PROJECT 3 SOLUTION SET

MGGG

Section 2. Explore Prepared Data

1. Example from suggested column combinations: Everett.csv/ P_SIMONELLI / P_HISP/ TOT_CCAL

Simonelli	Homogeneous precincts	Goodman ER	Ecol Inf
All but Hispanic support	0.073	0.085	0.099
Hispanic support	0.077	0.055	0.005

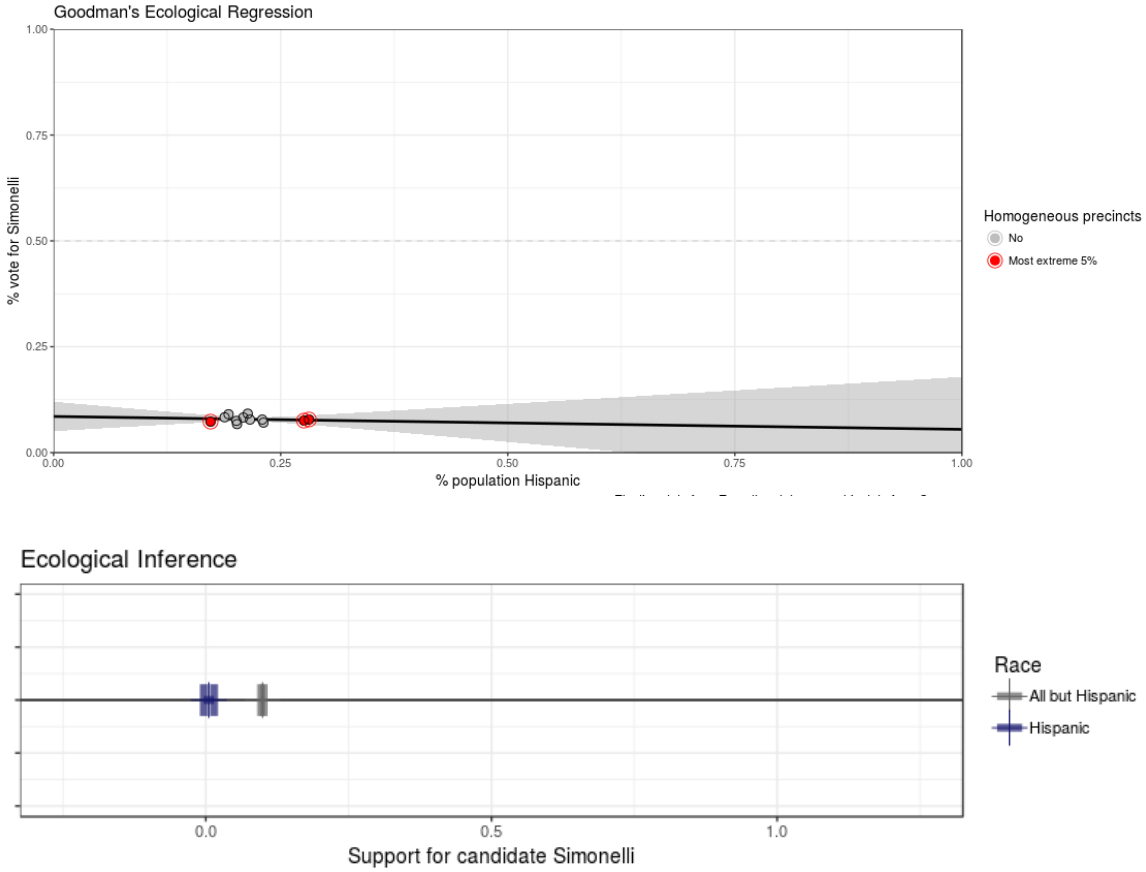


Figure 1. Results from running the Shiny app to assess Hispanic support for Simonelli in Everett. In the ER plot, red dots represent the most homogeneous precincts - the 5% with the highest Hispanic share of the population the 5% with the lowest Hispanic share. The EI figure shows confidence intervals around the estimated level of support among Hispanic and non-Hispanic voters.

Figure 1 shows very low support among both Hispanic voters and non-Hispanic voters, with ER and EI values under 10% for both groups. The EI confidence intervals suggest racial polarization, but only in the

sense that there is high confidence the groups have different levels of support for Simonelli. Because of the low levels of support overall, this does not mean that there isn't another candidate that Hispanics and non-Hispanics both heavily support.

