PSC 400
SYRACUSE UNIVERSITY

DATA ANALYTICS FOR POLITICAL SCIENCE

BIVARIATE RELATIONSHIPS

ASSIGNMENTS

- Data Analysis Memo 2 due on Friday
- Review Exercise 4 due today

UN VOTING

Table 3.8. United Nations Ideal Points Data.

Variable	*Description*
`CountryName`	name of the country
`CountryAbb`	abbreviated name of the country
`idealpoint`	its estimated ideal point
`Year`	year for which the ideal point is estimated
`PctAgreeUS`	proportion of votes that match with votes cast by the United States on the same issue
`PctAgreeRUSSIA`	proportion of votes that match with votes cast by Russia/the Soviet Union on the same issue

- unvoting.csv
- How have states' ideological positions changed since the end of the Cold War?
  - Ideological positions measured through voting patterns

UN VOTING

Table 3.8. United Nations Ideal Points Data.

Variable	*Description*
`CountryName`	name of the country
`CountryAbb`	abbreviated name of the country
`idealpoint`	its estimated ideal point
`Year`	year for which the ideal point is estimated
`PctAgreeUS`	proportion of votes that match with votes cast by the United States on the same issue
`PctAgreeRUSSIA`	proportion of votes that match with votes cast by Russia/the Soviet Union on the same issue

- idealpoint: captures countries "liberalism" (on e.g. political freedom, democratization, financial liberalization)
  - Higher values=more liberal
- PctAgreeUS/PctAgreeRUSSIA: share of votes that match how US/Russia vote

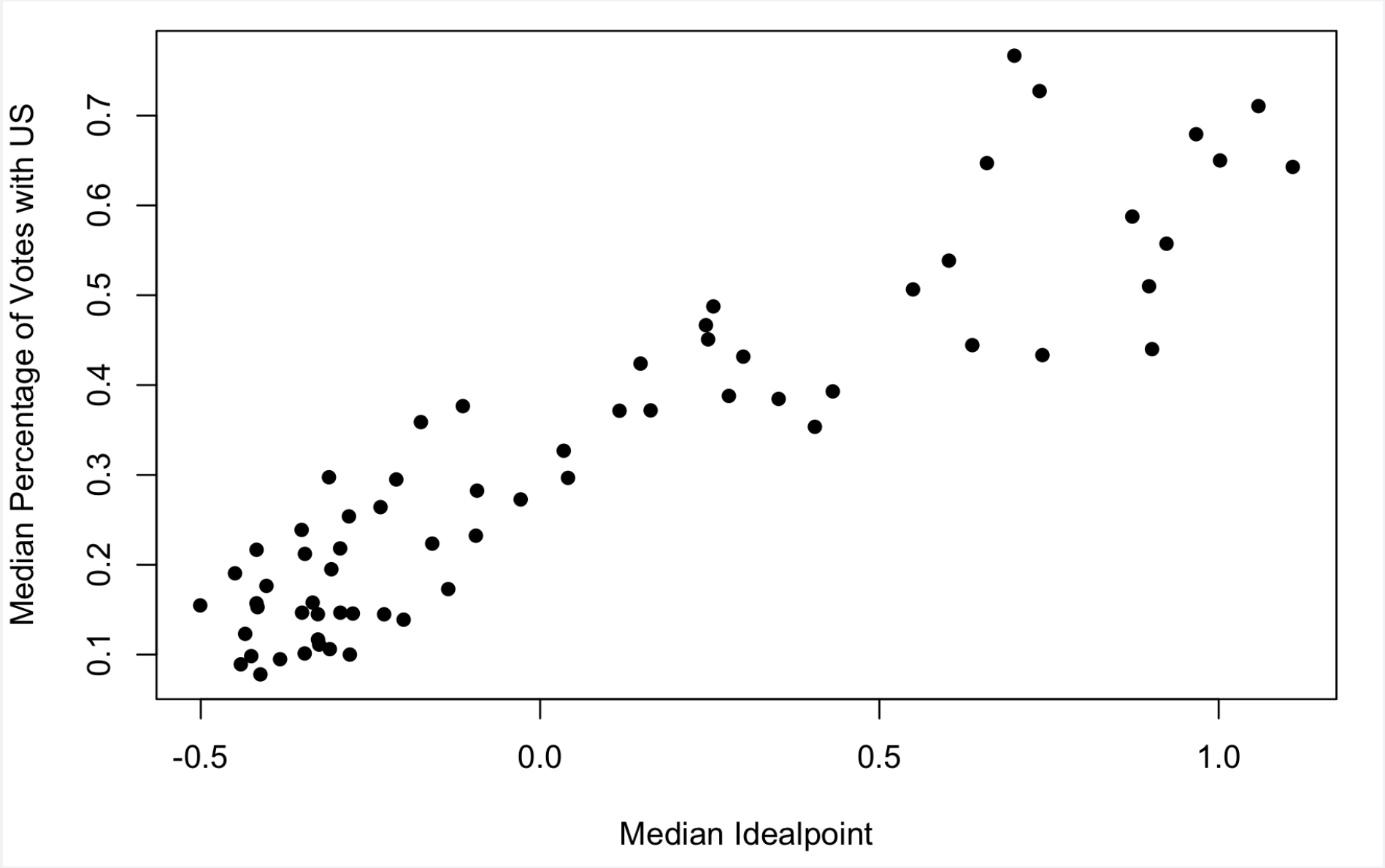
UN VOTING

Table 3.8. United Nations Ideal Points Data.

