A RETROSPECTIVE ON RETROSPECTIVE VOTING*

D. Roderick Kiewiet and Douglas Rivers

*Though we alone bear responsibility for the contents of this paper, we have benefitted from conversations over the past few years on this topic with Morris Fiorina, Douglas Hibbs, Donald Kinder, and Gerald Kramer. The National Science Foundation provided research support under grant SES-8309994.
ABSTRACT

This paper critically reviews the extensive literature on retrospective voting in response to economic conditions. Each of the major types of analyses which have been performed—time series analyses of national vote totals, presidential popularity, and cross-sectional analyses of individual survey responses—have raised several interesting and important questions. The answers which have been obtained, however, are only partial and limited, as each of these approaches entails serious problems of estimation and interpretation. Further progress in this area, we argue, requires explicit treatment of conceptual and statistical issues which have hindered previous research: the dynamic formulation of expectations and preferences, the incidence of policy (and non-policy) effects across the population, notions of incumbering and political responsibility.

A RETROSPECTIVE ON RETROSPECTIVE VOTING

D. Roderick Kiewiet and Douglas Rivers

And the whole congregation of the children of Israel murmured against Moses and Aaron in the wilderness: And the children of Israel said unto them, Would to God we had died by the hand of the Lord in the land of Egypt, when we sat by the flesh pots, and when we did eat bread to the full; for ye have brought us forth into this wilderness, to kill this whole assembly of people with hunger (Exodus 16:2-3).

1. INTRODUCTION

The idea that a leader's political fate turns upon the material well-being of the people he leads is, as this passage from the Old Testament indicates, an old one. To be sure, few economies have fluctuated as extremely as that of the Israelites in the wilderness, and few political figures have experienced as many large swings in their popularity as Moses. More recently, though, historians have traced the origin of the French Revolution to a series of bad harvests followed by increased prices, falling wages, and mass unemployment (Doyle, 1980). In Britain, following the expansion of the franchise in the early part of the nineteenth century, electoral fluctuations were regularly attributed to grain prices and other economic factors (Olney, 1973; Nossiter, 1974).

Political observers in the United States also have long believed that prevailing economic conditions exert a strong influence upon the choices voters make in national elections. Research in this
area makes up one of the oldest quantitative traditions in political science (for reviews of this early literature see Kramer, 1971, and Monroe, 1979). These studies examined several different conjectures: prosperity aids the Republicans; farm sector depressions lead to support for the populists; or conservative candidates fare better during good times. In short, the vague notion that economic conditions influence voting admits to a wide range of specific hypotheses.

The focus of research in this area narrowed considerably, however, following publication of Gerald Kramer’s seminal article, “Short-Term Fluctuations in U.S. Voting Behavior, 1896-1964,” in 1971. Following Kramer’s lead, most subsequent studies have concentrated on a few interrelated hypotheses: that voting in response to economic concerns is (1) retrospective, (2) incumbency-oriented, and (3) based upon the results of economic policies, and not upon the actual policies themselves. Taken together, these imply that voters give greater support to candidates of the incumbent party when the election is preceded by a period of prosperity than when times have been poor. We will henceforth refer to these interlocking hypotheses as the retrospective voting model.

In some respects the subsequent concentration of work on this one model is unfortunate, given that there are other plausible ways different types of economic concerns might influence voters’ decisions. On the whole, however, the narrowing of attention to a few potentially testable propositions has been beneficial. Retrospective voting is probably the simplest and most straightforward model of those which have been posited (perhaps also, as suggested above, the oldest as well), and it makes good sense to examine simple hypotheses before entertaining more complicated theories about political-economic relationships. Combined with the fact that most studies in this area have utilized much the same data, this concentration on retrospective voting has made it much more feasible than in other areas of political research to compare the results of different studies, and thus to identify the particular features of models and methods which lead to differences in findings.

In the next section of this paper we will review the many time series analyses which followed the publication of Kramer’s original article. We will then turn to an examination of the other major body of research on retrospective voting—analyses of individual level survey data. In both cases we will identify the more robust findings which have emerged from the various strands of research, those findings which have not been strongly substantiated, and those subjects about which we continue to know virtually nothing. (Although much of this research has concerned European and other western style democracies, our discussion will be confined to studies of American national elections.) In the fourth and final section, we consider pooling cross-section and time series data as a possible solution to the data problems encountered in previous studies.
2. TIME SERIES ANALYSES OF RETROSPECTIVE VOTING

As indicated earlier, the retrospective voting model which has informed most of the work in this field can actually be decomposed into three separate hypotheses: voting on the basis of economic concerns is in response to actual policy outcomes, retrospective, and incumbency-oriented. Thus it is useful to analyze these studies in terms of the evidence they bring to bear on the following three sets of questions:

1. What exactly are the economic outcomes to which voters respond?

2. If voters judge retrospectively, what is the nature of the dynamic relation between policy outcomes and electoral responses?

3. What is the proper incumbency concept? Is it the incumbent president, congressional candidates of his party, or some other concept of incumbency?

Retrospective voting can occur on noneconomic issues, but most of the literature focuses on economic performance. Employment, inflation, and real income are neither perfectly well-defined nor measured without error, but their meaning and measurement are better understood and less controversial than that of any noneconomic issues. So far as possible, we will also limit our attention here to economic performance, though this can cause difficulties since wars—even unpopular ones—stimulate the economy without benefiting the incumbent administration.

Kramer found that the share of the total national vote for congressional candidates of the incumbent president's party was influenced more strongly and consistently by changes in real per capita income than by several other economic variables. Aggregate time series analyses of retrospective voting which followed both corroborated and extended Kramer's findings. First, electoral support for the incumbent party is best predicted by one of a set of highly correlated measures of change in real output. Although most studies followed Kramer's lead and specified change in real per capita income (Bloom and Price, 1975; Tufte, 1975, 1978; Hibbing and Alford, 1981), unemployment or change in per capita real GNP appear to work nearly as well (Fair, 1978). Evidence from this class of studies on how voters react to price inflation is mixed, and appears to be sensitive to small changes in data and in specification.

