Psychological Review
1999, Vol. 106, No. 2, 362-384

0033-295X/99/$3.00

A Stochastic Model of Preference Change and Its Application to 1992 Presidential Election Panel Data

Michel Regenwetter
Duke University

Jean-Claude Falmagne and Bernard Grofman
University of California, Irvine

The authors present and test a model for the evolution of preferences. Personal preferences are represented by rankings with possible ties and are posited to change under the influence of "tokens" of information in the environment. These tokens may not be directly controlled or observed by the researcher. The authors apply the model to 1992 National Election Study panel data (W. E. Miller, D. R. Kinder, S. J. Rosenstone, & NES, 1993). The parameter estimates suggest that negative campaigning played a major role in the information flow. Democrats and Republicans experienced a barrage of contradicting information about Perot; Democrats, Republicans, and Independents each received or perceived different information. A shift in the perception of the candidates led the Republicans to evaluate Bush and Perot less favorably after the election. These results demonstrate the model's potential to analyze persuasion as a real-time stochastic process and without a media content analysis.

In the vast literature on persuasion and propaganda (see, e.g., Anderson, 1971; Latané, 1981; Petty & Cacioppo, 1981; Zaller, 1992), there are few quantitative models of attitude change that apply to panel data; the best known is the Converse "black and white" model (Converse, 1964, 1975; Converse & Markus, 1979; Markus, 1982; McPhee, Andersen, & Milholland, 1962). In a panel study, the same individuals in a large sample have been questioned repeatedly about their preference concerning a fixed set of alternatives. In addition to these questions, they may also have been asked more general questions about their social background or attitudes. Although the standard models for panel data have a probabilistic component (modeling, e.g., the measurement error), they typically do not cast such data as a manifestation of a stochastic process in the specific sense of this term in the theory of stochastic processes (e.g., Norman, 1972; Parzen, 1994). As a consequence, detailed temporal predictions cannot be computed.

In this article, we present a stochastic model of preference change that is an extension of a model developed by Falmagne, Regenwetter, and Grofman (1997). We also describe an application to an important set of data. The model is closely related to work in mathematical behavioral sciences and combinatorics (Doignon & Falmagne, 1997; Falmagne, 1996, 1997; Falmagne & Doignon, 1997; Regenwetter, 1997). The key idea is consistent with social choice theory in that, at any time, an individual's attitude is represented by a preference relation on the set of alternatives. These preference relations are formalized by "weak orders"; that is, rankings with possible ties. An individual's current ranking may be altered by "tokens" of information delivered by the environment. Each token contains information about a single alternative. The effect of a token is to move that alternative up or down in the current ranking or, in some circumstances, to leave the present ranking unchanged. We stress that these tokens are not necessarily identifiable or controllable by the behavioral scientist. They are theoretical constructs intended to represent the "particles" in the barrage of information to which an individual, say a potential voter in an election, may be submitted. A television program for example (or a conversation with a neighbor, or a newspaper article), may be the source of one or more tokens. Moreover, it is plausible that a Republican voter and his or her Democrat visitor watching the same television program may perceive different messages. This suggests that, by itself, a content analysis of the media gives only a partial picture of a marketing campaign and its effects on the consumers or voters. More specifically, a simple content analysis of the mass media misses the perceptual filter that individuals use to screen and interpret information. It also misses the way in which information flow is mediated by social interactions, not just by direct mass media exposure. Moreover, potential selective exposure effects lead different subsets of the population to sample different media sources.

Our model contains parameters that allow for the same message to be differentially perceived or received by different individuals. This model formalizes, in the guise of the unobservable tokens, a very general notion of information, not restricted to either mass media or social interaction, but rather combining all sources. It is a tool permitting a detailed statistical analysis of preference evolution in terms of inherently unobservable key psychological features of the information flow. The estimated parameters of the model actually provide a way to quantify the perceived content of information. We point out that the model is formalized as a

Michel Regenwetter, Decision Sciences Area, Fuqua School of Business, Duke University; Jean-Claude Falmagne, Cognitive Sciences Department, University of California, Irvine; Bernard Grofman, Department of Politics and Society, University of California, Irvine.

This research was supported by National Science Foundation (NSF) Grant SBR 9307423 and NSF Grant SBR 9730076. We are grateful to Joe Rodgers and Yung-Fong Hsu for their remarks.

Correspondence concerning this article should be addressed to Michel Regenwetter, Decision Sciences Area, Fuqua School of Business, Duke University, Box 90120, Durham, North Carolina 27708. Electronic mail may be sent to michel.regenwetter@duke.edu.

real-time stochastic process. In our case, that means a stochastic process in which the changes of preferences can occur at any time t, where t is any positive real number. The following quotation from Latané, Nowak and Liu (1994) suggests that the field is ripe for such a model. They write: "In order to predict an election three months hence, it is simply not enough to know the proportion of people favoring each candidate now. People may change their minds, not only in response to new information, but as a result of social interaction." Furthermore, they emphasize that "to fully understand how social systems organize and evolve, we need to further develop a dynamical view of group processes that fully takes into account the crucial dimensions of space and time."

We first review some relevant literature. Next, we describe the data to which the model is applied. An informal outline of the model follows, with a discussion of some key predictions. We then turn to the precise technical description of the model in the form of four axioms. The following sections sketch our statistical methods and provide the actual data analysis of the 1992 National Election Study (NES; Miller, Kinder, Rosenstone, & NES, 1993) panel. The concluding discussion summarizes our work and considers possible extensions and further applications of the model.

Related Literature

We divide our literature review into two parts. One examines three different possible mechanisms of persuasion. The other reviews two time-dependent social process models related to our own work.

Mechanisms of Persuasion

Theoretical approaches to persuasion fall into three broad categories. First are those models that focus on the biases in the information environment of the respondent. In particular, there is a vast literature in political science on ideological and partisan bias in the mass media (see, e.g., Ginsberg, 1986). Another very large, mostly experimental literature studies one-way versus two-way communication, recency effects, inoculation effects, the power of repetition, and so on (reviews can be found in McGuire, 1964, 1969, 1985; Zimbardo & Ebbesen, 1970).

In the second category, the respondents are not seen as passive recipients of propaganda. Rather they filter, categorize, and evaluate information in terms of their own values and biases. This strand of literature is summarized by Jowett and O'Donnell (1992). Among the classic psychology references are Abelson (1964), Anderson (1971), and Moscovici (1976, 1985). In political science, there are many references on party identification as a kind of cognitive bias. For example, it has been argued that partisans distort the information they hear, so as to construe it in a fashion more favorable to the candidates they support (Grofman, 1985). Perhaps the most important recent political science reference in this context is Zaller (1992).

Third, an important stream of research deals with the catalyzing role of the personal environment and the social context. One of the earliest works of importance is by Lazarsfeld, Berelson, and Gaudet (1948), who showed that the communications that people had with each other could be as important in shaping attitudes as the direct flow of information from the mass media. In particular, this research promoted the concept of a "two-step flow of information" in which "information leaders" act as intermediaries between the mass public and the mass media. That is, information leaders partially control the messages that the masses receive. Research such as that of Latané and his colleagues (Latané, 1981; Latané & Nida, 1979; Latané, Nowak, & Liu, 1994; Latané & Wolf, 1981; Nowak, Szamrej, & Latané, 1990) shows how the social networks in which respondents are embedded can dramatically change the impact of information on attitudes. For instance, Nowak et al. (1990) ran computer simulations of the change of attitudes in a social network based on Latané's (1981) "theory of social impact." The key variables of social impact theory are the strength, immediacy, and number of other people in the social environment. As Nowak et al. (1990) phrased it, "social impact theory concerns the magnitude of the impact that one or more people or groups (sources) have on an individual, and thus is a static theory of how social processes operate at the level of the individual at a given point in time." A comprehensive review of the psychology of attitudes can be found in Eagly and Chaiken (1993).

Two Time-Dependent Models of Attitude Change

Nowak et al. (1990) extended Latané's theory of social impact to attitude change by incorporating the individual's persuasiveness, that is, the ability to convince others of their own attitude. Attitude change takes place whenever the impact of individuals with a different opinion is greater than that of individuals with the same opinion. The impact of individuals on others also depends on their distance in the social network in which they are embedded. A computer simulation of a discrete time process based on these concepts shows how the distribution of attitudes in the social network can evolve. This study demonstrates how a microlevel theory of persuasion can lead to emerging macrolevel patterns, such as the polarization of attitudes, equilibria between competing attitudinal groups, and existence of minority subgroups. The models in these articles give a quantitative expression of the phenomenon of attitude changes. In their current formulation, however, they do not entail a straightforward statistical analysis of data sets, nor do they predict the evolution of attitudes in real time.

Among the methodologically most sophisticated models capable of statistical analyses and of sequential predictions is the epidemic model of the onset of social activities by Rodgers and Rowe (1993). These models predict the progress of the individual through a hierarchical stage process (e.g., adolescent sexual behavior) in discrete time. The key goal of the methodology is the prediction and fit of onset data. The model presented here resembles this work in that we specify the detailed statistical procedures for parameter estimation, and that substantive hypotheses can be cast as constraints on the parameters of the model.

Our model belongs to a general class studied by Falmagne (1997). It differs from all previous theorizations of preference change in two respects: (a) We formalize the information flow and its effects on the respondents as a real-time stochastic process; (b) we posit that the changes of preferences are due to tokens of information, which are inherently unobservable. This last feature is critical, because it renders the model applicable beyond the laboratory. Indeed, if the data are provided by respondents in their real-life environment (as in a standard polling situation), there is no practical way of controlling or measuring the panoply of messages bombarding them. That is, we cannot literally follow the respondent around and record the information flow to which she or he is exposed. Even if we could, we would still need to model the interaction between the individual respondent and that information flow.

Data

We first describe the type of data that the model is intended to explain. The exemplary data point is a sequence of rankings made by an individual at times $t_1, \ldots, t_n$. We suppose that at each time t_i three alternatives have been ranked by each respondent. These alternatives might be, for example, competing brands of a product, health plans, schools, music idols, or political candidates, as in the particular application analyzed here. The only requirement is that the alternatives can be ranked by the respondents according to the dimension under study, such as desirability, personal utility, trustworthiness, or trendiness. In our study, this dimension is evaluated by a rating scale measuring "how warm" the respondent feels toward the candidates (see the Data Analysis).

The model of this article is applied to panel data pertaining to the 1992 U.S. presidential election, with the three alternatives being Bush, Clinton, and Perot. There are compelling reasons for focusing on this case. For one, the NES panel (Miller, Kinder, Rosenstone, & NES, 1993) provides reliable information about evaluations of these three candidates for a substantial sample of the American electorate at two time points (one shortly before and one shortly after the election). For another, the respondents have classified themselves as Democrats, Republicans, or Independents, and this information can be used to investigate the differences between the three subpopulations. The partition of the general population into these three classes is based on the concept of party identification, which has been standard in the political science literature for more than 40 years. It can be expected that members of these three subpopulations receive different information during a campaign, may interpret the same information differently, and may initially have different evaluations of the candidates. All in all, by testing our model of preference change on a well-studied election, we can provide a good assessment of the validity of its substantive interpretations.

The axioms of the model ensure that detailed testable predictions can be obtained. In particular, we derive an exact expression for the asymptotic (i.e., long term) probability of any given ranking $\succ$ (see Theorem 3 in Appendix B). In the context of an election, this is the probability that a voter, sampled from a specified population after a long campaign, would rank the alternatives according to the relation $\succ$. We also compute the joint probability of observing rankings $\succ$ and $\succ'$ at times t and $t + \delta$, respectively, for each pair of rankings $\succ$ and $\succ'$ and for large t (see Theorem 5 in Appendix B). Moreover, the model permits to estimate, for each subpopulation, the positive and negative "bias" regarding each alternative. In the model, a particular bias results from a relative preponderance of certain tokens.

The data to be analyzed consist, for any subpopulation g and any pair $(\succ, \succ')$ of rankings of the alternatives, in the number $N(g, \succ, \succ')$ of respondents in the sample who belong to population g and have provided ranking $\succ$ at time t and ranking $\succ'$ at time $t + \delta$. Thus, summing over all the subpopulations and all the pairs of rankings yields the total number of participants (N) in the sample,

$$N = \sum_{g} \sum_{(\succ, \succ')} N(g, \succ, \succ'). \tag{1}$$

In our empirical example, $N = 2{,}024$. There are two basic types of theoretical results. One concerns the asymptotic probabilities of the preference relations at a single time point. This result can be used to predict the number $N(g, \succ) = \sum_{\succ'} N(g, \succ, \succ')$ of respondents in subpopulation g who have indicated a particular preference relation $\succ$ at time t. (This assumes that t is large enough to justify using the asymptotic result. This assumption was supported by the data, see Data Analysis, Part 2.) The second type concerns the joint probability of observing ranking $\succ$ at time t and ranking $\succ'$ at time $t + \delta$ for any pair $(\succ, \succ')$ of rankings and also for large t. These probabilities are used to predict the numbers $N(g, \succ, \succ')$ in Equation 1. (Predictions involving more than two time points can obviously also be derived using the methods of this article; see also Falmagne, 1997, or Falmagne et al., 1997). Both types of results are expressed in terms of parameters assessing the density and the type of information delivered by the medium, and the effects of this information on the preference relations of the respondents. These results allow us to study differences in the information flow to different subpopulations and to determine whether or not the election outcome—or possibly, the reentry of Perot in the race in October[1]—had an impact on the electorate's perception of the candidates.

The analysis of the data supports the following four conclusions: (a) negative campaigning seems to have played a major role in the information flow (cf. Ansalobehere, Iyengar, Simon, & Valentino, 1994; Garramone, 1985; Skaperdas & Grofman, 1995); (b) between the first (preelection) and the second (postelection) interviews, Democrats and Republicans appear to have been submitted to a barrage of contradicting information about Perot (negative vs. not so negative), revealing an unstable image of this candidate; (c) Democrats, Republicans, and Independents each received (or perceived) different information; and (d) there was a statistically significant shift between the two interviews in the evaluations of the candidates that led the Republicans to evaluate both Bush and Perot less favorably.

Some of these conclusions may appear trivial to a political scientist in view of the literature and the common wisdom in the field. Note, however, that this common wisdom resulted from the combination of a large number of methods of analysis. By contrast, we reach the four conclusions solely by reconstructing the information flow from the joint evolution of the individual preferences. In particular, these conclusions are obtained without relying on a conventional content analysis of the mass media. That we are able to reach accurate conclusions about perceptual biases, negativity of information flow, and postelection effects is encouraging. Had our analysis failed to reveal the volatility in the evaluations of Perot or failed to find differences in the actual or perceived information flow of Democratic or Republican partisans, we would have good reason to regard the model with suspicion.

[1] We are grateful to an anonymous reviewer for pointing out that, at the time of Perot's reentry into the race on October 1, about one third of the data from the first interviews had already been collected. Therefore, any difference between the first and the second interviews may be due in part to that reentry.

