Not Just for Cointegration: Error Correction Models with Stationary Data

Luke Keele
Department of Politics and International Relations
Nuffield College and Oxford University
Manor Road, Oxford
OX1 3UQ UK
Tele: +44 1865 278716
Email: luke.keele@politics.ox.ac.uk

Suzanna De Boef
Department of Political Science
Pennsylvania State University
State College, PA 16802
Tele: 814-863-9402
Email: sdeboef@psu.edu

December 16, 2004

Abstract

The error correction model is generally thought to be isomorphic to integrated data and the modeling of cointegrated processes, and as such, is considered inappropriate for stationary data. Given that many political time series are not integrated, analysts are unable to take advantages of the error correction model's ability to capture both long and short-term dynamics in a single statistical model. We use analytical results to demonstrate that error correction models are appropriate for stationary data. We use simulated data to then demonstrate the equivalency between auto-distributed lag models and error correction models. Finally, we reestimate a model of Supreme Court approval from the literature to demonstrate how the use of an error correction model enhances our understanding of political dynamics.

In 1993, *Political Analysis* published a series of articles designed to introduce cointegration and error correction models to the political science literature. This set of six articles is perhaps one of the best introductions to cointegration and error correction methods in print. But what is particularly interesting in this set of articles, besides the lucid introduction to cointegration, is the debate that develops among the authors on what is by-and-large considered to be a non-controversial statistical technique in disciplines outside political science - whether error correction models are suitable for stationary data. Two of the authors maintain that error correction models were developed prior to the theory of cointegration and are flexible enough to model stationary data that are long-memoried (Beck 1993; Williams 1993). Other authors argue that cointegration implies error correction and that error correction models in turn imply cointegration. As such, they see error correction models as unsuitable for stationary data (Durr 1993*a*,*b*; Smith 1993). As Smith (1993) argues, while error correction models predate the theory of cointegration, the data modeled with early error correction mechanisms were almost certainty cointegrated.

This debate over error correction and stationary data is peripheral to the central points in these articles, and the debate remains unresolved. But this controversy has substantive implications for not just how time series data are modeled, but also for how theories of political dynamics are developed. Apart from the desire to apply the most suitable modeling strategy, error correction models also offer the important benefit of allowing estimation of both short and long-term effects. For analysts, this allows consideration of theories where the dynamic effects include both short-term shocks and longer equilibrium forces. But given the paucity of true cointegrating relationships in political data, the power of error correction models is either lost to most applied analysts in political science or requires tortured arguments for cointegration. In this paper, we take up the debate raised in that issue of *Political Analysis*, in order to investigate the properties of error correction models when used with stationary data. In doing so, we offer applied analysts a new tool that can change the way we think about political dynamics.

1 Error Correction and Cointegration

The tight linkage between cointegration and error correction models stems from the Granger representation theorem. According to this theorem, two or more integrated time series that are cointegrated have an error correction representation, and two or more time series that are error correcting are cointegrated (Engle and Granger 1987). In short, the two concepts are isomorphic, as each implies the other. As such, any discussion of how to statistically model integrated data must make reference to error correction models, and most econometric textbooks treat error correction models and cointegration as the same topic, with cointegration as the precursor to any consideration of error correction models. But this isomorphism only holds with integrated time series processes, which implies that error correction models may have applications beyond estimating cointegrating relationships.

Moreover, there are reasons we might want to use error correction models with time series that are not integrated. We might want to model both long and short-term forces simultaneously in a single statistical model just as error correction models allow. To begin, we review how an error correction model works in order to demonstrate why we might want to use it more broadly.

