Finding The Swing Voter: Definitions and Survey
Methods for Voter Classification

by

Charlotte Swasey

Submitted to the Department of Political Science
in partial fulfillment of the requirements for the degrees of

Master of Political Science
and
Bachelor of Political Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2016

© Massachusetts Institute of Technology 2016. All rights reserved.

Author

Department of Political Science
August 5, 2016

Certified by

Adam Berinsky
Professor of Political Science
Thesis Supervisor

Accepted by

Ben Ross Schneider
Chair, Graduate Program Committee
Finding The Swing Voter: Definitions and Survey Methods for Voter Classification

by

Charlotte Swasey

Submitted to the Department of Political Science on August 5, 2016, in partial fulfillment of the requirements for the degrees of Master of Political Science and Bachelor of Political Science

Abstract

This thesis proposes a theory mapping emotional reactions to political information onto a theory of vote decisionmaking and then further onto measurable survey response. Using on-line processing based in emotion, voters form affective summaries about candidates, which store previous information as an emotional response. The act of voting is treated as a single realization of a probabilistic event, with the relative probabilities of each vote option being an expression of the affective summary. These summaries are expressed as warmness or feeling towards each candidate, which can be captured using the ANES Feeling Thermometer scales. A metric of the difference between the scores given to the Republican and Democratic candidates is used, based in the work of William Mayer. This metric suffers from significant survey error, but is related to party ID and expressed vote choice, as well as demographic factors and perceived efficacy. Feeling thermometer responses are found to carry meaningful information about a respondent’s relationship to the election and candidate preference.

Thesis Supervisor: Adam Berinsky
Title: Professor of Political Science
Chapter 1

Introduction

The study of voting behavior has long been a major topic for political science. Presidential elections are some of the largest political events in the US, and involve participation from a large portion of the population. The prediction and understanding of this involvement is of clear interest. Elections can be considered a phenomenon at two levels, the population and the individual. Population level study, such as that of ideological shifts among demographic groups, benefits from the aggregation of data and ability to detect trends in large samples. Individual level study, however, rapidly runs into a web of difficulty.

While it is entirely possible to describe an individual's political decisions, which in an election are as simple as a binary choice, explaining and predicting them is a much more difficult problem. The mechanisms of decision are internal and not easily observed, and their form is under heavy debate. How does an individual come to make a choice between candidates, and how can predict what choice they will make? While cognitive models of decision have been proposed and remain popular, they seem incomplete when applied to individuals with little political knowledge and interest. One major variety of alternative to these models involves the incorporation of emotion as a driving force.

Research on the role of emotion in politics has been strongly theoretical (see [42] for a review), and hampered by the difficulty of measuring a thing as vague as emotional reactions. Information processing theories based in emotion have provided
a substitute for more cognitively based theories of the intake and use of political information [40, 60] that reduce the cognitive burden on a voter. This is supported by research finding emotional reactions to potentially have primacy over cognitive processes[34]. These implicitly address the issues of low information and low interest faced by theories that posit strict rationality from voters, but currently fail to provide a link to measurement strategies.

This paper lays out a theory to map emotional reactions to political information onto a theory of vote decision making and then to measurable survey response. Grounded in research on the important role of emotion in understanding citizen’s relationship to political events [3, 12, 32], it attempts to bridge the gap between theories of opinion formation and measurement tools for vote prediction. Furthermore, it shows a method by which the task of classifying voters can be performed under an emotion-based theory of decisionmaking. Voter classification is key to the activities of election campaigns, as well as to the study of populations, and is at present poorly addressed by theories of emotional decision formation.

This paper works to develop a theory of the link between the formation of candidate preferences, their reporting on surveys and their expression in the voting booth in three major steps. First, it walks through an exploration of the role of emotion in vote choice. Rather than rely explicitly on the known importance of factors such as party ID, or expect high levels of information from the voter, the affective summary theory of voter information provides a mechanism for voters to aggregate information about candidate, potentially including such major cues as party ID and incumbency. Relying on emotional mechanisms, it reduces the burden of rationality and processing on the individual voter. While there is an ongoing debate over the dimensionality of emotion, with regards to political preferences, the relevant aspects for candidate and vote choice seem to be approximated by a single dimensional measure.

Second, it presents a proposal of the probabilistic model of voting, based in the emotional decision-making discussed in the first section. Rather than treating most vote-decisions as fixed (for example, a lifelong democrat voting a presidential election), this model treats every vote outcome as a single realization of a probabilistic
event. As such, what would previously be considered "fixed" votes are extremely high probability events, while less certain, or "swing", votes have probability distributions approaching 50/50. These probability distributions are a translation of the affective summaries discussed in the first section.

Finally, it proposes a measurement strategy for accessing both a voter's affective summary and probability distribution. This is based in literature around the ANES feeling thermometer, which asks for the reporting of emotional "warmth" towards candidates or objects. The link between this survey instrument and behavior is not immediately clear, given the lack of behavioral information in the ANES. Other influencers of thermometer scale scores are explored, notably self-reports of interest in the election and perceived efficacy.

Fundamentally, the goal of this work is to connect the study of emotion in politics with the practical study of voter classification, with an emphasis on swing voters. This linkage helps to ground the former in survey-based measurement strategies, and provides a set of observable behavioral implications on which to test the theory.
Chapter 2

Emotion, Decision-making, and Candidate Choice

2.1 Voter Decision-making

Election Day is an occasion when citizens who chose to vote (admittedly a subset of the population) have to make a clear-cut decision on a political issue. As fuzzy and unclear as their preferences may be, elections force their collapse to a simple, frequently effectively binary, choice between candidates. Given the candidates, how does a voter come to select an option?

While a hypothetical fully informed voter would make a judgment of the candidates’ policy position relative to their own (as under the spatial model of voting, [15]), and thus select a candidate who’s views most closely align with theirs, most voters seem to have little information about political issues, to a level such that they seem unable to reasonably cast issue-based votes [8]. This "Michigan voter" is seemingly unqualified for the task of political participation by their inability to (or lack of interest in) understanding political issues and candidate positions.

