Taking Time Seriously

Suzanna De Boef Pennsylvania State University
Luke Keele Ohio State University

Dramatic world change has stimulated interest in research questions about the dynamics of politics. We have seen increases in the number of time series data sets and the length of typical time series. But three shortcomings are prevalent in published time series analysis. First, analysts often estimate models without testing restrictions implied by their specification. Second, researchers link the theoretical concept of equilibrium with cointegration and error correction models. Third, analysts often do a poor job of interpreting results. The consequences include weak connections between theory and tests, biased estimates, and incorrect inferences. We outline techniques for estimating linear dynamic regressions with stationary data and weakly exogenous regressors. We recommend analysts (1) start with general dynamic models and test restrictions before adopting a particular specification and (2) use the wide array of information available from dynamic specifications. We illustrate this strategy with data on Congressional approval and tax rates across OECD countries.

Research in all subfields of political science is devoted in large measure to understanding the causes and consequences of events, opinions, behavior, and institutional change as they unfold over time. In recent years, we have seen a dramatic increase in the range of data available over time and with it increasing interest in and application of time series analysis. Once primarily limited to the analysis of the few time series in American politics that span significant amounts of political time, scholars of international political economy and comparative politics are increasingly conducting time series analysis. The interest in dynamics and the data needed to tackle these research questions can only continue to grow.

Political scientists have done a good job introducing and developing methods appropriate for different types of dynamic political data. Statistical models appropriate for durations (Box-Steffensmeier and De Boef 2006; Box-Steffensmeier and Jones 2004), counts (Brandt and Williams 2001), panels of time series (Beck and Katz 1995; Beck, Katz, and Tucker 1998; Stimson 1985), unit, near unit, and fractional roots (Box-Steffensmeier and Smith 1998; De Boef and Granato 1997; Ostrom and Smith 1992), and endogenous time series (Freeman, Williams, and Lin 1989), for example, have appeared in political science journals. This research recommends classes of estimators to ensure that the dynamic model is consistent with the empirical properties of the data.

But in the context of linear time series regression models for stationary time series and weakly exogenous regressors, political scientists are unfamiliar with important practices that allow a better understanding of political dynamics. General unfamiliarity with these techniques leads to common problems in three areas:

- First, analysts tend to adopt restrictive dynamic specifications on the basis of limited theoretical guidance and without empirical evidence that restrictions are valid, potentially biasing inferences and invalidating hypothesis tests.[1]
- Second, political scientists have linked the theoretical concept of equilibrium with the existence of cointegration and use of error correction models. As a result, analysts don't typically derive equilibrium conditions, estimate error correction rates, or use error correction models unless the data are found to be cointegrated. But, (1) error correction models represent a general class of regression

Suzanna De Boef is associate professor of political science, Pennsylvania State University, State College, PA 16802 (sdeboef@psu.edu). Luke Keele is assistant professor of political science, Ohio State University, Columbus, OH 43210 (luke.keele@polisci.sbs.ohio-state.edu).

Paper prepared for presentation at the 2005 annual meeting of the Society of Political Methodology, Florida State University. Thanks to Jim Stimson, our mentor in all things time series (and more), colleague, and friend, for the long-run multiplier effect his time series analysis continues to have in political science. Thanks also to Neal Beck, Daniel Kono, and Amber Boydstun for comments.

[1]Notable exceptions to this rule occur in the Box-Jenkins and VAR frameworks.

American Journal of Political Science, Vol. 52, No. 1, January 2008, Pp. 184–200

 ISSN 0092-5853

models appropriate for stationary data and (2) all regressions describe equilibrium conditions and error correction rates. Exclusively linking error correction models to cointegration limits our ability to use them effectively to understand the dynamics of politics.
- Third and finally, once analysts select and estimate a model, inferences are typically limited to short-run effects and interpretation follows that of a static model: "a unit change in X (at $t - s$) leads to an expected change in Y (at time t)." As such, analysts frequently fail to compute and interpret quantities such as long-run impacts of exogenous variables, mean and median lag lengths of effects, equilibrium conditions, and the rate of equilibrium correction. These quantities allow for statistical tests of the relevant theoretical model, and without them, the inferences we draw are incomplete.

Until we solve these problems, political theory and hypotheses will be undertested and our understanding of the temporal nature of politics incomplete. Here, we focus on the techniques needed to use dynamic regression models effectively. Specifically, we (1) recommend researchers begin analysis with the estimation of a general time series model (guided by theory) and test restrictions on that model, (2) demonstrate that error correction models are suitable for stationary data, (3) provide details on how to interpret a variety of quantities of interest from dynamic regression models that are seldom presented in applied work but that are informative theoretically, and (4) use empirical examples to illustrate the difference that good specification and interpretation can make for the kinds of inferences we draw. The techniques we present apply to time series cross-section data as well.

While others have visited some of these topics before, previous work does not provide systematic coverage of the necessary techniques. For example, there has been some presentation of various dynamic models and their interpretation (Beck 1985, 1991), but without attention to general-to-specific modeling techniques. And others have debated the nature of error correction models (Beck 1992; Durr 1992a, 1992b) but provided no definitive resolution to how and when they can be used. Moreover, as we will demonstrate, what coverage there has been has been largely ignored.

Careful time series analysis is critical to the study of change and its consequences. Estimating restricted models, misunderstanding equilibria, and poor interpretation undermine analysis. They lead, at best, to weakly connected theory and tests and limited understanding. Too often the costs include biased results and incorrect inferences as well. The recommendations that follow help analysts avoid these costs and are central to good time series practice.

Time Series Models in the Literature

We claim that being unaware of several time series techniques, analysts routinely make three mistakes: (1) they adopt restrictive specifications without evidence that restrictions are valid; (2) they link the theoretical concept of equilibrium with the existence of cointegration and use of error correction models (ECMs) and therefore do not derive equilibrium conditions, estimate error correction rates, or ECMs absent evidence of cointegration; and (3) they draw limited inferences about theoretical quantities of interest due to poor interpretation of results. To provide evidence for this claim, we searched articles appearing in *The American Political Science Review*, *The American Journal of Political Science*, and *The Journal of Politics* between 1995 and 2005, cataloging the types of dynamic analysis conducted, the models specified, the theoretical and statistical evidence presented for the specification, and the nature of interpretation of the results.

Between 1995 and 2005, 73 articles were published using time series regression in the context of stationary data. Of these only 10 either started with a general model or tested whether restrictions were empirically valid. Of these 10, only four expressly began with a general model; more than 85% of the articles examined did not test whether the restrictions used were empirically valid. Our review suggests not only that restrictive models are prevalent, but also that dynamic specifications are often selected ad hoc. In particular, frequently there is no discussion of the decision to include contemporaneous values of the exogenous variables as opposed to (or in addition to) lagged values. And no article reported tests for lag lengths.

Authors can be rather cavalier about how they deal with serial correlation. Two manuscripts that estimated static regressions did not discuss this specification choice at all. In one piece, the decision to include a lagged dependent variable was relegated to a footnote and the estimated coefficient and standard error were not reported. And the most common restriction imposed—restricting contemporaneous effects of the independent variables to be zero—is often not discussed at all. Quite often this specification or one with a lagged dependent variable is simply used as an afterthought to cure serial correlation.

The estimation of restricted models in and of itself is not evidence of poor specification. Combined with the lack of evidence for the validity of restrictions and justification for the specifications estimated, however, it suggests a potential for bias. The regularity with which we see this pattern of results is disturbing. It suggests that analysts are either unaware of options within the class of general time series regression specifications, ways to test restrictions on these specifications, or the consequences of estimating overly restrictive models.

The misunderstanding over ECMs is more likely to manifest itself by omission given stationary time series. Since most analysts associate error correction with cointegration, they won't use an ECM with stationary data. This appears to be the case: of the 73 articles that used dynamic regression, only eight used ECMs, and only two understood that cointegration is not necessary to justify their use. We presume that other pieces might have relied upon ECMs if they were better understood.

Last, we examined how many authors did more than interpret the estimated results as static effects. The results are better, but outside the context of ECMs, where equilibrium correction rates and long-run effects are typically interpreted, only 26% of the 65 articles examined provided any interpretation of dynamic quantities; well over a majority of the articles we examined treated dynamic quantities as static. Our review of the top journals in the discipline demonstrates that these errors are not isolated. We next examine in more detail why such errors matter and outline proofs and techniques to help analysts avoid them.

General Models and Perils of Neglecting Them

A General Model

Theories about politics typically tell us only generally how inputs relate to processes we care about. They are nearly always silent on which lags matter, whether levels or changes drive Y_t, what characterizes equilibrium behavior, or what effects are likely to be biggest in the long run. Occasionally theory tells us X_t affects Y_t with a distributed lag. Literature on the relationship between public opinion and Supreme Court decisions is typical. Mishler and Sheehan write: "Although each hypothesis predicts a lag in the impact of public opinion on judicial decisions, none is sufficiently developed in theory to specify the precise length of the expected lag" (1996, 175). Mondak and Smithey (1997) develop a detailed model specifically to predict and identify reasonable equilibrium behavior for support for the Court but are agnostic about the specific dynamics. The story is the same in other fields; theory tells us only generally that the (current and) past matters; it is mute on the specifics of specification.

