Expressive Bayesian Voters, their Turnout Decisions, and Double Probit: Empirical Implications of a Theoretical Model

Christopher H. Achen
Department of Politics and
Center for the Study of Democratic Politics
Princeton University
Princeton, NJ 08544
achen@princeton.edu

July 23, 2006

Abstract

Voting is an expressive act. Since people are not born wanting to express themselves politically, the desire to vote must be acquired, either by learning about the candidates, by using party identification as a cognitive shortcut, or by contact from a trusted source. Modeled as Bayesian updating, this simple explanatory framework has dramatic implications for the understanding of voter turnout. It mathematically implies the main empirical generalizations familiar from the literature, it predicts hitherto unnoticed patterns that appear in turnout data, it provides a better fitting statistical model ("double probit") for sample surveys of turnout, and it allows researchers to forecast turnout patterns in new elections when circumstances change. Thus the case is strengthened for the Bayesian voter model as a central organizing principle for public opinion and voting behavior.

In political life, noneconomic and/or nonegoistic motives seem to be even more important. Self–interest cannot explain even the very basic fact that most people choose to vote....In order to do justice to these empirical facts, in my opinion we need to replace the one–motive theory of purely egoistic economic motivation with less restrictive motivational assumptions.

—Harsanyi (1969, 519)

1   Introduction

Voter turnout in American elections is quite low by international standards. A little more than one half of the adult population participates in presidential elections. Midterm elections draw only about one third of the electorate, and some state–wide primaries and local elections see fewer than 10% at the polls. (A recent review is Patterson 2003, chap. 1.) Weak parties, very frequent elections, long ballots, high residential mobility, the absence of severe economic crises, and the legal disenfranchisement of felons are among the causes (Rosenstone and Hansen 1993; Boyd 1989; Squire *et al.* 1987; McDonald & Popkin 2001).

Low–turnout elections have disturbing consequences. First, they reduce the democratic credentials of the winners. Observers of the European Union speak of a "democratic deficit," due in part to its failure to attract voters to the polls. Similar remarks are often made about American elections.

Second, low–turnout elections tend to be unrepresentative of the populace, skewing resulting policies toward the ideological views and particular interests of active participants (Lijphart 1997; Griffin and Newman 2005; Hajnal and Trounstine 2006). American elec-

torates, for example, considerably over–represent the more prosperous, the better educated, and retirees.

Third and perhaps most important, a disengaged citizenry can no more be neglected than an inactive volcano. As the history of mass democracy has demonstrated repeatedly, a large pool of eligible but inactive citizens, unfamiliar with everyday political choices, may be seduced in times of crisis by demagogues, who then ride into office on a surge of turnout (McPhee and Ferguson 1962; Howard 1971, 232-236; Burnham 1972). Sometimes these officeholders are naive but serviceable, sometimes they are incompetent buffoons, and sometimes, as with Huey Long or, more tragically, Adolf Hitler, they turn out to be actively destructive of democratic government and human rights. In sum, voter turnout is a topic to be taken seriously, both the intellectual challenge of understanding it and the policy problem of raising it.

Because the problem is an old one, and because data have been plentiful, few political activities are as well studied as the decision to vote. Merriam & Gosnell's (1924), Gosnell's (1927), and Tingsten's (1937) pioneering investigations began a vast literature far too extensive to cite comprehensively here.[1] A substantial body of internationally well–tested empirical generalizations have emerged, such as the powerful correlation of information, political engagement, education, and age with the decision to vote. Yet theory relevant to the empirical literature is virtually non–existent, with depressing implications for the state of current knowledge. No one knows for certain how these variables relate causally, and so researchers are free to speculate and interpret liberally.

Consider the impact of age, for example. Blondel *et al.* (1998, 211-213) wonder why growing older increases turnout in every country. Is it a mere "habit of voting" without much cognitive or normative content? Or an acquisition of a social norm, "a sense of obligation to vote," perhaps with no real political comprehension? Or, more impressively, might the effect on the voters be due to learning "the knowledge and skills to enable them to cope with the political world"? Undoubtedly, age has all three effects to some degree. But which impacts are large and which small? No one knows. Wolfinger and Rosenstone (1980, 35-36) ask a parallel set of questions about education, "the surest single predictor of political involvement" in the United States (Campbell, 1962, 20).

The lack of theoretical structure leaves researches free to specify their statistical models arbitrarily, so that even closely related research teams make different choices. For example,

---

[1] Nearly all of the key ideas appeared in those three classics. Milbrath (1965) is an important early review of the literature. Sophisticated recent surveys include Teixeira (1992) for sound coverage and interpretation of empirical findings; and Blais (2000) for a wise theoretical overview. Verba *et al.* (1995) is the most comprehensive study of American political participation generally. Blondel *et al.* (1998) offer a wide-ranging investigation of European turnout in national and European Union elections, making clear what is universal and what is idiosyncratic about American experience. Highton (2004) focuses on the impact of American registration laws. Blais (2006) is a brief summary of recent developments.

Wolfinger and Rosenstone (1980, 124, 128-129) use both education and education squared in their probit equations. Squire, Wolfinger, and Glass (1987) employ only an education–squared term. Rosenstone and Hansen (1993, 273) use only education. These researchers all allow age to have curvilinear (usually quadratic) effects, but other researchers just enter age linearly (Ansolabehere *et al.* 1999). These modeling choices have substantial implications if we really mean them: If age measures learning, for example, it makes a difference whether over a lifetime, political learning accelerates, decelerates, or is constant. Alas, the theory that would provide the interpretation and structure our specifications is missing.

The weakness of our theoretical understanding of turnout can be seen most clearly by comparison with elementary microeconomics. When gasoline prices rise, economists predict that, all else equal, driving will decline in the short run, that the decline will be larger for those with less income or larger vehicles, and that if price increases persist, the market share of smaller and more efficient cars will increase. These predictions are almost always verified. By contrast, if we estimated our probit equations on data from a presidential year and then were asked to forecast which groups of presidential voters would drop out disproportionately in a midterm election, we would have only speculations (some of them contradictory, and many of them wrong, if my recent informal polls of political behavior specialists are any guide). For any serious answer, we would have to ask for a new dataset from the midterm election itself, so that we could begin running our probit equations all over again, no wiser than before, ransacking through the variables to maximize the fit, and congratulating ourselves that we had used maximum likelihood estimation.

The consequences of our theoretical ignorance became clear in the wake of recent American electoral reforms such as "motor voter," which allows citizens to register when they get their drivers licenses, and election day registration, which lets them register right at the polls. In the light of Wolfinger and Rosenstone's state–of–the–art statistical work a quarter century ago, ridding the voters of all restrictive registration laws was expected to improve turnout by nearly 10 percentage points. Actual experience with motor voter has been dramatically less impressive, as many studies, notably Highton and Wolfinger (1998) and Hanmer (2004), confirm. Clearly, there is something about turnout we do not understand.

While the behavioral literature has most often contented itself with correlational findings and informal causal interpretations, the topic of voter turnout has also spawned a large and rigorous theoretical literature, where one might look for explanations of the empirical results and the policy failures. Beginning from Downs (1957, 36-50, 260-276) and Riker & Ordeshook (1968), theorists have assumed that voting should be explained, at least to some degree, as instrumentally rational behavior. Thus voters should appear at the polls if the probability that they will break a tie (or create one) is high enough to outweigh the costs

of voting. Downs famously argued that the purely instrumental utility of voting is:

$$R = pB - c \tag{1}$$

where R is the "return" from voting, p is the probability that the citizen's vote will change the outcome, B is the benefit to the voter of changing the outcome, and $c > 0$ is the cost of voting. Unfortunately, however, since the probability p is negligible in electorates of any size, the inference follows immediately that the return from voting is negative, and hence almost all voters should stay away from the polls. Empirically, too, most studies have found only small or no effects on turnout due to the closeness of the election. (Blais 2000, chap. 3, gives a thorough review of the evidence.) Thus the existence of routinely large turnouts seemed to many scholars in this tradition "the paradox that ate rational choice theory" (Grofman, 1993).

Several *defensores fidei* have stepped forward. Perhaps the voters never catch on, going through life imagining that their one vote might actually change the state or national outcome (Riker & Ordeshook 1968, 38-39). Or perhaps the citizenry use an unusual or boundedly–rational decision rule that happens to make them vote (Ferejohn & Fiorina 1974; Diermeier & Van Mieghem 2000; Bendor *et al.* 2003). Still others noted that if the electorate were much smaller than actual electorates, instrumental voting *then* would be rational (Ledyard 1984; Palfrey & Rosenthal 1983; Palfrey & Rosenthal 1985). These exercises and many others made useful points, contributed to the body of technical tools available, and were widely cited in the instrumental rationality literature.

In addition to the technical contributions, however, one might hope that the formal theories about voter turnout would also be helpful in explaining the behavior of actual citizens. However, the authors' goals were often elsewhere, and most writers in this tradition have shown little interest in the problem of low youth turnout, or in the bias of the American electorate toward the better educated, or in the mobilization by political parties that gets people to the polls. Most of the theorizing overlaps very little with behavioral studies and often evinces little interest in them. In consequence, from Harsanyi onward, few scholars have believed that the instrumental rationality literature has relevance either for understanding actual votes or for informing public policy.

This paper takes up a different view of voter turnout, namely that the decision to vote is done primarily for normative or other non–instrumental reasons. That is, voting is not instrumental in the sense of being directed toward changing the election outcome. Moreover, since voting is not innate, citizens must learn to turn out. When written down formally, these simple ideas generate mathematical implications that map directly onto the behavioral literature and connect closely to what the voters are actually doing. They account for the

familiar empirical generalizations about turnout, and they also produce many new and non–obvious predictions that are confirmed by the evidence. Policy implications arise naturally and powerfully as well. Moreover, the model implies new functional forms for the statistical modeling of voter turnout. The resulting predictions go through the data points, while those from the most widely used statistical specifications in the behavioral literature do not.