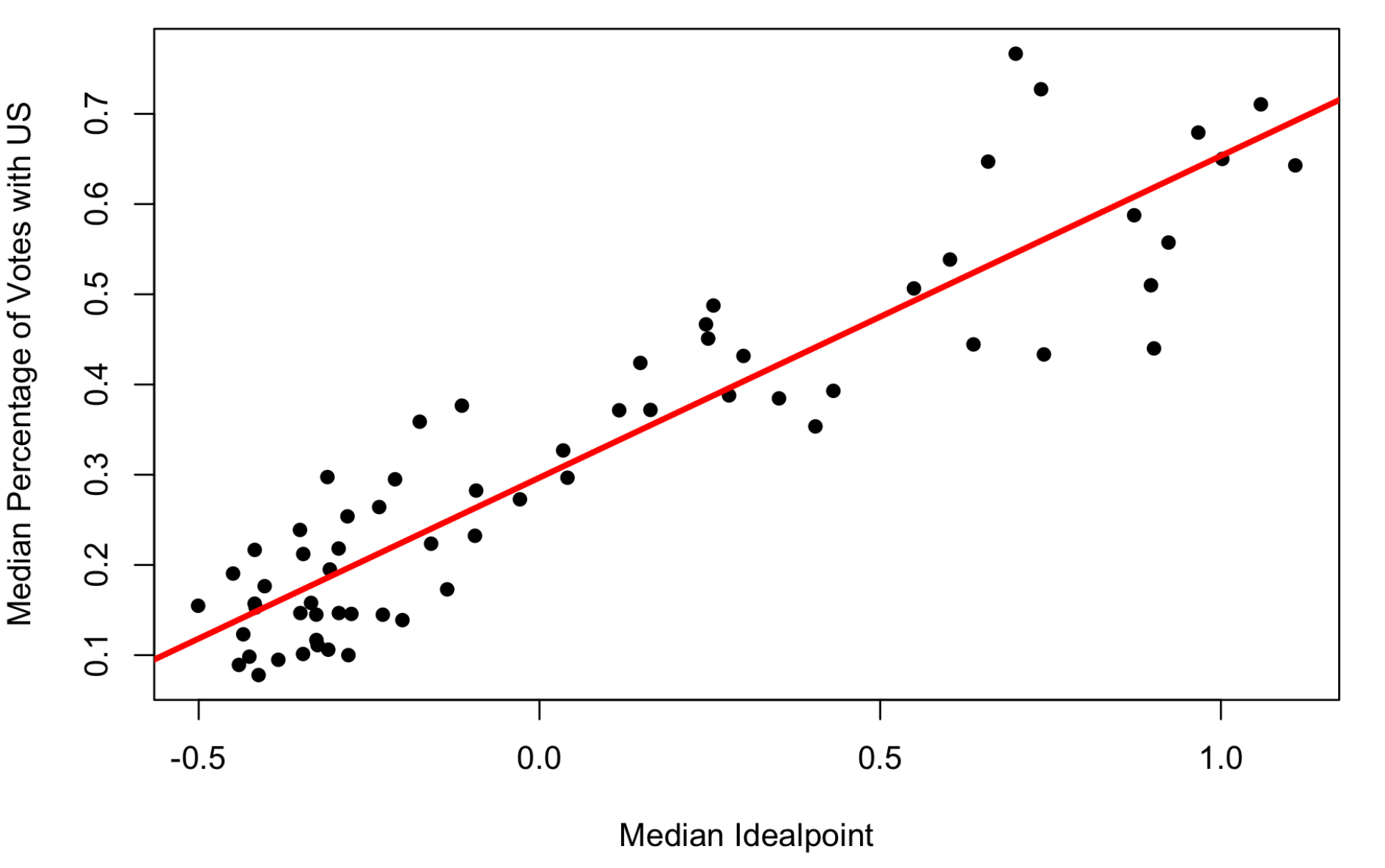
Variable	*Description*
`CountryName`	name of the country
`CountryAbb`	abbreviated name of the country
`idealpoint`	its estimated ideal point
`Year`	year for which the ideal point is estimated
`PctAgreeUS`	proportion of votes that match with votes cast by the United States on the same issue
`PctAgreeRUSSIA`	proportion of votes that match with votes cast by Russia/the Soviet Union on the same issue

- Compute the median of idealpoint for each year, and plot it
- Compute the median of PctAgreeUS for each year, and plot it
- Compute the correlation between the two variables

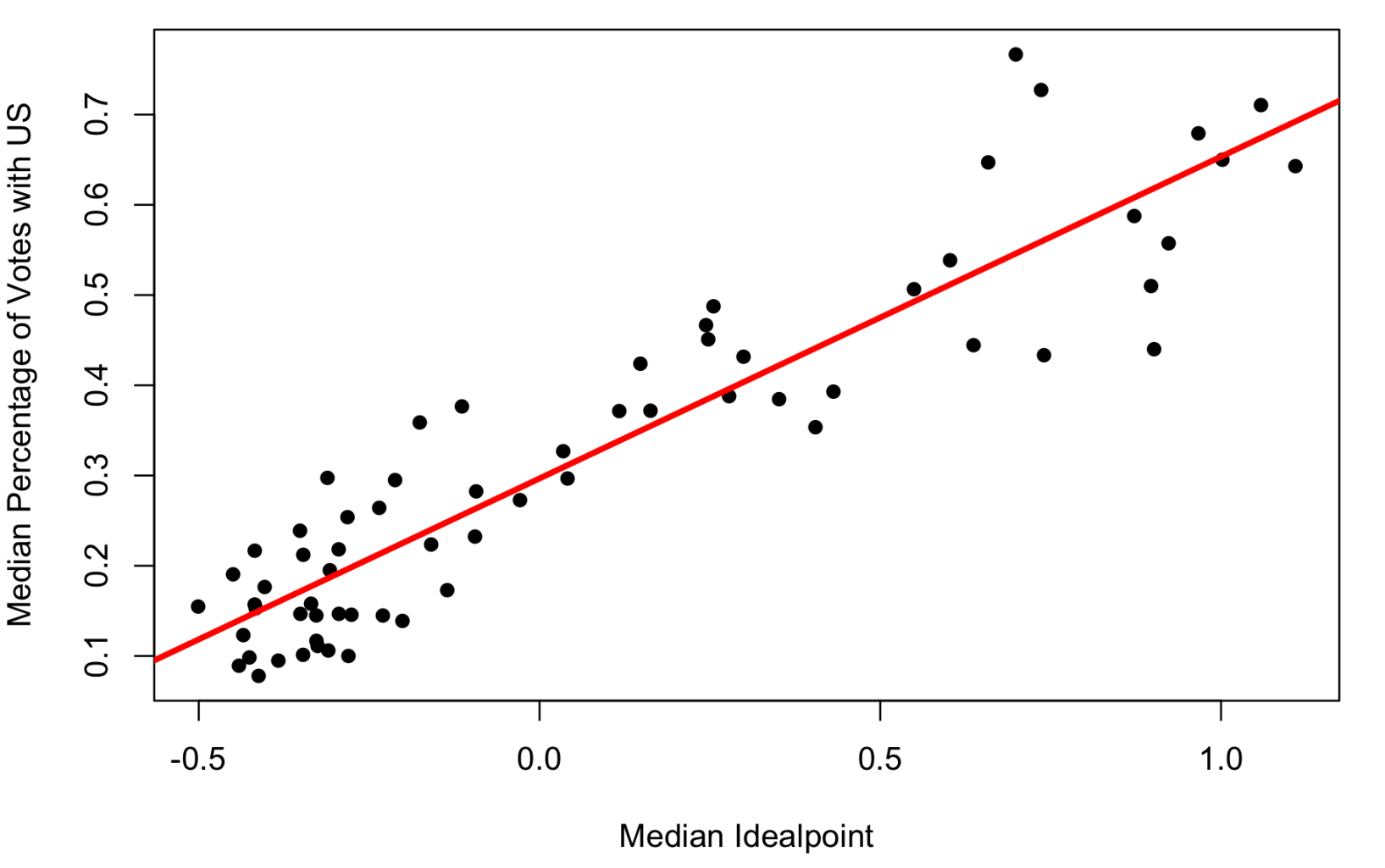
UN VOTING


Median Percentage of Votes with US
0.1
0.2
0.3
0.4
0.5
0.6
0.7
-0.5
0.0
0.5
1.0
Median Idealpoint

