

## Data Description

In order to explore possible relationships between incumbent expenditures, tenure and other assorted variables, we selected the Wooldridge dataset "Voter2". The set contains 186 observations of 1988 and 1990 House of Representatives races and measures many variables pertinent to our research question.

The variable within the set that concerns incumbents' campaign expenditures in 1990 (inexp90) is the dependent variable that our analysis focuses on. Although the data shows incumbents spending over \$430,000 on average, one incumbent spent 1.53 million dollars in their race, while another incumbent spent less than \$7,000. These figures suggest that expenditures vary widely depending on districts and local factors. That being said, looking at a histogram of incumbent expenditures shows that this \$1.5 million house race is an outlier, and that most campaign expenditures spend in amounts close to the \$430,000 average. In the following paragraphs I will discuss possible explanatory variables and their correlations with incumbent expenditures.

In the world of politics, there is a phenomenon called the "incumbent advantage" where incumbents, for a variety of reasons, are difficult to unseat or mount a formidable challenge against. To see if this is true, we used the "tenure" variable, which examines how long the incumbent had been in office at the time of the 1990 election. The table shows that the average time an incumbent has been in office is just over eleven years, or 5 election cycles. That's a long time, and shows that if an incumbent is elected in office, they are likely going to stay there for awhile. The histogram of tenure confirms this intuition, with the majority of incumbents falling between five and twenty years in office. If the incumbency advantage does exist, it's easy to hypothesize that incumbents will spend less money in their races as they become more established in congress and their districts. Creating a scatterplot with tenure on the X-axis and campaign expenditures on the Y-axis and creating a line of best fit confirms this relationship. The slope of the line is negative, but the severity of its incline is more slight than might have been expected. It appears that although it's significant, tenure does not have a dramatic influence on campaign expenditures.

The campaign expenditures of the incumbent's challenger (chexp90) may go a long way towards explaining the level of the incumbents' expensing. Examining the challengers' expenditures shows that, on average, challengers spend well over three times less than their incumbent opponents (\$128,900 versus \$430,000). Furthermore, although one challenger did spend over one million dollars, a histogram of this variable reveals that the vast majority of challengers spend around \$100,000. Logically, an incumbent's expenditures and their opponents' will likely mirror each other, reflecting the cost of the local media market (TV advertising, radio advertising, billboards) and rising concurrently as the two opponents attempt to out-organize one another with polling and canvassing efforts. A scatterplot of the two variables confirms this intuition. The upward sloping, sharply sloped line of best fit positive reflects a strong, positive correlation that's higher in magnitude than the relatively flat slope of the tenure scatterplot. It's also important to observe how many of the data points inhabit the lower left hand side of the graph and "hug" the y-axis. This area, with low challenger expenditures and high incumbent expenditures, suggests that a lot of districts had no

formidable, lucrative challenger in the 1990 election. The correlation between incumbent and challenger expenditures has one of the highest values of the explanatory variables we've examined.

The variable "Vote88" measures the incumbent's share of the two party vote in 1988, and is a good indicator of the incumbent's popularity within their district. Although useful, the variable's intense selection bias must be noted. The measure only examines candidates who won their 1988 election and, as a result, the average incumbent's share of the two party vote was 66.45%. Values below 50% reflect an independent incumbent that doesn't belong to either of the main parties. The histogram reflects this selection bias by showing a severe drop off in values below the fifty percent of the vote mark. The histogram also shows that most incumbents won the 1988 election handily with a vote share somewhere between 58 and 80%. Such a figure helps explain the "incumbent advantage" discussed earlier. Intuitively, there must be a negative relationship between vote share and campaign spending. If an incumbent won their last election by a double digit margin (higher than sixty percent share of the vote), they are more secure in their seat and thus less motivated to engage in the intensive, expensive campaigning that more insecure incumbents do. The scatterplot of expenditures and vote share confirm the negative correlation of these two variables. The line of best fit has a steep negative slope of a higher magnitude than the positive slope between challenger and incumbent expenditures, suggesting that the incumbent's popularity may be a better indicator of their campaign spending level than their challenger's expenditures.

The last variable is "prtystr" which measures the share of the vote that the incumbent's party's presidential candidate received in the incumbent's home district in the 1988 presidential election. For example, if a Republican congressman's district was won by the Republican nominee George H.W. Bush with 60% of the vote, then that incumbents "prtystr" measurement would be 60. If the same situation happened to a Democratic incumbent, they would have a score of 40. The average value for party strength was 53.2, showing that most incumbents represent districts who favor the political party of the incumbent. If an incumbent has a low party strength score, but a high voter share in the 1988 election, it could be because they reside in swing states and put forth positions not usually voiced by other members of their party. A so called "Red Democrat" or "Blue Republican" would fit this description. A low party strength could also signify that the incumbent's district is beginning to view the incumbent's party less favorably, which might translate into higher incumbent expenditures as the candidate spends more money to keep voters on their side. The scatterplot supports this hypothesis by generating a line of best fit with a negative slope. The slope of the line has a lower magnitude than the correlation of campaign expenditures and incumbent expenditures, but it is still a significant metric to analyze our data through.

Variable	Obs	Mean	Std. Dev.	Min	Max
inexp90	186	431787.8	249107.2	6768	1538945

Variable	Obs	Mean	Std. Dev.	Min	Max
tenure	177	11.29944	7.972656	1	49

Variable	Obs	Mean	Std. Dev.	Min	Max
chexp90	186	128940.4	189184	200	1067366

Variable	Obs	Mean	Std. Dev.	Min	Max
prtystr	186	53.22043	11.0287	27	95

Variable	Obs	Mean	Std. Dev.	Min	Max
vote88	186	66.45699	10.57845	26	97



















