	-1.0 ***
Center	(480)	7.0	-0.0	4.5	0.0	3.4	0.1	3.2	0.3 *
Right	(351)	7.4	0.4 **	5.2	0.7 ***	4.0	0.7 ***	4.1	1.1 ***
Religious	(556)	7.0	-0.0	4.7	0.3 *	3.5	0.2 *	3.3	0.4 ***
Non religious	(772)	7.0	0.0	4.2	-0.2 *	3.1	-0.1 *	2.7	-0.3 **
R2		0.02		0.05		0.08		0.16	

*** p<.001, ** p<.01, * p<.05

Multiple tidy marginal effects in mlogit regression

Tidy marginal effects of gender, age, ideology and religion **vs. main %**

Table: Margins and marginal effects **vs. main percentages** of voting

```
. dime mlogit GBvote i.gender age i.ideo i.religion
Table of marginals and global mean differences
```

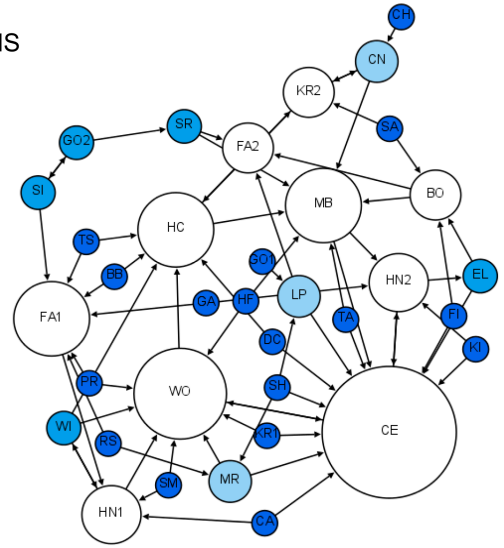
	Total	Conservative		Labour		Liberal Democrat		Others	
Total	(949)	37.2		38.5		11.3		13.1	
Male	(496)	38.7	1.5	34.8	-3.7	12.4	1.1	14.2	1.1
Female	(453)	35.5	-1.7	42.4	3.9	10.2	-1.1	12.0	-1.1
Age	(949)	37.5	0.3 ***	38.1	-0.3 ***	11.3	0.1	13.0	-0.1
Left	(374)	7.5	-29.7 ***	60.5	22.0 ***	15.3	4.1 *	16.6	3.6
Center	(315)	42.3	5.1	29.8	-8.7 ***	13.0	1.7	15.0	1.9
Right	(260)	72.2	35.0 ***	16.8	-21.6 ***	4.6	-6.7 ***	6.4	-6.6 ***
Religious	(412)	35.0	-2.2	38.6	0.1	12.0	0.7	14.4	1.3
Non religious	(537)	39.1	1.9	38.1	-0.4	10.7	-0.6	12.1	-1.0

```
*** p<.001, ** p<.01, * p<.05
Pseudo R2: 0.163; Nagelkerke´s R2: 0.363; chi2: 384.92; p: 1.013e-72
```