UN VOTING

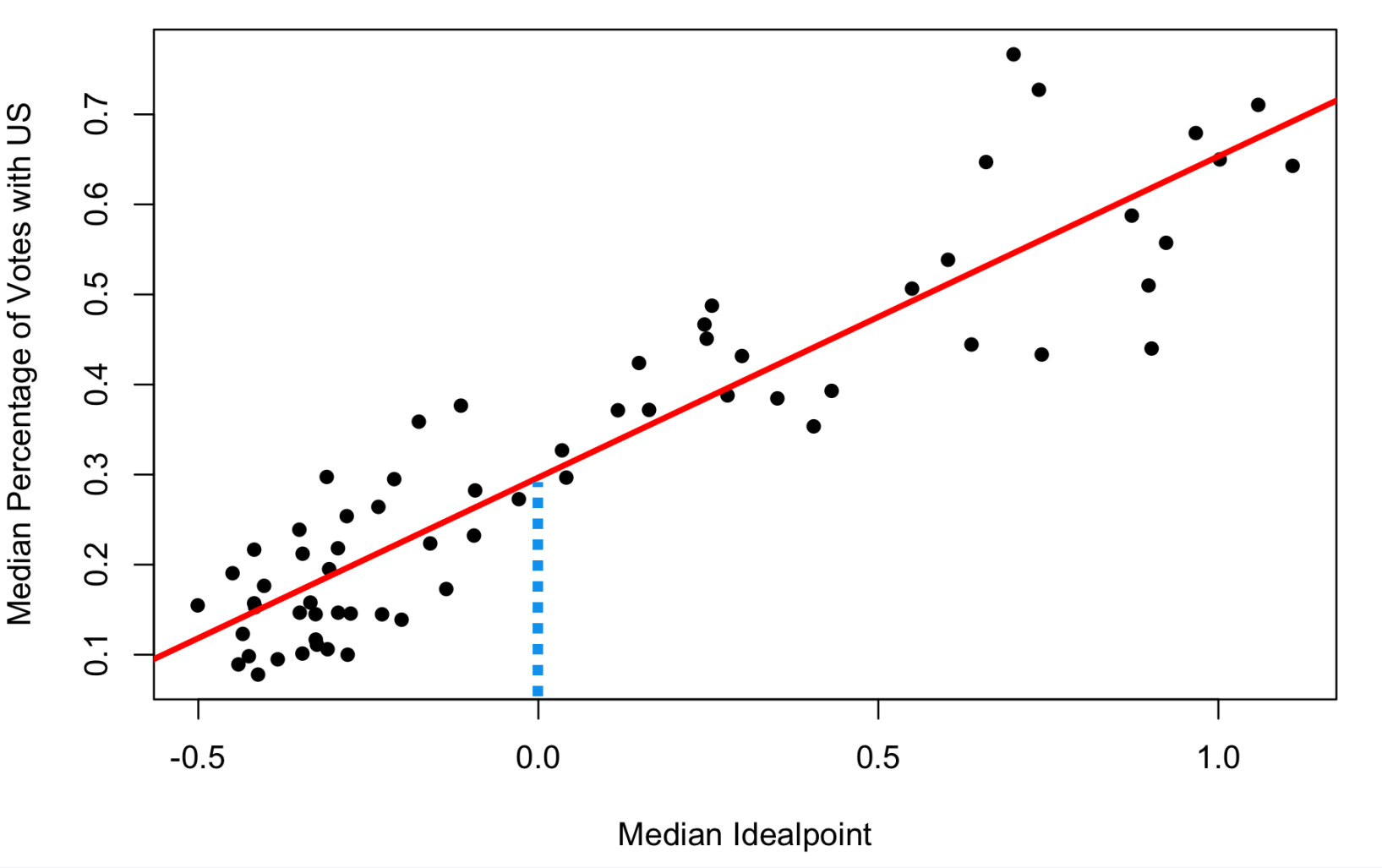

Median Percentage of Votes with US
0.1
0.2
0.3
0.4
0.5
0.6
0.7
-0.5
0.0
0.5
1.0
Median Idealpoint

UN VOTING



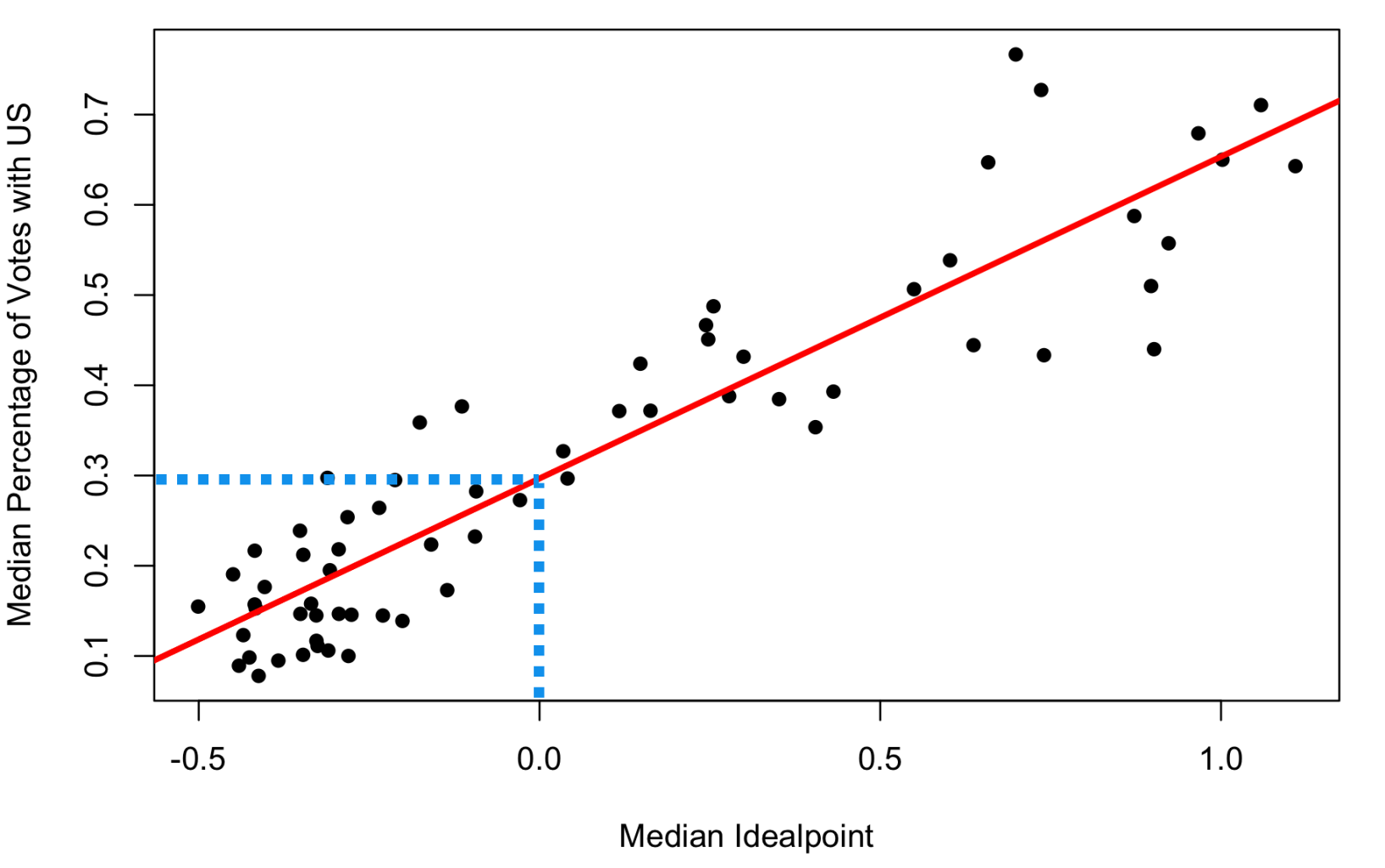
- On average, how much higher is the share of votes with the US for a country that is 1 on the ideal point scale, compared to one with a value of 0 on the ideal point scale?

