Chunky or Smooth?

New Methods for Analyzing Public Opinion and Information Flow

Abe Gong
University of Michigan
agong@umich.edu
March 2009

*** Draft prepared for presentation at the Annual Conference of the Midwest Political Science Association, 2-5 April 2009. Please do not reference or cite without express written consent from the author.***

Abstract
Where does the information that changes opinion come from? One conjecture states that public opinion changes through a top-down process of mass-mediated "elite discourse." A competing conjecture is that public opinion is a bottom-up process of individual "thermostatic" observation. Using statistical techniques from natural language processing and data from the 2000 and 2004 Annenberg Election Study, I examine whether changes in opinion occur in sudden chunks, consistent with the top-down claim, or smoothly over time, consistent with the bottom-up claim. The result: opinion chunkiness depends on the issue. For example, assessments of challengers seem to be top-down, but assessments of incumbents are bottom-up. These results have important implications for theories of aggregate opinion change, electoral learning, and societal information flow.

1. Introduction

Virtually every theory of public opinion formation is built on the premise that an opinion is a function of an individual's internal "predispositions" and external "information." Much has been written about the values, cues, models, group memberships, emotions, incentives, and genetics[1] that determine an individual's response to information. However, the nature of the "information flows" surrounding and connecting individuals has been treated much less thoroughly.

This imbalance has been driven largely by methodological constraints: individuals can be caught and interrogated; "flows" are abstract and almost invisible (Hovland, 1959). Nevertheless, the information context that makes the public the public is an important part of any theory of public opinion (Blumer, 1948). This paper demonstrates how empirical techniques from time series analysis and natural language processing can be used to approach the information half of the opinion equation.

The paper proceeds as follows. In the next section, I contrast two important theories of information flow. One posits a top-down process of mass-mediated "elite discourse." The other argues that public opinion is a bottom-up process of individual "thermostatic" observation. In section three, I argue that we can adjudicate these claims using time series of public opinion data. A top-down world is consistent with "chunky" changes in public opinion, and a bottom-up world would be characterized by smooth changes.

Section four describes a new statistical process -- I call it the "peanut butter test" -- for differentiating between chunky and smooth time series. In section five, I present results from application of the peanut butter test to 75 rolling cross sections from the 2000 and 2004 Annenberg election study. I find important differences in the "chunkiness" of opinion change: information about presidential challengers seems to flow top-down, but information about incumbent presidents flows bottom up; national questions are evaluated top-down, but personal questions are evaluated bottom-up. Section six discusses patterns in these results and directions for further research.

2. Literature Review

In this section I define information flow more precisely, explain how most existing theories of information flow can be classified broadly as "top-down" or "bottom-up," and argue that thinking about information flow in these terms can move us from open-ended conjecture to testable empirical work.

In *The Nature and Origins of Mass Opinion* (1992), Zaller outlines a framework in which "opinion is the marriage of information and predispositions." In this formulation, predispositions are personal and unchanging, but information is shared and dynamic. Building on this theoretical construct, I define information flow as *the process by which individuals in a given society come to share opinions*.

---

1. Examples abound. See Luker (1984), Lupia (1994), Zaller and Feldman (1992), Green et. al. (2002), Brader (2006), Downs (1957), and Alford, Funk and Hibbing (2005) for my favorites.

In democratic theory, we are particularly concerned with information flow between professional political actors ("elites") and the electorate at large ("the mass"), because claims about transparency, responsiveness, accountability, and legitimacy all depend directly on level of coincidence between elite opinion and mass opinion (Key, 1968)(Arnold, 2004). Many observers (e.g. Plato, Lippman, 1922) have been concerned that political elites may be able to exert undue influence over the mass public by controlling the flow of information. Theories about the exact nature of elite control vary, but it seems reasonable to gather them together under the heading of "top-down" theories of information flow (e.g. Parenti, 1992; Bennett et. al., 2008). These theories argue, essentially, that political elites are able to tell the mass what to think. These stand in opposition to "bottom-up" theories of information flow, which contend that the mass public is able to observe and assimilate new information independent of elites.

Within the field of modern public opinion research, Zaller's theory of "elite discourse" is an excellent example of a top-down model. In this model, all politically relevant information originates with elites and percolates down through a passive mass public. If elites disagree among themselves, the mass may polarize, but they will never do so on their own. The limitating variable in Zaller's theory is cognitive capacity -- the mass public lacks the initiative and deep understanding necessary to form meaningful opinions on their own. Although members of the mass have access to many of the same facts as political elites, they are not capable of assembling those facts into a coherent basis for personal opinions.

Other top-down theories argue that gate keeping (Gans, 1980), media bias (Patterson, 1994), etc. deny members of the mass public the raw information they would need to form opinions. In either case, citizens' lack of information makes them subject to persuasion by political elites. Generally speaking, mass media are assumed to play an important role in conveying information from the elite to the public. Research in top-down traditions often has a pessimistic quality: voters are portrayed as controlled by political elites through their televisions[2].

In contrast, citizens in bottom-up theories are usually portrayed as more capable and self-directed. Perhaps the best example of this theory in current opinion research is Stimson's “thermostatic democracy” hypothesis (2004, see also Erikson, Mackuen, and Stimson, 2002; Page and Shapiro, 1992). Stimson argues that individuals observe and react directly to the outcomes of policy. Like homeowners regulating room temperatures via thermostat, citizens do not need to exercise great foresight or posses detailed knowledge of the mechanisms of government. Instead, they notice the effects of policy (e.g. less pollution, higher crime) and vote for policies to achieve more or less of these things. Stimson’s evidence that approval of government is largely driven by economic performance and his somewhat weaker evidence of a liberal/conservative cycle in national opinion both suggest that elites do not control all relevant political information.

In a top-down society, mass opinion is based entirely on facts or interpretations supplied by elites. In bottom-up society, mass opinion is based on personal observation, preferences, and

2. Lupia and McCubbins's (1998) work is a notable exception. They construct a model based on top-down assumptions, and much of on cues explores conditions and limits under which top-down systems lead to socially suboptimal outcomes.

cognition. Which world better describes the one we live in? In principle, this should be an empirical question, but direct empirical tests have been lacking. In the next section, I outline a research strategy for determining the extent to which public opinion is influenced by elite manipulation of information. It is not a perfect approach, but it does provide new perspective and evidence for thinking about information flow.

3. Research Questions

As defined above, top-down and bottom-up information flow are mutually exclusive. We can imagine situations that split the difference between the two, but not situations in which both are true simultaneously. A natural next question is which world we actually live in -- and how we would know. Most of the straightforward approaches to this problem (e.g. continuous surveillance and brain-scans of a representative panel of likely voters) make unreasonable demands on data collection. Drawing conclusions about information flow from existing data requires a little more finese.

One approach is to look at the the timing of aggregate opinion change. Consider two scenarios. In the first, a nation with top-down information flow is involved in a bloody, losing war overseas. Because citizens only change their opinions in response to elite discourse, the bulk of the electorate doesn't realize the extent of the carnage until a one day when a prominent politician makes a widely publicized speech. In a persuasive address, she denounces the war, outlines its costs, and argues for a change in policy. Many citizens change their minds as a result of this speech. Importantly, they change their minds *at the same time*. If we were to field surveys the day before and the day after the speech, we would see substantial swings in attitudes about the war. The speech persuades a large chunk of the electorate, which shows up as a discontinuity in the data. Other later media events might lead to similar persuasive effects in one direction or the other.

