

A New Solution to the Collective Action Problem: The Paradox of Voter Turnout



Satoshi Kanazawa

American Sociological Review, Vol. 65, No. 3 (Jun., 2000), 433-442.

Stable URL:

<http://links.jstor.org/sici?sici=0003-1224%28200006%2965%3A3%3C433%3AANSTTC%3E2.0.CO%3B2-H>

American Sociological Review is currently published by American Sociological Association.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://uk.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://uk.jstor.org/journals/asa.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

A NEW SOLUTION TO THE COLLECTIVE ACTION PROBLEM: THE PARADOX OF VOTER TURNOUT

SATOSHI KANAZAWA

Indiana University of Pennsylvania

Macy's work offers a potential solution to the paradox of voter turnout. The stochastic learning theory of voter turnout (Kanazawa 1998) posits that citizens perceive a correlation between their behavior (voting versus abstention) and the outcome of collective action (win versus loss for their candidate), and that they interpret the outcome as a reinforcer or a punisher. The theory can solve the paradox of voter turnout because now p , the probability that one's vote is or appears decisive, equals approximately .500 in the calculus-of-voting model (instead of $p \cong 0$). I use General Social Survey data to test the theory. The empirical results indicate that citizens make their turnout decisions according to the "Win-Stay, Lose-Shift" pattern predicted by the stochastic learning theory, especially if there are no strong third-party candidates.

THE paradox of voter turnout is one of the most persistent and recalcitrant empirical puzzles for the rational choice theory of politics. The probability that one would cast a decisive vote is not significantly different from zero in large national elections, and the electoral outcomes are public goods that are enjoyed (or suffered) equally by voters and nonvoters alike. Why then would rational actors invest their personal time and energy into going to the polls to cast their ballots?

Voting in large national elections is an example of a collective action problem (for the provision of public goods). Rational choice theory predicts that actors free ride and do not voluntarily contribute to the production of public goods, yet millions of citizens vote at every election. This paradox of voter turnout is perceived to be so perplexing that Fiorina (1990) calls it "the paradox that ate

rational choice theory" (p. 334). In the most comprehensive and incisive critique to date, Green and Shapiro (1994) choose this paradox as one of the four areas in which rational choice theory has not performed well empirically.

In a previous article, I offered a new perspective on the paradox of voter turnout (Kanazawa 1998), which incorporates elements of stochastic learning into Riker and Ordeshook's (1968, 1973) calculus-of-voting model. The theory proposes that actors are backward-looking and adaptive, rather than forward-looking and utility maximizing. This model simultaneously redefines the " p " term and endogenizes the " D " term in the calculus-of-voting model (see equation 1). It posits that voters become more likely to vote again if they vote for the winning candidate and less likely to vote again if they vote for the losing candidate; abstainers become more likely to vote in the next election if they support the losing candidate and less likely to vote in the next election if they support the winning candidate. The panel data from the 1972–1974–1976 American National Election Study largely supported this reformulation of the calculus-of-voting model based on stochastic learning theory (Kanazawa 1998).

Direct all correspondence to Satoshi Kanazawa, Department of Sociology, Indiana University of Pennsylvania, Indiana, PA 15705-1087 (kanazawa@grove.iup.edu). I thank Morris P. Fiorina and Robert W. Jackman for their comments on early drafts; Hazem Ghobarah, Melvin J. Hinich, and Phillip Paolino for useful discussions; and Annie Gallup for musical and intellectual inspiration.

In the present study, I use pooled data from seven separate waves of the General Social Survey (GSS) to test the stochastic learning theory of voter turnout. The use of the GSS data from a 20-year period (1972–1992) compensates for a major weakness of my 1998 empirical test, which used data from the historically unique early 1970s.

STOCHASTIC LEARNING REFORMULATION OF THE CALCULUS-OF-VOTING MODEL

The calculus-of-voting model is probably the most influential theory of voting behavior. Originally proposed by Downs (1957), and later developed and refined by Riker and Ordeshook (1968, 1973), the model predicts that a citizen will turn out to vote if:

$$pB + D > C, \quad (1)$$

where B represents all the benefits that the voter will personally receive only if the voter's candidate of choice wins the election; D is the "citizen duty" term, which captures all the intrinsic psychic satisfaction the voter receives from the pure act of voting itself, regardless of who wins the election; C is the sum of all the personal costs of voting, both direct costs (the time and energy that one invests before voting in learning about issues and candidates, and that one expends on the day of the election by going to the polls and voting), and opportunity costs (forgone wages, etc.); and p represents the probability that one's vote will be decisive in that it either makes or breaks a tie in one's candidate's favor. The calculus-of-voting model predicts that a citizen will abstain if $pB + D < C$, and will be indifferent if $pB + D = C$.