UN VOTING



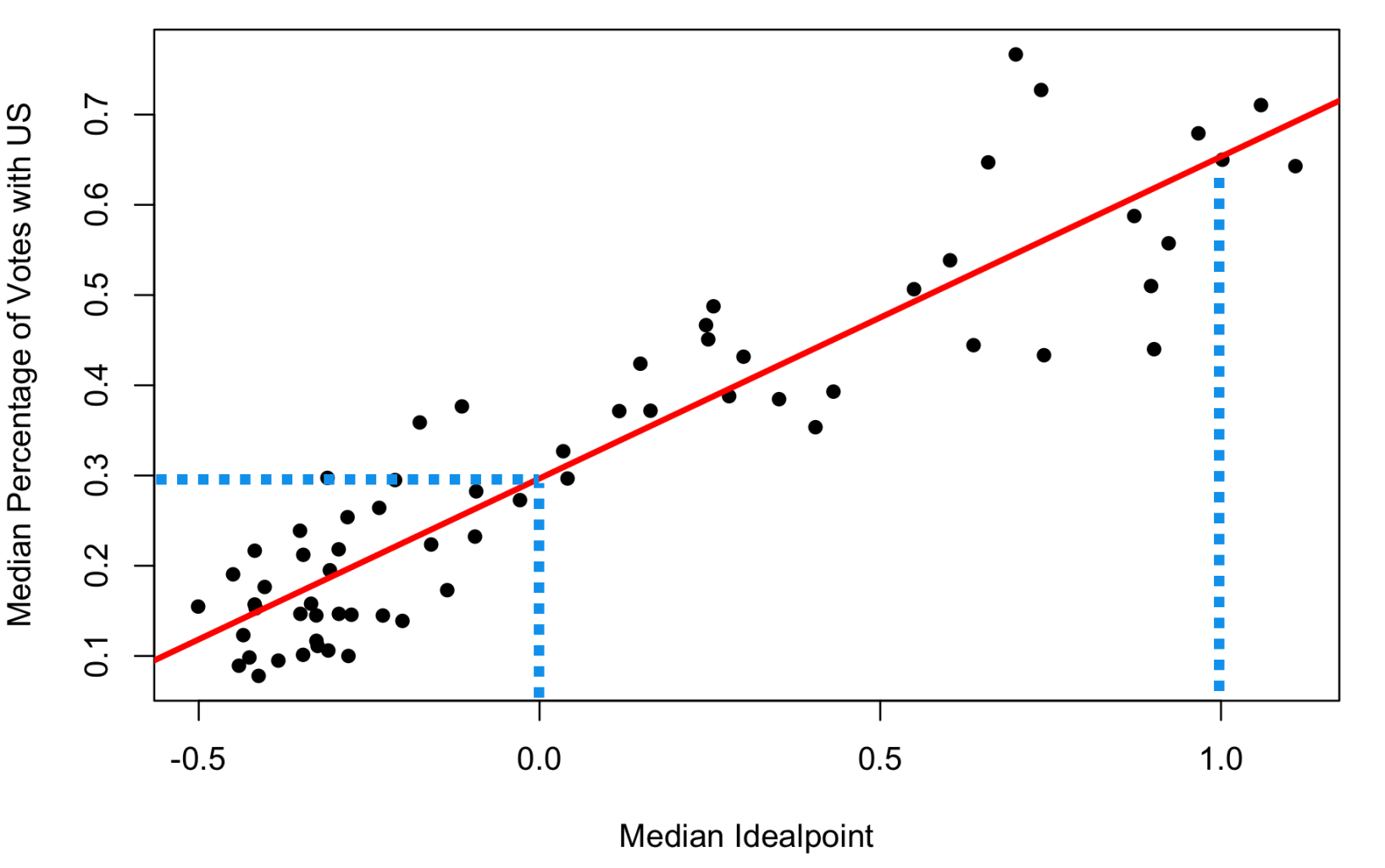
- On average, how much higher is the share of votes with the US for a country that is 1 on the ideal point scale, compared to one with a value of 0 on the ideal point scale?

UN VOTING



- On average, how much higher is the share of votes with the US for a country that is 1 on the ideal point scale, compared to one with a value of 0 on the ideal point scale?

