Newsle tter of the oliti cal Metho dology ecti on

Ame rican oliti cal cie nce Asso ciation
Volume 5, Number 2, Spring-Summer 1994

Contents

Notes From the Editors, New and Old

Charles H. Franklin
University of Wisconsin, Madison

R. Michael Alvarez
California Institute of Technology

Nathaniel Beck
University of California, San Diego

After a long absence, *TPM* is back. In this issue, we take up where we left off, with the 1992 election. This issue of *TPM* was co-edited by Charles Franklin, Michael Alvarez, and Nathaniel Beck. In the next issue, expected in the fall, the new editors will take over with Caltech kindly providing editorial facilities.

The new editors realize that Charles Franklin (and Gary King before him) set a very high standard for *TPM*, making it by far the best newsletter of all of the **APSA** organized subfields. We will try our best to keep to their standard, but we wish they hadn't made the task so difficult.

The new editors welcome submissions of all types, other than pure research articles which should be sent to John Freeman for consideration in *Political Analysis*. We would like submissions which discuss research methodologies, teaching methodology (both graduate and undergraduate), as well as book and software reviews. In short, we are looking for anything interesting which doesn't belong in *Political Analysis*.

Also, the new editors would like to maintain *TPM* as a forum for discussing issues of teaching methods in our graduate programs. Here we would like to regularly publish graduate syllabi and reviews of graduate methodology textbooks. But we also want to broaden this discussion to include our other vast constituency — undergraduate methods courses and texts.

Further, the new editors are very interested in finding articles about computer software. In the past, *TPM* has had several excellent discussions of the GAUSS software system. We would like to broaden this discussion to include other types of software used by political methodologists in their research and teaching. In our next issue we will compare GAUSS to another powerful statistical software system, S+. We would love to have more articles which either assess software packages from the standpoint of political methodology or show how to use these packages to do innovative things, in future issues of *TPM*.

In short, please send to the new editors things which you would like to read in *TPM*. We will strive to pull together thematic issues (e.g., undergraduate teaching, graduate teaching, semi-parametric models, high-performance computing). Each issue will contain book and software reviews and other items of general interest to political methodologists. If you are unsure about whether to send something, please send it!

We will return to a biannual publication schedule, with one issue in the early winter and the second in the early summer. However, we will be happy to take materials all year. The new editors would strongly prefer that items were submitted in LaTeX but we will take anything! Please send e-mail submissions to beck@uscd.edu **AND** rma@hss.caltech.edu. Hard copy submissions should be sent to R. Michael Alvarez, Division of Humanities and Social Sciences 228-77, California Institute of Technology, Pasadena CA 91125.

As we mentioned earlier, this issue contains a series of short pieces about forecasting presidential elections. This endeavor has become quite popular recently, and among political scientists has become a growth industry. Interestingly, the media has begun to take our forecasts seriously, at least enough to report them with a straight face.

In this issue we have nine different discussions of presidential forecasting models. One point must be kept in mind about these models. They all were submitted to the previous editors *before the 1992 presidential election had been decided*. In that sense, all are trying to predict the outcome of the past presidential race, with different degrees of success. Of course, we will let you decide which models predict presidential elections the best, but keep in mind the following three facts:

- Clinton won the election with 43% of the popular vote; Bush followed with 38% and Perot with 19%.
- For Clinton this translated into 370 electoral votes, or 68.8% of the votes cast in the Electoral College; Bush obtained 38 electoral votes and Perot none.
- Clinton obtained 53.4% of the two-party vote.

However, more important than which models correctly predicted the outcome of this election, each of these authors takes a different approach to the methodological and ethical issues associated with academic attempts to predict political events. As the 1994 midterm elections approach, and with the 1996 presidential election campaign only about 20 months away, it is imperative that many of these issues be debated in the months to come.

Also in this issue of *TPM* is the first of a two-part series on the use of Granger causality to make inferences about the validity of exogeneity assumptions, a further caution about the use of summary statistics, and the programs from both the upcoming summer methodology meetings and APSA.

Is It Time for Them to Go? Forecasting the 1992 Presidential Election

Alan I. Abramowitz
Emory University

Is it time for George Bush and Dan Quayle to go? This paper will attempt to answer this question by using a model developed for forecasting the outcomes of U.S. presidential elections. The model is a slightly revised version of one originally described in a 1988 article in *PS* (Volume 21, pp. 843–846.) It is based on three predictors: the Commerce Department's estimate of the change in real gross domestic product (GDP) between the first and second quarters of the election year, the president's Gallup Poll approval rating in July (APPROVAL) and a dummy variable that takes on the value 0 if a party has been in power for four years and 1 if a party has been in power for eight years or longer (TIME). The dummy variable, which is intended to capture the presence of "time for change" sentiment in the electorate, is unique to this model and distinguishes it from similar models such as that of Lewis-Beck and Rice.

The model, along with a model excluding the "time for change" dummy variable, was estimated with data on all 11 presidential elections since World War II. The dependent variable was the incumbent party's share of the major party vote. The results are shown in Table 1. The expanded model, including the "time for change" variable, performed substantially better than the basic model. However, the 1988 elections produced by far the largest error of any prediction made by the model for the 11 post-War elections — underestimating George Bush's vote by 2.9 percentage points (see Table 2). One possible explanation for the relatively inaccurate prediction in 1988 is that the model assumes that both major party candidates will run competent campaigns whose effects will cancel each other out. In 1988, however, Michael Dukakis may have violated this assumption. What does the model predict for 1992? During

Table 1: Forecasting models of Vote Share

	Basic Model	Time for Change Model
Constant	38.4	46.6
Approval	.28	.16
s.e.	.08	.04
GDP	1.5	2.5
s.e.	1.0	0.5
Time		-5.7
s.e.		1.1
Adjusted R^2	.74	.94
SEE	3.1	1.5

Table 2: Prediction errors for "Time for Change" Model

Election	Actual Vote	Predicted Vote	Error
1948	52.3	51.6	−0.7
1952	44.6	45.5	+0.9
1956	57.8	58.9	+1.1
1960	49.9	49.7	−0.2
1964	61.3	61.2	−0.1
1968	49.6	50.5	+0.9
1972	61.8	61.1	−0.7
1976	48.9	50.8	+1.9
1980	44.7	44.2	−0.5
1984	59.2	59.4	+0.4
1988	53.9	51.0	−2.9

the second quarter of 1992, the Commerce Department estimated that GDP increased by 0.35 percent. In early July, George Bush's approval rating in the Gallup Poll was 31 percent. Based on these results, the predicted vote for the Bush-Quayle ticket is 46.7 percent. Thus, according to the model, it is time for them to go.

Can Bush Hit a Home Run?

R. Michael Alvarez
California Institute of Technology

Brian Loynd
Duke University

Political scientists have not been involved in predicting presidential elections until relatively recently. And some of these predictions have been quite close to the mark — especially when they use information available *after the election* to predict the outcome. As noted by Lewis-Beck (1985), a number of the most influential predictors in the better-known models are not available until well after the election is over (Rosenstone 1983, Tufte 1978). Additionally, the accuracy of these models is contingent on a number of very questionable and ad hoc operationalizations of variables, including the inclusion of variables for the proportions of Catholics in states in the 1960 election (Rosenstone 1983), variables for region in certain elections (Campbell 1992), and different intercepts for Republican incumbent presidents (Alesina, Londregan and Rosenthal 1991).[1]

Our prediction of the 1992 presidential election is based on a very simple model, the data for which is readily available before the election. It relies upon no bizarre operationalizations of variables. And, surprise, surprise, it produces a relatively good fit to the data we have collected for the 11 presidential elections in our sample.

The model is based on a very simplistic political-economic interpretation of national elections. The forecast variable is the percentage national vote for the Republican candidate. We have two variables to account for macroeconomic performance: annual percentage change in GNP and annual percentage change in personal consumption expenditures, both measured in 1982 dollars.

Two variables account for past political events: the national vote share of the Republican party in the previous Congressional midterm election and the president's approval

[1] Alesina, A., J. Londregan, and H. Rosenthal. "A Model of the Political Economy of the United States." Manuscript, October 1991; Campbell, J. E. "Forecasting the Presidential Vote in the States." *American Journal of Political Science* 36: 386- 407; Lewis-Beck, M. "Election Forecasts in 1984: How Accurate Were They?" *PS* 18: 53-62; Rosenstone, S. J. *Forecasting Presidential Elections*. New Haven: Yale University Press, 1983; Tufte, E. R. *Political Control of the Economy*. Princeton: Princeton University Press, 1978.

rating as measured by the Gallup Polls in the period most proximal to the election (June or August). And our last variable accounts for the mood of the electorate: which league wins the World Series in October.

The Prediction Model is:

Predictors of Republican Party Vote:

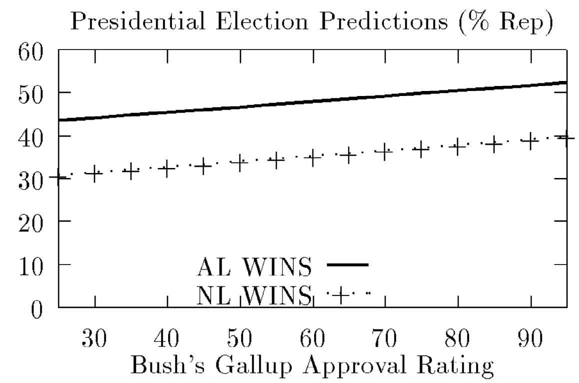
Independent Variables	Estimated Effect	Standard Error	T-test
Constant	126.3	22.4	5.65
World Series	12.7	3.36	3.80
Gallup Approval	0.13	0.10	1.17
Midterm Support	-1.77	0.47	-3.75
Economic Growth	1.51	1.44	1.05
Consumer Happiness	-3.55	1.63	-2.17

The dependent variable — the variable being "predicted" — is the Republican vote for president. The winner of the World Series is the first independent variable (1=American League winner, 0=National League Winner), the late summer Gallup percentage approving the president is the second variable, the Republican midterm Congressional vote share is the third variable. We use two economic indicators: Economic Growth, which is the annual change in economic growth, and what we call "Consumer Happiness", which is annual change in personal consumption expenditure. The model was estimated on the 11 presidential elections since 1948 (that being the first election for which we can obtain reliable approval ratings), and produces and R^2 of 0.80, with a standard error of 4.34 for a dependent variable with a mean of 50.96.

As is easily seen in the table, the model predicts the past 11 presidential elections very well. We will not engage in elaborate testing of the forecasting ability of the model on the past elections, in the sake of brevity. However, we will note that all of the variables in the model have standard errors smaller than the estimates, which we feel is quite a feat given only 11 observations in the dataset. Additionally, three of the variables have very precisely-estimated effects: Consumer Happiness (the happier consumers, the more they support the Democratic candidate), Republican Midterm Support (the better the Republicans do in the Congressional midterm election, the worst they do in the following presidential election), and the World Series Winner (if the American League wins the World Series, the Republican candidates average almost 13% more support, holding the other variables constant).

To use this model to predict the 1992 election, we simply substitute current values into the predicting model for the variables which are already fixed before the election: the Midterm Republican support (46.6% in 1990), Bush's late summer Gallup Approval rating (37%), Economic Growth

Figure 1: 1992 Forecasts



(1.4%), as well as Consumer Happiness (5.09%).[2]This model leads us to predict that the only scenario under which Bush can be re-elected in November is if he is able to increase his approval rating by approximately 40 points and if the American League wins the World Series. Our forecast can be seen in Figure 1.

The reader can use Figure 1 to come up with his or her own forecast for the election. One simply decides to what level Bush might be able to push his Gallup Approval rating to in the next few weeks. For example, take Bush's current Gallup approval rating of almost 40%. Based on the fixed variables (discussed above), the figure demonstrates that Bush is predicted to get only about 33% of the two-party vote, if the National League wins the World Series, and just over 45% two-party of the vote, if the American League wins the World Series. Keeping in mind that the standard error of the model is 4.34, that means the model predicts that Bush can obtain 41.7% of the two-party popular vote *at the most* if his Gallup approval rating remains at approximately 40% and the National League wins. If the American League wins, however, the model predicts that at best, Bush will obtain 53.7% of the two-party popular vote — enough to win.

We do not believe at this point that Bush will be able to change his Gallup approval rating. Nothing short of a *successful* major military action could possibly lead to a major increase in Bush's approval rating. We do not believe that even the bully in Baghdad will be silly enough to help Bush in his re-election bid.

Additionally, we predict that *there is no way in which an American League team can win the World Series this*

[2]The economic variables were taken from the current Department of Commerce forecast in the May 1992 *Survey of Current Business*.)

year.[3] Both authors agree that the World Series winner this year will come from the National League East: Alvarez's prediction is that Pittsburgh will win (since the Twins are unlikely to overcome their current deficit), while Loynd is confident that Atlanta will be the victors of the World Series (being a avid Durham Bulls fan).

And we do have empirical support for our baseball predictions. A semi-scientific poll of faculty and graduate students in the Duke Political Science Department last week yielded a 65.6% prediction of a National League win (37.9% picking Atlanta, 13.8% picking Pittsburgh; but 20.7% predicted that Toronto would win, producing the ironic twist that perhaps Bush's chances of re-election could be contingent on a World Series victory by a Canadian team — perhaps the rationale behind the election-year signing of the Free Trade Agreement?).[4]

Thus, we predict with great confidence that come January 21, 1993, Bill Clinton will be the occupant of the White House. Whether that is good or bad is left to the reader to decide. However, we would like to conclude by pointing out that our model is one of the few we have heard of which is currently *and clearly* predicting a Clinton victory using information available before the election. The other prediction models we have seen or heard of either are predicting a Bush re-election, or are predicting an election which is too close for their models to decide. All of this off of 11 cases, five variables, and a poll of our colleagues. So much for a predictive social science!

Forecasts of the 1992 Presidential Election

James Campbell
Louisiana State University

This article includes forecasts of both the 1992 presidential vote at both the national and state levels. I present two national level models, but the preferred model is the one that I published with Ken Wink in *American Politics Quarterly* in 1990. This forecasts a Clinton victory with Clinton receiving 52.9% and Bush getting 47.1% of the national popular two-party vote (see Tables 1 and 2). There are several minor differences in this model from that reported in *APQ*: (1) the early September trial-heat percentage divides "indecideds" and "others" evenly rather than proportionately between the two major party candidates, (2) the July rather than August *Survey of Current Business* second quarter economic reports are used, (3) because of the change in Commerce Department reporting emphases, the 1992 figure used is change in GDP rather than change in GNP (since these are percentage changes and since GNP is largely made up of GDP, the two numbers track pretty closely— though this change need to be examined more systematically), and (4) through an oversight, the *APQ* article erroneously reports that we used change in real GNP *per capita* rather than just real GNP change, the measure actually used.

