Methods and Initial Results

Julian McClellan

Due 5/17/17

Model Justification

In the subreddit /r/changemyview, Reddit users (Redditors) post their views as submissions on the subreddit, and other Redditors attempt to change these views. If the Redditor who originally posted their view has changed their view to some extent, they are able to award a "delta" to comment (and the Redditor who authored that comment). These deltas are explicitly coded in the submission and thus, one can say that for a given submission in /r/changemyview, whether or not the original poster (OP) changed changed their mind at least once; i.e. they awarded at least one delta.

Self-affirmation, a theory in psychology suggests:

> "[I]f global self-worth is temporarily bolstered by success in a second, unrelated domain, the individual should be more willing to tolerate a threat to the domain of interest."

(Correll, Spencer, and Zanna 2004)

In the context of submission in /r/changemyview, the "threats" to OP's domain of interest are the comments that attempt to change their views. For many of the OP's, their submissions to /r/changemyview are not their first submissions they have made in Reddit. Any prior submissions an OP made can be considered potential cases where his or her global self-worth may have been temporarily bolstered. Looking at these submissions prior to an /r/changemyview submissions, all of them receive a "score" of some number, the net of upvotes and downvotes from other members of the community could serve as proxy for the "degree" to which the individual's self-worth was bolstered.

Additionally, one might argue that the very act of posting a submission bolsters global self-worth on its own. Indeed, many psychological experiments testing self-affirmation theory focus on global self-worth bolstered by the person themselves (Cohen, Aronson, and Steele 2000).

It might also be justified to also look at the *total* score of the Redditor's prior submissions, to

explicitly see the lifetime score of their submissions prior to that /r/changemyview submission. In a generalized linear model, we can account for this with an interaction term between the number of submissions and the average score per submission.

Thankfully, the entire submission history of any Redditor is publically available, so all of these statistics can be gathered and calculated.

Past Submission Diversity

Additionally, one might also want to test for an sort of "echo" chamber effect. The subreddit in which submissions were made can be obtained, and there might be a difference in how likely a Redditor is to award a "delta" if that Redditor's past submissions took place across a relatively diverse set of subreddits versus a select few. Gini index can serve as suitable metric for this measure.

Model Definition

The binary nature of our outcome variable, whether or not OP awards a delta, makes the following logistic model viable:

$$prob\_OP\_award\_delta = \frac{1}{1 + exp(-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_1 X_2))}$$

where:
$$\begin{cases} X_1 &= & num\_prior\_subs \\ X_2 &= & avg\_sub\_score \\ X_3 &= & subreddit\_gini\_index \end{cases}$$

Data

The data includes all 5777 submissions made by 4152 unique authors. Of these 5777 submissions, 3298 of them, or 57.09 % did not have their OP award a delta. (The view of the OP was not changed). The other 2479 of the posts, or 57.09 % had their OP award a delta.

From this initial set of submissions, we are only interested in submissions in which the OP had prior submission history.

From the initial 5777 submissions, only 5326 (92.19 %) were made by OP's with prior submission history. Of these 5326 submissions, 3024 of them (56.78 %) did not have their OP award a delta. (The view of the OP was not changed). The other 2302 of the posts (56.78 %) of had their OP award a delta.

A summary table for:

Table 1: Independent Variable Summary Statistics

Statistic	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
num_prior_subs	134.343	209.315	1	16	56	153	1,839
avg_sub_score	73.032	39.474	0.000	62.667	69.212	77.500	500.000
gini_index	0.408	0.218	0.000	0.256	0.432	0.576	0.912

Initial Results

Now we compute the logistic model as described above:

$$prob\_OP\_award\_delta = \frac{1}{1 + exp(-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_1 X_2))}$$

where:

$$\begin{cases} X_1 &= num\_prior\_subs \\ X_2 &= avg\_sub\_score \\ X_3 &= subreddit\_gini\_index \end{cases}$$

