Inference of Candidate Qualities through Political Stances

Cindy Zhang
Department of Mathematics, MIT
cindyyz@mit.edu

Eileen Pan
Department of EECS, MIT
eileenp@mit.edu

Abstract

Because most voters are unable to interact directly with political candidates, many voters will instead infer personal qualities of political candidates through information such as the candidate's political platform. In this paper, we use Bayesian inference and linear regression models to predict how people infer personal qualities and traits from limited information about candidates' political stances. We test these models through an experiment where we ask subjects for their evaluations of candidates with varying political views. Our models are able to closely predict the distribution of personal traits ratings, which suggests that political stances determine a large part of people's perception of a candidate's personality.

1 Introduction

With the 2020 United States elections coming up, politics are once again at the center stage of our attention. What drives a candidate to victory? One main factor is the alignment of a candidate's political stances with the stances of the majority of voters. However, according to a poll conducted by the Pew Research Center, only 66% of those who supported Clinton and 68% of those who supported Trump said they knew a lot about the candidates' stances on political issues (Doherty et al., 2016). This tells us there's more than just political stances that draws voters to candidates.

Beyond political stances, candidate personality has been known to be an important factor in elections. Voters are inclined to vote for candidates they perceive as having traits they personally value in themselves (Koppensteiner and Stephan, 2014). Evidence suggests that voters also prefer candidates that they perceive to have a personality similar to their own (Caprara et al., 2007).

Each political party in the United States has their own trait attribution. Democrats are rated as more compassionate, empathetic, intelligent, and knowledgable, while Republicans are rated as more moral, decent, and stronger leaders (Hayes, 2005). Some of this difference can be attributed to party stances on political issues, as Democrats and Republicans have very different and generally invariant political positions. Survey result also support the idea that people infer a candidate's traits from a candidate's political platform, which implies that a candidate's personality is intrinsically linked with their issue stances.

The lay theory of dispositionism supposes that people have dispositions that change rather little over time. Those who subscribe to this mentality are entity theorists and those who believe dispositions are rather fluid and more a function of circumstance are termed incre-

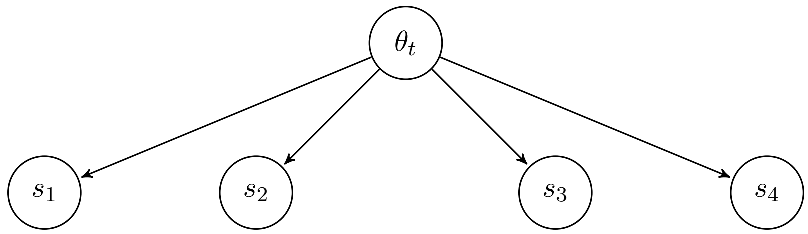

Figure 1: Model for a single trait given stances s_1, s_2, s_3, s_4. This structure is repeated for each candidate trait to obtain an independent probability distributions

mental theorists (Plaks et al., 2009). Entity theorists are more likely to attribute someone's actions to innate personality, as opposed to outstanding circumstance, put more weight on personality traits when evaluating a person, and are skeptical personalities can change (Levy et al., 1999). For entity theorists, the way they perceive a candidate's personality has a large impact on their voting decision.

We seek to develop a computational model for how people infer candidate traits from issue stances. One possible way to model trait inference is by using an inverse planning framework to capture how humans infer desires from observable actions by inverting a probabilistic generative model of goal-dependent plans (Baker et al., 2007). There have also been frameworks that use the naive utility calculus to model early social reasoning that is inferred from people's actions (Jara-Ettinger et al., 2016). The naive utility calculus theory suggests that agents people operate on personal expected cost and reward functions, and mistakes can be attributed to ignorance about true cost and reward functions. Using this idea, we developed three models for inferring candidate traits.

In this paper, we first present the computational models we chose for candidate trait inference. The first and third models assume that the ratings for each trait are independent, while the second model does not. We then describe the experiment we conducted to test how well the quantitative predictions of our model reflect human judgements of candidate traits. In the remainder of our paper, we discuss the advantages and disadvantages of each model and sketch out possible ideas to further explore.