Before introducing the error correction model, we define some basic terms. First, let's say we have a time series Y_t that is a function of past values and some random error:

$$Y_t = \rho Y_{t-1} + \varepsilon_t \tag{1}$$

The ρ term dictates how strongly the present value of Y is dependent on the prior value of Y while ε_t is an i.i.d. random variable. If $|\rho| = 1$, we say that Y_t is integrated, and the series has a permanent memory such that past shocks to the series cumulate. An integrated series has a mean and variance that depend on time (Bannerjee et al. 1993). If we have an explanatory variable that is also integrated and causally related to Y then we say the two series are cointegrated. Cointegration implies that the two integrated series never drift far apart from each other, that is they maintain an equilibrium. The classic example of a cointegrating relationship is that of short and long term interest rates (Engle and Granger 1987). It would be surprising if these two series drifted apart from one another over time as that would present opportunities for arbitrage. The error correction model is the preferred method for estimation when two integrated time series are statistically related or cointegrated since the error correction model can be formally derived from the properties of integrated time series.

The error correction model, however, is particularly powerful since it allows an analyst to estimate both short term and long run effects of explanatory time series variables. For example, let us consider a bivariate single-equation error correction model:[1]

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

In equation 2, current changes in Y are a function of current changes in X (the first difference of X) and the degree to which the two series are outside of their equilibrium in the previous time period. Specifically, β_0 captures any immediate effect that X has on Y, described as a contemporaneous effect or short-term effect. The coefficient, β_1, reflects the equilibrium effect of X on Y. It is the causal effect that occurs over future time periods, often referred to as the long-term effect that X has on Y. Finally, the long-term effect occurs at a rate dictated by the value of α_1. A simple example demonstrates the interpretation of all three coefficients in the model. Let's say we regress the first difference of a presidential approval time series on one lag of presidential approval, one lag of economic expectations, and the first difference of economic expectations as in Equation 2. The estimated coefficients are $\hat{\beta_0}$ = 0.5, $\hat{\alpha_1}$ = -0.5, and $\hat{\beta_1}$ = 1.0. If economic expectations were to increase five points, how will that affect presidential approval in the context of the error correction model? First, presidential approval will increase 2.5 points immediately (5 * 0.5, the coefficient of $\hat{\beta_0}$). But the error correction model implies that presidential approval and economic expectations also have an equilibrium relationship, where this increase in economic expectations disturbs the equilibrium, causing presidential approval to be too low. As a result, presidential approval will increase another five points (5 * 1.0, the coefficient for $\hat{\beta_1}$). But the increase in presidential approval (or re-equilibration, 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

[1]There are two different types of error correction model. First is the Engle-Granger two-step method, and the second, the one I consider, is the single-equation error correction model. The two-step estimator has been shown to be inferior to the single-equation estimator both empirically (DeBoef 2001; Durr 1993*a*), and theoretically (Beck 1993).

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.

When considering this model, analysts tend to focus on its statistical properties with integrated data. We, instead, emphasize the theoretical advantages of thinking in terms of error correcting processes. In cross-sectional data, without a temporal component to the data, all estimated effects are static, and we cannot assess whether causal effects have any residual component that persists into future time periods. This is not the case when dealing with time series data. A change in X may affect Y immediately, or the effect of X on Y may be delayed, occurring in the future across several time periods. This is a very different way to think about the causal effects between independent and dependent processes. With temporal data, there are at least three possible combinations of dynamic effects:

- An X variable may have only contemporaneous effects, where X affects Y immediately, but that effect does not persist into the future. In the presidential approval example, this occurs when β_1 is equal to 0.

- An X variable may have a contemporaneous effect as well as an equilibrium component that persists across future time periods and decaying at some rate. This is a situation identical to the presidential approval example above.

- An X variable may have no contemporaneous effect, but instead have an equilibrium effect, where the causal effect on Y only occurs across future time points. In the context of the presidential approval example, this would occur if β_0 is equal to 0.

The power of error correcting models is that they allow us to estimate and test for all three types of effects. The ability to estimate such effects ought to encourage the development of theories of political dynamics that incorporate these questions. For applied analysts, this means greater attention should be paid to whether their dynamic theory implies effects that are contemporaneous, equilibrium or both.

