

This is a preprint of an article published in [Green, Donald Philip and Eric Schickler. 1993. Multiple-Measure Assessment of Party Identification. *Public Opinion Quarterly* 57: 503-35]. Pagination of this preprint may differ slightly from the published version.

Multiple-Measure Assessment of Party Identification

Author(s): Donald Philip Green and Eric Schickler

MULTIPLE-MEASURE ASSESSMENT OF PARTY IDENTIFICATION

DONALD PHILIP GREEN
ERIC SCHICKLER

Abstract Recent research has focused attention on the importance of accounting for measurement error in party identification when modeling the stability of partisanship and the determinants of the vote. Measurement error estimates have in the past been based on a single measure of partisanship observed at multiple points in time, a test-retest methodology that requires fairly strong assumptions about the character of change over time. This article assesses the reliability of the Michigan party identification scale using multiple measures of partisanship at a single point in time. Our data not only corroborate previous test-retest results but also suggest that the accuracy with which partisanship predicts candidate preferences can be enhanced using multiple measures. One measure in particular, a labeled 7-point self-placement continuum, is found to hold significant potential to supplement and illuminate the Michigan scale.

For decades the conceptualization and measurement of party identification have received considerable attention (cf. Budge, Crewe, and Farlie 1976; Campbell et al. 1960; Converse and Pierce 1985; Dennis 1988; Green and Palmquist 1990; Miller 1991). At the same time, however, the way in which surveys have assessed party identification has remained more or less uniform. The vast majority of national surveys conducted over the past 30 years have employed either the Michigan or the Gallup party identification measures (see Appendix for question wording), with private polling firms often using the latter and academic survey organizations leaning heavily toward the former. While uniformity of method is of considerable convenience to those who wish to

DONALD PHILIP GREEN is associate professor of political science at Yale University. ERIC SCHICKLER is a graduate student in political science at Yale University. This research was supported by grants from the Institution for Social and Policy Studies at Yale, 111 Prospect St., New Haven, CT 06520. The authors are grateful to the Inter-University Consortium for Political and Social Research and the Times-Mirror Organization for supplying the data used in this article. Thanks to Paul Abramson, Joel Fetzer, and Brad Palmquist for their comments on an earlier draft.

track responses to the same question over a long span of time, it imposes certain limitations on other research agendas. In particular, it hampers investigation of the statistical reliability of the party identification measure.

Reliability assessment in this case is more than an arcane technical endeavor; it lies at the heart of the debate about the malleability of party identification. The imputed reliability of the party identification scale directly affects one's impression of the dynamics of party affiliation. Analyses that assume party identification to be measured perfectly have found it to be unstable over time and responsive to short-term political forces such as candidate evaluations (Page and Jones 1979), retrospective performance evaluations (Brody and Rothenberg 1988; Fiorina 1981), ideological proximity to the parties (Franklin and Jackson 1983; Jackson 1975), and past voting behavior (Markus and Converse 1979). In contrast, analyses that have taken into account the finding that 10-15 percent of the variance of the 7-point Michigan party identification scale is random noise have found partisanship to be highly stable over time (Achen 1975; Asher 1974; Green and Palmquist 1990, and in press; Krosnick 1991; Palmquist and Green 1992) and unresponsive to the aforementioned short-term forces (Green 1990; Green and Palmquist 1990; Schickler and Green 1993).

Although reliability assessments by different authors have tended to produce similar results across a range of data sets, the way in which reliability has been assessed is not above criticism. In every case of which we are aware, reliability estimation has relied on techniques designed for panel data in which a single indicator of party identification is repeated at different points in time (Heise 1969; Wiley and Wiley 1970, 1974). As noted by Palmquist and Green (1992), Sullivan and Feldman (1979), and others, this test-retest method rests on strong assumptions about the character of the causal process, measurement error, and structural change. In particular, it is assumed that the errors associated with a measure at any given wave are independent of errors that occur at any other point in time. It is also assumed that the factors that produce change in partisanship are independent over time. Finally, one must assume that measurement errors are independent of the factors that induce true change in partisanship (for an extensive review of these assumptions, see Palmquist and Green [1992]).

These are not necessarily invalid assumptions, but they are sufficiently strong to warrant an investigation of reliability that does not rely on test-retest techniques.¹ The purpose of this article is to assess

1. As Palmquist and Green (1992) point out, test-retest methods using single indicators offer certain advantages. The practical demands of survey design and space limitations are reduced, inasmuch as the researcher need only come up with one valid indicator of

the reliability of the Michigan party identification measure using an alternative approach, one that involves multiple measures of partisanship at a single point in time. The advantage of this approach is that it does not require us to assume anything about the character of change in partisanship over time.² Using a variety of different measures of party identification appearing in two surveys, the December 1973 National Opinion Research Center (NORC) Amalgam Survey and the Times-Mirror Spring 1990 Political Update, we obtain results that are broadly consistent with earlier analyses that utilize test-retest methods.

In addition to facilitating structural equation modeling with latent variables, the inclusion of multiple indicators of partisanship has potential implications for prediction and survey design. How can survey instruments be augmented so that the vote is better predicted? To address this question, we assess the reliability of the supplementary measures of party identification. Of the four auxiliary measures of long-term party identification we consider, one is sufficiently reliable to enhance the predictive power of a regression model that includes only the standard Michigan party identification scale. Since this measure, a labeled 7-point self-placement scale, is relatively easy to administer in personal interviews, we recommend its adoption in face-to-face voter surveys.

Conceptualization and Measurement of Party Identification

Party identification is the central theoretical component of the more general construct of partisanship. While partisanship connotes such dimensions as group membership, behavioral expressions of commitment, psychological closeness, policy agreement, and electoral support, the concept of party identification focuses on the individual's self-definition (Green and Palmquist 1990; Campbell et al. 1960, 1986). As such, a valid measure of party identification requires that the respondent identify himself or herself in relation to the political parties. Measures of affect toward the parties are not valid measures of party identification; a valid measure of an individual's attachment to a party

the latent construct. In addition, the administration of just one measure at a given time point avoids the sort of priming and recency effects that might cause nonrandom errors across multiple measures, a concern expressed by Bishop, Tuchfarber, and Smith (1992).² The disadvantage is that measurement errors associated with one question may be correlated with errors associated with other measures employed in the same survey. This problem can be minimized by placing different measures in different parts of the survey and changing the response formats offered. To their credit, the designers of the surveys used here made an effort to take both precautionary steps.

must distinguish his or her momentary likes and dislikes from “attributes felt to be part of his or her persona, or definition of the social self” (Converse and Pierce 1985, p. 144). Phrases such as “think of yourself,” “call yourself,” “place yourself,” or “regard yourself” tap this self-concept, while feeling thermometers or questions that elicit prospective performance evaluations do not.