2 Computational Models

Our first two models are based on the idea that social cognition is supported by a probabilistic generative model that determines how mental states lead to actions using the conditional probabilities of actions given mental states (Baker et al., 2017). These two models are able to predict the distribution of ratings that candidates receive for each trait. Our third model uses linear regression to predict the average ratings that candidates receive for each trait.

2.1 Model 1: Probabilistic Generative Model with Independent Traits

In this model, we assume that the effect each trait has on a candidate's political stance is independent of the effect of other traits. For each trait t, we have one free parameter θ_t that represents the distribution of how strongly the candidate has trait t.

Given a set of stances S for a particular candidate, the parameter θ_t is found for each trait of the candidate. This distribution is obtained using Bayes' rule:

$$P(\theta_t | S) \propto P(S | \theta_t) P(\theta_t). \tag{1}$$

Our experiment is designed to calculate the probability of the set of stances S occurring given a particular distribution θ, which is initialized as the uniform distribution. The marginal distribution for the output of the probabilistic generative model given these priors is the distribution of how strongly the candidate possesses the trait. The overall schematic of our model is pictured in Figure 1.

2.2 Model 2: Probabilistic Generative Model with Pairs of Dependent Traits

In this model, we consider pairs of traits that are dependent and the effect each pair of traits has on a candidates' political stance.

We chose this model in order to explore how multiple traits work together to impact a person's political stances. A correlation matrix to see how different traits are correlated with each other using data from our experiment is displayed as a heat map in Figure 2. This model is test with both pairs of traits that are highly correlated and pairs of traits that aren't highly correlated to evaluate how the correlation between the traits affects the predicted distributions.

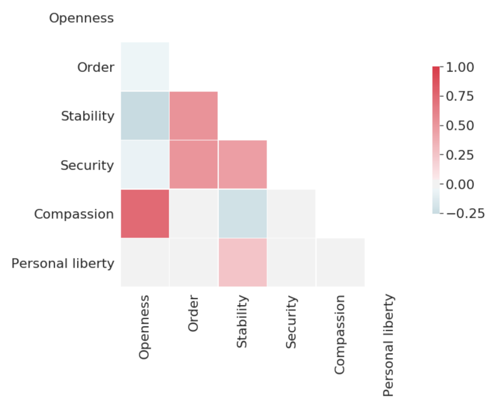

Figure 2: Heat map displaying the Pearson correlation coefficient between each pair of distinct candidate traits.

For each pair of traits t_1 and t_2, we have two free parameters θ_{t_1} and θ_{t_2} that represent the distributions of how strongly the candidate possesses t_1 and t_2 respectively. This model is based on the assumption that the distribution for trait t_2 is dependent on the distribution for trait t_1. It's difficult to know exactly which trait is dependent on the other one (or if they are dependent on each other), but in our model we chose to let the trait we are trying to infer be the one that is dependent on the second trait.

We generate the inference model with θ_{t_2} depending on θ_{t_1} as

$$P(\theta_{t_2} | \theta_{t_1}) = \frac{P(\theta_{t_1} \cap \theta_{t_2})}{P(\theta_{t_1})} \tag{2}$$

And the stances depending on θ_{t_1} and θ_{t_2} as

$$P(S | \theta_{t_1}, \theta_{t_2}) = \frac{P(S \cap \theta_{t_1} \cap \theta_{t_2})}{P(\theta_{t_1} \cap \theta_{t_2})} \tag{3}$$

The overall schematic of this model is pictured in Figure 3.

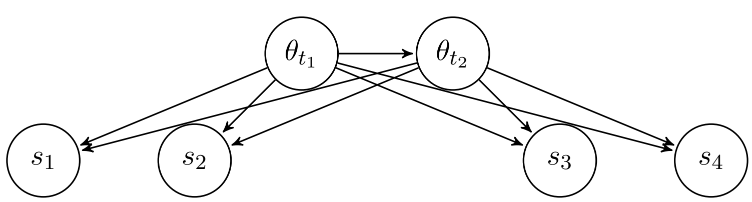

Figure 3: Model with a pair of dependent personality traits t_1 and t_2, where the distribution θ_{t_1} affects the distribution θ_{t_2}, and both affect the political stances s_1, s_2, s_3, s_4.