In response to this, and to account for the perceived ability of voters to react to changes and hold legislators accountable even without full information, theories of
retrospective voting [18] and pocketbook considerations [25] instead propose voters who are capable of drawing reasonable conclusions about politics from their own personal circumstances. These theories argue that voters are able to select candidates who are acceptably good for their own self interest by using simple cues from their surroundings, in a quality of life based, rather than issue based, process.

Their hypothetical voters answer simple questions such as "are you better off than you were four years ago?" [26] While this is far from a rational choice view [15], it preserves some confidence in the capabilities of voters, and explains how voters can seem to hold government accountable for outcomes without any significant knowledge about government. A variant of this theory is that of policy mood and thermostatic voting, which effectively apply the individual level theory of simple positive/negative reaction to the recent past to an aggregate level [56].

2.2 Information Processing Theories

For the purposes of voter decision prediction, we are working at the individual level. Retrospective voting is a form of "low information heuristic" to simplify the cognitive process needed to form opinions or evaluations [54]. These heuristics are used by voters as "substitutes for encyclopedic information" [38]. Use of heuristics can refer to cuetaking from parties or elites in various forms [15, 2], but they have also been used to refer to extrapolation from limited information, as in pocketbook voting, above. Thus, voters need only look at some easily accessible source of information, such as the party identification of a candidate or the economic state of their town, in order to form a political judgement. This integrates the existence of voter political opinions with the relative scarcity of voter political information.

Similarly, Popkin's low information rationality expands on the theory of cuetaking and extrapolation, presenting a number of "shortcuts and rules of thumb" to be used when evaluating politicians and their positions, which are all based in a process of con-
tinuous updating of old information with new [46]. Much like retrospective voting, low information rationality proposes that voters gather relevant information as a byproduct of their daily lives, with the addition of assuming exposure to campaigns. Thus, heuristics here are used to evaluate campaign information and candidate information.

While these models do provide compelling explanations for the formation of voter opinions in the absence of strong interest of complete information, they imply that voters are undertaking a cognitive process to form an opinion or evaluation at the moment it is needed. References are made to opinions being continuously updated over time, but no particular mechanism is proposed. Furthermore, these models position decision-making, even under serious constraints and with the minimum possible information, as rational. The addition of emotion to the opinion formation process, next, moves away from rationality and pure information-based decision-making into considering the emotional reactions which seem to be so important in politics.

2.3 Emotion as Evaluation

One of the primary contentions in the study of emotion in politics is that emotional reactions to political stimuli are as or more important that rational, logical reactions. Emotion shapes the way that individuals evaluate political issues at a level that is not well explained by rationality [52, 53, 12]. From campaign ads [4] to candidate behavior [57], it is further well documented that emotion is a strong determinant of an individual’s reaction to and assessment of political figures. The role of emotion is equally strong in voting and candidate choice [40, 44]. The exact components of this emotion, and the way they are activated by various types of appeals, is firmly still under debate.

Affect is also highly important in predictive work. Studies have found affective response to be more accurately predictive of opinion than cognitive judgement [12]. Other scholars have taken up the "affect-as-information" approach, which similarly
hypothesizes that individuals make decisions based on current affect [11].

2.4 Emotion as Processing and Storage

While it is clear that emotions are specifically targeted and aroused by political campaigners [4], the momentary existence of emotions in considering political issues and people does not necessarily imply that it will have a large impact on a voting decision made many months later. For that to be true, there needs to be some mechanism to turn a series of emotional and factual data points into a judgment, and ultimately a decision.

How does emotion persist from campaign events, such as the viewing of an anxiety-provoking advertisement, to the voting booth? Clearly voters cannot be expected to recall all the emotions they experience over the course of the campaign, or even to explicitly notice them at the time.

On-line processing, or "impression driven processing" uses emotion as a storage mechanism to reduce the cognitive load of decision making [36, 35]. Under this model, voters create and update emotional evaluative summaries about candidates as they are exposed to various pieces of information, and then forget the original information on which the evaluation was based. This is contrasted with memory-based models, which require a voter to hold in their head a full set of candidate information and create a preference from there [31]. Instead, on-line processing only requires voters to keep a "running tally" of their preference between candidates, which reduces the amount of information in working memory and preserves cognitive economy [35]. This type of processing can function in an environment where the informational content voters receive is either policy oriented (ads about each candidate's position on gun control, for example) or personal (as in retrospective voting).

Under impression driven processing, the "running tally" is counting positive and
negative impressions towards the candidate, or more simply likes and dislikes. This is a "summary affective evaluation" [36], a measure of how positively a voter feels towards them.

Operating under a similar hypothesis of affective evaluation, but with different causes, is the theory of "hot cognition" [45, 37, 9]. This proposes that the inputs, sociopolitical information and ideas, have a particular affective feel, and that this feeling is processed and stored by individuals as information. The result is very similar to the affective summary of on-line processing, but assumes that political concepts are inherently affectively charged, rather than that a cognitive process must occur to convert them to affect. The view of emotion as a tool of information processing, existing alongside cognition, is also echoed by Marcus and MacKeun [40]. Affective Intelligence theory also assumes that individuals tend to default to a stored reaction and evaluation, very similar to the summary affect evaluation of on-line processing, unless provoked to reconsider attitudes. It presents two basic emotional systems. The first monitors the environment for familiarity (positive feeling) or unfamiliarity (negative feeling). Familiar and positive environment's provoke reliance on habitual choice, whereas unfamiliar ones invoke deliberative choice [41]. Thus, anxiety cues rejection and reconsideration of established attitudes.

Supporters of this model have applied it to various aspects of political behavior, including processing of racial cues in campaigns [28], response to emotional appeals in campaign ads [3], and the learning process of campaigns and voting [10]. "Habitual choice" in this theory is defined as choice based on routines and heuristics. In a familiar situation, voters avoid the cognitive load of generating a new judgement, and rely on one previously developed. This is consistent with the on-line processing model, with habitual choice defining environments in which the current affective summary will be relied upon.

