

Las elecciones (y las elecciones) a través de datos

Kiko Llaneras

Kiko Llaneras | @kikollan | politikon.es

Tres temas:

Explicar cómo son los votantes

Seguir la evolución de mil encuestas

Predecir el próximo gobierno (o intentarlo)

Es análisis de datos:

Hay que tratar con números,
saber (algo de) estadística,
y tener una **actitud científica**

También es periodismo:

Recursos escasos y mucha prisa,
el lector no es ningún experto,
y tienes que **divertirlo**

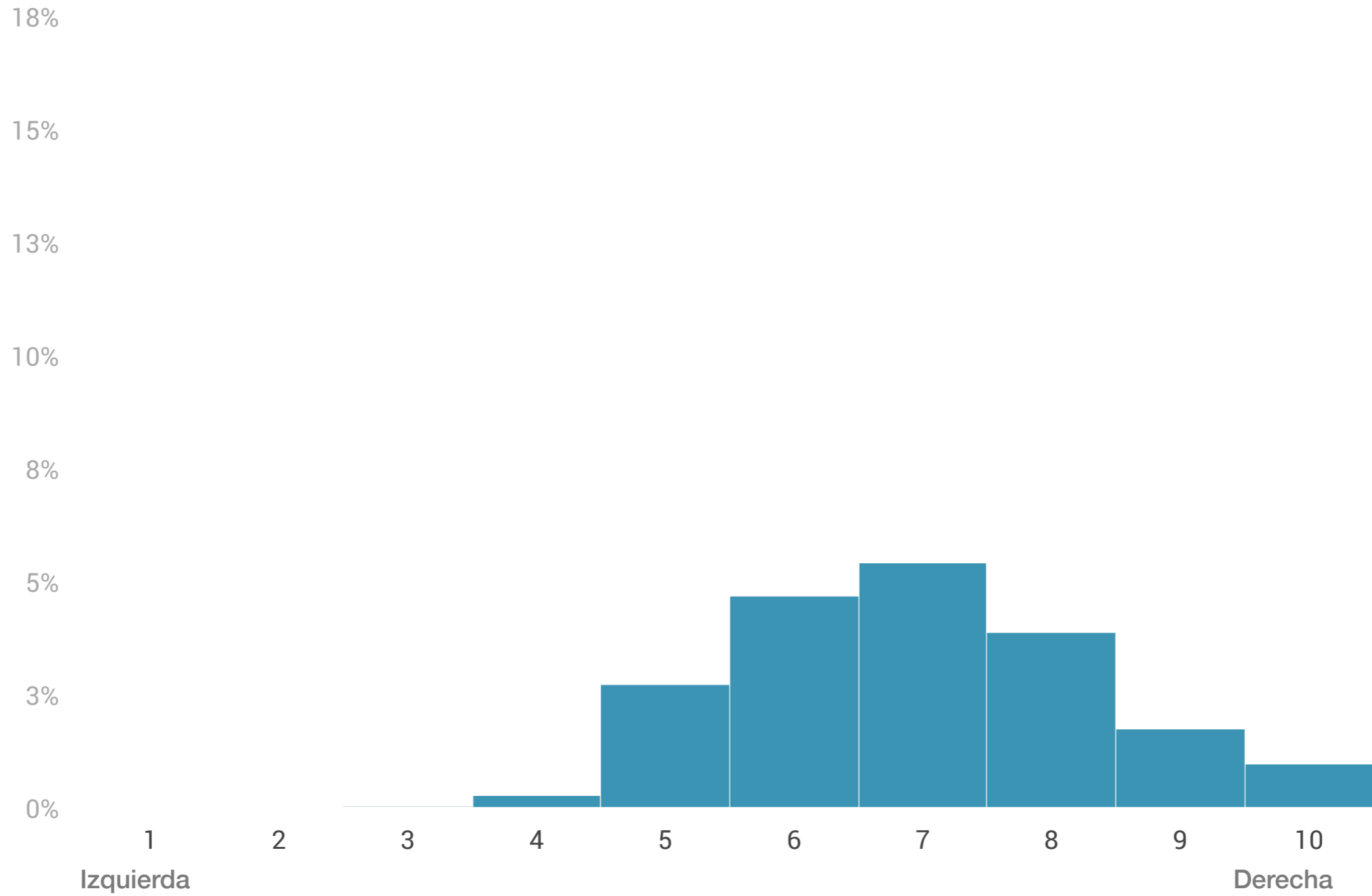
¿Cómo son los votantes?

Responder con datos

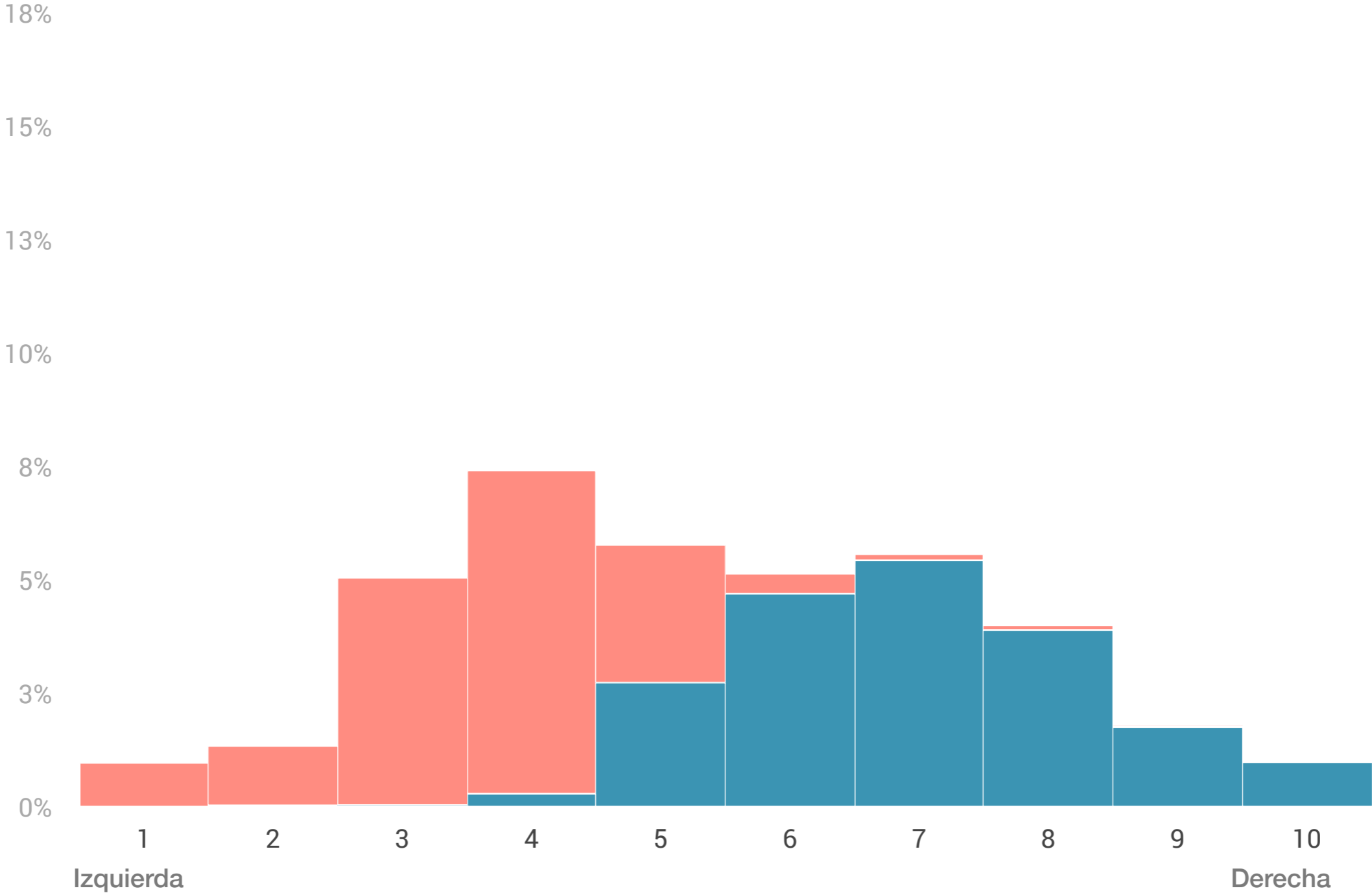
Distribución de los votantes por autoubicación en izquierda y derecha, y según el partido que votaron el 20D.



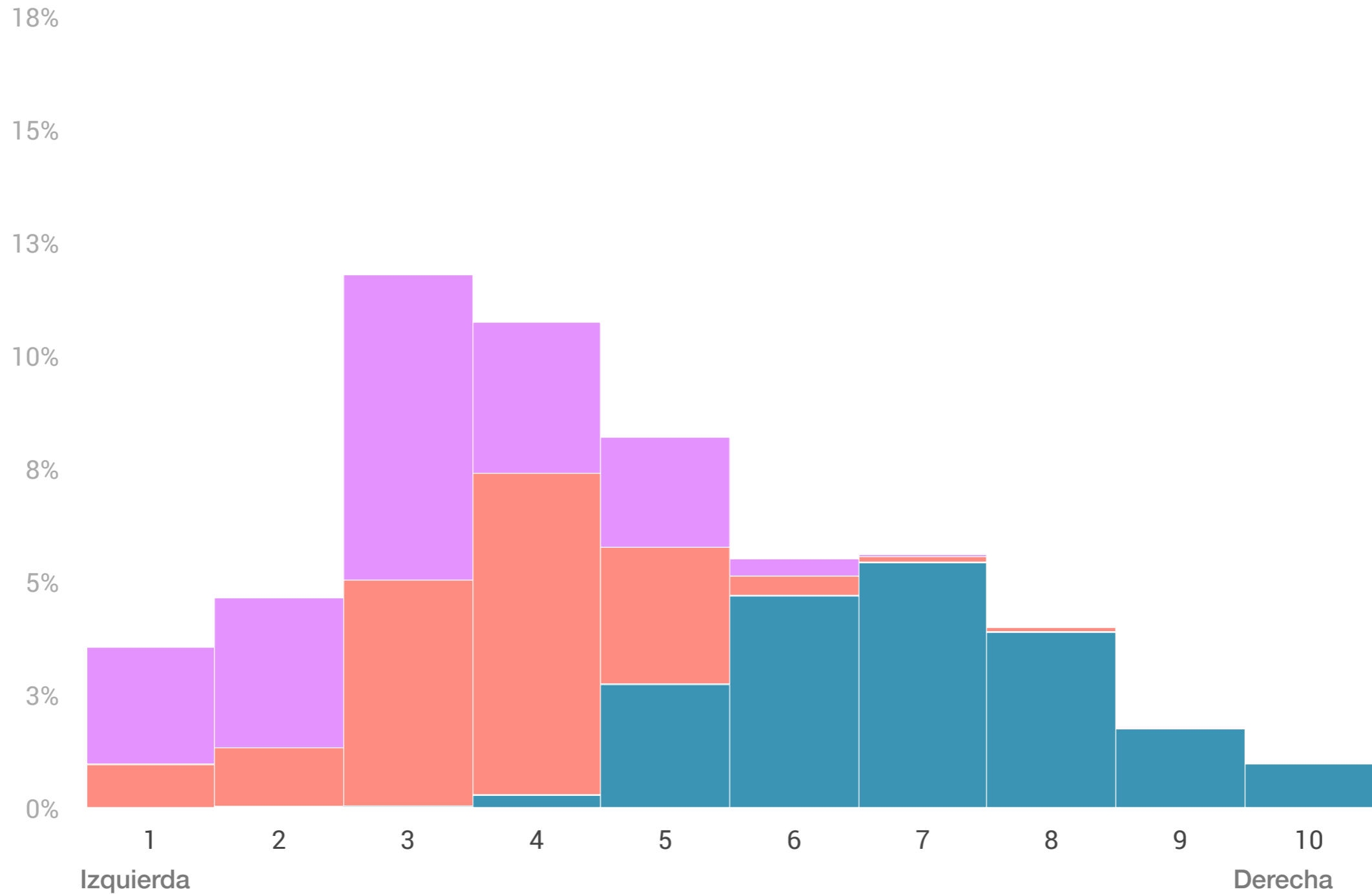
Distribución de los votantes por autoubicación en izquierda y derecha, y según el partido que votaron el 20D.



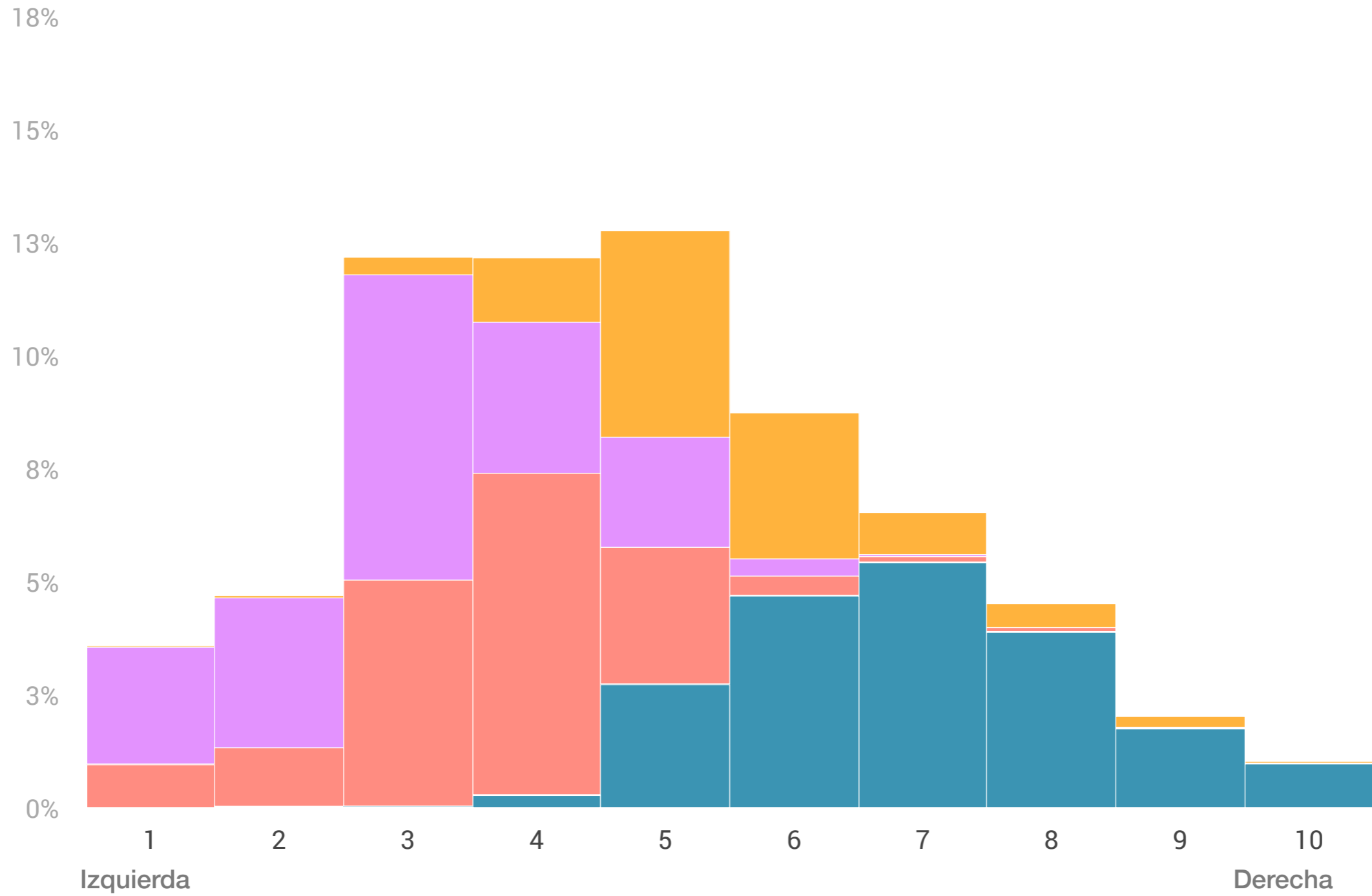
Distribución de los votantes por autoubicación en izquierda y derecha, y según el partido que votaron el 20D.



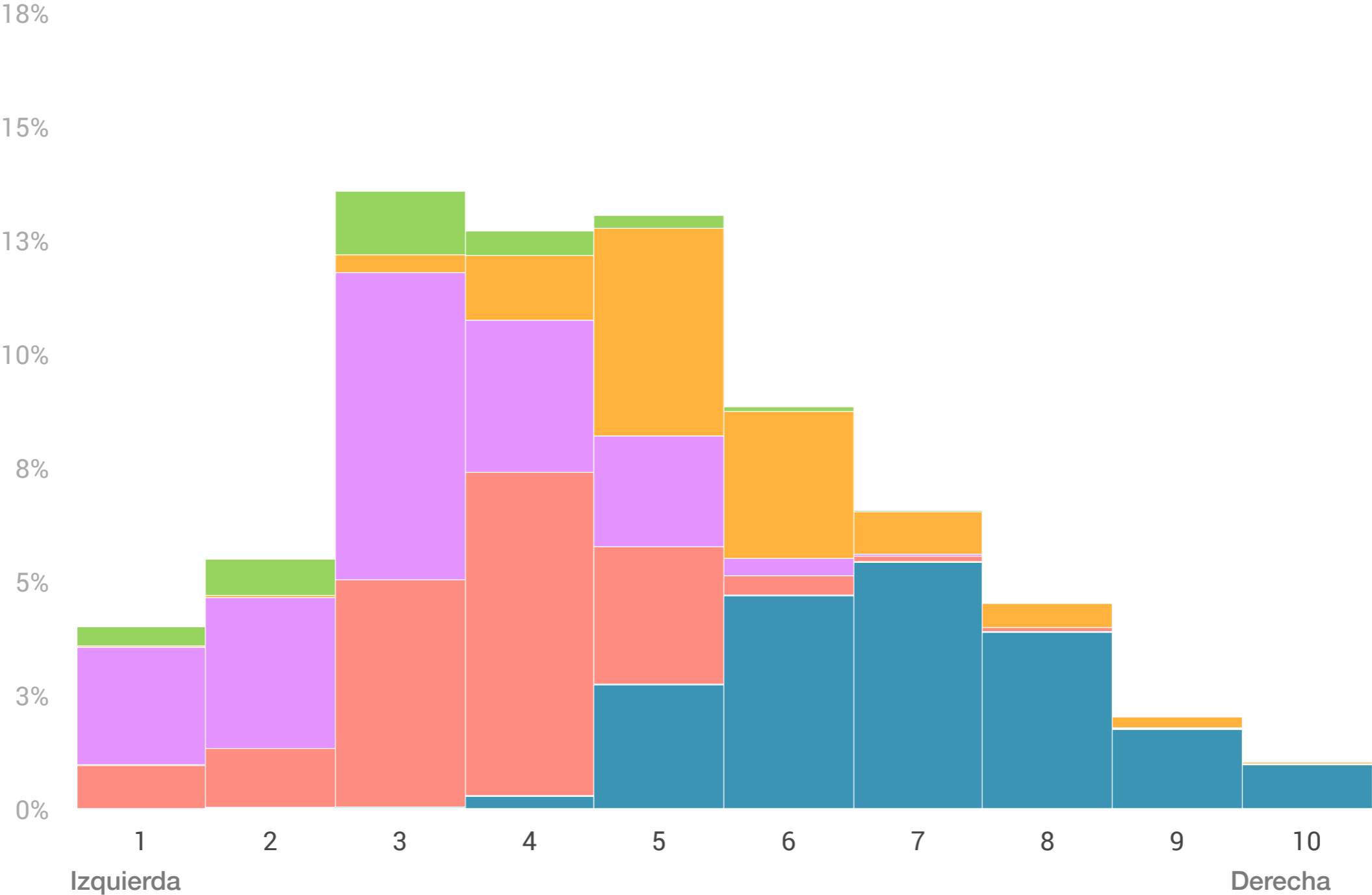
Distribución de los votantes por autoubicación en izquierda y derecha, y según el partido que votaron el 20D.



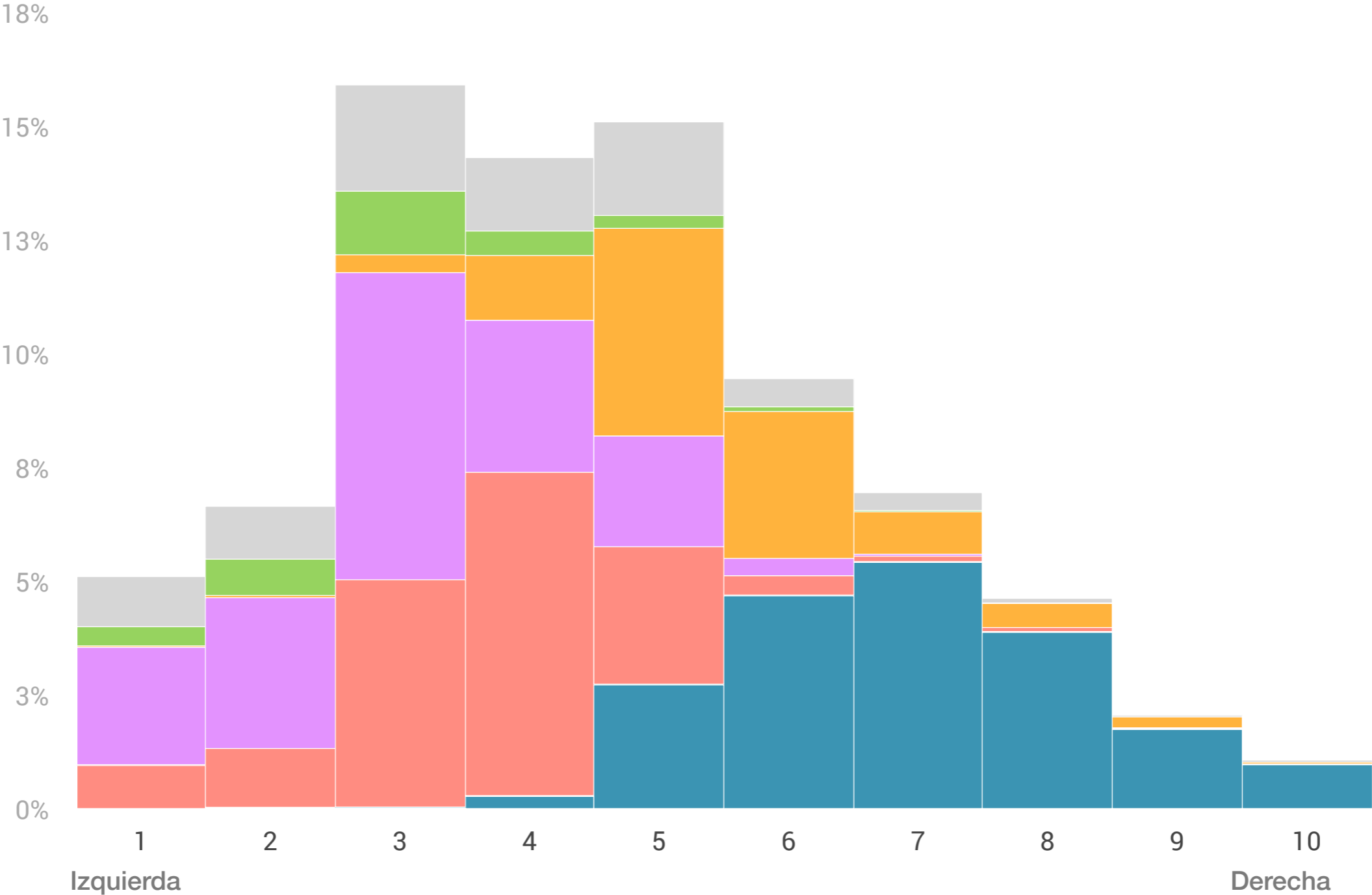
Distribución de los votantes por autoubicación en izquierda y derecha, y según el partido que votaron el 20D.



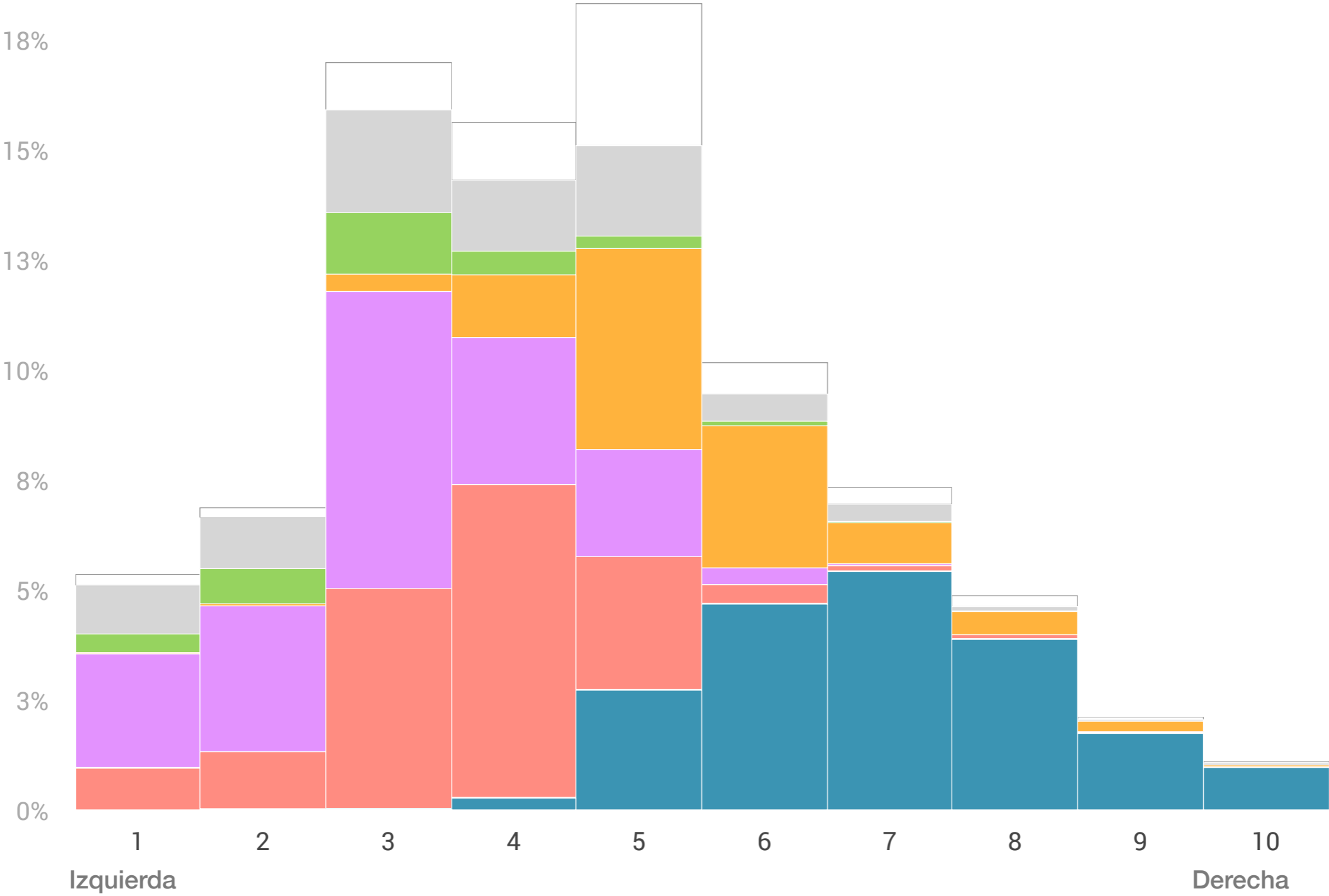
Distribución de los votantes por autoubicación en izquierda y derecha, y según el partido que votaron el 20D.



Distribución de los votantes por autoubicación en izquierda y derecha, y según el partido que votaron el 20D.