Substantive theory, then, typically does not provide enough guidance for precise dynamic specifications. Econometric theory is clear in this case: analysts should start with a general model—one that subsumes the data-generating process (Hendry 1995). Only then should we test restrictions, the exclusion of particular X_t (and Y_t) as well as lags. In the context of stationary data and weakly exogenous regressors, the following model fits our criteria:[2]

$$Y_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \sum_{j=1}^{n} \sum_{i=0}^{q} \beta_{jp} X_{jt-i} + \varepsilon_t \qquad (1)$$

where ε_t is white noise, $|\sum_{i=1}^{p} \alpha_i| < 1$ so that Y_t is stationary, and the processes generating X_j are weakly exogenous for the parameters of interest such that $E(\varepsilon_t, X_{js}) = 0 \; \forall \; t, s$, and j. This is an autoregressive distributed lag or ADL(p, q;n) model, where p refers to the number of lags of Y_t, q the number of lags of X_t, and n the number of exogenous regressors.[3]

Such a general model has much to recommend it. First, it makes no assumptions about the lags at which X_t influences Y_t. Second, it should be consistently estimated by ordinary least squares (Davidson and MacKinnon 1993).[4] Finally, the model nests many commonly estimated specifications and therefore can be used to test the appropriateness of the restrictions.

For simplicity, we refer to the case when $p = q = n = 1$, but the results generalize:[5]

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t. \qquad (2)$$

The estimated coefficients, β_0 and β_1, called short-run or impact multipliers, give the immediate effect on Y_t of a unit change in X_t at a given t. If X_t is a measure of economic expectations and data are quarterly, β_0 tells us how levels of economic expectations in 1991Q2 affect presidential approval in that same quarter. β_1 tells us how previous levels of economic expectations in 1991Q1 affect presidential approval in the subsequent quarter. The long-run effect, referred to as the long-run, dynamic, or total multiplier, is given by $k_1 = \frac{(\beta_0+\beta_1)}{(1-\alpha_1)}$.

[2]Box-Jenkins transfer function models are more general still, allowing unique patterns of decay for each X. Techniques such as VAR and dynamic simultaneous equations generalize to the case where regressors are not weakly exogenous.

[3]q gives the maximum lag length; any β_{jt-i} can be set to zero generalizing notation.

[4]The proof for the consistency of OLS assumes that ε_t is IID after the lag of Y_t is included in the model. See Keele and Kelly (2006) for a study of when this is not true.

[5]Longer lag lengths should be tested and eliminated via either an F-test or AIC.

TABLE 1 Restrictions of the ADL General Dynamic Model

Type	ADL Model	Restriction
General	$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t$	None
Partial Adjustment*	$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 X_t + \varepsilon_t$	$\beta_1 = 0$
Static[a]	$Y_t = \alpha_0 + \beta_0 X_t + \varepsilon_t$	$\alpha_1 = \beta_1 = 0$
Finite Distributed Lag[b]	$Y_t = \alpha_0 + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t$	$\alpha_1 = 0$
Differences[c]	$\Delta Y_t = \alpha_0 + \beta_0 \Delta X_t + \varepsilon_t$	$\alpha_1 = 1, \beta_0 = -\beta_1$
Dead Start	$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_1 X_{t-1} + \varepsilon_t$	$\beta_0 = 0$
Common Factor[d]	$Y_t = \beta_0 X_t + \varepsilon_t, \varepsilon_t = \beta_1 \varepsilon_{t-1} + u_t$	$\beta_1 = -\beta_0 \alpha_1$

*Also known as the Koyck model.
[a] $k_1 = \beta_0$; Dynamic effects at lags beyond zero constrained to be zero.
[b] $k_1 = \sum_{j=1}^{n} \sum_{i=0}^{q-1} \beta_{ji}$.
[c] Infinite mean lag length.
[d] $k_1 = \beta_0$, $\mu = 0$, EC rate 100%.

The ADL is a fully general dynamic model. Table 1 lists statistical models that are special cases of the ADL and the restrictions each imposes on parameters of the ADL. Each restriction can be tested in the context of the ADL using t-tests or F-tests. For example, to determine whether the partial adjustment model (PA) is consistent with the DGP, we estimate the ADL and conduct a t-test on $\beta_1 = 0$. If we cannot reject the null, then we can proceed to draw inferences from the PA model assured that the restriction is valid. If we reject the null, then we need to either test alternate restrictions or proceed with an analysis of the general model.

Each dynamic specification implies a specific lag distribution; the lag distribution is the calculation of how much of the effect of X_t on Y_t is distributed across time. This is best understood using a visual illustration, since the values of the lag distribution can be plotted for a visual representation of the dynamic effects. In the first panel of Figure 1, we plot the lag distribution for an ADL where $\beta_0 = .50$, $\beta_1 = .25$, $\alpha_0 = 0$, and $\alpha_1 = .75$. We see that there is a large initial effect that increases during $t + 1$, and then decays over future time periods. If the various restrictions are valid, a different lag structure is implied. The rest of the panels for Figure 1 contain the lag distributions when the restrictions listed in Table 1 are valid for the ADL in the first panel. The static and first differences models have lag distributions that differ most from the ADL since they constrain 100% of the effect to occur immediately (an effect equivalent to the lag zero effect in the ADL). For the other models, we see effects that persist into future time periods. Understanding these differences can be important for the formulation of theory.

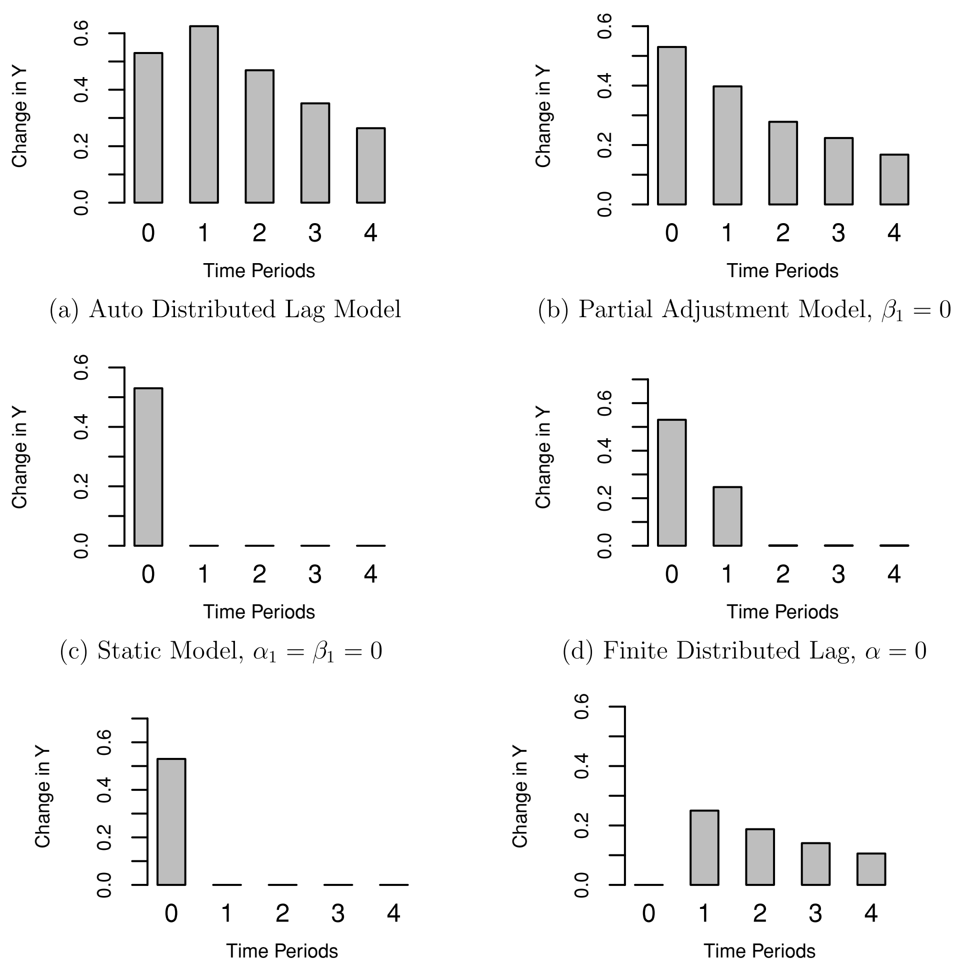
The Costs of Invalid Restrictions

What happens when analysts impose invalid restrictions? That is, what happens when analysts estimate a restricted model when a more general or alternative form is appropriate? Consider the common practice of restricting $\beta_1 = 0$. Analysts should only impose this restriction after testing its validity. The resulting PA model is, however, often estimated based only on an appeal to a broad theoretical justification that is necessary, but not sufficient, for a PA model: the effects of X are biggest in the current period and decay in subsequent periods. Occasionally, a more specific theoretical story is offered: individuals or governments pursue target values of Y_t given current values of X_t, but changing Y_t is costly so that immediate adjustment—change—in Y_t is slow or partial.[6] Nadeau et al. are unique in offering a theoretical justification for their empirical specification of models of presidential approval, economic forecasts, and economic attitudes, writing: "From a theoretical perspective, the 'partial adjustment' model . . . guided our model specification since individuals, whether they are masses or elites, adjust incompletely to new information that becomes available in a given period. This incomplete adjustment could be due to psychological or institutional constraints" (1999, 130). More typically, however, the justification doesn't appeal to theory at all, as the lagged dependent variable is included to clean up serial correlation. When this restriction is invalid, β_0 and α_1 will be biased. The degree and direction of bias are a function of the covariance of X_t and Y_{t-1}, respectively, with X_{t-1}. Because X_t tends to be highly autocorrelated, the bias in β_0 will tend to be large.[7]

[6] Note that such logic itself implies that with a long enough time horizon and given certain patterns in X_t, imbalance in X_t and Y_t is virtually guaranteed so that the model will make no sense empirically or theoretically. This is easily seen when the model is written in error correction form.