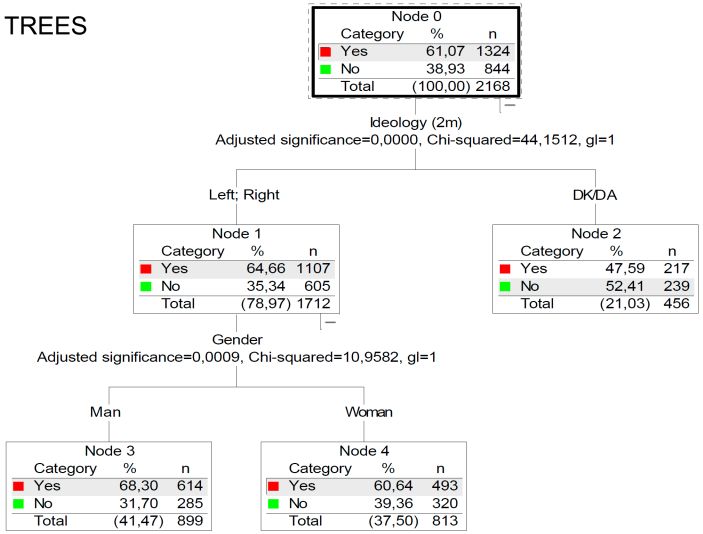
Graphs in social research

Examples of graphs

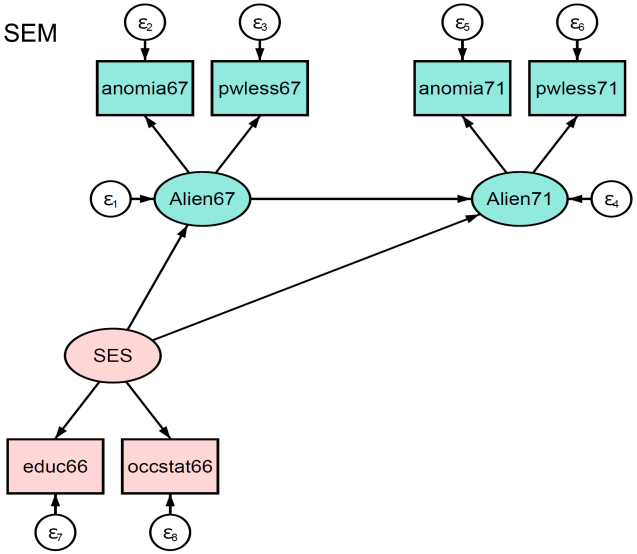
GRAPHS



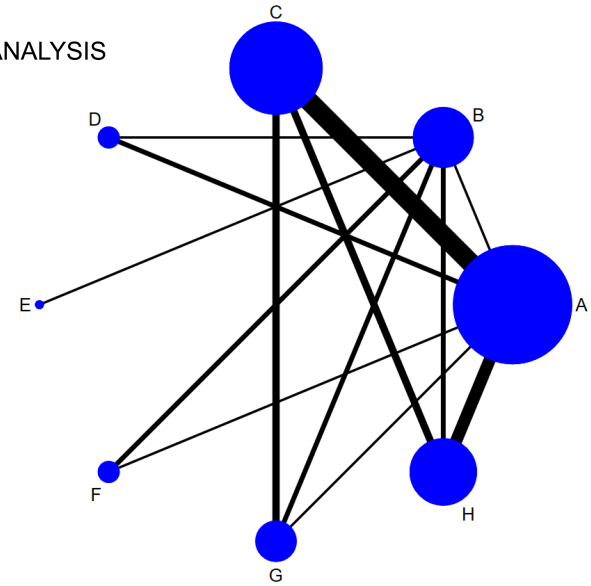
TREES



SEM



META-ANALYSIS



Regression graphs

With marginal effects

- As mentioned before, marginal effects represent the change in the dependent variable that accounts for an increase in the independent variables of a general linear model.
- Marginal effects are suitable (with certain precautions) to convert them into graphs.
- We are developing Stata and R programs to obtain these marginal effects and to construct two tables: one of nodes (variables or categories with their properties) and another of links (between variables with a significant (and positive) marginal effect).
- These two tables are the starting point for the elaboration of a network graph.

Multiple tidy marginal effects in sureg

Tidy marginal effects of gender, age, ideology and religion vs. mean

Table: Margins and marginal effects on mean of gov. satisfaction

```
. dime sureg (life_s i.gender age i.ideo i.religion) ///
> (democracy_s i.gender age i.ideo i.religion) ///
> (economy_s i.gender age i.ideo i.religion) ///
> (government_s i.gender age i.ideo i.religion),  graph
```

