Three Virtues of Panel Data for the Analysis of Campaign Effects

Larry M. Bartels

THE PRIMARY AIM OF PARTICIPANTS in election campaigns is to produce politically significant changes in the attitudes and perceptions of prospective voters. The primary aim of scholarly observers of election campaigns is to measure and explain those politically significant changes. Because campaigns are dynamic phenomena, good campaign studies must be dynamic too. Time must enter the analysis either directly or indirectly (as a proxy for campaign activities and events). Survey researchers can incorporate time most simply by interviewing different respondents at different times and incorporating the date of the interview in their analyses as a potential explanatory variable (or incorporating variables describing campaign events keyed to dates of interview). The result is a species of *longitudinal* study or (in the quasi-experimental design literature) *interrupted time-series design* (Cook and Campbell 1979, chap. 5). In the specific setting of survey research—especially within the compass of a single survey spanning days or weeks rather than years—it has come to be known as a *rolling cross-section design* (Johnston and Brady 2002).

The rolling cross-section design's primary virtue is its simplicity. It exploits the leverage for scientific inference of a brute fact about the logistics of (at least most academic) survey research: interviews take time to conduct. By treating variation in the date of interview as a potential explanatory variable, this administrative nuisance becomes an opportunity to learn something about the substance of the political or social processes underlying survey responses.

Of course, a rolling cross-section survey may be timed to coincide with especially interesting political or social processes, as with the American National Election Studies (NES) 1984 Continuous Monitoring Survey, which was designed to capture the effects of primary campaigns, conventions, and other political events outside the time frame of the traditional NES preelection surveys (Bartels 1988). Moreover, artifactual differences in the characteristics of respondents interviewed at different times may be minimized by conscious attention to sampling procedures in the course of data collection, as with the 1988 Canadian Election Study's daily release of fresh sample replicates (Johnston et al. 1992, appendix A). But even in the absence of conscious attention to the implications of temporal variation in the course of data collection, analysts may come to realize that they are analyzing rolling cross-sections in more or less the same way that speakers come to realize that they are speaking prose. Thus, for example, Sanders (1996), Vavreck (1997), and Bartels (2000c; essay in this volume) have treated the traditional NES preelection surveys as two-month rolling cross-sections, and Wright (1990, 1993) has used the date of NES postelection interviews to examine the deterioration of accuracy in vote reports, despite the fact that these surveys were designed and conducted with neither purpose in mind.

The other most common way to incorporate a dynamic element in survey design is to interview the same individuals at two or more points in time and attribute *changes* in their attitudes or perceptions to the effects of intervening events. This so-called *panel design* can provide more direct evidence of campaign effects than the rolling cross-section design, in the sense that change is observed in the responses of the same individuals rather than being inferred from comparisons of different survey respondents. On the other hand, the problem of attributing observed changes to *specific* intervening events is often more difficult with panel data, since many different events may intrude in the period between successive panel waves—from a few weeks to several months or much longer, as in the four-year panel studies conducted by NES in 1956-58-60, 1972-74-76, and 1992-94-96 or the even longer political socialization panels analyzed by Jennings and Niemi (1981).

Panel designs have been especially prominent in the field of electoral studies. Significant campaign studies based primarily upon election-year panel surveys include the classic Columbia studies of the 1940 and 1948 presidential campaigns (Lazarsfeld, Berelson, and Gaudet 1948; Berelson, Lazarsfeld, and McPhee 1954), the parallel television-era panel studies of Patterson and McClure (1976) and Patterson (1980), and several studies based upon the 1980 NES election-year panel (Markus 1982; Bartels 1993; Finkel 1993). The recent efforts of Just et al. (1996) suggest that panel designs continue to appeal to scholars interested in capturing campaign effects.

The variety of significant campaign analyses based upon rolling cross-section surveys on the one hand and panel surveys on the other should be sufficient to demonstrate that neither approach represents an exclusive path to scientific progress in the field of electoral studies. Both will, no doubt, continue to provide important insights regarding the reactions of prospective voters to campaign events. However, scholarly assessments of the advantages and disadvantages of the two approaches seem to me to have been skewed—and scholarly exploitation of existing panel data seems to me to have been hampered—by insufficient appreciation of some of the specific analytic virtues of panel designs. Thus, my aim here is to outline what I consider to be the three primary advantages of panel data for the analysis of campaign effects.

(1) Panel data facilitate adjustments for measurement error in survey responses using Wiley and Wiley (1970) or other measurement models. Given the embarrassingly low reliability of typical measures of political attitudes and perceptions, adjustments for measurement error are essential for achieving plausible estimates of the magnitudes of many campaign effects.

(2) Panel data permit analyses of opinion change in which prior opinions appear as explanatory variables. Given the stability of typical political opinions and their modest correlations with relevant explanatory variables, direct measurement of prior opinions substantially increases the efficiency of statistical estimation and provides crucial perspective on the relative political importance of precampaign and campaign events.

(3) Panel data facilitate analyses in which relevant explanatory variables are measured outside the immediate campaign setting. Given the susceptibility of political attitudes to rationalization in terms of causally irrelevant themes prominent in campaign discourse, prior measurement of potential explanatory variables provides important insights regarding the causal priority of specific political attitudes and perceptions in the evolution of candidate evaluations.

In addition to rehearsing these three advantages of panel data, I briefly address the two most distinctive inferential problems posed by panel designs: panel attrition and panel conditioning. I also consider possibilities for melding panel and rolling cross-section elements in a single survey design. Although such mixed designs raise significant unresolved problems of model specification and estimation, they offer the very attractive prospect of combining the best features of panel and rolling cross-section designs and thus of contributing further to our understanding of opinion change in campaign settings.

Allowances for Measurement Error

Analysts of opinion dynamics typically ask two kinds of questions: First, how stable are opinions over time? And second, to the extent that opinions change, what produces those changes?

Thanks primarily to the efforts of Achen (1975, 1983), the effects of measurement error on inferences about opinion stability are now widely recognized. In particular, it is clear that the low levels of apparent opinion stability in mass publics observed by Converse (1964) and others are attributable in large part to the effects of random fluctuation in survey responses. Although the nature and significance of this "measurement error" are a matter of vigorous theoretical debate (Achen 1975; Feldman 1990; Zaller and Feldman 1992; Zaller 1992; Brady 1993), analysts of political attitudes and belief systems now at least know better (or *should* know better) than to mistake opinion responses in surveys for the real political opinions underlying those responses.

Curiously, the same level of methodological sophistication that has become fairly commonplace in analyses of opinion stability is much less commonplace in analyses of opinion change. It is by no means unusual for analysts to regress opinion variables on prior opinions plus some other variables intended to capture potential causes of opinion change, making no allowance for error either in measured prior opinions or in the other measured variables associated with potential opinion change. Unfortunately, such an approach cannot, in general, lead to reliable inferences about the nature and causes of opinion change; nor does data analytic experience give us good reason to believe that the resulting biases are likely to be so minor that we can safely ignore them in practice.

The key to correcting the biases created by measurement errors in explanatory variables is to obtain estimates of the magnitudes of those measurement errors. Panel data facilitate corrections for measurement error because repeated measurement of the same opinion or behavior provides a check on the statistical *reliability* of observed responses. In two-wave panels, an obvious measure of reliability is the "test-retest" correlation of individual responses. However, it would be imprudent to interpret instability as measurement error in contexts where underlying "true" opinions may also be changing. The primary advantage of three-wave panels is that they provide leverage for distinguishing between random measurement errors and changes in underlying true opinions.