The state presidential vote forecasts are based on a model reported in *AJPS*, with a couple of minor changes noted at the bottom of Table 3. The equation is reported in Table 3. Table 4 reports the actual state vote forecasts and aggregated state and national electoral votes under two conditions: the left column assumes Clinton–Gore receive the full boost (nearly 8 percentage points) in southern states from having a Southern Democrat on the ticket; the right column assumes Democrats get half of this bonus. This is admittedly ad hoc but it recognizes the problem of using historical models to forecast. I think that it is implausible to assume either that Clinton in 1992 will get the same "Bubba vote" that Carter received in 1976 or that he will get no advantage from being a southerner— thus I split the difference in Table 4B. This still predicts a Clinton victory, through by this estimate the election will be extremely close.

Between the two models, I am somewhat more confident in the national model— even though it is only based on eleven data points. The confidence is based on the examination of out-of-sample "forecast" errors in the national model. I think the use of the distribution of out-of-sample errors is a more meaningful way to judge confidence in these models than setting up traditional 95% confidence intervals.

Readers may be interested in Figure 1 of the *APQ* article. It demonstrates that while the raw trial-heat polls in early September on their own are not very good predictors of the vote, the model built with them is and is actually more accurate (measured by the mean error) than either the later trial-heats even as late as November or forecasts based on models employing these later trial-heats.

[3] **Editor's Note:** In 1992 the Minnesota Twins (American League) won the World Series.

[4] The poll was conducted from September 10-11, and respondents were asked to give their predicted team and an estimate of their subjective probability of that team winning. As an inducement for participation, we offered a choice of: 1. Five American Dollars; 2. One weeks worth of statistics homework help for first and second year graduate students; 3. That Alvarez and Loynd would never again say "It is endogenous", "Vector-Autoregression", or "Let the data speak" for the rest of their careers. 31 responses were obtained, most choosing the last inducement.

Table 1: Campbell and Wink Forecasting Equations of the Popular Two-Party National Presidential Vote, 1948–1988
Dependent Variable: The National Two-Party Popular Presidential Vote for the Incumbent Presidential Party

Predictor Variables	Equations: Original	Equations: Alternative	1992 Value
Early September Gallup Two-Party Trial Heat %	.531 (9.170)	.478 (8.292)	42.5
July Report of 2nd Qtr GNP Economic Growth	2.191 (5.422)	2.412 (6.505)	.345
Elected Incumbent Running for Reelection		1.803 (1.885)	1
Constant	23.817 (8.025)	25.742 (9.265)	
Number of Cases	11	11	
R^2	.952	.968	
Adjusted R^2	.940	.955	
Standard Error	1.516	1.320	
1992 Bush Vote Forecast	47.1%	48.7%	
Forecast Winner	Clinton	Clinton	

Note: t-scores are in parentheses. Equation 1 is essentially that used in Campbell and Wink (1990), except that earlier reports (July rather than August) of economic growth are used when available (since 1960). Also, the two-party trial-heat rating is based on an equal division of respondents not indicating a preference for either of the major party candidates (rather than dividing them in proportion to those with preferences for either of the two major party candidates).

Table 2: Out-of-Sample Forecast Errors for Trial-Heat Equations in Table 1

Election	Actual Vote	Original Specification Expected Vote	Original Specification Error	Alternative Specification Expected Vote	Alternative Specification Error
1948	52.3	52.3	+.1	52.1	+.3
1952	44.6	46.2	−1.6	45.8	−1.2
1956	57.8	53.8	+4.0	54.3	+3.4
1960	49.9	51.7	−1.7	51.0	−1.1
1964	61.3	63.2	−1.8	59.7	+1.6
1968	49.6	50.1	−.5	50.0	−.4
1972	61.8	61.7	+.1	63.3	−1.5
1976	49.0	47.5	+1.4	47.5	+1.5
1980	44.7	44.2	+.5	47.1	−2.4
1984	59.2	59.5	−.3	61.0	−1.9
1988	53.9	54.1	−.2	53.3	+.6
Mean Absolute Error			±1.1		±1.4
Median Absolute Error			±.5		±1.5
Largest Absolute Error			±4.0		±3.4

Note: Based on out-of-sample errors, the original specification has been more accurate than the alternative (from Table 1) in six of the seven elections. Based on this, the difference in mean and median errors, and the one less parameter required in the original equation, the original specification is preferred.
Percentage of out-of-sample forecast errors exceeding the 1992 Forecast Margin in the original specification ($e \geq 2.9$): 1 of 11 (.09).
1992 National Vote Forecast: Clinton 52.9% and Bush 47.1%

Table 3: The Forecasting Equation of the Two-Party Presidential Popular Vote Percentage in the States, 1948–1988
Dependent Variable: Democratic Two-Party State Presidential Vote %

Predictor Variables	Estimate	t-ratio	beta
Constant	26.895	17.706	
National Variables			
Democratic Trial-Heat %	.432	13.526	.365
2nd Qtr GNP Growth x Incumbent Pty	2.482	14.163	.339
Elected Incumbent Seeking Reelection	2.831	6.033	.215
Incumbent Presidential Party	−1.162	3.771	−.119
State Variables			
Prior State Deviation from National Vote (t-4) (adjusted)	.326	9.779	.268
Prior State Deviation from National Vote (t-8) (adjusted)	.272	9.694	.252
Presidential Home State Advantage	4.648	5.501	.097
Vice-Pres. Home State Advantage	1.886	2.206	.039
State Legislature Party Division (t-2)	.041	3.955	.101
Standardized 1st to 1st Qtr. State Economic Growth x Incumbent Party	.299	1.743	.031
State Liberalism Index (ADA & ACA)	.035	6.476	.127
Regional Variables			
Pres. Home Region (Southern) Advantage	7.953	7.932	.163
Southern State, 1964	−7.120	4.112	−.094
Deep Southern State, 1964	−18.001	6.529	−.139
New England State (1960 & 1964)	8.015	6.710	.122
Rocky Mountain West State (1976 & 1980)	−6.528	6.630	−.125
North Central State (1972)	6.152	4.555	.082
Number of Cases	531		
R^2	.847		
Adjusted R^2	.841		
Standard Error of Estimate	3.877		

Note: Partisan divisions (e.g. presidential vote %) involve only the two major parties. Except for regional trends, positive values on each variable favor Democratic presidential candidates. Regional trends are dummy variables (1 for states in the region for the specified year(s) and 0 otherwise.) The adjustment to the prior presidential vote deviations include the temporary prior state and regional (southern) presidential and vice presidential advantages. Twelve variables come into play in the 1992 forecast. The five regional variables are zero for all states in 1992.

Table 4: State 1992 Presidential Vote Forecast

	Full Southern Democrat Home Region Effect				Half of Southern Democrat Home Region Effect		
	Predicted	Electoral Vote			Predicted	Electoral Vote	
State	Dem Vote	Clinton	Bush	State	Dem Vote	Clinton	Bush
Utah	39.0	0	5	Utah	39.0	0	5
Idaho	39.9	0	4	Idaho	39.9	0	4
Nebraska	40.1	0	5	Nebraska	40.1	0	5
Wyoming	42.0	0	3	Wyoming	42.0	0	3
Arizona	43.6	0	8	Arizona	43.6	0	8
Nevada	43.8	0	4	Nevada	43.8	0	4
New Hampshire	44.0	0	4	New Hampshire	44.0	0	4
Kansas	44.1	0	6	Kansas	44.1	0	6
Alaska	44.8	0	3	Alaska	44.8	0	3
Oklahoma	44.9	0	8	Oklahoma	44.9	0	8
Indiana	45.1	0	12	Indiana	45.1	0	12
North Dakota	45.5	0	3	North Dakota	45.5	0	3
Colorado	47.0	0	8	Colorado	47.0	0	8
South Dakota	47.3	0	3	South Dakota	47.3	0	3
Delaware	47.3	0	3	Texas	47.3	0	32
Kentucky	48.3	0	8	Delaware	47.8	0	3
Ohio	48.5	0	21	South Carolina	48.2	0	8
New Mexico	49.3	0	5	Kentucky	48.3	0	8
Michigan	49.5	0	18	Virginia	48.5	0	13
New Jersey	49.9	0	15	Ohio	48.5	0	21
Montana	50.3	3	0	Florida	48.6	0	25
Missouri	50.4	11	0	Mississippi	48.6	0	7
Maine	50.7	4	0	New Mexico	49.3	0	5
Vermont	50.7	3	0	North Carolina	49.4	0	14
Texas	51.3	32	0	Michigan	49.5	0	18
Illinois	51.4	22	0	New Jersey	49.9	0	15
Connecticut	51.5	8	0	Alabama	49.9	0	9
South Carolina	52.1	8	0	Montana	50.3	3	0
California	52.2	54	0	Missouri	50.4	11	0
Pennsylvania	52.3	23	0	Georgia	50.5	13	0
Virginia	52.4	13	0	Maine	50.7	4	0
Florida	52.6	25	0	Vermont	50.7	3	0
Mississippi	52.6	7	0	Illinois	51.4	22	0
Oregon	53.0	7	0	Connecticut	51.5	8	0
Washington	53.1	11	0	Louisiana	52.0	9	0
Wisconsin	53.1	11	0	California	52.2	54	0
North Carolina	53.4	14	0	Pennsylvania	52.3	23	0
Minnesota	53.5	10	0	Oregon	53.0	7	0
Alabama	54.0	9	0	Washington	53.1	11	0
New York	54.0	33	0	Wisconsin	53.1	11	0
Iowa	54.0	7	0	Tennessee	53.1	11	0
West Virginia	54.2	5	0	Minnesota	53.5	10	0
Maryland	54.2	10	0	New York	54.0	33	0
Georgia	54.5	13	0	Iowa	54.0	7	0
Massachusetts	55.6	12	0	West Virginia	54.2	5	0
Louisiana	56.0	9	0	Maryland	54.3	10	0
Hawaii	56.4	4	0	Massachusetts	55.6	12	0
Tennessee	57.1	11	0	Hawaii	56.8	4	0
Rhode Island	59.2	4	0	Arkansas	56.8	6	0
Arkansas	60.8	6	0	Rhode Island	59.2	4	0
Total	51.5%	392	146		50.5%	284	254
States Carried		30+DC	20			23+DC	27

Forecasting the Presidential Vote, 1992

Robert S. Erikson and Christopher Wlezien
University of Houston

This essay is an attempt to forecast the two-party presidential vote for 1992, based on Erikson's (1989) model. That model predicts the incumbent vote from two variables: a measure of per capita disposable income growth (IG) adapted from Hibbs (1987) and a measure of net candidate evaluations (NCA) using the National Election Study (NES) survey for the fall campaign. IG is the cumulative weighted average of quarterly income growth during the current presidential term, with each quarter weighted 1.25 times as much as the preceding quarter. In this way, the rate of income growth counts increasingly more as the election approaches. NCA, or "net candidate advantage," is the mean (for major party voters only) net number of in-party candidate likes and out-party candidate dislikes minus the out-party candidates likes and in-party candidate dislikes, based on the NES open-ended responses reflecting what respondents like and don't like about the candidates. Following Tufte (1978), only evaluations based on the personal qualities of the candidates are counted. As described, the model explicitly accounts for election outcomes after they have occurred. But the model also is useful for forecasting purposes, as will become evident below.

The Equation

To begin with, it is necessary to update Erikson (1989) for the 1988 election. The estimated equation (with standard errors in parentheses) for the eleven presidential elections between 1948 and 1988 is

$$\begin{array}{c} V = 45.07 + 2.73\,IG + 6.33\ NCA + e_t \\ (1.31)\quad (.54)\quad (1.53) \\ \text{Adjusted } R^2 = .868 \\ \text{Standard Error of estimate} = 2.26 \end{array}$$

where V is the incumbent presidential party's percentage of the two-party presidential vote.[1] IG and NCA account for 87 percent of the variance in the incumbent party vote during the period, with a standard error of the estimate of 2.26 percentage points. In theory, the equation should predict about as accurately as a pre-election trial heat with a confidence interval of about 4 points.

However, the model is explanatory and not expressly useful for the purposes of forecasting, because it relies on information that is not available until after the election. While much of the information about cumulative income growth for the Bush presidency already is available, information about candidate evaluations is not. Thus, in order to provide a forecast, we need a way of predicting the 1992 net candidate advantage in advance.

Forecasting 1992

Net candidate advantage can be partially estimated, or forecasted, using two indicators that often find their way into models of the presidential vote— presidential approval and incumbency. The equation is

$$\begin{array}{c} NCA = -1.263 + .025\ A + .468\ I + e_t \\ (.17)\quad (.003)\quad (.097) \\ \text{Adjusted } R^2 = .911 \\ \text{Standard Error of estimate} = .15 \end{array}$$

where A is the percentage approving of the job the president is doing in the August (or most recent) Gallup Poll; I is equal to 1 if an incumbent president is running for reelection, 0 otherwise. Both August approval and incumbency predict net candidate advantage fairly well over the period. Indeed, the two variables account for 91 percent of the variance in NCA.[2]

Next, we plug +1 for incumbency and the most recent Gallup approval rating (37 percent in July 1992) into the equation to produce a "forecast" for the 1992 net candidate advantage. The resulting estimate of .141 suggests a slight edge for Bush over Clinton in terms of candidate evaluations. Driven by the positive coefficient for incumbency, the forecasted value of NCA actually exceeds Bush's advantage over Dukakis (.101) but is smaller than the advantage Ford had over Carter (.214) and Carter had over Reagan (.279). To forecast the vote for 1992, the forecasted NCA and the most recent information about cumulative income growth, $-.03$ for the term of Bush's presidency (through the second quarter of 1992), are inserted into the vote model. This predicts a 45.9% vote for Bush.[3] If we take seriously our standard error of 2.26 points (from equation 1), we might consider calling the election for Clinton. That would be wrong, however, because it not only assumes the correct model specification but also that we have exact knowledge of the ultimate evaluations for Clinton and Bush.

Still, the forecast looks good for Clinton, largely because of the very low level of cumulative income growth during the Bush presidency. For the expected vote to go for Bush, a net candidate advantage of about .8 is necessary. This advantage is more than double the edge Reagan had over

[1] Also see Erikson (1989) for details about measurement of IF and NCA.

[2] When the measure of income growth (IF) also is included in the NCA equation, its coefficient is surprisingly negative (and has little effect on the other coefficients.) It appears, therefore, that NCA captures elements of candidate evaluations that are largely unrelated to economic conditions, at least as reflected in the measure of income growth.

[3] The prediction for Bush is just slightly higher (46.6%) using coefficients from a vote model where the effects of approval and incumbency are estimated directly.

Mondale and even greater than Nixon's edge over McGovern. Only a marked advantage on the order of Eisenhower's over Stevenson in 1956 or Johnson's over Goldwater would tip the expectation to Bush (see Appendix for details).