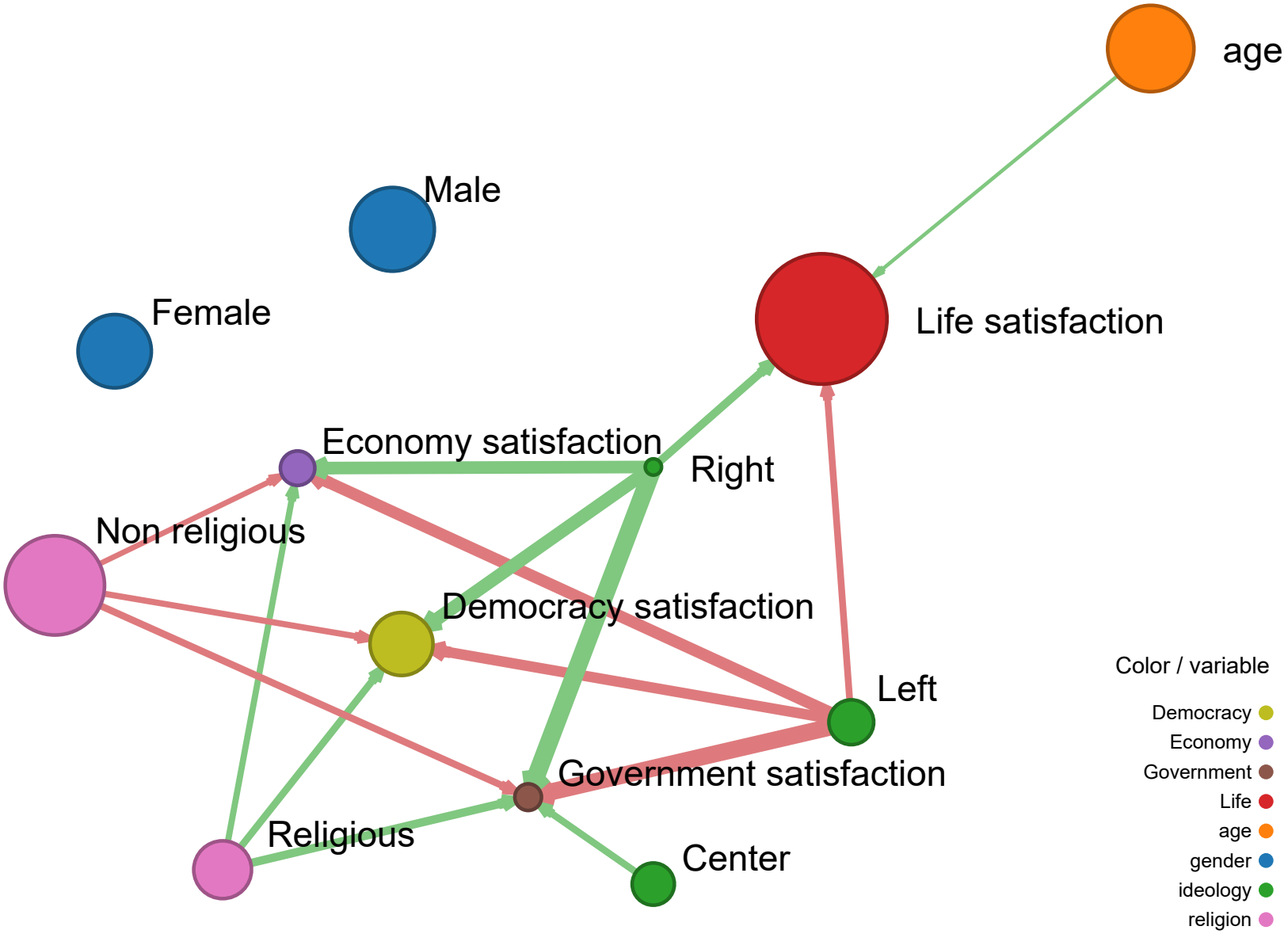
Table of marginals and global mean differences

	Total	Life	Democracy	Economy	Government
Total	(1,328)	7.0	4.4	3.3	3.0
Male	(694)	7.0 -0.0	4.6 0.1	3.4 0.1	3.0 0.0
Female	(634)	7.0 0.0	4.3 -0.1	3.1 -0.1	2.9 -0.0
Age	(1,328)	7.0 0.0 *	4.4 -0.0	3.3 -0.0	3.0 -0.0
Left	(497)	6.8 -0.2 **	3.9 -0.5 ***	2.6 -0.6 ***	1.9 -1.0 ***
Center	(480)	7.0 -0.0	4.5 0.0	3.4 0.1	3.2 0.3 *
Right	(351)	7.4 0.4 **	5.2 0.7 ***	4.0 0.7 ***	4.1 1.1 ***
Religious	(556)	7.0 -0.0	4.7 0.3 *	3.5 0.2 *	3.3 0.4 ***
Non religious	(772)	7.0 0.0	4.2 -0.2 *	3.1 -0.1 *	2.7 -0.3 **
R2		0.02	0.05	0.08	0.16