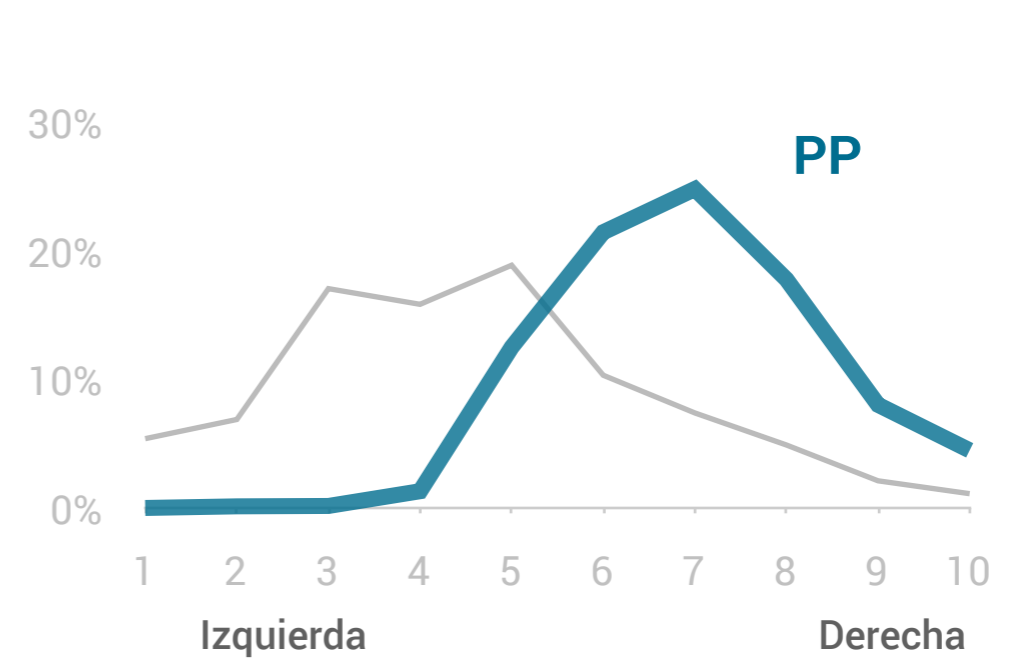
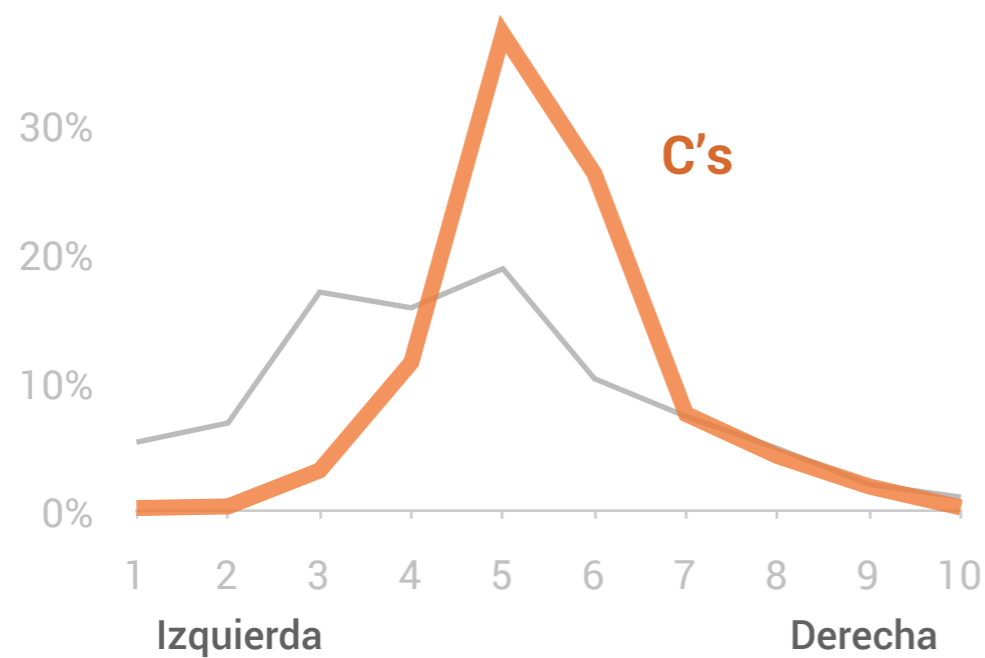
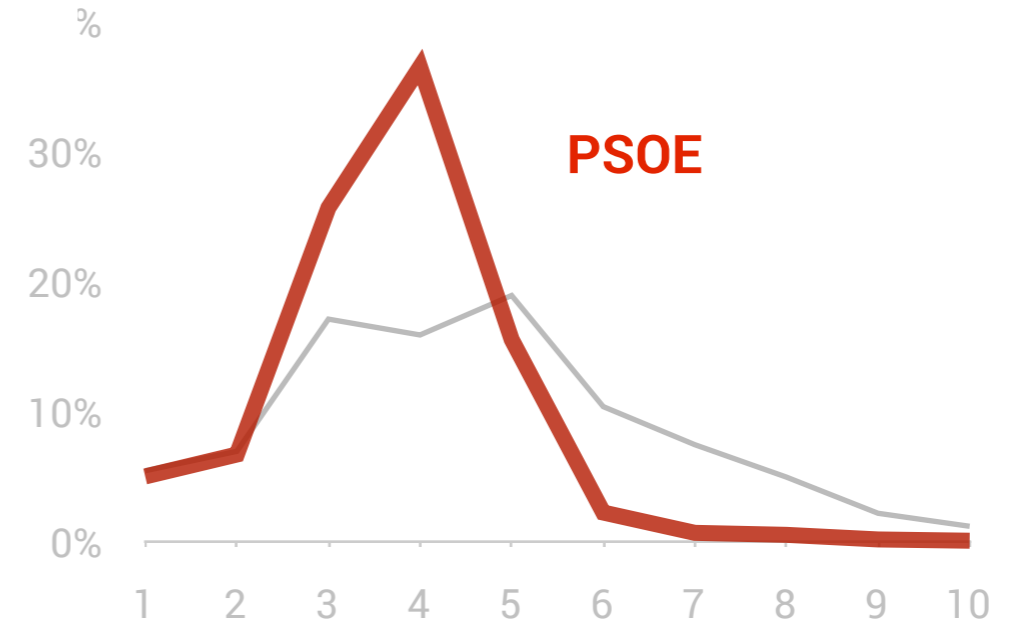
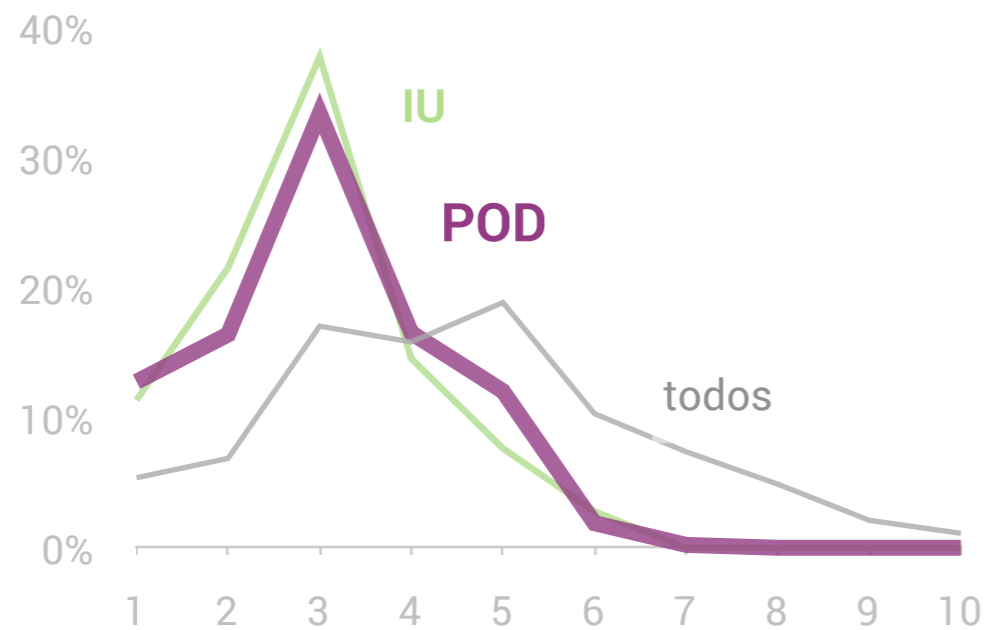


Distribución de los votantes por autoubicación en izquierda y derecha, y según el partido que votaron el 20D.



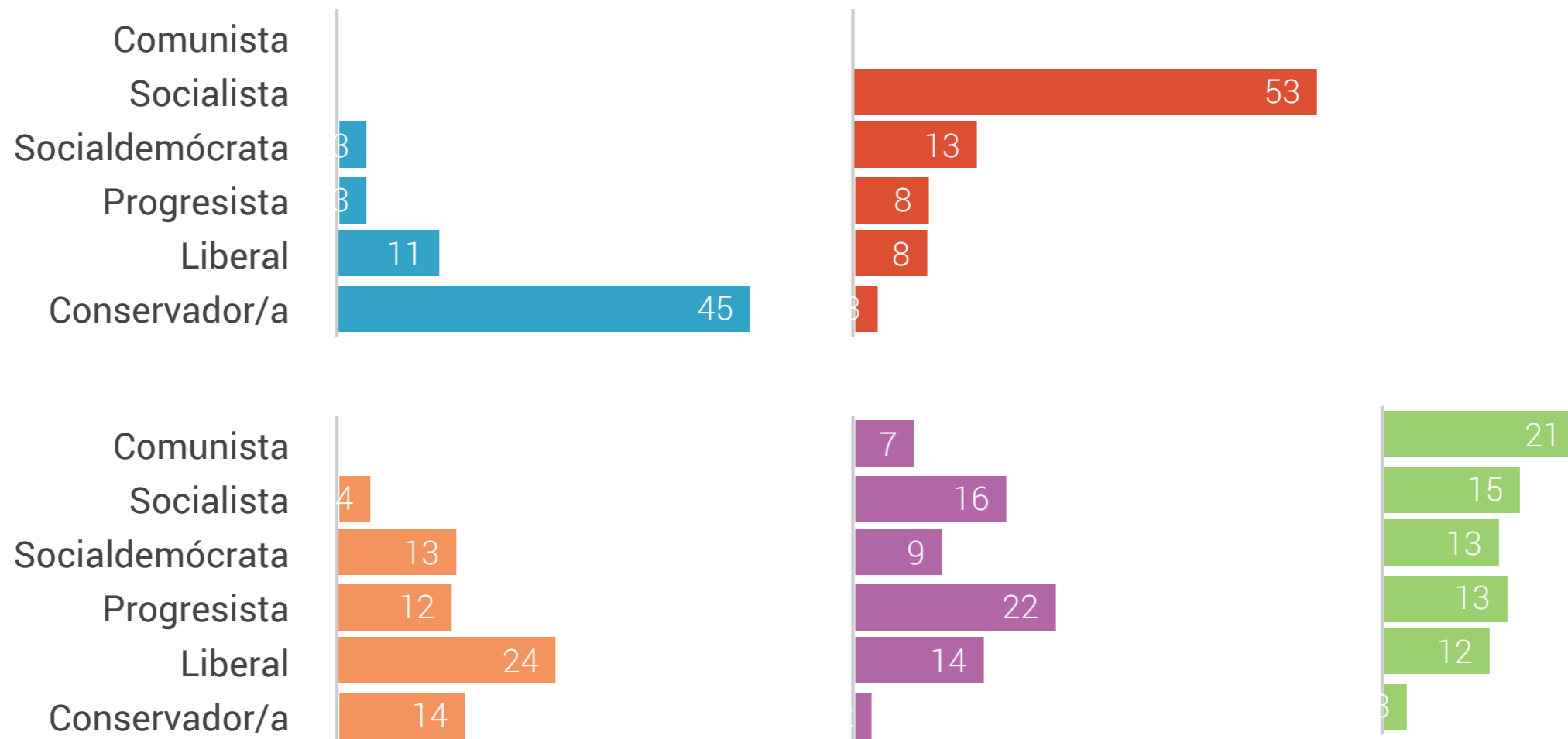
Ideología de los votantes

Autoubicación en izquierda/derecha de los votantes del 20D.



Ideologías en cada partido

Porcentaje de votantes de cada partido el 20D según la etiqueta con que se identifican.



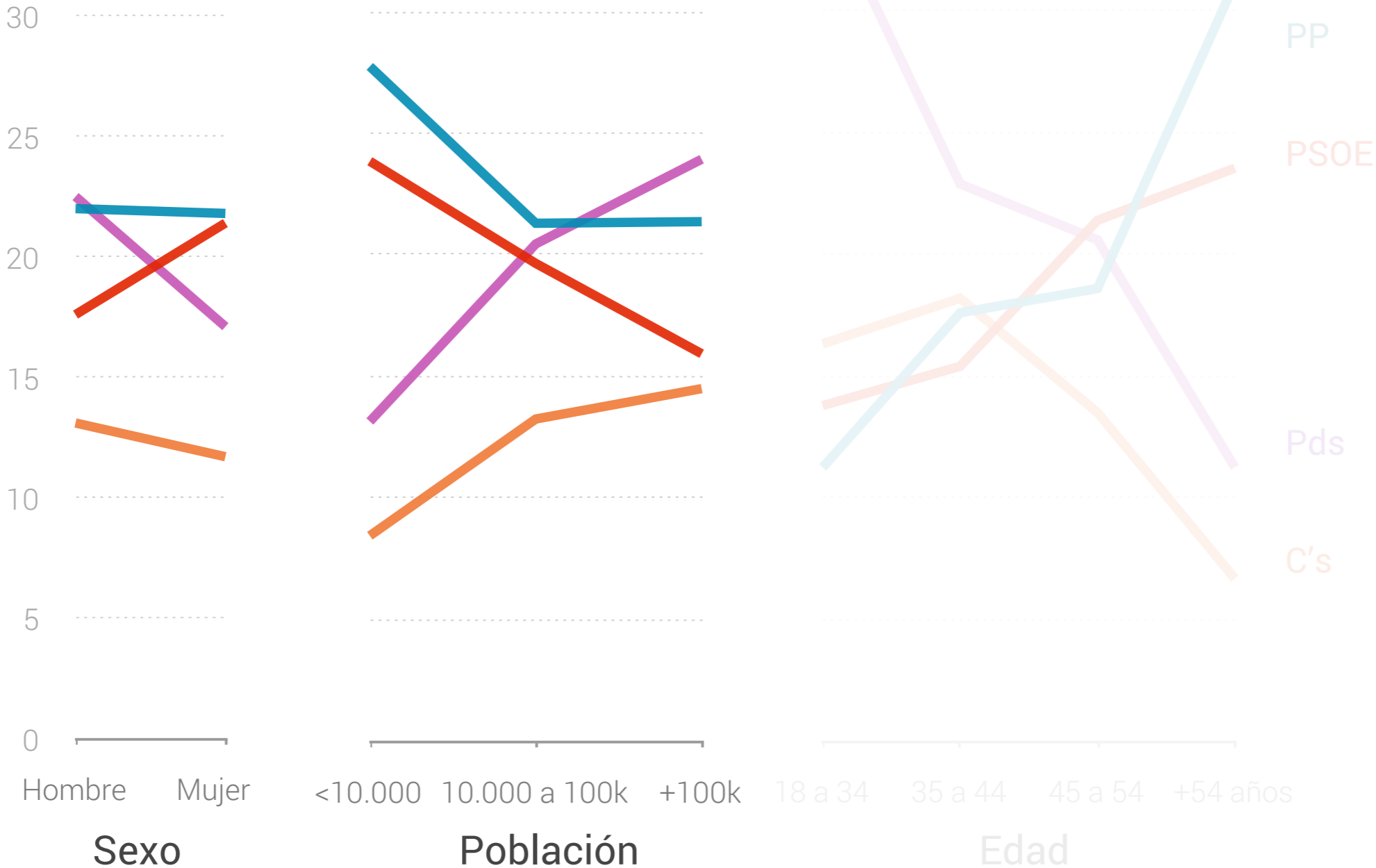
Votantes por sexo, hábitat y edad

Porcentaje de personas en la categoría que votaron a cada partido el 20D.



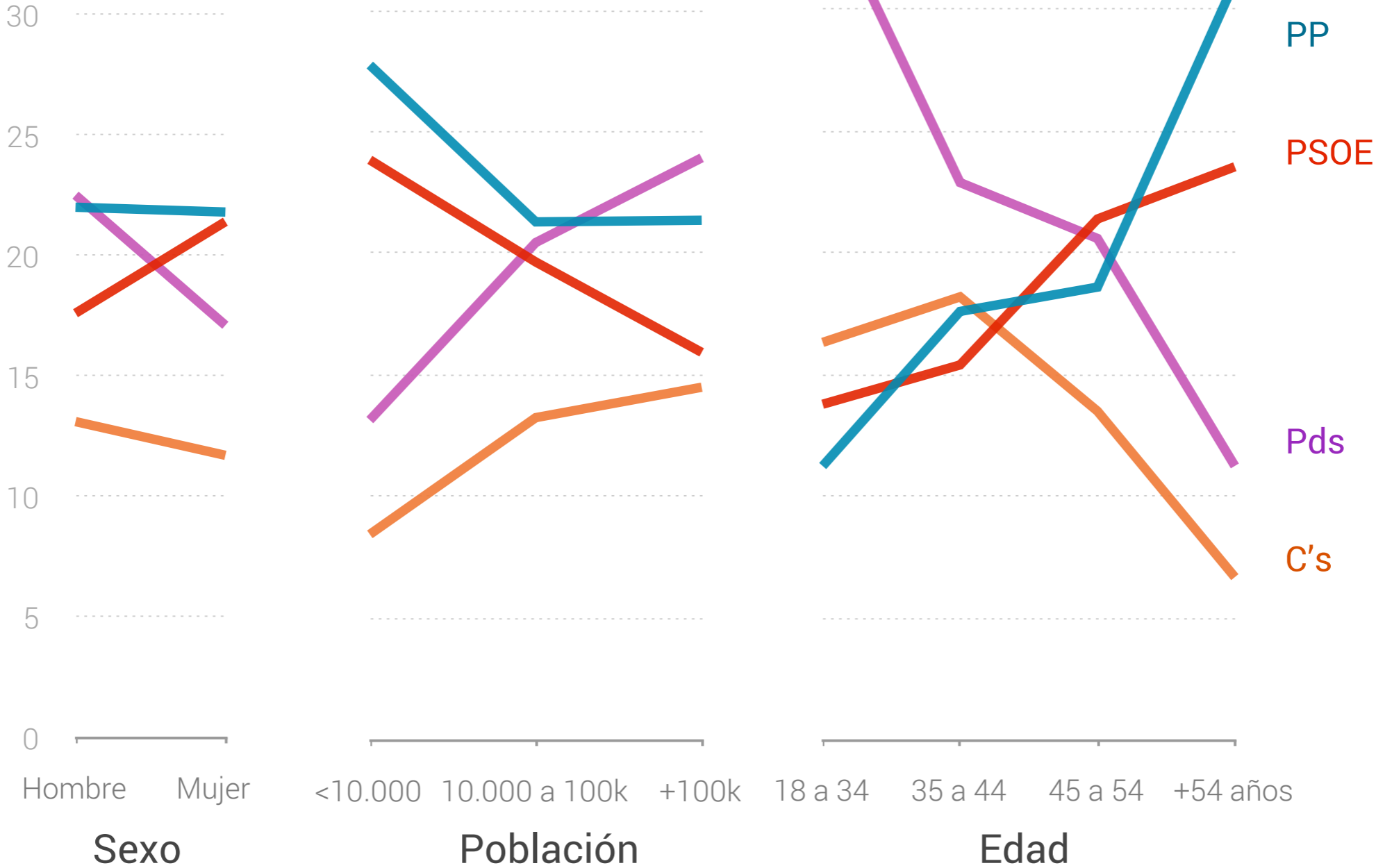
Votantes por sexo, hábitat y edad

Porcentaje de personas en la categoría que votaron a cada partido el 20D.



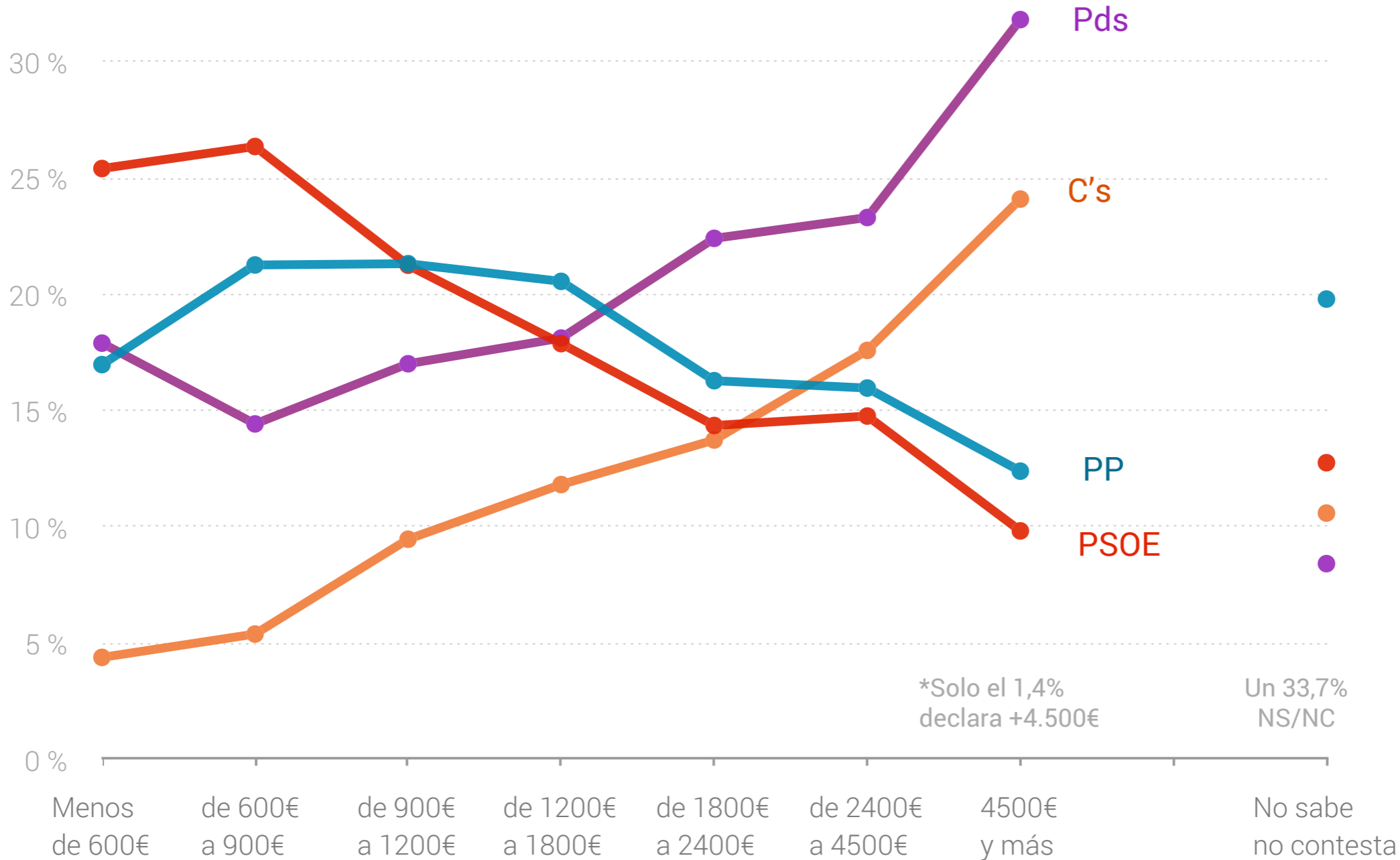
Votantes por sexo, hábitat y edad

Porcentaje de personas en la categoría que votaron a cada partido el 20D.



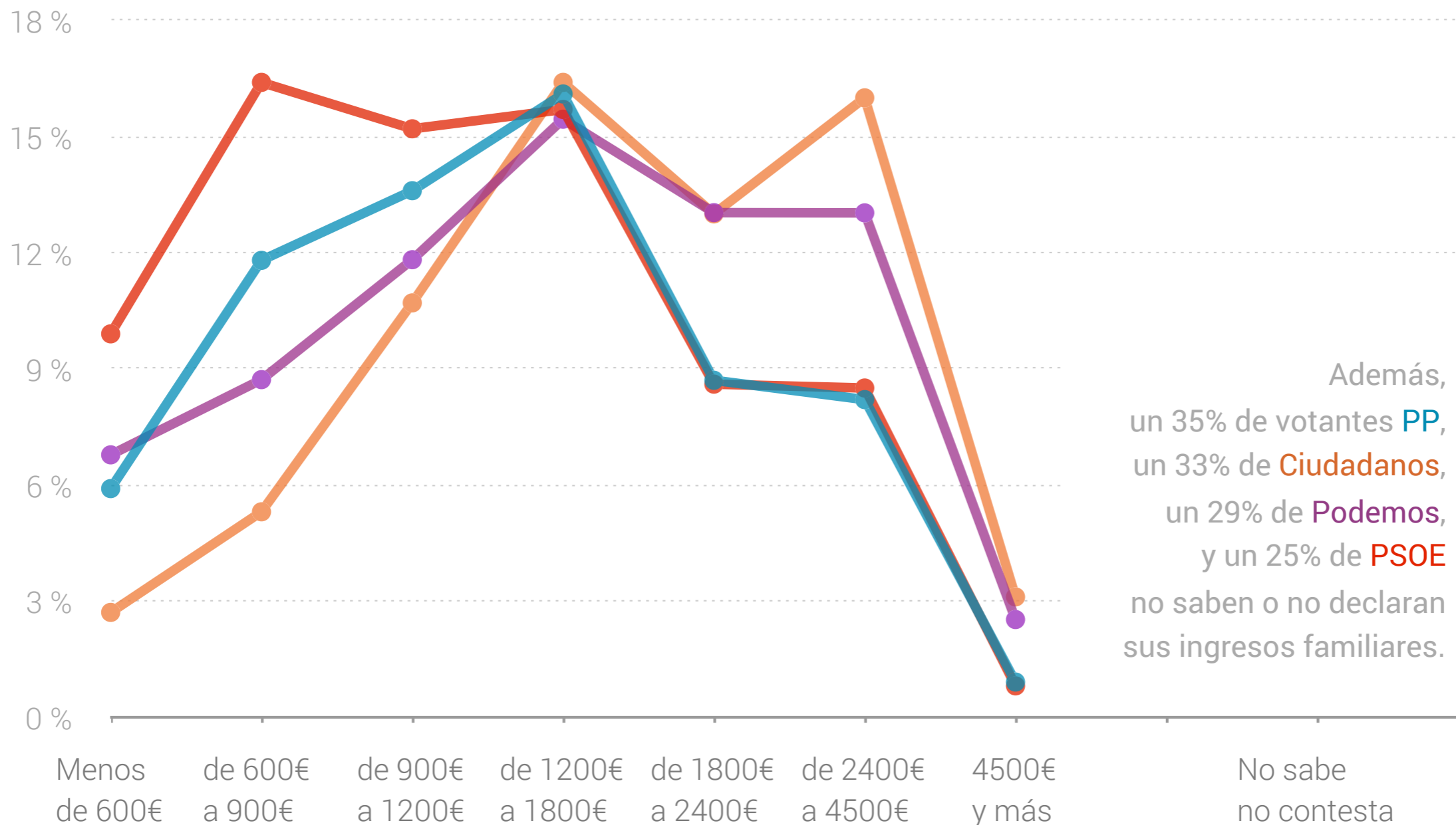
Votantes por ingresos del hogar

Personas en la categoría que votaron a cada partido el 20D.



Perfil de cada partido, por ingresos del hogar

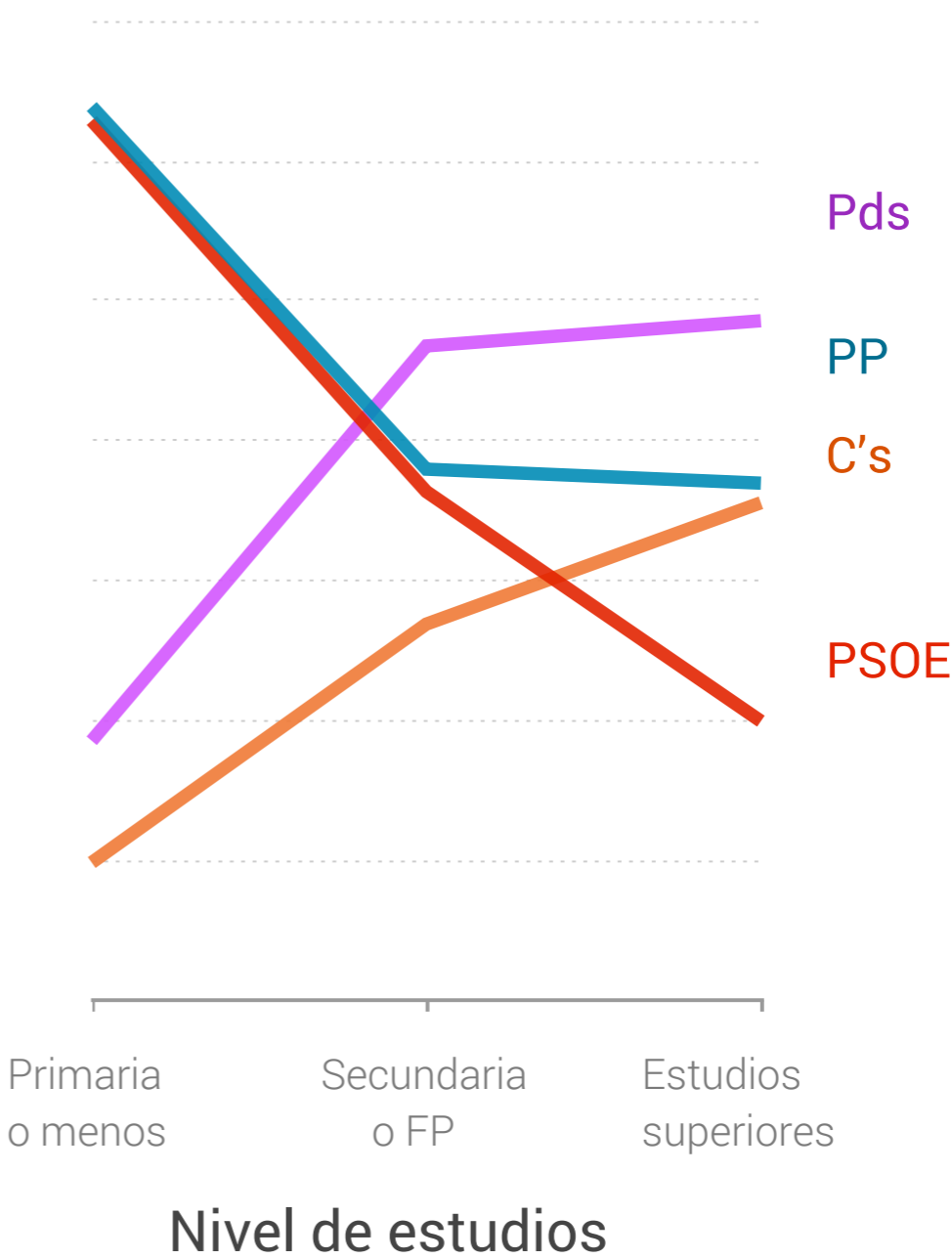
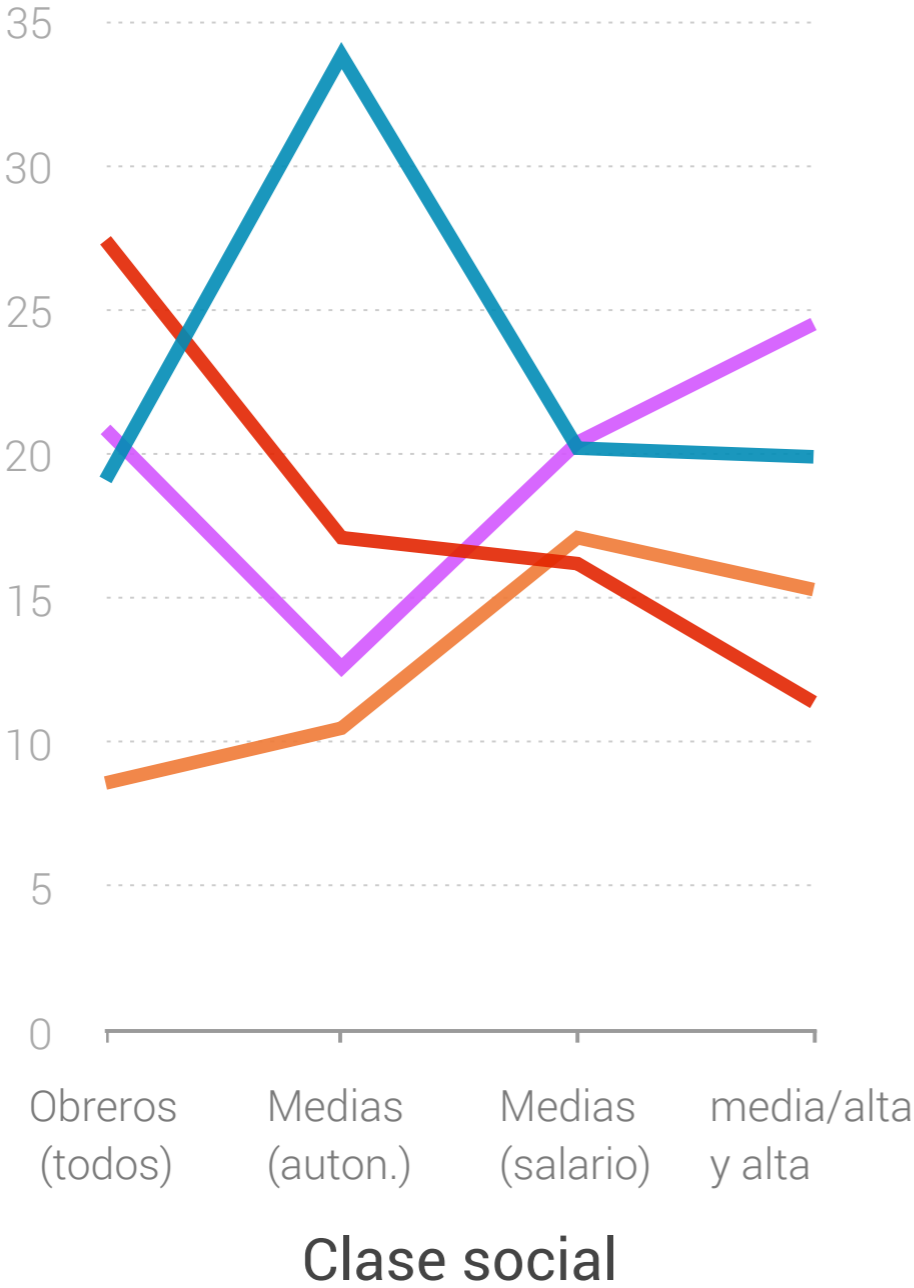
Distribución de los votantes de cada partido según su renta.