The Strategic Approach

A

One conservative modification of our forecasting attempt is to produce an estimate from a second-stage equation that forecasts the vote from income growth and the value of net candidate advantage *predicted* from variables known in advance: August approval and incumbency. The equation is:

$$V = 45.85 + 2.30\,IG + 6.87\,\widehat{NxA} + e_t$$
$$(1.45)\quad (.64)\quad (1.91)$$
$$\text{Adjusted } R^2 = .841$$
$$\text{Standard Error of estimate} = 2.48$$

Using the value of the income growth and the estimated net candidate advantage for 1992, we still forecast a Bush loss, but with Bush garnering 46.8 percent. The slightly expanded standard error (2.48) technically makes the election too close to call. Still, the probability of a Bush victory is between one chance in ten and one chance in twenty.

Appendix: Raw Data

Year	Hibbs PCDI Growth	Net Candidate Advantage	2-Party Pres. Vote	Incumbency	Aug. Pres. Appr.
1948	2.58	.093	52.37	1	39
1952	1.80	-.488	44.59	0	32
1956	2.63	1.025	57.76	1	68
1960	1.03	.376	49.91	0	62
1964	4.72	1.031	61.34	1	74
1968	2.60	-.438	49.59	0	35
1972	3.14	.745	61.79	1	56
1976	1.37	.214	48.89	1	45
1980	-.78	.279	44.70	1	32
1984	3.51	.358	59.17	1	54
1988	2.10	.101	53.80	0	53
1992	-.03			1	37

References

Erikson, Robert S. 1989. "Economic Conditions and the Presidential Vote." *American Political Science Review* 83:567–573.

Hibbs, Douglas A., Jr. 1987. *The American Political Economy.* Cambridge: Harvard University Press.

Tufte, Edward R. 1978. *Political Control of the Economy.* Princeton: Princeton University Press.

Forecasting the 1992 Election

Walter Hill
St. Mary's College of Maryland

My modeling work is inspired by Lewis-Beck and Fair. The latter is generally criticized for his exclusion of variables measuring the impact of political events, hence the most recent report has him giving Bush a landslide with 55%. I adopt the traditional multiple regression model, $v = x'\beta$, where x is a vector of variables and β is a vector of parameters to be estimated.

Candidate variables are

i : incumbent party, 1=dem, -1=rep

g : Second quarter gross national product

gi : g*i

dper : Democratic incumbent (1=dem, 0=no incumbent, -1=rep)

trend : 0 in 1916, 1 in 1920, 2 in 1924 etc.

trendfdr : Trend through 1964, but fixed at 12 after dealignment in 1965

auh2o : 1 in 1964, 0 otherwise

dpsouth : 1 if dem candidate is from a state which fought for the Confederacy during the War of the Rebellion (1861–1865), 0.5 if he is from a border state, 0 otherwise

dvpsouth : Like dpsouth for the dem vice presidential candidate

nh : Vote for the President as a proportion of the top two vote getters in the New Hampshire primary

nhquit : = 0 if the President did not run in NH, i otherwise

nh3 : i*(3*nh -2) A measure of dissension or challenge within the ruling party

Several alternative specifications are estimated in Table 1. Model 1 is my preferred model. It turns out to have the best F statistic. This model forecasts 48% of the vote for the Democrat. The standard error of the estimate is 1.4%. The spread in the scattergram is about 2%, giving me ample opportunity to hedge on the victor!

I have a few comments on the analysis. Despite widespread attention to the choice of Vice-President, none of these models finds the choice of running mate significant. The New Hampshire primary enters when deleting the "trend" term. The model seems too sensitive to this variable, though Model 4 does predict a comfortable Democratic victory (because of a non-cohesive Republican party.) If one sets $nh = 0$, it is not unreasonable that $nhquit = 2 * nh3$. But inserting a trend effect does replace the New Hampsire primary effect. Finally, it is a little disgruntling that the range of these models is as large as 6 percentage points.

Table 1: Alternative Model Specifications

Variable	Model 1	Model 2	Model 3	Model 4
dper	.01922	.01093	.01685	ns
s.e.	.00047	.00504	.00702	
gi	.01246	.01215	.01127	.01215
s.e.	.00071	.00074	.00099	.00073
dpsouth	.04182	.04384	.05149	.05415
s.e.	.01117	.01198	.01650	.01125
dvpsouth	ns	ns	ns	ns
trend	.00307			
s.e.	.00007			
trendfdr		.003692		.00398
s.e.		.000968		.00099
auh2o	.06681	.05909	.07102	.04359
s.e.	.01722	.01882	.02605	.01923
nh3			ns	.02114
s.e.				.00519
nhquit			ns	.04189
				.01279
constant	.41911	.41821	.44723	.41763
F	98.808	84.443	51.761	74.118
	5, 13	5, 13	4, 14	6, 12
sigma	.01426	.01539	.02160	.01520
Predicted Vote	.48251	.4703	.4661	.5355

1992 Presidential Election Predictions

Michael S. Lewis-Beck
University of Iowa

Tom W. Rice
University of Vermont

A recent quip in the *Economist* proclaimed that:

> For the first time, the nominees of both big parties are left- handed. America has had only four left-handed presidents– James Garfield, Harry Truman, Gerald Ford and George Bush. The bad news for Mr. Bush: of the other three, none won twice. (July 25, 1992:26)

We trust that our forecasting model is based on more than coincidence, which underpins this "left-handed" rule. In any case, our model predicts a different winner. Below we summarize the components of the model, review its precision, and issue our presidential forecast for 1992 (for full details, see our *Forecasting Elections*, CQ Press, 1992).

Stated simply, our model holds that presidential election outcomes are in large part a function of issues (economic and non- economic), the relative strength of the incumbent party among the electorate, and the appeal of the incumbent party candidate. Each of these factors has been linked to electoral behavior by decades of social science research. Our task was to devise measures for them that would be available in advance of an election and produce accurate forecasts.

For a measure of the economy, we use the percentage change in real GNP from the fourth quarter in the year before the election to the second quarter of the election year. To capture the influence of other issues, we employ the President's approval rating in the July Gallup Poll. As a gauge of party strength, we settled on the number of seats in the House of Representatives that the President's party won or lost in the preceding midterm election. To tap the appeal of the incumbent, we use the percentage of the total primary vote won by the incumbent party nominee (scored 1 if the nominee received at least 60 percent of the vote, 0 if the nominee received less than 60 percent).

When we assembled these variables for all the presidential elections from 1948 to 1988 and entered them in a regression equation to predict the incumbent party's percentage share of the electoral vote, the OLS results were as follows:

$$\underset{(0.50)}{V= 6.83} + \underset{(3.79)}{7.76\,G} + \underset{(3.39)}{0.86\,PP} + \underset{(2.87)}{0.52\,PS} + \underset{(3.30)}{19.66\,C}$$

R-squared = .95
Adj. R-squared = .92
N = 11
SEE = 9.10
D-W = 2.34

where V is the incumbent party's percentage share of electoral vote; G is the percentage change in real GNP from the fourth quarter before the election to the second quarter of the election year; PP is the Gallup approval percentage for the president in July; PS is the number of House seats incumbent party lost in last midterm election; X is the percentage of total primary vote incumbent party candidate received, scored 1 if at least 60 percent and 0 if less than 60 percent. The values in parentheses are t-ratios.

Obviously the model performs well. All the independent variables are statistically significant at the .05 level and the R-square is extremely high. Data to forecast the 1992 race were available in late July. According to Gallup Poll, President Bush's approval ratings stood at 32 percent in July. The Republicans had a net loss of eight seats in the midterm House elections. Bush easily captured over 60 percent of his party's primary vote. And the real GDP grew 1.1 percent from the fourth quarter to the second quarter (real GNP, available a month later, also registered 1.1 percent non-annualized growth over the period). Inserting these numbers into our equations yields the following:

$$\begin{aligned} V &= 6.83 + 7.76(1.1) + 0.86(32) + 0.52(-8) + 19.66(1) \\ &= 6.83 + 8.54 + 27.52 - 4.16 + 19.66 \\ &= 58.39 \end{aligned}$$

where V is the incumbent party's share of electoral vote.

Our model, then, forecasts a narrow Bush victory, with 58.39 percent of the electoral vote. To get an idea how close of a race we are predicting, consider that Kennedy won about the same percentage of the electoral vote in his razor-thin 1960 victory. Another way to comprehend the closeness is to translate the electoral vote figure to a popular vote share. When we do this, using the formula below (derived from regression the popular vote share on the electoral vote share), we show Bush winning 51.51 percent of the popular vote.

$$\begin{aligned} PV &= 38.66 + 0.22V \\ &= 38.66 + 0.22(58.39) \\ &= 38.66 + 12.85 \\ &= 51.51 \end{aligned}$$

where PV is the incumbent party's share of popular vote; V is the incumbent party's share of electoral vote.

How certain can we be of a Bush victory? Well, across the 11 presidential elections since 1948 our model has an average prediction error of 5.63 percent of the electoral vote. Thus, if prediction error does not exceed average, Bush will win with between 52.76 and 64.02 percent of the electoral vote (58.39 + 5.63). Applying a more stringent 95 percent confidence interval to our equation reveals that the Bush total should fall between 40.19 and 76.59 percent of the electoral vote (58.39 + 2(9.10)). Translating this to the popular vote share yields an interval between 47.50 and 55.51 percent.

In short, this interval estimate admits the possibility of a Clinton victory. However, our point estimate predicts Bush. And, if Bush prevails, he will have broken the aforementioned "left-handed" rule of presidential election forecasting!

Predicting the 1992 Election

Lee Sigelman
George Washington University

The two models considered here are lineal descendants of the simple one I once used to test Mueller's (1973: 202) contention that Gallup's presidential popularity index "has little direct relevance to the electoral result" (Sigelman, 1979). For the seven elections, 1940-1976, in which an incumbent president ran for re-election, I uncovered a strong connection between the incumbent's standing in the final pre-election popularity poll (which was usually conducted during the summer, but sometimes several months earlier) and the percentage of popular votes he won in the November election. Soon, more elaborate versions of this model began to appear (e.g., Abramowitz, 1988; Brody and Sigelman, 1983; Lewis-Beck and Rice, 1992).

Model 1 updates the original model to include the two subsequent elections (1980 and 1984) in which an incumbent sought re-election. For these nine elections this model expresses the incumbent's percentage of the popular vote as a linear function of the percentage of the public that voiced approval of his performance as president in the final pre-election Gallup Poll.[1]

Of those surveyed in the July 1992 Gallup Poll, 32% expressed approval of the way George Bush was handling his job. According to Model 1, this should translate into only 46.6% for Bush of the votes cast on November 9. However, the simple model on which this prediction is based does not display a very precise fit to the data, so it seems appropriate to try an alternative specification.

Model 2 adds dummy variables (an apt term in this case, I think) for the 1972 McGovern and 1984 Mondale campaigns to represent the unusually inept campaigns those two Democratic challengers waged against Republican incumbents. The fit of Model 2 is substantially better than that of Model 1 and the coefficients for both dummy variables are sizeable, denoting a six- to seven-point swing toward the Republican incumbent when the Democratic challenger

[1] The 1948-1988 data on presidential popularity and the incumbent's share of the popular votes are from Lewis-Beck and Rice (1992). Data for 1940 and 1944 are from Sigelman (1979).

Table 1: Predictions of Bush's 1992 Popular Vote Share

	1	2
Intercept	34.46	33.83
s.e.	3.97	1.75
T-value	8.67	19.30
Presidential Popularity	0.38	0.36
s.e.	0.07	0.03
T-value	5.08	11.02
McGovern dummy		6.73
s.e.		1.55
T-value		4.33
Mondale dummy		6.27
s.e.		1.55
T-value		4.06
N	9	9
Adjusted R2	.76	.95
Standard error	3.29	1.45
Prediction	46.6	45.4
Alternate prediction		51.9

mounts an unusually weak campaign, as occurred in 1972 and 1984. As of Labor Day, the traditional start of the general election campaign season, it cannot be foretold how effective a campaign the Democrats will wage this time. If they mimic the poor performances of McGovern and Mondale, then it seems appropriate, based on the coefficients for the McGovern and Mondale dummy variables, to add 6.5% to Bush's predicted vote share in Model 2, bringing his percentage to 51.9%. That is, if the Democrats turn out to be as dismal as campaigners in 1992 as they were in 1972 and 1984, then the best bet is that Bush will win a narrow majority of popular votes; but that best bet is none too good, for even under this worst-case scenario for the Democrats the forecast is close enough to 50% for a good probability to remain of a Democratic majority in the popular vote column. On the other hand, if Clinton and Gore somehow manage to dodge the bullet that past Democratic standard- bearers have so accurately aimed at themselves, then according to Model 2 they will hold Bush to only 45.4% of the popular vote.

In sum, according to the simple models considered here, the key is whether the Clinton campaign will rival the proclivity for self-destruction that the McGovern and Mondale campaigns displayed in such abundance. That is, the simple incumbent popularity-based models project a Democratic edge in popular votes unless the Democrats revert to their form of 1972 and 1984. By this reckoning, the election is the Democrats' to lose. At this writing, it seems fair to say that it will take a real effort on the Democrats' part to lose it. Of course, there is ample evidence that they are capable of mounting just such an effort, but if they uncharacteristically fail to do so (and if the electoral vote split reflects the popular vote split, itself a dicey proposition), then the models considered here suggest that the Democrats will soon find themselves in the unusual position of controlling the White House.

References

Abramowitz, Alan L. 1988. An improved model for predicting presidential election outcomes. *PS* 21: 843-847.

Brody, Richard, and Lee Sigelman. 1983. Presidential popularity and presidential elections: An update and extension. *Public Opinion Quarterly* 47: 325-328.

Lewis-Beck, Michael S., and Tom W. Rice. 1992. *Forecasting Elections.* Washington, DC: Congressional Quarterly Press.

Mueller, John. 1973. *War, Presidents and Public Opinion.* New York: Wiley.

Sigelman, Lee. 1979. Presidential popularity and presidential elections. *Public Opinion Quarterly* 43: 532-534.

The 1992 Iowa Political Stock Markets: SeptemI er Forecasts

Robert Forsythe
Forrest Nelson
George R. Neumann
University of Iowa

Jack Wright
The George Washington University

Introduction

The Iowa Political Stock Market (IPSM) is an experimental securities market that aims to predict election outcomes and track the dynamics of campaigns. It was designed and first implemented in June of 1988 to forecast the U.S. presidential election, and since has been applied to a variety of electoral settings, both domestically and internationally. The 1992 IPSM is far more extensive and elaborate than any of its predecessors. Following a decision by the Commodities Futures Trading Commission in February of 1992 to allow the IPSM to operate experimentally nationwide, the 1992 IPSM has attracted 579 traders— 281 of whom are non-Iowans— and a total investment of H38466.