[7] Assuming the true model is given by 2 and estimating $Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 X_t + e_t$. The estimate for β_0 is then given by $E(\beta_0) = \beta_0 + b_1 \beta_1$ where b_1 equals: $b_1 = \frac{\sum (X_t - \bar{X}_t)(X_{t-1} - \bar{X}_{t-1})}{\sum (X_t - \bar{X}_t)^2}$ so that b_1

FIGURE 1 Simulated Lag Distributions: The Effect of Restrictions



(a) Auto Distributed Lag Model

(b) Partial Adjustment Model, $\beta_1 = 0$

(c) Static Model, $\alpha_1 = \beta_1 = 0$

(d) Finite Distributed Lag, $\alpha = 0$

(e) Differences, $\alpha_1 = 1$, $\beta_0 = -\beta_1$

(f) Dead Start Model, $\beta_0 = 0$

Estimating a static model—regressing Y_t on X_t—is equivalent to imposing the joint restriction $\alpha_1 = \beta_1 = 0$. In this case, theory must specify that all movement in X_t translates completely and instantaneously to Y_t. This restriction may be valid with high levels of temporal aggregation, but in all cases it needs to be tested. If this restriction is invalid, the consequences when the ADL is the true model are severe: β_0 will neither capture the long-nor the short-run multiplier and calculations of the long-run equilibrium will be biased downward, unless X_t is a unit root process or $\beta_0\alpha_1 + \beta_1 = 0$. The static model constrains the lag lengths to be zero so that the bias in β_0 will be greater when the true lag lengths are longer. The errors of the equation will be autocorrelated. If this autocorrelation is positive, standard errors will be too small. Finally, testing for the constancy of parameters of the static model results in rejection of the null of constant parameters too often, providing misleading information about how to improve the specification.[8]

Unfortunately, postestimation diagnostic tests will be of little help in constructive model building when analysts begin with models that impose untenable restrictions. In particular, failing a specification test does not mean we've found *the*, or even *a*, problem with a given model specification. Consider an example. We can't be certain that autocorrelation in the residuals of an estimated PA model means that the underlying data-generating process is a PA

represents the bias caused due to estimating an overly restrictive dynamic model.

[8]One might object that few would estimate a static model with time series data. But this model is estimated when GLS estimators such as Prais-Winsten are used or OLS with Newey-West standard errors. And as we show, it is estimated with some frequency in applied work.

model containing autocorrelation. In fact using any single diagnostic test to build up a dynamic specification from a restricted (here the PA) model to a general model is dangerous. Diagnostic information is symptomatic of some ill only, but not necessarily of the particular ill or its cause. Autocorrelation may mask a functional form misspecification rather than serial correlation in the data-generating process, for example. Even an absence of assumption violations may merely mask other problems. In short, starting with a restricted model like the PA model can lead us down a road in which rejections of the null can mean many things—misspecification or data generated under the alternative hypothesis, but we cannot know which cause should be attributed because the tests are conditional on the model.

Imposing invalid restrictions biases inferences. The consequences extend beyond the model coefficients: If restrictions are invalid this leads to bias in the estimation of median and mean effects, as well as long-run effects, quantities that we often care about. Further, working backward from a restricted model is unlikely to get us to the best model. We demonstrate these pathologies more explicitly in the examples section. But the solution is simple: begin by estimating a general model, using theory and empirical evidence as the basis for estimating restricted models like those presented in Table 1.

ECMs—An Alternative General Model

The error correction model or ECM is an alternative class of models with a general form equivalent to the ADL. The term *error correction model* applies to any model that directly estimates the rate at which Y_t changes to return to equilibrium after a change in X_t. ECMs suffer from benign neglect in stationary time series applications in political science. Intimately connected with and applied almost exclusively to cointegrated time series, it seems analysts have concluded that ECMs are only suited to estimating statistical relationships between two integrated time series.[9] In fact, ECMs may be used with stationary data to great advantage. We return to this below. First, we review the use of ECMs in political science and the textbook treatment of ECMs, then we show that the ECM is equivalent to the ADL.

Political scientists first adopted ECMs when Ostrom and Smith (1992) introduced cointegration methods to the field. Since then, 82 articles have appeared using ECMs for the estimation of cointegrated relationships (JSTOR). In contrast, only five articles have used ECMs without explicit arguments of cointegration (JSTOR). A reading of articles in our database of five journals over the last 11 years shows that analysts are attracted by the behavioral story associated with cointegration of integrated series and associate it with ECMs: the behavior of Y_t is tied to X_t in the long run, and short-run movements in Y_t respond to deviations from that long-run equilibrium. In fact, there is some evidence that scholars have identified cointegration as a necessary condition for the use of ECMs and the existence of equilibrium. While Heo estimates the "long-term relationship" between defense expenditures and growth in South Korea using a distributed lag model of economic growth, he concludes that an ECM is inappropriate and that "no long-term equilibrium exists between these variables" because "in order to have that relationship between two variables, both must be cointegrated" (Heo 1996, 488–89). This conclusion is misguided; the appropriateness of ECMs need not be linked to cointegration and the existence of equilibrium is independent of the model. Indeed *all* stationary time series models specify an equilibrium relationship. The reason this is wrong is that if all series are stationary, the syllogism does not hold. Error correction implies cointegration only when all the time series in question each contain a unit root.

Confusion over the nature of ECMs is understandable (for an earlier debate on this topic see Beck 1992; Durr 1992a, 1992b; Smith 1992; Williams 1992). We surveyed 18 major econometrics texts, both general volumes and time series texts. Of those, eight discuss ECMs exclusively in the context of cointegration and five don't mention ECMs at all. Of the five that show the linkage between ECMs and stationary data, three are general texts. The text that is most clear about the nature of ECMs is Bannerjee et al. (1993), but this is a book about the analysis of cointegrated data. Statements such as this one from Enders (2001) are typical and contribute to the confusion over the general applicability of the ECM: "The error correction mechanism necessitates the two variables be cointegrated of order CI(1,1)"(2001, 366).

To prove the ECM is suitable for stationary data, we show the equivalence of the ADL (which has a stationarity condition) and ECM. Parts of the proof can be seen elsewhere (Bannerjee et al. 1993; Davidson and MacKinnon 1993). We describe the Bardsen transformation of the ECM, which is perhaps the most useful form of the ECM.[10] We start by subtracting Y_{t-1} from both sides of equation (1), the ADL (1,1;1):

$$\Delta Y_t = \alpha_0 + (\alpha_1 - 1)Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t. \quad (3)$$

[9]Exceptions include De Boef and Kellstedt (2004).

[10]There are many error correcting forms, all of which contain the same information.

Table 2 Restrictions of the Error Correction Model

Type	ECM Model	Restriction
General	$\Delta Y_t = \alpha_0 + \alpha_1^* Y_{t-1} + \beta_0^* \Delta X_t + \beta_1^* X_{t-1} + \varepsilon_t$	None
Partial Adjustment	$\Delta Y_t = \alpha_0 + \alpha_1^* Y_{t-1} + \beta_0^* X_t + \varepsilon_t$	$\beta_1^* - \beta_0^* = 0$
Static[a]	$\Delta Y_t = \alpha_0 + \alpha_1^* Y_{t-1} + \beta_0^* \Delta X_t + \beta_1^* X_{t-1} + \varepsilon_t$	$\alpha_1^* = -1, \beta_1^* + \alpha_1^* = \beta_0^* - 1$
Finite Distributed Lag[b]	$\Delta Y_t = \alpha_0 + \alpha_1^* Y_{t-1} + \beta_0^* \Delta X_t + \beta_1^* X_{t-1} + \varepsilon_t$	$\alpha_1^* = -1$
Differences[c]	$\Delta Y_t = \alpha_0 + \beta_0^* \Delta X_t + \varepsilon_t$	$\alpha_1^* = 0, \beta_1^* + \alpha_1^* = 0$
Dead Start	$\Delta Y_t = \alpha_0 + \alpha_1^* Y_{t-1} + \beta_1^* X_{t-1} + \varepsilon_t$	$\beta_0^* = 0$
Common Factor[d]	$\Delta Y_t = \alpha_0 + \beta_0^* \Delta X_t + \varepsilon_t, \varepsilon_t = \rho\, \varepsilon_{t-1} + u_t$	$\rho = 2\alpha_1^* + 1$

[a] $k_1 = \beta_0^*$.
[b] $k_1 = \beta_1^* + \alpha_1^* + 1$, EC rate 100%.
[c] Infinite mean lag length.
[d] $\mu = 0$, EC rate 100%, $\rho = \beta_1^* + \alpha_1^*$.

