Los Angeles County Racially Polarized Voting Analysis for 2021 Redistricting

Bruce Adelson, Esq. and Jonathan N. Katz, Ph.D.

October 19, 2021

Summary of Analysis

- Analyzed all contested Los Angeles Board of Supervisor elections from 2012 to 2020.
- The election analysis reveals some statistical evidence of **racially polarized voting** (RPV).
- However, the RPV is not legally cognizable because cohesion is inconsistent and non-minority voters do not vote as a bloc to usually defeat minority preferred candidates.

Outline

- Methods for Ecological Inference
- Defining Racially Polarized Voting
- Results of Analysis

Methods for Ecological Inference

- The problem of inferring voting behavior from aggregate information is known as ecological inference.

- We are interested in estimating how groups of voters, say Latinos and Others (i.e., non-Latinos), voted in a given election when all we observe are:
  - precinct-level returns
  - demographic make-up of the precincts

Homogenous Precincts and the Method of Bounds

- A common starting point is to consider only **homogeneous precincts**.
- For example, if a precinct is 100% Latino, then we know what fraction Latinos voted for each candidate there.
- The problem is that it throws out most of the data.

Homogenous Precincts and the Method of Bounds

- We can use the intuition from the homogeneous precincts to place bounds on the level of support each group gives a candidate.

- Consider the following equation, which is true by definition, that relates the vote share of given candidate to the voting behavior of Latinos and Others (i.e., non-Latinos):

$$V_i = \lambda_i^L X_i + \lambda_i^O (1 - X_i), \tag{1}$$

- V_i is candidate's vote share in precinct i, X_i is the fraction of Latinos in the precinct and $(1 - X_i)$ is the fraction of Other voters (assuming only two groups). λ_i^L is the fraction of Latinos voting for the given candidate and similarly λ_i^O is the fraction of Others voting for the given candidate.

Homogenous Precincts and the Method of Bounds

- Consider homogeneous Latino precincts again. In these precincts, $X_i = 1$, so that the equation simplifies to $V_i = \lambda_i^L$.
- However, from these Latino homogeneous precincts can not say anything about the voting behavior of Other voters.

Homogenous Precincts and the Method of Bounds

- Now suppose a precinct is 60% Latino ($X_i = 0.6$) and the vote share is 50% ($V_i = 0.5$). Then:
  - Then at most $\frac{5}{6}$ths of the Latino voters could have voted for the candidate.
  - On the other hand, even if all of the Others voted for the candidate then at least $\frac{1}{6}$th of the Latinos would have had to vote for the candidate as well.
- Thus we know that proportion of Latinos voting for the candidate, λ_i^L, must be greater than $1/6$ and less than $5/6$ and λ_i^O can take on any value between zero and one.
- We actually know more than this: we know that the feasible values for this district must lie on the line segment, called a constraint line, defined by the bounds $(\frac{1}{6}, 1)$ and $(\frac{5}{6}, 0)$.

Homogenous Precincts and the Method of Bounds

- Duncan and Davis (1953) fully developed the method of bounds outlined above to analyze ecological data.
- Unfortunately, with a large number of precincts, it is difficult to make much direct use of these bounds since we need a way to combine them to understand typical behavior in the district.
- These bounds do, however, provide useful information as we will see later.

Ecological or Goodman's Regression

- An alternative approach that examines all precincts simultaneously was developed by Goodman (1959).
- It is referred to in the literature as **ecological regression** or Goodman's regression.
- Like the method of bounds, it is based on the identity in Equation (1).

Ecological or Goodman's Regression

- Suppose that the fraction of support for a given candidate for both Others and Latinos was the same across all precincts in the district.
- Formally: $\lambda_i^A = \lambda^A$ and $\lambda_i^O = \lambda^O$ for every precinct i.
- Then we could estimate these fractions by choosing the best fitting line to the precinct-level data. This is just a standard linear regression, the most commonly used statistical procedure in the social sciences.

Ecological or Goodman's Regression

- It is important to note that ecological regression can produce widely inaccurate estimates of group voting behavior (King 1997). In fact, Goodman (1959) warned against it use.
- First, the assumption that the fraction of group support is constant across every precinct is highly implausible.
- Second, ecological regression does not use the bounds information at the precinct level. It does not even enforce the overall bounds that the average fraction of a group's support for a given candidate must be between 0 and 1.

Ecological Inference/EI

- King (1997) has developed an alternative approach called **Ecological Inference** or **EI**.
- While the technical details are complex, its advantage is that it uses all available information to generate more accurate estimates of voting behavior from aggregate data.
- It is basically a way to combine the regression approach of Goodman (1959) with the bounds from Duncan and Davis (1953).
- Further, it allows the estimates to vary (systematically) across precincts.

Ecological Inference/EI

The intuition behind is as follows:

- We calculate the constraint lines for every precinct.
- We then choose as our estimate for a given precinct a point on its constraint line near the center of the intersection of all of the other constraint lines.
- The actual point chosen is based on a standard statistical model.

More than Two Groups or Two Candidates

- The previous discussion on the development of methods for ecological inference assumed that we only had two groups and two candidates (or vote choices).
- Accommodating more than two groups is rather straight-forward, although notation and intuition become more complicated, especially for the constraint lines.