Además,
un 35% de votantes **PP**,
un 33% de **Ciudadanos**,
un 29% de **Podemos**,
y un 25% de **PSOE**
no saben o no declaran
sus ingresos familiares.

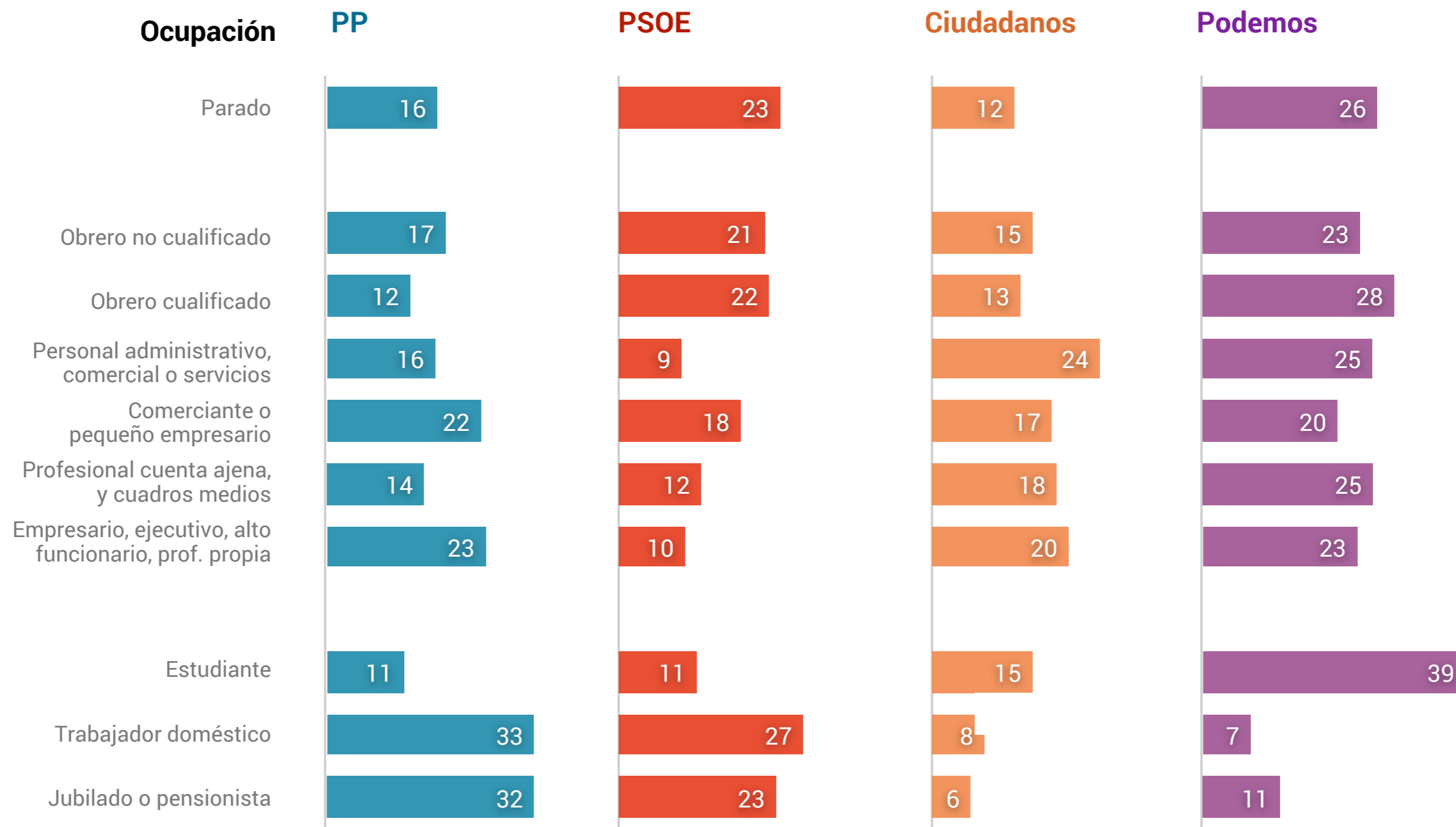
Votantes por clase social y estudios

Porcentaje de personas en la categoría que votaron a cada partido el 20D.



Votantes por situación laboral

Porcentaje de personas en la categoría que votaron a cada partido el 20D.



¡Está lleno de encuestas!

Distinguir entre ruido y señal

Se publican decenas de encuestas cada mes.

Un promedio es una forma de seguirlas.

Ventajas

- Reduce error de muestreo
- Cocina de “consenso”

GAD3	800	6/4/16	29,2 %	23,3 %	15,5 %	14,4 %	6,7 %	10,9 %
Metroscopia	800	5/4/16	29,0 %	20,1 %	17,0 %	17,7 %	6,7 %	9,5 %
DYM	1036	4/4/16	29,2 %	20,7 %	19,6 %	15,2 %	5,9 %	9,4 %
CELESTETEL	550	3/4/16	29,1 %	22,2 %	19,4 %	15,9 %	4,5 %	8,9 %
Metroscopia	400	29/3/16	27,7 %	21,0 %	15,9 %	18,8 %	6,9 %	9,7 %
EE	1002	28/3/16	27,5 %	19,3 %	17,3 %	16,9 %	7,3 %	11,7 %
Invymark	1200	30/3/16	28,6 %	21,6 %	19,9 %	15,2 %	4,1 %	10,6 %
Sigma Dos	800	30/3/16	30,2 %	22,0 %	16,7 %	16,2 %	4,5 %	10,4 %
My Word	1004	13/3/16	27,0 %	20,9 %	15,9 %	18,0 %	7,1 %	11,1 %
GAD3	800	9/3/16	28,2 %	24,2 %	18,7 %	14,1 %	3,2 %	11,6 %
NC Report	900	27/3/16	29,4 %	21,8 %	19,2 %	15,5 %	4,6 %	9,5 %
Metroscopia	1200	9/3/16	26,0 %	23,1 %	16,8 %	19,5 %	5,4 %	9,2 %
Invymark	400	8/3/16	28,5 %	20,8 %	21,5 %	14,3 %	4,3 %	10,6 %
Sondaxe	1002	8/3/16	28,6 %	22,0 %	19,1 %	15,6 %	4,2 %	10,5 %
Gesop	1000	7/3/16	26,4 %	21,9 %	18,1 %	17,8 %	5,9 %	9,9 %
Simple Lógica	1206	5/3/16	26,2 %	21,6 %	18,3 %	17,1 %	7,9 %	8,9 %
CELESTETEL	550	4/3/16	28,3 %	22,4 %	19,4 %	15,6 %	4,6 %	9,7 %
NC Report	700	1/3/16	29,2 %	21,8 %	20,2 %	14,9 %	4,2 %	9,7 %
Invymark	400	24/2/16	26,8 %	20,3 %	22,1 %	15,8 %	4,5 %	10,5 %
DYM	1071	21/2/16	25,6 %	22,1 %	19,8 %	17,6 %	5,4 %	9,5 %
Invymark	267	17/2/16	26,4 %	21,8 %	20,7 %	16,2 %	4,2 %	10,7 %
Sigma Dos	1000	17/2/16	27,8 %	23,1 %	18,8 %	15,3 %	4,0 %	11,0 %
NC Report	1400	15/2/16	28,9 %	21,4 %	21,1 %	14,2 %	3,4 %	11,0 %
Invymark	267	15/2/16	26,4 %	21,8 %	20,7 %	16,2 %	4,2 %	10,7 %
GAD3	800	10/2/16	27,9 %	20,8 %	21,4 %	14,9 %	3,3 %	11,7 %
Simple Lógica	1048	3/2/16	26,9 %	21,7 %	20,3 %	18,9 %	5,1 %	7,1 %
CELESTETEL	1100	3/2/16	28,3 %	22,2 %	20,2 %	15,0 %	3,6 %	10,7 %
Metroscopia	1000	3/2/16	24,0 %	23,3 %	19,9 %	18,5 %	4,3 %	10,0 %
Invymark	800	1/2/16	27,8 %	21,1 %	21,6 %	15,1 %	3,8 %	10,6 %
My Word	1000	18/1/16	28,0 %	21,3 %	20,3 %	14,4 %	4,7 %	11,3 %
GAD3	802	18/1/16	30,1 %	21,3 %	20,0 %	13,4 %	2,9 %	12,3 %
Metroscopia	1255	13/1/16	29,0 %	21,1 %	22,5 %	16,6 %	3,2 %	7,6 %
NC Report	1900	12/1/16	30,7 %	21,1 %	21,9 %	12,0 %	3,4 %	10,9 %
Invymark	800	11/1/16	30,8 %	20,8 %	21,6 %	12,6 %	3,5 %	10,7 %
CELESTETEL	1100	10/1/16	30,2 %	21,9 %	22,2 %	12,6 %	3,3 %	9,8 %
Llorente &	1000	10/1/16	31,4 %	23,5 %	19,5 %	11,8 %	3,5 %	10,3 %

Promedio de sondeos

Voto según 20 sondeos y un promedio ponderado.

30%

25%

20%

15%

10%

5%

0%

PP

PSOE

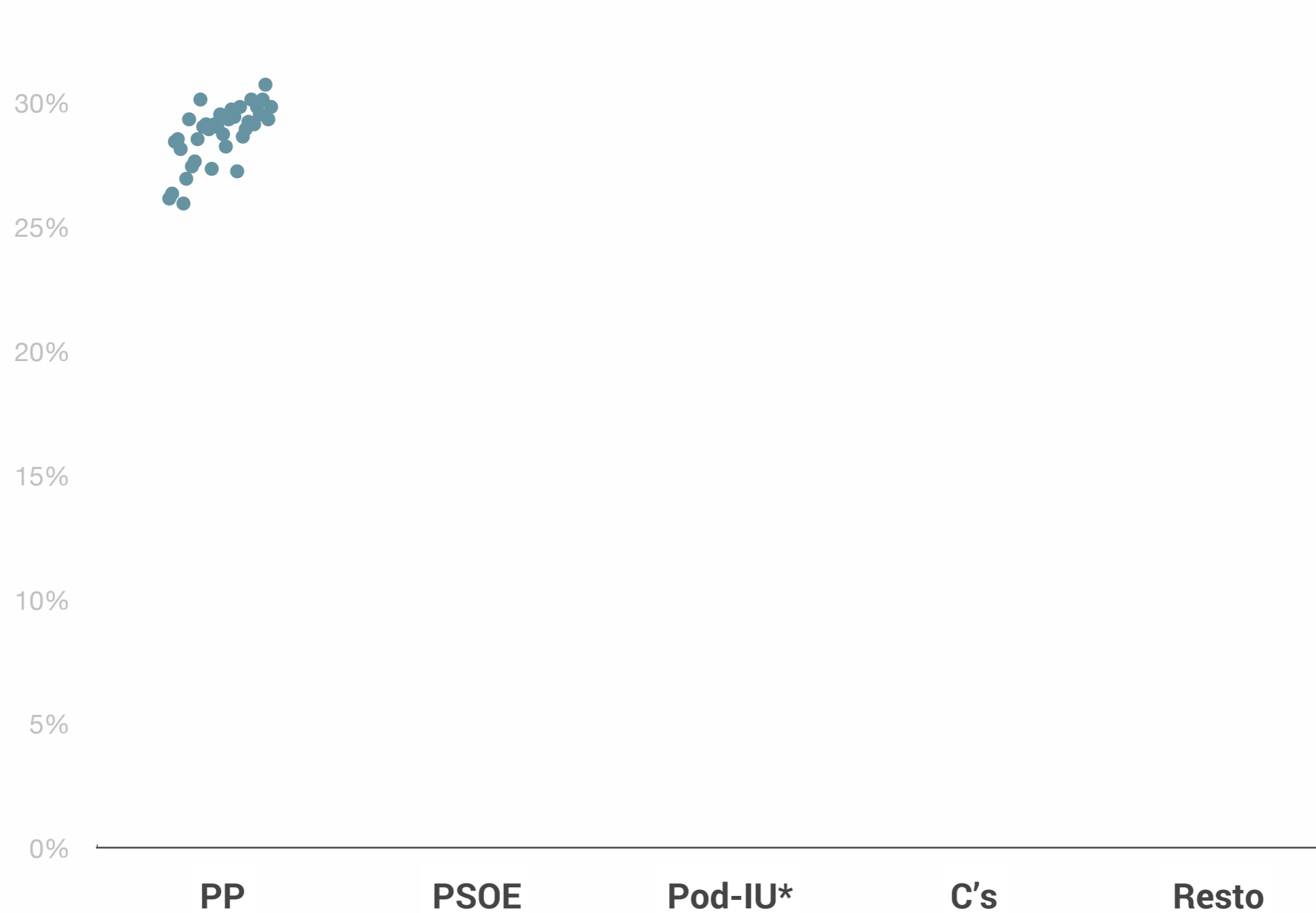
Pod-IU*

C's

Resto

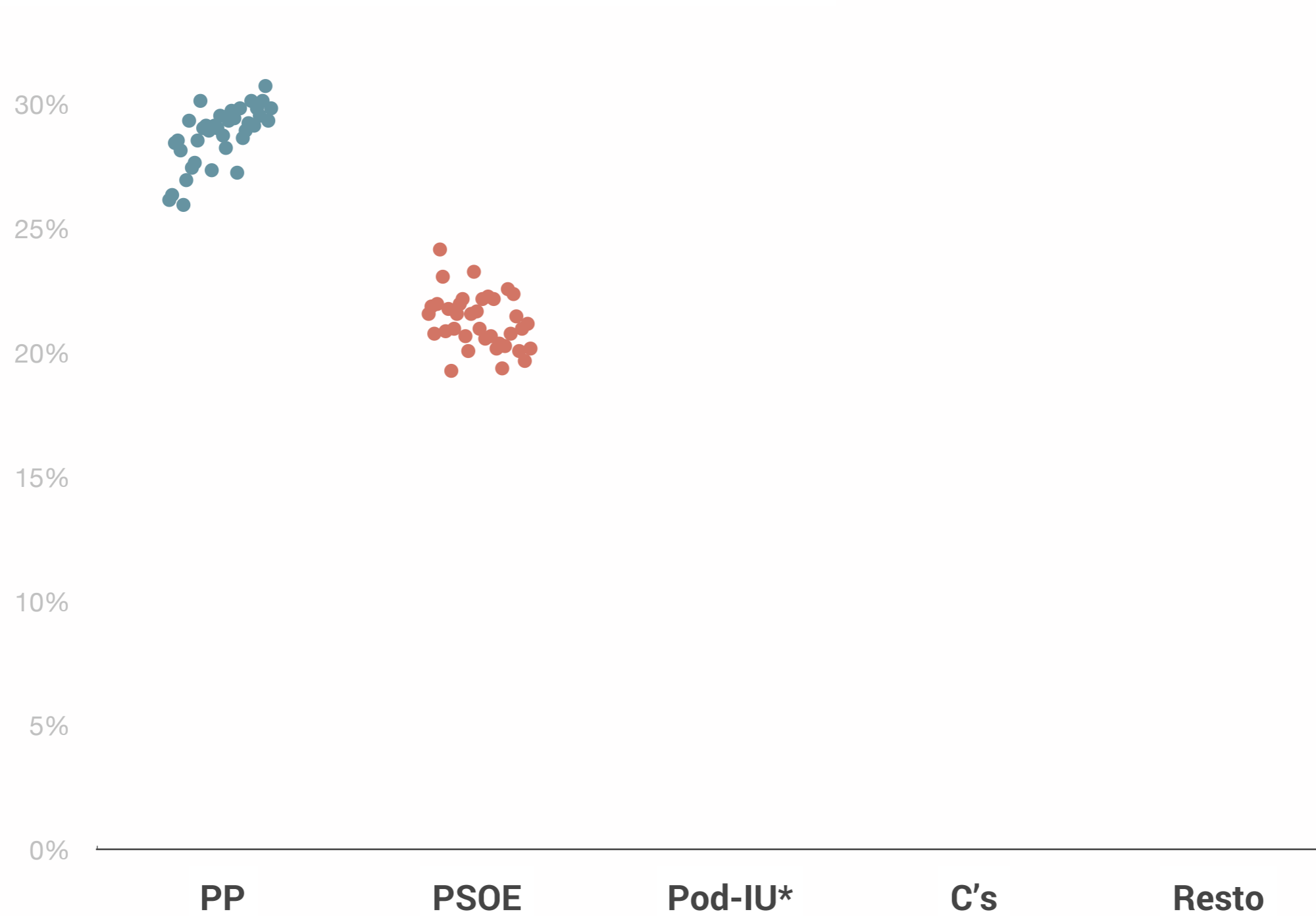
Promedio de sondeos

Voto según 20 sondeos y un promedio ponderado.



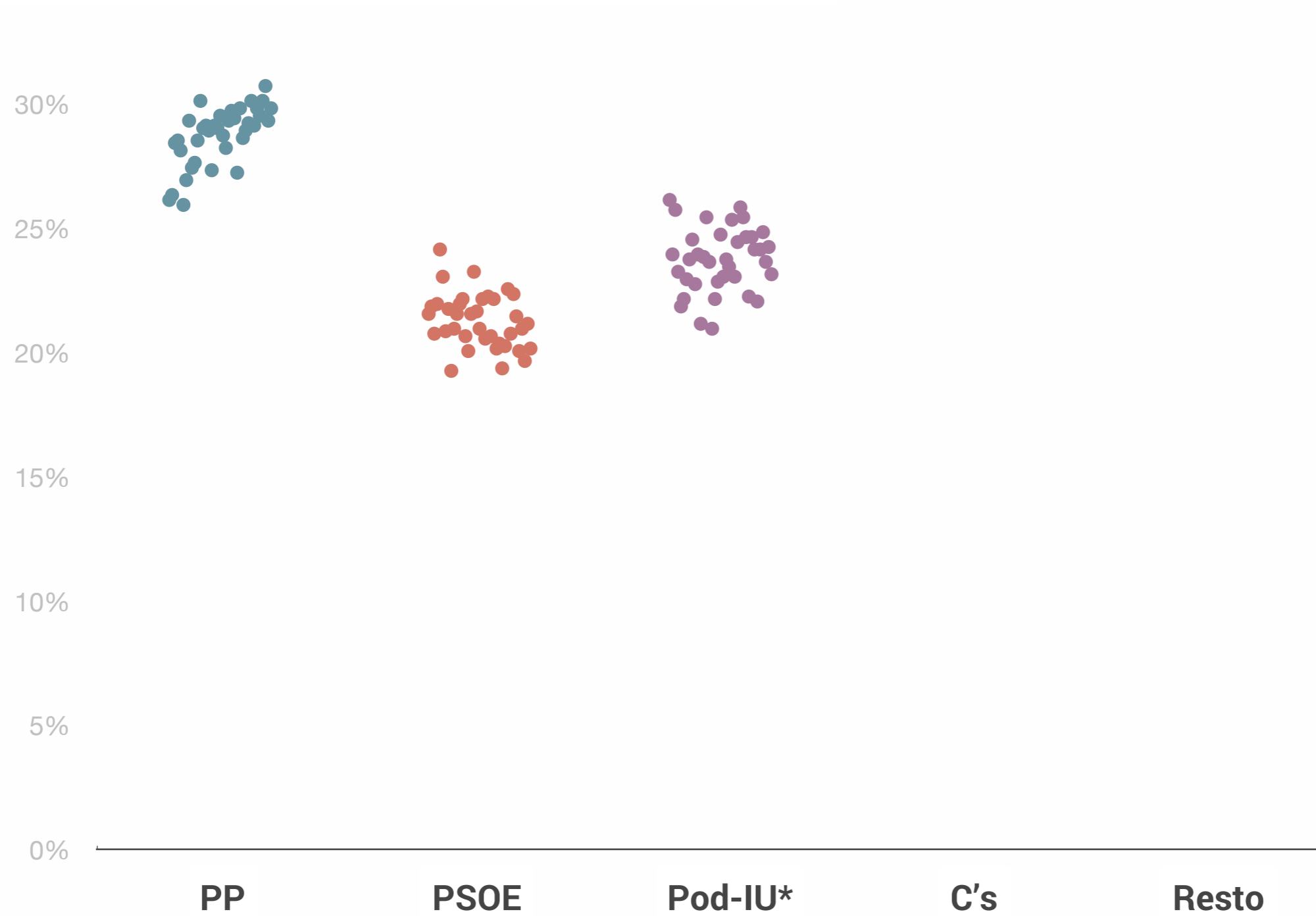
Promedio de sondeos

Voto según 20 sondeos y un promedio ponderado.



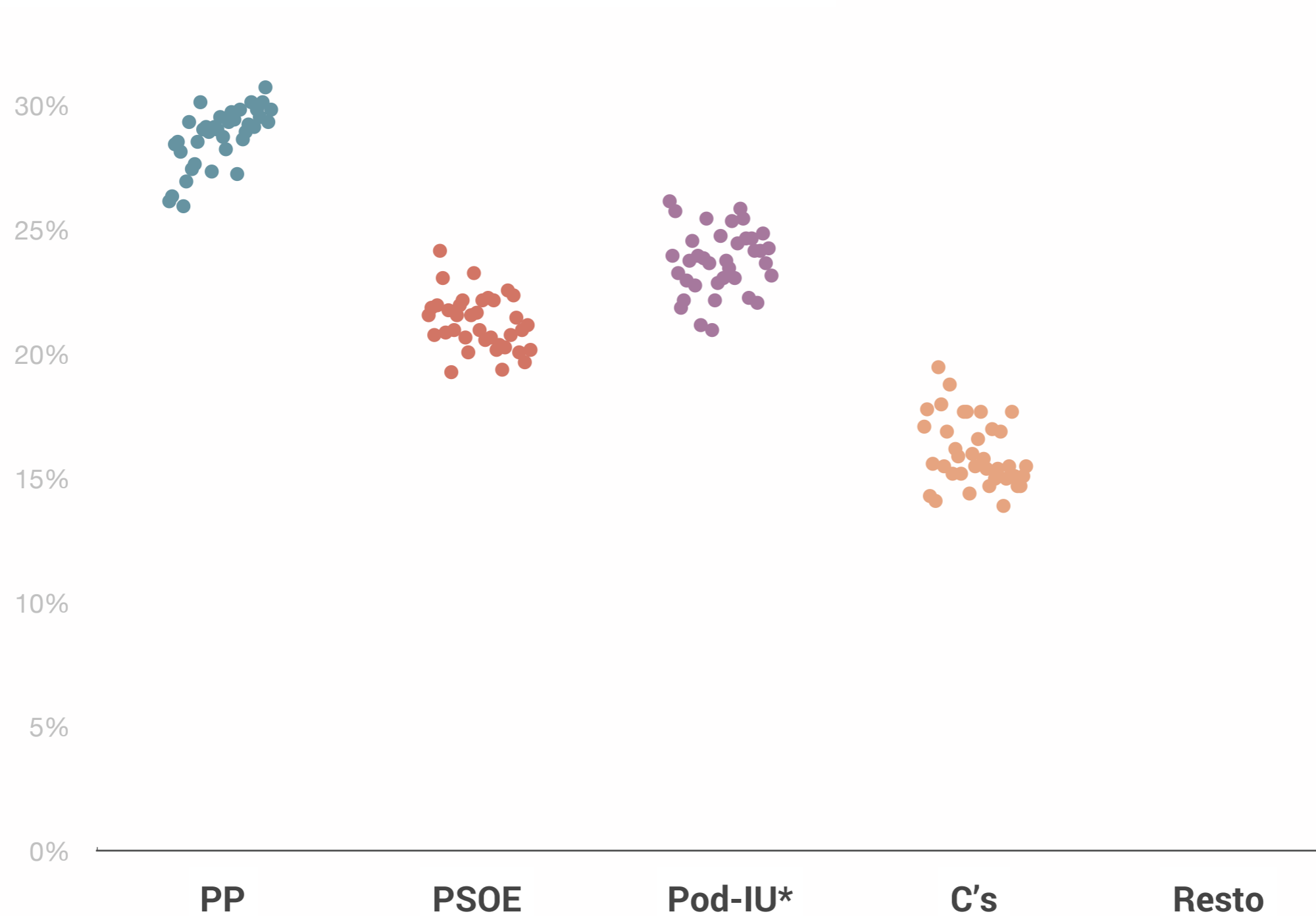
Promedio de sondeos

Voto según 20 sondeos y un promedio ponderado.



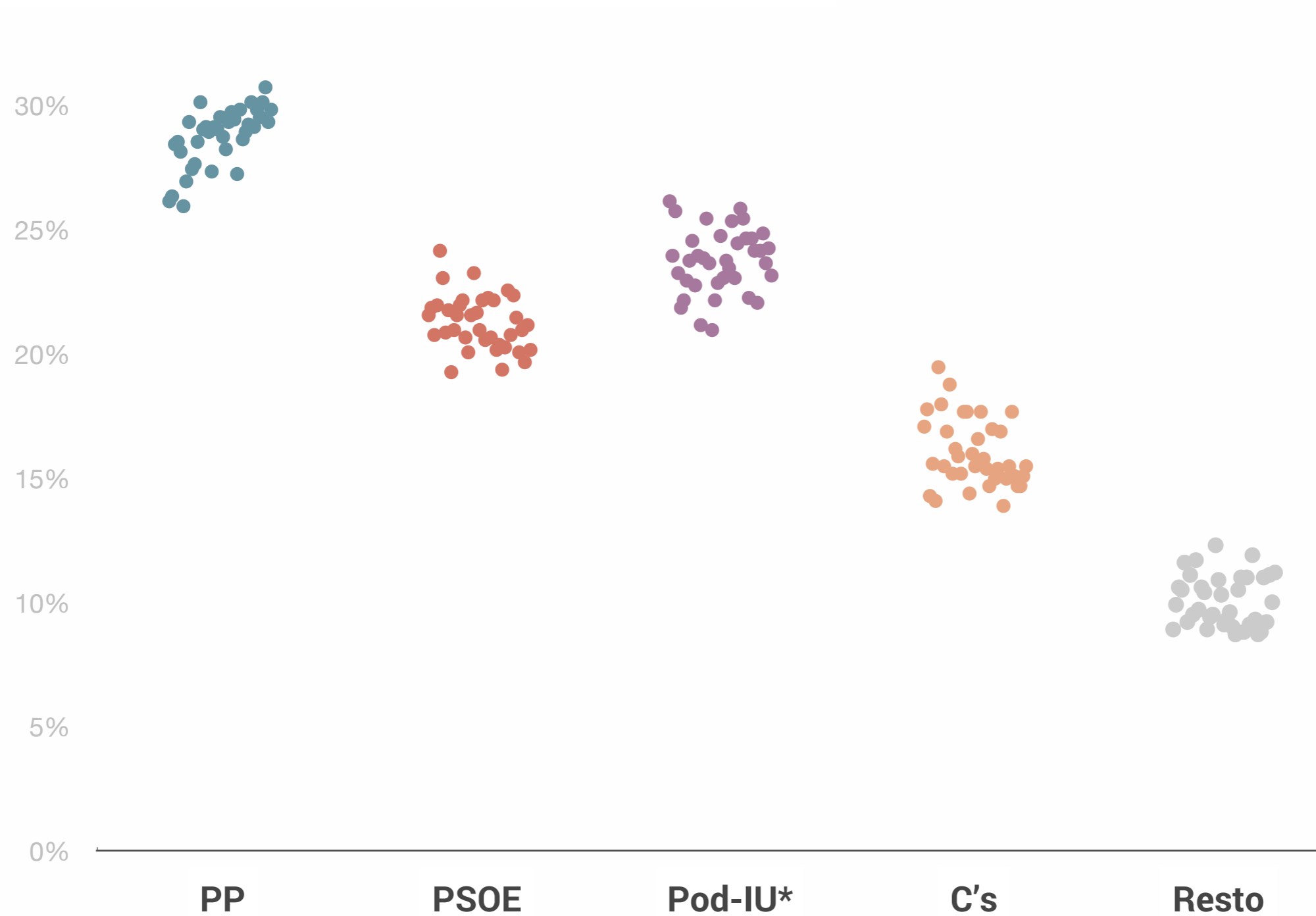
Promedio de sondeos

Voto según 20 sondeos y un promedio ponderado.



