Autocorrelation

Harold D. Clarke
University of Texas, Dallas, Richardson, Texas, USA

Jim Granato
National Science Foundation, Arlington, Virginia, USA

Glossary

autocorrelated errors Correlations between stochastic errors ordered over either time or space in a model, typically a linear regression model.

autoregressive process A data-generating process with "memory," such that the value of the process at time t reflects some portion of the value of the process at time $t-i$.

common factor restriction A restriction on model parameters, produced in the context of linear regression analysis by the use of a quasi-differencing procedure as a generalized least-squares "correction" for autocorrelated errors.

fractionally integrated process A data-generating process with signficant autocorrelations at long lags, such that the system has long memory and shocks erode very slowly.

integrated process An autoregressive process with perfect memory, such that shocks to the system at time t are not discounted in subsquent time periods.

minimum state variable (MSV) solution A solution procedure for rational expectations models that uses the simplest, least parameterized characterization.

near-integrated process An autoregressive process where a very large portion of the value of the process at time t carries over to time $t+1$.

rational expectations equilibrium (REE) Citizen expectations, based on all available information (in the model), about an outcome that equals the outcome on average.

stationary A time (data) series (or model) is stationary if there is no systematic change in the mean (e.g., no trend), if there is no systematic stochastic variation, and if strict periodic variations (seasonal) are stable. Time plays no role in the sample moments.

Autocorrelation exists when successive values of a random variable, ordered over either time or across space, have nonzero covariances. A variety of data-generating processes (DGPs) may produce autocorrelation. Although interest in spatial autocorrelation has grown, most analyses of autocorrelation involve the use of time series data and, accordingly, the time series case is the focus of this article. Over the past three decades, econometricians have made major advances in the analysis of one class of such data, namely, nonstationary series produced by DGPs involving stochastic trends. Particularly important have been the development of the concept of cointegration and the explication of the relationship between this concept and the specification of error correction models for analyzing nonstationary data. Related work has led to procedures for analyzing highly autoregressive and fractionally integrated data. The former manifest very strong, but imperfect, autocorrelations across adjacent time periods; the latter have statistically significant autocorrelations at very long lags and are said to have long memory. Existing side by side with these developments are a set of procedures—some now over a half-century old—that derive from a narrow (not unimportant) concern that inferences regarding parameter estimates generated by ordinary least-squares regression analysis will be thwarted by violations of the assumption that errors are uncorrelated. We argue that these traditional procedures are potentially misleading since they ignore the possibility that autocorrelated residuals constitute evidence of model misspecification. We also contend that there is a related, larger need to develop and test theories that account for the observed autocorrelation in series of interest. The argument is illustrated by the development of a theoretical model of macropartisanship.

Conventional Wisdom and Conventional Practice

Historically, interest in autocorrelation (often referred to as serial correlation) has been motivated by the use of ordinary least-squares (OLS) regression to estimate the parameters of models of the form

$$Y_t = \beta_0 + \sum \beta_{1-k} X_{1-k,t} + \epsilon_t, \qquad (1)$$

where Y_t is the dependent variable at time t; $X_{1-k,t}$ are independent variables at time t; β_0 and β_{1-k} are parameters to be estimated; and ϵ_t is the stochastic error term $\sim N(0, \sigma^2)$.

As is well known, inferences concerning OLS estimates of the parameters of this model are problematic if the stochastic errors, ϵ_t, are correlated [i.e., cov(ϵ_t, ϵ_{t-1}) $\neq 0$]. When errors are correlated, parameter estimates are unbiased, but standard errors are affected. Hence, t ratios (i.e., $\beta/s.e.$) are inflated or deflated (depending on the sign of the correlation) and the risk of making Type I or Type II errors is enhanced.

Generations of researchers have learned about this threat to inference and the advisability of conducting postestimation diagnostics to determine whether the threat obtains. Since the errors, ϵ_t, are not observed, these diagnostics are performed using the residuals from the regression analysis (i.e., $\hat{\epsilon}_t$). Proceeding in this fashion, analysts—either explicitly or, more frequently, implicitly—make the strong assumption that their model is correctly specified. If the assumption is invalid, autocorrelated residuals may be a consequence of model misspecification, rather than autocorrelated errors.

Some econometric texts have begun to emphasize this point, but many researchers still behave as if they were unaware of it. As a result, they continue to commit the fallacy of concluding that autocorrelated residuals constitute necessary and sufficient evidence of an autocorrelated error process.

Although various diagnostic procedures can be employed to detect autocorrelation in regression residuals, the standard one is the Durbin-Watson test. The Durbin-Watson test statistic is computed as:

$$d = \frac{\sum(\hat{\epsilon}_t - \hat{\epsilon}_{t-1})^2}{\sum \hat{\epsilon}_t^2}. \qquad (2)$$

As its formula indicates, the Durbin-Watson test considers only first-order autocorrelation (i.e., $\hat{\epsilon}_t$ and $\hat{\epsilon}_{t-1}$). The implication is that the analyst, either explicitly or implicitly, is entertaining the hypothesis that

$$\epsilon_t = \rho\epsilon_{t-1} + v_t, \qquad (3)$$

where $v_t \sim N(0, \sigma^2)$.