Now add and subtract $\beta_0 X_{t-1}$ from the right-hand side:

$$\Delta Y_t = \alpha_0 + (\alpha_1 - 1)Y_{t-1} + \beta_0 \Delta X_t + (\beta_0 + \beta_1)X_{t-1} + \varepsilon_t. \tag{4}$$

Regrouping terms leaves us with the following equation:

$$\Delta Y_t = \alpha_0 + \alpha_1^* Y_{t-1} + \beta_0^* \Delta X_t + \beta_1^* X_{t-1} + \varepsilon_t. \tag{5}$$

By substitution, we see the equivalence of the ADL and the Bardsen ECM: $\alpha_1^* = (\alpha_1 - 1)$, $\beta_0^* = \beta_0$, and $\beta_1^* = \beta_0 + \beta_1$. The short-run effects of X_t and X_{t-1}, respectively, in the ECM are given by β_0^* and by $\beta_1^* - \beta_0^*$. The ECM need not be linked with cointegration and is appropriate for use with stationary data. Table 2 shows how the restricted forms of the ADL are also special cases of the ECM. Thus any time we want to estimate a general model we can use either the ADL or ECM. Equally we can estimate an error correction form of the restricted models presented in Table 1 where these models imply the same restrictions as those in Table 1. This is an important point: imposing invalid restrictions on the ECM will have the same effect as on the ADL: biased estimates and invalid inferences.[11] Whether analysts choose to estimate the ADL or the ECM and test restrictions within either framework is largely a matter of "ease of use," depending on whether one prefers immediate access to short- or long-term quantities. This brings us to our final point: issues of interpretation. Armed with alternative forms of a general time series model, analysts can easily estimate and calculate a variety of dynamic quantities that may be of interest for drawing inferences about theories.

Interpretation

Time series analysis presents challenges to the interpretation of estimated models unlike those of cross-sectional data. In cross-sectional analysis, all estimated effects are necessarily contemporaneous and therefore static. Cross-sectional data do not allow us to assess whether causal effects are contemporaneous or lagged, let alone whether some component is distributed over future time periods.[12] In contrast, consider two types of effects we encounter in time series models:

- An exogenous variable may have only short-term effects on the outcome variable. These may occur at any lag, but the effect does not persist into the future. The reaction of economic prospections to the machinations of politicians, for example, may be quite ephemeral—influencing evaluations today, but not tomorrow. Here the effect of X_t on Y_t has no memory.
- An exogenous variable may have both short- and long-term effects. In this case, the changes in X_{t-s} affect Y_t, but that effect is distributed across several future time periods. Often this occurs because the adjustment process necessary to maintain long-run equilibrium is distributed over some number of time points. Levels of democracy may affect trade between nations both contemporaneously and into the future. These effects may be distributed across only a few or perhaps many future time periods.

[11] Not all of the forms of the ECM in Table 2 can be estimated with integrated data. When using an ECM with integrated data, analysts must ensure that all terms on the right-hand side of the equation are stationary.

[12] When lagged values of exogenous variables or indicators for observation-year are included, they are generally treated as controls and, absent a lagged dependent variable, are not dynamic models in the sense to which we refer.

How many time periods is an empirical question that can be answered with our data.

Dynamic specifications allow us to estimate and test for both short- and long-run effects and to compute a variety of quantities that help us better understand politics. Short-run effects are readily available in both the ADL and the ECM. We next review the mechanics of long-run effects.

Long-Run Effects

Two time series are in equilibrium when they are in a state in which there is no tendency to change. The "long-run equilibrium" defines the state to which the series converge over time. It is given by the unconditional expectations or the expected value of Y_t in equilibrium. Let $y^* = E(Y_t)$ and $x^* = E(X_t)$ for all t. If the two processes move together without error, in the long run they converge to the following equilibrium values for the ADL(1,1;1):

$$y^* = \alpha_0 + \alpha_1 y^* + \beta_0 x^* + \beta_1 x^*. \tag{6}$$

Solving for y^* in terms of x^* yields:

$$y^* = \frac{\alpha_0}{1-\alpha_1} + \frac{\beta_0 + \beta_1}{1-\alpha_1} x^* = k_0 + k_1 x^* \tag{7}$$

where $k_0 = \frac{\alpha_0}{(1-\alpha_1)}$ and $k_1 = \frac{(\beta_0+\beta_1)}{(1-\alpha_1)}$, and k_1 gives the long-run multiplier (LRM) of X_t with respect to Y_t. We can think of the LRM as the total effect X_t has on Y_t distributed over future time periods. In some cases, long-run equilibria and the LRM are of greater interest than short-run effects. Policymakers, for example, debate the optimal defense spending needed to generate a sustained peaceful equilibrium or the long-run effects of deficit spending on economic growth.

When the equilibrium relationship between two time series is disturbed, let's say between economic expectations and presidential approval, then $y^* - (k_0 + k_1 x^*) \neq 0$. In this case, we expect a change in the level of presidential approval in the next period back toward the equilibrium. Interest in the rate of return to equilibrium, also known as error correction, is often motivated by the desire to understand just how responsive a process is. Does consumer sentiment, for example, respond quickly to good economic news or are consumers more skeptical, responding slowly, in fact willing to tolerate sentiment too low for the long-run equilibrium in the interim? The ADL also provides us with information about the speed of this error correction. The speed of adjustment is given by $(1 - \alpha_1)$, as it dictates how much Y_t changes over each future period. If the LRM is 5.0, and the error correction rate is 0.50, Y_t will change 2.5 points in $t + 1$, and then another 1.25 points at $t + 2$ and then 0.625 in $t + 3$ and so on until the two series have equilibrated. Obviously increases in α_1 produce slower rates of error correction, with the reverse also being true.[13]

In the ECM, we directly estimate the error correction rate, α_1^*, the short-run effect of X_t, and their standard errors. The LRM, k_1, is more readily calculated using the ECM:

$$k_1 = \frac{\beta_1^*}{\alpha_1^*} = \frac{(\beta_1 + \beta_0)}{(\alpha_1 - 1)}. \tag{8}$$

A simple example demonstrates the interpretation of all three coefficients from the ECM. Let's say we regress the first difference of presidential approval on one lag of presidential approval, one lag of economic expectations, and the first difference of economic expectations as in equation 5. The estimated coefficients are $\hat{\beta}_0^* = 0.5$, $\hat{\alpha}_1^* = -0.5$, and $\hat{\beta}_1^* = 1.0$. If economic expectations increase five points, how will that affect presidential approval in the context of the ECM? First, presidential approval will increase 2.5 points immediately ($5 * 0.5$, the coefficient of $\hat{\beta}_0^*$). Because presidential approval and economic expectations also have an equilibrium relationship, this increase in economic expectations disturbs the equilibrium, causing presidential approval to be too low. As a result, presidential approval will increase an additional 7.5 points. But the increase in presidential approval (or reequilibration, in error correction parlance) is not immediate, occurring over future time periods at a rate dictated by $\hat{\alpha}_1^*$. The largest portion of the movement in presidential approval will occur in the next time period, when 50% of the shift will occur. In the following time period $(t + 1)$, presidential approval will increase 2.5 points, increasing 1.25 points at $t + 2$ and .63 points in $t + 3$ and so on, until presidential approval has increased five points. Thus, the economy has two effects on presidential approval: one that occurs immediately and another impact dispersed across future time periods. This example underscores how the ECM is a very natural specification in that the error correction rate, one short-term effect, and the long-term multiplier are directly estimated.

Neither the ECM nor ADL, unfortunately, provides a direct estimate of the standard error for the LRM. But since the LRM is the ratio of two coefficients in the ECM, $(\frac{\beta_1^*}{\alpha_1^*})$, the standard error can be derived from the ECM. In particular, the variance for the LRM is given by the

[13]When the maximum number of lags of Y_t exceeds 1, the sum of their coefficients minus one give the *cumulative adjustments*: $\sum_{i=1}^{p} \alpha_i - 1$ where p gives the maximum number of lags of Y_t in the model. The cumulative adjustment rate thus equals the rate of error correction when only one lag of the dependent variable is included.

formula for the approximation of the variance of a ratio of coefficients with known variances, in this case:

$$\mathrm{Var}(a/b) = (1/b^2)\mathrm{Var}(a) + (a^2/b^4)\mathrm{Var}(b) - 2(a/b^3)\mathrm{Cov}(a, b). \quad (9)$$

So using this formula, we can calculate the standard error for the LRM. Alternatively, we can directly estimate the LRM and its standard error using a transformation first proposed by Bewley (1979). The Bewley transformation is a computational convenience for calculating the LRM and its standard error and is not meant to serve as a representation of the underlying dynamics. The Bewley transformation takes the form of the following regression:

$$Y_t = \phi_0 - \phi_1 \Delta Y_t + \psi_0 X_t - \psi_1 \Delta X_t + \mu_t \quad (10)$$

where $\phi_0 = \eta\, \alpha_0$, $\phi_1 = \eta\, \alpha_1$, $\psi_0 = \eta(\beta_0 + \beta_1)$, $\psi_1 = \eta\, \beta_1$, $\mu = \eta\varepsilon_t$ and $\eta = \frac{1}{\alpha_1 - 1}$.

