The Varying Role of Voter Information Across Democratic Societies*

Jasjeet S. Sekhon †

Version 1.4

*Winner of the *Robert H. Durr Award* for "the best paper applying quantitative methods to a substantive problem" presented at MPSA. I thank Henry Brady, John Brehm, Scott Gehlbach, Sunshine Hillygus, Simon Jackman, Robert Luskin, Donald Rubin, Doug Rivers, Paul Sniderman, Jason Wittenberg, and especially Walter Mebane, Jr., Larry Bartels, Charles Franklin, Shigeo Hirano, Gary King, Suzanne Smith, and Jonathan Wand for valuable comments and advice. I thank Alexis Diamond for excellent comments and research assistance, Daniel Margolskee for research assistance, Jorge Domínguez for providing documentation for the Mexico 2000 Panel Study, John Zaller and German Social Science Infrastructure Services for providing data, and seminar participants for stimulating discussions at Stanford, Wisconsin and UC-Berkeley. Matching software can be downloading from `http://sekhon.berkeley.edu/match`. This work is supported in part by NSF grant SES-0214965. All errors are my responsibility.

†Associate Professor, Travers Department of Political Science, UC Berkeley, Survey Research Center 2538 Channing Way, Berkeley, CA 94720.

Abstract

Using new robust matching methods for making causal inferences from survey data, I demonstrate that there are profound differences between how voters behave in advanced democracies versus how they behave in new electoral democracies. The problems of voter ignorance and inattentiveness are not as serious in advanced democracies as many analysts have suggested but are of grave concern in new democracies. Citizens in advanced democracies are able to accomplish something that citizens in fledgling democracies are not: inattentive and poorly informed citizens are able to vote like their better informed compatriots and hence need to pay little attention to political events such as election campaigns in order to vote as if they were attentive. The results from the U.S. (which rely on various National Election Studies) and Mexico (2000 Panel Study) are reported in detail. Results from other countries are briefly reported.

"The people should have as little to do as may be about the government. They lack information and are constantly liable to be misled." —Roger Sherman, June 7, 1787 at the Federal Constitutional Convention (Collier 1971)

"In a crowd men always tend to the same level, and, on general questions, a vote, recorded by forty academicians is no better than that of forty water-carriers." —Gustave Le Bon, *The Crowd* (1896, 200)

1 Introduction

When James Wilson of Pennsylvania suggested that senators be directly elected by the people, not a single member of the Federal Constitutional Convention in Philadelphia rose to support him (Smith 1956, 236). Moreover, he was almost alone in supporting the direct election of the president. With few notable exceptions, such as Jefferson who was not at the Federal Convention, most of the American founding fathers had a deep suspicion, and in some cases fear, of popular sovereignty. There were several reasons for this suspicion, but a strong belief in the ignorance of the general population and hence the likelihood of them being misled was a prominent one.

Even in our demotic age, concerns remain about the ignorance and inattentiveness of voters and the ability of elites to manipulate them even though those concerns wax and wane. For example, in the aftermath of World War II, a war which resulted (in part) because of the *electoral* success of the Nazi Party, the concerns waxed. Notable and influential work highlighted the dangers of mass politics (e.g., Kornhauser 1959), and critics noted that public opinions are not generated by informed discussion and reasoned agreement but produced by modern communication technologies which are effective in manipulating poorly informed and disengaged citizens (e.g., Habermas 1996 [1964]).

These concerns are alive and well today—even in the aftermath of the "third wave" of democratization which has resulted, for the first time, in a majority of the world's population living under some kind of democratic regime (Bunce 2000; Huntington 1991). Analysts note that many of the new democracies are illiberal in the old sense of the word (Zakaria 1997), and are the kind of illiberal regimes which many of the American founding fathers feared would arise if mass politics were given too much of a role (Adams 1973 [1805]).

Mass society theorists have found their heirs in social capital theorists (Putnam 2000). And commentators continue to lament the low levels of political involvement, interest and information of voters in democratic societies (e.g., Patterson 2002). When it comes to political involvement in countries like the U.S., some have even come to see the 1950s and 60s as a relative golden age which is ironic given what commentators at the time were writing. These arguments are not just of a general nature. Specific policy consequences are thought to follow. Many argue that a large number of important policy outcomes occur because voters are too poorly informed and engaged to hold opinions which are in their interests (Althaus 2003). Both the American public's support for the war in Iraq post 9/11 (Pryor 2003) and the continuing support by poor southern whites of the Republican Party are examples to which commentators frequently point. Another is offered by Bartels (2003) who argues that the upward transfer of wealth over the past few decades in the U.S. has been broadly supported by Americans because they are misinformed about the role which public policy has played in fostering inequality.

Generally, the concerns analysts have about voters and democracy are different when they are discussing countries with long democratic traditions than when they are considering democratizing countries. These differences exist for various reasons including a belief that certain societies are especially unsuited for democracy because of poverty, rapid social change, a dysfunctional civil society and culture, the prevalence of illiberal values, too much or too little ethnic heterogeneity, bad institutions and a variety of other ills.[1] Many of these reasons lead to the belief that voters in such countries are unable to properly express their enlightened

[1] The democratization literature is vast and contradictory. For classic and influential discussions of the prerequisites for democracy see Lipset (1959) and Rustow (1970). More modern reviews and views are offered by Bunce (2000), Geddes (1999) and Snyder (2000).

preferences to the extent that democratic decisions in such societies are likely to undermine not only U.S. interests but their interests as well.[2] One prominent argument for why voters in such countries are unable to expresses their enlightened preferences is that if voters are unattached to groups or local leaders who can provide them with directions on how to behave in their own interests, voters can be manipulated and mobilized by mass movements which are inherently totalitarian in nature (Kornhauser 1959). Current versions of this argument emphasize the information and cues provided by democratic institutions such as news media, competitive parties, public interest groups and the presence of civic organizations such as unions.

I demonstrate that there are indeed profound differences between how voters behave in mature democracies versus how they behave in new ones. The problems of ignorance and inattentiveness are not as serious in mature democracies as many analysts have suggested but are of grave concern in new democracies. Citizens in mature democracies are able to accomplish something that citizens in fledgling democracies cannot: inattentive and poorly informed citizens are able to vote like their better informed neighbors and hence need not pay close attention to political events in order to vote as if they were attentive. A lot of institutional change has occurred since the time of the founding fathers, not the least of which is the creation of organized political parties, interest groups and public opinion polls. There are theoretical reasons to believe that these new institutions allow poorly informed citizens to behave as if they were better informed than they are (McKelvey and Ordeshook 1985a,b, 1986). But these are precisely the institutions which are either lacking or poorly developed in new democracies.

The results presented in this manuscript are striking and offer both good and bad news

[2] An extreme example of this view is offered by a statement made by Henry Kissinger during a policy meeting on whether the U.S. should fund an attempt to prevent Salvador Allende from winning the 1970 presidential election in Chile: "I don't see why we need to stand by and watch a country go communist due to the irresponsibility of its own people" (Hersh 1982). Richard Holbrooke was bothered by the issue on the eve of the 1996 elections in Bosnia: "Suppose the election was declared free and fair and those elected are racists, fascists, separatists, who are publicly opposed to [peace and reintegration]. That is the dilemma" (Zakaria 1998).

for the state of democratic society. On the one hand, the results suggest the concerns raised by scholars, politicians and activists about the inattentiveness of Americans and other democratic citizens to political events such as campaigns may be misplaced. In a mature democracy, citizens' lack of attention does not appear to harm the quality of their votes relative to those of their more attentive neighbors because their votes do not differ. On the other hand, the results support the concerns that many have about the rationality of new democracies. There is a qualitative difference in the roles which information and attentiveness play in elections in mature versus immature democracies.

It is important to note that this manuscript does not advance an argument that voters in mature democracies make choices which are in some general sense optimal or are the same choices they would make if they were "perfectly" informed. Many pathologies may exist, and those that do are shared by the well and poorly informed alike—by the attentive and the inattentive. Public opinion in mature democracies is often mistaken, but it is not mistaken because inattentive citizens are not watching the nightly news often enough.

This manuscript is organized into six sections. First, the role of political information in elections as measured by survey data is discussed. Second, I discuss the research design, causal model, estimator and corrections for panel attrition. The next two sections present results for the U.S. and Mexico. The penultimate section briefly describes results from other countries. The final section concludes.

2 Survey Data and Political Information

Survey research is strikingly uniform in its conclusions regarding the ignorance of the public. Berelson, Lazarsfeld, and McPhee (1954) argue voters fall short of classic standards of democratic citizenship. Campbell, Converse, Miller, and Stokes (1960) arrive at similar conclusions. Subsequent work did establish that information levels fluctuate over time (Bennett 1988), but no one disputes the long-established fact that most voters are politically ignorant (Althaus 1998, 2003; Converse 1964; Delli Carpini and Keeter 1996; Neuman 1986; Zaller 1992), and the argument that voters are inadequately informed given classical ideals of democratic citizenship has not been seriously challenged (e.g., Adams 1973 [1805]; Bryce 1995 [1888]; Habermas 1996 [1964]).

Although the fact of public ignorance has not been forcefully challenged, the meaning of this observation has been (Sniderman 1993). One approach is to claim that with limited information voters or collectives of voters can nevertheless make rational decisions. Some claim poorly informed individuals may still act in a sophisticated fashion because they make efficient use of signals (Alvarez 1997; Page and Shapiro 1992; Popkin 1991; Sniderman, Brody, and Tetlock 1991).[3] Even though individuals are poorly informed, political and electoral institutions may allow voters to make decisions that are much the same as they would make if they had better information. For instance, McKelvey and Ordeshook (1985a,b, 1986) suggest that polls and interest group endorsements may perform such cuing functions.

Notwithstanding the forgoing, the existing literature provides little empirical evidence for the contention that cues and heuristics are sufficient to produce election results that match what would happen if voters were well informed (Bartels 1996). My strategy for analyzing this question is to examine whether voters who change information state behave differently from voters who do not.

I show that in the United States, when voters are matched on baseline characteristics, voters who change information state vote no differently on election day than voters who do not. Before election day, however, American voters who change information state do differ in their vote intentions. This effect is present in September but insignificant by November. The election campaign appears to provide the necessary cues to make the differences vanish by election day. I am hardly the first to note people learn during the course of a campaign (e.g., Alvarez 1997; Johnston, Blais, Brady, and Crête 1992; Johnston, Hagen, and Jamieson

[3]Others rely on results related to the Condorcet Jury Theorem to claim that errors committed by individual voters will cancel out in the aggregate (Miller 1986; Wittman 1989). For a criticism of the usual jury theorem results see Austen-Smith and Banks (1996) and Berg (1993). For empirical results on the aggregate coherence of public opinion in the U.S. see Erikson, MacKuen, and Stimson (2002).

2004), but this is, to my knowledge, the first analysis to show that information effects on vote choice vanish by election day. In Mexico, however, differences based on information level persist even on election day. Mexican voters appear to have difficulty performing the information arbitrage.

2.1 Measuring Political Information

There is little agreement in the literature, which is largely based on American National Election Studies (NES), on the best way to measure levels of political information. The measures which are of most relevance for this project fall into three broad categories. The first type of measure relies on evaluations of the survey respondent made by the interviewer. The second type depends on factual general questions. These questions could have been answered even before the campaign started because they are of a general nature such as "who is the president of Russia" or "which political party controls the House of Representatives." The third type is specific to the campaign at hand and is arguably the most relevant for the vote decision being made in the current election. Measures of this type include the ability of respondents to place the candidates on issue dimensions, to accurately so place the candidates, and the ability to accurately identify specific campaign events.

The problem of measuring political information would be difficult if domain specific knowledge was important and separate from general political knowledge (Iyengar 1986). There is some evidence which supports the contention that domain specific knowledge is important (Holbrook, Berent, Krosnick, Visser, and Boninger 2004; Iyengar 1990; McGraw and Pinney 1990). But domain specific measures have difficulty producing effects which are not explained by general knowledge measures (Zaller 1986). And some argue that people who are informed about one political issue are also highly likely to be informed about other issues (Price and Zaller 1993).