For example, consider a second example in the context of trust in government. Perhaps we believe that trust in government responds to both presidential approval and social capital in the form of civic engagement and interpersonal trust. We might expect trust to quickly register the successes and failures of the president, giving presidential approval a contemporaneous effect on trust that may persist into the future. But a shock to social capital should only have an equilibrium (or long term) effect. A change in levels of civic engagement takes time to filter through society, and the impact on attendant attitudes of interpersonal trust will take time as well, effectively ruling out the presence of any contemporaneous effects. While we could estimate a simple model to test whether presidential approval and social capital have any effect on trust, the test of the theory is more powerful if we can posit both immediate and longer lasting effects and model them accordingly.

Generally, if one believed that an error correction model is appropriate for such a theory of trust in government, the methodological question would be whether trust in government, presidential approval, and social capital are all integrated (highly unlikely) and whether all three are cointegrated (also unlikely). The problem with following this route is that the analyst assumes that long-term effects only occur when the time series in question are integrated. In many cases the behavior story of the error correction model is so attractive that it drives analysts to assert cointegration in the processes they are studying miring them in the debate over the validity of unit root characterizations in political time series. In a search of the literature, error correction models appear with discussions of cointegration ## times, but without only ## times.

But time series can, in fact, have long memories and still be stationary. The methodological question is: can we use the error correction model to capture the long and short-term cycles that can occur in stationary data? More precisely, can we estimate the error correction model when ρ is less than one? At present, most analysts believe they cannot and there is little evidence to the contrary once we move out of the realm of integrated or near-integrated data.[2] In the next section, we present an analytical derivation of the single-equation error correction model that proves that the error correction model is for more than integrated data and cointegrating relationships.

2 The Analytics of Error Correction Models

The standard way to derive the error correction model is to show that if X and Y are linear functions of a latent integrated process, the residuals of Y regressed on X should be stationary. This derivation of the error correction model starts with the assumption that both Y and X are integrated and demonstrates that the error correction model captures the equilibrium causal movements between these two cointegrated processes. Occasionally, however, some authors derive the error correction model from a different and more promising starting point (Bannerjee et al. 1993; Davidson and MacKinnon 1993; Verbeek 2000).

In these derivations, the starting point for the error correction model is the autoregressive distributed lag (ADL) model. The ADL model is an extremely flexible model for time series data, and is often seen in the following bivariate form:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t \tag{3}$$

Specifically, this is an ADL(1,1) model, where the notation refers to the number of lags included in the model. It generalizes to an ADL(p,q) where p refers to the number of lags of Y and q refers to the number of lags of X included in the model. Given that the ADL (1,1) has a lagged dependent variable on the right side, it should be consistently estimated by OLS[3] and has a stationarity condition, that is Y_t must be stationary (Davidson and MacKinnon 1993). Now if one were to estimate an ADL(1,1) model with say presidential approval as Y_t and economic expectations as X_t, the short-run effects of economic expectation are readily estimated in the model by the coefficients β_0

[2] A near-integrated time series in one that is defined as $|\rho| = 1 + c$, where c is negative and small (DeBoef and Granato 1997).

[3] The proof for the consistency of OLS assumes that ε_t is IID after the lag of Y is included in the model.

and β_1, which give the immediate effect of a change in X at some given t. Any long-run equilibrium effects are given by the unconditional expectations or the expected value of Y_t. Let $y^* = E(Y_t)$ and $x^* = E(X_t)$ for all t. If the two processes moved together with out error, in the long-run, they would converge to the following equilibrium values:

$$y^* = \alpha_0 + \alpha_1 y^* + \beta_0 x^* + \beta_1 x^* \tag{4}$$

Solving for y_t^* in terms of x_t^*, yields:

$$y^* = \frac{\alpha_0}{1-\alpha_1} + \frac{\beta_0 + \beta_1}{1-\alpha_1} x^* \tag{5}$$