In the second secnario, imagine the same country in the same war, but with a bottom-up information process. Here, citizens experience the costs of the war directly -- they attend funerals for neighbors and relatives, debate wartime suppression of civil liberties, and watch as social programs are cut to fund the war -- and their assesments of government policy are updated thermostatically to reflect those individual experiences. Exactly the same information percolates through the electorate, with exactly the same persuasive effect. But in this scenario, changes in opinion are distributed across time. Instead of chunks of opinion change, we would expect to see smooth changes in aggregate opinion.

Figures 1 and 2 illustrate idealized chunky and smooth time series.

Fig 1. An idealized chunky time series

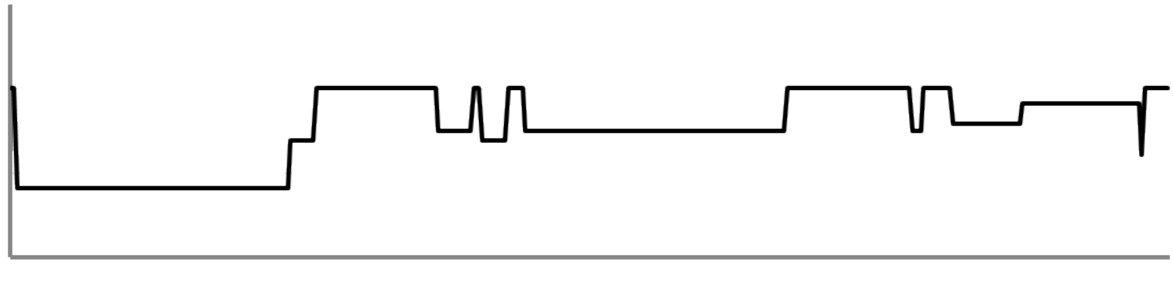
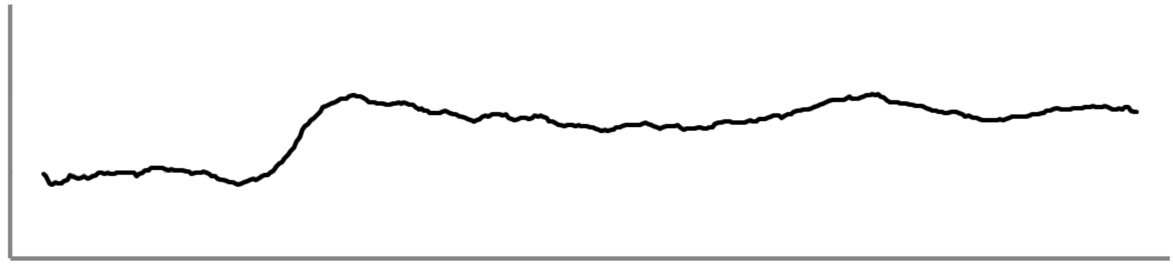


Fig 2. An idealized smooth time series



As suggested by these two scenarios, chunky series of opinion data seem broadly consistent with a top-down information process, and smooth series seem in line with bottom-up information flow. The correspondence is not perfectly one-to-one[3]. But all else equal, high chunkiness[4] would seem to be a good indicator of top-down flow.

In the next section, I describe a statistical procedure for identifying the level of chunkiness in a given time series. For reasons I will discuss, measuring the *absolute* level of chunkiness in a time series is difficult. Therefore, we cannot immediately answer the question, "Do we live in a top-down or bottom-up society?" However, we can measure the *relative* chunkiness of multiple time series. This enables us to makes comparisons among series, implicitly asking the question, "Are different types of opinion informed by different information processes?"

Specifically, I ask:

- · Are assessments of presidential challengers chunkier than assessments of sitting presidents?
- · Are opinions about personal issues chunkier than opinions about national issues?
- · Are personal decisions about political media consumption chunkier than participation in discussion about politics?

These questions are intended as a suggestive first cut of the data. Accordingly, I have chosen topics that are best suited to the properties of the test and the availability of data.

4. Methodology

In this section I describe a statistical procedure for identifying chunky versus smooth time series. I call it the peanut butter test, for obvious reasons. First I provide a non-technical overview of the test, followed by a description of the data I use to evaluate my research questions. These subsections are marked with asterisks because understanding them is essential to understanding the remainder of the paper. Because this method is new, I follow

---

3. From a logical perspective, a chunky persuasion process is a sufficient condition to infer top-down information flow. If many people change their minds simultaneously, they must have been simultaneously exposed to similar persuasive information. The reverse is not necessarily true. For example, if many elites make many persuasive speeches in a top-down system, the cumulative effects might still appear smooth instead of chunky.
4. I use the term chunkiness with apologies to Kleinberg, who coined the term "burstiness" to describe a similar phenomenon (see Kleinberg 2003). The conceptual underpinnings of our statistics are similar -- chunkiness was largely inspired by burstiness -- but the contexts, procedures, and interpretations are different enough to merit different terminology.

up with four additional subsections to provide additional technical details about the peanut butter test and its limitations.

** 4.1 Overview*

Although some of the technical details are quite involved, the overall strategy for the peanut butter test is straightforward. We imagine two extreme cases of the world: completely chunky or completely smooth. Accordingly, for any given time series, we can fit the data using a statistical model that assumes complete chunkiness, and a statistical model that assumes total smoothness. I use hidden Markov models and rolling averages to model chunky and smooth dynamics, respectively. Applying these two procedures to data yields two sequences of best fit.

It is easy to measure how well each sequence fits the data, then compare the two fit statistics. In a chunky (smooth) world, the sequence created using the chunky (smooth) model should fit the data better. By subtracting the fitness of the smooth sequence from the fitness of the chunky sequence, we arrive at a single value. This is the peanut butter score. It is lower for chunkier time series and higher for smoother time series. Additionally, peanut butter scores are comparable across time series of different lengths, samples, and topics.

** 4.2 Data*

Data for this analysis are drawn from the 2000 and 2004 National Annenberg Election Studies (NAES). These studies are almost indisputably the best publicly available rolling cross-sections of public opinion data. Each study was in the field daily for almost an entire year data leading up to the elections. Depending on the question, average daily sample sizes range from several dozen to the low triple digits, with cumulative samples on the order of 50,000 to 100,000 interviews. Moreover, the study is conducted using random digit dialing so as to be (roughly) nationally representative. Note that rolling cross-sections are not panels. Respondents in each time step are drawn independently from the population.

The Annenberg studies ask hundreds of questions, each of which could be analyzed in many different ways. For technical and practical reasons, I restricted my analysis to survey questions meeting three criteria. First, I excluded questions with categorical or open-ended answers (e.g. "What do you think is the most important issue facing the country today?"), and kept only questions with ordered answers (e.g. feeling thermometers about candidates, days of news consumption in the past week). The one exception to this rule was a handful of questions about political parties (e.g. party ID, intended vote choice for house elections). For these, I made the variable dichotomous by dropping everything except the major parties.

My second criteria for excluding questions was the length of the series. Largely for convenience, I limited my pool of time series to questions that were asked throughout the whole study. Longer series length and higher sample sizes make these questions the best place to validate the peanut butter test.

Finally, I excluded most of the demographic questions on the grounds that they don't vary and are therefore not useful for studying the dynamics of information flow.

These three cuts leave 77 time series, corresponding to 26 questions asked in the 2000 NAES and 51 questions asked in the 2004 NAES. These 77 question-series are the data points in my final analysis. Question wordings for these series are given in appendix A.

In addition to computing a peanut butter score for each question-series, I classified questions based on three categories corresponding to my research questions. First, the 2004 NAES includes a battery of questions about the attributes of the two candidates ("How much does the phrase 'cares about me'/'trustworthy'/'reckless'/etc. apply to George Bush/John Kerry?"). Assuming that the contest between Bush and Kerry was representative of typical challenger-incumbent dynamics, this set of questions (coded 1 for questions about Bush and 0 for questions about Kerry) offers a very clear opportunity for comparison. Exactly the same questions were asked about both candidates and the order of questioning was randomized, so the only difference between the questions was the names of the candidates.