General measures of information do almost as well as a special domain specific battery of 27 information questions which was used in the 1985 NES pilot study (Zaller 1986). These general measures include the correct placing of political parties and candidates on the left-right dimension and naming of public officials. Another general measure is to rely on the interviewer's judgment of how informed the respondent is. This is a five point scale which ranges from "very high" information to "very low." At least in face-to-face interviews with considerable political content, these ratings have been shown to have high reliability and comparability with more detailed measures (Zaller 1985, 1986).

Zaller (1986) concludes that interviewer ratings "are highly effective as measures of political information" (18). These ratings have an estimated reliability of 0.78 while the 1985 pilot study scale which consisted of 27 items, a scale Zaller calls the Cadillac information scale, has an estimated reliability of 0.89 (Zaller 1986). This Cadillac scale is not available outside of the pilot study, and most of the information scales available in the NES have reliabilities of about 0.8 to 0.85. These are very high reliabilities for, as Bartels (1996) points out, the usual seven-point issue scales have reliability coefficients of about .4 to .6. The interviewer scale apparently is not biased by the respondents' race, gender, education and income (Zaller 1985).[4]

Alvarez (1997) makes a case for why one may prefer the campaign specific measures of political information. These measures also appear to be the most consistent with the argument laid out by Holbrook et al. (2004). In contrast, Bartels (1996, 2003) relies on interviewer ratings.

I use all three kinds of measures. I use both the interviewer ratings of information and Zaller's information measure (Zaller 1992, forthcoming) which largely consists of factual information questions[5] and interviewer ratings. I also use measures which are specific to the current campaign such as the ability of a respondent to place the candidates in issue space and to accurately identify campaign events such as assigning campaign slogans to the correct candidate.

[4]If systematic biases like these did exist, they would not be a problem for my analysis because individuals are matched on all of these characteristics.

[5]These factual questions mainly involve placing candidates and parties on the left/right dimension and naming public officials and the majority parties in the House and Senate.

Notwithstanding the many arguments in the literature on how best to measure political information, my results are substantially the same regardless of which measure I use. Measurement details are presented in the country specific sections.

3 Research Design and Causal Inference

This section discusses four methodological issues: the research design, causal model, the estimator for the causal model and corrections for panel attrition.

3.1 Research Design

The traditional way of estimating the effect of political information on voting behavior is to use cross-sectional data. This approach is useful for descriptive analysis for it can answer questions such as "in the last election, did politically well informed voters vote differently than politically uninformed voters?" This kind of question is asked all of the time; the most common variant is "in the last election, how differently did men and women (or blacks and whites) vote?" But these questions are not causal. If the goal is to estimate the causal effect of being politically informed, it is unclear of how much use cross-sectional data are unless very strong modeling assumptions are made. Causal inference relies on a counterfactual of interest (Sekhon 2004a), and the one which is most obviously relevant for political information is "how would Jane have voted if she were better informed?" My goal here is to answer this counterfactual. There are other theoretically interesting counterfactuals which I do not know how to empirically answer such as "who would have won the last election if *everyone* were well informed?" The counterfactual I know how to approach is a local one about Jane. One may be able to find someone to compare Jane's voting behavior with who was comparable to Jane at the beginning of the campaign but who did become better informed. In this way we may compare the voting behavior of Jane and the counterfactual Jane who is better informed. The more general counterfactual about everyone being well informed would change almost everything about the political system as we know it including the behavior of candidates, the news media, advertisers, parties and other political activists and of course the voters themselves. I don't know how to answer such a general counterfactual using data and as far as I can tell neither does anyone else. Such general counterfactuals will have to be left to the theorists.[6] To the extent that the local counterfactual provides useful information about the more general one, the two counterfactuals are related. But, as is discussed in the next section, the second counterfactual violates a key identifying assumption of the causal estimator and this should be made clear at the outset.

It is easy enough using cross-sectional data to estimate the conditional probability of voting for the Republican candidate if a voter is informed and the probability if the voter is not informed. But it is unclear for what counterfactual this conditional probability is relevant. Although conditional probability is at the heart of inductive inference, by itself it isn't enough. Underlying conditional probability is a notion of counterfactual inference. It is possible to have a causal theory that makes no reference to counterfactuals,[7] but counterfactual theories of causality are by far the norm, especially in statistics.[8]

A randomized experiment follows the counterfactual reasoning outlined here. For example, we could either contact Jane to prompt her to vote as part of a turnout study or we could not contact her. But we cannot observe what would happen if we both contacted Jane and if we did not contact Jane—i.e., we cannot observe Jane's behavior both with and without the treatment. If we contact Jane, in order to determine what effect this treatment had on Jane's behavior (i.e, whether she voted or not), we still have to obtain some estimate of the counterfactual in which we did not contact Jane. We could, for example, seek to compare Jane's behavior with someone exactly like Jane whom we did not contact. In a randomized experiment we obtain a group of people (the larger the better) and we assign treatment to a

[6]Attempts to estimate how nonvoters would vote if they were to vote (e.g., Brunell and DiNardo 2004) face similar issues because the observed turnout is the result of specific appeals and resource allocations by political actors. In general, these appeals and allocations would be different if the actors could assume a given level of turnout irrespective of their behavior.

[7]See Dawid 2000 for an example and Brady 2002 for a general review of causal theories.

[8]Holland 1986; Rubin 1990, 1978, 1974; Splawa-Neyman 1990 [1923].

randomly chosen subset (to contact) and we assign the remainder to the control group (not to be contacted). We then observe the difference in turnout rates between the two groups and we attribute any differences to our treatment.

In principle the process of random assignment results in the observed and unobserved baselines variables of the two groups being balanced.[9] In observation research it is possible to only balance observed variables although it is possible to conduct sensitivity tests for the influence of unobserved bias (Rosenbaum 2002, 105–170). The issues involved with balancing observed covariates in observation data are discussed in detail in the subsequent sections.

One may conjecture that a match for Jane could be found using only cross-sectional data. There are several problems with this conjecture. For instance, without data over time, it is all but impossible to decide what is baseline and what is post treatment. An example of the problem can be seen by examining Bartels's (1996) cross-sectional analysis of the effect of the level of information on voting in U.S. presidential elections. In his models, he does not condition on variables such as partisanship and ideology which are usually included in vote models because he reasonably conjectures that these variables may be the result of information—i.e., be post-treatment variables. The model includes age, age squared and indicator variables for union membership, being black, sex, marital status, religious affiliation, profession, urban residence and region of the country. Years of education and income percentile are also included—variables which could be post-treatment. Because the estimated vote model excludes variables which are known to be important for the vote, such as partisanship, ideology and all other non-demographic variables, the model does not fit the data well and many extreme y-misclassification errors are present. With such outliers, inferences based on maximum-likelihood are generally inconsistent (Kunsch, Stefanski, and Carroll 1989; Hampel, Ronchetti, Rousseeuw, and Stahel 1986; Mebane and Sekhon 2004) while robust estimators support reliable inferences in either case (Cantoni and Ronchetti

[9]This occurs with arbitrarily high probability as the sample size grows. Balance implies that treatment assignment is independent of baseline covariates.

2001; Hampel et al. 1986; Huber 1981).[10] When the Bartels analysis is replicated using robust estimation, no information effects are found.[11] The robust estimator used is a new robust binary logistic model developed by Sekhon (2004b) which combines the down weighting of y-misclassification outliers of the conditionally unbiased bounded-influence approach of Kunsch et al. (1989) with a high breakdown point Mallows class estimator for down weighting x-outliers (Carroll and Pederson 1993). See Appendix A for details.

The point of the Bartels replication is not that doing the cross-sectional analysis using robust estimation is the correct way of proceeding. The point is that with cross-sectional data it is all but impossible to decide what is baseline and what is not. And if one tries to be principled about what is conditioned on and what is not, one is only able to condition on a handful of variables, variables which lead to a very poor model of voting. If one assumes that Bartels's parametric specification is correct, the implicit assertion is that once a handful of demographic variables have been (parametrically) balanced, informed and uninformed respondents are exchangeable aside from levels of information. This is a very strong assumption and one which is doubtful on theoretical grounds given the literature on voting showing that non-demographic variables significantly differentiate voters—unless one assumes that all of these non-demographic variables are *entirely* the result of information levels and the demographics conditioned on. This conjecture is shown to be empirically problematic by the robust estimation results which demonstrate that the model is not a good model for much of the data because the model produces a large number of y-misclassification outliers and that the inferences significantly change once the outliers are down-weighted.

When attempting to estimate causal effects, examining changes over time greatly reduces (but does not eliminate) the problem of deciding what is baseline and what is not. And as discussed in detail below, it becomes more tenable to argue that once all baseline variables

[10] Note that the extent of the y-misclassification issue cannot be determined by using maximum likelihood and simply examining the residuals because of the problem of masking (Atkinson 1986). One must use a robust procedure in order to determine the extent of the misclassification or outlier problem (Mebane and Sekhon 2004).

[11] The information effects also vanish if we include partisanship, ideology, first order interactions and simply use maximum likelihood.

have been conditioned on (such as voting behavior, ideology and partisanship), the respondents are exchangeable. Using changes in panel data to improve specification is old advice. Indeed, the inherent logic is present in Yule's (1899) classic analysis of pauperism in England which is, to my knowledge, the first recognizably modern use of regression in the social sciences. This is also the logic behind Difference-in-Difference estimators (Smith and Todd forthcoming).

The question then, is whether voters who change information level—i.e., voters who become either less or more informed—behave any differently than those who do not. If information is causally important, changes in the level of information should have an effect on behavior. The outcome of interest is change in vote from baseline.

3.2 Causal Inference

Making causal inferences from political opinion surveys is a difficult task. Many (if not most) of the measures known to be correlated with vote choice are largely endogenous. But one of the virtues of opinion surveys is that more than fifty years of experience has gone into constructing measures which are correlated with vote choice and other politically salient variables. And much of the variance of vote choice is soaked up by the usual kitchen sink models. Because soaking up variance is not the same as explanation, I use a methodological approach which relies on this feature of the surveys and coherently deals with the problems of endogeneity and selection associated with observational data.

Surveys lend themselves to analysis by the Rubin causal model and associated methods (Holland 1986; Rosenbaum 2002; Rubin 1974, 1978, 1990) if they contain, as the NES do, a large set of relevant baseline covariates which help to make one of the key identifying assumptions of the approach tenable.[12] The Rubin model conceptualizes causal inference in terms of potential outcomes under treatment and control, only one of which is observed for

[12] Although it is often called the Rubin causal model, it goes back to at least Neyman (e.g., Splawa-Neyman 1990 [1923]).

each unit. A causal effect is defined as the difference between an observed outcome and its counterfactual.

