Reprinted from THE REVIEW OF ECONOMICS AND STATISTICS *Published for Harvard University by the North-Holland Publishing Company Copyright, 1982, by the President and Fellows of Harvard College* Vol. LXIV, No. 2, May, 1982

NOTES

THE EFFECT OF ECONOMIC EVENTS ON VOTES FOR PRESIDENT: 1980 RESULTS

Ray C. Fair*

In Fair (1978) I developed a fairly general model of voting behavior. The model was tested using data on U.S. presidential elections. The results indicate that voters do not look back more than a year or two in judging the economic performance of an administration and that they look at the change in the economy rather than at its level in measuring economic performance. The basic sample period was 1916–1976.

From an econometric point of view, the 1980 election is a very good observation. The growth rate of real GNP six months prior to the election was much different from the growth rate one year prior. This helps break the collinearity between these two variables, and so the 1980 observation may help in deciding which of the two variables is the better measure of performance. No conclusions were drawn from the earlier results regarding this choice. Also, the rate of inflation prior to the election was high by historic standards, and this increased variance may help in deciding the effects of inflation on voting behavior. There was some evidence from the earlier results that inflation affected voting behavior, but it was not very strong.

The new results are presented in table 1. Equation 1 is the same as equation 4 in Fair (1978, table 2, p. 168). Equation 2 is the same equation estimated using the revised national income accounts data. (A major revision occurred in 1980.) The estimates of the two equations are quite similar, although the 1976 error for equation 2 is somewhat smaller (0.0566 vs 0.0607). As can be seen near the bottom of table 1, the growth rate for 1976 was revised downward, and this led to a larger predicted vote for Carter (0.4539 vs 0.4498).[1]

Equations 3 and 4 are the same as 1 and 2, respectively, except for the use of the growth rate in the second and third quarters of the election year (g^*_1) rather than the growth rate for the whole year (g_1). The estimates of the two equations are similar, although again the equation estimated using the revised data has a smaller error for 1976. The fits of equations 3 and 4 are somewhat better than the fits of equations 1 and 2, and one might conclude from these results that g^*_1 is a better measure of performance than is g_1. I did not, however, draw this conclusion in the earlier paper. Given the dangers of data mining in an exercise of this sort, the results did not seem conclusive enough to warrant this.[2] At any rate, it does now seem from the 1980 results that this conclusion can be drawn. This can be seen in two ways. First, the (outside sample) prediction error for 1980 for equation 4 is much smaller than that for equation 2 (−0.0226 vs −0.0693). Second, when both equations are estimated through 1980 (equations 5 and 6), the difference in fits is now greater (0.0352 vs 0.0415 compared to 0.0378 vs 0.0422). With respect to the effects of inflation, the t-statistic for the inflation variable increased in absolute value with the addition of the 1980 observation, and so somewhat more confidence than before can be placed on the hypothesis that inflation affects voting behavior. The inflation variable used, $|p_{2t}|$, is the absolute value of the inflation rate in the two-year period before the election.

The individual error terms for equation 6 are presented at the bottom of table 1. The largest error occurred in 1964, where the equation underpredicted the large vote for Johnson by 0.0563. The next largest error occurred in 1976, where the equation incorrectly predicted a Ford victory and made an error of 0.0372. Two other elections were predicted incorrectly regarding the winner: 1948 (Truman vs Dewey) and 1968 (Humphrey vs Nixon). These were, however, close elections, and the errors were moderate (0.0275 and −0.0340). The 1980 election was predicted quite well. The two main factors contributing to the low predicted vote for Carter were the low growth rate (−4.9%) and the high inflation rate (8.7%).

$\hat{\lambda}$ in table 1 is the estimate of the percentage of the variance of the total error term due to a candidate's unobserved independent vote getting ability or lack thereof (VGA). Because a number of people have run

Received for publication June 18, 1981. Revision accepted for publication September 8, 1981.

* Yale University.

The research described in this paper was financed by Grant SOC77-03274 from the National Science Foundation.

[1] I am making an implicit assumption here that the revised data better approximate the economic conditions known to the voters. The assumption is that the voters look not at the published GNP numbers in deciding how to vote, but rather at the actual conditions around them.