Due to the inclusion of ΔY_t on the right-hand side of the equation, instrumental variables regression must be used to obtain consistent estimates. A constant X_t, X_{t-1}, and Y_{t-1} should be used as instruments to estimate the model. To estimate the Bewley transformation between X_{1t}, X_{2t}, and Y_t, analysts first regress changes in Y_t on lagged Y_t, contemporaneous values of the X_{1t} and X_{2t}, and changes in X_{1t} and X_{2t}. Predicted values from this regression are included in the Bewley model in the following way: $Y_t = \phi_0 + \phi_1 \Delta \hat{Y}_t + \psi_0 X_{1t} + \psi_1 \Delta X_{1t} + \psi_2 X_{2t} + \psi_3 \Delta X_{2t}$. In spite of this added step, the Bewley transformation is appealing because the LRM is estimated as the coefficient on X_t in this specification: $\psi_0 = \frac{B(1)}{A(1)} = \eta(\beta_0 + \beta_1) = k_1$ as $\eta = \frac{1}{\alpha_1 - 1}$. Because it is estimated directly, the variance associated with the LRM is also directly estimated.[14,15] Analysts interested in long-run behavior thus have a way to estimate not only the total long-run effect but also the precision of that estimate.

Forgotten Dynamic Quantities: Mean and Median Lag Lengths

Other quantities that inform us about politics can be computed from dynamic regressions. In addition to knowing the magnitude of the total effect of a shock as measured by the long-run multiplier, it is often useful to know how many periods it takes for some portion of the total effect of a shock to dissipate or how much of the shock has dissipated after some number of periods. The mean and median of the lag distribution of X_t provide information about the pattern of adjustment a series Y_t makes to disequilibrium. The median lag tells us the first lag, r, at which at least half of the adjustment toward long-run equilibrium has occurred following a shock to X_t, providing information about the speed with which the majority of a shock dissipates. It is calculated by listing the effect of a unit change in X_t at each successive lag, standardizing it as a proportion of the cumulative effect, and then noting at which lag the sum of these individual effects exceeds half of the long-run effect. A median lag of 0 tells us that half of the effect is gone in the period it has occurred. We might expect such short median lags when equilibria are very "tight," that is, lagged Y_t has a small coefficient and X_t a large one. Although it will vary with the periodicity of the data, a median lag length of four periods would be quite long for most political processes measured quarterly or monthly. Given quarterly data on presidential approval, for example, empirical evidence suggests that a majority of the effects of a shock in inflation will be realized (well) within a year.[16] Mean lags tell us how long it takes to adjust back to equilibrium, the average amount of time for a shock to play out. In our experience, political processes tend to have mean lag lengths on the order of six quarters, perhaps less.

Median lag lengths are somewhat tedious to calculate and are typically given short shrift in political science. When deriving the formula for the median lag, it is useful to write the general model using lag polynomials:

$$A(L)Y_t = B(L)X_t + \varepsilon_t \quad (11)$$

where L is the lag operator: $L^i X_t = X_{t-i}$, $A(L) = 1 - \alpha_1 L - \alpha_2 L^2 - \cdots - \alpha_p L^p$ and $B(L) = \beta_0 + \beta_1 L + \beta_2 L^2 + \cdots + \beta_q L^q$.

We can calculate the median lag by computing m for successive values of r and recording the value of r when $m \geq .5$:

$$m = \frac{\sum_{r=0}^{R} \omega_r}{\sum_{r=0}^{\infty} \omega_r} \quad (12)$$

[14]Two things should be noted about this variance estimate. First, it is an approximation to the variance since instruments are used to estimate the model. But, second, it can be shown that this estimate is equivalent to that in equation 9 (Hendry 1995).

[15]We emphasize that the Bewley transformation is a statistical transformation for calculating the LRM and its standard error and does not represent a theoretical model. As such, the R^2 should not be interpreted; it tells us how the instruments account for variation in Y_t, rather than how the series of interest contribute to the explanatory power of the model.

[16]Experience suggests to us that a median lag length of 2 is common in the study of American public opinion when the periodicity is quarterly, as there is inertia in Y_t but X_t is often a strong predictor of Y_t as well. But this will depend on the periodicity of the data. Median lag lengths could be longer for monthly or weekly data.

where ω_r:

$$\omega_r = \frac{B(L)}{A(L)} \tag{13}$$

and $\sum_{r=0}^{\infty}\omega_r = \frac{B(1)}{A(1)}$, where $A(1) = 1 - \sum_{i=1}^{p}\alpha_i$ and $B(1) = \sum_{i=0}^{q}\beta_i$. ω_r represents the effect of a shock r periods after it occurs. The denominator summation is across all values of time and thus provides the familiar long-run effect, k_1. The numerator summation allows us to calculate effects through any number of periods, R. The division in equation 12 thus normalizes the adjustment as a proportion of the total adjustment through r periods. It is often useful to graph the standardized lag distribution to see patterns in effects, as we did in Figure 1. Graphs of lag distributions help answer questions such as "What proportion of the total effect has dissipated after four time periods?"[17] Unstandardized lag distributions or cumulative standardized lag distributions may also provide useful visuals for policy proscriptions.

In contrast, the mean lag length tells us how long it takes to adjust back to equilibrium. The mean lag for X_j is given by:

$$\mu = \frac{W(L)'}{W(L)} = \frac{B(L)'}{B(L)} - \frac{A(L)'}{A(L)} = \frac{\sum_{r=0}^{\infty} r\omega_r}{\sum_{r=0}^{\infty}\omega_r} \tag{14}$$

where $'$ denotes the derivative with respect to L.

Mean and median lags are useful when we wish to know how many periods it takes for a process to return to equilibrium. If the level of economic expectations increases, how long will it take to see movement in trust to its new equilibrium value? Is change fast or slow? A president looking ahead to his reelection campaign may be particularly interested in the length of time it takes the public to forget a recession. Where α_1 is large the mean lag length will be long. The median lag length will be short when β_0 approaches one-half of the long-run multiplier, $.5 \times k_1$, and will equal zero when it is greater than this quantity. When both the median and mean lag length are small, adjustment is fast; when the median is short and the mean long, a large part of the disequilibrium is corrected quickly while the long-run response takes some time to complete the adjustment.[18]

Analysts who estimate dynamic models should calculate all these quantities. A complete interpretation of dynamic linear models requires a careful explication of short- and long-run effects, error correction rates, long-run multipliers, and mean and median lag lengths. Moreover, with these quantities in hand, analysts can assess theory in greater detail.

A Simulated Example

We demonstrate the equivalence of the ADL and the ECM and calculation of the various dynamic quantities from the model using a simulated example. The DGP for Y_t is the ADL(1,1) model and X_t is a simple autoregressive process:

$$\begin{aligned} Y_t &= \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_{1t} \\ X_t &= \rho X_{t-1} + \varepsilon_{2t} \end{aligned} \tag{15}$$

where ε_{1t} and ε_{2t} are white noise, $cov(\varepsilon_{1t}, \varepsilon_{2t}) = 0$, ρ is 0.75, and we set the parameter values for the Y_t DGP as follows: $\alpha_0 = 0$, $\alpha_1 = 0.75$, $\beta_0 = 0.50$, and $\beta_1 = 0.25$. These values and the roots of this lag polynomial are less than one in absolute value, ensuring that the DGP is stationary. We estimated both an ADL and an ECM. The results appear in Table 3.

First, we examine the short-run effects of X_t on Y_t. For the ADL model in column 1, these are given explicitly by $\hat{\beta}_0$ and $\hat{\beta}_1$—the coefficients on X_t and X_{t-1}, here estimated as 0.53 and 0.25, respectively. For the ECM, $\hat{\beta}_1^*$—the coefficient on ΔX_t and its standard error—should be, and is, equal to that for $\hat{\beta}_0$, 0.53. For the second short-run effect—the immediate effect of X at $t-1$—the comparison is less obvious. In the ADL $\hat{\beta}_1 = 0.25$; for the ECM we use $\hat{\beta}_1^* - \hat{\beta}_0^*$ $(0.77 - 0.53)$, or 0.24. Differences in these estimates are due to rounding.

We calculated values for the LRM, k_1, for each model as well. The true value for the generated data is given by $\frac{(\beta_1+\beta_0)}{(1-\alpha_1)} = \frac{0.50+0.25}{0.25} = 3.00$. Using the values from the estimated model, we find that $\hat{k}_{1,ADL} = \frac{0.53+0.25}{(1-0.75)} = 3.12$ and $\hat{k}_{1,ECM} = \frac{\beta_1^*}{-\alpha_1^*} = \frac{0.77}{-(-0.25)} = 3.08$. We also calculated the LRM via a Bewley transformation, $k_{1,Bewley} = 3.06$, in order to obtain the standard error: 0.18. The difference between the estimated and true LRMs is caused by sampling error in the simulation.