Another study which examined pooled cross-section time series data—state level vote totals for president from approximately a half dozen presidential elections—reported similar findings. Although Meltzer and Wahl (1975) were not impressed by the strength of their findings, their analysis indicated that national unemployment and price inflation often influenced presidential vote totals. The economic data analyzed in all these studies, of course, are national level figures. In sharp contrast to the pattern of evidence which has emerged here, a study which estimated the effect of price and income changes at the level of the individual congressional district found
these variables to have virtually no impact upon district vote totals (Owens and Olsen, 1980).

Secondly, voters appear to base their decisions primarily upon the economic conditions of the recent past. Again, most studies simply followed Kramer's lead and used the year of the election as the time frame for their economic measures (this is obviously a bit crude; because elections are held in early November, approximately one sixth of the data summarized by such measures comes from after the election). Fair's (1978) study, however, experimented with several different lags. Although there were not enough observations in his time series to permit any firm conclusions, his evidence suggested that votes for president were best predicted by per capita change in GNP in the second quarter of the election year. It is also clear that averaging the performance indicators over a longer period improves the prospects for maintaining the null hypothesis. Stigler (1973), for instance, observed much weaker relationships between per capita income change and congressional vote totals when he adopted a two year retrospective time frame.

One problem with the aggregate presidential or congressional voting data is that one either has a very short time series (in the case of Tufte's (1975) analysis of post-1946 midterm elections, only eight observations!) or the data extend over a period so long that the assumed stability of the regression function becomes questionable. Over the past century there has been vast change in the structure of the economy, in the size and scope of the federal government, in information and communications technology, as well as in the nature of party competition that certainly must have affected the way voters hold political leaders accountable for economic performance. Even though we gain an additional electoral observation every two or four years, we also gain a lot of history along with it.

An alternative approach that allows estimation of more complex models is to use the monthly or quarterly Gallup presidential popularity series. In keeping with the aggregate voting studies, these studies have found that approval ratings of the incumbent president are adversely affected by declines in real output (Kernell, 1978; Rivers, 1980; Golden and Poterba, 1980; Hibbs and Vasilatos, 1982) and by price inflation (Monroe, 1978; Norpoth, 1984).

At the risk of over-simplification, the typical presidential popularity study models voters' "memories" as a distributed lag of past performance indicators. Specifications of the lag function vary from study to study. Hibbs et al. (1982) employ a fairly complicated geometric structure that involves inter-administration and inter-party comparisons. Golden and Poterba (1980) and Monroe (1978) use a polynomial distributed lag, while Norpoth (1984) tries short, unconstrained forms. Although each of these studies has found lagged effects, none, unfortunately, has been particularly informative about the nature of the dynamic relation between government performance and political evaluations. Part of the problem is in the data: fitting two or three different distributed lags to 80 or quarterly observations is guaranteed to produce unsatisfactory estimates. The
solution, of course, is to impose some *a priori* constraints on the retrospective model. Smoothness priors, of which geometric and polynomial lags are very strong forms, are one possibility, but we should insist that these have a sound theoretical basis (see Nerlove, 1977, for a good discussion of theoretically derived lag structures). Symptomatic of the theoretical sloppiness that characterizes much of this field is the fact that many analyses ignore changes in administrations altogether, with the result that poor performance by the previous administration hurts the popularity of its successor.

It appears that we are still a long way from answering important practical questions, such as how close to an election must an economic recovery occur to help the incumbent party. Certainly the lags which have been found are not long enough to make political business cycle manipulations entirely pointless (a rectangular lag over the entire term of office would of course eliminate any advantage to timing recoveries to coincide with elections). On the other hand, most findings in this area indicate that an incumbent would be well advised to have a recovery started before the election quarter. Whether the optimal timing is two quarters in advance, three quarters, or even earlier remains an open question.

Consideration of the third major aspect of the retrospective voting model—defining who the incumbents are—may seem a bit pedantic; the studies discussed above have concerned either the incumbent president or congressional candidates of his party, and this would seem like the obvious way to proceed. There are, however, several variations on the incumbency theme which are worth noting. First, Tufte (1975, 1978) found that change in per capita income consistently influenced the electoral fortunes of incumbent party congressional candidates in midterm elections. His findings thus refute the notion that midterm elections are no more than a regression toward the mean, or a return to the "normal vote" following the "surge" toward the party of the winning candidate in the preceding presidential election (Campbell, 1966).

Secondly, Hibbing and Alford (1981) argue that all congressional incumbents are not created equal; their analysis of aggregate time series data suggests that congressional candidates of the incumbent president’s party who are also incumbent congressmen (incumbent incumbents, if you will) are more strongly affected by economic fluctuations than are candidates of his party running in open seats. Because incumbent incumbents have a very high probability of winning reelection, this implies, as Hibbing and Alford note, that "the political effects of economic fluctuations, so clear when mean vote levels are examined, may lose a good deal of their force when the focus is shifted to actual seats won and lost" (p. 438). Although the robustness of their findings has been questioned (Fiorina, 1983), this remains an important point; for it is the subsequent composition of Congress, not national vote percentages, which is presumably of more interest as far as public policy is concerned.¹

What we have learned from time series analyses of retrospective voting is that the electoral fortunes of the incumbent
president and congressional candidates of his party, as well as the president's approval ratings, are influenced by fluctuations in employment, prices, and real output. To a skeptic, it might seem that this research has revealed nothing beyond what common sense should have told us in the first place. Perhaps, but propositions derived from common sense and political folk wisdom are often incorrect or contradictory.