If we group terms, the long-term value of y^* is:

$$y_t^* = k_0 + k_1 x^* \tag{6}$$

This equation represents the values for which Y and X are in equilibrium in the long run, and k_1 represents the long-run multiplier of X with respect to Y. Any deviation from equilibrium, $y^* - (k_0 \; + \; k_1 x^*) \neq 0$, should induce change back to the equilibrium in the next period. We cannot, however, directly estimate the rate of return to equilibrium in the ADL model. As such, analysts seldom draw inferences from ADL models about the rate of error correction implied by the model. However, such information about error correction is more readily available if a set of linear transformations are applied to the ADL model. This set of linear transformations also allow one to derive the error correction model from the ADL model. Doing so imposes no restrictions on the parameters in the model, and thus the two models contain the same information, implying the same behavioral relationship. Parts of this proof can be seen in other sources such as Davidson and MacKinnon (1993) and Bannerjee et al. (1993). We derive it in greater detail and with more attention paid to the equivalence between the ADL and error correction models. To see this, again consider an ADL(1,1) model:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t \tag{7}$$

First, we take the first difference of Y to produce:

$$\Delta Y_t = \alpha_0 + (\alpha_1 - 1)Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_t \tag{8}$$

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

$$\Delta Y_t = \alpha_0 + (\alpha_1 - 1)Y_{t-1} + \beta_0 \Delta X_t + (\beta_0 + \beta_1)X_{t-1} + \varepsilon_t \tag{9}$$

And next add and subtract $(\alpha_1 - 1)X_{t-1}$ from the right hand side and rewrite to produce the Generalized Error Correction Model (GECM):

$$\Delta Y_t = \alpha_0 + \gamma(Y_{t-1} - X_{t-1}) + \lambda_1 \Delta X_t + \lambda_2 X_{t-1} + \varepsilon_t \tag{10}$$

where $\gamma = (\alpha_1 - 1)$, $\lambda_1 = \beta_0$, and $\lambda_2 = \beta_1 + \beta_0 + \alpha_1 - 1$.

The GECM, unlike the ADL model, directly tells us how quickly the system reacts to any disequilibrium, as γ the coefficient on the lag of Y is the error correction rate. In other words, the term $(\alpha_1 - 1)$ can be interpreted as the speed at which Y adjusts to any discrepancy between Y and X in the previous period. One can see that γ must be negative since it is just $\alpha_1 - 1$. The term $(Y_{t-1} - X_{t-1})$ is zero when X and Y are in equilibrium and measures the extent to which the long-run relationship is not satisfied.

Does the GECM produce the same long and short-run effects observed in the ADL model? In a word, yes. Taking each in turn, the short-run effects are represented by λ_1 and $\lambda_2 - \lambda_1 - \gamma$ in the GECM. By substitution, these are $\lambda_1 = \beta_0$ and $\lambda_2 - \lambda_1 - \gamma = \beta_1 + \beta_0 + \alpha_1 - 1 - \beta_0 - (\alpha_1 - 1) = \beta_1$, which are the short-run effects estimated in the ADL model.

Next, one can also derive the long run multiplier from the GECM. Assume the equilibrium relationship is given by $y^* = k_1 x^*$, where again k_1 denotes the long run multiplier. The long-run effect of a change in X on Y is:

$$k_1 = -\frac{(\lambda_2 - \gamma)}{\gamma} \tag{11}$$

or by substitution:

$$k_1 = -\frac{(\lambda_2 - \gamma)}{\gamma} = -\frac{(\beta_1 + \beta_0 + \alpha - 1 - (\alpha_1 - 1))}{(\alpha_1 - 1)} = \frac{(\beta_1 + \beta_0)}{(1 - \alpha_1)} \quad (12)$$