We can also calculate the mean and median lag length from each specification. The median lag length in the generated data is 2. The easiest way to calculate this from

[17] These graphs make the relationship between the lag distribution, the impulse response function (IRF), and the cross correlation function (CCF) readily apparent. The IRF and CCF are familiar to those acquainted with the Box-Jenkins time series framework. The former gives the response of Y_t to an impulse shock in X_t in its natural units; the latter standardizes the effects. Thus the graph of the standardized lag distribution is equivalent to the CCF.

[18] If any of the lag weights are negative, then the mathematics used do not apply and mean lag lengths cannot be computed. In such cases, it is likely that the model is misspecified, meaning that negative lag weights can be a good diagnostic for dynamic specification (Hendry 1995, 216).

TABLE 3 ADL and ECM Estimates

	ADL Model	ECM
Y_{t-1}	0.75	−0.25
	(0.02)	
X_t	0.53	–
	(0.06)	
X_{t-1}	0.25	0.77
	(0.07)	
ΔX_t	–	0.53
		(0.06)
ΔY_t	–	–
k_1	3.12	3.08
N	249	249
Adj. R^2	.94	.45

Note: Simulated data. True $k_1 = 3.00$.
Using the Bewley transformation $k_1 = 3.06$ with a standard error of 0.18.

the estimated model is by substituting into the ADL and assuming equilibrium conditions except for a single unit shock in X_t so that in the first period ($r = 0$) the response in Y_t is given by:

$$Y_t = 0.53 X_t \tag{16}$$

Normalizing as a proportion of the estimated total or long-run effect we have 0.53/3.12 = .170 or 17.0% of the effect. The effect of that same shock one period later ($r = 1$) is given by:

$$Y_{t+1} = 0.75 Y_t + 0.25 X_t \tag{17}$$

$$= 0.75(0.53) + 0.25 = 0.648 \tag{18}$$

Divide by $\hat{k}_{1,ADL}$ to normalize: 0.648/3.12 = 0.208 or 20.8%. The cumulative portion of the effect expended is 0.208 + 0.170 or 37.8%. Continuing, the formula is simpler as no additional short-term effects of X_t enter the model.

$$Y_{t+2} = 0.75 Y_{t+1} \tag{19}$$

$$= 0.75(0.648) = 0.486 \tag{20}$$

0.486/3.12 = 0.156. Adding, we exceed one-half: 0.156 + 0.208 + 0.170 = 53.4%.[19]

It is almost always the case that the mean and median lags will need to be calculated by hand. While it is not always the case that the median (or mean) will be of theoretical interest, we generally care about the distribution of the effects so characterizing that distribution in some form is worth doing. While there are other ways to calculate medians or other features of the distribution, substitution is the most intuitive way to understand the patterns of decay.

The mean lag length for the simulated data can also be calculated readily from the ADL representation. The general formula for the mean lag for the ADL(1,1;1) is:

$$\mu = \frac{B(L)'}{B(L)} - \frac{A(L)'}{A(L)} = \frac{\beta_1}{\beta_0 + \beta_1} - \frac{-\alpha}{1-\alpha}$$
$$= \frac{.25}{.50 + .25} + \frac{.75}{1 - .75} = 3\frac{1}{3}. \tag{21}$$

It is particularly easy to see the effect of changes in the dynamic parameter α on the mean lag length; bigger α produce longer mean lag lengths. The mean lag length is relatively short in our example, although not overly so.

Finally, our example provides a useful reminder about the potentially misleading nature of R^2. The R^2 for the ADL is .94, while that for the ECM is less than half, $R^2 =$ .45. The two models fit the data *equally* well, but the fit appears to be much better with the ADL, where the dependent variable is in levels, as opposed to changes as in the ECM.

Analysts seldom draw inferences about either the equilibrium relationship, the LRM, the rate of error correction, or features of the lag distribution, yet often these quantities allow for richer interpretation of theory and may be of central importance for understanding politics. Important exceptions include Clarke and colleagues' (2005) analysis of approval of the British prime minister and Huber's (1998) study of health care costs. Each of the general models we have discussed will provide estimates of these quantities, and it is important to report and interpret them. In spite of their ready availability, however, discussions of these quantities are typically underdeveloped and are often omitted from the results altogether, limiting our understanding of the temporal nature of politics.[20]

[19] We can write this sequence of lag values in terms of the ω_r as well: $\omega_0 = \beta_0 X_0 = .53$.

$\omega_1 = \beta_1 X_0 + \alpha Y_0 = .25 + .75(.53) = .648$
$\omega_2 = \alpha Y_1 = .75(.648) = .486$
$\vdots$
$\omega_r = \alpha^{r-1} Y_1,\ r > 2.$

[20] Exceptions occur when analysts estimate VAR models and VECMs and report impulse response functions and variance decomposition. See, for example, Williams (1990) and Williams and Collins (1998).

Selecting a General Model

At this point, one might ask: Is there any reason to choose the ADL over the ECM (or the reverse)? Both are equally general and contain the same information. Both fit the data equally well and both are easily estimated with most any statistical software. The only empirical distinction is that differing quantities are directly estimated in each model. In the ADL, both short-term effects are directly estimated, while in the ECM only one is, but the error correction rate is directly estimated. In addition to these differences, each of the models has unique advantages, which we now review.

The ADL's greatest advantage may be the familiarity analysts tend to have with the model. Additionally, the ADL provides estimates of short-run effects (and their standard errors) in a transparent fashion. This gives us an explicit indicator of the "stickiness" of the process we care about.

ECMs, on the other hand, allow for a tighter link between theory and model. The behavioral story of error correction is broadly appealing: two (or more) processes are tied together in the long run such that when one process increases (or decreases), the other must adjust to maintain this long-run equilibrium relationship. In fact, the error correction model can be derived from formal theories of equilibrium behavior. Another important advantage of ECMs is that variables are parameterized in terms of changes, helping us to avoid spurious findings if the stationarity of the series is in question due to strongly autoregressive or near-integrated data, for example (De Boef 2001; De Boef and Granato 1997). Often political time series produce conflicting results to tests for integration. The ECM is useful in such ambiguous situations.[21]

So which model is the right one? There is no right answer. We prefer the ECM since it has a natural theoretical interpretation, and the error correction rate, short-run effects, and long-run multiplier are most readily available. The only situation where one would strongly prefer the ECM is if the data are strongly autoregressive. But so long as the analyst starts with a general model and fully interprets the final model, it matters little whether the starting point is an ADL or ECM. In the next section, we turn to two examples to demonstrate the need to use general models. We also apply our strategy for dynamic specification and show how consideration of different dynamic specifications can help us to gain a richer understanding of politics.

[21]If the data are integrated, an alternative form of the ECM must be estimated so that no regressors are integrated.

Two Examples

Taxation in OECD Countries

Our first example comes from a widely used type of time series cross-sectional (TSCS) data, a collection of macroeconomic and political indicators for a series of OECD nations. We use Swank and Steinmo's (2002, hereafter SS) data on taxation in OECD countries from 1981 to 1995. SS compare internal economic factors that affect taxation, such as structural unemployment, to external factors such as capital mobility and trade. SS estimate the same model across different forms of taxation; we focus on just one of these: the effective tax rate on labor. SS use dynamic models mixing partial adjustment and dead-start effects in the following way:

$$Y_{it} = \alpha_1 Y_{it-1} + \beta_{10} X_{1it} + \beta_{21} X_{2it-1} + \varepsilon_t \qquad (22)$$

The dynamics for X_1 are those of partial adjustment, while the dynamic effect of X_2 is that of a dead-start model. More formally, the restriction for X_1 is that $\beta_{11} = 0$ and for X_{2t} is $\beta_{20} = 0$. Importantly, they do not mention whether more general models were estimated to justify the constraints imposed by their specification. The authors do interpret their effects dynamically by calculating long-run multipliers for the restricted models. This is a rarity among dynamic models estimated with TSCS data. However, the long-run multipliers they calculate are affected by the model constraints and do not have standard errors. They assume that the long-run multipliers are statistically significant if the estimated coefficient for the variable in question is significant.

SS include 12 independent variables and find four statistically significant. Table 4 lists the variables included in the model estimated by SS. An asterisk in column 1 marks statistically significant variables reported in the original model. We reestimated their model with a more general ECM, first ensuring that we didn't need a higher order model using the AIC, and calculated standard errors for the long-run multipliers using the Bewley transformation.[22] We find that only one of the four variables the authors concluded had significant long-run effects remains significant. According to our estimates, the

[22]We also performed diagnostic tests for autocorrelated residuals using an LM test. We found no evidence of autocorrelation in the models we estimated. When using a Bewley transformation in this context, every X_t variable and its lag along with a lag of Y_t are used as instruments for the first difference of Y_t. The predicted values from this equation are placed on the right-hand side of the second stage estimating equation in place of the first difference of Y_t.