Second, I separated questions into a three-way categorization: questions phrased completely in terms of personal issues ("When you talked with family or friends [about politics in the last week] how much disagreement was there?") , questions phrased entirely in terms of national issues ("How would you rate economic conditions in this country today?"), and questions that involve some of both ("How well does 'shares my values' apply to Georg W. Bush?").

Finally, in order to capitalize on the many media use questions in the study, I split these questions into three categories: questions about news media consumption and questions about discussing politics.

Table 1. Opinion stream coding

Classification
Challenger/Incumbent
Personal / National issues / Mixed
News Consumption / Political Discussion

These three categorizations are the independent variables in my research questions. Peanut butter scores are the dependent variables. Testing for differences among the various categorizations is just a matter of applying two-sample t-tests. Descriptive statistics, peanut butter scores, and codes for all three categorizations are given in appendix B.

4.3 Notation

Here we plunge into the technical details of the peanut butter test, starting with notation. All of the later analysis begins with times series of aggregated responses to a dichotomous question, plus a corresponding series of poll sample sizes. When working with data, I will refer to the first series as proportion vector, P, the second series as the sample size vector, N, and both series together as a *question-series* or *opinion stream*.

Both P and N are of length l, where l is the number of time steps in the stream. Elements of P and N will be referenced as p_t and n_t. Note that on any given day, p_t is on the interval [0,1] and n_t is a natural number. Note also that opinion streams are the product of two levels of aggregation: first, survey responses on a given day, t, are aggregated up to generate p_t and n_t; second, values of p_t and n_t are aggregated across some time span, $t \varepsilon [a, b]$, to form opinion streams.

In order to fit the NAES data to these specification, I collapsed each question-series using a procedure similar to a median split. For each question, I collapsed answers so that everything below the center point of the scale was coded 0, and everything at the center point or above was coded 1. The one exception to this rule was for questions about news media. So few respondents report following politics in the news that I coded any media consumption as a 1, and no media use as a 0. For all questions, responses coded as "don't know," missing, and the like were dropped. Practically speaking, n_t is the count of 1's and 0's on day t, and p_t is the number of 1's divided by n_t.

For most questions in public opinion, we would want to know whether p_t is a valid measure of some opinion on day t. In this case, problems of sampling bias, nonresponse, and biased "doorstop opinions" would all have a strong tendency to corrupt the data. In this paper, we are not interested in accurately estimating public opinion or correlates of opinion -- we are interested in patterns of change in public opinion. Therefore, as long as the group contacted for interviews is representative of some population across time, we can safely infer that P is a good estimate of some set of opinions. Consequently, the various sources of bias are a much smaller threat to the internal validity of my results than they would be in a typical public opinion study. The possibility of external validity remains -- perhaps my results don't generalize to the population at large, or to opinions that actually influence voting -- but this is essentially a pilot study, so I am willing to leave those problems for later.

Underlying any opinion stream is a *true sequence of aggregate opinion*, Q. Conceptually, Q is the set of values we would observe if we interviewed "rolling census" of the population, rather than a rolling cross-section. Because of the statistical properties of random sampling, p_t is an unbiased estimate of q_t on any given day. However, because of sampling error $p_t \neq q_t$. The distribution of errors is determined by n_t. We wish to find maximally likely sequences for Q under two sets of assumptions: perfect chunkiness and perfect smoothness.

4.4 Hidden Markov Models

In order to estimate Q^C, the optimal sequence under chunky assumptions, I use a class of statistical models known as hidden Markov models (HMM). Since HMMs are largely (but not completely) unknown in political science, some description is in order.

Hidden Markov models are often employed in bioinformatics, voice recognition, and operations research (See Rabiner, 1990 for a good introduction). They are designed to explain the relationship between an observable *signal* (e.g. survey responses over time) and an underlying set of *states* (a range of percentages over which opinion may vary). Given a signal and some knowledge about the probabilities of *state transitions* (changes in aggregate opinion) and *signal generation* (sampling errors) by states, HMM-based algorithms can identify the maximally likely set of underlying states (actual public opinion over time). Complementary methods can estimate properties of state transitions (how often and how much opinion changes) for data sets of sufficient size.

In contrast to most conventional statistical approaches, HMMs can capture the effects of large, discrete events instead of just the accumulation of small changes. Most time series models, including the ARIMA specification and all of its subclasses (e.g. random walks or moving averages), are incremental. In other words, they assume that change occurs in a steady trickle of additive steps, perhaps with some trend or drift. HMMs are well suited to tackling systems characterized by long periods of stability and punctuated by occasional change -- exactly the kind of chunky outcome suggested by top-down models of information flow.

Any HMM can be completely characterized with four formal components. First, we need a set of discrete *states*. In this case, states are levels of true aggregate opinion -- the values that q_t can take. Strictly speaking, q_t should be a continuous variable on the interval [0,1]. We can approximate this using a lattice of k states covering the same interval. As k increases, errors from rounding decrease to zero. However, larger values of k come with a computational cost. Consequently, there is a tension between choosing small values of k for faster computation and larger values of k for greater precision.

The choice of values for the k states can alleviate this tension somewhat. Instead of using an evenly-spaced lattice of states, I broke the observed values of P into k quantiles and used the boundaries of those quantiles as my states. This procedure follows the intuition that the distribution of observed P should reflect the distribution of actual Q. Moreover, my tests suggest that it usually leads to better estimates in practice.

The second choice component of a hidden Markov model is the *transition matrix*. This k by k Markov matrix defines the probability of a change from one state to another in any given time period. Entries down the main diagonal define the probabilities for staying in each state. Rows must sum to one.

Where theory dictates, specifications of transition matrices can be very complicated. In this case, we have little theoretical grounding to encode in the matrix, so I invoke two simplifying assumptions. Assume that the probability that important news will become available on any given day is fixed. Call that probability B. This is the probability of a state transition in a given round. Conversely, the probability of no transition is $1-B$. Consequently, the values down the main diagonal of the transition matrix are all $1-B$. Assume further that an important story can change any number of minds in the electorate so the probability of moving to any given state is equal. Therefore, all of the off-diagonal entries in the matrix have value $B/(k-1)$. This fully specifies the transition matrix.

The third component of an HMM is an *emission probability function* that maps states onto signals. We know that our observed signals are noisy observations of the underlying states. In order to make inferences about the true sequence of states, we must know the probability of observing any given signal from any given state. In the context of this paper, that means that we must know the probability of observing p_t, given that the true state of the world is q_t.

Fortunately, sampling theory works very nicely here. The usual binomial statistics used to calculate margins of error allow us to calculate the probability of observing p_t, given a sample size of n_t and a true state of q_t. Note that although the Markov model requires us to use a discrete set of states to describe q_t, we can describe p_t in continuous space. I use the normal approximation of the binomial distribution. Since almost all of the daily samples employed here have at least a few dozen respondents and tend to be concentrated in the middle ranges of probability (.3 ~ .7), the approximation is quite close. Moreover, I have found that the normal approximation is preferable because it can be computed faster and does not contain gaps or discontinuities.

The final (and arguably least important) component of a Markov model specification is a stochastic vector of *initial state probabilities*. This vector describes our prior expectations about which state the sequence started in. In long sequences, this component can affect estimates of the first few states, but usually has little impact on the sequence as a whole. I use sampling error calculations from the first period to generate this vector. Another reasonable candidate would be a uniform distribution over all states.