Noting that $\sum \hat{\epsilon}_t^2$ and $\sum \hat{e}_{t-1}^2$ are approximately equal when N is large, the formula for d implies that $d \simeq 2(1-\rho)$. Thus, when there is perfect positive first-order autocorrelation (i.e., $\rho = 1.0$), $d = 0$, and when there is perfect negative first-order autocorrelation (i.e., $\rho = -1.0$), $d = 4$. Critical values for d vary by the number of regressors in a model and the number of data points and are characterized by an "indeterminate zone," where it is unclear whether the null hypothesis should be rejected.

Econometricians advise that the upper bound (d_u) of the critical values should be used in circumstances when regressors are changing slowly and caution that the test is not valid when one of the regressors is a lagged endogenous variable (e.g., Y_{t-1}). When a lagged endogenous variable is present, other tests (e.g., Durbin's h, Durbin's M) should be used.

If the null hypothesis of no (first-order) autocorrelation is rejected, the traditional response is to treat the autocorrelation as a technical difficulty to be "corrected," rather than evidence of possible model misspecification. The correction is to transform the data such that the error term of the resulting modified model conforms to the OLS assumption of no autocorrelation. This generalized least-squares (GLS) transformation involves "generalized differencing" or "quasi-differencing."

Starting with an equation such as Eq. (1), the analyst lags the equation back one period in time and multiplies it by ρ, the first-order autoregressive parameter for the errors [see Eq. (2) above]. Illustrating the procedure for a model with a single regressor, the result is

$$\rho Y_{t-1} = \rho\beta_0 + \rho\beta_1 X_{t-1} + \rho\epsilon_{t-1}. \qquad (4)$$

Then, (4) is subtracted from (1) to yield

$$Y_t - \rho Y_{t-1} = \beta_0 - \rho\beta_0 + \beta_1 X_t - \rho\beta_1 X_{t-1} + \epsilon_t - \rho\epsilon_{t-1}. \qquad (5)$$

This operation produces $\epsilon_t - \rho\epsilon_{t-1} = v_t$, which, by assumption, is a white-noise error process [$\sim N(0, \sigma^2)$]. The quasi-differencing operation involves computing $Y_t^* = Y_t - \rho Y_{t-1}$ and $X_t^* = X_t - \rho X_{t-1}$, and the resulting model is $Y_t^* = \beta_0 - \rho\beta_0 + \beta_1 X_t^* + v_t$, the parameters of which can be estimated by OLS.

Since the value of ρ typically is unknown, it must be estimated. Although various methods of doing so are available, the most popular (Cochrane-Orcutt) first uses OLS to estimate the parameters in model (1). The residuals then are used to estimate ρ. This $\hat{\rho}$ is used to transform Y_t and X_t to produce Y_t^* and X_t^*, respectively. The latter are then used in a second-round estimation and a new $\hat{\rho}$ is produced. This $\hat{\rho}$ then is used to transform the (original) data and another regression is performed. The process (called feasible GLS because ρ is estimated) continues until the value of ρ converges.

Common Factors and Dynamic Misspecification

Regardless of how ρ is estimated, the quasi-differencing procedure introduces a common factor restriction. This may be seen by rewriting Eq. (5) as

$$(1-\rho L)Y_t = (1-\rho L)\beta_0 + (1-\rho L)\beta_1 X_t + (1-\rho L)\epsilon_t, \quad (6)$$

where L is the lag operator (i.e., $LY_t = Y_{t-1}$ and $L^k Y_t = Y_{t-k}$). Thus, Eq. (6)—the model implied by the conventional wisdom that first-order autoregessive residuals entail a first-order autoregressive error process—is a restricted version of an autoregressive distributed lag model. An example of such a model is

$$Y_t = \alpha_0 + \lambda_1 Y_{t-1} + \gamma_1 X_t + \gamma_2 X_{t-1} + \nu_t. \quad (7)$$

The requirement that $\gamma_2 = -\lambda_1\gamma_1$ in Eq. (7) produces a common-factor (COMFAC) restriction similar to that in model (6)—the model implied by the quasi-differencing procedure typically utilized to deal with first-order autocorrelation in the residuals.

The validity of the restriction is an empirical question. Some econometricians stress the importance of COMFAC testing to help guard against misspecifying the dynamics in one's model. The test is easily implemented and can help one avoid the potentially misleading *non sequitur* of concluding that autocorrelated residuals imply autocorrelated errors.

Integrated, Near-Integrated, and Fractionally Integrated Processes

Modeling dynamic processes explicitly involves more than simply putting a lagged endogenous variable on the right-hand side of a regression equation. There are important theoretical and technical issues that should be addressed. One set of technical issues concerns the stationarity of the variables of interest. A variable Y_t is said to be (weakly) stationary if it has a constant mean $[E(Y_t) = \mu]$, a constant variance $[E(Y_t - \mu)^2 = \gamma_0]$, and a constant covariance $[E(Y_t - \mu)(Y_{t-k} - \mu) = \gamma_k]$ at lag $t-k$ regardless of the values of t or k.