Although specific models of measurement error in panel data may be complex (Heise 1969; Wiley and Wiley 1970; Jöreskog 1979; Achen 1983), the intuition underlying this leverage is simple: to the extent that "we can predict t_3 issue positions of individuals fully as well from a knowledge of their t_1 positions alone as we can from a knowledge of their t_2 positions alone" (Converse 1964), the apparent changes between t_1 and t_2 are interpreted as random measurement error rather than changes in underlying true positions. Conversely, to the extent that observed response instability varies inversely with the temporal proximity of the responses, the instability is attributed to real opinion change rather than to random measurement error.

My own work on the political impact of media exposure (Bartels 1993) provides a useful illustration of the importance of allowing for measurement error in analyses of campaign effects. My aim was to estimate the effects of television news exposure, newspaper exposure, and partisan predispositions on a variety of politically relevant attitudes and perceptions during the 1980 presidential campaign, using data from the NES election-year panel survey. The data consisted of a variety of opinion readings for 758 panel respondents (the survivors from a first-wave sample of 1,008) at three points in the 1980 campaign. The first wave of interviews was conducted in late January and February (before the first primary voting in New Hampshire), the second wave in June (between the end of the primary season and the national nominating conventions), and the third wave in September (during the first month of the general election campaign).

The dependent variables in the analyses summarized here include "thermometer" ratings of the competing candidates, assessments of Carter's performance as president, and ratings on a battery of trait items tapping the candidates' character, leadership, and competence.[1] To facilitate comparison, all of the original responses were recoded to range from 0 to 100, with 0 denoting the most negative possible opinion and 100 denoting the most positive possible opinion. For each perception, I estimated the effects of prior opinions, partisan predispositions, television news exposure, and newspaper exposure separately in each half of the election-year panel (February to June and June to September).[2]

I estimated the magnitudes of measurement errors using a variant of the Wiley and Wiley (1970) model.[3] The measurement error estimates calculated from the modified Wiley and Wiley model indicated that about 25 percent of the observed variance in thermometer ratings and about 40 percent of the observed variance in the job approval and trait ratings represented random noise. The estimated measurement reliabilities for the media exposure variables were somewhat higher—.75 for television news exposure and .78 for newspaper exposure—and the apparent measurement reliability of party identification was higher still, at .88.

The inferential consequences of these measurement errors are highlighted in table 1, which summarizes the estimated effects of television news exposure reported by Bartels (1993). The first and third columns of

TABLE 1. Impact of Measurement Error on Estimates of Television News Effects in the 1980 Presidential Campaign

	OLS June	EV June	OLS September	EV September
Carter thermometer rating	1.0	−2.6	2.3	2.4
	(2.4)	(3.1)	(2.2)	(2.9)
Reagan thermometer rating	3.4	5.2	1.4	1.2
	(2.2)	(3.0)	(2.4)	(3.3)
Carter job approval (5 items)	−0.7	−5.2	1.6	2.0
	(3.8)	(5.2)	(3.6)	(5.3)
Carter character (3 items)	0.5	−0.1	2.9	4.1
	(2.9)	(4.1)	(2.9)	(4.2)
Carter leadership (3 items)	0.5	−2.7	2.2	1.4
	(2.9)	(3.9)	(2.7)	(4.1)
Carter competence (3 items)	0.9	−2.9	3.8	4.5
	(2.7)	(3.9)	(2.7)	(4.1)
Reagan character (3 items)	2.4	0.2	3.2	1.4
	(2.6)	(3.7)	(2.8)	(4.1)
Reagan leadership (3 items)	0.3	0.2	−3.6	−8.5
	(2.6)	(3.6)	(2.9)	(4.3)
Reagan competence (3 items)	2.8	3.5	−4.5	−9.6
	(2.5)	(3.4)	(2.7)	(4.0)
Average (absolute values)	1.4	2.5	2.8	3.9
	(2.7)	(3.8)	(2.8)	(4.0)

Note: Average OLS and EV parameter estimates with average standard errors in parentheses, calculated from Bartels 1993. $N = 753$.

the table show ordinary least squares (OLS) parameter estimates for the effects of television news exposure in the period from February through June and from June through September, respectively. The second and fourth columns of the table show the corresponding errors-in-variables (EV) parameter estimates of the same effects.

As the average standard errors presented in table 1 make clear, most of the individual parameter estimates summarized in the table are too imprecise to be "statistically significant." However, the number and magnitude of "significant" effects—most notably, the negative impact of television news exposure between June and September on Reagan's "leadership" and "competence" ratings—are too large to be due to chance and suggest that television news coverage of political campaigns does sometimes produce large, politically consequential shifts in prospective voters' perceptions and evaluations of the candidates.[4]

For purposes of the present argument, what is most worth noting in table 1 is that these media effects are seriously underestimated in the OLS analysis, which takes no account of measurement error in the explanatory variables. The average magnitude of the OLS estimates is 2.1 points on the 100-point attitude scales; the corresponding average of the EV estimates, which allow for measurement error in the explanatory variables, is 50 percent larger.[5]

Table 2 shows a similar comparison of estimated effects of preexisting partisan loyalties on campaign season opinion changes. Here, in contrast to the case of media exposure, the OLS estimates significantly *overstate* the impact of partisan predispositions on opinion change during the campaign period, by an average of more than 50 percent. Although the EV estimates of the impact of party identification are still quite substantial, they suggest that partisan bias in campaign-season opinion changes is markedly less pervasive than naive analysis would suggest. This example illustrates Achen's (1983) warning that measurement error does *not* necessarily bias ordinary regression parameter estimates toward zero, making them "conservative" estimates of the corresponding true effects, as analysts sometimes seem to assume. In most cases of practical interest, the direction of biases produced by measurement error is simply not predictable a priori.

If the comparisons presented in tables 1 and 2 are typical—and there is no obvious reason to suppose they are not—then many estimates of campaign effects derived from statistical analyses without adjustments for measurement error are likely to be seriously misleading. If the prospect of positive or negative biases of 50 percent or more in our parameter estimates is not very worrisome, we should probably not be estimating the parameters in the first place. If that prospect *is* very worrisome, we should do whatever we can to incorporate adjustments for measurement error in our analyses.

Reasonable adjustments for measurement error do not necessarily require panel data. Analysts could, and should, make much more routine use of instrumental variables estimation strategies when data are sufficiently plentiful to allow for the detection of politically significant effects using the sorts of instruments typically available in opinion surveys (Bartels 1991). Moreover, estimates of measurement error variance from any source can be employed to produce consistent parameter estimates using either maximum likelihood or adjusted least squares approaches (Fuller 1987); measurement models employing redundant measures of key variables at a single point in time can serve as well as those employing panel data (Bollen 1989).

TABLE 2. Impact of Measurement Error on Estimates of Partisan Activation in the 1980 Presidential Campaign

	OLS June	EV June	OLS September	EV September
Carter thermometer rating	8.0	5.9	7.3	4.8
	(1.1)	(1.4)	(1.0)	(1.3)
Reagan thermometer rating	8.3	7.8	5.8	2.6
	(1.1)	(1.3)	(1.1)	(1.6)
Carter job approval (5 items)	6.9	4.5	11.3	6.5
	(1.8)	(2.2)	(1.6)	(2.1)
Carter character (3 items)	3.5	1.9	3.6	1.2
	(1.4)	(1.7)	(1.3)	(1.7)
Carter leadership (3 items)	7.6	5.6	6.6	1.1
	(1.3)	(1.7)	(1.3)	(1.8)
Carter competence (3 items)	5.4	2.4	7.5	2.9
	(1.3)	(1.8)	(1.2)	(1.8)
Reagan character (3 items)	3.9	2.3	7.6	5.4
	(1.2)	(1.5)	(1.2)	(1.6)
Reagan leadership (3 items)	6.0	4.9	7.0	3.3
	(1.2)	(1.5)	(1.3)	(1.7)
Reagan competence (3 items)	6.5	5.1	8.2	5.0
	(1.2)	(1.5)	(1.3)	(1.8)
Average	6.2	4.5	7.2	3.6
	(1.3)	(1.6)	(1.3)	(1.7)

Note: Average OLS and EV parameter estimates with average standard errors in parentheses, calculated from Bartels 1993. $N = 753$.