In any large electorate, however, with millions of potential voters, p is essentially zero, because even the closest contests are decided by a margin of tens of thousands of votes and the probability that one's vote proves to be decisive is infinitesimal. If p is essentially zero, then pB is essentially zero, however large the personal benefits that one receives from the election of one candidate as opposed to the other(s). Thus equation 1 reduces to $D > C$, and the model predicts that a citizen will vote if one's satisfaction from fulfilling the "citizen duty"—performing

one's share to uphold democracy and all the other intangible psychic benefits from the act of voting itself—outweighs all the tangible personal costs of voting. In other words, the decision-theoretic and rational choice model of calculus of voting must explain citizens' decisions to vote in terms of the noninstrumental D term (Aldrich 1993: 257–58; Barry 1970:13–19). The paradox of voter turnout remains unresolved in the rational choice theory of politics.

Two problems emerge from the past literature on the paradox (Kanazawa 1998). First, while some, in their attempt to save the calculus-of-voting model, have argued that D (citizens' sense of duty to vote) is large or that C (personal costs of voting) is small, none has questioned the size of p . It is this unquestioned assumption that $p \cong 0$ that leads to $pB \cong 0$, and then to the necessity of explaining rational actors' decisions to vote by demonstrating that $D > C$. However, if p were not infinitesimal (as has hitherto been assumed), and if p were large, as large as, say, .500, then any modest preference for one candidate over another (B) will lead to a sufficiently large pB , which can easily outweigh C ; thus $pB > C$, obviating the need to explain the decision to vote in terms of the noninstrumental D term. The paradox of voter turnout could potentially be solved if $p \cong .500$.

Second, in Riker and Ordeshook's (1968, 1973) calculus-of-voting model, as well as its later refinements, the D term remains completely exogenous. While it seems reasonable enough to assume that citizens in democratic nations feel certain obligations to vote, where does this sense of obligation come from? Why do some citizens comply with this norm more strictly than others? Where does the individual variation in the size of D come from?

In my earlier study, (Kanazawa 1998), I used Macy's (1989, 1990, 1991a, 1991b, 1993, 1995) work on the stochastic learning model simultaneously to propose a possible solution to the paradox of voter turnout and to endogenize D . The calculus-of-voting model is a decision-theoretic model that conceives of rational actors as subjective expected utility maximizers. Actors in the subjective expected utility maximization theory look forward, evaluate all options available

to them within informational and structural constraints, and assess their potential consequences. They evaluate each consequence of a given contemplated course of action in terms of its utility; weigh it by the subjective probability that a particular consequence will happen; sum across all potential consequences; and derive subjective expected utility for each contemplated course of action. Then rational actors choose the course of action that carries the highest subjective expected utility.

Macy proposes an alternative model of human behavior: the stochastic learning model. This model derives from the earlier work of Bush and Mosteller (1955) on individual behavior; Macy applies their learning model to large-N collective action problems. In this model, actors are backward-looking, adaptive learners, rather than forward-looking utility maximizers. They do not perceive the *causal* link between their contribution and the collective action outcome, but merely the *correlational* links. They interpret the success and the failure of the collective action as reinforcers and punishers from the environment and associate these outcomes with their own behavior. If they contribute and the collective action succeeds, their contribution is reinforced, and (following the principles of operant conditioning) they become somewhat more likely than before to contribute again in the future. If they contribute and the collective action fails, their contribution is punished, and they become somewhat less likely than before to contribute in the future. The same operant logic applies when actors do *not* contribute. If the collective action succeeds their defection is reinforced, and they become even less likely to contribute. If the collective action fails their defection is punished, and they become somewhat more likely to contribute.

In a recent Axelrod-style tournament (a round-robin tournament in which each strategy plays an iterated Prisoner's Dilemma game against every other strategy), Nowak and Sigmund (1993) show that PAVLOV, a game strategy based on the same principle of stochastic learning, typically outperforms one of the most successful of all game strategies: "Tit-For-Tat." PAVLOV's strategy is usefully summarized as "Win-Stay, Lose-Shift." If the actor makes a choice and the

joint outcome of the game results in a "win" from the actor's individual perspective (by receiving either "Temptation" or "Reward" payoff), then the actor stays with the same choice on the next round. If the actor "loses" (by receiving either "Punishment" or "Sucker" payoff), then the actor shifts to the other option (from "Cooperation" to "Defection," or from "Defection" to "Cooperation"). Macy's (1995) laboratory experiment strongly supports the hypothesis that humans behave as backward-looking adaptive learners and engage in the Win-Stay, Lose-Shift pattern.

