Journal of Economic Perspectives—Volume 10, Number 3—Summer 1996—Pages 89–102

Econometrics and Presidential Elections

Ray C. Fair

At the beginning of the 1970s, Kramer (1971) wrote an influential paper on voting behavior, which concluded that votes depend on economic events in the year of the election. My interest in this topic was piqued when Orley Ashenfelter in June 1971 used a Kramer-type equation and my then-current prediction of the 1972 growth rate of real output to predict that Nixon would win the U.S. presidential election with a little over 60 percent of the vote. Nixon actually received 61.8 percent.[1] I began work on a model of how economic events affect voting behavior that I argued encompassed the theories of Kramer (1971), Stigler (1973), who believed that well-informed voters would look back more than a year, and the earlier theory of Downs (1957). This work was eventually published in Fair (1978).

The general theory behind the model is that a voter evaluates the past economic performances of the competing parties and votes for the party that provides the highest expected future utility. Within the context of the model, one can test both how far back voters look in evaluating the economic performances of the parties and what economic variables they use in their evaluations. Many tests were performed in Fair (1978) using data on U.S. presidential elections. The results supported the view that voters look only at the economic performance of the current party in power, not also, for example, the performance of the opposition party the last time it was in power. Furthermore, the most important economic variable

[1] Alas, four years later Ashenfelter's equations predicted that Ford would be elected in 1976, which was not to be. This is the first of a number of warnings in this paper not to become too confident with any of the equations.

■ *Ray C. Fair is Professor of Economics, Yale University, New Haven, Connecticut.*

was the growth rate of real per capita output in the year of the election, suggesting that voters look back only about a year.[2]

The equations in Fair (1978) were estimated through the 1976 election, and I have updated the equation after each election since (Fair, 1982, 1988, 1990, 1996). This paper reviews the voting equation, with particular emphasis on the update made after the 1992 election. For a complete discussion of everything that was tried for the 1992 update and the reasons for the final choices, see Fair (1996). Forecasts of the 1996 election are also made, conditional on a forecast of the economy. All the data used for the estimation and forecasts are presented in Table A in the Appendix. The focus here is on the empirical specifications; the reader is referred to Fair (1978) for the details of the theory.

The main interest in this work from a social science perspective is how economic events affect the behavior of voters. But this work is also of interest from the perspective of learning (and teaching) econometrics. The subject matter is interesting; the voting equation is easy to understand; all the data can be put into a small table; and the econometrics offers many potential practical problems. In fact, Ashenfelter developed an interest in this area in part because he needed a problem assignment that had both economics and politics in it for an econometrics class he was teaching at Princeton University. Thus, this paper is aimed in part at students taking econometrics, with the hope that it may serve as an interesting example of how econometrics can be used (or misused?). Finally, this work is of interest to the news media, which every fourth year becomes fixated on the presidential election. Although I spend about one week every four years updating the voting equation, some in the media erroneously think that I am a political pundit—or at least they have a misleading view of how I spend most of my days.

A Review of the Voting Equation

The task of the equation is to explain the Democratic party's share of the two-party vote. The sample period begins in 1916.[3] Two types of explanatory variables are used: incumbency variables and economic variables. Until the 1992 changes, the basic equation was as follows:

[2] Although Kramer (1971) used both data on congressional and presidential elections, he did not find that the presidential vote was very responsive to economic conditions. This may be because he constrained the coefficient estimates in the equation explaining the presidential vote to be the same as the coefficient estimates in the equation explaining the congressional vote. The results in Fair (1978) are only for presidential elections, and one of the maintained assumptions is that voters hold the party in the White House responsible for the state of the economy.

[3] I have collected the data back to 1880, and some experimentation was done using observations prior to 1916. As was the case for the original work in Fair (1978), however, the results using the elections before 1916 were not as good, and so the sample period was chosen to begin in 1916. The data prior to 1916 are presented in Fair (1996).

$$V = \alpha_1 + \alpha_2 t + \alpha_3 I + \alpha_4 DPER + \alpha_5 g \cdot I + \alpha_6 p \cdot I + u$$

where t is a time trend that takes a value of 8 in 1916, 9 in 1920, and so on; I is 1 if the Democrats are in the White House at the time of the election and -1 if the Republicans are; *DPER* is 1 if the president himself is running and is a Democrat, -1 if the president himself is running and is a Republican and 0 otherwise; g is the growth rate of real per capita GDP over some specified period prior to the election; and p is the absolute value of the inflation rate over some specified period prior to the election. Whenever "growth rate" is used in this paper, it always refers to the growth rate of real, per capita GDP at an annual rate. Likewise, "inflation rate" refers to the absolute value of the growth rate of the GDP price index at an annual rate.