Beyond documenting common sense, however, we now have a rough sense, at least, of the magnitude of effects of real income fluctuations on voting. Kramer (1971) initially estimated approximately that a 0.5% (±0.2%)^2 decline in the incumbent party's congressional vote share would follow a one percent loss of real income. Bloom and Price (1975) found a slightly larger effect for income declines (approximately 0.7% ±0.2%), but no evidence of any effect for income gains. Tufte (1975), controlling for presidential popularity which also depends on real income changes, estimated a 0.35% (±0.1%) vote loss for every one percent of real income lost. Somewhat larger effects are found on presidential votes. Fair (1978) estimated roughly point for point (1.2% ±0.4%) changes in presidential vote shares and percentage real income growth rates. With stable prices, Frey and Schneider find that a one percent increase in real income raises presidential popularity about 0.85% (±0.3%) while Hibbs et al. (1982), controlling for unemployment, find that a one percent loss in real income eventually costs the incumbent about 0.8% in popularity. Of course, Stigler (1973) and Arcelus and Meltzer (1975) have demonstrated that enough tampering with the specification (period dummies, time trends, adding and deleting variables without rhyme or reason) and odd choice of variables can destroy these findings. In our view, however, a judicious summary of the empirical literature might run: A one percent decline in real income will cost the incumbent party between one-half and one percent of its vote share in the last election.

We lack, at present, any clear sense of how particular aspects of economic performance—unemployment and inflation, for example—contribute to electoral outcomes. The aggregate studies offer conflicting evidence on their effects. Some find unemployment, others inflation, some both, others neither, as significant determinants of election outcomes and popular support for political leaders. In part, this reflects heterogeneity in the population, as Hibbs (1977, 1979) has stressed, over the relative importance of full employment and price stability. Moreover, public attitudes on unemployment and inflation appear to shift over time, so there is little hope of resolving these issues with the aggregate time series data.

The time series models have not been perfected to the point necessary to produce useful and reliable forecasts. High R^2s—in the .7 or .8 range—tend to impress those nurtured on survey research, but they often provide vote forecasts with standard errors of 5% or more which do not make them very helpful in predicting election outcomes. This problem is confounded by the fact that time series analyses of retrospective voting have estimated the impact of economic conditions
upon changes in vote totals and not seat totals. (For a discussion of the shortcomings of forecasting on the basis of these sorts of models, especially those which involve calculation of swing ratios, see Rivers, 1984). One task for future work in this area is to rigorously model those factors (e.g., the nature of party competition within individual congressional districts) which determine how performance at the national level is translated into seat changes at the district level. The subsequent composition of Congress, after all, is of more direct consequence for public policy than the national vote total.

3. SURVEY ANALYSES OF RETROSPECTIVE VOTING

Time series analyses of aggregate voting and popularity data constitute only one part of the large body of research instigated by Kramer’s 1971 paper. Dozens of other researchers turned instead to survey data for evidence on how economic concerns influence the choices made by individual voters in national elections with the intention of explaining the aggregate level time series findings. Implicit in most of these analyses was the assumption that survey data could resolve questions that the aggregate data could not, that the effects found in survey data were somehow more real than those in aggregate data since they were not subject to the “ecological fallacy,” and that, equipped with the right survey questions, we could uncover the psychological motivations underlying electoral behavior.

In fact, for the most part survey analyses have provided very limited and partial answers to the questions posed. Interesting hypotheses have been proposed, but theoretical progress has been slow.

Since 1956, the American National Election studies have included an item which has frequently been used as a micro-level measure of real income change:

We are interested in how people are getting along financially these days. Would you say that you (and your family) are better off or worse off financially than you were a year ago, or about the same?

Most survey-based investigations have proceeded on the basis that evidence of retrospective voting at the individual level would consist of voters in the “better off” category exhibiting greater support for incumbents than those in the “worse off” category and with those in the “same” category falling somewhere in between. Aggregate changes in real income are the sum of individual household changes which should be reflected by responses to this item. Incumbents’ loss of electoral support during recessions would result from the fact that larger numbers of people were suffering financial hardships (and ending up in the “worse off” category).

What we have here is, in its simplest form, a linear aggregation problem. Suppose \( y_{it} \) is a measure of the \( i \)th voter’s support for the incumbent administration at time \( t \) (which we will refer to, for convenience, as his or her vote, though the measure will be treated as continuous rather than discrete) and let \( x_{it} \) denote the change in the voter’s income (usually real disposable income) in the period prior to the election. The assumed micro model for retrospective voting is:
\[ v_{it} = \beta x_{it} + u_{it} \]  
(1)

where \( u_{it} \) captures the effects of all other variables on voting. Let \( v_t, x_t, \) and \( u_t \) denote the averages of \( v_{it}, x_{it}, \) and \( u_{it} \) over all voters in election \( t \). Then it follows directly from (1) that:

\[ v_t = \beta x_t + u_t \]  
(2)

If \( u_{it} \) and \( x_{it} \) are uncorrelated during election \( t \), then the cross-sectional estimate of \( \beta \) obtained from (1) will be consistent. If \( u_t \) and \( x_t \) are uncorrelated over a certain time period, then the time series estimate of \( \beta \) from (2) will also be consistent. That is, under these conditions, the cross-sectional (micro) estimates and the time series (macro) estimates should not conflict. There are, however, several reasons why the micro and macro estimates might differ. The most obvious reason, which we will return to later, is that either (1) or (2) is misspecified in the sense that \( x \) and \( u \) are correlated. We temporarily ignore this possibility and proceed on the assumption that (1) is the correct micro specification.