One of the strengths of Macy's stochastic learning model is that it explains not only instrumental behavior but also normative, habitual behavior as well. In fact, "although learning theory can be used to model consciously instrumental (but backward-looking) behavior, it is typically applied to behavior that is unthinking or habitual" (Macy 1990:811; Macy 1991a:812). The same operant conditioning takes place when actors engage in behavior in order to comply with norms, and the resultant reinforcement or punishment strengthens or weakens their attachment to norms and the extent to which they will comply with them in the future. "The attachment to prosocial norms increases when those who comply are repeatedly rewarded and when those who disregard social obligations and disdain collective welfare are penalized. Conversely, the attachment declines when compliance is penalized and deviance is rewarded" (Macy 1990:811). As in instrumental behavior, normative behavior is reinforced when the collective action to which the actors contribute (through their normative compliance) succeeds, and it is punished when the collective action fails.

However, Macy notes an important distinction between instrumental and normative behavior in his stochastic learning model. While instrumental learning can take place rather quickly, possibly in response to a single reinforcer (successful collective action) or punisher (failed collective action), as in the case of pure Win-Stay, Lose-Shift like PAVLOV, normative learning can take longer because norm compliance tends to be habitual and unthinking. "While pragmatists may change tack after every wind shift, hab-

its are slow to change" (Macy 1991:813). Normative learning thus lags behind instrumental learning, and it takes a larger number of reinforcers and punishers to change normative behavior than it does to change instrumental behavior.

What implications do the stochastic learning model and Win-Stay, Lose-Shift have for the paradox of voter turnout? Voting in a large national election is a quintessential example of collective action. There is clearly winning (success) and losing (failure): Individuals (voters and nonvoters) win if their candidate of choice wins the election; they lose if their candidate loses the election. Thus learning can take place over a series of elections. And Macy's model, applied to Riker and Ordeshook's calculus of voting, can endogenize, and explain individual variations in, both p and D .

P

Macy points out that it is virtually impossible to assess one's marginal contribution to the outcome of a large- N collective action. All that individuals can accurately assess when voting in a large electorate are their own action (voting or not voting) and the collective outcome (win or loss for their candidate of choice). When prospective voters look backward, the p in the calculus of voting no longer measures the probability that their vote will be decisive *in the future*; this will not be a variable in their calculus because there is no way to compute it before the fact. *For the backward-looking adaptive learners, p represents the probability that one's vote was associated with a win in the past.*¹

¹ Grafstein's (1991) evidential decision theory of voter turnout also assumes that voters engage in magical thinking—they believe that other supporters of their candidate of choice will be more likely to vote if they voted (thereby leading to their candidate's victory) and that fellow supporters will be less likely to vote if they didn't vote (thereby leading to their candidate's loss). Grafstein's voters therefore perceive an illusory correlation between their present behavior (voting versus abstention in the current election) and the simultaneous behavior of other voters (*their* voting or abstention). In my model, voters perceive an illusory correlation between their *past* behavior (voting versus abstention in the *last*

Macy's stochastic learning model defines the individual's propensity toward cooperation as follows:

$$\begin{aligned}
 p_{i+1,j} &= p_{ij} + [O_{ij}(1-p_{ij})C_{ij}] \\
 &\quad - [O_{ij}(p_{ij})(1-C_{ij})], \text{ if } O_{ij} > 0; \\
 &= p_{ij} + [O_{ij}(p_{ij})C_{ij}] \\
 &\quad - [O_{ij}(1-p_{ij})(1-C_{ij})], \text{ if } O_{ij} < 0; \quad (2)
 \end{aligned}$$

where p_{ij} is the probability that actor j will contribute toward the collective action on the i th round, O_{ij} is a positive constant if the collective action succeeds on the i th round (and therefore reinforces j 's choice) and a negative constant if it fails (and therefore punishes j 's choice), and $C_{ij} = 1$ if j contributes on the i th round and $C_{ij} = 0$ if j defects on the i th round. $|O_{ij}|$ captures the magnitude of reinforcement or punishment in the learning process.