It's important to note characteristics of the data set we're using. It's a small data set (only 12 points) and its clustered around precincts with about 20% Hispanic-share of the population. That means we are trying to predict what would happen in a 100% Hispanic precinct based only on precincts that are 20% Hispanic! This is a common problem with ER and EI; a small data set, or one that is clustered towards low population-share precincts can make for poor model fitting.

2. Example not included in the suggested column combinations: Lowell_EI_17.csv/ PSC_LAY / CVAPASIAN/ TOT_SC

Lay	Homogeneous precincts	Goodman ER	Ecol Inf
All but Asian support	0.104	0.117	0.068
Asian support	0.106	0.088	0.305

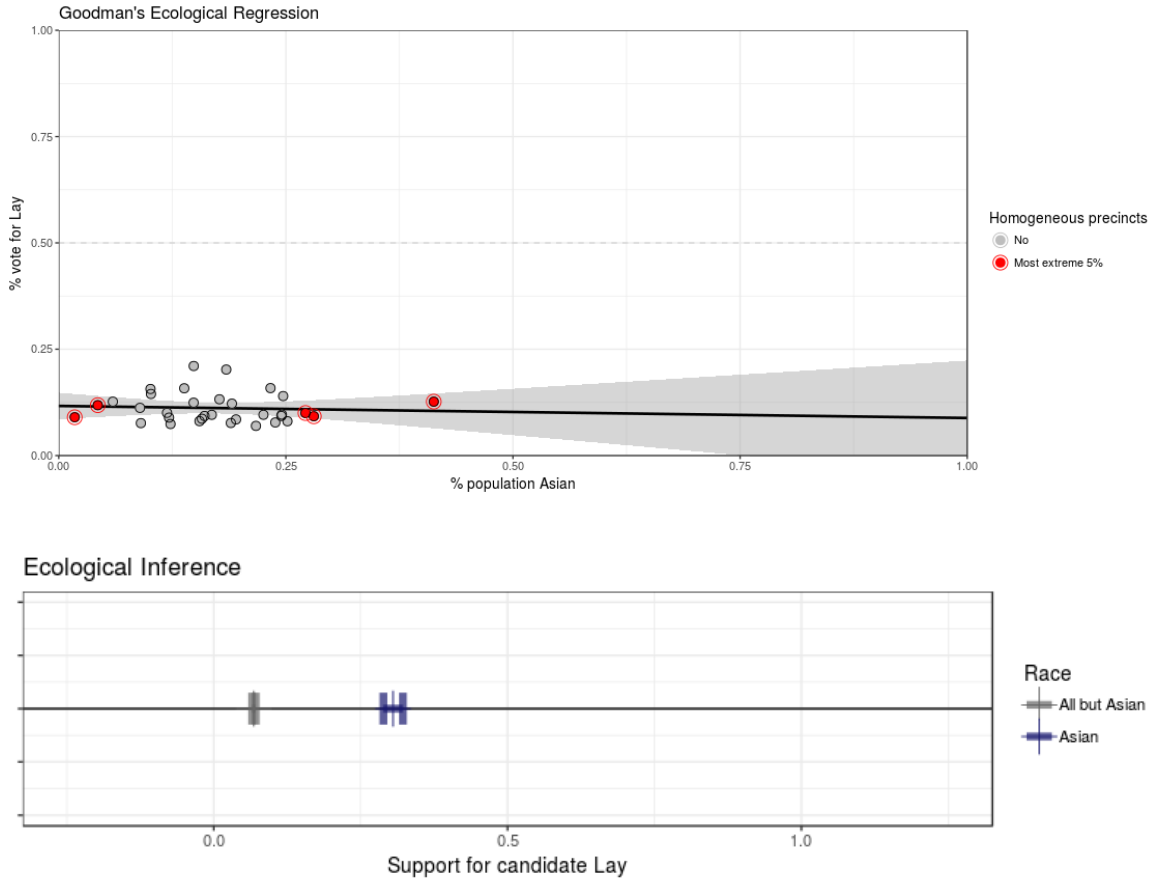


FIGURE 2. Results from running the Shiny app to assess Asian support for Lay in Lowell. In the ER plot, red dots represent the most homogeneous precincts - the 5% with the highest Asian share of the population the 5% with the lowest Asian share. The EI figure shows confidence intervals around the estimated level of support among Asian and non-Asian voters.