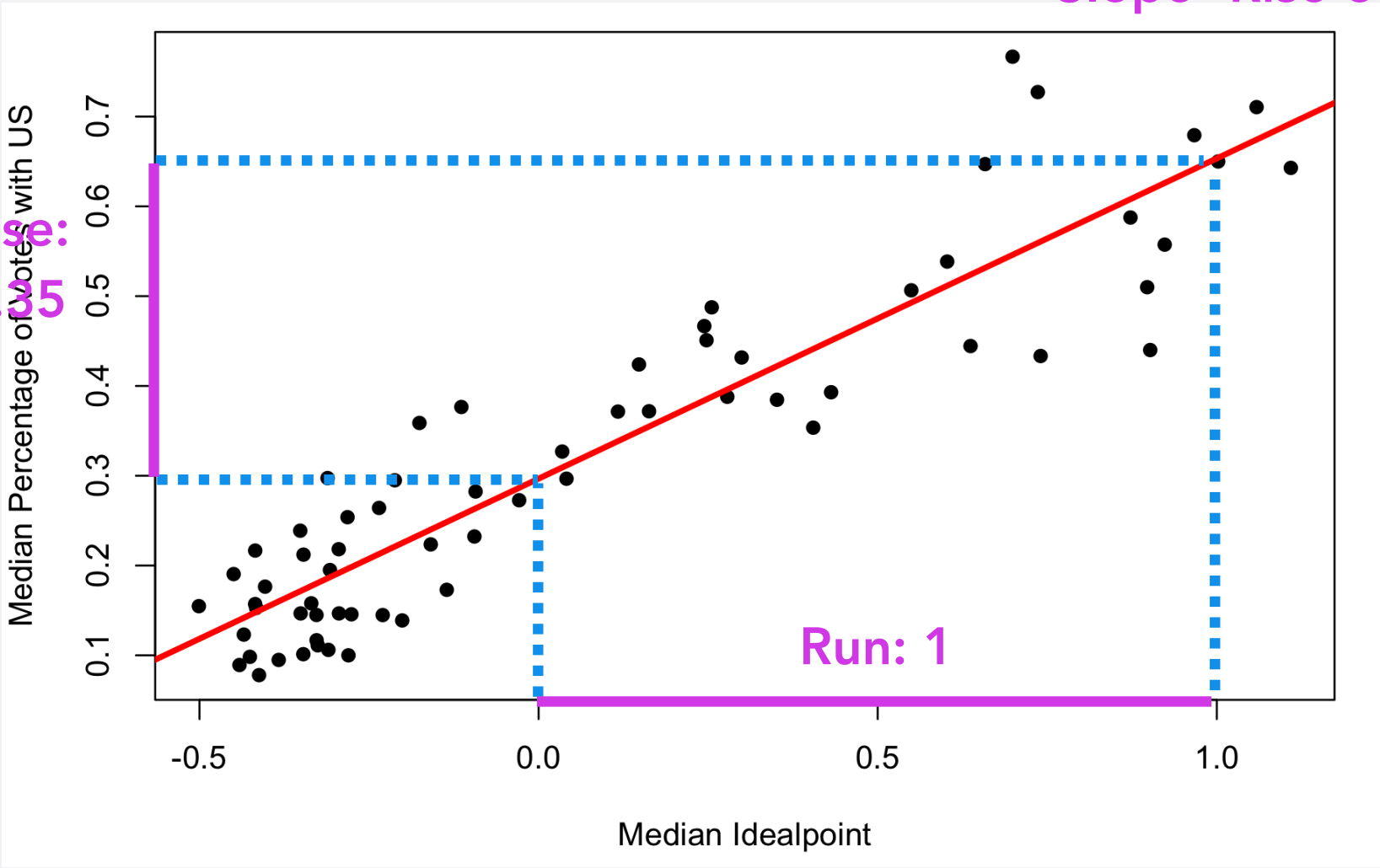
UN VOTING



- On average, how much higher is the share of votes with the US for a country that is 1 on the ideal point scale, compared to one with a value of 0 on the ideal point scale?