Outline of the Model

As in Falmagne et al. (1997), we write $\mathscr{A}$ for the set of alternatives. The symbols i, j, k denote variables referring to the elements of $\mathscr{A}$. In our examples and in the data analysis section, $\mathscr{A}$ = {Bush, Clinton, Perot}. (We will often abbreviate "Bush," "Clinton," and "Perot" by B, C, and P, respectively.)

The model presented in this article is built on the following theoretical constructs: the subpopulations of the respondents, their latent preferences, their initial state, and the information medium, which is represented by a stochastic flow of information "tokens." Our mathematical model will relate these concepts to observable entities, namely the respondents' evaluation of the alternatives at time points $t_1, \ldots, t_n$, and the partition of the sample of respondents into subsamples corresponding to party identification. Note for further reference that, in our application, the respondents' evaluations will be derived from the so-called thermometer ratings (see later discussion), and the partitioning will be based on the self-classification of the respondents into Democrats, Republicans, and Independents.

The Subpopulations

We assume that the sample of respondents can be partitioned into three subsamples corresponding to three subpopulations according to party identification (for instance). We suppose that this information is provided by the respondents (as was the case for the data analyzed here). It seems reasonable to suppose that the respondents in each subsample have access to possibly different channels of information and may have different prior evaluative biases. In the application, we use the three subpopulations: Democrats, Republicans, and Independents. We generically refer to subpopulation g, with g = d, r, u (for Democrat, Republican, and Uncommitted, respectively). This classification is assumed to be invariant in the course of the study. In other words, even though respondents may change their preferences over time, they do not change their underlying political orientation. (This assumption is consistent with the fact that our 1992 election data cover only a time interval of a few months.)

The Latent Preferences

We suppose that the responses of an individual to some questions of a survey are governed by a latent personal preference relation, which we call the *state* of that individual. By a *preference relation* on the set $\mathscr{A}$, we mean a binary relation in the usual sense of set theory (i.e., a set of ordered pairs of elements of $\mathscr{A}$). We restrict consideration to strict weak orders, that is, rankings with ties. (A formal definition of the concept of a strict weak order is in Appendix A.) We write $i \succ j$ to mean that alternative i is strictly preferred to alternative j. This family of relations contains as a special element the empty relation $\varnothing$, representing the situation in which no alternative is strictly preferred to any other. The relation $\varnothing$ is referred to as the *neutral state.*

In the case of $\mathscr{A}$ = {B, C, P}, there are 13 possible states. The set of all states will be denoted by $\mathscr{S}$. These 13 states are represented by their graphs in the 13 rectangles of Figure 1. Notice that the neutral state is represented by the empty rectangle in the middle of the figure. (Ignore for the moment the arrows in the figure and the reference to the random walk in the caption.) Figure 1 is designed so as to facilitate the visualization of the process in three dimensions, each of which corresponds to one of the candidates. The third dimension (Perot's dimension) is captured by representing elements in the background (i.e., Perot at the bottom) smallest and elements at the foreground (i.e., Perot at the top) largest.

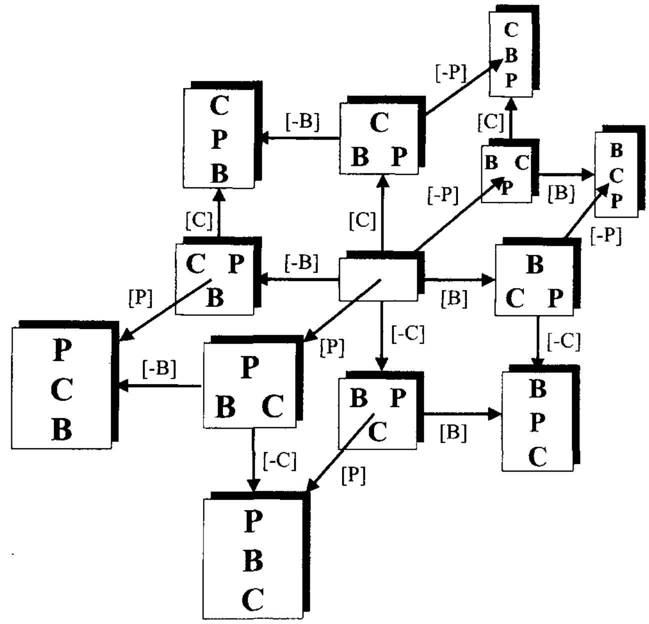


Figure 1. Three-dimensional transition diagram of the random walk on the set $\mathscr{S}$ of states (strict weak orders). The positive or negative tokens producing a transition are marked along the arrows of the diagram. Note, for further reference, that parallel arrows pointing in the same direction correspond to the same token. To simplify the graph, only the centrifugal transitions (i.e., away from the neutral state) are indicated. The centripetal transitions can be obtained by reversing the arrows and by capping the symbols representing the tokens by tildes. Note that the probabilities of the transitions depend on the subpopulation. B = Bush; C = Clinton; P = Perot.

The Initial State

At the beginning of the process, any individual of the population of reference is in some state, which may be the indifference relation or some other state. In our empirical example, this initial state may be thought of as the preference ranking of an individual when first informed about the list of presidential candidates. For instance, Democrats might start more favorable to Clinton and Republicans more favorable to Bush. Here, we simply assume that an initial distribution on the set of states exists, which may depend on party identification. For data collected early enough in the process, the initial distribution can, in principle, be reconstructed (i.e., the parameters can be estimated). In the data analysis presented here, however, this initial distribution only plays a technical role because we assume (and also test) that the respondents have reached asymptote at the time of the first evaluation.

The Stochastic Environment

Starting from the initial state, successive transformations may take place over time. Specifically, we assume that the individual is immersed in a stochastic environment delivering at certain random times $t_1, \ldots, t_n, \ldots$ "tokens" of information regarding the alternatives. These tokens represent events occurring in the environment and having a positive or negative connotation regarding particular alternatives. We assume that the stochastic environment of respondents in different subpopulations may be different, either because the information they sample is different or because they evaluate information differently. For instance, a television report that depicts Bush as a "true Republican" may be positive information for most Republicans but negative information for most Democrats. Thus, both the sources and the interpretation of information from the same source may differ between subpopulations. People in different subpopulations may nevertheless overlap in their sources of information. All these notions are operationalized by assuming that both the rate of occurrence of the tokens and their nature may differ depending on the subpopulation to which a respondent belongs. An important feature of the analysis is that the density of the tokens and their respective probabilities can be estimated for each subpopulation. These estimates can be compared, and may reveal critical differences among Democrats, Republicans, and Independents concerning the exposure to the information available in the medium.

Many studies of persuasion use an experimental or quasiexperimental design (Iyengar & Kinder, 1987; Iyengar, Peters, & Kinder, 1982). In our model, however, the occurrence of the tokens need not be either observed or controlled by the social scientist in order for the model to apply. Rather, we infer statistical properties of the token flow from the application of the model to the observed data on preferences and preference changes. This aspect of the model makes it applicable to nonexperimental settings.

The Tokens and Their Effects on the State

As indicated, the particular tokens delivered by the environment are not recorded or controlled by the social scientist and thus are not identified in the model.[2] Nevertheless, the following list is suggestive of the possible sources of tokens: television programs, newspaper articles, campaign ads, conversations with acquaintances, and so on. In the model discussed here, four types of tokens are considered for each alternative: A token can be *positive* or *negative,* and to each of these two cases corresponds its *opposite* token. For convenience, we refer to the class of tokens providing positive information about i as "token $[i]$." We assume that each token of a given type about some alternative is equivalent in effect to any other token of the same type and about the same alternative. Thus, in our example $[B]$ denotes any positive token for Bush.

Table 1 shows the four possible types of tokens and the notation used for each. The occurrence of a particular token does not necessary modify the state of an individual. For example, a positive token $[i]$ has no effect if there already exists a unique best alternative in the current state or if alternative i is currently viewed as the unique worst. Otherwise, the positive token $[i]$ modifies the current state so as to move alternative i to the top position of the strict weak order, that is, the position in which i strictly dominates the two other alternatives. There are three such states in which i is the unique best (i-Best), namely

Table 1
The Four Types of Tokens and Their Notation

Types of tokens	Representing symbol
Alternative i is the best	$[i]$
i is not the best	$[\tilde{i}]$
i is the worst	$[-i]$
i is not the worst	$[\widetilde{-i}]$

$$[i > j,\, i > k],\ [i > k > j], \quad \text{and} \quad [i > j > k].$$

(The notation should be self-explanatory. For example, $[i > j, i > k]$ denotes the preference relation of a person who prefers i to both j and k but is indifferent between the latter two.[3]) An individual in state $[i > k, j > k]$—that is, preferring both i and j to k—and perceiving token $[i]$ would end up in state $[i > j > k]$. This transition is represented by the upward arrow in the upper right corner of Figure 2.

All the possible transitions are represented in this figure. This generic graph will be useful for the rest of this section. For each preference relation, we indicate the subpopulation g of the respondent.[4] The reader should also keep in mind that the processes described here are taking place according to the same mechanism in different subpopulations, but possibly with different token frequency patterns among subpopulations.

We previously indicated that a positive token $[i]$ has no effect on the current state if a unique best alternative already exists, or if alternative i is regarded as the unique worst. Similarly, a negative token $[-i]$ has no effect if there already exists a unique worst alternative or if alternative i is currently viewed as the unique best. Otherwise, the effect of a negative token $[-i]$ is to move alternative i to the bottom of the weak order. The opposite of a positive token $[i]$, denoted by $[\tilde{i}]$, has an effect only if i is currently viewed as the unique best. As indicated in Figure 2, the occurrence of such a token will modify the current state by removing alternative i from its top position. Two instances of such a transition occur in Figure 2, namely, the two counterclockwise downward arrows. The downward arrow in the lower left corner represents, for instance, the transition from the state $[i > j, i > k]$ to the neutral state. Similarly, the opposite of a negative token $[-i]$ is denoted by $[\widetilde{-i}]$ and removes i from the bottom position if i was there. Note that the opposite tokens, when they are effective, always transform a state into one nearer the neutral state.

Tokens play such a crucial role in the stochastic mechanism we posit as responsible for the evolution of preferences that they

[2] The possibility of such an identification is discussed in the last section of this article.

[3] Note the convenient abuse of notation committed in i-Best: Because each of the three formulas specifies a different strict weak order, different symbols (e.g., $>$, $>'$ and $>''$) should have been used. Our convention simplifies the writing and will be used whenever the context makes clear what is intended.

[4] We only have the information concerning party membership at the time of the first interview. We assume that the respondents did not change parties in the time period covered by this study. If the party identification had been asked at each interview, the model could have been elaborated to cover this case.

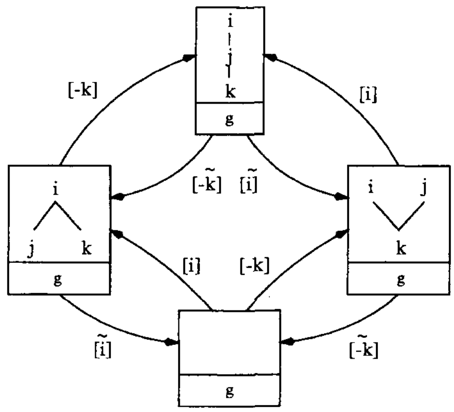


Figure 2. Effects of the tokens on the various possible states of the respondents in subpopulation g. Six different instances of this graph are found in Figure 1, obtained by setting i, j, and k equal to B (Bush), C (Clinton), and P (Perot) in the six possible ways.

deserve close attention. There are various reasons motivating the particular choice of token types made in this article.[5] These tokens formalize the concept of an atomic unit of information concerning a single alternative. They capture the natural intuition that an item of information can be either favorable or unfavorable toward a given alternative. Note that we allow for intensity of message in that we distinguish between a positive token $[i]$ which says that i is the best, and a token $[\widetilde{-i}]$, which says that i is not the worst (and similarly for the other two tokens). In the framework of a weak order representation of preferences, the effects of the types of tokens used here have a convenient geometric interpretation that, in turn, induces a particular form of asymptotic probability distribution on the states. These aspects will be laid out later. A key feature is that the effect of a single token on an individual's preference state critically depends on that individual's current state. In particular, the transition mechanisms that we postulate instantiate the intuition that dramatic changes of preference cannot be caused by single items of information (see Mackelprang, Grofman, & Thomas, 1975, in this connection).

Indeed, the states are endowed with some rigidity in the sense that (as already mentioned) a token is not always effective. For example, the occurrence of token $[i]$ has no effect on state $[j > k > i]$ because each of the three states having i in the top position (the states in i-Best) is far removed from $[j > k > i]$. In general, transformations only take place between adjacent states (i.e., states linked by an arrow in Figure 2).

The effect of a token ζ on a state $>$ will be captured by an operation $\bigcirc$, which is defined by the graph of Figure 2. Thus, the operation $\bigcirc$ maps the pair $(>, \zeta)$ to some strict weak order $>' = > \bigcirc \zeta$. In other terms, an individual in state $>$ and receiving the token ζ ends up in state $>'$. (Note that in some cases, $>$ and $>'$ may denote the same state, i.e., the token need not always have an effect.) We shall make this more concrete and illustrate the basic concepts with a hypothetical token sequence involving the 1992 presidential candidates. This is depicted in Figure 3.

The time axis is in the first column, flowing from the top to the bottom of the figure. The horizontal bars mark the occurrence of the tokens. The tokens themselves are indicated in the second column. The current state is pictured in the third column by its graph (the Hasse diagram of combinatoric theory). The respondent in this example is a Republican. We suppose that, at Time 0, the respondent prefers Bush to the other two between whom she or he is indifferent. The respondent remains in that state until the occurrence of the first token $[B]$ at time t_1. This token may arise, for instance, from a newspaper article depicting Bush as a "true Republican," which the respondent understands as a positive token about Bush. (Note that the same newspaper article may generate a negative token $[-B]$ from the viewpoint of a Democrat.) This token does not change the respondent's preferences, as it is already in line with the current state. Thus,

$$[B > C, B > P] = [B > C, B > P] \bigcirc [B].$$

Next comes the negative token $[-C]$ at time t_2 (e.g., a token that sheds negative light on Clinton). Its effect is to move alternative C to the bottom of the preference relation, resulting in the state

$$[B > P > C] = [B > C, B > P] \bigcirc [-C].$$

Token $[\tilde{B}]$ occurs at time t_3, say, in a television report that shows Bush in confrontation with some fellow Republicans. The respondent in our example does not agree with Bush and displaces him from the top position, moving to a preference state closer to the neutral state

$$[B > C, P > C] = [B > P > C] \bigcirc [\tilde{B}].$$

The token $[C]$ occurring at time t_4, has no effect on the current state $[B > C, P > C]$. The reason is that changing that state into a state having alternative C in the top position would be a major transformation, which cannot be realized by a single token. (An interpretation is that this positive token $[C]$, being in conflict with the individual's current state, is discarded as propaganda; cf. Zaller, 1992.) We leave it to the reader to verify the effects of tokens $[\widetilde{-C}]$ and $[\tilde{B}]$ occurring at times t_5 and t_6. In particular,

$$\varnothing = \varnothing \bigcirc [\tilde{\boldsymbol{B}}].$$

At first blush, the opposite tokens may appear superfluous. Intuitively, it may perhaps seem that their role could be reassigned—by some appropriate modification of the rules of the model—to the positive and negative tokens. However, they were introduced to give the model some flexibility in capturing important effects. For example, a prevalence of opposite tokens in the environment would induce a high probability of the neutral state. In particular, in a political campaign, a high proportion of opposite tokens of the types $[\tilde{i}]$ or $[\widetilde{-i}]$ could yield a large number of undecided voters. It may also seem that the principles of the model preclude a drastic change of state in a small amount of time. Such a change is, in fact, feasible as a result of a volley of tokens, an event of relatively low probability.