Thus modeling voter turnout using the tools of economic reasoning but not the erroneous instrumentalist assumptions gives the benefits of rigor without the loss of empirical verisimilitude. The remainder of this paper shows how this might be done, not with a pretense of complete comprehensiveness, mathematical generality, or perfect empirical harmony, but rather with a relatively simple model that makes a start. The model takes some space to lay out, since nothing shorter has hope of integrating the sprawling behavioral literature. Happily, though, the argument is modular, with the same Bayesian micro–model used repeatedly. Once its logic is grasped, the rest of the argument is meant to be uncomplicated.

Expressive Voting and Social Approval

If the decision to vote is to make rational sense, it must be *expressive*—a decision to do one's duty or take pleasure in a collective enterprise or cheer for one's team without imagining that one might personally determine the outcome of the game (Milbrath,1965, 12-13). In recent years, the notion that people carry out many political acts for expressive reasons has attracted prominent advocates (Brennan and Lomasky 1993; Brennan & Hamlin 2000; and Schuessler 2000). And for voter turnout, expressive motives have probably always been the standard explanation among political scientists (Merriam and Gosnell 1924, chap. 7; Lazarsfeld *et al.* 1944, chap. 5; Berelson *et al.* 1954, 24-29); Teixeira 1992, chap. 5). Riker and Ordeshook themselves added an expressive D term to their equation for the benefit of voting, and while they defined it as the citizen's duty to vote, it can be interpreted to include partisan enthusiasm as well (Fiorina 1976).

Recognizing the power of expressive rationality does not force us into the sterile tautology that people vote because they want to. Expressive voting can be examined theoretically just as instrumental voting has been. Thus we need not abandon rational choice modeling of turnout because some initial attempts have failed, throwing out the baby with the bathwater as some have urged (for example, Green and Shapiro, 1994, chap. 4).

Expressive voting is a version of social conformity. Confident public support of group norms reassures other group members of one's reliable judgment and trustworthiness. Even keeping silent in social gatherings about one's secret ballot may not enable the confused

voter to escape others' doubts about her dependability and suitability as a friend or business associate.[2]

People are remarkably susceptible to these social pressures to conform, as social psychologists have long known (Lane and Sears 1964, 34, 35). The following remark, directed specifically at voter turnout, may be taken as summarizing the findings of a great many experiments and case studies on how groups influence their members:

> The social environment rewards people when they live up to the requirements imposed upon "their kind of people," whether with respect to etiquette or morality or politics; and it punishes them, however subtly, when they do not (Berelson *et al.* 1954, 25; similar remarks appear in Rosenstone and Hansen 1993, 23).[3]

The overwhelming majority of people internalize family and group norms, thinking of them as their own. Internalization is particularly likely to occur in the political realm, since virtually all political information comes from trusted groups, particularly political parties, rather than personal experience. Thus the voter will usually not feel explicit social pressure. "The obligation to vote becomes almost an aspect of the citizen's national identity" (Butler and Stokes 1969: 37).

Spelling out precisely how social groups apply the subtle pressure and how norms are internalized is no simple matter, though progress is visible from several directions (for example, Ansolabehere and Brady 1989; Uhlaner 1989; Bendor & Mookherjee 1990; Hollander 1990; Morton 1991; Kandori 1992; Bernheim 1994; Putnam 2000; Brennan and Pettit 2004). However, the mechanics of identity transmission, norm enforcement, and cultural socialization, important as they are, constitute large and quite challenging research topics that lie outside the focus of this paper. Here it will simply be assumed that expressing solidarity with an attractive candidate during the campaign and at the polls is socialized into people and/or rewarded sufficiently that many people maintain the norm of doing so. Adding Bayesian learning theory then imparts a dynamic over the citizen's life course. The result is a theoretical integration of the behavioral literature and a new statistical specificaiton for turnout research.

The Voter's Uncertainty

People will vote if they see a difference between the candidates and if they have enough confidence that their choice is right for them. That is, people will go to the polls to cheer

---

[2] "The person who appears in public and 'doesn't say much' is a difficult character to understand: 'He's a good fellow, but you never know what he's thinking' " (Vidich and Bensman, 1960, p. 39, describing life in small town "Springdale," New York).

[3] Additional dramatic evidence comes from many communist regimes, which systematically exploited this human vulnerability in "thought–reform" camps, with morally appalling results (for example, Lifton, 1957).

for their side if they know enough to be sure which side they are on.[4]   Formalizing these two variables—perceived candidate difference and the confidence in that judgment—is the task of this section.

We begin with the perceived candidate difference, confining the discussion to a two–party system. In that case, the voter can focus just on the utility difference between the parties, and strategic voting is irrelevant. For a voter who has experienced n prior elections, denote the true value of her difference between the two parties in the next period by u_{n+1}.   Of course, the citizen does not know her future party benefits and must estimate them based on her information set I at time n. The subjective distribution of the voter's expected utility difference between the parties in the following period is taken to have a density denoted by $f(u_{n+1}|I)$. Let the corresponding cumulative distribution function (cdf) be $F(u_{n+1}|I)$, and denote the mean by $\hat{u}_{n+1}$ and the variance by σ_{n+1}^2.   For analytic convenience, the utility difference between the parties, $\hat{u}_{n+1}$, is scaled so that it is always non–negative: Thus positive values favor the party that seems better to the voter, zero represents indifference between the parties, and by definition, negative values never occur.

Now consider the citizen's expressive utility of voting. We begin by setting to zero the pB term in Riker and Ordeshook's formulation, for the reasons given above. We keep their c term. What remains, then, is to replace the D term with a more nuanced and explicitly dynamic treatment of the benefits and costs of expressive voting.

Suppose, then, that the citizen derives utility from voting "correctly," that is, in accordance with her true partisan utility difference in the next period, and that this payoff rises with the perceived importance of the election to the voter, denoted by $\alpha > 0$.   We also suppose that the voter receives disutility proportional to $-\alpha$ from voting "incorrectly," that is, from voting for the party that actually would give her less utility.[5]   Hence the expressive benefit of voting is proportional to a weighted average of these two utilities, where the weights are the subjective probabilities of being right and wrong.

The probability of a correct judgment is $1 - F(0|I)$, the area under the posterior distribution to the right of zero. Since $\hat{u}_{n+1} > 0$, this probability exceeds one half. Similarly, $F(0|I) < \frac{1}{2}$ is the probability of being wrong. Hence the voter's estimated benefit D_{n+1} is:

---

[4]A recent experimental confirmation of the importance of information to turnout is Lassen (2005). The main alternative approaches emphasize citizens' learning the norm of voting (Blais 2000) or, especially in international comparisons, electoral competition and the meaningfulness of the vote (Gosnell 1930; Tingsten 1937, 223-225; Franklin 2004).   It seems difficult in those frameworks to account for the heavy turnout of partisans in uncompetitive districts or for the sudden surges in turnout that accompany charismatic candidacies in difficult times, even when the electoral outcome is not in doubt. Nonetheless, more needs to be done to assess the relative explanatory power of civic norms, competitive elections, and cheering for one's side. A brave attempt to sort out several competing theories is Clarke *et al.* (2004, chaps. 7-8).

[5]If being wrong has a disutility different from the utility of being right, then the double probit model set out below continues to apply, though with a modified interpretation of the parameters δ and γ.

$$E(D_{n+1}) = \alpha[1 - F(0|I)] - \alpha F(0|I)$$
$$= \alpha[1 - 2\,F(0|I)] \tag{2}$$

Thus the utility of voting, U_v, may be written as:

$$U_v = E(D_{n+1}) - c$$
$$= \alpha[1 - 2\,F(0|I)] - c \tag{3}$$

with the interpretation that the citizen turns out to vote if $U_v \geq 0$, and abstains otherwise. If $\hat{u}_{n+1}$ is normally distributed, then the usual transformation to z–scores gives:

$$U_v = \alpha[1 - 2\,F(0|\hat{u}_{n+1}, \sigma^2_{n+1})] - c$$
$$= \alpha[1 - 2\,\Phi(-\hat{u}_{n+1}/\sigma_{n+1})] - c \tag{4}$$

where Φ is the cdf of the standard normal distribution.

This equation incorporates the effects on turnout of both apathy and ignorance. To see this, note first that voters for whom $\hat{u}_{n+1}$ is small care little about the election; they are apathetic. Hence $\Phi(-\hat{u}_{n+1}/\sigma_{n+1}) \approx \Phi(0) = \frac{1}{2}$, so that the right–hand–side of the previous equation becomes negative, and the citizen does not appear at the polls. Second, the less informed the voter, the more difficult she finds discriminating between the parties, and thus the larger is σ_{n+1}. But the larger it is, the more negative is the right–hand–side of the equation, so that abstention becomes more likely. Thus in contrast to the hopes of many reformers, sufficiently unengaged and poorly informed voters will not appear at the polls for most elections, no matter how convenient election registration becomes. Hence the poor results of "motor voter." In short, this framework implies that political systems that leave voters in a state of apathy and ignorance will have low voter turnout, regardless of the convenience of voting, as practical political experience and much prior research have demonstrated.

Long-Term Forces: Learned Partisanship

How then does the successful voter decide which choice she should make? By middle age, most adults vote in consequential elections. They do so primarily because they see that they have a stake in the election at hand. That grasp of politics has been learned.

The central argument of this paper is that the voter learns using three pieces of information—

first, her partisanship, second, her information from the current campaign, and third, contact by a trusted source. To model this process of acquiring a stake in the election, we begin with partisanship.