In my prior study, I proposed to substitute the p term in the original calculus-of-voting model with Macy's $p_{i+1,j}$ (Kanazawa 1998). In other words, p_{ij} is the probability that individual j has voted for a winning candidate prior to the i th election; $O_{ij} > 0$ if j 's preferred candidate wins the i th election and $O_{ij} < 0$ if the candidate loses; and $C_{ij} = 1$ if j votes in the i th election, and $C_{ij} = 0$ if j abstains in the i th election. In a typical presidential election, about half of the voters vote for the winning candidate, and about half of the nonvoters support the winning candidate. Thus, in the stochastic-learning model of voter turnout, on average over a series of elections, $p \cong .500$.²

election) and the past electoral outcomes (Kanazawa 1998). In this sense, Grafstein's voters are still looking forward (albeit diagnostically rather than causally), while my voters are looking backward. Further, my model is a modification of the original calculus-of-voting model, while Grafstein works completely outside of it.

² This, of course, assumes that no single party remains in office election after election and that citizens do not have strong party loyalty. The value of p can be much larger than .500 if one's party of choice keeps winning elections, and it could be much less than .500 if one's party of choice keeps losing elections; p will remain close to .500, however, if the incumbent party alternates or if citizens vote for different parties in different elections.

While Macy's original model of behavior is stochastic and his p_{ij} is the probability of cooperation, my *stochastic learning reformulation of the calculus-of-voting model is still deterministic overall* (as in the original formulation by Riker and Ordeshook [1968, 1973]). The theory predicts that an individual would vote if $pB + D > C$ and abstain otherwise. In my theory, $p_{i+1,j}$ merely substitutes for the p term in the calculus-of-voting model, and does not determine the probability of voting. The stochastic components in my reformulation are the new definitions of the p and D terms.

D

For the first time in the calculus-of-voting literature, the stochastic learning model of voter turnout can endogenize the citizen duty term (D) and parsimoniously explain why some citizens have larger D s than others (without resorting to an ad hoc list of psychic benefits, as in Riker and Ordeshook [1968:28; 1973:63]). The D term fluctuates as a result of the stochastic learning of normative behavior. Citizens' attachment to the norm of civic duty will strengthen if their voting is associated with a successful election of their candidate, and similarly it will decline if their voting is associated with a defeat of their candidate. Conversely, their sense of citizen duty will decline if their candidate wins when they didn't vote, and strengthen if their abstention is associated with the defeat of their candidate. In the stochastic learning model of voter turnout, therefore, the p term (redefined as the past correlation between individual behavior and collective action) and the D term (attachment to the prosocial norm of civic duty) have the same source of variation.

The rates at which p and D respond to collective action outcomes differ, however, because instrumental learning (changes in p) occurs faster than normative learning (changes in D). If citizens vote for a candidate who wins the election, then both the instrumental and normative components of their behavior are reinforced, and they become somewhat more likely to vote again in the next election because their p and D are both larger now than before. If citizens vote for a candidate who loses the election, then

both the instrumental and normative components of their voting behavior are punished. However, the former (p) is quicker to respond to punishment than the latter (D), and while the pragmatist in the voters immediately becomes somewhat less likely to vote in the next election, the normativist in them might still choose to vote out of habit or attachment to the norm of civic duty. It would take a longer series of punishers to stop the normativist in them from voting. After one failed collection action, they become somewhat less likely to vote than had their collective action succeeded (with the candidate of choice winning the election), but not as much as if voting was purely instrumental, with no normative component.

The reverse is true of nonvoters. If citizens abstain and their candidate of choice wins the election, then both the instrumental and normative components in their behavior are reinforced and they become somewhat less likely to vote in the next election because both their p and D are smaller than before. If citizens abstain and their candidate of choice loses the election, then both their instrumental and normative behaviors are punished. However, the former (p) is quicker to respond to the punishment than the latter (D), and while the pragmatist in the citizens immediately becomes somewhat more likely to vote in the next election, the normativist in them lags behind and still might not want to vote because their attachment to the prosocial norm of civic duty has been considerably weakened by a long history of past reinforcement contingency. It would take a longer series of punishers for the normativist in them to (re)build strong attachments to the prosocial norm. After one failed collective action, they become somewhat more likely to vote than had the collective action succeeded without their vote, but not as much as if voting was purely instrumental.