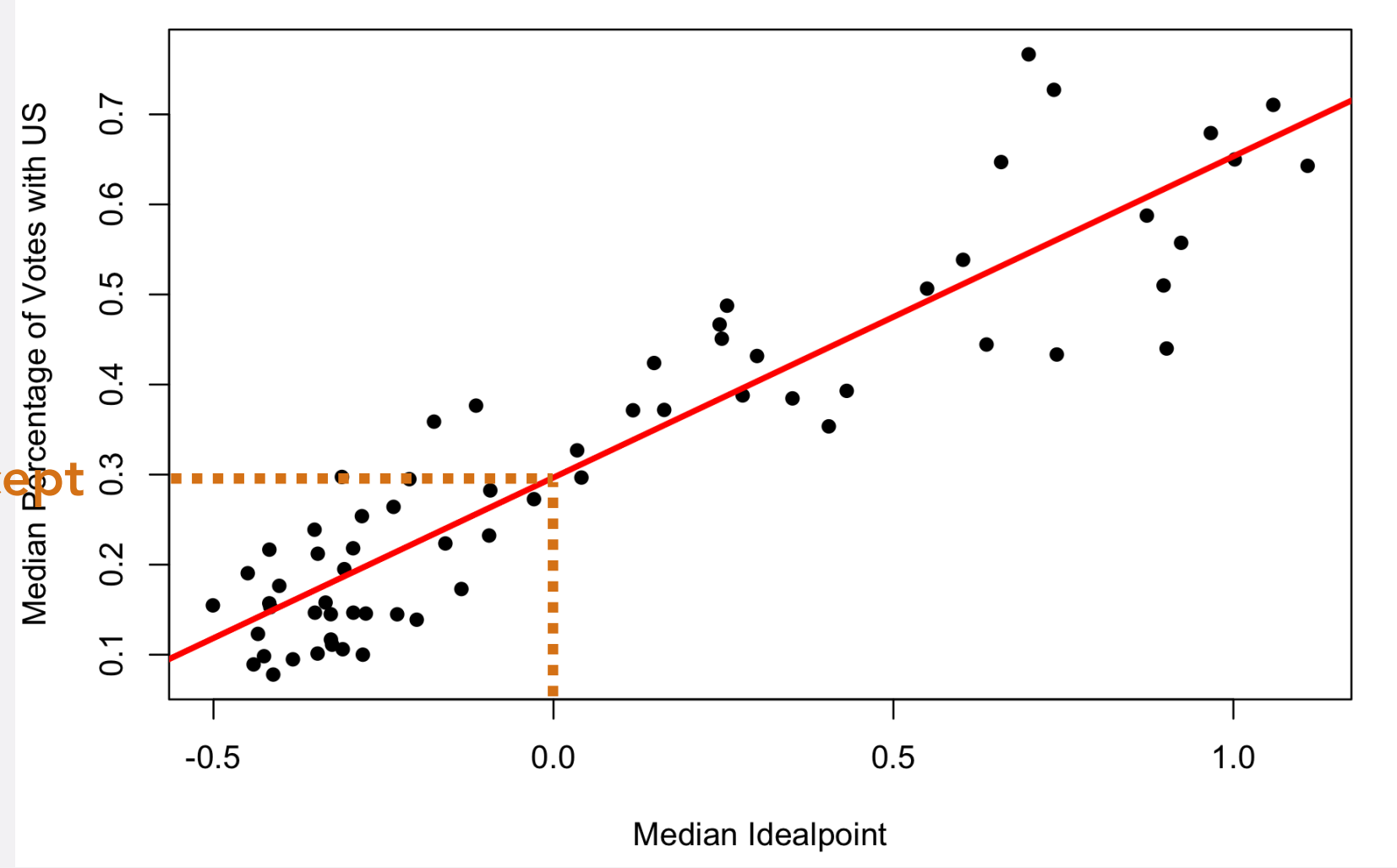
UN VOTING

Slope=Rise over run



Slope=Rise over run=0.35/1=0.35

UN VOTING



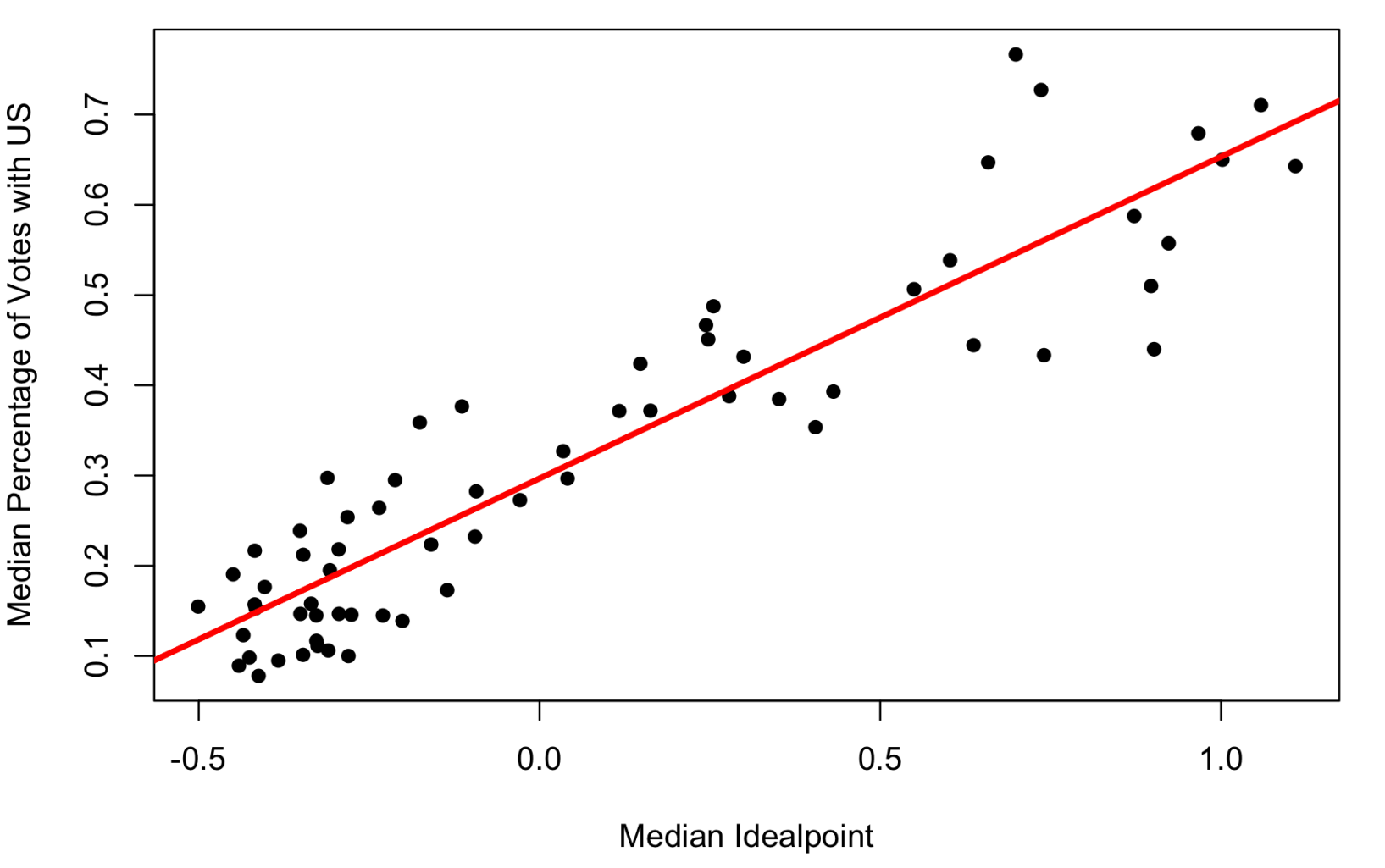
Intercept

Intercept=0.3

LINEAR REGRESSION

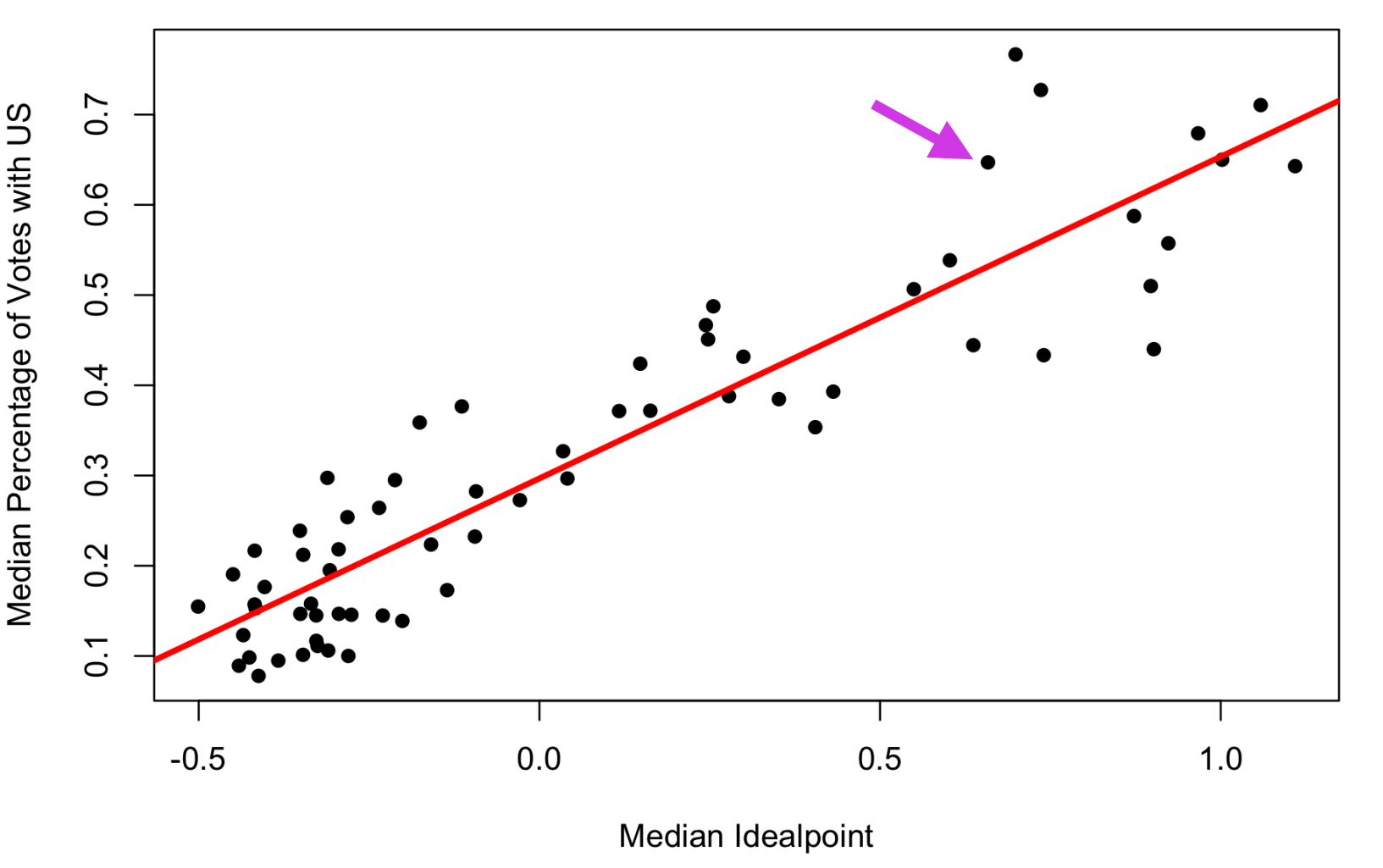
- Linear regression: Equation that tells us *direction* and *size* of relationship between independent variable (IV) and dependent variable (DV)
- DV = Intercept + Slope * IV + error