Table 1 presents four versions of this equation: the original and three updates. The 1976 equation is in Fair (1978, Table 2, equation 4), the 1980 equation in Fair (1982), the 1984 equation in Fair (1988) and the 1988 equation in Fair (1990). The general form of the equation remained unchanged over this time, although there were some modest changes in the definition of the variables. For example, from the 1980 equation on, g was changed from being the GDP growth rate in the year of the election to the growth rate in the second and third quarters in the year of the election. From the 1984 equation on, p was taken to be the absolute value of the inflation rate in the eight quarters prior to the election, with the last quarter being the third quarter of the election year. Earlier, the last quarter had been taken to be the fourth quarter of the election year. Finally, for the 1988 equation the time trend was stopped in 1976 and its value after 1976 was taken to be the 1976 value.[4]

The estimates in Table 1 are based on a small number of observations, and a number of the coefficients are not precisely estimated. The growth rate is always significant, but the inflation rate is not. However, the coefficients on the inflation rate do have the expected sign. The time trend is meant to pick up possible trend effects on the Democratic share of the vote since 1916, and it has a t-statistic close to 2 by the fourth version. The variable reflecting whether the president himself is actually running for reelection (*DPER*) has a t-statistic greater than or equal to 2 after the first version. The standard errors of the 1984 and 1988 equations are about 3 percentage points.[5]

[4] For the estimation of the 1976 and 1980 equations, Ford was counted as an incumbent running again in the construction of *DPER*. From the 1984 equation on he was not so counted, which improves the fit of the equation. Excluding Ford in *DPER* is justified(?) on the grounds that, unlike other vice presidents who became president, he was appointed rather than elected.

[5] An attempt was made in Fair (1978) to account for the independent vote-getting ability of someone who ran more than once. This was done by postulating certain restrictions on the covariance matrix of the error term when a person had run before. In the econometric work, a parameter of the covariance matrix was estimated along with the structural coefficients by a nonlinear procedure. The 1976 and 1980 equations in Table 1 are estimated using this procedure, but from the 1984 equation on the restrictions were not imposed and the equations were just estimated by ordinary least squares. I have always liked this treatment of the covariance matrix, and I hope that some students may look it up; but it is probably too clever by half given the limited number of observations. The estimates of the parameter in the covariance matrix were never very precise.

Table 1
Previous Versions of the Voting Equation

Eq. Name: *Sample:*	*1976* *1916–1976*	*1980* *1916–1980*	*1984* *1916–1984*	*1988* *1916–1988*
cnst	.401	.418	.4073	.4021
	(6.45)	(10.08)	(11.73)	(11.70)
t	.00474	.00346	.0033	.0036
	(1.29)	(1.52)	(1.80)	(1.97)
I	.0043	.0147	.0049	.0053
	(0.16)	(0.81)	(0.29)	(0.34)
DPER	.0485	.0415	.0449	.0424
	(1.69)	(2.00)	(2.69)	(2.74)
$g \cdot I$	.0088	.0098	.0102	.0104
	(2.12)	(3.50)	(4.99)	(5.30)
$p \cdot I$	−.0055	−.0068	−.0034	−.0031
	(−0.98)	(−1.79)	(−1.13)	(−1.07)
SE	.0422	.0352	.0310	.0296
R^2	NA	NA	.887	.890
No. obs.	16	17	18	19

t-statistics are in parentheses.
See the text and the Appendix for the definitions of the variables: *g* is *gYR* for the 1976 equation and *g2* for the others; *p* is *p2YR* for the 1976 and 1980 equations and *p8* for the others.