The 1992 IPSM opened for trading on January 10, 1992. Initially, two separate markets, a "nomination market" and a "presidential market", were available for trading. The nomination market, designed to predict the winner of the Democratic nomination campaign, closed at the end of the Democratic National Convention. The presidential market, however, continues to offer trading in shares of the major party candidates and to generate day-to-day predictions of the share of the two-party vote each candidate will receive. In response to strong interest in the third-party candidacy of H. Ross Perot, a "Perot market" was opened on May 19, and then on July 10, a "plurality market" was opened. This plurality market was designed to predict the presidential winner, as defined by the candidate receiving the most popular votes. We provide specific details about the operation of each of these markets, as well as a summary of how the IPSM operates generally, below. Then, we review market prices to date and discuss predictions for the 1992 campaign.

How the IPSM Works

The IPSM operated in 1988 and 1990 as a computerized, double-auction market, running 24-hours a day from the time it opened until it closed, the morning after the election. In 1988, the market operated on a mainframe computer at the University of Iowa, but since then it has operated on a network of personal computers with off-campus access. Individuals receive trading rights and computer access upon deposits to a cash account. This cash account provides the liquidity for stock transactions— purchases of shares are charged to it, and sales of shares credited to it. In general, participants acquire stock by purchasing individual shares of a security at the current market price or, for a fixed price, purchase a basic portfolio consisting of one share in each candidate on the slate. Conversions of cash account deposits into basic portfolios or basic portfolios into cash are allowed at any time, and withdrawals from cash accounts are allowed at periodic intervals throughout the campaign. New traders are allowed to enter the market at any time, and existing traders are allowed to add to their cash accounts at any time.

Shares are given value by the dividends paid after the election, with the dividend on each share determined as the candidate's fraction of the vote times the portfolio price. In 1992, dividends in the presidential market will be paid on the basis of a candidate's share of the two-party vote, and in the Perot market, dividends on Perot's stock will be based on his share of the popular vote. For the nomination and plurality markets, which are a winner-take-all, winning shares will earn exactly the portfolio price (one dollar) and losing shares will earn zero. Because the vote shares necessarily sum to one across all candidates in the presidential and Perot markets, the total dividend paid on a basic portfolio of one share in each candidate will just match the fee charged for that portfolio. This payoff rule provides a direct translation of market prices into estimates of vote shares as follows:

$$\text{Expected Vote Share} = \text{Price}/\text{Portfolio Price}.$$

In the cases of the nomination and plurality markets, prices reflect directly the probabilities of winning. Table 1 summarizes the essential aspects of the 1992 markets.

Participants trade in the double auction markets by issuing offers to buy (bids) or offers to sell (asks). There can be many bid and ask prices in the system at any time; they are maintained in bid and ask queues ordered first by offer price and then by time of issuance. When an offer is entered into the bid or ask queue, it remains there until: (a) it is withdrawn by the bidder, (b) it reaches the top of the bid queue (bottom of the ask queue) and is found to be infeasible as described below, or (c) it reaches the top of the bid queue (bottom of the ask queue) and is matched with an opposing offer. The actual transactions are executed by the system when it finds overlapping bid and ask prices in the respective queues.

The computerized market provides facilities for obtaining information about the trader's account and the market, as well as for issuing bids and offers. Available account information includes the number of shares held in each candidate, the balance in the cash account, a list of outstanding offers, and a list of transactions. Available market information also includes current high bid, low ask and last transaction prices, and a record of the previous day's activity including opening and closing bid and ask prices, the last transaction price, the average transaction price, and the number of shares traded in each security. As on most stock markets, information on the depth of the bid and ask queues is never revealed.

Short sales and purchases on margin are disallowed; offers to buy with insufficient funds in the cash account of the buyer or offers to sell when the seller's portfolio contained no shares in that candidate are ruled infeasible. Checks for feasibility are made only when an offer reaches the top of its queue (high bid or low ask). If an offer fails the feasibility check it is withdrawn from the queue. Thus the system accepts offers that are not feasible at the time of issuance. The intention here is to enrich the set of strategies a trader might adopt while at the same time preventing the market itself from becoming a net creditor. Since only the high bid and low ask prices are revealed to traders, the system also avoids giving false information; traders are guaranteed that at least one share is available at quoted prices.

When a new feasible bid enters with a price equal to or exceeding the current minimum price in the ask queue, the system records a trade at the ask price. Likewise, if a new ask is entered with a price equal to or less than the current maximum price in the bid queue, a trade is recorded at that bid price. Such trades are executed one share at a time, regardless of the number of shares bid or asked, with renewed checks for feasibility of both the bid and the ask after each one-share transaction. The recording of a transaction includes notes in the transaction logs of the two traders involved, a credit to the cash account of the seller, a debit to the cash account of the buyer, and a transfer of the share of stock from the seller's portfolio to that of the buyer. The principles followed for the execution of trades are: (a) offers to buy are processed "high-prices first"; (b) offers to sell are processed "low-prices first"; (c) in the case of ties (two offers at the same price), the earliest offer to arrive on the market is processed first; and (d) when an overlap between bid and ask prices is found, the trade is executed at the price of the older of these two offers.

Traders in the IPSM— in contrast to most national opinion polls— are not representative of the national electorate. But representativeness in this sense does not matter in a market where participants are rewarded for successfully anticipating the true election outcome, not their preferred outcome. Respondents in an opinion poll— even if queried about who would win and by how much— have little incentive to acquire good (i.e., costly) information about the eventual outcome; in contrast, a financial market provides investors with the incentives to seek out the relevant information about the outcome.

Table 1: 1992 Political Markets

Market	Securities	Portfolio Price	Payoff Rule	Dates
Nomination	Brown Clinton Harkin Kerrey Tsongas Rest-of-field	H1.0	Shares of nomination winner = H1.0 each Shares of all other candidates = H0.0	1/10-7/16
Presidential	Bush	H1.0	1 share Bush stock = (Bush's share of two-party vote) (H1.0)	1/10-11/3
	Clinton		1 share Clinton stock = (Clinton's share of two party vote) (H1.0)	
Perot	Perot	H1.0	1 share Perot stock = (Perot's share of popular vote) (H1.0)	5/19-11/3
	D & R		1 share D&R stock = (sum of Dem. and Rep. share of popular vote) (H1.0)	
Plurality	Bush	H1.0	Shares of candidate receiving most popular votes = H1.0 each	7/10-11/3
	Clinton		Shares of all other candidates = H0.0	

Figures go here

A unique feature of the IPSM as a forecasting device is that its predictions are dynamic, responding throughout the course of the campaign to new information about the election outcome. Any unanticipated rise or fall in the fortunes of the candidates are clearly marked by daily prices. For the 1992 Democratic nomination campaign, fluctuations in the daily price of Clinton's stock are plotted in Figure 1. Since the nomination market was winner-take-all with a portfolio price of one dollar, Clinton's price reflects his probability of winning the nomination.

After a brief period of instability early in the market, Clinton's price climbed to 68 cents by January 23, but then began to fall as information surfaced about his alleged affair with Gennifer Flowers. The largest single drop in price occurred on January 25, the day before Clinton appeared on CBS's 60 Minutes, and his price continued to fall for the following two days before turning around. Clinton soon recovered, though, and his price increased steadily for roughly two weeks until it reached 56 cents on February 8 and 9, at which time news of his efforts to avoid the Vietnam draft led to a sharp downturn. His stock then fell for five consecutive days, finally bottoming out at 31 cents on February 14. Then, on February 18, the day of the New Hampshire primary, Clinton's stock jumped dramatically by 10 cents. Clinton's better-than-expected performance in the nation's first primary resulted in a steady climb in the value of his stock over the next three weeks of the campaign. Finally, for all practical purposes, he locked up the nomination on Super Tuesday, March 10, when his price jumped a full nine cents to a campaign high of H.90

As Clinton struggled through the nomination campaign, George Bush's expected share of the two-party vote in November grew steadily. June 30 marks the high point in Bush's November fortunes at 58 percent. But by the end of the following week, clearly one of the most critical of the campaign, Bush's value plunged to 51 cents. The first week of July brought bad economic news for the Bush Campaign. On July 3, the Labor Department reported that unemployment had risen to 7.8 percent, and on the same day, the Federal Reserve announced that it was lowering the discount rate to three percent, a sure sign that the economy was still sluggish. Also significant during the first week of July was the announcement that Mario Cuomo, who earlier in the campaign had been critical of Clinton, would deliver the nominating speech at the Democratic convention. The news about Democratic unity and the economy, combined with earlier events in June— the generally favorable reception of Clinton's economic proposal, contrasted with the sharp attacks and counterattacks of the Bush and Perot campaigns over alleged surveillance activities— were sufficient to reverse Bush's momentum and significantly narrow the gap between the two candidates.

It was not until early August, however, that Clinton finally surpassed Bush, in both the presidential market and in the plurality market. Bush's price in the plurality market is plotted against days in Figure 3, where a sharp decline in price begins on August 3 and continues through August 15, just two days before the Republican Convention. This trend is also reflected, but not so dramatically, in the presidential market. A flurry of distinct events early in this period help account for the additional problems of the Bush campaign. On August 1, the Orange County Register called for Bush's withdrawal from the campaign, and the following day, conservative activist Richard Viguerie added his voice to those clamoring for a new Republican ticket. Then, on August 3, deputy campaign manager Mary Matalin released her derisive memorandum on the Clinton campaign's efforts to control "bimbo eruptions," but found herself apologizing the following day as the White House disavowed the attack. Finally, on August 8, Clinton called for Bush to act more forcefully on the Balkans, a position that appeared to have widespread popular support. Together, these events triggered a decline in Bush's stock price that did not stabilize until the Republican convention.

As of Labor Day, the traditional start of the Fall campaign, the IPSM markets predict a Clinton victory, as summarized in Table 2. In the presidential market, Clinton's predicted share of the two-party vote is 52.9 percent, and his overall probability of winning a larger share of the total vote is .60. H. Ross Perot, however, is predicted to receive 9.4 percent of the total popular vote, making the final predictions of the popular vote shares: Clinton, 49.5 percent; Bush, 44.1 percent; and Perot, 9.4 percent. These prices are subject to change, though, as unanticipated events unfold over the remainder of the 1992 campaign.

References

Forsythe, Robert, Forrest Nelson, George R. Neumann, and Jack Wright. 1991a. "Forecasting the 1988 Presidential Election: A Field Experiment" in R. M. Isaac, ed., *Research in Experimental Economics*, vol. 4: 1-44 .

Forsythe, Robert, Forrest Nelson, George R. Neumann, and Jack Wright. 1991b. "The Explanation and Prediction of Presidential Elections: A Market Alternative to Polls," in *Contemporary Laboratory Experiments in Political Economy*, T.R. Palfrey, ed., Ann Arbor: University of Michigan Press.

Forsythe, Robert, Forrest Nelson, George R. Neumann, and Jack Wright. 1992. "Anatomy of a Market," *American Economic Review*, forthcoming, December.

Table 2: IPSM Forecasts of the 1992 Presidential Election, September 7, 1992

Market	Prediction	
Presidential		
Share of Two-Party Vote	Bush	47.1%
	Clinton	52.9%
Perot		
Share of Popular Vote	Perot	9.4%
	Bush + Clinton	93.6%
Plurality		
Probability of Winning	Bush	.38
	Clinton	.60
	Others	.02

We Should Be Modest: Forecasting the 1992 Presidential Election

Nathaniel Beck
University of California, San Diego

Introduction

As I write this late September of 1992, most commentators find George Bush well behind in his drive for reelection; polls consistently have Bill Clinton leading by at least ten points. This does not mean that Bush will not be reelected, but only that he is not doing well in the month following the Republican Convention. But academic forecasting models tell a different story. The two most popular models, those of Fair (1978) and Lewis-Beck and Rice (1992) have Bush well ahead. The Fair model predicts that Bush will get about 56% of the popular vote; the Lewis-Beck and Rice model predicts that he will get 58% of the *electoral vote*.[1]

It is easy to understand why Fair would have Bush ahead. His model forecasts popular vote (as a percentage of the two party vote) based on incumbency and the pre-election economy. His prediction equation (Fair, 1990) for 1992 (based on 19 elections from 1916 through 1988) is (in rounded form)

$$\text{Bush Pop. Vote} = 56\% + \text{Real GNP Growth (Per Cap)}$$

[1] Lewis-Beck announced that he stood by his forecast at a round table discussion at the American Political Science Association's Annual Meeting in early September.

$$-\frac{\text{Inflation}}{3}$$

where the growth in GNP Per Capita is measured in the first half of 1992 and the inflation rate is from July, 1990 to June, 1992. Real economic growth (per capita, annualized) is about 1% and inflation is about 3%, yielding the prediction.

Fair's prediction uses no information about the candidates or the campaign; it can be read as saying that, all other factors being equal, a Republican incumbent running in an economy showing no real growth but little inflation should win fairly easily. This is a good description of our state of knowledge well before the beginning of the 1992 campaign. The Fair model does not take into account current information about the campaign and candidates. In many elections these other factors will net out; in September, it appears that the other factors will tip towards Clinton, but September isn't November.

The Lewis-Beck and Rice (hereafter 'LBR')forecast is more surprising, since their forecast uses Bush's popularity in June and a measure of Bush's 'candidate appeal,' as well as GNP growth. Their model should be able to take into account the Bush weaknesses that others see, yet, according to Tom Rice, "what we're saying is that given the long-term trends and the relationship of these key factors to election outcomes, our best guess is that Bush is going to win this election" (Shogan, 1992).

LBR forecast the Bush proportion of the total electoral vote. Their forecasting model (Lewis-Beck and Rice, 1992), based on 11 elections from 1948 through 1988, is

$$\begin{aligned}\text{Bush Electoral Vote} &= 6.8\% + 7.8 * \text{Real GNP Growth}\\ &\quad +0.9 * \text{Pop} + 0.5 * \text{PartyStr}\\ &\quad +19.7 * \text{CandAppeal}\end{aligned}$$

where Real GNP Growth (not per capita, not annualized) is measured in the first half of 1992 (just under 1%), Pop is Gallup Poll approval measured in June, 1992 (33%), PartyStr is the Republican gain in House Seats in the 1990 Election (-8) and CandAppeal is a dummy variable measuring Bush's performance in the primaries (1). These figures yield a forecast that Bush will get slightly more than 58% of the *electoral vote.*

It is hard to interpret the LBR forecast in terms of electoral votes. Is 58% of the electoral vote a landslide or a close election? Electoral votes are obviously more dispersed than are popular votes. A candidate getting 60% of the popular vote will get almost all the electoral votes. Since LBR's theory is in terms of popular vote, I can estimate a model identical to the LBR model, but predicting the incumbent's percentage of the two party vote, rather than his electoral vote.[2] Results of this estimation are in Table 1.

