

## Assignment 6 – Solutions

1. Since our goal is prediction we should include all variables that are available at the time we run the analysis and that have some relationship with our dependent variable.
  - (a) Include age
  - (b) Include whether the individual contributed money to a political candidate.
  - (c) Do not include. We cannot include a variable on whether the individual watched the 2008 debate, since presumably it hasn't happened yet.
  - (d) Include variable on political preferences.
  - (e) Include. This variable can be included on the grounds that it is a proxy for "green" political preferences. Environmentalists may be less likely to vote Republican.
  - (f) Do not include. This variable is not available and probably not related to whether a person is a Republican or not.
  
2. Iyengar and Ansolabehere's goal is causal inference, thus they would include variables that are expected to affect turnout.
  - (a) Include. Demographics, such as age, are common control variables in regression models. Studies on turnout also found that older voters are more likely to show up at the polls, so this variable should be included.
  - (b) Include. This variable can be seen as a proxy to the voter's level of interest and engagement in politics, which, in turn, may help explain higher motivation to show up at the polls, even if negative advertising took place.

- (c) Include. The argument is similar to the one made above. Also, if a person is more knowledgeable about politics, she is probably less likely to be affected by negative campaigning.
- (d) Do not include. There is no straightforward theory linking being moderate with higher (or lower) probability of turning out to vote.
- (e) Do not include. There is no theory connecting turnout with choice of vehicle.
- (f) Do not include. It is difficult to think of a compelling straightforward argument relating a person's belief about the honesty of candidates to her likelihood of voting. Some may take this belief as a reason not to vote, since most politicians are honest. Others see it the other way round. Some may argue these people are highly affected by negative campaigning, since its contents go against their beliefs. Others may think they are not affected by the contents of the advertisements since their opinions about the integrity of the candidates are already formed.

3.

- (a) Based on the bivariate regression output there was no statistically significant difference in turnout between states that had initiatives on the ballot and states that did not.

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.566	.012		48.845	.000
	Initatives on the Ballot in 2004	.029	.020	.200	1.431	.159

a. Dependent Variable: Turnout in 2004 (VAP)

(b) Based on this same regression output we can see that the coefficient is small (2.9%) and not statistically significantly different from zero. A 95% confidence interval around that point estimate is given by :  $[.029 \pm 1.96 \cdot .02]$  that is  $[-.0102, .0682]$ . As we can see .05 is contained in this interval, thus we cannot reject the null that  $\beta_{init} = 5\%$ .

(c) Swing states with initiative on the ballot:

$$y_{pred} = .546 + .055 + .077 - .087 = .591$$

Swing states with no initiative:

$$y_{pred} = .546 + .077 = .623$$

Safe states with initiative:

$$y_{pred} = .546 + .055 = .601$$

Safe states with no initiative:

$$y_{pred} = .546$$

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.546	.013		43.602	.000
	Initatives on the Ballot in 2004	.055	.027	.382	2.050	.046
	Swing2004	.077	.024	.551	3.170	.003
	Init_Swing_04	-.087	.039	-.510	-2.211	.032

a. Dependent Variable: Turnout in 2004 (VAP)

4.

(a) A 100% increase in  $x_n$  leads to a .26 decrease in  $y_n$ . The effect is statistically significant at the 5% level.

(b) A 100% increase in  $x_n$  leads to a 170% increase in  $y_n$ . However, this effect is not statistically significant at the 5% level.

(c)  $x_n$  has a nonlinear statistically significant effect on  $y_n$ . The higher the value taken by  $x_n$ , the stronger its negative effect on  $y_n$ .

(d) A one unit increase in  $x_n$  leads to a 492% increase in  $y_n$ . The effect is statistically significant at the 1% level.

5.

(a)

**Coefficients(a)**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	.653	.018		36.321	.000
	Incumbent Position	.005	.017	.018	.315	.753
	Incumbent Position ^ 2	-.144	.057	-.142	-2.549	.011
	State Ideology	-8.61E-005	.000	-.010	-.184	.854
	State Unemployment	.002	.002	.042	.755	.451
	Challenger Quality	-.025	.003	-.431	-7.860	.000
	Incumbent is a Freshman	-.015	.010	-.082	-1.495	.136

a Dependent Variable: Incumbent Two-Party Vote Share

The coefficients on incumbent position and incumbent position-squared indicate that incumbent position non-linearly impacts incumbent vote share. More specifically, for low levels, more conservatism increases the vote share, but as conservatism becomes more extreme, it actually reduces the vote share. The relationship between vote share and incumbent position can be represented by an inverse-U shape.

(b)

**Coefficients(a)**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta	B	Std. Error
1	(Constant)	.704	.025		27.754	.000
	Incumbent Position	.188	.055	.657	3.390	.001
	State Ideology	.000	.000	-.047	-.823	.411
	State Unemployment	.001	.002	.025	.447	.655
	Challenger Quality	-.024	.003	-.416	-7.595	.000
	Incumbent is a Freshman	-.016	.010	-.087	-1.585	.114
	Incumbent is a Republican	-.060	.023	-.322	-2.656	.008
	Incumbent Position / Republican Interaction	-.202	.068	-.413	-2.994	.003

a Dependent Variable: Incumbent Two-Party Vote Share

The table shows that a Republican incumbent with a position of zero will lower the vote share by 0.06. Also, when ‘Incumbent is a Republican’ is held at zero, i.e., when the incumbent is not a Republican, a unit increase in the incumbent’s position will increase the vote share by 0.188.

The impact of incumbent position, conditional upon a republican incumbent is:

$$0.188 + (-0.202) = -0.014$$

The impact of a republican incumbent, conditional upon incumbent position is:

$$(-0.06) + (-0.202) = -0.262$$

(c) The predicted vote share for John Ashcroft of MO is 61.72% with a 95% confidence interval of (59.05, 64.39). The actual vote share was 49%. The prediction differs from John Ashcroft’s actual vote share because this model posits that higher conservatism will increase the vote share, whereas we saw in part (a) that the

relationship between incumbent position and vote share is curvilinear. Therefore, it is likely that his reasonably high score on incumbent position actually hurt him in terms of gaining votes, something that the model in part (b) overlooks.