The mechanisms of attitude change depicted in Figure 2 as a function of tokens of information are assumed identical for all respondents (i.e., even though different subpopulations may be

[5] In other models based on similar principles but relying on different representations of the preferences (total orders in Falmagne, 1996; semiorders and other relations in Falmagne & Doignon, 1997), the tokens are of a comparative nature (e.g., "i is better than j").

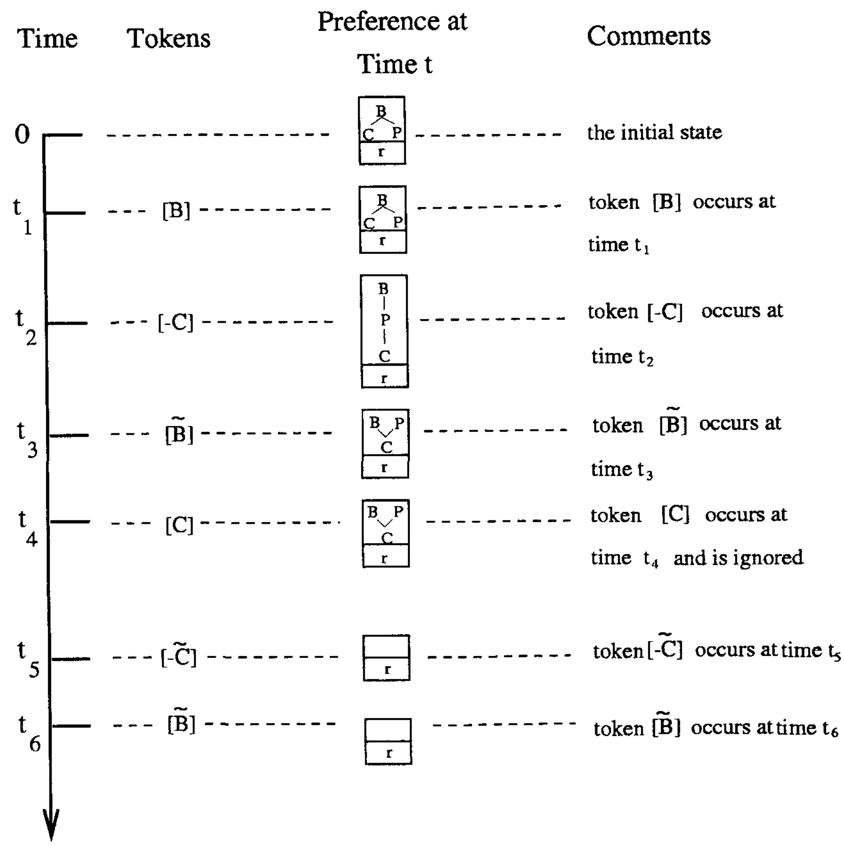


Figure 3. Illustrative hypothetical sequence of tokens occurring at times $t_1, \ldots, t_6, \ldots$ and the resulting preference relations. The hypothetical respondent is a Republican. This is illustrated by the label r under each state. B = Bush; C = Clinton; P = Perot.

exposed to the same type of token with different probabilities, the effect of a given type of token is the same for everyone regardless of his or her subpopulation). This assumption is not shocking because, as indicated earlier, the same physical piece of information may, for individuals belonging to different subpopulations, translate into different tokens.

Finally, unless specified otherwise, we suppose that the delivery of the tokens is a stable process in the sense that the probabilities of occurrence of the tokens in any interval of time $[t, t + \delta]$, $t > 0$, $\delta > 0$ do not vary with t. (However, these probabilities are allowed to differ among subpopulations.) Note that this temporal stability is only critical for the short-term aspects of our predictions (see the remark after Theorem 3 in Appendix B).

An important consequence of the axioms stated later is that the temporal succession of states for respondents in any given subpopulation g is a homogeneous random walk on the family of all states; the transition probabilities are governed by the parameters corresponding to channel g. The graph in Figure 1 displays the transitions of this random walk. To simplify the graph, only the centrifugal transitions (i.e., away from the neutral state) are indicated. The centripetal transitions can be obtained by reversing the arrows and capping each of the symbols representing tokens by a $\tilde{}$ sign. To understand the functioning of the random walk, it may be useful to reexamine the sequence in Figure 3, and follow the transitions between the states on the graph of Figure 1.

One interpretation for the transformations illustrated by Figures 1 and 2 is that the alternatives are implicitly evaluated by the respondents on a 3-point scale having -1, 0, and $+1$ as possible values, with 0 serving as a reference point. Equivalent alternatives are always rated 0. The value 1 corresponds to the top position of the state, when only one alternative occupies that position (cf. alternative i in i-Best). A value of -1 corresponds to the bottom position, with the same proviso. For example, a value of -1 on the third dimension corresponds to Perot being at the bottom of the ranking. The value 0 given to some alternative j corresponds to the neutral state and any of the following three pairs of cases, namely:

$$[i > j > k], \quad [k > j > i],$$
$$[j > k, i > k], \quad [k > i, j > i],$$
$$[i > j, i > k], \quad [k > j, k > i].$$

The effect of a positive (negative) token is to add (subtract) 1 to (from) the value of an alternative if the present value is 0 and no other alternative presently has value 1 (-1). The effect of an opposite of a positive (negative) token is to subtract (add) 1 from (to) the value of an alternative if the present value is 1 (-1). An arrow between adjacent elements in Figure 1 corresponds to adding or subtracting 1 to or from the value of one of the alternatives.

Sketch of Some Mathematical Results

One important feature of the random walk is that asymptotic results can be obtained. In particular, it can be shown that the asymptotic probabilities of the states satisfy the following regularity condition (Invariance of Ratios), which follows easily from results in Falmagne et al. (1997; see Theorem 3 in Appendix B and the comments that follow it). In words,

> *The asymptotic probabilities of two distinct adjacent states $\succ$ and $\succ \bigcirc \zeta$ differ by a factor that depends on the token ζ and on the subpopulation g,* but not on the state $\succ$.

In other words, the ratio of these two probabilities does not depend on the state $\succ$. Let us write $\pi(g, \succ)$ to denote the long-term probability of sampling an individual that belongs to subpopulation g and is in state $\succ$. As an illustration of the Invariance of Ratios property, consider the three transitions of Figure 1, which are marked by an arrow pointing to the upper right corner. All three transitions result from the occurrence of the same token $[-P]$. According to the Invariance of Ratios property, the following three ratios must be equal:

$$\begin{aligned}\frac{\pi(g, [B \succ C \succ P])}{\pi(g, [B \succ C, B \succ P])} &= \frac{\pi(g, [B \succ P, C \succ P])}{\pi(g, \varnothing)} \\ &= \frac{\pi(g, [C \succ B \succ P])}{\pi(g, [C \succ B, C \succ P])} \\ &= \mathcal{B}_g[-P]. \end{aligned} \tag{2}$$

The last inequality defines the quantity $\mathcal{B}_g[-P]$, which may be estimated from the data and may be regarded as an index measuring the "bias" against Perot in subpopulation g (i.e., the tendency, in that subpopulation g, to rank Perot last[6]). Similarly, the three arrows marked $[B]$ pointing toward the right in Figure 1 yield

$$\begin{aligned}\frac{\pi(g, [B \succ P \succ C])}{\pi(g, [B \succ C, P \succ C])} &= \frac{\pi(g, [B \succ C, B \succ P])}{\pi(g, \varnothing)} \\ &= \frac{\pi(g, [B \succ C \succ P])}{\pi(g, [B \succ P, C \succ P])} \\ &= \mathcal{B}_g[B], \end{aligned} \tag{3}$$

with the index $\mathcal{B}_g[B]$ reflecting the bias favoring Bush in subpopulation g. More generally, for any subpopulation g, any state $\succ$, and any token ζ such that $\succ \neq \succ \bigcirc \zeta$, we have

$$\mathcal{B}_g(\zeta) = \frac{\pi(g, \succ \bigcirc \zeta)}{\pi(g, \succ)}. \tag{4}$$

We indicate, in passing, an immediate consequence of Equation 4 marking a trade-off between the tendencies of judging an alternative best, or worst, and their respective opposites. If $\succ' = \succ \bigcirc \zeta$, then $\succ' \bigcirc \tilde{\zeta} = (\succ \bigcirc \zeta) \bigcirc \tilde{\zeta} = \succ$, and if $\succ$ and $\succ'$ are distinct states, then

$$\mathcal{B}_g(\zeta) = \frac{\pi(g, \succ \bigcirc \zeta)}{\pi(g, \succ)} = \frac{\pi(g, \succ \bigcirc \zeta)}{\pi(g, (\succ \bigcirc \zeta) \bigcirc \tilde{\zeta})} = \frac{1}{\mathcal{B}_g(\tilde{\zeta})},$$

yielding

$$\mathcal{B}_g(\zeta) \cdot \mathcal{B}_g(\tilde{\zeta}) = 1. \tag{5}$$

(Note that we use the convention: $\tilde{\tilde{\zeta}} = \zeta$ for any token ζ.) The bias indices have an interesting relationship to the probabilities of the tokens and their opposites. Writing $\theta_g(\zeta)$ for the probability of token ζ in subpopulation g, we have

$$\mathcal{B}_g(\zeta) = \frac{\theta_g(\zeta)}{\theta_g(\tilde{\zeta})}$$

(see Theorem 4 in Appendix B). The two ratios

$$\mathcal{B}_g[i] = \frac{\theta_g[i]}{\theta_g[\tilde{i}]}, \quad \mathcal{B}_g[-i] = \frac{\theta_g[-i]}{\theta_g[\widetilde{-i}]}$$

play an important role in the model and will be referred to as the *positive bias (ratio)* and the *negative bias (ratio)* concerning alternative i in subpopulation g. For example, the positive bias toward Clinton for a Democrat is represented by

$$\mathcal{B}_d[C] = \frac{\theta_d[C]}{\theta_d[\tilde{C}]}.$$

The invariance condition described by the Invariance of Ratios property and symbolized by Equations 2 and 3 is a consequence of the particular form of the asymptotic probabilities of the states, which we indicate here by three examples, in terms of the function $\mathcal{B}$.

Let us write $\kappa(g)$ for the probability that a randomly sampled respondent is a member of subpopulation g. With $\pi(g, \succ)$ as in Equations 2 to 4, we obtain for the asymptotic probability that a randomly sampled respondent is a Democrat who likes Bush best, Clinton second best, and Perot least:

$$\pi(d, [B \succ C \succ P]) = \frac{\kappa(d)}{Q_d} \mathcal{B}_d[B]\mathcal{B}_d[-P], \tag{6}$$

where Q_d is a normalizing factor. Notice that the functional form of the right side is very suggestive: Equation 6 can be put into words as stating the probability that "the respondent is a Democrat with a positive bias toward Bush and a negative bias against Perot." Similarly, the asymptotic probability of a randomly sampled respondent to be an Independent and to be indifferent between Bush and Clinton, but to prefer both to Perot, is

$$\pi(u, [B \succ P, C \succ P]) = \frac{\kappa(u)}{Q_u} \mathcal{B}_u[-P], \tag{7}$$

with Q_u a normalization factor. In words, the event is as follows: "the respondent is an Independent whose sole bias is negative and against Perot." Finally, the probability that a randomly sampled respondent is a Republican with no biases toward any of the three candidates has the form

$$\pi(r, \varnothing) = \frac{\kappa(r)}{Q_r}, \tag{8}$$

[6] A rule of thumb concerning notation is that the variable g representing the subpopulation is indicated as an index for conditional probabilities or quantities related to such probabilities, and otherwise in parentheses.

with Q_r a normalizing factor. These results are special cases of Theorem 3 in Appendix B. We see in the data analysis that this type of prediction fits the data of the first interview (i.e., shortly before the election) very well. Because the number of parameters is large relative to the number of degrees of freedom in the data, this favorable result may not strike the reader as impressive (despite the large number of respondents).

Another type of result deals with two successive polls separated by a time interval δ and makes more serious demands on the model. This result is too technical to state here (see Theorems 2 and 5 in Appendix B). It gives an exact expression for the joint probability that a randomly sampled respondent belongs to population g and is in state $\succ$ at time t and state $\succ'$ at time $t + \delta$, for large t and for any pair of preference states $\succ$, $\succ'$. This type of result provides a quantitative prediction of the effect of the passage of time on the correlation between successive judgments given by the same individuals.

It should be pointed out that this model is only one of a large class based on similar ideas. In all these models, the successive preference relations of an individual are seen as a realization of a stochastic process. The models differ by the type of tokens considered, by the particular type of preference relations adopted for the states, and by the transformations of the states induced by the tokens (i.e., the operation $\bigcirc$ of this article). Even though all these models are conceptually related, the technical differences between them are not trivial. We chose to present the strict weak order case in view of its successful fit to important data. A different model, centered on semiorders (cf. Falmagne & Doignon, 1997; see also Doignon & Falmagne, 1997) was applied to the same data but proved less successful (see the discussion section).

There are two differences between the model presented here and that in Falmagne et al. (1997). First, the present model does not assume that all respondents begin the process indifferent among the alternatives. This difference is a minor one because the initial distribution on the set of states plays no role in the predictions tested here. (In our data set, the first interview took place late in the campaign.) Second, more importantly, the model of this paper allows for different subpopulations to be exposed to different information and/or to interpret the same information differently.

These modifications are important despite the increase of the number of parameters that they entail because they also increase the explanatory potential of the model.[7] In many important settings, the information delivery varies among subpopulations. Moreover, different subpopulations may start with different prior evaluations of the alternatives and may also interpret the same information from a different perspective. We wanted our model to be capable of accounting for such factors.

We will see that these conjectures concerning the differences between populations were confirmed by the analysis of the data. We now turn to the technical presentation of the model.

Formal Statement of the Model

Basic Concepts

We consider four basic sets. We recall that $\mathcal{A}$ denotes the set of alternatives (i.e., in our application $\mathcal{A} = \{B, C, P\}$). We denote by $\mathcal{G}$ the set of subpopulations (constituencies, partisanships), which are identified with their corresponding channels of information (i.e., in the application $\mathcal{G} = \{d, u, r\}$). Each of the 13 strict weak orders on $\mathcal{A}$ is a possible state of an individual, and we write $\mathcal{S}$ for the set of all such states. The set of tokens (cf. Table 1) is represented by the symbol $\mathcal{T}$, where $\zeta \in \mathcal{T}$ means that ζ is either $[i]$, $[\tilde{i}]$, $[-i]$, or $[\widetilde{-i}]$, for some $i \in \{\mathrm{B, C, P}\}$. The effect of a token ζ on a state $\succ$ is captured by the operation $\succ \bigcirc \zeta = \succ'$ already encountered and defined by the graph of Figure 2.