A related point is that identification is the facet of partisanship that is most likely to represent an enduring, underlying trait. Noting that the stability of partisanship over time is what lends the construct its explanatory power, Converse and Pierce (1985) observe that the standard Michigan partisanship question is particularly suited to this task because it taps self-identity with an extended time horizon. For unlike the wording used by the Gallup organization, which asks respondents about how they think of themselves “in politics as of today,” the Michigan measure is prefaced with the phrase “generally speaking” or “in general.” A good deal of recent scholarship has focused on the extent to which this difference in wording actually affects the over-time stability of partisanship, with conflicting results (Abramson and Ostrom 1991, 1992, 1993; Bishop, Tuchfarber, and Smith 1992; MacKuen, Erikson, and Stimson 1989, 1992). Common to both sides of this debate, however, is a recognition that identification is the key construct of interest.

Many surveys have included measures of different aspects of partisanship, ranging from identification to electoral support, but only a small handful have made serious efforts to measure party identification redundantly (see Abramson and Ostrom 1993). Here, we focus our attention on two such surveys, the December 1973 NORC Amalgam Survey and the Times-Mirror Spring 1990 Political Update.³ The NORC survey used a national multistage area probability sample. At the block level, quota sampling was used. The 1,489 respondents were interviewed in person. The survey was designed in large part to test the effects of question wording. As a result, after the first several questions, respondents were randomly assigned to one of three groups.

3. We elected not to analyze the Political Action Panel Study (Barnes et al. 1988), which contains two measures of partisanship, the Michigan scale and a self-placement scale on which the respondent states to which party he or she feels “closer.” Although the “closeness” question is similar to the second branch of the Michigan scale (directed toward self-identified independents), it is not clear whether it validly measures party identification when not prefaced with a trunk question that specifically calls attention to sense of identity. Also unsuitable were the items available in the 1980 National Election Study (NES), which are either invalid measures of party identification (e.g., party support) or illuminate only specific ranges of the partisan spectrum, such as independence (Dennis 1988). Finally, we chose not to analyze the 1987 NES Pilot Study (Niemi, Reed, and Weisberg 1991) because the items it contains are less diverse than the NORC study we examine here. It should be noted that where the 1987 NES and 1973 NORC studies overlap, the results turn out to be extremely similar.

The three groups were asked identical first and third questions about their party identification, but each group was asked a different second question. The virtue of this design is that any one set of respondents was confronted with only three party identification questions dispersed throughout a lengthy survey, yet data from five measures of partisanship were obtained in all. In the Times-Mirror survey, 3,004 respondents were interviewed in person. The sample used for the survey was a multistage probability sample, consisting of 100 primary sampling units and 1,600 secondary sampling units. Interviewing was conducted in 193 sampling locations. The Times-Mirror study included numerous innovative items designed to gauge various dimensions of partisanship, although only two of the items meet the criteria for the valid measurement of party identification discussed above.

The NORC survey includes five measures of party identification. The first is the standard 7-point Michigan scale. The second item asks the respondent to place him- or herself on a fully labeled 7-point scale, with strong Democrats and strong Republicans on either end and independents in the middle (cf. Krosnick and Berent 1993). We refer to this item as the *7-point self-placement scale*. The third item, referred to as the *forced-pair comparison item*, is identical to the first part of the branched Michigan question but requires the respondent to volunteer independent partisanship. The fourth item contains an even more explicit long-term prompt than the Michigan item: respondents are asked how they “basically” think of themselves, regardless of recent voting behavior or future vote intentions. We refer to this as the *basic self-regard item*. The final item uses the standard Gallup “politics-as-of-today” question wording but does not probe the strength of partisanship or attempt to identify leaning partisans. We refer to this item as the *3-point Gallup question*. Unlike the other items in the NORC survey, the prompt in this item is explicitly short-term.

The Times-Mirror survey includes two measures of party identification. The first is a 7-point scale using the Gallup politics-as-of-today prompt. The second is the difference between the respondent's view of the suitability for him or her of the descriptions “a Republican” and “a Democrat.” This party difference score, which arguably taps a longer-term self-conception, is a 19-point scale that ranges from -9 to +9.

Comparing Alternative Measures of Party Identification

UNIVARIATE STATISTICS

The marginal distributions of the seven measures of partisanship are presented in tables 1-4. Looking first at the five NORC questions, we

Table I. Response Distribution of Michigan 7-Point Party Identification and 7-Point Self-Placement Scale: NORC

Identification	Michigan (%)	Self-Placement (%)
Strong Democrat	18.4	13.8
Weak Democrat	19.9	...
Prefer Democrats	...	14.8
Democrat, strength unclear	1.7	...
Leaning Democrat	13.5	18.9
Independent	9.4	22.7
Leaning Republican	7.5	13.0
Republican, strength unclear	1.1	...
Prefer Republicans	...	8.7
Weak Republican	12.4	...
Strong Republican	8.1	4.1
Other party	3.5	...
Don't know/refused	4.5	4.0
Total	100.0	100.0
<i>N</i>	1,489	507

Table 2. Response Distribution of Forced-Pair Comparison, Basic Self-Regard, and Gallup 3-Point Scale: NORC

Identification	Forced Pair (%)	Basic Self-Regard (%)	Gallup (%)
Democrat	47.9	43.8	40.5
Republican	23.3	19.3	21.0
Independent		27.8	33.4
Independent (volunteered)	18.7		
Other (volunteered)	3.2	1.8	2.1
Don't know/refused	6.8	7.3	3.0
Total	99.9	100.0	100.0
<i>N</i>	497	493	1,489

Table 3. Response Distribution of Gallup 7-Point Scale: Times-Mirror 1990 Study

Identification	Percentage
Strong Democrat	17.3
Weak Democrat	18.5
Lean Democrat	11.9
Independent	11.3
Lean Republican	12.0
Weak Republican	15.2
Strong Republican	12.3
Refused or not applicable	1.6
Total	100.1
<i>N</i>	3,004

find that the responses are fairly consistent across the Michigan, Gallup, and basic self-regard measures. The contrast in wording between Michigan and Gallup (“generally speaking” as opposed to “in politics as of today”) seems to have had little effect on the marginal distribution of responses (see Bishop, Tuchfarber, and Smith [1992] for a similar result; and Abramson and Ostrom [1993] for a somewhat contrasting finding). When asked to ignore their past voting tendencies, respondents in 1973 tended to report greater Democratic affiliation. Even so, erasing the partisan coloration of two recent Republican presidential victories had only modest effects on respondents' reported party affiliation. Pooling the three items together—the 3-point Michigan classification, the Gallup measure, and the basic self-regard item—we find no support for the hypothesis that differences in wording affected the frequency of responses. A χ^2 (4) test of independence yields a statistic of 5.62, well below the critical value needed to reject the null hypothesis of no differences due to question wording ($\alpha = .05$). Not surprisingly, the proportion of the sample identifying with the two major parties goes up significantly when the trunk question of the Michigan measure omits reference to independents. If anything, the surprise is that even when independent identification is a response option that must be volunteered, it still accounts for 19 percent of the sample. In contrast, when respondents are presented with a visual 7-point continuum on which they must place themselves, the availability of a middle option attracts 23 percent of the sample, more than twice the proportion identifying as pure independents on the Michigan scale. Evidently, as others (Barnes et al. 1988; Keith et al. 1992; Ken-