Under all these models, assuming that a voting booth is a reasonably familiar
environment and unlikely to provoke serious reconsideration of beliefs, the affective summary is the key metric in determining candidate choice. The candidate choice process can be modeled as a comparison of two (or more) relative affects. The use of an affective summary as a decision-making tool makes sense from an efficiency standpoint. As discussed above, processing and integration of information is cognitively expensive, whereas affect is precognitive and faster than logical reasoning [14, 60]. Thus, emotional evaluation simplifies cognition, and a voter’s affect functions as mechanism for storing an evaluation of each of the candidates, for comparison on Election Day.

2.5 Dimensionality of Emotion

Assuming that affect as a summary of information about candidates is the key component of a voting decisions, the question becomes how to understand and measure it.

The simple positive/negative "summary affect evaluation" from Lodge et al. implies a collapse of complicated evaluations to a single dimensional scale of affect and emotional feeling towards a candidate. This is assumed in their model to be simplest information for an individual to store. However, researchers of emotion have proposed three basic models of how emotional evaluation functions: valence, discrete, and multidimensional [9].

The first model, valence, is analogous to the positive/negative tally, placing affect on a single axis running from positive to negative. This considers emotional response to effectively collapse to a single dimension [22]. While valence models do not necessarily refute the existence of multiple types of emotion (enthusiasm vs happiness, for example), they do propose that the most important component of these emotions is their positivity or negativity. Hot cognition and on-line processing both rely on this structure of emotional response. It is also the most common implicitly referenced
model, likely due to its simplicity. This model also makes the comparison of affect for choice process intuitive, since feelings towards each candidate can be ranked at a point in a single dimensional scale, and trivially compared.

Discrete models separate affect into some number of different elements which can vary independently, and cannot be summarized as negative/positive [42]. The exact element identities vary between users, one example being the Richard Lazarus's model using Anger, Guilt, Fear-anxiety, Sadness, Hope-challenge, and Happiness. Kinder and Abelson, in measuring emotions induced by political candidates in particular, have focused on hope, pride, sympathy, disgust, anger, fear, and uneasiness, which have become folded into the context of the National Election Study [1]. These models are differentiated from pure valence models by the assumption that each emotional element carries different information, above its negative/positive characteristics, and that emotions cannot all be arrayed as opposites on a negative/positive scale (think, for example, of calmness and excitement, which are seemingly opposites but both seemingly positive) [49, 34, 11].

Some scholars have moved from consideration of a discrete model to explicit focus on the predictive power of individual emotions, notably fear, anger, and anxiety [51, 55]. Under AI, anxiety is theorized to induce individuals to reconsider existing beliefs and preferences. These particular emotions are studied partially because of their clear use in election campaign materials [4]. However, multidimensional models further complicate by theorizing that the multiple element of emotion covary, and are better represented as dimensions on some sort of emotional space, commonly a circle or "circumplex", reviewed in Marcus, Neuman, and MacKeun [39]. Users of this model tend to be more interested in representing psychological processes than predicting behavior, and their measures of emotion are necessarily very complex.

The theories of dimensionality of emotion specify how many emotions or dimensions need to be recorded in order to have a complete summary of a respondent’s
emotional state. Even for scholars operating fully under the multidimensional model, there is a tradeoff between capturing more dimensions and reducing the accuracy of measurement by asking too many questions (or too complicated questions) of a respondent. This choice depends to a large degree on the use case of affective measures. The case specified in this paper, candidate choice prediction, is discussed below.

2.6 Use Case: Candidate Choice Prediction

Affect measures have been specifically applied to the domain of candidate evaluation and candidate choice [1]. The decision between candidates is theorized as the process of comparing affective summaries for each, and determining which is most positive. Questions asking about voter's affect towards each candidate should be activating the same process as the actual voting decision.

The theories of information processing themselves do not inherently make claims about what the summary affect evaluation will be based on, other than "information and cues" and "personal circumstances". The components put in to this summary judgment are important for contexts where we are concerned about representation or normative questions of voter candidate preferences, but unimportant if we are only concerned about predicting voting behavior from survey responses.

Likewise, much of the potential detail of affect is not necessary for this use case. While on an abstract level, we might wish to record and understand the full complexity of emotional reaction to a candidate, for this application we need only the elements of emotion that are significantly predictive. Additional dimensions or complexity that do not cause a significant increase in accuracy of prediction can be safely left out, and their inclusion may even harm prediction by increasing the complexity of the required measurement strategies. As shown below, the full complexity of emotion can be safely collapsed into a more measurable form.
2.7 Implications of Vote Choice as Affective Summary Comparison

Under the model proposed above, vote choice is the comparison of the relative affect towards each candidate. Thus, a voter's decisions is based on an emotionally stored summary of information gathered over the course of the campaign (or potentially earlier, including partisan preference and demographic factors). This theory indicates that voter-reported reasons for their vote choice, for example the lists of likes and dislikes for each candidate frequently gathered in election surveys, are justification of a decisions that is, at that point, being made on an emotional basis.

Further, this defines undecided or uncertain voters as those who have similar affect towards both candidates. Given the processing theory presented earlier, this state could occur one of three ways. A voter could feel positively towards both candidates, negatively towards both, or be roughly neutral towards both. These options provide no hypothesis about an undecided voter's level of information, as there is no immediate reason a lower- or higher-than-usual information candidate could not fall into any of those categories. Increased information going in to the affective summary does not automatically mean greater differentiation between candidates.

However, this still leaves open the question of how best to model the decision itself.
Chapter 3

Voting Behavior

3.1 Probabilistic Vote Choice

There are many conceptions of the source of voting behavior, and how it should be best modeled. Given this project's focus on uncertainty of vote choice, and those with unsettled preferences, I propose and rely on a probabilistic model of voting behavior.