The resulting forecast model for 1992 is

$$\begin{aligned}\text{Bush Popular Vote} &= 42.8\% + 1.6 * \text{Real GNP Growth}\\ &\quad +0.2 * \text{Pop} + 0.1 * \text{PartyStr}\\ &\quad +2.4 * \text{CandAppeal.}\end{aligned}$$

This model forecasts that Bush will receive just under 52% of the two party popular vote, that is, a good but not overwhelming victory for the Republican Party.

The Fair and LBR models have been used to forecast elections over the last decades. The LBR model for popular vote produces forecasts that are comparable to those of Fair. How well have these models predicted past elections? Which one is superior? This issue is dealt with in the next section.

Once these questions are answered, we can turn to 1992. Both Fair and LBR have provided forecasts for 1992. These are point forecasts, that is, the single number that best represents their prediction about the outcome of the 1992 election. But all forecasts have some uncertainty, and so forecasters should provide a range of possible outcomes which they feel are likely. This range is computed in the third section.

Finally, why are the forecasts so at odds with what appears to be the conventional wisdom in September 1992? In particular, why is the LBR forecast, which takes into account information about the incumbent as a candidate, so favorable to Bush? In the fourth section we shall see that the culprit is their measure of Candidate Appeal. Ho

Both Fair and LBR report that their models do quite well. Fair reports that his predictions were off by just above one point for the last six elections; the corresponding figure for LBR, is 1.3 points.[3] Based on these figures, it appears as though Bush should be quite confident of his reelection.

But this would be a false confidence. The errors reported above are based on using all information available in 1992 to make 'forecasts' for the earlier years. We should not be so impressed with a model that can 'forecast' well for 1988 when the 1988 election is almost 10% of the data used to fit the model. The true test of a forecasting model is how well

[2] LBR argue that one should use electoral vote percentages since we are interested in forecasting who wins. They note that in three of the 11 most recent elections the winner received less than half the *total* popular vote. They also argue that there is a strong linear relationship between popular and electoral vote, so that they can use ideas derived from studying the individual vote decision. Using the percentage of the two party popular vote, as Fair does, avoids the problem of winners receiving less than half the vote. This also avoids adding the uncertainty of translating popular vote into electoral vote, a process that is hard to model. If I were interested in predicting electoral votes, I would use state level data, as in the forecasting models of Rosenstone (1983) and Campbell (1992).

[3] LBR are off by about 5 points in their prediction of electoral vote for the last six elections. Remember that electoral vote is much more dispersed than is the popular vote.

it forecasts elections based on the information that would have been available before that forecast was made.

Table 2 reports the true forecast errors for both models. The 1976 forecast error, for example, is computed by estimating the model with data through 1972 and then using those estimates to compute a true forecast for 1976, with the error being the difference between the Democratic percentage of the two party vote and the forecast value.[4]

Not surprisingly, the true forecast errors for both models are much larger than the errors reported by Fair and LBR. For the last six elections, Fair's average forecast error was just under 2 points.[5] The LBR average forecast error (for popular vote) for those six elections is 3.5%. *In four of the last six elections, the LBR forecast error exceeded their 1992 predicted margin of victory for Bush.*[6]

It is clear that the Fair model forecasts election outcomes better than the LBR model, at least through 1988. This is surprising, given that LBR use more current information. The superiority of the Fair forecasts is not due to his using a larger number of elections, since the method used here to compare the two forecasts gives no necessary advantage to forecasts based on a large amount of data. The one variable in the Fair model that is not in the LBR model is whether the incumbent is running for reelection. Including that variable in the LBR model allows it to forecast as well or better than the Fair model. Since my interest is in evaluating the published models, I do not pursue this modification further.

The large forecast errors for the LBR model mean that Bush should take little reassurance from the LBR forecast of his victory. Fair's forecast that Bush will get 56% of the popular vote might appear more heartening to Bush; the predicted margin of victory exceeds the Fair model's typical forecast error for the last six elections. But even the Fair forecast should not make Bush too sanguine. In the three elections contested during the 1960's, Fair's *minimum* forecast error was just under 5%. And had Fair not fudged in 1976 (see note 5) he would have over-predicted the Ford vote by almost 6%. So it surely would not be unprecedented if Bush were to lose in spite of the margin of victory predicted by Fair. The next section makes this notion more precise.

[4] The LBR model forecasts the incumbent party's vote. All re-estimations of the LBR model used that convention. The LBR forecasts were then transformed to make them consistent with Fair's.

[5] This average is based on Fair's (1988) acknowledged 'fudge' for 1976. After observing the election, he decided not to count Ford as an incumbent running for reelection, reducing his (true) forecast for Ford by 3.8-. Had he not fudged, his average prediction error would have been half a point higher. He still would have outperformed LBR.

[6] The corresponding figures for electoral vote are an average error of almost 20 points, with four of the last six elections having an error exceeding Bush's eight point predicted victory in the Electoral College.Confidence

Forecasts based on econometric models have some uncertainty. There is obviously uncertainty due to lack of knowledge about whether the environment has changed, whether the model is Vcorect,' whether the model has been honed by Vdatamining,' to say nothing of whether errors follow a normal distribution with such small samples. But for present purposes I assume away all these sources of uncertainty to simply focus on uncertainty in the parameter estimates and residual uncertainty of the model (the Verro' term in the regression). As we shall see, these two sources of uncertainty lead to huge uncertainty about the vote forecasts; but even this huge uncertainty is an underestimate of true uncertainty about the forecasts.

The usual measure of forecast uncertainty, based only on uncertainty about the true parameters of the model and residual model error, is the standard error of the forecast. After this statistic is computed, a 95% confidence interval for the forecast can then be constructed. We would be willing to make a bet at 19:1 odds that the election outcome will lie in the 95% confidence interval. In practice, we take the confidence interval as giving the range of forecast values that we would be willing to entertain.

The well known formula for the standard error of a forecast is

$$\widehat{\sigma_f} = \hat{\sigma}\sqrt{1 + x_f(X'X)^{-1}x_f\prime}$$

where X is the matrix of observations on the independent variables, x_f is the vector of independent variables for the forecast and $\hat{\sigma}$ is the usual unbiased estimate of σ. A 95% confidence interval for a forecast is

$$\widehat{Y_f} \pm\ t_{.025}\widehat{\sigma_f}$$

where $\hat{Y_f}$ is the forecast value of Y and the t statistic has (n-k) degrees of freedom, with n being the number of observations and k the number of independent variables in the regression.[7]

[7] For small sample sizes the assumption that the forecasts follow a t distribution depends on the normality of the error term in the regression. This means that, in practice, we will be underestimating the true width of the confidence interval if, as is likely, the error process is longer tailed than the normal.

Table 1:

Table 1: Estimates for LBR Forecasting Models, 1948–1988[a]

	I: Electoral Vote		II: Popular Vote		III: Popular Vote	
Variable	b	se	b	se	b	se
Constant	6.8	13.6	42.8	3.1	41.8	3.1
GNP(Growth)	7.8	2.0	1.6	0.5	1.6	.5
Pop	0.86	0.25	0.18	0.06	0.19	0.06
PartyStr	0.52	0.18	0.10	0.04	0.10	0.04
CandAppeal	19.7	6.0	2.4	1.3		
CandAppeal(%)[b]					0.035	0.019
$\hat{\sigma}$	9.1		2.1		2.0	

[a]n=11

[b]Percentage of primary vote received

Table 2:

Table 2: Forecast Errors of Fair and LBR Models[a]

	Actual	Fair	LBR (Popular)	LBR (Electoral)
1948	52.4	-1.2		
1952	44.6	-2.4		
1956	42.2	-2.2		
1960	50.1	5.2		
1964	61.3	11.1		
1968	49.6	-4.7	7.0	47.1
1972	38.2	-2.7	5.1	-31.7
1976	51.1	2.0	0.5	9.6
1980	44.7	-0.2	3.3	7.1
1984	40.8	-1.4	0.0	13.1
1988	46.1	-0.9	5.0	-5.6

[a]Differences between Democratic Percentage of Vote and Predicted Value

The 95% confidence interval for Fair's forecast is (48%,63%), that is, we can be quite sure that the 1992 election will be somewhere between a slight Bush loss and a Bush victory by an unprecedented landslide. The corresponding interval for the LBR forecast (of popular vote) is (45%,58%), that is, we can be quite sure that the outcome will be between a Bush win by a near landslide and a Bush loss by a wide margin. (The interval for the LBR electoral vote forecast is (29%,86%), again not very helpful.) When we take into account all the other sources of uncertainty that are not included in the standard calculations, we see that the LBR forecast is of little value, and the Fair forecast would allow for either Bush or Clinton to win the election. Neither campaign would pay a lot for such forecasts.[8]

It is important to use the correct formula for the confidence interval of a forecast. Many practitioners, including all the participants at the APSA Round Table on Forecasting the 1992 Election, incorrectly compute this confidence interval as $\pm\ 2\hat{\sigma}$. This leads to an underestimate of our uncertainty (for the LBR model) of about 50%. Half of this underestimate is due to the correct value of t being about 2.5 instead of 2, given the small number of degrees of freedom in the LBR model; the other half is due to ignoring the square root term in Equation 1. We know that forecasts become more and more inaccurate as the independent variables for the forecast move further and further from the mean; it is the square root term that measures this increase in uncertainty. In the LBR case, the independent variables for 1992 (especially presidential approval) are quite far from their historical means, and hence this term is of considerable importance. Should we really forecast Bush in the lead

The LBR model uses both Bush's June approval and his performance in the primaries to predict the outcome. Bush's approval was unprecedentedly low and Bush clearly had difficulties in the primary season. Why do LBR still forecast him ahead? The answer lies in the Candidate Appeal variable. This is a dummy variable, which is scored as one if the candidate of the incumbent party receives over 60% of the primary vote. Bush is measured as having a score of one on the Candidate Appeal variable using this measurement strategy. He is thus classed with Nixon and Reagan rather than Ford and Carter. The coefficient on Candidate Appeal in Model II is about 2.5 points. *It is this scoring of Bush on the Candidate Appeal variable that leads to the forecast of his victory.*

One solution might be to simply classify Bush with Ford and Carter. This would decrease the point forecast for Bush's electoral vote to 38%, while his popular vote forecast would decrease to just under half. This procedure seems a bit *ad hoc*; it cannot be right to use measurement to make a forecast seem reasonable.[9]

But there is a better alternative here. LBR report the actual percentage of the primary vote received by the incumbent. If I use that measure instead of the dummy variable version, I get the results shown in Table 1, Model III. Since there is little, if any, reason to throw away information by dichotomizing, Model III should be superior to Model II; Table 1 shows that it is.

Now suppose Bush is scored at 70% on this measure, which is consistent with his performance in the primaries. This leads to a forecast for Bush of 51% ($\pm 7\%$). The corresponding forecast for electoral votes is also 51% ($\pm 30\%$). In other words, *the best forecast for 1992 is that the race is a toss-up, with results ranging from an impressive Bush victory to a major defeat not out of the question.*[10]

Fair does not have any campaign variables in his model. We can think of Fair's forecast as saying that a normal Republican incumbent should win relatively easily in 1992. But forecasters might then ask whether Bush is a normal Republican incumbent. Answering no, we might then adjust Fair's forecast of the Bush vote downward. Given the large confidence intervals for the forecasts, we might conclude that a Bush reelection is far from assured.

The Fair model does have a variable about whether an incumbent president is running for reelection. This factor increases the Bush margin by over 4 points. Fair has already fudged on this variable for Ford (see note 5). If we judge that Bush does not enjoy an incumbency advantage, then the Fair model forecasts a Bush victory with 52% of the popular vote. A confidence interval for this forecast margin is (44%,58%). Incumbents have done very well in most of the 19 elections studied by Fair. If we think that Bush should not be classed, using recent examples, with Eisenhower, Johnson, Nixon and Reagan, then the seemingly odd Fair forecast of a strong Bush victory turns into a more reasonable forecast of a very close election. In short,

[8] To see just how uncertain the forecasts are, Fair's results would allow us to take a bet at *even* odds that Bush will get between 53- and 58- of the vote; LBR would allow us to take a bet at even odds that Bush will achieve an outcome somewhere between a very close loss and a landslide victory (50- to 67- of the electoral vote). No statistician recommends that we use this very optimistic measure of uncertainty.

[9] Such a strategy is not unknown. In the forecasting literature, it is called 'judgmental' forecasting. Note 5 is one use of judgmental scoring. It should be noted that the measurement of Candidate Appeal cries out for some judgment. As just one example, LBR score Johnson as zero on this variable for 1964. This is because he received only 17- of the primary vote. This figure is totally misleading. The 1964 nomination was not contested, and Johnson's name appeared on the ballot in only two states. Most (70-) of the total primary vote went to unpledged slates and favorite sons. Almost all of the unpledged vote was in California, which saw two unpledged slates in opposition; California cast about two fifths of the total primary vote. There is no question that Johnson should have been scored one on the Candidate Appeal Measure.

[10] This conclusion holds for a wide range of scores for Bush's candidate appeal on the continuous measure.

Bush cannot take any comfort from either the Fair or LBR models.Conclusion

It is clearly good that we put our theories to the test by making forecasts. Fair, Lewis-Beck and Rice have been in the forefront of this effort. But it is critical to report confidence intervals as well as point predictions. The confidence intervals on the Fair and LBR forecasts are huge. There is little that can be done about this so long as we restrict ourselves to national level data. LBR have only 11 post-war elections, and cannot create more very quickly. We should be modest about our forecasts when they are based on such a modest amount of data. Alternatively, we should seek out other strategies to increase the amount of available data. The state level models of Rosenstone (1983) and Campbell (1992) obviously make sense from this perspective.

Do the forecasts overstate Bush's prospects? I think the answer is yes. The culprit for LBR is their measurement of Bush's candidate appeal. Moving to LBR's alternative continuous measurement of that variable shows that the race is a toss-up, not a clear Bush victory. The culprit for Fair is the historical advantage enjoyed by incumbent presidents seeking reelection. Removing this advantage for Bush would lead the Fair model to predict a close Bush victory. But the critical message is that the national level data leave us very uncertain about any point forecasts, and that forecasters must report confidence intervals as well as point forecasts. The message is in those intervals, and the message from both the Fair and LBR models is that we really do not know who will win the 1992 election.Refere nces

Campbell, J. 1992. "Forecasting the presidential vote in the states." *American Journal of Political Science* 36:386–407.

Fair, R. 1978. "The effect of economic events on votes for the president." *Review of Economics and Statistics* 60:159–73.

Fair, R. 1988. "The effect of economic events on votes for the president: 1984 update." *Political Behavior* 10:168–79.

Fair, R. 1990. "The effect of economic events on votes for the president: 1988 update." Cowles Foundation, Yale University.