2.3 Model 3: Linear Regression

The final model we consider assumes there is a linear relationship between a candidates' political stances and how strongly the candidate possesses each personal trait. In this model, we also return to the assumption that each trait is independent of the other traits.

For the model, we express the candidate's stances and traits in matrices and find a linear hypothesis that minimizes mean squared error for each trait. This linear hypothesis captures whether each political stance increases or decreases how strongly people perceive whether a candidate has a particular trait.

3 Experiment

3.1 Survey

We designed a survey asking subjects to rate candidates on a scale of 1-5 (very low to very high) how much of certain traits a candidate possesses based on their political stances. We provided primers to in an attempt to keep semantic meaning consistent:

1. Openness: Willingness to try new things
2. Order: Desire to maintain structure
3. Stability: Desire to keep things the same
4. Security: Desire to eliminate threats
5. Compassion: Desire to support those who are struggling
6. Personal Liberty: How unwilling to compromise personal liberty for things such as the public good

The six candidates are all real candidates from the 2020 election, selected from various parts of the political spectrum to give a realistic snapshot of current platforms. The candidates were anonymized as 'Candidate X' and issue stances were collected from their campaign websites and various third-party political stance aggregators. In an attempt to avoid any potential bias from our subjective rating of how extreme a candidate's stance on an issue was, we binarized views as for or against. The issues and stances we selected were:

1. Supports or opposes Medicare for All
2. Supports or opposes school choice
3. Supports or opposes immediate government action to protect the environment
4. Supports or opposes easier pathway to citizenship
5. Supports or opposes strong gun control laws
6. Supports or opposes free trade
7. Supports or opposes increase in military spending
8. Supports or opposes increase in minimum wage

We selected these particular political issues and stances in order to provide subjects a general snapshot of each candidates' platform, but also leave blanks for the subject to fill in about what the candidate is like as a person.

3.2 Data

We had 31 participants in our survey. Most of the participants were students in the MIT community, which contributes to a partisan bias. For a particular issue i and trait t, the

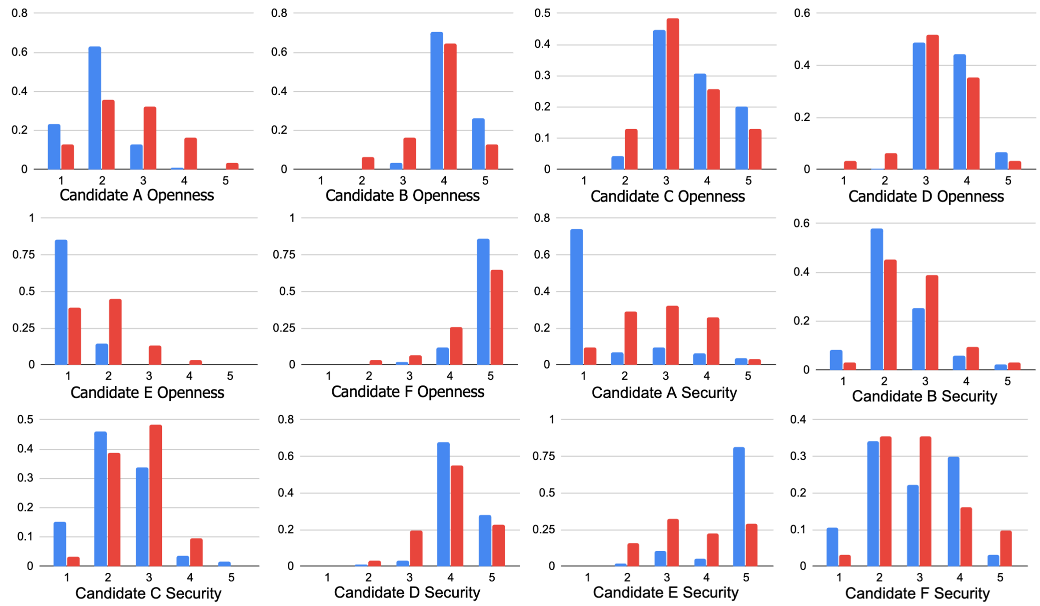

Figure 4: Graphs show distributions obtained from Model 1 (blue bars) versus distributions from survey responses (red bars) for the traits openness and security. Model predictions closely match the distributions obtained from polling.