To recap so far: by defining four components, we can specify a hidden Markov model and estimate a sequence of best fit under chunky statistical assumptions. I have outlined procedures for defining all four components, based only on available data (P and N), and two parameters (k and B). To specify the Markov models, I used the values $k = 15$ and $B = .14$. In other words, given an opinion stream, k and B are sufficient to fully specify the model.

With an HMM defined, we can bring to bear a well-developed set of analytical machinery. Of particular interest is a recursive procedure known as the Viterbi algorithm. The Viterbi algorithm uses a set of data (e.g. P and N) and the four components of a hidden Markov model (e.g. as defined by k and B), and finds the maximally likely set of underlying states (e.g. our goal: Q^C). In principle, this problem is a search over kl possible state sequences, where k is the number of possible states and l is the number of time steps in the sequence. For nontrivial state sequences, a brute force search of this potential sequence space would take an astronomical amount of computing power. Fortunately, the Viterbi algorithm allows us to exploits the Markov property of transitions to compute the maximally likely state sequence in $O(l\,k^2)$ time. This means that the computational demands of the algorithm scale linearly with the length of the sequence, and quadratically with the number of states.

For a full description of the Viterbi algorithm, I refer readers again to Rabiner's (1990) excellent tutorial paper. For now, suffice to say that it allows us to realize our original goal: solving for the optimal state sequence, Q^C, for a given opinion stream (P, N) and parameters k and B.

4.5 Rolling Averages

Compared to finding the optimal state sequence under chunky assumptions, finding an optimal sequence, Q^S, under smooth assumptions is simple. Borrowing from a long tradition among pollsters, I use a rolling average as a smoothing function. In this case, I use a window of plus or minus three days around each observation. Within that window, all survey responses are weighted equally. So if 400 of 1,000 respondents answered a given question positively between time t-3 and t+3, then I estimate q_t at .4, regardless of how the responses were distributed within that period. Using our notation to say the same thing, $q_t = \sum_{i \in t-3,t+3} p_i n_i / \sum_{i \in t-3,t+3} n_i$. To accommodate the window, I drop the first and last three observations in each series.

In many ways, rolling averages are the default for the way we think about trends in time series data. Before accepting the estimates they make as the "right" numbers, it is important to understand the assumptions and limitations behind the method. At a conceptual level, rolling averages assume that opinion has a specific kind of continuity across time. Effectively, this lets us inflate our daily sample sizes by borrowing respondents from the past and/or future.

Mathematically, the general case of rolling averages can be written as $q_t = \Sigma\, p_i\, f(t\text{-}i) n_i \,/\, \Sigma\, f(t\text{-}i) n_i$, with the summation carried over all available data. $f(t\text{-}i)$ is a weighting function describing the relationship between data on different days. (The 3-day window can be expressed as a step function with $f(x) = 1$ for $-3 \leq x \leq 3$, and $f(x) = 0$ for all other values of x.) Generally speaking, we assume that the recent past is more closely related to today than the distant past -- in other words, $f(x)$ declines as x moves away from zero. However, specifying the right relationship between past and present is difficult because it depends fundamentally on the process driving aggregate opinion change.

In particular, if discrete events sometimes cause chunky opinion change, then any rolling average will mis-estimate the sequence of true opinions close to the time of the event[5]. Rolling averages implicitly assume a smooth process for opinion change -- they are the wrong model for estimating a chunky world. Since we have ample reason to believe that many important aspects of our world are driven by discrete events, we should be careful before blindly applying rolling averages to data.

4.6 Peanut Butter Scores

Using the procedures described above, I estimated two sequences, Q^C and Q^S, for each of the 77 NAES opinion streams. Examples of optimal sequences are given in Figures 3 and 4. These sequences are the optimal chunky and smooth sequences for the question-series "How well would you rate John Kerry?"

Fig 3. HMM optimal state sequence for Kerry favorability

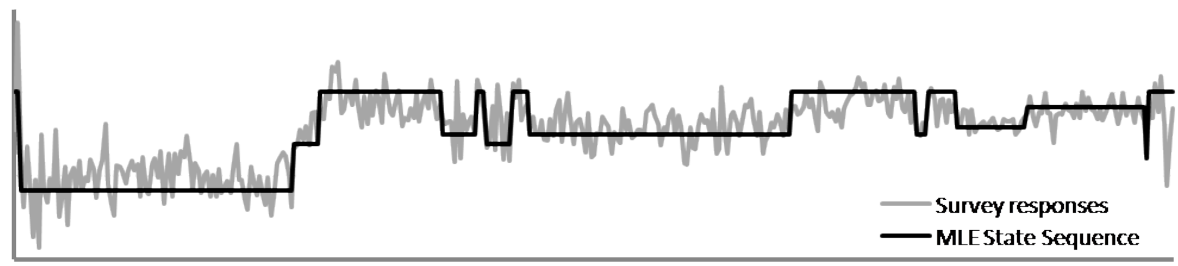
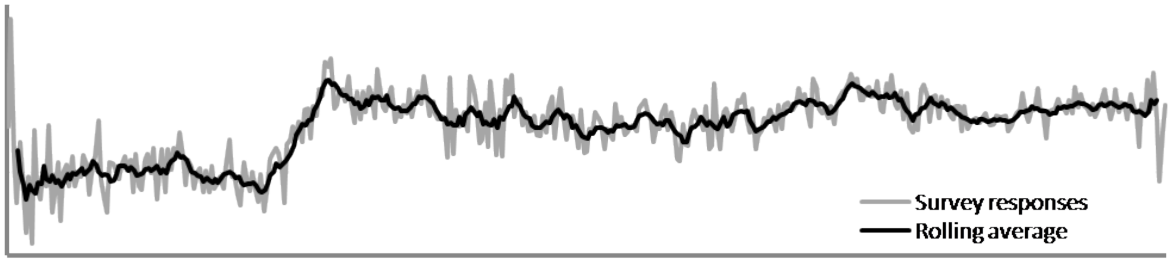


Fig 4. Rolling average optimal state sequence for Kerry favorability

---

5. The sole exception is the family of trivial rolling averages in which f(x) = c if x = 0 and 0 otherwise. But this is just the original day-by-day estimate, not an average.



Our final task is to compare the fitness of these sequences. This is a straightforward computation along the lines of the likelihood ratios test for nested models and the Bayesian information criterion. For both Q^C and Q^S, we compute the log likelihood of observing P, yielding LL^C and LL^S. These are measures of goodness of fit, with a large negative values indicating relatively poor fit and smaller negative values indicating relatively good fit. We then subtract LL^C from LL^S, giving us a single peanut butter score for the series. A high score indicates a smooth time series and a low score indicates a chunky series.

4.7 Limitations

As it stands, peanut butter scores are *ordinal* measures of chunkiness. Consequently, we can use them to make comparisons among time series, but we cannot make any inferences about the absolute chunkiness of any given set of series. Ideally, we would like to create a *cardinal* measure of chunkiness. However, three problems make this difficult.

First, HMMs and rolling averages both require us to specify some parameters before generating maximally likely sequences. For HMMs, we must specify k and B. For rolling averages, we must choose the size of the averaging window, or more generally, $f(x)$. In each case, the choice of parameters is somewhat arbitrary. It is easy to choose parameters to fit the data better. In general, models generated with higher k, higher B, or a narrower averaging window will all yield better fit statistics. But we have no clear criteria for determining the "right" parameters. Consequently, our choice of parameters probably favors one model over the other, but we can't assess which.