LINEAR REGRESSION



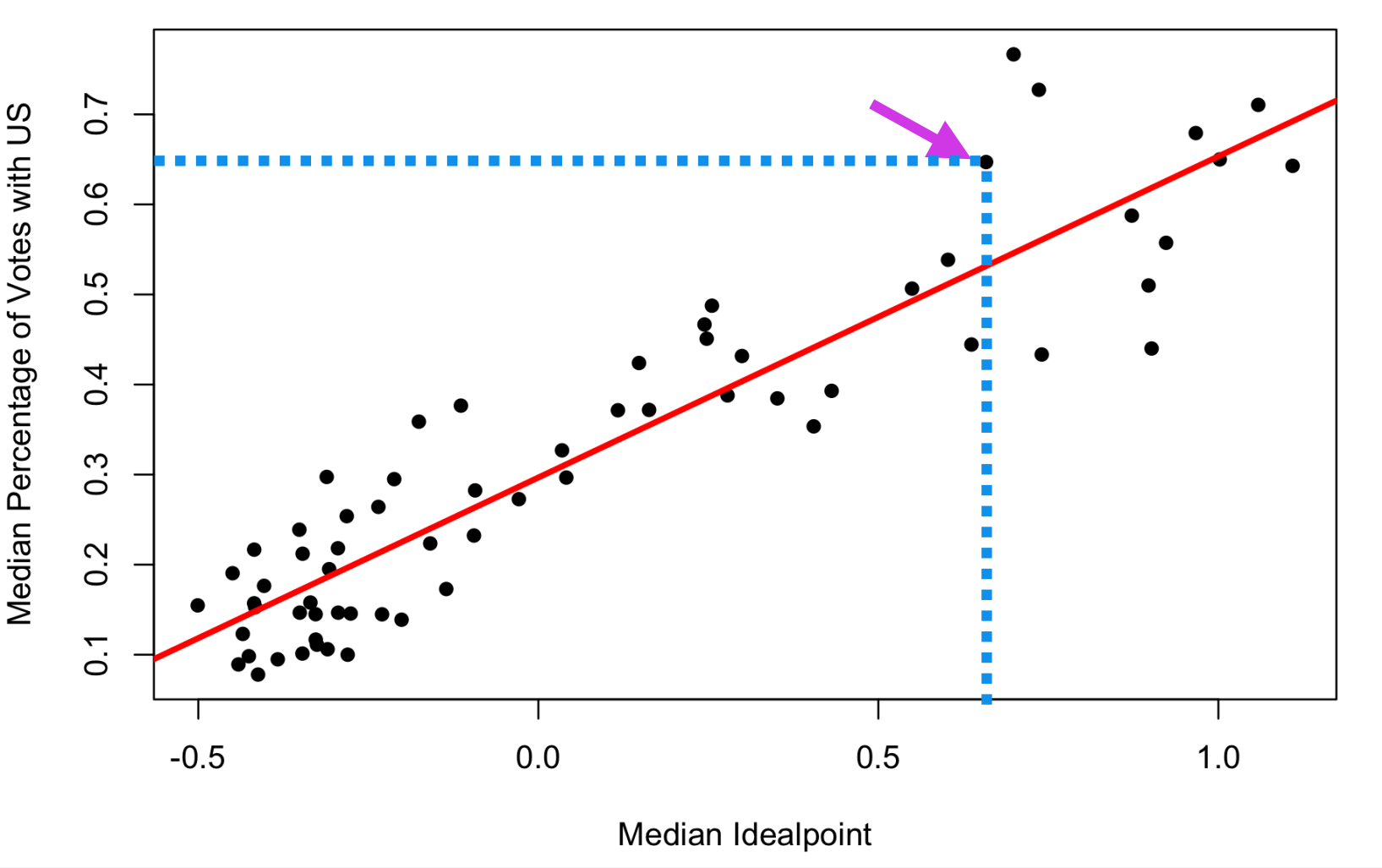
- Voting with US = 30 + 0.35 * Idealpoint + error

LINEAR REGRESSION



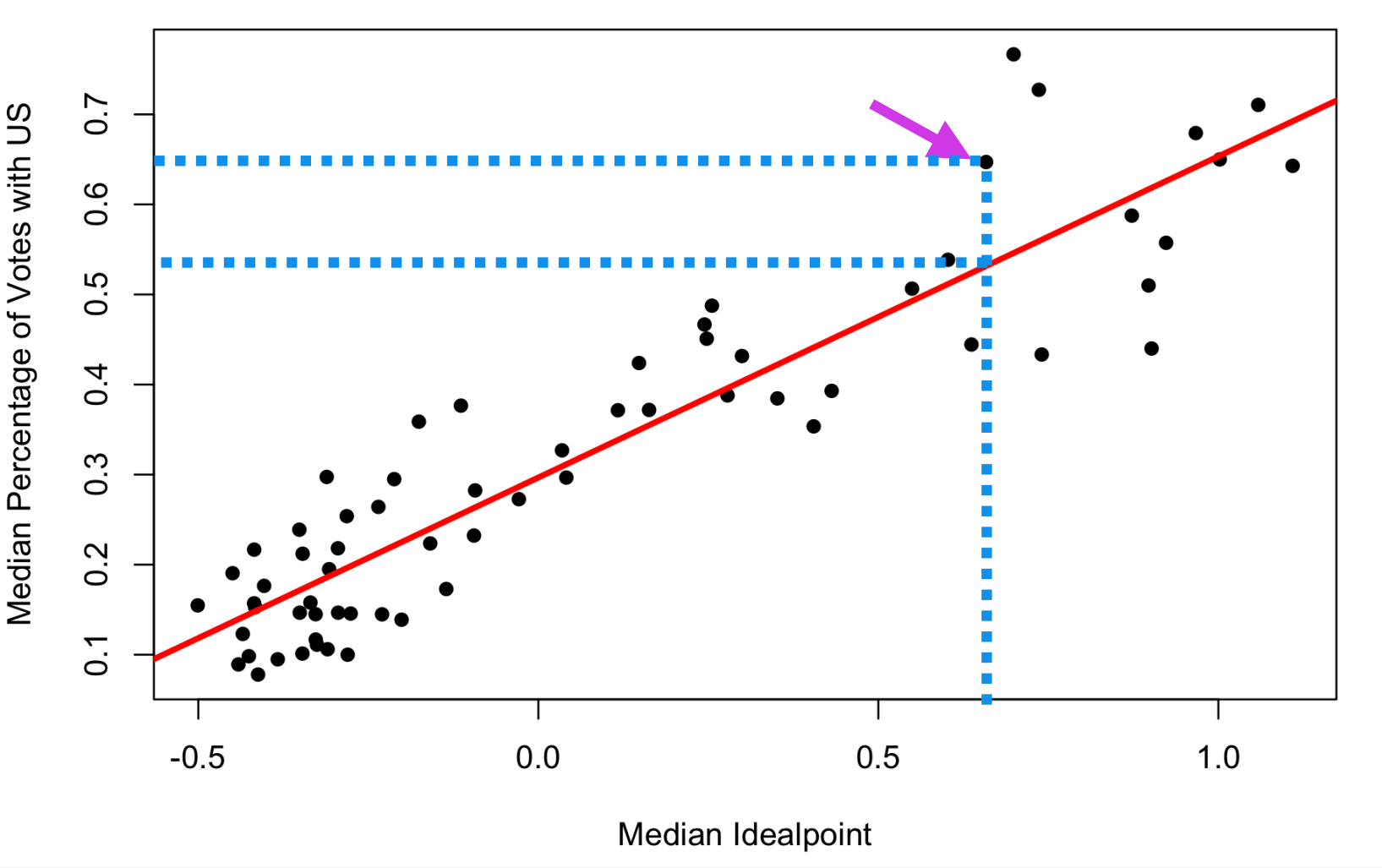
- Voting with US = 30 + 0.35 * Idealpoint + error