Each voter is conceived of as having some probability of making a particular choice on election day. For purposes of discussion, let us define the outcome of interest as voting for the Democratic candidate. An individual voter might have, for example, a 60 percent probability of voting for the Democrat and a 40 percent probability of not doing so, conditional on their making it to the voting booth on election day.

These probabilities are the result of a number of factors affecting an individual voter. Under this model factors such as party ID, income, and race serve to shift a voter's odds of voting a particular way. A strongly Democratic identified voter might then have a 99 percent chance of voting Democratic, while a centrist might have only a 50 percent chance of doing so.

On election day, each person's vote is a single random draw based on their underlying vote probability. Thus, for any given group of voters, the rate at which they vote Democratic can be considered an approximate average of the probabilities of individual members to vote Democratic.

Individual probability of voting Democratic is from hereon referred to as "vote
probability". This should be distinguished theoretically from probability of voting *at all*, that is, making it to the polls on election day. Probability of voting at all will be briefly addressed later on.

Under this conception of election day vote choice as a single realization of a probabilistic outcome, without additional information, a voter’s choice one election tells you nothing about their likely choice next election.

However, given observations about the relative stability of vote choice by voters over multiple elections, we can safely say that the majority of voters have probabilities of voting democratic that are on the extreme ends of a normal distribution. Under the probabilistic outcome model, by observing many voters "drawing" Democratic again and again across time with high consistency, we can infer that their underlying probability of voting Democratic is high, even if we cannot determine it precisely. As such, the distribution of Democratic vote probabilities across the voting population is likely strongly bimodal, but the stability of individual vote choice does not invalidate the model.

### 3.2 Proxies for Vote Probability

We have no direct way of getting at an individual’s vote probability. There is no reason to believe that a individual should be able to accurately report it (as confounded by response bias as such a measure would be), and in the absence of many iterations of the same election, there is no real way to calculate it from behavior. As a proxy, we need a voter’s evaluation of a candidate, on some consistent scale, relative to their evaluation of the other candidate. This at least provides an indication of which way their vote probability is skewed, and likely some indication of the magnitude of that skew.

As discussed in the previous section, voter evaluations of candidates are stored as affective summaries, or feeling towards each candidate. This summary can be measured using the ANES feeling thermometers. While the precise correlation between the thermometer scale measurement and vote probability is unclear, one appears to
exist, and will be discussed later.

3.3 Persuadability

Whatever the model, it is clear that most voters have a strong tendency to vote one way constantly, even across elections. Persuadable voters are those whose vote choice can be shifted through techniques such as advertising and canvassing. Under the probabilistic model, persuasion is shifting a voter's vote probability. Persuasion may "succeed" by shifting a person's vote probability from 80 percent to 75 percent. However, as this example shows, small magnitude persuasion can only likely impact effective vote choice in the marginal population who do no have a particularly high probability of voting for either candidate. Assuming that it is highly unlikely for a particular persuasion strategy (say, a mailer) to be to able to shift someone's vote probability by more than a small amount (arbitrarily, say, 15 percent), for maximum effectiveness any such outreach would need to be targeted at voters with nearly 50 percent vote probability. These people are what is traditionally referred to as "swing voters": very likely to turn out, uncertain about their candidate choice. Their identification is interesting theoretically but also of practical use to campaigns and organizers.

3.4 Population Probability and Survey Response

As mentioned above, probabilistic voting conceives of election day vote choice as a random draw. On a deeply simplified level, for a given subpopulation, the election day vote share is the average of the vote probabilities of the component individuals. In working with the data we have, we can attempt to work out the vote probabilities of group members from that group's overall vote share. In term of swing voter identification, we should expect a group of swing voters to have a roughly 50-50 vote share.

The survey data in the ANES provides two measures of vote behavior. The first is a
prediction of eventual vote behavior, asked pre-election. The second is a self-report of actual vote choice, asked post-election. These two reports can both be conceptualized as random draws in the same way as actual vote on election day can be.

As such, we should expect most individuals to report the same vote choice in both waves. This matches with the data, where 97.3 percent of respondents report the same candidate choice in the pre and post election waves. Of respondents who reported both intent to vote and having voted, only 53 reported changing presidential candidate choice between survey waves.

3.5 Nonvoters

The other major component of election day vote share is voter turnout. Likelihood of turning out on election day, along with vote probability, provides a more complete measure of a person's expected added vote-value on election day. However, turnout is very difficult to predict. Direct questions about intent to vote (such as are included on the ANES) produce very high percentages of "likely voters", much higher than are observed on election day. Post-election measures also suffer from this problem, with high percentages of survey respondents reporting having voted. Only a small percentage of survey respondents report intent to vote pre-election and then report not having voted post-election. These people will be referred to as "dropouts". While not expected to be a complete set of those who actually failed to vote on election day, their characteristics can give clues as to the actual characteristics of the nonvoting population. Specifically, their distribution over the thermometer scores will be discussed in detail later.
Chapter 4

Research Design

4.1 ANES and Measurement

The ANES is a national election study administered every election year. A variant of it, the time series study, has a pre-election and post-election section, administered to the same panel of respondents. The pre-election survey is conducted over the course of 2 months before the election, while the postelection part begins on November 7th. This study is extremely useful for the study of voting behavior, since it gives both a snapshot of expected or predicted behavior before the election and the ability to cross check predictions with reported behavior. The post-election survey is soon enough after the actual election that the error should mostly be intention misreporting respondents, and not forgetfulness.

The study covers a number of questions, dealing with attitudes about politics in general, as well as issue specific ones. Most of the questions have be run in a similar form for many years, some beginning as early as 1948, which makes it very easy to compare results across the years. It also provides a fairly comprehensive profile of respondents, since it contains demographic data as well as questions about personal situation. Of particular interest to the study of swing voters is the pre-election questions about probably candidate and party choice, as well as party affiliation and intention to vote. A number of traditional measures of ideology are asked of each re-
spondent, which provides a useful ability to compare between reported partisanship, ideology, and party/candidate preference.

