Comparing Dynamic Specifications: The Case of Presidential Approval

Nathaniel Beck

Abstract

This article compares a variety of models of presidential approval in terms of their dynamic properties and their theoretical underpinnings. Exponential distributed lags, partial adjustment, error correction, and transfer function models are considered. The major difference between the models lies in interpretation rather than statistical properties. The error correction model seems most satisfactory. Approval models based on individual level theories are examined, and found to give no additional purchase.

Introduction

It is possible to view political time-series as "merely" presenting a series of interesting technical problems, ranging from serially correlated errors and autoregressive conditional heteroskedasticity to unit roots and cointegration. The technical details of time-series become formidable, indeed, when we move to Kalman filters and the state-space form, nonlinear models, or the frequency domain. Over the last decade, political science as a discipline has started to deal with these technical issues. Much of the credit for this must go to Douglas Hibbs (1974), who carefully showed us the pitfalls, both substantive and technical, of ignoring the structure of the error process. We now seldom see papers that fail to test for, and correct, serially correlated errors. Recent applications of cointegration (Ostrom and Smith 1990), vector auto-

This project was started while I was visiting the Government Department at Harvard University. Jim Alt, Gary King, Michael MacKuen, and Doug Rivers provided helpful comments. I owe a special debt to Rob Engle, who pointed out to me that what I thought was an obscure empirical article about the UK consumption function might really be quite interesting. The first version of this paper was given at the annual meeting of the Political Methodology Society, Duke University, 1987.

regression (Freeman, Williams, and Lin 1989) and Kalman filters (Beck 1989) show that complicated technical questions have entered the discourse of political science.

In addition to being highly technical, time-series analysis offers an astonishing variety of possible specifications. Both cross-sectional and time-series analysts must choose a set of independent variables and a functional form relating those variables to the dependent variable. But the time-series analyst also must choose from a wide variety of dynamic specifications for each variable as well as the "error" process.[1] Choice of a dynamic specification can be guided by statistical theory. There are a variety of tests of misspecification to ensure that model residuals are appropriate (uncorrelated and homoskedastic), that the model coefficients are stable over time, and that the functional form is appropriate; all specifications should always be subjected to a battery of misspecification tests. Choice among alternative specifications can be guided by the standard F-test if one specification is nested inside another, or by a variant of the Cox test for evaluating nonnested models.[2] But statistics alone is often not sufficient to lead us to a single, "best" specification. We then must choose a specification on other grounds, and one way of choosing is by comparing the dynamic properties of a specification with the dynamics implied by theories. It is this task that is the subject of this article.[3]

The purpose of this article is to elucidate questions of choice of dynamic structure by looking at the substantive properties of various structures in the context of one substantive arena, the modeling of presidential approval. Approval was the context of Hibbs's pioneering effort, and has probably received more analysis than any other single political time-series. While I hope that the lessons learned from the present analysis generalize to other areas of interest, those interested in generalization will obviously have to proceed cautiously.

1. The Judge et al. treatise (1985), for example, devotes about a sixth of its thousand pages to issues of dynamic specification.

2. A specification is nested inside another if the first is a specialization of the second, that is, the first specification is the second with some parameters constrained. The Cox procedure deals with nonnested models by constructing a likelihood function that is a linear combination of the likelihood functions for each model; the test is of which likelihood contributes more to this combined likelihood. Unfortunately, the Cox test is often inconclusive; this is the case for comparing the better exemplars of the models estimated in this article. Some easily computable Cox tests are reported in Davidson and MacKinnon 1981. An alternative to the Cox test is to use one of a number of criteria, such as the Akaike Information Criterion (AIC) or the Schwartz Criterion (Judge et al. 1985, 869–73). These are based on minimizing the standard error of estimate plus some penalty for lack of parsimony (measured by degrees of freedom). Since the models estimated in this article are quite similar in parsimony, these criteria are not helpful for distinguishing between them. Thus, only the standard error of estimate, σ, is reported in the tables.

3. There are many other nonstatistical methods for choosing between specifications. An enlightening discussion of this whole area may be found in Leamer 1978.

There are myriad studies of presidential approval (sometimes misleadingly called popularity), but only a limited number of dynamic specifications have been used. In this article I compare and estimate, using a common set of variables and a linear, functional relationship, several of the most important of these specifications. While different scholars have used different variables to explain approval, the variables used here are the ones most frequently used. In discussing the various specifications I also simplify matters by referring to only a very small number of substantive applications. While these were chosen for my methodological purposes and discussion of the applications focuses on issues of dynamics specification, the applications examined are among the more important contributions to the study of presidential approval. But this article should not be read as a bibliographic essay on presidential approval.

The generic (linear) approval function relates measured approval at time t, denoted A_t, to previous values of approval, a vector of current exogenous variables, denoted X_t, and its lags, and ϵ_t, the "error" term.[4] Harvey (1990) calls this type of model an autoregressive distributed lag (AD) model.

The AD model can be written

$$A_t = \sum_{i=0}^{L} X_{t-i}\beta_i + \sum_{i=1}^{M} \phi_i A_{t-i} + \epsilon_t , \qquad (1)$$

where values corresponding to time periods 0 and before are not observed, and either summation may be infinite. The error process is a sequence of random variables; the distribution of these random variables may be specified in a number of ways. Differing assumptions about the β's, ϕ's and the error process lead to different models of approval, which has both substantive and statistical consequences. If only current X's are in equation 1, that is, if L is zero and the second summation is excluded, we have a static model of approval; if only a finite number of lagged X's are in the equation, that is, if L is finite and the second summation is excluded, we have a finite distributed lag model; if L is infinite or lagged approvals are in the equation, we have an infinite distributed lag model. In general, these models contain too many parameters to be estimated straightforwardly; thus, some relationship among

4. I use the word *error* in quotes since the term is only an error from the standpoint of the analyst, who has either ignored or mismeasured some variables or otherwise failed to completely model the approval process. It might be better to call the error term an unmeasured shock, to remind us that there is little theoretical distinction between X and ϵ, but I follow convention and use the word *error*. There is some truly random error in the approval function, because it is measured by survey and hence contains sampling error.

the parameters is usually assumed, leading to various specific forms of the AD model. Any of these may be combined with any model of the error process.

In this article, the dynamic properties of several of the most important and commonly used specifications of the approval function are compared. Since the data are common across all the models, the next section treats data and some other preliminary issues with the following section dealing with common issues about the error process. The next four sections examine specific AD models: the static model, the exponentially distributed lag (EDL) model, the partial adjustment (PA) model and the error correction (EC) model. The following section examines an alternative to the AD model, the Box and Jenkins transfer function (1976). While these alternatives are not inherently different, their underlying philosophies do differ. The penultimate section treats the somewhat different question of specifying the dynamics of aggregate approval based on theories of individual approval. The conclusion sums up the strengths and weaknesses of the various models and approaches.

Preliminaries

This article uses monthly data. Most approval studies use quarterly data in the belief that quarterly averaging simplifies the model and reduces the effect of measurement error. But as various studies of temporal aggregation have shown, aggregation makes the dynamic model more complicated and produced incorrect estimates of dynamic parameters (Beck 1988; Freeman 1989). The complete sample period begins in April, 1953, and ends in October, 1988, yielding a maximum of 426 observations (fewer if more lags are required) for analysis.

The dependent variable in most analyses is the level of presidential approval, as measured by the Gallup Poll; a few analyses use the monthly change in approval, taken as a simple first difference.[5] The most important exogenous variables are two economic measures, the rate of change in unemployment (ΔU) and the rate of (consumer price) inflation, (I).[6] Both inflation and unemployment are assumed to affect approval with a lag of at least one month; this assumed lag is taken into account in the notation, so, for example, ΔU_t refers to the change in unemployment in time period $t - 1$. Many studies use the level of unemployment (U) instead of its first difference, but Kernell (1978) has made a cogent argument for using the first difference of unemployment. The error correcting model, as we shall see, makes sense of the question of using levels versus differences as explanatory variables.[7]

Specifications also include a constant term (C) and three "events" variables, marking the long-term effect of the Vietnam War on President Johnson's approval rating (V), the long-term effects of the Watergate affair (W), and a variable designed to account for the short-term effects of a number of dramatic events (E).[8]

All models are estimated so that events in one administration do not affect approval in a subsequent administration. This was done, in general, by treating the first month of each administration as missing data.[9] Some models require other procedures to eliminate leakage; this is discussed in conjunction with those models.

All models were estimated with RATS 386 Version 3.11, using either ordinary least squares (OLS), nonlinear least squares (NLLS) or the Hildreth-Lu (HL) grid search technique for models with an autoregressive error structure. None of the estimations were time consuming, and so issues of computational convenience are of little importance here.

The notation used in this article tries to be standard, though there is no real standard. X_t will always refer to a vector of exogenous variables (including a one for the constant term) and β is vector of parameters conformable

5. The data are from King and Ragsdale 1988, table 6.2, supplemented with various issues of the *Gallup Reports*. Approval is the proportion of the sample approving of the president's performance, so those with no opinion are lumped with those who disapprove of presidential performance. A few missing months were interpolated; in Beck 1989 I show that such interpolation is benign. For months with more than one survey, the last survey of the month was used.

6. Inflation is measured by the monthly percentage change, annualized, in the Consumer Price Index. Both unemployment and the CPI are from Citibase (LHUR and PUNEW).

7. Another reason for using ΔU instead of U is that unemployment has a unit root. (A series has a unit root if, loosely speaking, it is not stationary but its first difference is stationary. A series of stationary if its stochastic properties are time invariant. Harvey [1990, 23–30] presents a good introduction to these issues.) Regressors with unit roots can cause statistical problems (Stock and Watson 1988). Neither inflation nor approval have unit roots. See eq. 11 for this test.