Table 4. Response Distribution of Party Difference Score: Times-Mirror 1990 Study

Party Difference Score	Percentage
-9	18.5
-8	2.8
-7	3.4
-6	3.4
-5	3.2
-4	4.2
-3	2.0
-2	3.4
-1	2.5
0	17.4
1	2.9
2	3.3
3	2.9
4	3.7
5	3.0
6	2.8
7	3.2
8	2.0
9	9.3
Don't know (for Republican description)	.5
Don't know (for Democratic description)	2.9
Not applicable	2.8
Total	100.1
<i>N</i>	3,004

NOTE.—See Appendix for scoring.

ney and Rice 1988) have suspected, the ratio of partisans to independents is a malleable quantity, sensitive to the format of survey items. One further point concerning the NORC data deserves mention. Each respondent addressed three party identification questions, beginning with the branched Michigan scale and ending with the trunk of the Gallup scale. In between, respondents were assigned to one of three experimental questions. Thus, the Gallup question may be regarded as a posttest in a priming experiment, where the self-placement scale, the forced-pair question, and the basic self-regard questions serve as the primes. Interestingly, there is no evidence of any priming

effect on either the mean ($\chi^2 [4] = 8.8, p > .05$) or variance ($F[2, 1,272] = 1.48, p > .05$) of responses to the Gallup item. While this finding, by itself, does not show that priming effects are of no concern in multiple-measure assessments of party identification, it does at least indicate that when spaced a few minutes apart, consecutive measures do not bleed into one another.

Turning to the Times-Mirror data, we consider the 19-point scale constructed by taking the difference between “self-regard” ratings of the two parties. Like the Michigan or Gallup 7-point scales, the party difference score is nonnormal in appearance, with clusters at both endpoints. Despite the length of the scale, the endpoints account for 28 percent of the valid responses. In addition, like the self-placement scale examined above, approximately one-fifth of the sample ends up at the midpoint. Thus, while the many gradations of partisanship offered by the scale make it a potentially attractive item, the pattern of responses suggest that it fails to overcome some of the limitations of other measures.

BIVARIATE RELATIONSHIPS

When one cross-tabulates responses to the Michigan scale with answers to the alternative measures used in the NORC survey, the results indicate quite clearly that the standard party identification measure is not an errorless indicator of partisanship (see tables 5-8). Response variation is particularly evident in the more articulated scales, such as self-placement, where an occasional respondent identified as a strong partisan by the Michigan measure surfaces in the middle of the partisan continuum. But even when the 7-point scale is arrayed against a crude 3-point classification, “party switching” occurs. For example, when offered a forced-pair choice between the Democratic and Republican parties, three of the 198 respondents dubbed Democrats by the Michigan measure called themselves Republicans. Conversely, four of 94 ostensible Republicans called themselves Democrats. Thus, the total rate of interparty conversion is 7/292 (or 2.4 percent), a rate that exceeds what one typically observes in panel studies where respondents are interviewed several months apart (Green and Palmquist, *in press*). Somewhat less interparty conversion is evident when the comparison item is the 3-point Gallup measure, which furnishes a middle response option. Nevertheless, as Nie, Verba, and Petrocik (1979) note, 2.0 percent of the 917 Michigan partisans underwent “conversions.”⁴

4. The corresponding rates for the basic self-regard item and the self-placement scale are 1.9 percent (6/322) and 1.0 percent (3/309). The rate for the latter is depressed due to the relative paucity of respondents choosing to call themselves strong or weak partisans when presented with a 7-point continuum.

Table 5. Cross-Tabulation of Michigan 7-Point Party Identification and 7-Point Self-Placement

Self-Placement 7-Point Scale	Michigan Identification											
	Strong Democrat	Weak Democrat	Democrat, Unclassified	Lean Democrat	Independent	Lean Republican	Republican, Unclassified	Weak Republican	Strong Republican	Other	Don't Know or Refused	Total
Strong Democrat	58	6	2	0	0	0	0	0	0	2	2	70
Weak Democrat	22	36	6	4	0	0	0	2	0	2	3	75
Lean Democrat	8	40	1	37	2	2	0	3	0	2	1	96
Independent	4	2	1	29	36	17	2	9	1	8	6	115
Lean Republican	0	0	0	0	3	23	4	27	7	2	0	66
Weak Republican	0	1	0	0	0	1	1	21	20	0	0	44
Strong Republican	0	0	0	0	0	0	2	2	17	0	0	21
Other/don't know	0	1	0	0	3	1	0	2	1	7	5	20
Total	92	86	10	70	44	44	9	66	46	23	17	507

Table 6. Cross-Tabulation of Michigan 7-Point Party Identification and Forced-Pair Comparison

Forced-Pair Comparison (Volunteer “Independent”)	Michigan Identification											Total
	Strong Democrat	Weak Democrat	Democrat, Unclassified	Lean Democrat	Independent	Lean Republican	Republican, Unclassified	Weak Republican	Strong Republican	Other	Don't Know or Refused	
Democrat	75	102	7	34	5	1	0	4	0	4	6	238
Independent (volunteered)	0	4	0	28	39	15	0	0	0	1	6	93
Republican	0	3	0	1	2	18	4	45	41	1	1	116
Other (volunteered)	0	0	0	0	5	0	0	0	0	10	1	16
Don't know/refused	2	5	0	2	6	1	0	0	0	1	17	34
Total	77	114	7	65	57	35	4	49	41	17	31	497

Table 7. Cross-Tabulation of Michigan 7-Point Party Identification and Basic Self-Regard

Basic Self-Regard	Michigan Identification											Total
	Strong Democrat	Weak Democrat	Democrat, Unclassified	Lean Democrat	Independent	Lean Republican	Republican, Unclassified	Weak Republican	Strong Republican	Other	Don't Know or Refused	
Democrat	98	82	7	17	0	0	0	3	1	2	6	216
Independent	5	11	0	49	33	23	0	9	1	2	4	137
Republican	0	2	0	0	0	7	3	53	29	1	0	95
Other (volunteered)	0	0	0	0	1	0	0	2	0	6	0	9
Don't know/refused	2	7	1	0	6	4	0	3	3	1	9	36
Total	105	102	8	66	40	34	3	70	34	12	19	493

Table 8. Cross-Tabulation of Michigan 7-Point Party Identification and Gallup 3-Point Scale

Gallup 3-Point Scale	Michigan Identification											Total
	Strong Democrat	Weak Democrat	Democrat, Unclassified	Lean Democrat	Independent	Lean Republican	Republican, Unclassified	Weak Republican	Strong Republican	Other	Don't Know or Refused	
Democrat	264	254	21	26	5	0	0	5	2	12	14	603
Independent	7	33	2	173	124	95	3	28	4	10	18	497
Republican	2	8	1	1	3	16	13	151	114	2	2	313
Other (volunteered)	0	1	0	1	4	0	0	0	0	23	3	32
Don't know/refused	1	1	1	0	4	1	0	1	0	5	30	44
Total	274	297	25	201	140	112	16	185	120	52	67	1,489

That responses to the different party identification queries vary could of course be due to the unreliability of the auxiliary measures, not the traditional party identification scale. This possibility, however, is ruled out by the pattern of correlations among the measures of party identification (see tables 9-11). Notice that all but one of the interitem correlations (Pearson r) for the NORC data fall between .85 and .89, the lone outlier being a .92 correlation between the 3-point Gallup measure and “basic self-regard” (table 11).⁵ Of particular interest is the fact that the correlations between the Michigan scale and the other four items only vary between .87 and .89. In other words, casual inspection of the correlation matrices seems to suggest that the reliabilities are roughly the same for each of the measures. We will return momentarily to a more rigorous assessment of reliability using these patterns of interitem relationships.