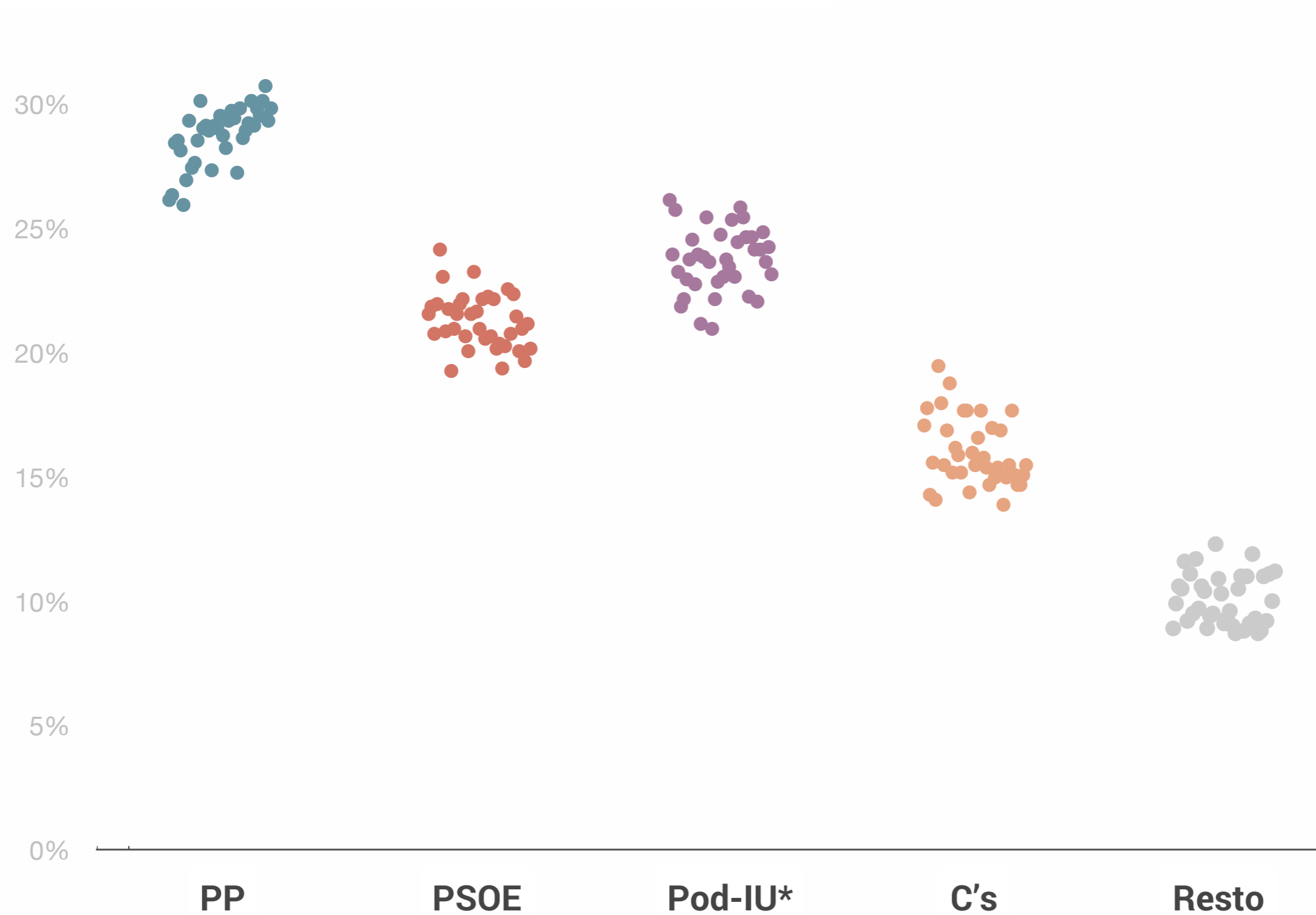
Promedio de sondeos

Voto según 20 sondeos y un promedio ponderado.



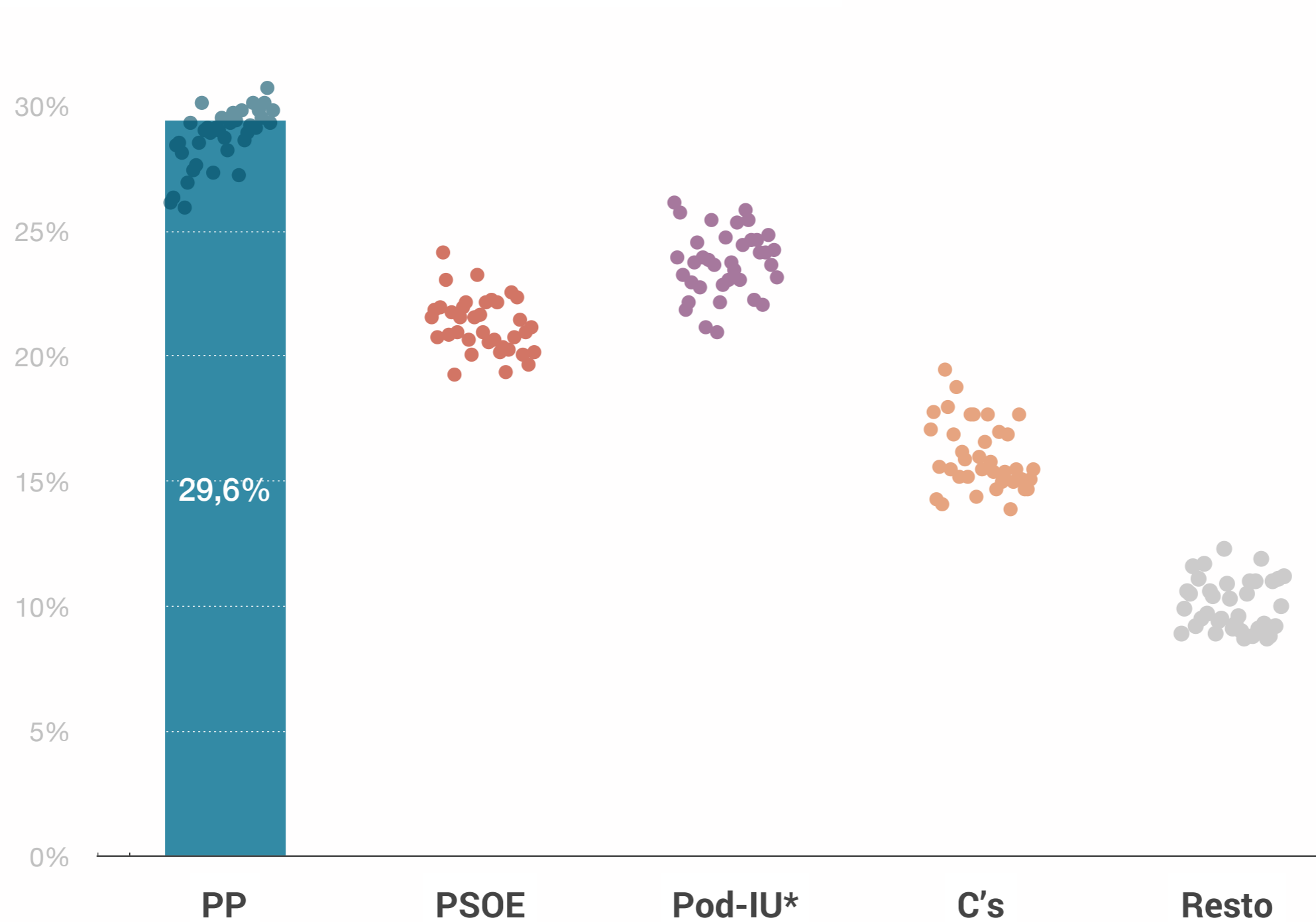
Promedio de sondeos

Voto según 20 sondeos y un promedio ponderado.



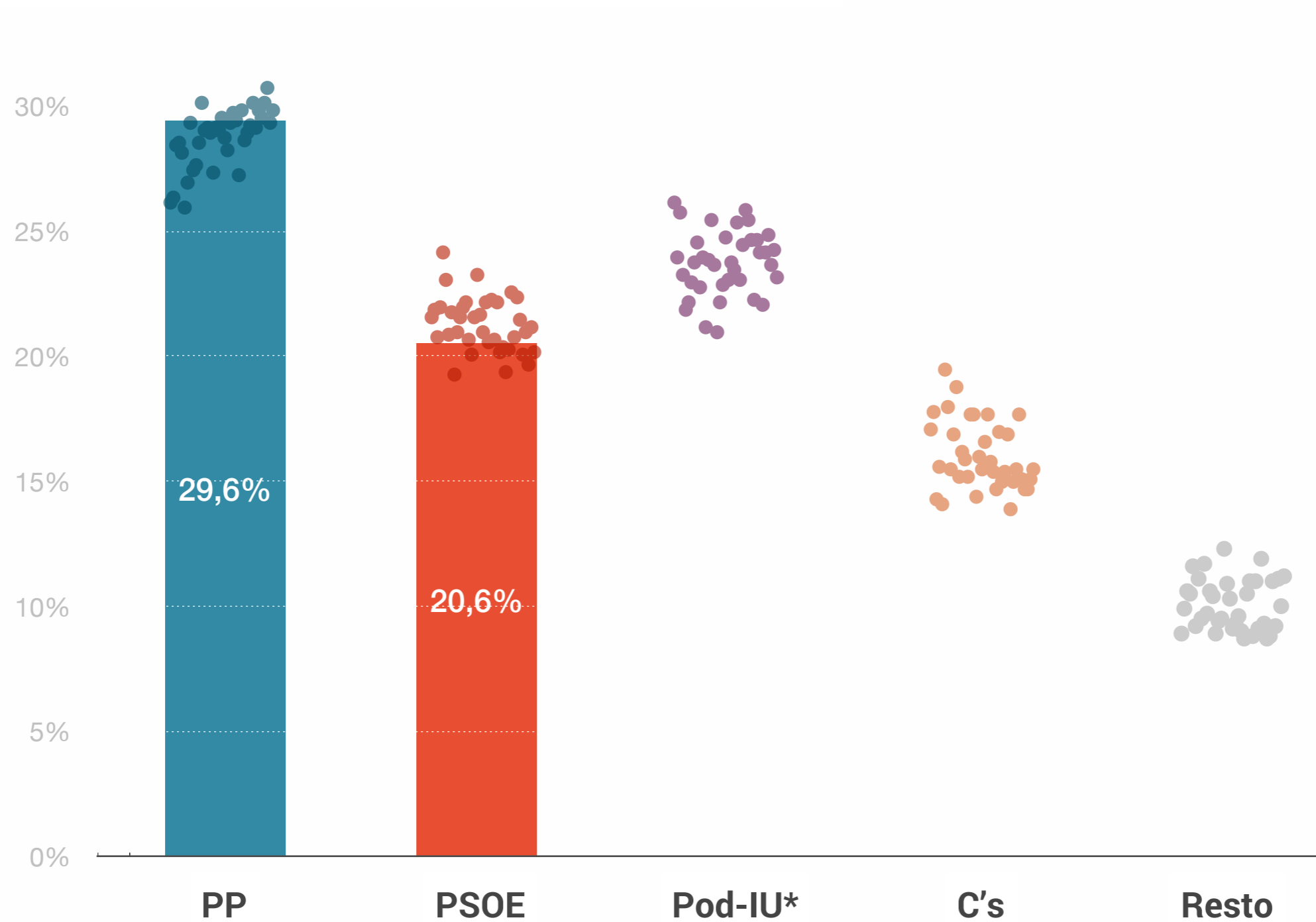
Promedio de sondeos

Voto según 20 sondeos y un promedio ponderado.



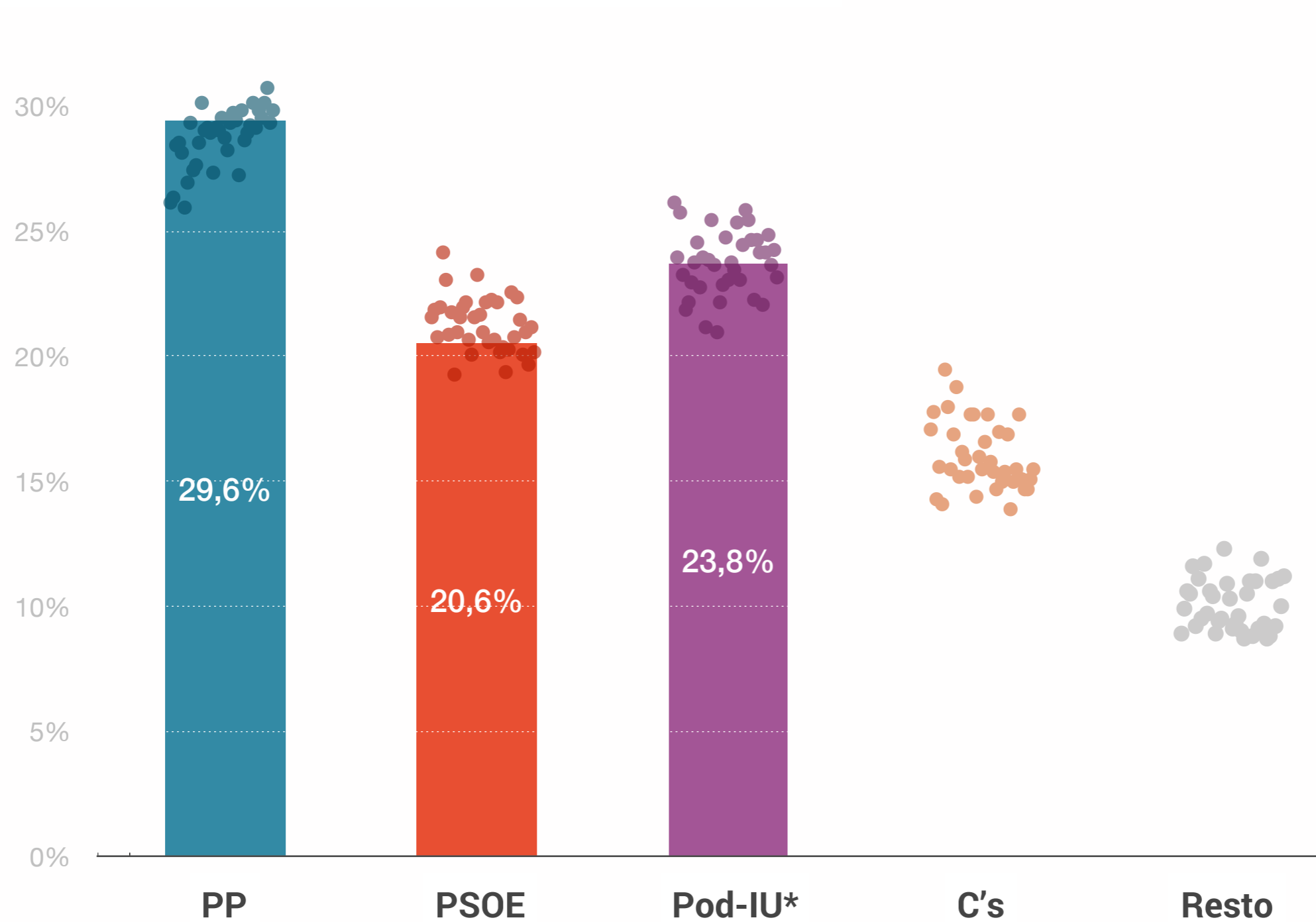
Promedio de sondeos

Voto según 20 sondeos y un promedio ponderado.



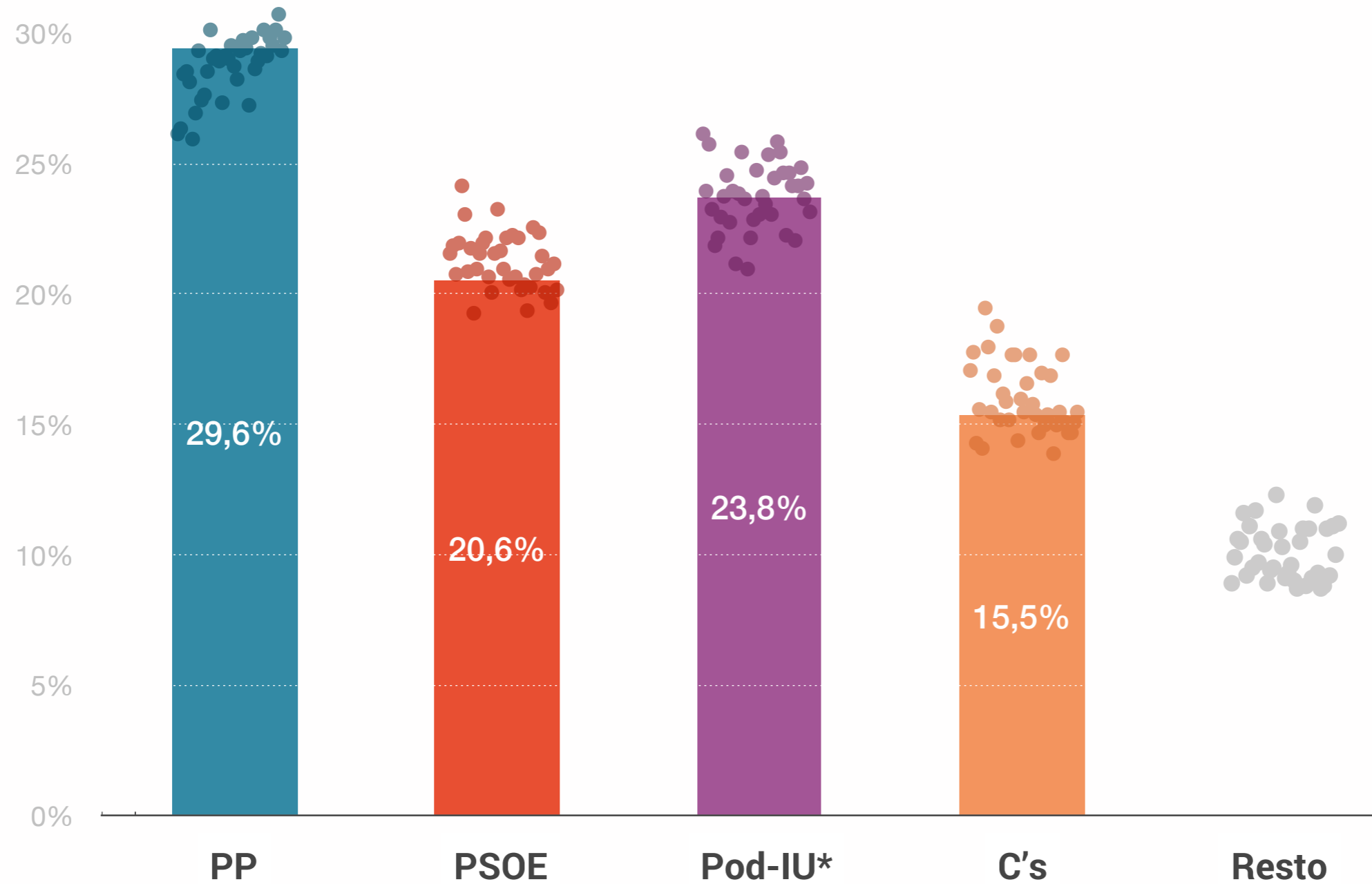
Promedio de sondeos

Voto según 20 sondeos y un promedio ponderado.



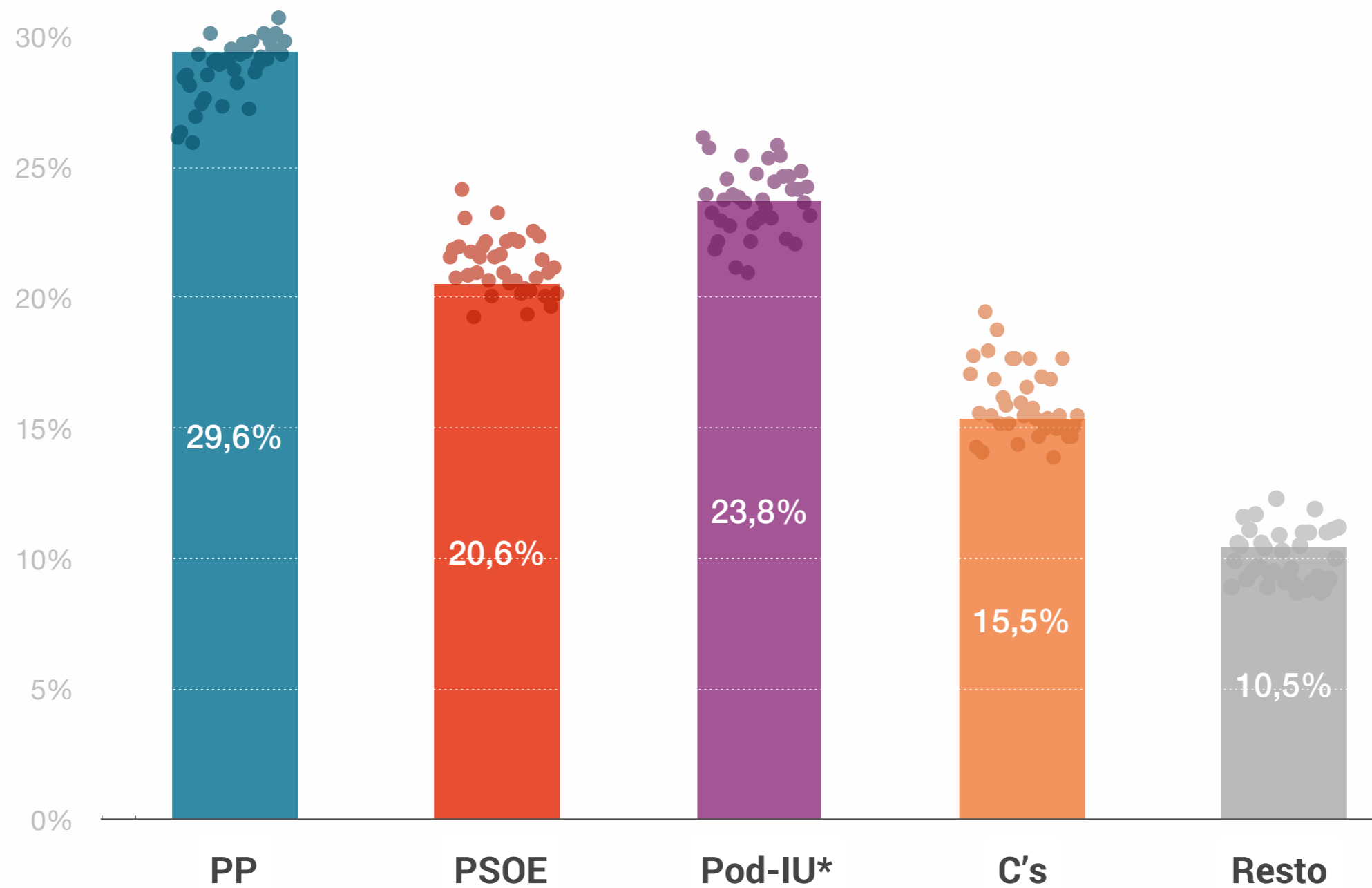
Promedio de sondeos

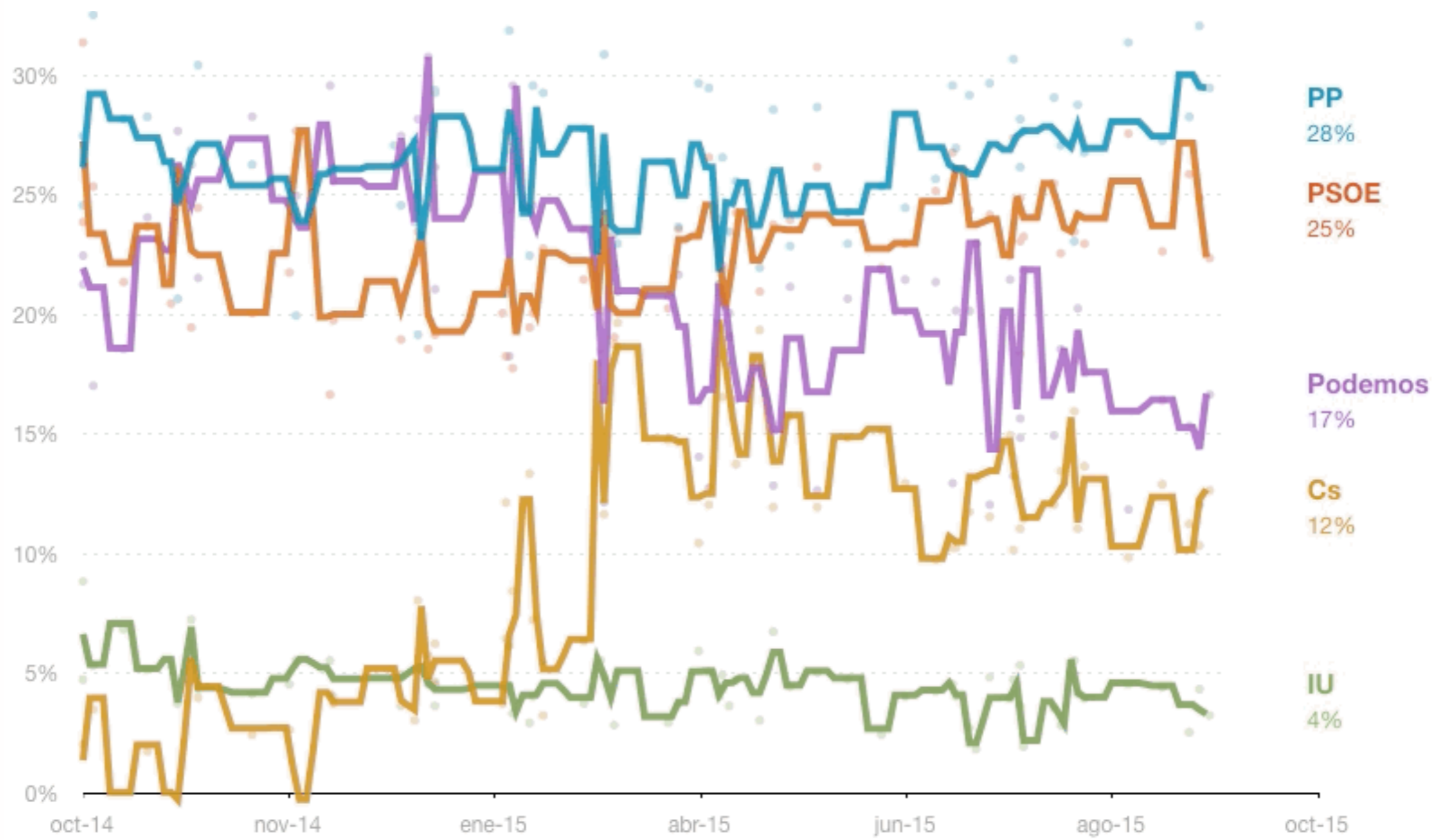
Voto según 20 sondeos y un promedio ponderado.



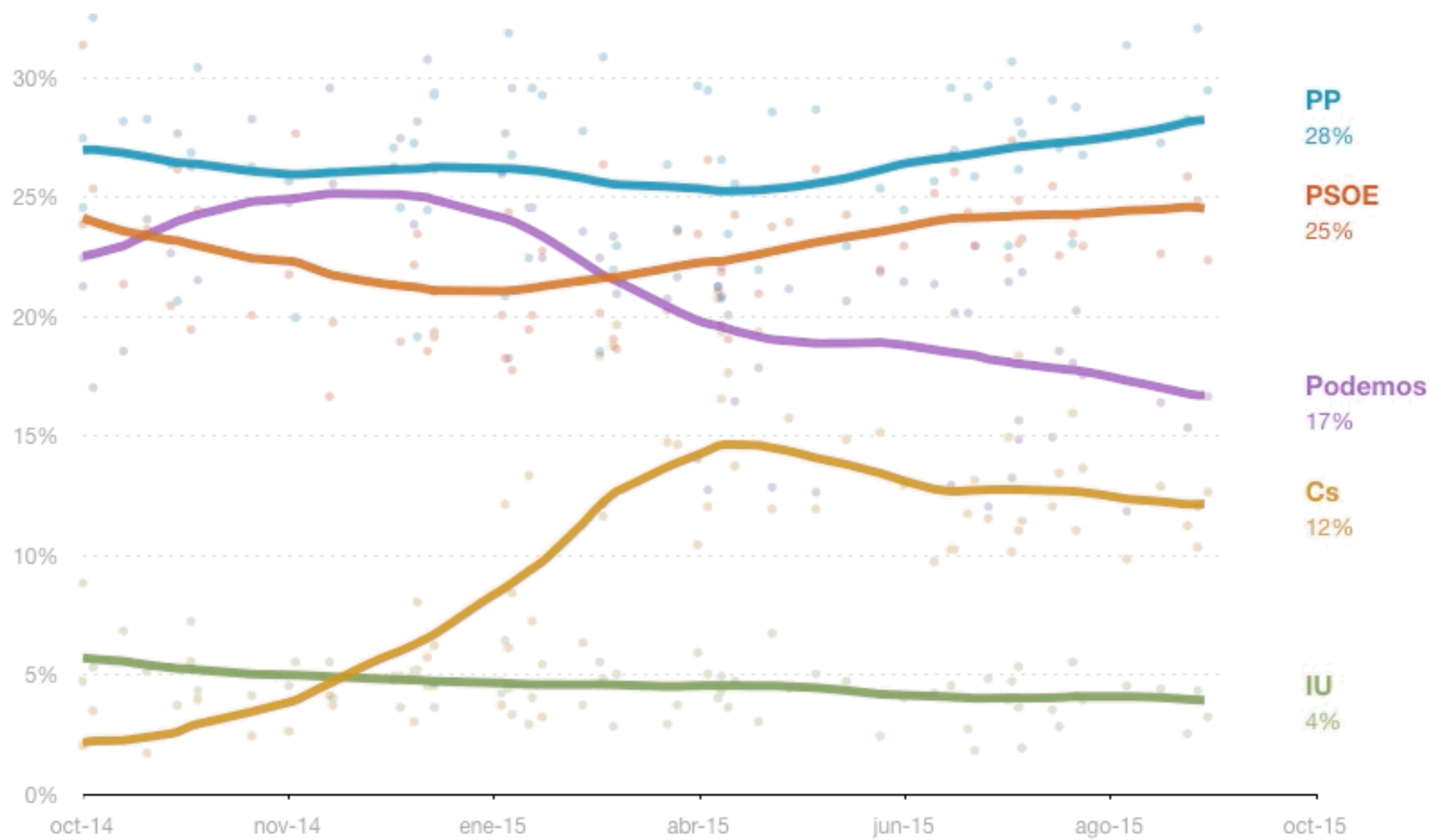
Promedio de sondeos

Voto según 20 sondeos y un promedio ponderado.





Fuente: Wikipedia. Encuestas de Metroscopia, Celestetel, Sigma Dos, Invymark, Simple Lógica, My Word, NC Report, Encuestamos, DYM, GAD3, GESOP, GETS, JM&A, TNS Demoscopia y CIS.