We suppose that there exists a probability distribution $\kappa : g \mapsto \kappa(g)$, with $\kappa(g) > 0$ on the set of subpopulations. That is, an individual sampled from the population at large will belong to subpopulation g with a probability equal to $\kappa(g)$.

We also assume that the occurrence of the tokens is monitored by two related probabilistic mechanisms: one concerning the times of occurrence of the tokens and the other concerning their nature. The times of occurrence of tokens in channel g, regardless of their nature, are governed by a Poisson process with intensity λ_g (see Axiom T). Thus, if λ_g is large, there will be many tokens delivered to subpopulation g. As to the nature of the tokens, it is controlled by three probability distributions θ_g: $\zeta \mapsto \theta_g(\zeta)$ $(g \in \mathcal{G})$, with $\theta_g > 0$ on the set $\mathcal{T}$ of all tokens: If a token is delivered at any time t in channel g, then this token is equal to ζ with probability $\theta_g(\zeta)$.

The model is cast in terms of random variables. We write:

$\mathbf{G} = g$ to signify that a randomly sampled individual belongs to subpopulation g.

(We recall that party identification is assumed to be constant through the study).

The stochastic part of the model is expressed in terms of several collections of random variables indexed by the time of the event under consideration. We write:

$\mathbf{S}_t = \succ$ to specify that $\succ$ is the state at time $t \geq 0$ of the sampled individual.

The values of the random variables $\mathbf{S}_t$ are governed by other random variables describing the occurrence of the tokens in the three channels. We write:

$\mathbf{N}_{t,t+\delta} = k$ to mean that k tokens have occurred during the (half open) interval of time $]t, t + \delta]$, with $t > 0$. Note that for simplicity, we use the abbreviation $\mathbf{N}_t = \mathbf{N}_{0,t}$.

We also write:

$\mathbf{T}_t = \zeta$ to indicate that ζ was the last token presented before or at time t. We set $\mathbf{T}_t = 0$ if no tokens were presented (i.e., if $\mathbf{N}_t = 0$).

Thus, $\mathbf{S}_t$ takes its values in the set $\mathcal{S}$ of states, each value of $\mathbf{N}_{t,t+\delta}$ is a nonnegative integer, and each value of $\mathbf{T}_t$ is in $\mathcal{T} \cup \{0\}$. It turns out that if we conditionalize by the subpopulation g (in other terms, given $\mathbf{G} = g$), $\mathbf{N}_t$ is the "counting random variable" of a Poisson process for channel g, specifying the number of Poisson events occurring in g during the interval $[0, t]$. The first axiom that follows specifies the probability that a respondent belongs to subpopulation g. The three remaining axioms recur-

[7] From a statistical viewpoint, the increase in the number of parameters goes hand in hand with an increase in the number of degrees of freedom in the data.

sively define, for each channel g, a stochastic process $(\mathbf{N}_t, \mathbf{T}_t, \mathbf{S}_t; g)$. The parameters of the model are the probabilities $\kappa(g)$ of the channels, the densities λ_g of three Poisson processes governing the occurrence of the tokens in the three channels, and the probabilities $\theta_g(\zeta)$ of the tokens ζ in channel g. We denote by $\mathscr{E}_t$ any arbitrarily chosen history of the process before time $t \geq 0$; $\mathscr{E}_0$ stands for the empty history.

Axioms

[G] (Subpopulations). The probability that a sampled individual belongs to subpopulation g in $\mathscr{G}$ is equal to $\mathrm{P}(\mathbf{G} = g) = \kappa(g)$. Thus, $\kappa(g) \geq 0$, $\Sigma_{g\in\mathscr{G}}\, \kappa(g) = 1$. In fact, we assume that $\kappa(g) > 0$.

[I] (Initial state). The initial states of the individuals are governed by some probability distribution (which we leave unspecified). An individual remains in his or her initial state until the occurrence of the first token of information, that is,

$$\sum_{\succ\in\mathscr{S}} \mathrm{P}(\mathbf{S}_0 = \succ) = 1,$$

$$\mathrm{P}(\mathbf{S}_t = \succ | \mathbf{N}_t = 0, \mathbf{S}_0 = \succ, \mathscr{E}_t, \mathbf{G} = g)$$

$$= \mathrm{P}(\mathbf{S}_t = \succ | \mathbf{N}_t = 0, \mathbf{S}_0 = \succ) = 1.$$

[T] (Occurrence of the tokens). The occurrence of the tokens in each channel g is governed by a homogeneous Poisson process of intensity λ_g. When a Poisson event is realized, the token ζ occurs with probability $\theta_g(\zeta)$, regardless of past events. Thus, for any nonnegative integer k, any real numbers $t \geq 0$ and $\delta > 0$, any channel g and any history $\mathscr{E}_t$,

$$\mathrm{P}(\mathbf{N}_{t,t+\delta} = k | \mathscr{E}_t, \mathbf{G} = g) = \frac{(\lambda_g \delta)^k e^{-\lambda_g \delta}}{k!},$$

$$\mathrm{P}(\mathbf{T}_{t+\delta} = \zeta | \mathbf{N}_{t,t+\delta} = 1, \mathscr{E}_t, \mathbf{G} = g)$$

$$= \mathrm{P}(\mathbf{T}_{t+\delta} = \zeta | \mathbf{N}_{t,t+\delta} = 1, \mathbf{G} = g) = \theta_g(\zeta).$$

[L] (Change of state). If an individual in subpopulation g is in state $\succ$ at time t, and a single token ζ arises in channel g between times t and $t + \delta$, then the individual will be in state $\succ \bigcirc \zeta$ at time $t + \delta$, regardless of past events before time t. Formally,

$$\mathrm{P}(\mathbf{S}_{t+\delta} = \succ' | \mathbf{T}_{t+\delta} = \zeta, \mathbf{N}_{t,t+\delta} = 1, \mathbf{S}_t = \succ, \mathscr{E}_t, \mathbf{G} = g)$$

$$= \mathrm{P}(\mathbf{S}_{t+\delta} = \succ' | \mathbf{T}_{t+\delta} = \zeta, \mathbf{N}_{t,t+\delta} = 1, \mathbf{S}_t = \succ)$$

$$= \begin{cases} 1 & \text{if } \succ' = \succ \bigcirc \zeta, \\ 0 & \text{if } \succ' \neq \succ \bigcirc \zeta. \end{cases}$$

The implications of Axioms G, I, T, and L are contained in five theorems that are stated precisely in Appendix B. The intuitive content of Theorems 1, 3, and 4 has been discussed in Sketch of Some Mathematical Results. Theorem 2 describes the transition probabilities between states as a homogeneous Markov process (essentially a random walk on the set of strict weak orders). Theorem 5 provides the key result needed for our analysis, namely, the asymptotic (i.e., for $t \to \infty$) joint probabilities of observing state $\succ$ at time t and state $\succ'$ at time $t + \delta$. Thus, these probabilities are asymptotic in t but not in δ. In practice, δ could be very small (i.e., a few days).[8]

Data Analysis

The data analyzed here are from the 1992 NES (Miller et al., 1993). More specifically, we will use the sampled respondents' self-evaluation on a partisanship scale (strong Democrat, weak Democrat, independent Democrat, independent Independent, independent Republican, weak Republican, or strong Republican) plus their Feeling Thermometer ratings of Bush, Clinton, and Perot at two time points: one shortly before and one shortly after the election. In the so-called Feeling Thermometer, the respondents are asked to rate on a scale ranging from 0 to 100 how "warm" they feel toward the candidates; 50 represents "indifference," 100 is equivalent to "very warm," and 0 is equivalent to "very cold." Thus, each respondent's data consist of a septuple $\langle o, \phi_B, \phi_C, \phi_P, \phi'_B, \phi'_C, \phi'_P\rangle$, where o is their stated political orientation on the 7-point scale and $\phi_i, \phi'_i \in \{0, 1, \ldots, 100\}$ are the Feeling Thermometer ratings of candidate $i \in \{B, C, P\}$ before and after the election, respectively. We call *Democrats* those respondents who rated themselves as strong or weak Democrats, *Republicans* those who rated themselves as strong or weak Republicans, and *Independents* those who rated themselves as independent Democrats, independent Independents, or independent Republicans. (In our analysis, a partition into more than three subpopulations would require a much larger sample.) Those who did not rate themselves on the partisanship scale, approximately 40 of 2,064 respondents, are left out of our analysis.

We indulge in some idealizations by assuming that each of the two interviews was simultaneous for all the respondents.[9] Each person's data vector is recoded in a straightforward fashion as a triple $\langle g, \succ, \succ'\rangle$ representing their political orientation and their pre- and postelection preferences through the assignment

$$g = \begin{cases} d \Leftrightarrow o \in \{\text{strong Democrat, weak Democrat}\} \\ u \Leftrightarrow o \in \{\text{independent Democrat, independent} \\ \qquad\qquad \text{Independent, independent Republican}\} \\ r \Leftrightarrow o \in \{\text{weak Republican, strong Republican}\} \end{cases}$$

$$i \succ j \Leftrightarrow \phi_i > \phi_j,$$
$$i \succ' j \Leftrightarrow \phi'_i > \phi'_j.$$

The raw data are reported in Figures E1, E2, and E3. The set of respondents to the panel are being viewed as a random sample of size $N = 2{,}024$ from the population.

The statistical tests reported in Table 2 support the conclusion that, at the time of the first interview, the distribution of preferences among the electorate has stabilized. In contrast, for the data of the second (postelection) interview, the hypothesis that the

[8] Note that similar predictions can be obtained for small t, at the cost of extra parameters representing some initial distribution on the set of states.

[9] The preelection data were collected almost entirely in September and October, with a few interviews in early November. There were about one-and-one-half times as many interviews in October as in September. The postelection interviews took place almost entirely in November and December, with a few in January. About equal numbers of interviews were conducted in November and December.

Table 2
Likelihood-Ratio Tests of the Asymptotic Model and its Submodels

Likelihood-ratio test	*df*	G^2	*p*
Preelection AM vs. multinomial	18 (38 − 20)	21.6	.25
Postelection AM vs. multinomial	18 (38 − 20)	36.5	.006
Preelection submodel: $\mathcal{B}_g[\zeta]$ invariant with g vs. AM	12 (20 − 8)	950	$<10^{-6}$
Preelection submodel: $\mathcal{B}_g[P]$, $\mathcal{B}_g[-P]$ invariant with g vs. AM	4 (20 − 16)	12	.017
Preelection submodel: $\mathcal{B}_d[P] = \mathcal{B}_r[P]$, $\mathcal{B}_d[-P] = \mathcal{B}_r[-P]$ vs. AM	2 (20 − 18)	5.6	.06
Preelection submodel: $\mathcal{B}_u[B] = \mathcal{B}_u[-B]$, $\mathcal{B}_u[P] = \mathcal{B}_u[-P]$ vs. AM	2 (20 − 18)	2.67	.26

Note. G^2 = Log-likelihood ratio; AM = asymptotic model.

distribution of preferences is at asymptote was rejected (see Table 2). This suggests that the information flow must have changed between the preelection interview and the postelection interview. (In particular, for Perot it is plausible that his controversial reentry in the race initiated a new process.) In turn, this means that the combined data had to be analyzed with a stochastic model involving three political groups and different parameters before and after the election. In fact, such a model, which is consistent with the hypothesis that the preferences were not at asymptote at the second time point, explains the data very well. It also gives some insight into the differences between the information flow to different political subpopulations. The details are given in this section.

Remarks on Statistical Methods

The task is to estimate the underlying parameters of the model, to evaluate whether the data are well explained by the best fitting set of parameters, and to test various hypotheses formulated in terms of the stochastic model. The parameters were estimated through maximum likelihood estimation (MLE) and chi-square minimization. We only report goodness-of-fit statistics and parameter values for the likelihood ratio method. Chi-square tests were performed whenever appropriate and led to similar conclusions. The results of those latter tests and the corresponding parameter estimates are not given here.

As can be seen in Figures E1, E2, and E3, many empirical cells are sparsely populated for the joint data of the two interviews. All cells with a frequency of zero were grouped with other cells according to the following method. Suppose $N(g, \succ, \succ') = 0$. Then the cell $\langle g, \succ, \succ' \rangle$ was pooled with that cell with the highest observed frequency, which differed from $\langle g, \succ, \succ' \rangle$ only by subpopulation or only by the starting state or only by the ending state. With this method of grouping, we keep as many degrees of freedom in the data as possible. As can be easily checked from the three tables, this yields 318 degrees of freedom in the data after grouping. Several quantities had to be numerically approximated in the computer implementation of the statistical test. These are given in Appendix C. Some technical issues with the reliability of the parameter estimation in the full stochastic model are reported and discussed in the Appendix D.

Statistical Tests: The Asymptotic Predictions

We began by testing Equation B1 (see Theorem 3 in Appendix B) on the data for the two interviews separately. This equation, which is referred to as the "asymptotic model," predicts the long-term probabilities $\pi(g, \succ)$ that a sampled individual belongs to a particular subpopulation g and is in state $\succ$. Testing this equation on the first interview data was a natural first step. Equation B1 has $20 = 2 + 2 \times 9$ free parameters for $38 = 39 - 1$ degrees of freedom in the data. The parameters are the $3 - 1$ subpopulation parameters $\kappa(g)$, and the 2×9 bias ratios $\mathcal{B}_g[i]$, $\mathcal{B}_g[-i]$, $g \in \mathcal{G}$, $i \in \mathcal{A}$ entering in the expressions of the $H_{i;g}(\succ)$ and $L_{i;g}(\succ)$ of Equation B1. The respondents are classified into 39 cells: 3 subpopulations and 13 possible rankings. Because the number of respondents is fixed (at 2,024), the number of degrees of freedom in the data is $39 - 1$. The test statistics (likelihood ratio and chi square) are approximately distributed χ^2 with $18 = 38 - 20$ degrees of freedom.

Test of the asymptotic model on the first interview data. We first tested the asymptotic model against the trivial multinomial model, in which each empirical cell has an estimated probability equal to its observed relative frequency. This model has 38 parameters. The outcome of the likelihood ratio test is shown in the first row of Table 2, which contains all the statistical results for the asymptotic model. We see that the model fits the data well, with a significance level of .25.

The effect of the election. We also tested the asymptotic model on the postelection data. As in the case of the preelection data, we tested the asymptotic model against the multinomial model. The results in the second row of Table 2 indicate that the model is sharply rejected by the likelihood ratio statistic. We tentatively concluded that, although the preferences were at asymptote at the time of the first interview, this was no longer the case when the second interviews were collected, either because of the election outcome or for other reasons, such as the reentry of Perot in the race.