Another pattern of note in the cross-tabulation between the Michigan scale and alternative measures is the differential rate of dispersion across the seven partisanship categories. Returning momentarily to table 7, consider the cross-tabulation of the 3-point basic self-regard measure with the Michigan scale. Leaving aside “don't know/other” responses, we find that five of the 160 weak partisans switch parties, as opposed to zero of the 96 leaning independents. None of the pure independents budge at all. The same pattern emerges in the cross-tabulation between the Michigan item and the 7-point self-placement scale (table 5). Again, weak partisans appear more heterogeneous than pure independents, with leaning partisans falling somewhere in between.

Nevertheless, it should be noted that the cross-tabulations presented in tables 5-8 in every case show the party identification scale to be ordinal in nature. None of the five criterion measures for the three NORC subsamples ever produce an “intransitivity” of the sort detected by Petrocik (1974) and documented more recently by Keith et al. (1992). Indeed, in each instance leaning partisans are quite different from weak partisans in terms of the ways in which they respond to alternative measures of partisanship, as are strong and weak partisans. That the ordered series of partisan categories in the Michigan scale do not always produce a monotonic relationship with presidential vote

5. Apparently, when respondents are asked to consider their party identification apart from their past voting behavior (see wording to the basic self-regard item in the Appendix), they tend to focus on the aspect of their partisanship that is akin to party affiliation in politics-as-of-today. Perhaps this explains why political evaluations, such as presidential approval, do not always correlate differently with the Gallup and Michigan measures asked at the same point in time (cf. Abramson and Ostrom [1993] and Bishop, Tuch-farber, and Smith [1992] for opposing findings on this matter). While politics-as-of-today picks up short-run considerations and is apparently more unstable over time (Abramson and Ostrom 1993), it also may shed considerations that were part of past voting decisions, factors that may be correlated with contemporary political evaluations.

Table 9. Means, Standard Deviations, and Correlation Matrix for Experimental Group 1 ($N = 438$)

Variable	Mean	Standard Deviation	Michigan Party ID	7-Point Self-Placement	Gallup 3-Point Scale
Michigan 7-point party ID	3.525	2.052	1.000		
7-point self-placement scale	3.527	1.674	.886	1.000	
Gallup 3-point scale	2.676	1.529	.880	.850	1.000

NOTE.—Data were deleted in listwise fashion in computing the correlations.

Table 10. Means, Standard Deviations, and Correlation Matrix for Experimental Group 2 ($N = 415$)

Variable	Mean	Standard Deviation	Michigan Party ID	Forced-Pair Comparison	Gallup 3-Point Scale
Michigan 7-point party ID	3.398	1.979	1.000		
Forced-pair comparison	2.470	1.705	.874	1.000	
Gallup 3-point scale	2.590	1.590	.873	.857	1.000

NOTE.—Data were deleted in listwise fashion in computing the correlations.

Table 11. Means, Standard Deviations, and Correlation Matrix for Experimental Group 3 ($N = 422$)

Variable	Mean	Standard Deviation	Michigan Party ID	Basic Self-Regard	Gallup 3-Point Scale
Michigan 7-point party ID	3.256	2.012	1.000		
Basic self-regard	2.483	1.580	.871	1.000	
Gallup 3-point scale	2.493	1.571	.878	.924	1.00

NOTE.—Data were deleted in listwise fashion in computing the correlations.

preference may say less about defects of the partisanship measure than about the deficiencies of focusing on the simple bivariate relationship between vote preference and party identification.

The cross-tabulation between the 7-point Gallup item and the party difference scale (table 12) reveals both the transitive ordering of party identification categories and the heterogeneity within each category. Again, weak partisans and leaning independents differ substantially in terms of their locations on the party difference score, although considerable dispersion is evident in both categories. One defect of the party difference scale, as suggested earlier, seems to be the encouragement of indifferent responses. While this accurately gauges the sentiments of those who rate as pure independents by the Gallup measure (despite the fact that the party difference questions make no mention of "independent" identification), it also attracts many respondents in other partisan categories, generating a lumpy distribution in several of the columns of table 12. Nonetheless, the bivariate correlation between the two partisanship measures is .873.

CORRELATIONS WITH DEMOGRAPHIC VARIABLES AND CANDIDATE PREFERENCES

If the various measures of party identification tap the same underlying trait with random error, we should expect to find that they are similarly correlated with determinants and consequences of party affiliation. The only systematic difference we would expect to observe between party identification measures is that less reliable items have more attenuated correlations with assorted demographic variables and vote preferences.⁶

The pattern of correlations is reported in tables 13 and 14. On the whole, the Michigan item is not very different from the other party identification measures in terms of its relationship to the seven criterion variables. The self-placement item is on par with the Michigan measure; the other three items in the NORC study rate slightly lower than the Michigan item in terms of reliability, although the comparison

6. The demographic variables are years of education, dummy variables for female, nonwhite, Catholic, Southern residence, and age. (In the statistical analysis that follows, age is entered as a series of three dummy variables.) For the NORC survey, the dependent variable is the sum of responses to five hypothetical 1976 presidential election match-ups. Each response is coded "-1" if the respondent would vote for the Democrat, "1" if he or she would vote for the Republican, and "0" if he or she cannot decide. The five trial heats are Humphrey-Percy, McGovern-Goldwater, Kennedy-Percy, Humphrey-Reagan, and Kennedy-Reagan. The Times-Mirror survey contains only one vote preference question, regarding the intended vote for U.S. House of Representatives in the upcoming Congressional election. This variable is scored "-1" if the respondent intends to vote for the Democratic candidate, "1" if Republican.

between Michigan and Gallup items may be confounded by the long-term versus short-term referent of the questions. The Times-Mirror data provide only a 7-point Gallup measure as a yardstick for comparison, but by this standard it would appear that the 19-point party difference measure is a shade worse in terms of reliability.