*** p<.001, ** p<.01, * p<.05

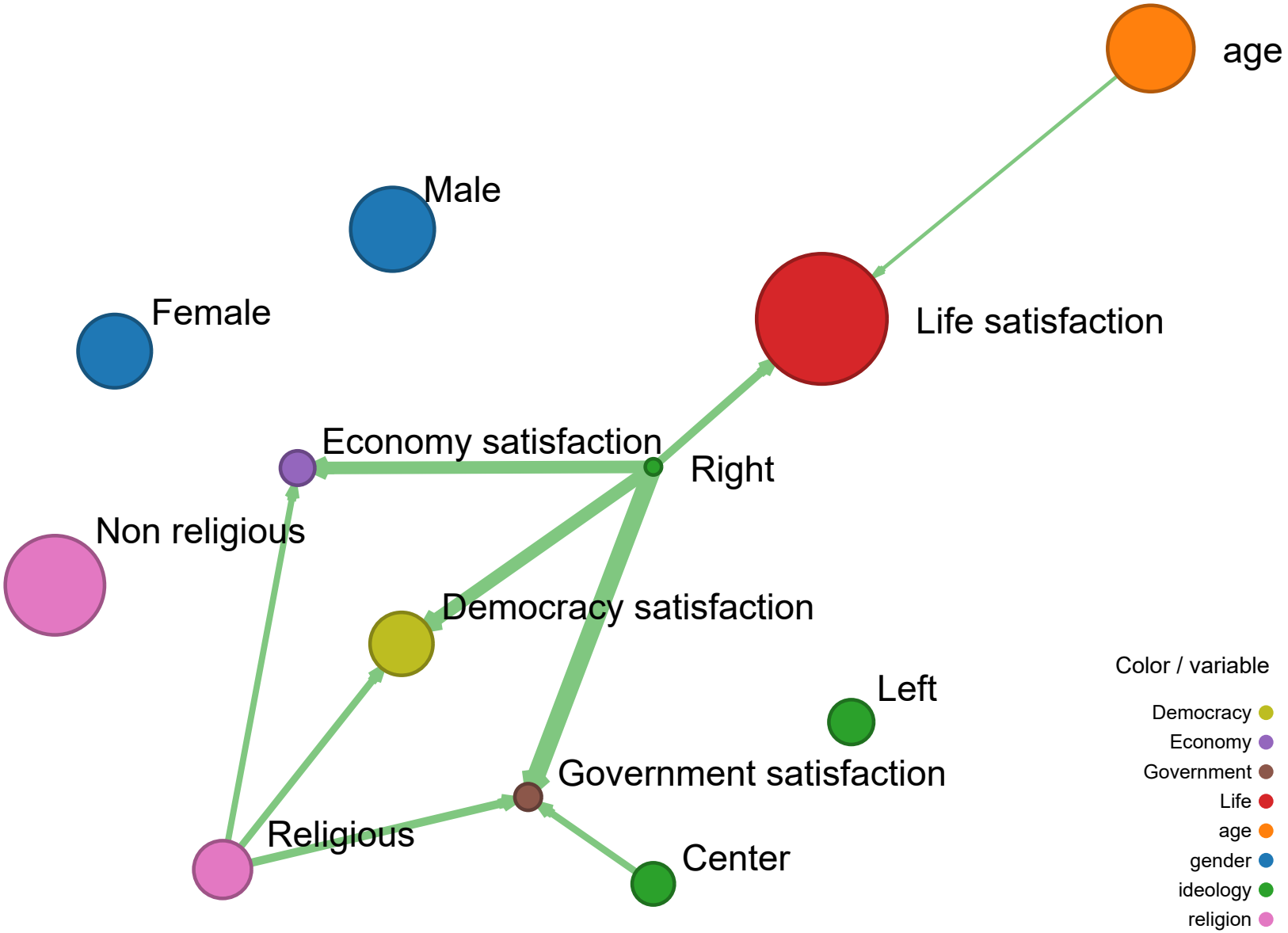
Regression graph

Regression of p(satisfaction) on gender, age, ideology and ideology (bipolar



Regression graph

Regression of p(satisfaction) on gender, age, ideology and ideology



Multiple tidy marginal effects in mlogit regression