LINEAR REGRESSION



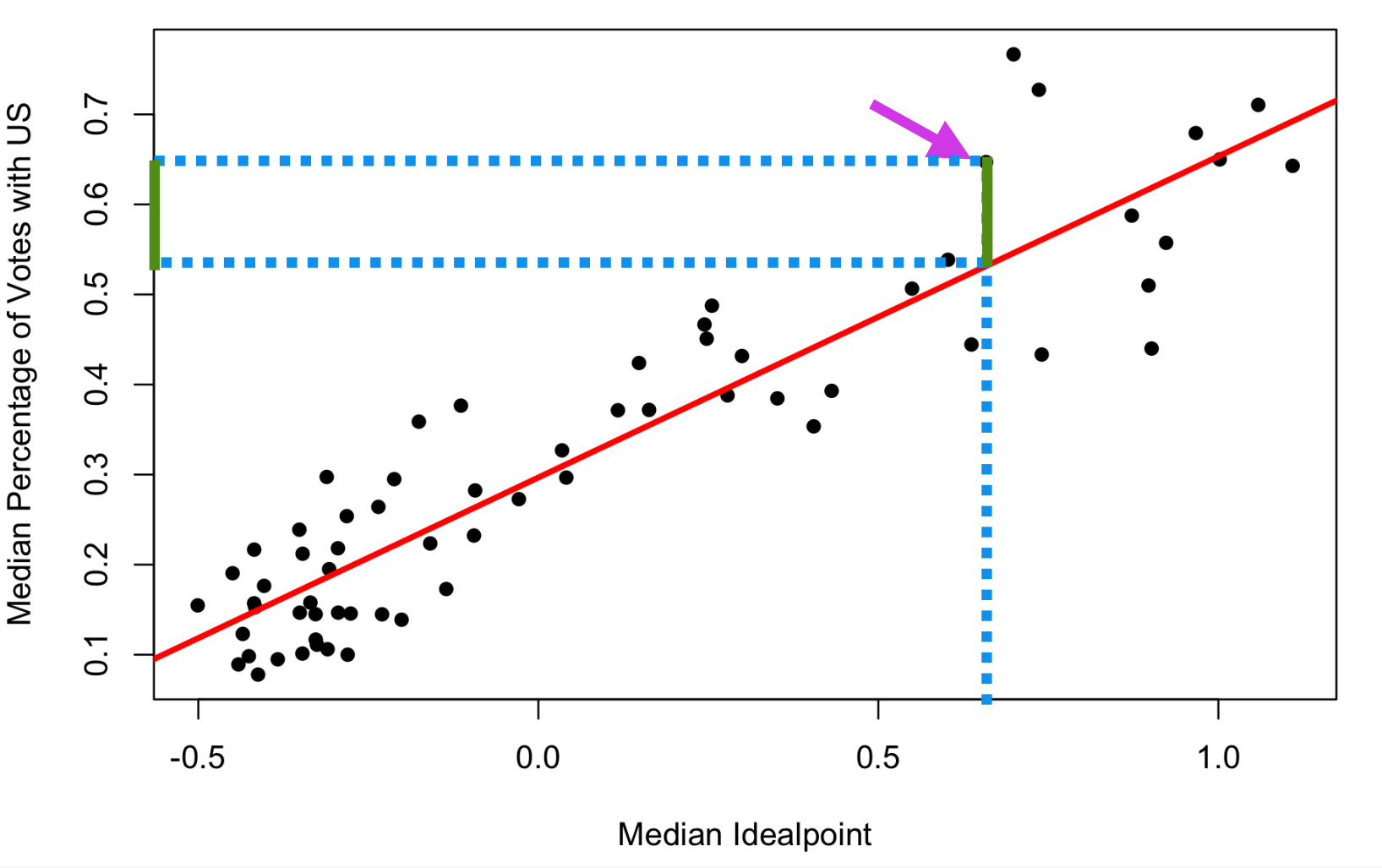
- Voting with US = 30 + 0.35 * Idealpoint + error

LINEAR REGRESSION



- Voting with US = 30 + 0.35 * Idealpoint + error

LINEAR REGRESSION



- Voting with US = 30 + 0.35 * Idealpoint + error

PREDICTION ERROR

- For each observation, we have a prediction error: $y - \hat{y}$
  - y: actual observed value
  - $\hat{y}$: predicted value (by regressions line)
  - $y - \hat{y}$: prediction error, residual
- We square the prediction errors: $(y - \hat{y})^2$
  - Squared prediction errors especially large for predictions that are way off
    - e.g. prediction error 2 vs. 20
    - squared prediction errors will be 4 vs. 400

BEST LINE

- The best line is the one with the smallest sum of squared prediction errors
- “Ordinary Least Squares” (OLS) Linear Regression

EXAMPLE

Which person is the most competent?

Figure 4.2. Example Pictures of Candidates Used in the Experiment. Source: A. Todorov et al. (2005) *Science*, vol. 308, no. 10 (June), pp. 1623–1626.

EXAMPLE

Table 4.4. Facial Appearance Experiment Data.

Variable	*Description*
`congress`	session of Congress
`year`	year of the election
`state`	state of the election
`winner`	name of the winner
`loser`	name of the runner-up
`w.party`	party of the winner
`l.party`	party of the loser
`d.votes`	number of votes for the Democratic candidate
`r.votes`	number of votes for the Republican candidate
`d.comp`	competence measure for the Democratic candidate
`r.comp`	competence measure for the Republican candidate

- face.csv

EXAMPLE

Table 4.5. 2012 US Presidential Election Data.

Variable	*Description*
`state`	abbreviated name of the state
`Obama`	Obama's vote share (percentage)
`Romney`	Romney's vote share (percentage)
`EV`	number of Electoral College votes for the state

- pres12.csv
- How does Obama's vote share in 2012 depend on his 2008 vote share?

EXAMPLE

Table 4.1. 2008 US Presidential Election Data.

Variable	*Description*
`state`	abbreviated name of the state
`state.name`	unabbreviated name of the state
`Obama`	Obama's vote share (percentage)
`McCain`	McCain's vote share (percentage)
`EV`	number of Electoral College votes for the state

- pres08.csv