Consider the case of a simple first-order autoregressive DGP

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t, \quad (8)$$

where ε_t is a white-noise error process and $\text{cov}(\varepsilon_t, \varepsilon_{t-k}) = 0$ for all t and k. If $|\phi_1| < 1.0$, the process is stationary. However, if $\phi_1 = 1.0$, it is nonstationary. The defining characteristic of such a process, commonly called a random walk, is that it "never forgets." Since the ϕ_1 coefficient for Y_{t-1} equals 1.0, stochastic shocks (ε_t) are not discounted as the process evolves; rather, they accumulate over time at their full value.

One can see this by repeated recursive substitution for Y_{t-1} in Eq. (8). Doing so shows that $Y_t = Y_0 + \sum \varepsilon_i (i = 1, t)$, with Y_0 being the initial value of Y. This process has a nonconstant variance and is said to exhibit a stochastic trend. Modifying Eq. (8) by introducing a constant β_0 yields $Y_t = \beta_0 + Y_{t-1} + \varepsilon_t$. Repeated recursive substitution shows that this latter model, called a random walk with drift, produces nonstationary variables. After substitution, the model contains both $\sum \varepsilon_i$ and $t\beta_0$, with the latter being the "drift" component that introduces a deterministic trend.

To illustrate the behavior of these nonstationary processes, a random walk and a random walk with drift (with $\beta_0 = 0.05$) are simulated and then compared with a stationary first-order autoregression (with $\phi_1 = 0.50$). The errors for the three variables are white noise, with constant variances of 1.0. As Fig. 1 shows, the random walk and random walk with drift soon depart from their initial value (0) and never return during the 240 periods for which the data are generated. In contrast, the stationary first-order autoregressive process with an initial value of 0 repeatedly moves from positive to negative values and vice versa and shows no tendency to "trend" in any direction.

In a classic article from 1974, Granger and Newbold demonstrated that the threats to inference posed by nonstationary variables are decidedly nontrivial. Employing Monte Carlo simulation methods, they generated a large number of independent random walk variables and then performed regressions of the form $Y_t = \beta_0 + \beta_1 X_t + \mu_t$. These regressions rejected the (true) null hypothesis that $\beta_1 = 0$ at alarming rates.

Since the publication of these findings, researchers have come to appreciate the danger of such "spurious regressions." It has become routine to use unit-root tests and other, less formal, diagnostic procedures (graphs, autocorrelation functions) to determine whether variables of interest are nonstationary.

If nonstationarity is suspected, conventional practice is to difference the variables prior to conducting regression analyses. The implicit assumption is that the data-generating process producing the nonstationarity is a random walk or a random walk with drift, rather than a deterministic trend model of the form $Y_t = \pi_0 + \pi_1 t + \eta_t$, where t is time and the π values are model parameters. If the assumption is valid, the variable is said to be integrated of order 1 [$\boldsymbol{I}(1)$] because differencing it once will produce stationarity. Thus, in the random walk case, one has

$$Y_t - Y_{t-1} = 1.0Y_{t-1} - Y_{t-1} + \varepsilon_t. \quad (9)$$

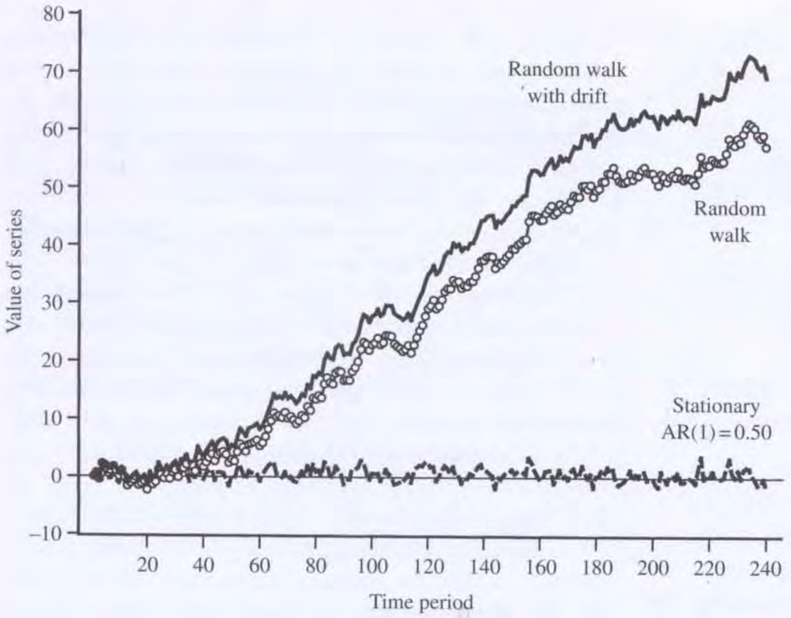

Figure 1 Simulated stationary first-order autoregressive, random walk, and random walk with drift processes.

