

Comments on  
“Unfamiliar Names and  
Congressional Elections”

Spring 2012

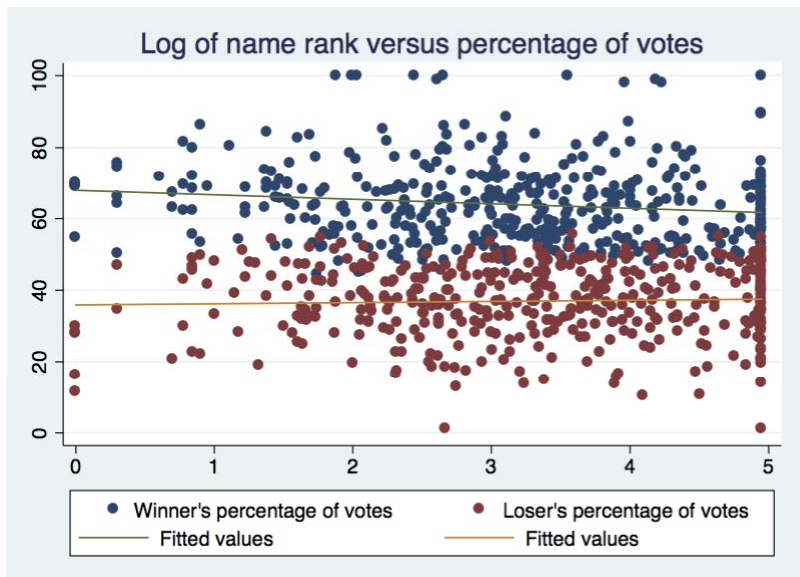
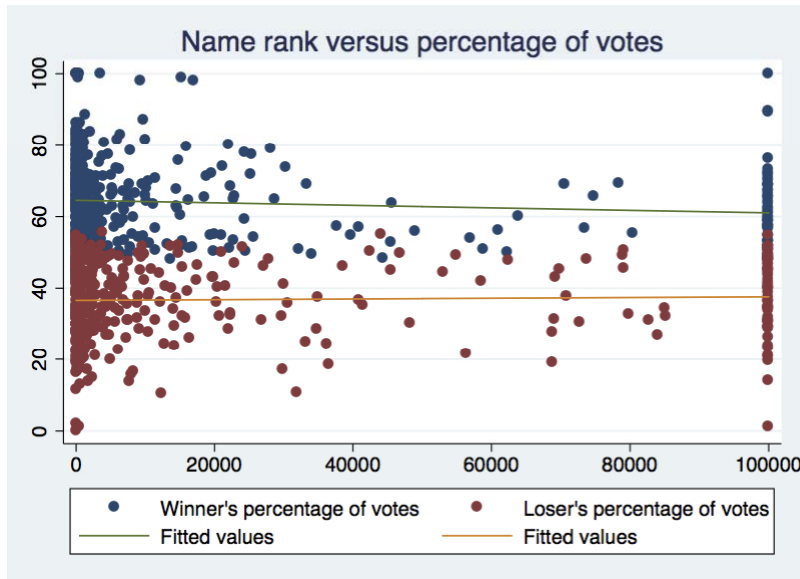
# Question

- Does having a less familiar name hurt candidates in House elections?

# Basic set-up

- Case selection: all House elections in the 2010 general election
- Unit of analysis: congressional election (435 cases)
- Main analysis: graph of winner's share vs measures of last name frequency.
- Data sources:
  - Vote share: official returns reported by House Clerk (excellent!)
    - However, a double check of the data revealed some errors, especially in coding open seat status.
  - Census Bureau last name data file (double excellent!)
    - However, I could not find the 2010 data from the link given