The citizen is primarily a retrospective thinker where party identification (PID) is concerned, as many empirical researchers have found (Key, 1966; Jackson, 1975; Fiorina, 1981; Franklin and Jackson, 1983). Hence we suppose that each citizen is a Bayesian learner in a two-party system, as in Achen (1992). She begins knowing that the citizenry as a whole has balanced mean partisanship, that is, that Democrats and Republicans are equally well represented in the population, and that the distribution of partisanship over the population is normal (Gaussian). Using the customary Bayesian logic and knowing that the voter herself is a member of that population, her prior is distributed $N(0, \sigma_0^2)$, where σ_0^2 is known.[6] Beginning with this prior at time 0, she then uses her accumulated knowledge at subsequent periods to learn about the parties. At each time period t, she experiences a party benefit differential. That is, she learns retrospectively whether the Democrats or the Republicans have been better for her in the previous period. As an analytic convenience, this utility difference between the parties is assumed distributed independently and normally over time, with known constant variance ω^2 but unknown mean δ, so that $u_t \sim IN(\delta, \omega^2)$ and:

$$u_t = \delta + v_t \tag{5}$$

where v_t is a normally distributed disturbance term with mean zero and variance ω^2.

Thus everyone experiences pleasures and pains from each administration, and different parties benefit different people. However, it is assumed here that ω^2 has no subscript denoting the individual: Experiences with each administration vary over time for everyone, and the variance is the same for the rich and the poor, the wise and the foolish. Thus people of different educational and informational levels update their partisanship at similar rates as they acquire experience. Implicitly, too, we are assuming that people know what their experiences have been. Indeed, one need not have a Ph.D. to notice that the economy was bad under Herbert Hoover, that taxes were cut under Ronald Reagan, or that William Clinton had a mistress. Some such assumption is needed to account for the empirical fact

---

[6]In Achen (2002a), citizens inherited a (weakened) PID from their parents because they anticipated that their own position in the social structure would be somewhat correlated with that of their parents. The standard findings about political socialization were shown to follow from that framework. Here, however, the influence of parents on partisanship is set aside because it would clutter the presentation substantially without adding much insight. In effect, we are dropping a few initial unusually noisy observations from the citizens' updating procedures.

Similarly, the possibility that the electorate as a whole may lean slightly Democratic or Republican, so that the prior mean is not zero, is of a little informational value to the citizen, but very little under normal circumstances in two–party systems, and so it is omitted here. (See Achen, 1992, for a fuller treatment). Multi–party systems would require a more detailed treatment of both parental socialization and party shares.

that many less educated and not particularly well informed individuals are firm partisans who vote regularly.

Formally, the best estimate of the citizen's mean utility difference between the parties, $\hat{\delta}_n$, is interpreted as the voter's current party identification (PID), and she updates in each period according to the Bayesian model corresponding to the estimate of a normal mean with known variance. Thus suppose that u_t is the normally–distributed difference in party benefits she has received during each period t, with those benefits being independently distributed with unknown mean δ and known over–time variance ω^2. Let $\bar{u}_n = \sum u_t / n$. Set $h_1 = n/\omega^2$ and $h_0 = 1/\sigma_0^2 = n_0/\omega^2$, so that the population distribution of PID carries as much information for the voter as n_0 electoral periods. Then the relevant updating theorem is familiar from the early pages of any Bayesian statistics text, and the application to PID here yields:

$$\hat{\delta}_n = \frac{h_1 \bar{u}_n}{h_0 + h_1} \tag{6}$$

This estimate is normally distributed with posterior precision $h_n = h_0 + h_1 = (n_0 + n)/\omega^2$.

This simple framework has been shown to logically imply the main empirical findings of the socialization and party identification literature (Achen 1992, 2002a). The model has been applied and extended by Bartels (1993) and Gerber & Green (1998), and received support in an empirical assessment by Bartels (2002). This model also implies Converse's (1969) empirical finding that on average, partisanship strengthens over the lifecycle, a point that becomes important below.

Short-Term Forces: Learning about Candidates

In the course of a campaign, the voter gets some additional information c_{n+1} beyond her prior party ID about what u_{n+1} is likely to be in the next period:

$$c_{n+1} = u_{n+1} + \theta_{n+1} + \epsilon_{n+1} \tag{7}$$

where θ_{n+1} and ϵ_{n+1} are normally distributed, independently of $\hat{\delta}_n$, u_{n+1} and each other, with mean zero and known variances φ^2 and τ^2/m, respectively, where m is an index of the depth of campaign information. The precision $(\varphi^2 + \tau^2/m)^{-1}$ will be denoted by h_τ. Thus campaign information reduces ϵ_{n+1} but not θ_{n+1}: Even the most well informed citizen cannot forecast future candidate performance perfectly.

Following this logic, since the variation of the next period's benefit around the long–run PID is independent with variance ω^2, we can simply add that variance to the variance in the forecast of the long–run PID to get the error in current PID as a forecast of u_{n+1}. Therefore the current PID has a variance for this purpose equal to the sum of $1/h_n$ and ω^2,

which will be denoted by $1/h_c = [\omega^2/(n_0 + n) + \omega^2] = (n_0 + n + 1)\omega^2/(n_0 + n)$.

Finally, we add one additional source of knowledge about the voter's choice—contact by a trusted source (a political party, interest group, or spouse) who encourages the citizen to vote is treated as a short-cut supply of campaign information (Rosenstone and Hansen 1993, 27).[7] We take this information q_{n+1} to be distributed normally with mean u_{n+1} and known precision h_q. Implicitly, we are assuming that the voter can use the information gleaned from personal contacts to form an unbiased (though perhaps very noisy) estimate of her future experience with the parties.

Thus the citizen's best final estimate of her party difference u_{n+1} is given by combining the PID, campaign information, and personal contact with weights equal to their precisions, according to the same Bayes formula[8]:

$$\hat{u}_{n+1} = \frac{h_c \hat{\delta}_n + h_\tau c_{n+1} + h_q q_{n+1}}{h_c + h_\tau + h_q} \tag{8}$$

This perceived party difference has posterior precision:

$$h_u = h_c + h_\tau + h_q = (n_0 + n)/[(n_0 + n + 1)\omega^2] + (\varphi^2 + \tau^2/m)^{-1} + h_q \tag{9}$$

It follows that $\sigma_{n+1} = (h_c + h_\tau + h_q)^{-1/2}$. Since $\hat{\delta}_n$ and c_{n+1} are each normally distributed and everything else is fixed, $\hat{u}_{n+1}$ is normally distributed.

Finally, substituting into Equation (4) from the previous equation, using the value of σ_{n+1} just derived, and employing the standard result that $\Phi(-z) = 1 - \Phi(z)$ gives the utility of voting as:

$$
\begin{aligned}
U_v &= \alpha \left\{ 1 - 2\,\Phi \left[-\hat{u}_{n+1}(h_c + h_\tau + h_q)^{1/2} \right] \right\} - c \\
&= \alpha \left\{ 2\,\Phi \left[\hat{u}_{n+1}(h_c + h_\tau + h_q)^{1/2} \right] - 1 \right\} - c \tag{10} \\
&= \alpha \left[2\,\Phi \left(\frac{h_c \hat{\delta}_n + h_\tau c_{n+1} + h_q q_{n+1}}{(h_c + h_\tau + h_q)^{1/2}} \right) - 1 \right] - c \tag{11}
\end{aligned}
$$

where as before, the first term on the right is the estimated utility of choosing the better candidate and thus is necessarily non–negative.[9]

---

[7] As Lupia and McCubbins (1998) stress, the citizen must trust the source of information for party contacts to be successful. Hence parties contact primarily their own partisans.

[8] First, $\hat{\delta}_n$ plays the role of prior and c_{n+1} is the new information corresponding to $\bar{x}_n$ in the theorem. Then at the second round, the posterior is updated again in the same way, this time with q_{n+1} as the new information.

[9] A note for methodologists: Since $\hat{u}_{n+1} > 0$, it follows that U_v is concave as a function of $\hat{u}_{n+1}$: $\partial U_v/\partial \hat{u}_{n+1} > 0$ and $\partial^2 U_v/\partial \hat{u}_{n+1}^2 < 0$. When as usual, U_v becomes the argument of a probit function, and if c is not very large, then the probit function will be confined almost entirely to the positive, concave part of its domain. Thus for nearly all voters, the probability of voting will be a concave function Φ of an

Auxiliary Assumptions and Empirical Implications

Connecting the previous equation to the empirical literature requires both some minor additional assumptions and an empirical interpretation of the parameters:

- Age or "systemtime" (= age – 18 years) proxies for political experience with the parties, and thus for strength of partisanship $\hat{\delta}_n$ (Converse 1969).[10] The theory implies that age should enter linearly, not with the powerful quadratic term seen in current specifications.

- Education is treated as a measure of "intellectual capital" (both training and native endowments), so that it measures how easily the citizen acquires campaign information c_{n+1} (Delli Carpini and Keeter 1996, 182-184; 188-199). Other effects of education are treated as minor and are set aside.[11]

- Self-assessed "interest in the election" is regarded as the respondent's assessment of the social importance of the election α.

- Self-assessed "caring about the outcome" proxies for the respondent's judgment of the utility difference between the candidates $\hat{u}_{n+1}$.

- All aspects of ease of voting—Sunday voting, longer poll hours, convenient polling locations, availability of postal/absentee voting, and so on—are treated as reducing the cost of voting c (Wolfinger and Rosenstone 1980, chap. 4).

- Low–salience elections will be defined as those with a lower flow of campaign information.

increasing concave function U_v, and hence concave. Hence in a homogeneous sample with no countervailing effects, when turnout is graphed against explanatory factors we will see something resembling the concave upper half of the customary probit ogive (S–shape), not the convex lower half. This accounts for Glaser's (1962, 20-21) intriguing but puzzling empirical finding that half of the usual probit shape was missing in his plots of turnout relationships.

Of course, if we model voter turnout in a two–party system as an ordered trichotomy—voted left party, abstained, voted right party—which is also consistent with the theory, then the full probit shape will return (Sanders 2001, following Downs 1957, chap. 14).

[10]Other interpretations of the effects of age, such as the recurrent notion that young people do not vote because they are busy with establishing families and careers, not only fail to explain why turnout continues to rise through late middle age, but also do not correctly predict turnout differences among young people (Wolfinger and Rosenstone 1980, 55-58; Highton and Wolfinger 2001). Similarly, the idea that well educated young citizens do not vote because they have not learned "civic skills" about the process of casting a vote encounters related difficulties, notably in explaining why young voters who successfully turn out for presidential elections undergo a dramatic loss of civic skills every two years during midterm elections.