This suggests a variation of the model in which the events of the final weeks of the campaign or the election outcome correspond to a new version of the stochastic process, involving new parameters measuring the token probabilities. We tested such a model on the full panel data. The results were successful and are reported later in this article.

Differences between subpopulations. To test whether or not we need three different sets of parameters for the three subpopulations at the first time point, we performed a nested likelihood ratio test of a submodel of the asymptotic model. In this submodel, the bias ratios $\mathcal{B}_g(\zeta) = \theta_g(\zeta)/\theta_g(\tilde{\zeta})$ do not vary with the subpopulation $g \in \mathcal{G}$, reducing to six the number of parameters for the asymptotic probabilities of the states, with eight parameters overall. The hypothesis that the three subpopulations have the same bias ratio is overwhelmingly rejected by the likelihood ratio test (see row 3 of Table 2). This means that the asymptotic probabilities of the preferences vary with the subpopulation. However, because the starting states play no role in the asymptotic distribution, the asymptotic distribution depends only on the information flow. We can, therefore, reason backward and infer that the three groups have been exposed to markedly different information during the campaign or have interpreted the same information differently. This finding is, of course, hardly surprising because we should have expected the Democrats' view of Clinton and Bush to be very different from the Republicans' view of them. However, it is less clear whether we should expect differences between subpopulations in their perceptions of Perot.

The perception of Perot. Political scientists would expect that Perot, as a non-major party candidate, is evaluated differently by respondents classified as Democrats or Republicans on the one hand and by Independents on the other hand. However, it is not clear whether or not Democrats would perceive Perot differently from the Republicans. Actually, it is wellknown in the political science literature that partisan identification was not a good predictor of the support for Perot (Wattenberg, 1994). In any event, we first tested the hypothesis that all three subpopulations perceive Perot in the same way. Specifically, we tested against the asymptotic model, the submodel in which the bias ratios concerning Perot are the same for Democrats and Republicans. This submodel is represented by the two sets of equations:

$$\mathcal{B}_d[P] = \mathcal{B}_r[P] = \mathcal{B}_u[P],$$

and

$$\mathcal{B}_d[-P] = \mathcal{B}_r[-P] = \mathcal{B}_u[-P].$$

This test allows us to reject the hypothesis that Perot is viewed identically by all subpopulations at the .017 level of significance (see row 4 of Table 2).

We then tested the other variant, namely that Perot was perceived essentially the same way by Democrats and Republicans. This hypothesis translates into a submodel specified by the hypotheses

$$\mathcal{B}_d[P] = \mathcal{B}_r[P], \quad \mathcal{B}_d[-P] = \mathcal{B}_r[-P].$$

A nested test of this model against the asymptotic model of row one is reported in row 5 of the table, and yields a significance level of .06. Even though the hypothesis is technically not rejected, this low value suggests that Democrats and Republicans may perceive Perot differently. (These data certainly contribute to the rejection of the model in row 4 of the table.) We postpone for the moment the discussion of row 6.

Negativity. For any channel $g \in \mathcal{G}$ and alternative $i \in \mathcal{A}$, the two bias ratios,

$$\mathcal{B}_g[i] = \frac{\theta_g[i]}{\theta_g[\tilde{i}]}, \quad \text{and} \quad \mathcal{B}_g[-i] = \frac{\theta_g[-i]}{\theta_g[\widetilde{-i}]},$$

can be regarded as measuring, respectively, the net amounts of positive and negative information flowing in that channel for that alternative. The bar graph displayed in Figure 4 gives the values of all the ratios estimated on the preelection data, for the asymptotic model. The values for the positive information are gathered on the left side of the graph and those for the negative information on the right side.

As shown by the estimated values of these ratios, there has been a substantial amount of negative information in the campaign (i.e., information that tends to move candidates to the bottom of the preference ranking). It is clear and not surprising that Democrats perceived more favorable than unfavorable information about

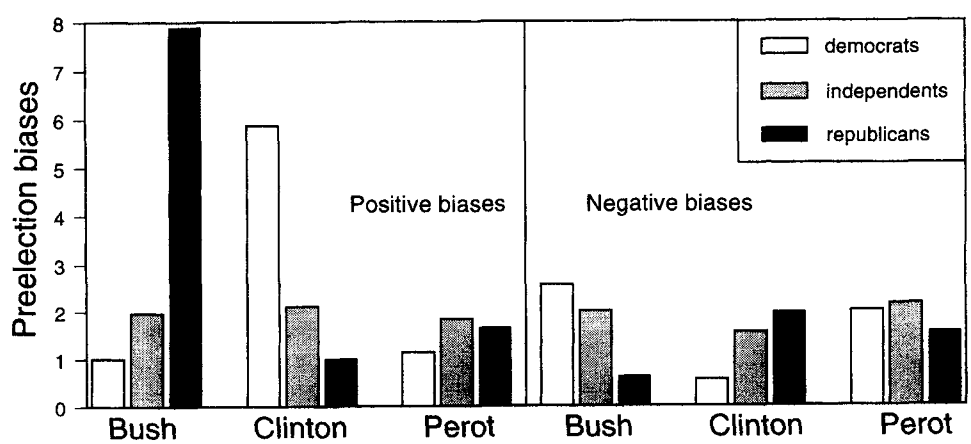


Figure 4. Maximum likelihood parameter estimates of the bias ratios $\mathcal{B}_g(\zeta) = \theta_g(\zeta)/\theta_g(\tilde{\zeta})$ for the first interview. The left half of the graph displays the values of the positive biases $\mathcal{B}_g[i] = \theta_g[i]/\theta_g[\tilde{i}]$; the right half displays the values of the negative biases $\mathcal{B}_g[-i] = \theta_g[-i]/\theta_g[\widetilde{-i}]$.

Clinton and that a similar pattern holds for Bush and the Republicans. The estimates suggest, however, that Independents perceived favorable and unfavorable information in approximately equal amounts, at least as far as Bush and Perot are concerned.

This can be translated into a statistical hypothesis through the submodel specified by the two equalities

$$\mathcal{B}_u[B] = \mathcal{B}_u[-B], \quad \mathcal{B}_u[P] = \mathcal{B}_u[-P].$$

The last row of Table 2 shows that this hypothesis could not be rejected in a nested test against the general asymptotic model. On the other hand, the hypothesis that the Independents received equal amounts of favorable and unfavorable information about each of the three candidates was rejected. The quantitative results are not reported here for reasons of brevity.

Note that our investigation of the negativity in the campaign was performed in terms of the bias ratios $\mathcal{B}_g(\zeta) = \theta_g(\zeta)/\theta_g(\bar{\zeta})$. With the asymptotic predictions for a single interview considered here, the token probabilities themselves cannot be estimated because they appear in the asymptotic formula only through the $\mathcal{B}_g(\zeta)$. This limitation is partly eliminated in the full stochastic model discussed later, which permits the estimation of each of the token probabilities after the election.

Test of the Full Stochastic Model

This model explains the combined data of the two interviews, under the assumptions that the token parameters may change some time after the first interview, with the subpopulation probabilities remaining constant. (Whether or not the Poisson densities vary after the first interview cannot be assessed, because these densities up to the first interview play no role in the predictions; cf. our remark after Theorem 3.) This model has $2 + 3 \times (6 + 11 + 1) = 56$ free parameters ($3 - 1$ parameters for the three subpopulations; and for each subpopulation: 6 parameters explaining the state probabilities before the election, and for the postelection period, $12 - 1$ independent token probabilities and one Poisson density). There are $3 \times 13 \times 13 = 507$ cells in the data, but many of them have a count of zero and grouping is in order (see Remarks on Statistical Methods discussed previously). This leaves us with 318 degrees of freedom for the multinomial model. The test statistics (likelihood ratio and chi square) are approximately distributed χ^2 with $262 = 318 - 56$ degrees of freedom. The results for the likelihood ratio test are in Table 3.

As the first row of Table 3 shows, the likelihood ratio test yields a good fit of the full stochastic model, at a significance level of .384. However, this good fit of the model is obtained through a fairly large number of parameters. This suggests investigating whether an equally good fit could be achieved more economically, that is, in the form of a submodel.

Two submodels of the full stochastic model. We first tested whether the Poisson delivery rates differ significantly among political subpopulations. This submodel is specified by the equations $\lambda_d = \lambda_u = \lambda_r$, which reduce the model to 54 parameters. (Note that these parameters affect only the postelection predictions.) A nested likelihood ratio test of that model against the full stochastic model rejects the hypothesis on the .03 significance level (see Row 2 of Table 3). This suggests that the three groups were exposed to different amounts of information between the two interviews.

We also tested the submodel given in Theorem 5 against the full stochastic model. This submodel has 38 degrees of freedom, and assumes that the token probabilities at the two time points (before and after the election) are the same. The nested likelihood ratio very convincingly rejects the model at the .0001 significance level, as reported in the last row of Table 3. This result confirms our earlier conclusion that the election (or Perot's reentry) had a strong effect on the information flow.

Asymptotic distributions of the states, based on the full data. As emphasized earlier, an important aspect of the model lies in the bias ratios $\mathcal{B}_g[i]$ and $\mathcal{B}_g[-i]$, which govern the asymptotic distribution of the states. Under Statistical Tests: The Asymptotic Predictions discussed previously, the data analyses chiefly relied on those parameters, which were estimated from the (one time point) preelection data. We have compared these estimates with those obtained from fitting the full stochastic model to the joint data of the two interviews. The values of the estimates for the same bias ratios are similar, thus confirming the pattern already revealed by our analysis of the first interview in terms of the asymptotic model.

We recall that, in the full stochastic model, the token probabilities before and after the election are not necessarily the same. In fact, a systematic change in the bias ratios is noticeable. The three graphs in Figure 5 display the estimated values of the preelection and postelection bias ratios. Each graph concerns one candidate. The preelection estimates are on the abscissa and the postelection on the ordinate. The representing points of each graph are white for the Democrats, gray for the Independents, and black for the Republicans. Any point above the diagonal indicates an increased bias ratio (positive for the circles and negative for the squares). It can be seen how the election changed the bias ratios. The most dramatic impact of the election concerns the Republicans' attitude toward Bush (see the black circle and the black square in Bush's graph). The positive bias is much decreased, and the negative bias slightly increased. On the other hand, the attitude toward Clinton is not much changed: All the points are near the diagonal. As for Perot, the positive bias is markedly decreased for the Republicans and to a lesser extend for the Independents. The negative bias is also slightly increased for all subpopulations.

An illustration of the fit of the full stochastic model to the first interview data is given in Figure 6, which displays the graphs of the 13 possible weak orders with, next to each graph, the percentages of such a weak order predicted by the model in the three political categories and (in parentheses) the corresponding observed percentages in the data. The good fit of the model indicates,

Table 3
Tests of the Full Stochastic Model and Two Submodels

	df	G^2	p
FSM vs. panel multinomial	262 ($318 - 56$)	268.2	.384
Submodel: $\lambda_d = \lambda_u = \lambda_r$ vs. FSM	2 ($56 - 54$)	7.01	.03
Submodel of Theorem 5 vs. FSM	18 ($56 - 38$)	47.9	.0001

Note. G^2 = Log-likelihood ratio; FSM = full stochastic model.

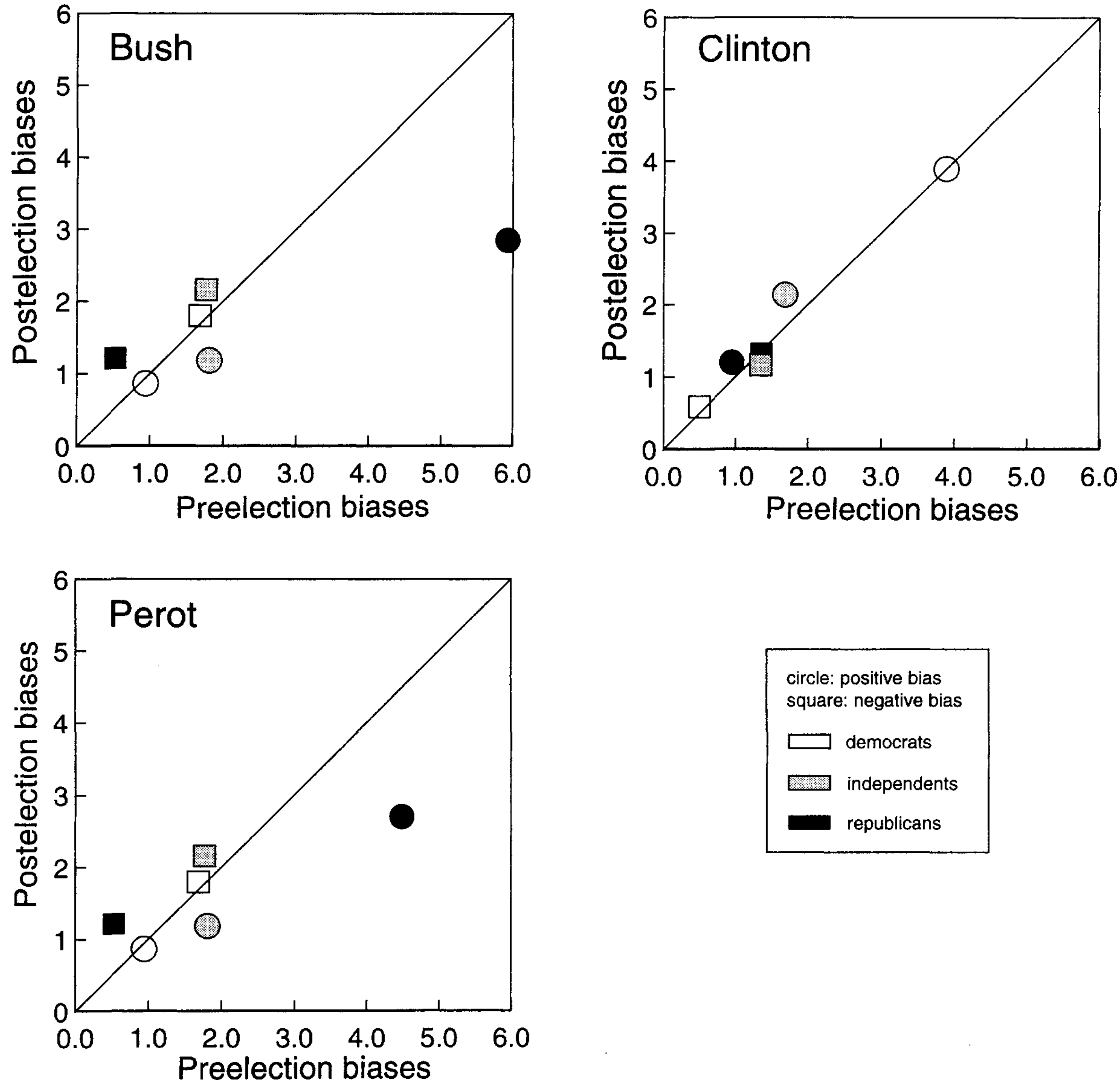
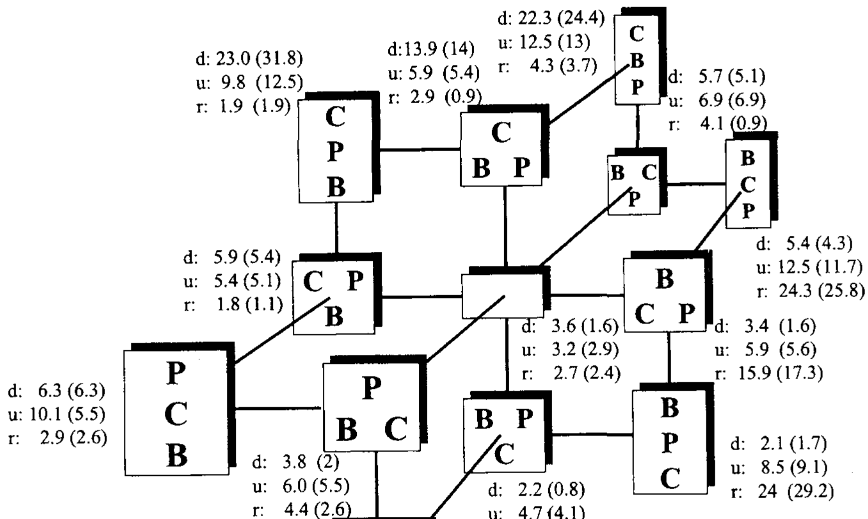


Figure 5. Three graphs representing the estimates of the positive and negative bias ratios $\mathcal{B}_g(\zeta) = \theta_g(\zeta)/\theta_g(\tilde{\zeta})$. Each graph concerns one candidate. The preelection estimates are on the abscissa, and the postelection on the ordinate.

in particular, that the "invariance of ratios" property of Theorem 4, illustrated by Equations 2 and 3 is supported by the data.