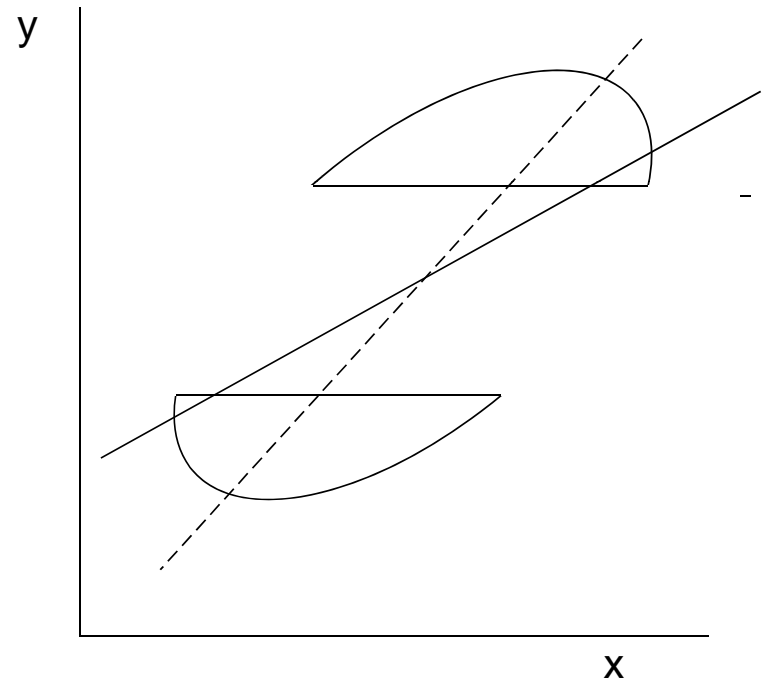
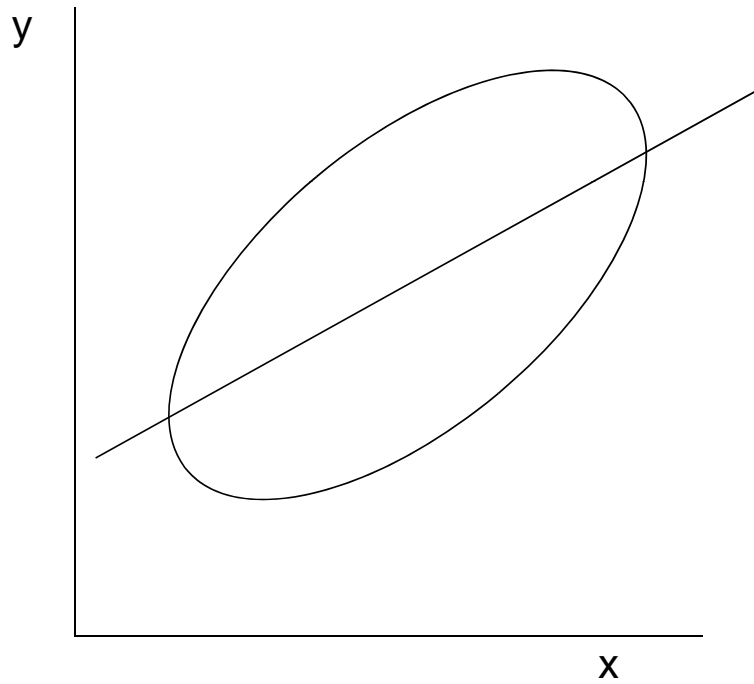
# Basic Analysis