The resulting variable $(1-L)Y_t = \varepsilon_t$ is stationary. The result for a random walk with drift is similar, except that the expected value of the differenced variable is β_0 rather than zero.

Differencing nonstationary $\boldsymbol{I}(1)$ variables solves the problem of spurious regressions, but does not take into account possible long-run relationships between Y_t and X_s. It is possible to study long-run relationships between nonstationary [$\boldsymbol{I}(1)$] variables if those variables cointegrate. For example, two nonstationary variables, Y_t and X_t, are said to cointegrate if a linear combination of the variables constitutes a stationary variable.

Cointegration cannot be assumed, but must be demonstrated empirically. In 1987, Engle and Granger proposed doing this by regressing Y_t on X_t and testing the residuals (by definition a linear combination of Y and X) for stationarity. Determining that the residuals of this "cointegrating regression" constitute a stationary series warrants the inference that Y and X cointegrate. If so, Y and X can be modeled in error correction form, with the error correction mechanism capturing the long-run relationship between them:

$$(1-L)Y_t = \beta_0 + \beta_1(1-L)X_t - \alpha(Y_{t-1} - cX_{t-1}) + v_t. \tag{10}$$

In this model, the expression $Y_{t-1} - cX_{t-1}$ constitutes the error correction mechanism. Given the negative feedback properties of error correction, the parameter α is expected to carry a negative sign. Its absolute value calibrates the speed with which shocks to the system are reequilibrated by the cointegrating relationship between Y and X. Rearranging terms shows that models such as Eq. (10) are variants of the more familiar autoregressive distributed lag form. For example, Eq. (10) may be written as:

$$Y_t = \beta_0 + (1-\alpha)Y_{t-1} + \beta_1 X_t + (\alpha c - \beta_1)X_{t-1} + v_t. \tag{11}$$

Note also that since all variables in a model such as Eq. (10) are stationary, the spurious regression problem does not arise. Thus, if other conventional assumptions hold, the parameters in model (10) may be estimated via OLS. Engle and Granger suggest a two-step process, where step 1 is to regress Y on X in levels. Assuming that the residuals from this regression are stationary, step 2 is to estimate the parameters in an error correction model such as Eq. (10). Other analysts have advocated a one-step method in which all coefficients in an error correction model are estimated simultaneously. If Y and X do not cointegrate, α will not differ significantly from zero.

An error correction model specification is attractive because it enables one to study both short- and long-run relationships among nonstationary variables. However, establishing that a variable is nonstationary in the classic sense can be difficult. This is because the principal statistical tool for this purpose, unit-root tests, has low statistical power in the face of alternative DGPs that produce highly persistent data. Two such alternatives are the near-integrated and fractionally integrated cases. A near-integrated variable is the product of an autoregressive process [e.g., model (8) above], where the ϕ_1 parameter is slightly less than 1.0 (e.g., 0.95). In this case, unit-root tests are prone to fail to reject the null hypothesis of nonstationarity, even though ϕ_1 actually is less than 1.0.

Unit-root tests also are apt to prove misleading when the DGP produces a fractionally integrated, or long-memory, variable. As its name implies, the concept of fractional integration relaxes the assumption that variables must have integer orders of integration. Rather than being integrated at orders, say 0 or 1, a variable may be generated by an autoregressive, fractionally integrated, moving average (ARFIMA) process such as:

$$(1 - L)^d Y_t = \frac{\varphi(L)}{\phi(L)} \omega_t. \quad (12)$$

This is a generalization of the familiar autoregressive, integrated, moving average (ARIMA) class of models, with ϕ and φ representing autoregressive and moving average parameters.

The fractional differencing parameter d can vary from -0.5 to 1.0. When it is ≥ 0.5, the series is nonstationary although, unlike a random walk, shocks do eventually decay. The simplest member of this model class is called fractional Gaussian noise [i.e., $(1-L)^d Y_t = \omega_t$]. The left-hand side may be expanded as an (infinite) autoregression:

$$1 - dL - (1/2)d(1-d)L^2 - (1/6)d(1-d)(2-d)L^3 - \cdots - (1-j!)d(1-d)(2-d)\cdots[(j-1)-d]L^j - \cdots.$$

Figure 2 presents simulated examples of near-integrated ($\phi_1 = 0.95$) and fractionally integrated ($d = 0.95$) data in comparison with the aforementioned typical first-order autoregressive series where $\phi_1 = 0.50$. Although the near-integrated and fractionally integrated series do not exhibit the obvious trending behavior of the random walk depicted in Fig. 1, they clearly behave much differently than the garden variety AR(1), moving away from their initial value (0) for very long periods.