By endogenizing the citizen duty term (D), the stochastic learning reformulation of the calculus-of-voting model contributes toward the discussion of how citizens form their political identities. Someone who votes at every national election obviously has a different political identity from someone who never votes; their conceptions of themselves as citizens of democratic society are quite divergent. How do citizens within the same

society come to form different political identities? The stochastic learning theory of voter turnout suggests that the determinants of political identity might be the environmental reinforcers and punishers in response to their political activities in the past. Citizens form stronger political identities when their past "contributions" are associated with a "success," and they form weaker political identities when their past "contributions" are associated with a "failure." One might further extend this explanation to the formation of other identities and the propensity toward compliance with other norms. The stochastic learning model suggests that individuals are more likely to conform to norms and assume the pertinent identity when environmental responses reward their compliance and punish their noncompliance.

Extending Macy's original model, I define (Kanazawa 1998) the D term as:

$$D_{i+1,j} = D_{ij} + k([O_{ij}(D_{ij})C_{ij}] - [O_{ij}(D_{ij})(1-C_{ij})]), \quad (3)$$

where D_{ij} is individual j 's magnitude of normative attachment to voting prior to the i th election, O_{ij} and C_{ij} are as defined in equation 2, and k varies between 0 and 1 and sets the rate of learning, which is always slower for normative behavior than for instrumental behavior (and hence $k < 1$). Once again, $D_{i+1,j}$ simply substitutes for the D term in the calculus of voting model, and does not determine the individual's propensity to vote by itself (Kanazawa 1998). In my model, a citizen will still turn out to vote only if $pB + D > C$ (using the redefined p and D terms).³

³ Fiorina's (1981) retrospective theory of voter choice maintains that voters largely disregard the promises that candidates make for the future and instead base their choice among candidates mostly on the incumbent's past performance. If the incumbent administration has performed well, then the voters reward the incumbent by voting for him and against challengers. If the incumbent administration has not performed well, then the voters punish the incumbent by voting against him and for a challenger. My theory of voter turnout (Kanazawa 1998) and Fiorina's (1981) theory of voter choice have several features in common. First, both my citizens (contemplating a decision between voting and not voting) and Fiorina's voters (contemplating a decision between voting for the incumbent and voting for a challenger) under-

The logic of the stochastic learning model of voter turnout thus leads to the following complementary hypotheses:

Hypothesis 1: Individuals who vote for the winning candidate at time t_0 will be significantly more likely to vote at time t_1 (vote \times win $\rightarrow + +$).

Hypothesis 2: Individuals who vote for the losing candidate at time t_0 will be less likely to vote at time t_1 than those who vote for the winning candidate, but still more likely to vote than those who do not vote at time t_0 (vote \times lose $\rightarrow +$).

Hypothesis 3: Individuals who do not vote but support a losing candidate at time t_0

take actions whose utilities for the actors depend on the future state of the world. Yet both use past information to make their decisions; the actors in both theories rely on past information as indicative of the future state of the world. My citizens use the past correlation between their own actions and the collective action outcomes to decide whether to vote. Fiorina's voters use the past performance of the incumbent administration to decide whether to vote for the incumbent or a challenger.

Second, in their attempt to use the past information to make their decisions, both my citizens and Fiorina's voters rely on largely illusory correlations. My citizens discern a correlation between their past actions (voting versus not voting) and the outcome of the collective action (win or loss for their candidate of choice). This correlation is largely illusory, in that one person's vote (or abstention) makes virtually no difference for the collective outcome. Fiorina's voters judge the incumbent administration's performance by the state of the economy prior to the election. To the extent that even the president cannot directly and strongly influence the state of the economy (such as unemployment and inflation), and to the extent that other factors beyond the president's control also influence it, voters' perceptions of association between the incumbent administration's performance and the state of the economy in retrospective voting are largely illusory. Nonetheless, Fiorina's voters use the readily available macroeconomic indicators to pass judgment on the incumbent administration because this information is less costly than the careful analyses of the administration's policies and their effects. The complexities of our society (including its polity and economy), coupled with the uncertainty of the future, force both my citizens and Fiorina's voters to use easily available correlational information to make their choices.

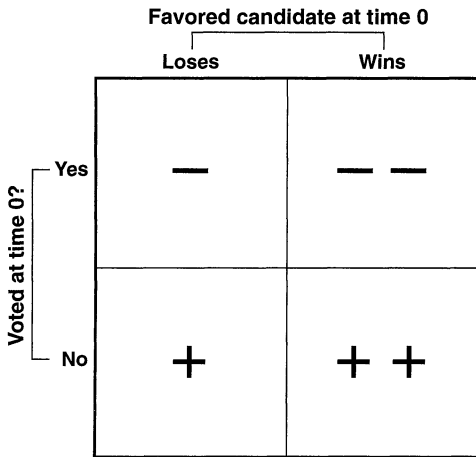


Figure 1. Relative Likelihood of Voting at Time 1

will be more likely to vote at time t_1 than those who do not vote but support a winning candidate, but still less likely to vote than those who vote at time t_0 (nonvote \times lose \rightarrow -).