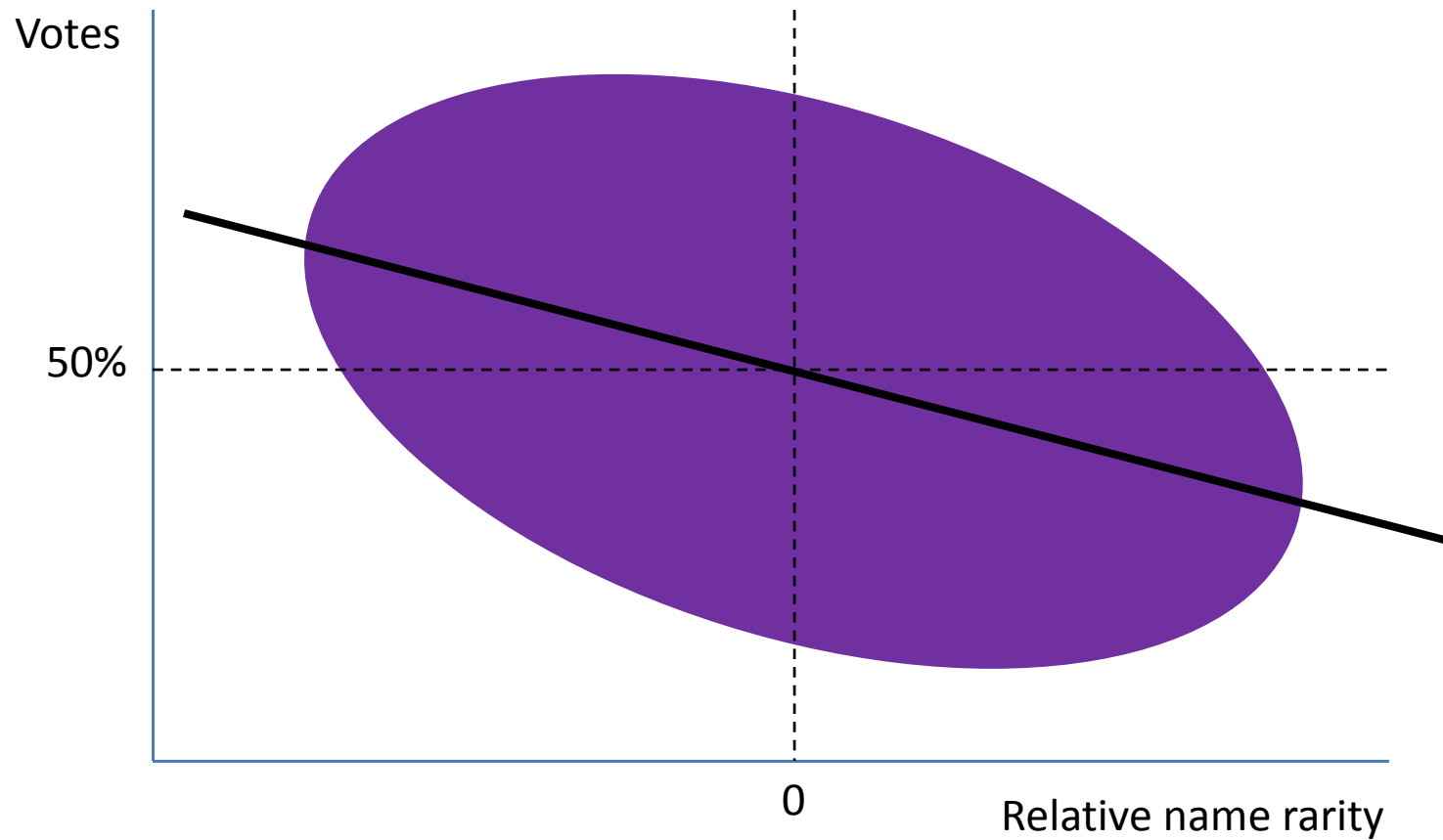
- Handling of names not on list
  - Try sensitivity analysis
  - Try omitting these
- Tiny coefficients:  $-0.0000351$  &  $0.0000102$
- *But* multiply these by 100k:  $-3.51$  &  $1.02$
  
- Transformed ind. var. to deal with skew
- Graph in terms of  $\log_{10}(x)$ , but regression in terms of  $\ln(x)$
- With  $\log_{10}$ , the slope coeff. Interpreted as  $\partial Y$  given an order-of-magnitude change in  $x$
- Either way, multiplied across the range of  $X$ , the effect is significant:  $-1.25 \times 5 = -6.25$  and  $0.328 \times 5 = 1.64$