There is a philosophical problem here as well. Suppose we live in a world that is entirely chunky. That implies that the model based on rolling averages is wrong. In that case, it is hard to defend a process in which we must pick the right parameters for a model that we know is wrong.

Finally, computational constraints make it impractical to arrive at a truly optimal estimate of the chunky state sequence. As discussed earlier, we would like to choose a large value of k in order to reduce "rounding" error in our estimates. In principle, a Markov model with a large number of states ($\lim k \to \infty$) can approximate the optimal state sequence to arbitrary precision. But as the number of states increases, the amount of computing time required to execute the Viterbi algorithm increases quadratically. If it takes a few minutes to analyze the data when $k = 7$ and a few hours when $k = 15$, it could take days to conduct analysis for $k = 50$.

Despite these limitations, the peanut butter test provides some useful insights into public opinion and information flow. In the next section, I present descriptive statistics and results from my hypothesis tests.

5. Results

The methodology described above condenses hundreds of thousands of survey responses down to a small dataset of 77 question-streams. All of the following analysis is based on the properties of those streams. I begin with descriptive statistics and robustness checks for the peanut butter test. I cover substantive results in the second half of the section.

Descriptive statistics

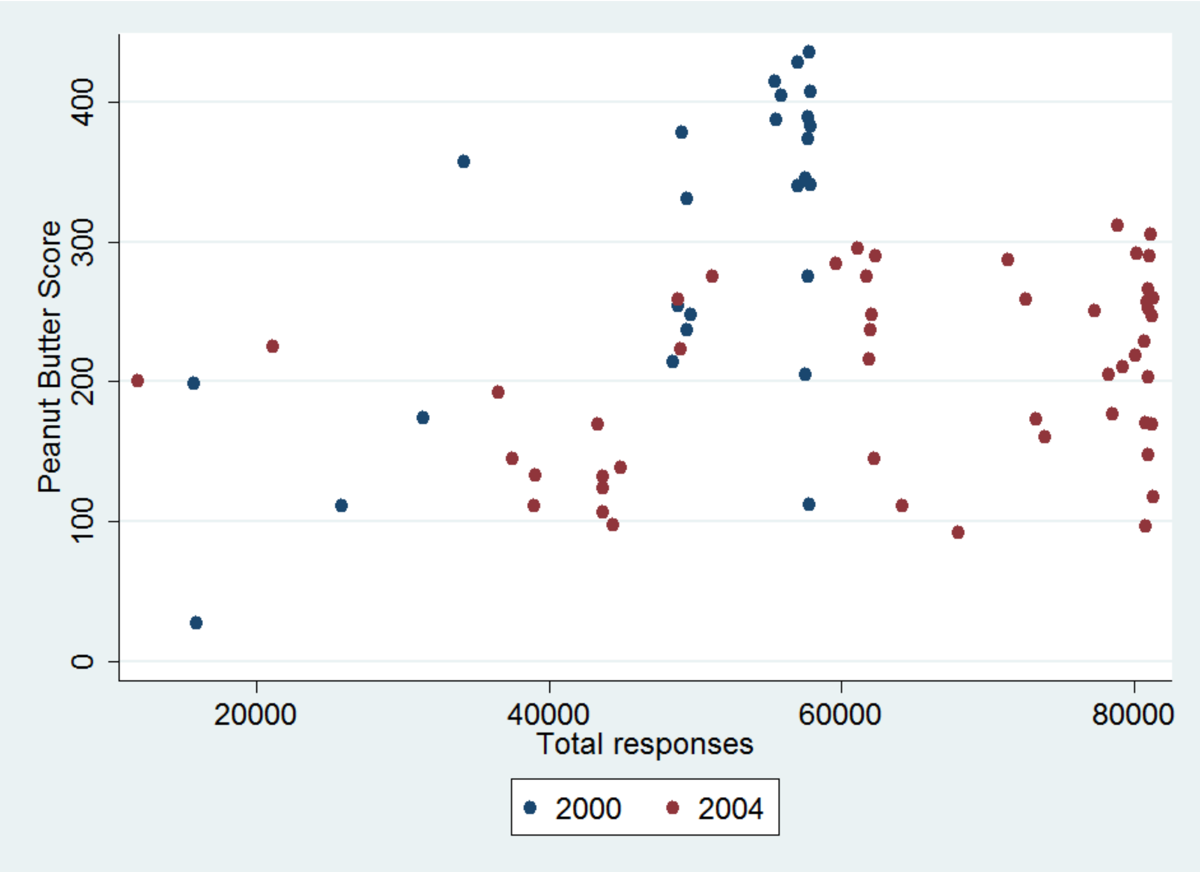
Peanut butter scores had mean 237.67, with a standard deviation of 93.97. Recall that low (high) scores corresponds to chunkier (smooth) opinion streams. The lowest score was 27.29 for the question "When you talked [about politics] at work or on the internet [in the last week], how much disagreement was there?", and the highest score was 435.31 on the question "How many days in the past week did you listen to [talk radio]?". As discussed previously, the relative ordering of the scores is meaningful, but the individual values are not. These numbers are intended simply to help readers get acquainted with the measure. Figure 5 shows the full distribution of peanut butter scores.

Fig 5. Histogram of peanut butter scores

One initial worry was that peanut butter scores would be biased based on sample sizes or sequence lengths. The test was constructed to be agnostic to these factors, but I was not certain how they would play out in practice. As it turned out, peanut butter scores were not correlated with total responses (the total number of respondents who answered a given question at any time) or sequence length. The test exhibited some heteroskedasticity with respect to sequence length -- shorter sequences had higher variance[6].

These results are illustrated in Figure 6. Blue points are taken from 2000, where all sequences were 390 days long. Red points are taken from 2004, where sequences had length 399. Overall, all of these results strongly suggest that the peanut butter test performed as intended, without bias from its major parameters.

Fig 6. Robustness checks for peanut butter scores



One side product of the peanut butter test is an estimate of the total change in an opinion stream over the time period of interest. That is, by summing up the day-to-day changes in either (or both) of the optimal sequences, we can arrive at an estimate of the total number of people who changed their minds over the course of the times series. This estimate has several limitations: it is almost certainly biased upwards, making it useful only as an ordinal measure; it doesn't tell us which parts of the population are changing their minds; the potential for bias with respect to sample size remains.