The results in Figure 2 show low support for Lay among non-Asian voters from both ER and EI. Among Asian voters, EI estimates about 30% support for Lay, and ER estimates about 9% support; this is a scenario

where EI and ER differ significantly. This is often an indication of poor and/or sparse data. This can cause poor predictions in both models, and not necessarily the same poor predictions.

Although not reflected in these results, we know there has historically been low turnout among Asian voters in Lowell. But we also know from research on the ground, that Asian voters that do turn out typically support Asian candidates. The ER results of course do not show this. This is because there is little, if any, upward movement in Lay's support as data points move to the right in the plot, or as the Asian share of precinct-population increases. That means the regression line won't slope upward (or if so, not by very much). Because support for Lay is low in precincts with a small Asian population, this results in a low final ER estimate (or the estimate in a 100% Asian precinct). Why might this be? We explore this phenomenon more in Section 4.

Section 3. From Raw Data to RPV

Denham Springs. According to ACS data, Denham Springs' Citizen Voting Age Population (CVAP) is 86% white, 11% black, 2% Hispanic and 1% other. Figure 3 shows the demographic layout across Denham Springs. We can see significant racial segregation, particularly between white and black neighborhoods. Table 1 has EI and ER results for Denham Springs' 2018 City Council election for white, Hispanic and

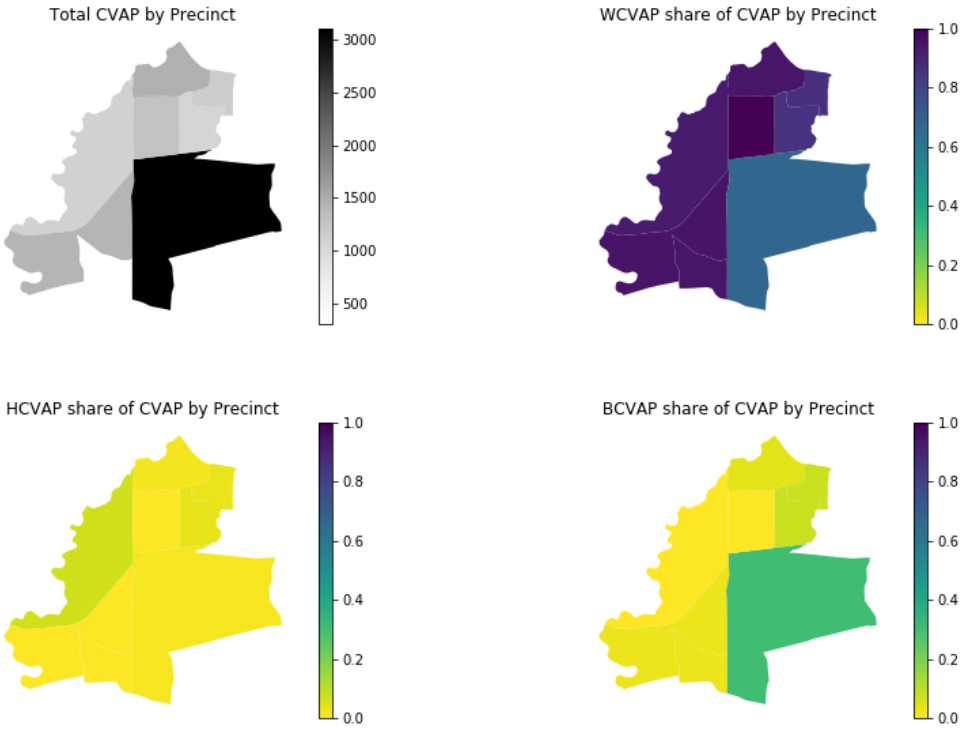


Figure 3. Total CVAP by precinct in Denham Springs as well has WCVAP, HCVAP and BCVAP share of CVAP by precinct

black voters. Yellow cells indicate a lack of racial polarization (as determined by overlapping EI confidence intervals in the Shiny app). Gray cells indicate ER and EI results significantly differed. Notice that this happened frequently when estimating Hispanic candidate support. Both EI and ER suffer greatly from low populations - in this case HCVAP is only 2% of CVAP. However, the models don't always perform badly in the same ways! In these instances its best to call the results inconclusive as opposed to trying to glean information from these model results.