Tidy marginal effects of gender, age, ideology and religion vs. mean

Table: Margins and marginal effects vs. main percentages of voting

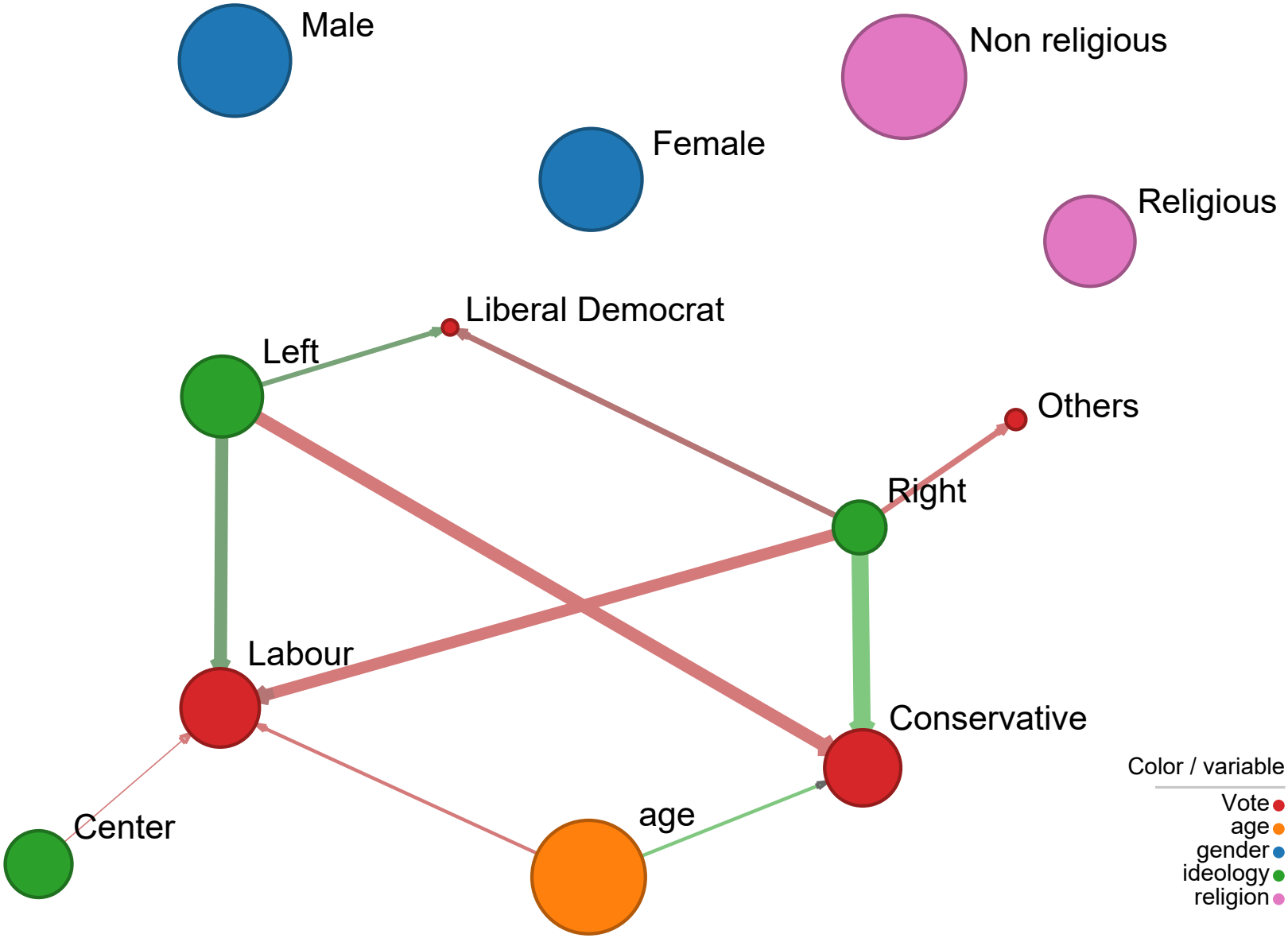
```
. dime mlogit GBvote i.gender age i.ideo i.religion,  graph
Table of marginals and global mean differences
```