Lewis-Beck, M., and T. Rice. 1992. *Forecasting Elections.* Washington, D.C.: Congressional Quarterly Press.

Rosenstone, S. 1983. *Forecasting Presidential Elections.* New Haven: Yale University Press.

Shogan, R. 1992. "Scholars disdain the polls, point to Bush trump cards." *Los Angeles Times* August 3, 1992:1,13–14.

Exogeneity, Inference, and Granger Causality: Part 1, The Stationary Case

Jim Granato
Michigan State University

Renée M. Smith
University of Rochester

This is the first of a two-part discussion on the use of Granger causality to determine the validity of exogeneity assumptions. In Part I, we assume that the variables under consideration are stationary, with constant means and finite variances; in Part II, to appear in the next issue of *TPM*, we will discuss these tests for the case of nonstationary, integrated regressors. The results below will be familiar to time series analysts, but we introduce them for purposes of comparison with the nonstationary case.

Beginning with Freeman's (1983) discussion of Granger causality, time series analysts in political science have relied more frequently on these statistical tests to investigate the validity of a priori "exogeneity" assumptions and to justify claims about causal orderings (e.g.; Alt (1985), MacKuen, Erikson, and Stimson (1989; 1992)). Sometimes these tests have been employed in less restrictive vector autoregression (VAR) applications (Williams (1990)). According to Freeman (1983: 329), Granger causality tests "offer qualitative characterizations of the relationships under study." In addition, Freeman (1983: 329) argues that Granger causality tests are important precursors to obtaining consistent coefficient estimates and drawing valid inferences from statistical models because "if variables are not, in fact, exogenous, the conditions for identification will be misstated and parameter estimates will not be consistent." However, as we discuss below, Granger causality tests do not provide unambiguous confirmation of such "exogeneity" assumptions (Cooley and LeRoy (1985)). Furthermore, they shed no light on the conditions that lead to parameter consistency. To make these points clearer, we begin by discussing various definitions of exogeneity.[1]

[1] Throughout this paper, we draw from examples and discussions in Cooley and LeRow (1985) and Engle, Hendry and Richard (1983).

In the 1940s and 1950s, the Cowles Commission[2] discussed issues of identification and exogeneity by distinguishing between variables that were predetermined and those that were strictly exogenous within the context of a particular structural model or system of simultaneous equations. According to the commission's definitions, a variable is predetermined if it is independent of the contemporaneous and future disturbances in the equation in which it appears, while a variable is strictly exogenous if it is independent of the contemporaneous, future, and past errors in an equation. In static models or in models with non-autocorrelated errors, this distinction is of no consequence because there are no effects of history to worry about.

More recently, Engle, Hendry, and Richard (1983) distinguish between weak and strong exogeneity.[3] According to these authors, weak exogeneity implies that we can draw inferences from a single equation model provided that (a) the parameter(s) of interest rest solely in that model and (b) there is an absence of cross-equation restrictions between a (marginal) model for the process generating a regressor, say X_t, and the structural (conditional) model of interest for the endogenous variable, Y_t.[4] When these restrictions are absent, the likelihood functions for the marginal and structural (conditional) models (i.e.; their densities) become independent.[5] In the context of the specific parameters of interest, this independence implies that knowledge of a parameter value in the marginal model provides no information about the range of values for parameters in the structural (conditional) model. Hence, no information is lost if the data-generating process for X_t is ignored during estimation of the structural (conditional) model for Y_t.

In linear models, weak exogeneity is closely associated with the concept of predeterminedness since, under certain circumstances, both imply that a regressor is independent of contemporaneous and future values of the error process in the structural (conditional) model of interest. However, one difference between the two definitions is that weak exogeneity is always defined relative to a particular parameter or set of parameters of interest.

For instance, consider the following two equation model:

$$Y_t = \beta X_t + \epsilon_{1t} \quad (1)$$

$$X_t = \tau Y_{t-1} + \epsilon_{2t} \quad (2)$$

where the standard classical linear assumptions hold and $\mathrm{E}(\epsilon_{1t}\epsilon_{2t}) = 0$. In this case, Equation 1 represents the structural (conditional) model from which we wish to draw inferences, while Equation 2 represents a marginal model for the process generating the regressor X_t. Now suppose we are interested in investigating the stability of our structural (conditional) model in Equation 1. We begin by specifying the reduced from for Y_t as:

$$Y_t = \Gamma Y_{t-1} + \xi_t \quad (3)$$

where $\Gamma = \beta\tau$. By estimating Γ in this reduced form, we can learn about the stability of the model in Equation 1; that is, Γ is the parameter of interest when questions of stability arise. Notice, however, that estimation of Γ requires knowledge of both β and τ. Therefore, ignoring the marginal model (i.e., density) for X_t results in a loss of information during estimation of Γ. Furthermore, the parameters, β, τ, and Γ in these models are not variation free since the stability of Equation 1 requires that $\mid \Gamma \mid < 1$ and more importantly, that $\mid \beta\tau \mid < 1$.

In short, X_t is NOT weakly exogenous with respect to the parameter Γ despite the fact that X_t IS predetermined for the system (i.e.; $\mathrm{E}(X_t\epsilon_{1t}) = 0$). Thus, we can see that in some cases, there is an advantage to using the definition of weak exogeneity rather than that of predeterminedness, although in other instances these definitions will coincide.

Indeed, the conditions of predeterminedness and weak exogeneity coincide if we are interested in conditioning on Equation 1 alone. Since we already know that 1 is predetermined when can we say it is also weakly exogenous? In 1 make β the parameter of interest. Specifically, is X_t weakly exogenous for this particular parameter? Based on the conditions above, we can see that the parameter β is solely in the structural (conditional) model 1. In addition, the parameters in the marginal model 2 provide no information concerning the range of values that β can take. The variable, X_t, is therefore weakly exogenous for this particular parameter of interest.

---

[2] See Koopmans (1950) for example.

[3] Engle, Hendry and Richard (1983) also define a third category— super exogeneity— which is not relevant for our purposes here. Super exogeneity conditions are useful when analysts are interested in policy simulations.

[4] Throughout we make an attempt to bridge the differences in semantics between the various authors. In fact, this may be a reason for the existing conceptual confusion. The word structural and conditional model refer to the particular equation or model that contains the dependent variable of interest. The marginal model defines the "omitted" equations or models that generate the regressors in the structural (conditional) model. Ideally, we would like to restrict our attention to the conditional (structural) model without violating basic estimation criteria (e.g. consistency).

[5] Fenerally speaking all statistical models that are of reduced form, can be broken down—read factorized— into joint probability distributions. In the case above we can construct a joint density function for X and Y. In addition, this joint density function can be partitioned into a conditional distribution and a marginal distribution (see Spanos (1986)). Consider the following:

$$D(Y, X) = D(Y \mid X)D(X)$$

or alternatively

$$D(Y, X) = D(X \mid Y)D(Y)$$

The issue here that Engle, Hendry and Richard (1983) nicely isolate is whether we can safely ignore the marginal density functions ($D(X)$ or $D(Y)$) and focus our attention on the parameters that interest us in the conditional density or single equation represented by $D(Y \mid X)$ and $D(X \mid Y)$.

In addition to their discussion of weak exogeneity, Engle, Hendry, and Richard (1983) argue that a variable is strongly exogenous to a particular parameter if it is *both* weakly exogenous and is not "Granger caused" by the endogenous variable. In linear models, strong exogeneity is equivalent to the Cowles definition of strict exogeneity (Cooley and LeRoy (1985)); that is, both imply that a regressor is independent of the contemporaneous, future, and past values of the disturbances in the model of interest. Notice that strong (hereafter strict) exogeneity requires that the endogenous variable, Y_t, does not Granger cause the regressor, X_t. Given weak exogeneity, a finding of Granger causality from Y_t to X_t can provide evidence against the assumption that X_t is strictly exogenous. However, as Cooley and LeRoy (1985) note, findings of Granger non-causality are only consistent with the existence of strict exogeneity, but do not unambiguously confirm such an assumption.

Engle, Hendry, and Richard (1983) argue that the relevant exogeneity assumption for valid inference is that of weak exogeneity since this implies that the parameters of the marginal and structural (conditional) models are variation free. This in turn implies that the parameters in the structural (conditional) model for Y_t will be invariant to structural changes or interventions in the data-generating process for X_t. Therefore, whenever weak exogeneity holds, the correlation between the regressor and disturbance converges in probability to zero, so that parameter estimates are consistent. Hence, if the purpose of estimating a structural (conditional) model is solely to draw inferences from the model's parameters, then the weak exogeneity of X_t is sufficient. However, if interest lies in forecasting Y_t conditional on X_t, then an analyst needs strict exogeneity to hold, so that feedback from Y_t to X_t is ruled out. Engle, Hendry, and Richard (1983: 286) emphasize this point:

> ...[I]t is misleading to emphasize Granger noncausality when discussing exogeneity. The two concepts serve different purposes: weak exogeneity validates conducting inference conditional on G while Granger noncausality validates forecasting G and then forecasting Y conditional on the future G's. As is well known, the condition that Y does not Granger cause G is neither necessary nor sufficient for the weak exogeneity of G. Obviously, if estimation is required before conditional predictions are made, then strong exogeneity which covers both Granger noncausality and weak exogeneity becomes the relevant concept.

Or using the Cowles Commission terminology, Cooley and LeRoy (1985: 297) note:

> Plainly Granger non-causality is neither necessary nor sufficient for predeterminedness ... Since predeterminedness is the exogeneity concept relevant for the analysis of interventions, it follows that the Granger and Sims tests are irrelevant to whether a causal interpretation of a conditional correlation is justified. Further, predeterminedness is also the exogeneity concept relevant for econometric estimation, implying that the Granger and Sims tests are equally irrelevant to the question of whether a model is consistently estimated.

Note, too, that the remarks of these authors are made within the context of stationary regressors.

To show the intuition behind the statements above, consider a situation where we have two variables Y_t and X_t. In this case, X_t Granger causes Y_t if Y_t is better predicted by including past values of X_t than by predicting Y_t exclusively on the basis of its own past values; that is, X_t Granger causes Y_t if including lagged values of X_t reduces the variance of the disturbance in the autoregressive representation of Y_t. To see why Granger causality is not relevant for testing assumptions of weak exogeneity or predeterminedness, suppose that Y_t and X_t can be represented as follows:[6]

$$Y_t = \alpha X_t + \delta_{11} Y_{t-1} + \delta_{12} X_{t-1} + \mu_{1t} \quad (4)$$
$$X_t = \Theta Y_t + \delta_{21} Y_{t-1} + \delta_{22} X_{t-1} + \mu_{2t} \quad (5)$$

Now consider the reduced form for the system of equations above:

$$Y_t = \Pi_{11} Y_{t-1} + \Pi_{12} X_{t-1} + \eta_{1t} \quad (6)$$
$$X_t = \Pi_{21} Y_{t-1} + \Pi_{22} X_{t-1} + \eta_{2t} \quad (7)$$

Solving for Y_t and X_t in 4 and 5 and breaking down the reduced form further in 6 and 7 yields:

$$Y_t = \frac{\alpha\delta_{21} + \delta_{11}}{1 - \alpha\Theta} Y_{t-1} + \frac{\alpha\delta_{22} + \delta_{12}}{1 - \alpha\Theta} X_{t-1} + \frac{\alpha\mu_{2t} + \mu + 1t}{1 - \alpha\Theta} \quad (8)$$
$$X_t = \frac{\Theta\delta_{11} + \delta_{21}}{1 - \alpha\Theta} Y_{t-1} + \frac{\Theta\delta_{12} + \delta_{22}}{1 - \alpha\Theta} X_{t-1} + \frac{\Theta\mu_{1t} + \mu_{2t}}{1 - \alpha\Theta} \quad (9)$$

The relevant (for Granger causality purposes) reduced form coefficients are equivalent to:

$$\Pi_{12} = \frac{\alpha\delta_{22} + \delta_{12}}{1 - \alpha\Theta}$$
$$\Pi_{21} = \frac{\Theta\delta_{11} + \delta_{21}}{1 - \alpha\Theta}$$

[6] The intercept is dropped for computational convenience.

Recalling Granger's (1969) definition discussed above, we see that testing Granger causality centers on estimating π_{12}, the coefficient on X_{t-1} in Equation 6, and on estimating π_{21}, the coefficient on Y_{t-1} in Equation 7. In this example, X_t Granger causes Y_t if $\pi_{12} \neq 0$, while Y_t Granger causes X_t if $\pi_{21} \neq 0$. Now suppose that an analyst is interested in obtaining a consistent estimate of α in the system of equations above (4, 5). The well-known consistency results dictate that α can be consistently estimated provided that $\mathrm{E}(\mu_{1t}) = 0$ and that X_t is uncorrelated with μ_{1t} in Equation 4.

Notice that X_t will be contemporaneously uncorrelated with μ_{1t} when $\Theta = 0$. And now suppose that we try to test whether feedback exists from Y_t to X_t using a Granger causality test. Unfortunately, such a test allows us only to investigate whether $\pi_{21} = 0$, yet a finding that $\pi_{21} = 0$ provides no information about whether $\Theta = 0$. In fact, Granger causality tests can lead to Type I and Type II errors when drawing inferences about Θ. For instance, if $\delta_{21} \neq 0$, then an analyst is likely to reject incorrectly the null that $\Theta = 0$. Furthermore, if $\delta_{21} = -\Theta\delta_{11}$, then a Granger causality test can lead an analyst to conclude incorrectly that $\Theta = 0$ since $\pi_{21} = 0$. Thus, Granger causality tests lack power and shed no light on whether α can be consistently estimated because they do not allow one to test assumptions of weak exogeneity.[7]

We are not arguing that Granger causality tests are useless since once weak exogeneity has been confirmed, they can give us some insight into the existence of strict exogeneity; that is, they bear not on inference, but on forecasting. When our concern centers on forecasting from a single equation, then a lack of Granger causality is needed to ensure that the linear predictor, Y-hat or X-hat, is unbiased.

For instance, suppose we want to use the model in Equation 2 to forecast X_t. Since we first need an unbiased forecast of X_t, this implies that we need $\pi_{12} = 0$ so that Y_t is unaffected by feedback from X_t. Notice that when $\pi_{12} = 0$, then δ_{12} and $\alpha = 0$. Hence, when $\pi_{12} = 0$, the reduced forms in Equations 8 and 9 become equivalent to the model in Equation 5. This, in turn, implies that estimating the single model in Equation 5 involves no loss of information.