We see some significant differences in candidate support between black and white voters, namely for Riley (where black support is estimated to be between 53% and 77% and white support is estimated between 6.6%

and 8%.) and Williams. However, we can be less confident in the latter because of overlapping confidence intervals in white voters' EI estimate. In the case of Riley, notice that black support is estimated to be greater than 50%, so by some measures Riley would be considered black voters' "candidate of choice.' Because Riley has the lowest support among white voters of all the candidates, we can say Riley is not white-preferred. Circumstances like these - minority and majority voters having different preferred candidates - are necessary conditions for bringing VRA claims.

Other factors - such as additional elections and the success of minority-preferred candidates - would have to be analyzed to see if a VRA claim is warranted in Denham Springs. But assessing minority vs. majority preferences in as many instances possible can help bolster a claim of racially polarized voting.

Race	Methodology	Dugas	Williams	Poole	Riley	Smith	Wesley
WCVAP	ER	0.173	0.182	0.185	0.08	0.191	0.189
WCVAP	EI	0.157	0.154	0.18	0.066	0.191	0.189
BCVAP	ER	0.09	0.078	0.104	0.528	0.1	0.1
BCVAP	EI	0.056	0.139	0.058	0.774	0.057	0.096
HCVAP	ER	0.27	0.356	0.277	0	0.132	0.102
HCVAP	EI	0.995	0.343	0.543	0.984	0.3	0.199

TABLE 1. Estimates of white, Hispanic and black support for 2018 City Council candidates in Denham Springs. Cells are yellow if confidence intervals overlapped in the Shiny app (meaning we are not confident there are different support levels between a given race and all other voters) and gray cells highlight where ER and EI are significantly different.

Santa Monica. According to ACS data, Santa Monica's CVAP is 73% white, 12% Hispanic, 3.4% black and 11.5% other. It's City Council has 7 members. If minority voters have similar preferences, proportional representation on the City Council would be about 2 minority-preferred candidates.

Figure 4 shows the demographic breakdown across Santa Monica. We can see some racial segregation: there are some dense white neighborhoods, and the densest Hispanic and black neighborhoods are where the white share of the population is smallest.

The estimates show that several candidates had similar levels of support across racial groups. The yellow cells in Table 2 similarly suggest low confidence that different races had strongly different preferences.

Three seats were up for election in the 2018 City Council race. They were filled by Himmerlich, Morena and McKeown. They joined the 4 existing members of the council, one of whom was Hispanic (Tony Vasquez, who was replaced by Ana Maria Jara in January 2019 when Vasquez was elected to the State Board of Equalization). The 2018 City Council election came in the midst of a lawsuit filed by the Pico

Race	Methodology	Himmerlich	Powell	Neri	McKeown	Morena	Oconnor	Bellomo
WCVAP	ER	0.242	0.1	0.076	0.178	0.192	0.108	0.096
WCVAP	EI	0.238	0.1	0.08	0.179	0.189	0.109	0.102
BCVAP	ER	0.312	0.2	0	0.36	0.298	0.17	0
BCVAP	EI	0.25	0.8	0	0.351	0.306	0.123	0.386
HCVAP	ER	0.274	0.165	0	0.24	0.228	0.142	0
HCVAP	EI	0.263	0.145	0	0.237	0.238	0.143	0.02

TABLE 2. Estimates of white, Hispanic and black support for 2018 City Council candidates in Santa Monica. Cells are yellow if confidence intervals overlapped in the Shiny app (meaning we are not confident there are different support levels between a given race and all other voters) and gray cells highlight where ER and EI are significantly different.

Neighborhood Association, among other plaintiffs, against the City of Santa Monica. Plaintiffs alleged that Santa Monica's at-large election system diluted Latino voting power in violation of the California Voting Rights Act (CVRA). The trial court ruled in favor of the plaintiffs and the City appealed the decision.