While panel data are by no means necessary to make allowances for measurement error in studies of opinion change, they remain the most common and best-tested source of measurement error estimates. These estimates need not come from the data set actually being analyzed; benchmark estimates of measurement error in a variety of frequently used explanatory variables could and should be calculated from available panel data and then routinely employed in settings where panel data are unavailable.[6] Nevertheless, for analyses employing explanatory variables whose measurement properties are not already familiar, panel data are likely to provide invaluable leverage for taking serious account of the substantial threats to statistical inference posed by measurement error in survey data.

Prior Measurement of Dependent Variables

The second primary virtue of panel data is that they allow for direct observation of individual-level change in attitudes and perceptions in response to campaign events. "Change scores" may be calculated from observed responses at two points in time and subjected directly to statistical analysis, or (more generally) new opinions may be analyzed as a function of old opinions and intervening characteristics or events.[7] This is in one sense a very familiar virtue of panel data, since an explicit focus on individual-level change is practically a defining feature of panel analysis. However, as with many familiar virtues, the very familiarity of this point may obscure our understanding of *why* direct measurement of prior attitudes and perceptions is so valuable.

My argument here is that direct measurement of prior opinion has two distinct but interrelated advantages, one essentially methodological and the other more substantive. On the one hand, given the relative stability of most political attitudes under most circumstances, including prior opinions among our explanatory variables will significantly increase the precision of our statistical analyses. Every student of regression analysis learns that the precision of regression parameter estimates is inversely proportional to the residual variance in the dependent variable, which reflects the impact of potential explanatory factors that are not included (either explicitly or by proxy) as explanatory variables. If there is any panel study of political campaigns in which prior opinions do not turn out to be strongly correlated with subsequent opinions, I have not met it. Conversely, if there is any cross-sectional study of political campaigns in which current opinions are so well accounted for by contemporaneous factors that adding lagged dependent variables would not significantly reduce the residual variance, I have not met it.[8] Thus, even if our real interest is in the causes of short-term opinion *change* rather than in opinion *stability,* panel data will greatly facilitate our ability to isolate the effects of specific campaign events or processes.

On the other hand, even if our real interest is in the causes of short-term opinion change, our substantive understanding of the specific events or processes that produce that change may be greatly enriched by juxtaposing change with stability. This point could also be expressed in methodological terms by reference to potential biases in parameter estimates stemming from the "omission" of prior opinions that "belong" in a "well-specified" regression model. However, it may be expressed more substantively by simply noting that, in situations where prior opinions *are* observed and their effects are sensibly estimated, the effects of even the most salient campaign events are likely to be modest by comparison.

It may seem odd to begin by arguing from "the relative stability of most political attitudes under most circumstances," when Converse (1964) and many others have emphasized the remarkable *in*stability of most survey responses in the political realm. The key to this apparent contradiction is the distinction between attitudes and survey responses, which brings us squarely back to the issue of measurement error. If we define a political attitude as the mean of a distribution of potential survey responses, and the variation around that mean as "measurement error," it may make perfect sense to say that underlying attitudes are quite stable while survey responses are quite unstable. Nothing in the formulation or estimation of the measurement error models described in the previous section ensures that underlying attitudes will, in fact, turn out to be quite stable once random measurement error in observed survey responses is taken into account. Nevertheless, that is the nearly invariable pattern in the empirical analyses carried out by Achen (1975), Feldman (1990), and many other scholars applying measurement models to panel data.

To illustrate this point without multiplying examples, table 3 provides a comparison of OLS and EV parameter estimates of opinion stability in the 1980 presidential campaign, again from Bartels 1993. This comparison parallels the comparisons of unadjusted and adjusted parameter estimates for media exposure and partisan effects presented in tables 1 and 2. As in tables 1 and 2, there is strong evidence in table 3 of biases in the ordinary regression parameter estimates due to measurement error in the various explanatory variables included in the analysis. As in the case of media exposure (but in contrast to the case of partisan predispositions), the consequence of measurement error is to greatly understate the effect of prior opinions on current opinions in each wave of the 1980 panel study. Significant biases are apparent in every one of the separate parameter estimates, and the average magnitude of these biases is even larger than for the media exposure effects in table 1.

What is more, since similar biases are apparent in both waves of the 1980 panel study, the cumulative stability of opinions over the course of the entire election year is even more distorted in the ordinary regression estimates. The stability coefficient for each political attitude from February through September is the product of its separate stability coefficients from February through June and from June through September. Since each component of this product is badly underestimated in the ordinary regression analysis, the product is even more seriously underestimated. This com-

TABLE 3. Impact of Measurement Error on Estimates of Opinion Stability in the 1980 Presidential Campaign

	OLS June	EV June	OLS September	EV September
Carter thermometer rating	.62	.81	.63	.78
	(.03)	(.06)	(.03)	(.04)
Reagan thermometer rating	.49	.67	.61	.89
	(.03)	(.05)	(.03)	(.07)
Carter job approval (5 items)	.39	.68	.52	.95
	(.03)	(.06)	(.03)	(.06)
Carter character (3 items)	.40	.77	.42	.79
	(.04)	(.12)	(.03)	(.10)
Carter leadership (3 items)	.42	.70	.46	.89
	(.03)	(.07)	(.03)	(.07)
Carter competence (3 items)	.42	.82	.40	.86
	(.03)	(.11)	(.03)	(.09)
Reagan character (3 items)	.38	.72	.43	.88
	(.03)	(.09)	(.03)	(.10)
Reagan leadership (3 items)	.37	.71	.43	.92
	(.03)	(.10)	(.04)	(.03)
Reagan competence (3 items)	.37	.70	.44	.84
	(.03)	(.10)	(.04)	(.10)
Average	.43	.73	.48	.87
	(.03)	(.08)	(.03)	(.07)

Note: Average OLS and EV parameter estimates with average standard errors in parentheses, calculated from Bartels 1993. $N = 753$.

pounded bias is evident in the first two columns of table 4, which compare estimates of the cumulative stability of campaign-related opinions between February and September derived from the OLS and EV estimates in table 3, respectively.

The OLS estimates in the first column of table 4 suggest that, on average, only a little more than 20 percent of respondents' opinions at the time of the first NES interview in February carried over to the third interview in September. Thus, the apparent stability of these opinions is quite low, even by comparison with the fairly arcane policy issues examined over a two-year interval by Converse (1964). The implication of these estimates is that impressions formed during the election year dominate electoral politics, at least at the presidential level. By contrast, the EV estimates in the second column of table 4 suggest that, on average, more than 60 percent of respondents' February attitudes and perceptions carried over to September—a threefold increase in apparent stability over the uncorrected estimates. By this somewhat less naive reckoning, most of what people believed about both Carter and Reagan in the midst of the general election campaign was already fixed months earlier, before the public phase of the campaign had even begun.