[2] Equation 3 was not presented in Fair (1978). Most of the estimated equations in the paper excluded the inflation variable. With the inflation variable excluded, the results for g_1 and g^*_1 in the earlier paper are virtually identical with respect to fit (0.0421 vs 0.0419).

TABLE 1.—ESTIMATES OF THE VOTING EQUATION
(DEPENDENT VARIABLE IS THE DEMOCRATIC SHARE OF THE TWO-PARTY VOTE)

Eq. No.	Old or New Data	Sample Period	Explanatory Variables g_{u}	g^*_{u}	$\lvert p_{2t}\rvert$	$DPER_t$	t	$CNST_1$	I_t	$\hat{\lambda}$	SE	1976 Act. Pred. Error	1980 Act. Pred. Error
1	Old	1916–1976	0.0088 (2.12)		−0.0055 (−0.98)	0.0485 (1.69)	0.00474 (1.29)	0.401 (6.45)	0.0043 (0.16)	.07 (.15)	.0422	.5105 .4498 .0607	—
2	New	1916–1976	0.0092 (2.35)		−0.0052 (−0.95)	0.0487 (1.81)	0.00484 (1.40)	0.399 (6.75)	0.0025 (0.10)	.09 (.19)	.0408	.5105 .4539 .0566	.4470 .5163 −.0693
3	Old	1916–1976		0.0099 (3.35)	−0.0062 (−1.42)	0.0374 (1.63)	0.00429 (1.48)	0.406 (8.19)	0.0146 (0.76)	.31 (.81)	.0378	.5105 .4636 .0469	—
4	New	1916–1976		0.0099 (3.48)	−0.0061 (−1.43)	0.0416 (1.91)	0.00407 (1.49)	0.408 (8.59)	0.0123 (0.64)	.29 (.69)	.0367	.5105 .4748 .0357	.4470 .4696 −.0226
5	New	1916–1980	0.0079 (2.61)		−0.0080 (−1.73)	0.0496 (1.98)	0.00267 (1.03)	0.433 (8.91)	0.0131 (0.54)	0[a]	.0415	.5105 .4530 .0575	.4470 .4812 −.0342
6	New	1916–1980		0.0098 (3.50)	−0.0068 (−1.79)	0.0415 (2.00)	0.00346 (1.52)	0.418 (10.08)	0.0147 (0.81)	.26 (.56)	.0352	.5105 .4733 .0372	.4470 .4595 −.0125
7	New	1916–1980		0.0080 (2.36)	−0.0096 (−2.74)	0.0479 (2.15)	—	0.478 (38.70)	0.0230 (1.25)	.25 (.52)	.0373	.5105 .4627 .0478	.4470 .4378 .0092
8[b]	New	1916–1980		0.0095 (3.27)	−0.0055 (1.69)	0.0493 (2.45)	0.00290 (1.37)	0.419 (11.24)	0.0093 (0.51)	.04 (.06)	.0317	.5105 .5001 .0104	.4470 .4530 −.0060
9	New	1916–1980		0.0073 (3.64)	−0.0074 (2.56)	0.0637 (3.35)	—	0.476 (43.30)	—	—	.0365	.5105 .4530 .0575	.4470 .4396 .0073

Notes: g_{u} = growth rate of real per capita GNP in the year of the election, multiplied by −1 if the incumbent party is Republican.
g^*_{u} = growth rate of real per capita GNP in the second and third quarters of the year of the election (annual rate), multiplied by −1 if the incumbent party is Republican.
$\lvert p_{2t}\rvert$ = absolute value of the average growth rate of the GNP deflator in the two-year period before the election (annual rate), multiplied by −1 if the incumbent party is Republican.
$DPER_t$ = 1 if there is a Democratic incumbent and he is running for election, −1 if there is a Republican incumbent and he is running for election, 0 otherwise.
t = time trend: 1 in 1888, 2 in 1892, ... 24 in 1980.
$CNST$ = constant term.
I_t = 1 if there is a Democratic incumbent, −1 if there is a Republican incumbent.
$\hat{\lambda}$ = estimate of the percentage of the variance of the total error term due to a candidate's unobserved vote getting ability or lack thereof.
SE = estimated standard deviation of the total error term.

t-statistics are in parentheses. Equation 1 is equation 4 in table 2 in Fair (1978, p. 168).
[a] Corner solution value.
[b] For equation 8, $DPER_t$ was taken to be 0 rather than −1 in 1976.