8. The three events variables are conventionally used in approval studies. W is a dummy variable used from March, 1973, through August, 1974; V is the number of U.S. soldiers killed (in thousands) in Vietnam during the Johnson administrations. Variable E consists of ones, zeros, and minus ones to control for a series of short "dramatic" events, based primarily on MacKuen, Erickson, and Stimson 1989. The series is always coded so that a positive number increases approval. The series codes for Eisenhower's heart attack (one in July, 1954; minus one in August, 19~~84~~ 54), Khruschev's visit to the United States (one in December, 1959; minus one in January, 1960), the Cuban missile crisis (one in November, 1962), the Johnson era urban riots (minus one in July and August, 1967), the major Vietnam antiwar march (one in November, 1969; minus one in December, 1969), the mining of Haiphong harbor (one in May, 1972; minus one in June, 1972), the *Mayaguez* incident (one in May, 1975; minus one in July, 1975), the Iran hostage affair (one in December, 1979 and January, 1980; minus one in February and March, 1980), the Reagan assassination attempt (one in April, 1981; minus one in June, 1981) and the Iran-Contra affair (minus one in December, 1986). The series is empirically, not theoretically, derived. My defense for using all three series is conventional; failure to use event data leaves a lot of the action in the error process.

9. The Johnson administration begins in December, 1963, and the Ford administration begins in September, 1974. The Eisenhower, Johnson, Nixon, and Reagan administrations were treated as a single administration, with no special demarcation of the second term.

with X_t. To avoid transpositions, the approval function will be written $X_t\beta$. L will refer to the lag (or backshift) operator,[10] that is, $L^k(X_t) = X_{t-k}$. Finally, Δ will always refer to the first difference operator, that is, $\Delta = 1 - L$.

Modeling the Error Process

In cross-sectional analysis, we model the error term for each individual observation; in time-series, the errors are in general interdependent, hence the stochastic properties of the entire error process must be specified. The simplest assumption is that the errors are independent and identically distributed (iid). In this case, we can proceed as in cross-sectional analysis. It is well known that using OLS causes severe estimation problems if the error process is not iid (Harvey 1990, 195–98); tests for whether the error process is iid are easy to construct and now almost universally used by time-series analysts.[11]

We sometimes think of non-iid errors as a nuisance that makes OLS inappropriate (Beck 1985). But, if we take dynamics seriously, and if we think of the error process as being made up of unmeasured shocks that are not fundamentally different from the measured independent variables, then we would expect the error process to, in general, take a form like equation 1. We usually simplify a bit and assume either that the error process is autoregressive (AR) with

$$\epsilon_t = \sum_{i=1}^{p} \rho_i \epsilon_{t-i} + \nu_i \,, \qquad (2)$$

or a moving average (MA) with

$$\epsilon_t = \sum_{i=1}^{p} \theta_i \nu_{t-i} + \nu_t \,, \qquad (3)$$

where, in either case, the ν's are iid and p determines the order of the process.

10. While the algebra of lag operators and polynomials appears formidable, in practice it simplifies the analysis of time-series; Harvey (1990, 26–27) provides a good introduction to this algebra.

11. The critical i in iid, from a time-series perspective, is for independent, and so I here refer to the tests for serially correlated errors, the most well known of which is the Durbin-Watson test. The easiest and most general tests for independent errors for Lagrange multiplier tests, which regress OLS residuals on lags of those residuals and any lagged dependent variables (Harvey 1990, 278). The advantage of the Lagrange multiplier tests is that they can easily be designed to pick up error processes that show complicated forms of interdependence. Both Durbin's h-test (for first-order serial correlation with a lagged dependent variable) and the Box-Ljung Q-test, a χ^2

The AR error process almost completely dominates applied work for a very simple reason: it is easy to estimate models under this assumption. There is, in general, no theoretical reason to assume that error processes are AR, and, indeed, some theoretical reasons to believe they are MA. Measured approval contains iid sampling error; Granger's lemma (Granger and Morris 1976) shows that the error process for measured popularity must contain an MA term (Beck 1989). The MA form is probably also easier to fit into a general theoretical framework (King 1988, chap. 7).

The dynamics of AR and MA errors differ. Shocks in the AR error model persist forever, dying out exponentially (assuming that $|\rho| < 1$, that is, the error process is stationary). In the MA(1) model, shocks persist exactly one period. While, in principal, we ought to be able to use this information to guide our choice of error process, in practice this is difficult. An AR(1) error with, say, $\rho = .4$ looks very much like an MA(2) process because only about 5 percent of the AR shock persists more than two periods. (AR[1] processes with larger values of ρ mimic higher order MA processes.) In general, any stationary AR process may be represented by a higher order MA process and vice versa.

It is also empirically difficult to discriminate between AR and MA errors. The standard Lagrange multiplier test for an MA error process of order p is identical to the test for an AR process of that order (Harvey 1990, 278). We are thus fairly free to specify an AR or MA error, and our choice can be guided by convenience or the desire for a model with a small number of parameters.[12] In this article I usually use an AR error process because it is easier to ensure that shocks from one administration do not leak into the subsequent administration if errors are AR; this is done by treating the first period of each administration as missing.[13] It is much harder, as we shall see, to prevent leakage if errors are MA. Models with AR errors can also be estimated more easily. It is not very time consuming to use NLLS to estimate models with MA errors, but in some of the computationally intensive estimations, particularly the exponentially distributed lag models, it would be incon-

test for whether the residuals are serially correlated up to some given order) are Lagrange multiplier tests. There are also tests for whether the error process is identically distributed over time, the White test for general heteroskedasticity or Engle's test for autoregressive conditional heteroskedasticity (Harvey 1990, 172, 221–23). All specifications used in this article use these tests to make sure that the error process has the appropriate properties; given the purpose of this article, these tests are not stressed, and the heteroskedasticity tests are not reported.

12. The only approval issue that I know of where it is important to use the theoretically specified MA error is in testing hypotheses derived from the rational expectations hypothesis (Beck 1989).

13. This assumption about missing data is why I use the Hildreth-Lu (HL) grid search method of estimation instead of the more common Cochrane-Orcutt iterative procedure. The two procedures are asymptotically equivalent.

venient if the errors were MA. Let us now turn to the dynamics of the variables that are measured by the analyst.

Static Models

The simplest, and earliest, specification is that of no dynamics. Such a model was used in Mueller's path-breaking study of presidential popularity (1970). In the static model, only current economic terms enter the approval function; nothing this month has an effect on anything next month. Changes in X's immediately show up in approval, but have an effect for only one month. This seems unreasonable. Should one good month after two years of economic disaster lead to the same high level of approval as twenty-five months of economic boom? In the static model, approval fluctuates as freely as the business cycle.

The static specification also has severe econometric problems. In particular, as Hibbs demonstrated for the Mueller model (1974), the residuals from the static model will, in general, not be independent; thus, the coefficient estimates will not be fully efficient and the estimated standard errors will not even be consistent.

OLS estimates of the static equation in table 1 seem to show that both unemployment and inflation have a significant immediate impact on approval (with the one month lag built into the measures). While the coefficients are not overwhelming in absolute size, the unemployment coefficient is more than

TABLE 1. Static Model of Approval

	OLS		HL	
Variable	β	SE	β	SE
Constant	62.64	.67	55.53	3.19
V	−13.69	1.82	−2.59	3.30
W	−17.44	2.39	−17.66	3.46
E	5.10	2.14	4.94	.65
ΔU	−4.43	2.15	.12	.78
I	−1.26	.11	.05	.05
ρ	—	—	.94	.02
σ[a]	9.57		4.20	
df	415		413	
Q[b]	1734.74		60.21	
DW[c]	.56		—	

[a]Standard error of estimate.
[b]Ljung-Box statistic with $df = 60$.
[c]Durbin-Watson statistic.

twice its estimated standard error, and the inflation coefficient exceeds its standard error by a factor of ten. The Durbin-Watson statistic clearly shows that the model is misspecified, and that the error process is not independent. When corrected for serial correlation through the HL grid search technique, the seemingly significant economic effects disappear, and most of the "action" in the model appears to be in the correlated errors (with the error correlation exceeding .9). But the problems with the static model are more than econometric; the underlying logic of the model is flawed.

The static model claims that approval instantaneously adjusts to new information, and that prior information is of no consequence. The first claim is an argument for no "stickiness" while the second is for no "memory." Rejecting the no stickiness assumption argues for putting lagged approval on the right-hand side; rejecting the no memory assumption argues for putting lagged economic variables there. Rejecting either (or both) assumption(s) leads to the estimation of a distributed lag model. Almost all modern work on approval uses a distributed lag model. Such models are broken into two subtypes: infinite distributed lag models, where effects persist forever, and finite distributed lag models where lagged effects disappear in a finite number of periods. I start with the latter.

Exponentially Distributed Lags

A finite distributed lag model of approval has the form

$$A_t = \sum_{i=0}^{L} X_{t-i}\beta_i + \epsilon_t \,, \qquad (4)$$

where L is finite (L is the time it takes for all lag effects to disappear). Any of the various error processes may be adjoined to this model. Equation 4 requires the estimation of many parameters, too many if L is at all large; multicollinearity between the different lags of X also makes estimation of the model difficult. The typical solution is to constrain the β's in some manner, which both decreases the number of parameters to be estimated and makes the parameter estimates more stable.

There are a variety of parameterizations of the β's that are commonly used. These range from ad hoc constraints, such as the Almon polynomial method, which assumes the β's lie on a low-order polynomial, to constraints that are theoretically based. The most common method used in the approval literature is the exponentially distributed lag (EDL), where the effects of past economic events become less and less important (either because of forgiving or forgetting), dying off at a constant rate each month.

The EDL model can be written

$$A_t = \sum_{i=0}^{L} (X_{t-i}\lambda^i)\beta + \epsilon_t \,.^{14} \qquad (5)$$

The value of λ must be strictly between zero and one.[15] The EDL model has been used by Chappell and Keech (1985) and, in modified form, by Hibbs (1987). In specifying equation 5, it is important to make sure that the summation extends back only to the beginning of an administration and to omit the first observation of any administration.[16]

The EDL model (with appropriate finite summation) can be estimated most easily by using a grid search to estimate λ. Since λ must be between zero and one, we can impose a grid of appropriate fineness over the unit interval and then do OLS on the approval function for each value of λ. The value of λ that minimizes the sum of squared errors (SSE) is the estimate for λ, with the OLS estimates for the other parameters coming from the regression using this λ. The procedure yields maximum likelihood estimates. If the error structure is AR1 the procedure is similar, substituting the Hildreth-Lu algorithm for OLS. In table 2, column 1, the errors are assumed to be independent and so the grid search is combined with OLS. The Durbin-Watson and Q-statistics show that this assumption is untenable, and so, in column 2, the results of reestimating assuming an AR1 error process are shown. Diagnostic checks show that this model is consistent with the data.