Estimated average number of tokens delivered between interviews. Figure 7 contains the estimated average number of tokens of each type—that is, $[i]$, $[\tilde{i}]$, $[-i]$, and $[\widetilde{-i}]$, with $i \in \{B, C, P\}$—delivered to each subpopulation between the two interviews. The estimated number of tokens of type ζ perceived by subpopulation g is obtained from the estimated average number of Poisson events perceived by subpopulation g, weighted by the estimated proportion of tokens of type ζ in that subpopulation. The estimated average numbers of tokens concerning Bush are represented in the top graph of Figure 7: The blank bars refer to the Democrats, the shaded bars to the Independents, and the black bars to the Republicans. The middle and lower graphs are similar, and refer to Clinton and Perot, respectively.

Not surprisingly, Democrats receive more negative information regarding Bush than regarding Clinton, and Republicans receive more negative information regarding Clinton than regarding Bush. The most noticeable result, however, is that the number of tokens concerning Perot, especially of type $[-P]$ and $[\widetilde{-P}]$ for both Democrats and Republicans, is considerably larger than all the other numbers displayed in Figure 7. This is especially remarkable on the background of the analysis of the bias ratios represented in Figure 5. There, the graphs concerning Bush and Perot are very similar. In particular, both graphs display a decrease of the positive bias of the Republicans toward these candidates between the two



	Frequencies	
	Estimated	Observed
Dem.	698	702
Ind.	781	784
Rep.	545	538
Total:	2,024	

Figure 6. Three-dimensional graph of all the possible states. Next to each state, we indicate for each subpopulation (d, u, and r) the estimated preelection percentage of that state based on the full panel data and (in parentheses) the corresponding observed percentage. The table at the lower right indicates the predicted and observed frequencies of the three subpopulations in our sample of 2,024 respondents. Thus, in the framework of the model, 698/2024 ≈ 0.345 is an estimate of sampling (based on our data set) a Democrat in the population of reference. B = Bush; C = Clinton; d = Democrat; Dem. = Democrat; Ind. = Independent; P = Perot; r = Republican; Rep. = Republican; u = Independent.

interviews; this decrease is more extreme for Bush than for Perot. A tentative interpretation of the data of Figures 5 and 7 is that many Republicans received a few negative tokens about Bush, with a mild, but widespread, adverse effect on their rankings of Bush. By contrast, the large numbers of mutually opposite tokens $[-P]$ and $[\widetilde{-P}]$ received by both Democrats and Republicans indicate a wild swing in the opinions regarding Perot, which reveals the unstable image of this candidate in the minds of many voters. A technical point worth mentioning is that, because of some unreliability of the estimates of the densities λ_g of token deliveries, the average number of tokens $[-P]$ and $[\widetilde{-P}]$ perceived by the voters after the election could not be estimated very reliably. Nevertheless, despite this unreliability, the overall picture previously given can be regarded as valid. In particular, the comparatively large number of tokens $[-P]$ and $[\widetilde{-P}]$ used in Figure 7 are conservative estimates. (See Appendix D for details.)

Discussion

We have described a stochastic model of persuasion organized around two fundamental ideas: (a) At any time t, each individual in the relevant population is in a state that can be represented by a strict weak order, that is, a ranking with possible ties; (b) these individuals are subjected to a stochastic stream of elementary messages, which we have called tokens. The characteristics of the stream may vary with the subpopulation to which the individual belongs. We have postulated four types of tokens for each of the alternatives in the choice set: the positive and negative tokens (those tokens that improve or hurt the image of an alternative) and their respective opposites. The potential effect of a token is to change the position of the alternative in the current ranking (see Figure 2). We do not assume that the tokens can be directly observed by the researcher. However, their combined effect can be assessed through a statistical analysis of the data in terms of a model formalizing the two fundamental ideas.

We have developed such a model in the case of three alternatives. The model presented in this article is an extension of a model by Falmagne et al. (1997) and (essentially) a special case of that in Falmagne (1997). Models based on similar ideas can be found in Falmagne (1996) and Falmagne and Doignon (1997). The axioms cast the model as a real-time stochastic process. A basic result of this article is that the succession of states is a random walk on the collection of all weak orders on the set of alternatives. The article spells out quantitative predictions depending on a number of parameters, namely, for each subpopulation the probabilities of the tokens and the density of their overall occurrence. A key concept of the article lies in the positive and negative bias ratios (of token probabilities)

$$\mathcal{B}_g[i] = \frac{\theta_g[i]}{\theta_g[\tilde{i}]} \quad \text{and} \quad \mathcal{B}_g[-i] = \frac{\theta_g[-i]}{\theta_g[\widetilde{-i}]}.$$

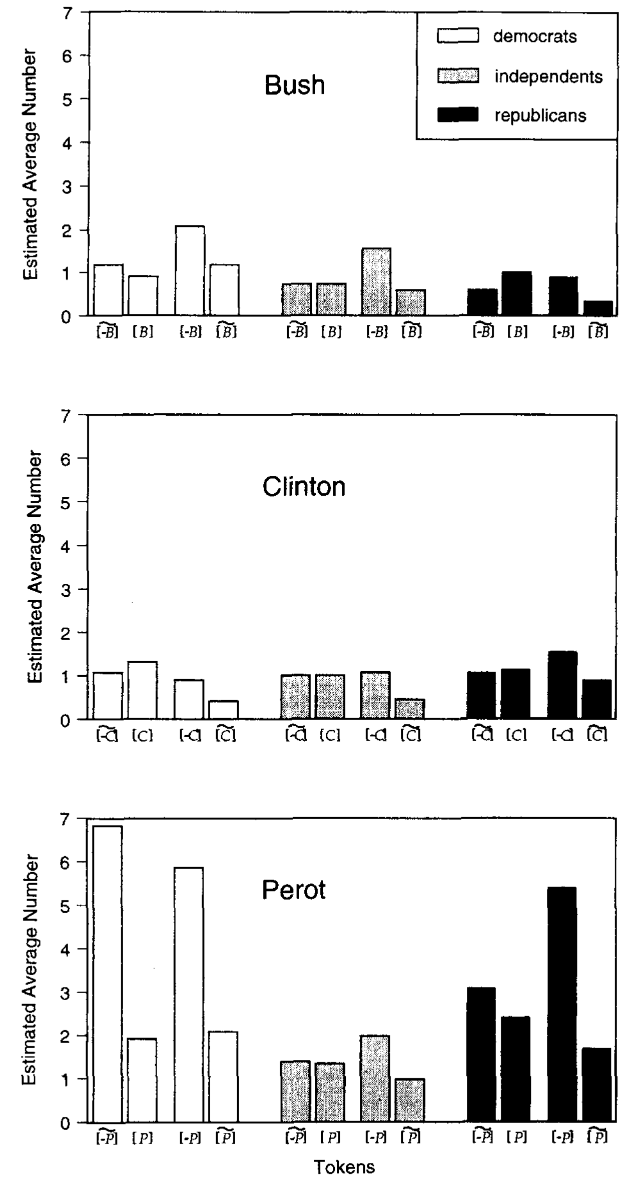


Figure 7. Estimated average number of tokens of each type about the three candidates delivered to the three groups between the interviews based on maximum likelihood estimation. B = Bush; C = Clinton; P = Perot.

This model can be applied to any set of ranking data. As indicated earlier, the only requirement is that the alternatives can be ranked by the respondents according to the dimension under study, such as desirability, personal utility, trustworthiness, or trendiness.

The model has been applied to the 1992 NES data, which have been recoded into two rankings of the three main candidates (Bush, Clinton, and Perot) by each of 2,024 respondents. The two rankings were provided before and after the 1992 election, respectively. On the basis of the information given by the respondents concerning their political affiliations, they were classified into three categories: Democrats, Independents and Republicans. A very good fit of the model was obtained. We have reached the following conclusions.

The preelection preference probabilities are different for respondents with different political orientations. More precisely, the estimates of the bias ratios $\mathcal{B}_g(\zeta)$ vary with party identification (see Table 2). Reasoning backward from this result, we conclude that, during the campaign, the three subpopulations have been exposed to different information or have interpreted differently the information received, which should not come as a surprise. More interesting is the observation that, in the framework of this model, because the asymptotic distribution is unaffected by the initial distribution on the set of states, the information flow alone has to account for these differences.

Negative campaigning appears to have played a major role in the information flow (see Figure 4). However, as expected, for Democrats, favorable information about Clinton outweighed unfavorable information about him and the same holds for Bush and the Republicans.

There was a significant variation in the estimates of the token probabilities between Interview 1 and Interview 2, which may be due to the election results or to the reentry of Perot in the race on October 1. For the same period, the information processes also differed for the three subpopulations. In particular, the token channels of the Democrats, Independents, and Republicans have different token delivery rates, again a result to be expected.

Political scientists expect a "rally-around-the-winner" effect, leading to a "presidential honeymoon" period. This suggests that the president-elect is perceived more positively. We find that Clinton was perceived essentially the same way in the two interviews, but that, on the other hand, there was a shift in the perception of the candidates that led the Republicans to evaluate both Bush and Perot less favorably. During the same period, Democrats and Republicans appear to have been submitted to a barrage of contradicting information about Perot (negative vs. not so negative), revealing an unstable image of this candidate.

Another model, by Falmagne and Doignon (1997), has been tried on the same data but led to a statistically very significant failure. This model uses semiorders (Krantz, Luce, Suppes, & Tversky, 1971; Roberts, 1979) rather than weak orders to represent the preferences. Our interpretation of this failure is that, although the Feeling Thermometer ratings have a natural transformation into strict weak orders (retaining only the order of the ratings), no such transformation is available for the semiorders, because each respondent may have a personal threshold. The transformation we used assumes that the threshold is the same for every respondent. For an analysis in terms of semiorders, ordinal data should be collected, involving three possible responses for each pair of alternatives: i is better than j, j is better than i, neither i nor j is better than the other.

We have made clear in our presentation that the tokens are not regarded as identifiable. Rather, they are hypothetical constructs intended to subsume and formalize the multiple, varied pieces of information that may reach an individual concerning the alternatives but which are inaccessible to the researcher. Nevertheless, steps toward identifying the tokens can be taken in some situations. We mention two examples. The first is a large-scale controlled experiment. The researcher would present the participants with a set of descriptions of candidates for a specified political office (say, president of the United States). The participants would be asked to rank these candidates on the basis of these descriptions. In a second phase of the experiment, the participants would receive a carefully drafted supplement of information, and a second ranking would be required. The researcher would analyze the data using the methods presented here. The number of tokens of each type, for each supplemental description, would then be estimated. Such an analysis may contribute to elucidating the nature of the tokens. It can also serve as a measurement technique in that the effect of each piece of supplemental information can be evaluated by the average number of tokens that it represents. Note that for an analysis as elaborate as ours and involving a control group, a sample size of at least 4,000 would be required.

The second example is a realistic one. The effect of an important event in a political campaign could be evaluated by the model in the following way. For instance, a sufficiently large sample of viewers of a televised debate would be asked to rank the candidates before and after that debate. An analysis in the style of this article could be carried out, which would provide estimates of the average numbers of tokens of each type. These numbers may be helpful in parsing out and measuring the impact of features of the debate on the viewers, offering an appealing alternative to conventional content analysis.

This work can be extended in various directions. The model should be applied to panel data from other elections using similar types of surveys, whether from the U.S. NES or from other countries. As pointed out by a reviewer, an especially promising data set on which to test this model would be the 1980 NES major panel study, which interviewed the same voters four times. This data set is especially interesting because it provides multiple interviews within an ongoing election campaign, which can then be matched against specific events and speeches in the campaign, plus one interview after the election. Other important applications mentioned earlier are consumer choice-preference panel data, for instance brand preference from market surveys or product choice from scanner panel data. Minor alterations of the model of this article would render it appropriate for such data (see the remarks in Falmagne, 1996, pp. 80–81, in this connection).

Interesting theoretical developments are also possible. We mention only one of them. As mentioned earlier, an important issue concerns the effect of the social context on the evolution of preferences (cf. the work of Lazarsfeld et al., 1948, and that of Latané and his colleagues cited in our review of the literature). A possible approach in the framework of this article would be to collect the preferences expressed by couples of individuals living together, say husband and wife. Thus, at times $t_1, \ldots, t_n$, each couple would be asked to provide a pair of rankings $(\succ_h, \succ_w)$ (where the indices refer to "husband" and "wife," respectively). The model could be elaborated to predict the probabilities of the transitions from any pair of rankings $(\succ_h, \succ_w)$ to any other pair $(\succ'_h, \succ'_w)$ under the effect of the tokens. We could then formalize and analyze the possible dependence or independence of the rankings and their transitions within the pairs. We do not, however, pursue these developments here.

A cross-validation on 1996 NES data is in progress. Preliminary results indicate that the model also fits, but suggest different substantive conclusions. The full results will not be available for some time, however.

References

Abelson, R. P. (1964). Mathematical models of the distribution of attitudes under controversy. In N. Frederiksen & H. Gulliksen (Eds.), *Contributions to mathematical psychology* (pp. 142–160). New York: Holt, Rinehart & Winston.