The survey evidence on the effect of personal financial conditions on voting is mixed. Votes for president do appear to reflect people's recent financial fortunes. With only a few exceptions, those who reported being better off financially were found to be more likely to vote for the incumbent party in presidential elections than those who reported being worse off or the same (Wides, 1976; Fiorina, 1978, 1981; Klorman, 1978; Tufte, 1978; Kinder and Kiewiet, 1979, 1981; Kiewiet, 1983). Similarly, the more favorable recent trends in family finances have been, the higher the marks individuals give to the president for his handling of the economy (Fiorina, 1981) and for the overall job he is doing as president (Wides, 1976; Kinder, 1981). Much weaker effects or even no effects at all were found in election for other offices. Although a few studies found evidence for personal financial effects in the 1950s, few congressional elections after 1960 registered this effect (Ben-Gera, 1977; Fiorina, 1978; Kinder and Kiewiet, 1979, 1981; Weatherford, 1979). Nor was the evidence more supportive in either senatorial or gubernatorial elections (Ben-Gera, 1977; Klorman, 1978).

Several problems of interpretation plague most cross-sectional analyses. First, a tendency to "control" for other attitudinal variables which may also depend on economic conditions leads to an underestimate of economic effects. Second, beyond reporting that the coefficient of the family finances item was statistically significant or not, most studies have paid little attention to the meaning of the estimated coefficients. Estimates of, for example, the effect of a percent change in real income have not been computed, so although the cross-sectional estimates have generally been viewed as inconsistent with the aggregate time series estimates, no serious systematic attempt has been made to compare the two.

The use of the survey personal finances item, of course, does not permit direct comparisons of cross-sectional estimates with the time-series estimates discussed in the previous section. In fact, the meaning of the personal finances item is a matter of considerable
FIGURE 1
SURVEY REPORTS OF PERSONAL FINANCIAL SITUATION AND REAL INCOME GROWTH RATES
controversy. Rosenstone et al. (1983) report correlations ranging from 0.16 to 0.57 between it and other more specific personal financial items, so, despite the question's rather vague wording, it does appear to measure concrete aspects of the respondent's financial situation. It is also the case that high rates of real income growth are associated with larger fractions of the population reporting themselves to be "better off" financially as illustrated by the scatterplot in Figure 1. Regressing the percentages responding "better," "same," and "worse" on the quarterly percentage rate of growth of real personal disposable income per capita (annualized), we obtain:

\[
\text{BETTER} = 31.9 + 1.3 \text{ AINC} \\
\text{(0.4)} \\
R^2 = 0.51
\]

\[
\text{SAME} = 37.8 + 0.3 \text{ AINC} \\
\text{(0.9)} \\
R^2 = 0.03
\]

\[
\text{WORSE} = 30.2 - 1.8 \text{ AINC} \\
\text{(1.1)} \\
R^2 = 0.21
\]

\[T = 13 \text{ (1956–1980)} \text{ (Standard errors in parentheses)}\]

The responses to the personal finances item vary over a relatively narrow range for the most part and, when the two extreme observations (1964 and 1974) are deleted, only a tenuous correlation remains between it and the real income growth rate. This suggests that only when times are bad enough to become a topic of general conversation will declines in real income be reflected in the survey item and that it is a rather crude instrument for measuring smaller income fluctuations. Clearly the personal finances item measures real income change rather imprecisely which means that survey-based estimates of retrospective voting models are biased. In the simple case of equation (1), the direction of the bias is toward zero with the degree of attenuation proportional to the squared correlation between the personal finances item and the micro-level changes in real income.

Neglecting the unreliability of the personal financial situation item as a measure of real income change, Table 1 reports an estimate of the effect of a one percent change in real income based on the regression reported above and Kiewiet's (1983) presidential vote equations. The income effects range from a 0.08 loss in the incumbent party vote share to a 0.34 loss for every percentage point decline in real income. On average the estimated vote loss for a one percent real income decline is about 0.2%. Not only are the estimated effects small relative to those found in the aggregate time series studies, but the estimates are unstable. (Similar calculations could be made for congressional elections, but the estimates are frequently perversely signed so that the exercise would be pointless.)

Various explanations have been proposed to account for the discrepancies between the cross-sectional and time series estimates. One approach employed to find a micro-level foundation for the aggregate relationship between economic conditions and voting was to
<table>
<thead>
<tr>
<th>Year</th>
<th>Personal Finances</th>
<th>Real Personal Disposable Income Growth Rate</th>
<th>Incumbent Vote Loss for 1% Real Income Decline</th>
</tr>
</thead>
<tbody>
<tr>
<td>1956</td>
<td>39</td>
<td>18</td>
<td>3.0</td>
</tr>
<tr>
<td>1960</td>
<td>33</td>
<td>47</td>
<td>0.09</td>
</tr>
<tr>
<td>1964</td>
<td>46</td>
<td>40</td>
<td>0.08</td>
</tr>
<tr>
<td>1968</td>
<td>34</td>
<td>47</td>
<td>0.19</td>
</tr>
<tr>
<td>1972</td>
<td>34</td>
<td>42</td>
<td>0.25</td>
</tr>
<tr>
<td>1976</td>
<td>34</td>
<td>35</td>
<td>0.20</td>
</tr>
<tr>
<td>1980</td>
<td>32</td>
<td>25</td>
<td>0.18</td>
</tr>
</tbody>
</table>

a. See footnote 5 for explanation.
identify subsets of voters whose political behavior was more sensitive to their personal economic situation. Weatherford (1978) found that working class voters who had just experienced the hardship of a recession composed just such a group. In a study discussed earlier, Hibbing and Alford (1981) argue that incumbent congressmen of the president’s party are more easily credited or blamed by voters in their districts for fluctuations in their personal economic conditions. Finally, according to Feldman (1982), voters’ decisions in presidential elections reflect the state of their families’ finances only to the extent “there is a perception of social (economic, governmental) responsibility for financial well-being” (p. 449). Most of these studies are based on data from one or two elections, so there is some doubt about the generalizability of their findings (Fiorina, 1983). In any event, even if these findings are to be believed, in no way do they resolve the discrepancy between the cross-sectional and time series estimates. That they have identified groups of voters for whom personal economic considerations are relatively more important in deciding how to vote has to be balanced against the fact that each of these groups of voters constitute a distinct minority of the electorate.