The difference between these two sets of estimates is of obvious significance for any general understanding of the electoral process. The apparent stability of candidate evaluations, once measurement error is taken into account, highlights the political significance of what Box-Steffensmeier and Franklin (1995) have referred to, in the context of Senate elections, as "the

TABLE 4. Impact of Measurement Error on Estimates of Cumulative Opinion Stability and "Distinctive Messages"

	OLS Stability	EV Stability	OLS Messages	EV Messages
Carter thermometer rating	.394	.628	4.8	1.0
Reagan thermometer rating	.296	.592	5.0	14.4
Carter job approval (5 items)	.205	.643	1.4	−8.0
Carter character (3 items)	.170	.617	3.4	7.5
Carter leadership (3 items)	.196	.626	3.0	−2.4
Carter competence (3 items)	.166	.695	4.7	−2.1
Reagan character (3 items)	.168	.634	4.9	1.8
Reagan leadership (3 items)	.172	.664	−4.3	−26.8
Reagan competence (3 items)	.163	.584	−3.9	−16.2
Average (absolute values)	.214	.631	3.9	8.9

Note: Averages based upon OLS and EV parameter estimates, calculated from Bartels 1993. $N = 753$.

long campaign." Their analysis suggests that job approval measured two years before the 1992 Senate elections had more impact on the 1992 election outcomes than changes in job approval between 1990 and 1992, and they concluded (1995, 314) that "the time of governing strongly affects public perceptions and evaluations, and so provides a key linkage between representatives and their constituents." By the same token, the results in table 4 suggest in the context of presidential elections that attitudes toward the candidates are more a product of long-term political assessments than of short-term reactions to campaign events. In an important sense, most of what is important in electoral politics happens before the election year even begins. If one of our aims is to gauge the broad political significance of campaigns, this is a crucial fact to keep in view.

In addition to being important in its own right, the stability of prior opinions in campaign settings has more subtle but equally important implications for our understanding of how prospective voters interpret and respond to campaign events. As I have argued elsewhere (Bartels 1993), reasoning from the logic of Bayesian opinion change, new information must compete with a relatively larger mass of prior beliefs than has generally been supposed and thus must itself be much more distinctive than has generally been supposed in order to produce the changes in opinion that we actually observe. Thus, somewhat counterintuitively, evidence that preexisting opinions are quite stable suggests, albeit indirectly, that the new information absorbed during campaigns must, at least occasionally, be quite distinctive.

This second-order implication of opinion stability is illustrated in the third and fourth columns of table 4, which presents estimates of the "distinctive messages" that would be necessary to account for observed changes in candidate evaluations and trait ratings among prospective voters exposed to television news reports during the 1980 presidential campaign. Again, the results are summarized from Bartels 1993 and are reported and interpreted in more detail there. For present purposes, it may be sufficient to explain that the "distinctive message" estimates summarized in table 4 represent differences (on the 100-point attitude scales) between the new information apparently absorbed over the eight months of the 1980 NES panel study by respondents maximally exposed to network television news and those completely unexposed to network television news. Thus, for example, the estimate of −16.2 for Reagan competence in the fourth column of the table suggests that the new information regarding

Reagan's competence absorbed by regular viewers of network television news between February and September was, on average, 16 points more negative (on the 100-point scales) than the corresponding new information absorbed by nonviewers.

The comparison of "distinctive message" estimates derived from ordinary regression analysis (in the third column of table 4) and EV analysis (in the fourth column of the table) indicates that here, too, adjustment for measurement error has substantial implications for the nature of our substantive conclusions about how campaigns matter. In the ordinary regression analysis, where the stability of prior opinions and the effects of television news exposure are both significantly underestimated, the apparent distinctiveness of the campaign messages received from television news never exceeds 5 points on the 100-point attitude scales. In the EV analysis, half of the estimates exceed 5 points and the *average* estimate is almost 9 points, suggesting that the campaign messages received by television news viewers were more than twice as distinctive as the uncorrected estimates suggest.

This analysis of "distinctive messages" provides an unusually clear example of the implications of opinion stability for our understanding and interpretation of campaign processes. The Bayesian model of opinion change on which the analysis is based highlights the interconnection of preexisting opinions and new information and also provides an explicit framework for using evidence of opinion stability to specify quantitatively the implications and significance of observed opinion change. However, even in the absence of a specific formal model of opinion change, it seems likely that the direct evidence of opinion stability provided by panel data will be of considerable value in informing and enriching our understanding of campaigns as dynamic phenomena.

Prior Measurement of Explanatory Variables

The third significant virtue of panel data is that they allow us to assess the effects of explanatory variables measured outside the immediate campaign period. Contemporaneous correlations among attitude variables are seldom subject to straightforward causal interpretation, since reciprocal causation and rationalization can seldom be ruled out. This general problem may be especially severe in campaign settings, since campaigners may actually strive to provide rationalizations and reinforcement for existing attitudes as much as they strive to create new, more favorable attitudes. If a prospective voter tells us that the candidate she intends to support is "moral," or has issue preferences similar to her own, or will keep the country out of war, are her perceptions causes or effects of her vote intention? In the absence of strong theoretical preconceptions about what causes what, we are likely to be left with a morass of intercorrelations, impossible to disentangle in any very convincing way, even using sophisticated simultaneous-equation techniques. What we need, and what panel data provide, are baseline measurements of preexisting opinions and attitudes, so that *temporal* priority can inform our conclusions about *causal* priority even in the absence of strong theory.

Thus, one familiar application of panel data is to estimate so-called *cross-lag* models, in which each of two interrelated variables is regressed on lagged values of both variables. Marked asymmetries in the apparent effects of the two variables are interpreted as evidence of causal priority: if changes in B can be well accounted for by previous values of A, *but not vice versa,* then A seems more likely to be a cause of B than an effect of B.[9]

Rahn, Krosnick, and Breuning (1994) used panel data in this way to assess the causal significance of responses to open-ended questions tapping prospective voters' "likes" and "dislikes" of the competing candidates in a gubernatorial campaign. They found that overall "thermometer" ratings of the candidates at the beginning of the campaign were strongly related to the balance of likes and dislikes at the end of the campaign, even after controlling for prior likes and dislikes, but that likes and dislikes at the beginning of the campaign had only modest effects on thermometer ratings at the end of the campaign after controlling for prior thermometer ratings. They concluded that the likes and dislikes were essentially rationalizations of evaluations formed earlier on other grounds rather than causes of those evaluations.

More generally, prior measurement of explanatory variables may shed significant light on issues of causality even when cross-lag analysis is impossible or uninteresting. The basic rationale in this case is similar: prior values of explanatory variables are less likely than current values to incorporate reciprocal effects of the presumed dependent variable arising from rationalization and the like. In the language of statistical analysis, they may more plausibly be considered "exogenous" rather than "endogenous," making interpretations of their apparent effects a good deal more straightforward and compelling.

Zaller and Hunt's (1995) analysis of support for Ross Perot in the 1992 presidential election provides a useful illustration of the significant differences that may appear in comparing exogenous and endogenous correlates of vote choices. Table 5 presents two distinct sets of parameter estimates for a logit model of support for Perot. The parameter estimates in the first column of the table are derived from an analysis using explanatory variables measured in the fall of 1992, at the same time as the vote choices they are used to explain. The parameter estimates in the second column of the table are derived from an analysis using the same explanatory variables measured two years earlier, when the same respondents were interviewed in the 1990 NES survey.

The parameter estimates in the first column of table 5 suggest that support for Perot in 1992 was strongly related to partisan independence, ideological centrism, worry about the federal budget deficit, and distrust of government—essentially the same set of issues stressed by Perot in his on-again, off-again presidential campaign and by political pundits attempting then and later to explain how Perot succeeded in attracting more popular support than any other challenger to the two-party system in eighty years.

TABLE 5. Correlates of Support for Ross Perot in 1992

	Explanatory Variables Measured in 1992	Explanatory Variables Measured in 1990
Intercept	−1.40	−.37
Strength of party attachment (0–3, pure independent	−.51	−.37
to strong partisan)	(.11)	(.11)
Ideological centrism (centrist = 1; other = 0)	.52	.18
	(.22)	(.23)
Distrust of government (1–5)	.36	−.01
	(.14)	(.12)
Worry about budget deficit (0–1)	.38	.19
	(.21)	(.22)
Disapproval of Congress (1–5)	.11	.12
	(.09)	(.08)
White	1.78	1.88
	(.61)	(.62)
Male	.36	.33
	(.21)	(.21)
Age (in years, logged)	−1.97	−1.73
	(.63)	(.62)
N of cases	901	903

Note: Logit parameter estimates with standard errors in parentheses, from Zaller and Hunt 1995.