TABLE 4 Replication of Swank and Steinmo

Independent Variable	Original Model[a]	General Model[b]
Liberalization of Capital Controls	*	
Trade		*
Structural Unemployment	*	
Public Sector Debt		
Elderly Population		*
Growth		*
% Change in Real Profits		
Domestic Investment		
Inflation		
Unemployment	*	
Left Government	*	*
Christian Democratic Govt.		*

[a]*significant at .05 level for reported coefficient.
[b]*significant at .05 level for long-run multiplier.

long-run multipliers for four additional variables, however, are now significant (see column 2).

Let's look more closely at the effects of two variables. SS hypothesize that liberalization of capital controls decreases labor tax rates and find evidence of an effect. Using estimates from their model, a unit increase in capital controls decreases labor tax rates 88 points in the long run. Using a more general model, we find that the long-run multiplier is about 4 points and is not statistically significant. While liberalization of capital controls is no longer significant in the more general model, trade matters. The authors conclude trade has no effect, while we find that the long-run multiplier for trade is highly statistically significant. This example underscores the importance of estimating a general model, testing whether restrictions are empirically valid, and estimating the standard error on the long-run multiplier.

Dynamic Specifications of Congressional Approval

The next example relies on standard time series data. We use Durr, Gilmour, and Wolbrecht's (1997, hereafter DGW) analysis of Congressional approval from 1974 to 1993. DGW test whether Congressional approval is a function of both the actions of Congress in the form of the passage of major legislation, veto overrides, and internal discord, and external factors such as presidential approval, the state of the economy, and media coverage.

The authors specify a PA model, but do not report results for a more general model. Specifically, they regress Congressional approval on a lag of Congressional approval, measures of institutional activity, scandals, presidential approval, economic expectations, and media coverage. The measure of presidential approval fails to achieve standard levels of statistical significance and is dropped from the model. We reestimate their model but include the measure of presidential approval and report the results in the first two columns of Table 5.[23]

The model estimates in column 1 of Table 5 are those the authors published, and as we see in column 2, presidential approval is not statistically significant, just as they reported. Economic expectations and media coverage are both statistically significant. The computed long-run effects are given by $\frac{.07}{1-.80} = .35$ for economic expectations and $\frac{.21}{1-.80}$ or about 1.0 for media coverage. DGW do not calculate a long-run effect for presidential approval given the lack of significance for the levels term.

The specification presented in column 1 of Table 5 is a restricted form of the ADL in which lagged values of the exogenous variables are restricted to have no effect ($\beta_1 = 0$ in terms of Table 1). The model in column 2 imposes the additional restriction that presidential approval has no effect (in terms of Table 1, $\beta_0 = 0$). If these restrictions are invalid, we expect some level of bias. However, the amount and direction is difficult to predict in a multivariate setting. If the restrictions are valid, then the restricted coefficients will not be significantly different from zero in an ADL or ECM specification.

The results from the two general models (see the last two columns of Table 5) emphasize two points.[24] They illustrate first how invalid restrictions can change the inferences we draw and second how alternative general models help us understand dynamic relationships.

Consider first the ADL model (see column 3). At first glance—if the effects were interpreted as if it were a static model—the statistical results would appear to be exactly opposite of those published. That is, both short-run effects for presidential approval are significant, while neither of the coefficients for economic expectations approaches standard levels of statistical significance. The results from the ECM in column 2 more readily reveal the long- and short-run dynamic patterns: The effect of presidential approval is one that is almost entirely immediate,

[23]We make one change to the authors' specification. The authors use a measure of presidential approval that is purged of its economic variance and adjusted for divided government. We use a measure of presidential approval that is purged of its economic variance, but we remove the adjustment for divided government. The results are the same but the interpretation is more straightforward without the adjustment.

[24]For these more general models, we again selected the lag length via the AIC.

TABLE 5 Replication of Durr, Gilmour, and Wolbrecht with General Models

	Original Model 1	Original Model 2	ADL	ECM
Congressional Approval$_{t-1}$	0.80*	0.77*	0.77*	−0.22*
	(0.05)	(0.06)	(0.06)	(0.06)
Presidential Approval	–	0.05	0.11*	–
		(0.04)	(0.06)	
ΔPresidential Approval	–	–	–	0.11*
				(0.06)
Presidential Approval$_{t-1}$	–	–	−0.09*	0.02
			(0.06)	(0.04)
Economic Expectations	0.07*	0.08*	0.02	–
	(0.03)	(0.03)	(0.07)	
ΔEconomic Expectations	–	–	–	0.02
				(0.07)
Economic Expectations$_{t-1}$	–	–	0.06	0.08*
			(0.07)	(0.03)
NY Times Coverage	0.21*	0.20*	0.18*	–
	(0.07)	(0.07)	(0.07)	
Δ*NY Times* Coverage	–	–	–	0.28*
				(0.07)
NY Times Coverage$_{t-1}$	–	–	0.04	0.22*
			(0.07)	(0.09)
Constant	9.38*	10.14*	9.92*	9.92*
	(3.33)	(3.38)	(3.59)	(3.59)
LM Test χ^2	0.19	0.50	0.17	0.17
p-value	0.66	0.48	0.68	0.68
Long-run multipliers				
Presidential Approval				0.10
				(0.18)
Economic Expectations				0.37**
				(0.13)
NY Times Coverage				1.00**
				(0.43)

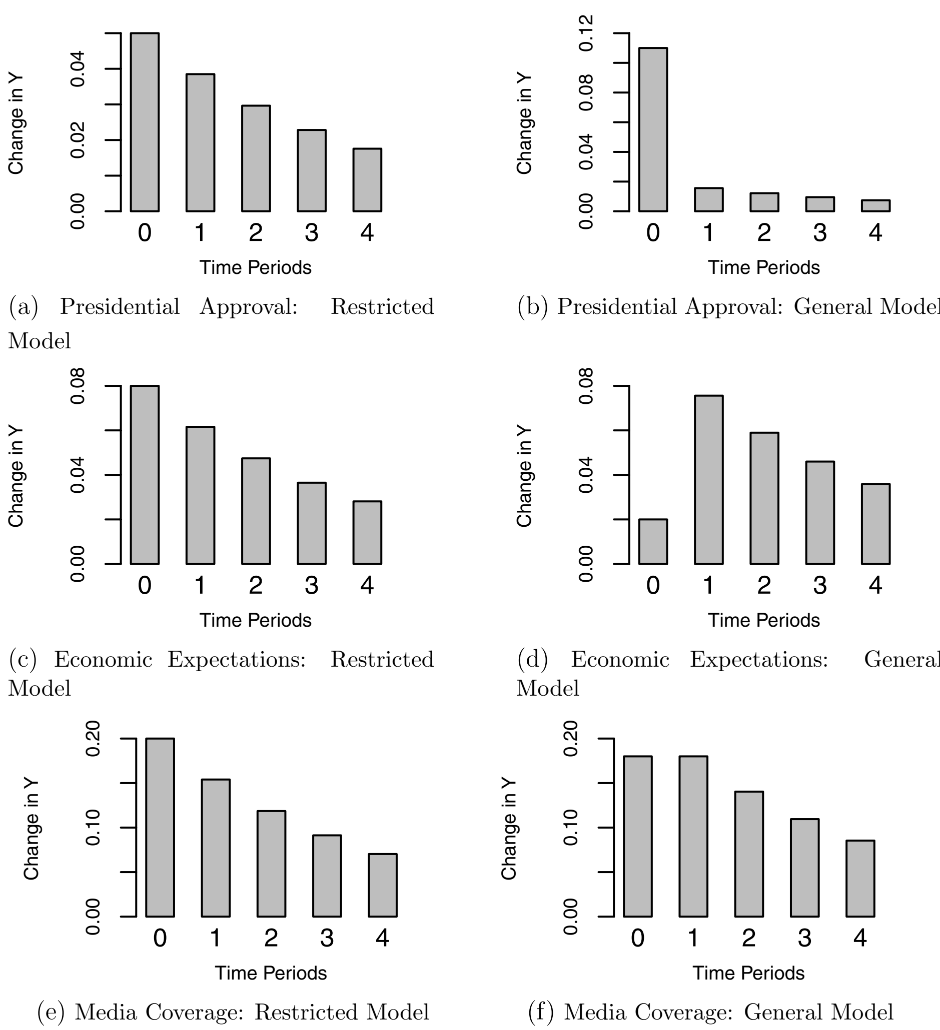
Note: Other variables from the analysis are included in the estimating equation but omitted from the table. Standard errors in parentheses. One-tailed tests. Long-run multiplier estimated via Bewley model. Lag length selected via AIC.
*p-value < 0.05, ** < 0.01.

while the effect of economic expectations is entirely across future quarters. The effect of media coverage, on the other hand, is nearly evenly balanced across the short and long run.

The long-run effects for the results in Table 5 using the ADL are as follows: presidential approval $\frac{(.11+(-.09))}{1-.77} = .08$; economic expectations $\frac{(.02+.06)}{1-.77} = .34$; and media coverage $\frac{(.18+.04)}{1-.77} = .96$. And using the ECM, the results are as follows: presidential approval $\frac{.02}{.22} = .09$; economic expectations $\frac{.08}{.22} = .36$; and media coverage $\frac{.22}{.22} = 1$. The minor differences are due to rounding. The long-run multipliers for two of the three covariates are statistically significant. The move to the general model tells us nothing more about the long-run effects. This is not particularly surprising in this case as the long- and short-run effects are not highly correlated (our omitted variable was uncorrelated with the included variables). The cost paid in this case was an incorrect inference about the short-run dynamics for presidential approval.