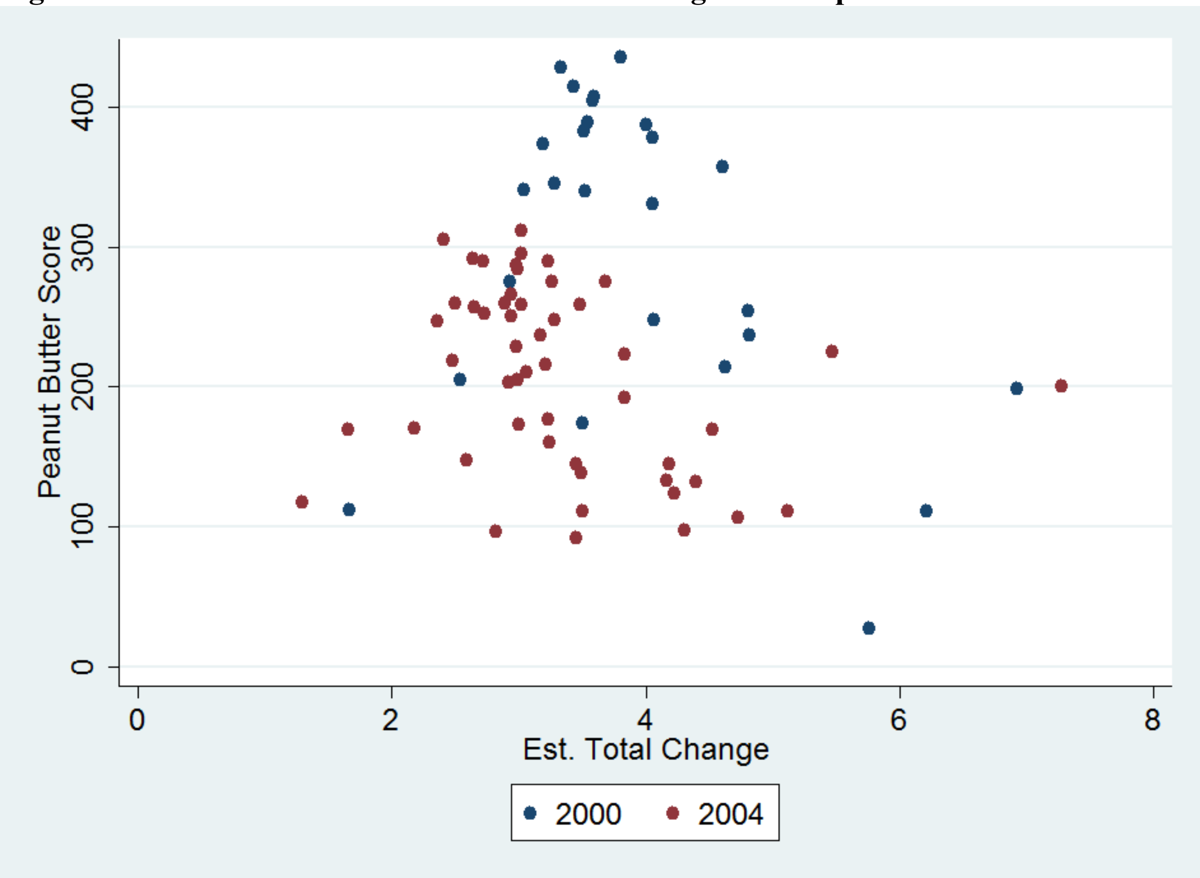
With these limitations in mind, it is still interesting to have a sense of which opinion streams were subject to the most change over time. Figure 7 plots peanut butter scores and estimated

---

6. It would be useful to know the functional form of this heteroskedasticity. This was the first of several properties of the peanut butter test I encountered that call for additional statistical work.

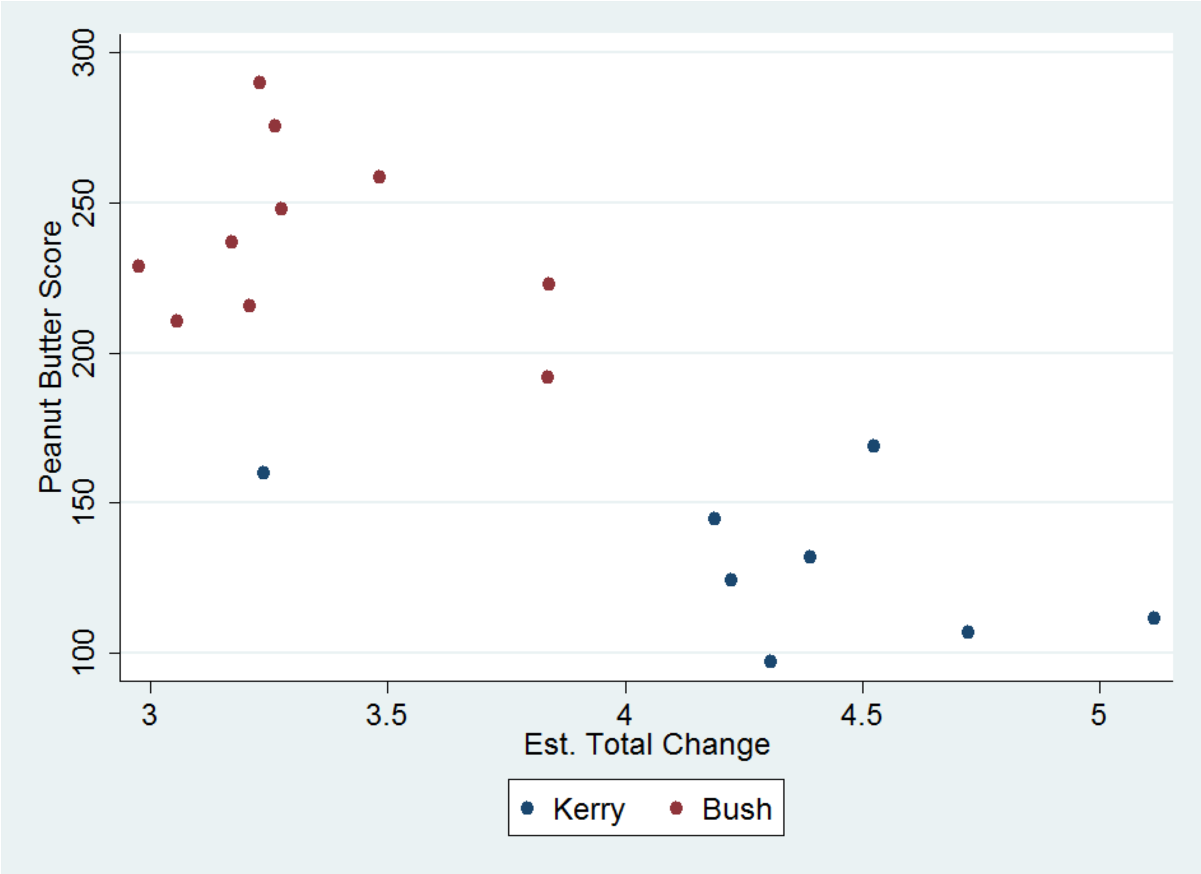
total change for all streams. Note that the two are not strongly correlated -- the chunkiness of opinion change has little to do with the total volume of opinion change.

Fig 7. Peanut butter scores and estimate total change for all opinion streams



Hypothesis tests
With these basic diagnostics in hand, I turn to direct tests of my original research questions. For the first test, I compared ten opinion streams asking questions about President Bush and eight streams asking identical questions (except for the two) about Senator Kerry in 2004. Peanut butter scores for the two groups differed by more than 100 points and 13 (sample) standard errors, a strongly significant ($t = 7.95$, $p < 0.01$) result. Figure 8. shows the peanut butter scores and estimated total change for all 18 streams.

Fig 8. Peanut butter scores and estimate total change for Bush / Kerry streams



This is strong evidence that opinions about Kerry were the result of a more top-down process than opinions about Bush. Although I don't have the evidence to fully support the claim, I am inclined to explain this result in terms of presidential challengers and incumbents. At the start of the campaign, Bush had been in the public eye for almost four years. Not only were potential voters familiar with him, they had had time to link other observations (e.g. economic performance, the war in Iraq) with their assessments of him. Kerry was not as well known. Especially at the beginning of the campaign, most potential voters knew of Kerry only through the mass media -- exactly the conditions under which we would expect top-down information flow to be most at work. This is conjecture, of course, but seems reasonable.

The other tests did not yield significant results[7]. Mean peanut butter scores (245.86, 246.76, and 235.50, respectively) for personal, national, and mixed opinion streams barely differed at all. Mean scores for political discussion (245.27) and political news consumption (179.88) were not significantly different either ($t = -1.22$, $p=0.237$). Evidently, these constructs are not important determinants for top-down versus bottom-up information flow[8]. In one sense, this is surprising (Voters' opinions about personal matters are just as chunky as opinions about

7. A technical note: all of these tests are based on standard errors estimated from the relevant sample of streams. This test is a faithful execution of a thought experiment in which we compare representative groups of question-streams to see if their average chunkinesses differ. A different statistical test would be to compare individual streams using standard errors derived from the peanut butter procedure itself. This is a promising area for future work.
8. The subjective nature of the coding could also be a factor.

national matters?), but in another sense, it is not surprising that my original typologies of opinion streams do not match empirical reality well. This is the first time the data have been cut this way -- we should expect to be surprised by some of the results.