By contrast, the parameter estimates in the second column of table 5 for these same explanatory variables measured in 1990 are much smaller—half as large in the case of the budget deficit, one-third as large in the case of ideological centrism, and nonexistent in the case of distrust of government. Only the coefficient for partisan independence remains statistically significant, and even it is almost 30 percent smaller when partisan independence is measured outside the immediate campaign context than when it is measured simultaneously with vote intentions. (By contrast, the coefficients for demographic variables remain essentially unchanged between the two columns, with whites, young people, and perhaps also men being especially likely to support Perot.)

The political implications of these differences should be obvious. The parameter estimates in the first column of table 5 suggest that Perot's independent campaign tapped a variety of strong currents of public dissatisfaction with the prevailing political system: distrust of government, dissatisfaction with the ideological extremism and rigidity of the major parties and their leadership cadres, and concern about the apparent inability or unwillingness of traditional politicians to put the government's fiscal house in order. However, most of these specific currents of public dissatisfaction evaporate in the second column of table 5, leaving lack of prior attachment to the party system (as measured by strength of party identification and less directly by age) as the sole apparent basis of Perot's support. Rather than mobilizing a preexisting complex of specific political discontents, Perot appears to have provided a convenient political rationale for rather free-floating alienation from the prevailing party system.

It seems likely that a good many vote equations would look significantly different if the relevant explanatory variables were measured outside of the immediate campaign setting—and that these differences would alter our interpretations of how and why the relevant explanatory variables seem to matter. Here I provide one more example, derived from Keith et al.'s (1992) analysis of the relationship between party identification and presidential votes. A key piece of evidence offered by Wattenberg (1990) and others in support of the hypothesis of declining partisanship in American electoral politics is that survey respondents are somewhat less likely than in the 1950s to say that they think of themselves as Republicans or Democrats and more likely to claim independence from the parties. However, Keith et al. (1992) pointed out that most of these "independents" acknowledge in response to the standard NES follow-up question that they are "closer" to one party or the other and that these "independent leaners" are, in fact, just as loyal in their voting behavior as self-acknowledged party identifiers.

Table 6 reports parameter estimates from simple probit analyses of three separate NES surveys with (major-party) presidential votes as the dependent variable and categories of party identification as the explanatory variables. The results are reproduced from Bartels 2000a, and more detailed discussions of the data and analysis are presented there. As in table 5, the

TABLE 6. Party Identification and Presidential Votes

	Current Party Identification	Lagged Party Identification	Instrumental Variables
1960 (N = 1,057)			
"Strong" identifiers	1.634	1.250	1.578
	(.103)	(.082)	(.155)
"Weak" identifiers	.866	.804	.669
	(.073)	(.070)	(.200)
Independent "leaners"	1.147	.546	1.185
	(.141)	(.119)	(.601)
Republican bias	.289	.251	.227
	(.054)	(.048)	(.052)
Log-likelihood, pseudo-R^2	−418.0, .43	−506.4, .31	−506.4, .31
1976 (N = 799)			
"Strong" identifiers	1.450	1.224	1.577
	(.117)	(.107)	(.188)
"Weak" identifiers	.684	.707	.491
	(.080)	(.081)	(.243)
Independent "leaners"	.781	.545	.848
	(.109)	(.104)	(.413)
Republican bias	.103	.141	.103
	(.053)	(.051)	(.052)
Log-likelihood, pseudo-R^2	−376.8, .32	−418.4, .24	−418.4, .24
1992 (N = 729)			
"Strong" identifiers	1.853	1.311	1.622
	(.146)	(.109)	(.176)
"Weak" identifiers	.948	.761	.745
	(.099)	(.088)	(.284)
Independent "leaners"	1.117	.530	1.092
	(.122)	(.105)	(.499)
Republican bias	−.073	−.072	−.045
	(.065)	(.057)	(.059)
Log-likelihood, pseudo-R^2	−236.9, .52	−343.1, .30	−343.1, .30

Note: Probit parameter estimates with standard errors in parentheses, from Bartels 2000a. Dependent variable is Republican presidential vote (major-party voters only); party identification variables scored +1 for Republican Party identifiers of indicated strength, −1 for Democratic Party identifiers of indicated strength, and 0 otherwise.

first column reports parameter estimates derived from analyses using party identification measured during the heat of each presidential campaign, and the second column reports comparable parameter estimates derived from analyses using party identification measured two years earlier.

The results in the first column of table 6 are strongly consistent with Keith et al.'s claim that independent leaners are closet partisans; while the leaners are notably less loyal in their presidential vote choices than strong party identifiers, they appear in each of the three election years to be somewhat *more* loyal than weak party identifiers. By comparison, the results in the second column are much less consistent with the notion that independent leaners are partisans at heart. Almost all of the estimated effects in the second column are weaker than those in the first column—not surprising, given the fact that partisan loyalties are here being measured two years before the vote choices they are used to explain. However, the decline from the first column to the second is especially steep in the case of independent leaners (a decline of about 45 percent, as against 23 percent for strong identifiers and 8 percent for weak identifiers).

The final column of table 6 provides comparable parameter estimates from analyses in which lagged party identifications are used as instruments for current party identifications in each election year. This instrumental variables analysis simultaneously addresses the problems of endogeneity and measurement error, which turn out in this instance to have largely offsetting effects—endogeneity biases all of the coefficients upward, whereas measurement error biases them all downward to a roughly similar extent. As a result, the instrumental variables parameter estimates are generally quite similar in magnitude to the simple probit parameter estimates in the first column of the table, except in the case of weak party identifiers. But that correspondence is a happy accident, not a reason to conclude that either problem can safely be ignored.

In general, the availability of panel data seems to provide significant leverage on questions of causal priority that could not be addressed convincingly with cross-sectional data, even from an unusually long and rich rolling cross-section. Similar questions of causal priority arise in every field of nonexperimental research, but they are likely to be especially prevalent in the field of campaign studies, where our aim is to understand complex psychological processes operating under the influence of strenuous efforts by campaigners to forge or dissolve links among a wide variety of potentially relevant political attitudes and perceptions.

Consequences of Panel Attrition and Conditioning

The preceding discussion emphasizes three attractive features of panel designs for the study of campaign effects but does not address the corresponding drawbacks of panel data. The most serious of these is that the sample surveyed in successive panel waves may become increasingly unrepresentative of the original population, either because some of the first-wave respondents cannot be relocated or refuse to be reinterviewed (panel attrition) or because the experience of being interviewed itself affects the subsequent behavior and responses of those who are reinterviewed (panel conditioning). In either case, observed opinion change in the surviving panel sample may provide a biased estimate of the corresponding opinion change in the relevant population.

Failure to reinterview some panel respondents may result in selection bias if the probability of being reinterviewed is correlated with substantively relevant characteristics of the respondents (Heckman 1976). However, selection bias due to panel attrition seems likely to be a relatively minor problem in practice, at least in cases where panel reinterview rates are not too low. In carefully conducted surveys they need not be low. For example, 75 percent of respondents were retained through three waves of the 1980 NES election-year panel, 78 percent over the two years of the 1990–92 NES panel, and about 90 percent over the six weeks or so between typical NES pre- and postelection interviews. Given these sorts of reinterview rates, the respondents who drop out of a panel would have to be quite different from the survivors in order for their absence to produce serious selection bias.