Hypothesis 4: Individuals who do not vote but support a winning candidate at time t_0 will be significantly less likely to vote at time t_1 (nonvote \times win \rightarrow --).

Figure 1 graphically represents these four hypotheses. If these four hypotheses (and the stochastic learning theory of voter turnout from which they are derived) are true, then the following empirical predictions can be made: (1) The dummy variable measuring whether the respondents voted in the previous election should be positive and highly significant. (2) The dummy variable measuring whether the respondents supported the winner or the loser should not be significant. (3) Most important, the interaction term between the two dummy variables should be positive and significant. In the following empirical analysis, Hypotheses 1 through 4 are tested en masse using these two dummy variables and their interaction term.

EMPIRICAL RESULTS

DATA

I use the General Social Survey (GSS), conducted by the National Opinion Research

Center (NORC) at the University of Chicago, to test my theory of voter turnout. The NORC has administered the GSS either annually or biennially since 1973. Personal interviews are conducted with a nationally representative sample of noninstitutionalized adults 18 and over in the United States. The sample size is about 1,500 for each annual survey, and about 3,000 for each biennial survey.

At each survey, respondents are asked three questions about the last presidential election: whether they voted; if so, who they voted for; if not, who they supported. In some years (1973, 1977, 1982, 1985, 1987, 1989, 1993), however, the GSS asked the respondents the same questions about the last two consecutive presidential elections. For instance, respondents were asked about the 1968 and 1972 presidential elections in the 1973 survey. I pool respondents from these seven surveys for the empirical tests of the stochastic learning theory of voter turnout.

The GSS data overcome one of the primary weaknesses of my earlier empirical test. It used the 1972–1974–1976 panels of the American National Election Study because these were the only panel data available on three consecutive national elections with vote validation. For each election, not only did the interviewers ask the respondents whether they voted, they also went to respondents' precincts to examine the official voter record and validate respondents' responses. While the vote validation allowed me to analyze respondents' true voting behavior (rather than what they claim to have done in their survey responses), it is nevertheless unfortunate that these data came from the early 1970s. Watergate and the Vietnam War mark the 1970s as an unusual period in American political history. It is therefore not an ideal period in which to test a general theory of voter turnout. My use of the GSS data here overcomes this problem because the data come from early 1970s to early 1990s, covering all six presidential elections. If the GSS data support the stochastic learning theory of voter turnout, it would be difficult to attribute the result to any particular historical events or circumstances.

While the GSS data overcome one major problem, however, they introduce another problem. Unlike the 1972–1974–1976 panels of the American National Election Study

data, the GSS does not conduct vote validation (primarily because the GSS is about a wide range of attitudes and behavior, not specifically about political behavior). I therefore must take respondents' verbal responses to voting questions at their face value. Respondents' misreporting of their voting behavior could therefore be a potential problem.

I attempt to minimize the possibility of misreporting as much as possible by increasing the cost of misreporting. As I note above, the GSS asks two separate questions of voters: Whether they voted (to which the respondents can answer "yes" or "no"); and for which presidential candidate they voted (to which the respondents can answer by mentioning the specific candidate or by saying "I don't know" or "I don't remember"). While most respondents who answer "yes" to the first question go on to mention the specific candidates they voted for, some respondents avoid mentioning the candidate's name. I assume that it is more costly psychologically to lie by making up the candidate's name for whom they supposedly voted (when they didn't in fact vote) than by simply responding "yes" to the binary question. I therefore use the respondents' response to the second question as my dependent measure. I count respondents as voting only if they mentioned the candidate for whom they voted; I assume they did not vote if they were vague about this question, even though they responded "yes" to the binary question.

There is some empirical support from the American National Election Study data for the validity of this safeguard against misreporting. For the 1972 presidential election, for instance, among all respondents, only 70.9 percent reported their voting behavior accurately. In contrast, among those respondents who explicitly state which candidate they voted for or supported, 92.3 percent reported their voting behavior accurately. It thus appears that one eliminates some (but not all) misreporting of voting behavior by using the alternative measure. In the absence of vote validation, however, I never truly know if I succeed in reducing the amount of misreporting by this method.⁴

⁴ I hasten to add, however, that very few social science data are validated—respondents can mis-

ANALYSIS OF PRESIDENTIAL ELECTIONS, 1972 TO 1992

Table 1 presents the results of logistic regression analysis using pooled data from seven different GSS surveys covering six consecutive presidential elections. The logistic regression equation contains the stochastic learning predictors (the two dummy variables and their interaction), along with controls for party ID (Democrat or Republican, with independent as the reference category) and standard demographic characteristics (age, race, sex, education, and income).