Therefore the GECM produces the same value for k_1 as the ADL does. However, there is a more convenient form for estimating error correction models (which we denote as the Error Correction Model (ECM)). Instead of explicitly including an error correction term in the model in the form of $(Y_{t-1} - X_{t-1})$, we can estimate the following regression:

$$\Delta Y_t = \alpha_0 + \gamma Y_{t-1} + \eta_1 \Delta X_t + \eta_2 X_{t-1} + \varepsilon_t \quad (13)$$

where $\gamma = (\alpha_1 - 1)$, $\eta_1 = \beta_0 = \lambda_1$, and $\eta_2 = \beta_1 + \beta_0$. This model can be rewritten in the following error correcting form:

$$\Delta Y_t = \alpha_0 + \gamma(Y_{t-1} - \eta_2 X_{t-1}) + \eta_1 \Delta X_t + \varepsilon_t \quad (14)$$

The long run multiplier, k_1, is even more readily calculated from this form of the ECM:

$$k_1 = \frac{\eta_2}{\gamma} = \frac{(\beta_1 + \beta_0)}{(\alpha_1 - 1)} \quad (15)$$

Again, the term $(Y_{t-1} - \eta_2 X_{t-1})$ is zero when X and Y are in equilibrium and measures the extent to which the long-run relationship is not satisfied. The term γ is still interpreted as the speed at which Y adjusts to any discrepancy between Y and X in the previous period. It is this component of the model that captures long-term effects. But as we have seen in the earlier examples, the error correction model is interesting not just because it can model such equilibrium behavior, but also because it captures any contemporaneous effects that may occur. In equation 14, the terms, η_1 is equal to β_0 in the ADL and $\eta_2 - \eta_1$ is equal to β_1. both of which capture any immediate effect X may have on Y independent of the equilibrium relationship between the two processes.

What distinguishes the error correction model from the ADL(1,1)? In the error correction model, the speed of adjustment in the equilibrium relationship appears directly in the model (Bannerjee et al. 1993) while in the ADL model the long-run multiplier must br calculated. Another advantage of the error correction model is that since the dependent variable in the model is differenced, any danger of estimating a spurious regression with near-integrated data is eliminated. And finally, given that the error correction model is a linear reparameterization of the ADL, it can be estimated with OLS. The error correction model, then, offers a means of testing theories that formulate subtle political dynamic mechanisms.

To better demonstrate the equivalence between the ADL model, GECM, and ECM, we next use simulated data to estimate each model and calculate the short and long run effects of each. After this example, we explore how using an error correction model can change our understanding of political dynamics by estimating a model from the literature on approval of the Supreme Court. This model was estimated with standard techniques for stationary data that are technically correct, but the model lacks the dynamic richness the error correction model can provide.

2.1 An Example of Model Equivalence

While we have demonstrated that the ADL model, GECM, and ECM all estimate the same quantities just in different forms, an example brings this point into sharper focus. For the example, we use simulated data. The data generation process (DGP) for Y is the an ADL(1,1) model and the DGP for X is a simple autoregressive process:

$$\begin{aligned} Y_t &= \alpha_0 + \alpha_1 Y_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \varepsilon_{1t} \\ X_t &= \rho X_{t-1} + \varepsilon_{2t} \end{aligned} \quad (16)$$

where ρ is 0.75 making the model clearly stationary. We set the parameter values for the Y DGP as follows: $\alpha_0 = 0$, $\alpha_1 = 0.75$, $\beta_0 = 0.50$, and $\beta_1 = 0.25$. We, then, estimated an ADL model, a GECM, and an ECM. The results appear in Table 1.

Table 1: ADL and ECM Estimates

	ADL Model	GECM	ECM
Y_{t-1}	0.75	–	-0.25
	(0.02)		(0.02)
Error Correction Term	–	-0.25	
		(0.02)	
X_t	0.53	–	–
	(0.06)		
X_{t-1}	0.25	0.52	0.77
	(0.07)	(0.05)	(0.07)
ΔX_t	–	0.53	0.53
		(0.06)	(0.06)
N	249	249	249
R^2	.94	.45	.45
Note: Simulated Data.			