The table below shows statistical signifiance at the `0.01` level for the number of prior submissions and the total prior submissions score we created as an interaction term. Now let's evaluate the model accuracy with a look at the area under the ROC curve (AUC).
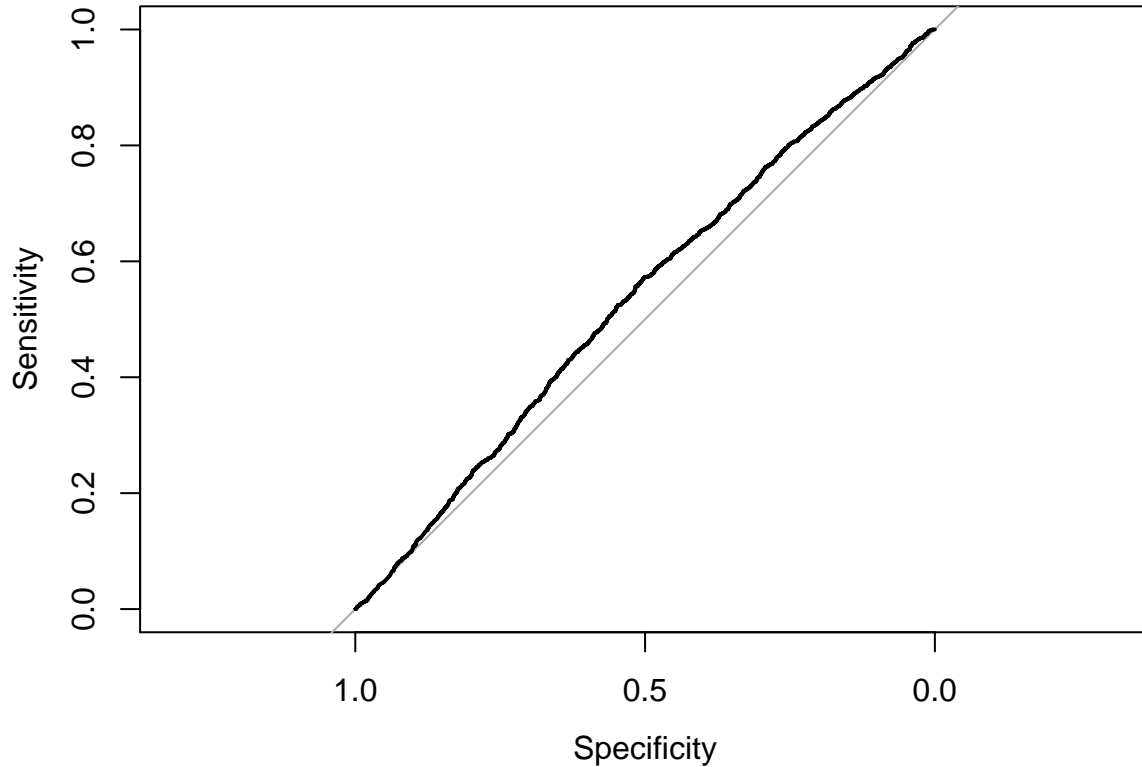
Table 2: Initial Logistic Regression Results

	Dependent variable:
	OP_gave_delta
Number of prior Submissions	0.002***
	(0.001)
Average prior submission score	0.00002
	(0.0002)
Gini Index	−0.077*
	(0.040)
Total Prior Submission Score (Interaction)	−0.00003***
	(0.00001)
Constant	0.478***
	(0.021)

Note: *p<0.1; **p<0.05; ***p<0.01

ROC Curve for Initial Logistic Model | AUC = 0.537

Although we only get an AUC of 0.5374808 this is still very promising. Tan et al. (2016) were only able to get an AUC of `.54`, by looking at a large number of features extracted from the langauge of the CMV submission itself. In the logistic model given above, we use much more primitive features, simply looking at the OP's past submissions, their scores, and a simple measure of subreddit diversity in these submissions. Additionally, Tan et al. utilized a filtering procedure, based on OP interaction and the number of unique respondents, to remove submissions that did not have enough debate (in their view) to warrant a delta; no such filtering procedure has been used here.

The statistical signifiance of some of our features provides some evidence that self-affirmation plays a role in whether or not the original poster of /r/changemyview submission will change their view (award a delta). Additionally, because the original text of each individual submission is also available for natural language processing, this logistic model can be improved by adding these sorts of language features as well. For instance, Tan et al. utilized first person pronoun useage as a proxy for self-affirmation in the /r/changemyview submission, and first-person pronoun useage in prior submission content can be be extracted as well.

References

Cohen, Geoffrey L, Joshua Aronson, and Claude M Steele. 2000. "When Beliefs Yield to Evidence: Reducing Biased Evaluation by Affirming the Self." *Personality and Social Psychology Bulletin* 26 (9). Sage Publications Sage CA: Thousand Oaks, CA: 1151–64.

Correll, Joshua, Steven J Spencer, and Mark P Zanna. 2004. "An Affirmed Self and an Open Mind: Self-Affirmation and Sensitivity to Argument Strength." *Journal of Experimental Social Psychology* 40 (3). Elsevier: 350–56.

Tan, Chenhao, Vlad Niculae, Cristian Danescu-Niculescu-Mizil, and Lillian Lee. 2016. "Winning Arguments: Interaction Dynamics and Persuasion Strategies in Good-Faith Online Discussions." In *Proceedings of the 25th International Conference on World Wide Web*, 613–24. International World Wide Web Conferences Steering Committee.