Let Y_{i1} denote the change in voter i's vote intention from baseline when voter i changes information state (i.e., is in the treatment regime), and let Y_{i0} denote the change in vote choice when voter i does not change information state (i.e., is in the control regime). The causal inference problem is a missing data problem because both Y_{i1} and Y_{i0} cannot be observed—a given voter cannot both change and not change information state at the same time. Let T_i be a treatment indicator: 1 when i is in the treatment regime and 0 otherwise. The observed outcome for observation i is then $Y_i = T_i Y_{i1} + (1 - T_i) Y_{i0}$. The treatment effect for observation i is defined by $\tau_i = Y_{i1} - Y_{i0}$.[13]

In principle, if assignment to treatment is randomized, the inference problem is straightforward because the two groups are by construction drawn from the same population. The observed and unobserved baseline variables of the two groups are balanced; treatment assignment is independent of all baseline variables. This occurs with arbitrarily high probability as the sample size grows. With the independence assumption, the missing data problem is simple to resolve because treatment assignment is independent of Y_0 and Y_1—following Dawid's (1979) notation, $\{Y_{i0}, Y_{i1} \perp T_i\}$. Hence, for $j = 0, 1$

$$E(Y_{ij}|T_i = 1) = E(Y_{ij}|T_i = 0) = E(Y_i|T_i = j)$$

Therefore the average treatment effect (ATE) can be estimated by:

$$\begin{aligned} \tau &= E(Y_{i1}|T_i = 1) - E(Y_{i0}|T_i = 0) \\ &= E(Y_i|T_i = 1) - E(Y_i|T_i = 0) \qquad (1) \end{aligned}$$

It is possible to estimate Equation 1 because observations in the treatment and control groups are exchangeable. In the simplest experimental setup, individuals in both groups are

[13] For comparability with the literature, the notation here is the same as in Dehejia and Wahba (1999).

equally likely to receive the treatment, and hence assignment to treatment will not be associated with anything which may also affect one's propensity to vote in a particular fashion. Even in an experimental setup, much can go wrong which requires statistical correction (e.g., Barnard, Frangakis, Hill, and Rubin 2003; Imai forthcoming). In an observational setting, unless something special is done, the treatment and non-treatment groups are almost never balanced.

With observational data, the treatment and control groups are not drawn from the same population. Thus, a common quantity of interest, and the one I use, is the average treatment effect for the treated (ATT):

$$\tau|(T=1) \quad = \quad E(Y_{i1}|T_i=1) - E(Y_{i0}|T_i=1). \tag{2}$$

Unfortunately, Equation 2 cannot be directly estimated because Y_{i0} is not observed for the treated. Progress can be made if we assume that the selection process is the result of only observable covariates denoted by X. Following Rosenbaum and Rubin (1983), one can assume that once one conditions on X, treatment assignment is unconfounded ($\{Y_0, Y_1 \perp T\} \, |X$) and that there is overlap: $\delta < Pr(T=1|X) < 1-\delta$, for some $\delta > 0$. Unconfoundedness and overlap together define strongly ignorability. These conditions are required to identify ATE. Heckman, Ichimura, Smith, and Todd (1998) demonstrate that to identify ATT the unconfoundedness assumption can be weakened to mean independence: $E\left[Y(w)|T,X\right] = E\left[Y(w)|X\right]$, for $w = 0, 1$.[14] The overlap assumption for ATT only requires that the support of X for the treated is a subset of the support of X for the control observations.

Then, following Rubin (1974, 1977) we obtain

$$E(Y_{ij}|X_i, T_i=1) = E(Y_{ij}|X_i, T_i=0) = E(Y_{ij}|X_i, T_i=j). \tag{3}$$

[14] Also see Abadie and Imbens (2004).

Therefore, conditioning on the observed covariates, X_i, the treatment and control groups have been balanced. The average treatment effect for the treated can then be estimated by

$$\tau|(T=1) \quad = \quad E\left\{E(Y_i|X_i, T_i=1) - E(Y_i|X_i, T_i=0) \mid T_i=1\right\}, \tag{4}$$

where the outer expectation is taken over the distribution of $X_i|(T_i = 1)$ which is the distribution of baseline variables in the treated group.

The most straightforward and nonparametric way to condition on X is to exactly match on the covariates. This is an old approach going back to at least Fechner (1966 [1860]), the father of psychophysics. This approach of course fails in finite samples if the dimensionality of X is large or if X contains continuous covariates. Alternatively, one may employ nearest-neighbor matching based on some distance metric, such as Mahalanobis distance (Rubin 1980). If X consists of more than one continuous variable, nearest-neighbor matching is inefficient and has a bias term which does not asymptotically go to zero at $\sqrt{n}$ (Abadie and Imbens 2004).

An alternative way to condition on X, is to match on the probability of being assigned to treatment—i.e., the propensity score.[15] If one matches on the propensity score, one will obtain balance on the vector of covariates X (Rosenbaum and Rubin 1983).

Let $p(X_i) \equiv Pr(T_i = 1|X_i) = E(T_i|X_i)$. This defines $p(X_i)$ to be the propensity score. If we assume that $0 < Pr(T_i|X_i) < 1$ and that $Pr(T_1, T_2, \cdots T_N|X_1, X_2, \cdots X_N) = \Pi_{i=1}^{N} p(X_i)^{T_i}(1 - p(X_i))^{(1-T_i)}$, then as Rosenbaum and Rubin (1983) prove,

$$\tau|(T=1) \quad = \quad E\left\{E(Y_i|p(X_i), T_i=1) - E(Y_i|p(X_i), T_i=0) \mid T_i=1\right\},$$

where the outer expectation is taken over the distribution of $p(X_i)|(T_i = 1)$.

Since $p(X_i)$ is generally unknown, it must be estimated. While the logistic model is often used, a variety of semi-parametric approaches may be used instead. But it is important

[15] The first estimator of treatment effects to be based on a weighted function of the probability of treatment was the Horvitz-Thompson statistic (Horvitz and Thompson 1952).

to keep in mind that none of the coefficients of this model are of interest and individual coefficients do not need to be estimated consistently. Indeed, the propensity model does not need to be uniquely identified—i.e., the coefficient estimates need not be unique for a given propensity value. Furthermore, it is straightforward to prove that for discrete X (such as the X observed from survey data) the weighting function does not matter asymptotically as long as it is nondegenerate—i.e., the weighting function maps every unique combination of X to a unique propensity so there are no zero weights assigned to any given X.

There are five important implications of the forgoing that I wish to highlight. First, the modeling portion of the estimator is limited to the model of $p(X_i)$. Estimation of this model requires **no** knowledge of the outcome. Once balance has been obtained, any outcome can be estimated. Hence, one can conduct model selection (which centers on achieving balance in the X variables) without ever observing what the estimated outcome is under the various models. This is an incredibly important virtue of this approach, and it cannot be stressed enough given the problems of data mining in observational work. Unlike in the regression case, there is a clear standard for choosing an optimal model; it is the model which balances the covariates, X. All of the propensity models in this paper were chosen before any outcome was estimated.

Second, the key assumption required is that no variable has been left unobserved which is correlated with T_i *and* with the outcome. This is called the unconfoundedness assumption. If such a variable exists, there is hidden bias. The rich set of covariates in the surveys analyzed make this assumption more plausible than in most observational datasets. But all observational work is open to the criticism of hidden bias. Fortunately, rigorous sensitivity tests are available to determine how robust the results are to hidden bias (Rosenbaum 2002, 105–170).

Third, once the propensity score has been matched on (or if one uses direct matching), any outcome quantity of interest can be simply calculated, be it the mean or any quantile (Rosenbaum 1999). Fourth, no functional form is implied for the relationship between treatment and outcome. No homogeneous causal effect assumption has been made; the causal effect may vary with the propensity score (and hence with values of X_i).

Finally, use of this approach (unless considerable adjustments are made) requires the assumption that "the observation on one unit should be unaffected by the particular assignment of treatments to the other units" (Cox 1958, §2.4). Rubin (1978) calls this "no interference between units" the stable unit treatment value assumption (SUTVA). SUTVA implies that the potential outcomes for a given voter do not vary with the treatments assigned to any other voter, and that there are no different versions of treatment. The first part of the SUTVA assumption is true by construction in the NES because the sampling procedure all but ensures that none of the individuals in the sample know each other. For example, only one person is interviewed in each household.[16]

The SUTVA assumption is violated by the general counterfactual mentioned in the previous section; the counterfactual about everyone being well informed. As more and more people become informed, it is likely that voters in the control group will be influenced because of the changing behavior of political activists and other actors (including the newly informed voters themselves). This issue could be dealt with if one were to have a good model of the resulting interference among voters. But, alas, I know of no such model.

Because we are taking a conditional expectation, we need to decide what to condition on. If we condition on too little, our estimates are confounded and therefore biased. If, however, we condition on variables which are not baseline variables but the result of treatment, we obtain biased estimates because of post-treatment bias (Rosenbaum 2002). The panel nature of the research design greatly aids this issue. Every single variable observed at baseline is balanced (matched on). For most datasets this results in balance being obtained for several hundred variables. As will be discussed in the country specific sections, balancing on so many variables is possible because variables in survey data are highly correlated. One isn't balancing several hundred nearly orthogonal variables. The actual dimensionality of the data

[16]There is the issue that cluster sampling is used instead of simple random sampling, but that surely is an unimportant distinction given the size of the sample and population.

is greatly less.

3.3 Estimation of Causal Model

One-to-one nearest-neighbor matching with replacement is used because it provides the best balance, and because theory dictates that one-to-one matching minimizes the expected bias. The actual matching is done using "Matching" which is an **R**[17] package developed by me and available at `http://jsekhon.fas.harvard.edu/matching`. The standard errors produced by "Matching" coherently and deterministically take into account ties and sampling controls with replacement when finding matches—for details see Abadie and Imbens (2004).[18] A combination of multivariate and propensity score matching is done the details of which are explained in the country specific chapters. All closed form questions in the baseline survey are matched on. Exact matching is done on partisan identification and previous vote and everything else is matched on using propensity score matching. The results do not significantly change if only propensity score matching is used, but the observed balance across covariates is worse. A caliper is used when needed.

Because of the highly correlated nature of survey data, it was found that the optimal propensity model was based not on the raw data but on principal components of the baseline variables. Principal components were taken of all of the observed variables and no ordering is assumed (e.g., seven point scales are turned into separate dummy variables for each category including an eighth category for no response). All first order interactions are included. And for continuous variables, such as age, squared terms are also included. Observations with missing values are not dropped, but matched on. Thus, the usual seven point scales are really eight point scales. The principal components are then entered into the propensity model which is estimated by both ML and the robust logistic model previously discussed.

[17]See `http://www.r-project.org/`.

[18]The estimated treatment effect and standard errors are based on linear regression adjustment (based on the propensity score) on matched-pair differences. This has been shown to help reduce bias even with incorrect specification (Abadie and Imbens 2004; Rubin 1973, 1979).

The principal components greatly reduce the dimensionality of the data. The results based on the ML models are reported because the robust logistic model did not significantly change inferences.

When reviewing the matching strategy, it is important to keep in mind the result previously mentioned that for discrete X the weighting function does not matter asymptotically as the weighting function maps every unique combination of X to a unique propensity. Also note that use of the principal component approach is *not* necessary to obtain the reported results, but it does greatly simplify the propensity model.

Both univariate and multivariate balance is evaluated. Balance is evaluated using the observed data and *not* the principal components. First, univariate balance is judged by two non-parametric tests. The first, the Wilcoxon rank sum test (Wilcoxon 1945), is well suited for testing for differences in the first moment and its properties have been extensively studied (Hettmansperger 1984). This test is equivalent to the Mann and Whitney test (Mann and Whitney 1947). For paired binary data, I use the McNemar test of marginal homogeneity (McNemar 1947).

Univariate tests are not sensitive to the relationships between variables. So the main test for balance is to compare the two propensity distributions (for treated and control) using the Kolmogorov-Smirnov test for equality. This is a nonparametric test based on the Kolmogorov distance between the two empirical distribution functions (Knuth 1998; Wilcox 1997). A bootstrapped Kolmogorov test is used because it provides consistent tests levels even in the presence of point masses (Abadie 2002) which are a concern given the categorical nature of survey data.[19] Another multivariate test is offered by running a logistic model in which the dependent variable is treatment assignment and all of the baseline variables are

[19]To be clear, only the Kolmogorov test is bootstrapped (conditioning on the matched data) not the matching procedure because there is concern that bootstrapping the latter does not provide consistent estimates. No theorems exist demonstrating that bootstrapping matching estimators produces consistent coverage, and Monte Carlos by both Sekhon and Abadie and Imbens (personal communications) show that bootstrapping in fact does not provide consistent coverage for confidence intervals. One-thousand resamples are used when the Kolmogrov test is bootstrapped.

entered on the right hand side.[20] If the residual deviance of this model is significantly less than the null deviance, there is evidence against balance.