Anderson, N. H. (1971). Integration theory and attitude change. *Econometrica, 78,* 171–206.

Ansalobehere, S., Iyengar, S., Simon, A., & Valentino, N. (1994). Does attack advertising demobilize the electorate? *American Political Science Review, 88,* 829–838.

Converse, P. E. (1964). The nature of belief systems in mass publics. In D. E. Apter (Ed.), *Ideology and discontent* (pp. 207–261). New York: Free Press.

Converse, P. E. (1975). Public opinion and voting behavior. In F. L. Greenstein & N. Polsby (Eds.), *Handbook of political science* (Vol. 4, pp. 75–169). Reading, MA: Addison-Wesley.

Converse, P. E., & Markus, G. B. (1979). Plus ça change . . . : The new CPS Election Study Panel. *American Political Science Review, 73,* 32–49.

Doignon, J.-P., & Falmagne, J.-C. (1997). Well graded families of relations. *Discrete Mathematics, 173,* 35–44.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes.* Fort Worth, TX: Harcourt Brace Jovanovich.

Falmagne, J.-C. (1996). An ergodic theory for the emergence and the evolution of preferences. *Mathematical Social Sciences, 31,* 63–84.

Falmagne, J.-C. (1997). Stochastic token theory. *Journal of Mathematical Psychology, 41,* 129–143.

Falmagne, J.-C., & Doignon, J.-P. (1997). Stochastic evolution of rationality. *Theory and Decision, 43,* 107–138.

Falmagne, J.-C., Regenwetter, M., & Grofman, B. (1997). A stochastic model for the evolution of preferences. In A. A. J. Marley (Ed.), *Choice, decision and measurement: Essays in honor of R. Duncan Luce* (pp. 113–131). Mahwah, NJ: Erlbaum.

Garramone, G. M. (1985). Effect of negative political advertising: The role of sponsor and rebuttal. *Journal of Broadcasting and Electronic Media, 29,* 147–159.

Gegenfurtner, K. R. (1992). PRAXIS: Brent's algorithm for function minimization. *Behavior Research Methods, Instruments & Computers, 24,* 560–564.

Ginsberg, B. (1986). *The captive public: How mass opinion promotes state power.* New York: Basic Books.

Grofman, B. (1985). Models of voting. In S. Long (Ed.), *Micropolitics annual* (pp. 31–64). Greenwich, CT: JAI Press.

Iyengar, S., & Kinder, D. R. (1987). *News that matters: Television and American opinion.* Chicago: University of Chicago Press.

Iyengar, S., Peters, M. D., & Kinder, D. R. (1982). Demonstration of the "not so minimal" consequences of television news. *American Political Science Review, 76,* 848–858.

Jowett, G., & O'Donnell, V. (1992). *Propaganda and persuasion* (2nd ed.). Newbury Park, CA: Sage.

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement* (Vol. 1). San Diego: Academic Press.

Latané, B. (1981). The psychology of social impact. *American Psychologist, 36,* 343–365.

Latané, B., & Nida, S. (1979). Social impact theory and group influence. In P. Paulus (Ed.), *Psychology of group influence* (pp. 3–34). Hillsdale, NJ: Erlbaum.

Latané, B., Nowak, A., & Liu, J. H. (1994). Measuring emergent social phenomena: Dynamism, polarization, and clustering as order parameters of social systems. *Behavioral Science, 39,* 1–24.

Latané, B., & Wolf, S. (1981). The social impact of majorities and minorities. *Psychological Review, 88,* 438–453.

Lazarsfeld, P. L., Berelson, D., & Gaudet, H. (1948). *The peoples' choice: How the voter makes up his mind.* New York: Duell, Sloan and Piarce.

Mackelprang, A. J., Grofman, B., & Thomas, N. K. (1975). Electoral change and stability: Some new perspectives. *American Politics Quarterly, 3,* 315–339.

Markus, G. B. (1982). Political attitudes during an election year: A report on the 1980 NES Panel Study. *American Political Science Review, 76,* 538–560.

Matsunawa, T. (1986). Poisson distribution. In S. Kotz, L. J. Norman, & C. B. Read (Eds.), *Encyclopedia of statistical sciences* (Vol. 7, pp. 23–25). New York: Wiley.

McGuire, W. J. (1964). Inducing resistance to persuasion: Some contemporary approaches. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 1). New York: Academic Press.

McGuire, W. J. (1969). The nature of attitudes and attitude change. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology,* (Vol. 3, pp. 136–314). Reading, MA: Addison-Wesley.

McGuire, W. J. (1985). Attitudes and attitude change. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (Vol. 2, pp. 233–346). New York: Random House.

McPhee, W. N., Andersen, B., & Milholland, H. (1962). Attitude consistency. In W. N. McPhee & W. A. Glaser (Eds.), *Public opinion and congressional elections* (pp. 78–120). New York: Free Press.

Miller, W. E., Kinder, D. R., Rosenstone, S. J., & NES. (1993). *American national election study, 1992: Pre- and post-election survey.* Ann Arbor, MI: Inter-University Consortium for Political and Social Research.

Moscovici, S. (1976). *Social influence and social change.* London: Academic Press.

Moscovici, S. (1985). Social influence and conformity. In G. Lindzey & E. Aronson (Eds.), *The Handbook of social psychology* (Vol. 2, pp. 347–412). Reading, MA: Addison-Wesley.

Norman, F. (1972). *Markov processes and learning models.* New York: Academic Press.

Nowak, A., Szamrej, J., & Latané, B. (1990). From private attitude to public opinion: A dynamic theory of social impact. *Psychological Review, 97,* 362–376.

Parzen, E. (1994). *Stochastic processes.* San Francisco: Holden-Day.

Peizer, D. B., & Pratt, J. W. (1968). The normal approximation for binomial, F, beta, and other common, related tail probabilities. *Journal of the American Statistical Association, 63,* 1417–1456.

Petty, R. E., & Cacioppo, J. T. (1981). *Attitudes and persuasion: Classic and contemporary approaches.* Dubuque, IA: William C. Brown.

Powell, M. J. D. (1964). An efficient method for finding the minimum of a function in several variables without calculating derivatives. *Computer Journal, 7,* 155–162.

Regenwetter, M. (1997). Probabilistic preferences and topset voting. *Mathematical Social Sciences, 34,* 91–105.

Roberts, F. S. (1979). *Measurement theory.* London: Addison-Wesley.

Rodgers, J. L., & Rowe, D. C. (1993). Social contagion and adolescent sexual behavior: A developmental EMOSA model. *Psychological Review, 100,* 479–510.

Skaperdas, S., & Grofman, B. (1995). Modelling negative campaigning. *American Political Science Review, 89,* 49–61.

Wattenberg, M. (1994). *The decline of American political parties* (4th ed.). Cambridge, MA: Harvard University Press.

Zaller, J. (1992). *The nature and origins of mass opinions.* New York: Cambridge University Press.

Zimbardo, P., & Ebbesen, E. E. (1970). *Influencing attitudes and changing behavior.* Reading, MA: Addison-Wesley.

Appendix A

Strict Weak Orders

A relation $\succ$ is a (strict) weak order or a "ranking with ties" on a set $\mathcal{A}$ (cf. Roberts, 1979) if for all i, j, and k in $\mathcal{A}$,

$$\text{If } i \succ j \text{ then } \begin{cases} \text{not } j \succ i \text{ and} \\ \text{either } i \succ k \text{ or } k \succ j \text{ (or both).} \end{cases}$$

Note that the empty relation $\varnothing$ (i.e., the relation containing no ordered pairs) is a strict weak order because the prior condition holds vacuously, (i.e., the case $i \succ j$ never arises). A weak order on a finite set $\mathcal{A}$ induces a numerical order on $\mathcal{A}$ in the sense that there exists for every i in $\mathcal{A}$ a number $u(i)$ such that

$$i \succ j \quad \text{if and only if} \quad u(i) > u(j)$$

(Krantz et al., 1971). Note that the scale u is only defined up to an arbitrary strictly increasing transformation. This equivalence justifies our recoding of the Feeling Thermometer ratings into weak orders.

Appendix B

Basic Theoretical Results

In Falmagne et al. (1997), four theorems were proven for the case of a stochastic process $(\mathbf{S}_t, \mathbf{N}_t, \mathbf{T}_t)$ concerning a single population. Essentially the same theorems are restated here as Theorems 1, 2, 3, and 5 with the additional subdivision into subpopulations. We only state the theorems and do not go into any proofs because they would be repetitious. Theorem 4 is new and deals with the "invariance of ratios" discussed under Sketch of Some Mathematical Results. The first result that follows is basic to all further developments.

Theorem 1

For any subpopulation g, the sequence of states is a random walk on the set $\mathcal{S}$ of all states. The one-step transition probability $p_g^*(\succ, \succ')$ that an individual in population g moves from any state $\succ$ to any state $\succ'$ is given by

$$p_g^*(\succ, \succ') = \begin{cases} \theta_g(\zeta) & \text{if } \succ' = \succ \odot \zeta \\ 0 & \text{otherwise.} \end{cases}$$

A graph of this random walk is pictured in Figure 1. We now give three examples of one-step transition probabilities $p^*_g(>, >')$ for this random walk, all concerning an individual in subpopulation d who happens to be in state [C > B, C > P]:

$$p^*_d([\mathrm{C} > \mathrm{B}, \mathrm{C} > \mathrm{P}], [\mathrm{C} > \mathrm{P} > \mathrm{B}]) = \theta_d[-\mathrm{B}],$$

$$p^*_d([\mathrm{C} > \mathrm{B}, \mathrm{C} > \mathrm{P}], \varnothing) = \theta_d[\bar{\mathrm{C}}],$$

$$p^*_d([\mathrm{C} > \mathrm{B}, \mathrm{C} > \mathrm{P}], [\mathrm{B} > \mathrm{P}, \mathrm{C} > \mathrm{P}]) = 0.$$

We write $p^*_{k;g}(>, >')$ to denote the k-step transition probability that an individual in population g moves from state $>$ to state $>'$ in exactly k steps of the random walk. Using these transition probabilities for the random walk, it is easy to derive the probabilities of moving from state $>$ to state $>'$ in δ units of time. The following result holds.

Theorem 2

For individuals in subpopulation g, the occurrence of the states is a homogeneous Markov process. In this process, the transition probabilities $p_{\delta;g}(>, >')$ of moving from state $>$ to state $>'$ in δ units of time are specified by the equation

$$\underbrace{p_{\delta;g}(>, >')}_{\substack{\text{Prob. in } g \text{ to go}\\ \text{from } > \text{ to } >'\\ \text{in } \delta \text{ units of time}}} = \sum_{k=0}^{\infty} \left(\underbrace{p^*_{k;g}(>, >')}_{\substack{\text{Prob. in } g \text{ to go}\\ \text{from } > \text{ to } >'\\ \text{in } k \text{ steps}}} \times \underbrace{\frac{(\lambda_g \delta)^k e^{-\lambda_g \delta}}{k!}}_{\substack{\text{Prob. in } g \text{ to get}\\ k \text{ tokens within}\\ \delta \text{ units of time}}} \right).$$

The captions should help the reader to parse this equation, which is easy to apply. The only tricky part lies in the k-step transition probabilities $p^*_{k;g}(>, >')$. These are obtained by computing the successive powers of the transition matrix of the random walk. In practice, only a few dozen powers need to be computed before reaching an acceptable approximation to the stationary distribution of the states.

We now turn to the "long-term" predictions, which are relevant whenever it can be assumed that the sample of respondents has been exposed to the flow of tokens for a long time. (This is the case in the data analyzed here.) We first consider the asymptotic probabilities $\pi(g, >)$ of sampling, in the population at large, an individual in subpopulation g and in state $>$. These asymptotic probabilities depend on the probabilities of the tokens only through the bias ratios

$$\mathcal{B}_g(\zeta) = \frac{\theta_g(\zeta)}{\theta_g(\tilde{\zeta})}$$

(cf. Equation 5). We define:

$$H_{i;g}(>) = \begin{cases} \mathcal{B}_g[i] & \text{if } i > j \text{ and } i > k \\ 1 & \text{otherwise,} \end{cases}$$

$$L_{i;g}(>) = \begin{cases} \mathcal{B}_g[-i] & \text{if } j > i \text{ and } k > i \\ 1 & \text{otherwise.} \end{cases}$$

(The choice of letters H and L is meant to evoke a high position and a low position in the ranking.) The next theorem uses this notation.

Theorem 3

The asymptotic probabilities $\pi(g, >)$ that a sampled individual is both in subpopulation g and in state $>$ exist and are specified by the equation

$$\underbrace{\pi(g, >)}_{\substack{\text{Asymptotic prob.}\\ \text{of subpopulation } g\\ \text{and state } >.}} = \underbrace{\kappa(g)}_{\substack{\text{Probability of}\\ \text{subpopulation } g}} \times \underbrace{\frac{\prod_{i \in \mathcal{A}} H_{i;g}(>) L_{i;g}(>)}{\sum_{>' \in \mathcal{S}} \prod_{i \in \mathcal{A}} H_{i;g}(>') L_{i;g}(>')}}_{\substack{\text{asymptotic probability of}\\ \text{state } > \text{ for subpopulation } g}} \tag{B1}$$

Notice that the density λ_g of the Poisson process governing the density of occurrence of the tokens in channel g does not appear in Equation B1. In fact, Theorem 3 would also hold under much more general assumptions concerning the delivery of the tokens. For example, it would be sufficient to require that the occurrence of the tokens be governed by a renewal process.

Three cases of Theorem 3 have been encountered before, namely Equations 6, 7, and 8. We now prove Equation 6. Suppose that $>$ represents the strict weak order $[B > C > P]$. Then,

$$L_{B,d}(>) = H_{C,d}(>) = L_{C,d}(>) = H_{P,d}(>) = 1,$$

$$H_{B,d}(>) = \mathcal{B}_d[B],\ L_{P,d}(>) = \mathcal{B}_d[-P].$$

Writing Q_g for the denominator in Equation B1, we obtain from Theorem 3,

$$\pi(d, [B > C > P]) = \frac{\kappa(d)}{Q_d} H_{B,d}(>) L_{B,d}(>) H_{C,d}(>) L_{C,d}(>) H_{P,d}(>) L_{P,d}(>)$$

$$= \frac{\kappa(d)}{Q_d} \mathcal{B}_d[B] \mathcal{B}_d[-P],$$

that is, Equation 6. Equations 7 and 8 are obtained from similar arguments.

A straightforward consequence of Theorem 3 is the "invariance of ratios" property discussed under Sketch of Some Mathematical Results.

Theorem 4

For any subpopulation g, any state $>$, and any token ζ such that $> \neq > \bigcirc \zeta$, we have

$$\frac{\pi(g, > \bigcirc \zeta)}{\pi(g, >)} = \frac{\theta_g(\zeta)}{\theta_g(\tilde{\zeta})}.$$

Accordingly, this ratio does not depend on the state $>$.