4.2 Validating The Feeling Thermometer

The feelings thermometer has been present in the ANES since 1964, originally only covering demographic groups such as Protestants and Southerners. In 1968, it was expanded to include various public figures, notably political candidates and prominent politicians. The thermometer asks respondents to rate their feelings for a group or person on a scale from 0 to 100, with 100 being "very warm", 50 being "neutral" and 0 being "very cold or "not warm at all". It is an emotional scale, which captures an individual's "affective sentiment" [58], allowing a respondent to collapse multiple possible dimensions of candidate or group evaluation into a single measure.

4.3 Foundation: Mayer

Mayer provides an operationalization of the concept of the swing voter based in the feeling thermometer method [43]. He subtracts the Democrat’s thermometer score from the Republican’s, creating a "difference in feeling" metric, scaled from -100 to 100. Negative scores represent preference for the Democratic candidate, while positive scores represent preference for the Republican.

Mayer uses an absolute value scaled version of these difference scores. He argues that a voter with low differentiation (a low absolute difference score) between the candidates is likely to be unconvinced that one candidate is better than the other, and therefore fit his definition of a "swing voter". Thus, he defines those within the 0-15 degree of difference range as likely swing voters, given the relative similarity of their feeling towards the two candidates. The 15 point mark was chosen based on his research into the percent of the population who voted the way they indicated they likely would in the pre-election survey. These people make up about 27 percent of the 2008 ANES sample, an impressively (and worryingly) high percentage for a
population traditionally thought to be very small and very marginal. I theorize that his definition is encompassing people not traditionally thought of as swing voters, leading to the inflated percentage.

4.4 Defining Voter Population

The 2008 ANES Time Series study is done in a two-wave panel design. The pre-election survey includes the feeling thermometer questions used for this analysis, and a question asking about respondents' intent to vote in the upcoming election. The post-election survey is run within two months after the election, and asks respondents to report how they voted in November.

As should be expected, there is a significant difference between the number of respondents reporting intent to vote and the number reporting having voted, 57 and 60 percent of the sample, respectively. Since the ANES is not validated against voter file data, it is difficult to know precisely how accurate the post-election report of voting is. Some preliminary analysis involving the use of proxies for nonvoting behavior will be discussed later.

4.4.1 Revalidation of Mayer on 2008 Data

Mayer’s work averages over many years of ANES studies. This increases his sample size, and allows for examination of the usefulness of the scale across time. This paper seeks to apply his methods to a single election, that of 2008, between Barack Obama and John McCain.

The distribution of difference scores is shown below. There is a significant spike at 0, which likely reflects survey bias instead of the true incidence of indifference between the candidates.

One of the main drivers of Mayer’s theory is the correlation between difference scores and reported candidate choice. Due to the way the difference metric is constructed, very negative scores mean that the respondent felt more warmly towards
the Democratic candidate, and very positive scores mean the respondent felt more warmly towards the Republican. Mid-range scores imply that the respondent felt similarly about both candidates, or exhibited a low degree of differentiation between the two. As such, one would expect to see very high and very low difference score populations voting almost entirely for the respective candidate. This is reflected in the data, where nearly 100 percent of the very extreme (over 70 or under -70) difference score populations indicate intent to vote for their higher-rated candidate. This positions these populations as those with very skewed vote probabilities, being highly likely to vote for one candidate.

If any preference for a candidate was enough to ensure a vote for them, even the mid range scores should give a clear indication of which candidate the respondent will vote for. However, this is not the case for scores at the middle of the scale.

The percent of respondents indicating intent to vote for their higher rated candidate stays very high (over 90 percent) below -15 and above 20. Between those points, the odds of voting for the higher rated candidate drop dramatically. The population scoring between -15 and 20 thus likely has relatively equal vote probabilities, and the distribution of reported vote choice reflects nearly coin-flip vote events.
4.5 Use of the Feeling Thermometer

While in theory the feeling thermometer is a continuous 0-100 scale, in practice some points get selected far more often than others. Most people, when confronted such a scale, choose round numbered points. This feeds into the intuition that data from the thermometer is representative of rank order information than interval information. A plot of scale points against selection frequencies is below.

Immediately visible in this data are a handful of commonly selected points. 0, 15, 30, 40, 50, 60, 70, 85, 100. The percentages choosing the extreme two points (0 and 100) are somewhat lower. Thus, this scale seems to function as a 9 point Likert scale.

These points are overwhelmingly more popular than the rest of the scale, and the vast majority of the scale points are not used at all.

4.6 Alternate Way of Considering Scale

Given the distribution of scale use as discussed above, and the significant tendency of respondents to choose one of 9 main points, the scale closely resembles a 9 point Likert
Thermometer Scale Use

![Bar chart showing frequency distribution of difference scores.](image)

Figure 4-3: Distribution of difference scores

Table 4.1: Most Selected Points by Percentage of Total Respondents

<table>
<thead>
<tr>
<th>Score</th>
<th>Obama</th>
<th>McCain</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>6.9</td>
<td>6.8</td>
</tr>
<tr>
<td>15</td>
<td>7.2</td>
<td>7.1</td>
</tr>
<tr>
<td>30</td>
<td>9.3</td>
<td>6.7</td>
</tr>
<tr>
<td>40</td>
<td>11.1</td>
<td>7.7</td>
</tr>
<tr>
<td>50</td>
<td>16.8</td>
<td>13.6</td>
</tr>
<tr>
<td>60</td>
<td>14.9</td>
<td>12.5</td>
</tr>
<tr>
<td>70</td>
<td>14.1</td>
<td>14.2</td>
</tr>
<tr>
<td>85</td>
<td>12.9</td>
<td>16.3</td>
</tr>
<tr>
<td>100</td>
<td>3.9</td>
<td>11.4</td>
</tr>
<tr>
<td>Other</td>
<td>2.8</td>
<td>3.7</td>
</tr>
<tr>
<td>Candidate Avg</td>
<td>51.9</td>
<td>57.7</td>
</tr>
</tbody>
</table>

24
scale. As such, a difference score of under 15 points effectively means "selected two adjacent scale points". This provides some validation to Mayer’s selection of 15 points difference as the benchmark for swing voter status. Assuming voters are functionally treating this as a 9 point scale, selection of two adjacent, non-identical points would give a max difference score of 15 points. "Low difference score" is hereafter defined as exhibiting that under-15 point difference between candidate evaluations.