The persistence (memory) in the near-integrated and fractionally integrated data means that they have significant autocorrelations at very long lags. This point is illustrated clearly by calculating autocorrelation functions (ACFs) for these series (see Fig. 3). Although the ACFs for the near-integrated and fractionally integrated series decline more quickly than does the ACF for the random walk, they are very different than the ACF for the conventional AR(1) series. For these simulated data, the latter becomes statistically insignificant ($p > 0.05$) after 2 lags, whereas the near and fractionally integrated series remain significant through 18 and 15 lags, respectively.

Near-integrated variables and fractionally integrated variables with d values in the nonstationary range (i.e., >0.5) can create threats to inference similar to those posed by random walks. In this regard, DeBoef has demonstrated that single-equation error correction models are useful for analyzing near-integrated variables. In 2003, Clarke and Lebo cited several procedures for estimating the d parameter (and associated standard error) and discussed how the ideas of cointegration and error correction may be generalized to handle the cases where variables of interest are fractionally integrated and fractionally cointegrated.

The preceding discussion has emphasized statistical issues and techniques that are relevant for the analysis of autocorrelated data. These matters are very important but, as emphasized, a crucial consideration is theoretically guided model specification. This is the topic of the next section.

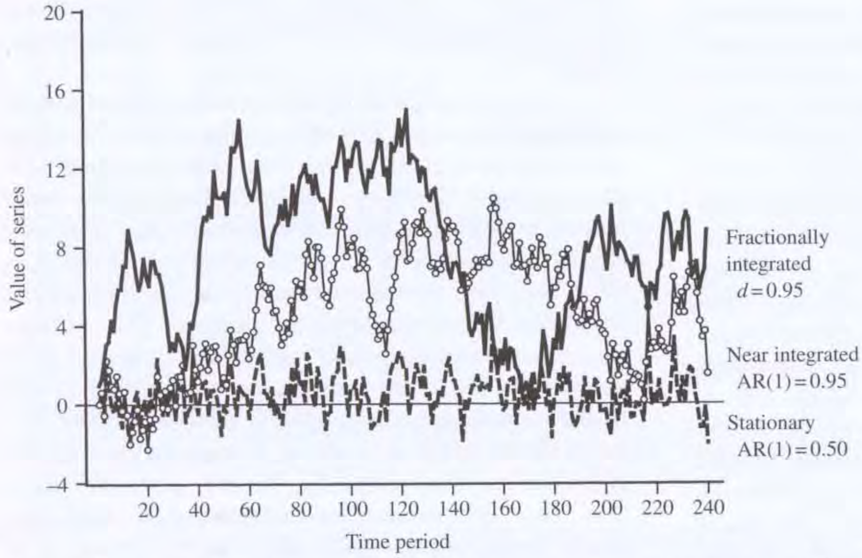

Figure 2 Simulated stationary first-order autoregressive, near-integrated, and fractionally integrated processes.

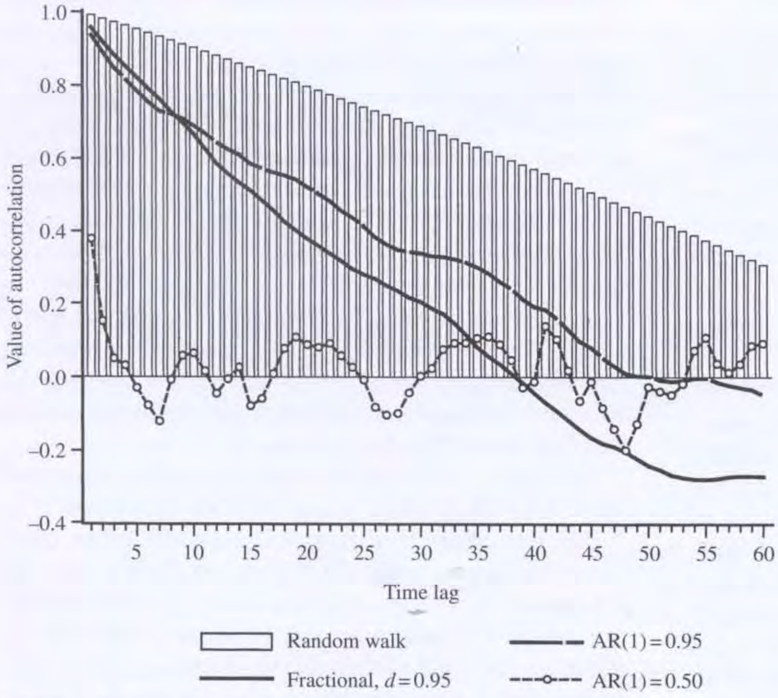

Figure 3 Autocorrelation functions for simulated autoregressive, near-integrated, fractionally integrated, and random walk processes.

From Practice to Theory: A Model of Macropartisanship

Here it is demonstrated that persistence can be explained theoretically. The example given is a simple model of party identification—so-called macropartisanship. The model makes no claim to be a powerful explanation of macropartisanship. Rather, the weaker claim is made that it serves to illustrate that its highly autoregressive behavior is a matter of specification and theoretical development. Strictly mechanical approaches, which use a variety of statistical patches, fail to address these theoretical concerns.