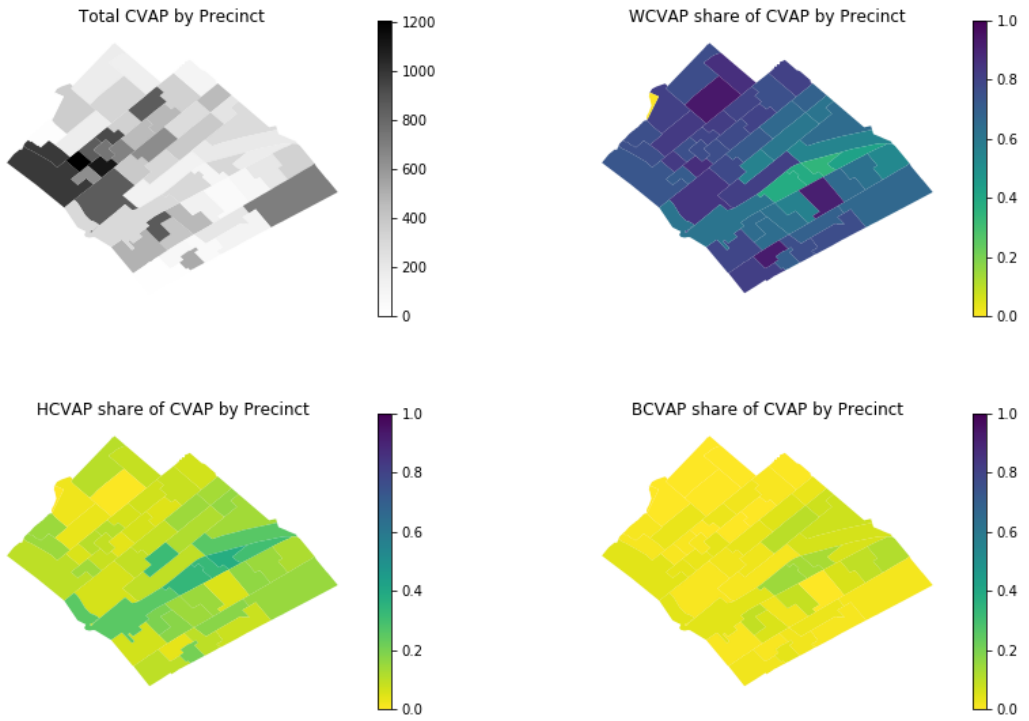


FIGURE 4. Total CVAP by precinct in Santa Monica as well has WCVAP, HCVAP and BCVAP share of CVAP by precinct

The ER/EI results in Table 2 should not be taken as the only indicator of whether there is Racially Polarized Voting in Santa Monica. For one, the lack of candidates of color doesn't offer ample opportunity for minority voters to show different preferences. Furthermore, the attention that legal challenges can bring to an election-system can change voter engagement and behavior. It's also important to do ER/EI analysis on elections well before the challenge took place. Results suggesting polarized voting might include much different candidate-support levels across races (particularly with minority-candidates) and fewer overlapping confidence intervals in EI results.

However, hopefully this has provided some guidance on how to conduct such an analysis across elections and consider their results holistically.

SECTION 4. EXTRAS: THINGS THAT CAN GO WRONG

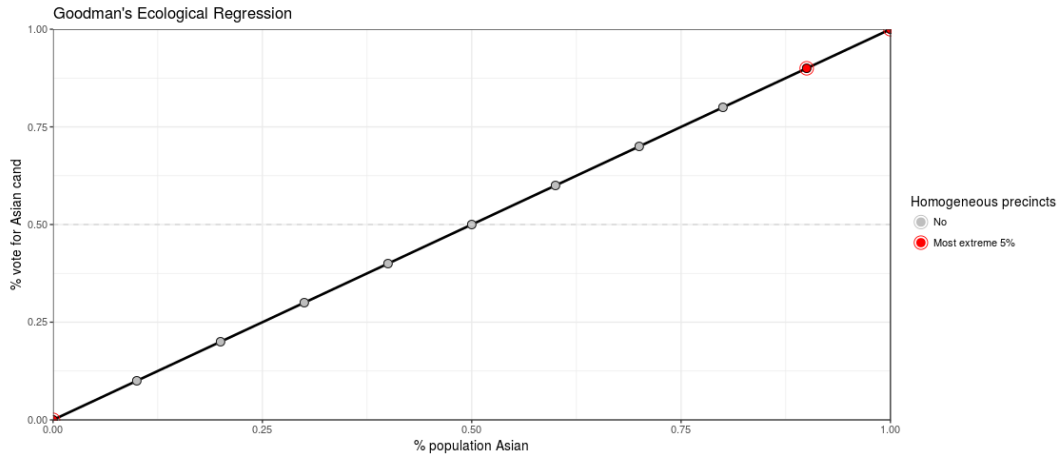
Exploring differential turnout (prompt 3). Consider the simple example of a district with 11 precincts, each with 100 people. Everyone is either Asian or white and there is one Asian candidate and one white candidate. There is full racial polarization: All Asian voters vote for the Asian candidate and all white voters vote for the white candidate.