4.7 Zeros

As mentioned above, there is an extremely large portion of the population reporting zero difference between the two candidates. These people are distributed across original thermometer scores (D or R scores) as such:

![Distribution of Zero Difference Score Respondents](image)

The large spike at 50 indicates that a large portion of these people are reporting zero difference between candidates as a result of impatience with the survey itself, not as an accurate report of opinion.
4.7.1 Dropouts

Respondents who reported intent to vote pre-election and failed to report voting post-election (hereafter called dropouts) are unevenly distributed over the range of difference scores.

Working as he is off of self reported survey data, Mayer has to make assumptions about the accuracy of his data. He is also working with a sample of the general population, not one that has been filtered to include only the voting population. This filtering (or attempt at filtering) is seen in many horse-race type polls, where the sample population is limited to either self-reported eligible, registered, or likely voters. His statements about the population frequency of various groups must be assumed to be among the general population, not the voting population. This may somewhat explain his high estimated swing voter percentage.
4.8 Switchers

The most significant barrier to analysis based on those who reported a change in candidate choice between the pre- and post- election samples is a numerical one. There are only 64 respondents who reported a different vote between the pre and post election waves, with 11 of those reporting a change in vote from a major party candidate to "someone else". This is about 4 percent of the total population reporting having voted post-election. These respondents are distributed across therm scores as shown in figure 4-1.

31 of these respondents are within the +/- 15 point difference score range, with 15 of them at 0.

4.9 Switchers to "Swing Voters"

Making the simplifying assumption that in aggregate swing voters have a 50 percent chance of changing their vote, the number of people actually exhibiting switching behavior would indicate that swing voters as a whole are about 8 percent of the voting
population. This maps with previous work on other definitions of swing voters. Given the probabilistic model underlying this paper, it is possible that some subset of these "switchers" are actually non-swing-voters with very high probabilities of voting for a given candidate, and that the random draw which caused them to switch their vote was a highly unusual event. Looking at the distribution of difference scores among this population, the small cluster of switchers around 100 are likely a result of survey error, and the cluster with non-100 very high difference scores and more likely the result of a low probability draw.

However, the ability to draw conclusions from this population is seriously hampered by their small size.

4.10 Proposed Solutions for Behavioral Measures

While the theoretical link between switchers, swing voters, and low difference scores is strong, the evidence provided by a survey such as the ANES is weak. This is because is a survey of this sort is necessarily loosely tied to actual behavior. Current advances in voter classification, such as tying survey results directly to voter file data, have the possibility to overcome this limitation, but are difficult to access because of cost and confidentiality concerns. However, these methods would allow for significantly more accurate testing of this survey instrument. As a first pass, being able to screen out those who actually failed to vote on election day would be extremely informative, and might help validate the intuition that disengaged groups are falsely reporting voting more than others.
Chapter 5

Predicting Low Difference Scores

Even if the behavioral measures in the ANES are less than ideal, the phenomenon of low reported differentiation between candidates remains interesting. While we cannot say from the data available that low difference scores necessarily represent swing voters, they do represent respondents who are reacting unusually to the pressures of an election cycle. This survey was taken within 2 months of the 2008 election, a period of higher than usual exposure to candidate-related materials for most citizens. The very presence of low difference scores is somewhat surprising in light of this.

5.1 Demographic Predictions

Given the prior information about variations in scale use, it seems plausible that low difference scores are more a result of variation in individual understanding of the scale than they are of underlying differences in candidate evaluation. Different demographic groups should be expected to have notably different partisanship, and as such different median scores for the Democratic and Republican candidates. It would also make sense for certain groups to be more moderate and thus feel more similarly about both candidates. These groups should then also exhibit relatively even Democratic and Republican vote share.

Breakouts of median Democratic candidate score, median difference score, and democratic vote share by demographic group are presented below.
Table 5.1:

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean Difference Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asian</td>
<td>25</td>
</tr>
<tr>
<td>Black</td>
<td>48</td>
</tr>
<tr>
<td>Hispanic</td>
<td>36</td>
</tr>
<tr>
<td>Native American</td>
<td>42</td>
</tr>
<tr>
<td>White</td>
<td>37</td>
</tr>
<tr>
<td>Male</td>
<td>36</td>
</tr>
<tr>
<td>Female</td>
<td>40</td>
</tr>
<tr>
<td>&gt;36</td>
<td>35</td>
</tr>
<tr>
<td>37-55</td>
<td>40</td>
</tr>
<tr>
<td>56-75</td>
<td>39</td>
</tr>
<tr>
<td>&lt;75</td>
<td>40</td>
</tr>
</tbody>
</table>

Clearly, demographics provide important distinctions in difference scores, and in vote share. To examine the impact of demographics on individual difference score, a generalized linear model was fit, predicting difference score as a combination of party ID, age, gender, and race.

All of these demographic variables but race are significant, and thus to some extent explain the variation in difference score. Notably, Democrats exhibited greater differentiation between candidates, which is likely related to the overall higher ratings by Democrats of the democratic candidate than by Republicans of the republican candidate.

5.2 Interest and Efficacy

As shown above, some of the difference in difference scores are the result of demographic factors. However, these things are not enough to explain the full spectrum of scores seen in the sample.