The behavioral assumptions are primitive, but the model does have the potential to be enriched by further theoretical (formal) revisions that would incorporate findings in, for example, economics, cognitive psychology, communications, and political science. Although these modifications are welcome, the larger point is to incorporate these modifications within a framework that merges both formal and empirical analyses. When this is achieved, persistence—autocorrelation—can be identified at its source and researchers can use theory to assist in differentiating between various types of persistence (i.e., near-integrated versus fractionally integrated processes) that may characterize an empirical realization of a time series of interest.

In the model demonstrated it is assumed that political campaign advertisments influence the public. It is argued that the persistence of rival political parties' political advisors to target and influence (through the use of political advertisement resources) rival party voters reduces the well-known persistence in macropartisanship. As a consequence, shocks to macropartisanship can either be amplified or die out quickly depending on rival political advisor behavior. Moreover, the linkage between the formal and empirical models identifies the persistence as near-integrated.

The model comprises three equations. Each citizen (i) is subject to an event (j) at time (t). The model aggregates across individuals and events so the notation will have only the subscript t. The first equation [Eq. (13)] specifies what influences aggregate party identification (M_t). The variable M_{t-1} accounts for the possibility of persistence. Citizens also have an expectation of what portion of the population will identify with the party ($E_{t-1}M_t$). It is assumed that, in forming their expectations, citizens use all available and relevant information as specified in this model (rational expectations). It is further assumed that party identification depends on how favorably a citizen views the national party (F_t). Finally, party identification can be subject to unanticipated stochastic shocks (realignments) (u_{1t}), where $u_{1t} \sim N(0, \sigma^2_{u_{1t}})$. It is assumed the relations are positive—$a_1, a_2, a_3 \geq 0$.

Macropartisanship

$$M_t = a_1 M_{t-1} + a_2 E_{t-1} M_t + a_3 F_t + u_{1t}. \quad (13)$$

Equation (14) represents citizens' impression ("favorability") of a political party (F_t). In this model, favorability is a linear function of the lag of favorability (F_{t-1}) and an advertising resource variable (A_t). There are many ways to measure political advertising resources. These measures include but are not limited to the total dollars spent, the dollars spent relative to a rival party (parties), the ratio of dollars spent relative to a rival party (parties), and the tone, content, timing, and geographic location of the advertisements (on a multinomial scale). Data have been collected for individual races but have the potential to be aggregated along partisan lines. For more details on measurement issues consult the Wisconsin Advertising Project Web site at: http://www.polisci.wisc.edu/tvadvertising. u_{2t} is a stochastic shock that represents unanticipated events (uncertainty), where $u_{2t} \sim N(0, \sigma^2_{u_{2t}})$. The parameter $b_1 \geq 0$, while $b_2 \gtreqqless 0$ depending on the tone and content of the advertisement.

Favorability

$$F_t = b_1 F_{t-1} + b_2 A_t + u_{2t}. \quad (14)$$

Equation (15) presents the contingency plan or rule that (rival) political advisors use. It is argued that political advisors track their previous period's advertising resource expenditures (A_{t-1}) and react to that period's favorability rating for the (rival) national party (F_{t-1}). Political advisors also base their current expenditure of advertisement resources on the degree to which macropartisanship (M_t) approximates a prespecified and desired target (M^*). Ideally, political advisors want $(M_t - M^*) = 0$ since it is unrealistic and far too costly to make $M_t = 0$.

It is assumed that the parameters c_1 and c_3 are positive. The parameter c_2 is countercyclical $(-1 \leq c_2 \leq 0)$. This reflects political advisors' willingness to increase or conserve their advertising resources depnding on whether macropartisanship is above (decrease advertising) or below (increase advertising) the target.

Rival Political Advisor

$$A_t = c_1 A_{t-1} + c_2(M_t - M^*) + c_3 F_{t-1}. \quad (15)$$

To obtain the reduced form for macropartisanship, substitute (15) into (14):

$$F_t = b_1 F_{t-1} + b_2[c_1 A_{t-1} + c_2(M_t - M^*) + c_3 F_{t-1}] + u_{2t} \quad (16)$$

and

$$F_t = (b_1 + b_2 c_3) F_{t-1} + b_2 c_1 A_{t-1} + b_2 c_2 (M_t - M^*) + u_{2t}. \quad (17)$$

Now substitute Eq. (17) into Eq. (13):

$$M_t = a_1 M_{t-1} + a_2 E_{t-1} M_t + (b_1 + b_2 c_3) F_{t-1} + b_2 c_1 A_{t-1} + b_2 c_2 (M_t - M^*) + u_{2t} + u_{1t}. \quad (18)$$

Collect terms and divide through by $(1 - b_2 c_2)$:

$$M_t = \frac{a_1}{(1 - b_2c_2)} M_{t-1} + \frac{a_2}{(1 - b_2c_2)} E_{t-1} M_t + \frac{b_2 c_1}{(1 - b_2c_2)} A_{t-1} + \frac{(b_1 + b_2 c_3)}{(1 - b_2c_2)} F_{t-1} - \frac{b_2 c_1}{(1 - b_2c_2)} M^* + \frac{u_{2t} + u_{1t}}{(1 - b_2c_2)}. \quad (19)$$