6. Conclusion

In this paper, I have tried to bring together a theoretical and empirical framework for researching top-down versus bottom-up information flow. I see this framework and the empirical strategy for testing it as my primary contributions. The results described in the previous section are suggestive, but barely scratch the surface of what we will eventually need to explore in order to thoroughly understand societal patterns of information flow. At this point, I would characterize them as a successful construct validation of the peanut butter test. Following are ideas for future related work.

First, the peanut butter test remains a work in progress. Throughout the paper I have highlighted limitations of this estimation strategy: it is ordinal; it is heteroskedastic with respect to stream length; its variance is unknown; it produces biased estimates of total opinion change. These do not strike me as insurmountable challenges. Monte Carlo simulations followed by analytic derivations seem like a promising way to move forward.

Second, more exploring is in order. The 75 opinion streams used in this paper represent only a fraction of the questions asked in the Annenberg studies. Any question with sufficient sample size and sequence length is a candidate for analysis by this method. In addition, subsample analysis would be interesting where possible. Do moderates move in chunks? What about voters in swing states? Are the early weeks of campaigns chunkier than the final weeks? Taking different elections and subgroups as units of analysis would greatly multiply our set of counterfactual scenarios. Other time series could also be analyzed using this approach, although the theoretical interpretation of chunkiness in other contexts (e.g. tone of political content, volume of online donations) might need to be modified.

Finally, a great deal of theoretical work remains to be done. I stand by the assertion that information flow remains understudied as a topic and fragmented as a field. Most theories that touch on information flow do so from the perspective of individuals or messages, but rarely both, and almost never in full social context. These will need to be synthesized, often across subfield lines, before we arrive at systemic understanding of political information flow.

Information flows remain one of the great unexplored frontiers of public opinion research. For some time now, the field has focused on opinion formation, relying on simplifying assumptions about mass media when thinking about sources of political information. As information technologies and modes of communication proliferate, these oversimplified models of information flow are likely to become less and less accurate. Understanding opinion formation in a world of asychronous, microtargeted, and participatory media will require the develoment of new methodological and theoretical tools. I hope this paper makes a contribution on both counts.

Bibliography

Alford, John R., Carolyn L. Funk, and John R. Hibbing. "Are political predispositions genetically transmitted?." APSR 92.2 (2005): 153-167.

Arnold, R.. Congress, the Press, and Political Accountability. Princeton: Princeton University Press, 2004.

Blumer, Herbert. "Public Opinion and Public Opinion Polling." American Socialogical Review 13.5 (1948): 542-549.

Brader, Ted. Campaigning for Hearts and Minds How Emotional Appeals in Political Ads Work.. Chicago: University Of Chicago Press, 2006.

Downs, Anthony. An Economic Theory of Democracy. Toronto: Addison Wesley, 1957.

Erikson, Robert S., Michael B. Mackuen, and James A. Stimson. The Macro Polity (Cambridge Studies in Public Opinion and Political Psychology). New York: Cambridge University Press, 2002.

Gans, Herbert J.. Deciding What's News : A Study Of CBS Evening News, NBC Nightly News, Newsweek,. New York: Pantheon Books, 1979.

Green, Donald, Bradley Palmquist, and Eric Schickler. Partisan Hearts and Minds: Political Parties and the Social Identity of Voters. New Haven: Yale University Press, 2002.

Hovland, Carl. "Reconciling Conflicting Results from Experimental and Survey Studies of Attitude Change." American Psychologist 14 (1959): 8-17.

Key, V.O.. Public Opinion and American Democracy. New York: Alfred A. Knopf, 1968.

Kleinberg, Jon. "Bursty and Hierarchical Structure in Streams." Data Mining and Knowledge Discovery V7.4 (2003): 373-397.

Lippman, Walter. Public Opinion. San Diego: Harcourt Brace & Company, 1922.

Lukers, Kristin. Abortion & the Politics of Motherhood. Los Angeles: Univ. Of California Press, 1984.

Lupia, Arthur, and Mathew D. Mccubbins. The Democratic Dilemma: Can Citizens Learn What They Need to Know? (Political Economy of Institutions and Decisions). New York: Cambridge University Press, 1998.

Lupia, Arthur. "Shortcuts versus Encyclopedias: Information and Voting Behavior in California Insurance Reform Elections." APSR 88.1 (1994): 63-76.

Page, Benjamin I., and Robert Y. Shapiro. The Rational Public: Fifty Years of Trends in Americans' Policy Preferences (American Politics and Political Economy Series). Chicago: University Of Chicago Press, 1992.

Parenti, Michael. Inventing Reality: The Politics of News Media. California: Wadsworth, 1992.

Patterson, Thomas E.. Out of Order: An incisive and boldly original critique of the news medias domination of Americas political process. New York: Vintage, 1994.

Stimson, James A.. Tides of Consent: How Public Opinion Shapes American Politics. New York: Cambridge University Press, 2004.

Wolfsfeld, Gadi. "When the press fails: Political power and the news media from Iraq to Katrina - by W. Lance Bennett, Regina G. Lawrence, & Steven Livingston." Journal of Communication 58.1 (2008): 198-199.

Zaller, John R.. The Nature and Origins of Mass Opinion (Cambridge Studies in Public Opinion and Political Psychology). New York: Cambridge University Press, 1992.

Zaller, John, and Stanley Feldman. "A Simple Theory of the Survey Response: Answering Questions versus Revealing Preferences." American Journal of Political Science 36.3 (1992): 579-616.

Appendix A: Question Wordings

2000 ANES Questions

ca01	Bush favorability	On a scale of zero to 100, ho
ca11	Gore favorability	On a scale of zero to 100, ho
ca52	Clinton favorability	On a scale of zero to 100, ho
ce20	Have online access	Do you have access to the I
ck04	Care which party wins presidential election	Generally speaking, do you
cv01	Party ID	Generally speaking, do you
cv02	Party ID Strong	Do you consider yourself a
cv03	Lean Republican or Democrat	Do you think of yourself as
cw01	Sex	
cw04	Of Hispanic or Spanish origin	Are you of Spanish or Hispa
cw07	US citizen	Are you a citizen of the Uni
ca02	Bush cares about people like me	Does the phrase "really care
ca03	Bush honest	Does the word "honest" des
ca04	Bush inspiring	Does the word "inspiring" d
ca05	Bush knowledgeable	Does the word "knowledgea
ck01	Interested in government	Some people seem to follow
cm01	Trust federal government	Thinking about the federal g
ck07	Disagreed when discussing politics with family or friends in the past week	When you talked with famil
ck11	Disagreed when discussing politics at work or online in last week	When you talked at work or
cv04	Conservative or liberal	Generally speaking, would y
ce01	Watched TV news in past week	How many days in the past
ce02	Watched cable news in past week	How many days in the past
ce06	Watched local TV news in past week	How many days in the past
ce13	Read newspaper in past week	How many days in the past
ce18	Listened to talk radio in the past week	How many days in the past
ck05	Discussed politics with family or friends in past week	How many days in the past