Other likely explanation for variation in degree of perceived difference between candidates is interest in the election. Previous work has suggested that disinterested voters may have a relatively undifferentiated view of the candidates, due to lack of information or disinterest. If difference scores increase dramatically with personal
Table 5.2:

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>therm</th>
</tr>
</thead>
<tbody>
<tr>
<td>age</td>
<td>0.077**</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
</tr>
<tr>
<td>is_male</td>
<td>-4.185***</td>
</tr>
<tr>
<td></td>
<td>(1.314)</td>
</tr>
<tr>
<td>is_white</td>
<td>-1.974</td>
</tr>
<tr>
<td></td>
<td>(1.610)</td>
</tr>
<tr>
<td>is_dem</td>
<td>10.518***</td>
</tr>
<tr>
<td></td>
<td>(1.426)</td>
</tr>
<tr>
<td>Constant</td>
<td>35.171***</td>
</tr>
<tr>
<td></td>
<td>(2.290)</td>
</tr>
</tbody>
</table>

Observations 1,977

Note: *p<0.1; **p<0.05; ***p<0.01
interest in the election, that change would be explainable as additional information allowing a greater distinction to be drawn between candidates, in line with the affective summary hypothesis proposed earlier.

The 2008 ANES contains a set of questions designed to test respondent’s interest in politics (and the upcoming election, specifically) as well as their perception of their own ability to impact government and politics.

The first question asks respondents how much they have thought about the upcoming election, with response options being "quite a lot" and "only a little". 75 percent of respondents report having thought about the election "quite a lot".

The second type is called efficacy questions.

These questions are divided into two variants, each asked of half the sample. The first variant asks respondents to rate how much they agree or disagree with a series of statements (below). The second asks respondents about the same basic ideas in a slightly different way, avoiding the implicit bias of the statements. Each question is scored on a 5 point scale, from "Strongly Agree" to "Strongly Disagree" or "All of the time" to "Never", depending on version.

Exact wording and sample means for these questions are presented below.

First Set:
A: Sometimes, politics and government seem so complicated that a person like me can’t really understand what’s going on.
B: I feel that I have a pretty good understanding of the important political issues facing our country.
C: Public officials don’t care much what people like me think.
D: People like me don’t have any say about what the government does

Second Set:
A: How often do politics and government seem so complicated that you can’t really understand what’s going on?
B: How well do you understand the important political issues facing our country?
C: How much do public officials care what people like you think?
D: How much can people like you affect what the government does?
For the purposes of analysis, the two sets of efficacy questions were aggregated into a single metric, representing the respondent's average score (out of 5), with component questions recoded so higher scores always indicated greater efficacy.

### 5.3 Demographics and Interest/Efficacy

It is possible that interest and efficacy are really just themselves results of demographic variation. To test this, a model was fit predicting efficacy score from the demographics used above.

![Table 5.3:](image)

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Efficacy Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
</tr>
<tr>
<td>Gender (binary, male=1)</td>
<td>0.137***</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
</tr>
<tr>
<td>Race (binary, white=1)</td>
<td>-0.061</td>
</tr>
<tr>
<td></td>
<td>(0.046)</td>
</tr>
<tr>
<td>Party (binary, democrat=1)</td>
<td>0.053</td>
</tr>
<tr>
<td></td>
<td>(0.041)</td>
</tr>
<tr>
<td>Constant</td>
<td>2.823***</td>
</tr>
<tr>
<td></td>
<td>(0.065)</td>
</tr>
</tbody>
</table>

*Observations: 2,008

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

Only gender was a significant predictor of efficacy, with being male increasing a respondent's efficacy score.

Similarly, for level of though about the election:
Table 5.4:

<table>
<thead>
<tr>
<th></th>
<th>Dependent variable:</th>
<th>Thought About Election</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.004***</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Gender</td>
<td>0.031</td>
<td>(0.020)</td>
</tr>
<tr>
<td>Race</td>
<td>0.030</td>
<td>(0.024)</td>
</tr>
<tr>
<td>Party ID</td>
<td>0.056***</td>
<td>(0.021)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.509***</td>
<td>(0.034)</td>
</tr>
</tbody>
</table>

Observations 2,008

*Note:* *p<0.1; **p<0.05; ***p<0.01
Age and party ID are both predictors of having thought about the election (as perhaps makes sense), but the coefficients on both are extremely small.

5.4 Interest, Efficacy, and Therm Scores

As is perhaps expected, those who report having thought about the election have higher difference scores than those who do not, by a significant margin.

| Table 5.5: |
|---|---|
| Thought About Election | Mean Difference Score |
| Only a little | 25.388 |
| Quite a lot | 42.422 |

Those with higher efficacy scores also have higher difference scores.

| Table 5.6: |
|---|---|
| Efficacy Score Bucket | Mean Difference Score |
| 1 | 28.051 |
| 2 | 32.073 |
| 3 | 36.991 |
| 4 | 41.910 |
| 5 | 51.121 |

The causality here cannot be proven from the given data, and is even theoretically unclear. Respondents with a greater feeling of efficacy could be driven to acquire more information, increasing candidate differentiation, or candidates who perceive a strong difference between candidates could be driven to acquire more information and increase perceived efficacy. Both of these explanations are consistent with the probabilistic voting model.
Table 5.7:

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Difference Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.021 (0.037)</td>
</tr>
<tr>
<td>Male</td>
<td>-5.161*** (1.277)</td>
</tr>
<tr>
<td>White</td>
<td>-2.156 (1.559)</td>
</tr>
<tr>
<td>Democrat</td>
<td>9.410*** (1.384)</td>
</tr>
<tr>
<td>Efficacy Score</td>
<td>4.112*** (0.832)</td>
</tr>
<tr>
<td>Thought About Election</td>
<td>13.235*** (1.574)</td>
</tr>
<tr>
<td>Constant</td>
<td>16.957*** (3.170)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,977</td>
</tr>
</tbody>
</table>

* Note: *p<0.1; **p<0.05; ***p<0.01
5.5 Interest and Efficacy as Prediction of Difference Score

Efficacy and interest in the election are both significant predictors of difference score, with very high coefficients. Thought about election is the largest factor, possibly reflecting the fact that only the least interested respondents would be willing to report having thought about the 2008 presidential election only a little, less than 2 months before election day. Being a Democrat is also a large factor in high difference score, which further makes sense given the higher interest and involvement of Democrats in the 2008 election (as shown by the democratic victory).
Chapter 6

Conclusions and Results

6.1 Emotion and Decisionmaking

An emotional theory of political information processing is well supported in the literature, if difficult to explicitly measure or prove. The on-line processing model describes the translation of political information (of indeterminate source, not explicitly differentiating between personal circumstances or more traditional information sources like advertising) into emotional reactions, which are then stored as an affective summary. Therefore, measurement strategies that ask for reasoning or particular information remembered in support of either candidate are unlikely to be detecting the true mechanism of decision making. It is thus better to rely on a measurement strategy with an emotional basis.