Column 2 shows that both unemployment and inflation have significant impacts on approval. A tenth of a point increase in unemployment decreases approval the next month by about 0.15 percentage points; the estimate for λ (.94) shows that this effect dies out slowly, at a rate of 6 percent per month. A half-point increase in the inflation rate (inflation is much more volatile than unemployment) decreases approval by almost a point the next month. The EDL model builds in the assumption that the effects of unemployment and inflation die out at exactly the same rate, $1 - \lambda$.

14. In the approval literature, event variables such as V, W, and E, have the lag structure built in and hence are assumed to affect approval only contemporaneously. The estimations reported here use that convention. Thus, X contains only the economic variables for unemployment and inflation. The event variables are omitted from the general equations.

15. If λ is negative, then lagged X's alternatively increase and decrease approval, which makes no sense. If λ is greater than one, then lagged effects dominate contemporaneous effects, with the effects of older and older X's becoming larger and larger. Such a model makes no sense (and also violates stationarity).

16. In the standard EDL setup, the summation is infinite. Since there is only a finite amount of data used, the summation is broken up into a finite sum and the so-called truncation remainder, which consists of all the unobserved data (multiplied by λ^t if data before time period 1 are unobserved). In the approval model, the truncation remainder is simply set to zero, and truncation is assumed to occur at the beginning of each administration. This is easy to set up in RATS; it may be more difficult in other statistics packages.

TABLE 2. Estimations of Exponentially Distributed Lag Model of Approval

	OLS (1)		HL (2)		HL (3)		HL (4)	
Variable	β	SE	β	SE	β	SE	β	SE
Constant	68.94	.52	65.06	1.65	65.01	1.67	65.06	1.68
V	−14.90	1.25	−4.00	2.64	−3.72	2.74	−3.67	2.73
W	−12.98	1.60	−10.72	3.08	−10.49	3.09	−10.38	3.08
E	4.90	1.42	4.98	.61	4.97	.61	4.98	.60
ΔU	−2.23	.36	−1.44	.77	−1.68	.84	−2.19	1.07
I	−.26	.01	−.18	.02	−.16	.02	−.28	.03
ΔU_1	—	—	—	—	—	—	1.26	.94
I_1	—	—	—	—	—	—	.18	.05
λ	.93	—	.94	—	—	—	.90	—
λ_U	—	—	—	—	.87	—	—	—
λ_I	—	—	—	—	.95	—	—	—
ρ	—	—	.86	.03	.86	.03	.87	.03
σ[a]	6.34		3.76		3.75		3.72	
df	415		413		413		411	
Q[b]	974.35		51.92		51.64		51.10	
DW[c]	.42		—		—		—	

[a]Standard error of estimate.
[b]Ljung-Box statistic with $df = 60$.
[c]Durbin-Watson statistic.

This is a very strong assumption. It is possible to estimate equation 5 with different λ's for unemployment and inflation. This model is estimated with a two-dimensional grid search over both λ_U and λ_I, combined with a grid search over the AR1 parameter (results are shown in col. 3). The immediate impacts of both economic variables are similar to those from the single λ case; the values for λ_U and λ_I are also quite close, suggesting that the effects of both employment and inflation die out at about the same rate.[17]

Another possibility is that while, in general, the lagged effects decline exponentially, the process for the first (or first few) month(s) differs from the overall process. It is hard to specify a theoretical reason why this should be so, but it clearly may happen in practice. To check for this, X_t (and possibly its first few lags) can be added to equation 5 with free parameters on those variables. The null hypothesis that those free parameters are zero then can be

17. With the grid search, the standard F-test to test the equality of λ_U and λ_I is not strictly applicable. If we compute the statistic ignoring the issue of grid search, we cannot reject the null hypothesis of parameter equality.

tested to see if the early lag structure follows the overall exponentially declining lag structure.

Column 4 of table 2 reports the results of such a procedure to see if the initial effect of unemployment and inflation on approval is different than the subsequent lagged effects. (The two free parameters are in the rows labeled ΔU_1 and I_1 to remind us that each variable has its initial effect with a lag of one month.) We cannot reject the hypothesis that the coefficient of ΔU_1 is zero, that is, the lag structure on the unemployment terms is a declining exponential lag for all lags including the first. The story for inflation is different. The estimated coefficient is positive and significant. The effect of a one-point increase in inflation is a tenth of a point decline in next month's approval (the sum of .28 and −.18), with the subsequent effect being a quarter point decline (the produce of .28 and .90), with further effects declining at the rate of 10 percent per month. Thus, the initial effect of inflation on approval is less than that specified by the simple EDL model, but lagged effects are about half again as large as in the EDL model.

To see the dynamics of the EDL model, we need to look at the effect of both transitory and permanent changes in the economic variables. Using the estimates from column 2, a one-time, one-point decrease in the inflation rate (with a subsequent increase the next month) leads to about a fifth of a point increase in approval the following month. Of this increase, 94 percent persists the following month, and so forth. Thus the effect of a one-month decrease in inflation persists for over four years, with half the effect persisting for over a year (see the solid line in fig. 1).

If the first lagged effect is left free, using the estimates from column 4, the initial impact of inflation on approval is less, but the effect after two months is greater, with a more rapid subsequent decline. The combination of the greater early effect and the faster rate of decline means that the impact of inflation on approval is lower in the EDL model in the first year, but greater in subsequent years; with the first lag free, about two-thirds of the initial transitory impact of inflation dissipates within a year (see the long dashed line in fig. 1, labelled EDLFREE).

The effect of a sustained decrease in the inflation rate of one point is shown in figure 2. The first month's effect is identical to the transitory effect, but the permanent effect continues to build. In the EDL model (solid line) the effect increases to about a third of a point after two months, half a point after three months and about a point after six months. The long-run effect of this increase in inflation is a gain of about four points in approval. It takes a long time (over four years) for this effect to be fully realized. With the first lag free (long dashed line labeled EDLFREE), as in the transitory analysis, the initial impact of the permanent change is less than in the EDL model, but the two-month impact is greater. The long-run effect of inflation on approval with the

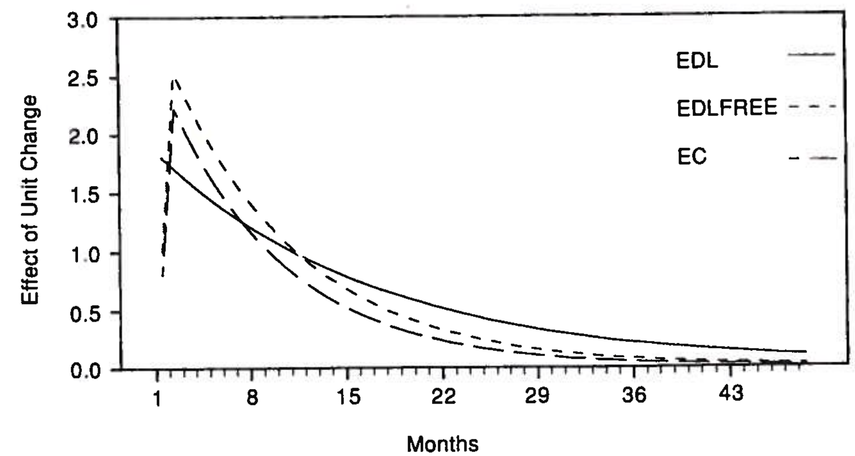

Fig. 1. Dynamic effects of a transitory unit decrease in inflation

first lag free is about two-thirds that of the EDL model; equilibrium, however, is reached much faster (essentially within two years). The effect of inflation on approval is not huge (one point is a large change to be sustained) but it is not insubstantial either.

It is interesting to note that the estimates of λ and ρ are similar. This means that the effect on approval of both the economic variables and unmeasured shocks die out at about the same rate. The similarity of all the

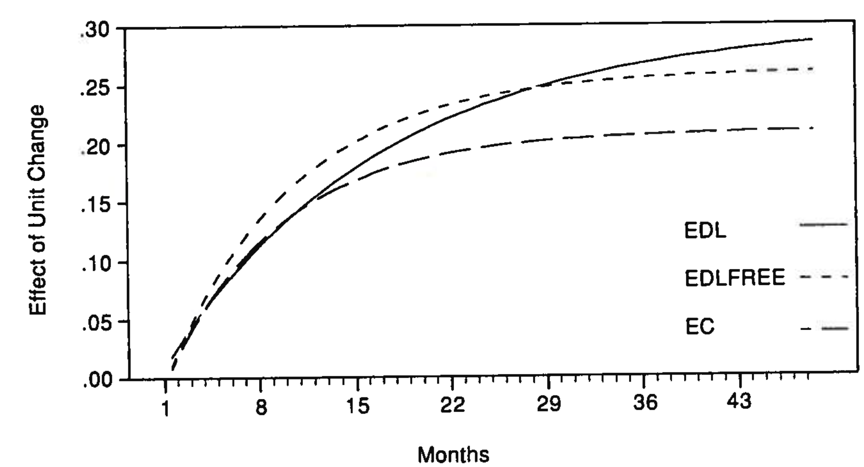

Fig. 2. Dynamic effects of a permanent unit decrease in inflation

dynamics in the EDL model allows for radical simplification of the EDL model. This simplification is essentially the partial adjustment model.

Partial Adjustment

Both sides of the equation 5 can be multiplied by the so-called Koyck (1954) transformation, $(1 - \lambda L)$. After simplification, this yields

$$A_t = X_t\beta + \lambda A_{t-1} + \epsilon_t - \lambda\epsilon_{t-1}, \qquad (6)$$

which looks very much like what is known as the partial adjustment (PA) model,

$$A_t = X_t\beta + \lambda A_{t-1} + \epsilon_t, \qquad (7)$$

except for the MA error process (with parameter λ).[18] Note that the MA parameters is λ, which is also the "speed of adjustment" parameter, so estimating equation 6 is a problem in constrained optimization.