Simplifying the notation shows that there is an autoregressive component in the reduced form for macropartisanship

$$M_t = \Theta_0 + \Theta_1 M_{t-1} + \Theta_2 E_{t-1} M_t + \Theta_3 A_{t-1} + \Theta_4 F_{t-1} + \varepsilon^*_t, \quad (20)$$

where $\Theta_0 = (b_2 c_1 Y^*)/(1 - b_2 c_2)$, $\Theta_1 = a_1/(1 - b_2 c_2)$, $\Theta_2 = a_2/(1 - b_2 c_2)$, $\Theta_3 = b_2 c_1/(1 - b_2 c_2)$, $\Theta_4 = (b_1 + b_2 c_3)/(1 - b_2 c_2)$, and $\varepsilon^*_t = (u_{2t} + u_{1t})/(1 - b_2 c_2)$.

The system is now simplified to a model of macropartisanship that depends on lagged macropartisanship and also the conditional expectation at time $t-1$ of current macropartisanship. The prior values of advertising and favorability also have an effect.

To close the model, the rational expectations equilibrium can be solved by taking the conditional expectation at time $t-1$ of Eq. (20) and then substituting this result back into Eq. (20)

$$M_t = \Pi_1 + \Pi_2 M_{t-1} + \Pi_3 A_{t-2} + \Pi_4 F_{t-2} + \xi'_t, \quad (21)$$

where

$$\Pi_1 = \left[\frac{\Theta_0}{1 - \Theta_2} - \left(\frac{\Theta_3}{1 - \Theta_2} - \frac{\Theta_4}{1 - \Theta_2} b_2\right) c_2 Y^*\right],$$

$$\Pi_2 = \left[\frac{\Theta_1}{1 - \Theta_2} + \left(\frac{\Theta_3}{1 - \Theta_2} + \frac{\Theta_4}{1 - \Theta_2} b_2\right) c_2\right],$$

$$\Pi_3 = \left[\left(\frac{\Theta_3}{1 - \Theta_2} + \frac{\Theta_4}{1 - \Theta_2} b_2\right) c_1\right],$$

$$\Pi_4 = \left[\frac{\Theta_3}{1 - \Theta_2} c_3 + \frac{\Theta_4}{1 - \Theta_2}(b_1 + b_2 c_3)\right],$$

and

$$\xi'_t = \left(\frac{\Theta_4}{1 - \Theta_2} u_{2t} + \varepsilon^*_t\right).$$

Equation (21) is the minimum state variable solution for macropartisanship. Macropartisanship (M_t) depends, in part, on the lag of macropartisanship (M_{t-1}). More importantly, the persistence of macropartisanship (Π_2) is now shown to depend on the persistence and willingness of rival political advisors to maintain a rival macropartisanship target (c_2).

This can be shown by examining the reduced form AR(1) coefficient expression Π_2:

$$\Pi_2 = \frac{a_1 + b_2c_2(c_1 + b_1 + b_2c_3)}{1 - b_2c_2 - a_2}. \tag{22}$$

The derivative of Eq. (22) with respect to (c_2) is taken and the following relationship is obtained:

$$\frac{\partial \Pi_2}{\partial c_2} = \frac{b_2[a_1(-1 + a_2)(b_1 + c_1 + b_2c_3)]}{(-1 + a_2 + b_2c_2)^2}. \tag{23}$$

The denominator is always positive since it is squared. Given the assumptions about the signs of the coefficients in the model, the numerator is positive as long as $a_2 < 1$. Therefore, under these conditions, it is known that the relationship is positive ($\partial\Pi_2/\partial c_2 > 0$).

The final step is to determine whether the model reaches an equilibrium when it starts from a point of reference that contains nonequilibrium values. Because the model presented has a first-order autoregressive component, the result is straightforward. The stability condition is summarized in the following proposition:

PROPOSITION 1. *Equation* (21) *is a uniquely stationary MSV solution if* $|\Pi_2| < 1$.

The relationship between c_2 and Π_2 is demonstrated in Fig. 4. The following values are used for Eq. (22): $a_1 = a_2 = b_1 = b_2 = c_1 = c_3 = 0.5$. The parameter c_2 ranges from 0.0 to −1.0. As the value of c_2 is varied between 0.0 and −1.0, it is found that the persistence (autocorrelation) in macropartisanship (Π_2)—all things being equal—is zero when $c_2 = -0.5$. On the other hand, macropartisanship becomes highly autoregressive ($\Pi_2 \to 1.0$) when rival political advisors fail to react ($c_2 \to 0.0$) to deviations from their prespecified target. The conclusion from this model is that negative advertisements from rival political parties can influence the persistence of their opponent's national party identification.