2004 ANES Questions

caa01	Bush favorability	On a scale of zero to 10, how would y
caa04	Bush cares about people like me	On a scale of zero to 10, how well do
caa05	Bush inspiring	On a scale of zero to 10, how well do
caa06	Bush strong leader	On a scale of zero to 10, how well do
caa07	Bush trustworthy	On a scale of zero to 10, how well do
caa08	Bush shares my values	On a scale of zero to 10, how well do
caa09	Bush knowledgeable	On a scale of zero to 10, how well do
caa10	Bush reckless	On a scale of zero to 10, how well do
cab01	Kerry favorability	On a scale of zero to 10, how would y
cab02	Kerry cares about people like me	On a scale of zero to 10, how well do
cab03	Kerry inspiring	On a scale of zero to 10, how well do
cab04	Kerry strong leader	On a scale of zero to 10, how well do
cab05	Kerry trustworthy	On a scale of zero to 10, how well do
cab06	Kerry shares my values	On a scale of zero to 10, how well do
cab07	Kerry knowledgeable	On a scale of zero to 10, how well do

cab08	Kerry reckless	On a scale of zero to 10, how well do
cac01	Cheney favorability	On a scale of zero to 10, how would y
cca01	Country going in right direction	Do you feel things in this country are
ccb09	Approve of Bush handling economy	Do you approve or disapprove of the
ccd19	Approve of Bush handling Iraq	Do you approve or disapprove of the
ccd61	Approve of Bush handling terrorism	Do you approve or disapprove of the
cea21	Have online access	Do you have access to the Internet or
cma01	Party ID	Generally speaking, do you usually th
cma02	Strength of party ID	Do you consider yourself a strong or
cma03	Lean Republican or Democrat	Do you think of yourself as closer to
cra01	Registered to vote	These days, many people are so busy
crb01	Democratic primary election status	cRB01 indicates if the respondent wa
crd01	Voted in 2000 general election	Did you vote in the 2000 presidential
cua07	Voted in 2002 election	In talking with people about politics a
cua09	House candidate voted for in 2002	In that election, you were able to vote
cwa01	Sex	The interviewer observed the respond
caa02	Bush job approval	Do you approve or disapprove of the
caa03	Bush favorability as a person	I'd like to ask you how you feel abou
ccb01	Country's economy today	How would you rate economic condi
ccb03	Country's economy next year	Thinking about one year from now, d
ccb04	Personal economic situation today	How would you rate your own person
cea05	Attention to political news on network or cable TV in past week	During the past week, how much atte
cea07	Attention to political news on local TV in past week	During the past week, how much atte
cea13	Attention to news in newspapers in the past week	During the past week, how much atte
cma06	Conservative or liberal	Generally speaking, would you descr
cea01	Watched network news inn the past week	How many days in the past week did
cea03	Watched cable news in the past week	How many days in the past week did
cea06	Watched local news in the past week	How many days in the past week did
cea08	Watched Spanish news in the past week	How many days in the past week did
cea10	Read newspaper in past week	How many days in the past week did
cea14	Listened to NPR in past week	How many days in the past week did
cea15	Listened to talk radio other than NPR in past week	If listened in NPR in cEA14: Apart fr
cea18	Watched late-night comedy in past week	How many days in the past week did
cea22	Accessed political information online in past week	How many days in the past week did
ckb01	Discussed politics with family or friends in past week	How many days in the past week did
ckb03	Discussed politics at work in past week	How many days in the past week did

Appendix B: Descriptive Statistics and Codes

Name	Year	Length	Total Responses	Est. Total Change (RA)	Est. Total Change (HMM)	Number of Chunks	Pea
ca01	2000	390	55425	3.44	2.10	16	
ca11	2000	390	55493	4.00	4.82	30	
ca52	2000	390	56979	3.34	2.10	17	
ce20	2000	390	57871	3.51	3.26	24	
ck04	2000	390	57046	3.53	1.18	10	
cv01	2000	390	34188	4.61	3.60	21	
cv02	2000	390	49082	4.06	3.69	22	

cv03	2000	390	15726	6.93	4.66	16
cw01	2000	390	57929	3.60	3.60	24
cw04	2000	390	57534	2.54	2.81	37
cw07	2000	390	57808	1.67	2.21	40
ca02	2000	390	48804	4.81	3.88	24
ca03	2000	390	48489	4.63	3.14	19
ca04	2000	390	49427	4.82	5.53	28
ca05	2000	390	49658	4.07	1.72	11
ck01	2000	390	57552	3.28	2.86	20
cm01	2000	390	25798	6.21	6.04	27
ck07	2000	390	31386	3.50	4.49	31
ck11	2000	390	15906	5.77	5.54	22
cv04	2000	390	55905	3.58	3.00	22
ce01	2000	390	57670	3.19	1.41	12
ce02	2000	390	57704	3.54	2.53	20
ce06	2000	390	57669	2.93	2.31	25
ce13	2000	390	57845	3.04	2.77	23
ce18	2000	390	57777	3.80	3.58	26
ck05	2000	390	49477	4.06	2.01	15
caa01	2004	399	80672	2.98	1.73	17
caa04	2004	399	61904	3.21	2.49	21
caa05	2004	399	62059	3.28	2.38	21
caa06	2004	399	62348	3.23	1.57	13
caa07	2004	399	62019	3.17	1.55	13
caa08	2004	399	61739	3.26	2.15	18
caa09	2004	399	48981	3.84	3.57	23
caa10	2004	399	48804	3.48	2.85	22
cab01	2004	399	73911	3.24	1.74	15
cab02	2004	399	43661	4.22	1.71	10
cab03	2004	399	44354	4.31	1.96	17
cab04	2004	399	43714	4.72	3.05	21
cab05	2004	399	43354	4.52	2.32	15
cab06	2004	399	43646	4.39	2.55	19
cab07	2004	399	38937	5.11	6.01	25
cab08	2004	399	37465	4.19	3.54	18
cac01	2004	399	77351	2.94	0.84	10
cca01	2004	399	73321	3.00	3.03	20
ccb09	2004	399	78304	2.99	2.90	20
ccd19	2004	399	78521	3.23	2.84	20
ccd61	2004	399	39080	4.16	2.69	18
cea21	2004	399	81224	2.50	2.59	29
cma01	2004	399	51154	3.68	3.95	25
cma02	2004	399	71402	2.98	2.68	22
cma03	2004	399	21130	5.47	4.50	21
cra01	2004	399	80837	2.18	2.53	26
crb01	2004	399	81310	1.29	2.00	7

crd01	2004	399	59624	2.99	2.83	24
cua07	2004	399	61124	3.02	2.95	25
cua09	2004	399	11903	7.28	7.67	29
cwa01	2004	399	81310	2.89	2.50	24
caa02	2004	399	79208	3.06	2.72	21
caa03	2004	399	36581	3.84	2.30	15
ccb01	2004	399	80071	2.48	1.99	21
ccb03	2004	399	72578	3.02	3.29	21
ccb04	2004	399	80179	2.64	2.03	19
cea05	2004	399	67975	3.45	1.78	18
cea07	2004	399	64138	3.50	1.48	13
cea13	2004	399	62267	3.45	1.71	17
cma06	2004	399	78925	3.02	4.06	34
cea01	2004	399	80937	2.92	2.29	20
cea03	2004	399	80992	2.73	3.76	33
cea06	2004	399	81017	2.59	2.91	30
cea08	2004	399	81217	1.66	2.23	32
cea10	2004	399	81199	2.36	3.24	32
cea14	2004	399	80873	2.65	3.15	31
cea15	2004	399	81086	2.72	1.53	15
cea18	2004	399	81163	2.41	3.03	30
cea22	2004	399	44872	3.49	1.50	10
ckb01	2004	399	80804	2.82	2.11	25
ckb03	2004	399	80996	2.94	1.78	17