First, let us examine the short run effects of X on Y. For the ADL model in column 1 these are given explicitly by the estimated coefficients β_0 and β_1, which are 0.53 and 0.25 respectively. For the GECM and ECM, λ_1 and η_1 are both equal to β_0, which is what we find as both of these estimated coefficients are 0.53 (notice that the estimates of the standard errors are also the same). For the second short-run effect, the comparison is less obvious. In the ADL, $\beta_1 = 0.25$. To calculate this effect in the GECM, we use $\lambda_2 - \lambda_1 + \gamma$, which equals again 0.25. Finally, for the ECM we just use $\eta_2 - \eta_1$, again this is 0.25.

We can also calculate the exactly same values for the long-run multiplier across all the models. Recall that for the ADL model, the long-run multiplier, k_1, is $\frac{(\beta_1+\beta_0)}{(1-\alpha_1)}$. Using the values from the estimated model, we find that $\hat{k_1}$ is 3.06. In the GECM, the long-run multiplier is: $\frac{(\lambda_2-\gamma)}{\gamma}$. Using the coefficients from the model, $\hat{k_1}$ is again 3.06. And this value is most easily calculated in the ECM where it is simply: $\frac{\eta_2}{\gamma}$ again 3.06. In short, the three estimated models are all equivalent. The only difference being that different steps are required to recover the various values implied by the DGP. The ECM offers the information in the most accessible form. Which model is the right model? In point of fact, they all are. The ECM and GECM are perhaps most attractive in that in these models the theory of equilibrium behavior is explicit instead of implicit. But more to the point, there is no need for elaborate arguments that two processes are cointegrated. If the theory dictates error correcting behavior the model can be estimated with a this theoretical justification. In the next section, we re-estimate a model of Supreme Court approval from literature using an error correction model in order to demonstrate how thinking about long and short-term effects changes how we consider about political dynamics.

3 An Example of an ECM with Stationary Data

We consider a time series model from the applied literature on public opinion to demonstrate how using an error correction model changes our understanding of political dynamics over and above using standard techniques for stationary data. The model is from Durr, Martin and Wolbrecht (1993). The authors argue that public support for the Supreme Court over time is a function of three factors: (1) how far the ideological position of the Court diverges from the ideological position of the public, (2) countermajoritarian behavior by the Court, and (3) general evaluations of government. Their dependent variable is a semi-annual time series measure of public support for the Supreme Court built from a variety of survey items. They find in their tests that the series is stationary, as we would expect. It seems unlikely that the Dred Scott case still influences public opinion about the Supreme Court, but it is not unreasonable to think that *Roe v. Wade* still matters, implying the presence of long-term (but not permanent) memory in support for the Supreme Court. The properties of this time series are similar to many in applied work: surely not integrated, but likely to be long-memoried.

The authors contend that support for the Supreme Court is a dynamic process, where attitudes toward the Supreme Court are a function of past attitudes toward the Supreme Court as modified by new information on ideological divergence, countermajoritarian behavior, and general evaluations of government. As such, they expect past shocks to feed forward into the future, decaying at an exponential rate, therefore shocks from the last period will matter half as much in the present time period, and so on. To capture these dynamics, the authors use a lagged dependent variable. Given the theory, the dynamic specification used in by the authors is entirely reasonable.

The authors find that ideological divergence affects public support for the Court as does congressional approval, an indicator of general government approval. Countermajoritarian behavior and presidential approval have little effect. However, the coefficient for past support for the Supreme Court fails to be significant, suggesting that some other dynamic process may be at work.