[11]Many authors have speculated that the young and poorly educated may find getting to the polls difficult, and that education predicts voting for that reason. Yet they are not surprised when young men of modest educational attainment drive several hundred miles to an unfamiliar football stadium and arrive on time and suitably provisioned, all without difficulty.

For simplicity in the derivations, we assume that the behavioral measures are simple linear functions of the underlying concept. However all but the last three results hold under weaker conditions such as positive monotonicity (of age on partisan strength, for example) or diminishing marginal productivity (of education on knowledge, for example).

2　Derivations from the Model

Now to assess the impact of any factor included in the model, we simply take partial derivatives of U_v with respect to the corresponding factor. Note that we do not focus on the probability of turning out to vote itself, namely $\Pr(U_v > 0)$, since it is subject to floor and ceiling artifacts (because probabilities cannot exceed one nor fall below zero). Instead we work on the probit scale U_v, just as we would in carrying out a standard probit analysis of a voter survey.

We have from elementary calculus the following well known results, which have been verified repeatedly in the U.S. and elsewhere (along with references to early and/or prominent studies in which they appeared).[12]

Proposition 1. Turnout is generally higher in elections for more important offices and in more consequential or polarized times (Holls 1889, 589-590; Bean 1948, 31-49; Key 1964 [1942], chap. 21; Campbell 1966; Hetherington 2001): $\partial U_v / \partial \alpha > 0$.

Proposition 2. Those interested in the election are more likely to vote (Merriam and Gosnell 1924, 159; Lazarsfeld *et al.* 1944, 46; Berelson *et al.* 1954, 24-33): $\partial U_v / \partial \alpha > 0$.

Proposition 3. The educated generally vote more than the less educated (Arneson 1925; Lazarsfeld *et al.* 1944, chap. 5): $\partial U_v / \partial h_\tau > 0$.

Proposition 4. The better informed are more likely to vote, whether information is measured directly (Merriam and Gosnell 1924, 183-194; Hastings 1956; Delli Carpini and Keeter 1996, 224-27) or indirectly, as media consumption or political discussion (Berelson *et al.* 1954, 118-122, 248-251; Glaser 1962, 23-28): $\partial U_v / \partial h_\tau > 0$.

Proposition 5. Older citizens vote more than the young (Arneson 1925; Tingsten 1937, chap. 2): $\partial U_v / \partial h_n > 0$.

---

[12]In effect, we are taking partial derivatives of expected outcomes on the probit scale, holding all else constant. Thus the propositions say what usually happens, all else equal, not what always happens. Second, converting the first–order partials to turnout *probabilities* is straightforward. Some qualifications enter with the second–order partials due to floor and ceiling effects.

Proposition 6. Stronger partisans are more likely to vote than weak partisans or Independents (Campbell *et al.* 1960, 96-101): $\partial U_v / \partial h_n > 0$.

Proposition 7. Those contacted by a trusted source are more likely to vote (Gosnell 1927; Berelson 1954, 176; Rosenstone and Hansen 1993; Goldstein 1999): $\partial U_v / \partial h_q > 0$.

Proposition 8. Those who care more about the outcome of the election are more likely to vote (Merriam and Gosnell 1924, 167-168; Glaser 1962, 27): $\partial U_v / \partial \hat{u}_{n+1} > 0$.

Proposition 9. Turnout is higher when voting is convenient or administratively easy (Merriam and Gosnell 1924, 95-102; Kelley *et al.* 1967; Wolfinger and Rosenstone 1980, chap. 4; Brady and McNulty 2004): $\partial U_v / \partial c < 0$.

Thus the first validation of the model is that it reproduces what we already know descriptively, and that it organizes that knowledge theoretically.[13] However, the model also implies a good many other, less–familiar and more interesting propositions:

Proposition 10. If the costs of voting are small, non-voters will consist of those who do not care very much about the election, or those who care but have little information, or, in smaller numbers, of well-informed, caring citizens who find the candidates equally appealing (Lazarsfeld *et al.* 1944, chaps. 5 and 6): From Equation (10), $U_v < c \Rightarrow$ either $\alpha \approx 0$, or $(h_c + h_\tau + h_q) \approx 0$, or else $\hat{u}_{n+1} \approx 0$. .

Proposition 11. Those whose turnout rates drop the most from high–salience to low–salience elections will be disproportionately young, newly enfranchised, or otherwise weak in their partisanship (Tingsten 1937, 230; Glaser 1962, 35-44): $\partial^2 U_v / \partial h_n \partial h_\tau < 0$.

Proposition 12. Young people who begin voting will vote more often in the future, even though there is no causal relationship between the two events, thus inducing a pseudo–habit–formation feature to the data which may have mislead scholars (Plutzer 2002; Gerber *et al.* 2003): $\partial U_v / \partial h_n > 0$.

Proposition 13. Contact by a trusted source will have a greater effect on the less engaged and less informed (Berelson *et al.* 1954, 174-177; Glaser 1962, 34): $\partial^2 U_v / \partial h_q \partial h_\tau < 0$. Thus there is no number that is "the effect" of contacting voters; everything depends on the population contacted.

---

[13]In an earlier version of this paper, another Bayesian module explained how the voter learns the importance of the election from the information provided by her social ties. It follows that those who are better connected socially (married or church attenders) would vote more often, which is a familiar empirical result. That part of the model is not needed for the empirical work below, and so it is omitted to save space. The impact of social ties on turnout deserves further theoretical study.

Still other results of an interactive character, most of them not familiar from the literature but confirmed in survey data, are implied by the model:

Proposition 14. Party identification (or age) has greater effects on the turnout of poorly informed (less educated) citizens: $\partial^2 U_v / \partial h_n \partial h_\tau < 0$.

Proposition 15. Information (or education) has a greater effect on the voting rates of less partisan (or younger) citizens: $\partial^2 \mathrm{E}_t(U_v) / \partial h_\tau \partial h_c < 0$.

Note that Propositions 13-15 are second–order derivatives—quantities that should be zero on the probit scale if our standard probit and logit specifications with no interaction terms are correct.[14] Thus the theory implies that our current statistical models for turnout are incorrect and may induce biased estimates.

Theory Testing in Turnout Data

This section shows that patient theoretical work has important consequences for empirical modeling. We focus primarily on the 1998 and 2000 Current Population Surveys (CPS) of voter turnout by the American population, conducted by the Census Bureau. These large studies contain nearly 100,000 observations at each biennial time period. They also contain a wealth of demographic information about each respondent, though no attitudinal and vote choice questions. In consequence, we study the two key theoretical variables, partisanship and information, proxied by age and education. We also discuss briefly the Annenberg 2000 presidential election study conducted by the University of Pennsylvania. For greater details and international comparative analyses, see Achen and Sinnott (2007).

Figure 1 displays what is to be explained—American voter turnout by age and education for the 2000 election. With a few exceptions, such as the lightly–populated categories at the bottom left (young people with very little education), each of the points in these graphs represents several thousand citizens, so that their turnout is known with unusual reliability. Vote percentages are displayed on a probit scale to get rid of misleading floor and ceiling effects, and to parallel the probit specification. Thus linear or quadratic specifications within a probit link function will appear as straight lines or parabolas, respectively, on this scale.

---

[14]That is, on the probit scale, statistical specifications have zero second derivatives with respect to all pairs of non–interacted variables. That is, they have no interaction effects except those imposed by the shape of the probit fuction.
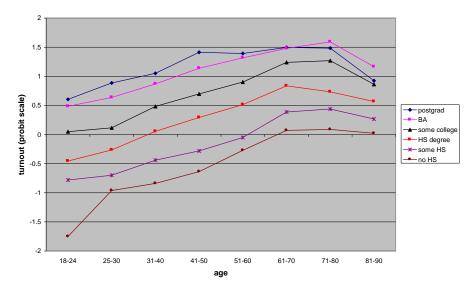
Figure 1: U.S. 2000 Turnout by Age and Education

Two features of the graph are immediately apparent. First, our usual voter turnout probit specifications are quadratic in age, so that the shapes in the figure should all be quadratic and equidistant from each other at all points. They are neither. Second, note the convergence of the lines at the upper right—a violation of conventional probit and logit specifications for turnout, but just what Proposition 15 above predicts (that is, the older you are, the less education matters). Similar remarks apply to the effect of education.

Does the model of this paper fit the data better than our conventional specifications? Answering this question with a reliable specification can easily involve the researcher in all the conceptual problems of turnout research—the distinctive histories of African–Americans, Latinos, and other minorities; the enculturation of immigrants; the consequences of high residential mobility; the problems of low youth turnout; and the disabilities of old age. For example:

- Education as a proxy works differently among young people. Many eighteen–year–olds who have not finished high school are not dropouts; they are still in school. A large fraction will eventually finish college or graduate work. Their ability to learn about the campaign is not well proxied by their current educational status.[15]

- Young people are highly mobile. Controlling for age but not mobility, as so many studies have done, attributes too much of the drop in youth turnout to the wrong factor.

---

[15] For example, in the 2000 CPS among those who have not finished high school, eighteen–year–olds vote considerably more than nineteen–year–olds. Most of the nineteen–year–olds are high school dropouts. Most of the eighteen–year–olds are not.

16

- Older people's turnout rates decline steadily for reasons of health and disability, a process political scientists know little about. By trying to model it, we have been practicing medicine without a license.

To minimize these biases, the data analysis here focuses on native–born white Americans between the ages of 25 and 70 who have lived at their current address five years or more.[16] At a stroke, this decision controls for a long series of difficult–to–specify variables and lets us focus on a large but relatively causally homogeneous subsample. There the central ideas of the model can be tested without blindly assuming that, for example, black people are just the same as white people except for a dummy variable.