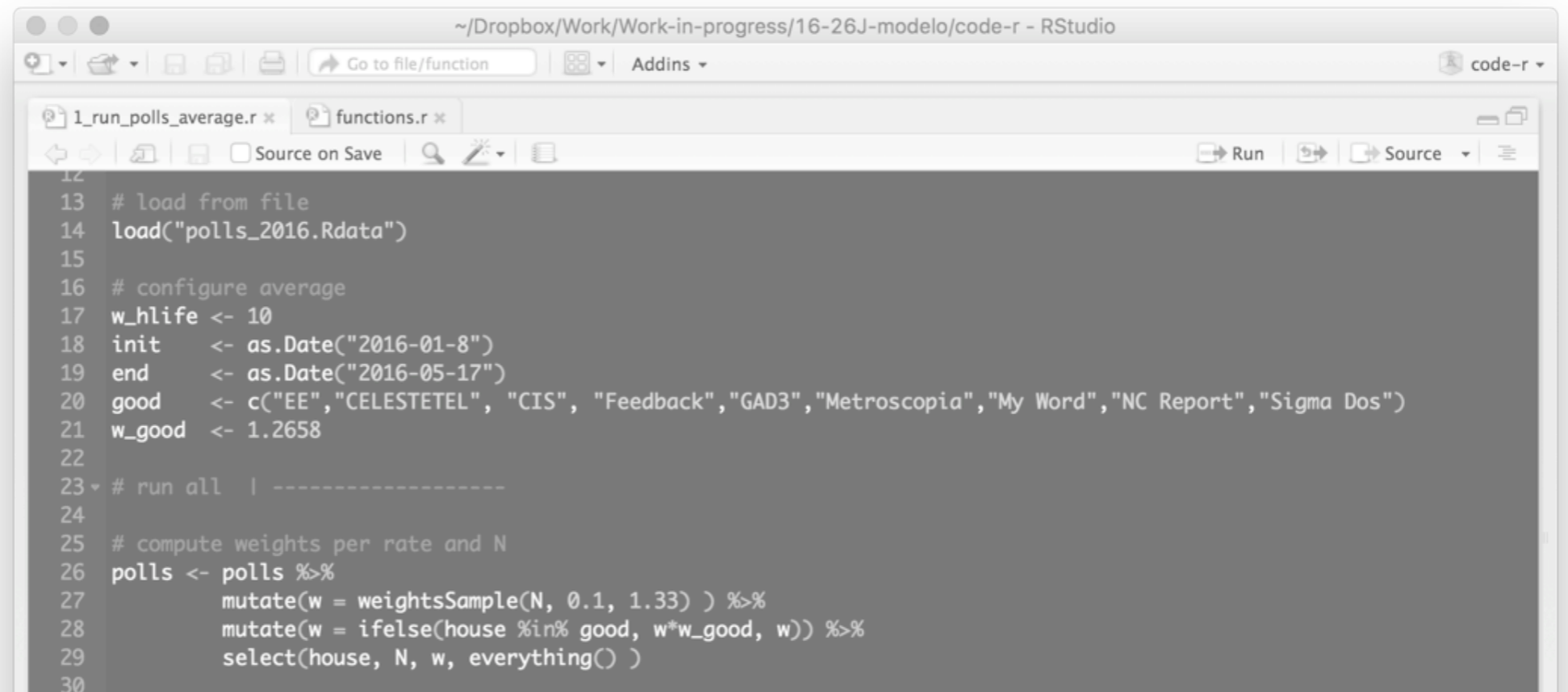


Fuente: Wikipedia. Encuestas de Metroscopia, Celestetel, Sigma Dos, Invymark, Simple Lógica, My Word, NC Report, Encuestamos, DYM, GAD3, GESOP, GETS, JM&A, TNS Demoscopia y CIS.

Kiko Llaneras | EL ESPAÑOL

¿Cómo pondero?

- » Por fecha (*exponential decay*)
- » Por tamaño muestra (entrevistas)
- » Por encuestador (*house effect y rating*)



```
~/Dropbox/Work/Work-in-progress/16-26J-modelo/code-r - RStudio
12
13 # load from file
14 load("polls_2016.Rdata")
15
16 # configure average
17 w_hlife <- 10
18 init <- as.Date("2016-01-8")
19 end <- as.Date("2016-05-17")
20 good <- c("EE", "CELESTETEL", "CIS", "Feedback", "GAD3", "Metroscopia", "My Word", "NC Report", "Sigma Dos")
21 w_good <- 1.2658
22
23 # run all | -----
24
25 # compute weights per rate and N
26 polls <- polls %>%
27   mutate(w = weightsSample(N, 0.1, 1.33) ) %>%
28   mutate(w = ifelse(house %in% good, w*w_good, w)) %>%
29   select(house, N, w, everything() )
30
```

¿Quién va a gobernar?

Tratar de hacer predicciones

En realidad, 52 elecciones.

Problemas

Pocos sondeos provinciales.

Pocos datos de minoritarios.

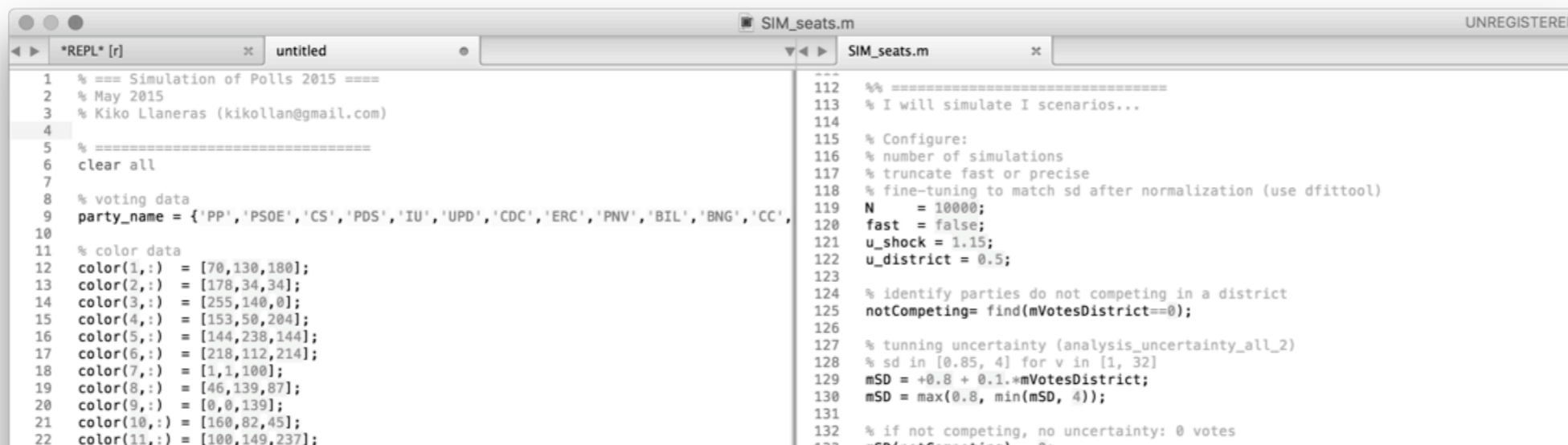
GAD3	800	6/4/16	29,2 %	23,3 %	15,5 %	14,4 %	6,7 %	10,9 %
Metroscopia	800	5/4/16	29,0 %	20,1 %	17,0 %	17,7 %	6,7 %	9,5 %
DYM	1036	4/4/16	29,2 %	20,7 %	19,6 %	15,2 %	5,9 %	9,4 %
CELESTETEL	550	3/4/16	29,1 %	22,2 %	19,4 %	15,9 %	4,5 %	8,9 %
Metroscopia	400	29/3/16	27,7 %	21,0 %	15,9 %	18,8 %	6,9 %	9,7 %
EE	1002	28/3/16	27,5 %	19,3 %	17,3 %	16,9 %	7,3 %	11,7 %
Invymark	1200	30/3/16	28,6 %	21,6 %	19,9 %	15,2 %	4,1 %	10,6 %
Sigma Dos	800	30/3/16	30,2 %	22,0 %	16,7 %	16,2 %	4,5 %	10,4 %
My Word	1004	13/3/16	27,0 %	20,9 %	15,9 %	18,0 %	7,1 %	11,1 %
GAD3	800	9/3/16	28,2 %	24,2 %	18,7 %	14,1 %	3,2 %	11,6 %
NC Report	900	27/3/16	29,4 %	21,8 %	19,2 %	15,5 %	4,6 %	9,5 %
Metroscopia	1200	9/3/16	26,0 %	23,1 %	16,8 %	19,5 %	5,4 %	9,2 %
Invymark	400	8/3/16	28,5 %	20,8 %	21,5 %	14,3 %	4,3 %	10,6 %
Sondaxe	1002	8/3/16	28,6 %	22,0 %	19,1 %	15,6 %	4,2 %	10,5 %
Gesop	1000	7/3/16	26,4 %	21,9 %	18,1 %	17,8 %	5,9 %	9,9 %
Simple Lógica	1206	5/3/16	26,2 %	21,6 %	18,3 %	17,1 %	7,9 %	8,9 %
CELESTETEL	550	4/3/16	28,3 %	22,4 %	19,4 %	15,6 %	4,6 %	9,7 %
NC Report	700	1/3/16	29,2 %	21,8 %	20,2 %	14,9 %	4,2 %	9,7 %
Invymark	400	24/2/16	26,8 %	20,3 %	22,1 %	15,8 %	4,5 %	10,5 %
DYM	1071	21/2/16	25,6 %	22,1 %	19,8 %	17,6 %	5,4 %	9,5 %
Invymark	267	17/2/16	26,4 %	21,8 %	20,7 %	16,2 %	4,2 %	10,7 %
Sigma Dos	1000	17/2/16	27,8 %	23,1 %	18,8 %	15,3 %	4,0 %	11,0 %
NC Report	1400	15/2/16	28,9 %	21,4 %	21,1 %	14,2 %	3,4 %	11,0 %
Invymark	267	15/2/16	26,4 %	21,8 %	20,7 %	16,2 %	4,2 %	10,7 %
GAD3	800	10/2/16	27,9 %	20,8 %	21,4 %	14,9 %	3,3 %	11,7 %
Simple Lógica	1048	3/2/16	26,9 %	21,7 %	20,3 %	18,9 %	5,1 %	7,1 %
CELESTETEL	1100	3/2/16	28,3 %	22,2 %	20,2 %	15,0 %	3,6 %	10,7 %
Metroscopia	1000	3/2/16	24,0 %	23,3 %	19,9 %	18,5 %	4,3 %	10,0 %
Invymark	800	1/2/16	27,8 %	21,1 %	21,6 %	15,1 %	3,8 %	10,6 %
My Word	1000	18/1/16	28,0 %	21,3 %	20,3 %	14,4 %	4,7 %	11,3 %
GAD3	802	18/1/16	30,1 %	21,3 %	20,0 %	13,4 %	2,9 %	12,3 %
Metroscopia	1255	13/1/16	29,0 %	21,1 %	22,5 %	16,6 %	3,2 %	7,6 %
NC Report	1900	12/1/16	30,7 %	21,1 %	21,9 %	12,0 %	3,4 %	10,9 %
Invymark	800	11/1/16	30,8 %	20,8 %	21,6 %	12,6 %	3,5 %	10,7 %
CELESTETEL	1100	10/1/16	30,2 %	21,9 %	22,2 %	12,6 %	3,3 %	9,8 %
I lorente &	1000	10/1/16	31,4 %	23,5 %	19,5 %	11,8 %	3,5 %	10,3 %

Modelo de predicción electoral

```
SIM_seats.m UNREGISTERED
*REPL* [r] x untitle SIM_seats.m x
1 % == Simulation of Polls 2015 ==
2 % May 2015
3 % Kiko Llaneras (kikollan@gmail.com)
4
5 % =====
6 clear all
7
8 % voting data
9 party_name = {'PP', 'PSOE', 'CS', 'PDS', 'IU', 'UPD', 'CDC', 'ERC', 'PNV', 'BIL', 'BNG', 'CC',
10
11 % color data
12 color(1,:) = [70,130,180];
13 color(2,:) = [178,34,34];
14 color(3,:) = [255,140,0];
15 color(4,:) = [153,50,204];
16 color(5,:) = [144,238,144];
17 color(6,:) = [218,112,214];
18 color(7,:) = [1,1,100];
19 color(8,:) = [46,139,87];
20 color(9,:) = [0,0,139];
21 color(10,:) = [160,82,45];
22 color(11,:) = [100,149,237];
---
112 %% =====
113 % I will simulate I scenarios...
114
115 % Configure:
116 % number of simulations
117 % truncate fast or precise
118 % fine-tuning to match sd after normalization (use dfittool)
119 N = 10000;
120 fast = false;
121 u_shock = 1.15;
122 u_district = 0.5;
123
124 % identify parties do not competing in a district
125 notCompeting= find(mVotesDistrict==0);
126
127 % tuning uncertainty (analysis_uncertainty_all_2)
128 % sd in [0.85, 4] for v in [1, 32]
129 mSD = +0.8 + 0.1.*mVotesDistrict;
130 mSD = max(0.8, min(mSD, 4));
131
132 % if not competing, no uncertainty: 0 votes
133 mSD(notCompeting) = 0;
```

Modelo de predicción electoral

1. Promediar sondeos (a nivel nacional)



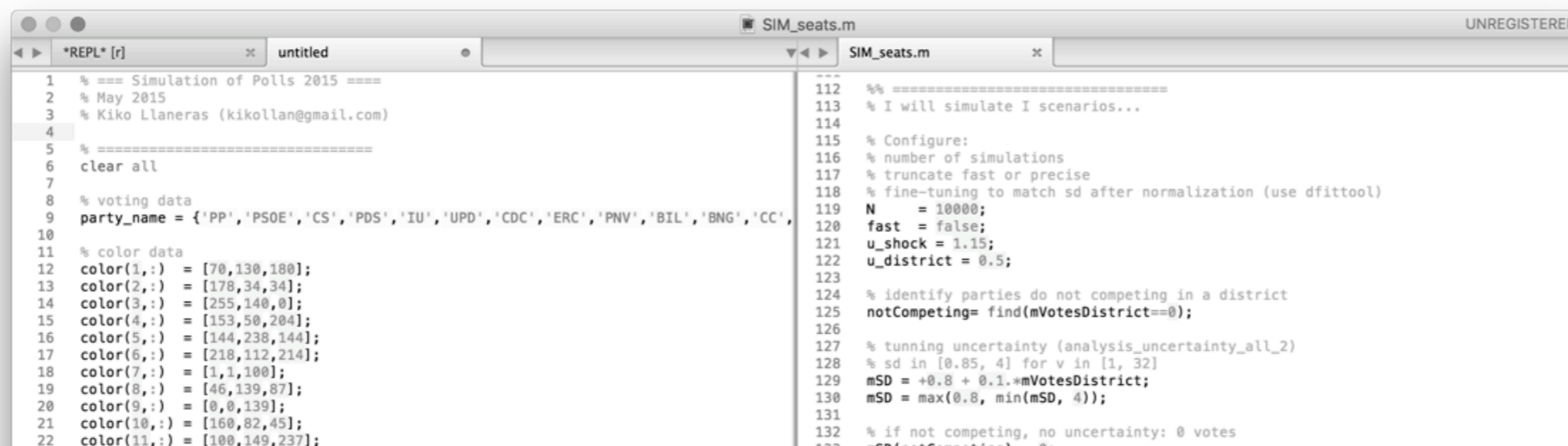
```
1 % == Simulation of Polls 2015 ==
2 % May 2015
3 % Kiko Llaneras (kikollan@gmail.com)
4
5 % =====
6 clear all
7
8 % voting data
9 party_name = {'PP', 'PSOE', 'CS', 'PDS', 'IU', 'UPD', 'CDC', 'ERC', 'PNV', 'BIL', 'BNG', 'CC',
10
11 % color data
12 color(1,:) = [70,130,180];
13 color(2,:) = [178,34,34];
14 color(3,:) = [255,140,0];
15 color(4,:) = [153,50,204];
16 color(5,:) = [144,238,144];
17 color(6,:) = [218,112,214];
18 color(7,:) = [1,1,100];
19 color(8,:) = [46,139,87];
20 color(9,:) = [0,0,139];
21 color(10,:) = [160,82,45];
22 color(11,:) = [100,149,237];
---
```

```
112 %% =====
113 % I will simulate I scenarios...
114
115 % Configure:
116 % number of simulations
117 % truncate fast or precise
118 % fine-tuning to match sd after normalization (use dfittool)
119 N = 10000;
120 fast = false;
121 u_shock = 1.15;
122 u_district = 0.5;
123
124 % identify parties do not competing in a district
125 notCompeting= find(mVotesDistrict==0);
126
127 % tuning uncertainty (analysis_uncertainty_all_2)
128 % sd in [0.85, 4] for v in [1, 32]
129 mSD = +0.8 + 0.1.*mVotesDistrict;
130 mSD = max(0.8, min(mSD, 4));
131
132 % if not competing, no uncertainty: 0 votes
133 mSD(notCompeting) = 0;
```

Modelo de predicción electoral

1. Promediar sondeos (a nivel nacional)

2. Aproximar voto en provincias (con históricos)

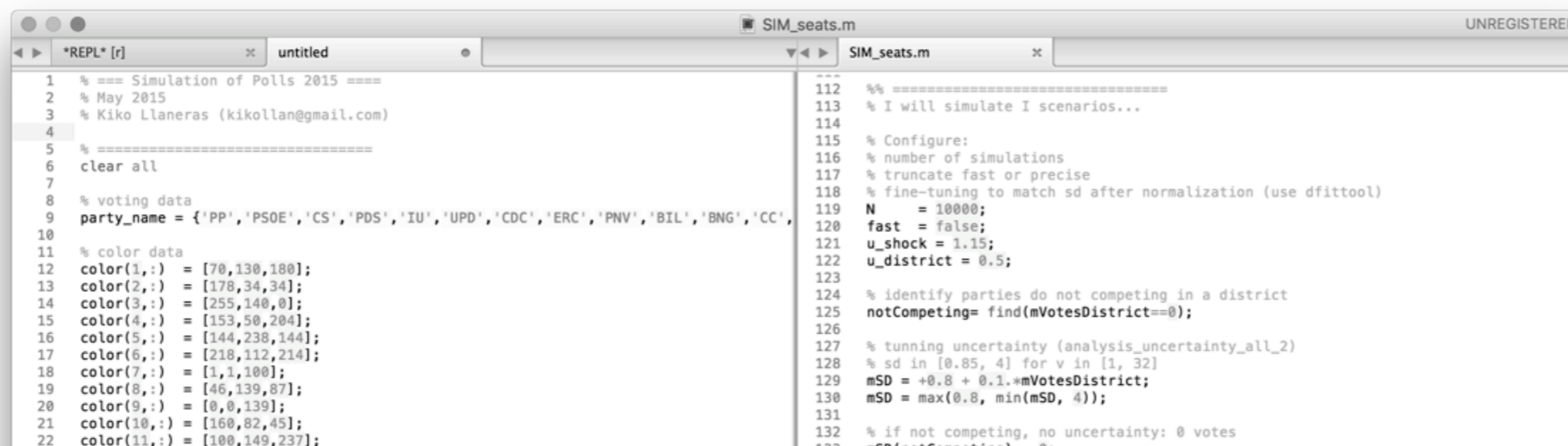


```
1 % == Simulation of Polls 2015 ==
2 % May 2015
3 % Kiko Llaneras (kikollan@gmail.com)
4
5 % =====
6 clear all
7
8 % voting data
9 party_name = {'PP', 'PSOE', 'CS', 'PDS', 'IU', 'UPD', 'CDC', 'ERC', 'PNV', 'BIL', 'BNG', 'CC',
10
11 % color data
12 color(1,:) = [70,130,180];
13 color(2,:) = [178,34,34];
14 color(3,:) = [255,140,0];
15 color(4,:) = [153,50,204];
16 color(5,:) = [144,238,144];
17 color(6,:) = [218,112,214];
18 color(7,:) = [1,1,100];
19 color(8,:) = [46,139,87];
20 color(9,:) = [0,0,139];
21 color(10,:) = [160,82,45];
22 color(11,:) = [100,149,237];
---
```

```
112 %% =====
113 % I will simulate I scenarios...
114
115 % Configure:
116 % number of simulations
117 % truncate fast or precise
118 % fine-tuning to match sd after normalization (use dfittool)
119 N = 10000;
120 fast = false;
121 u_shock = 1.15;
122 u_district = 0.5;
123
124 % identify parties do not competing in a district
125 notCompeting= find(mVotesDistrict==0);
126
127 % tuning uncertainty (analysis_uncertainty_all_2)
128 % sd in [0.85, 4] for v in [1, 32]
129 mSD = +0.8 + 0.1.*mVotesDistrict;
130 mSD = max(0.8, min(mSD, 4));
131
132 % if not competing, no uncertainty: 0 votes
133
```

Modelo de predicción electoral

1. Promediar sondeos (a nivel nacional)
2. Aproximar voto en provincias (con históricos)
3. Añadir incertidumbre a los votos (calibrada con históricos)

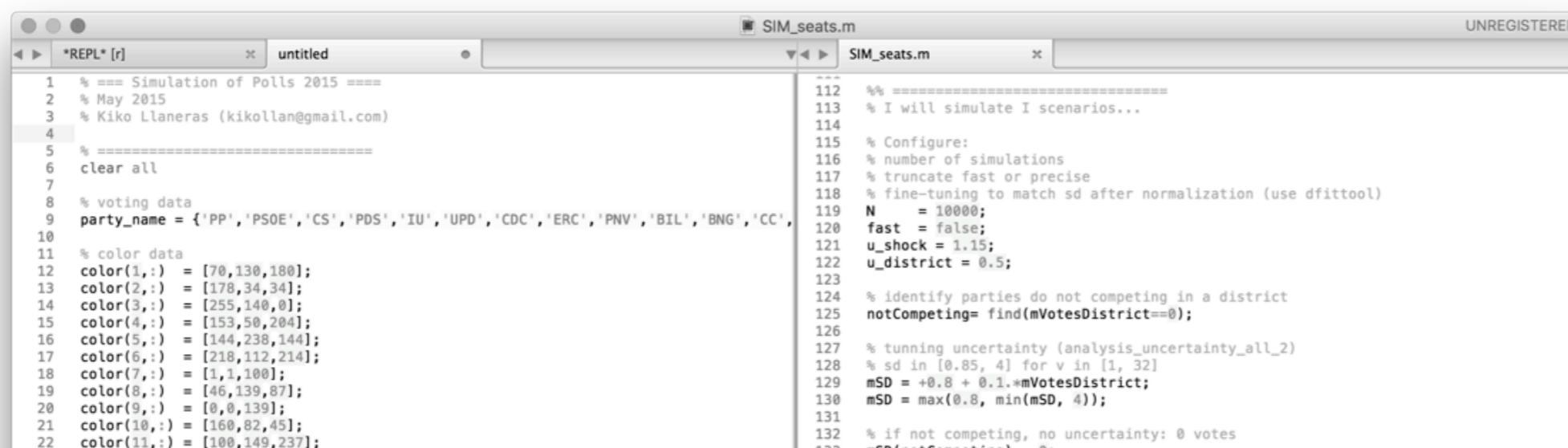


```
1 % == Simulation of Polls 2015 ==
2 % May 2015
3 % Kiko Llaneras (kikollan@gmail.com)
4
5 % =====
6 clear all
7
8 % voting data
9 party_name = {'PP', 'PSOE', 'CS', 'PDS', 'IU', 'UPD', 'CDC', 'ERC', 'PNV', 'BIL', 'BNG', 'CC',
10
11 % color data
12 color(1,:) = [70,130,180];
13 color(2,:) = [178,34,34];
14 color(3,:) = [255,140,0];
15 color(4,:) = [153,50,204];
16 color(5,:) = [144,238,144];
17 color(6,:) = [218,112,214];
18 color(7,:) = [1,1,100];
19 color(8,:) = [46,139,87];
20 color(9,:) = [0,0,139];
21 color(10,:) = [160,82,45];
22 color(11,:) = [100,149,237];
---
```

```
112 %% =====
113 % I will simulate I scenarios...
114
115 % Configure:
116 % number of simulations
117 % truncate fast or precise
118 % fine-tuning to match sd after normalization (use dfittool)
119 N = 10000;
120 fast = false;
121 u_shock = 1.15;
122 u_district = 0.5;
123
124 % identify parties do not competing in a district
125 notCompeting= find(mVotesDistrict==0);
126
127 % tuning uncertainty (analysis_uncertainty_all_2)
128 % sd in [0.85, 4] for v in [1, 32]
129 mSD = +0.8 + 0.1.*mVotesDistrict;
130 mSD = max(0.8, min(mSD, 4));
131
132 % if not competing, no uncertainty: 0 votes
133
```

Modelo de predicción electoral

1. Promediar sondeos (a nivel nacional)
2. Aproximar voto en provincias (con históricos)
3. Añadir incertidumbre a los votos (calibrada con históricos)
4. Simular 15.000 elecciones (d'hondt en cada provincia)

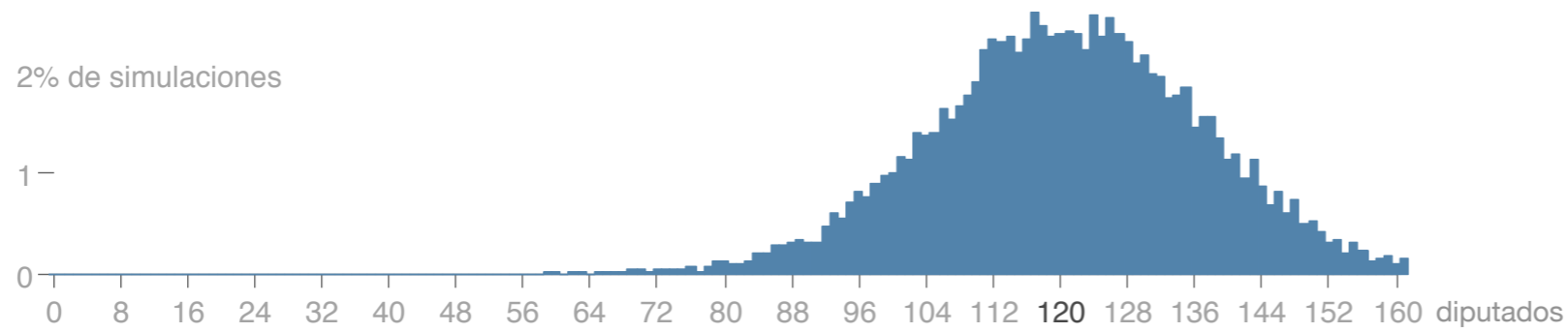


```
1 % == Simulation of Polls 2015 ==
2 % May 2015
3 % Kiko Llaneras (kikollan@gmail.com)
4
5 % =====
6 clear all
7
8 % voting data
9 party_name = {'PP', 'PSOE', 'CS', 'PDS', 'IU', 'UPD', 'CDC', 'ERC', 'PNV', 'BIL', 'BNG', 'CC',
10
11 % color data
12 color(1,:) = [70,130,180];
13 color(2,:) = [178,34,34];
14 color(3,:) = [255,140,0];
15 color(4,:) = [153,50,204];
16 color(5,:) = [144,238,144];
17 color(6,:) = [218,112,214];
18 color(7,:) = [1,1,100];
19 color(8,:) = [46,139,87];
20 color(9,:) = [0,0,139];
21 color(10,:) = [160,82,45];
22 color(11,:) = [100,149,237];
112 % =====
113 % I will simulate I scenarios...
114
115 % Configure:
116 % number of simulations
117 % truncate fast or precise
118 % fine-tuning to match sd after normalization (use dfittool)
119 N = 10000;
120 fast = false;
121 u_shock = 1.15;
122 u_district = 0.5;
123
124 % identify parties do not competing in a district
125 notCompeting= find(mVotesDistrict==0);
126
127 % tuning uncertainty (analysis_uncertainty_all_2)
128 % sd in [0.85, 4] for v in [1, 32]
129 mSD = +0.8 + 0.1.*mVotesDistrict;
130 mSD = max(0.8, min(mSD, 4));
131
132 % if not competing, no uncertainty: 0 votes
133 mSD(notCompeting) = 0;
```



Estimación de escaños

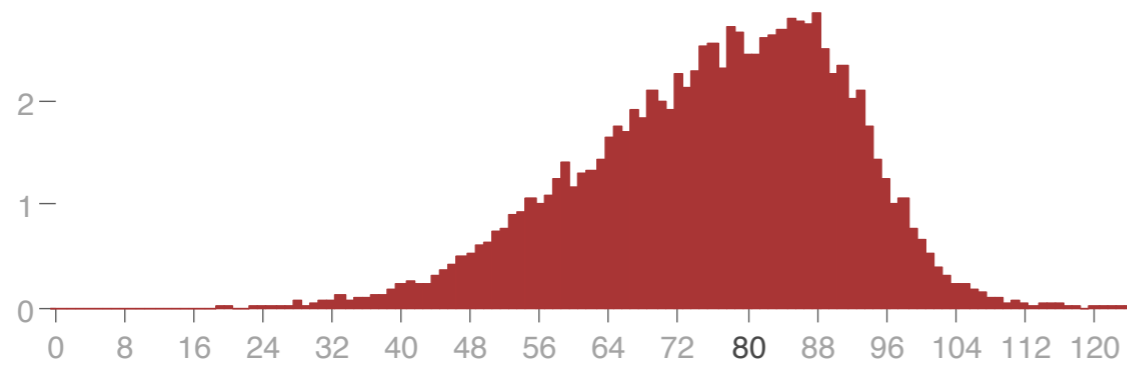
Estimación de escaños en 15.000 simulaciones.