The discussion so far has assumed that the errors in the original EDL model are iid. But this means that shocks to the system (unmeasured effects on approval) last only one period. Why should unmeasured effects die out any differently than measured effects? Given the time-series nature of the data, one would not expect the errors to be independent. Perhaps the most likely assumption is that the errors are generated by an AR1 process with parameter λ, that is, shocks die out at the same rate as lagged values of the measured variables. This is the message of table 2. But in that case, the Koyck transformed model is exactly equation 7 with iid errors and can be estimated easily via OLS.

A slight generalization of this is to assume correlated errors in equation 5 with $\rho \neq \lambda$. This leads to a model like equation 7 with an AR1 error process that can be estimated using the HL technique. The advantage of the PA over the EDL setup is that the PA model is easier to estimate with standard software (there is no grid search and no computing of the summations; leakage between administrations can be prevented by treating the first observation of each administration as missing). OLS and HL estimations of equation 7 are shown in table 3, columns 1 and 2. There is a slight amount of autocorrelation in the OLS errors, indicating that the autocorrelation in the EDL errors is not quite λ. Using OLS would be a lot better than estimating equation 6, but the HL estimates in column 2 appear best.

The coefficient on lagged approval in table 3 should be very close to the

18. If the error process in eq. 5 is ARMA(p,q), the process in eq. 6 is, by Granger's lemma, ARMA$(p, q + 1)$.

TABLE 3. Partial Adjustment Model of Approval

		OLS (1)		HL (2)		HL (3)	
Variable	Lag	β	SE	β	SE	β	SE
Constant		5.20	1.08	4.77	.98	6.11	1.07
V		−.98	.69	−.87	.61	−1.17	.62
W		−3.55	.88	−3.25	.77	−3.29	.78
E		9.43	.76	9.44	.75	9.40	.75
ΔU		−1.59	.75	−1.71	.72	−.94	.77
ΔU_1	1	—	—	—	—	−1.34	.76
I		−.104	.045	−.118	.043	−.056	.048
I_1	1	—	—	—	—	−.136	.048
A	1	.91	.017	92	.02	.90	.02
ρ		—	—	−.14	.05	−.11	.05
σ[a]		3.35		3.33		3.30	
df		414		412		409	
Q[b]		68.31		63.21		66.39	
$AR1$[c]		6.81					

[a]Standard error of estimate.
[b]Ljung-Box statistic with $df = 60$.
[c]Lagrange multiplier test for AR1 error, χ^2 with $df = 1$.

estimate of λ in table 2, and it is (.92 versus .94). The coefficients on inflation and unemployment are also reasonably close in the two tables, with the effect of inflation being estimated less precisely in the PA model. This is as it should be, and shows that the partial adjustment model can be used to more easily estimate the EDL model (so long as the unrealistic assumption of iid errors is not made for the EDL model).

The PA and EDL models are not identical. In the EDL model, only lagged economic variables have an effect on current approval; in the PA model, anything that affects prior approval affects current approval through the lagged approval term. Thus, for example, shocks have a different effect in the two models, having an effect only through the autocorrelated errors in the EDL model but also having an effect through lagged approval in the PA model. This accounts for the different estimates in the tables 2 and 3. Since the two models are not nested, it is hard to say which is superior, though the smaller standard error of estimate does give the nod to the PA model. But no substantive conclusion hangs on this choice.[19] In some cases the EDL model

19. Column 3 of table 3 is the analog of Column 4 in table 2. Both columns show that the effects of unemployment and inflation on approval vary nonsystematically over the first few periods.

would have an advantage in that separate speeds of adjustment for inflation and unemployment (table 2, col. 3) can be set up more easily in that model.[20]

Thus far I have used the PA model to simplify estimation of the EDL model. Can the PA model be considered as a model on its own? (This is an issue in interpretation because, obviously, the Koyck transformation ties the estimates of the two models.) The standard story behind the PA model is that it is costly (to some optimizing agent) to adjust the dependent variable. The X's indicate the optimal A, but the realized A is an average of its optimal value and its realized lagged value. This model makes sense for inventory control or reaction functions, where it is costly to move the dependent variable. The alternative story is one of inertia. The X's again indicate how much A should move, but, because of inertia or sluggishness, the realized A is an average of the optimal A and its past value. Kernell (1978 and 1986) uses a similar argument to justify his PA approval function. He claims that

> [t]he president's current popularity reflects the level of approval during the preceding month. This proposition suggests that the president's popularity will respond sluggishly to environmental forces. During the brief interval between observations, many citizens will maintain their assessments of the president's performance regardless of intervening events. The built-in inertia of popularity is revealed by the fact that the best information available for predicting an individual's future evaluation of the president is his or her current evaluation. (1978, 515)

This is a justification for writing a model in terms of change in approval (i.e., putting lagged approval on the right-hand side with a coefficient of one) but it is not a justification for the partial adjustment model. Why should there be inertia in public opinion? What are the costs of rapid swings in one's evaluation of the president? Why should (individual) approval not adjust fully each month? Just because the PA story has proven useful in economics does not make it a natural story for political science.

If we take Kernell's information argument seriously, we end up not with a PA model but a Bayesian updating model, where voters update their opinions of the president as new information becomes available.[21] Consider a single individual and let A_t^i be that individual's approval rating (on a continuous scale) of the president. Suppose now that individuals are not sure of

20. The model in table 2, column 3 can be transformed into a model that looks like the partial adjustment model through two applications of the Koyck transformation, one with each λ. In practice, the PA model leads us to ignore the issue of different speeds of adjustment.

21. I only sketch the model here. Details and tests of the model are in Beck 1991.

their approval rating, but have instead a probability density over approval, $g_t(A_t^i)$.

Suppose there is some relationship between the state of the economy, X_t, and approval. Suppose also that presidents who are more highly approved will, in general, provide better economic outcomes. Bayes's theorem says that the presidential approval density at time $t + 1$, $g_{t+1}(A_{t+1}^i)$, will be updated proportionally to $f(X_t|A_t^i)g_t(A_t^i)$ where f is the conditional density of economic outcomes given approval. (The assumption here is that one supports competent presidents and competent presidents do a better job of managing the economy.)

Observed approval in the survey, A_t, is just the average A_t^i. Thus, Bayesian updating, with aggregate data, yields

$$A_{t+1} = \beta f(X_t/A_t)A_t + \epsilon_t, \tag{8}$$

where ϵ represents unmeasured factors that affect approval. (This assumes that approval is the mean of the approval density, g.) The difference between this form and the partial adjustment form is that the economy enters multiplicatively, and its effect depends on whether economic outcomes are consonant with prevailing presidential approval (representing prevailing beliefs about presidential competence). Only "surprising" economic outcomes should modify current beliefs.

It is hard to give empirical content to this notion of surprise. The point here is that the PA model cannot be justified by an inertial argument without some indication about why approval should move sluggishly and some model of information processing. It seems hard to justify the partial adjustment model of approval except as a Koyck transformation of an EDL model.[22] Thus, I think it best to conceive of the PA model as a simplified way of estimating the EDL model; at that point, the small differences between the two models should be borne in mind.

The First Difference Model

What about using Kernell's story to justify a model with change in approval as dependent variable and changes in the economy as independent variables? The first difference model has

$$\Delta A_t = \Delta X_t + \epsilon_t. \tag{9}$$

22. It must be stressed that this is a statement about approval, not partial adjustment. The partial adjustment model has a good theoretical basis in many other areas of interest.

Results based on such a model are shown in table 4. Empirically, the model does not perform very well. Both economic coefficients have the wrong sign and are small as compared to their standard errors. (If we think that U belongs in the model for approval, so that ΔU belongs in equation 9, the coefficient of ΔU at least has the right sign, but is not statistically significant.)

While we often choose specifications because they give us the coefficients we (or our referees or editors) want, the lack of positive findings for the first difference model can hardly be decisive against it. There are, however, important theoretical arguments against the simple model in first differences. A technical argument is that if the error process in the equation for the level of approval is iid, the error in the first difference model will be MA1. It is hard to design a plausible error process for the level of approval that leads to iid errors for equation 9.

A more serious problem is that equation 9 says that short-term changes in the economy lead to short-term changes in popularity, but that there are no effects of the economy on approval that last longer than one month. According to this model, voters respond to a small improvement in the economy at the bottom of a depression in the same way as they do at the top of a boom. Note also that if the first difference model gets "off track" in that high approval coexists with a weak economy or vice versa, there is no way of it ever getting back on track. In the former case, as the economy improves approval only increases. One good-sized shock can throw the first difference model off track forever. This problem was noticed by Davidson et al. in their model of

TABLE 4. First Difference Model of Approval

	OLS		OLS	
Variable	β	SE	β	SE
Constant	−.29	.18	−.27	.18
ΔV	1.76	3.18	1.61	3.18
ΔW	−9.89	3.77	−9.71	3.74
ΔE	4.92	.57	4.90	.56
$\Delta^2 U$	.14	.67	—	—
ΔU	—	—	−1.13	.84
ΔI	.04	.04	.04	.04
σ[a]	3.75		3.74	
df	414		414	
Q[b]	58.12		57.11	
DW[c]	2.06		2.07	

[a]Standard error of estimate.
[b]Ljung-Box statistic with $df = 60$.
[c]Durbin-Watson statistic.

the consumption function (1978). To deal with the problem they devised a model that has come to be known at the "error correction" (EC) model.