The coefficients for the stochastic learning predictors support all three predictions: The dummy variable measuring whether the respondents voted in the previous election is strongly positive and significant; the dummy variable measuring whether the respondents supported the winner or the loser is not significant; and, most important, the interaction term between these two dummy variables is positive and marginally significant ($p < .06$). The significance of the interaction term increases to $p < .008$, however, if I exclude the two presidential elections (1980 and 1992) in which there were strong third-party candidates (John B. Anderson in 1980; H. Ross Perot in 1992).

This empirical pattern supports the stochastic learning theory of voter turnout and its hypotheses on the relative likelihood of voting among citizens. The analysis of the GSS data further suggests that the theory holds more strongly if there are no strong third-party candidates. While there are minor third-party candidates in every presidential election (e.g., Libertarian or Green Party candidates), during the study period, only in 1980 and 1992 did a serious third-party candidates run. More important, these are the

report their age or educational attainment. Further, Traugott, Traugott, and Presser (1992) discover that the process of vote validation is highly inaccurate and introduces new errors (for instance, only 57.5 percent of those who claimed to vote in 1988 but were determined by vote validation not to have voted were reconfirmed not to have voted when their voting was revalidated in 1991), and the National Election Studies have discontinued vote validation in 1992. I thank one anonymous reviewer for pointing this out to me.

Table 1. Coefficients from the Regression of the Likelihood of Voting on Selected Independent Variables: Presidential Elections, 1972 to 1992

Independent Variable	All Elections	No Third-Party Candidates
<i>Stochastic Learning Predictors</i>		
Voted in previous election (yes = 1)	2.564*** (.084)	2.401*** (.094)
Supported the winner before (yes = 1)	-.131 (.087)	-.133 (.109)
Interaction	.219 (.115)	.368** (.138)
<i>Party Identification</i>		
Democrat	.432*** (.069)	.455*** (.082)
Republican	.486*** (.079)	.460*** (.097)
<i>Demographic Controls</i>		
Age	.003 (.002)	.004 (.002)
Race (nonblack = 1)	.130 (.080)	.085 (.098)
Sex (male = 1)	-.139* (.058)	-.207** (.068)
Years of education	.060*** (.010)	.069*** (.011)
Income	-.002 (.002)	-.002 (.002)
Constant	-1.782 (.184)	-1.850 (.218)
-2 log likelihood	7,931.99	5,648.74
χ^2 (d.f. = 10)	3,007.82***	1,989.49***
Percentage correctly classified	82.83	82.22
Number of cases	9,404	6,554

Note: Numbers in parentheses are standard errors.

* $p < .05$ ** $p < .01$ *** $p < .001$ (two-tailed tests)

only presidential elections for which the GSS has a separate category for the third-party candidate in the multiple-choice question asking which presidential candidate the respondent voted for or supported. In other words, Anderson and Perot are the only third-party candidates in the 20-year period that the respondents could mention by name (rather than by saying "other").

There appears to be something about presidential elections with strong third-party candidates that disturbs the citizens' stochastic learning process. This is possibly because popular third-party candidates tend to galvanize the electorate and mobilize citizens who might otherwise not vote, much like a booming economy increases the size of the labor force by mobilizing those who otherwise would not work. The stochastic learning theory of voter turnout is inherently binary: Citizens make a binary choice (vote or abstain) and the collective action outcomes are binary (success or failure). The theory therefore might be unsuitable for predicting voter turnout in a trinary choice situation. The true reason why strong third-party candidates present an obstacle for the stochastic learning theory of voter turnout, however, remains unclear.

CONCLUSION

I have provided a further empirical test of my stochastic learning reformulation of the calculus of voting (Kanazawa 1998). The analyses of the GSS data from seven different surveys in two decades and covering six different presidential elections support the hypotheses derived from the stochastic learning theory of voter turnout. Citizens who vote for the winning candidate or abstain but support the losing candidate become somewhat more likely to vote in the subsequent election. Those who vote for the losing candidate or abstain but support the winning candidate become somewhat less likely to vote in the subsequent election. Those who vote for the winner may be more likely to vote again because they feel more efficacious (Acock and Clarke 1990; Mirowsky, Ross, and van Willigen 1996) or because they derive utility from voting for the winner and disutility from voting for the loser (Hinich 1981). A sense of political efficacy or personal satisfaction may underlie the process of stochastic learning in the calculus of voting.