3.4 Panel Attrition

Given the use of panel data some may be concerned that panel attrition is a significant issue. And indeed there are reasons to be concerned that panel attrition is correlated with information measures. But concern about panel attrition should not be exaggerated. Bartels (1999) presents evidence that panel attrition only rarely leads to significant bias with NES data.

In any case, multiple imputation provides a straightforward way in which to correct for panel attrition. Because the NES and Mexico surveys include fresh (non-panel) samples, a relatively general correction for attrition is possible. This correction, developed by Hirano, Imbens, Ridder, and Rubin (1998) and called the additive nonignorable (AN) model, allows for attrition based on lagged variables and for attrition based on contemporaneous variables. The former type of attrition is called *selection on observables* because it relies on the *missing at random* assumption of Rubin (1976) and Little and Rubin (1987), and the latter type of attrition is called *selection on unobservables* because attrition partly depends on variables that are not observed for respondents who drop out. The former was developed by Rubin (1976) and the latter by Hausman and Wise (1979). When the AN correction is applied, my substantive results do not significantly change in the U.S. and Mexico. Because imputation does not change the substantive results, the non-imputed results are presented in this paper.[21]

[20]The test was also run including all first order interactions, and the inferences were the same.

[21]Imputation is done using **Splus** code developed by me which adapts Schafer's (1997a; 1997b) imputation code for the purpose of correcting for panel attrition using the AN model for categorical data. Schafer's code is available at `http://www.stat.psu.edu/~jls/misoftwa.html`.

4 Political Information in the United States

4.1 Data

Since we are interested in estimating the causal effect of changes in the level of information, we require panel data to conduct the analysis. We also obviously need good measures of information, the outcome of interest, and a set of baseline variables sufficient to make the unconfoundedness assumption reasonable. Recent NES provide these measures but panel studies are rare. Moreover, before 1968, the information questions are of poor and highly variable quality. In particular, the 1956-1958-1960 panel is unusable because the 1956 survey does not contain sufficient information indicators.

With these restrictions in mind, three different NES datasets are analyzed: the 1980, 1972-1974-1976 and 1992-1994-1996 panel studies. The 1980 panel study allows one to estimate the causal effects of changes in information levels during an election campaign—i.e., the effects of learning during a campaign. Many commentators have lamented what they consider the low level of attention which voters give to even presidential campaigns (Patterson 2002). This dataset allows us to see what effects, if any, such inattentiveness has. The two multi-year panel studies allow us to examine if changes in information states over a medium period of time have causal effects.[22]

The 1980 panel study consists of four waves: (1) the primary wave, from January 22 through February 25; (2) post primary wave, from June 4 through July 13; (3) post convention wave, from September 2 through October 1; (4) post-election wave, from November 5 through November 25. Baseline is defined as the first wave. Two different treatments are of interest. The first is defined as change in information state between the first and second wave and the second as change in information state between the first and third wave. For the first definition of treatment, there are two outcomes of interest: the third (September)

[22]The counterfactual is somewhat clearer in the 1980 panel study than the multi-year panels. In the 1980 panel study, the counterfactual is one of a given voter paying more or less attention to the campaign. Although the data show that information states change in the multi-year panels (even as judged by answers to objective unchanging questions), it is less clear exactly what the counterfactual at work is.

and fourth waves (November). For the second treatment only one outcome period can be considered: November.[23]

For both of the multi-year panels, baseline is defined as the first presidential election observed which is either 1972 or 1992 (post-election survey). Two different treatments are considered. The first is the change in information from the first presidential election to the midterm election (post-election survey). And the second is the change in information from baseline to the pre-election survey of the second presidential election. For the first treatment, three different outcomes are of interest: September, October and November of the second presidential year. For the second treatment only November of the second presidential election year can be considered

In the 1996 survey, whether a respondent was interviewed in September or October was randomly assigned. More precisely, four random replicates were used in the pre-election survey. They were released as follows: September 3, September 12, September 26, and October 10. Therefore, I consider the September observations for the 1996 survey to be those which come from the first two replicates and the October observations to be those which come from the last two replicates.[24] In the 1980 survey, the September observations are from the third panel wave which was conducted between September 2 and October 1. There were no panel interviews in October. The 1976 survey poses a problem because the month respondents were questioned was not randomized in the pre-election survey, and randomization tests show (p=0.000) that whether a respondent was interviewed in September or October was *not* random. Moreover, the 1976 pre-election survey did not begin interviewing respondents until well into September: September 17. Therefore, for the 1976 survey, pre-election results broken down by month are not presented.

[23] If one insists on examining the outcome in the third wave even though treatment is also measured in that wave, the results don't substantively change from those obtained by examining the causal effect on the third wave of information changes between the first and second.

[24] If respondents from the first two replicates were not interviewed by September 26, they are not included in the data analysis. If these deleted respondents are included, the results do not significantly change. I thank Charles Franklin for noting that there is some concern that the randomization failed. And balance tests indeed show that there is some imbalance. But the observations are being matched on baseline covariates so a balance adjustment is already being made.

4.2 Measurement

Let $I_{i,t}$ denote the information state of voter i at time t. Information is measured in five different ways: the ability to spatially locate the candidates in issue space, interview ratings, the Zaller information measure,[25] the Alvarez (1997, 69–72) measure[26] and the Alvarez measure scaled using the Aldrich and McKelvey (1977)[27] scaling method. Note that the Zaller measure is only used for the multi-year panels and the measure based on the ability to spatially locate the candidates is only used in the 1980 intra-election panel.

All of these measures provide substantively similar results. Hence, results for the Zaller measure, the Alvarez measure and the Alvarez measure scaled using the Aldrich and McKelvey measurement model are not presented. Results based on the interview ratings are presented first for comparability with Bartels (1996). Results are then presented for the spatial location measures.

The interviewer information measure is a five point scale which ranges from "very low" to "very high". Change in treatment is defined as changes in this scale from baseline. This is a non-binary treatment regime because one could move from category 1 to 2 or from 1 to 3 or from 4 to 2, etc. All of the multinomial treatment effects have been estimated using the approach of Imbens (2000).[28] But the following simple binary treatment definition is representative of the multinomial results. Voter i is considered to be in the treatment regime if $I_{i,t_1} \neq I_{i,t_2}$, then $T_i = 1$. In other words, the respondent is considered to be in the treatment regime if there was any change in his or her information level. Table 1 presents the total

[25] Because I need to make comparisons over time, the question arises of how to compare the Zaller information measure across time. The simple approach of making the Zaller index comparable over time by looking at what quantile of the measure the respondent falls into has been tried, and the substantive results are the same as those reported here. I have also used only the subset questions of Zaller's information measures which are repeatedly asked in the time-periods in question. This modified Zaller measure yields the same substantive results as those reported here based on the interviewer information ratings.

[26] The Alvarez measure consists of estimating voter i's uncertainty about a given candidate as the squared difference between voter i's placement of a given candidate on an issue dimension and the mean location all voters assign the candidate. Alvarez uses the mean dispersion across the issue questions. My implementation uses the median instead of the mean and maintains the discreteness of the data. Alvarez assigns those with no opinion to the category which results in the maximum uncertainty.

[27] Also see Palfrey and Poole (1987).

[28] Also see Rosenbaum (2002, 300–302).

number of respondents who increased or decreased information state in the various datasets under consideration. The table does not include observations lost due to panel attrition but recovered by imputation.

The interviewer information measure is highly correlated with the Zaller method and is hence a measure of general political knowledge. Both measures are somewhat distant from the task at hand which is casting a vote in a particular election. Therefore, results for an alternative more specific measure are also presented. This more specific measure is preferred because based on it there is clear evidence of learning during the campaign.

Using the ability to locate candidates measure, a voter is considered informed if she can place a given candidate on an issue space. If she cannot, she is considered less informed.[29] Results for this measure are only presented for the 1980 panel study. Table 2 presents relevant descriptive statistics including the percentage of voters who could place Reagan and Carter on various issue positions during the three pre-election waves of the 1980 panel.[30] The fourth column displays the proportion of all respondents who placed Carter to the left of Reagan. And the fifth column of the table displays the proportion of respondents who placed Carter to the left of Reagan of those who could place both candidates.

It is apparent from Table 2 that the proportion of respondents who could place Reagan on the issue dimensions greatly increased between the first and third panels and that most of the increase occurred between the first and second panels. For example, in the first panel (conducted in January) 62% of respondents could place Reagan on the usual liberal-conservative dimension. By the second panel (June) 71% of respondents could place him and a statistically indistinguishable 69% could in the third wave (September). In January, only 50% of respondents could place Reagan on the dimension asking about the trade-off between unemployment and inflation, but 60% could in June. For both of the spending questions

[29] Bartels (1986) presents a model of nonresponse which measures voter uncertainty by estimating a linear probability model where the dependent variable is nonresponse on an issue scale. Unfortunately, the parametric model explains relatively little of the variance in nonresponses—only about a quarter. Nevertheless, in a future draft, I will, in addition to all of the current measures, estimate the causal effects of this parametric measure.

[30] The issue placements were not asked in the fourth wave.

significant learning continued to occur between June and September. 60% of respondents could place Reagan on the issue of increasing or decreasing defense spending in January, 72% in June and 80% in September. 61% of respondents could place Reagan on the issue about the trade-off between government spending and (domestic) government services in January. This increased to 69% in June and 75% in September.

Compared with Reagan, the proportion of respondents who placed Carter doesn't change much during the course of the campaign. This is not surprising given that Carter was the incumbent president. At the start of the campaign, Carter is a relatively known quantity and Reagan is not. What does also change significantly during the course of the campaign is the percentage of respondents who can successfully place Carter to the left of Reagan in issue space. The fifth column in the table reports the proportion of respondents who can do this of those who also placed both candidates. In January, 65% of those who could place both candidates placed Carter to the left of Reagan on the liberal-conservative dimension. By June 74% did so and a statistically indistinguishable 77% did so in September. There is also a significant change for the defense spending issue: in January 58% of placers located Carter to the left of Reagan, by June it was 69% and by September it was 79%.

As determined both by the ability to place the candidates at all and by the ability to place Carter to the left of Reagan, there was learning during the campaign.

Treatment effects have been estimated for each of the five measures separately where treatment has been defined in two different ways. The first treatment of interest is going from being unable to place Reagan to being able to place him. Given that the Carter placement proportions do not change much (since he is the incumbent), Carter is not considered. The second treatment is defined as going from being unable to place Carter to the left of Reagan (either because the respondent could not place the candidate(s) at all or because the respondent did not place them correctly) to being able to do so. The second measure requires more from the respondent than the first. Hence, there are 10 different treatment definitions. For all ten no significant treatment effect is found in November and a significant effect is found in September for most of them.

Instead of presenting a blizzard of numbers, the results based on a treatment definition which combines the five measures is presented. The results of this combined measure are representative of the individual measure results. The combined measure is defined as follows: a respondent is considered to be uninformed if he or she is unable to place Reagan on at least one of the five issue questions and informed if he or she is able to place him on all of the questions. The treatment of interest is if a respondent goes from being uninformed to being informed.

This dichotomous combined issue space information measure is correlated with the interviewer information measure to a reasonable degree, and the correlation increases during the campaign. The correlation between the interviewer rating and this measure in wave one is 0.301, in wave two it is 0.356 and in wave three it is 0.400.

Many different outcomes can be considered. Although many have been estimated, the results of two representative outcomes are presented. For the interviewer rating measure, the outcome is defined in a fashion which parallels the treatment definition. Let $V_{i,t}$ denote voter i's vote intention or choice at time t. V equals 1 if the voter prefers the Republican presidential candidate, V equals 2 if the voter prefers the Democratic candidate, V equals 3 if the voter prefers a third party candidate, and V equals 4 if the voter is undecided or if the voter did not vote. The outcome is denoted by $Y_{i,t}$, and $Y_{i,t} = 1$ if $V_{i,1} \neq V_{i,t}$. This outcome simply measures if the vote intention of the respondent changed in any fashion.