probability of a candidate having stance $s \in \{0, 1\}$ on the issue given they possess value $v \in \{1, 2, 3, 4, 5\}$ of the trait is determined from Bayes' rule:

$$P(i = s | t = v) = \frac{P(i = s \cap t = v)}{P(t = v)}. \tag{4}$$

$P(i = s \cap t = v)$ was determined as the count of responses for candidates with stance s that were given a rating of v for trait t. The number of responses for a trait sums to the product of the number of candidates and the number of responses because each respondent rates each candidate on each trait. The resulting empirical probabilities were used in Model 1.

For Model 2, we also calculated the probability of a trait given another trait and the probability of an issue given a pair of traits using Bayes' rule. For traits t_1 and t_2, we calculate

$$P(t_1 = v_1 | t_2 = v_2) = \frac{P(t_1 = v_1 \cap t_2 = v_2)}{P(t_2 = v_2)}. \tag{5}$$

We then find the probability of the candidate having stance s on issue i using the equation

$$P(i = s | t_1 = v_1, t_2 = v_2) = \frac{P(i = s \cap t_1 = v_1 \cap t_2 = v_2)}{P(t_1 = v_1 \cap t_2 = v_2)}. \tag{6}$$

4 Results

Figure 4 shows the results from Model 1. Our model was able to closely predict the distribution of a candidates' traits using the empirical probabilities calculated from survey responses. The KL divergence between the distribution r of survey responses and the distribution m predicted by the model of the values of a particular trait t for a particular candidate is calculated with the formula

$$D_{KL}(r||m) = \sum_{i=1}^{5} r(x_i) \log\left(\frac{r(x_i)}{m(x_i)}\right), \tag{7}$$

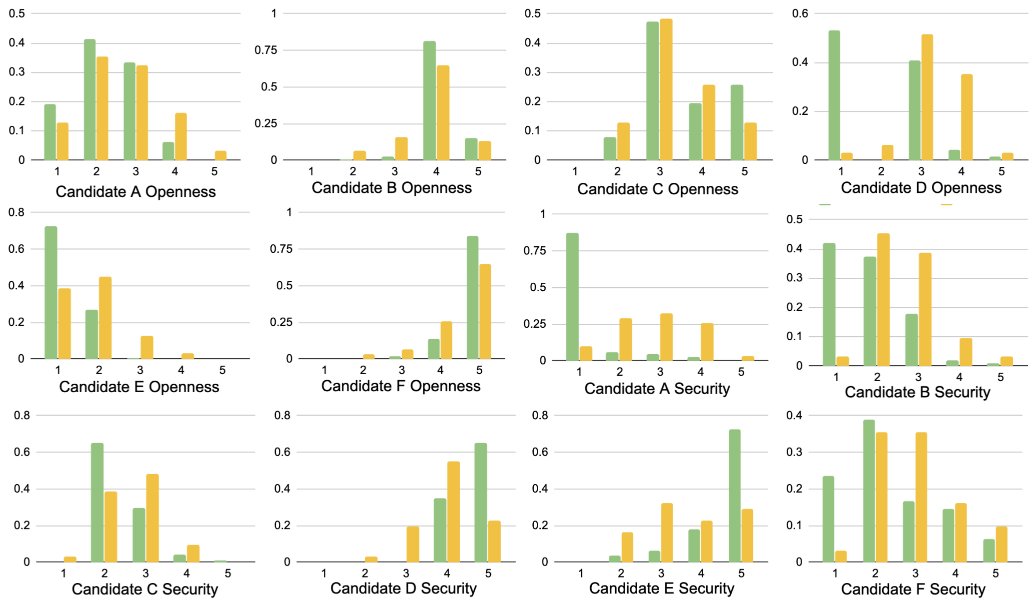

Figure 5: Graphs show distributions obtained from Model 2 (green bars) versus distributions from subjects' responses (gold bars) for the traits openness and security. Both traits were conditioned on the candidates' compassion distribution.

where x_i is the event that trait t has value i. The average KL divergence of the distribution predicted by the model and the actual distribution obtained by the survey for the openness of each candidate is 0.62735. This is lower than the average KL divergence between the uniform distribution and the survey distribution for openness of each candidate, which is 0.70217.