To see this, impose the $\pi_{12} = 0$ restriction on Equation 8 to obtain:

$$Y_t = \delta_{11} Y_{t-1} + \mu_{1t} \tag{10}$$

which simplifies to:

$$Y_t - \mu_{1t} = \delta_{11} Y_{t-1} \tag{11}$$

Now, substituting 11 into 9 and again imposing the restriction $\pi_{12} = 0$, we obtain:

$$X_t = \delta_{22} X_{t-1} + \Theta Y_t - \Theta\mu_{1t} + \delta_{21} Y_{t-1} + \Theta\mu_{1t} + \mu_{2t} \tag{12}$$

The simplification in 12 is trivial:

$$X_t = \Theta Y_t + \delta_{21} Y_{t-1} + \delta_{22} X_{t-1} + \mu_{2t} \tag{13}$$

Note that Equation 13 is equivalent to 5. This equivalence could not be achieved unless $\pi_{12} = 0$; that is, unless Y_t is strictly exogenous. Thus, given weak exogeneity, Granger causality can provide evidence that the strict exogeneity of Y_t does hold.

Conclusion and Discussion

When regressors are stationary and the disturbances in a structural (conditional) model are non-autocorrelated, consistent estimation requires assumptions of weak exogeneity. Granger causality tests, however, are not directly relevant for testing such assumptions, although, given the existence of weak exogeneity, they do provide some information about a lack of strict exogeneity. But too often, analysts are willing to use findings of Granger non-causality as evidence of the type of exogeneity that rules out simultaneity and allows consistent estimation of parameters within the context of a single equation. Or as Cooley and LeRoy (1985: 299) phrase it:

> Characteristically, ...the statement that exogeneity is testable is accompanied neither by any showing as to why the relevant exogeneity concept is strict exogeneity rather than predeterminedness, nor by the proviso that the interpretation of the test is unambiguous only in the case of rejection of Granger non-causality.

A political science example of this is found in MacKuen, Erikson, and Stimson (1992), who conclude from a series of Granger causality tests that economic perceptions are "exogenous" to presidential approval. Perhaps their conclusion is correct, but it amounts to an overstatement of their empirical evidence. We do not mean to suggest that Granger causality tests are useless since they can be employed in

[7] In the case of a least squares estimator, consistency is predicated on the behavior of the covariance matrix. Consider the following standard definition for an OLS estimate (β):

$$\hat{\beta} = \beta + (X'X)^{-1} X'\epsilon$$

For consistency the covariance matrix must vanish asymptotically:

$$\hat{\beta} = \beta + (X'X)^{-1} X'\epsilon = \beta + \frac{(X'X)^{-1}}{T} \frac{X'\epsilon}{T}$$

where T is the sample size. As long as the X and ϵ (or in our case, X_t and μ_{1t}) are uncorrelated we know that the sample covariance matrix $\rightarrow 0$ as $T \rightarrow \infty$. Therefore, our estimate is consistent: it converges to its true value as time progresses.

conjunction with tests for weak exogeneity to investigate the validity of exogeneity assumptions. However, we want to encourage analysts to be cautious about the conclusions they reach on the basis of Granger causality tests alone.

Since weak exogeneity justifies estimation and inference from a single structural equation model when regressors are stationary,[8] analysts interested in testing their a priori exogeneity assumptions will want to use tests for weak exogeneity. Unfortunately, no direct test exists. Common sense suggests that one should construct various marginal models and look for significant parameters that are not part of the structural (conditional) model, but this is not a practical strategy for two reasons. First, the marginal model may be misspecified, thereby leading to spurious conclusions about the regressors in the structural (conditional) model, and second, constructing marginal models defeats the whole purpose of trying to condition on a set of variables in a single equation.

An indirect alternative for testing weak exogeneity relies on parameter constancy tests of the Dufour (1982) variety. Tests for parameter constancy can shed light on a lack of weak exogeneity because coefficient estimates that are unstable over time are often indicative of a structural change in the process generating the X_t regressors (i.e.; in the marginal model). When weak exogeneity holds, the structural (conditional) model will be invariant to such changes; hence, evidence of parameter instability provides some indirect evidence that weak exogeneity fails to hold.

At one time, parameter constancy tests involved cumbersome calculations, but with the advent of new statistical software such as PC-Give, MicroTSP, and SHAZAM, there are now easily executable parameter constancy tests, including recursive residual tests, cumulative sums of residuals tests, and one and n-step ahead forecasting tests. We urge analysts to implement them more frequently.

Next up: In Part II, we will examine exogeneity, inference, and Granger causality tests for nonstationary, integrated regressors.

References

Alt, James. 1985. "Political Parties, World Demand, and Unemployment: Domestic and International Sources of Economic Activity." *American Political Science Review,* 79: 1016-1040.

Cooley, Thomas., and Stephen LeRoy. 1985. "Atheoretical Macroeconometrics: A Critique." *Journal of Monetary Economics* 16: 283-308.

Dufour, Jean-Marie. 1982. "Recursive Stability of Linear Regression Relationships." *Journal of Econometrics,* 19: 31-76.

[8] As we will see in Part II, strict exogeneity becomes the relevant concept for the case of nonstationary regressors.

Engle, Robert, David Hendry, and Jean-Francois Richard. 1983. "Exogeneity." *Econometrica,* 51: 277-304.

Freeman, John. 1983. "Granger Causality and the Time Series Analysis of Political Relationships." *American Journal of Political Science,* 27: 327-358.

Granger, C. W. J. 1969. "Investigating Causal Relations by Economic Models and Cross Spectral Methods." *Econometrica,* 37: 424-438.

Granger, C. W. J. 1980. "Testing for Causality: A Personal Viewpoint." *Journal of Economic Dynamics and Control,* 2: 329-352.

Kmenta, Jan. 1971. *Elements of Econometrics.* New York: Macmillan.

Koopmans, Tjalling. 1950. "When is an Equation System Complete for Statistical Purposes?" in *Statistical Inference in Dynamic Economic Models,* ed. by Tjalling Koopmans. New York: John Wiley and Sons.

MacKuen, Michael, Robert Erikson, and James Stimson. 1989. "Macropartisanship." *American Political Science Review,* 83: 1125-1142.

MacKuen, Michael, Robert Erikson, and James Stimson. 1992.. "Peasants or Bankers? The American Electorate and the U.S. Economy." *American Political Science Review.* 86: 597-611.

Spanos, Aris. 1986. *Statistical Foundations of Econometric Modelling.* Cambridge: Cambridge University Press.

Williams, John. 1990. "The Political Manipulation of Macroeconomic Policy." *American Political Science Review,* 84: 767-795.

Consensus, Idiosyncrasy and Why Your R-Square Seems Low

Eric Plutzer
Iowa State University

Statistics, as a field of mathematics, is objective in the sense that conclusions can be verified, proofs can be replicated, and mistakes can be uncovered by peers who begin with the same axiomatic assumptions. In contrast, the application of statistics by scientists involves a mix of arbitrary decisions, consensus and idiosyncrasy.

Arbitrary decisions

Political scientists, just like our counterparts in the other social and natural sciences, are obliged to make arbitrary decisions to interpret statistical findings. We must decide what level of statistical significance is needed to reject a null hypothesis, what degree of precision is needed before accepting the estimates from a particular model iteration, and what level of response rate is acceptable for a sample survey, to mention just a few. Unlike those made by statisticians, these assumptions are not axiomatic. That is, they are not made in order to see what logically follows from them. Rather, they are like the decision rules reached in engineering and operations research: they provide arbitrary decision rules that permit us to move on to interpretation and, sometimes, policy recommendations.Consensu

For some of these arbitrary decisions, a consensus has been reached. For example, it is conventional for sample statistics to be taken as different from some fixed constant (usually zero) if the difference is so large that it would occur no more than 5% of the time by chance. While no one can argue that the .05 probability level is better than .06, .04, or .01, the consensus on .05 permits scientists to engage in a dialogue among one another. Though such dialogues may be hampered by a lack of consensus on a number of methodological and theoretical issues, few fight over significance levels.Idiosyncrasy

Yet there are a number of arbitrary decision rules that are applied idiosyncratically. Among these are decision rules to determine the number of latent factors underlying a set of indicators, lower bounds of goodness of fit statistics for maximum likelihood estimates, and adequate levels of explained or unexplained variance in models estimated by least squares.

Among the group of practicing political scientists, the most common source of disagreement concerns the appropriate size of standard errors of regression estimates and it's mathematically equivalent expression as R-square. To add a personal note, I recently had a paper rejected by a major journal. There were a number of criticisms with which I agreed (and addressed in a subsequent draft). In addition, however, one referee and the editor both noted that my R-square was too low.

Readers of /tpm know two important sides to this argument and I won't repeat them here. Rather, I wish to inform the debate by showing why apparently low R-squares (or big standard errors of estimate) are inherent in survey research and several other areas of statistical application. Wider understanding of this may provide some stimulus to approaching consensus on what constitutes a model which "explains a lot" (which, of course, is not necessarily the same as a good model) or whether we can even make such an assessment.Limits

Most fundamentally, R-squares computed from models of survey data are low because of low reliability of dependent variables. While this may seem intuitive to most practitioners, I have found it surprising how few political scientists can explain why this is the case and how even fewer can express this formally. Measurement is given short shrift in methods classes and is rarely mentioned in intermediate statistics classes where it should be taught (for an expanded version of this argument as well as an excellent primer to social measurement, see Bohrnstedt 1969). Although less likely in practice, a large class of survey problems also run the risk of inflated R-square values deriving from invalid measurement. My purpose here is to provide relatively simple illustrations of both phenomena, both of which suggest using structural equations methods whenever possible, and for minimizing reliance on R-square (of low reliability) when this is not.Consequenc

Consider that we want to explain variation in some underlying attitude trait. A dependent variable I have examined recently is pro/anti feminism. I use a scale based on six items in Euro-barometer #18. One frequently used measure of reliability is the internal consistency measure Cronbach's alpha. This is commonly interpreted as measuring the lower bound of the constructed index's correlation with the latent variable. There is a loose consensus that a minimally reliable scale should have an alpha level of .65 (although some say .6, others .7, and many psychologists and educational testers prefer higher levels). My scale has a reliability level of .72.

Now let us assume, that I have a perfectly specified model. That is, I have a set of regressors, X, that completely explain the variance in feminist attitudes (the latent trait) both in the population and in the sample. While impossible in practice, such a model presents a useful limiting case. We also assume that although the latent trait is measured without reliability, the index is valid. Our model is illustrated in Figure 1, below. Only two parameters are needed for purpose of illustration. The first is a structural parameter (beta) which is equal to one because the linear combination of the X's perfectly predicts the latent variable. The second, is the measurement parameter Lambda, which is .72, indicating how reliably the constructed scale measures the underlying trait.

to R-square

Figure 1

```
   beta=1.0
X ---------> LATENT ATTITUDE
                   |
                   |lambda=.72
                   |
                   v
            OBSERVED SCALE
```

In this circumstance, the multiple R is Lambda, or .72 and R-square is .52. What this simple illustration shows is that in properly specified models with valid measures, R-square has an upper limit equal to the square of the dependent variable's reliability.

R-square is interpreted in a number of ways but the most common seems to be as a flag for underspecification. In this usage, R-square is being used to get some idea if relevant predictors are present in the model; presumably, lots of unexplained variance means something has been left out. In this light, apparently low levels of R-square should be interpreted not as proportions of variance explained, but as proportions of all variance that could possibly be explained. Comparing R- square to a value of 1.0 makes no sense when a perfectly specified model can do no better than half that amount.

When the reliability can be estimated we can certainly compare our observed R-square to the theoretical maximum but it will normally be better to estimate the structural portion of the model using techniques designed for models with latent variables. In some cases, latent variable models cannot be estimated, however. An important class of such cases are when the dependent variable has only a single indicator (e.g., reported voter turnout). In such cases we cannot estimate reliability and there is no way to estimate the attenuatiand inRv-alidarmBasqurem en t

Unfortunately poor measurement can also inflate the value of R-square. Those who use R-square as a rough indicator of whether all relevant regressors are included may again come to the wrong conclusion. An important class of situations are those where the dependent variable is invalid by contamination with one or more regressors. Long a problem in intelligence testing (to IQ tests measure social background as well as some undefined mental quality we call intelligence?) there are cases familiar to most political scientists. Consider the family of tolerance indexes that are included on many sample surveys. These typically ask respondents if they would find a number of civil liberty restrictions (e.g., banning public demonstrations) for each of a number of groups. The goal of many studies is to identify the social background characteristics which predict (or determine) intolerance. Yet if an individual belongs to a group that is mentioned in the scale, a built in dependence exists between X and one component of the dependent measure. Such a situation is easily depicted by the MIMIC class of models as in Figure 2, below.

Figure 2

```
  beta = 1.0
X ------------> LATENT ATTITUDE
|                     |
|                     |lambda=.72
|                     |
| gamma =/= 0         v
-------------->   OBSERVED SCALE
```

Here, gamma represents a structural parameter between one or more regressors and one or more indicators of the latent trait. Once again, our model is specified perfectly with a linear combination of X's explaining all the variation in the latent trait. But our observed R-square will be inflated by residual correlation between X and one indicator. Estimating this model by least squares produces a single fit statistic when in fact it contains two components: the variance in the latent trait explained (presumably out theoretical interests lie here) and additional variance explained in our measure of the latent trait.

Obviously, latent variable models such as those estimable by LISREL and similar programs can clarify matters and should be used when appropriate. Yet again, available measures may not permit decomposition into component parts. In such cases, explained variance in the observed measure can again be very misleading.Summary

R-square usually reports variance explained in observed dependent measures. When our interest lies in a latent trait, as it often does in survey research, R-square is always attenuated due to unreliability and sometimes inflated due to built in dependencies between the left and right hand side of the equations. When problems do not permit estimating the precise amounts of attenuation and inflation, R-square should be interpreted with extreme caution.

Coming next time in *TPM*:

Part II of "Exogeneity, Inference and Granger Causality" by Jim Granato and Renée M. Smith.

"Bootstrapping a Regression Model" (with GAUSS code!) by Christopher Z. Mooney.

Special Issue on statistical software: a comparison of GAUSS and S+ by Simon Jackman; reviews of SAS, SHAZAM, RATS and other popular software packages.

And more of the things you expect from *TPM*!

Eleventh Annual Political Methodology Conference Program

The Eleventh Annual Political Methodology Conference will be held July 21-24, 1994, at the University of Wisconsin, Madison. An anonymous ftp site (at *hss.caltech.edu*, in the directory pubpol_methods_1994) will be maintained for access to the papers presented at this meeting. For information about anonymous ftp access to these papers, send an e-mail message to rma@hss.caltech.edu and nagler@wat2213.ucr.edu.

Measuring Bias in Political Participation, by Henry Brady, *University of California, Berkeley*. Disc: John Brehm, *Duke University*.

An Exploratory Analysis of Public Opinion Dimensionality, by Robert Durr and Andy Whitford, *Washington University*. Disc: Robert Erikson, *University of Houston*.

Statistical Models of Path Dependence, by John Jackson, *University of Michigan*. Disc: Christopher Achen, *University of Michigan*.