For the Reagan placement measure, results are presented for the previous outcome—i.e., for any change. And for a change specific to voting for Reagan. This second outcome is defined as one if the voter changed from or to the Republican candidate and zero otherwise. Results for this second outcome are presented because most of the movement in votes between January and September which account for the first significant effect are due to people moving in the Republican direction as they become more informed. This parallels the descriptive fact that, in the 1980 panel, informed voters moved away from Carter earlier than uninformed voters.

4.3 Balance

In order to correct for confounding of observed variables in treatment assignment, a combination of exact and propensity score matching is done. Because most all of the variables measured in the NES are discrete, it is theoretically possible to use exact matching only, but the curse of dimensionality soon gets one: there simply are too many variables and too few observations to do this. So instead, a hybrid approach is used. Exact matching is used for baseline partisan identification and whether the respondent voted for the Republican presidential candidate in the baseline survey. Both variables are highly correlated so exact matching on both is not an unrealistic demand. Balance can then be significantly improved by using propensity score matching for the other variables in the baseline NES survey.

It is worth noting that although all of the closed-form baseline questions are used in the propensity setup described in Section 3.3, this is not necessary to obtain the reported results nor to obtain good balance. It has been found that the more conventional approach of using the following baseline variables achieves balance on not only these variables but for most all of the other variables in the NES. These variables are: the 7-point liberal-conservative ideology scale, education, home ownership status, retired, housewife, union membership, Hispanic, black, east, south, west, religion, age, family income, the changing list of seven point policy scales and retrospective economic evaluations. All variables except for age, education and income are included in the propensity model as fully factored indicator variables.[31]

The multivariate balance tests for the interviewer based information measure are reported in Table 3. The table shows that before matching there is a great deal of imbalance and

[31] Age is included as is the square of age. Education is included (as measured by years of schooling) and an indicator variable for whether the respondent has more than the median years of schooling. Family income is ranked and divided up into five quantiles, and an indicator variable is included for each quantile. Observations with missing values are not dropped, but matched on. Thus, the usual seven point scales are really eight point scales. First order interactions are added for those variables found to be unbalanced without them.

that after matching balance has been obtained. The table presents the p-values for both the Kolmogorov-Smirnov and likelihood-ratio tests for equality of the propensity score densities of those who changed information state and of those who did not. For the 1990s panel and information change between $t = 1$ (1992 post-election survey) and $t = 2$ (1994 post-election survey), the Kolmogorov-Smirnov test yields a p-value of 0.00 before matching and a p-value of 0.981 after matching. The ridiculously low p-value before matching should not be interpreted literally, but the conclusion is clear: the two groups are profoundly imbalanced before matching and are balanced afterwards. For the 1980s panel, the test yields a p-value of 0.00 before matching and a p-value of 0.910 after matching. For the 1970s panel, the test yields a p-value of 0.00 before matching and a p-value of 0.746 after matching. The likelihood ratio tests tell the same substantive story. The results are the same for changing information between $t = 1$ and $t = 3$.

Figure 1 displays the densities of the propensity scores for the voters who changed information level (as measured by interviewer ratings) and for those who did not. The figure visually confirms what the formal tests just demonstrated: the densities of the two groups are markedly different before matching and after matching are almost exactly the same.

Table 4 presents the balance tests for the issue scale measure. Like the previous table, it shows that there is profound imbalance before matching and balance afterwards.

4.4 Results

Table 5 presents the sample average treatment effect (ATT) estimates for the effect of changing interviewer information levels on changing vote intention or choice. There are significant effects in September but none in November. In all surveys in November, the standard errors are larger than the estimated effects. The estimated effect in September is largest in the 1990s panel in which voters whose level of information changed between 1992 and 1994 were 34% more likely to change their partisan vote intention between 1992 and September 1996 than voters whose level of information did not change. The estimated causal effect in this panel is more than three times larger than the causal effect estimated in the 1980 panel, but these two effects are impossible to compare because the changes being considered in the 1980 panel occur over a much shorter time-period.

Although the September effect for the 1980 panel is smaller than for the 1990s panel, it is still substantively large. In 1980, voters whose information level changed between January and June were 11% more likely to change their partisan vote intention from January to September than voters whose information level did not change. Recall that because the 1976 NES survey did not randomize whether respondents were interviewed in September or October, monthly pre-election estimates are not presented for 1976. For completeness, I have estimated for the 1976 pre-election survey (which interviewed respondents between September 17 and November 1) the effect of changing information between 1972 and 1974. This estimate is 0.006 with a standard error of 0.051. Given the results for the other panels, an insignificant estimate is to be expected because most of the pre-election respondents were interviewed in October and because no interviews at all were conducted in the first half of September.

All of these results tell a consistent story. The effect of changing information state between $t = 1$ and $t = 3$ in November all are insignificant, just like the November effects of changing information between $t = 1$ and $t = 2$. By election day, there is no effect of information on changing one's vote. These results show that between September and election day something significant changes in American political life which negates the effects of information on voting behavior.

Table 6 presents the results for the issue scale measure of information. The two outcomes are any change in the Republican vote and any change at all in vote intention. The table shows that regardless of whether information changes between January and June (1980) or January and September are considered, there are no significant information effects by election day. In September, however, respondents with increasing information are 18% more likely to change their vote preference than people whose information level did not change. This result is being driven by the fact that as people became informed, they started switching to the Republican presidential candidate, Ronald Reagan. If the outcome is restricted to moving from *not* intending to vote for Reagan in January to intending to vote for Reagan in September, the effect is about 19% (t-value=2.23). By November there is no treatment effect. As people became informed, they moved before their less poorly informed compatriots, but their less poorly informed compatriots also eventually switched their vote intentions.

5 Political Information in Mexico

The analysis in this section is based on the Mexico 2000 Panel Study which was explicitly designed to measure campaign effects and voting behavior in Mexico's 2000 presidential election.[32] The survey consists of about 7,000 interviews and four panels. The first survey was conducted just after the beginning of the campaign in February 19–27 and consisted of a national sample of 2,400 adults. A random half of the first survey was reinterviewed in the second wave which was in the field from April 28 to May 7. The second wave consisted of about 950 respondents. The third wave was conducted from June 3 to 18 and consisted of panel respondents from the second wave plus some randomly chosen respondents from the first. The fourth (post-election) wave was conducted between July 7 and 16. The last wave included a refreshment sample of about 1,200 respondents. This refreshment sample is used to correct for panel attrition as described in Section 3.4

The July 2, 2000 election was an important one for Mexico for it brought to an end the world's oldest one-party regime. Vicente Fox of the center-right National Action Party (PAN) defeated Francisco Labastida of the long ruling Institutional Revolutionary Party (PRI). Many analysts have concluded that the campaign itself was important for Fox's

[32]Participants in the Mexico 2000 Panel Study included (in alphabetical order): Miguel Basañez, Roderic Camp, Wayne Cornelius, Jorge Domínguez, Federico Estévez, Joseph Klesner, Chappell Lawson (Principal Investigator), Beatriz Magaloni, James McCann, Alejandro Moreno, Pablo Parás, and Alejandro Poiré. Funding for the study was provided by the National Science Foundation (SES-9905703) and *Reforma* newspaper.

victory because at the beginning of the campaign it appeared that Labastida would win easily (Domínguez 2004; Lawson 2004).

The data show that a significant amount of learning occurred during the campaign not only concerning the particular issue positions of the candidates and parties but also about the very nature of the Mexican political system. Most of the learning appears to have occurred between the first (February) and third panels (June). Table 7 presents descriptive statistics for some opinion and information measures over the campaign. At the beginning of the campaign, only 40% of Mexicans responded affirmatively when asked if Mexico is a democracy. By July, the percentage had increased markedly to 63%. Not only did Mexicans generally feel better about their political institutions, they also learned details about them. For example, at the beginning of the election campaign only 37.8% of respondents could name the judiciary when asked to name the three branches of government. By July, the number had climbed to 50.4%. This election had profound effects on public opinion in Mexico; some commentators have called it a revolution because it significantly changed attitudes towards democracy and the country's political institutions (Camp 2004).

There are three types of information measures one could use in the Mexico 2000 Panel Study. The first concerns general political information such as the ability to name the three branches of government. The second involves the ability to place the parties and candidates in issue space, and the third is the ability to recall specific campaign events such as the ability to correctly associate a slogan with the party which used it. Table 7 shows that by all three types of measures, the information level of the electorate increased during the campaign. For example, in February 67% of respondents could place PAN on the liberal-conservative space. This increased to 74% in June and to 81% in July. The numbers for PRI are similar: 68% of voters could place PRI on the liberal-conservative space in February while 74% could in June and 80% in July.

Many respondents also learned to identify the slogans used by the parties. For example, the Fox slogan "Enough Already" was accurately identified as such by 50% of respondents in February, 73% in June and 80% in July. 40% of respondents could correctly identify the slogan "Power should serve the people" as a Labastida slogan in February while 70% could in June and 72% in July.

For this survey baseline, $t = 1$, is considered to be the first wave (February). And the treatment of interest consists of change in information from baseline to the third wave (June). The outcome of interest is any change in voting behavior in the fourth wave (post-election) from baseline.

Table 8 presents the multivariate balance test results for three representative information measures before and after matching. The three measures are the ability to place PAN on the liberal-conservative scale, the ability to so place PRI and the ability to mention the judiciary when asked to name the three branches of government. It is clear, as it was in the U.S. case, that there is profound imbalance before matching and good balance afterwards. For example, for the measure based on the ability to place PAN, the bootstrapped Kolmogorov-Smirnov test has a p-value of 0.00 before matching and 0.806 afterwards. The χ^2 test goes from 0.019 before matching to 0.901 afterwards. The results for the other measures are similar.

Table 9 presents ATT estimates for increasing information level (between panels 1 and 3) on changing one's vote in the post-election survey from baseline. For all three measures, there is a significant treatment effect even though we are analyzing the outcome from the post-election study. For example, respondents who are able to place PAN on the liberal-conservative space in the third wave but not the first are 21.4% more likely to change their vote on election day from baseline than respondents who continue to be unable to place PAN. This result is largely driven by the fact that as voters learned about the PAN they were more likely to defect from the PRI. By election day, respondents who had expressed an intention of voting for the PRI at baseline were 14% more likely to defect to another party if they learned to place the PAN than voters who were still unable to place PAN.

The results for learning to place PRI on the liberal-conservative space are similar to the results for PAN. Voters who learned to place PRI were 17% more likely to change their vote from baseline than voters who did not learn to do so. And as before, this result is driven by the fact that voters who expressed a PRI vote intention at baseline were 15% more likely to defect to another party by election day if they had learned to place PRI than voters who were still unable to place the party.

The results for learning about the judiciary are consistent with those of the previous two measures. Voters who learned about the judiciary were 21% more likely to change their baseline vote. And such voters were 12% more likely to switch from the PRI than PRI voters who were still unable to mention the judiciary when asked to name the three branches of government. Once again, as voters became better informed, they were more likely to move away from PRI—the party which long dominated Mexican politics and the party which in February looked likely to win the 2000 presidential election.

The results in Mexico are sharply different from those obtained in the U.S. Unlike in the U.S., information effects are found on election day. This difference is sharpened by the fact that the Mexico 2000 Panel Study is remarkably similar in design to the 1980 NES Panel Study.

6 Political Information in Other Countries

********* Included in Book Manuscript *********

7 Conclusion

Although I present no evidence of the mechanisms by which voters are able to perform information arbitrage in the U.S., there are many implications of their ability to do so. This is an important demonstration of the power of electoral institutions in the U.S. which along with markets are the primary methods by which individual preferences are aggregated. I conjectured that information effects would be present even on election day if the electoral institutions which provide cues were less mature. This conjecture has been borne out by the results in Mexico where information results are present even on election day. The implication is that as democracies develop, citizens no longer need to stay informed in order to act as if they are. Our current state of relative political disengagement is a high achievement and not a social failing to be lamented as is so often done. Given opportunity costs, it is unreasonable to advocate that all citizens be politically well informed especially since they can rely on institutions, such as polls and interest groups (McKelvey and Ordeshook 1985a,b, 1986), to make choices as if they are informed.