Proof

Because of Equation 5 we need to consider only the cases in which ζ is of the form $[j]$ or $[-j]$. Suppose that $\zeta = [j]$. Take any state $>$ such that $> \neq > \bigcirc \zeta$. Then, by Equation B1 we obtain

$$\frac{\pi(g, > \bigcirc [j])}{\pi(g, >)} = \frac{\kappa(g) Q_g \prod_{i \in \mathcal{A}} H_{i;g}(> \bigcirc [j]) L_{i;g}(> \bigcirc [j])}{\kappa(g) Q_g \prod_{i \in \mathcal{A}} H_{i;g}(>) L_{i;g}(>)}. \tag{B2}$$

However, the only difference between $>$ and $> \bigcirc [j]$ is that alternative j is the unique best in $> \bigcirc [j]$ but not in $>$. Therefore, the right member of Equation B2 simplifies into

$$\frac{H_{j;g}(> \bigcirc [j])}{H_{j;g}(>)} = \mathcal{B}_g[j] = \frac{\theta_g[j]}{\theta_g[\tilde{j}]}.$$

Similarly, for any $>$ with $> \neq > \bigcirc [-j]$,

$$\frac{\pi(g, \succ \bigcirc[-j])}{\pi(g, \succ)} = \frac{L_{j;g}(\succ \bigcirc[-j])}{L_{j;g}(\succ)} = \mathcal{B}_g[-j] = \frac{\theta_g[-j]}{\theta_g[\widetilde{-j}]}.$$

□

Finally, we compute the long-term predictions that are instrumental in the case of two interviews carried out at different times, say t and $t + \delta$, when t can be regarded as large.

Theorem 5

The asymptotic probabilities that a sampled individual is in subpopulation g and in states $\succ$ *and* $\succ'$ *at times t and* $t + \delta$*, respectively, for* $t \to \infty$ *exist and are specified by the equation*

$$\lim_{t\to\infty} p(\mathbf{G} = g, \mathbf{S}_t = \succ, \mathbf{S}_{t+\delta} = \succ') = \pi(g, \succ)p_{\delta;g}(\succ, \succ'), \quad \text{(B3)}$$

where the quantities in the right member are specified by Theorems 3 and 2, respectively.

This follows immediately from Theorem 3 and 2. From probability theory, we know that

$$\begin{aligned}&\mathrm{P}(\mathbf{G} = g, \mathbf{S}_t = \succ, \mathbf{S}_{t+\delta} = \succ')\\ &\quad = \mathrm{P}(\mathbf{G} = g, \mathbf{S}_t = \succ)\mathrm{P}(\mathbf{S}_{t+\delta} = \succ' | \mathbf{G} = g, \mathbf{S}_t = \succ).\end{aligned}$$

Taking limits for $t \to \infty$ on both sides of this equation and using Theorems 3 and 2, Theorem 5 obtains.

In some situations, it makes sense to generalize Theorem 5 by supposing that a critical event has occurred between the times t and $t + \delta$, which may have altered the parameter values of the processes, that is, the token probabilities and the densities of the Poisson processes for some or all subpopulations. In the case of the data analyzed here, the critical events are the election itself and the reentry of Perot into the race, both of which took place between the two interviews and may have drastically changed the flow of information. Theorem 5 generalizes by letting the parameter values used in computing $\pi(g, \succ)$ (before the election) differ from those used in computing $p_{\delta;g}(\succ, \succ')$. This generalization of Theorem 5 is used in the data analysis.

Appendix C

Poisson Approximation

The values $p_{\delta;g}(\succ, \succ')$ are approximated through

$$p_{\delta;g}(\succ, \succ') \approx \sum_{k=0}^{150} \left(p^*_{k;g}(\succ, \succ') \times \frac{(\lambda_g\delta)^k e^{-\lambda_g\delta}}{k!}\right) + \frac{\pi(g, \succ')}{\kappa(g)} \times \sum_{k=151}^{\infty} \frac{(\lambda_g\delta)^k e^{-\lambda_g\delta}}{k!}.$$

We set $\delta = 1$ and computed the right tail of the Poisson distribution by an approximation of its cumulative probability by $\Phi(z)$ with:

$$z = \frac{\sqrt{(2k+1) \times ln\left(\frac{k+\frac{1}{2}}{\lambda_g}\right) + 2 \times \lambda_g - 2k - 1} \times \left(k + \frac{2}{3} - \lambda_g + \frac{0.02}{k+1}\right)}{|k + \frac{1}{2} - \lambda_g|}.$$

This formula was suggested by Peizer and Pratt (1968) and is recommended by Matsunawa (1986). The optimizations for the MLE and minimum total chi square were computed with a conjugate gradient search algorithm by Powell (1964), which is available in form of the C subroutine PRAXIS (Gegenfurtner, 1992).

Appendix D

Reliability and Robustness

An analysis of the reliability of the parameter estimates was carried out, which is too extensive to be given in detail here. We only report the main points.

The likelihood maximization routine was run many hundred times with different starting points to guard against local optima. The final values reported were those obtained in virtually all runs. As mentioned earlier here, parallel estimations were computed using chi-square minimizations and yielded results consistent with those obtained by the maximum likelihood technique. To obtain a more detailed picture of the goodness-of-fit pattern, we also tabulated individual chi-square terms. These are included in Tables E1, E2, and E3. The data in these tables and in Figure 6 give no indication of a systematic deviation. In fact, a plot of the predicted probabilities against the observed relative frequencies practically lies on a straight line.

We also investigated the contribution of the individual parameters to the overall goodness of fit of the model. To this end, we fixed the value of a particular parameter arbitrarily and fitted the model by reestimating the remaining parameters. This procedure was repeated for a range of fixed values around the maximum likelihood estimate of that parameter. The reliability of that estimate was evaluated by the evolution of the goodness-of-fit around the maximum likelihood value. This procedure was used in particular for the parameters λ_g describing the average number of tokens delivered to subpopulation g. These parameters could not be estimated very reliably for Republicans and Democrats. We discovered that a fairly large increase in λ_d or λ_r can be compensated, without affecting the goodness-of-fit, by a decrease in all the token probabilities, except for those of tokens [−P] and $[\widetilde{-\mathrm{P}}]$, which increase. Fortunately, a related statistic can be estimated with good reliability, namely, the average number of tokens of a given type delivered to each constituency. An exception is the average number of tokens [−P] and $[\widetilde{-\mathrm{P}}]$ perceived by the Democrats and the Republicans. This lack of reliability does not affect our conclusions, however, because the MLEs for these numbers are conservative: Any decrease results in a decrease of the quality of the fit.

Appendix E

Row Data

	B P C	B P C	P B C	P B C	P C B	C P B	C P B	C B P	C B P	B C P	B C P	B C P	
B P C	3 2 0	4 1 -2	- -	- -	- -	3 0 -12	- -	- -	- -	1 0 0	- -	1 1 0	- -
B P C	1 1 0	- -	- -	2 1 0	- -	- -	- -	- -	- -	1 0 0	- -	2 0 -1	- -
P B C	2 1 0	- -	1 2 1	1 1 0	1 1 0	- -	1 0 0	- -	- -	- -	1 0 0	- -	- -
P B C	- -	- -	1 1 0	- -	5 3 0	2 2 0	1 3 1	- -	2 2 0	1 1 0	- -	1 0 0	1 1 0
P C B	- -	- -	1 1 0	3 3 0	14 8 -3	11 6 -3	7 7 0	3 2 0	4 2 0	1 2 0	- -	- -	- -
C P B	- -	- -	- -	1 2 0	3 5 1	9 5 -3	13 9 -1	2 3 0	5 4 0	2 1 0	1 1 0	1 0 0	1 1 0
C P B	- -	- -	- -	1 2 0	9 6 -1	7 9 0	138 176 8	33 25 -2	33 30 0	- -	1 1 0	- -	1 3 1
C B P	- -	2 0 -1	- -	1 1 0	5 2 -2	5 4 0	26 27 0	29 18 -5	27 25 0	3 3 0	- -	1 1 0	- -
C B P	- -	1 1 0	- -	- -	7 3 -3	9 5 -1	42 40 0	23 30 2	76 76 0	6 6 0	3 2 0	- -	4 4 0
B C P	- -	- -	2 1 0	- -	1 2 0	1 2 0	5 5 0	7 4 -1	9 6 -1	4 2 0	4 2 0	1 2 0	2 2 0
B C P	4 2 0	1 1 0	- -	- -	- -	- -	- -	2 2 0	5 3 0	3 3 0	7 7 0	6 5 0	2 2 0
B C P	2 1 0	- -	- -	- -	1 0 0	1 1 0	- -	1 1 0	- -	- -	1 4 2	5 3 -1	- -
	- -	- -	2 0 -1	- -	- -	- -	1 3 2	2 2 0	- -	3 1 -1	- -	1 1 0	2 1 0

Figure E1. Data and predictions for the Democrats. The rows are the starting states, and the columns are the ending states. Each entry consists of the empirical frequency (top), the predicted frequency (middle, after pooling and rounded to the closest integer), and individual χ^2 value (bottom, rounded to the closest integer). A minus sign in front of a χ^2 indicates underestimation. Cells marked with a dash had empirical frequency 0 and were pooled. Transitions from a state to itself or to a neighboring state (linked to it by an arrow in Figure 1) are framed by a tiled border. B = Bush; C = Clinton; P = Perot.

	B P C	B P C	P B C	P B C	P C B	C P B	C P B	C B P	C B P	B C P	B C P	B C P	
B P C	26 21 0	3 7 2	6 4 0	- -	- -	2 1 0	1 0 0	- -	2 1 0	4 2 0	10 9 0	14 10 -1	3 2 0
B P C	4 4 0	5 5 0	10 7 0	1 3 1	- -	2 1 0	- -	- -	1 1 0	4 1 -3	1 1 0	3 2 0	1 1 0
P B C	8 4 -2	10 10 0	17 20 0	10 9 0	10 8 0	1 3 2	2 1 0	- -	- -	2 2 0	- -	- -	1 3 1
P B C	3 1 -1	3 3 0	7 7 0	10 6 -2	7 11 2	5 5 0	- -	3 1 -1	- -	2 1 0	1 1 0	1 1 0	1 2 0
P C B	1 0 0	1 2 0	3 5 0	9 8 0	38 30 -2	12 14 0	12 9 0	- -	- -	2 1 0	2 0 -1	2 0 -1	- -
C P B	1 0 0	1 1 0	- -	1 2 1	8 10 0	6 7 0	13 10 0	3 2 0	2 2 0	2 1 0	- -	3 0 -8	- -
C P B	- -	- -	- -	1 1 0	6 5 0	11 8 0	58 71 2	9 10 0	9 9 0	- -	2 0 -3	- -	2 1 0
C B P	- -	- -	- -	- -	2 1 0	3 3 0	13 14 0	10 7 0	10 11 0	- -	1 1 0	- -	3 1 -1
C B P	1 0 0	- -	- -	1 1 0	3 2 0	3 4 0	33 19 -9	15 18 0	37 35 0	5 8 1	3 3 0	1 1 0	- -
B C P	- -	2 2 0	3 1 0	3 2 0	2 2 0	5 3 -1	2 5 1	1 5 3	11 10 0	17 7 -13	3 7 2	2 3 0	3 2 0
B C P	13 11 0	2 3 0	4 1 -2	3 2 0	1 1 0	- -	- -	3 2 0	10 6 -2	11 10 0	32 32 0	11 18 2	2 4 1
B C P	9 8 0	- -	2 1 0	1 1 0	1 1 0	- -	1 1 0	1 1 0	2 2 0	3 3 0	6 11 2	15 8 -5	3 2 0
	1 1 0	2 1 0	1 1 0	2 1 0	- -	3 2 0	2 3 0	3 1 0	- -	3 1 0	1 1 0	1 1 0	4 1 -5

Figure E2. Data and predictions for the Independents. The rows are the starting states, and the columns are the ending states. Each entry consists of the empirical frequency (top), the predicted frequency (middle, after pooling and rounded to the closest integer), and individual χ^2 value (bottom, rounded to the closest integer). A minus sign in front of a χ^2 indicates underestimation. Cells marked with a dash had empirical frequency 0 and were pooled. Transitions from a state to itself or to a neighboring state (linked to it by an arrow in Figure 1) are framed by a tiled border. B = Bush; C = Clinton; P = Perot.

	B P C	B P C	P B C	P B C	P C B	C P B	C P B	C B P	C B P	B C P	B C P	B C P	
B P C	70 105 12	10 7 0	11 7 -2	- -	1 1 0	- -	1 0 0	3 1 -2	- -	6 4 0	33 32 0	21 21 0	1 2 1
B P C	5 3 0	5 2 -1	6 4 0	2 2 0	- -	- -	- -	1 0 0	- -	1 1 0	2 2 0	3 1 -1	- -
P B C	7 4 -1	5 5 0	15 9 -2	3 4 0	2 2 0	1 1 0	- -	1 0 0	- -	1 2 0	4 2 -1	2 1 0	- -
P B C	- -	3 2 0	4 4 0	3 2 0	2 2 0	1 1 0	- -	1 0 0	- -	- -	- -	- -	- -
P C B	- -	- -	- -	- -	6 5 0	2 2 0	2 2 0	- -	- -	- -	- -	- -	- -
C P B	- -	- -	- -	1 0 0	2 2 0	- -	- -	1 0 0	1 0 0	- -	- -	- -	1 0 -1
C P B	- -	- -	- -	2 0 -7	1 1 0	- -	6 3 -1	- -	1 1 0	- -	- -	- -	- -
C B P	- -	- -	- -	- -	- -	1 0 0	1 1 0	2 2 0	- -	1 1 0	- -	- -	- -
C B P	- -	- -	1 0 0	- -	2 0 -2	1 0 0	5 2 -2	2 3 0	7 7 0	- -	1 1 0	- -	1 1 0
B C P	- -	- -	2 1 0	- -	- -	- -	1 1 0	1 1 0	2 3 0	2 2 0	4 3 0	1 2 0	- -
B C P	30 27 0	2 3 0	2 3 0	- -	1 1 0	- -	- -	1 2 0	5 4 0	13 7 -3	56 51 0	28 27 0	1 3 1
B C P	27 21 -1	3 2 0	1 2 0	- -	- -	- -	- -	- -	2 2 0	3 4 0	29 31 0	24 17 -2	4 2 -1
	1 1 0	- -	1 1 0	1 1 0	1 1 0	- -	- -	1 1 0	- -	2 1 0	1 2 0	1 1 0	4 0 -15

Figure E3. Data and predictions for the Republicans. The rows are the starting states, and the columns are the ending states. Each entry consists of the empirical frequency (top), the predicted frequency (middle, after pooling and rounded to the closest integer), and individual χ^2 value (bottom, rounded to the closest integer). A minus sign in front of a χ^2 indicates underestimation. Cells marked with a dash had empirical frequency 0 and were pooled. Transitions from a state to itself or to a neighboring state (linked to it by an arrow in Figure 1) are framed by a tiled border. B = Bush; C = Clinton; P = Perot.

Received November 21, 1995
Revision received May 20, 1998
Accepted September 25, 1998 ■