Assessing the Measurement Properties of the Michigan Scale

Having examined the relationships between the Michigan party identification measure and alternative measures of party affiliation, we now take up the task of assessing the reliability of the Michigan measure using the NORC data.⁷ In order to gauge the degree of measurement error in the Michigan item, we constructed a structural equations model with latent variables (Bollen 1989; Hayduk 1987; Jöreskog and Sörbom 1988) that integrates three sets of cross-sectional information: (i) redundant measures of partisanship, (ii) demographic correlates of partisanship, such as age cohort, race, religion, education, and region, and (iii) presidential candidate preferences across a series of trial heats. Our structural model assumes that partisanship and various demographic variables are mutually intercorrelated causes of candidate preferences. In this respect, our analysis mimics a simple regression model in which candidate preferences are regressed on a list of predictors that includes demographic variables and partisanship (cf. Shanks and Miller 1990). The main departure from standard regression analysis is that partisanship is measured twice in each of the three experimental groups.⁸

Redundant measures of party identification within each experimental group enable us to identify the key parameters of interest and thereby calculate the reliability with which party identification is measured. The availability of three experimental groups, all of which confronted a common set of demographic, party identification, and candidate preference questions, improves the precision with which these parameters may be estimated. Using a multiple-group analysis (Alwin and Jackson 1981; Hayduk 1987; Lomax 1983), we estimated the variance of the

7. Since the Times-Mirror study does not contain this item, we will focus our attention on the NORC data. In a separate analysis not reported here, we analyzed the reliabilities of the two-party identification measures available in the Times-Mirror study. These results, which are available from us on request, show these items to be less reliable than the alternative measures developed by NORC.

Note that our analysis uses only the long-term partisanship questions. To have added the Gallup questions would have risked confounding measurement error with potentially meaningful differences between long-term and short-term partisan self-regard.

Table 12. Cross-Tabulation of Party Difference Score and Gallup 7-Point Scale

Party Difference Score	Gallup 7-Point Scale								
	Strong Democrat	Weak Democrat	Lean Democrat	Independent	Lean Republican	Weak Republican	Strong Republican	Refused or No Answer	Total
-9	315	191	34	5	1	1	2	6	555
-8	44	29	7	3	0	0	0	0	83
-7	47	41	12	2	0	0	0	1	103
-6	25	54	18	3	0	1	0	0	101
-5	19	39	33	1	0	1	0	3	96
-4	18	49	45	4	3	6	1	1	127
-3	6	26	20	4	2	0	0	2	60
-2	10	27	47	9	3	3	2	0	101
-1	7	15	29	11	4	4	2	2	74
0	6	49	83	202	118	38	6	22	524
1	2	4	2	10	38	23	7	2	88

2	1	4	1	4	46	38	4	0	98
3	1	3	1	6	33	31	11	0	86
4	1	1	0	2	37	55	15	1	112
5	0	0	0	1	18	43	25	2	89
6	1	1	1	2	11	49	20	0	85
7	0	0	0	0	8	42	47	0	97
8	0	0	0	0	6	22	31	0	59
9	2	0	3	2	13	76	181	2	279
Don't know (Republican)	2	6	2	2	1	1	2	0	16
Don't know (Democrat)	3	4	9	56	6	4	2	2	86
Refused or no answer	11	13	9	9	10	19	10	3	85
Total	521	556	356	338	359	457	368	49	3,004

NOTE.—See Appendix for scoring.

Table 13. Correlations of NORC Party Identification Measures with Criterion Variables

Criterion Variable	Group 1 (<i>N</i> = 361)		Group 2 (<i>N</i> = 351)		Group 3 (<i>N</i> = 345)		Total (<i>N</i> = 1,086)	
	Michigan Party ID	Self-Placement	Michigan Party ID	Forced Pair	Michigan Party ID	Basic Self-Regard	Michigan Party ID	Gallup 3-Point
Vote index	.672	.715	.696	.657	.668	.643	.670	.653
Nonwhite	-.294	-.321	-.290	-.292	-.212	-.237	-.253	-.272
Education	.160	.212	.095	.073	.192	.200	.149	.152
South	.025	.031	.016	.015	-.002	-.005	.015	.001
Catholic	-.167	-.139	-.175	-.134	-.216	-.170	-.187	-.153
Female	.007	.027	-.048	-.038	.081	.053	.015	.010
Age	-.007	-.056	.115	.107	-.029	-.032	.033	.002

Table 14. Correlations of Times-Mirror Party Identification Measures with Criterion Variables ($N = 2,073$)

Criterion Variable	Gallup 7-Point	Party Difference Score
Vote intention	.834	.803
Nonwhite	-.288	-.292
Education	.145	.131
South	-.048	-.032
Catholic	-.024	-.013
Female	-.072	-.079
Age	-.038	-.028

NOTE.—Data were deleted in a listwise fashion in computing the correlations.

latent partisan factor, its effect on the vote, and the measurement error variance associated with each of the party identification measures. The model's critical assumption is that the measurement errors associated with the party identification items are random, which is to say that error in each measure of party identification bears no systematic relationship to the other variables in the model or the measurement errors associated with them.⁹ The precise specifications of our structural equations with latent variables model are described in a technical appendix available from us.¹⁰

Since the redundancy with which partisanship is measured produces 8 *df* in each of the three experimental groups and the constraints imposed across groups produces an additional 8 *df*,¹¹ the entire model

9. The success of the model does not hinge on the assumption that the demographic variables or candidate preference measure are error free. The reliability of partisanship will be estimated consistently provided these variables are measured with random error.

10. Since the partisanship measures are not continuous and are arguably ordinal rather than metric, we experimented with ordinal factor analysis (Jöreskog and Sörbom 1988). Reestimating the models with weighted least-squares using polychoric correlation matrices as input produce results similar to the maximum likelihood estimates we report (see Green and Palmquist [in press] for a similar finding). For the sake of simplicity, therefore, we rely on more standard techniques for the analysis of covariance structures (Bollen 1989; Long 1983). All of the results reported here are obtained using maximum likelihood estimation.

11. The variance of party identification, the measurement error associated with the Michigan scale, and the structural disturbance term of the candidate preference equation are constrained to be equal across groups. A similar across-group constraint is imposed on the unstandardized effect of party identification on candidate preference, for a total of four constrained parameters. Since these four parameters are estimated once for three experimental subgroups, there is a savings of 8 *df*.

has a total of 32 *df*. The over identification of the model has two benefits. First, it affords us the opportunity to assess the fit between model and data by means of a χ^2 goodness-of-fit test. Second, the large stock of surplus information contributes to the robustness of the estimates in the face of sampling error and certain forms of model mis specification (Green, Goldman, and Salovey 1993). Put simply, the correlations among the partisanship measures, the demographic variables, and candidate preferences enable us to triangulate on the measurement error variance associated with the different measures of party affiliation. To the extent that these correlations reveal that different partisanship measures are associated differently with other variables in the analysis, the statistical fit of the model will be poor, and we will be forced to question whether the assumptions underlying the analysis are valid.