Data (not multiplied by −1 in 1976)

	Old g_{u}	Old g^*_{u}	Old $\lvert p_{2t}\rvert$	New g_{u}	New g^*_{u}	New $\lvert p_{2t}\rvert$
1976	5.4	3.4	7.2	4.6	1.7	7.2
1980	—	—	—	−1.1	−4.9	8.7

Actual and Predicted Values for Equation 6

	1916	1920	1924	1928	1932	1936	1940	1944	1948	1952	1956	1960	1964	1968	1972	1976	1980
Actual	.5168	.3613	.4568	.4124	.5914	.6247	.5500	.5377	.5237	.4460	.4224	.5009	.6134	.4960	.3821	.5105	.4470
Predicted	.5114	.3486	.4221	.4528	.6022	.6329	.5805	.5471	.4962	.4696	.4378	.5056	.5571	.5299	.4089	.4733	.4595
Error	.0054	.0127	.0347	−.0404	−.0108	−.0082	−.0305	−.0094	.0275	−.0236	.0155	−.0047	.0563	−.0340	−.0268	.0372	−.0125

more than once, it is possible to obtain this estimate under the assumptions (1) that a particular individual's *VGA* is the same in all elections in which he is a candidate and (2) that the variance of the *VGA*s across Democratic candidates is the same as the variance of the *VGA*s across Republican candidates (see Fair (1978, p. 163)). The estimate for equation 6 is 0.26, which means that about one-fourth of the variance of the total error term is estimated to be due to the Democratic candidate's *VGA*, about one-fourth to the Republican candidate's *VGA*, and about one-half to other factors.

If an individual has run more than once, it is possible to determine the effects of his *VGA* on the predicted vote. This can be done by making two predictions for a given election, one using the estimated value of λ and one using a value of zero. In Carter's case, for example, the predicted value for 1980 using equation 6 and the estimated value of λ is 0.4595 (bottom of table 1). When a value of zero is used, the predicted value is 0.4494. Because the equation underpredicted the vote for Carter in 1976, his estimated *VGA* is positive, and so the equation predicts more votes for Carter in 1980 when the estimated value of λ is used. The estimated *VGA* for Reagan is small because the estimated error term for 1980 is small.

Alternative measures of the growth rate and the rate of inflation were tried using the new data, and none proved to be as good as g^*_1 and $|p_{2t}|$. In particular, the results using the growth rate in the first three quarters of the year were not as good. They were in between the results using the growth rate in the second and third quarters and the results using the growth rate in the entire year. Also, the results using the one-year rather than the two-year rate of inflation were not as good.

As mentioned above, there is a serious danger of data mining in a study like this, and one is reluctant to try too many versions. Three other versions are, however, of some interest. First, the use of the time trend in equation 6 is somewhat unsatisfying, and it is clearly not sensible to extrapolate this variable indefinitely into the future. Fortunately, the results are not very sensitive to the exclusion of the time trend. Equation 7 is the same as equation 6 without the time trend. The estimates of the two equations are close. The main difference is an increase in absolute value in the size of the coefficient estimate of the rate of inflation variable (−0.0068 vs −0.0096).

Second, the question of how to treat Ford with respect to the $DPER_t$ variable is a difficult one. Since he was appointed rather than elected, should he be counted as a Republican incumbent running for reelection? For the first 7 equations in table 1 he has been so counted; for equation 8 he has not been. The estimates for equation 8 are similar to those for equation 6 except for the estimate of λ, which is much smaller. The fit of equation 8 for 1976 is much better, and the election is