Bigger issue: Sampling on the  
Dependent Variable, a Different Way

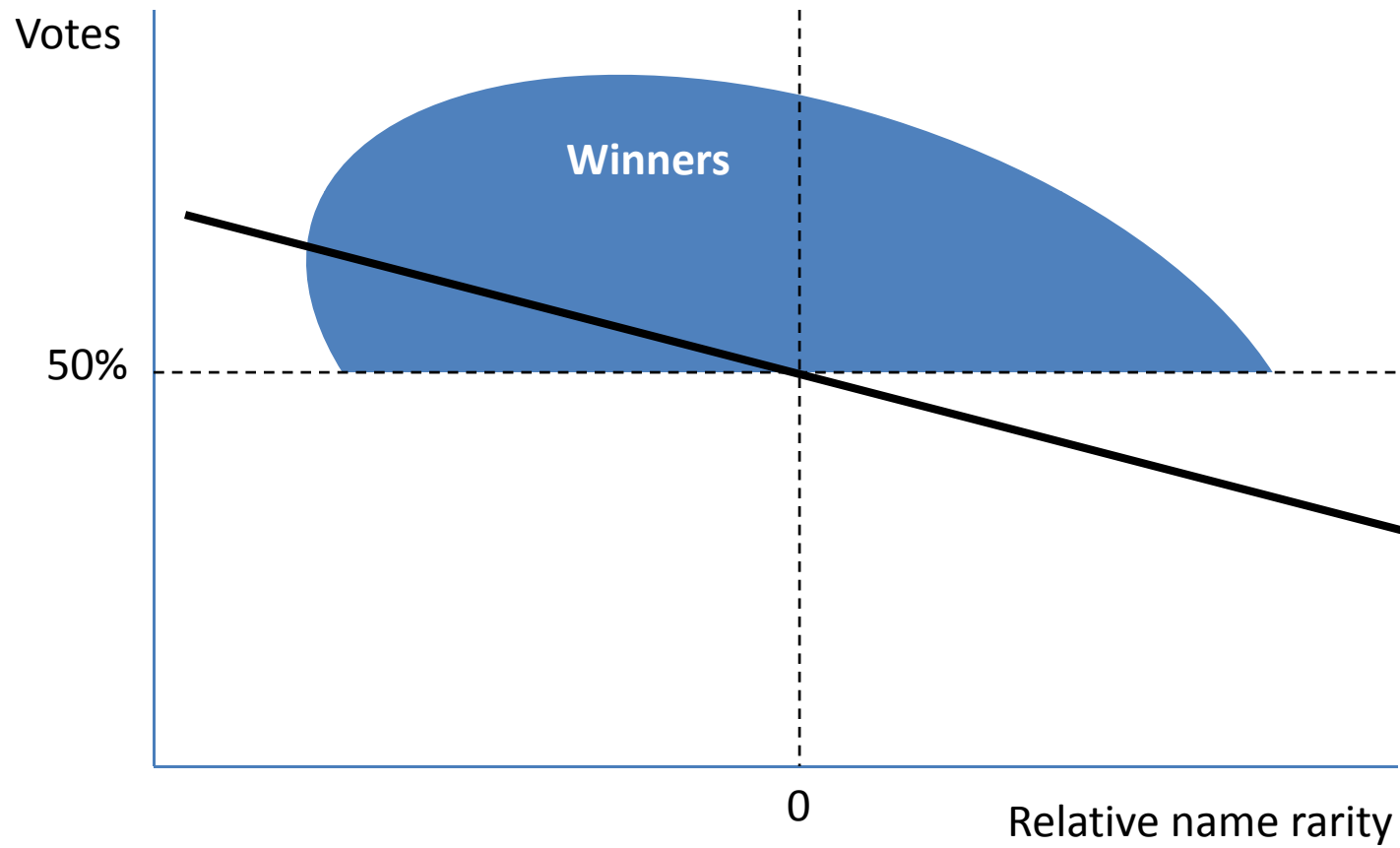
# Previous SODV Picture:



# A Different Kind of SODV Issue:

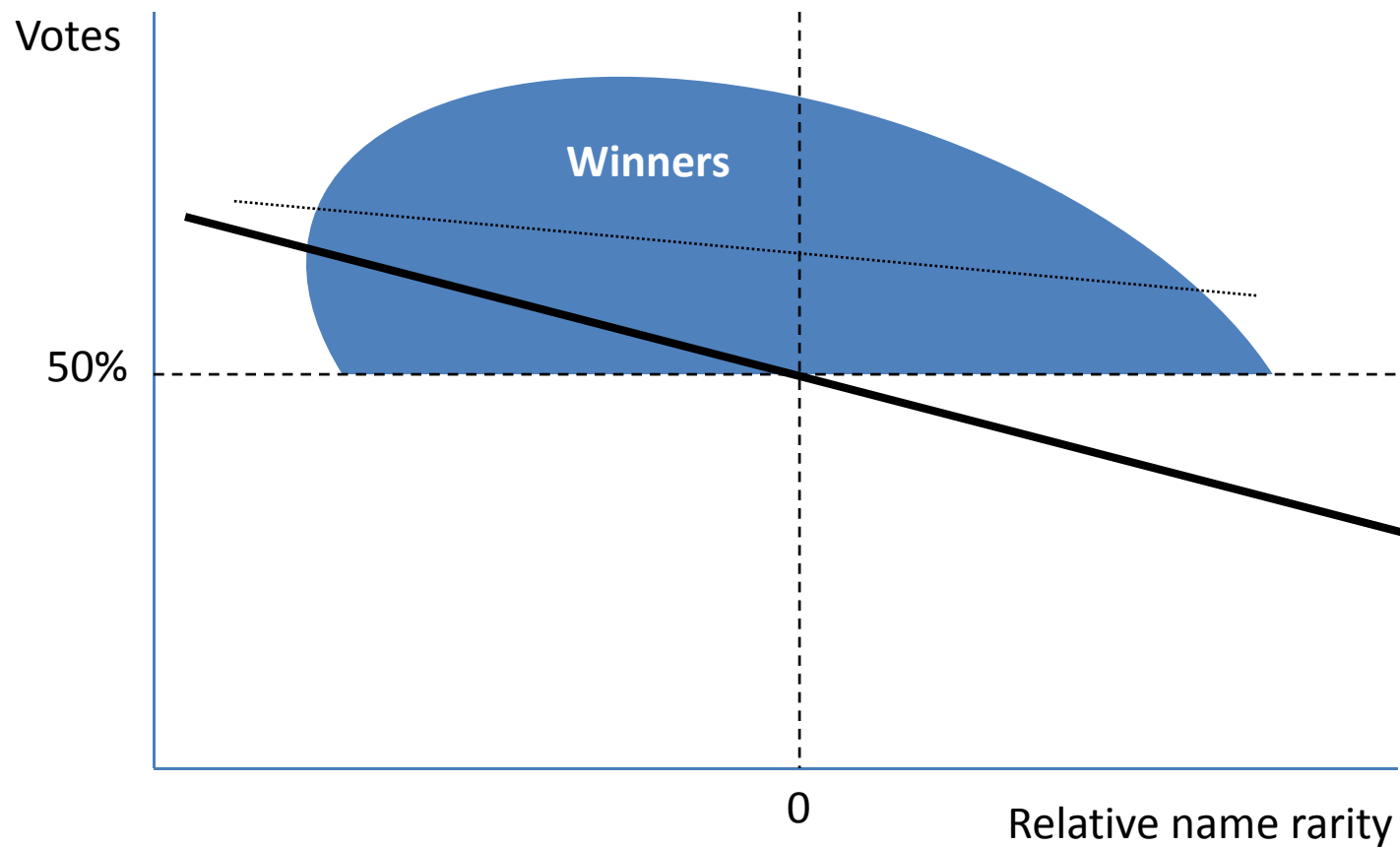


# A Different Kind of SODV Issue:

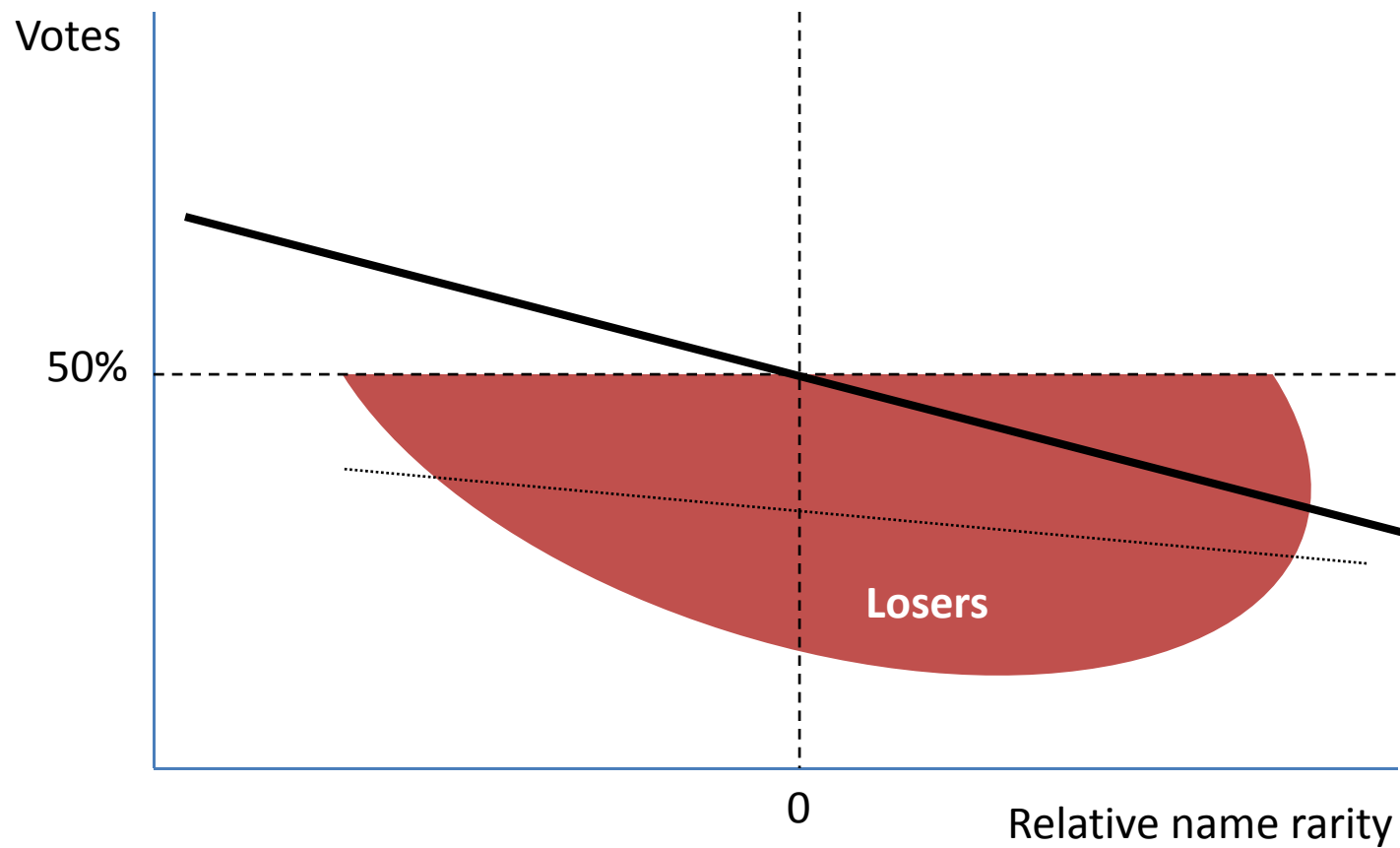




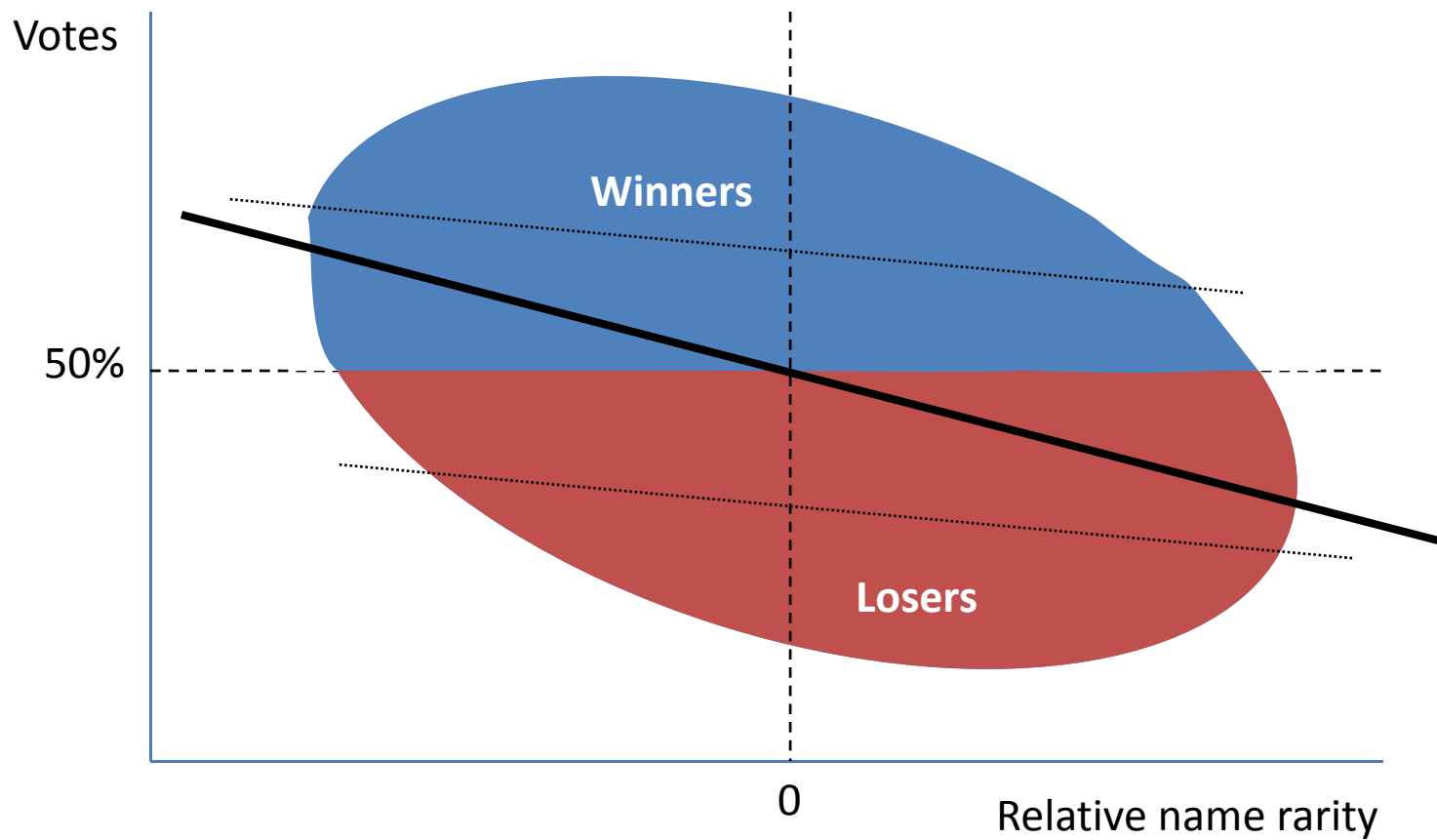
# A Different Kind of SODV Issue:



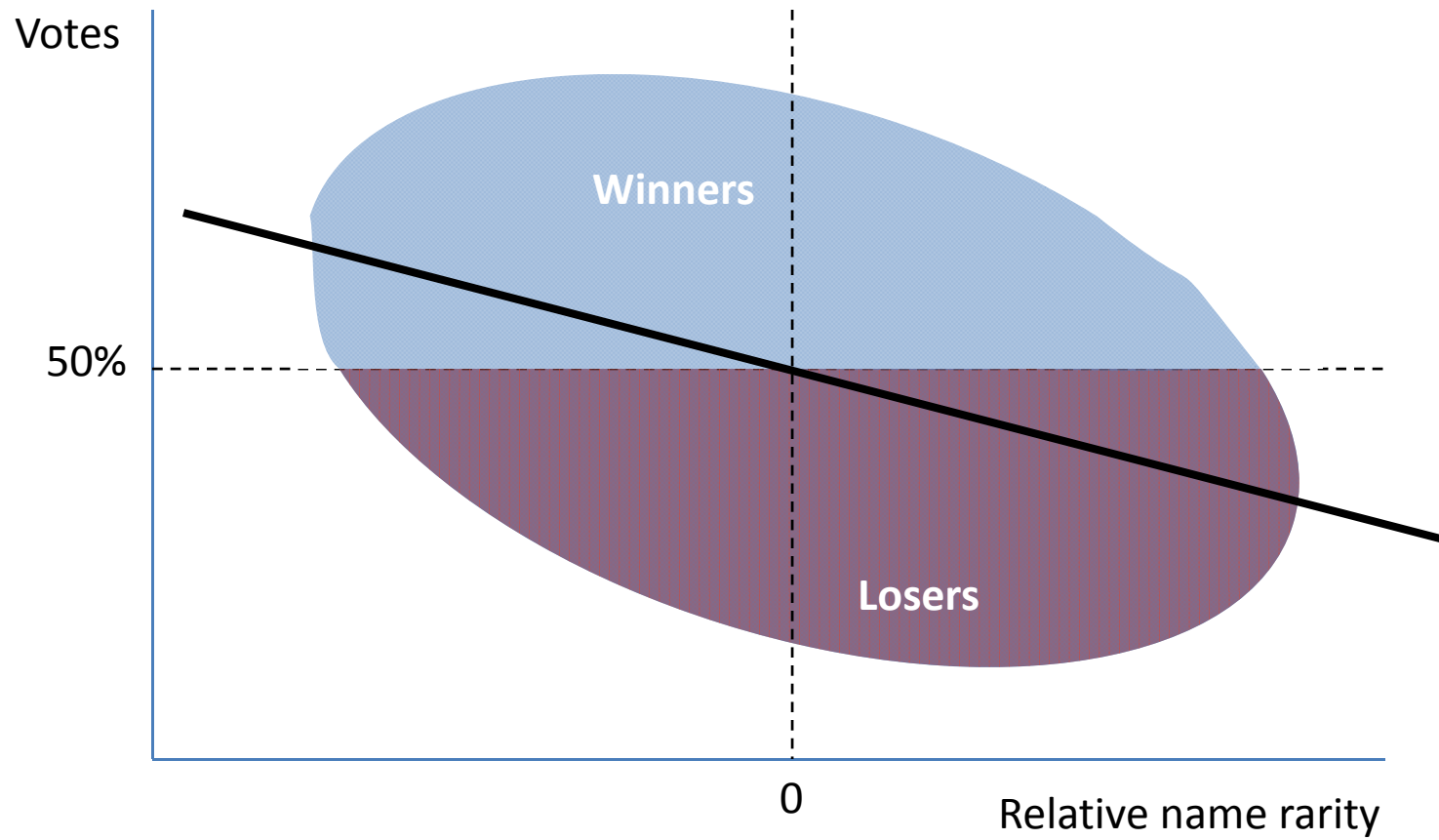
# A Different Kind of SODV Issue:



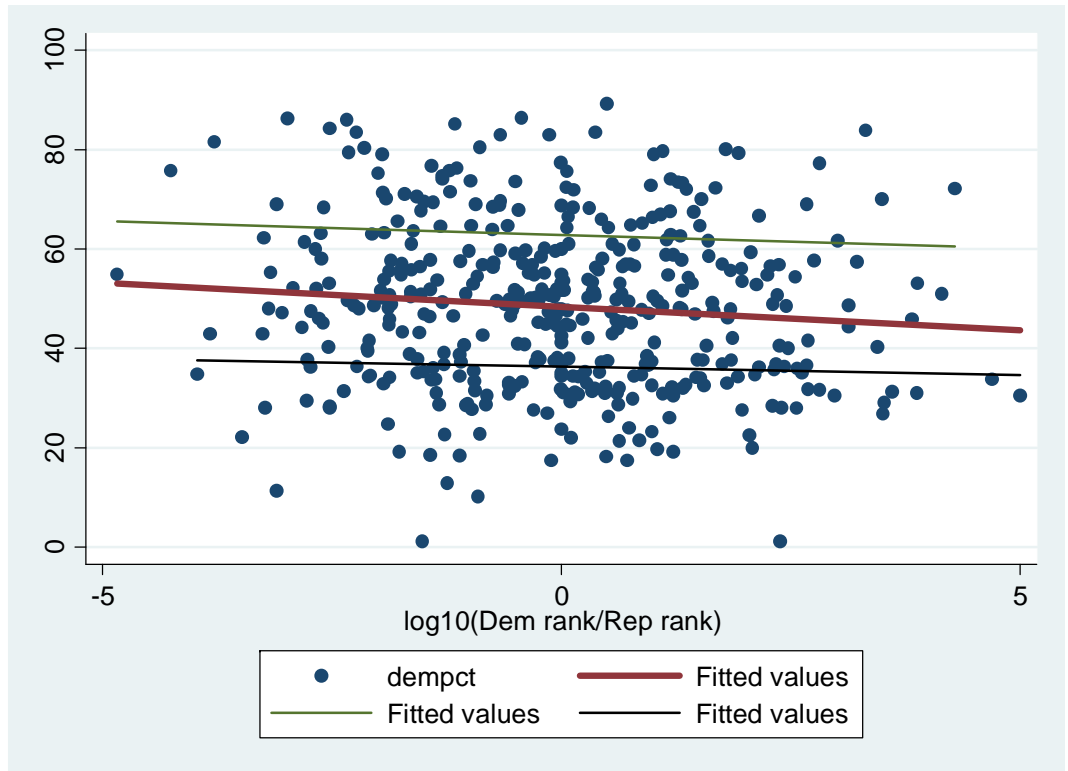
# A Different Kind of SODV Issue:



# The Data You Want



# The Unbiased Regression



	All	Winners	Losers
$\text{Log}_{10}(\text{Dem rank/Rep rank})$	-0.95 (0.47)	-0.55 (0.44)	-0.33 (0.36)
Constant	48.43 (0.80)	62.88 (0.73)	36.35 (0.62)
N	419	191	228
R <sup>2</sup>	.01	.01	.004
RMSE	16.4	10.0	9.4

# Deal with Unmeasured Rank Differently

	Original	Missing to 200k	Drop missing
$\text{Log}_{10}(\text{Dem rank/Rep rank})$	-0.95 (0.47)	-0.89 (0.45)	-0.98 (0.59)
Constant	48.43 (0.80)	48.44 (0.80)	48.19 (0.93)
N	419	419	320
R <sup>2</sup>	.01	.01	.01
RMSE	16.4	16.4	16.6

# What about incumbency?

	Original	Control for incumbency	Incumbent running	Open seat
$\log_{10}(\text{Dem rank/Rep rank})$	-0.95 (0.47)	-0.29 (0.31)	-0.10 (0.31)	-2.06 (1.22)
Democrat incumbent*	---	13.46 (0.56)	13.53 (0.54)	---
Constant	48.43 (0.80)	46.25 (0.53)	46.05 (0.54)	47.78 (2.00)
N	419	419	373	46
R <sup>2</sup>	.01	.59	.63	.06
RMSE	16.4	10.6	10.2	13.6

=1 if Dem. incumbent, -1 if Rep., 0 if open seat