6.2 Probabilistic Voting

Moving away from the hard division of voters into "swing" and "nonswing", this theory instead proposes that there is something of a continuum of voters, ranging from those with the most certain voting behavior to the least certain. This internal probability results from the summaries discussed above. The probability of voting democratic (referred to as "vote probability") can be estimated on a population level by the percent of members reporting having voted for the democrat, with the as-
sumption that on a large-N level that percentage will reflect the average probability of each member voting democratic.

Under the probabilistic model, swing voters are those who have vote probabilities in the range close to 50. Persuasion strategies seek to shift vote probability, but will have a small impact at best (consistent with the overall stability of vote choice in the face of overwhelming quantities of advertising). These people are an extremely small group, as evidenced by the small number of individuals who report changing their vote decision before an election or between elections. Their targeting is important for election campaigns and candidates seeking to alter vote share in their region, as well as interesting theoretically in contrast with the typical stability of vote choice.

### 6.3 ANES Thermometer

The ANES feeling thermometer is hypothesized to be recording affect towards political candidates, the expression of the summary formed by on-line processing. Using a mechanism laid out in William Mayer’s previous work on swing voters, creating a difference between candidates scores from the two feeling thermometers, voters can be divided into categories expressing different degrees of differentiation between the candidates. Those with low differentiation exhibit distinct characteristics in response to other questions on the survey. They have a more nearly equal probability of reporting belonging to either political party, as well as reporting having voted for either candidate. This is the behavior that would be expected from a group with nearly even vote probabilities, providing some backing to the use of the thermometer scale as a proxy for vote probability and affective summary.

There is significant error in the thermometer scale use, with a large number of respondents reporting no difference between the candidates, despite a clear candidate preference expressed in other questions. This makes it more difficult to infer results from the scale, since it seems reasonable that a larger portion of these false-zeros would be respondents with low interest in the election.

The other significant issue is the lack of verifiability of the thermometer scale in-
struments against other sources of data. These instruments are extremely rarely used outside the ANES, and then typically only in studies making explicit use of ANES measures. Thus, the conclusions about scale use are not generalizable, and may be very different in another (shorter) survey. Similar testing could be performed on a modified thermometer instrument, one framed more like a Likert scale, which would likely give a similar degree of accuracy. This structure would be more compatible with modern surveys and be less difficult to field.

6.4 Behavioral Measures

Perhaps the clearest result from this study is a reinforcement of the difficulty of predicting behavioral results from a survey instrument not closely tied to behavioral measures. While the ANES is extremely useful given its large sample size and accessibility, it is by design entirely self-report. The behavior of most interest to this study, voting, is impossible to cross-check against actual voter history on the given data, which compromises its prediction.

The mechanism of reported vote switching had to be entirely excluded from this analysis given the very low sample size and clear presence of survey error. Voters who functionally change their mind about which candidate to vote for between two months before election day and election day itself are clearly and intuitively a small portion of the population, and thus difficult to produce meaningful results for in a non-targeted study.

6.5 Exploring the Feeling Thermometer Results

Given the difficulties in tying survey responses to behavior, this aper instead chooses to explore in-survey causes of variation in reported differentiation between candidates. This metric is, by necessity, one step away from voting behavior. However, proving
that factors such as party ID and interest in the election, which have been traditionally considered to have a strong impact on vote choice, have an impact on the difference metric provides some backing to the theory that the difference metric is comparable to other predictions of vote choice.

6.6 Trends in Reported Differentiation

A large portion of the surveyed population reports seeing little difference between the two candidates in terms of emotional affinity. While some of this lack of differentiation can be explained as survey error (such as the high number of zeros clustered around 50/50 ratings), some of it reflects real indifference between candidates. Certain demographic groups are more likely to exhibit low differentiation, in patterns that reflect common intuition about demographic groups more involved in the political process.

However, measurements of interest and efficacy provide an improvement in difference score prediction over demographics alone. Thus, it is likely that interest and efficacy factors are determining in part the extent to which a respondent has an affective summary for each candidate. Respondents with low efficacy may be sufficiently disengaged from the political process to have not built up a difference in feeling between the two candidates.

6.7 Implications

The link between the affective summary model and the probabilistic model of voting behavior shown in this paper is an important departure from other models of voting behavior. It accounts for the low level of interest and information that most voters seem to have, as well as accounting for their inability to recall the formation of their preferences. Overall, the model proposed in this paper contains a high degree of uncertainty. This is appropriate for an attempt to model individual preferences, which are often vague. While frustrating from a statistical angle, this uncertainty
allows for a more accurate detection of voter decision making. The emotional storing mechanism in particular implies a unique measurement strategy, one designed to pick up on emotional reaction instead of asking a voter to offer a definite opinion. The ANES feeling thermometer uses this sort of strategy, and is perhaps worthy of more research as a predictive tool.
Bibliography


[12] Pamela Johnston Conover and Stanley Feldman. Emotional reactions to the economy: I'm mad as hell and i'm not going to take it anymore. 30(1):50–78.


[59] Nicholas Winter and Adam Berinsky. What’s your temperature? thermometer ratings and political analysis.