The results of our analysis, presented in table 15, suggest not only that the fit between model and data is very good ($\chi^2 [32] = 30.2, p = .56$) but also that the reliability of party identification is almost exactly what has been obtained through test-retest methods. The error variance associated with the Michigan party identification scale is .49 (SE = .06), which implies a reliability of approximately .89. By comparison, Green and Palmquist's (in press) analysis of eight multiwave panel studies produced a median error variance of .52 and a median reliability of .87. That the two methodological approaches produce similar findings—a close fit between theory and data and reliability estimates indicating that slightly more than 10 percent of the observed variance in the Michigan scale is random noise—does not prove that these models are correct in their fundamental assumptions. At the same time, the fact that two analyses that rest on different statistical premises produce consistent results means that those who have been skeptical of earlier reliability results (Franklin 1992) can only sustain their skepticism by adopting a significantly more elaborate set of objections to the way in which measurement error is modeled.

As far as the alternative measures are concerned, the only measure whose reliability is at least as great as that of the Michigan scale is the self-placement item. In fact, this item slightly outperformed the Michigan scale in terms of reliability (.92 versus .89), although the difference falls short of statistical significance. None of the other alternative measures of party identification had a reliability exceeding .87.

Selecting Companion Measures to Improve Prediction

In addition to facilitating reliability assessment and parameter estimation in structural equation models, the inclusion of redundant measures

Table 15. LISREL Parameter Estimates for NORC Model

	Unstandardized Factor Loadings	Error Variance	Reliabilities			Estimate	R^2	
			Group 1	Group 2	Group 3			
Michigan 7-point party ID	1.000	.486 (.060)	.887	.880	.887	
7-point self-placement scale	.837	.247 (.051)	.915			
Forced-pair comparison	.850	.388 (.061)		.870		
Basic self-regard	.760	.422 (.057)			.838	
Additional parameters of interest:								
Party identification trait variance				3.733	...	
Unstandardized effect of partisanship				1.149 (.044)	...	
Standardized effect of partisanship651	...	
Experimental group:								
Group 1 (self-placement)	58.4	
Group 2 (forced-pair comparison)	60.6	
Group 3 (basic self-regard)	56.1	

NOTE.—Dependent variable is an index of preferences concerning five hypothetical 1976 presidential election match-ups. Fit of the model: $\chi^2(32) = 30.17$ ($p = .559$). Numbers in parentheses are standard errors.

of partisanship offers an opportunity to improve the accuracy with which party identification predicts vote preferences. To the extent that party identification is measured with random error, the bivariate relationship between it and the vote will be attenuated. One way to recapture some of the lost predictive capacity of the Michigan party identification measure is to add one or more supplementary measures of partisanship to the regression equation.¹²

One inference to be drawn from the cross-tabulations presented earlier in tables 5–8 is that the most useful supplementary measures are those that can clarify the heterogeneity within the categories of “weak partisans” and discern gradations in partisanship even among those classified by the Michigan measure as “strong partisans.” There are, however, practical limitations to the sorts of questions that can be asked in the context of a single survey. It seems awkward to ask both parts of the branched Michigan question twice in the same survey.¹³ At most, the Michigan measure may be supplemented by the trunk question of the Gallup scale, as in the NORC study (and even here, there is an uncomfortable similarity in question wording and response format). As we saw in tables 5–8, supplementing the Michigan scale with an additional 3-point Gallup measure does little to illuminate the differences among self-professed partisans. The party difference measure employed in the Times-Mirror study is a step in the right direction, in that it does not duplicate the format of the Michigan questions. Even better, because the responses it elicits do not cluster at the end points of the scale, is the 7-point self-placement scale used in the NORC study.

How much variance in candidate preferences do the auxiliary measures predict over and above the standard Michigan scale? In order to give all of the measures their best opportunity to account for variance, we divide each of the variables into k dummy variables, where k is one less than the number of valid response categories. The baseline regression model uses only the 7-point Michigan scale as a predictor. Here, candidate preferences are regressed on six dummy variables representing the various categories of the Michigan scale. The alternative regression model adds a series of dummy variables representing particular auxiliary measures of partisanship. The two regression mod-

12. In doing so, one must be careful about interpreting the regression estimates. As Achen (1985) points out, the slope coefficients are biased when redundant measures are introduced. For purposes of prediction, the relevant statistics are the proportion of variance explained (R^2) and the standard error of estimate. The latter statistic penalizes the regression model for the number of predictors it uses. See Johnston (1984).

13. In the 1987 NES Pilot Study, repeated fully-branched Michigan items caused one-fourth of the sample to comment spontaneously to the interviewers that they had just been asked that question.

els are then compared by means of an F -test to determine whether the inclusion of the supplementary measures contributed significantly to the predictive accuracy of the model.

The NORC data reveal that the forced-pair comparison, the basic self-regard item, and the 3-point Gallup measure provide modest improvement in predictive accuracy over and above what was obtained using the Michigan scale alone. A series of F -tests reveal that each of the three alternative measures, when added to the Michigan 7-point scale, contributes significantly ($p < .05$) to the predictive accuracy of the model. Yet in no case did the inclusion of one or more of these three measures improve the R^2 by more than .03. The only NORC measure to make an appreciable and statistically significant contribution to predictive accuracy is the 7-point self-placement measure, which raised the R^2 from .46 to .53. The success of this measure is consistent with our earlier finding suggesting that the self-placement scale is the most reliable of the alternative measures of party identification.

Turning to the Times-Mirror study, we find that the party difference index provides a marginally improved forecast of intended House of Representatives vote. Together with the 7-point Gallup measure, this item generates an R^2 of .76; when the 7-point Gallup item is the sole regressor, the R^2 drops to .74, a statistically significant ($p < .05$) but small change. In one sense, this disappointing result underscores the limitations of the party difference measure noted above. In addition, these results together with those obtained from the NORC study show how hard it is to improve on the predictive power of the Michigan scale, given its reliability of approximately .89. In order to have an appreciable effect on the coefficient of determination, an auxiliary measure of partisanship must be at least as reliable, a criterion met only by the NORC self-placement scale.

Before drawing this exercise to a close, let us consider one additional predictive model that takes into account the specific pattern of responses to alternative measures of party identification. Earlier, we included dummy variables for the Michigan scale and a supplementary measure on the assumption that the two scales contributed additively to the observed measure of candidate preference. We now create dummy variables for each observed combination of responses to the different party identification questions so as to allow for the possibility that different patterns of answers reflect distinct attitudes. This specification, while a bit eccentric, is not without theoretical significance. To find that this model produces a significant improvement in fit would run counter to the basic premise of our earlier reliability analysis, namely, that responses to various party identification items represent linear manifestations of a single underlying party identification factor

measured with error. If responses to different questions interact with one another, then part of what our measurement model dubs random error is really substantively meaningful variation.

Again, *F*-tests were employed to compare the predictive accuracy of two models, one with dummy variables for each configuration of responses and another with dummy variables just for main effects. In each instance, this test furnishes no evidence that interaction terms enhance the model's accuracy ($p < .05$). The only potential exception is the Gallup 3-point scale, which yields statistically significant improved prediction when interactions are included. Although this finding may provide some support for those who argue that the Gallup and Michigan items measure different traits, it is important to note that the interaction effect is partially an artifact of sample size. When the regression results are examined for each experimental group separately, one finds no evidence of any interaction effects.¹⁴ The results, in other words, are compatible with the notion that the various measures of party identification, with the possible exception of the Gallup item, tap a single underlying trait with random error.