Data mining is a potentially serious problem in the present context, given the small number of observations. Much searching was done in arriving at the final specification, and it may be that an equation was found that fits the historical data well, but that is, in fact, a poor approximation of the way that voters actually behave. Put another way, the equation may be overparameterized: since there are a relatively high number of parameters for the number of observations, small changes in the data or the specification can lead to substantial changes in the estimates. How can one test for this?[6] The most straightforward test is to see how the equation predicts outside the estimation period. If the equation is badly misspecified, it should not predict the future well even if the actual values of the economic variables are used.

Prior to 1992, the equation looked good. Using the actual economic values, the 1980 equation makes a prediction error of only .028 for the 1984 election and the 1984 equation makes a prediction error of only .017 for the 1988 election. Also, the coefficient estimates do not change much as new observations are added. But the 1992 election was a bad one for the equation. The 1992 error for the 1988 equation, for example, is .098 (using actual economic values), which is over three times the estimated standard error of the equation. In 1992 President Bush had the incumbency advantage, the inflation rate was modest and the growth rate was

[6] With so few observations, structural stability tests are not really practical.

not too far below average; and so the 1988 equation predicted that he should have had an easy victory, which he did not.

The 1992 Update

The main concern of the most recent revision was trying to account for the large error in predicting the 1992 election. Several possibilities deserve consideration: how the presence of the Perot vote disrupted the equation, whether recent revisions to economic data might offer a more precise fit and whether to add one or two new variables to the basic equation. Certain adjustments do improve the fit of the equation for the 1992 election, but the need for such adjustments provides less confidence in future predictions.

Treatment of Third-Party Votes

Except for the 1924 election, where the votes for Davis and LaFollette have been added together and counted as Democratic,[7] no adjustments have been made in my work for third-party votes: V is the Democratic share of the *two*-party vote. By not making an adjustment, it is implicitly assumed that the percentage of the third-party votes taken from the Democrats is the same as the Democratic share of the two-party vote. For example, President Clinton got 53.5 percent of the two-party vote in 1992, and there were 20.4 million third-party votes, mostly for Perot. If it is assumed that Clinton would have received 53.5 percent of the third-party votes had there been no third-party candidates, his share of the *total* vote would also have been 53.5 percent. Haynes and Stone (1994, p. 125) cite exit polls suggesting that Perot took about equal amounts from both Clinton and Bush, which is close to the implicit assumption made here of 53.5 percent being taken from Clinton. However, Ladd (1993) argues that Perot may have taken most of his votes from Bush. If this is true and one were to allocate most of Perot's votes to Bush, then the equation no longer would show a large prediction error for 1992. This would be an easy way of rescuing the equation, but I have chosen to stay with the assumption that Perot took roughly equal amounts of votes from Clinton and Bush.

New Economic Data

In calculating rates of past GDP growth, it is necessary to use a price index. Earlier work had relied on fixed-weight price indexes, but these have well-known problems associated with using fixed weights over long periods of time: the weights become less representative of actual spending patterns over time, until they are abruptly updated, at which point their quality again begins decaying. However, chain-link price indexes dating back to 1959 have now become available, which

[7] A slightly different procedure, based on the analysis in Burner (1971), was used for the 1992 update. LaFollette was assumed to have taken only 76.5 percent, rather than 100 percent, from Davis.

avoid many of the problems of fixed-weight indexes. The update after the 1992 election was able to use the GDP chain-link price index to deflate nominal GDP for the 1959:1–1992:4 period.[8] The other major data change was to use quarterly GDP data prior to 1946, as constructed by Balke and Gordon (1986). In the earlier work, only annual data were used.

A key question when dealing with revised data is whether one should use the latest revised data or the data as it was known at the time. I have always used the latest revised data in this context, based on the view that voters look at the economic conditions around them—how their friends and neighbors and employers are doing—and not at the numbers themselves.

The use of the updated data made a noticeable difference to the equation. When the 1988 equation was estimated over the original period, 1916–1988, using the new data, the time trend became insignificant and the coefficient on the growth rate fell by more than half, while the coefficient on the inflation rate more than doubled.[9] The other three coefficient estimates had noticeable changes as well. The fit of the equation using the updated data was not as good, with a standard error of .0325, and it had a larger outside-sample prediction error for 1992—.120 versus .098. This degree of sensitivity to the use of revised data is, of course, of some concern. As noted earlier, it may be a sign that the equation is, in fact, overparameterized, so that even small changes in the data can lead to large changes in the coefficient estimates.