In any case, well-developed econometric techniques exist for analyzing data in the presence of selection bias (Heckman 1979; Achen 1986). Although these techniques have not, as far as I am aware, been applied to the specific problem of panel attrition, Brehm's (1993) analysis of selection bias due to nonresponse in cross-sectional surveys provides a useful (and, for the most part, reassuring) parallel. Brehm used data from the 1986 and 1988 NES surveys to estimate a variety of familiar regression models (of turnout, candidate evaluation, economic voting, issue preferences, and so on) with and without Heckman-style corrections for selection bias. He concluded (1993, 158) that most of the models "escaped with only small changes to the coefficients."[10]

Unfortunately, biases due to panel conditioning are less well understood and techniques to correct them are less well developed. The most straightforward way to investigate the potential effects of panel conditioning is to compare panel responses with responses from a parallel fresh cross-section unexposed to the conditioning process. I provided some comparisons of this sort in a study entitled "Panel Effects in the American National Election Studies" (Bartels 2000b), exploiting the combination of panel reinterviews and fresh cross-section interviews in the 1992 and 1996 NES surveys. As with Brehm's (1993) analysis of nonresponse bias, the results are mostly, though not uniformly, reassuring. Analyses of candidate trait ratings, ideological placements, economic perceptions, and vote choices were largely unaffected by panel conditioning. On the other hand, analyses of campaign interest and turnout produced rather different results among the panel respondents than among respondents in the parallel fresh cross-sections, suggesting that panel conditioning (or panel attrition or both) significantly diluted the value of the panel data for the purposes of those analyses.

Having assessed the magnitude of panel effects in NES surveys, I also proposed some methods for addressing those effects in contexts where they are large enough to be worrisome. Given appropriate auxiliary data, it may be possible to mitigate panel biases using a "semi-pooled" model, including a simple indicator for panel respondents (if the available data include both panel and fresh cross-section components), a proxy for selection bias à la Heckman (if data are available from panel dropouts), or some other fairly parsimonious representation of potential panel effects. If data from a parallel fresh cross-section are sufficiently plentiful, they may facilitate more elaborate "panel adjusted" analyses using a straightforward variant of Franklin's (1990) "auxiliary instrumental variables" estimator. Finally, if auxiliary data and firm prior beliefs about the specific nature of likely panel effects are both lacking, it may be sensible simply to discount the inferential weight of the panel data using "fractional pooling" (Bartels 1996). Examples derived from NES surveys suggest that discounts on the order of 10 to 15 percent may often be appropriate, except for analyses of campaign interest or turnout (Bartels 2000b, 13).

Combining Panel and Rolling Cross-Section Designs

Another potentially serious limitation of traditional panel designs is that they may provide less leverage than rolling cross-sections do on the impact of specific campaign events, if those events happen to occur between successive panel waves. In the 1980 NES panel, for example, observed opinion changes between June and September might in principle be attributable to any event that occurred in the intervening period, including both parties' nominating conventions, economic and foreign policy developments, Reagan's string of campaign gaffes, the emergence of John Anderson as an independent candidate, and media coverage of these and other events. Of course, some attention to *what* opinions changed (Reagan competence ratings? economic perceptions?) and *whose* opinions changed (committed partisans? television news viewers?) may go a long way toward compensating for a lack of precise information about *when* opinions changed. Nevertheless, the sort of simple and elegant interrupted time-series analysis of the impact of specific campaign events facilitated by rolling cross-section designs is likely to be much less feasible with typical panel data.

This limitation of panel data is exacerbated by the inclination of panel study designers to concentrate each wave of interviews in a relatively short and, if possible, relatively tranquil period, in order to minimize heterogeneity in political contexts within each panel wave. Thus, for example, the first two waves of the 1980 NES panel study were conducted in the relatively quiet periods just before and just after the spring primary season. The resulting data are simple to analyze in the traditional framework of panel analysis, since "t_1" and "t_2" may be treated for most purposes as constants rather than variables. However, the cost of this simplicity is that most of the real political action occurs between panel waves and is not amenable to detailed dynamic analysis.

An obvious and potentially attractive solution to this dilemma would be to combine panel and rolling cross-section elements in the same survey design, with each panel wave spanning a variety of potentially significant campaign events. In principle, a *rolling panel design* of this sort might span the entire campaign season, with second-wave interviews beginning as soon as (or even before) first-wave interviews end and so on.

When the separate waves of a panel survey are of short duration and widely separated in time, little is lost by treating all of the interviews in a single wave as though they were conducted at the same time. But the longer each panel wave is in the field—and the higher the proportion of potentially important campaign events that occurs within panel waves rather than between panel waves—the more important and potentially interesting it will be to exploit the date of interview as an explanatory variable within each wave, in essentially the same way that rolling cross-section analysis exploits the date of interview as an explanatory variable.

The unresolved problem is how to specify and estimate a model that melds the distinct approaches to dynamic analysis of traditional panel and rolling cross-section designs. Some of the difficulties may be illustrated by considering a general framework for temporal analysis:

$$Y_{it} = Y_{i0} + \sum_{w=0 \ldots t-1} Y_{iw} \Gamma_w + \sum_{w=0 \ldots t-1} X_i \mathrm{B}_w + E_{it}, \qquad (1)$$

where Y_{it} is a $(1 \times \mathrm{M})$ vector of politically relevant attitudes and perceptions of respondent *i* on date *t*, Y_{i0} is the corresponding set of attitudes and perceptions on an arbitrary "date 0" representing the beginning of the campaign, X_i is a $(1 \times \mathrm{K})$ vector of fixed characteristics of respondent *i*, and E_{it} is a $(1 \times \mathrm{M})$ vector of respondent- and date-specific idiosyncratic components in the set of relevant attitudes and perceptions of respondent *i* measured on date *t*. The parameters to be estimated in this model include *T* distinct $(\mathrm{K} \times \mathrm{M})$ matrices B_w relating the evolution of attitudes and perceptions to fixed characteristics of respondents—one for each day of the campaign—and *T* distinct $(\mathrm{M} \times \mathrm{M})$ matrices Γ_w relating the evolution of attitudes and perceptions to previous attitudes and perceptions—again, one for each day of the campaign.

This model provides ample scope for potential campaign dynamics. Each relevant attitude or perception may evolve on each day of the campaign in a way that depends upon both exogenous characteristics of respondents and prevailing values of endogenous variables. Thus, in particular, reactions to specific campaign events may be conditioned by respondents' current attitudes (for example, candidate evaluations), fixed characteristics (for example, location in a specific media market), or both.

Unfortunately, this model is much *too* general as it stands. Even in a very simple system with five exogenous characteristics, five endogenous attitudes, and sixty campaign days, there would be three thousand distinct parameters to be estimated. Moreover, even if all these parameters could be estimated, they could not be estimated straightforwardly, because most of the previous values of endogenous variables represented in equation (1) by Y_{iw} (for $w = 0 \ldots t$) are unobserved.

In the present context, the only difference between rolling-cross section and panel data is that in a rolling cross-section *all* of the previous values of endogenous variables are unobserved, whereas with panel data at least *some* of these values are observed. In either case, the model must be simplified in a way that finesses the problem of missing data while doing as little violence as possible to the political processes at work in actual campaign settings.[11]

If Y_{it} represents attitudes measured in the second wave of a panel study, we do have one previous set of observations for each respondent, Y_{is}, with $s < t$. Applying the same model to Y_{is} as to Y_{it} and differencing produces

$$Y_{it} - Y_{is} = \sum_{v=s\,...\,t-1} Y_{iv}\,\Gamma_v + \sum_{v=s\,...\,t-1} X_i\,\mathrm{B}_v + (E_{it} - E_{is}). \quad (2)$$

One thing to note about equation (2) is that the original disturbance term E_{it} in equation (1) now becomes $(E_{it} - E_{is})$. This change is likely to be a blessing, since the differencing eliminates any individual-specific fixed effects in the idiosyncratic components.[12] However, the second thing to note about equation (2) is that observed opinion change between the first and second waves still depends upon unobserved intervening values of the endogenous variables. We might deal with this complication in any of (at least) three ways.