PP

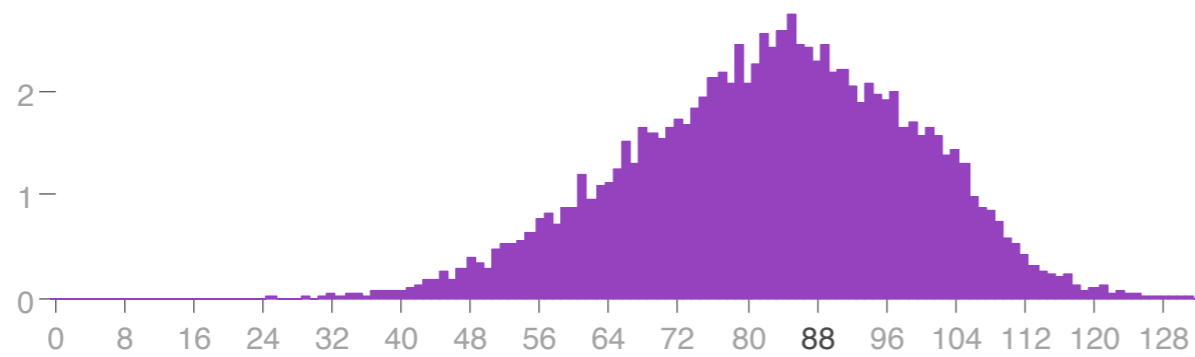
110-132 diputados

(Intervalo del 50% de probabilidad)



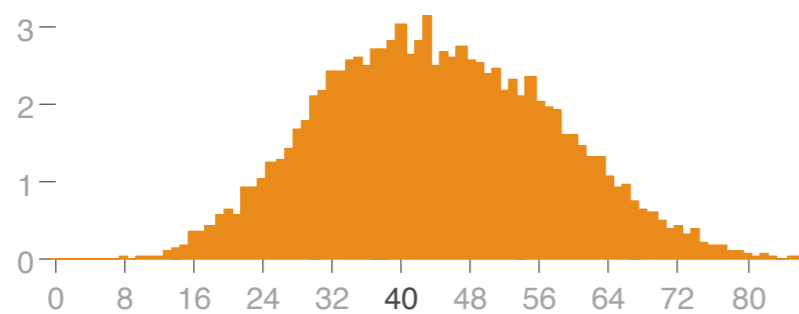
PSOE

67-87



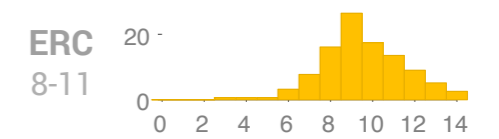
U. Podemos

72-95

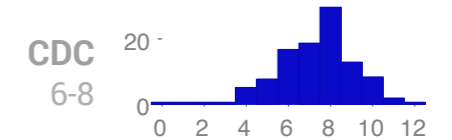


Ciudadanos

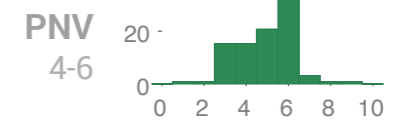
34-54



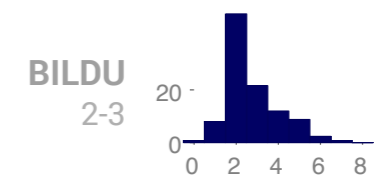
ERC
8-11



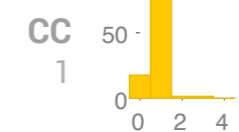
CDC
6-8



PNV
4-6

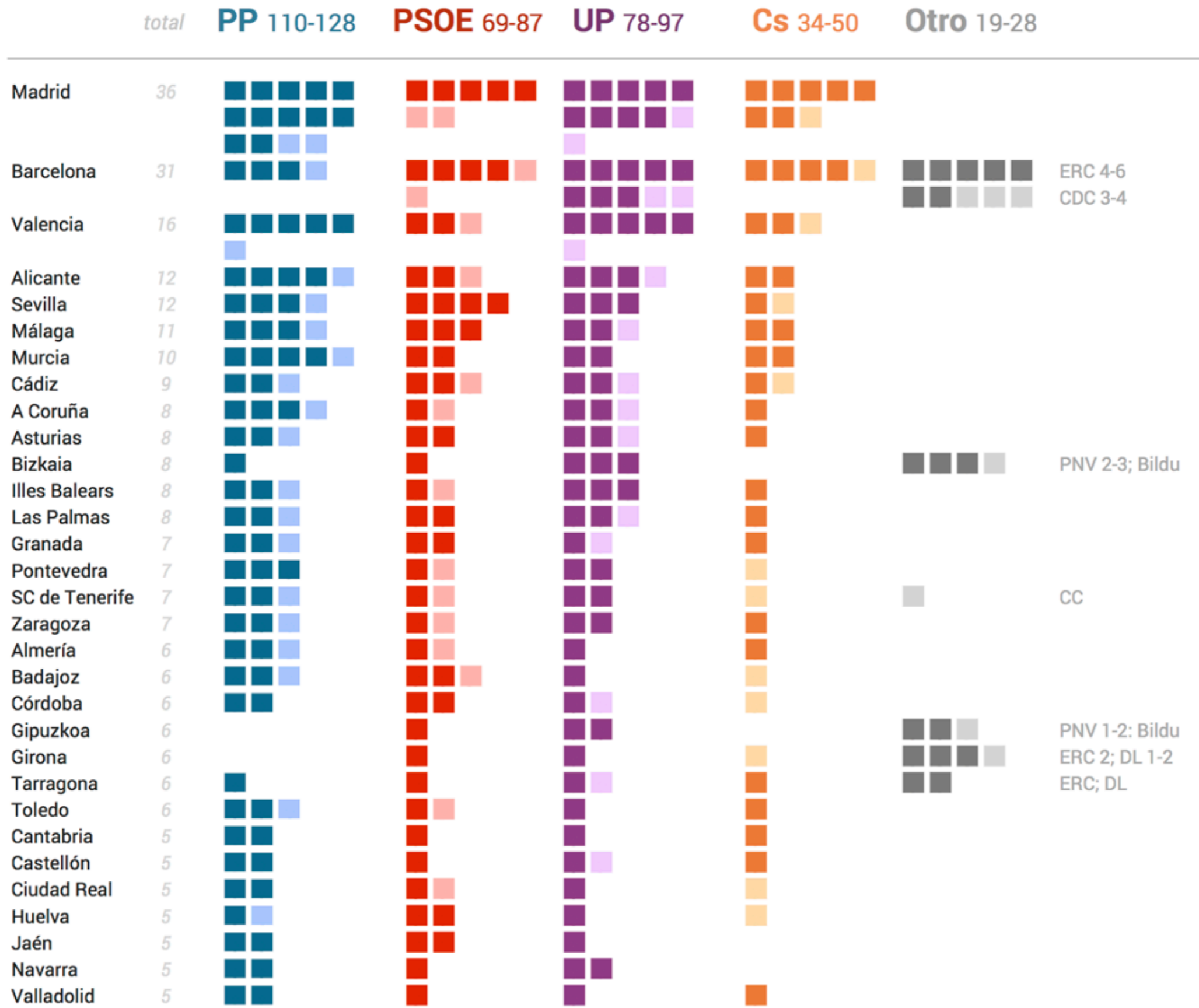


BILDU
2-3



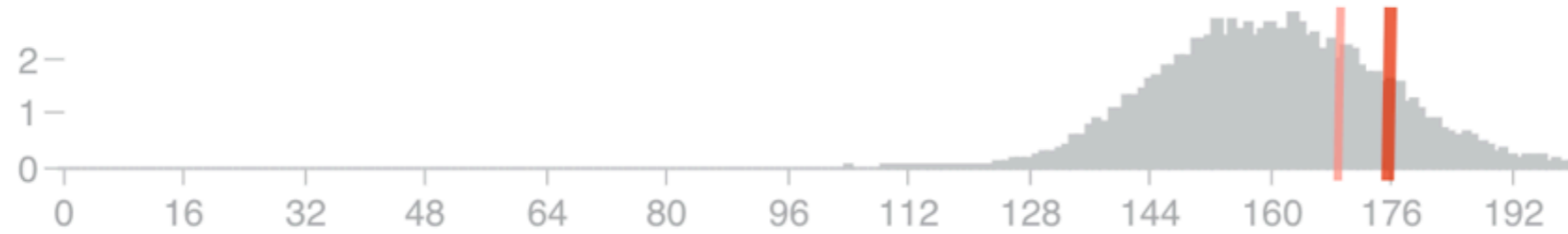
CC
1

The likely seats are in dark (75% chance); possible ones in light (around 50%).

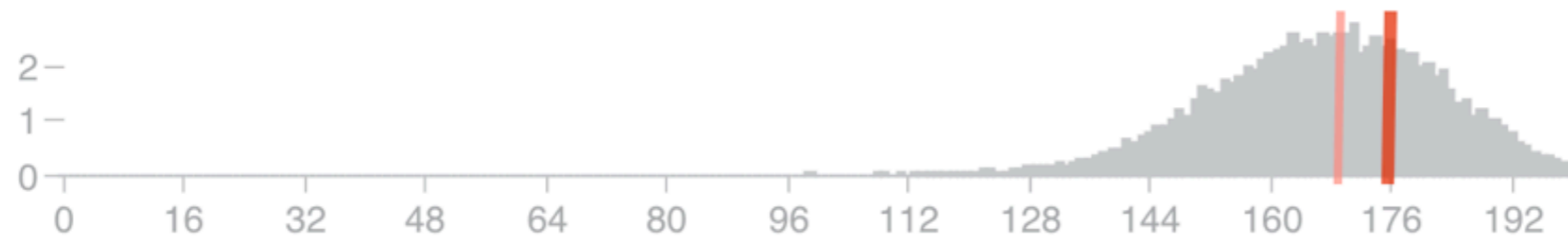


Possible coalitions with 176 seats

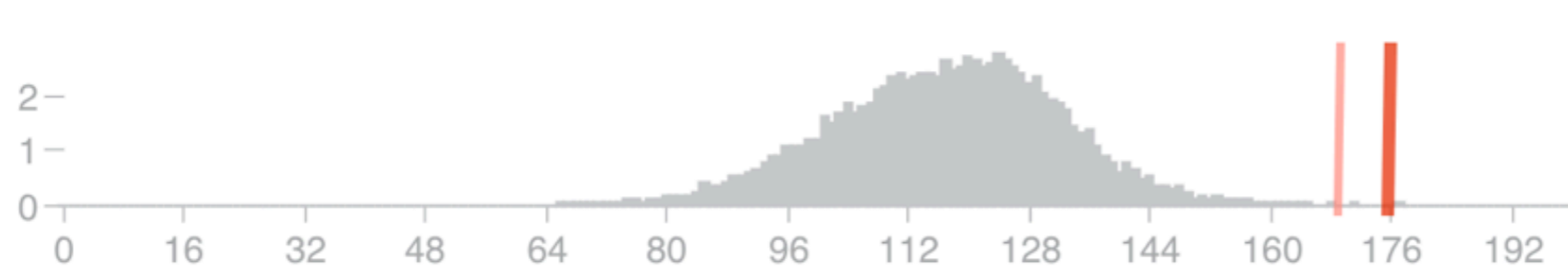
Sum of seats based on 15.000 simulations



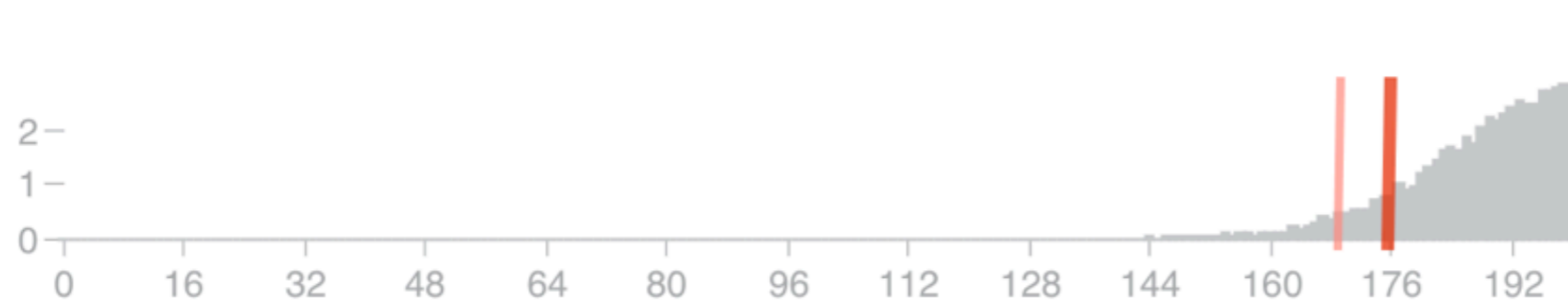
PP + C's
16% of 176
28% of 170



UP + PSOE
30% of 176
45% of 170



PSOE + C's
<1% of 176
<1% of 170



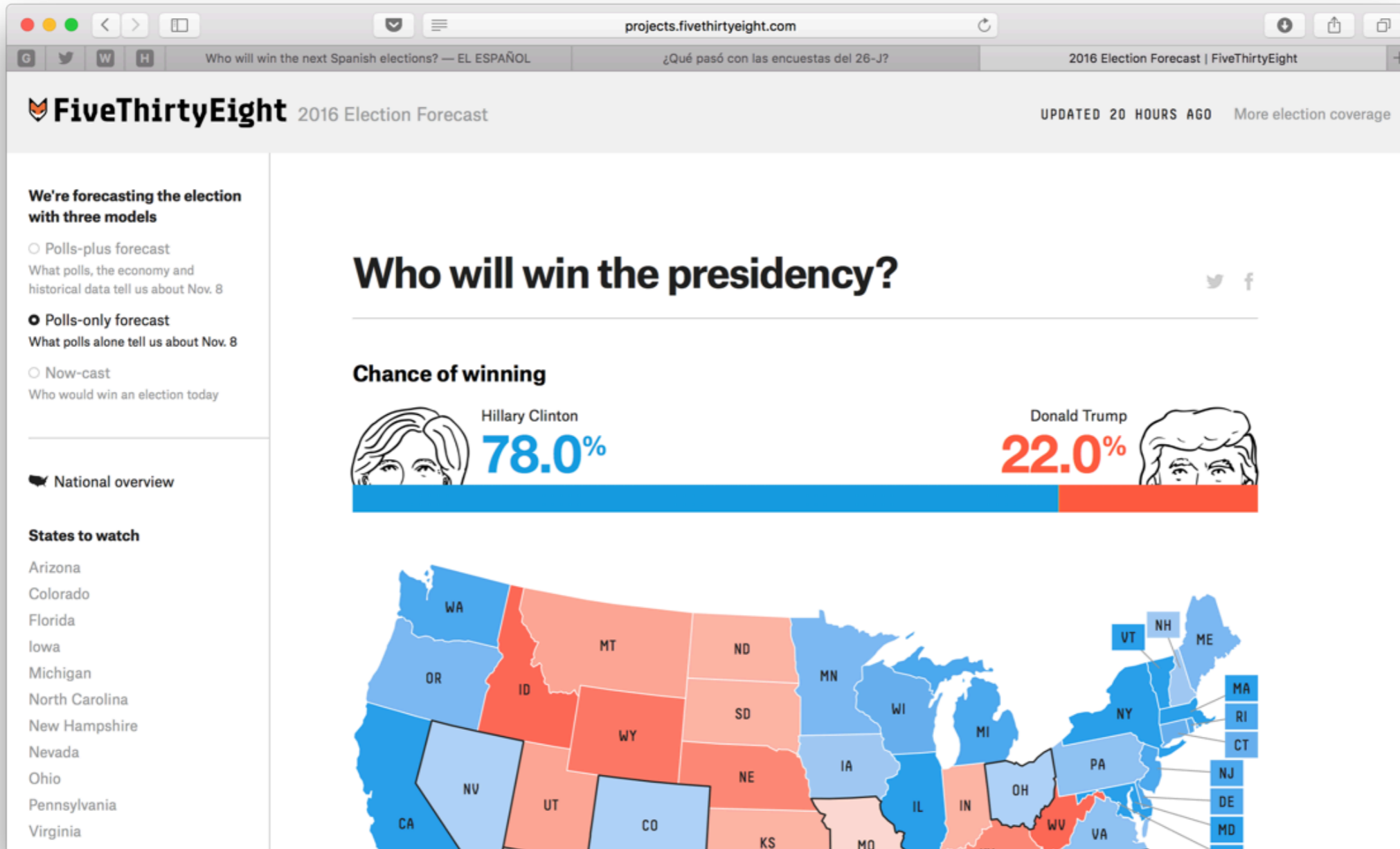
PP + PSOE
93% de 176
97% de 170

Ayer fueron elegidos estos 100 diputados

No tendrían ni que hacer campaña. Un centenar de candidatos tiene el puesto asegurado al 99% y otros tantos al 90%. En total, 200 escaños. El sistema electoral y las encuestas hacen posible predecirlo.



Otro modelo de predicción (referencia)



¿Y ahora?

Los datos sobreviven

Tema

¿Fallaron las encuestas?

Tema

¿Fallaron las encuestas?

Tema

¿Por qué fallaron?

Tema

¿Fallaron las encuestas?

Tema

¿Por qué fallaron?

Tema

¿Y ahora qué...? Los acuerdos

Tema

¿Fallaron las encuestas?

Tema

¿Por qué fallaron?

Tema

¿Debió ser tanta sorpresa?

Tema

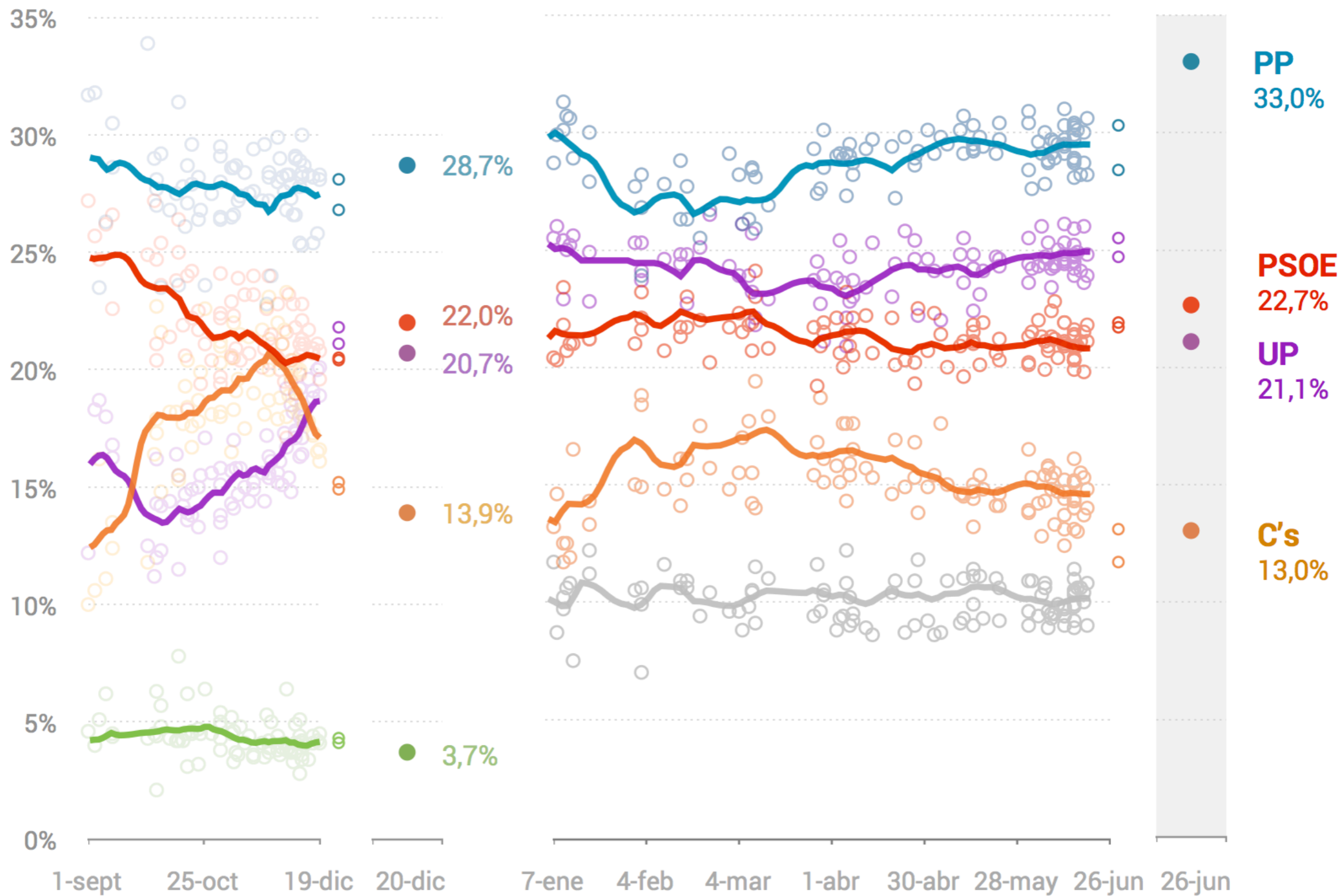
¿Y ahora qué...? Los acuerdos

Encuestas 2015

20D

Encuestas 2016

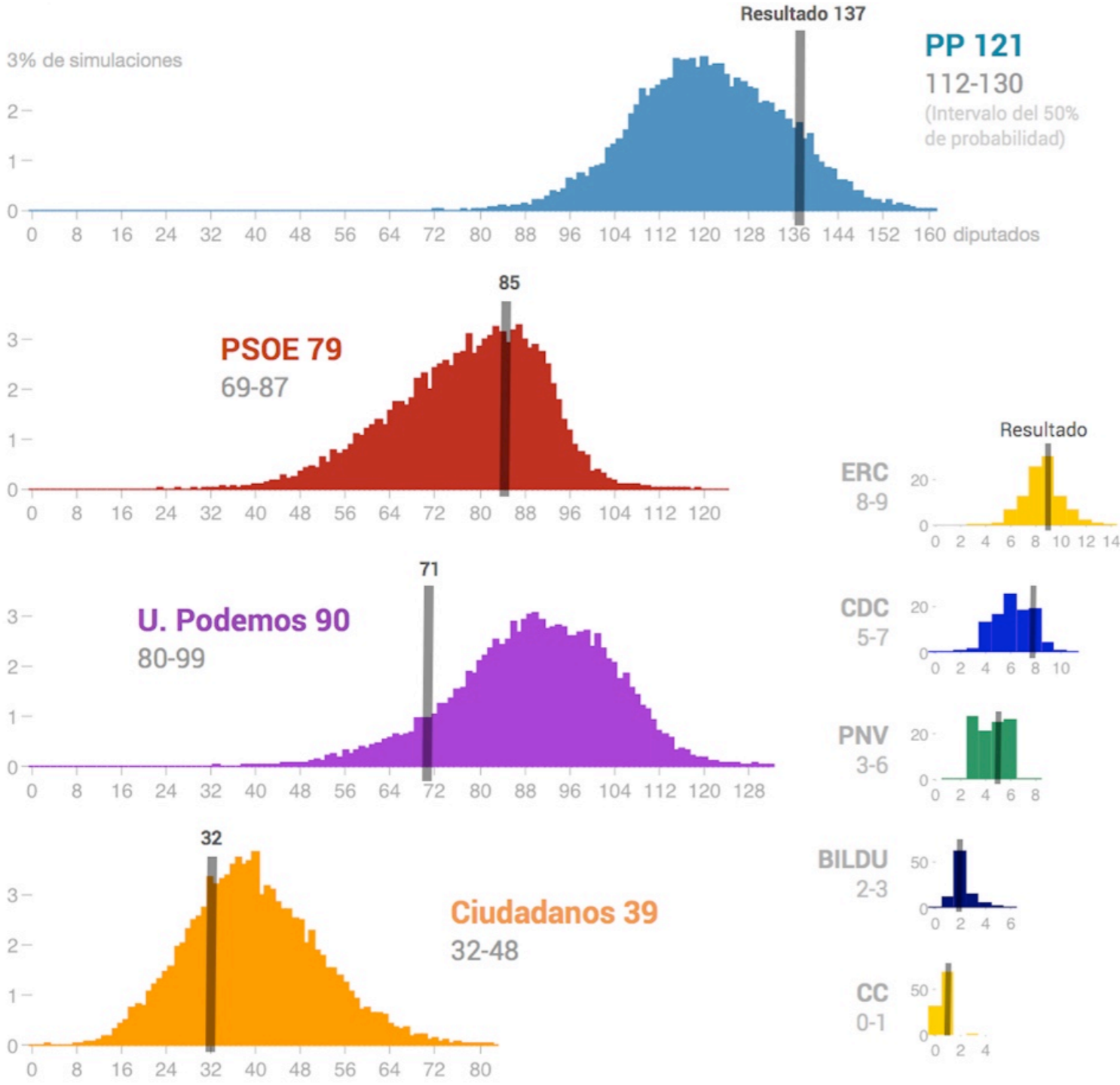
26J



No es raro que las encuestas 'fallen'

1. En **2011**, el PSOE logró +2 o +3 puntos.
2. En **2014**, Podemos entró al PE con +5 puntos; mientras el PP se quedaba con 7 menos.
3. En andaluzas **2015**, Podemos obtuvo +2 puntos.
4. En catalanas **2015**, C's mejoró en +3 puntos.
5. Colau batió por 3 puntos y Carmena por 10.

De hecho, para el modelo no gran sorpresa



El mayor reto:

Comunicar (bien) la
incertidumbre

El reto: comunicar la **incertidumbre**

El reto: comunicar la **incertidumbre**

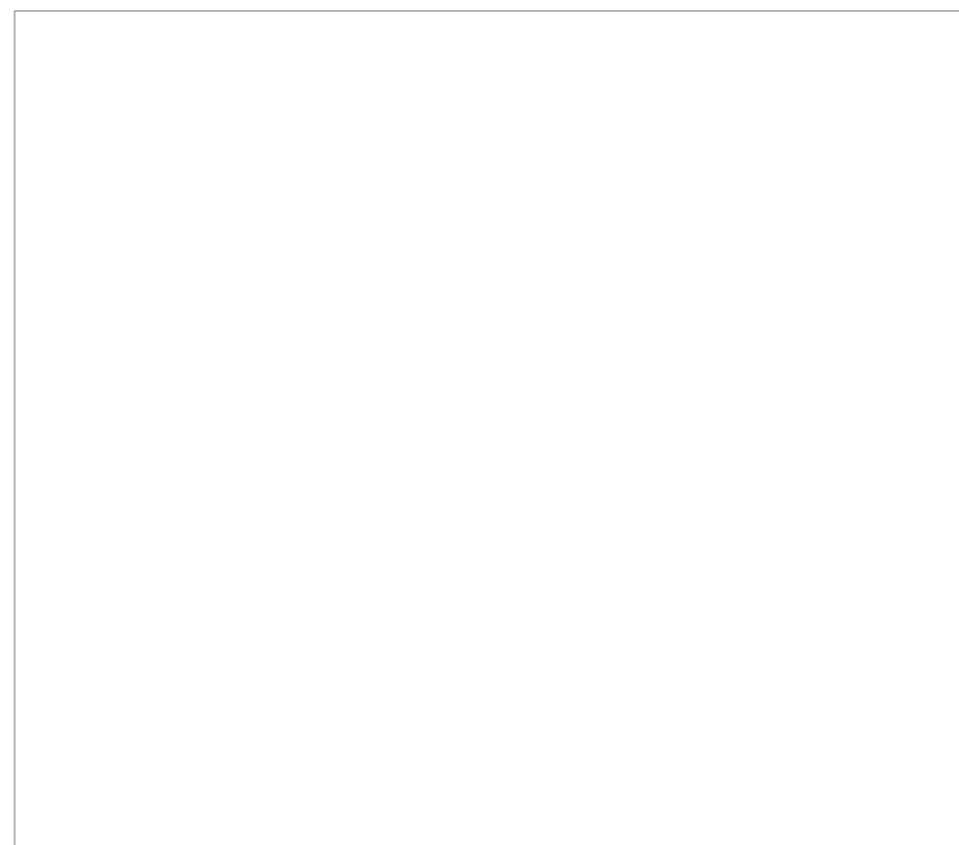
UP segundo

PSOE segundo

PP + C's suman 170 diputados

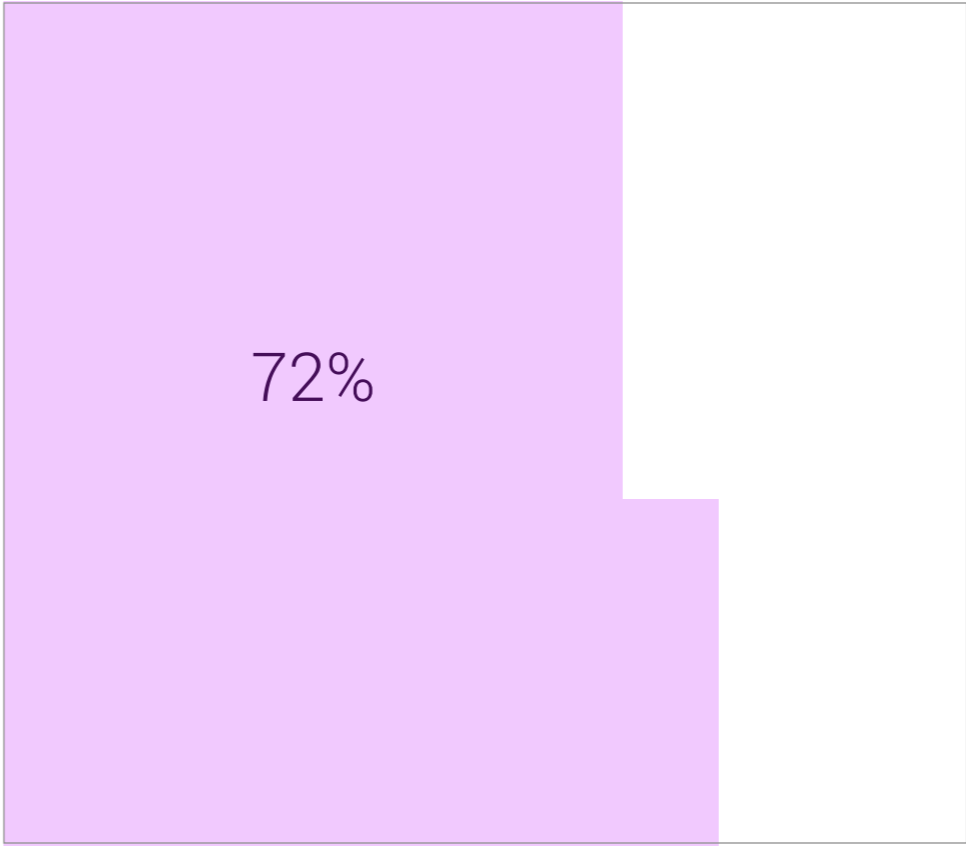
No suma ni uno ni otro

UP + PSOE suman 170 diputados

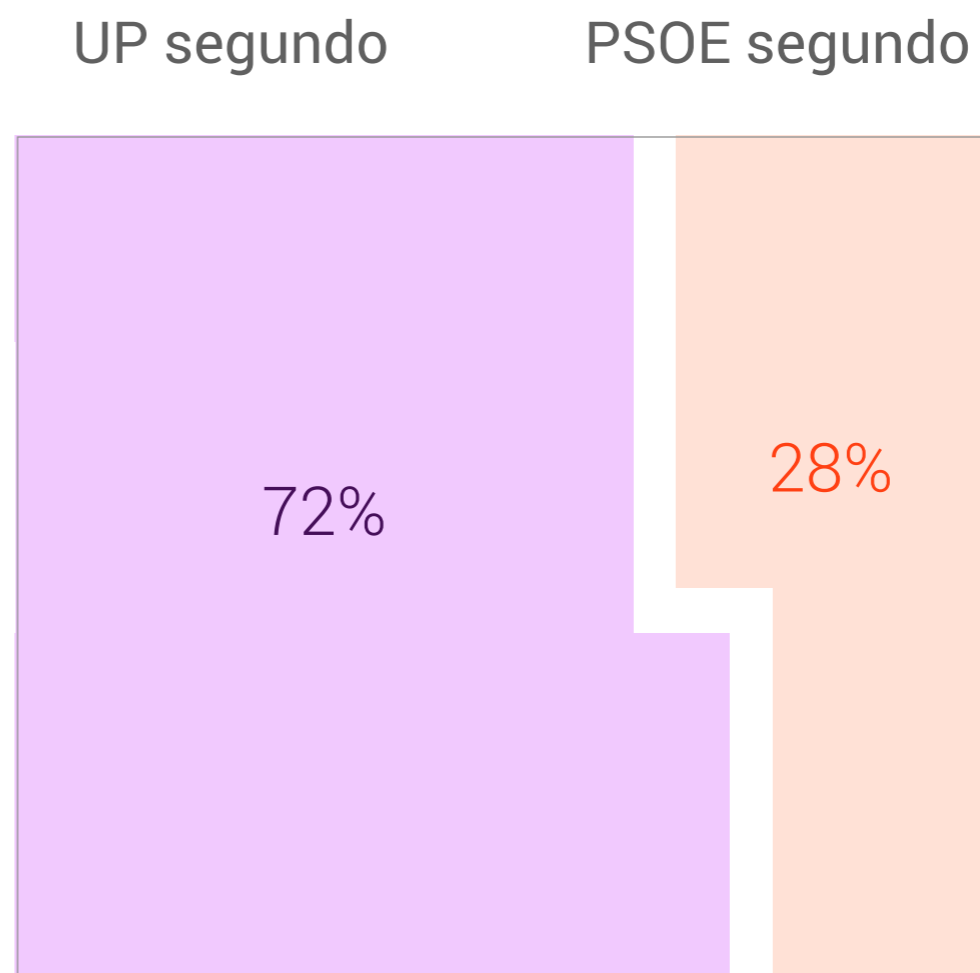


El reto: comunicar la **incertidumbre**

UP segundo



El reto: comunicar la **incertidumbre**



El reto: comunicar la **incertidumbre**

PP + C's suman 170 diputados



27 %

El reto: comunicar la **incertidumbre**

PP + C's suman 170 diputados

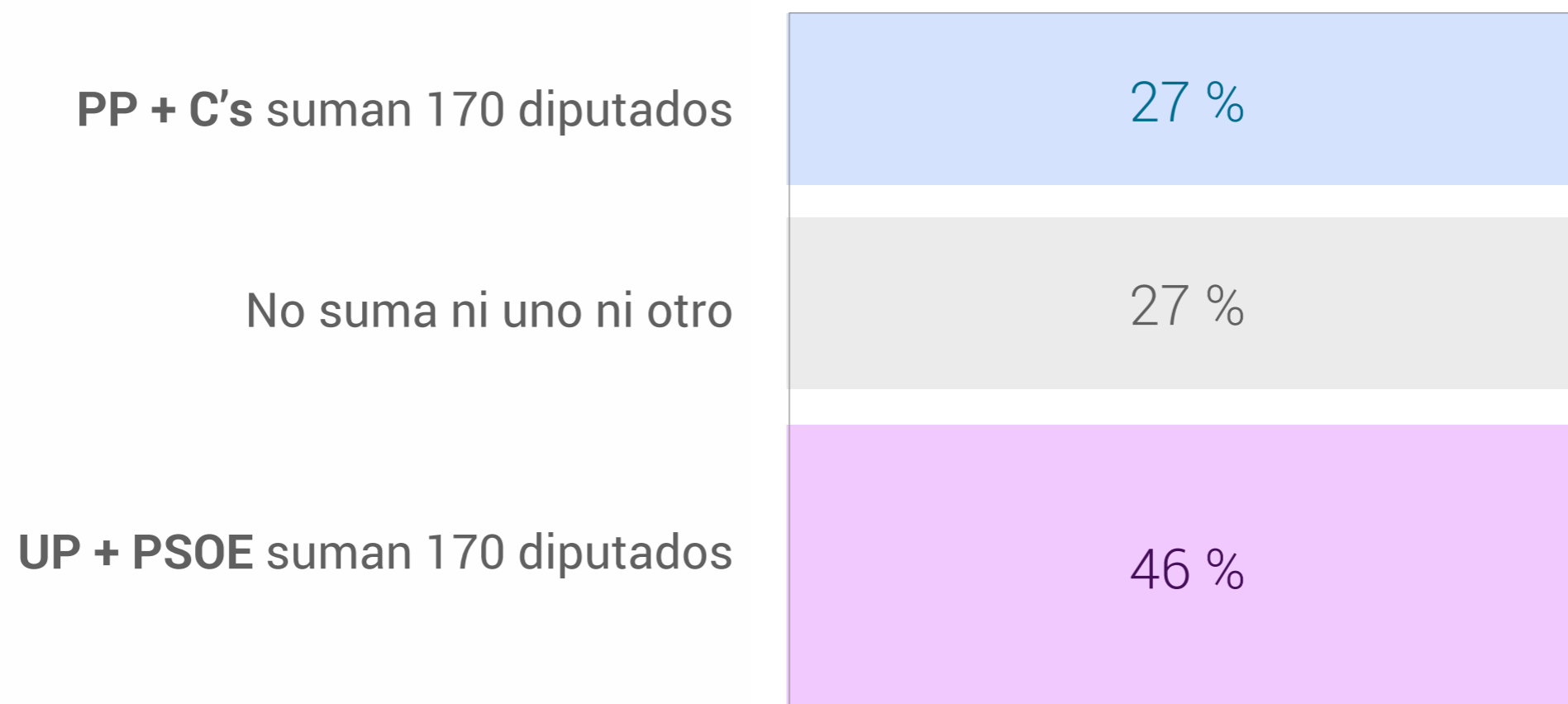
27 %

No suma ni uno ni otro

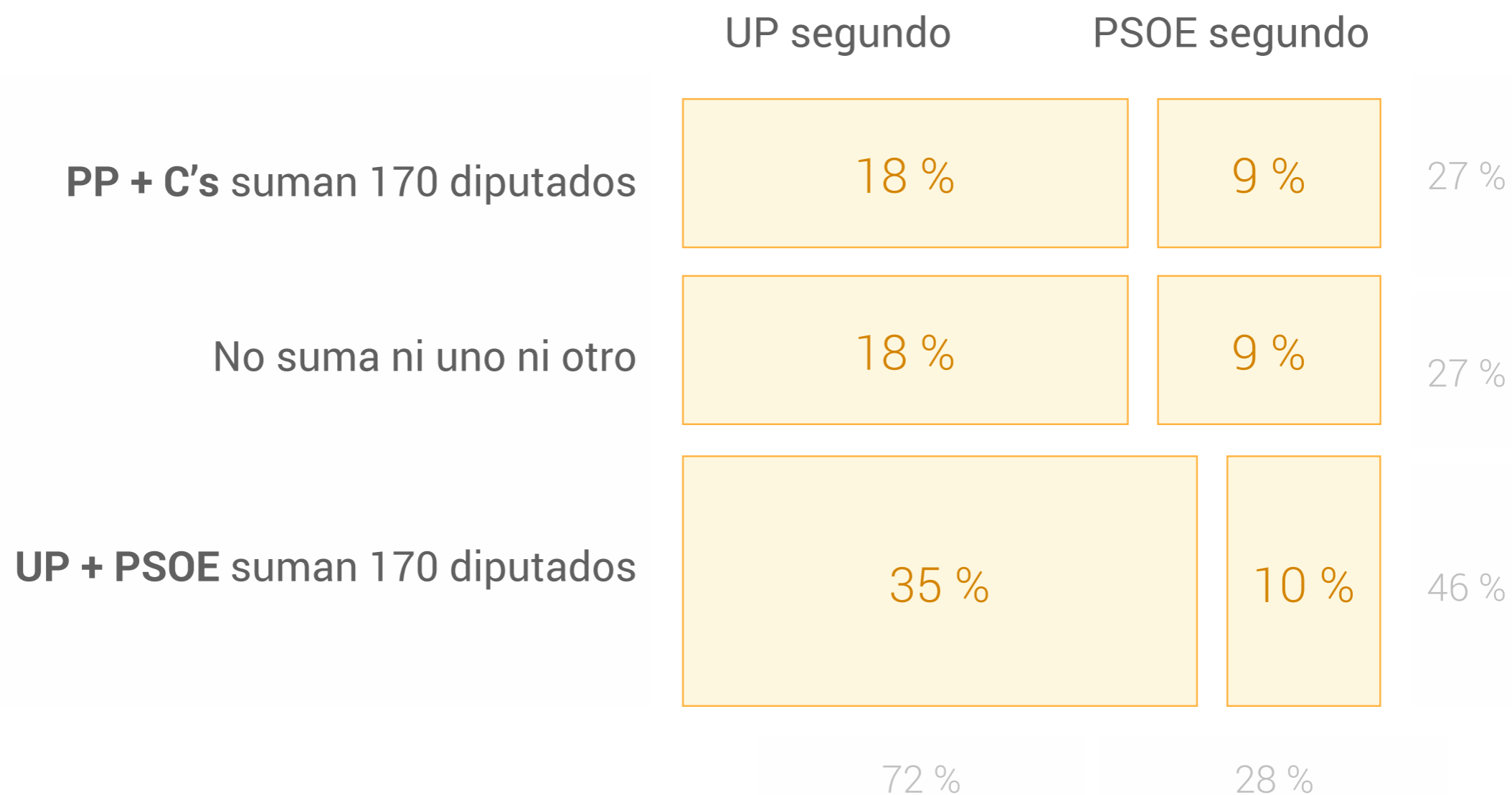
27 %

|

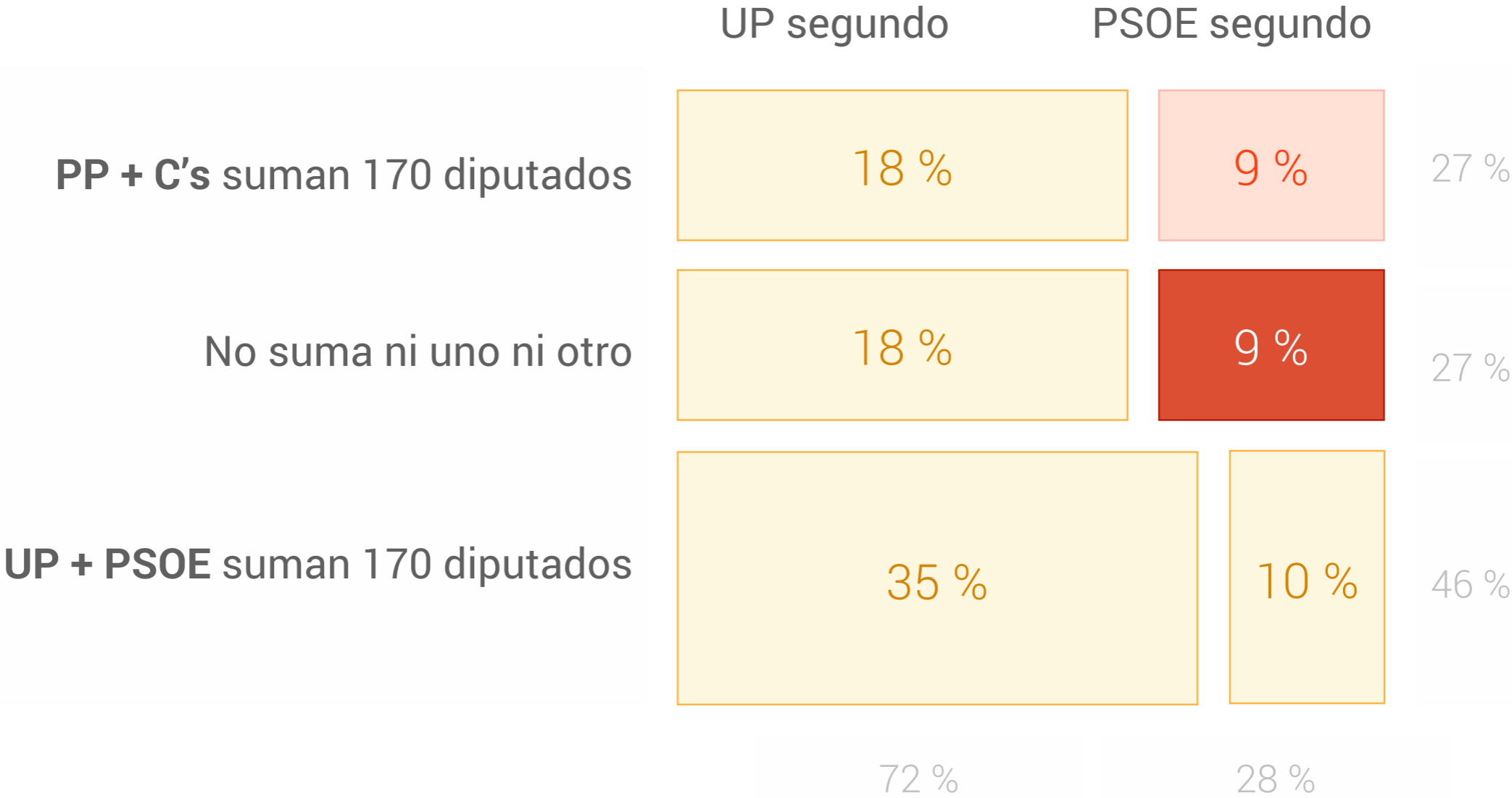
El reto: comunicar la **incertidumbre**



El reto: comunicar la **incertidumbre**



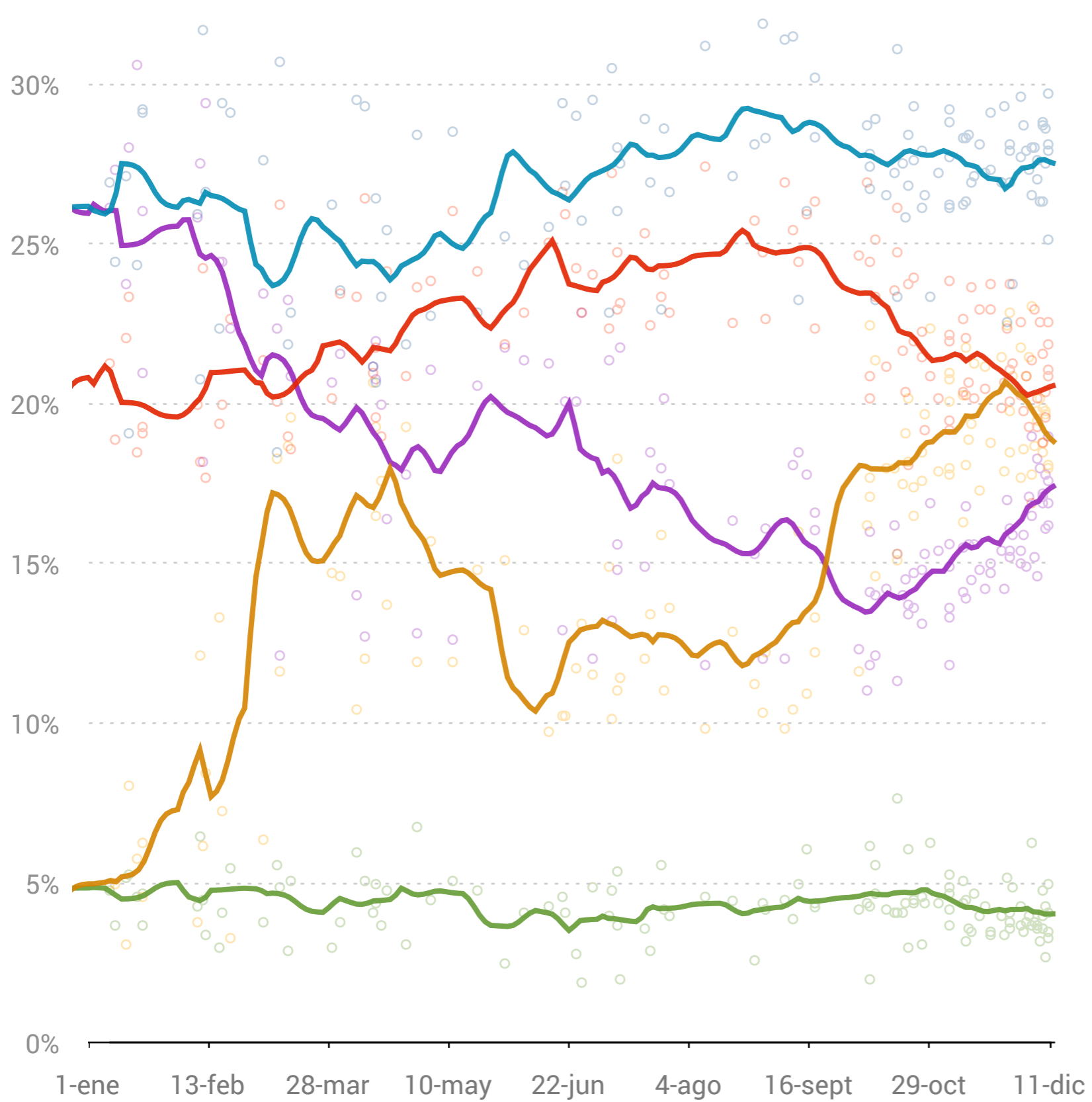
El reto: comunicar la **incertidumbre**



Gracias

20-D

Encuestas

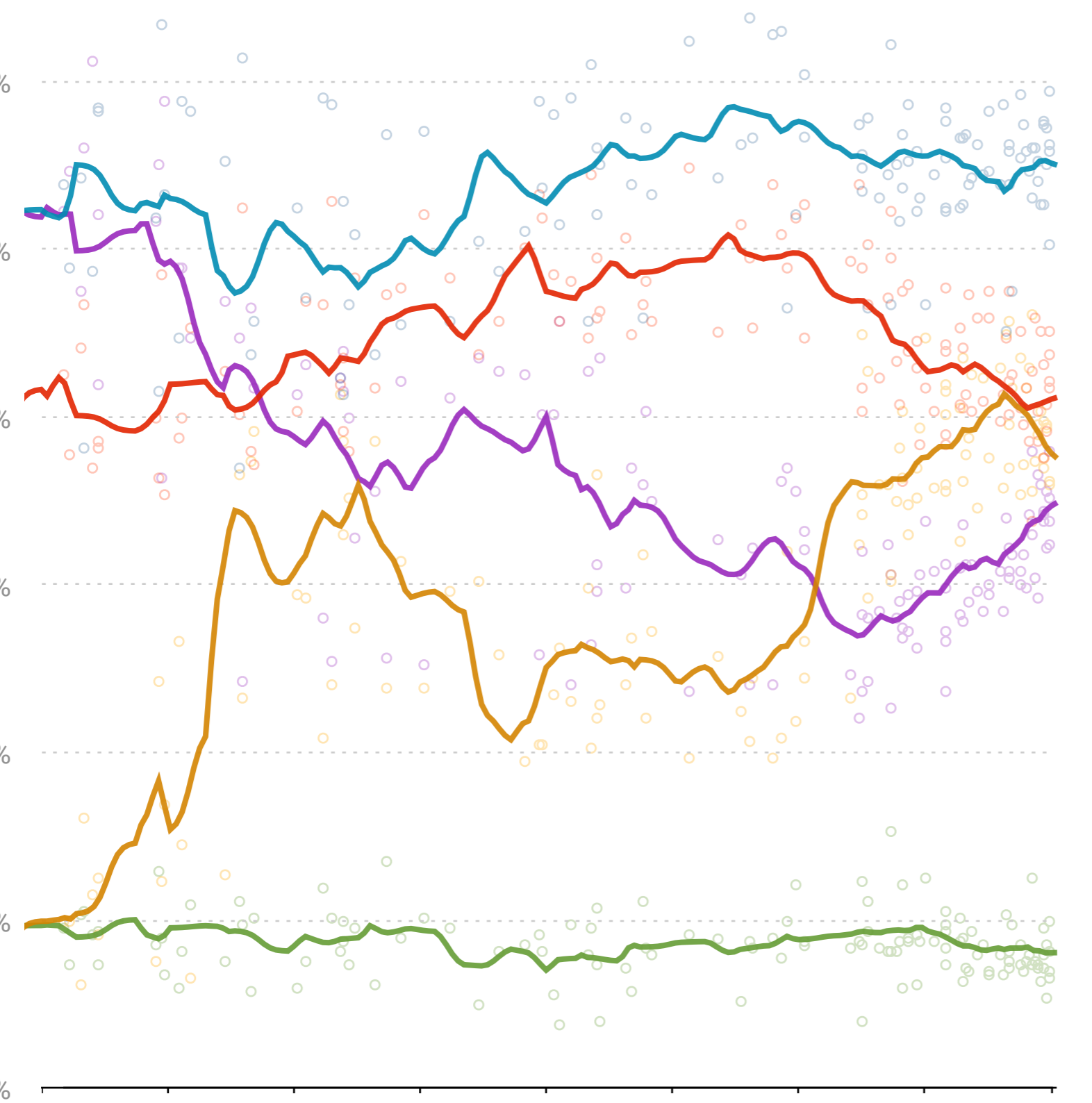


Encuestas

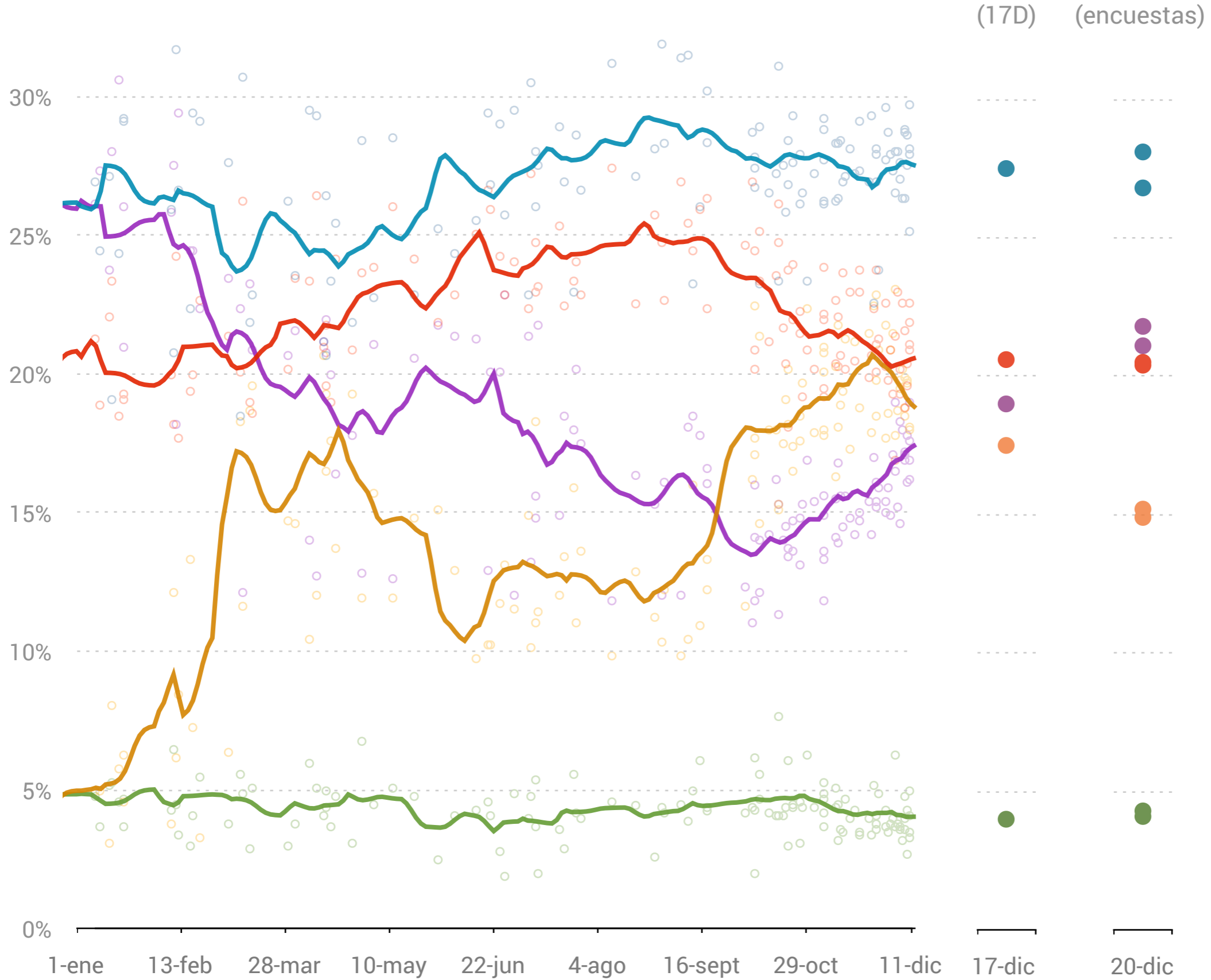
EE
(17D)

30%
25%
20%
15%
10%
5%
0%

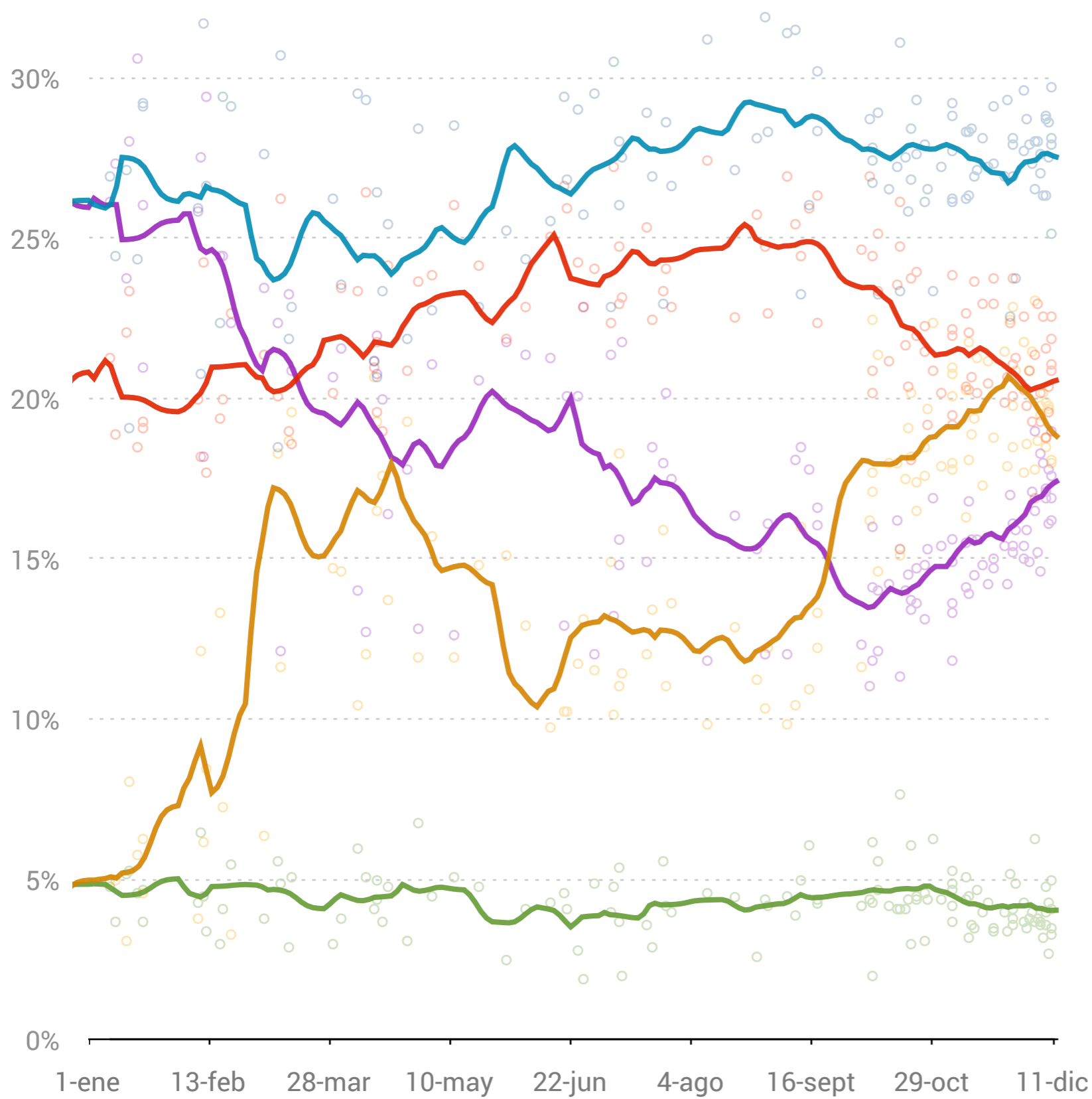
1-ene 13-feb 28-mar 10-may 22-jun 4-ago 16-sept 29-oct 11-dic 17-dic



Encuestas



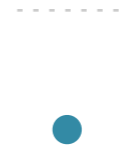
Encuestas



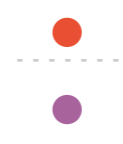
EE
(17D)

20D
(encuestas)

Resultado
(al 100%)