Given the updated data and the new observation for 1992, searching was done to see which set of economic variables led to the best fit, and several changes were made. The growth rate in the three quarters before the election did better than the growth rate in only the last two. The inflation rate over the whole 15-quarter period before the election did better than the inflation rate only over the last eight. The time trend was dropped from the equation because it was clearly not significant. However, even with these changes, the basic equation still led to a large error in predicting the 1992 election.

A New Variable: The Number of Quarters of Good News

At the time of the election in 1992, the inflation rate was modest and the growth rate was not too bad. One might have thought that people would have been at least neutral about the economy, but surveys of consumer sentiment and voter attitudes in 1991 and 1992 revealed that people were quite pessimistic. Many possible reasons have been suggested: Bush wasn't interested enough in the economy; foreign competition seemed threatening; white collar workers were hit harder than

[8] Some of the early data are data on GNP, gross national product, rather than GDP, gross domestic product. The differences between GDP and GNP are trivial for the early years, and for ease of reference GDP will always be used in referring to the national output data.

[9] Specifically, the coefficient on the time trend fell from .0036 with a t-statistic of 1.97 to −.0007 with a t-statistic of −0.35. The coefficient estimate for the growth rate went from .0104 with a t-statistic of 5.30 to .0042 with a t-statistic of 2.49, and the coefficient estimate for the inflation rate went from −.0031 with a t-statistic of −1.07 to −.0070 with a t-statistic of −2.12.

usual in the 1990–1991 recession; the press was overly negative; or people were worried about growing income inequality and a lack of "good jobs at good wages."

Answers like the above are all plausible, but for a testable explanation, one needs a variable for which observations can be collected back to the election of 1916. With hindsight, what struck me about the 1989–1992 period was there was no quarter within the overall 15-quarter period before the 1992 election in which the growth rate was especially strong. The news was either bad (as during the 1990–1991 recession) or just OK. Maybe the lack of good news began to wear on people and led to their gloom.

To test this idea, a "good news" variable, denoted n, was constructed. This variable is the number of quarters of the first 15 quarters of each period of a presidential administration in which the growth rate is greater than 2.9 percent (which is the value that gave the best fit). By this measure, the Bush administration experienced zero quarters of good news. It is the only administration since 1916 for which this is true (as shown in the Appendix in Table A), and this obviously helps explain the 1992 result.[10]

An Additional Incumbency Variable

It has been argued that voters eventually get tired of a party if it has been in power a long time. A number of authors have used some measure of how long a party has been in the White House without a break to help explain votes for president (Abramowitz, 1988; Campbell and Wink, 1990; Haynes and Stone, 1994; Fackler and Lin, 1994). For the work here, five versions of a duration variable, denoted *DUR*, were tried. The general version of *DUR* was taken to be 0 if the incumbent party has been in power for only one or two consecutive terms, 1 [-1] if the Democratic [Republican] party has been in power for three consecutive terms, $1 + k$ [$-(1 + k)$] if the Democratic [Republican] party has been in power for four consecutive terms, $1 + 2k$ [$-(1 + 2k)$] if the Democratic [Republican] party has been in power for five consecutive terms, and so on. Values of k of 0, .25, .5, .75 and 1.0 were tried, and *DUR* is defined here for $k = .25$, where the best results were obtained.

The Final Version

The final variables in the equation are listed in the first column of Table 2. The equation differs from the 1988 equation in ways that have already been

[10] The use of n, which pertains to the entire 15-quarter period before the election, and the use of the inflation rate over this same period brings up the question of how to treat the war years. The 15-quarter period before the 1920 election is dominated by World War I, and the 15-quarter periods before the 1944 and 1948 elections are dominated by World War II. These periods may differ in kind from the other periods. To try to account for this problem, the assumption was made that the coefficients for n and for inflation are zero for these three elections. Voters are assumed to consider the other variables in the equation, including the growth rate in the year of the election, but not n and the inflation rate. This assumption leads to one extra coefficient being estimated. The new variable introduced is denoted d, which is 1 for the 1920, 1944 and 1948 elections and 0 otherwise.