Discussion

The reliability assessment undertaken here, in keeping with previous results, indicates that although the Michigan scale is highly reliable compared to other attitude measures (Achen 1975; Krosnick 1991; Krosnick and Berent 1993), it is nonetheless fallible. In the NORC sample roughly 89 percent of the observed variance in the Michigan scale may be attributed to variation in partisanship; the remaining 11 percent is some combination of interviewer biases, misread or misunderstood questions, incomplete translation of the partisanship continuum into a discrete 7-point categorization, coding mistakes, and an inability on the part of respondents to gauge their own partisan sentiment.

To say that the reliability of the Michigan scale is roughly .89 does not appear to pose much of a challenge to previous studies that have presumed this measure to have a reliability of 1.00. But in fact the difference between .89 and 1.00 is sufficiently large to undermine many of the "revisionist" claims against the original notion (Campbell et al. 1960) that partisanship is a trait that remains stable over the course of adulthood. The empirical foundation for many revisionist challenges to the original conception of party identification is that a substantial

14. The Times-Mirror study reveals a small but statistically significant gain from the inclusion of interaction terms between the Gallup measure and the Party Difference Score. This finding, like the NORC results, may be attributable to sample size.

proportion of a national sample changes its party identification score over a span of a few months (Brody and Rothenberg 1988) or years (Brody 1977) or that the test-retest correlation differs from unity (Allsop and Weisberg 1988; Dreyer 1973). This observation, however, is misleading insofar as it confounds real change in partisanship with change due to response error. When party identification data are collected only minutes apart, thus ruling out the possibility of real change, the same pattern of apparent instability emerges. Our cross-sectional results in tandem with the test-retest results presented by Green and Palmquist (in press) suggest that mismeasurement alone is sufficient to account for much of what passes for real change in partisanship.

One implication of our analysis is that measurement error must be made an explicit part of models involving partisanship. To do this effectively requires some redundancy in the way partisanship is measured, whether in the form of multiple indicators at a given point in time or repeated measurement across successive waves of a panel study or, ideally, some combination of the two. The present study provides some insights into how alternative measures of party identification might be drafted. While several of the items considered here seemed to meet the criteria of face validity laid out by Converse and Pierce (1985) and Campbell et al. (1960), only the 7-point self-placement measure proves to be relatively immune to “don't know” responses, easily administered in face-to-face interviews, capable of producing a unimodal distribution of responses, and as reliable as the Michigan 7-point scale. Moreover, unlike the other alternative measures, the self-placement scale discerns gradations of partisanship that exist within the “strong” and “weak” categories of the Michigan partisanship classification. While one data set alone cannot firmly establish the attractiveness of a measure, our results suggest that this item deserves a closer look. Given the pivotal role partisanship plays in analyses using data from the American National Election Study and General Social Survey, these face-to-face surveys would do well to supplement and illuminate the traditional Michigan party identification scale.

Appendix

Question Wording

NORC SURVEY

Seven-point Michigan scale. “Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or what?” If Republican or Democrat: “Would you call yourself a strong [PARTY CHOSEN] or

a not very strong [PARTY CHOSEN]?" If Independent: "Do you think of yourself as closer to the Republican or to the Democratic party?"

Seven-point self-placement scale. "On this card is a scale with strong Democrats on one end and strong Republicans on the other, and with independents in the middle. Where would you place yourself on this scale?" The 7-point response scale was marked with the following labels at each scale point: (1) I strongly prefer the Democrats, (2) I prefer the Democrats, (3) I am basically independent but lean toward the Democrats, (4) I am independent, I have no preference for either party, (5) I am basically independent but lean toward the Republicans, (6) I prefer the Republicans, and (7) I strongly prefer the Republicans.

Forced-pair comparison item. "Generally speaking, do you usually think of yourself as a Democrat or a Republican?" Independent was a possible option if volunteered by the respondent.

Basic self-regard item. "No matter how you voted in the last couple of national elections or how you think you might vote in next November's national election—do you basically think of yourself as a Republican, a Democrat, an Independent, or what?"

Three-point Gallup item. "In politics, *as of today* do you consider yourself a Republican, Democrat, or Independent?"

Although our cross-tabulations include all respondents, the correlations presented exclude respondents who answer "other" or "don't know." In addition, those respondents who could not decide if they were strong or not very strong partisans in the Michigan question were excluded. The Michigan question and the self-placement were scored as 7-point scales, ranging from strong Democrats to strong Republicans, with independents who do not lean toward either party in the middle. The other questions were scored as 5-point scales with Democrats scored as 1, independents scored as 3, and Republicans scored as 5.

TIMES-MIRROR SURVEY

Seven-point Gallup scale. "In politics, *as of today*, do you consider yourself a *Republican*, a *Democrat*, an *Independent*, or what?" If Democrat or Republican: "Would you call yourself a strong [PARTY CHOSEN] or a not very strong [PARTY CHOSEN]?" If "independent," "no preference," "other," or "don't know": "Would you say you lean more to the Republican party or more to the Democratic party?"

Parry difference score. "Now I'd like to ask you a question about how you regard yourself. On a scale from 1 to 10, where '10' represents a description that is perfect for you, and '1' a description that is totally wrong for you, how well do each of the following describe you? ... A Republican. ... A Democrat." Each respondent's party difference score is computed by subtracting the score given to the Democrats from the score given to the Republicans. Respondents who declined to provide a score for either party are excluded from the statistical analysis.