	Total	Conservative		Labour		Liberal Democrat		Others	
Total	(949)	37.2		38.5		11.3		13.1	
Male	(496)	38.7	1.5	34.8	-3.7	12.4	1.1	14.2	1.1
Female	(453)	35.5	-1.7	42.4	3.9	10.2	-1.1	12.0	-1.1
Age	(949)	37.5	0.3 ***	38.1	-0.3 ***	11.3	0.1	13.0	-0.1
Left	(374)	7.5	-29.7 ***	60.5	22.0 ***	15.3	4.1 *	16.6	3.6
Center	(315)	42.3	5.1	29.8	-8.7 ***	13.0	1.7	15.0	1.9
Right	(260)	72.2	35.0 ***	16.8	-21.6 ***	4.6	-6.7 ***	6.4	-6.6 ***
Religious	(412)	35.0	-2.2	38.6	0.1	12.0	0.7	14.4	1.3
Non religious	(537)	39.1	1.9	38.1	-0.4	10.7	-0.6	12.1	-1.0

```
*** p<.001, ** p<.01, * p<.05
Pseudo R2: 0.163; Nagelkerke´s R2: 0.363; chi2: 384.92; p: 1.013e-72
```

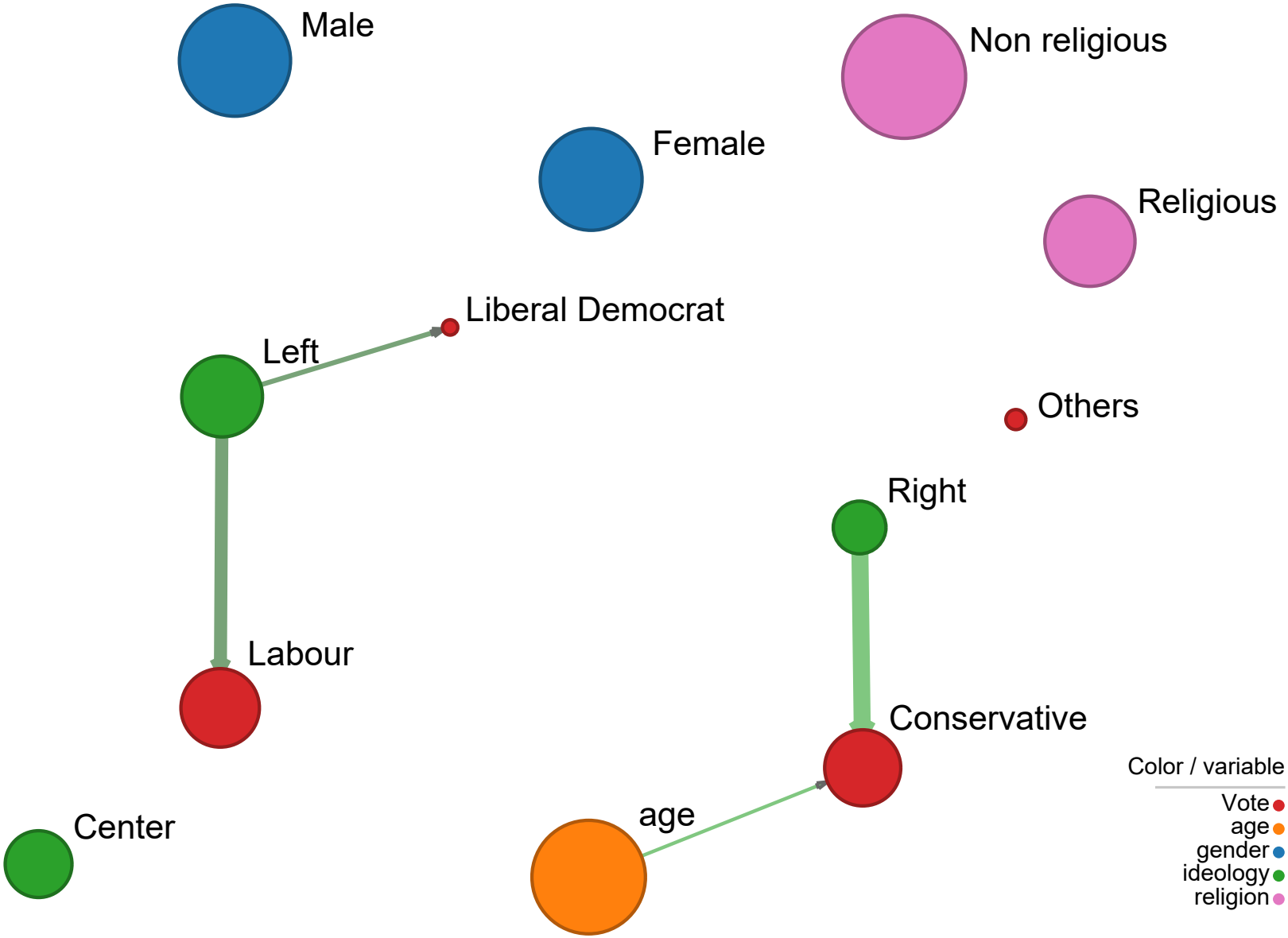
Multinomial graph

Multilogit regression of government satisfaction (bipolar)



Multinomial graph

Multilogit regression of government satisfaction



Rules for net-representation of regressions

- **Regression graphs** are proposed as a visual analytical framework to represent dependency models between variables and categories.
- In these representations the **nodes** are the variables or categories and their size represents their importance.
  - By frequency (or percentage) in the case of categorical variables.
  - By the normalization on a scale from 0 to 100 in the case of numerical variables.
- In these graphs, the width of the **links** represents the association between variables or categories..
  - With certain precautions, the **marginal effects** with respect to the mean value are used for them.
  - In the case of having many variables or categories, it is recommended to express only the **significant** (and **positive**) effects.

Remarks