(1) If the endogenous variables are believed to have only modest effects on reactions to campaign events, we might assume that $\Gamma_v = 0$ for all v. In that case, we are left with observed changes in the endogenous variables as time-varying functions of exogenous respondent characteristics. Even this simplest approach is likely to represent a considerable improvement over straightforward cross-sectional analysis, since only the incremental impact of endogenous variables, rather than their total impact, is omitted (or absorbed in the effects of the exogenous characteristics).

(2) If the endogenous variables evolve fairly slowly, we might be tempted to simply replace each intervening value Y_{iv} with the corresponding first-wave value Y_{is}. In that case, we are left with observed changes in the endogenous variables as time-varying functions of both exogenous respondent characteristics and first-wave values of endogenous variables. In exchange for the inelegance and added complexity of the resulting model, we at least allow for some possibility of investigating how attitudes at each point in the campaign condition responses to subsequent campaign events.

(3) If the endogenous variables do not evolve slowly and their

dynamic effects are non-negligible, it may be necessary to construct instruments for the intervening values of endogenous variables Y_{iv}, using the exogenous characteristics X_i or the observed first-wave values Y_{is} or both. Obviously, this is a much more daunting approach, since it requires both careful attention to identifying restrictions and a great deal of data. Whether the possibility is even worth pursuing seems to me to be an open question.[13]

Even if it is possible somehow to finesse the problem of missing data for intervening values of the endogenous variables, we are left with the problem of further simplifying the model to reduce the number of parameters to be estimated. A moment's reflection will confirm that distinct date-specific effects cannot be identified for any day on which there are no respondents. Thus, if all first-wave interviews occurred on a single day s and all second-wave interviews on a single day t, only a single $(K \times M)$ matrix B and a single $(M \times M)$ matrix Γ could be estimated for the intervening period $t - s$. Then we are back in the traditional panel framework, with no hope of distinguishing date-specific effects and with the problem of missing data for intervening values of endogenous variables suppressed but not eliminated.

Even if there are panel respondents on every day of the campaign, it will obviously not be feasible or desirable to estimate a separate $(K \times M)$ matrix B_v and a separate $(M \times M)$ matrix Γ_v for every day. This embarrassment of parameters might be reduced in a variety of ways, including

- grouping days into weeks or other potentially meaningful time periods,
- imposing linear or other constraints on the daily effects over part or all of the campaign period,[14] or
- setting daily effects to zero for days with no salient campaign events, such as major speeches, scandals, or debates.

Which of these simplifications (or others or combinations of these and others) will turn out to be most fruitful is a substantive—theoretical and empirical—question. Previous work with rolling cross-section models (e.g., Johnston et al. 1992; Blais and Boyer 1996) provides some relevant suggestions, but insight and experience will be required to adapt these suggestions in order to develop tractable models that fully exploit the distinctive virtues of panel data in complex dynamic settings.

NOTES

News about research design and data analysis have been significantly shaped by the precepts and examples provided over many years by Christopher Achen and Henry Brady. Achen, Kathryn Cirksena, Charles Franklin, Shanto Iyengar, Simon Jackman, John Jackson, Norman Nie, G. Bingham Powell, John Zaller, and anonymous referees provided helpful comments on three previous pieces (Bartels 1993, 2000a, 2000b) from which some of the arguments and data analysis presented here are adapted. Additional arguments and examples derive from discussions with Steven Rosenstone and John Zaller. All of my empirical analysis is based upon data from NES, originally collected by the Center for Political Studies, University of Michigan, and made available through the Inter-University Consortium for Political and Social Research. I am grateful to the Woodrow Wilson School of Public and International Affairs at Princeton University, the John Simon Guggenheim Memorial Foundation, and the Pew Charitable Trusts for generous financial support of the research reported here.

1. The job performance measures included a general approval item plus specific items on Carter's handling of the Iranian hostage crisis, inflation, unemployment, and the energy crisis. The specific trait measures for each candidate included three items each tapping character ("power-hungry," "moral," and "dishonest"), leadership ("inspiring," "provide strong leadership," and "weak"), and competence ("develop good relations with other countries," "solve our economic problems," and "knowledgeable"). Issue preferences and candidate issue placements were also analyzed by Bartels (1993) but are omitted here.

2. To guard against the possibility of estimating spurious partisan or media exposure effects, all of the analyses included age, education, and race as additional control variables.

3. The main assumptions underlying the Wiley and Wiley model are that the measurement process produces constant error variance in each wave of the panel and that measurement errors for the same respondent in different waves of the panel are uncorrelated. I augmented the standard Wiley and Wiley model to incorporate the effects of demographic characteristics—age, education, race, and party identification—on newspaper and television news exposure in each wave of the panel and to allow for possible correlations between unmeasured factors affecting newspaper exposure and television news exposure in each wave (though these correlations turned out to be small: .03 in June and −.13 in September). I also explored a variety of other generalizations of the standard Wiley and Wiley assumptions—for example, by allowing measurement error variances to differ across panel waves or by allowing measurement errors for different responses by the same respondent to be correlated; none produced more than marginal improvements in the statistical fit of the model, and none appreciably changed the substantive results. These results are consistent with those reported by Feldman (1990, 33, 38), who applied the Wiley and Wiley model to a variety of political survey items (party identification, issue positions, and candidate evaluations) using data from a five-wave panel, for which the model is, as here, overidentified. He concluded that "the simple measurement model fits very well."

4. For more on the interpretation and political significance of these media effects, see Bartels 1993; Zaller 1996.

5. A similar pattern appears in estimated newspaper exposure effects not reported here.

6. Green, Palmquist, and Schickler's (2002) work on the measurement properties of party identification provides a good example of benchmark analysis that could easily and fruitfully be drawn upon in subsequent analyses in which party identification appears as an explanatory variable. Of course, some experience will be required to gauge how much the measurement properties of different attitudes and perceptions vary across survey settings. Ironically, we know much more about the measurement properties of party identification than of many other common explanatory variables that are much less reliable.

7. Simply analyzing "change scores" as a function of intervening characteristics or events imposes the restrictive assumption that the magnitude and direction of change is uncorrelated with the original attitude or perception. This restrictive assumption is violated by many reasonable models of opinion change, including Bayesian models of the sort motivating the analysis of Bartels (1993) described here. Analyzing "change scores" with original attitudes as explanatory variables is equivalent to analyzing subsequent attitudes with original attitudes as explanatory variables, except for a one-unit shift in the coefficient attached to original attitudes.

8. I exclude from consideration here analyses in which current opinions are impressively "explained" by essentially redundant contemporaneous factors, as when vote intentions are "explained" by evaluations of the competing candidates.

9. The concept of "Granger causality" in time-series analysis is essentially identical. See, e.g., Freeman 1983; Freeman, Williams, and Lin 1989.

10. The two most notable exceptions to this generalization were a significant overestimate of the intercept in a model of turnout and a significant underestimate of the impact of political information (as measured by familiarity with the candidates) in a model of congressional voting. The latter bias is especially worrisome given the prominence of political information as a conditioning variable in recent campaign studies.

11. Brady and Johnston's (1996) discussion of rolling cross-section models is very much in this spirit. Specific comparisons between their models and the more general formulation offered here are left as an exercise for the reader.

12. In effect, this is another aspect of the increased efficiency of panel estimation referred to in the section "Prior Measurement of Dependent Variables."

13. I note in passing that, well before this point in their parallel discussion of cross-sectional models, Brady and Johnston (1996, 9–10) warn that "[i]f this kind of change occurs, then the estimation problem is very difficult, and the simple methods we are exploring in this paper may become problematic. We hope to explore this problem in detail in other papers, but for the moment we simply note that this poses some substantial difficulties."