Error Correction

The EC models starts with equation 9, but adds another term to make sure that approval and the state of the economy stay "on track." The EC model is thus based on an long-run equilibrium notion, in this case that the economy and approval are in equilibrium, and so approval cannot remain high for long in the face of a poor economy. The EC model thus adds an "error correction mechanism" (ECM) to equation 9. The ECM measures how far the economy is out of equilibrium with approval; the coefficient on the ECM measures the speed at which the approval returns to its equilibrium value (in terms of the economy). The model is thus

$$\Delta A_t = \Delta X_t \beta + \gamma(A_{t-1} - X_{t-1}\nu) + \epsilon_t. \tag{10}$$

The term in parentheses is the ECM, which keeps approval "on track."[23]

The EC model allows for a much more sensible treatment of levels and changes. We have already seen that, in previous models, it is not clear whether levels or changes belong in the approval function, and it is perhaps bothersome to have an approval function based on the level of inflation but the change in unemployment. The EC model allows both levels and changes of the economy to affect approval in what seems like a reasonably satisfying manner: changes in the economy have a short-run effect on approval, but in the long run the level of the economy should effect the level of approval. Sensible treatment of levels and changes is one of the major advantages of the EC model.

The EC model is now very popular because any set of variables that are cointegrated can be represented in the EC form (Engle and Granger 1987). Two variables are cointegrated, loosely speaking, if, even though neither is stationary, the two are in an equilibrium relationship, so neither wanders far from the other for very long. Cointegration is only of interest if the dependent variable is nonstationary. Approval, in my long sample, is stationary.[24] The

23. The error process may be iid, AR, or MA. It is usually written as an MA process for notational convenience (Granger and Newbold 1986, 224–26), but for consistency with the other models in this article I assume that the errors are AR if they are not iid. In estimation, I assume that the event variables belong in the short-run portion of the model, and that all error correction is based on the economic variables only.

24. Ostrom and Smith (1990) have used cointegration to study approval. They work with only the Reagan administration and, in such a short sample, cannot reject the null hypothesis that approval is a random walk.

appropriate test for stationarity (against the null that approval is a random walk) is the Dickey-Fuller test (1979). This test is done by estimating

$$\Delta A_{t+1} = \beta A_{t-1} + c + \epsilon_t, \tag{11}$$

and then examining the ratio of the estimate of β to its standard error, the Dickey-Fuller statistic.[25] For the entire sample period (omitting the first month of each administration), the Dickey-Fuller statistic is 3.21. Using MacKinnon's (1991) correct tables, this statistic is significant with a P-value of about .03, so we can reject the null hypothesis that approval is a random walk in favor of the hypothesis that approval is a stationary process. But error correction is still of interest. The ECM enforces a tighter relationship between the economy and approval than does the simpler first difference model. Thus, error correction is relevant even if cointegration is not.

Thinking about error correction rather than cointegration also improves modeling. The EC model is asymmetric between variables, with a clear distinction between left- and right-hand side variables; the relationship between cointegrated variables is more symmetric. Thus, in the EC model for approval, approval adjusts if it is out of equilibrium with the economy, but the economy does not adjust to move into equilibrium with approval. This seems sensible. If we think in cointegration terms, it is too easy to allow equilibrium by adjustment of all variables simultaneously.[26] This may be the right assumption for some situations, but seems incorrect for modeling approval.

The results of estimating the EC model are shown in table 5. The OLS estimation shows a bit of autocorrelation. The second column (HL) shows reestimates of the model using an AR1 error process. The two columns give similar estimates; I work with those in the HL column, which, correcting for autocorrelation, are slightly superior. The estimations in table 5 drop the first two observations for each administration to cut the linkage between administrations.

All of the economic variables are estimated with the correct sign and all economic coefficients are (at least marginally) statistically significant. The estimated coefficient for ΔU is similar to that in the EDL model (table 2, col. 2); a monthly increase in unemployment of, say, one tenth of a point decreases approval the next month by about two tenths of a point. Again, the economy has some, but not a huge, effect on approval.

The coefficients on the level of inflation are also similar in both EDL and EC models. But note that the interpretation of these coefficients is different. In

25. This test assumes that approval does not show a linear trend and the ϵ's are iid. Tests show that both assumptions are consistent with the data.

26. Ostrom and Smith, for example, allow the economy to adjust to approval.

TABLE 5. Estimations of Error Correcting Model (dependent variable = first difference of approval)

		OLS		HL	
Variable	Lag	β	SE	β	SE
Constant		9.78	1.78	8.52	1.62
V		−2.16	.81	−1.85	.73
W		−3.95	.92	−3.64	.83
E		9.20	.74	9.26	.74
ΔU		−1.62	.76	−1.72	.74
ΔI		−.084	.047	−.081	.048
		ECM			
A	1	−.13	.02	−.11	.02
U	1	−.25	.12	−.21	.11
I	1	−.25	.06	−.23	.05
ρ		—	—	−.12	.05
σ[a]		3.29		3.28	
df[b]		406		404	
Q[b]		60.73		58.13	
$AR1$[c]		3.87		—	

Note: Equation estimated is
$\Delta A_{t+1} = C + \beta_1 V + \beta_2 W + \beta_3 E + \beta_4 \Delta U_{t-1} + \beta_5 \Delta I_{t-1} + \beta_6 A_{t-1} + \beta_7 U_{t-2} + \beta_8 I_{t-2} + \epsilon_t$.

[a]Standard error of estimate.

[b]Ljung-Box statistic with $df = 60$.

[c]Lagrange multiplier test for AR1 error, χ^2 with $df = 1$.

the EDL model, the level of inflation is a direct stimulus to popularity, while the EC model it works through the equilibrium mechanism. Note also that inflation enters with a one-month lag in the EDL model but with a two-month lag in the EC model.

A few examples can help give a feeling for the workings of the EC model, as well as a comparison of the EC and EDL models. To interpret the ECM we need to transform the results of table 5. The ECM can be written as

$$.11(A_{t-1} - 1.9U_{t-2} - 2.1I_{t-2} + C), \tag{12}$$

where C is an undetermined constant. This constant means that we cannot estimate the equilibrium level of approval for a given economic situation.[27]

27. Davidson et al. (1978) get around this problem by assuming that their economic variables show long-run proportionality. Thus, they need not estimate C. Having no natural zero, we in political science are not so fortunate.

Suppose that at some time, say December, approval is consistent with the overall state of the economy, so there is no error correction in January. Now suppose there is a one-time decrease in inflation of one point in January. The effects on subsequent approval are shown in figure 1 (long dashed line, labeled EC). February would show a 0.08 point increase in approval. The equilibrium level of approval in February will have increased two points and, hence, approval will be about 1.90 points below equilibrium. In March there will be two effects. Since by assumption the decrease in inflation was temporary, it must increase in February by a point. This increase leads directly to a 0.08 point decrease in approval in March. But this is offset because 11 percent of the February error is corrected in March, leading to a 0.20 point net increase in approval. Since the economy has now returned to its December state, equilibrium approval in March is about a quarter of a point too high. This error is corrected at the rate of 11 percent per month after March, so approval decreases by about 0.025 points in April, another tenth of a point through September and another 0.05 points through the next March, with 75 percent of the error having been corrected by then. It takes about two years for the temporary decrease in inflation to no longer have any noticeable effect on approval.

Note the pattern here is more complicated than the pattern shown for the EDL model. That model showed an initial increase in approval of a fifth of a point in February compared to the 0.08 point increase in the EC model. This increase disappears exponentially at a rate of 6 percent per month in the EDL model. In the EC model, on the other hand, approval continues to increase in March. Only then does the effect of the transient change exponentially decline. The EC model has a slower initial increase in approval than the EDL model, with a more rapid subsequent erosion of that effect in the EC model. The transitory impact of inflation disappears from the system about a year more quickly in the EC model.

Figure 2 (long dashed line labeled EC) shows the effect of a sustained one point decrease in inflation. Again, assume the change occurs in January and that initially approval was in equilibrium. February shows the short-run effect of a 0.08 point increase in approval. Since the decrease in inflation is sustained, there are no subsequent short-run effects, but there are continued error corrections as approval slowly moves to a new equilibrium. In February, equilibrium approval will be about two points above actual approval. Error correction eats away at this difference at the rate of 11 percent per month, so March shows an increase in approval of about a fifth of a point. After six months, approval will have increased one point, and, after a year, a point and a half. The new equilibrium level of approval is two points higher than before. It takes about a year and a half to reach this new equilibrium. This compares to a long-run gain in approval in the EDL model of three points, two of which are gained in the first two years. The new equilibrium is reached considerably more quickly in the EC model.

The error correction model and the exponentially distributed lag model provide similar results about the effect of the economy on approval. This should not be very surprising because both the EDL model and the EC model can be transformed into something that looks quite similar to the partial adjustment model. If we take the EC model and write out the change variables explicitly, we get

$$A_t = (1 + \gamma)A_{t-1} + X_t\beta - X_{t-1}(\gamma\nu + \beta) + \epsilon_t. \tag{13}$$

This looks like the partial adjustment variable with the addition of economic variables lagged an extra period. That addition is unlikely to make a major difference empirically. Thus, both the EDL and EC models may fit the data about equally well. But the interpretation of the results is quite different depending upon which model we think actually generated that data. The advantage of error correction lies not in obtaining very different estimates, but in providing a meaningful framework for thinking about both short- and long-run movements in approval, or, alternatively, the role of levels and differences in the approval function.

Transfer Function Models

A somewhat different approach to the study of approval uses the Box-Jenkins transfer function methodology (1976). This methodology comes from industrial engineering, where the task is to derive the relationship between the inputs and output of some industrial process. The transfer function approach has been heavily used in the study of approval, with its foremost exponents being Norpoth (1986 and 1991) and MacKuen (1983; MacKuen, Erickson, and Stimson 1989).

It is easiest to present the transfer function model in terms of lag polynomials. If L is the lag operator, the lag polynomial $A(L)$ is just

$$A(L) = \sum_{i=0}^{T} a_i L^i . \tag{14}$$

The transfer function model (for two independent variables, W and Z) is

$$A_t = \frac{B(L)}{C(L)} W_{t-r} + \frac{D(L)}{E(L)} Z_{t-s} + \frac{\Theta(L)}{\Phi(L)} \epsilon_t . \tag{15}$$

By convention, the denominator polynomial is $1 - c_1L - \ldots - c_TL^T$; r and s are delay parameters, giving the time it takes for a change in an input to show up in approval.