Table 2
Estimates of the 1992 Update

Sample:	*1916–1992*	*1916–1988*	*1916–1960*
cnst	.468	.466	.463
	(90.62)	(124.05)	(88.08)
I	−.034	−.015	−.028
	(−1.26)	(−0.75)	(−1.31)
$I \cdot d$	.047	.016	.031
	(2.09)	(0.88)	(1.50)
$g3 \cdot I$	.0065	.0070	.0076
	(8.03)	(11.60)	(8.95)
$p15 \cdot I \cdot (1 - d)$	−.0083	−.0093	−.0066
	(−3.40)	(−5.21)	(−1.98)
$n \cdot I \cdot (1 - d)$	.0099	.0064	.0068
	(4.46)	(3.40)	(3.10)
DPER	.052	.061	.063
	(4.58)	(7.10)	(5.80)
DUR	−.024	−.017	−.016
	(−2.23)	(−2.14)	(−1.98)
SE	.0190	.0138	.0133
R^2	.960	.981	.990
No. obs.	20	19	12
$\hat{V}_{1992}$	.501	.467	.463

t-statistics are in parentheses.
See the text and the Appendix for the definitions of the variables.

mentioned. The time trend is dropped. The growth rate in the three quarters before the election replaces the growth rate in just the two quarters before. The inflation rate over the entire 15-quarter period replaces the inflation rate over only the last eight quarters. The good news variable *n* is added. The coefficients of the inflation rate and *n* are assumed to be 0 for the 1920, 1944 and 1948 elections (the "war" elections, as discussed in note 10). Finally, the duration variable *DUR* is added. One interesting implication of lengthening the time period for the inflation variable and adding the good news variable is that voters are being assumed to look back further than they did in previous versions of this model.

Ordinary least squares estimates for the three sample periods, 1916–1992, 1916–1988 and 1916–1960, are presented in Table 2. The coefficient estimates for the growth rate, the inflation rate and *n* are all significant and of the expected sign. Based on the coefficients for the first sample period (through 1992), one sees: an increase of 1 percentage point in the growth rate in the three quarters before the election increases the vote share by .65 percentage points; an increase of 1 percentage point in the inflation rate over the 15-quarter period decreases the vote share by .83 percentage points; and each quarter in which the growth rate is greater than 2.9 percent adds .99 percentage points to the vote share. The coefficient estimates of *DPER* and *DUR* are of the expected signs, positive and negative respectively. The estimated standard error of the equation is less than two percentage points at .0190, and the (within-sample) prediction for 1992 actually has Clinton winning with 50.1 percent of the two-party vote!

The second sample period in Table 2 drops the 1992 observation, which has a noticeable effect on some of the coefficient estimates. The coefficient estimate for the good news variable falls, which makes sense because it was important in helping to explain Bush's low share of the vote. The coefficient estimate for *DPER* rises, which makes sense because Bush was an incumbent running again. The (outside-sample) prediction for 1992 is .467, which, given that Clinton actually received 53.5 percent of the two-party vote, is a prediction error of .068. The estimated standard error of the equation is only .0138 (which then rises to .0190 when the 1992 observation is added).

The third sample period in Table 2 ends in 1960. The main result here is that the coefficient estimates for this sample period are very similar to the coefficient estimates for the 1916–1988 period, except perhaps for the coefficient estimate for inflation. The equation is quite stable in this respect.

The predictions and prediction errors for the equations estimated for the first and third sample periods in Table 2 are presented in Table 3. The errors for the 1916–1992 equation are all within-sample, but the errors for the 1916–1960 equation are outside-sample from 1964 on. All the predictions used the actual values of the economic variables. As expected, given the small estimated standard errors, the prediction errors are generally small in Table 3. The largest error for each equation occurs in 1992.

Perhaps the most remarkable feature of the errors in Table 3 is the string of very small errors between 1964 and 1988 for the equation estimated only through 1960. These are all outside-sample errors, and, for example, the error for the 1988 election is outside sample by 28 years. The mean absolute error for these seven errors is only .014. If the 1992 error of .072 is added, the mean absolute error rises to .021. This seems to me to be the strongest evidence in the paper in favor of the new voting equation.