Of course, the aspects of American voter turnout set aside by limiting the sample are interesting and important—perhaps more interesting and important than those the sample can address. However, it is not the purpose of this paper to explain all aspects of turnout. Rather, the goal is to develop tools to replace those that have failed us, and to replace them with something better so that we can get back to answering the important questions.

Theory–Based Probit Specifications for Voter Turnout

Having in effect controlled for many of the key variables affecting turnout, we are left with strength of partisanship, level of campaign information, and persuasive contact from a trusted source. Contact is not common in the U.S. and has relatively small effects (Rosenstone and Hansen 1993); it is also not measured in the CPS survey.[17] Hence we confine the included variables to partisanship and information, which drops the term h_q from Equation (10).

Since neither partisanship nor information is measured directly in the CPS samples, we begin by proxying them by systemtime (age - 18 years) and education, for the reasons given above. Neither of these variables is much correlated with merely having a preference in an election, since people of all ages and educations do that. Instead, it is the strength and unshakability of the electoral preference that age and education proxy for. That is, in Equation (10) these two variables affect primarily $(h_c + h_\tau + h_q)^{1/2}$, which increases with age and education, but they affect $\hat{u}_{n+1}$ much less, since it is relatively constant. With

---

[16]In the 2000 CPS, turnout begins to decline very slowly in people's early seventies and then decreases more rapidly after eighty. Failure to register after five years cannot be due to the burdens of a recent move, and it means that the citizen has passed through at least one prior presidential campaign (and usually one automobile license renewal with a painless opportunity to register) without getting onto the voter rolls. Hence to the extent possible, in the sample used here, failure to vote is due to lack of motivation, not health or registration laws.

We also found no systematic effect of self-reporting vs. proxy reporting for CPS respondents, and so both groups are included here.

[17]Similarly, gender has no effect and is omitted.

personal contact dropped, $(h_c + h_\tau + h_q)^{1/2}$ will be approximately linear in systemtime (or age) and education.[18]

If we add a normally distributed (Gaussian) error term to Equation (10), it follows that for the causally homogeneous subset of the CPS data, the specification implied by the theory of this paper is approximately:

$$\Pr(\text{vote}) = \Phi\left\{\delta + \gamma\left[2\Phi(\beta_1\text{systemtime } + \beta_2\text{education}) - 1\right]\right\} \tag{12}$$

where as before, $\Phi(.)$ is the cdf of a standard normal variable.[19] This setup is rather like a conventional probit equation, which with the same two variables would read:

$$\Pr(\text{vote}) = \Phi(\beta_0 + \beta_1\text{systemtime } + \beta_2\text{education}) \tag{13}$$

The difference is that the Bayesian–derived Equation (12) implies that the probit functional form enters twice, one nested inside the other. This is an entirely new statitistical specification, to be called "double probit."

How might the double probit coefficients be interpreted? Note from Equation (10) that a voter who knew nothing about the election would have zero precision in her posterior estimates, so that $(h_c + h_\tau + h_q)^{1/2} = 0$. Hence the probability that a wholly uninformed person would vote is given by $\Phi\left\{\delta + \gamma\left[2\Phi(0) - 1\right]\right\} = \Phi(\delta)$. Thus we expect δ to equal -2 or -3; that is, only a percent or two of the population would vote in the completely uninformed condition. On the other hand, complete information would have infinite precision, meaning

---

[18]By the usual Bayesian logic in the normally distributed case, precisions can be written as the equivalent of a certain number of observations. For both PID and campaign information, however, there is an upper bound on the precision, since there is a random draw on incumbent performance after the campaign that the voter does not see in advance. Under the assumptions, the PID precision h_c has a form like $n/(nk_1^{-1}+k_2) = k_1[1 - k_2/(nk_1^{-1} + k_2)]$, where n is the number of units of information available to the voter and the k_j's are constants. Approximating $k_2/(nk_1^{-1} + k_2)$ by $k_4\exp(-k_5 n) \approx$ -$k_4(1 - k_5 n)$ and substituting shows that h_c is approximately linear in n. By the same logic, information is approximately linear in m. Hence setting aside the effect of personal contact h_q, the term $(h_c + h_\tau + h_q)^{1/2}$ takes the approximate form $[K - k_4(1 - k_5 n) -k_6(1 - k_7 m)]^{-1/2}$. For K large relative to the other two terms, this expression is approximately $[K - k_4(1 - k_5 n)/2 -k_6(1 - k_7 m)/2]/K^{1/2}$, which is again linear in n and m (i.e., in PID and information). Making our standard substitution of systemtime for n and education for m then produces a positive linear function of both variables. The essential condition to make this approximation work is that inherited partisanship and the inability to forecast incumbent performance are sufficiently dominant that systemtime and education can make only limited improvements in predictive power. In practice, the linear approximation worked well and other transformations and functional forms were statistically rejected.

[19]Note that the constant is suppressed in the $\Phi(.)$ inside the brackets. If we had entered a constant β_0, then since the zero values of systemtime and education represent no information, for uninformed voters $\Phi(\beta_0 + \beta_1\text{systemtime } +\beta_2\text{education}) = \Phi(\beta_0)$. But complete lack of information is the condition $\Phi(0)$. Hence $\beta_0 = 0$. In practice, this equality does not hold exactly due to our inability to proxy perfectly for "zero knowledge." However, since $\Phi(.)$ is nearly perfectly linear around 0, this small difference is smoothly absorbed into the intercept δ. (One can estimate double probit with β_0 entered to assess how well the approximation works, but the task is often numerically difficult due to near–unidentification, it results in almost precisely the same forecasts, and it makes the interpretation of δ more elaborate.)

that $(h_c + h_\tau + h_q)^{1/2} = \infty$, so that the probability that a fully informed person would vote is $\Phi\{\delta + \gamma[2\Phi(\infty) - 1]\} = \Phi(\delta + \gamma)$. Thus we foresee γ equal to 4 or 5, making $\delta + \gamma$ approximately 2, meaning that in the fully informed condition, all but 2-3% of the population would vote. (At any given time, approximately that many are unexpectedly out of town, sick, having a baby, attending a family funeral, and so on, as many surveys have shown). Thus in spite of their unconventional appearance, the extra coefficients δ and γ in double probit have a straightforward and intuitive interpretation.

We now proceed to test double probit against the conventional probit alternative.[20] We also compare the fit using the scobit estimator proposed by Nagler (1994), who demonstrated statistically what Glaser (1962) had found graphically, that something was amiss with conventional probit specifications for voter turnout. Since age and systemtime produce exactly the same fit and receive the same coefficient in both ordinary probit and scobit, and since age is more conventional, we use age in those equations. Systemtime is theoretically more appropriate in double probit and not quite equivalent to age there, so it is retained in that model.[21] Since double probit in Equation (12) and its scobit equivalent each have four coefficients, we add the usual age^2 term to the probit specification, so that it has four coefficients as well. The model comparison then is fair—same data, same dependent variable, same number of coefficients for all three models.[22]

Some advance insight into the comparison follows from Equation (12) after we replace systemtime by its near–equivalent, age. Then a second–order Taylor series expansion of $\delta + \gamma[2\Phi(\beta_1 \text{age} + \beta_2 \text{education}) - 1]$ produces an ordinary probit specification with terms in age, age^2, education, education^2, an interaction term age*education, plus a constant. Minus the interaction term, this is the specification of age and education that Wolfinger and Rosenstone (1980) used, as well as many careful researchers since then. Thus by running hundreds of probits over several decades and by using additional parameters, careful empirical work has approximated what theory suggests. To honor that tradition, we include its version of the probit specification in the following tables, even though it has one more parameter than the other models. And because the traditional empirical fits are well honed, we expect log–likelihoods to improve when double probit is used, but not dramatically. The larger gains will come instead in theoretical interpretability, in the abolition of ad–hoc specification fixes such as age^2 and education^2, and in predictive portability across contexts,

---

[20]Of course, logit gives fits virtually identical to those of probit apart from a change of scale for the coefficients, and so it was not tested separately. However, logit is a special case of scobit, so logit would always have a poorer log–likelihood than scobit in the following tables.

[21]If double probit is run with age instead of systemtime, the log likelihoods change by less than a hundreth of a percentage point, leaving the model comparisons unchanged in every respect. The coefficients, too, change only trivially, though the change of explanatory variable alters the scale of δ and γ and makes their intepretation much less intuitive.

[22]Age was entered as age in years; systemtime as age minus eighteen years. Education was coded linearly from one to six, using the same categories as in Figure 1.

as we show below.

The results of the first comparison are given in Table 1.[23] To repeat, all the samples studied here consist of native–born white Americans between the ages of 25 and 70 who have lived at their current address five years or more. As expected, the log–likelihood is somewhat better for double probit, meaning of course that it fits better. Its t–ratios are also generally better than those of the other models, and a great deal better than scobit's. The standard five–parameter version of probit with two quadratic terms is not as good as double probit with just four coefficients, as the fourth column shows. Even scobit with age, age^2, education, and $education^2$, giving it six coefficients, fits these data slightly less well (log–likelihood = -11751.4) than double probit with just four parameters (coefficients not shown).