TABLE 2.—PREDICTED DEMOCRATIC SHARE OF THE TWO-PARTY VOTE FOR PRESIDENT FOR 1984 USING EQUATION 6 IN TABLE 1

	$\lvert p_{2t}\rvert$									
g^*_{1t}	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0	11.0	12.0
					Reagan Running					
−6.0	.525	.532	.539	.546	.553	.561	.568	.575	.582	.589
−5.0	.515	.522	.529	.536	.543	.550	.557	.565	.572	.579
−4.0	.504	.512	.519	.526	.533	.540	.547	.554	.561	.569
−3.0	.494	.501	.509	.516	.523	.530	.537	.544	.551	.558
−2.0	.484	.491	.498	.505	.513	.520	.527	.534	.541	.548
−1.0	.474	.481	.488	.495	.502	.509	.517	.524	.531	.538
0.0	.464	.471	.478	.485	.492	.499	.506	.514	.521	.528
1.0	.453	.461	.468	.475	.482	.489	.496	.503	.510	.518
2.0	.443	.450	.457	.465	.472	.479	.486	.493	.500	.507
3.0	.433	.440	.447	.454	.461	.469	.476	.483	.490	.497
4.0	.423	.430	.437	.444	.451	.458	.466	.473	.480	.487
5.0	.413	.420	.427	.434	.441	.448	.455	.462	.470	.477
6.0	.402	.409	.417	.424	.431	.438	.445	.452	.459	.466
					Reagan Not Running					
−6.0	.569	.576	.583	.589	.596	.603	.610	.617	.624	.631
−5.0	.559	.566	.573	.580	.586	.593	.600	.607	.614	.621
−4.0	.549	.556	.563	.570	.577	.583	.590	.597	.604	.611
−3.0	.539	.546	.553	.560	.567	.574	.581	.587	.594	.601
−2.0	.530	.536	.543	.550	.557	.564	.571	.578	.584	.591
−1.0	.520	.527	.533	.540	.547	.554	.561	.568	.575	.581
0.0	.510	.517	.524	.530	.537	.544	.551	.558	.565	.572
1.0	.500	.507	.514	.521	.527	.534	.541	.548	.555	.562
2.0	.490	.497	.504	.511	.518	.524	.531	.538	.545	.552
3.0	.480	.487	.494	.501	.508	.515	.522	.528	.535	.542
4.0	.470	.477	.484	.491	.498	.505	.512	.519	.525	.532
5.0	.461	.468	.474	.481	.488	.495	.502	.509	.516	.522
6.0	.451	.458	.465	.471	.478	.485	.492	.499	.506	.513

Notes: See table 1 for the definition of g^*_{1t} and p_{2t}.

now predicted correctly with respect to the win for Carter. It is not clear, however, that much confidence should be placed on this result. Each election has special features, and without too much work one could rig the equation to predict most of the observations perfectly within sample.

Third, the editor[3] has expressed concern about the number of coefficient estimates in, say, equation 6 that are insignificant by conventional standards. Given the small number of observations, this is not necessarily surprising, and my general procedure is to retain insignificant variables if their coefficient estimates are of the right sign. I would rather risk the loss of some efficiency by incorrectly including a variable than risk the loss of consistency by incorrectly excluding it. In the present case there is no question that the likelihood function is fairly flat in the vicinity of the estimate of λ, although the size of the estimate seems reasonable. The only observations that really matter for the estimate of λ are those in which at least one of the candidates has run before, and so the number of observations involved in estimating λ is very small. There is also considerable collinearity among the estimates of λ and the coefficients of $DPER_t$ and I_t. (The coefficient of I_t is merely the constant term in the definition of the measure of performance (see Fair (1978, p. 164)), and it is not very important.) There is also some collinearity between the time trend and the inflation variable. Equation 9 is the same as equation 6 except that the time trend, I_t, and λ have been excluded. The main change is that the estimate of the coefficient of $DPER_t$ has increased from 0.0493 to 0.0637. The results are otherwise quite similar. The 0.0637 estimate of the coefficient of $DPER_t$ in equation 9 seems somewhat high to me, and I am inclined to stay with equation 6 even though there are insignificant estimates. Tastes differ on this, however, and there is clearly no right answer.

As a final exercise, equation 6 was used to predict the 1984 election for alternative values of g^*_t and $|p_{2t}|$. The results are in table 2. For an inflation rate of, say, 8.0%, the break-even point (0.5) corresponds to a growth rate of about 0.0%. One standard deviation (0.0352) below 0.5 corresponds to a growth rate of slightly more than 3.0% for an inflation rate of 8.0%. If Reagan does not run, the Republicans do not have the person advantage, and the break-even points are less favorable for them.

[3] I am grateful to the editor for helpful discussions regarding this point. I, however, am responsible for all errors of fact or judgment in the following.

REFERENCE

Fair, Ray C., "The Effect of Economic Events on Votes for President," this REVIEW 60 (May 1978), 159–173.