Filtering as a Means of Theory Construction and Interdisciplinary Bridge Building, by John Freeman and James Stimson, *University of Minnesota*. Disc: Nathaniel Beck, *University of California, San Diego*.

Induced Ideal Points and the Nature of Legislator Motivations, by William Bianco, *Duke University*. Disc: John Londregan, *Princeton University*.

Revenge of the Slurge: On Measuring the Incumbency Advantage, by Jonathan Katz, *University of California, San Diego*. Disc: Douglas Rivers, *Stanford University*.

Government Duration, Censoring, and Competing Risks, by Randolph Stevenson, *University of Rochester*. Disc: Gary King, *Harvard University*.

The Microfoundations of Macropartisanship, by Janet Box-Steffensmeier, *Ohio State University*, and Renee Smith, *University of Rochester*. Disc: W. Phillips Shively, *University of Minnesota*.

Correlated Errors and the Incorrect Signs Phenomenon in Proxy Variable Regression, by Robert Luskin and Joe TenBarge, *University of Texas*. Disc: Bradley Palmquist, *Harvard University*.

Pooling Disparate Observations, by Larry Bartels, *Princeton University*. Disc: R. Michael Alvarez, *Caltech*.

Genetic Monte Carlo Cross-validation, by Walter Mebane and Gregory Wawro, *Cornell University*. Disc: Simon Jackman, *University of Chicago*.

Electoral Laws, Run-off Requirements, and Minority Representation, by Elisabeth Gerber, *Caltech*, and Rebecca Morton, *University of Iowa*. Disc: Philip Schrodt, *University of Kansas*.

A GAUSS Mailing List has been created for the discussion of virtually anything relating to the GAUSS software system. The goal of the Mailing List is to provide a vehicle for GAUSS users to share information, experiences, and even frustration. Unlike other similar e-mail forums, the folks at Aptech even monitor the various mailings, and from time to time do respond. For subscriptions to the list, send e-mail to "gaussians-request@uclink.berkeley.edu". To post to the list, send your message to the list address "gaussians@uclink.berkeley.edu."

1994 APSA Annual Meeting, Political Methodology Division Preliminary Program

Walter R. Mebane, Jr.
Cornell University.

"Pooled and Nonparametric Specifications"

Chair: Janet Box-Steffensmeier, *Ohio State University*

"Pooling Disparate Observations" Larry Bartels, *Princeton University*

"Nonparametric Analysis of Changes in the Income Distribution, 1979–1987" Edward J. Bird, *University of Rochester*

"Nationalism, Racism and Intolerance in the European Community" Jasjeet Sekhon, *Cornell University*

Discussants: Janet Box-Steffensmeier, *Ohio State University* and Theresa Marchant-Shapiro, *Union College.*

"Advances in Estimation"

Chair: W. Phillips Shively, *University of Minnesota*

"Revenge of the Slurge: On Measuring the Incumbency Advantage" Gary Cox and Jonathan Katz, *University of California, San Diego*

"Respecification Approaches to Ecological Inference: A Comparison of Control Variables and Breakage Effects" Bradley Palmquist, *Harvard University*

"Incidental Truncation in Event History Models" Christopher J. Zorn, *Ohio State University*

Discussants: Gary King, *Harvard University* and W. Phillips Shively, *University of Minnesota.*

"Ideological Scaling and Measurement"

Chair: William G. Jacoby, *University of South Carolina*

"Scaling Environmental Ideologies of Groups Involved in Environmental Policy" Scott P. Hays, Michael V. Esler and Carol E. Hays, *Southern Illinois University at Carbondale*

"Disaggregating Reliability for a Substantive Measure" Igor Philip Matkovsky, *American University*

"Looking at Liberal Discourse in Cold War: Context and Content Analysis" Joanna Scott, Karen Schauman, Adrian Lottie, Miles McNiff, and Claudia Dahlerus, *Eastern Michigan University*

Disc.: William G. Jacoby, *University of South Carolina.*

"Stochastic Choice Models"

Chair: John Londregan, *Princeton University*

Papers: "A Comparison of Multinomial Probit and GEV for Estimating Models with Multiple Candidates" R. Michael Alvarez, *California Institute of Technology* and Jonathan Nagler, *University of California, Riverside*

"Optimizing the Classification Performance of Probit Models in Political Science Research" Barbara M. Yarnold, *Florida International University* and Paul R. Yarnold, *Northwestern University*

Discussants: John Londregan, *Princeton University* and Langche Zeng, *University of Oregon.*

"Aggregate Models of Elections, Popularity and Opinion"

Co-sponsored with Elections and Electoral Behavior (Political Methodology is the primary sponsor).

Chair: Regina Baker, *University of Michigan*

"Economic Perceptions and Economic Reality: Modeling Voter Information" R. Michael Alvarez, *California Institute of Technology* and Dean Lacy, *Duke University*

"A Heteroskedastic Time Series Model of Volatility in Presidential Approval Ratings" John Brehm and Paul Gronke, *Duke University*

"Exploring the Myth of Nationalization: The Growing Importance of Constituency Factors in Congressional Elections" David A. Scocca, *University of North Carolina*

"Explaining Congressional Approval" Robert H. Durr, John B. Gilmour and Christina Wolbrecht, *Washington University*

Discussants: Regina Baker, *University of Michigan* and Simon Jackman, *Princeton University.*

"Advances in Simulation"

Chair: Marcia Whicker, *Rutgers University*

"Estimating Continuous Time Nonlinear Systems with Chaotic Potential" Courtney Brown, *Emory University*

"Emergence of Political Elites by a Genetic Algorithm" Thad A. Brown, *University of Missouri* and Michael D. McBurnett, *University of Illinois*

"Emergence as a Paradigm for the Study of Political Phenomena: Models of Bottom-up Processes in Politics" David Lazer, *University of Michigan*

Disc: Walter Hill, *University of 7orth Carolina* and Marcia Whicker, *Rutgers UniversityP*

"Resampling Methods for Bayesian and Robust Estimation"

Chair: Jonathan Katz, *University of California, San Diego*

"Imputations and Inferences with Weak Data: An Exposition of the Gibbs Sampler" Simon Jackman, *Princeton University*

"Genetic Monte Carlo Cross-validation" Walter R. Mebane, Jr. and Gregory J. Wawro, *Cornell University*

Discussants: Jonathan Katz, *University of California, San Diego* and Douglas Rivers, *Stanford UniversityP*

"Empirical Tests of Formal Theories of Legislatures and Elections"

Co-sponsored with Formal Political Theory (Political Methodology is the primary sponsor).

Chair: Scott Ainsworth, *University of Georgia*

"The Rise of Seniority in Legislative Systems: A Definition and Comparative Analysis" David Epstein, *Columbia University* and David Brady, *Stanford University*

"A Model and Test of the Effect of Elections on Retirements and the Effect of Retirements on Elections" Timothy Groseclose, *Carnegie-Mellon University*

"Position-taking in Dynamic Elections: An Empirical Test" Ken Kollman, *University of Michigan*

"Ideological Conflict in the Chilean Legislature" John Londregan, *Princeton University*

Disc: Jonathan Nagler, *University of California, RiversideP9*

"A vMeet the Authors' Roundtable on King, Keohane and Verba's *Designing Social Inquiry: Scienti0c Inference in Qualitative Research*"

Chair: David Collier, *University of California, Berkeley*

Authors: Gary King, *Harvard University* and Robert O. Keohane, *Harvard University*

Discussants: Larry Bartels, *Princeton University*, Henry Brady, *University of California, Berkeley*, Peter Lange, *Duke University*, and Ronald Rogowski, *University of California, Los AngelesP*

"Roundtable: Statistical Reporting, Archiving and Replication— Norms for Publication"

Chair: James A. Stimson, *University of Minnesota*

Panelists: Charles Franklin, *University of Wisconsin*, Walter R. Mebane, Jr., *Cornell University*, Philip A. Schrodt, *University of Kansas*, and B. Dan Wood, *Texas A&M UniversityP*

"Roundtable on Formal Political Theory: Mathematics, Mechanics, or Magic?"

Co-sponsored with Formal Political Theory (Panel 4–1; Political Methodology is the secondary sponsor).

Don't forget — *TPM* has moved from the Midwest to California! Please send all correspondence related to *TPM* to the new editors. Hardcopies should be sent to R. Michael Alvarez, Division of Humanities and Social Sciences, MC 228-77, California Institute of Technology, Pasadena, CA 91125. All e-mail should be sent to **both the editors** at rma@hss.caltech.edu and beck@ucsd.edu. Our rule is anything that should not be sent to *Political Analysis* belongs in *TPM*!

anel 1

Contents of *Political Analysis*, T olumes 4 and V

Political Analysis is the journal of the Political Methodology Section of the American Political Science Association. Serving as the primary outlet for research in political methodology, *Political Analysis* publishes research articles in all areas of political methodology. As it is the primary vehicle for the publication of research in political methodology, please continue to support *Political Analysis* by ordering a copy for yourself (from The University of Michigan Press) and getting your university library to place a standing order for the journal of our organized section. Don't forget, in addition, that special discount pricing is available for members of the methodology section!

Below are the contents of Volume 4 (now available), and the contents of Volume 5 (available within the next 18 months). If you are interested in submitting an article to *Political Analysis*, please consult the manuscript format information in previous issues and send your materials to: John R. Freeman, Editor, *Political Analysis*, Department of Political Science, 1414 Social Sciences Building, University of Minnesota, Minneapolis, MN 55455.

Contents of Volume 4

Contents of Volume 5

TextI ook Review: *Statistical Methods in Econometrics*, Ramu Ramanathan, Academic Press, 1993. $VV.

Nathaniel Beck
University of California, San Diego

Advanced work in either econometrics or statistics requires some understanding of mathematical statistics and at least a grounding in mathematical probability. Such courses are not prerequisite for admission to most Ph.D. programs in political science; I dare say that even a few political methodologists might be lacking in this type of training. Ramanathan's text is ideal for efficiently making up for these deficiencies.

A good graduate econometrics text (e.g. Johnston's *Econometric Methods* or Greene's *Econometric Analysis*) will have one long chapter on mathematical statistics. This chapter makes the book self-contained, but one chapter, no matter how good (and both the Johnston and Greene chapters are very good) cannot teach all the mathematical statistics that is needed. The result is that the material is often presented without precise definitions and that theorems are stated but not proven. The consequence is that the student (or faculty member) may know, for example, that the log of the likelihood ratio is distributed as chi square, but have no real insight into why this is the case.

One alternative is to take a serious course (or read a full text) in mathematical probability and statistics. But this alternative is unrealistic. And it is also inefficient, since courses in mathematical statistics and probability are not tailored to the needs of those interested in econometrics, to say nothing of those interested in political methodology. Ramananthan's text is tailored exactly for this audience, and the fit is admirable.

The book is in three sections: mathematical probability, mathematical statistics and econometrics. The first two consist of a well laid out treatment of those parts of probability and statistics that are necessary for further econometric work. The theorems are carefully stated and the proofs, while not always complete, are sketched carefully enough to allow the reader to have insight into what is going on. I regard the proofs as the major selling point of the book. Theorems without proofs give no understanding, but, alas, many mathematical proofs make it difficult for the reader to get the essential insight of the theorem. Ramanathan has done an incredible job at steering between this Scylla and Charybdis.

Part I, on probability, is probably less useful than Part II, on statistics. The reader of Part I will get a careful introduction to measure theory and the Stieljes integral, as well as a careful presentation of characteristic functions. Ramanathan also presents a variety of distributions, distributions that are of great use in applied work but often neglected in standard econometrics texts. (The pictures of distributions in Rothschild and Logothetis' *Probability Distributions* would be an excellent supplement to this section.) But it is Part II that sells the book.

Here we have carefully laid out treatment of the statistical underpinnings of what we do. The chapter on asymptotics is as clear a statement about the modes of convergence and the various central limit theorems as I have ever seen. The reader will get a good treatment of the relationship among the different modes of convergence and there is a nice sketch of a proof of the Lindberg-Levy version of the central limit theorem.

There is a similarly good treatment of estimation (and particularly maximum likelihood) and hypothesis testing (in the Neymann-Pearson framework). Many know the asymptotic properties of maximum likelihood; the reader of this book will also understand why maximum likelihood has these properties. As befits a book at this level, most attention is paid to the multiple parameter case.

Part III of the book, on econometrics, is disappointing. The topics covered here are covered in more detail in a good econometrics text. Many of the topics, such as generalized least squares, do not even build on the impressive foundation laid in the first two parts. My advice to readers and students would be to treat the book as ending with Part II.

I would have preferred that Part III be replaced with a good treatment of linear algebra. Ramanathan provides a good list of the theorems of linear algebra in an appendix, but no one will get any intuition about linear algebra from this appendix. The reader who has either forgotten or never learned linear algebra will have to go to another source. (The linear algebra chapter in Greene is quite good, but the best intuitive introduction is in Weintraub's fabulous *Mathematics for Economists*).

Who should get this book? Most of our graduate students will find it requires more mathematics than they know. But students have to learn this material before they can be turned loose on White's *Asymptotic Theory for Econometricians*. Such a student will be well served by Ramanathan's text. But we have few such graduate students. The major beneficiaries of this book will be faculty teaching advanced methods courses. I think that we should know why the central limit theorem works and why all the asymptotics of maximum likelihood are as they are. We may never (and possibly should never) present a proof of these theorems in a class, but we should present the intuitions behind them. I know of no better text for efficiently sharpening these intuitions than the Ramanathan text. He has performed a real service.

In forthcoming issues of *TPM* we are planning on including syllabi from political methodology courses. If you have any syllabi you would like to share, please send them to the editors. Other than graduate methodology syllabi, we are interested particularily in syllabi:

From undergraduate methodology courses;

Which approach the subject of methodology in new and innovative ways;

With a disciplinary focus (i.e., graduate syllabi on comparative politics or international relations methodology;

Which take on a non-traditional topic, like experimental methods or computer simulations.

The Political Methodologist, c/o ... ,
Maib de
alioia ,titte o Techolog
Pa,adea

The Political Methodologist is the newsletter of the Political Methodology Section of the American Political Science Association. We gratefully acknowledge the support of the California Institute of Technology in helping to defray the editorial and production costs of the newsletter.

Subscriptions to *TPM* are free to members of the APSA's Methodology Section. Please contact APSA to join the section. Dues are H8.00 per year.

Submissions to *TPM* are welcome. Articles should be sent to the co-editors if possible, by e-mail to rma@hss.caltech.edu and beck@ucsd.edu. Alternatively, submissions can be made on diskette as plain ascii files sent to R. Michael Alvarez, Division of Humanities and Social Sciences 228-77, California Institute of Technology, Pasadena, CA 91125. LaTeX format files are especially encouraged. The deadline for submissions for the next issue is August 1, 1994.