Probably the most interesting feature of these studies, however, is that they all share the same key assumption that economic conditions influence voting decisions only to the extent that voters attribute responsibility for these conditions to incumbent politicians. In each of these studies the pattern of evidence which is reported is interpreted in terms of the differential attribution of responsibility. Thus, voters’ decisions in presidential elections are influenced much more strongly by trends in family finances because the president is assigned the lion’s share of credit or blame:

It is the president who is primarily responsible for the general thrust of macroeconomic policy, whether it be the “guns and butter” policies of Lyndon Johnson or the “Reaganomics” of the current administration. It is also the president who shoulders most of the credit or blame for the ultimate success or failure of the policies pursued by his administration. Consequently, any sort of economic problem which voters might be concerned about will exert a larger influence upon their choice between presidential candidates than upon that between the candidates for Congress (Kiewiet, 1983, p. 126).

Hibbing and Alford also point to differential attributions of responsibility to account for the greater influence of economic concerns in elections involving incumbent incumbents:

If anyone is held responsible for current financial conditions, it should be those who are in the best position to take credit or receive blame for these conditions—that is senior incumbents of the in-party incumbents, those gradations should be related in some positive fashion to tenure (Hibbing and Alford, 1981, p. 432).

Feldman (1982) invokes the same line of reasoning:

Personal economic conditions will influence voting behavior only when there is a perception of social (economic, governmental) responsibility for financial well-being (Feldman, 1982, p. 449).

In short, the assumption that economic conditions influence voting decisions only to the extent that voters attribute responsibility for these conditions to incumbent politicians is ubiquitous in this
literature. That this is so, however, raises the crucial question of exactly how such attributions should be modeled.

More precisely, do voters, having experienced a stream of economic outcomes, attempt to discriminate between that part of it which is properly attributable to the actions of incumbent policymakers and that part of it which is not? Or do they simply take this stream of outcomes at face value and evaluate the incumbents more or less favorably on the basis of it? Strict notions of voter rationality would certainly suggest that they would attempt to differentiate between income change which is "government-induced" and that which is not; after all, why should an individual choose between competing candidates on the basis of things that neither could possibly control? On the other hand, this task may well place large information costs and unreasonable demands on the inferential powers of the typical voter. The avoidance of these costs was one of the major features of Downs' (1957) model of retrospective voting.

Similarly, Butler and Stokes (1969) argue that "the technical difficulties of assigning responsibility for past government action or inaction" force voters to adopt a satisficing strategy:

Modern electorates tend to "solve" this problem of causal reasoning by assuming that certain causal relationships must exist rather than by discerning what they are. Electors focus their attention primarily on certain conditions which they value positively or negatively and simply assume that past or future governments affect them. The public can call for a government's dismissal in economic hard times just as it calls for a team manager's dismissal in a losing season, in each case concluding that causal relationships must exist without knowing in detail what they are (p. 25).

But are such distinctions really so difficult to make? There are many predictable changes in a person's financial situation over the life cycle which voters understand and anticipate. Some are favorable and improve one's financial state (e.g., finishing graduate school and taking a full-time teaching position) while others decrease one's income (e.g., having a child enter college or retiring). It seems far-fetched that voters would credit the incumbent party for changes of this kind. Or to give a more extreme example: suppose a distant relative dies, leaving a substantial inheritance. Does the lucky recipient attribute his or her good fortune to whoever happens to be in the White House at that moment?

Moreover, there is a good deal of evidence that voters rarely associate changes in their personal financial situation with government policy. When asked why their family financial situation had changed, Kinder and Weisner (1983) report that "virtually no one sees government policy contributing to their family's economic achievements or setbacks. In any single national survey, no more than 1% of those interviewed point directly to government." This is not to say that voters place no responsibility on government for their personal economic fortunes, but only that not all changes in a family finances have political effects. Schlozman and Verba's (1981) interviews with several hundred unemployed individuals yielded very similar findings, as did Brody and Sniderman's (1977a, 1977b) studies of "coping" and the ethic of self-reliance.

The question of whether or not voters hold government
responsible for all changes in their personal financial situation concerns the specification of equation (1). Kramer (1983) argues that voters respond not to changes in their real income, but instead to government-induced changes in real income. He partitions the change in a voter’s real income $x_{it}$ into a component $g_{it}$ which can be attributed to the actions of the incumbent government and an idiosyncratic component $e_{it}$:

$$x_{it} = g_{it} + e_{it}$$  \hspace{1cm} (3)

Kramer does not claim that voters are able to precisely distinguish which fraction of their income change is actually government-induced, but only that some substantial part of the change in a family’s financial situation is caused by forces clearly outside the government’s control.

Replacing $x_{it}$ by $g_{it}$ in (1) yields:

$$v_{it} = g_{it} + u_{it}$$  \hspace{1cm} (4)

If (4) is the correct specification for the vote equation (i.e., if $g_{it}$ and $u_{it}$ are uncorrelated), then the cross-sectional estimate of $\beta$ obtained by regressing vote on change in family income will (under the conditions of White, 1982) converges to:

$$\hat{\beta} = \beta \left[ \frac{\text{var}(g_{it}) + \text{cov}(g_{it}, e_{it})}{\text{var}(x_{it})} \right] + \frac{\text{cov}(u_{it}, e_{it})}{\text{var}(x_{it})}$$ \hspace{1cm} (5)

with probability one as sample size increases.

From (5) there are several obvious sources of bias in the cross-sectional estimates:

1. If the government is responsible for a relatively small fraction of the cross-sectional variation in family finances, then $\text{var}(g_{it}) < \text{var}(x_{it})$.

2. If government-induced income changes tend to compensate for nongovernment-induced income changes, then $\text{cov}(g_{it}, e_{it}) < 0$.

3. If $u_{it}$ includes party or other effects on voting which are correlated with nongovernment-induced income changes, then the sign of $\text{cov}(u_{it}, e_{it})$ will vary from sample to sample, depending on, among other things, which party is in power at the time of the sample.