While this model appears different from the AD models previously considered, this difference is superficial. At first glance it appears as though transfer functions are similar to finite distributed lag models, with only current and lagged explanatory variables, but not lagged approval, explaining current approval. This is true if the denominator polynomials (C, E, and Φ) are all one. Then equation 15 reduces to

$$A_t = B(L)W_t + D(L)Z_t + \Theta(L)\epsilon_t, \tag{16}$$

which is just the finite distributed lag model with a moving average error process.

But if there are denominator lag polynomials (that are not one) in either the transfer function for the explanatory variables or the error process, then the transfer function becomes an infinite distributed lag model. For example, assume that D is zero, Θ and Φ are both one, and C is a first-order polynomial. Then equation 15 reduces to

$$(1 - dL)A_t = B(L)W_t + (1 - dL)\epsilon_t. \tag{17}$$

This is exactly the Koyck transformation of the EDL model (with a slightly more complicated lag structure on W). We can also transform equation 15 by multiplying all sides by the denominator fractions (first eliminating common factors, if any). This gives

$$C(L)E(L)\Phi(L)A_t = B(L)E(L)\Phi(L)W_t + D(L)(CL)\Phi(L)Z_t + \Theta(L)C(L)E(L)\epsilon_t, \tag{18}$$

which is known as an ARMAX (AutoRegressive Moving Average with eXogenous variables) model. In this model, approval is a function of several of its own lags, a complicated finite distributed lag of W and Z and a complicated moving average error. Thus, all standard dynamic models can be seen as imposing some constraints on equation 15, which makes the transfer function model the most general of all dynamic models (Harvey 1990, 264–66).

Equation 15 can be estimated by standard methods, usually NLLS. There is some ambiguity about what assumptions to make about data points in the model that are not observed (because they occur prior to the sample period),

but the particular assumption is not critical in large samples.[28] All standard statistical tests and methods may be used to evaluate transfer function models; there is no fundamental dichotomy between transfer function models and the other models considered previously.

There does appear to be a difference in the practice of transfer function methodology and more traditional autoregressive distributed lag methodology. The latter uses standard econometric methods, starting with a general model and then testing restrictions, moving to a more parsimonious model if the data so indicate. This is easy to do in the AD setup. The initial model can have long dynamics (in X_t, lagged approval and the error process). Constraints can then be imposed on this general model, with simplification accepted unless the data indicate such simplification is too costly (in terms of an increase in the estimated standard error of estimate). All this can be done through the usual sequence of nested F-tests.

Transfer function modelers usually do not follow this sequence, but instead try to let the data indicate a single model. This model is suggested by first transforming each input (explanatory variable) to an iid process ("prewhitening"), transforming approval (using the same transformation applied to the input), and then examining the cross-correlations at all lags between the whitened input and transformed approval. These cross-correlations suggest a form for the transfer function (Norpoth 1986). This sequence is repeated for each explanatory variable in the model. The model is estimated (by NLLS) and checked for adequacy; at a minimum, the model residuals must be iid. Given their tools, the Box-Jenkins modelers seem to be more careful about the fit of the model to the data than do AD modelers, who often seems happy to assume that only relatively short lags should be examined. Transfer function modelers, on the other hand, check for long lags in their dynamics, letting the data indicate whether such lags should be included in the model.[29]

28. Programs from the Box-Jenkins tradition engage in "backcasting," that is, using the model to estimate unobserved data and then reestimating with the new data; those with an econometric orientation find it easiest to treat the first several observations as fixed, estimating conditional on those first fixed observations (Harvey 1990, 241–42). The latter assumption makes maximum likelihood estimation fairly straightforward, with NLLS providing the maximum likelihood estimates.

29. A perhaps extreme example of this is the work of Whitely (1984). His approval function for the United States has the first difference of approval being affected by the first difference in unemployment lagged 14, 24, and 30 months, the first difference in inflation lagged 2 months, and an MA error process with lags of 1 and 15 months. These lags are indicated by the data and are appropriately chosen with the Box-Jenkins methodology. An AD modeler would almost certainly choose a simpler model, with a gain in parsimony but a loss of in-sample fit. It would seem that Whitely's result is very sample specific.

Transfer function modelers usually do not start with a general model and then test restrictions on that model. They certainly could, since AD models are a subset of transfer function models. But complicated transfer function analysis does not lead the modeler in this direction. Part of the problem is that transfer functions do not present the obvious nesting of models as in the AD framework; another part of the problem is that tests for common factors in the lag polynomials are harder to construct in the transfer function framework (see note 31); a third part of the problem is that concerns for picking up all the dynamics featured in a sample set of data cuts against a concern for parsimony.

The transfer function methodology looks very arcane. It is best suited to tease out the complicated relationship between a single input and a single output (Norpoth 1986). With more than a single input, the prewhitening and cross-correlation methodology becomes much more difficult, being similar to trying to specify a multiple regression model by looking only at bivariate correlations. While transfer function modelers often make inferences from their model selection diagnostics, in the end they have a statistical model and can draw all relevant inferences from the estimated coefficients. Thus, there is no reason to worry about the whole debate about prewhitening and the examination of bivariate cross-correlations (Harvey 1990, 248). The Box-Jenkins took kit is a useful, but not the only, way to specify a transfer function model.

Transfer function analysis is ideal for estimating the exact relationship between a single input and approval. This is because the transfer function, $B(L)/C(L)$, gives the precise relationship between input and approval; if all other inputs (including the error) are set at zero, then this transfer function completely specifies the relationship between input and approval, making interpretation straightforward. This is particularly useful when the input is a "pulse," that is, a dummy variable that marks some specified event, such as a war. It is often of interest to know how quickly approval responds to such an event, the shape of the response (that is the pattern of lagged effects), and the length of time it takes before the effect dies out. Norpoth (1991) has probably made the best use of transfer functions in this manner.

His interest was the effect of the Falklands War on Prime Minister Thatcher's reelection. This was studied by estimating a transfer function of a pulse that marked the month of the war. Norpoth showed that the war had a great affect on Mrs. Thatcher's approval, but that most of that affect dissipated before the election. Specifically, he showed that the war increased Mrs. Thatcher's approval by 5 points in April, 1982, and 15 points in May, 1982, to be eroded at a rate of about 7 percent per month thereafter, leaving a small residual impact on her approval for the election of 1983. Norpoth is here working like an historian. He was not attempting a general theory of the effect of war on approval, but a specific study of how the Falklands War affected Mrs. Thatcher's approval. For this he wanted an estimate of the transfer function that fit the data as well as possible (while retaining some parsimony). There was no interest in extrapolating these results to other samples.

An alternative is to use transfer functions in a more theoretical manner, as in the various works of MacKuen. MacKuen takes advantage of the fact that each separate transfer function can be used to theoretically understand the relationship between a single input variable and approval. His general approach is that a temporary, one-time change in an input should have an effect that died out slowly. Thus, there is an equilibrium relationship between each input and approval. The transfer function for any input can then be modeled as $\beta/1 - \delta L$, where different parameters β and δ are estimated for each input variable. If $|\delta| \leq 1$, then the transfer function can be rewritten as

$$\beta \sum_{t=0}^{\infty} \delta^t , \tag{19}$$

so the effect of the input dies out an exponential rate, δ. (Note the similarity to the EDL story. If all transfer functions have the same denominator, this transfer function model is exactly the EDL model.)

The advantage of such a model is that it is possible to examine the relationship between any input and approval by simply looking at the estimated transfer function.[30] Unlike, say, the PA model, no dynamics are contained in lagged approval terms. This makes interpretation straightforward. Thus, for example, MacKuen (1983) used transfer functions to study the immediate and long-term impacts of various types of events (including economic events) on approval. The economy is seen to have a substantial effect on approval, but it has a smaller, albeit more long-lived effect than do some more dramatic events.

The difficulty of this type of model is that it is much less parsimonious than, say, the EDL model, since the relationship between each input and approval has a separate parameterization. (Of course, this may be a strength if the EDL model imposes an incorrect constraint.) The transfer function model is also more costly to compute, though this is not a major factor. More seriously, if the transfer functions for the different inputs contain approximate common factors, then it is very hard to get precise estimates of the parameters

30. This approach depends heavily on the inputs being independent, so it is possible to examine the effect of a single input setting all other inputs to zero.

in those factors.[31] It is also very difficult (if not impossible) to ensure that events in one administration do not affect approval in a subsequent administration; effects persist forever if there are nay denominator polynomials in the transfer functions. But perhaps most important, there is no well-defined methodology for choosing a parsimonious but empirically valid transfer function in the multivariate case.

I estimate a series of transfer function models using the common data set. Each contains a dummy variable to mark the start of a new administration; these cannot prevent the leakage of information from the previous administration, but they lessen its impact.[32] Results of the estimation are shown in table 6. Models A through C deal with the effect of the economy on approval, while models D through F show the impact of the Cuban Missile Crisis on President Kennedy's approval. (To save space, extraneous coefficients are omitted in models D through F; estimates are similar to model A.)

The best-fitting transfer function for the error process is ARMA(1,1) as in models A and C; the AR1 model B is distinctly inferior, with a very high Q-statistic. The resulting errors in models A and C appear to be independent and the models appear adequately specified (with a standard error of estimate slightly higher than in the EDL model, with most of this increase being caused by the inability to discard the first observation of each administration). We can get a feeling for the meaning of the ARMA error by ignoring all the other inputs. This gives (from model A)

$$A_t = \frac{1 + .12L}{1 - .85L}\epsilon_t , \tag{20}$$

which can be written

$$A_t = .85A_{t-1} + \epsilon_t + .12\epsilon_{t-1}. \tag{21}$$

This is a simple ARMA representation of approval.

31. Each lag polynomial can be factored into a series of linear terms. If all polynomials contain the same factor, this term should be factored out. But since the coefficients of the polynomial are estimated, the estimates will never coincide exactly. This is what causes the problem. Ignoring common factors causes large standard errors. But Box and Jenkins note "[i]n practice we shall be dealing with estimated coefficients which may be subject to rather large errors, so that only approximate factorization can be expected, and considerable imagination may be needed to spot a possible factorization" (1976, 387). Sargan (1980) has developed a methodology to test for common factors, but this methodology seems to be used primarily by AD modelers.

32. This inability to prevent leakage from administration to administration is a major weakness of the transfer function model.