Table 1. CPS 2000 Turnout Study
(Standard errors in parentheses. N = 25,666)

	probit	scobit	double probit	probit
age/systemtime	.0355 (.00676)	.0610 (.00686)	.00957 (.000671)	.0337 (.00677)
age^2	-.000126 (.0000689)			-.000101 (.0000691)
education	.447 (.00915)	1.199 (.131)	.189 (.0126)	.705 (.0422)
$education^2$				-.0355 (.00563)
constant	-2.242 (.161)	-4.437 (.356)		-2.649 (.174)
scobit α		.547 (.0788)		
δ			-2.424 (.141)	
γ			4.865 (.0976)	
log-likelihood	-11780.1	-11752.0	-11750.5	-11760.46
parameters	4	4	4	5

The double probit estimates make substantive sense. The values of δ and γ are just what

---

[23] All computations were carried out with STATA 9, using the **ml model lf** procedure for double probit. (See the Appendix.) Thus asymptotic standard errors were computed numerically. The model conditions on age and education and assumes fixed coefficients, so that survey sampling weights were not used except in purely descriptive presentations like Figure 1. Examination of all age and education groups showed extremely small differences between weighted and unweighted turnouts and no systematic effects.

they should be, implying that very few citizens would vote with no information, while all but a few would vote with perfect information. Pleasingly, too, the age and education variables enter the double probit link function linearly: There is no need for quadratic terms. (In fact, when those quadratic terms are added to double probit, they take on substantively tiny and highly statistically insignificant coefficients.) Put the other way around, the quadratic terms in conventional probit specifications are the statistically necessary distortions and loss of degrees of freedom induced by using an inappropriate model.[24]

The same pattern occurs when the 1998 CPS is analyzed with the same sample definition and the same variables (not shown). Double probit's log–likelihood is again better than those for the four–parameter versions of probit and scobit by approximately the same proportion as in the 2000 presidential election (-15975.08 vs. -16003.22 and -15975.20, respectively).[25] Probit with five parameters is again inferior to double probit with four. Double probit's coefficients have the same sensible interpretations as before. The implied fully–informed turnout is somewhat lower, as is appropriate for a midterm election.

A third comparison of fits was done with the 2000 Annenberg survey, chosen because it has a considerably larger sample than the 2000 American National Election Survey. We applied the same models to these data, with the statistical sample defined in the same way.[26] However, the Annenberg sample's turnout rate is nine percentage points higher in the group analyzed, probably due to the Census Bureau's higher response rate and thus the presence of more non–voters in their sample.

Table 2 shows the results. Again double probit gives a slightly better fit than probit or scobit, but the more interesting object of study is the coefficients. Double probit finds very similar coefficients to those in Table 1, with all t–ratios exceeding 3.5 here, as befits the fact that this sample is drawn from the same population of people in the same election as Table 1 and that the sample size is relatively large. Probit, on the other hand, suddenly finds that

---

[24]Only when probit and scobit are given all the Taylor series expansion terms, including both quadratic terms and an interaction term between education and age, do they match the double probit performance (though with twice as many parameters). This implies that empirical researchers who want to avoid tying themselves to a particular theoretical framework and those who have only probit or logit software should at minimum control for age, age^2, education, education^2, an interaction term age*education, mobility, mobility^2, the variable of interest and its square, plus all the first–order interaction terms such as mobility*age, variable–of–interest*age, variable–of–interest*education, and so on. (Dummy variables might be used for the categories of age and education instead of linear and squared terms, but then each dummy must be interacted.) The proliferation of variables is the price paid for using a less appropriate estimator, but the estimated impact of the variable of interest, appropriately calculated using all the terms in which it appears, has a much better chance of being correct than in the usual garbage–can probits and logits.

[25]Apart from some instability in its coefficients, scobit performs well here and in all the comparisons of Tables 1-3. With its $\alpha < 1$, it diminishes the effect of large values of the explanatory variables, just as double probit does. Thus scobit again proves its value as an exploratory tool, finding probit and logit specifications that are inadequate.

[26]The one exception is that Annenberg has no variable for being native–born, so that non–native–born U.S. citizens are included in this sample. With Latinos and Asians already excluded, this difference will affect very few observations.

the age term is probably small (its coefficient has dropped to 2% of its former value with a wide standard error), and scobit's standard error for education has gone up by a factor of ten. The four–parameter version of probit has only one statistically significant coefficient (for education), and scobit has none at all. These anomalies with probit and scobit in comparing Tables 1 and 2 might lead to all sorts of queries about sampling biases, survey house effects, and so on. Instead, the fault lies first with conventional probit specifications: As a careful look at the Taylor series expansion mentioned above will demonstrate, the probit coefficients will move around in different sampling frames and different contexts for reasons having nothing to do with the underlying causal reality. And for scobit, the estimator is often simply noisy (Hanmer 2003).

Table 2. Annenberg 2000 Turnout Study
(Standard errors in parentheses. N = 2765)

	probit	scobit	double probit	probit
age/systemtime	.000486	.0792	.00973	-.0000585
	(.0262)	(.0615)	(.00220)	(.0263)
age^2	.000193			.000204
	(.000265)			(.000266)
education	.419	1.757	.220	.737
	(.0328)	(1.336)	(.0447)	(.165)
education^2				-.0414
				(.0209)
constant	-.856	-4.803		-1.419
	(.628)	(3.323)		(.691)
scobit α		.405		
		(.333)		
δ			-2.086	
			(.572)	
γ			4.571	
			(.363)	
log-likelihood	-820.1	-817.3	-817.0	-818.2
parameters	4	4	4	5

The Annenberg study can also be used to test the effects of partisanship and information directly, since unlike the CPS, it has survey questions addressing those topics, albeit with the usual survey noise.[27] Table 3 gives the results. Again double probit does better than

---

[27] Information levels were coded linearly from the interviewer summary ratings: A=4, B=3, C=2, D or F=1. For partisanship, the coding was: strong=4, weak=3, leaning independent=1, true independent=0.

probit or scobit, and its coefficients retain sensible interpretations. Scobit in particular fails to achieve a reliable set of estimates; its standard errors have exploded. As expected, the imprecise questions mean that not everyone who appears to be at the low end of information and partisanship actually is poorly informed, and so the intercept δ is somewhat larger than when age and education are used. But the theoretical expectations generated by the Bayesian model are again confirmed.

Table 3. Annenberg 2000 Turnout Study
(Standard errors in parentheses. N = 2765)

	probit	scobit	double probit
PID strength	.309	5.718	.223
	(.121)	(26.39)	(.0345)
(PID strength)^2	-.00735		
	(.0358)		
information	.380	8.407	.368
	(.0384)	(38.80)	(.0613)
constant	-.263	-1.849	
	(.129)	(8.683)	
scobit α		.0770	
		(.355)	
δ			-1.212
			(.326)
γ			3.112
			(.256)
log-likelihood	-835.7	-830.1	-827.8
parameters	4	4	4

Finally, we assess the ability of double probit to forecast across different contexts. As noted at the beginning, good theory should be able to do so, while ad hoc specifications should fail. To carry out the test, the 2000 presidential election parameter estimates were computed. Then one extra parameter λ was added to each of the four–parameter versions of double probit, probit, and scobit, representing the effect of a midterm election. For probit and scobit, since the conventional models contain no theoretical expectations about what happens to their coefficients in lower turnout elections, the parameter λ was added

---

Of course, being a "strong" partisan is ambiguous. It conflates position with firmness, $\hat{\delta}_n$ with h_n. In this and other ways—interviewer measurement error, respondent vagueness, and the simple linear codings of categories—these measures are very noisy and the estimates cannot be taken at face value for any of the estimators. Moreover, with the weak data, all the estimators struggled with poor standard errors—scobit immediately and double probit with the addition of quadratic terms. Probit's sampling errors held up a bit better, although its log–likelihood never matched that of double probit with equal numbers of parameters.

linearly inside their link functions in the customary way. In effect, it simply modifies the intercept. For double probit, the λ parameter was interacted with the effect of education, since in the theory, education represents the effect of campaign information, which is what changes at midterms. We do not expect a similar change in party identification.

Thus using the same homogeneous sample as before, the forecasting procedure was as follows:

1. Estimate each model in the 2000 CPS sample with age and education as predictors, plus age^2 for probit to keep the number of parameters equal across models. To improve the fit, dummy variables were used for the various categories of education. (The six education categories do not quite produce equally spaced effects.) Thus all models had eight parameters, including five dummy variables for education. Double probit again fit best, though the log–likelihood differences were small.

2. Add the parameter λ and modify it until the average forecast from the 2000 model matches the overall turnout in the midterm. For probit, for example, this means keeping the education dummies, age, and age^2 coefficients but lowering the 2000 intercept until the predicted turnout is that of 1998. The same approach was used for scobit, though without the age^2 term. For double probit, all the education dummy coefficients were reduced proportionately until the 1998 turnout was matched.

3. Applying the coefficients from the 2000 fit along with the estimated λ to the respondents in the 1998 election sample, compute their forecasted turnout. Then compare these 1998 forecasts with the actual 1998 outcomes.

The result is that all three models predict extremely well in the middle–education categories—usually within a percentage point or two—and reasonably well in the lowest education categories, usually within a few percentage points. The only substantial and consistent difference among the model forecasts emerged among those citizens with a college degree or graduate work. That difference is displayed in Figure 2. There it is clear that double probit is substantially better, as its forecasts generally go through the data points while probit and scobit are consistently too high by 5-10 percentage points. In particular, as expected from the Bayesian model, double probit predicts that turnout among the well educated will drop more at midterms than our conventional specifications expect. That prediction is amply confirmed.
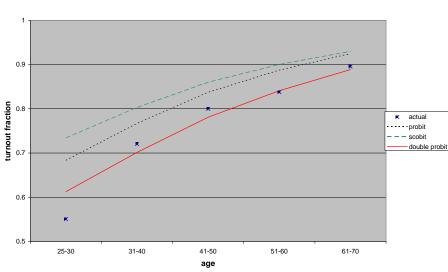
24

**Figure 2: Forecasts of 1998 Turnout from 2000 Data
for Native-Born Whites of Stable Residence with BA or more**

Thus double probit models allow us to forecast accurately across contexts, while the conventional probit and scobit specifications do not. Bayesian theory provides the specification guidance that makes better forecasts possible.

Conclusion

This paper has set out a simple Bayesian model to unify the study of voter turnout. It suggests three different paths by which voters get to the polls. First, they can follow the current campaign and deduce the appropriate choice for themselves. Second, they can rely on their own cognitive shortcuts, typically partisan identification. In some countries with weakly rooted parties, PID may be replaced by ideology, religious preference, or strong, continuing candidate attachments. Finally, voters may be encouraged to vote by a trusted source—a party, an interest group, a church, or a spouse. The model lays out how these paths relate to other causal factors and to each other.