More than Two Groups or Two Candidates

- Allowing for more than two candidates or vote choices is a bit more complicated.
- In the special case of only two choices, we only need to model the vote share going to one of them since we then automatically know the fraction going to the other candidate.
- If we add a third choice, then we need to model the vote share going to any two of the options and then we get third by subtraction.
- There will always be more than two vote choices because some individuals will choose not to vote in particular race.
- We need to account for this abstention in order to make proper inferences although we condition them out in the presentation of results.

Defining Racially Polarized Voting

From a statistical perspective, we can define evidence of racially polarized voting as:

- In the case of two groups: RPV occurs when λ_i^L and λ_i^O are on opposite sides of 0.5 — e.g., $\lambda_i^L > 0.5$ and $\lambda_i^O < 0.5$
- That is, a majority of one group voting for one candidate and the majority of the other group voting for the opposite candidate.
- If this holds, then the larger the difference between support levels, the greater the level of polarization.
- Since we are dealing with statistical estimates, this difference must be greater than the statistical uncertainty in the estimates.

Defining Racially Polarized Voting

- Generalizing to the case of more than two choice is more complicated because we need to consider **cohesion**.
- We only two choices we get cohesion for free, since one of the choices must garner a majority of support of the group and is the clearly preferred group choice.
- This is not the case when we have more than two choices. That is, the preference of the group may split across several choices with no clearly preferred candidate.

Defining Racially Polarized Voting

- The previous discussion, however, does not establish whether the statistical evidence racially polarized voting is legally cognizable.

- For that, we need to look at the *Gingles* test:
  - Is the minority group cohesive in its voting behavior?
  - Do White/Other voters cohesively oppose the minority preferred candidate and do these voters vote as a bloc to usually defeat the minority preferred candidate?

Results of Analysis

- Examined all 13 contested elections for the Los Angeles Board of Supervisors from 2012 to 2020
- There is some statistical evidence of racially polarized voting in one election, the 2016 General Election for 5th district.
- However, the RPV is not legally cognizable because cohesion is inconsistent and White/Other voters do not vote as a bloc to usually defeat minority preferred candidates.

Results of Analysis

Interesting elections:

- 2016 General Election, 5th District: Park was preferred candidate of Latino voters, but Barger was preferred by Asian and Other voters.
- 2020 Primary, 2nd District: Jeong was clearly the preferred candidate of Asians, but neither Others nor Latinos were cohesive in their voting choice.
- 2020 Primary, 5th District: Barger was the clearly the preferred candidate for Other voter, but neither Latinos nor Asians were cohesive in their voting choices.
- 2016 Primary, 4th district: Hahn was the clearly the preferred candidate of Other voters, but a plurality of Latino voters preferred Pacheco. However, a third of Latinos also voted for Hahn and some for Napolitano.

Results of Analysis

EI Results for 2020 General for Supervisor 2nd District

	Latino	Asian	Other
Wesson, Jr.	45.6	11.6	37.3
	(44.4, 46.9)	(8.0, 15.2)	(36.4, 38.0)
Mitchell	54.4	88.4	62.7
	(53.1, 55.6)	(84.8, 92.0)	(62.0, 63.6)

Results of Analysis

EI Results for 2020 Primary for Supervisor 2nd District

	Latino	Asian	Other
Perry	6.1 (5.3, 6.8)	3.5 (2.6, 4.9)	15.4 (15.0, 15.8)
Nuño	20.1 (19.6, 20.5)	1.8 (1.5, 2.1)	1.1 (0.9, 1.2)
Mitchell	10.7 (9.9, 11.6)	19.7 (15.4, 24.0)	37.9 (37.4, 38.4)
Rigard	13.5 (13.0, 14.0)	3.0 (2.3, 3.7)	2.2 (2.0, 2.4)
Wesson Jr.	17.6 (16.4, 18.6)	2.8 (2.2, 3.7)	37.9 (37.4, 38.5)
Jeong	5.8 (5.4, 6.2)	64.7 (61.4, 68.2)	1.3 (1.1, 1.5)
Robles	26.3 (25.5, 27.0)	4.6 (2.7, 6.5)	4.2 (3.9, 4.6)

Results of Analysis

EI Results for 2020 Primary for Supervisor 4th District

	Latino	Asian	Other
Hahn	72.6	85.6	76.5
	(71.6, 73.7)	(81.5, 89.8)	(75.8, 77.2)
Washington	27.4	14.4	23.5
	(26.3, 28.4)	(10.2, 18.5)	(22.8, 24.2)

Results of Analysis

EI Results for 2020 Primary for Supervisor 5th District

	Latino	Asian	Other
Barger	44.5	40.6	65.0
	(42.4, 46.6)	(36.2, 44.2)	(64.4, 65.8)
Harabedian	14.9	39.0	20.2
	(13.3, 16.8)	(35.6, 42.5)	(19.6, 20.7)
Park	40.6	20.4	14.8
	(38.8, 42.5)	(17.5, 23.6)	(14.2, 15.3)