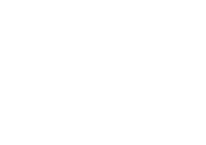
PP
28,7%



PSOE
22,0%



Pods
20,7%



Cs
13,9%



IU
3,7%

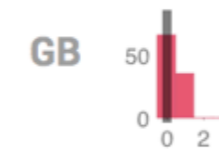
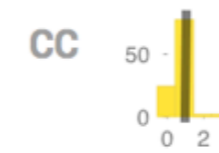
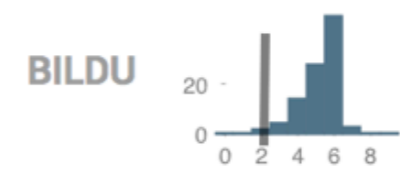
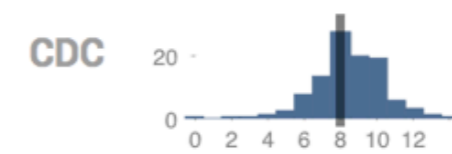
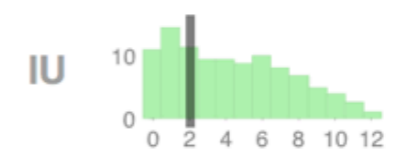
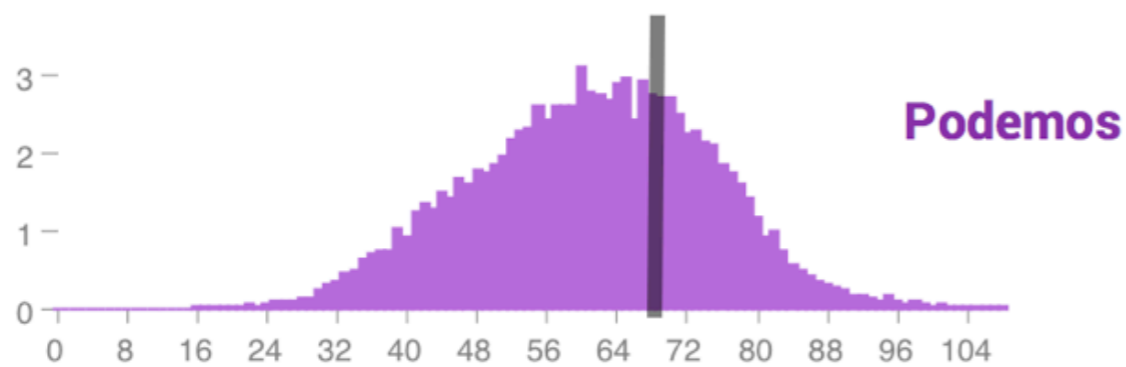
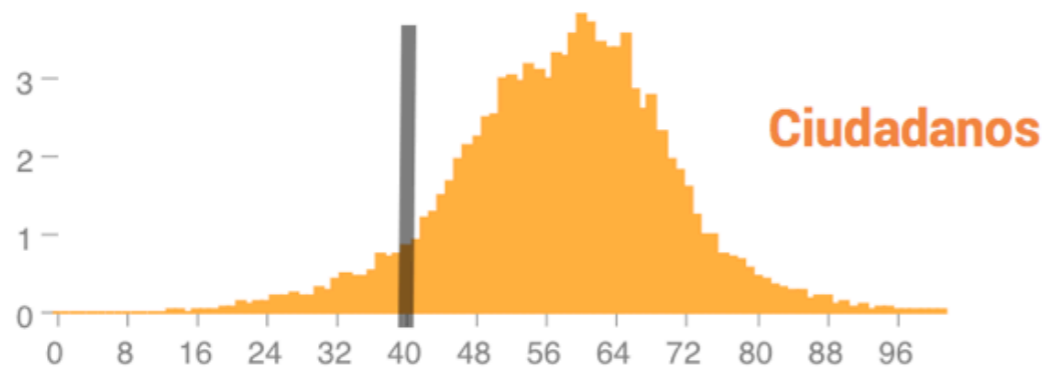
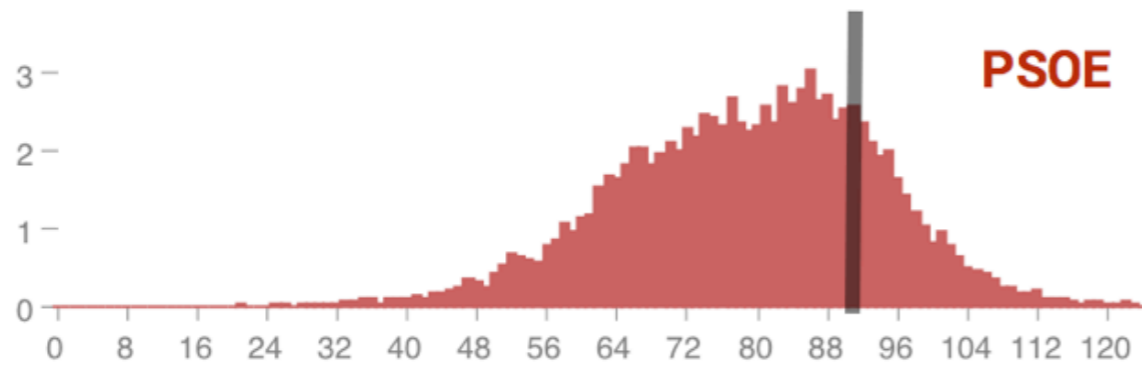
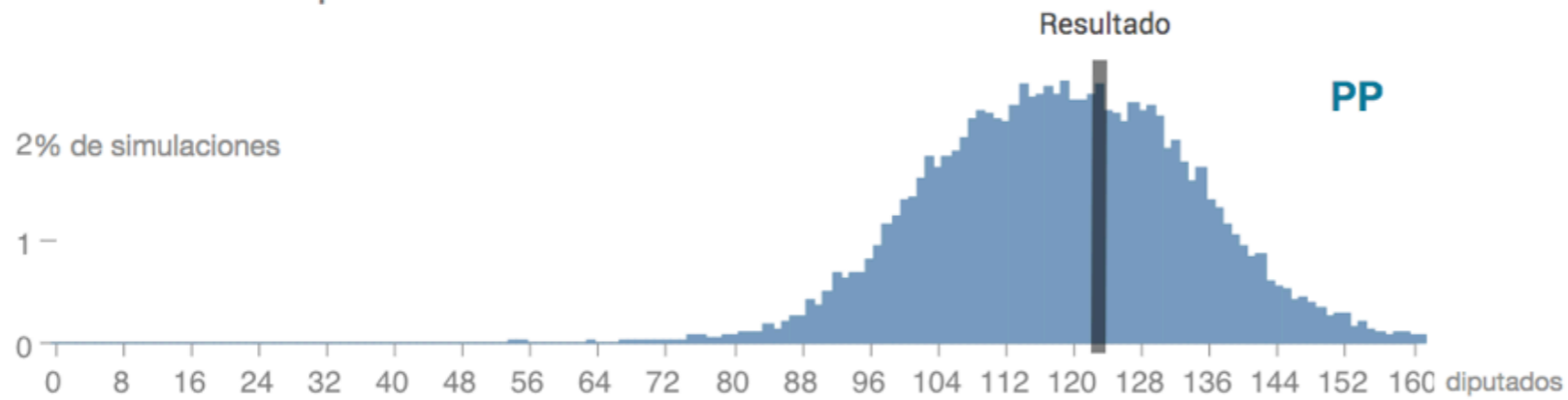
17-dic

20-dic

20-dic

Resultado vs. modelo 14D con tendencias

Distribuciones de probabilidad



Resultado vs. modelo 20D

(Promedio de sondeos, distribuido con nuestro modelo)

Partido	Resultado	Intervalo 50%		Intervalo 90%	
PP	123	108	129	94	143
PSOE	90	69	90	54	101
Podemos	69	52	71	38	82
Ciudadanos	40	50	66	37	77
IU	2	2	7	0	10
CDC	8	8	10	6	11
ERC	9	8	10	6	11
PNV	6	5	6	3	6
Bildu	2	5	6	3	6
CC	1	1	1	0	1

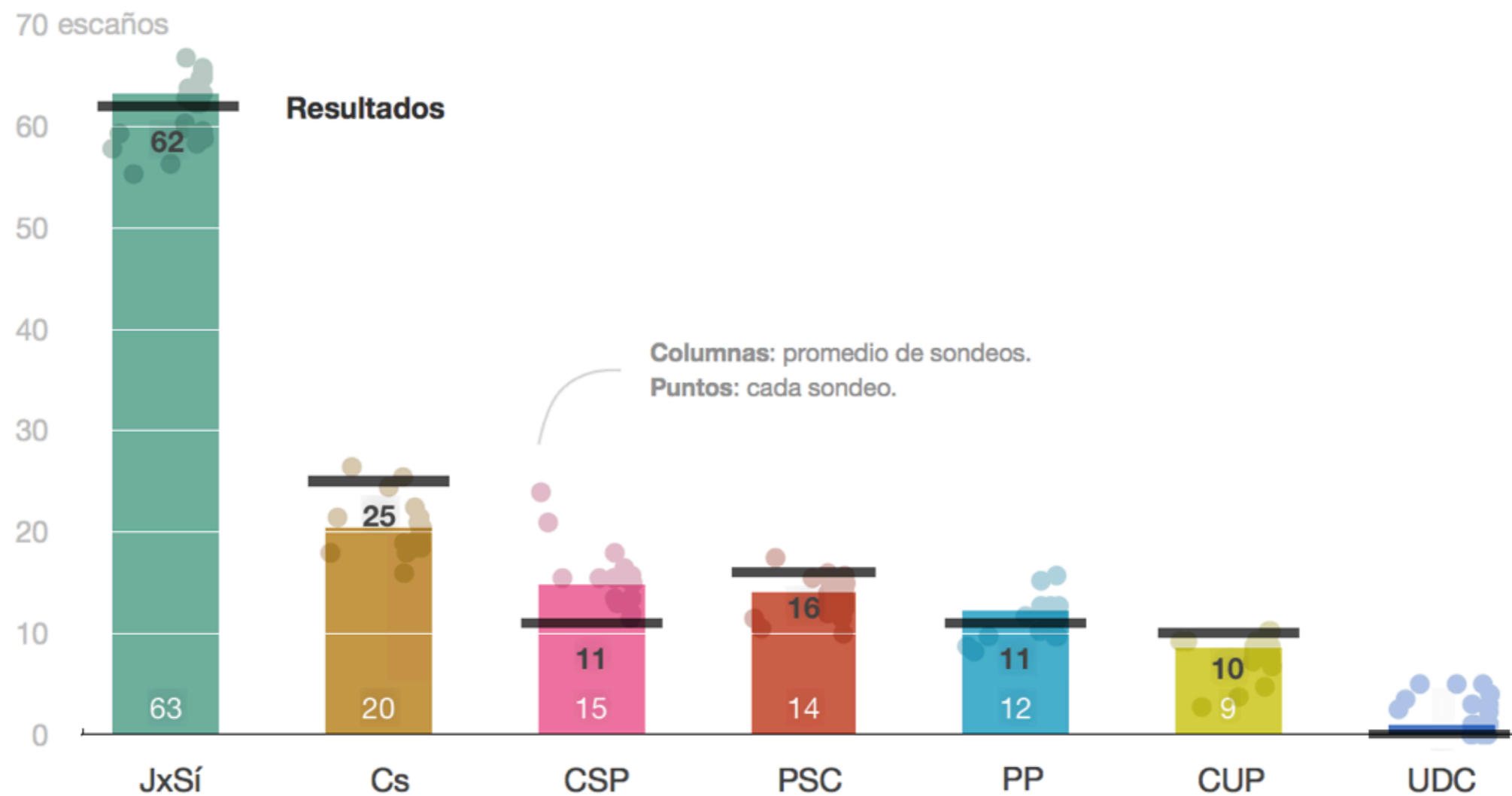
Partido	2ª posición	3ª	4ª
PSOE	70 %	20 %	9 %
Ciudadanos	10 %	37 %	50 %
Podemos	17 %	42 %	39 %

27-S

Parlament 27S

27S: encuestas vs resultados

Los escaños según las encuestas y los resultados



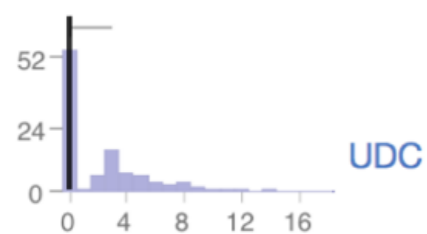
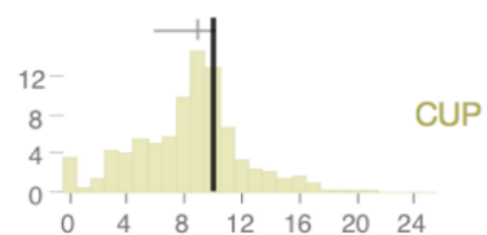
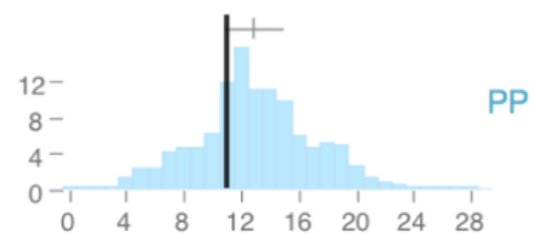
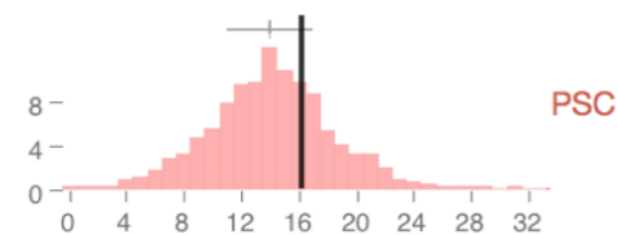
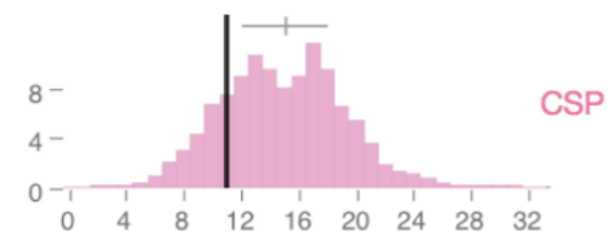
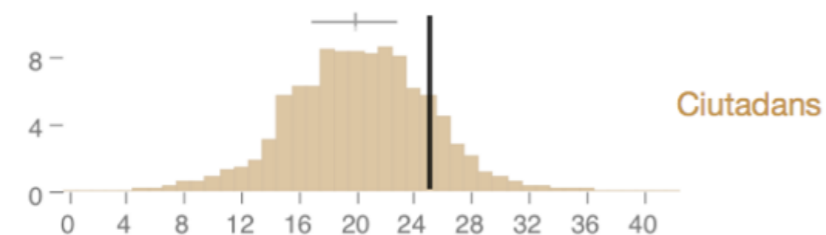
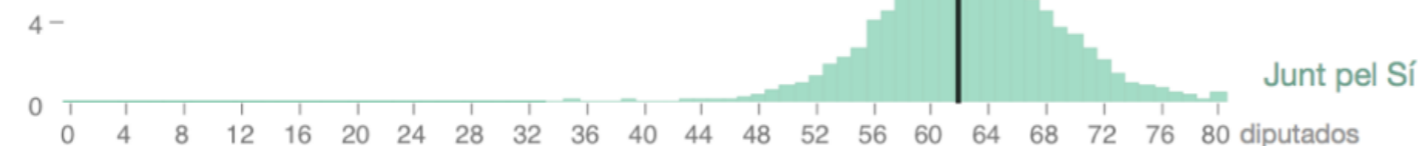
Sigma Dos, GAPS, GESOP, NC Report, Invymark, IBES, Feedback, CIS, Metroscopia, DYM, GAD3, My Word, Celeste Tel.

Kiko Llaneras | EL ESPAÑOL

Predicción vs resultados

Nuestro modelo contra los resultados del 27S

8% de las 15.000 simulaciones.



Resultado

Predicción

El histograma de color representa la probabilidad de cada resultado.

La cruz sobre cada histograma representa la predicción mediana y el intervalo del 50% de probabilidad. Es decir, es esperable que los resultados caigan fuera de ese intervalos la mitad de las veces.

Detalles técnicos del modelo

Resultados de 15.000 simulaciones a partir de un promedio ponderado de 20 encuestas. Se considera una incertidumbre normal (con una desviación estándar de entre 1.5% y 4%).

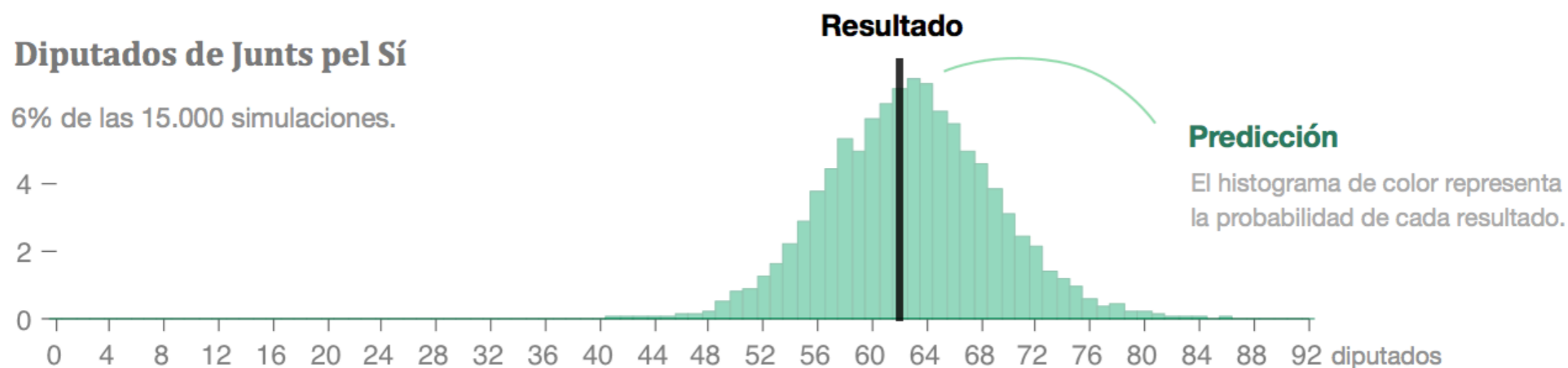
Detalles y fuente completa en EL ESPAÑOL.

Parlament 27S

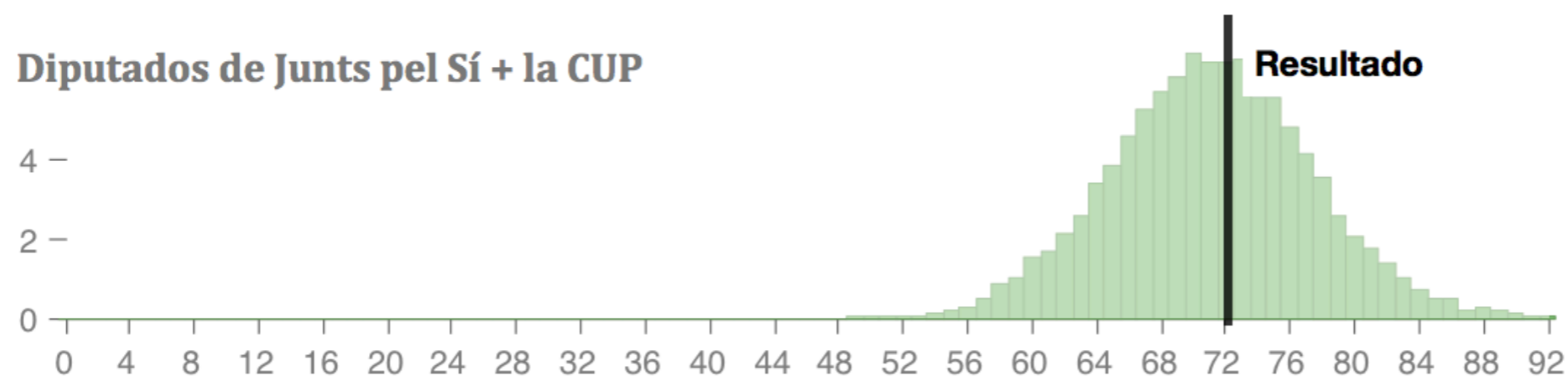
Predicción vs resultados

Diputados de Junts pel Sí

6% de las 15.000 simulaciones.



Diputados de Junts pel Sí + la CUP

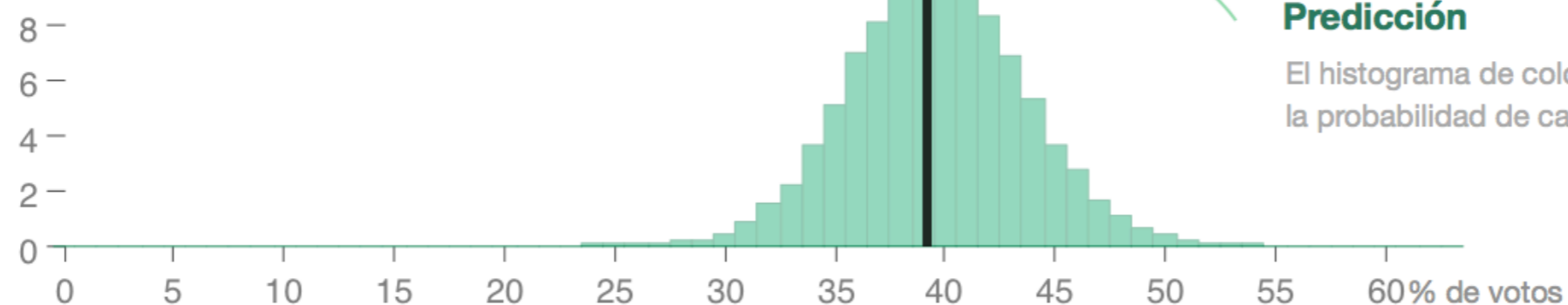


Parlament 27S

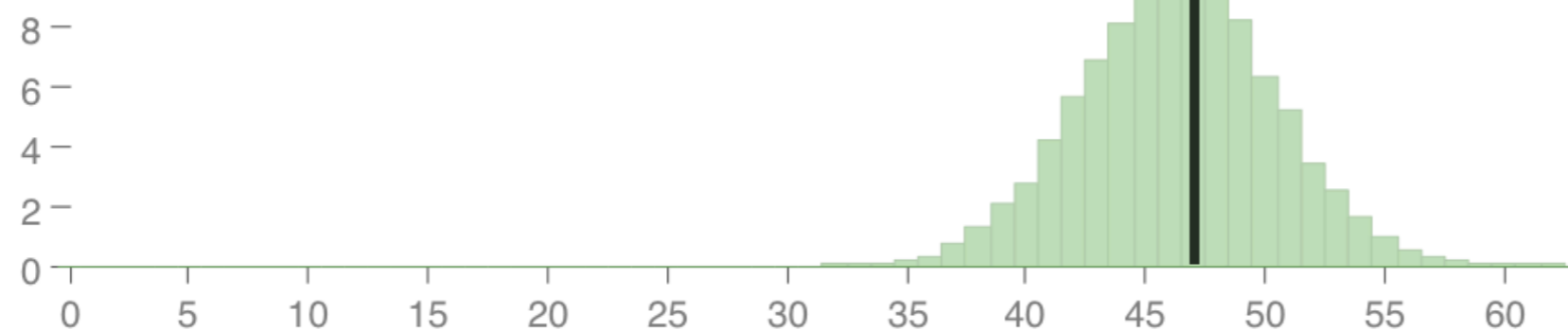
Predicción vs resultados

Votos de Junts pel Sí

10% de las 15.000 simulaciones.



Votos de Junts pel Sí + la CUP

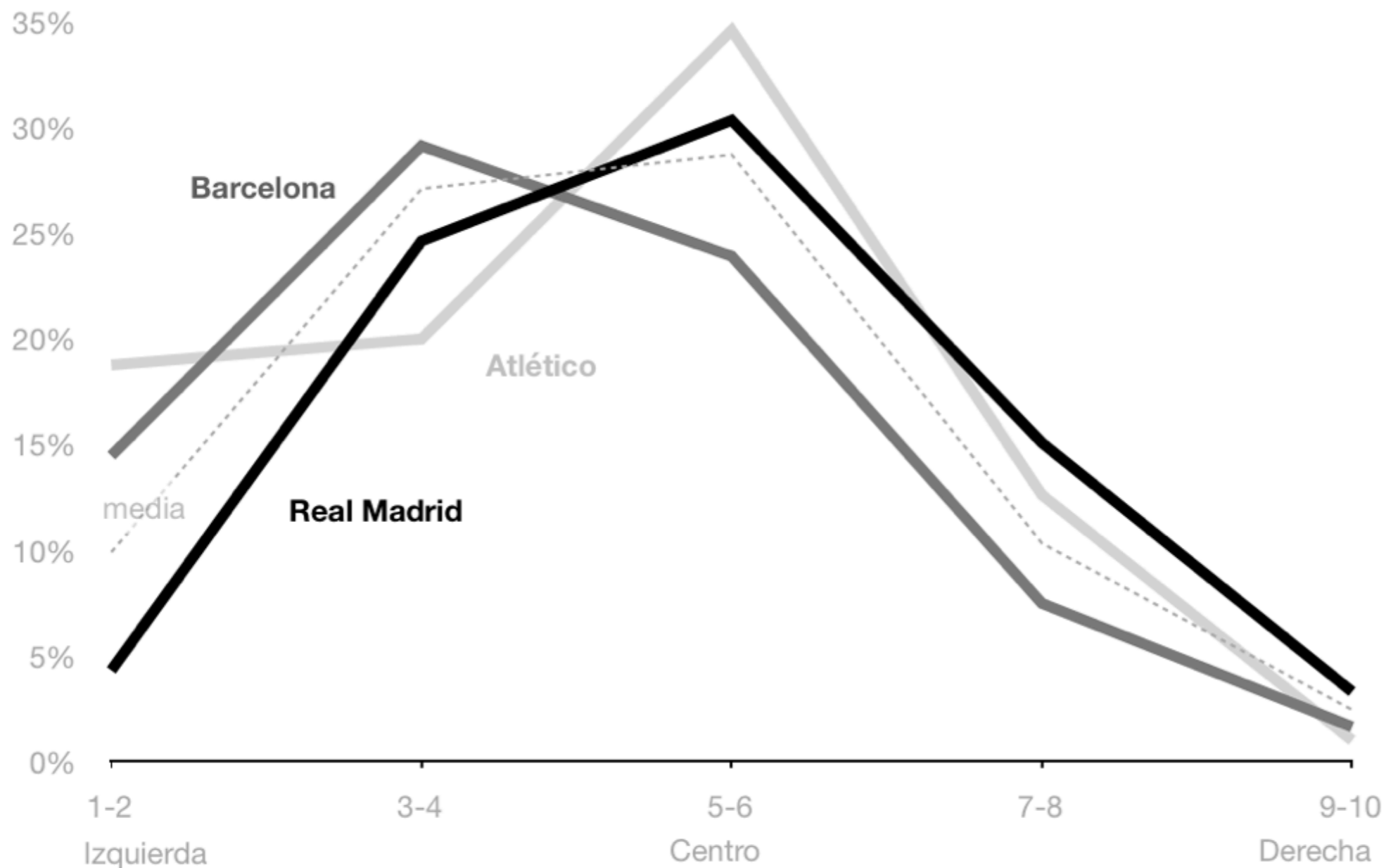


Datos random

Perfil ideológico de cada equipo

CIS (2014)

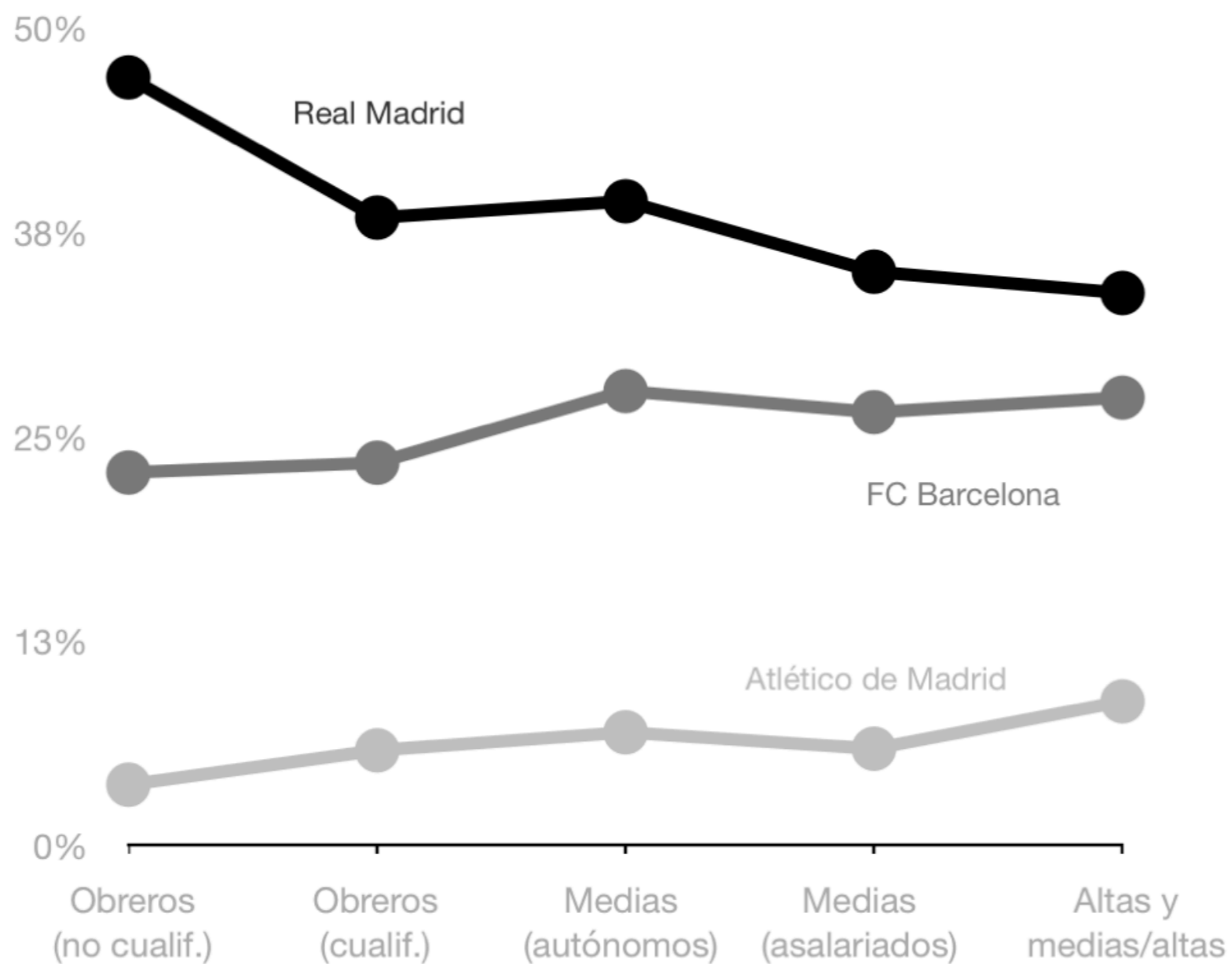
Simpatizantes de cada equipo según su autoubicación en izquierda/derecha.



Fútbol y clase social

CIS (2014)

¿Cuál es el equipo por el que siente más simpatía?
Respuestas según clase social.



¿Provincias del FC Barcelona o del Real Madrid?

Diferencia entre el porcentaje de seguidores del Real Madrid y el Barcelona.