All the main results of the empirical literature follow from the model, as do several new and non-obvious implications, which the evidence supports. Moreover, the model implies a new functional form for probit models of voter turnout. These new models require a bit more thinking than the old garbage–can probit specifications with every variable dumped in linearly, but the new specifications make theoretical sense, fit better, eliminate specification ad–hockeries, and exhibit coefficient stability across sampling frames, which the older approaches do not. More dramatically, they predict outside the original sample, while the old–fashioned atheoretical specifications fail. (Further detail and additional applications

25

will appear in Achen and Sinnott 2007.)

Policy prescriptions follow from the theoretical framework, too. For example, it suggests that reformers should now focus less on making registration easier and more on civic education, particularly education that teaches students what the stakes are in political life and how they they might learn to take a side (Niemi and Junn 1998, chap. 7). It also suggests that shorter ballots and less frequent elections would reduce the information demands on American voters. Bayesian theory should help assess these policy proposals.

Bayesian theories have begun to provide a unified theoretical framework for the study of voting behavior. Competing formal theoretic frameworks, based in behavioral economics, have begun to emerge as well (Bendor et al. 2003; Collins 2005). Although most of our classroom readers and textbooks on voting behavior continue to present the subject in the older behavioral tradition, rapid change is afoot. While the profession will continue to learn from careful study of data without theoretical preconceptions, it seems clear that the future version of the subject will look more like consumer theory in economics than like our current presentations to students and our current texts in public opinion and voting.

Methodology, too, has to keep up (Achen 2002b). The days are gone in which we could say, "The dependent variable is dichotomous, so I used probit (or logit)." As this paper has tried to demonstrate, the measurement level of a dependent variable has no necessary implication for which estimator should be used. Researchers can be systematically misled by atheoretical choice of an estimator. Work of that kind can sometimes be suggestive, but the results are always under suspicion.

The model of this paper is just a start, of course. It is easy to think of ways in which the current simplified setup could be extended theoretically–ARMA(p,q) time series assumptions for the citizens' retrospective evaluations instead of the white noise assumptions, unknown variances instead of known, non–normal distributions, formal treatment of survey sampling weights, and so on. Hierarchical models could be tried on the full CPS series of biennial surveys, and similarly for the growing number of comparative international surveys, particularly now that theoretical advances have made the analysis so much easier (Jusko and Shively, 2005). Most importantly, various additions and revisions will be needed to cope with the details of electoral history for young people, old people, those who went through defining political events such as the Depression and World War II (Miller and Shanks 1996), minorities, immigrants, newcomers to a jurisdiction, citizens in countries with non–democratic interludes, and much else. Without doubt, many such steps will be required before turnout is genuinely understood.

The point of this paper, then, is not to tout a statistical "solution" for voter turnout studies. Dumping a dozen variables willy–nilly into double probit is no better than doing so in ordinary probit. There is no escape from detailed substantive knowledge, extensive

graphical analysis, and genuine development of mathematical theory for each context and each key variable. No one estimator fits all. The point is rather to propose a redirection of the literature. Too much behavioral research on voting has consisted of disconnected local findings that make no case for generality and make no connection to rigorous theory. Too much formal theoretical work has bypassed the task of real science—accounting for the actual political world and making non-obvious behavioral predictions that data could verify. Setting up a closer connection between formal theory and data has become the central goal for contemporary quantitative political science (Granato and Scioli, 2004). This paper is meant to take a step in that direction for the study of voter turnout.

3 Appendix: STATA Code for Double Probit

This appendix gives the STATA code for running double probit. The first six lines define the program. They are standard and are not modified for different applications. The final three lines generate the estimates for a particular model.

In the **ml model** statement, age and education are the explanatory variables listed here as an example, but of course different variables with different names can be added as appropriate. Nothing else in the statement would change. As usual, a variety of options are available with this command that are not listed here. Note that the variable on the left–hand–side of the equal sign is assumed to be the dependent variable.

The **ml init** statement makes double probit converge more quickly. The starting values given here seem to work well in practice. In particular, initializing each independent variable with a small positive coefficient is helpful. The likelihood function need not be globally concave, so that plausible starting values and the sophisticated numerical maximization routines in a reliable statistical package like STATA are particularly desirable.

```
program define doubleprobit
version 8
args lnf theta1 theta2 theta3
quietly replace 'lnf' = ln(norm('theta1' +'theta2'*(2*norm('theta3')-1))) if $ML_y1==1
quietly replace 'lnf' = ln(1-norm('theta1' +'theta2'*(2*norm('theta3')-1))) if $ML_y1==0
end

ml model lf doubleprobit () () (vote = age education, noconstant)
ml init eq1:_cons= -2 eq2:_cons= 4 eq3:age= .1 eq3:education= .1
ml max
```

References

[1] Achen, Christopher H. 1992. Social Psychology, Demographic Variables, and Linear Regression. *Political Behavior* 14,3 (Sept.): 195-211.

[2] Achen, Christopher H. 2002a. Parental Socialization and Rational Party Identification. *Political Behavior* 24,2 (June): 151-170.

[3] Achen, Christopher H. 2002b. Toward a New Political Methodology: Microfoundations and ART. *Annual Review of Political Science* 5: 423-450.

[4] Achen, Christopher H., and Richard Sinnott, eds. 2007. *Voter Turnout in Multi–Level Systems.* Book manuscript in preparation.

[5] Aldrich, John H. 1993. Rational Choice and Turnout. *American Journal of Political Science* 37,1 (Feb.): 246-278.

[6] Ansolabehere, Stephen D., Shanto Iyengar, Adam Simon, and Nicholas Valentino. 1994. Does Attack Advertising Demobilize the Electorate? *American Political Science Review* 88,4 (Dec.): 829-838.

[7] Arneson, Ben A. 1925. Non-Voting in a Typical Ohio Community. *American Political Science Review* 19,4 (Nov.): 816-825.

[8] Bartels, Larry M. 1993. "Messages Received." *American Political Science Review* 41, 2: 641-674.

[9] Bartels, Larry M. 1998. *Campaign Reform.* Princeton: Woodrow Wilson School, Princeton University.

[10] Bartels, Larry M. 2002. Beyond the Running Tally: Partisan Bias in Political Perceptions. *Political Behavior* 24: 117-150.

[11] Bean, Louis. 1948. *How to Predict Elections.* New York: Knopf, 1948.

[12] Bendor, Jonathan, and Dilip Mookherjee. 1990. Norms, Third-Party Sanctions, and Cooperation. *Journal of Law, Economics, and Organization* 6,1 (Spring): 33-63.

[13] Bendor, Jonathan, Daniel Diermeier, and Michael Ting. 2003. A Behavioral Model of Turnout. *American Political Science Review* 97,2 (May): 261-280.

[14] Berelson, Bernard R., Paul F. Lazarsfeld, and William N. McPhee. 1954. *Voting.* Chicago: University of Chicago Press.

[15] Bernheim, B. Douglas. 1994. A Theory of Conformity. *Journal of Political Economy* 102, 5 (Oct.): 841-877.

[16] Blais, André. 2000. *To Vote or Not to Vote.* Pittsburgh: University of Pittsburgh Press.

[17] Blais, André. 2006. What Affects Voter Turnout? *Annual Review of Political Science* 9: 111-125.

[18] Blondel, Jean, Richard Sinnott, and Palle Svensson. 1998. *People and Parliament in the European Union.* Oxford: Clarendon.

[19] Boyd, Richard W. 1989. The Effects of Primaries and Statewide Races on Voter Turnout. *Journal of Politics* 51, 3 (Aug.): 730-739.

[20] Brady, Henry E., and Stephen Ansolabehere. 1989. The Nature of Utility Functions in Mass Publics. *American Political Science Review* 83, 1 (Mar.): 143-163.

[21] Brady, Henry E., and John E. McNulty. 2004. The Costs of Voting: Evidence from a Natural Experiment. Paper prepared for delivery at the 2004 Annual Meeting of the Society for Political Methodology, Stanford University, July 29-31.

[22] Brennan, Geoffrey, and Loren Lomasky. 1993. *Democracy and Decision.* Cambridge, England: Cambridge University Press.

[23] Brennan, Geoffrey, and Alan Hamlin. 2000. *Democratic Devices and Desires.* Cambridge, England: Cambridge University Press.

[24] Brennan, Geoffrey, and Philip Pettit. 2004. *The Economy of Esteem.* Oxford: Oxford University Press.

[25] Burnham, Walter Dean. 1972. Political Immunization and Political Confessionalism. *Journal of Interdisciplinary History* 3,1 (Summer): 1-30.

[26] Butler, David, and Donald Stokes. 1969. *Political Change in Britain.* New York: St. Martin's.

[27] Cain, Bruce E. 1978. Strategic Voting in Britain. *American Journal of Political Science* 22, 3 (Aug.): 639-655.

[28] Campbell, Angus, Gerald Gurin, and Warren E. Miller. 1954. *The Voter Decides.* Evanston, Illinois: Row, Peterson.

[29] Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. New York: Wiley.

[30] Campbell, Angus. 1962. The Passive Citizen. *Acta Sociologica* 6,1-2: 9-21.

[31] Campbell, Angus. 1966. Surge and Decline. In Angus Campbell, Philip E. Converse, Warren E. Miller, and Donald E. Stokes, *Elections and the Political Order.* New York: Wiley. Pp. 40-62.

[32] Clarke, Harold D., David Sanders, Marianne C. Stewart, and Paul Whitely. 2004. *Political Choice in Britain.* Oxford: Oxford University Press.

[33] Converse, Philip. E. 1966b. "Information Flow and the Stability of Partisan Attitudes." In Angus Campbell, Philip E. Converse, Warren E. Miller, and Donald E. Stokes, *Elections and the Political Order.* New York: Wiley. Pp. 136-57.

[34] Collins, Nathan A. 2005. A Behavioral Model of Age and Turnout. Manuscript, Department of Political Science, Stanford University.

[35] Converse, Philip. E. 1969. "Of Time and Partisan Stability." *Comparative Political Studies* 2:139 71.