The transfer function for inflation is $-.16/(1 - .94L)$ with a delay for two months. Thus, inflation affects approval after two months (shades of error correction, but without the underlying story). One strength of the transfer function methodology is that it forces us to confront the issue of delay from input to output, though no theoretical reason is offered for that delay. A one-

TABLE 6. Transfer Function Model of Approval

Variable (Delay)	Model A β^a	Model A SE	Model B β^a	Model B SE	Model C β^a	Model C SE
Constant	66.81	2.47	68.81	.78	127.15	2011.33
JFK1[b]	6.62	2.81	3.05	3.28	6.48	2.80
LBJ[b]	9.74	2.79	8.96	3.25	10.02	2.81
RMN1[b]	5.26	2.94	3.63	3.27	5.25	2.97
GKF1[b]	1.82	3.19	0.99	3.37	1.96	3.21
JEC1[b]	4.51	2.78	8.74	3.23	4.50	2.80
RWR1[b]	6.87	2.78	9.73	3.23	6.60	2.90
V	−5.41	2.92	−16.79	1.74	−5.45	2.95
W	−17.00	3.12	−19.20	2.11	−16.78	3.14
E	5.07	.61	5.40	.73	5.04	.61
$\Delta U(2)$	−1.37	.70	−2.09	.90	0.12	.28
δ	—	—	0.956	.007	1.000	.010
I(2)	−.16	.05	−.12	.02	−.16	.06
δ	.937	.023	.956	.007	.916	.038
ϕ_1	.85	.03	—	—	.84	.03
θ_1	.120	.057	.770	.030	.130	.060
σ^c	3.87		5.04		3.89	
df	411		410		410	
Q^d	58.68		717.41		61.79	
$AR1^e$	—		63.97		—	

Variable (Lag)	Model D^f β^a	Model D^f SE	Model E^f β^a	Model E^f SE	Model F^f β^a	Model F^f SE
CUBA[g]	4.57	2.78	11.22	3.86	7.63	3.35
δ	—	—	.913	.122	—	—
CUBA(1)	—	—	—	—	5.500	3.353
σ^c	3.88		3.86		3.87	
df	410		409		409	
Q^d	58.27		62.64		60.59	

[a]Unless otherwise noted in the variable label.
[b]Dummy variable, set to 1 in first month of administration.
[c]Standard error of estimate.
[d]Ljung-Box statistic with $df = 60$.
[e]Lagrange multiplier test for AR1 error, χ^2 with $df = 1$.
[f]All other variables as in model A.
[g]Cuban missle crisis dummy (November, 1962).

point, one-time decrease in inflation leads to a .16 point increase in approval two months later; this impact dies away at a rate of about 6 percent per month. This estimate is very close to that obtained in the EDL or PA analyses.

The effect of unemployment is a bit harder to determine. The assertion of model A is that unemployment affects approval with a delay of two months; a one-time, one-point decrease in unemployment leads to a 1.37 point increase in approval two months later. This immediate effect is similar to that seen the EDL or PA models. But the dynamics here are different. The transfer function for unemployment lacks a denominator polynomial. This means that the effect of unemployment on approval only lasts for one period, disappearing completely after that.[33] The transfer function analysis shows that the dynamic effect of inflation on approval is different than the dynamic effect of unemployment. This is a strength of transfer function analysis.

Models D through F study the impact of the Cuban missile crisis on President Kennedy's approval. The input here is a pulse that is one in November, 1962, and zero otherwise. (The missile crisis is obviously omitted from the general event variable.) In model D, the assumption is that the crisis simply increase approval in November (that is, the transfer function is simply β). Two alternatives appear superior. In model E, the transfer function is $\beta/(1 - \delta L)$, that is, the crisis had an immediate impact on approval that died out exponentially. In model F, the transfer function is simply $\beta_1 + \beta_2 L$, that is, the crisis had an effect distributed over exactly two months. (Combining the two transfer functions made the estimated standard errors huge, indicating an overparametarized model.) The exponential model E is slightly superior to the finite distributed lag model F. Thus, the effect of the Cuban missile crisis was an immediate increase in President Kennedy's approval of over 11 points. This increase died out slowly, at a rate of about 10 percent per month. One defect of this transfer function model is that the effect of the Cuban missile crisis lingered (a bit) into President Johnson's approval rating.

Students of approval seem firmly divided into two camps: transfer function modelers and AD modelers. There is no mathematical justification for this; transfer function models can easily be transformed into AD models. The transfer function approach comes complete with its own took kit. But there is nothing different about transfer functions that make the AD tool kit useless, and the transfer function tools (prewhitening and cross-correlation analysis) are hard to use for multivariate analysis. The two approaches seem to lead researchers in somewhat different directions, with transfer function modelers

33. Model C estimates a transfer function for unemployment that shows the same dynamics as that for inflation. In that model, unemployment has no statistically significant impact on inflation (the estimate of the numerator β is well under its standard error). With an insignificant numerator, the dynamics implied by the denominator polynomial are irrelevant.

being more data driven. The transfer function seems ideal for studying the impact of specific events on approval, but, as MacKuen has shown, it can work well for theoretical studies of approval. Transfer function models are generally less parsimonious than their AD counterparts, but the other side of that coin is that transfer function models are more flexible.

How should we choose whether to use a transfer function or AD model? The AD model leads to a more well-defined sequence of tests for model specification, from general to specific (Hendry, Pagan, and Sargan 1984), whereas the transfer function model is at its best in estimating a complicated transfer function for a single input. Neither approach is generally superior, but the transfer function approach comes into its own where interest focuses on a single determinant of approval, as in Norpoth's work. Where interest is in the full multivariate analysis of approval, the econometric tool kit that comes with the AD approach makes it, for me, superior. Some models (such as error correction) also arise more naturally in the AD context. But, in the end, the two approaches are mathematically equivalent and so must yield statistically equivalent results. Choice between the two approaches must then be driven by differences in modeling philosophy, not the data.

Modeling Individuals or Aggregates

Thus far I have looked at the dynamics of aggregate models of approval. Can anything be gained by starting with a dynamic model of individual approval? Such a modeling strategy, which appears to be based in the theory of rational choice, is today very popular. Such a strategy also appears promising for dealing with some important controversies; the arguments over, for example, "sociotropic" voting (Feldman 1985) are framed at the individual level and so models based on individuals should be better suited to deal with this controversy. The important recent work by Hibbs (1987, chap. 5) is based on a model of individual approval; his grouped logit analysis appears to give some remarkable results. But are models that start with individual approval superior to those that start directly with aggregate approval? If we had individual-level data, the answer would be obvious; with only aggregate data the answer is much less clear.[34]

Let A_t^i represent the approval level of a representative individual at time t. Assume this is measured on a continuum and is converted into an approve-disapprove dichotomy depending on whether it exceeds some threshold. Let X_t^i be out individual's perception of economic (and other conditions) and let ϵ_t^i

34. I exclude here analyses based on large subsets of individuals, say all Democrats or all Republicans. Such analyses are highly instructive, but, methodologically, they are no different than fully aggregated analyses.

reflect all other, unmeasured variables. Then the individual approval function has

$$A_t^i = X_t^i\beta_i + \epsilon_t^i. \tag{22}$$

If we average this equation over the whole population or sample, we get the basic approval function,

$$A_t = X_t\beta + \epsilon_t, \tag{23}$$

where A_t and X_t are averages of the individual-level variables. So far starting with an individual level model has not gotten us very far.

Suppose we break up the independent variables in equation 22 into those that are measures of the overall economy (perceived identically by all) and those that are idiosyncratic to individuals (say unemployment status or income). Let these be denoted X_t^i and Z_t (since Z is perceived identically by everyone). The relative impact of X and Z is of great interest, since X corresponds to self-interested voting (or approval) and Z seems to have sociotropic connotations (people approve of the president because they are doing well versus people approve of the president because the economy is doing well). Suppose we then estimate an aggregate approval function, which can only contain Z (since the X's vary from individual to individual, and can best be thought of as independent deviations around Z). If the coefficient on Z is large, does that not tell us something about sociotropic voting?

Unfortunately the answer is no. Granger (1990) has shown that Z may have essentially no impact on individual approval but a huge impact on aggregate approval. This is because Z is what he calls a "common factor," so even if Z has a very small impact on each individual's approval, when we add these effects over thousands of people we find Z to have a large effect in the aggregate model. Granger has shown that it is very difficult to make inferences from individual to aggregate models, and vice versa, when there are common factors in the individual model. So starting with an individual-level model tells us little about sociotropic approval.

Now suppose that each individual perceives the same economic situation and converts it into approval in the same manner. This is the approach of Hibbs (1987, chap. 5). In this case, we have a grouped logit (or probit) analysis. Each month presents a grouping of over a thousand people, all manifesting the same values for X, and only differing in whether they approve of the president. Under this assumption, we have not a time-series of 400 points, but individual data on over 400,000 individuals. This increase in information is not quite a factor of one thousand, since the individual-level data have a binary dependent variable and, thus, less information than would

be contained if we actually observed individual approval ratings. The only source of error in the grouped logit analysis, that is, the only reason that Hibbs's estimated standard errors are not zero, is that approval is measured by survey, and so we have an estimate of approval in any period that differs from true approval by sampling error. But with the size of the Gallup Poll's samples, this sampling error is trivial.

It is this fantastic assumption that gives Hibbs's analyses such small standard errors and high t-ratios. In the general literature, and the results reported here, t-ratios on the economic variables in the approval function range from two to four (and are often much less than two). Hibbs's t-ratios are 28, 11, and 15 on inflation, real income, and unemployment. If we really believe that individuals perceive exactly the same economy in any month, and hence the X_t in the approval function also belongs in each individual's approval function, then these results are correct. If, as I do, we believe that the X_t in the approval function represents an average of the (varying) X_t^i, then Hibbs's statistical results are wrong, dramatically wrong.