The only obstacle to the theory appears to be popular third-party candidates. The GSS data support the theory more strongly if presidential elections with strong third-party candidates (1980 and 1992) are excluded. The reason for this is not clear and presents

the next challenge for the theory of voter turnout. Whether the stochastic learning theory of voter turnout can successfully account for turnout decisions in presidential elections with popular third-party candidates or such elections must be placed outside the scope of the theory remains to be seen. One way to explore this question might be to test the theory in other political systems (such as in Italy and Germany), where multiple viable candidates routinely compete for political office.

Satoshi Kanazawa is Assistant Professor of Sociology at Indiana University of Pennsylvania. His current work focuses on the introduction of evolutionary psychology into rational choice theory and sociology in general. His recent articles have appeared in Mathematical Social Sciences, Sociological Theory, Journal of Politics, Social Science Research, Journal of Political and Military Sociology (with Debra Friedman), Social Forces (with Mary C. Still), Rationality and Society, and Evolution and Human Behavior. He recently published a book (with Alan S. Miller) titled Order by Accident: The Origins and Consequences of Conformity in Contemporary Japan (Westview Press, 2000).

REFERENCES

- Acock, Alan C. and Harold D. Clarke. 1990. "Alternative Measures of Political Efficacy: Models and Means." *Quality and Quantity* 24:87-105.
- Aldrich, John H. 1993. "Rational Choice and Turnout." *American Journal of Political Science* 37:246-78.
- Barry, Brian. 1970. *Sociologists, Economists and Democracy*. Chicago, IL: University of Chicago Press.
- Bush, Robert R. and Frederick Mosteller. 1955. *Stochastic Models of Learning*. New York: Wiley.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper and Row.
- Fiorina, Morris P. 1981. *Retrospective Voting in American National Elections*. New Haven, CT: Yale University Press.
- . 1990. "Information and Rationality in Elections." Pp. 329-32 in *Information and Democratic Processes*, edited by J. A. Ferejohn and J. H. Kuklinski. Urbana, IL: University of Illinois Press.
- Grafstein, Robert. 1991. "An Evidential Decision Theory of Turnout." *American Journal of Political Science* 35:989-1010.
- Green, Donald P. and Ian Shapiro. 1994. *Pathologies of Rational Choice Theory: A Critique of Applications in Political Science*. New Haven, CT: Yale University Press.
- Hinich, Melvin J. 1981. "Voting as an Act of Contribution." *Public Choice* 36:135-40.
- Kanazawa, Satoshi. 1998. "A Possible Solution to the Paradox of Voter Turnout." *Journal of Politics* 60:974-95.
- Macy, Michael W. 1989. "Walking Out of Social Traps: A Stochastic Learning Model for Prisoner's Dilemma." *Rationality and Society* 1:197-219.
- . 1990. "Learning Theory and the Logic of Critical Mass." *American Sociological Review* 55:809-26.
- . 1991a. "Learning to Cooperate: Stochastic and Tacit Collusion in Social Exchange." *American Journal of Sociology* 97:808-43.
- . 1991b. "Chains of Cooperation: Threshold Effects in Collective Action." *American Sociological Review* 56:730-47.
- . 1993. "Backward-Looking Social Control." *American Sociological Review* 58:819-36.
- . 1995. "PAVLOV and the Evolution of Cooperation: An Experimental Test." *Social Psychological Quarterly* 58:74-87.
- Mirowsky, John, Catherine E. Ross, and Marieke van Willigen. 1996. "Instrumentalism in the Land of Opportunity: Socioeconomic Causes and Emotional Consequences." *Social Psychology Quarterly* 59:322-37.
- Nowak, Martin and Karl Sigmund. 1993. "A Strategy of Win-Stay, Lose-Shift that Outperforms Tit-For-Tat in the Prisoner's Dilemma Game." *Nature* 364(July):56-8.
- Riker, William H. and Peter C. Ordeshook. 1968. "A Theory of the Calculus of Voting." *American Political Science Review* 62:25-42.
- . 1973. *An Introduction to Positive Political Theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Traugott, Michael W., Santa Traugott, and Stanley Presser. 1992. "Revalidation of Self-Reported Vote." Technical Report #42, National Election Studies, Center for Political Studies, University of Michigan, Ann Arbor, Michigan.