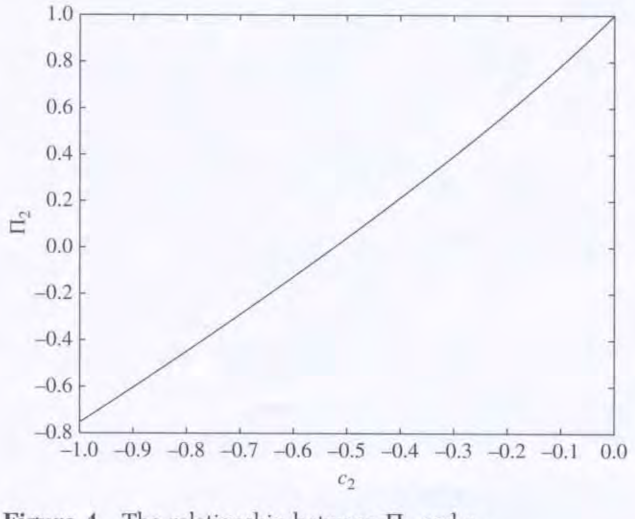

Figure 4 The relationship between Π_2 and c_2.

An Agenda for Future Research

Social scientists' interest in autocorrelation is longstanding. This interest reflects a scientifically unassailable concern to make sound inferences regarding forces driving the dynamics of a variable of interest. As discussed above, researchers historically have interpreted the remit implied by this concern very narrowly, focusing their attention exclusively on autocorrelation in the residuals of an OLS regression analysis. Construed so narrowly, autocorrelation is—to use Hendry's characterization—a "nuisance," with traditional "corrections" constituting possibly unwarranted and misleading model simplifications.

Since the early 1990s, advances in the analysis of integrated, near-integrated, and fractionally integrated data have helped to relieve this situation by explicitly articulating technical concerns with the diagnosis of various forms of autocorrelation and the implications thereof for successful inference to questions of model specification. Although these developments constitute a major step in the right direction, the analytic stance remains highly reactive. Model specification is governed by *ad hoc* inferences regarding characteristics of the data and relatively little attention is given to the theoretical foundations of the model. Above it is argued that it is desirable to redress the balance, by giving greater priority to the development of theoretical models that can generate various forms of autoregressive behavior in a variable being studied.

The argument has been illustrated by developing a small theoretical model of macropartisanship—a topic traditionally of great concern to students of voting, elections, and the stability of democratic political systems—and demonstrating how it implies autoregressive, quite possibly highly autoregressive, behavior in observed time series data. Following Granger, this model could be elaborated in various ways, for example, by specifying autoregressive heterogeneity in the electorate as a means of generating the nonstationary fractional dynamics observed in macropartisanship in mature democracies. The general point is that various forms of autocorrelation can and should be seen as empirical implications of theoretical models. Developing and testing such models constitutes a challenging agenda for future research.

See Also the Following Articles

Spatial Autocorrelation • Time-Series–Cross-Section Data

Further Reading

Beran, J. (1994). *Statistics for Long Memory Processes.* Chapman and Hall, New York.

Box, G. E. P., and Jenkins, G. (1976). *Times Series Analysis: Forecasting and Control,* revised Ed. Holden Day, Oakland, CA.

Clarke, H. D., and Lebo, M. (2003). Fractional (co)integration and governing party support in Britain. *Br. J. Polit. Sci.* **33,** 283–301.

DeBoef, S. (2001). Modeling equilibrium relationships: Error correction models with strongly autoregressive data. *Polit. Anal.* **9,** 78–94.

DeBoef, S., and Granato, J. (2000). Testing for cointegrating relationships with near-integrated data. *Polit. Anal.* **8,** 99–117.

Erikson, R. S., MacKuen, M., and Stimson, J. A. (2002). *The Macro Polity.* Cambridge University Press, Cambridge, UK.

Franses, P. H. (1998). *Time Series Models for Business and Economic Forecasting.* Cambridge University Press, Cambridge, UK.

Ghysels, E., Swanson, N. R., and Watson, M. W. (2001). *Collected Papers of Clive W. J. Granger, Volume II, Causality, Integration and Cointegration, and Long Memory.* Cambridge University Press, Cambridge, UK.

Godfrey, L. G. (1988). *Misspecification Tests in Econometrics.* Cambridge University Press, Cambridge, UK.

Gujarati, D. (2003). *Basic Econometrics,* 4th Ed. McGraw-Hill/Irwin, New York.

Hendry, D. (1995). *Dynamic Econometrics.* Oxford University Press, Oxford, UK.

Hendry, D., and Doornik, J. (2001). *Empirical Econometric Modelling Using PcGive,* Vol. 1. Timberlake Consultants, London.

Maddala, G. S., and Kim, I. M. (1998). *Unit Roots, Cointegration, and Structural Change.* Cambridge University Press, Cambridge, UK.

McCallum, B. T. (1983). On nonuniqueness in linear rational expectations models: An atttempt at perspective. *J. Monetary Econ.* **11,** 134–168.