References

- Abramson, Paul R., and Charles W. Ostrom, Jr. 1991. "Macropartisanship: An Empirical Reassessment." *American Political Science Review* 85:181–92.
- _____. 1992. "Question Wording and Macropartisanship: Response." *American Political Science Review* 86:481–86.
- _____. 1993. "Question Wording and Partisanship: Change and Continuity in Party Loyalties during the 1992 Election Campaign." Paper presented at the annual meeting of the American Political Science Association, Washington, DC, September 2–5.
- Achen, Christopher H. 1975. "Mass Political Attitudes and the Survey Response." *American Political Science Review* 69:1218–31.
- _____. 1985. "Proxy Variables and Incorrect Signs on Regression Coefficients." *Political Methodology* 11:299–316.
- Allsop, Dee, and Herbert F. Weisberg. 1988. "Measuring Change in Party Identification in an Election Campaign." *American Journal of Political Science* 32:996–1017.
- Alwin, Duane F., and David J. Jackson. 1981. "Applications of Simultaneous Factor Analysis to Issues of Factorial Invariance." In *Factor Analysis and Measurement in Sociological Research: A Multi-dimensional Perspective*, ed. David J. Jackson and Edgar F. Borgatta, pp. 249–79. Beverly Hills, CA: Sage.
- Asher, Herbert B. 1974. "Some Consequences of Measurement Error in Survey Data." *American Journal of Political Science* 18:469–85.
- Barnes, Samuel H., M. Kent Jennings, Ronald Inglehart, and Barbara Farah. 1988. "Party Identification and Party Closeness in Comparative Perspective." *Political Behavior* 10:215–231.
- Bishop, George F., Alfred J. Tuchfarber, and Andrew E. Smith. 1992. "Question Wording and Partisanship: An Experimental Test of the Hypothesis." Paper presented at the annual meeting of the American Political Science Association, Chicago, September 3–6.
- Bollen, Kenneth A. 1989. *Structural Equations with Latent Variables*. New York: Wiley & Sons.
- Brody, Richard A. 1977. "Stability and Change in Party Identification: Presidential to Off-years." Paper presented at the annual meeting of the American Political Science Association, Washington, DC, September 1–4.
- Brody, Richard A., and Lawrence S. Rothenberg. 1988. "The Instability of Partisanship: An Analysis of the 1980 Presidential Election." *British Journal of Political Science* 18:445–65.
- Budge, Ian, Ivor Crewe, and Dennis Farlie, eds. 1976. *Party Identification and Beyond: Representations of Voting and Party Competition*. London and New York: Wiley.
- Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. New York: Wiley.
- Campbell, James E., Mary Munro, John R. Alford, and Bruce A. Campbell. 1986. "Partisanship and Voting." In *Research in Micropolitics*, ed. Samuel Long, 1:99–126. Greenwich, CT: JAI Press.
- Converse, Philip E., and Roy Pierce. 1985. "Measuring Partisanship." *Political Methodology* 11:143–66.
- Dennis, Jack. 1988. "Political Independence in America. I. On Being an Independent Partisan Supporter." *British Journal of Political Science* 18:77–109.
- Dreyer, Edward C. 1973. "Change and Stability in Party Identifications." *Journal of Politics* 35:712–22.
- Fiorina, Morris P. 1981. *Retrospective Voting in American National Elections*. New Haven, CT: Yale University Press.
- Franklin, Charles H. 1992. "Measurement and the Dynamics of Party Identification." *Political Behavior* 14:297–309.

- Franklin, Charles H., and John E. Jackson. 1983. "The Dynamics of Party Identification." *American Political Science Review* 77:957–73.
- Green, Donald Philip. 1990. "The Effects of Measurement Error on Two-Stage Least-Squares Estimates." *Political Analysis* 2:57–74.
- Green, Donald Philip, Susan Lee Goldman, and Peter Salovey. 1993. "Measurement Error Masks Bipolarity of Affect Ratings." *Journal of Personality and Social Psychology* 64:1029–41.
- Green, Donald Philips, and Bradley Palmquist. 1990. "Of Artifacts and Partisan Instability." *American Journal of Political Science* 34:872–902.
- _____. In press. "How Stable Is Party Identification?" *Political Behavior*.
- Hayduk, Leslie A. 1987. *Structural Equation Modeling with LISREL*. Baltimore and London: John Hopkins University Press.
- Heise, David R. 1969. "Separating Reliability and Stability in Test-Retest Correlation." *American Sociological Review* 34:93–101.
- Jackson, John E. 1975. "Issues, Party Choices, and Presidential Votes." *American Journal of Political Science* 19:161–85.
- Johnston, J. 1984. *Econometric Methods*. 3d ed. New York: McGraw-Hill.
- Jöreskog, Karl G., and Dag Sörbom. 1988. *LISREL VII*. Mooresville, IN: Scientific Software.
- Keith, Bruce E., David B. Magleby, Candice J. Nelson, Elizabeth Orr, Mark C. Westley, and Raymond E. Wolfinger. 1992. *The Myth of the Independent Voter*. Berkeley and Los Angeles: University of California Press.
- Kenney, Patrick, and Tom Rice. 1988. "The Evaporating Independents." *Public Opinion Quarterly* 52:231–93.
- Krosnick, Jon A. 1991. "The Stability of Political Preferences: Comparisons of Symbolic and Nonsymbolic Attitudes." *American Journal of Political Science* 35:547–76.
- Krosnick, Jon A., and Matthew K. Berent. 1993. "Comparisons of Party Identification and Policy Preferences: The Impact of Survey Question Format." *American Journal of Political Science* 37:941–64.
- Lomax, Richard G. 1983. "A Guide to Multiple-Sample Structural Equation Modeling." *Behavior Research Methods and Instrumentation* 15:580–84.
- Long, J. Scott. 1983. *Confirmatory Factor Analysis*. Beverly Hills, CA: Sage.
- MacKuen, Michael B., Robert S. Erickson, and James A. Stimson. 1989. "Macropartisanship." *American Political Science Review* 83:1125–42.
- _____. 1992. "Question Wording and Macropartisanship." *American Political Science Review* 86:475–81.
- Markus, Gregory B., and Philip E. Converse. 1979. "A Dynamic Simultaneous Equation Model of Electoral Choice." *American Political Science Review* 73:1055–70.
- Miller, Warren E. 1991. "Party Identification, Realignment, and Party Voting: Back to the Basics." *American Political Science Review* 85:557–68.
- Nie, Norman H., Sidney Verba, and John R. Petrocik. 1979. *The Changing American Voter*, expanded ed. Cambridge, MA: Harvard University Press.
- Niemi, Richard G., David R. Reed, and Herbert F. Weisberg. 1991. "Partisan Commitment: A Research Note." *Political Behavior* 13:213–21.
- Page, Benjamin L., and Calvin C. Jones. 1979. "Reciprocal Effects of Policy Preferences, Party Loyalties, and the Vote." *American Political Science Review* 73:1071–89.
- Palmquist, Bradley, and Donald P. Green. 1992. "Estimation of Models with Correlated Measurement Errors from Panel Data." *Sociological Methodology 1992*, ed. Peter V. Marsden, pp. 119–46. Washington, DC: American Sociological Association.
- Petrocik, John R. 1974. "An Analysis of Intransitivities in the Index of Party Identification." *Political Methodology* 1(3):31–47.
- Schickler, Eric, and Donald Philip Green. 1993. "Issues and the Dynamics of Party Identification: A Methodological Critique." Revised version of paper presented at

- the annual meeting of the Midwest Political Science Association, Chicago, April 16–19.
- Shanks, J. Merrill, and Warren E. Miller. 1990. "Policy Direction and Performance Evaluation: Complementary Explanations of the Reagan Elections." *British Journal of Political Science* 20:143–235.
- Sullivan, John L., and Stanley Feldman. 1979. *Multiple Indicators: An Introduction*. Beverly Hills, CA: Sage.
- Wiley, David E., and James A. Wiley. 1970. "The Estimation of Measurement Error in Panel Data." *American Sociological Review* 35:112–17.
- Wiley, James A., and Mary Glenn Wiley. 1974. "A Note on Correlated Errors in Repeated Measurements." *Sociological Methods and Research* 3:172–88.