Notwithstanding the foregoing, the results in this paper should not be taken to imply that voters make choices which are in some general sense optimal. Voting is a simple act which in American politics is, for most voters, a binary choice: Republican or Democrat. Political parties work hard to project a consistent "brand image" which further simplifies the decision for voters. Also, the information measures available in surveys do not discriminate between individuals who are *highly* expert and informed about narrow issues (say issues related to the regulation of particular companies and industries) and those who are not. In any case, such expert voters are an extremely small proportion of the population; almost certainly too small of a proportion for samples of the size of the NES to include a significant number of such voters.

Much of the work of government involves highly technical issues which are not discussed in general political discourse—such as the issue of who benefits from the government debt guarantee for Fannie Mae and Freddie Mac and the extent to which such a guarantee constitutes a moral hazard. There is no evidence in this paper that the public understands, or even knows of, such issues. Thus, one should not conclude from these results that voter behavior, political rhetoric, and policy outcomes would be no different if every American voter knew the details of such issues. But such a level of expertise and information is so unrealistic that it is nearly impossible to conjecture what politics would look like with such people or how such a world could come about. This paper does answer the tractable question of whether changes in the level of information of the magnitude commonly observed have any effect on vote intentions on election day in the U.S. The answer to this question, contrary to the existing literature, is a resounding no.

A Replication of Bartels's "Uninformed Votes"

Table 10 is a replication of Bartels's (1996) Table 2 (p. 209). I extend Bartels's analysis to include results for the 1968 and 1996 NES surveys.[33] When I use maximum likelihood (ML) estimation, the substantive inferences from the replication are exactly the same as those made by Bartels although the exact numbers do differ. Using ML, I find significant information effects in the same years that Bartels does. The table also displays estimates from robust estimation. The robust estimator used is a new robust binary logistic model developed by Sekhon (2004b) which combines the down weighting of y-misclassification outliers of the conditionally unbiased bounded-influence approach of Kunsch et al. (1989) with a high breakdown point Mallows class estimator for down weighting x-outliers (Carroll and Pederson 1993).[34] Model selection is done using the theory developed by Cantoni and Ronchetti (2001). My robust estimation computer code is written in **R** (`http://www.r-project.org/`) and is available upon request.

Using ML, one finds significant information effects in four of the eight elections examined. Using robust estimation, however, significant information effects are never found.

As stated in the main text, Bartels does not include variables such as partisanship, ideology and all other non-demographic variables which are usually included in vote models because these variables may be the result of information. The estimated vote models leave out variables which are known to be important for the vote. It is not surprising then that the models do not fit the data well and that many extreme y-misclassification errors are present. In the presence of such outliers, inferences based on ML are generally inconsistent (Kunsch et al. 1989; Hampel et al. 1986; Mebane and Sekhon 2004) while the robust estimator supports reliable inferences (Cantoni and Ronchetti 2001; Hampel et al. 1986; Huber 1981). Hence, it is not surprising that the ML and robust estimation results differ.

Note that the extent of the y-misclassification issue cannot be determined by using max-

[33] The interviewers' ratings of respondents' levels of political information are not available before 1968.

[34] ML binary logistic models yield the same substantive inferences as obtained by Bartels's binary ML Probit models.

imum likelihood and simply examining the residuals because of the problem of masking (Atkinson 1986). One must use a robust procedure in order to determine the extent of the misclassification or outlier problem (Mebane and Sekhon 2004). The third column in Table 10 lists the proportion of observations in each dataset which received a weight less than 0.5. On average, about 10% of the observations were given weights of less than 0.5.

References

Abadie, Alberto. 2002. "Bootstrap Tests for Distributional Treatment Effect in Instrumental Variable Models." *Journal of the American Statistical Association* 97 (457): 284–292.

Abadie, Alberto and Guido Imbens. 2004. "Large Sample Properties of Matching Estimators for Average Treatment Effects." Working Paper.

Adams, John. 1973 [1805]. *Discourses on Davila: A Series of Papers on Political History*. New York: Da Capo Press.

Aldrich, John H. and Richard D. McKelvey. 1977. "A Method of Scaling with Applications to the 1968 and 1972 Presidential Elections." *American Political Science Review* 71 (1): 111–130.

Althaus, Scott L. 1998. "Information Effects in Collective Preferences." *American Political Science Review* 92 (3): 545–558.

Althaus, Scott L. 2003. *Collective Preferences in Democratic Politics: Opinion Surveys and the Will of the People*. New York: Cambridge University Press.

Alvarez, R. Michael. 1997. *Information and Elections*. Ann-Arbor: University of Michigan Press.

Atkinson, A. C. 1986. "Masking Unmasked." *Biometrika* 73 (Dec.): 533–541.

Austen-Smith, David and Jeffrey S. Banks. 1996. "Information Aggregation, Rationality, and the Condorcet Jury Theorem." *American Political Science Review* 90: 34–45.

Barnard, John, Constantine E. Frangakis, Jennifer L. Hill, and Donald B. Rubin. 2003. "Principal Stratification Approach to Broken Randomized Experiments: A Case Study of School Choice Vouchers in New York City." *Journal of the American Statistical Association* 98 (462): 299–323.

Bartels, Larry M. 1986. "Issue Voting Under Uncertainty: An Empirical Test." *American Journal of Political Science* 30 (4): 709–728.

Bartels, Larry M. 1996. "Uninformed Votes: Information Effects in Presidential Elections." *American Journal of Political Science* 40 (1): 194–230.

Bartels, Larry M. 1999. "Panel Effects in the National Election Studies." *Political Analysis* 8 (1).

Bartels, Larry M. 2003. "Homer Gets a Tax Cut: Inequality and Public Policy in the American Mind." Annual Meeting of the American Political Science Association, Philadelphia, August.

Bennett, Stephen Earl. 1988. "Know-nothings' Revisited: The Meaning of Political Ignorance Today." *Social Science Quarterly* 69: 476–490.

Berelson, Bernard R., Paul F. Lazarsfeld, and William N. McPhee. 1954. *Voting: A Study of Opinion Formation in a Presidential Campaign*. Chicago: University of Chicago Press.

Berg, Sven. 1993. "Condorcet's Jury Theorem, Dependency Among Jurors." *Social Choice and Welfare* 10: 87–95.

Brady, Henry. 2002. "Models of Causal Inference: Going Beyond the Neyman-Rubin-Holland Theory." Paper presented at the 19th Annual Summer Political Methodology Meetings. Seattle, WA.

Brunell, Thomas L. and John DiNardo. 2004. "A Propensity Score Reweighting Approach to Estimating the Partisan Effects of Full Turnout in American Presidential Elections." *Political Analysis* 12 (1): 28–45.

Bryce, James. 1995 [1888]. *The American Commonwealth*. Indianapolis: Liberty Fund.

Bunce, Valerie. 2000. "Comparative Democratization: Big and Bounded Generalizations." *Comparative Political Studies* 33 (6/7): 703–724.

Camp, Ai Roderic. 2004. "Citizen Attitudes toward Democracy and Vicente Fox's Victory in 2000." In Jorge I. Domínguez and Chappell Lawson, editors, *Mexico's Pivotal Democratic Election* Stanford: Stanford University Press. pages 25–46.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. New York: John Wiley & Sons.

Cantoni, Eva and Elvezio Ronchetti. 2001. "Robust Inference for Generalized Linear Models." *Journal of the American Statistical Association* 96: 1022–1030.

Carroll, R. J. and Shane Pederson. 1993. "On Robustness in the Logistic Regression Model." *Journal of the Royal Statistical Society, Series B* 84 (3): 693–706.

Collier, Christopher. 1971. *Roger Sherman's Connecticut: Yankee Politics and the American Revolution*. Middletown, Conn.: Wesleyan University Press.

Converse, Phillip. 1964. "The Nature of Belief Systems in Mass Publics." In David Apter, editor, *Ideology and Discontent* New York: Free Press. pages 240–268.

Cox, David R. 1958. *Planning of Experiments*. New York: Wiley.

Dawid, A. Phillip. 1979. "Conditional Independence in Statistical Theory." *Journal of the Royal Statistical Society, Series B* 41 (1): 1–31.

Dawid, A. Phillip. 2000. "Causal Inference without Counterfactuals (with Discussion)." *Journal of the American Statistical Association* 95 (450): 407–424.

Dehejia, Rajeev and Sadek Wahba. 1999. "Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs." *Journal of the American Statistical Association* 94 (448): 1053–1062.

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

Domínguez, I. Jorge. 2004. "Why and How Did Mexico's 2000 Presidential Election Campaign Matter?" In Jorge I. Domínguez and Chappell Lawson, editors, *Mexico's Pivotal Democratic Election* Stanford: Stanford University Press. pages 321–344.

Erikson, Robert S., Michael B. MacKuen, and James A. Stimson. 2002. *The Macro Polity*. New York: Cambridge University Press.

Fechner, Gustav Theodor. 1966 [1860]. *Elements of psychophysics, Vol 1.*. New York: Rinehart & Winston. Translated by Helmut E. Adler and edited by D.H. Howes and E.G. Boring.

Geddes, Barbara. 1999. "What Do We Know about Democratization." *Annual Review of Political Science* 2: 129–148.

Habermas, Jürgen. 1996 [1964]. *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society*. Cambridge, MA: MIT Press. Translated by Thomas Burger.

Hampel, Frank R., Elvezio M. Ronchetti, Peter J. Rousseeuw, and Werner A. Stahel. 1986. *Robust Statistics: The Approach Based on Influence Functions*. New York: Wiley, John and Sons.

Hausman, Jerry A. and David A. Wise. 1979. "Attrition Bias in Experimental and Panel Data: The Gary Income Maintenance Experiment." *Econometrica* 47: 455–473.

Heckman, James J., Hidehiko Ichimura, Jeffrey Smith, and Petra Todd. 1998. "Characterizing Selection Bias Using Experimental Data." *Econometrica* 66 (5): 1017–1098.

Hersh, Seymour M. 1982. "The Price of Power: Kissinger, Nixon, and Chile." *The Atlantic Monthly* (12).
URL `http://www.theatlantic.com/issues/82dec/hersh.htm`

Hettmansperger, T. 1984. *Statistical Inference Based on Ranks*. New York: Wiley.

Hirano, Keisuke, Guido W. Imbens, Geert Ridder, and Donald B. Rubin. 1998. "Combining Panel Data Sets with Attrition and Refreshment Samples." *Econometrica* 69: 1645–1659.

Holbrook, Allyson L., Matthew K. Berent, Jon A. Krosnick, Penny S. Visser, and David S. Boninger. 2004. "Attitude Importance and the Accumulation of Attitude-Relevant Knowledge in Memory." Working Paper.

Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81 (396): 945–960.

Horvitz, D. G. and D. J. Thompson. 1952. "A Generalization of Sampling without Replacement from a Finite Universe." *Journal of the American Statistical Association* 47: 663–685.

Huber, Peter J. 1981. *Robust Statistics*. New York: Wiley, John and Sons.

Huntington, Samuel P. 1991. *The Third Wave: Democratization in the Late Twentieth Century*. Norman and London: University of Oklahoma Press.

Imai, Kosuke. forthcoming. "Do Get-Out-The-Vote Calls Reduce Turnout? The Importance of Statistical Methods for Field Experiments." *American Political Science Review*.

Imbens, Guido W. 2000. "The Role of the Propensity Score in Estimating Dose-Response Functions." *Biometrika* 87 (3): 706–710.

Iyengar, Shanto. 1986. "Whither Political Information." Report to the NES Board of Overseers. Center for Political Studies, University of Michigan.