Results of Analysis

EI Results for 2018 Primary for Supervisor 3rd District

	Latino	Asian	Other
Glaser	15.1	5.6	9.7
	(14.4, 15.9)	(4.2, 7.6)	(9.5, 9.9)
Kuehl	57.8	78.5	79.4
	(56.6, 59.0)	(72.5, 84.8)	(79.1, 79.8)
Preven	27.0	15.9	10.9
	(26.1, 28.0)	(9.5, 22.1)	(10.5, 11.2)

Results of Analysis

EI Results for 2016 General for Supervisor 4th District

	Latino	Asian	Other
Hahn	60.9	51.3	54.4
	(60.0, 61.8)	(48.4, 54.7)	(53.7, 55.1)
Napolitano	39.1	48.7	45.6
	(38.2, 40.0)	(45.3, 51.6)	(44.9, 46.3)

Results of Analysis

EI Results for 2016 General for Supervisor 5th District

	Latino	Asian	Other
Barger	36.9	57.2	66.1
	(35.6, 38.3)	(54.2, 60.0)	(65.5, 66.6)
Park	63.1	42.8	33.9
	(61.7, 64.4)	(40.0, 45.8)	(33.4, 34.5)

Results of Analysis

EI Results for 2016 Primary for Supervisor 4th District

	Latino	Asian	Other
Pacheco	45.1	7.0	2.4
	(44.5, 45.8)	(5.0, 8.9)	(2.0, 2.8)
Hahn	36.7	46.6	52.6
	(35.8, 37.7)	(41.6, 50.5)	(51.8, 53.4)
Napolitano	18.1	46.5	45.0
	(17.2, 19.0)	(42.7, 50.5)	(44.3, 45.8)

Results of Analysis

EI Results for 2016 Primary for Supervisor 5th District

	Latino	Asian	Other
Kahlon	2.6 (2.3, 2.8)	1.8 (1.4, 2.2)	0.8 (0.8, 0.9)
Najarian	3.8 (3.1, 4.7)	10.5 (8.1, 13.0)	15.8 (15.4, 16.1)
Malone	6.8 (6.5, 7.1)	1.8 (1.4, 2.1)	1.2 (1.1, 1.3)
Park	26.4 (25.3, 27.4)	21.3 (18.5, 23.6)	11.6 (11.3, 12.0)
Barger	27.8 (26.6, 29.3)	36.8 (33.5, 40.0)	29.6 (29.1, 30.0)
Englander	7.3 (6.3, 8.5)	6.0 (4.2, 8.3)	14.0 (13.7, 14.3)
Carr	19.4 (18.2, 20.5)	5.9 (4.0, 7.8)	9.5 (9.2, 9.9)
Huff	5.9 (4.8, 7.3)	16.0 (13.1, 18.6)	17.3 (16.9, 17.7)

Results of Analysis

EI Results for 2014 General for Supervisor 3rd District

	Latino	Asian	Other
Kuehl	50.9	57.4	53.1
	(49.7, 52.2)	(47.4, 66.0)	(52.7, 53.5)
Shriver	49.1	42.6	46.9
	(47.8, 50.3)	(34.0, 52.6)	(46.5, 47.3)

Results of Analysis

EI Results for 2014 Primary for Supervisor 1st District

	Latino	Asian	Other
Hood	8.8	19.7	24.8
	(8.2, 9.4)	(17.2, 22.1)	(23.8, 25.8)
Solis	77.2	61.4	62.7
	(76.5, 78.0)	(58.2, 64.6)	(61.5, 64.0)
Gomez	14.0	18.9	12.5
	(13.4, 14.6)	(16.1, 21.5)	(11.5, 13.5)

Results of Analysis

EI Results for 2014 Primary for Supervisor 3rd District

	Latino	Asian	Other
Preven	1.9	4.4	1.5
	(1.5, 2.3)	(3.6, 5.9)	(1.4, 1.6)
Melendez	22.0	5.2	0.8
	(21.2, 22.8)	(3.9, 7.2)	(0.7, 0.9)
Fay	4.7	15.6	2.8
	(4.0, 5.4)	(12.7, 18.8)	(2.6, 2.9)
Conley Ulich	15.2	28.0	4.1
	(14.3, 16.2)	(23.7, 32.2)	(3.8, 4.3)
Kremer	2.7	3.8	3.4
	(2.3, 3.3)	(3.1, 4.7)	(3.3, 3.5)
Duran	13.2	8.9	17.2
	(12.1, 14.2)	(6.5, 12.7)	(16.9, 17.5)
Shriver	23.0	21.4	30.4
	(21.6, 24.4)	(16.0, 26.1)	(30.0, 30.8)
Kuehl	17.3	12.6	39.9
	(15.9, 18.8)	(7.6, 16.9)	(39.5, 40.2)

Results of Analysis

EI Results for 2012 Primary for Supervisor 5th District

	Latino	Asian	Other
Antonovich	71.9	85.3	80.2
	(68.9, 74.8)	(80.9, 89.2)	(79.7, 80.7)
Kahlon	28.1	14.7	19.8
	(25.2, 31.1)	(10.8, 19.1)	(19.3, 20.3)