Precinct 1 has 0 Asian people and 100 white people, Precinct 2 has 10 Asian people and 90 white people etc. ... Precinct 11 has 100 Asian people and 0 white people.

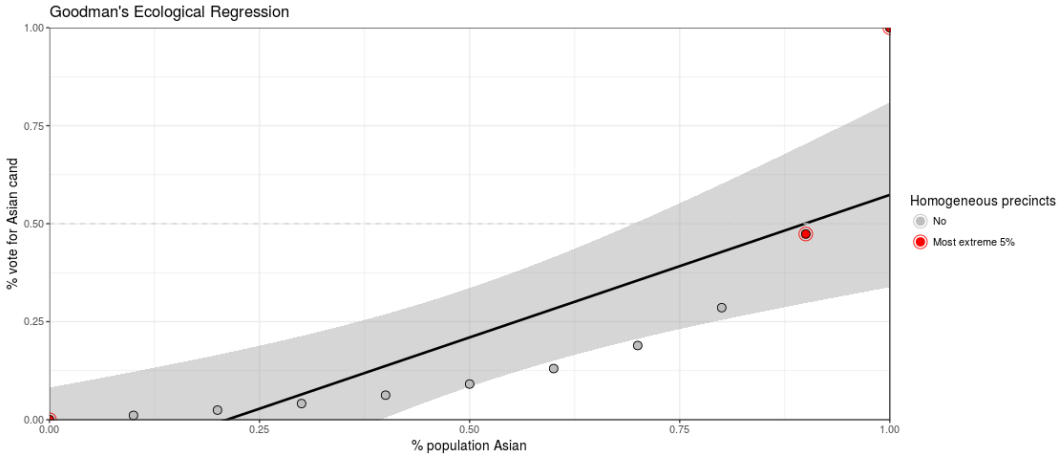
Consider two scenarios: The first has full turnout for both groups, the second has 10% Asian turnout and 100% white turnout. We'd like to see if ER can pick up on the "ground truth" - in particular that Asian voters fully support the Asian candidate - in these very different turnout scenarios.

Figure 5 shows ER results for both cases. In both plots, the Asian share of population is on the x-axis and the vote share for the Asian candidate is on the y-axis. Figure 5a gives the "correct" answer: it predicts that in a 100% Asian precinct, 100% of the votes cast would go to the Asian candidate. Figure 5b predicts that in a 100% Asian precinct, about 60% of the vote share would go to the Asian candidate. Without accounting for turnout differences, and instead assuming equal turnout, this version of ER would produce

the Figure 5b plot, estimating a much lower level of support among Asians for the Asian candidate than it is in reality.



(A) Full turnout for Asian and white voters



(B) 10% Asian turnout

FIGURE 5. ER for simple example in cases of full turnout and 10% Asian turnout

Intuitively, this is because data points in Figure 5b are much "lower" than their counterparts in Figure 5a. This is because of how much the white vote dominates in any precinct in this case. For example, in the precinct with 40 Asian people and 60 white people, the Asian candidate vote share is $40/100 = 40\%$ in the full turnout case, but only $4/64 = 6.25\%$ in the 10% Asian turnout case. This in turn "drags down" the data in Figure 5b, causing a lower predicted value in a 100% Asian precinct, or the result of ER.