A voting equation like the present one should be judged according to the size of its errors and not according to how many winners it correctly predicted. From a least squares point of view, a close election predicted incorrectly as to winner but with a small error is better than a landslide predicted correctly as to winner but with a large error.[11]

Of course, most people can't resist pointing out the elections in which the winner was not predicted correctly. For the 1916–1992 equation, the elections

[11] If, on the other hand, the aim is not to explain vote share, but to predict the winner correctly, a different procedure from least squares may be desirable. There are some interesting econometric issues here, but these are beyond the scope of the present paper.

Table 3
Prediction Errors for Table 2 Equations

			1916–1992 eq.		1916–1960 eq.	
Winner		V	$\hat{V}$	$V-\hat{V}$	$\hat{V}$	$V-\hat{V}$
Wilson	1916	.517	.495	.022	.507	.010
Harding	1920	.361	.382	−.021	.363	−.002
Coolidge	1924	.418	.419	−.001	.424	−.006
Hoover	1928	.412	.427	−.015	.426	−.014
Roosevelt	1932	.592	.607	−.015	.591	.001
Roosevelt	1936	.625	.629	−.004	.633	−.008
Roosevelt	1940	.550	.553	−.003	.551	−.001
Roosevelt	1944	.538	.522	.016	.531	.007
Truman	1948	.524	.518	.006	.528	−.004
Eisenhower	1952	.446	.449	−.003	.446	−.000
Eisenhower	1956	.422	.417	.005	.413	.009
Kennedy	1960	.501	.494	.007	.489	.012
Johnson	1964	.613	.617	−.004	.603	.010
Nixon	1968	.496	.504	−.008	.495	.001
Nixon	1972	.382	.392	−.010	.376	.006
Carter	1976	.511	.507	.004	.491	.020
Reagan	1980	.447	.446	.001	.453	−.006
Reagan	1984	.408	.387	.021	.373	.035
Bush	1988	.461	.489	−.028	.480	−.019
Clinton	1992	.535	.501	.034	.463	.072

that were predicted incorrectly as to the winner are the elections of 1916 (error of .022), 1960 (error of .007) and 1968 (error of −.008). For the 1916–1960 equation, the elections are 1960 (error of .012), 1976 (error of .020) and 1992 (error of .072). The errors for these elections are all small—except the error for the 1992 election.

Evaluation

What judgment should one make of the equation? If one just looks at the final equation estimated for the 1916–1992 period, it does a remarkable job in explaining votes for president. The estimated standard error is less than 2 percentage points, and the largest within-sample error is only 3.4 percentage points. Also, when the equation is estimated only with data through 1960, it does a good job of predicting the elections outside the sample from 1964 through 1988. In this sense the equation is very stable. The fact that the good news variable, n, is significant even when the equation is estimated only through 1960 suggests that it is not merely a dummy variable for the 1992 election.

On the other hand, there is plenty of reason to be cautious. The estimates are based on only 20 observations, and much searching was done in arriving at the "final" equation. This included searching for the best variables, for the best threshold values for the good news n and duration DUR variables, and restricting the sample period to elections from 1916 on. Strategic decisions that helped the statistical fit were made about how to treat the inflation and good news variables in war years (see note 10) and about not categorizing Ford as incumbent running again because he had not been elected vice president (see note 4). Finally, the outside-sample prediction errors for 1992 are large, and adding the 1992 observation to the estimation period results in fairly large changes in some coefficient estimates.

It may be that the equation is better at explaining the past than the future. Time will tell. If the equation predicts the next two or three elections within two or three percentage points, there may be something to it. Otherwise, I will have to keep searching or do something else in my updating week every four years.

Conditional Predictions of the 1996 Election

Any of the equations discussed in this paper can be used to make a prediction of the 1996 election conditional on the economy. Since Clinton is running for reelection, all the incumbency variables are known. For example, given the incumbency information, the equation in Table 2 estimated for the 1916–1992 period is: $V = .4859 + .0065g3 - .0083p15 + .0099n$. The first term is calculated by plugging in the coefficients for the known variables, multiplied by their coefficients, and adding the intercept term. In applying this equation, remember that the growth rate $g3$ and quarters of good news n pertain to growth rates of *per capita* real GDP. Since the U.S. population is currently growing at an annual rate of about 1 percent, the growth rates to use for the present calculations are 1 less than the growth rates for the aggregate economy normally quoted in the press.

