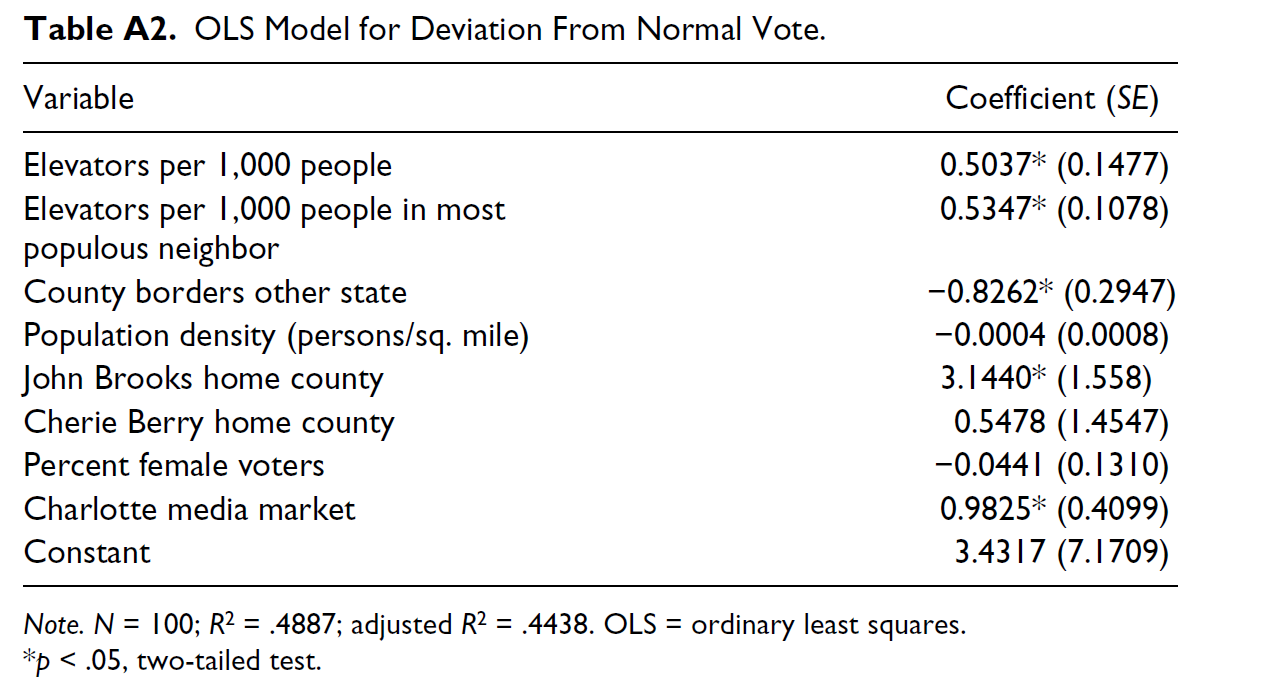
**Bonus Problem Set**

This bonus problem set covers material from Friday July 31, Monday August 3, and Tuesday August 4. It is due on Tuesday August 4 at 11:59 pm.

This bonus problem set focuses on this table from Jacob (and Neil Weinberg’s) paper [The Elevator Effect: Advertising, Priming, and the Rise of Cherie Berry](https://journals.sagepub.com/doi/abs/10.1177/1532673X15602755?journalCode=aprb). The dependent (response) variable here is Cherie Berry’s deviation from the average Republican vote in a county for other Republicans run for state offices elected statewide (e.g., Governor, Secretary of State). For example, if she were to do 2 percentage points better, this value would be equal to 2.

The main independent (explanatory) variable here measures the number of elevators per 1,000 people in a county. Its mean value is 1.75 and its standard deviation is 1.23. Other important variables include the number of elevators in the most populous neighboring county and a dummy variable for whether the county borders another state. Control variables include population density, dummy variables for the counties where each candidate is from, the percent of voters in a county who are female, and a dummy variable measuring whether a county is located in the Charlotte media market (where Cherie Berry ran television advertisements).



1. Please interpret the coefficient for elevators per/1,000 people using both frameworks we discussed in class and discuss its statistical significance. Given that Cherie Berry received about 53% of the vote in 2012, winning by a bit over 6 percentage points, is this a substantively large effect? How do you know? (Hint: think about the scale of the variables and the fact that in a two candidate race a vote added to one candidate is also in effect a vote subtracted from the other candidate.) (4 pts.)
2. Please interpret the R2 value for this model. What might be contributing to the fact that the adjusted R2 is about 0.045 lower than the R2? Please be specific. (2 pts.)
3. This model includes a dummy variable for Cherie Berry’s home county and for John Brooks’ (her Democratic opponent) home county. Why does it not include a dummy variable for counties that neither candidate is from? What happens if you include it? (2 pts.)
4. Let’s say that you wanted to predict if Cherie Berry won or lost a county rather than her deviation from the average Republican vote. What type of regression might you use? Why? (2 pts.)