Iyengar, Shanto. 1990. "Shorts to Political Knowledge: Selective attention and the accessibility bias." In John Ferejohn and James Kuklinski, editors, *Information and the Democratic Process* Urbana: University of Illinois Press.

Johnston, Richard, André Blais, Henry E. Brady, and Jean Crête. 1992. *Letting the People Decide: Dynamics of a Canadian Election*. Montreal: McGill-Queen's University Press.

Johnston, Richard, Michael G. Hagen, and Kathleen Hall Jamieson. 2004. *The 2000 Presidential Election and the Foundations of Party Politics*. New York: Cambridge University Press.

Knuth, Donald E. 1998. *The Art of Computer Programming, Vol. 2: Seminumerical Algorithms*. Reading: MA: Addison-Wesley 3rd edition.

Kornhauser, William. 1959. *The Politics of Mass Society*. New York: The Free Press.

Kunsch, H. R., L. A. Stefanski, and R. J. Carroll. 1989. "Conditionally Unbiased Bounded-Influence Estimation in General Regression Models." *Journal of the American Statistical Association* 84: 460–466.

Lawson, Chappell. 2004. "Introduction." In Jorge I. Domínguez and Chappell Lawson, editors, *Mexico's Pivotal Democratic Election* Stanford: Stanford University Press. pages 1–21.

Le Bon, Gustave. 1896. *The Crowd: A Study of the Popular Mind*. New York: The Macmillan Co.

Lipset, Seymour Martin. 1959. "Some Social Requisites of Democracy: Economic Development and Political Legitimacy." *American Political Science Review* 53 (1): 69–105.

Little, Roderick J. A. and Donald B. Rubin. 1987. *Statistical Analysis with Missing Data*. New York: J. Wiley & Sons.

Mann, H. and D Whitney. 1947. "On a Test of Whether One of Two Random Variables is Stochastically Larger than the Other." *Annals of Mathematical Statistics* 18: 50–60.

McGraw, Kathleen and Neil Pinney. 1990. "The Effects of General and Domain-Specific Expertise on Political Memory and Judgement." *Social Cognition* 8: 9–30.

McKelvey, Richard D. and Peter C. Ordeshook. 1985a. “Elections with Limited Information: A Fulfilled Expectations Model Using Contemporaneous Poll and Endorsement Data as Information Sources.” *Journal of Economic Theory* 36: 55–85.

McKelvey, Richard D. and Peter C. Ordeshook. 1985b. “Sequential Elections with Limited Information.” *American Journal of Political Science* 29 (3): 480–512.

McKelvey, Richard D. and Peter C. Ordeshook. 1986. “Information, Electoral Equilibria, and the Democratic Ideal.” *Journal of Politics* 48 (4): 909–937.

McNemar, Q. 1947. “Note on the Sampling Error of the Differences Between Correlated Proportions or Percentage.” *Psychometrika* 12: 153–157.

Mebane, Walter R. Jr. and Jasjeet S. Sekhon. 2004. “Robust Estimation and Outlier Detection for Overdispersed Multinomial Models of Count Data.” *American Journal of Political Science* 48 (2): 391–410.

Miller, Nicholas R. 1986. “Information, Electorates, and Democracy: Some Extensions and Interpretations of the Condorcet Jury Theorem.” In Bernard Grofman and Guillermo Owen, editors, *Information Pooling and Group Decision Making* Greenwich, CT: JAI.

Neuman, W. Russell. 1986. *The Paradox of Mass Politics: Knowledge and Opinion in the American Electorate*. Cambridge, MA: Harvard University Press.

Page, Benjamin I. and Robert Y. Shapiro. 1992. *The Rational Public: Fifty Years of Trends in Americans’ Policy Preferences*. Chicago: University of Chicago Press.

Palfrey, Thomas R. and Keith T. Poole. 1987. “The Relationship between Information, Ideology, and Voting Behavior.” *American Journal of Political Science* 31 (3): 511–530.

Patterson, Thomas E. 2002. *The Vanishing Voter: Public Involvement in an Age of Uncertainty*. New York: Alfred A. Knopf.

Popkin, Samuel L. 1991. *The Reasoning Voter: Communication and Persuasion in Presidential Campaigns*. Chicago: University of Chicago Press.

Price, Vincent and John Zaller. 1993. "Who Gets the News? Alternative Measures of News Reception and Their Implications for Research." *Public Opinion Quarterly* 57 (2): 133–164.

Pryor, Kane. 2003. "A national state of confusion." *Salon.com* (6). Accessed 07/05/2004. URL `http://www.salon.com/opinion/feature/2003/02/06/iraq_poll`

Putnam, Robert D. 2000. *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon & Schuster.

Rosenbaum, Paul R. 1999. "Using Quantile Averages in Matched Observational Studies." *Applied Statistics* 48 (1): 63–78.

Rosenbaum, Paul R. 2002. *Observational Studies*. New York: Springer-Verlag 2nd edition.

Rosenbaum, Paul R. and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70 (1): 41–55.

Rubin, Donald B. 1973. "The Use of Matching and Regression Adjustment to Remove Bias in Observational Studies." *Biometrics* 29: 185–203.

Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66 (5): 688–701.

Rubin, Donald B. 1976. "Inference and Missing Data." *Biometrika* 63: 581–592.

Rubin, Donald B. 1977. "Assignment to a Treatment Group on the Basis of a Covariate." *Journal of Educational Statistics* 2: 1–26.

Rubin, Donald B. 1978. "Bayesian Inference for Causal Effects: The Role of Randomization." *Annals of Statistics* 6 (1): 34–58.

Rubin, Donald B. 1979. "Using Multivariate Sampling and Regression Adjustment to Control Bias in Observational Studies." *Journal of the American Statistical Association* 74: 318–328.

Rubin, Donald B. 1980. "Bias Reduction Using Mahalanobis-Metric Matching." *Biometrics* 36 (2): 293–298.

Rubin, Donald B. 1990. "Comment: Neyman (1923) and Causal Inference in Experiments and Observational Studies." *Statistical Science* 5 (4): 472–480.

Rustow, Dankwart A. 1970. "Transitions to Democracy: Toward a Dynamic Model." *Comparative Politics* 2 (3): 337–363.

Schafer, Joseph L. 1997a. *Analysis of Incomplete Multivariate Data*. London: Chapman & Hall.

Schafer, Joseph L. 1997b. "Imputation of missing covariates under a general linear mixed model." Technical report, Dept. of Statistics, Penn State University.

Sekhon, Jasjeet S. 2004a. "Quality Meets Quantity: Case Studies, Conditional Probability and Counterfactuals." *Perspectives on Politics* 2 (2): 281–293.

Sekhon, Jasjeet S. 2004b. "Robust Alternatives to Binary Logit and Probit: With reanalysis of Fearon and Laitin's (2003) "Ethnicity, Insurgency and Civil War" and Bartels's (1996) "Uninformed Votes: Information Effects in Presidential Elections." Working Paper.

Smith, Charles Page. 1956. *James Wilson, Founding Father, 1742-1798*. Chapel Hill: University of North Carolina Press.

Smith, Jeffrey and Petra Todd. forthcoming. "Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?" *Journal of Econometrics*.

Sniderman, Paul M. 1993. "The New Look in Public Opinion Research." In Ada W. Finifter, editor, *Political Science: The State of the Discipline II* Washington, DC: American Political Science Association.

Sniderman, Paul M., Richard Brody, and Philip E. Tetlock. 1991. *Reasoning and Choice: Explorations in Political Psychology*. New York: Cambridge University Press.

Snyder, Jack. 2000. *From Voting to Violence: Democratization and Nationalist Conflict*. New York: W. W. Norton.

Splawa-Neyman, Jerzy. 1990 [1923]. "On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9." *Statistical Science* 5 (4): 465–472. Trans. Dorota M. Dabrowska and Terence P. Speed.

Wilcox, Rand R. 1997. *Introduction to Robust Estimation*. San Diego, CA: Academic Press.

Wilcoxon, F. 1945. "Individual Comparisons by Ranking Methods." *Biometrics* 1: 8083.

Wittman, Donald A. 1989. "Why Democracies Produce Efficient Results." *Journal of Political Economy* 97: 1395–1424.

Yule, Undy G. 1899. "An Investigation into the Causes of Changes in Pauperism in England, Chiefly During the Last Two Intercensal Decades (Part I.)." *Journal of the Royal Statistical Society* 62 (2): 249–295.

Zakaria, Fareed. 1997. "The Rise of Illiberal Democracy." *Foreign Affairs* 76 (6): 249–295.

Zakaria, Fareed. 1998. "Doubts About Democracy." *Newsweek* (January).

Zaller, John R. 1985. "Proposal for the Measurement of Political Information." Report to the NES Board of Overseers. Center for Political Studies, University of Michigan.

Zaller, John R. 1986. "Analysis of Information Items in the 1985 NES Pilot Study." NES Pilot Study Report. Center for Political Studies, University of Michigan.

Zaller, John R. 1992. *The Nature and Origins of Mass Opinion*. New York: Cambridge University Press.

Zaller, John R. forthcoming. "Floating Voters in U.S. Presidential Elections, 1948–1996." In Willem Saris and Paul Sniderman, editors, *Studies in Public Opinion: Gauging Attitudes, Nonattitudes, Measurement Error and Change* New Jersey: Princeton University Press.

Table 1: People Who Changed Information State Between $t = 1$ and $t = 2$ (no imputation)

Year	September	October	November
1976			
Total observations	345	449	794
Change Information State	99	122	221
1980			
Total observations	563		563
Change Information State	126		126
1996			
Total observations	98	148	246
Changed Information State	21	46	67

The table does not include observations lost due to panel attrition but recovered by imputation. Each cell displays the number of respondents in that category. The counts are restricted to respondents from the panel surveys. November observations are from the post-election surveys. In 1980, no interviews were conducted in October. $t = 1$ is, respectively, the 1992 post-election, 1980 January wave and 1972 pre election surveys. $t = 2$ is the 1994 post-election, 1980 June wave and 1974 post-election surveys. And $t = 3$ is the 1996 pre-election, 1980 September wave and 1976 post-election surveys.

Table 2: Issue Placements During the 1980 Panel Study

Issue/panel	Reagan Placement Proportion	Carter Placement Proportion	Carter < Reagan All Respondents	Carter < Reagan Of Placers
Liberal-Conservative:				
panel 1 (Jan)	0.615	0.664	0.409	0.650
panel 2 (June)	**0.707**	**0.709**	**0.527**	**0.735**
panel 3 (Sept)	0.691	0.698	0.531	0.765
Defense Spending:				
panel 1	0.598	0.829	0.351	0.578
panel 2	**0.715**	0.831	**0.496**	**0.690**
panel 3	**0.802**	0.862	**0.638**	**0.794**
Government Services/Spending:				
panel 1	0.605	0.789	0.591	0.667
panel 2	**0.687**	0.804	0.567	0.652
panel 3	**0.747**	0.802	0.560	0.672
Unemployment vs. Inflation:				
panel 1	0.504	0.653	0.453	0.597
panel 2	**0.595**	0.695	**0.373**	**0.447**
panel 3	0.556	**0.591**	**0.320**	0.503
"Get Along" with Russia:				
panel 1	0.658	0.885	0.380	0.576
panel 2	**0.707**	0.851	**0.444**	0.626
panel 3	0.742	0.833	**0.502**	0.673

Numbers in bold are significantly different from the same measure in the previous panel at the 0.05 level. The McNemar test is used for all tests except for those in the fifth column where the t-test is used because of unequal observation numbers. The table does not include observations lost due to panel attrition but recovered by imputation. Note that the unemployment vs. inflation and government services/spending scales are reversed from the question wordings in order to make the Carter position the left position.