However, the average KL divergence between Model 1 predicted distributions and survey distributions for security is 0.54997, which is slightly higher than 0.53949, the average KL divergence between the uniform distribution and the survey distribution for security. This indicates that the Model 1 predictions are a better approximation of the survey distributions than the uniform distribution for openness, but are a slightly worse approximation than the uniform distribution for security.

Figure 5 shows the results from Model 2. This model was able to predict the distribution of a candidates' traits while conditioning on the distribution of another trait. The average KL divergence of the distribution predicted by Model 2 for openness ratings and the distribution obtained by the survey is 0.56354. This indicates that Model 2 provides a better approximation for the openness distribution than the uniform distribution, which has an average KL divergence of 0.70217.

On the other hand, when conditioning the distribution of security on the distribution of compassion, the average KL divergence of Model 2's predictions compared to the survey distributions of security is 1.07342. This is much higher than the KL divergence of the uniform distribution and the survey distribution of security, which is only 0.53949. Possible reasons for the lack of consistency in Model 2's predictions are discussed in the following section

The coefficients and intercepts obtained from Model 3, linear regression, are displayed in Table 1. We conducted linear regression using the average ratings for each candidate's traits and binarized matrices of each candidate's stances, with a 0 indicating opposition and a 1 indicating support. The linear hypotheses were able to closely fit the data points and provide further insight into the impact of each political stance.

Table 1: Coefficients and intercepts of each trait obtained from performing linear regression. The scale was 1 for low in a trait and 5 for high in a trait.

Trait	Openness	Order	Stability	Security	Compassion	Personal Liberty
Medicare	0.72125	-0.05977	-0.40645	-0.42960	0.72675	-0.25389
Public School	0.27226	0.11211	-0.04516	0.08773	0.42188	-0.20499
Environment	0.07922	0.01983	0.17742	-0.15091	0.05009	0.10816
Citizenship	0.13216	-0.37201	-0.12580	0.08513	0.12979	0.42656
Gun Control	0.50528	0.36884	-0.15161	0.26578	0.68678	-0.60139
Free Trade	0.17213	-0.20755	-0.00967	0.02454	0.02289	0.34332
Military	-0.52821	0.15205	0.18387	0.66824	-0.35496	-0.05926
Min. Wage	0.27226	0.11211	-0.04516	0.08772	0.42188	-0.20499
Intercept	2.36154	3.38127	3.31612	2.96508	2.08829	3.22593

5 Discussion

We found that some traits are better inferred from political stances than others. The distribution of ratings for openness fit well to our probabilistic generative models. Using uniform as a baseline, both Model 1 and Model 2 performed better, which implies that there is some relationship between political stances and perceived candidate openness. Model 2, which conditioned the distribution of stances on both compassion ratings and openness ratings, was able to produce the best approximations of the survey distribution. We think this may be because because a candidate that is both open and compassionate behaves more stably on an issue. It could also be an artifact of over-fitting to the data. It's unclear without a larger scale survey with more candidates if the improvement is due to better representation of the structure of the problem.

On the other hand, the uniform distribution proved to be a better approximation of the survey distribution for security ratings than both Model 1 and Model 2's predicted distributions. This suggests that survey participants didn't have a consensus on how much someone valued security based on the political issues presented. This is surprising given military spending was included, which logically should be a strong indicator of how much someone valued security. It's possible that the political leanings of our survey participants biased their perception of what military funding meant to security. In Model 2, the distribution of security ratings was conditioned on the distribution of compassion ratings. Compassion was purposely selected to have very little correlation with security. Using uncorrelated attributes didn't seem to help, but we think this may be due to insufficient data that leads to omission of some trait value pairs that interfered with the calculation of our trait joints.