The individual approach also leads to another problem. Hibbs uses an EDL model, so only X's and lagged X's are in his analysis. This is fortunate, since we do not really know how to do logit analysis with lagged dichotomous approval as an explanatory variable. Thus, the grouped logit analysis limits us to the EDL model, excluding partial adjustment or error correction. Hibbs's analysis also assumes that the error term in the approval function is uncorrelated over time. This assumption is implicit because, in aggregating from individual to monthly approval, the individual error term is integrated out. We have seen that the EDL model with uncorrelated errors is untenable. Thus, starting with a specification of individual-level approval may well lead to an inferior aggregate specification. In short, there seems little to be gained in practice from modeling individual approval and then aggregating. At best we end up with the same aggregate model we would have written directly, and, at worst, we fool ourselves. Among the mistakes that can be induced are believing that there is some direct correspondence between the estimated aggregate parameters and individual-level parameters, that we have considerably more information than, in fact, we do, and that there are no error dynamics.

Conclusions

What is the appropriate dynamic form for an approval function? Does the answer to this question tell us anything about the more general question of specification of time-series models in political science? Obviously, specification in any substantive area will depend on substantive knowledge in that area, but some of the lessons learned about the approval function do generalize.

Starting with the more specific question, it seems clear that static models cannot explain approval; it is a dynamic process. The partial adjustment model also seems an inadequate description of the approval process, unless a better case is made for why approval should move sluggishly. The exponentially distributed lag model (corrected for correlated errors) does seem like a plausible model of approval; the partial adjustment model performs well in practice because it is a transform of this model. The error correction model seems to do the best job of providing an approval function that corresponds with the approval process; at least it makes sense of short- and long-run issues and the importance of levels and differences in explaining approval. In analyses of approval over shorter time periods than used in this article, approval may not be stationary; at that point, error correction would be even more useful.

The transfer function approach can contain all the other approaches, so it is impossible to say that approach is inferior. On the other hand, the transfer function is just another statistical model, and so all standard statistical tools and tests are relevant in assessing these models. Transfer function modelers often seem to limit themselves to the Box-Jenkins methodology; this is unfortunate, especially since this methodology is less useful for multivariate analysis. The transfer function approach does have its advantages; it is ideal for assessing the impact of a specific historical event on approval.

The various models often resemble each other in terms of fit to the data, so we cannot choose a specification on purely empirical grounds. The advantage of the error correction model over the partial adjustment model is that it gets us to think about updating approval in a meaningful way and aids our thinking about the relationship between levels and changes, a distinction that other models blur. The advantage of the exponential distributed lag model is that it brings memory (or forgiving) into the system in a reasonable way. But these advantages are theoretical, not empirical.

The partial adjustment model works, in practice, because the exponentially distributed lag model often shows substantial autocorrelation. This indicates that we cannot simply choose a specification because it fits the data well. It also shows that we cannot simply treat matters such as autocorrelation as technical questions affecting the quality of estimates but not the substance of models.

Going beyond presidential approval, the partial adjustment model may often be preferred over the exponentially distributed lag model. Sometimes there are costs of adjustment. This is most likely to be the case for reaction functions, where the EDL story makes little sense but the partial adjustment story is quite plausible. (It is costly for the Federal Reserve to change interest rates, and even more costly for it to undo mistaken changes.)

In other areas, both the partial adjustment story and the exponentially distributed lag story may both be plausible. I think that this is the case for models assessing the impact of parties or elections on economic outcomes. Political changes work their way slowly through the system (EDL), but it makes just as much sense to think of outcomes adjusting sluggishly (partial adjustment).

Error correction may be a very general process that commends itself in many situations. While it comes into its own in dealing with nonstationary processes, it is relevant wherever a series of variables are in an equilibrium relationship. It is surely worth considering as a candidate model in any situation involving dynamics.

Finally, the transfer function modelers and the autoregressive distributed lag modelers should again start talking to each other. While the models arise in different traditions and come with different tool kits, statistically the two models are not very different. Transfer function modelers could benefit by going beyond the standard Box-Jenkins set of tools, particularly in terms of model specification; autoregressive distributed lag modelers could gain by adding transfer function methods and tools, especially those dealing with lag structure determination, to their already impressive set of methods. Both sides could benefit by considering a wide variety of possible dynamic specifications for any process, rather than sticking to one favored type of specification.

REFERENCES

Beck, N. 1985. "Estimating Dynamic Models Is Not Merely a Matter of Technique." *Political Methodology* 11:71–90.

Beck, N. 1988. "Time Aggregation and Dynamic Estimation." Paper presented at the annual meeting of the Political Methodology Society, UCLA.

Beck, N. 1989. "Estimating Dynamic Models Using Kalman Filtering." *Political Analysis* 1:121–56.

Beck, N. 1991. "The Economy and Presidential Approval: An Information Theoretic Perspective." In *Economics and Politics: the Calculus of Support*, ed. H. Norpoth, M. Lewis-Beck, and J.-D. Lafay, 85–101. Ann Arbor: University of Michigan Press.

Box, G., and G. Jenkins. 1976. *Time Series Data Analysis, Forecasting, and Control.* Rev. ed. San Francisco: Holden Day.

Chappell, H., and W. Keech. 1985. "A New View of Political Accountability for Economic Performance." *American Political Science Review* 79:10–27.

Davidson, J., D. Hendry, F. Srba, and S. Yeo. 1978. "Econometric Modeling of the Aggregate Time-Series Relationship between Consumers' Expenditures and Income in the United Kingdom." *Economic Journal* 88:661–92.

Davidson, R., and J. MacKinnon. 1981. "Several Tests for Model Specification in the Presence of Alternative Hypotheses." *Econometrica* 49:781–93.

Dickey, D., and W. Fuller. 1979. "Distribution of the Estimators for Autoregressive Times Series with Unit Root." *Journal of the American Statistical Association* 74:427–31.

Engle, R., and C. Granger. 1987. "Cointegration and Error Correction: Representation, Estimation, and Testing." *Econometrica* 55:251–76.

Feldman, S. 1985. "Measuring Economic Self-Interest and the Vote: Evidence and Meaning." In *Economic Conditions and Electoral Outcomes: The United States and Western Europe*, ed. H. Eulau and M. Lewis-Beck, 141–66. New York: Agathon.

Freeman, J. 1989. "Systematic Sampling, Temporal Aggregation, and the Study of Political Relationships." *Political Analysis* 1:61–98.

Freeman, J., J. Williams, and T. M. Lin. 1989. "Vector Autoregression and the Study of Politics." *American Journal of Political Science* 33:842–77.

Granger, C. 1990. "Aggregation of Time-Series Variables: A Survey." In *Disaggregation in Econometric Modeling*, ed. T. Parker and H. Pesaran, 17–34. London: Routledge.

Granger, C., and M. Morris. 1976. "Time-Series Modeling and Interpretation." *Journal of the Royal Statistical Society*, ser. A 139:246–57.

Granger, C., and P. Newbold. 1986. *Forecasting Economic Time-Series*. 2d ed. Orlando: Academic Press.

Harvey, A. 1990. *The Econometric Analysis of Time-Series*. 2d ed. Cambridge, Mass.: MIT Press.

Hendry, D., A. Pagan, and J. Sargan. 1984. "Dynamic Specification." In *Handbook of Econometrics*, ed. Z. Griliches and M. Intriligator, 2:1023–1100. Amsterdam: North-Holland.

Hibbs, D. 1974. "Problems of Statistical Estimation and Causal Inference in Time-Series Regression Models." In *Sociological Methodology 1973–1974*, ed. H. Costner, 252–308. San Francisco: Jossey-Bass.

Hibbs, D. 1987. *The American Political Economy: Macroeconomics and Electoral Politics in the United States*. Cambridge, Mass.: Harvard University Press.

Judge, G., W. Griffiths, R. Hill, H. Lutkepohl, and T.-C. Lee. 1985. *The Theory and Practice of Econometrics*. 2d ed. New York: Wiley.

Kernell, S. 1978. "Explaining Presidential Popularity." *American Political Science Review* 72:506–22.

Kernell, S. 1986. *Going Public*. Washington, D.C.: CQ Press.

King, G. 1988. *Unifying Political Methodology: The Likelihood Theory of Inference*. New York: Cambridge University Press.

King, G., and L. Ragsdale. 1988. *The Elusive Executive*. Washington, D.C.: CQ Press.

Koyck, L. 1954. *Distributed Lags and Investment Analysis*. Amsterdam: North-Holland.

Leamer, E. 1978. *Specification Searches: Ad Hoc Inference with Nonexperimental Data*. New York: Wiley.

MacKinnon, J. 1991. "Critical Values for Cointegration Tests." In *Long-run Economic Relationships: Readings in Cointegration*, ed. R. Engle and C. Granger, 267–76. New York: Oxford University Press.

MacKuen, M. 1983. "Political Drama, Economic Conditions, and the Dynamics of Presidential Popularity." *American Journal of Political Science* 28:164–92.

MacKuen, M., R. Erickson, and J. Stimson. 1989. "Macropartisanship." *American Political Science Review* 83:1125–42.

Mueller, J. 1970. "Presidential Popularity from Truman to Johnson." *American Political Science Review* 64:18–34.

Norpoth, H. 1986. "Transfer Function Analysis." In *New Tools for Social Scientists: Advances and Applications in Research Methods*, ed. W. Berry and M. Lewis-Beck, 241–73. Beverly Hills, Calif.: Sage.

Norpoth, H. 1991. "The Popularity of the Thatcher Government: A Matter of War and Economy." In *Economics and Politics: The Calculus of Support*, ed. H. Norpoth, M. Lewis-Beck, and J.-D. Lafay, 141–60. Ann Arbor: University of Michigan Press.

Ostrom, C., and R. Smith. 1990. "Cointegration and Error Correction: Examining the Presidential Approval–Economy Connection." Paper presented at the annual meeting of the Political Methodology Society, Washington University, St. Louis.

Sargan, D. 1980. "Some Tests of Dynamic Specification for a Single Equation." *Econometrica* 48:879–97.

Stock, J., and M. Watson. 1988. "Variable Trends in Economic Time-Series." *Journal of Economic Perspectives* 2 (3): 147–74.

Whitely, P. 1984. "Inflation, Unemployment, and Government Popularity—Dynamic Models for the United States, Britain, and West Germany." *Electoral Studies* 3:3–24.