We develop a different dynamic theory within an error correction framework. First, we argue that the ideological position of the Court is salient only for a brief period in any given year–the late Spring and early Summer when decisions are announced. The Court mostly escapes the public eye for the rest of year, with only brief attention paid to the cases the Court accepts and to oral arguments. These events provide little, if any, information about the Court's ideological position relative to that of the public. Arguably, any divergence between the Court's and the public's ideological position is only briefly salient and likely recedes. Thus, the effect of ideological divergence is probably immediate, but also short-lived.

The effect of congressional approval, however, arguably takes a different trajectory. Congressional approval should affect support for the Supreme Court, since it is a proxy for general feelings about the government. Any change in general feelings about the government should matter immediately, as new information alters the status quo evaluation. The effect, however, should persist in that the public will be able to readily generalize their feelings about government into a specific opinion about the Supreme Court. As such, there should be an equilibrium relationship between the two as both are indicators of a general government evaluation.

In the context of an error correction model, this suggests two types of effects. For ideological divergence, the effect on feelings toward the Supreme Court should be contemporaneous with the announcement of decisions, but given the brief salience of these decisions, ideological divergence will be unlikely to have long-term effects. Congressional approval, however, should have a contemporaneous effect as well as a long-term impact, as general feelings of government move together and equilibrate over time. Specifically, we expect that the short-term effect of ideological divergence will be statistically significant, while both the short and long-term effects of congressional approval will be significant.

Table 2: Model Comparison of Public Support for the Supreme Court

	Supreme Court Support_t	ΔSupreme Court Support_t
Supreme Court Support_{t-1}	0.24	$-0.72**$
	(0.14)	(0.16)
Ideological Divergence_t	$-5.48*$	–
	(2.58)	
Congressional Approval_t	$0.47**$	–
	(0.13)	
ΔIdeological Divergence_t	–	$-5.97*$
		(2.77)
Ideological Divergence_{t-1}	–	-4.17
		(3.39)
ΔCongressional Approval_t	–	$0.46*$
		(0.21)
Congressional Approval_{t-1}	–	$0.44**$
		(0.15)
Constant	$41.63**$	$40.25**$
	(11.74)	(12.24)
N	41	41
Adjusted R^2	0.47	0.43
Box-Ljung Q Test	8.20	9.54
χ^2 p-value	0.97	0.95

Note: OLS Estimates. Standard Errors in Parentheses.
Data are semi-annual, 1973:1 to 1993:2
Two tailed tests.
* $p < .05$
** $p < .01$

In the first column of Table 2, we include a model with the original dynamic specification as estimated by the authors. We estimate a more parsimonious version of the model, dropping presidential approval and the measures of countermajoritarian behavior, which were insignificant in all specifications. The omission of these variables has no effect on the results of either the original model or the new error correction model. Again, we see that both ideological divergence and Congressional approval are significant predictors, while past support for the Supreme Court is not.

The second column of Table 2 contains the new error correction specification. The error correction model matches the dynamics of these process to a greater extent than the lagged dependent variable model. The error correction coefficient and the lag of congressional approval are both highly significant, indicating that error correcting long-term equilibrium behavior occurs. As expected, ideological divergence only has a contemporaneous effect. The difference between the Supreme Court's ideology and the public matters briefly, but does not have a long-term effect as the Court's actions fade from people's memory. However, an equilibrium exists between general government feelings and support for the Supreme Court, and if one changes, the other will as well over a longer time horizon than the current time period.

Moreover, with the error correction model, we find a larger effect for congressional approval. A one point increase in congressional approval moves support Supreme Court up 0.46 points immediately, with another 0.44 point increase over future time periods for a total effect of 0.90 points. In the authors' model, the total effect of a one point increase in Congressional approval increases support for the Supreme Court about 0.62, a difference of nearly 0.30 points. The effect, then, is 1/3 larger than before. It is also important to note that the error correction model fits much better in terms of adjusted R^2. While the R^2 is slightly higher for the published model, it is quite low given the inclusion of a lagged dependent variable. In the error correction model, we expect a lower R^2, given that the dependent variable is differenced. Given the parsimonious specification, the size of the R^2 is impressive.