The first two effects discussed above will bias the cross-sectional estimates downward (i.e., toward zero). The third effect can bias the cross-sectional estimates either upward or downward, but, in any event, will make the estimates vary from one cross-section to another. Kramer argues that these biases account for the findings of the cross-sectional studies reviewed above: unstable coefficient estimates which, on average, are somewhat smaller than those found in aggregate time-series studies.

Whether retrospective voting should be specified as a function of total income change ($x_{it}$) or government-induced income change ($g_{it}$) is seen to be a question of central importance. If the latter specification is the correct one, then absence of income effects found
in many cross-sectional studies is no more than a statistical artifact. It will be difficult to resolve this issue with current survey measures. Nor is posing a question to survey respondents such as "What percentage of the increase or decrease in your income over the past year do you attribute to the effects of government policy?" likely to be very informative. Such calculations, if they occur at all, are probably subconscious and the survey responses may be tainted by rationalization. The most promising strategy in our view would be to identify—as objectively as possible—different sources of income change and then to estimate the separate effect of each type of income change on voting.

In view of the poor results obtained with the personal finances item, attention has shifted to other forms of economic effects that might provide a link between micro and macro-level findings. The sociotropic voting hypothesis of Kinder and Kiewiet (1979, 1981) is one effort in this direction. They argued that it is difficult for individuals in this country to disentangle the effect of current governmental policies upon their own economic fortunes from a whole host of other factors which affect the demand for their labor and the value of their assets. The state of the nation's economy, however, reflects upon the performance of the incumbent party in a far more direct fashion:

Conditions in the national economy, of course, are probably not as personally salient to most individuals as their own financial situation. But national economic assessments are, by definition, of general, widespread phenomena. Consequently, in most people's minds national economic conditions reflect upon the performance and policies of the incumbent administration much more directly than the conditions of their own lives . . . what national economic assessments lack in personal relevance, they make up for by being of more obvious political relevance (Kiewiet, 1983, p. 130).

Their hypothesis is that perceptions of national economic conditions and events determine the degree to which voters support incumbent candidates. In other words, the aggregate level time series findings were generated not so much by voters responding to their own individual economic circumstances, but rather by their response to the aggregate level economic data themselves. Kinder and Kiewiet's analyses were based primarily upon responses to a question which closely resembled the family finances item:

Now turning to business conditions in the country as a whole, would you say that at the present time business conditions are better or worse than they were a year ago, or about the same?

Employing a fairly simple model, they found that the more favorably voters viewed recent trends in national business conditions, the higher the probability of their voting for incumbent candidates. As with the family finances measure, however, the influence of this variable upon voting decisions was considerably stronger in presidential elections than in congressional races.

It is also the case that the distribution of responses to the national business conditions item swing much more sharply between good years and bad years than the distribution of responses to the family finances item. Furthermore, the results of Kiewiet's (1983) simulations indicated that voting on the basis of recent trends in
national business conditions could account for roughly twice as much variation in the level of support received by incumbent presidential and congressional candidates than voting in response to shifting family financial fortunes. He calculated that the effects of both variables taken together accounted for about half of the aggregate change in votes implied by Kramer's and Fair's findings.

Analyses of sociotropic voting are also, as Kramer (1983) has pointed out, subject to serious problems of evidence and interpretation. If a strict interpretation of the sociotropic hypothesis is taken, then the relevant variable is either the average income change \( \Delta I_t \) or its government-induced counterpart (denoted \( G_t \) in Kramer's analysis). Neither of these variables varies in a cross-section, so it is hopeless to obtain any evidence about this form of sociotropic voting from a cross-sectional analysis. Any observed cross-sectional variation, according to Kramer, will reflect either measurement error in the survey instrument or differences in voter perceptions.

We suspect that cross-sectional variation in perceptions of national economic trends arises from many sources. Some of it will just be partisan rationalization, but some of it may reflect different sources of information available to voters. For example, in depressed areas voters may perceive national conditions to be worse than voters in booming areas.\(^5\) Other variation may depend on the specific form of voters' utility functions. Different individuals may focus on different economic indicators, weight the same indicators differently, or exhibit different rates of time preference. Having acknowledged that survey measures of national economic conditions may be more than an amalgamation of rationalization and measurement error, we should caution that we have very little understanding of precisely what they do measure (see Kiewiet, 1983, ch. 6).

The second major difficulty with analyzing sociotropic voting is one of interpretation. Kiewiet (1983, p. 131) points out that "voting in response to national economic assessments could reflect very different motivations. It could have its basis in a purely patriotic or altruistic concern for the interests of all Americans. Alternatively, it could be entirely motivated by a self-interested concern for one's economic well-being; in this case voters simply use information about national economic conditions as an indicator of how well the incumbent administration has promoted their own (and their fellow citizens') welfare." In fact, if voters wish to evaluate incumbent performance on the basis of government-induced changes in their own income, the national average real income change provides a reasonable basis to distinguish idiosyncratic income changes from those which are government induced. Voters' responses to the national economic conditions question may therefore be a better indicator of their estimate of government-induced income change in their personal income than the personal finances item.
4. POOLING CROSS-SECTIONS AND TIME SERIES

The data problems which trouble cross-sectional analyses of retrospective voting are not solely statistical in nature. Inasmuch as the performance indicators used to estimate retrospective voting models are subject to error, the nature of this error is very different in the cross-section than in the time series. More specifically, government policy will cause different kinds of income change in cross-sections and time series. Depending upon where one learned their macroeconomics, national output and employment levels are either viewed as quite sensitive to fiscal and monetary policies (Kramer, having spent some time at the Cowles Commission, attributes roughly half the variance in the rate of real income growth to government policy) or as completely irrelevant (as is believed in some sectors of Chicago and Minneapolis).