14. Some guidance along these lines might be provided by the literature on time-varying parameter models in time-series analysis (Beck 1983; Newbold and Bos 1985).

REFERENCES

Achen, Christopher H. 1975. "Mass Political Attitudes and the Survey Response." *American Political Science Review* 69:1218–23.

———. 1983. "Toward Theories of Political Data." In *Political Science: The State of the Discipline,* ed. Ada W. Finifter. Washington, DC: American Political Science Association.

———. 1986. *The Statistical Analysis of Quasi-Experiments.* Berkeley: University of California Press.

Bartels, Larry M. 1988. *Presidential Primaries and the Dynamics of Public Choice.* Princeton: Princeton University Press.

———. 1991. "Instrumental and 'Quasi-Instrumental' Variables." *American Journal of Political Science* 35:777–800.

———. 1993. "Messages Received: The Political Impact of Media Exposure." *American Political Science Review* 87:267–85.

———. 1996. "Pooling Disparate Observations." *American Journal of Political Science* 40:905–42.

———. 2000a. "Partisanship and Voting Behavior, 1952–1996." *American Journal of Political Science* 44:35–50.

———. 2000b. "Panel Effects in the American National Election Studies." *Political Analysis* 8:1–20.

———. 2000c. "Campaign Quality: Standards for Evaluation, Benchmarks for Reform." In *Campaign Reform: Insights and Evidence,* ed. Larry M. Bartels and Lynn Vavreck. Ann Arbor: University of Michigan Press.

Beck, Nathaniel. 1983. "Time-Varying Parameter Regression Models." *American Journal of Political Science* 27:557–600.

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.

Blais, André, and Martin Boyer. 1996. "Assessing the Impact of Televised Debates: The Case of the 1988 Canadian Election." *British Journal of Political Science* 26: 143–64.

Bollen, Kenneth A. 1989. *Structural Equations with Latent Variables.* New York: Wiley.

Box-Steffensmeier, Janet M., and Charles H. Franklin. 1995. "The Long Campaign: Senate Elections in 1992." In *Democracy's Feast: Elections in America,* ed. Herbert F. Weisberg. Chatham, NJ: Chatham House.

Brady, Henry E. 1993. "It's Not Just Measurement Error: Guessing and Interpersonal Incomparabilities in the Survey Response." Paper presented at the annual meeting of the Midwest Political Science Association, Chicago.

Brady, Henry E., and Richard Johnston. 1996. "Statistical Methods for Analyzing Rolling Cross-Sections with Examples from the 1988 and 1993 Canadian Election Studies." Paper presented at the annual meeting of the Midwest Political Science Association, Chicago.

Brehm, John. 1993. *The Phantom Respondents: Opinion Surveys and Political Representation.* Ann Arbor: University of Michigan Press.

Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent,* ed. David E. Apter. New York: Free Press.

Cook, Thomas D., and Donald T. Campbell. 1979. *Quasi-Experimentation: Design and Analysis Issues for Field Settings.* Boston: Houghton Mifflin.

Feldman, Stanley. 1990. "Measuring Issue Preferences: The Problem of Response Instability." *Political Analysis* 1:25–60.

Finkel, Steven E. 1993. "Reexamining the 'Minimal Effects' Model in Recent Presidential Campaigns." *Journal of Politics* 55:1–21.

Franklin, Charles H. 1990. "Estimation across Data Sets: Two-Stage Auxiliary Instrumental Variables Estimation (2SAIV)." *Political Analysis* 1:1–24.

Freeman, John R. 1983. "Granger Causality and the Time Series Analysis of Political Relationships." *American Journal of Political Science* 27:327–58.

Freeman, John R., John T. Williams, and Tse-min Lin. 1989. "Vector Autoregression and the Study of Politics." *American Journal of Political Science* 33:842–77.

Fuller, Wayne A. 1987. *Measurement Error Models.* New York: Wiley.

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

Heckman, James J. 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models." *Annals of Economic and Social Measurement* 5:475–92.

———. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47: 153–61.

Heise, David R. 1969. "Separating Reliability and Stability in Test-Retest Correlation." *American Sociological Review* 34:93–101.

Jennings, M. Kent, and Richard G. Niemi. 1981. *Generations and Politics: A Panel Study of Young Adults and Their Parents.* Princeton: Princeton University 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, and Henry E. Brady. 2002. "The Rolling Cross-Section Design." *Electoral Studies* 21:283–95.

Jöreskog, Karl G. 1979. "Statistical Models and Methods for Analysis of Longitudinal Data." In *Advances in Factor Analysis and Structural Equation Models,* ed. Karl G. Jöreskog and Dag Sörbom. Cambridge, MA: ABT Books.

Just, Marion R., Ann N. Crigler, Dean E. Alger, Timothy E. Cook, Mantague Kern, and Darrell M. West. 1996. *Crosstalk: Citizens, Candidates, and the Media in a Presidential Campaign.* Chicago: University of Chicago Press.

Keith, Bruce E., David B. Magleby, Candice J. Nelson, Elizabeth Orr, Mark C. Westlye, and Raymond E. Wolfinger. 1992. *The Myth of the Independent Voter.* Berkeley: University of California Press.

Lazarsfeld, Paul F., Bernard Berelson, and Hazel Gaudet. 1948. *The People's Choice.* New York: Columbia University Press.

Markus, Gregory B. 1982. "Political Attitudes during an Election Year: A Report on the 1980 NES Panel Study." *American Political Science Review* 76:538–60.

Newbold, Paul, and Theodore Bos. 1985. *Stochastic Parameter Regression Models.* Newbury Park, CA: Sage.

Patterson, Thomas E. 1980. *The Mass Media Election: How Americans Choose Their President.* New York: Praeger.

Patterson, Thomas E., and Robert D. McClure. 1976. *The Unseeing Eye: The Myth of Television Power in National Elections.* New York: Putnam.

Rahn, Wendy M., Jon A. Krosnick, and Marijke Breuning. 1994. "Rationalization and Derivation Processes in Survey Studies of Political Candidate Evaluation." *American Journal of Political Science* 38:582–600.

Sanders, Mitch S. 1996. "Beneath Stormy Waters: The Evolution of Individual Decision Making in the 1984 and 1988 Presidential Elections." PhD diss., Department of Political Science, University of Rochester.

Vavreck, Lynn. 1997. "More than Minimal Effects: Explaining Differences between Clarifying and Insurgent Campaigns in Strategy and Effect." PhD diss., Department of Political Science, University of Rochester.

Wattenberg, Martin P. 1990. *The Decline of American Political Parties, 1952–1988.* Cambridge, MA: Harvard University Press.

Wiley, David E., and James A. Wiley. 1970. "The Estimation of Measurement Error in Panel Data." *American Sociological Review* 35:112–17.

Wright, Gerald C. 1990. "Misreports of Vote Choice in the 1988 NES Senate Election Study." *Legislative Studies Quarterly* 15:543–63.

———. 1993. "Errors in Measuring Vote Choice in the National Election Studies, 1952–1988." *American Journal of Political Science* 37:291–316.

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

———. 1996. "The Myth of Massive Media Impact Revived: New Support for a Discredited Idea." In *Political Persuasion and Attitude Change,* ed. Diana C. Mutz, Paul M. Sniderman, and Richard A. Brody. Ann Arbor: University of Michigan Press.

Zaller, John, and Stanley Feldman. 1992. "A Simple Theory of the Survey Response: Answering Questions vs. Revealing Preferences." *American Journal of Political Science* 36:579–616.

Zaller, John, and Mark Hunt. 1995. "The Rise and Fall of Candidate Perot: The Outsider versus the Political System—Part II." *Political Communication* 12:97–123.