We can more clearly see this cost in the lag distribution implied by the restricted (PA) model estimated by DGW for these three variables compared with the lag distributions from the general model we estimated (see Figure 2). The PA model the authors estimate constrains

FIGURE 2 Comparison of Estimated Lag Distributions for Congressional Approval



each lag distribution to a pattern of geometric decay, as can be seen in the top row of Figure 2. The lag distributions from the more general model (in the second row of Figure 2) demonstrate how each variable has a different dynamic impact on congressional approval. The effect of presidential approval is almost entirely immediate, while the effect of economic expectations is mostly felt at $t+1$ and beyond. The effect of media coverage has both a large contemporaneous and long-run component. By calculating the median lag length, we get a better sense of how much of the total effect is short versus long. For presidential approval, the median lag is one, as 61% of the effect occurs in the first lag. Compare that to the effect of the economy, where 52% of the effect occurs by the third quarter. We find a similar pattern with the mean lags. The mean lag for presidential approval is just over one lag, while that for economic expectations is over four lags. So congressional approval takes eight months longer to adjust to a change in the economy than a change in presidential approval.

Implications for Dynamic Analysis

Time series analysis is an important part of statistical analysis in political science, and linear models are a critical tool in the analyst's tool box. As the number of time series datasets and the interest in political change continue to grow, so too will the number of dynamic linear models we estimate. If we are to make the most of the growth in data sources and the attendant body of research growing up around it, it is imperative that we revisit dynamic specification, paying attention to basic dictums of econometric time series analysis and drawing detailed interpretations about dynamics from the models we estimate. Until and unless we do, our knowledge of political change will suffer. We summarize our work by offering a strategy for the specification and interpretation of dynamic regressions for stationary and (at least) weakly exogenous regressors that we think will help build better models and in turn lead to a greater accumulation of theory, enhancing knowledge about political change.

First, applied analysts should begin with general models. While theory is a necessary condition for building good dynamic specifications, it is seldom sufficient. Given that caveat, good econometric practice is to start with general models and test using *t* or *F* tests or the AIC to be sure any restrictions imposed are consistent with the data. This will avoid biased estimates of short- and long-run coefficients as well as equilibria, error correction rates, and mean and median lag lengths. The analyst has an obligation to inform readers about this process. The task is also a simple one, requiring only OLS and due diligence.

Second, analysts should consider whether ECMs provide dynamic quantities in a form that is more natural or useful than the ADL while remembering that the two are equivalent. Alternately, as the ECM is useful for stationary and integrated data alike, analysts need not enter debates about unit roots and cointegration to discuss long-run equilibria and rates of reequilibration. As we have shown, these dynamic quantities are implied by virtually all dynamic regressions involving stationary data, although some specifications impose restrictions on these quantities. In fact, analysts are under some obligation to discuss equilibration and error correction rates when dynamic regressions of any form are estimated.

Third, and finally, regardless which general model is used, analysts should extract all the information available to them in the model. We should report and interpret error correction rates, long-run multipliers, lag distributions, and mean and median lags. These quantities tell us how the dynamics of politics unfold; they reveal the timing of responses and patterns of change that characterize much of what we study. Taken together, careful specification and interpretation of dynamic linear models will enable us to have confidence in the estimates from our models and to draw more complete interpretations from them so that we will better understand change in the world around us.

References

Bannerjee, Anindya, Juan Dolado, John W. Galbraith, and David F. Hendry. 1993. *Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data.* Oxford: Oxford University Press.

Beck, Nathaniel. 1985. "Estimating Dynamic Models Is Not Merely a Matter of Technique." *Political Methodology* 11: 71–89.

Beck, Nathaniel. 1991. "Comparing Dynamic Specifications: The Case of Presidential Approval." *Political Analysis* 3: 27–50.

Beck, Nathaniel. 1992. "The Methodology of Cointegration." *Political Analysis* 4: 237–48.

Beck, Nathaniel, and Jonathan N. Katz. 1995. "What to Do (and Not to Do) with Time-Series cross-Section Data." *American Political Science Review* 89 (September): 634–47.

Beck, Nathaniel, Jonathan N. Katz, and Richard Tucker. 1998. "Taking Time Seriously: Time-Series Cross-Section Analysis with a Binary Dependent Variable." *American Journal of Political Science* 42 (October): 1260–88.

Bewley, R. A. 1979. "The Direct Estimation of the Equilibrium Response in a Linear Model." *Economic Letters* 3: 357–61.

Box-Steffensmeier, Janet M., and Suzanna De Boef. 2006. "Repeated Events Survival Models: The Conditional Frailty Model." *Statistics in Medicine* 25 (December): 3518–33.

Box-Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*. New York: Cambridge University Press.

Box-Steffensmeier, Janet M., and Renée M. Smith. 1998. "Investigating Political Dynamics Using Fractional Integration Methods." *American Journal of Political Science* 42 (April): 661–89.

Brandt, Patrick, and John T. Williams. 2001. "A Linear Poisson Autoregressive Model: The Poisson AR(p)." *Political Analysis* 9 (Spring): 164–84.

Clarke, Harold D., Marianne C. Stewart, Michael Ault, and Euel Elliot. 2005. "Men, Women, and the Dynamics of Presidential Approval." *British Journal of Political Science* 35 (January): 31–51.

Davidson, Russell, and James G. MacKinnon. 1993. *Estimation and Inference in Econometrics*. New York: Oxford University Press.

De Boef, Suzanna. 2001. "Testing for Cointegrating Relationships with Near-integrated Data." *Political Analysis* 9 (Winter): 78–94.

De Boef, Suzanna, and Jim Granato. 1997. "Near-integrated Data and the Analysis of Political Relationship." *American Journal of Political Science* 41 (April): 619–40.

De Boef, Suzanna, and Paul Kellstedt. 2004. "The Political (and Economic) Origins of Consumer Confidence." *American Journal of Political Science* 38 (October): 633–49.

Durr, Robert H. 1992a. "An Essay on Cointegration and Error Correction Models." *Political Analysis* 4: 185–228.

Durr, Robert H. 1992b. "Of Forests and Trees." *Political Analysis* 4: 255–58.

Durr, Robert H., John B. Gilmour, and Christina Wolbrecht. 1997. "Explaining Congressional Approval." *American Journal of Political Science* 41 (January): 175–207.

Enders, Walter. 2001. *Applied Econometric Time Series*. 2nd ed. New York: Wiley and Sons.

Freeman, John R., John T. Williams, and Tse-min Lin. 1989. "Vector Autoregression and the Study of Politics." *American Journal of Political Science* 33 (November): 842–77.

Hendry, David F. 1995. *Dynamic Econometrics*. Oxford: Oxford University Press.

Heo, Uk. 1996. "The Political Economy of Defense Spending in South Korea." *Journal of Peace Research* 33 (November): 483–90.

Huber, John D. 1998. "How Does Cabinet Instability Affect Political Performance? Portfolio Volatility and Health Care Cost Containment in Parliamentary Democracies." *American Political Science Review* 92 (September): 577–91.

Keele, Luke J., and Nathan J. Kelly. 2006. "Dynamic Models for Dynamic Theories: The Ins and Outs of Lagged Dependent Variables." *Political Analysis* 14 (Spring): 186–205.

Mishler, William, and Reginald S. Sheehan. 1996. "Public Opinion, the Attitudinal Model, and Supreme Court Decision Making: A Micro-Analytic Perspective." *The Journal of Politics* 58 (February): 169–200.

Mondak, Jeffery J., and Shannon Ishiyama Smithey. 1997. "The Dynamics of Public Support for the Supreme Court." *The Journal of Politics* 59 (November): 1114–42.

Nadeau, Richard, Richard G. Niemi, David P. Fan, and Timothy Amato. 1999. "Elite Economic Forecasts, Economic News, Mass Economic Judgments, and Presidential Approval." *The Journal of Politics* 61 (February): 109–35.

Ostrom, Charles W., Jr., and Renée Smith. 1992. "Error Correction, Attitude Persistence, and Executive Rewards and Punishments: A Behavioral Theory of Presidential Approval." *Political Analysis* 4: 127–84.

Smith, Renée. 1992. "Error Correction, Attractors, and Cointegration: Substantive and Methodological Issues." *Political Analysis* 4: 249–54.

Stimson, James A. 1985. "Regression Models in Space and Time: A Statistical Essay." *American Journal of Political Science* 29 (November): 914–47.

Swank, Duane, and Sven Steinmo. 2002. "The New Political Economy of Taxation in Advanced Capitalist Democracies." *American Journal of Political Science* 46 (July): 642–55.

Williams, John T. 1990. "The Political Manipulation of Macroeconomic Policy." *The American Political Science Review* 84 (September): 767–95.

Williams, John T. 1992. "What Goes Around Comes Around: Unit Root Tests and Cointegration." *Political Analysis* 4: 229–36.

Williams, John T., and Brian K. Collins. 1998. "The Political Economy of Corporate Taxation." *American Journal of Political Science* 41 (January): 208–44.