However, one stands on the new classical macroeconomics, all can agree that the government redistributes income among citizens in myriad ways: taxes, transfers, regulation, inflation, etc. In the cross-section, all of the government induced income variation is of the redistributive variety. As indicated earlier, Kramer argues that this variation makes up a relatively small fraction of the change in incomes for individuals in any period. Moreover, knowing that an individual is the beneficiary of a particular government redistributive policy will also indicate the likely partisanship of that individual. This relationship will not be stable over time or across party administrations (more Democrats receive welfare benefits than oil depreciation allowances, so changing one policy affects Democrats and Republicans differently) and, as a consequence, little systematic can be learned from the personal finances item in a cross-section. By aggregating, Kramer argues, it is possible to average out most idiosyncratic and redistributive income effects.

One hypothesis that has not been investigated here is that voting on aggregate performance and individual redistribution are of a different character. It seems quite plausible that voting based on aggregate performance might be incumbency-oriented; that is, voters could ignore the policies used to achieve macroeconomic goals, support any ones which seem to work, and oppose any which fail. After all, both parties claim to have the secret to macroeconomic success and if economists can't agree on who is right, why should voters try? On redistributional issues, however, the parties do take different positions and voters have little difficulty determining which policies benefit them. It would be unwise for the recipients of government benefits to indiscriminately credit the incumbent party—which may be trying to cut those benefits—for its supposed largesse. Aggregate performance levels would indicate the skill of the incumbent party, while redistributional effects would represent the policy position taken by the incumbent party. A purely self-interested voter might prefer a less-skilled party which was willing to redistribute income in his or her direction (see Rivers, 1983, for analysis of a model of this type).

Combining cross-section and time-series data permits, in
principle, investigation of these and other interesting hypotheses. Rosenstone (1983), for example, has improved on aggregate electoral forecasts by disaggregating to the state level in a pooled cross-sectional analysis. Markus (1984) has tried pooling seven of the Michigan surveys to test the so-called self-interest and sociotropic voting hypotheses and, before long, many others will probably turn to pooling cross-sections in the face of the Kramer critique. While panel data is potentially quite useful for some questions, it is important to understand what it possibly can or cannot tell us. It is shown in the appendix to this paper that the pooled cross-section regression estimate is a weighted average of the relevant cross-section estimates and the aggregate time-series estimate. Since, under Kramer’s assumptions, both the cross-sectional and time series estimates are biased toward the origin, the pooled cross-section estimate will also be biased since it lies in the convex hull of these estimates. In fact, if (as is the case) the typical cross-sectional variation in real income growth is larger than the average time-series variation, the pooled cross-section estimate will lie closer to the average cross-section estimate than to the time-series estimate. Hence, no progress has been made toward alleviating the errors in variables problem by pooling, nor as yet do we have a feasible way of attacking cross-sectional and time-series issues simultaneously. Pooling, as Hausman and Taylor (1981) have shown, can provide a natural solution to the errors-in-variables problem, but this requires an explicit treatment of measurement errors that has not yet been attempted in retrospective voting models.

5. WHAT IS TO BE DONE?

The cautionary tone of our survey should not be taken as unduly pessimistic. The retrospective voting literature has raised, we think, a large number of interesting questions about the relationship between elections and public policy. If the answers which have been obtained are partial and limited, there is nonetheless an identifiable body of findings contained in this literature. The growth of knowledge about retrospective voting has been facilitated by the existence of a simple, clearly specified model. Difficulties have arisen when some concepts have not been clearly defined or when the relationship between empirical measures (particularly survey measures) and theoretical concepts have been treated too casually.

Further progress is unlikely to come by continued mining of the same types of data in traditional ways. More promising, in our view, will be efforts to identify new sources and types of data that are potentially informative about outstanding theoretical issues. If a variable of interest, national economic conditions for example, does not vary in a particular sample (in a cross-section for example), one must resort to another source of data (such as panel data). The new source of data will not be automatically free of the problems of the old data (such as measurement error, as discussed in Section 4) and it may entail new problems (e.g., sample selection bias) that could be ignored before. While we do not believe that attempts to solve these
problems are guaranteed to be successful, the continued vitality of research in this area depends on the attempts being made.

The other major problem facing those interested in retrospective voting is one of interpretation. By now the proposition that voters will punish incumbents for poor performance should not be controversial. Attempts to elaborate upon this relationship may, however, raise interpretive issues that will require reformulation of the simple retrospective voting model or its extension to behaviors other than voting. For example, partisan voting was frequently viewed as a competitor of the retrospective voting hypothesis until Fiorina (1981) showed how partisanship naturally fitted within the retrospective voting framework. This makes the interpretation of party effects within these models as problematic, particularly in a field where it has been routine practice to "control" for party identification. Similarly, we have argued that the existence of aggregate economic effects on individual political behavior are consistent with either self-interested or sociotropic voting. In this case, it may be necessary to examine behavior on other issues where different decision processes can be more easily distinguished. Here, too, we believe the issues raised in the retrospective voting literature are important and worthy of further attention.

**APPENDIX**

Let \( i = 1, \ldots, N \) index the observations in samples \( t = 1, \ldots, T \) so that \((x_t, y_t)`composes the data from sample \( t \) with
\[
x_t = (x_{1t}, \ldots, x_{Nt})'\quad \text{and} \quad y_t = (y_{1t}, \ldots, y_{Nt})'.
\]
(The restriction that the samples be of the same size is unnecessary and can easily be eliminated.) Let \( x = (x_1', \ldots, x_T') \) and \( y = (y_1', \ldots, y_T') \) and define the orthogonal projection operators:
\[
P_N = \frac{1}{N}1_N 1_N' \\
Q_N = I_N - P_N
\]
where \( 1_N \) denotes an \( N \times 1 \) vector of ones. Note that \( P_N y_t = \bar{y}_t 1_N \) where
\[
\bar{y}_t = (1/N) \sum_{i=1}^{N} y_{it} \quad \text{and similarly for } x_t.
\]
With this notation established, consider the following three regressions:

1. **Cross-section.**
   \[
b_t = x_t' Q_N y_t / x_t' Q_N x_t \quad (t = 1, \ldots, T)
   \]

2. **Average Time Series.**
   \[
   \bar{b} = x' (I_T \otimes P_N) Q_{NT} (I_T \otimes P_N) y / x' (I_T \otimes P_N) Q_{NT} (I_T \otimes P_N) x
   \]

3. **Pooled Cross-section.**
   \[
b = x' Q_{NT} y / x' Q_{NT} x
   \]

That is, for each cross-section \( t = 1, \ldots, T \), \( y_t \) can be regressed on \( x_t \) yielding an estimate \( b_t \). Alternatively, \( \bar{y}_t \) can be regressed on \( \bar{x}_t \)
yielding the average time series estimate \( \bar{b} \). Instead of averaging, the cross-sections could be pooled and the NT observations used to produce the pooled cross-section estimate \( b \). (\( \otimes \) denotes the Kronecker product.) It is easily verified that:
\[
Q_{NT} = (I_T \otimes Q_N)' + (Q_T \otimes P_N)
\]
Using $Q_N P_N = 0$ gives:

\[(I_T \otimes P_N) Q_N (I_T \otimes P_N) = Q_T \otimes P_N\]

so that:

\[
b = x'(Q_T \otimes P_N) y / x'(Q_T \otimes P_N)x
\]

With some rearrangement and substitution we find:

\[
b = [\bar{b} x'(Q_T \otimes P_N)x + \sum_{t=1}^{T} b_t(x_t' Q_N x_t)] / [x'(Q_T \otimes P_N)x + \sum_{t=1}^{T} x_t' Q_N x_t]
\]

We have shown that the pooled cross-section estimator $\bar{b}$ is a weighted average of the average time series estimator $\bar{b}$ and the $T$ cross-section estimates $b_1, \ldots, b_T$:

\[
b = w_0 \bar{b} + \sum_{t=1}^{T} w_t b_t \quad (\sum_{t=0}^{T} w_t = 1, \ w_t \geq 0 \ for \ all \ t)
\]

where:

\[
w_0 = T s_0^2 / \left( \sum_{t=1}^{T} s_t^2 + Ts_0^2 \right)
\]

\[
w_t = s_t^2 / \left( \sum_{t=1}^{T} s_t^2 + Ts_0^2 \right) \quad (t = 1, \ldots, T)
\]

Where $s_t^2$ is the variance of $x$ in sample $t$ and $s_0^2$ is the variance of $\bar{x}$, i.e.,

\[
\bar{s}^2 = \frac{1}{T} \sum_{t=1}^{T} (x_t - \bar{x})^2 \quad \bar{x} = \frac{1}{T} \sum_{t=1}^{T} x_t.
\]

---

**NOTES**

1. Fiorina (1983) performs a test of equality for the coefficients of the "Better," "Same," and "Worse" dummies in Democratic, open, and Republican seats. Apart from problems with the survey data covered later, Fiorina's test lacks power since it involves open seats and voters whose financial position is unchanged which are peripheral to the Hibbing-Alford hypothesis. In fact, Fiorina's estimates provide some support for the different treatment of incumbents in midterm but not in presidential election years.

2. Approximately 95% confidence intervals listed in parentheses.

Most articles cited below report a variety of specifications with a corresponding range of estimates. The estimates reported here are selected to be representative of the results reported. Readers are encouraged to refer to the original papers rather than rely on our somewhat subjective summary.

3. The variable most highly correlated with the personal finances item is the reported direction of income change relative to the cost of living. For reasons not apparent to us, political scientists tend to prefer questions with better/worse response categories to, for example, reported dollar amounts of income change. Even if respondents do not know exactly how much their income has changed, it is difficult to see how forcing them into vague response categories will reduce measurement error.
4. Kiewiet (1983) reports probit estimates of the form:

\[ \text{Prob(Republican Vote)} = \Phi(\beta(\text{Better}) + \gamma(\text{Worse}) + \text{other terms}) \]

If we evaluate the above expression for a voter whose personal financial situation is unchanged (i.e., Better = Worse = 0) and set the "other terms" equal to zero, then we have (approximately):

\[ \text{Prob(Republican Vote)} = 0.5 + \beta(0.4)(\text{Better}) + \gamma(0.4)(\text{Worse}) \]

using an approximation suggested by Amemiya (1981). Using the regression estimates reported in the text, the change in the Republican vote share resulting from a one percent increase in the real income growth rate can be estimated by:

\[ \Delta \text{Prob(Republican Vote)} = \text{Prob(Republican Vote|Better)}\Delta \text{Prob(Better)} + \text{Prob(Republican Vote|Same)}\Delta \text{Prob(Same)} + \text{Prob(Republican Vote|Worse)}\Delta \text{Prob(Worse)} \]

\[ \approx (0.4)(0.013\beta - 0.018\gamma) \]

The approximate change in the incumbent party (rather than Republican) vote share is reported in the text.

5. This suggests using the state unemployment rate as another indicator for the business conditions item in an errors-in-variables model.

6. The distinction between incumbent-oriented and policy-oriented voting is discussed in Kiewiet (1981). Instead of blaming the incumbents for all forms of economic difficulties, policy-oriented voters support the party which places a higher priority on attacking the problems they find most worrisome. Assuming Democrats place higher priority on full employment and Republicans on price stability (Hibbs, 1977), policy-oriented voting in response to inflation would result in more support for the Republicans, while concern over unemployment would help Democratic candidates, regardless of who the incumbents happened to be.

7. For static models without unobservable individual effects, repeated independent cross-sections are equivalent to a panel.

8. Gerald Kramer has suggested that preferences over tax schedules would be a situation where sociotropic behavior can be easily distinguished from narrowly self-interested behavior.
REFERENCES