Table 4 presents predictions of the 1996 election using the four equations in Table 1 and the three equations in Table 2. The economic forecasts used for these predictions are my own, and they are presented in the Appendix in the last row of Table A. These forecasts differ little from the "consensus" view.

The four equations from Table 1 predict a substantial Democratic victory. Since these equations do not take into account the defeat of incumbent Bush, it makes sense that they offer a larger incumbency advantage for Clinton. However, all three equations from Table 2 are predicting a close election. The equation estimated through 1992 has a predicted Democratic share of the two-party vote of .495 (which is a narrow Republican victory), and the other two equations have the predicted value slightly above .5. If the economic forecasts are accurate and if the election is close, then the Table 2 equations will have done well, regardless of which party wins. On the other hand, if the economic forecasts are accurate but the election is not close, then the equations will not have done well, again regardless of which party wins.

Table 4
Forecasts of the 1996 Election Democratic Share of the Two-Party Vote

Equation	$\hat{V}$
Table 2, 1916–1992	.495
Table 2, 1916–1988	.512
Table 2, 1916–1960	.508
Table 1, 1916–1988	.549
Table 1, 1916–1984	.564
Table 1, 1916–1980	.578
Table 1, 1916–1976	.588

So as not to be accused of presenting so many predictions that one of them is bound to be right, let me say that I take the equation from Table 2 estimated through 1992 to be my "final" choice. However, if one felt that Perot contaminated the 1992 election so much that the observation should not be used, then the second equation in Table 2 is a possible choice. It also predicts a close election. One would only use the previous version of the equation (like the fourth equation in Table 1) if it was felt that none of the changes for the 1992 update were any good.

Appendix
Data for Econometrics and Presidential Elections

Let Y be real GDP divided by population and let P be nominal GDP divided by real GDP. The construction of Y and P is explained in Fair (1996). Let subscript k denote the kth quarter within the 16-quarter period of an administration and let (-1) denote the variable lagged one 16-quarter period. Finally, let q_k be the growth rate of Y in quarter k (at an annual rate), which is $((Y_k/Y_{k-1})^4 - 1) \times 100$ for quarters 2 through 16 and $((Y_1/Y_{16}(-1))^4 - 1) \times 100$ for quarter 1. Then:

$$g2 = ((Y_{15}/Y_{13})^{(4/2)} - 1) \times 100$$
$$g3 = ((Y_{15}/Y_{12})^{(4/3)} - 1) \times 100$$
$$gYR = ((Y_{16} + Y_{15} + Y_{14} + Y_{13})/(Y_{12} + Y_{11} + Y_{10} + Y_9) - 1) \times 100$$
$$p8 = |((P_{15}/P_7)^{(4/8)} - 1) \times 100|$$

$p15 = |((P_{15}/P_{16}(-1))^{(4/15)} - 1) \times 100|$
$p2YR = |(((P_{16} + P_{15} + P_{14} + P_{13})/(P_8 + P_7 + P_6 + P_5))^{.5} - 1) \times 100|$
n = Number of quarters in the first 15 in which q_k is greater than 2.9.