[36] Delli Carpini, Michael X., and Scott Keeter. 1996. *What Americans Know about Politics and Why It Matters.* New Haven: Yale University Press.

[37] Diermeier, Daniel, and Jan. A. Van Mieghem. 2002. Coordination and Turnout in Large Elections. Manuscript, Northwestern University, Evanston, Illinois.

[38] Downs, Anthony. 1957. *An Economic Theory of Democracy.* New York: Harper and Row.

[39] Ferejohn, John A., and Morris P. Fiorina. 1974. The Paradox of Not Voting. *American Political Science Review* 68: 525-546.

[40] Fiorina, Morris. 1981. *Retrospective Voting in American Elections.* New Haven: Yale University Press.

[41] Fiorina, Morris. 1976. The Voting Decision: Instrumental and Expressive Aspects. *Journal of Politics* 38,2 (May): 390-413.

[42] Franklin, Charles H., and John E. Jackson. 1983. "The Dynamics of Party Identification." *American Political Science Review* 77:957-73.

[43] Gerber, Alan, and Donald P. Green. 1998. "Rational Learning and Partisan Attitudes." *American Journal of Political Science* 42, 3: 794-818.

[44] Gerber, Alan S., Donald P. Green, and Ron Shachar. 2003. Voting May Be Habit Forming. *American Journal of Political Science* 47: 540-550.

[45] Glaser, William A. Fluctuations in Turnout. 1962. In Willliam N. McPhee and William A Glaser, *Public Opinion and Congressional Elections*. New York: Free Press, pp. 19-51.

[46] Goldstein, Kenneth M. 1999. *Interest Groups, Lobbying, and Participation in America.* Cambridge, England: Cambridge University Press.

[47] Green, Donald P., and Ian Shapiro. 1994. *Pathologies of Rational Choice Theory.* New Haven: Yale University Press.

[48] Griffin, John D., and Brian Newman. 2005. Are Voters Better Represented? *Journal of Politics* 67,4 (Nov.): 1206-1227.

[49] Gosnell, Harold. 1927. *Getting Out the Vote.* Chicago: University of Chicago Press.

[50] Granato, Jim, and Frank Scioli. 2004. Puzzles, Proverbs, and Omega Matrices: The Scientific and Social Significance of Empirical Implications of Theoretical Models (EITM). *Perspectives on Politics* 2, 2: 313-323.

[51] Grofman, Bernard. 1993. Is Turnout the Paradox that Ate Rational Choice Theory? In *Information, Participation, and Choice*, ed. Bernard Grofman. Ann Arbor: University of Michigan Press.

[52] Hajnal, Zoltan, and Jessica Trounstine. 2006. Turnout Matters: Voter Turnout and City Spending. Manuscript, Princeton University.

[53] Hanmer, Michael J. 2003. An Investigation of Scobit. Manuscript. Department of Political Science, Georgetown University.

[54] Hanmer, Michael J. 2004. From Selection to Election and Beyond: Understanding the Causes and Consequences of Electoral Reform in America. Doctoral dissertation, Political Science, University of Michigan.

[55] Harsanyi, John C. 1969. Rational–Choice Models of Human Behavior vs. Functionalist and Conformist Theories. *World Politics* 21: 513-538.

[56] Hastings, Philip K. 1956. The Voter and the Non–Voter. *American Journal of Sociology* 62, 3: 302-307.

[57] Hetherington, Marc J. 2001. Resurgent Mass Partisanship: The Role of Elite Polarization. *American Political Science Review* 95: 619-631.

[58] Highton, Benjamin, and Raymond E. Wolfinger. 1998. Estimating the Effects of the National Voter Registration Act of 1993. *Political Behavior* 20 (June): 79-103.

[59] Highton, Benjamin, and Raymond E. Wolfinger. 2001. The First Seven Years of the Political Life Cycle. *American Journal of Political Science* 45 (Jan.): 202-209.

[60] Highton, Benjamin. 2004. Voter Registration and Turnout in the United States. *Perspectives on Politics* 2 (Sept.): 507-515.

[61] Hollander, Heinz. 1990. A Social Exchange Approach to Voluntary Cooperation. *American Economic Review* 80, 5 (Dec.): 1157-1167.

[62] Holls, Frederick W. 1889. Compulsory Voting. *American Political Science Review* 1 (Apr.): 586-607.

[63] Howard, Perry H. 1971. *Political Tendencies in Louisiana*, rev. ed. Baton Rouge: Louisiana State University Press.

[64] Jackson, John E. 1975. "Issues, Party Choices and Presidential Votes." *American Journal of Political Science* 19:161-85.

[65] Janowitz, Morris. 1983. *The Reconstruction of Patriotism.* Chicago: University of Chicago Press.

[66] Jennings, M. Kent, and Richard G. Niemi. 1981. *Generations and Politics.* Princeton: Princeton University Press.

[67] Jusko, Karen Long, and W. Phillips Shively. 2005. Applying a Two–Step Strategy to the Analysis of Cross-National Public Opinion Data. *Political Analysis* 13,4: 327-344.

[68] Kandori, Michihiro. 1992. Social Norms and Community Enforcement. *Review of Economic Studies* 59,1 (Jan.): 63-80.

[69] Kelley, Stanley, Richard E. Ayres, and William G. Bowen. 1967. Registration and Voting: Putting First Things First. *American Political Science Review* 61,2 (Jun.): 359-379.

[70] Key, V.O. 1964 [1942]. *Politics, Parties, and Pressure Groups.* New York: Crowell.

[71] Key, V.O. 1966. *The Responsible Electorate.* New York: Vintage.

[72] Lazarsfeld, Paul, Bernard Berelson, and Hazel Gaudet. 1944. *The People's Choice.* New York: Columbia University Press.

[73] Lane, Robert E., and David O. Sears. 1964. *Public Opinion.* Englewood Cliffs, New Jersey: Prentice–Hall.

[74] Lassen, David Dreyer. 2005. The Effect of Information on Voter Turnout. *American Journal of Political Science* 49,1: 103-118.

[75] Ledyard, John O. 1984. The Pure Theory of Large Two-Candidate Elections. *Public Choice* 44: 7-41.

[76] Lifton, Robert Jay. 1961. *Thought Reform and the Psychology of Totalism.* New York: Norton.

[77] Lijphart, Arend. 1997. Unequal Participation: Democracy's Unresolved Dilemma. *American Political Science Review* 91,1 (Mar.): 1-14.

[78] Lupia, Arthur, and Matthew D. McCubbins. 1998. *The Democratic Dilemma.* Cambridge, England: Cambridge University Press.

[79] McDonald, Michael P., and Samuel L. Popkin. 2001. The Myth of the Vanishing Voter. *American Political Science Review* 95: 963-974.

[80] McPhee, William N., and Jack Ferguson. 1962. Political Immunization. In William N McPhee and William A. Glaser, *Public Opinion and Congressional Elections.* New York: Free Press, pp. 155-79.

[81] Merriam, Charles E. and Harold F. Gosnell. 1924. *Non-Voting: Causes and Methods of Control.* Chicago: University of Chicago Press.

[82] Merriam, Charles E. 1931. *The Making of Citizens.* Chicago: University of Chicago Press.

[83] Milbrath, Lester W. 1965. *Political Participation.* Chicago: Rand McNally.

[84] Miller, Warren E., and J. Merrill Shanks. 1996. *The New American Voter.* Cambridge, Massachusetts: Harvard University Press.

[85] Morton, Rebecca B. 1991. Groups in Rational Turnout Models. *American Journal of Political Science* 35, 3 (Aug.): 758-776.

[86] Niemi, Richard G., and Jane Junn. 1998. *Civic Education.* New Haven: Yale University Press.

[87] Palfrey, Thomas R., and Howard Rosenthal. 1983. A Strategic Calculus of Voting. *Public Choice* 41: 7-53.

[88] Nagler, Jonathan. 1994. Scobit: An Alternative Estimator to Logit and Probit. *American Journal of Political Science* 38,1 (Feb.): 230-255.

[89] Palfrey, Thomas R., and Howard Rosenthal. 1985. Voter Participation and Strategic Uncertainty. *American Political Science Review* 79: 62-78.

[90] Patterson, Thomas E. 2003. *The Vanishing Voter.* New York: Vintage.

[91] Plutzer, Eric. 2002. Becoming a Habitual Voter. American Political Science Review 96: 41-56.

[92] Putnam, Robert D. 2000. *Bowling Alone.* New York: Simon and Schuster.

[93] Riker, William H., and Peter C. Ordeshook. 1968. A Theory of the Calculus of Voting. *American Political Science Review* 62,1 (Mar.): 25-42.

[94] Rosenstone, Steven J., and John Mark Hansen. 1993. *Mobilization, Participation, and Democracy in America.* New York: Macmillan.

[95] Sanders, Mitchell S. 2001. Uncertainty and Turnout. *Political Analysis* 9, 1: 45-57.

[96] Schuessler, Alexander A. 2000. *A Logic of Expressive Choice.* Princeton: Princeton University Press.

[97] Squire, Peverill, Raymond E. Wolfinger, and David P. Glass. Residential Mobility and Voter Turnout. *American Political Science Review* 81 (Mar.): 45-66.

[98] Teixeira, Ruy A. 1992. *The Disappearing American Voter.* Washington, D.C.: Brookings.

[99] Tingsten, Herbert. 1937. *Political Behavior.* London: P. S. King.

[100] Uhlaner, Carole J. 1989. Rational Turnout: The Neglected Role of Groups. *American Journal of Political Science* 33,2 (May): 390-422.

[101] Verba, Sidney, Kay Lehman Schlozman, and Henry E. Brady. 1995. *Voice and Equality.* Cambridge, Massachusetts: Harvard University Press.

[102] Vidich, Arthur J., and Joseph Bensman. 1960. *Small Town in Mass Society.* Garden City, New York: Doubleday, 1960

[103] Wolfinger, Raymond E. and Steven J. Rosenstone. 1980. *Who Votes?* New Haven: Yale University Press.