From the coefficients of our linear hypotheses in Model 3, we are able to gain new insight into the relationship between political stances and perception of candidate traits. For openness, we see that supporting Medicare for All and stronger gun control laws contributes the most to a higher openness rating, while supporting increased military spending contributes the most to a lower openness rating. We also see that a candidate supporting an increase in military funding does make participants more likely to rate them as valuing security. Most of these relationships make sense intuitively, but we would need a larger scale survey with more candidates with assorted stance positions to be more sure on each issue's individual impact on each trait.

Our current models demonstrate a clear relationship between a candidate's political stances and the public perception of the candidate's personal traits. There are some interesting ideas in this area that have the potential to be further explored, as detailed in the following section.

6 Reflection

It feels natural that some traits are more correlated than others and a model that predicts someone cares about stability should also be more likely to predict someone values security, and so we tried a model that allowed for dependence between traits. However, we didn't see much of an improvement because we were lacking in candidate combinations. When we loosen the mutual independence assumption between issues and traits, we found some of the candidates had pretty unexpected distributions because some pairings of candidate trait values, such as openness = 4 and stability= 5, did not exist in our dataset.

We also considered trying a model that assumed issue stances were a function of all the traits, but it would require far more candidates to avoid interference between stances. Our current scope would not allow us to get reasonable results from this model because we do not have enough candidates to calculate reliable conditional probabilities for $P($ trait = value $\mid$ issue1 = 0, issue2 = 1 $\ldots)$. The potential framework for this model is displayed in Figure 6.

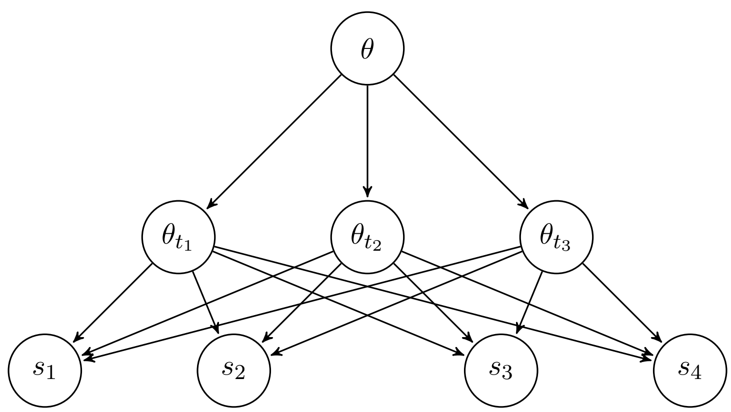

Figure 6: Model with personality parameter θ that affects personality traits $\theta_{t_1}, \theta_{t_2}, \theta_{t_3}$, which in turn affect political stances s_1, s_2, s_3, s_4. This structure does not assume that each trait has an independent distribution.

References

Baker, C. L., Jara-Ettinger, J., Saxe, R., and Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires, and percepts in human mentalizing. *Nature Human Behaviour*, 1(64).

Baker, C. L., Tenenbaum, J. B., and Saxe, R. R. (2007). Goal inference as inverse planning. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 29(29):779–784.

Caprara, G. V., Vecchione, M., Barbaranelli, C., and Fraley, C. (2007). When likeness goes with liking: The case of political preference. *Political Psychology*, 28(5):609–632.

Doherty, C., Kiley, J., and Johnson, B. (2016). 2016 campaign: Strong interest, widespread dissatisfaction. Technical report, Pew Research Center.

Hayes, D. (2005). Candidate qualities through a partisan lens: A theory of trait ownership. *American Journal of Political Science,*, 49(4):908–923.

Jara-Ettinger, J., Gweon, H., Schulz, L. E., and Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences*, 20(8).

Koppensteiner, M. and Stephan, P. (2014). Voting for a personality: Do first impressions and self-evaluations affect voting decisions? *Journal of Research in Personality*, 51(100):62–68.

Levy, S. R., Stroessner, J. S., and Dweck, C. S. (1999). Modes of thought: Implicit theories and social understanding. *Journal of Personality and Social Psychology*, 74:1421–1436.

Plaks, J. E., Levy, S. R., and Dweck, C. S. (2009). Lay theories of personality: Cornerstones of meaning in social cognition. *Social and Personality Psychology Compass*, 3(6):1069–1081.