Table A
The Data

	V	*I*	*DPER*	*DUR*	*g3*	*p15*	*n*	*g2*	*gYR*	*p8*	*p2YR*
1916	0.5168	1	1	0.00	2.229	4.252	3	−1.213	6.035	6.752	8.183
1920	0.3612	1	0	1.00	−11.463	16.535	5	−14.496	−8.147	17.240	16.800
1924	0.4176	−1	−1	0.00	−3.872	5.161	10	−9.777	−1.009	0.737	1.231
1928	0.4118	−1	0	−1.00	4.623	0.183	7	6.043	−0.225	0.242	0.753
1932	0.5916	−1	−1	−1.25	−15.574	6.657	3	−16.249	−14.369	10.127	10.177
1936	0.6246	1	1	0.00	12.625	3.387	9	18.765	12.933	0.928	1.349
1940	0.5500	1	1	1.00	2.420	0.553	8	10.350	6.708	0.542	0.765
1944	0.5377	1	1	1.25	2.910	6.432	13	3.317	5.768	3.737	3.965
1948	0.5237	1	1	1.50	3.105	10.369	3	3.976	2.289	8.852	9.520
1952	0.4460	1	0	1.75	0.910	2.256	7	0.632	1.935	3.542	4.160
1956	0.4224	−1	−1	0.00	−1.479	2.132	6	−0.576	0.338	2.958	2.674
1960	0.5009	−1	0	−1.00	0.020	2.299	5	−2.940	−0.038	2.314	2.260
1964	0.6134	1	1	0.00	4.950	1.201	11	3.334	4.374	1.375	1.288
1968	0.4960	1	0	1.00	4.712	3.160	9	4.125	3.599	3.775	3.748
1972	0.3821	−1	−1	0.00	5.716	4.762	6	5.216	4.247	4.611	4.629
1976	0.5105	−1	0	−1.00	3.411	7.604	6	0.923	4.406	7.288	7.650
1980	0.4470	1	1	0.00	−3.512	7.947	5	−5.494	−1.330	8.718	8.722
1984	0.4083	−1	−1	0.00	5.722	5.296	7	3.965	6.114	3.692	3.877
1988	0.4610	−1	0	−1.00	2.174	3.392	5	2.285	2.943	3.737	3.509
1992	0.5345	−1	−1	−1.25	1.478	3.834	0	1.102	0.979	3.325	3.461
1996[a]		1	1	0.00	2.100	3.000	2	2.200	1.500	2.200	2.200

[a] Economic values are forecasts made May 3, 1996.

References

Abramowitz, Alan I., "An Improved Model for Predicting Presidential Election Outcomes," *Political Science & Politics,* 1988, *21,* 843–47.

Balke, Nathan S., and Robert J. Gordon, "Appendix B Historical Data." In Gordon, Robert J., ed., *The American Business Cycle: Continuity and Change.* Chicago: University of Chicago Press, 1986, pp. 781–810.

Burner, David, "Election of 1924." In Schlesinger, Arthur M., Jr., ed., *History of American Presidential Elections 1789–1968, Volume III.* New York: McGraw-Hill Book Co., 1971, pp. 2459–2553.

Campbell, James E., and Kenneth A. Wink, "Trial-Heat Forecasts of the Presidential Vote," *American Politics Quarterly,* 1990, *18,* 251–69.

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

Fackler, Tim, and Tse-min Lin, "Political Corruption and Presidential Elections." Paper presented at the 1994 meetings of the Midwest Political Science Association, Chicago, Illinois, April 14–16, 1994.

Fair, Ray C., "The Effect of Economic Events on Votes for President," *Review of Economics and Statistics,* 1978, *60,* 159–73.

Fair, Ray C., "The Effect of Economic Events on Votes for President: 1980 Results," *Review of Economics and Statistics,* 1982, *64,* 322–25.

Fair, Ray C., "The Effect of Economic Events on Votes for President: 1984 Update," *Political Behavior,* 1988, *10,* 168–77.

Fair, Ray C., "The Effect of Economic Events on Votes for President: 1988 Update." Yale University, New Haven, Conn., 1990. Mimeographed.

Fair, Ray C., "The Effect of Economic Events on Votes for President: 1992 Update," *Political Behavior,* 1996, *18,* 119–39.

Haynes, Stephen E., and Joe A. Stone, "Why Did Economic Models Falsely Predict a Bush Landslide in 1992?" *Contemporary Economic Policy,* 1994, *12,* 123–30.

Ladd, Everett Carll, "The 1992 Vote for President Clinton: Another Brittle Mandate?" *Political Science Quarterly,* 1993, *108,* 1–28.

Kramer, Gerald H., "Short-Term Fluctuations in U.S. Voting Behavior, 1896–1964," *American Political Science Review,* 1971, *65,* 131–43.

Stigler, George J., "General Economic Conditions and National Elections," *American Economic Review,* 1973, *63,* 160–67.