Error correction models, then, allow us to understand political dynamics as rich causal processes. The error correction model does not change the basic finding, but does allow for a richer understanding of support for the Supreme Court as a dynamic process.

4 Conclusion

Many political time series are in an awkward position. It is quite rare to find true unit roots in political science. With finite sample sizes, political time series often look like unit roots, but given their theoretical properties, it is highly unlikely, they are unit roots. Analysts often forget that the presence of a unit root in economic data is more than just an empirical property; it also implies theoretical properties that are rarely applicable in political science (Beck 1992).

While we do not often work with unit roots, we quite often have stationary data that contains long cycles or has some component that responds to both short and long-term forces. For now, most applied work has ignored such a possibility. The error correction model is both a theoretically desirable and empirically feasible approach to such data.

The analytical evidence that we have presented clearly demonstrates that it is entirely appropriate to estimate an error correction model with stationary data. The estimates provided will be no different from those estimated with an ADL model.

And error correction model will often provide additional theoretical tractability since there is no longer any reason to lump short and long-term dynamics together. The analyst can formulate and test theories that are better able to discriminate between two processes that have short versus long-run behavior.

The question for analysts then is: when do I use a single-equation model? The answer is let theory be the guide. When theory dictates there might be different long and short-term effects, then error correction may be the appropriate model. Testing whether the error correction model is appropriate is simple. Anyone who suspects error correcting behavior between two time series should first estimate the ADL(1,1) model and then test the the following restriction: $\alpha_1 - 1 = -(\beta_0 + \beta_1)$. In a single step, the coefficients for the lag of the Y variable and the lag of the X variable must be statistically significant and when regressed on the first difference of Y. If not, error correcting behavior does not occur and the analyst should consider some other dynamic specification.

Conveniently, the error correction model requires no special software. Since error correction models are estimated with OLS, one only needs a difference and lag operator to estimate an error correction model, which can be done in a package like Stata. Of course, as with any time series model, the analyst should ensure that the residuals of the estimated error correction model are not auto-correlated.

In the *Political Analysis* cointegration issue, one author thought error correction models were entirely appropriate for stationary data, but worried that they might become hegemonic. In retrospect, that appears to be far from the case. With any luck, error correction models will enjoy at least a half measure of success in the future.

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. 1992. "Comparing Dynamic Specifications: The Case of Presidential Approval." *Political Analysis* 3:27–50.

Beck, Nathaniel. 1993. "The Methodology of Cointegration." *Political Analysis* 4:237–248.

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

DeBoef, Suzanna. 2001. "Testing for Cointegrating Relationships with Near-integrated Data." *Political Analysis* 9:78–94.

DeBoef, Suzanna and Jim Granato. 1997. "Near-integrated Data and the Analysis of Political Relationship." *American Journal of Political Science* 41:619–640.

Durr, Robert H. 1993*a*. "An Essay on Cointegration and Error Correction Models." *Political Analysis* 4:185–228.

Durr, Robert H. 1993*b*. "Of Forests and Trees." *Political Analysis* 4:255–258.

Durr, Robert H., Andrew D. Martin and Christina Wolbrecht. 1993. "Ideological Divergence and Public Support for the Supreme Court." *American Journal of Political Science* 44:768–776.

Engle, Robert F. and C. W. J. Granger. 1987. "Co-integration and Error Correction: Representation, Estimation, and Testing." *Econometrica* 55:251–276.

Smith, Renee. 1993. "Error Correction, Attractors, and Cointegration: Substantive and Methodological Issues." *Political Analysis* 4:249–254.

Verbeek, Marno. 2000. *A Guide to Modern Econometrics.* New York: John Wiley and Sons.

Williams, John T. 1993. "What Goes Around Comes Around: Unit Root Tests and Cointegration." *Political Analysis* 4:229–236.