About regression graphs

- In this presentation we have proposed a way of presenting margins and marginal effects **tables**.
- This may be useful to represent friendly and tidily several (multiple) regressions and (ordinal or multinomial) logistic or probit **models**.
- These marginal effects can be also graphed in a **network** way with programs we are developing for Stata (dime in coin) and R languages (glmCoin in netCoin).
- New models as poisson and negative binomial are going to be incorporated.
- Among other possible **challenges** would be the consideration of interaction terms and other models as mixed and longitudinal models.

Thanks for your attention!

modesto@usal.es
cristinacalvolopez@usal.es

References on Analytical networks

- Escobar, M. (2009). Redes semánticas en textos periodísticos: propuestas técnicas para su representación. *Empiria*, 17, 13-39.
- Escobar, M., and Gómez Isla, J. (2015). "La expresión de la identidad a través de la imagen: los archivos fotográficos de Miguel de Unamuno y Joaquín Turina". *Revista Española de Investigaciones Sociológicas*, 152, 23-46.
- Escobar, M. (2015). "Studying Coincidences with Network Analysis and Other Multivariate Tools". *The Stata Journal*, 15(4), 1118-1156.
- Escobar, M. (2016). "Ensayo sobre las coincidencias". In A. Almarcha, P. González, and L. Román (Eds.), *Donde la Sociología te lleve*. A Coruña: Universidad de A Coruña.
- Escobar, M., and C. Tejero (2018). "El análisis reticular de coincidencias". *Empiria*, 39, 129-148.
- Escobar, M. and L. Martínez (2020). "Network Coincidence Analysis: the netCoin R Package". *International Journal of Software*, 93, 11.