Table 3: Multivariate Tests for Balance for Interviewer Information Measure, Before and After Matching

Data	Before Matching		After Matching	
	K-S	LR	K-S	LR
Treatment: information changes from $t = 1$ to $t = 2$				
1992-94-96	0.00	0.00257	0.981	0.779
1980	0.00	0.0310	0.910	0.865
1972-74-76	0.00	0.0102	0.746	0.888
Treatment: information changes from $t = 1$ to $t = 3$				
1992-94-96	0.00	0.0072	0.912	0.718
1980	0.00	0.0393	0.704	0.891
1972-74-76	0.00	0.0126	0.449	0.550

The p-values in the K-S column are for the Kolmogorov-Smirnov test for equality of the propensity score densities of those who changed information state and of those who did not. The K-S test results are bootstrapped using 1000 resamples. The p-values in the LR columns are for the LR test of deviance (null deviance minus residual deviance) based on all baseline covariates. $t = 1$ is, respectively, the 1992 post-election, 1980 January wave and 1972 pre election surveys. $t = 2$ is the 1994 post-election, 1980 June wave and 1974 post-election surveys. And $t = 3$ is the 1996 pre-election, 1980 September wave and 1976 post-election surveys.

Table 4: Multivariate Tests for Balance for Issue Information Measure, Before and After Matching

Data	Before Matching		After Matching	
	K-S	LR	K-S	LR
Treatment: information changes from $t = 1$ to $t = 2$				
1980	0.00	4.02×10^{-3}	0.591	0.843
Treatment: information changes from $t = 1$ to $t = 3$				
1980	0.00	6.06×10^{-3}	0.762	0.698

See Table 3 and text for details.

Table 5: Estimated Effect of Changing Information State (Interviewer Measure)

Data	Information Changes from $t = 1$ to $t = 2$			Information Changes from $t = 1$ to $t = 3$
	September	October	November	November
1992-94-96	0.338	0.116	−0.0161	0.0146
	(0.107)	(0.0703)	(0.0828)	(0.0761)
1980	0.106		0.0253	−0.0451
	(0.0520)		(0.0676)	(0.0829)
1972-74-76	0.00618*		0.0190	0.0134
	(0.0505)		(0.0391)	(0.0401)

Table 5 presents ATT estimates of change in information level on changing one's vote from baseline. In 1980, no interviews were conducted in October. $t = 1$ is, respectively, the 1992 post-election, 1980 January wave and 1972 pre election surveys. $t = 2$ is the 1994 post-election, 1980 June wave and 1974 post-election surveys. And $t = 3$ is the 1996 pre-election, 1980 September wave and 1976 post-election surveys. *The 1976 pre-election survey did not randomize the month respondents were interviewed so September and October observations are combined.

Table 6: Estimated Effect of Changing Information State (Issue Scale Measure)

Data	Information Changes from Jan to June		Information Changes from Jan to Sept
	September	November	November
	—outcome: change in Republican vote—		
1980	0.179	0.0577	−0.0347
	(0.0801)	(0.0849)	(0.0898)
	—outcome: any change in vote—		
1980	0.178	−0.0554	0.0255
	(0.0967)	(0.0851)	(0.0810)

Table 6 presents ATT estimates of change in information level on changing one's vote from baseline. For the 1980 panel study, January was the first wave, June and early July the second and September the third.

Table 7: Issue Placements During the 2000 Mexico Panel Study

Survey Wave			
1st (February)	2nd (April)	3rd (June)	4th (July)
Is Mexico a democracy? [proportion answering yes]:			
0.400	0.446	0.4810	0.6310
Proportion who could name the judicial branch of government:			
0.378	0.424	0.507	0.504
Proportion who could place self on liberal-conservative scale:			
0.717	0.794	0.775	0.812
Proportion who could place PAN on liberal-conservative scale:			
0.673	0.758	0.739	0.804
Proportion who could place PRI on liberal-conservative scale:			
0.683	0.761	0.739	0.795
Proportion who could place PRD on liberal-conservative scale:			
0.663	0.760	0.739	0.795
Proportion who could correctly identify Fox slogan:			
0.500	0.605	0.725	0.774
Proportion who could correctly identify Labastida slogan:			
0.404	0.604	0.698	0.722

The table does not include observations lost due to panel attrition but recovered by imputation.

Table 8: Multivariate Tests for Balance for Mexico 2000 Panel Study, Before and After Matching

Before Matching		After Matching	
K-S	LR	K-S	LR
Treatment: Learning to place PAN on lib-con:			
0.00	0.019	0.806	0.901
Treatment: Learning to place PRI on lib-con:			
0.00	6.744×10^{-7}	0.877	0.817
Treatment: Learning to mention the judicial branch:			
0.00	9.316×10^{-9}	0.685	0.727

The p-values in the K-S column are for the Kolmogorov-Smirnov test for equality of the propensity score densities of those who changed information state and of those who did not. The K-S test results are bootstrapped using 1000 resamples. The p-values in the LR columns are for the LR test of deviance (null deviance minus residual deviance) based on all baseline covariates.

Table 9: Estimated Effect of Increasing Information Level, Mexico 2000 Panel Study

Estimate	SE
Treatment: Learning to place PAN on lib-con:	
0.214	0.0718
Treatment: Learning to place PRI on lib-con:	
0.1691	0.0798
Treatment: Learning to mention the judicial branch:	
0.208	0.0859

Table 9 presents ATT estimates of increasing information level (between waves 1 and 3) on changing one's vote (wave 4) from baseline (wave 1). The first wave was conducted in February, the 2nd in April, the 3rd in June and the 4th in July.

Table 10: Maximum and Robust Likelihood Ratio Tests for Deviations from Fully Informed Voting, 1968–1996

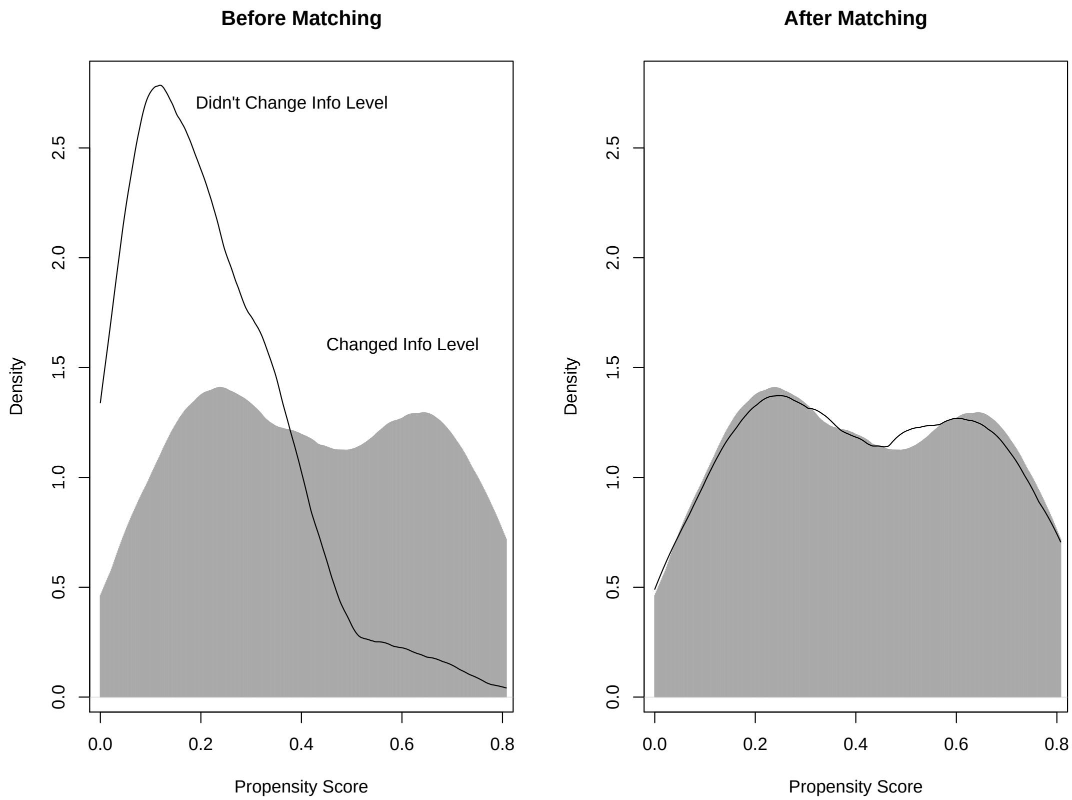
Year	p-value for LR test for likelihood with and without information effects	p-value for robust LR test for robust likelihood with and without information effects	Percentage of observations w/weights $< .5^*$
1968	0.394	0.387	0.144
1972	0.0118	0.117	0.126
1976	0.218	0.682	0.0865
1980	0.537	0.656	0.119
1984	0.0224	0.381	0.114
1988	0.287	0.346	0.0959
1992	0.0216	0.485	0.0847
1996	0.00549	0.209	0.133

Replication of Bartels's (1996) Table 2 plus replication with robust estimation. All p-values are obtained from χ^2 statistics with 21 degrees of freedom.

* The percentage of observations which the robust estimator gives a weight of less than 0.5 The ML estimator gives a weight of 1 to all observations.

Figure 1: Densities of Propensity Scores for Changes in the Interviewer Information Measure between $t = 1$ and $t = 2$, Before and After Matching

1992-1994-1996 Panel



1980 Panel

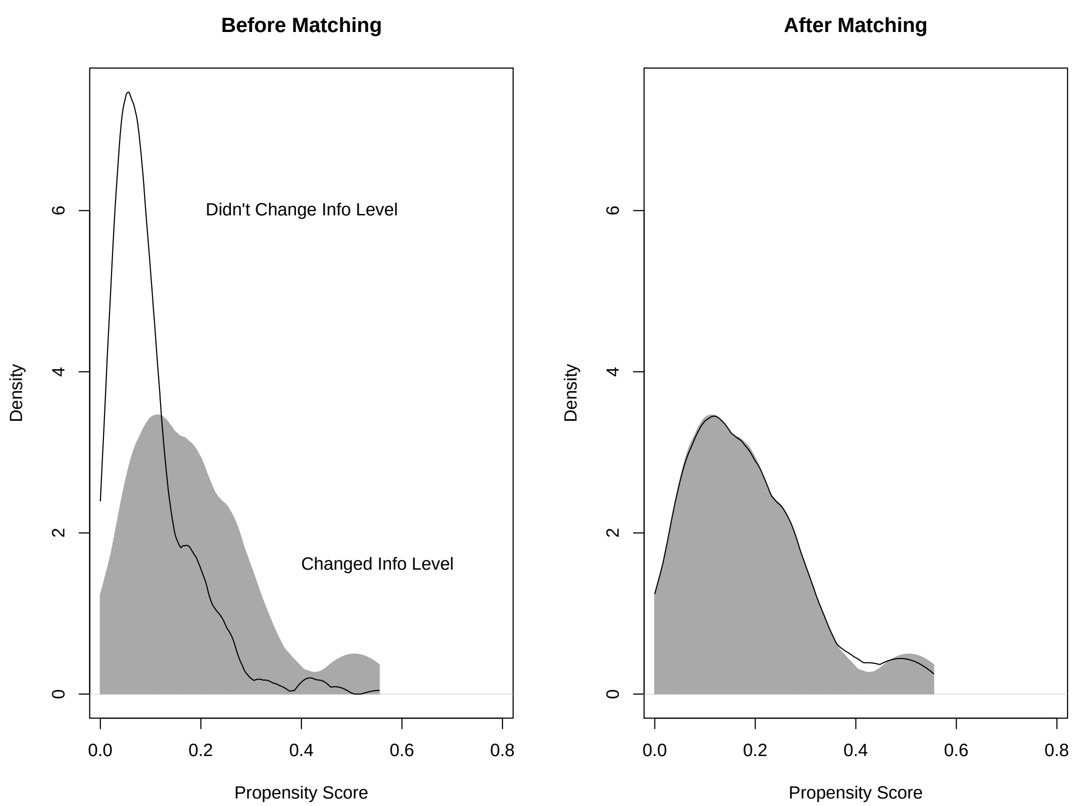


Figure 1 (continued)

1972-1974-1976 Panel